



Departamento de Física, Facultad de Ciencias, Universidad de Chile.  
Las Palmeras 3425, Ñuñoa. Casilla 653, Correo 1, Santiago  
FONO: 562 678 7276      FAX: 562 271 2973  
E-MAIL: [secretaria@fisica.ciencias.uchile.cl](mailto:secretaria@fisica.ciencias.uchile.cl)

---

*Apuntes de un curso de*  
**FÍSICA MATEMÁTICA**

José Rogan C.  
Víctor Muñoz G.



# Índice

<b>I</b>	<b>Análisis Vectorial</b>	<b>3</b>
<b>1</b>	<b>Análisis vectorial en coordenadas curvilíneas y tensores.</b>	<b>5</b>
1.1	Coordenadas ortogonales. . . . .	5
1.2	Operadores diferenciales vectoriales. . . . .	8
1.3	Sistemas especiales de coordenadas: introducción . . . . .	11
1.4	Coordenadas circulares cilíndricas $(\rho, \varphi, z)$ . . . . .	12
1.5	Coordenadas polares esféricas $(r, \theta, \varphi)$ . . . . .	14
1.6	Análisis tensorial. . . . .	17
1.7	Contracción y producto directo. . . . .	21
1.8	Regla del cociente. . . . .	23
1.9	Pseudotensores y tensores duales. . . . .	24
1.10	Tensores no cartesianos, diferenciación covariante. . . . .	30
1.11	Operadores diferenciales de tensores. . . . .	35
<b>2</b>	<b>Determinantes y matrices.</b>	<b>39</b>
2.1	Determinantes. . . . .	39
2.2	Matrices. . . . .	46
2.3	Matrices ortogonales. . . . .	53
2.4	Matrices Hermíticas, matrices unitarias. . . . .	62
2.5	Diagonalización de matrices. . . . .	65
2.6	Matrices normales. . . . .	72
<b>3</b>	<b>Teoría de grupo.</b>	<b>79</b>
3.1	Introducción. . . . .	79
3.2	Generadores de grupos continuos. . . . .	83
3.3	Momento angular orbital. . . . .	96
3.4	Grupo homogéneo de Lorentz. . . . .	100
3.5	Covarianza de Lorentz de las Ecuaciones de Maxwell. . . . .	103
<b>4</b>	<b>Series infinitas.</b>	<b>109</b>
4.1	Conceptos fundamentales . . . . .	109
4.2	Pruebas de Convergencia . . . . .	112
4.2.1	Pruebas de comparación. . . . .	112
4.2.2	Prueba de la raíz de Cauchy. . . . .	113
4.2.3	Prueba de la razón de D' Alembert o Cauchy. . . . .	114

4.2.4	Prueba integral de Cauchy o Maclaurin. . . . .	115
4.2.5	Prueba de Kummer. . . . .	117
4.2.6	Prueba de Raabe. . . . .	117
4.2.7	Prueba de Gauss. . . . .	118
4.2.8	Mejoramiento de convergencia. . . . .	119
4.3	Series alternadas. . . . .	120
4.3.1	Criterio de Leibniz. . . . .	120
4.3.2	Convergencia absoluta. . . . .	121
4.4	Álgebra de series. . . . .	122
4.4.1	Mejoramiento de la convergencia, aproximaciones racionales. . . . .	123
4.4.2	Reordenamiento de series dobles. . . . .	124
4.5	Series de funciones. . . . .	126
4.5.1	Convergencia uniforme. . . . .	127
4.5.2	Prueba $M$ de Weierstrass. . . . .	128
4.5.3	Prueba de Abel. . . . .	129
4.6	Expansión de Taylor. . . . .	130
4.6.1	Teorema de Maclaurin. . . . .	132
4.6.2	Teorema Binomial. . . . .	134
4.6.3	Expansión de Taylor de más de una variable. . . . .	136
4.7	Series de potencias. . . . .	136
4.7.1	Convergencia. . . . .	136
4.8	Convergencia uniforme y absoluta. . . . .	137
4.8.1	Continuidad. . . . .	137
4.8.2	Diferenciación e integración. . . . .	137
4.8.3	Teorema de la singularidad. . . . .	137
4.8.4	Inversión de series de potencia. . . . .	139
4.9	Integrales elípticas. . . . .	140
4.9.1	Definiciones. . . . .	141
4.9.2	Expansión de series. . . . .	142
4.9.3	Valores límites. . . . .	143
4.10	Números de Bernoulli. . . . .	144
4.10.1	Funciones de Bernoulli. . . . .	146
4.10.2	Fórmula de integración de Euler-Maclaurin. . . . .	147
4.10.3	Función zeta de Riemann. . . . .	148
4.10.4	Mejoramiento de la convergencia. . . . .	151
4.11	Series asintóticas o semiconvergentes. . . . .	151
4.11.1	Función gama incompleta. . . . .	152
4.11.2	Integrales coseno y seno. . . . .	154
4.11.3	Definición de series asintóticas. . . . .	155
4.11.4	Aplicaciones a cálculo numérico. . . . .	156
4.12	Productos infinitos. . . . .	157
4.12.1	Convergencia de un producto infinito. . . . .	157
4.12.2	Funciones seno, coseno y gama. . . . .	158

<b>5</b>	<b>Funciones de una variable compleja I.</b>	<b>161</b>
5.1	Algebra compleja. . . . .	162
5.1.1	Conjugación compleja. . . . .	164
5.1.2	Funciones de una variable compleja. . . . .	165
5.2	Condiciones de Cauchy-Riemann. . . . .	166
5.3	Teorema integral de Cauchy. . . . .	169
5.3.1	Integrales de contorno. . . . .	169
5.3.2	Prueba del teorema de Stoke. . . . .	170
5.3.3	Prueba de Cauchy-Goursat. . . . .	172
5.3.4	Regiones múltiplemente conexas. . . . .	174
5.4	Fórmula integral de Cauchy. . . . .	175
5.4.1	Derivadas. . . . .	176
5.4.2	Teorema de Morera. . . . .	177
5.5	Expansión de Laurent. . . . .	179
5.5.1	Expansión de Taylor. . . . .	179
5.5.2	Principio de Reflexión de Schwarz. . . . .	180
5.5.3	Continuación analítica. . . . .	181
5.5.4	Permanencia de la forma algebraica. . . . .	182
5.5.5	Serie de Laurent. . . . .	183
5.6	Mapeo. . . . .	185
5.6.1	Traslación. . . . .	186
5.6.2	Rotación. . . . .	186
5.6.3	Inversión. . . . .	187
5.6.4	Puntos de ramificación y funciones multivaluadas. . . . .	189
5.7	Mapeo conforme . . . . .	193
<b>6</b>	<b>Funciones de variable compleja II.</b>	<b>195</b>
6.1	Singularidades. . . . .	195
6.1.1	Polos. . . . .	195
6.1.2	Puntos de ramificación. . . . .	196
6.2	Cálculo del residuo. . . . .	198
6.2.1	Teorema del residuo. . . . .	198
6.2.2	Valor principal de Cauchy. . . . .	200
6.2.3	Expansión en polos de funciones meromórficas. . . . .	202
6.2.4	Expansión en producto de funciones enteras. . . . .	203
6.2.5	Evaluación de integrales definidas. . . . .	204
6.2.6	Evaluación de integrales definidas: $\int_0^{2\pi} f(\text{sen } \theta, \text{cos } \theta)d\theta$ . . . . .	205
6.2.7	Evaluaciones de integrales definidas: $\int_{-\infty}^{\infty} f(x)dx$ . . . . .	206
6.2.8	Evaluación de integrales definidas: $\int_{-\infty}^{\infty} f(x)e^{iax} dx$ . . . . .	207
6.2.9	Evaluación de integrales definidas: formas exponenciales. . . . .	213
6.2.10	Residuos de un polo de orden $m$ . . . . .	216
6.3	Relaciones de dispersión. . . . .	216
6.3.1	Relaciones de Simetría. . . . .	218
6.3.2	Dispersión óptica. . . . .	218
6.3.3	La relación de Parseval. . . . .	219

6.3.4	Causalidad. . . . .	220
6.4	El método de <i>steepest descents</i> . . . . .	221
<b>7</b>	<b>Ecuaciones diferenciales.</b>	<b>229</b>
7.1	Ecuaciones diferenciales parciales . . . . .	229
7.1.1	Ejemplos de PDE. . . . .	230
7.1.2	Clases de PDE y característica. . . . .	232
7.1.3	Las PDE no lineales. . . . .	234
7.1.4	Condiciones de borde. . . . .	235
7.2	Ecuaciones diferenciales de primer orden. . . . .	235
7.2.1	Variables separables. . . . .	236
7.2.2	Ecuaciones diferenciales exactas. . . . .	237
7.2.3	Ecuaciones diferenciales ordinarias de primer orden lineales. . . . .	238
7.2.4	Conversión a una ecuación integral. . . . .	241
7.3	Separación de variables. . . . .	241
7.3.1	Coordenadas cartesianas. . . . .	241
7.3.2	Coordenadas cilíndricas circulares. . . . .	243
7.3.3	Coordenadas polares esféricas. . . . .	244

# Índice de Figuras

1.1	Elemento de volumen curvilíneo. . . . .	9
1.2	Elemento de área curvilíneo con $q_1 = \text{constante}$ . . . . .	10
1.3	Coordenadas circulares cilíndricas. . . . .	12
1.4	Vectores unitarios en coordenadas circulares cilíndricas. . . . .	13
1.5	Elementos de área en coordenadas polares esféricas. . . . .	15
1.6	Coordenadas polares esféricas. . . . .	16
1.7	Inversión de coordenadas cartesianas, vector polar. . . . .	24
1.8	Inversión de coordenadas cartesianas, vector axial. . . . .	25
1.9	(a) Espejo en el plano $xz$ ; (b) Espejo en el plano $xz$ . . . . .	26
2.1	Sistemas de coordenadas cartesianos. . . . .	54
2.2	Sistemas de coordenadas rotados en dos dimensiones. . . . .	56
2.3	(a) Rotación respecto al eje $x_3$ en un ángulo $\alpha$ ; (b) Rotación respecto a un eje $x'_2$ en un ángulo $\beta$ ; (c) Rotación respecto a un eje $x''_3$ en un ángulo $\gamma$ . . . . .	59
2.4	Vector fijo con coordenadas rotadas. . . . .	61
2.5	Elipsoide del momento de inercia. . . . .	66
2.6	Vector fijo con coordenadas rotada. . . . .	75
3.1	Ilustración de la ecuación (3.13). . . . .	84
3.2	Ilustración de $M' = U M U^\dagger$ ecuación (3.42). . . . .	89
3.3	Octeto bariónico diagrama de peso para SU(3). . . . .	94
3.4	Separación de masa bariónica. . . . .	95
3.5	Separación de masa bariónica. . . . .	96
4.1	Test de comparación. . . . .	113
4.2	Comparación de integral con suma de bloques . . . . .	116
4.3	Rearreglo de serie armónica . . . . .	124
4.4	Series dobles. . . . .	125
4.5	Series dobles. . . . .	125
4.6	Series dobles. . . . .	126
4.7	Convergencia uniforme. . . . .	127
4.8	Péndulo simple . . . . .	140
4.9	Integrales elípticas. . . . .	142
4.10	Función zeta de Riemman. . . . .	149
4.11	Sumas parciales. . . . .	153
5.1	Plano complejo. . . . .	162

5.2	Mapeo en el plano complejo . . . . .	164
5.3	Complejo conjugado . . . . .	165
5.4	Aproximaciones a $z_0$ , . . . . .	167
5.5	Camino de integración . . . . .	170
5.6	Dominio simplemente conexo. . . . .	171
5.7	Contorno de Cauchy-Goursat. . . . .	173
5.8	Contorno cerrado en región múltiplemente conexa. . . . .	174
5.9	Conversión a simplemente conexa. . . . .	174
5.10	Exclusión de un punto singular. . . . .	176
5.11	Centro de expansión y puntos singulares. . . . .	179
5.12	Reflexión de Schwarz. . . . .	181
5.13	Continuación analítica. . . . .	181
5.14	Región anular. . . . .	183
5.15	Traslación. . . . .	186
5.16	Rotación. . . . .	187
5.17	Inversión. . . . .	188
5.18	Inversión, línea-círculo. . . . .	189
5.19	Mapeo en coordenadas hiperbólicas. . . . .	190
5.20	Línea de corte. . . . .	191
5.21	Superficie de Riemann para $\ln z$ . . . . .	192
5.22	Mapeo conforme. . . . .	194
6.1	Contorno en torno a punto de ramificación. . . . .	197
6.2	Singularidades aisladas. . . . .	199
6.3	<i>Bypass</i> de puntos singulares. . . . .	200
6.4	Cerrando el contorno. . . . .	201
6.5	Valor principal de Cauchy. . . . .	202
6.6	Contorno de integración. . . . .	206
6.7	Desigualdad. . . . .	208
6.8	Contorno de integración. . . . .	210
6.9	Contorno de integración. . . . .	211
6.10	Contorno de integración. . . . .	213
6.11	Contorno de integración. . . . .	215
6.12	Contorno de integración. . . . .	217
6.13	Punto de ensilladura. . . . .	223
6.14	Contornos para las funciones de Hankel. . . . .	224

*Segundo Curso*

**MÉTODOS DE LA FÍSICA  
MATEMÁTICA I**



**Parte I**  
**Análisis Vectorial**



# Capítulo 1

## Análisis vectorial en coordenadas curvilíneas y tensores.

versión final 1.7-020401<sup>1</sup>

Anteriormente nos hemos restringido casi por completo a coordenadas cartesianas. Estas coordenadas ofrecen la ventaja de que los tres vectores unitarios,  $\hat{x}$ ,  $\hat{y}$ ,  $\hat{z}$ , son constantes tanto en dirección como en magnitud. Infortunadamente, no todos los problemas físicos se adaptan bien a una solución en coordenadas cartesianas. Por ejemplo, si tenemos un problema de fuerzas centrales,  $\vec{F} = \hat{r}F(r)$ , tal como una fuerza gravitacional o electrostática, las coordenadas cartesianas son particularmente inapropiadas. Tales problemas llaman a usar un sistema de coordenadas en el cual la distancia radial es tomada como una de las coordenadas, es decir, coordenadas polares esféricas.

El sistema de coordenadas debe ser elegido apropiadamente para el problema, explotando cualquier restricción o simetría presente en él.

Naturalmente, hay un precio que pagar por usar un sistema no cartesiano de coordenadas. Debemos desarrollar nuevas expresiones para el gradiente, la divergencia, el rotor y el laplaciano en estos nuevos sistemas de coordenadas. En este capítulo desarrollaremos tales expresiones de una manera muy general para luego particularizarla a los sistemas de coordenadas más usados: coordenadas circulares cilíndricas y coordenadas polares esféricas.

### 1.1 Coordenadas ortogonales.

En coordenadas cartesianas nosotros tratamos con tres familias de planos mutuamente perpendiculares:  $x = \text{constante}$ ,  $y = \text{constante}$ , y  $z = \text{constante}$ . Imaginemos que imponemos sobre este sistema otras tres familias de superficies. Las superficies de una familia no necesariamente son paralelas unas con otras y ellas pueden no ser planos. Las tres nuevas familias de superficies no necesariamente son mutuamente perpendiculares, pero por simplicidad, rápidamente impondremos esta condición. Podemos describir cualquier punto  $(x, y, z)$  como la intersección de tres planos en coordenadas cartesianas o como la intersección de las tres superficies que forman nuestras nuevas coordenadas curvilíneas. Describiendo las superficies de las coordenadas curvilíneas por  $q_1 = \text{constante}$ ,  $q_2 = \text{constante}$ ,  $q_3 = \text{constante}$ , podemos identificar nuestro punto por  $(q_1, q_2, q_3)$  tanto como por  $(x, y, z)$ . Esto significa que en

---

<sup>1</sup>Este capítulo está basado en el segundo capítulo del libro: *Mathematical Methods for Physicists, fourth edition* de George B. Arfken & Hans J. Weber, editorial ACADEMIC PRESS.

principio podemos escribir

Coordenadas generales curvilínea,      Coordenadas circulares cilíndricas

$$\begin{array}{ll}
 q_1, q_2, q_3 & \rho, \varphi, z \\
 x = x(q_1, q_2, q_3) & -\infty < x = \rho \cos \varphi < \infty \\
 y = y(q_1, q_2, q_3) & -\infty < y = \rho \operatorname{sen} \varphi < \infty \\
 z = z(q_1, q_2, q_3) & -\infty < z = z < \infty,
 \end{array} \tag{1.1}$$

especificando  $x, y, z$  en términos de los  $q_i$  y las relaciones inversas,

$$\begin{array}{ll}
 q_1 = q_1(x, y, z) & 0 \leq \rho = (x^2 + y^2)^{1/2} < \infty \\
 q_2 = q_2(x, y, z) & 0 \leq \arctan(y/x) < 2\pi \\
 q_3 = q_3(x, y, z) & -\infty \leq z = z < \infty.
 \end{array} \tag{1.2}$$

Como una ilustración específica de los abstractos  $(q_1, q_2, q_3)$  incluimos las ecuaciones de transformación para las coordenadas circulares cilíndricas. Para cada familia de superficies  $q_i = \text{constante}$ , podemos asociar un vector unitario  $\hat{e}_i$  normal a la superficie  $q_i = \text{constante}$  y en la dirección de crecimiento de  $q_i$ . Entonces un vector  $\vec{V}$  puede ser escrito

$$\vec{V} = \hat{e}_1 V_1 + \hat{e}_2 V_2 + \hat{e}_3 V_3. \tag{1.3}$$

Los  $\hat{e}_i$  están normalizados por  $\hat{e}_i^2 = 1$  y forman un sistema de coordenadas diestro con volumen  $\hat{e}_1 \cdot (\hat{e}_2 \times \hat{e}_3) > 0$ .

Diferenciando a  $x$  en (1.1) conduce a

$$dx = \frac{\partial x}{\partial q_1} dq_1 + \frac{\partial x}{\partial q_2} dq_2 + \frac{\partial x}{\partial q_3} dq_3, \tag{1.4}$$

de la misma manera diferenciando  $y$  y  $z$ . A partir del teorema de Pitágoras en coordenadas cartesianas, el cuadrado de la distancia entre dos puntos vecinos es

$$ds^2 = dx^2 + dy^2 + dz^2. \tag{1.5}$$

El cuadrado del elemento de distancia en nuestras coordenadas curvilíneas puede ser escrito

$$\begin{aligned}
 ds^2 = & g_{11} dq_1^2 + g_{12} dq_1 dq_2 + g_{13} dq_1 dq_3 + g_{21} dq_2 dq_1 + g_{22} dq_2^2 \\
 & g_{23} dq_2 dq_3 + g_{31} dq_3 dq_1 + g_{32} dq_3 dq_2 + g_{33} dq_3^2 = \sum_{ij} g_{ij} dq_i dq_j,
 \end{aligned} \tag{1.6}$$

donde<sup>2</sup>

$$g_{ij} = \frac{\partial x}{\partial q_i} \frac{\partial x}{\partial q_j} + \frac{\partial y}{\partial q_i} \frac{\partial y}{\partial q_j} + \frac{\partial z}{\partial q_i} \frac{\partial z}{\partial q_j}. \tag{1.7}$$

---

<sup>2</sup>Los  $dq_i$  son arbitrarios. Por ejemplo, haciendo  $dq_2 = dq_3 = 0$  aislamos  $g_{11}$ .

Estos coeficientes  $g_{ij}$  pueden verse como especificando la naturaleza del sistema de coordenadas  $(q_1, q_2, q_3)$ . Colectivamente estos coeficientes son referidos como la métrica. Mostraremos más adelante que forman un tensor simétrico de segundo rango. En Relatividad General los componentes son determinados por las propiedades de la materia. La Geometría se mezcla con la Física.

En este punto nos limitamos a sistemas de coordenadas ortogonales (superficies mutuamente perpendiculares)

$$g_{ij} = 0, \quad i \neq j, \quad (1.8)$$

y  $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$ . Veremos algo de sistemas no ortogonales en la sección del análisis tensorial. Para simplificar la notación escribimos  $g_{ii} = h_i^2$  tal que

$$ds^2 = (h_1 dq_1)^2 + (h_2 dq_2)^2 + (h_3 dq_3)^2 = \sum_i (h_i dq_i)^2. \quad (1.9)$$

Los específicos sistemas de coordenadas ortogonales son descritos en las secciones siguientes, especificando estos factores de escala  $h_1, h_2$  y  $h_3$ . Por otra parte, los factores de escala pueden ser identificados por la relación

$$ds_i = h_i dq_i, \quad (1.10)$$

para algún  $dq_i$  dado, manteniendo los otros  $q_i$  constantes. Notemos que las tres coordenadas curvilíneas  $(q_1, q_2, q_3)$  no necesariamente son una longitud. Los factores de escala  $h_i$  pueden depender de las  $q_i$  y pueden tener dimensiones. Los productos  $h_i dq_i$  deben tener dimensiones de longitud y ser positivos. El vector de desplazamiento diferencial  $d\vec{s}$  puede ser escrito

$$d\vec{s} = h_1 dq_1 \hat{e}_1 + h_2 dq_2 \hat{e}_2 + h_3 dq_3 \hat{e}_3 = \sum_i h_i dq_i \hat{e}_i.$$

Usando las componentes curvilíneas, encontramos que la integral de línea llega a ser

$$\int \vec{V} \cdot d\vec{s} = \sum_i \int V_i h_i dq_i.$$

De la ecuación (1.10) podemos inmediatamente desarrollar los elementos de área y de volumen

$$d\sigma_{ij} = ds_i ds_j = h_i h_j dq_i dq_j \quad (1.11)$$

y

$$d\tau = ds_1 ds_2 ds_3 = h_1 h_2 h_3 dq_1 dq_2 dq_3. \quad (1.12)$$

Estas expresiones coinciden con los resultados obtenidos al usar las transformaciones explícitamente y el Jacobiano.

A partir de la ecuación (1.11) un elemento de área puede ser expandido:

$$\begin{aligned} d\vec{\sigma} &= ds_2 ds_3 \hat{e}_1 + ds_3 ds_1 \hat{e}_2 + ds_1 ds_2 \hat{e}_3 \\ &= h_2 h_3 dq_2 dq_3 \hat{e}_1 + h_3 h_1 dq_3 dq_1 \hat{e}_2 + h_1 h_2 dq_1 dq_2 \hat{e}_3. \end{aligned}$$

Una integral de superficie llega a ser

$$\int \vec{V} \cdot d\vec{\sigma} = \int V_1 h_2 h_3 dq_2 dq_3 + \int V_2 h_3 h_1 dq_3 dq_1 + \int V_3 h_1 h_2 dq_1 dq_2 . \quad (1.13)$$

Debemos dejar claro que el álgebra vectorial es la misma en coordenadas curvilíneas ortogonales que en coordenadas cartesianas. Específicamente, el producto punto

$$\vec{A} \cdot \vec{B} = A_1 B_1 + A_2 B_2 + A_3 B_3 , \quad (1.14)$$

donde los subíndices indican componentes curvilíneas. Para el producto cruz

$$\vec{A} \times \vec{B} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad (1.15)$$

## 1.2 Operadores diferenciales vectoriales.

### Gradiente.

El punto de partida para desarrollar los operadores gradiente, divergencia y el rotor en coordenadas curvilíneas es nuestra interpretación del gradiente como un vector que tiene la magnitud y dirección de la máxima razón de crecimiento. A partir de esta interpretación las componentes de  $\vec{\nabla}\psi(q_1, q_2, q_3)$  en la dirección normal a la familia de superficies  $q_1 = \text{constante}$  es dado por

$$\hat{e}_1 \cdot \vec{\nabla}\psi = \vec{\nabla}\psi|_1 = \frac{\partial\psi}{\partial s_1} = \frac{\partial\psi}{h_1 \partial q_1} , \quad (1.16)$$

ya que ésta es la razón de cambio de  $\psi$  para variaciones de  $q_1$ , manteniendo  $q_2$  y  $q_3$  fijos. La cantidad  $ds_1$  es un diferencial de longitud en la dirección de incremento de  $q_1$ . El vector unitario  $\hat{e}_1$  indica la dirección. Si repetimos este cálculo para las otras componentes  $q_2$  y  $q_3$  y sumamos vectorialmente obtenemos la expresión del gradiente

$$\begin{aligned} \vec{\nabla}\psi(q_1, q_2, q_3) &= \hat{e}_1 \frac{\partial\psi}{\partial s_1} + \hat{e}_2 \frac{\partial\psi}{\partial s_2} + \hat{e}_3 \frac{\partial\psi}{\partial s_3} \\ &= \hat{e}_1 \frac{\partial\psi}{h_1 \partial q_1} + \hat{e}_2 \frac{\partial\psi}{h_2 \partial q_2} + \hat{e}_3 \frac{\partial\psi}{h_3 \partial q_3} \\ &= \sum_i \hat{e}_i \frac{1}{h_i} \frac{\partial\psi}{\partial q_i} . \end{aligned} \quad (1.17)$$

### Divergencia.

El operador divergencia puede obtenerse de una de sus definiciones

$$\vec{\nabla} \cdot \vec{V}(q_1, q_2, q_3) = \lim_{\int d\tau \rightarrow 0} \frac{\int \vec{V} \cdot d\vec{\sigma}}{\int d\tau} , \quad (1.18)$$

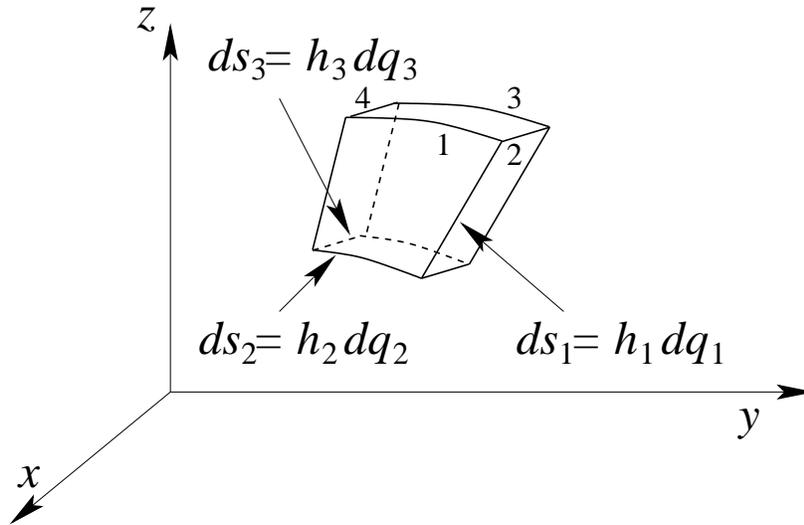


Figura 1.1: Elemento de volumen curvilíneo.

con un volumen diferencial  $h_1 h_2 h_3 dq_1 dq_2 dq_3$  (figura 1.1). Notemos que la dirección positiva ha sido escogida tal que  $(q_1, q_2, q_3)$  o  $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$  formen un sistema diestro,  $\hat{e}_1 \times \hat{e}_2 = \hat{e}_3$ .

La integral de área sobre las dos caras  $q_1 = \text{constante}$  es dada por

$$\left[ V_1 h_2 h_3 + \frac{\partial}{\partial q_1} (V_1 h_2 h_3) dq_1 \right] dq_2 dq_3 - V_1 h_2 h_3 = \frac{\partial}{\partial q_1} (V_1 h_2 h_3) dq_1 dq_2 dq_3, \quad (1.19)$$

considerando las otras componentes

$$\int \vec{V}(q_1, q_2, q_3) \cdot d\vec{\sigma} = \left[ \frac{\partial}{\partial q_1} (V_1 h_2 h_3) + \frac{\partial}{\partial q_2} (V_2 h_3 h_1) + \frac{\partial}{\partial q_3} (V_3 h_1 h_2) \right] dq_1 dq_2 dq_3. \quad (1.20)$$

Dividiendo por la diferencial de volumen produce

$$\vec{\nabla} \cdot \vec{V}(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial q_1} (V_1 h_2 h_3) + \frac{\partial}{\partial q_2} (V_2 h_3 h_1) + \frac{\partial}{\partial q_3} (V_3 h_1 h_2) \right]. \quad (1.21)$$

En la ecuación anterior,  $V_i$  es la componente de  $\vec{V}$  en la dirección  $\hat{e}_i$  en la que crece  $q_i$ . Esto es,  $V_i = \hat{e}_i \cdot \vec{V}$  es la proyección de  $\vec{V}$  sobre la dirección  $\hat{e}_i$ .

Podemos obtener el Laplaciano combinando las ecuaciones (1.17) y (1.21), usando  $\vec{\nabla} = \vec{\nabla}\psi(q_1, q_2, q_3)$ . Esto produce

$$\begin{aligned} \vec{\nabla} \cdot \vec{\nabla}\psi(q_1, q_2, q_3) = \\ \nabla^2 \psi = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial q_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial \psi}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left( \frac{h_3 h_1}{h_2} \frac{\partial \psi}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial \psi}{\partial q_3} \right) \right]. \end{aligned} \quad (1.22)$$

**Rotor.**

Finalmente, para desarrollar  $\vec{\nabla} \times \vec{V}$ , apliquemos el teorema de Stokes y, como con la divergencia, tomemos el límite en que el área de la superficie llega a ser despreciablemente pequeña. Trabajando sobre una componente a la vez, consideremos un diferencial de área en la superficie curvilínea  $q_1 = \text{constante}$ . A partir de

$$\int_S \vec{\nabla} \times \vec{V} \cdot d\vec{s} = \hat{e}_1 \cdot (\vec{\nabla} \times \vec{V}) h_2 h_3 dq_2 dq_3, \quad (1.23)$$

el teorema de Stokes produce

$$\hat{e}_1 \cdot (\vec{\nabla} \times \vec{V}) h_2 h_3 dq_2 dq_3 = \oint \vec{V} \cdot d\vec{s}, \quad (1.24)$$

con la integral de línea sobre la superficie  $q_1 = \text{constante}$ . Siguiendo el *loop* (1, 2, 3, 4) en la figura 1.2,

$$\begin{aligned} \oint \vec{V}(q_1, q_2, q_3) \cdot d\vec{s} &= V_2 h_2 dq_2 + \left[ V_3 h_3 + \frac{\partial}{\partial q_2} (V_3 h_3) dq_2 \right] dq_3 \\ &\quad - \left[ V_2 h_2 + \frac{\partial}{\partial q_3} (V_2 h_2) dq_3 \right] dq_2 - V_3 h_3 dq_3 \\ &= \left[ \frac{\partial}{\partial q_2} (V_3 h_3) - \frac{\partial}{\partial q_3} (V_2 h_2) \right] dq_2 dq_3. \end{aligned} \quad (1.25)$$

Tomamos el signo positivo cuando vamos en la dirección positiva sobre las partes 1 y 2 y negativa sobre las partes 3 y 4 porque aquí vamos en la dirección negativa. Los términos más altos en las expansiones de Taylor son omitidos. Ellos desaparecerán en el límite en que la superficie llega a ser despreciablemente pequeña ( $dq_2 \rightarrow 0, dq_3 \rightarrow 0$ ).

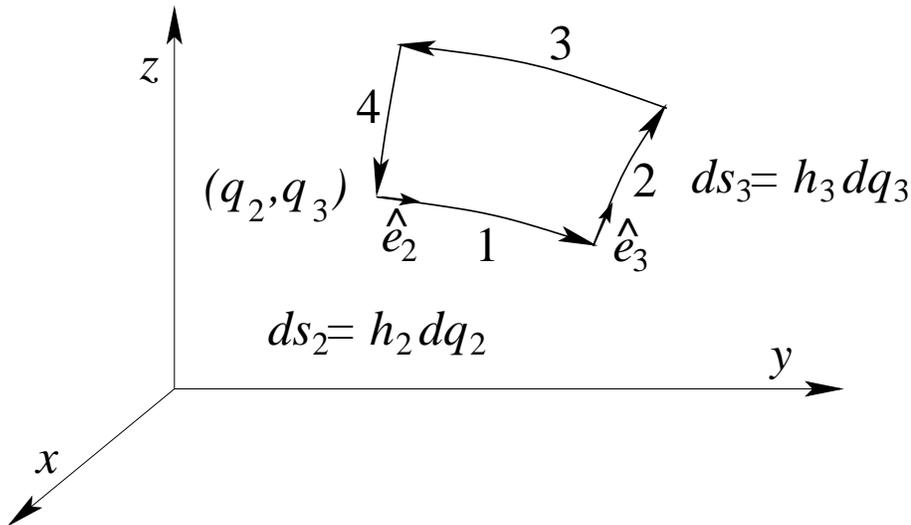


Figura 1.2: Elemento de área curvilíneo con  $q_1 = \text{constante}$ .

Desde la ecuación (1.24)

$$\vec{\nabla} \times \vec{V}|_1 = \frac{1}{h_2 h_3} \left[ \frac{\partial}{\partial q_2} (V_3 h_3) - \frac{\partial}{\partial q_3} (V_2 h_2) \right]. \quad (1.26)$$

Las dos componentes restantes pueden ser evaluadas tomando una permutación cíclica de los índices. Finalmente, podemos escribir el rotor en forma de determinante

$$\vec{\nabla} \times \vec{V} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} \hat{e}_1 h_1 & \hat{e}_2 h_2 & \hat{e}_3 h_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 V_1 & h_2 V_2 & h_3 V_3 \end{vmatrix}. \quad (1.27)$$

Notemos que esta ecuación no es idéntica a la forma del producto cruz de dos vectores, ya que  $\vec{\nabla}$  no es un vector ordinario, es un vector operador.

### 1.3 Sistemas especiales de coordenadas: introducción

Hay once sistemas de coordenadas en los cuales la ecuación tridimensional de Helmholtz puede separarse en tres ecuaciones diferenciales ordinarias<sup>3</sup>. Algunos de estos sistemas adquirieron importancia en el desarrollo histórico de la Mecánica Cuántica. Otros como el bipolar satisfacen necesidades específicas. Debido al desarrollo de computadores de alta velocidad y a la eficiencia de las técnicas de programación se ha reducido la necesidad de utilizar estos sistemas. Nos limitaremos a coordenadas cartesianas, coordenadas polares esféricas, y coordenadas circulares cilíndricas.

#### Coordenadas cartesianas rectangulares.

Este es el sistema más simple de todos:

$$\begin{aligned} h_1 &= h_x = 1, \\ h_2 &= h_y = 1, \\ h_3 &= h_z = 1. \end{aligned} \quad (1.28)$$

Las familias de superficies coordenadas son tres conjuntos de planos paralelos:  $x = \text{constante}$ ,  $y = \text{constante}$ , y  $z = \text{constante}$ . Las coordenadas cartesianas es el único sistema en que todos los  $h_i$  son constantes. Notemos también que los vectores unitarios,  $\hat{e}_1$ ,  $\hat{e}_2$ ,  $\hat{e}_3$  o  $\hat{x}$ ,  $\hat{y}$ ,  $\hat{z}$  tienen direcciones fijas.

Los operadores vectoriales corresponden a

$$\vec{\nabla} \psi = \hat{x} \frac{\partial \psi}{\partial x} + \hat{y} \frac{\partial \psi}{\partial y} + \hat{z} \frac{\partial \psi}{\partial z}, \quad (1.29)$$

---

<sup>3</sup>Ver por ejemplo, Morse and Feshbach.

$$\vec{\nabla} \cdot \vec{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}, \quad (1.30)$$

$$\vec{\nabla} \cdot \vec{\nabla} \psi = \nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}, \quad (1.31)$$

$$\vec{\nabla} \times \vec{V} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_x & V_y & V_z \end{vmatrix}. \quad (1.32)$$

## 1.4 Coordenadas circulares cilíndricas $(\rho, \varphi, z)$ .

En el sistema de coordenadas circulares cilíndricas las tres coordenadas curvilíneas  $(q_1, q_2, q_3)$  son reetiquetadas por  $(\rho, \varphi, z)$ . Las superficies coordenadas mostradas en la figura 1.3 son:

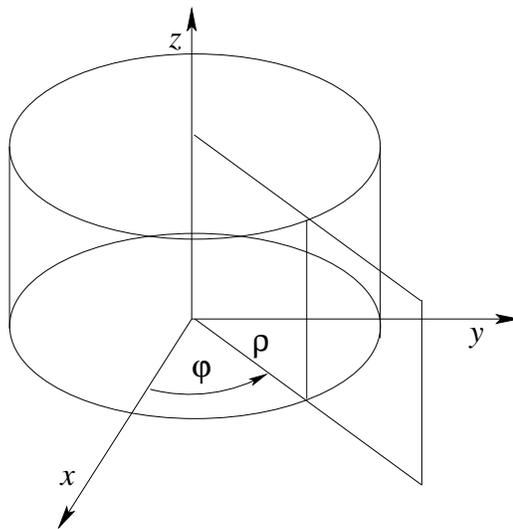


Figura 1.3: Coordenadas circulares cilíndricas.

1. Cilindros circulares derechos que tienen el eje- $z$  como eje común,

$$\rho = (x^2 + y^2)^{1/2} = \text{constante.}$$

2. Semiplanos a través del eje- $z$ ,

$$\varphi = \tan^{-1} \left( \frac{y}{x} \right) = \text{constante.}$$

3. Planos paralelos al plano  $xy$  como en el sistema cartesiano

$$z = \text{constante}.$$

Los límites para  $\rho$ ,  $\varphi$  y  $z$  son

$$0 \leq \rho < \infty, \quad 0 \leq \varphi \leq 2\pi, \quad \text{y} \quad -\infty < z < \infty.$$

Notemos que estamos usando  $\rho$  como la distancia perpendicular al eje  $z$ . Las relaciones de transformación inversas

$$\begin{aligned} x &= \rho \cos \varphi, \\ y &= \rho \operatorname{sen} \varphi, \\ z &= z, \end{aligned} \tag{1.33}$$

de acuerdo a las ecuaciones anteriores los factores de escala son

$$\begin{aligned} h_1 &= h_\rho = 1, \\ h_2 &= h_\varphi = \rho, \\ h_3 &= h_z = 1. \end{aligned} \tag{1.34}$$

Los vectores unitarios  $\hat{e}_1, \hat{e}_2, \hat{e}_3$  son reetiquetados  $(\hat{\rho}, \hat{\varphi}, \hat{z})$ , figura 1.4. El vector unitario  $\hat{\rho}$  es normal a la superficie cilíndrica apuntando en la dirección de crecimiento del radio  $\rho$ . El vector unitario  $\hat{\varphi}$  es tangencial a la superficie cilíndrica, perpendicular al semiplano  $\varphi = \text{constante}$  y apuntando en la dirección de crecimiento del ángulo  $\varphi$ . El tercer vector unitario,  $\hat{z}$ , es el vector unitario cartesiano usual.

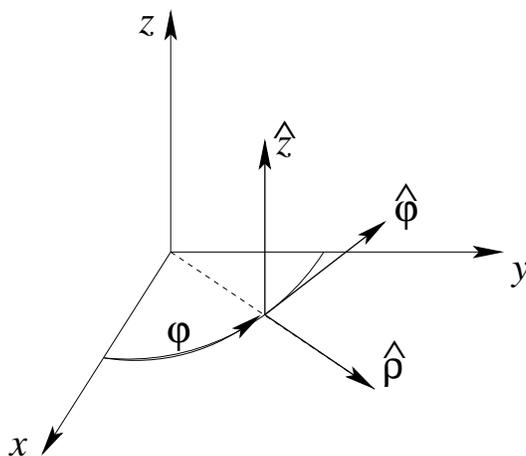


Figura 1.4: Vectores unitarios en coordenadas circulares cilíndricas.

Un vector diferencial de desplazamiento  $d\vec{s}$  puede ser escrito

$$d\vec{s} = \hat{\rho}ds_\rho + \hat{\varphi}ds_\varphi + \hat{z}dz = \hat{\rho}d\rho + \hat{\varphi}\rho d\varphi + \hat{z}dz. \tag{1.35}$$

Los operadores diferenciales que involucran  $\vec{\nabla}$

$$\vec{\nabla}\psi(\rho, \varphi, z) = \hat{\rho}\frac{\partial\psi}{\partial\rho} + \hat{\varphi}\frac{1}{\rho}\frac{\partial\psi}{\partial\varphi} + \hat{z}\frac{\partial\psi}{\partial z}, \quad (1.36)$$

$$\vec{\nabla} \cdot \vec{V} = \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho V_\rho) + \frac{1}{\rho}\frac{\partial V_\varphi}{\partial\varphi} + \frac{\partial V_z}{\partial z}, \quad (1.37)$$

$$\nabla^2\psi = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\psi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{\partial^2\psi}{\partial z^2}, \quad (1.38)$$

$$\vec{\nabla} \times \vec{V} = \frac{1}{\rho} \begin{vmatrix} \hat{\rho} & \rho\hat{\varphi} & \hat{z} \\ \frac{\partial}{\partial\rho} & \frac{\partial}{\partial\varphi} & \frac{\partial}{\partial z} \\ V_\rho & \rho V_\varphi & V_z \end{vmatrix}. \quad (1.39)$$

## 1.5 Coordenadas polares esféricas $(r, \theta, \varphi)$ .

Reetiquetando  $(q_1, q_2, q_3)$  como  $(r, \theta, \varphi)$ , vemos que el sistema de coordenadas polares esféricas consiste en lo siguiente:

1. Esferas concéntricas centradas en el origen,

$$r = (x^2 + y^2 + z^2)^{1/2} = \text{constante.}$$

2. Conos circulares centrados sobre el eje  $z$  y vertice en el origen

$$\theta = \arccos\left(\frac{z}{(x^2 + y^2 + z^2)^{1/2}}\right) = \text{constante.}$$

3. Semiplanos a través del eje- $z$ ,

$$\varphi = \tan^{-1}\left(\frac{y}{x}\right) = \text{constante.}$$

Por nuestra elección arbitraria de la definición de  $\theta$ , el ángulo polar, y  $\varphi$ , el ángulo azimutal, el eje  $z$  queda particularizado para un tratamiento especial. las ecuaciones de transformación son

$$\begin{aligned} x &= r \text{ sen } \theta \cos \varphi, \\ y &= r \text{ sen } \theta \text{ sen } \varphi, \\ z &= r \cos \theta, \end{aligned} \quad (1.40)$$

midiendo  $\theta$  desde el eje  $z$  positivo y  $\varphi$  en el plano  $xy$  desde el eje  $x$  positivo. Los intervalos de valores son  $0 \leq r < \infty$ ,  $0 \leq \theta \leq \pi$ , y  $0 \leq \varphi \leq 2\pi$ . A partir de la ecuación (1.7)

$$\begin{aligned} h_1 &= h_r = 1, \\ h_2 &= h_\theta = r, \\ h_3 &= h_\varphi = r \operatorname{sen} \theta. \end{aligned} \tag{1.41}$$

El elemento de arco

$$d\vec{s} = \hat{r}dr + \hat{\theta}rd\theta + \hat{\varphi}r \operatorname{sen} \theta d\varphi.$$

En este sistema de coordenadas esféricas el elemento de área (para  $r = \text{constante}$ ) es

$$dA = d\sigma_{\theta\varphi} = r^2 \operatorname{sen} \theta d\theta d\varphi, \tag{1.42}$$

el área sombreada en la figura 1.5. Integrando sobre el azimutal  $\varphi$ , encontramos que el elemento de área llega a ser un anillo de ancho  $d\theta$ ,

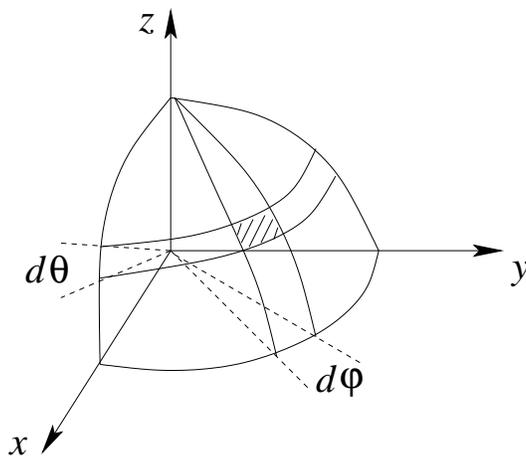


Figura 1.5: Elementos de área en coordenadas polares esféricas.

$$dA = 2\pi r^2 \operatorname{sen} \theta d\theta. \tag{1.43}$$

Esta forma aparece con frecuencia en problemas en coordenadas polares con simetría azimutal. Por definición el stereradian, un elemento de ángulo sólido  $d\Omega$  es dado por

$$d\Omega = \frac{dA}{r^2} = \operatorname{sen} \theta d\theta d\varphi. \tag{1.44}$$

Integrando sobre la superficie esférica completa, obtenemos

$$\int d\Omega = 4\pi.$$

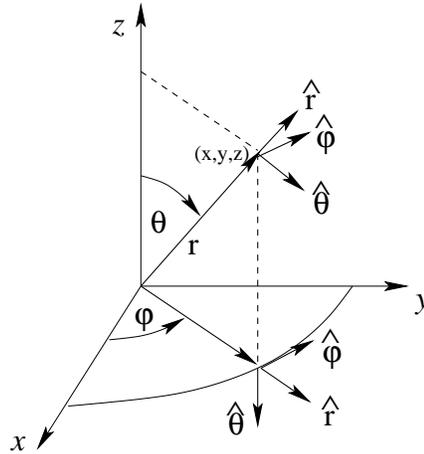


Figura 1.6: Coordenadas polares esféricas.

A partir de la ecuación (1.12) el elemento de volumen es

$$d\tau = r^2 dr \sin \theta d\theta d\varphi = r^2 dr d\Omega . \quad (1.45)$$

Los vectores unitarios en coordenadas polares esféricas son mostrados en la figura 1.6 Debemos enfatizar que los vectores unitarios  $\hat{r}$ ,  $\hat{\theta}$  y  $\hat{\varphi}$  varían en dirección cuando los ángulos  $\theta$  y  $\varphi$  varían. Cuando diferenciamos vectores en coordenadas polares esféricas (o en cualquier otro sistema no cartesiano) estas variaciones de los vectores unitarios con respecto a la posición no deben ser despreciadas. Escribamos los vectores unitarios  $\hat{r}$ ,  $\hat{\theta}$  y  $\hat{\varphi}$  en término de los vectores unitarios cartesianos de dirección fija  $\hat{x}$ ,  $\hat{y}$ ,  $\hat{z}$

$$\begin{aligned} \hat{r} &= \hat{x} \sin \theta \cos \varphi + \hat{y} \sin \theta \sin \varphi + \hat{z} \cos \theta , \\ \hat{\theta} &= \hat{x} \cos \theta \cos \varphi + \hat{y} \cos \theta \sin \varphi - \hat{z} \sin \theta , \\ \hat{\varphi} &= -\hat{x} \sin \varphi + \hat{y} \cos \varphi . \end{aligned} \quad (1.46)$$

Notemos que un vector dado puede ser expresado en un número de diferentes (pero equivalentes) maneras. Por ejemplo, el vector posición  $\vec{r}$  puede ser escrito

$$\begin{aligned} \vec{r} &= \hat{r} r \\ &= \hat{r} (x^2 + y^2 + z^2)^{1/2} \\ &= \hat{x} x + \hat{y} y + \hat{z} z \\ &= \hat{x} r \sin \theta \cos \varphi + \hat{y} r \sin \theta \sin \varphi + \hat{z} r \cos \theta . \end{aligned} \quad (1.47)$$

Podemos seleccionar la forma que nos es más útil para nuestro problema particular.

A partir de la sección 1.2 reetiquetando los vectores unitarios en coordenadas curvilíneas  $\hat{e}_1, \hat{e}_2, \hat{e}_3$  como  $\hat{r}, \hat{\theta}, \hat{\varphi}$  tenemos

$$\vec{\nabla} \psi(\rho, \varphi, z) = \hat{r} \frac{\partial \psi}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial \psi}{\partial \theta} + \hat{\varphi} \frac{1}{r \sin \theta} \frac{\partial \psi}{\partial \varphi} , \quad (1.48)$$

$$\vec{\nabla} \cdot \vec{V} = \frac{1}{r^2 \operatorname{sen} \theta} \left[ \operatorname{sen} \theta \frac{\partial}{\partial r} (r^2 V_r) + r \frac{\partial}{\partial \theta} (\operatorname{sen} \theta V_\theta) + r \frac{\partial V_\varphi}{\partial \varphi} \right], \quad (1.49)$$

$$\nabla^2 \psi = \frac{1}{r^2 \operatorname{sen} \theta} \left[ \operatorname{sen} \theta \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \frac{\partial}{\partial \theta} \left( \operatorname{sen} \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\operatorname{sen} \theta} \frac{\partial^2 \psi}{\partial \varphi^2} \right], \quad (1.50)$$

$$\vec{\nabla} \times \vec{V} = \frac{1}{r^2 \operatorname{sen} \theta} \begin{vmatrix} \hat{r} & r\hat{\theta} & r \operatorname{sen} \theta \hat{\varphi} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \varphi} \\ V_r & rV_\theta & r \operatorname{sen} \theta V_\varphi \end{vmatrix}. \quad (1.51)$$

## 1.6 Análisis tensorial.

### Introducción, definiciones.

Los tensores son importantes en muchas áreas de la Física, incluyendo relatividad y electrodinámica. Escalares y vectores son casos especiales de tensores. Como ya vimos, una cantidad que no cambia bajo rotaciones del sistema de coordenadas en un espacio tridimensional, un invariante, fue etiquetado como un escalar. Un escalar es especificado por un número real y es un *tensor de rango cero*. Una cantidad cuyas componentes transforman bajo rotaciones como las de la distancia de un punto a un origen elegido fue llamado un vector. La transformación de las componentes del vector bajo una rotación de las coordenadas preserva el vector como una entidad geométrica (tal como una flecha en el espacio), independiente de la orientación del sistema de referencia. En un espacio tridimensional, un vector es especificado por  $3 = 3^1$  números reales, por ejemplo, sus componentes, y es un *un tensor de rango uno*. En un espacio  $N$  dimensional un tensor de rango  $n$  tiene  $N^n$  componentes los cuales transforman de una manera definida.

Hay una posible ambigüedad en la ley de transformación de un vector

$$A'_i = \sum_j a_{ij} A_j, \quad (1.52)$$

en la cual  $a_{ij}$  es el coseno del ángulo entre el eje  $x'_i$  y el eje  $x_j$ .

Si partimos con un vector diferencial de distancia  $d\vec{r}$ , entonces, tomando  $dx'_i$  como una función de las variables sin primas,

$$dx'_i = \sum_j \frac{\partial x'_i}{\partial x_j} dx_j, \quad (1.53)$$

tomando la diferencial. Si fijamos

$$a_{ij} = \frac{\partial x'_i}{\partial x_j}, \quad (1.54)$$

las ecuaciones (1.52) y (1.53) son consistentes. Cualquier conjunto de cantidades  $A_j$  que transforman de acuerdo a

$$A'_i = \sum_j \frac{\partial x'_i}{\partial x_j} A_j, \quad (1.55)$$

es definido como un vector contravariante.

Sin embargo, hemos ya encontrado un tipo de transformacion vectorial un poco diferente. El gradiente de un escalar,  $\vec{\nabla}\varphi$ , definido por

$$\vec{\nabla}\varphi = \hat{x} \frac{\partial\varphi}{\partial x_1} + \hat{y} \frac{\partial\varphi}{\partial x_2} + \hat{z} \frac{\partial\varphi}{\partial x_3}, \quad (1.56)$$

usando  $(x_1, x_2, x_3)$  para  $(x, y, z)$ , transforma como

$$\frac{\partial\varphi'}{\partial x'_i} = \sum_j \frac{\partial\varphi}{\partial x_j} \frac{\partial x_j}{\partial x'_i}, \quad (1.57)$$

usando  $\varphi = \varphi(x, y, z) = \varphi(x', y', z') = \varphi'$ ,  $\varphi$  definida como una cantidad escalar. Notemos que esta difiere de la ecuación (1.55) en que tenemos  $\partial x_j / \partial x'_i$  en vez de  $\partial x'_i / \partial x_j$ . Tomemos la ecuación (1.57) como la definición de un vector covariante y al gradiente como el prototipo.

En coordenadas cartesianas

$$\frac{\partial x_j}{\partial x'_i} = \frac{\partial x'_i}{\partial x_j} = a_{ij}, \quad (1.58)$$

y no hay diferencia entre transformaciones covariantes y contravariantes. En otros sistemas la ecuación (1.58) en general no se aplica, y la distinción entre covariante y contravariante es real y debe ser observado. Esto es de primera importancia en espacios curvos de Riemann en relatividad general.

En lo que resta de esta sección las componentes de un vector contravariante son denotados por superíndices,  $A^i$ , mientras un subíndice es usado para las componentes de un vector covariante  $A_i$ .<sup>4</sup>

### Definición de tensores de rango dos.

Ahora procedemos a definir *tensores de rango dos contravariantes, mixtos y covariantes* por las siguiente ecuaciones para sus componentes bajo transformación de coordenadas:

$$\begin{aligned} A'^{ij} &= \sum_{kl} \frac{\partial x'_i}{\partial x_k} \frac{\partial x'_j}{\partial x_l} A^{kl}, \\ B'^i_j &= \sum_{kl} \frac{\partial x'_i}{\partial x_k} \frac{\partial x_l}{\partial x'_j} B^k_l, \\ C'_{ij} &= \sum_{kl} \frac{\partial x_k}{\partial x'_i} \frac{\partial x_l}{\partial x'_j} C_{kl}. \end{aligned} \quad (1.59)$$

---

<sup>4</sup>Esto significa que las coordenadas  $(x, y, z)$  deben ser escritas  $(x^1, x^2, x^3)$  ya que  $\vec{r}$  trasforma como un vector contravariante. Ya que nos restringiremos rápidamente a tensores cartesianos (donde la distinción entre contravariante y covariante desaparece) continuaremos usando subíndices para las coordenadas. Esto evita la ambigüedad de  $x^2$  representa el cuadrado de  $x$  o  $y$ .

Claramente, el rango va como el número de derivadas parciales (o cosenos directores) en la definición: cero para un escalar, una para un vector y dos para un tensor de segundo rango, y así sucesivamente. Cada índice (subíndice o superíndice) corre sobre el número de dimensiones del espacio. El número de índices (rango del tensor) es independiente de las dimensiones del espacio. Vemos que  $A^{kl}$  es contravariante con respecto a ambos índices,  $C_{kl}$  es covariante respecto a ambos índices, y  $B_i^k$  transforma contravariantemente con respecto al primero de los índices ( $k$ ) y covariantemente con respecto al segundo índice ( $l$ ). Si usamos coordenadas cartesianas, las tres formas de tensores de segundo rango, contravariante, mixto y covariante son la misma.

Como con las componentes de un vector, las leyes de transformación para las componentes de un tensor, ecuación (1.59), produce entidades (y propiedades) que son independientes de la elección de un sistema de referencia. Esto es lo que hace el análisis tensorial importante en Física. La independencia del sistema de referencia (invariancia) es ideal para expresar afirmaciones en Física.

El tensor de segundo rango  $\mathbf{A}$  (de componentes  $A^{kl}$ ) puede ser convenientemente escrito por sus componentes en un arreglo cuadrado (de  $3 \times 3$  si estamos en un espacio tridimensional),

$$\mathbf{A} = \begin{pmatrix} A^{11} & A^{12} & A^{13} \\ A^{21} & A^{22} & A^{23} \\ A^{31} & A^{32} & A^{33} \end{pmatrix}. \quad (1.60)$$

Esto no significa que cualquier arreglo de números o funciones formen un tensor. La condición esencial es que las componentes transformen de acuerdo a la ecuación (1.59).

### Suma y resta de tensores.

La suma y resta de tensores está definida en términos de los elementos individuales de la misma manera que para vectores. Para sumar o restar dos tensores, debemos sumar o restar sus correspondientes elementos. Si

$$\mathbf{A} + \mathbf{B} = \mathbf{C}, \quad (1.61)$$

entonces

$$A^{ij} + B^{ij} = C^{ij}.$$

Naturalmente,  $\mathbf{A}$  y  $\mathbf{B}$  deben ser tensores del mismo rango y ambos expresados en un espacio del mismo número de dimensiones.

### Convención de suma.

En análisis tensorial es costumbre adoptar una convención de suma para poner la ecuación (1.59) y las subsecuentes ecuaciones tensoriales en forma más compacta. Siempre y cuando distingamos entre covariante y contravariante, acordemos que cuando un índice aparezca a un lado de una ecuación, una vez como superíndice y una vez como subíndice (excepta para las

coordenadas donde ambos son subíndices), automáticamente sumaremos sobre aquel índice. Entonces podemos escribir la segunda expresión en la ecuación (1.59) como

$$B_j^i = \frac{\partial x'_i}{\partial x_k} \frac{\partial x_l}{\partial x'_j} B_l^k, \quad (1.62)$$

con las sumas sobre  $k$  e  $i$  implícitas en el lado derecho. Esta es la convención de suma.

Para ilustrar el uso de la convención de suma y algunas técnicas del análisis tensorial, mostremos que la familiar delta de Kronecker,  $\delta_{kl}$ , es realmente un tensor mixto de rango dos,  $\delta_l^k$ .<sup>5</sup> la pregunta es: ¿La  $\delta_l^k$  transformará de acuerdo a la ecuación (1.59)? Este es nuestro criterio para llamarlo un tensor. Usando la convención de suma tenemos

$$\delta_l^k \frac{\partial x'_i}{\partial x_k} \frac{\partial x_l}{\partial x'_j} = \frac{\partial x'_i}{\partial x_k} \frac{\partial x_k}{\partial x'_j}, \quad (1.63)$$

por definición de la delta de Kronecker. Ahora

$$\frac{\partial x'_i}{\partial x_k} \frac{\partial x_k}{\partial x'_j} = \frac{\partial x'_i}{\partial x'_j}, \quad (1.64)$$

a partir de la regla de la cadena. Sin embargo,  $x'_i$  y  $x'_j$  son coordenadas independientes, y por tanto la variación de una respecto a la otra debe ser cero si ellas son diferentes y uno si ellas coinciden, esto es

$$\frac{\partial x'_i}{\partial x'_j} = \delta_j^i. \quad (1.65)$$

Por tanto

$$\delta_j^i = \frac{\partial x'_i}{\partial x_k} \frac{\partial x_l}{\partial x'_j} \delta_l^k,$$

mostrando que  $\delta_l^k$  son realmente las componentes de un tensor mixto de rango dos. Notemos que el resultado es independiente del número de dimensiones de nuestro espacio.

La delta de Kronecker posee además una interesante propiedad. Tiene las mismas componentes en todos los sistemas de coordenadas rotados y por lo tanto es llamada isotrópica. Tensores de primer rango (vectores) no isotrópicos existen.

### Simetría–Antisimetría.

El orden en el cual los índices aparecen en nuestra descripción de un tensor es importante. En general,  $A^{mn}$  es independiente de  $A^{nm}$ , pero hay algunos casos de interés especial. Si, para todo  $m$  y  $n$ ,

$$A^{mn} = A^{nm}, \quad (1.66)$$

---

<sup>5</sup>Es práctica común referirse a un tensor  $\mathbf{A}$  especificando una componente típica  $A_{ij}$ .

lo llamaremos un tensor simétrico. Si, por otra parte,

$$A^{mn} = -A^{nm} , \quad (1.67)$$

el tensor lo llamaremos antisimétrico. Claramente, cada tensor (de segundo rango) puede ser resuelto en una parte simétrica y otra antisimétrica por la identidad

$$A^{mn} = \frac{1}{2}(A^{mn} + A^{nm}) + \frac{1}{2}(A^{mn} - A^{nm}) , \quad (1.68)$$

el primer término de la derecha es un tensor simétrico y el segundo un tensor antisimétrico. Una similar resolución de las funciones en una parte simétrica y otra antisimétrica es de extrema importancia en mecánica cuántica.

### Spinores.

Podríamos pensar que un sistema de escalares, vectores, tensores ( de rango dos) y superiores forman un sistema matemático completo, uno que es adecuado para describir una Física independiente de nuestra elección de sistema de coordenadas. Pero el universo y la física matemática no es así de simple. En el reino de las partículas elementales, por ejemplo, partículas de spin cero (mesones  $\pi$ , partículas  $\alpha$ ) pueden ser descritos con escalares, partículas de spin 1 (deuterones) por vectores, y partículas de spin 2 (gravitones) por tensores. Esta lista omite a las partículas más comunes: electrones, protones y neutrones, todas ellas con spin  $\frac{1}{2}$ . Estas partículas son propiamente descritas por spinores. Un spinor no es un escalar, vector o tensor.

## 1.7 Contracción y producto directo.

### Contracción.

Cuando tratamos con vectores, formamos un producto escalar sumando el producto de las componentes correspondientes:

$$\vec{A} \cdot \vec{B} = A_i B_i . \quad (1.69)$$

La generalización de esta expresión en el análisis tensorial es un proceso conocido como contracción. Dos índices, uno covariante y el otro contravariante, se igualan uno con otro y luego sumamos sobre ese índice repetido. Por ejemplo, veamos la contracción del tensor mixto de segundo orden  $B_j^i$ .

$$B_j^i \rightarrow B_i^i = \frac{\partial x'_i}{\partial x_k} \frac{\partial x_l}{\partial x'_i} B_l^k = \frac{\partial x_l}{\partial x_k} B_l^k , \quad (1.70)$$

por la ecuación (1.64) y luego por la ecuación (1.65)

$$B_i^i = \delta_k^l B_l^k = B_k^k . \quad (1.71)$$

Nuestro tensor contraído es invariante y por lo tanto un escalar. Esto es exactamente lo que obtuvimos anteriormente para el producto punto de dos vectores y para la divergencia de un vector. En general, la operación de contracción reduce el orden de un tensor en 2.

### Producto Directo.

Las componentes de un vector covariante (tensor de rango uno)  $a_i$  y aquellos de un vector contravariante (tensor de rango uno)  $b^j$  puede ser multiplicado componente a componente para dar el término general  $a_i b^j$ . Esto, por la ecuación (1.59) es realmente un tensor de segundo orden, por

$$a'_i b'^j = \frac{\partial x_k}{\partial x'_i} a_k \frac{\partial x'_j}{\partial x_l} b^l = \frac{\partial x_k}{\partial x'_i} \frac{\partial x'_j}{\partial x_l} (a_k b^l) . \quad (1.72)$$

Contrayendo, obtenemos

$$a'_i b'^i = a_k b^k , \quad (1.73)$$

como en las ecuaciones (1.70) y (1.71) para dar el producto escalar regular.

La operación de asociar dos vectores  $a_i$  y  $b^j$  como en el último párrafo es conocido como el producto directo. Para el caso de dos vectores, el producto directo es un tensor de segundo rango. En ese sentido podemos atribuir significado a  $\vec{\nabla} \vec{E}$ , el cual no estaba definido dentro del esquema del análisis vectorial. En general, el producto directo de dos tensores es un tensor de rango igual a la suma de los dos rangos iniciales; esto es,

$$A_j^i B^{kl} = C_j^{ikl} , \quad (1.74)$$

donde  $C_j^{ikl}$  es un tensor de cuarto rango. A partir de la ecuación (1.59)

$$C_j^{ikl} = \frac{\partial x'_i}{\partial x_m} \frac{\partial x_n}{\partial x'_j} \frac{\partial x'_k}{\partial x_p} \frac{\partial x'_l}{\partial x_q} C_n^{mpq} . \quad (1.75)$$

El producto directo aparece en física matemática como una técnica para crear nuevos tensores de mayor rango.

Cuando  $\mathbf{T}$  es un tensor cartesiano de  $n$ -ésimo rango,  $\partial/\partial x_i T_{jkl} \dots$ , un elemento de  $\vec{\nabla} \mathbf{T}$ , es un tensor *cartesiano* de rango  $n + 1$ . Sin embargo,  $\partial/\partial x_i T_{jkl} \dots$  no es un tensor bajo transformaciones más generales. En sistemas no cartesianos  $\partial/\partial x'_i$  actuará sobre las derivadas parciales  $\partial x_p \partial x'_q$  y destruirá la relación simple de transformación tensorial.

Hasta aquí la distinción entre una transformación covariante y una contravariante ha sido mantenida ya que existe en espacio no cartesiano y porque es de gran importancia en relatividad general. Ahora, sin embargo, nos restringimos al tensor cartesiano. Como se hizo notar, en el caso cartesiano, la distinción entre contravariancia y covariancia desaparece y todos los índices están a partir de ahora en adelante mostrado como subíndice.

### Convención de la suma.

Cuando un subíndice (letra, no número) aparece dos veces sobre un lado de una ecuación, implica que se suma con respecto a este subíndice.

### Contracción.

La contracción consiste en fijar dos índices distintos (subíndices) igualar uno a otro y luego sumarlos según la convención de suma.

## 1.8 Regla del cuociente.

Si  $A_i$  y  $B_j$  son vectores, podemos fácilmente mostrar que  $A_i B_j$  es un tensor de segundo rango. Aquí, estamos interesados en una variedad de relaciones inversas. Consideremos tales ecuaciones como

$$K_i A_i = B \quad , \quad (1.76a)$$

$$K_{ij} A_j = B_i \quad , \quad (1.76b)$$

$$K_{ij} A_{jk} = B_{ik} \quad , \quad (1.76c)$$

$$K_{ijkl} A_{ij} = B_{kl} \quad , \quad (1.76d)$$

$$K_{ij} A_k = B_{ijk} \quad . \quad (1.76e)$$

En cada una de esas expresiones  $\mathbf{A}$  y  $\mathbf{B}$  son tensores conocidos cuyo rango está indicado por el número de índices y el tensor  $\mathbf{A}$  es arbitrario. En cada caso  $K$  es una cantidad desconocida. Deseamos establecer las propiedades de transformación de  $K$ . La regla del cuociente afirma que si la ecuación de interés se mantiene en todos los sistemas (rotados) de coordenadas cartesianas,  $K$  es un tensor del rango indicado. La importancia en física teórica es que la regla del cuociente puede establecer la naturaleza tensorial de las cantidades. La regla del cuociente (ecuación 1.76b) muestra que la matriz de inercia que aparece en la ecuación de momento angular  $\vec{L} = I\vec{\omega}$ , es un tensor.

Para probar la regla del cuociente, consideremos la ecuación (1.76b) como un caso típico. En nuestro sistema de coordenadas prima

$$K'_{ij} A'_j = B'_i = a_{ik} B_k \quad , \quad (1.77)$$

usando las propiedades de transformación vectorial de  $\vec{B}$ . Ya que la ecuación se mantiene en todos los sistemas de coordenadas rotados,

$$a_{ik} B_k = a_{ik} (K_{kl} A_l) \quad . \quad (1.78)$$

Ahora, transformando  $\vec{A}$  regresamos al sistema de coordenadas prima (compare la ecuación (1.55)), tenemos

$$K'_{ij} A'_j = a_{ik} K_{kl} a_{jl} A'_j \quad . \quad (1.79)$$

Reordenando, tenemos

$$(K'_{ij} - a_{ik} a_{jl} K_{kl}) A'_j = 0 \quad . \quad (1.80)$$

Esto debe mantenerse para cada valor del índice  $i$  y para cada sistema de coordenadas prima. Ya que  $A'_j$  es arbitrario, concluimos

$$K'_{ij} = a_{ik} a_{jl} K_{kl} \quad , \quad (1.81)$$

la cual es nuestra definición de un tensor de segundo rango.

Las otras ecuaciones pueden ser tratadas en forma similar, dando origen a otras formas de la regla del cuociente. Un peligro menor debería ser tomado en cuenta, la regla del cuociente no se aplica necesariamente si  $\vec{B}$  es igual a cero. Las propiedades de transformación de cero son indeterminadas.

## 1.9 Pseudotensores y tensores duales.

Hasta aquí nuestras transformaciones de coordenadas han sido restringidas a rotaciones puras. Ahora consideramos el efecto de reflexiones o inversiones. Si tenemos coeficientes de transformaciones  $a_{ij} = -\delta_{ij}$ , entonces por la ecuación (1.53)

$$x_i = -x'_i, \quad (1.82)$$

la cual es una inversión o transformación de paridad. Notemos que esta transformación cambia nuestro sistema de coordenadas inicialmente diestro a un sistema de coordenadas siniestro.<sup>6</sup> Nuestro vector prototipo  $\vec{r}$  con componentes  $(x_1, x_2, x_3)$  se transforma al vector  $\vec{r}' = (x'_1, x'_2, x'_3) = (-x_1, -x_2, -x_3)$ . Este nuevo vector  $\vec{r}'$  tiene componentes negativas, relativas al nuevo conjunto de ejes transformados. Como se muestra en la figura 1.7, invirtiendo las direcciones de los ejes de coordenadas y cambiando los signos de las componentes da  $\vec{r}' = \vec{r}$ . El vector (una flecha en el espacio) permanece exactamente como estaba antes de que la transformación se llevara a cabo. El vector posición  $\vec{r}$  y todos los otros vectores cuyas componentes se comporten de esta manera (cambiando el signo con una inversión de los ejes de coordenadas) son llamados vectores polares y tienen paridad impar.

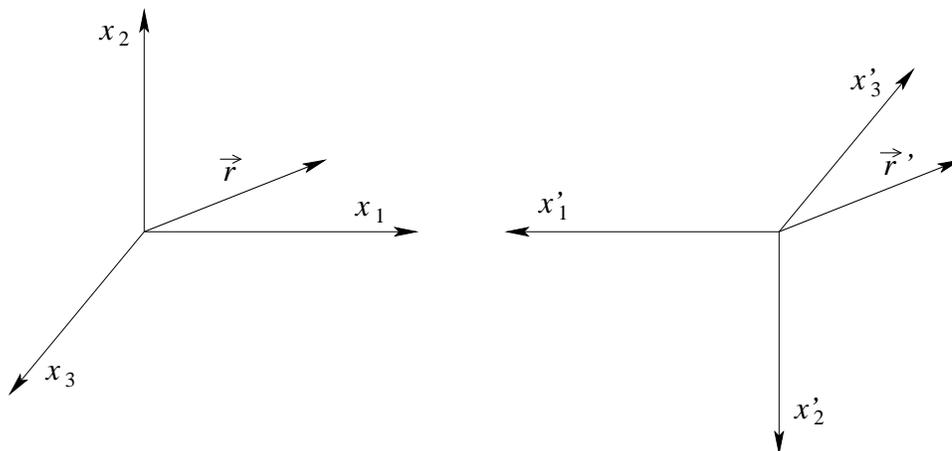


Figura 1.7: Inversión de coordenadas cartesianas, vector polar.

Una diferencia fundamental aparece cuando encontramos un vector definido como el producto cruz de dos vectores polares. Sea  $\vec{C} = \vec{A} \times \vec{B}$ , donde tanto  $\vec{A}$  como  $\vec{B}$  son vectores polares. Las componentes de  $\vec{C}$  están dadas por

$$C_1 = A_2B_3 - A_3B_2. \quad (1.83)$$

y así sucesivamente. Ahora cuando los ejes de coordenadas son invertidos,  $A_i \rightarrow A'_i$ ,  $B_j \rightarrow B'_j$  pero de su definición  $C_k \rightarrow +C'_k$ ; esto es, nuestro vector producto cruz, vector  $\vec{C}$ , no se comporta como un vector polar bajo inversión. Para distinguir, lo etiquetamos como pseudo vector o vector axial (ver figura 1.8) que tiene paridad par.

<sup>6</sup>Esta es una inversión del sistema de coordenadas, los objetos en el mundo físico permanecen fijos.

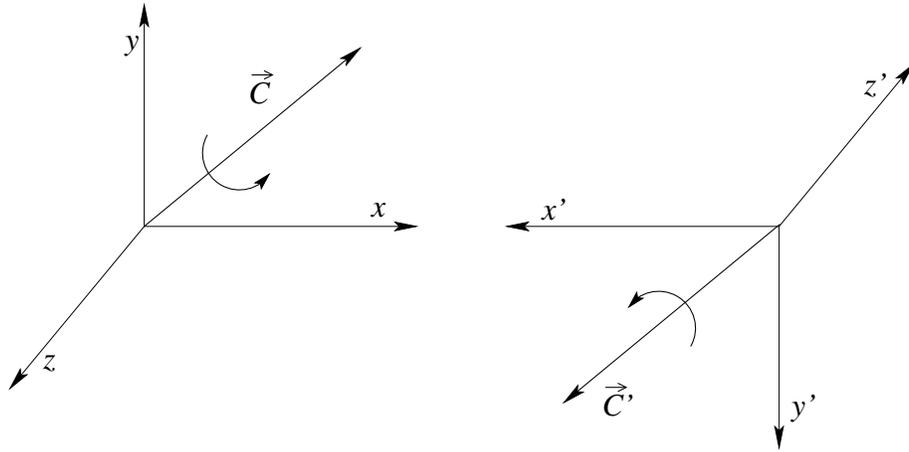


Figura 1.8: Inversión de coordenadas cartesianas, vector axial.

Ejemplos son:

$$\begin{aligned}
 \text{velocidad angular,} & \quad \vec{v} = \vec{\omega} \times \vec{r} , \\
 \text{momento angular,} & \quad \vec{L} = \vec{r} \times \vec{p} , \\
 \text{torque,} & \quad \vec{\tau} = \vec{r} \times \vec{F} , \\
 \text{campo magnético,} & \quad \frac{\partial \vec{B}}{\partial t} = -c \vec{\nabla} \times \vec{E} .
 \end{aligned}$$

En  $\vec{v} = \vec{\omega} \times \vec{r}$ , el vector axial es la velocidad angular  $\vec{\omega}$ , y  $\vec{r}$  y  $\vec{v} = d\vec{r}/dt$  son vectores polares. Claramente, los vectores axiales ocurren frecuentemente en Física elemental, aunque este hecho usualmente no es recalado. En un sistema de coordenadas de mano derecha un vector axial  $\vec{C}$  tiene un sentido de rotación asociado con el lado por la regla de la mano derecha. En el sistema invertido de mano izquierda el sentido de la rotación es una rotación de la mano izquierda. Esto es indicado por las flechas curvadas en la figura 1.8.

La distinción entre vectores polares y axiales también pueden ser ilustradas por una reflexión. Un vector polar se refleja en un espejo como un flecha física real, figura 1.9a. En la figuras 1.7 y 1.8 las coordenadas están invertidas; el mundo físico permanece fijo. Aquí los ejes de coordenadas permanecen fijos; el mundo es reflejado como en un espejo en el plano  $xz$ . Específicamente, en esa representación mantenemos los ejes fijos y asociamos un cambio de signo con la componente del vector. Para un espejo en el plano  $xz$ ,  $P_y \rightarrow -P_y$ . Tenemos

$$\begin{aligned}
 P &= (P_x, P_y, P_z) , \\
 P' &= (P_x, -P_y, P_z) , \quad \text{vector polar.}
 \end{aligned}$$

Un vector axial tal como un campo magnético  $\vec{B}$  o el momento magnético  $\mu$  (= corriente  $\times$  área de loop ) se comporta muy diferentemente bajo reflexión.

Consideremos el campo magnético  $\vec{B}$  y el momento magnético  $\mu$  como producidos de una carga eléctrica moviendose circularmente. La reflexión revierte el sentido de la rotación de la

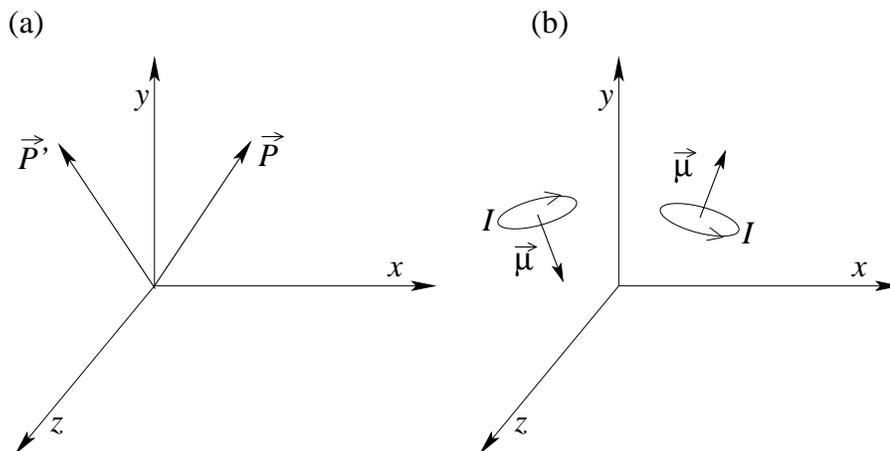


Figura 1.9: (a) Espejo en el plano  $xz$ ; (b) Espejo en el plano  $xz$ .

carga. Los dos loops de corriente y el resultante de los momentos magnéticos son mostrados en la figura 1.9b. Tenemos

$$\begin{aligned}\mu &= (\mu_x, \mu_y, \mu_z) , \\ \mu' &= (-\mu_x, \mu_y, -\mu_z) \quad \text{vector axial.}\end{aligned}$$

Si concordamos que el universo no se preocupa si usamos sistemas de coordenadas ya sea mano derecha o mano izquierda, luego no tiene sentido sumar un vector axial a un vector polar. En la ecuación vectorial  $\vec{A} = \vec{B}$ , ambos  $\vec{A}$  y  $\vec{B}$  son tanto vectores polares como vectores axiales<sup>7</sup>. Restricciones similares se aplican a los escalares y pseudoescalares y, en general, a los tensores y pseudotensores.

Usualmente, los pseudoescalares, pseudovectores, y pseudotensores transforman como

$$\begin{aligned}S' &= |a| S \\ C'_i &= |a| a_{ij} C_j , \\ A'_{ij} &= |a| a_{ik} a_{jl} A_{kl} ,\end{aligned}\tag{1.84}$$

donde  $|a|$  es el determinante de un arreglo de coeficientes  $a_{mn}$ . En nuestra inversión el determinante es

$$|a| = \begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -1\tag{1.85}$$

Para la reflexión de un eje, el eje  $x$ ,

$$|a| = \begin{vmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -1\tag{1.86}$$

<sup>7</sup>La gran excepción a esto está en el decaimiento beta, interacciones débiles. Aquí el universo distingue entre sistemas derechos y sistemas izquierdos.

y nuevamente el determinante es  $|a| = -1$ . Por otra parte, para todas las rotaciones puras el determinante  $|a|$  es siempre  $+1$ . A menudo las cantidades que transforman de acuerdo a la ecuación (1.84) son conocidas como densidad tensorial. Ellos son tensores regulares en cuanto a rotaciones se refiere, diferenciándose de los tensores solamente en reflexiones o inversiones de las coordenadas, y luego la única diferencia es la aparición de un signo menos adicional del determinante  $|a|$ .

Anteriormente mostramos que el producto escalar triple  $S = \vec{A} \times \vec{B} \cdot \vec{C}$  es un escalar (bajo rotaciones). Ahora considerando la transformación de paridad dada por la ecuación (1.82), vemos que  $S \rightarrow -S$ , probando que el producto escalar triple es un pseudoescalar. Este comportamiento fue presagiado por la analogía geométrica de un volumen. Si los tres parámetros del volumen, longitud, ancho, profundidad, cambian de distancias positivas a distancias negativas, el producto de los tres será negativo.

### Símbolo de Levi-Civita.

Para usos futuros es conveniente introducir el símbolo tridimensional de Levi-Civita  $\varepsilon_{ijk}$  definido por

$$\begin{aligned} \varepsilon_{123} &= \varepsilon_{231} = \varepsilon_{312} = 1 , \\ \varepsilon_{132} &= \varepsilon_{213} = \varepsilon_{321} = -1 , \\ \text{todos los otros } \varepsilon_{ijk} &= 0 . \end{aligned} \tag{1.87}$$

Note que  $\varepsilon_{ijk}$  es totalmente antisimétrico con respecto a todo par de índices. Suponga que tenemos un pseudovector de tercer grado  $\delta_{ijk}$ , el cual en un sistema de coordenadas particular es igual a  $\varepsilon_{ijk}$ . Luego

$$\delta'_{ijk} = |a| a_{ip} a_{jq} a_{kr} \varepsilon_{pqr} , \tag{1.88}$$

por definición de pseudotensor. Ahora

$$a_{1p} a_{2q} a_{3r} \varepsilon_{pqr} = |a| , \tag{1.89}$$

por directa expansión del determinante, mostrando que  $\delta'_{123} = |a|^2 = 1 = \varepsilon_{123}$ . Considerando las otras posibilidades una a una, encontramos

$$\delta'_{ijk} = \varepsilon_{ijk} , \tag{1.90}$$

para rotaciones y reflexiones. Luego  $\varepsilon_{ijk}$  es un pseudotensor. Además, se ve como un pseudotensor isotrópico con las mismas componentes en todos los sistemas de coordenadas cartesianos rotados.

### Tensores duales.

A cualquier tensor antisimétrico de segundo rango  $C_{jk}$  (en el espacio tridimensional) podemos asociarle un pseudo vector dual  $C_i$  definido por

$$C_i = \frac{1}{2} \varepsilon_{ijk} C_{jk} . \tag{1.91}$$

Aquí el antisimétrico  $C_{jk}$  puede ser escrito

$$C_{jk} = \begin{pmatrix} 0 & C_{12} & -C_{31} \\ -C_{12} & 0 & C_{23} \\ C_{31} & -C_{23} & 0 \end{pmatrix}. \quad (1.92)$$

Sabemos que  $C_i$  debe transformar como un vector bajo rotaciones a partir de la doble contracción del (pseudo) tensor de quinto rango  $\varepsilon_{ijk}C_{mn}$  pero éste es realmente un pseudovector desde la naturaleza pseudo de  $\varepsilon_{ijk}$ . Específicamente, los componentes de  $\vec{C}$  están dados por

$$(C_1, C_2, C_3) = (C_{23}, C_{31}, C_{12}). \quad (1.93)$$

Note que el orden cíclico de los índices que vienen del orden cíclico de los componentes de  $\varepsilon_{ijk}$ . Esta dualidad, dada por la ecuación (1.93), significa que nuestro producto vectorial tridimensional puede, literalmente, ser tomado ya sea como un pseudovector o como un tensor antisimétrico de segundo rango, dependiendo de como escogamos escribirlo.

Si tomamos tres vectores (polares)  $\vec{A}$ ,  $\vec{B}$ , y  $\vec{C}$ , podemos definir

$$V_{ijk} = \begin{vmatrix} A_i & B_i & C_i \\ A_j & B_j & C_j \\ A_k & B_k & C_k \end{vmatrix} = A_i B_j C_k - A_i B_k C_j + \dots \quad (1.94)$$

Por una extensión del análisis de la sección 1.6 cada término  $A_p B_q C_r$  es visto como un tensor de tercer rango, haciendo  $V_{ijk}$  un tensor de rango tres. A partir de su definición como un determinante  $V_{ijk}$  es totalmente antisimétrico, cambiando signo bajo el intercambio de cualquier par de índices, esto es, el intercambio de cualquier par de filas del determinante. La cantidad dual es

$$V = \frac{1}{3!} \varepsilon_{ijk} V_{ijk}, \quad (1.95)$$

claramente un pseudo escalar. Por expansión se ve que

$$V = \begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix}, \quad (1.96)$$

nuestro conocido producto escalar triple.

Para poder escribir las ecuaciones Maxwell en forma covariante, necesitamos extender este análisis vectorial dual a un espacio de cuatro dimensiones y, en particular, indicar que el elemento de volumen cuadrídimensional,  $dx_1, dx_2, dx_3, dx_4$ , es un pseudoescalar.

Introducimos el símbolo Levi-Civita  $\varepsilon_{ijkl}$ , el análogo cuadrídimensional de  $\varepsilon_{ijk}$ . Esta cantidad  $\varepsilon_{ijkl}$  es definida totalmente antisimétrica respecto a sus cuatro índices. Si  $(ijkl)$  es una permutación par de  $(1,2,3,4)$ , luego  $\varepsilon_{ijkl}$  es definido como  $+1$ ; si es una permutación impar, luego  $\varepsilon_{ijkl}$  es  $-1$ . El  $\varepsilon_{ijkl}$  de Levi-Civita puede probarse que es un pseudotensor de rango cuatro por análisis similar al usado para establecer la naturaleza de  $\varepsilon_{ijk}$ . Introduciendo un tensor de rango cuatro,

$$H_{ijkl} = \begin{vmatrix} A_i & B_i & C_i & D_i \\ A_j & B_j & C_j & D_j \\ A_k & B_k & C_k & D_k \\ A_l & B_l & C_l & D_l \end{vmatrix}, \quad (1.97)$$

construyendo a partir de los vectores polares  $\vec{A}$ ,  $\vec{B}$ ,  $\vec{C}$  y  $\vec{D}$  podemos definir la cantidad dual

$$H = \frac{1}{4!} \varepsilon_{ijkl} H_{ijkl} . \quad (1.98)$$

Realmente tenemos una cuádruple contracción la cual reduce el orden a cero. A partir de la naturaleza pseudo de  $\varepsilon_{ijkl}$ ,  $\vec{H}$  es un pseudo escalar. Ahora tomemos  $\vec{A}$ ,  $\vec{B}$ ,  $\vec{C}$  y  $\vec{D}$  como desplazamientos infinitesimales a lo largo de cuatro ejes de coordenadas (espacio de Minkowski),

$$\vec{A} = (dx_1, 0, 0, 0) , \quad \vec{B} = (0, dx_2, 0, 0) , \quad \vec{C} = (0, 0, dx_3, 0) , \quad \vec{D} = (0, 0, 0, dx_4) , \quad (1.99)$$

y

$$H = dx_1 dx_2 dx_3 dx_4 . \quad (1.100)$$

El elemento de volumen cuadrimensional está ahora identificado como un pseudo escalar. Este resultado era de esperarse a partir de los resultados de la teoría especial de relatividad. La contracción de Lorentz-Fitzgerald de  $dx_1$ ,  $dx_2$ ,  $dx_3$  justo se compensa con la dilatación temporal de  $dx_4$ .

Nos introducimos a este espacio cuadrimensional como una simple extensión matemática del espacio tridimensional y, por supuesto, podríamos haber discutido fácilmente espacios de 5, 6 o  $N$  dimensiones. Esto es típico del poder del análisis de componentes. Físicamente, este espacio tetradimensional puede ser tomado como un espacio de Minkowski,

$$(x_1, x_2, x_3, x_4) = (x, y, z, ict) , \quad (1.101)$$

donde  $t$  es el tiempo. Esto es la mezcla de espacio y tiempo realizada en la relatividad especial. Las transformaciones que describen las rotaciones en un espacio cuadrimensional son las transformaciones de Lorentz de la relatividad especial.

### Tensores irreducibles.

Para algunas aplicaciones, particularmente en la teoría cuántica del momento angular, nuestros tensores cartesianos no son particularmente convenientes. En el lenguaje matemático nuestro tensor general de segundo orden  $A_{ij}$  es reducible, lo que significa que puede ser descompuesto en partes de un tensor de menor orden. En efecto, ya hemos hecho esto. De la ecuación (1.71)

$$A = A_{ii} , \quad (1.102)$$

es una cantidad escalar, la traza de  $A_{ij}$ .

La parte antisimétrica

$$B_{ij} = \frac{1}{2} (A_{ij} - A_{ji}) \quad (1.103)$$

han sido mostrada como equivalente a un (pseudo) vector, o

$$B_{ij} = C_k , \quad \text{permutación cíclica de } i, j, k. \quad (1.104)$$

Sustrayendo el escalar  $A$  y el vector  $C_k$  de nuestro tensor original, tenemos un tensor de segundo orden irreducible, simétrico y de traza nula,  $S_{ij}$ , en el cual

$$S_{ij} = \frac{1}{2}(A_{ij} + A_{ji}) - \frac{1}{3}A \delta_{ij} , \quad (1.105)$$

con cinco componentes independientes. Luego, finalmente, nuestro tensor cartesiano original puede ser escrito

$$A_{ij} = \frac{1}{3}A \delta_{ij} + C_k + S_{ij} . \quad (1.106)$$

Las tres cantidades  $A$ ,  $C_k$ , y  $S_{ij}$  forman el tensor esférico de orden 0, 1 y 2 respectivamente, transformando como los armónicos esféricos  $Y_L^M$  para  $L=0, 1$  y  $2$ .

Un ejemplo específico de la reducción anterior está dada por el tensor cuadrupolo eléctrico simétrico

$$Q_{ij} = \int (3x_i x_j - r^2 \delta_{ij}) \rho(x_1, x_2, x_3) d^3x .$$

El término  $-r^2 \delta_{ij}$  representa una resta de la traza escalar (los 3 terminos  $i = j$ ). El resultante  $Q_{ij}$  tiene traza cero.

## 1.10 Tensores no cartesianos, diferenciación covariante.

La distinción entre transformaciones contravariante y transformaciones covariante fueron establecidas en la sección 1.6. Luego, por conveniencia, nos restringimos a coordenadas cartesianas (en la cual la distinción desaparece). Ahora en estas dos secciones volvemos a las coordenadas no cartesianas y resurge la distinción entre contravariante y covariante. Como en la sección 1.6, un superíndice será usado para denotar la dependencia contravariante y un subíndice para diferenciar la covariante. El tensor métrico de la sección 1.1 será usado para relacionar los índices contravariante y covariante.

El énfasis en esta sección es sobre diferenciación, culminando en la construcción de una derivada covariante. Vimos en la sección 1.7 que la derivada de un campo vectorial es un tensor de segundo orden en coordenadas cartesianas. La derivada covariante de un campo vectorial es un tensor de segundo orden en sistemas de coordenadas no cartesianas.

### Tensor métrico, subida y bajada de índices.

Empecemos con un conjunto de vectores base  $\vec{\epsilon}_i$  tal que un desplazamiento infinitesimal  $d\vec{r}$  podría estar dado por

$$d\vec{r} = \vec{\epsilon}_1 dq^1 + \vec{\epsilon}_2 dq^2 + \vec{\epsilon}_3 dq^3 . \quad (1.107)$$

Por conveniencia tomamos  $\vec{\epsilon}_1, \vec{\epsilon}_2, \vec{\epsilon}_3$  formando un sistema diestro. Estos tres vectores no son necesariamente ortogonales. También, se requerirá una limitación al espacio tridimensional solamente para la discusión de los productos cruz y rotores. De lo contrario estos  $\vec{\epsilon}_i$  pueden

estar en un espacio de  $N$ -dimensiones, incluyendo la cuarta dimensión espacio-tiempo de la relatividad especial y general. Los vectores base  $\vec{\varepsilon}_i$  pueden ser expresados por

$$\vec{\varepsilon}_i = \frac{\partial \vec{r}}{\partial q^i}, \quad (1.108)$$

Note, sin embargo, que los  $\vec{\varepsilon}_i$  no tienen necesariamente magnitud uno. Se puede probar que los vectores unitarios son

$$\hat{e}_i = \frac{1}{h_i} \frac{\partial \vec{r}}{\partial q^i}, \quad (\text{no hay suma}),$$

y por lo tanto

$$\vec{\varepsilon}_i = h_i \hat{e}_i, \quad (\text{no hay suma}). \quad (1.109)$$

Los  $\vec{\varepsilon}_i$  están relacionados a los vectores unitarios  $\hat{e}_i$  por el factor de escala  $h_i$  de la sección 1.2. Los  $\hat{e}_i$  no tienen dimensiones, los  $\vec{\varepsilon}_i$  tienen dimensiones de  $h_i$ . En coordenadas polares esféricas, como un ejemplo específico,

$$\begin{aligned} \vec{\varepsilon}_r &= \hat{e}_r, \\ \vec{\varepsilon}_\theta &= r \hat{e}_\theta, \\ \vec{\varepsilon}_\phi &= r \sin \theta \hat{e}_\phi. \end{aligned} \quad (1.110)$$

Como en la sección 1.1, construimos el cuadrado de un desplazamiento diferencial

$$(ds)^2 = d\vec{s} \cdot d\vec{s} = (\vec{\varepsilon}_i dq^i)^2 = \vec{\varepsilon}_i \cdot \vec{\varepsilon}_j dq^i dq^j. \quad (1.111)$$

Comparando esto con  $(ds)^2$  de la sección 1.1, ecuación (1.6), definimos  $\vec{\varepsilon}_i \cdot \vec{\varepsilon}_j$  como el tensor métrico covariante

$$\vec{\varepsilon}_i \cdot \vec{\varepsilon}_j = g_{ij}. \quad (1.112)$$

Claramente  $g_{ij}$  es simétrico. La naturaleza del tensor  $g_{ij}$  se deduce a partir de la regla del cociente. Tomemos la relación

$$g^{ik} g_{kj} = \delta_j^i, \quad (1.113)$$

para definir el correspondiente tensor contravariante  $g^{ik}$ . El contravariante  $g^{ik}$  entra como el inverso<sup>8</sup> del covariante  $g_{kj}$ . Usemos este contravariante  $g^{ik}$  para subir índices, convirtiendo un índice covariante en uno contravariante, como se muestra más adelante. Así mismo el covariante  $g_{kj}$  será usado para bajar índices. La elección de  $g^{ik}$  y  $g_{kj}$  para esta operación de subir y bajar es arbitraria. Cualquier tensor de segundo orden (y su inverso) podría hacerlo. Específicamente, tenemos

$$\begin{aligned} g^{ij} \vec{\varepsilon}_j &= \vec{\varepsilon}^i, & \text{relación de covariante y} \\ & & \text{contravariante de vectores bases} \\ g^{ij} \vec{F}_j &= \vec{F}^i, & \text{relación de covariante y} \\ & & \text{contravariante de componentes vectoriales.} \end{aligned} \quad (1.114)$$

---

<sup>8</sup>Si el tensor  $g_{ij}$  es escrito como una matriz, el tensor  $g^{ij}$  está dado por la matriz inversa.

Entonces

$$\begin{aligned} g_{ij}\vec{\varepsilon}^j &= \vec{\varepsilon}_i, \quad \text{son las correspondientes relaciones} \\ g_{ij}\vec{F}^j &= \vec{F}_i, \quad \text{de bajada de índices.} \end{aligned} \quad (1.115)$$

Como un ejemplo de esas transformaciones comenzaremos con la forma contravariante del vector

$$\vec{F} = F^i \vec{\varepsilon}_i. \quad (1.116)$$

De las ecuaciones (1.114) y (1.115)

$$\vec{F} = F_j g^{ji} g_{ik} \vec{\varepsilon}^k = F_j \vec{\varepsilon}^j, \quad (1.117)$$

la igualdad final viene de la ecuaciones (1.113). La ecuación (1.116) da la representación contravariante de  $\vec{F}$ . La ecuación (1.117) da la correspondiente representación covariante del mismo  $\vec{F}$ .

Deberíamos enfatizar de nuevo que  $\vec{\varepsilon}_i$  y  $\vec{\varepsilon}^j$  no tienen unidades. Esto puede verse en la ecuaciones (1.110) y en el tensor métrico  $g_{ij}$  para coordenadas polares esféricas y su inverso  $g^{ij}$ :

$$(g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \quad (g^{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2 \theta} \end{pmatrix}.$$

### Derivadas y símbolos de Christoffel.

Formemos la diferencial de un escalar

$$d\psi = \frac{\partial \psi}{\partial q^i} dq^i. \quad (1.118)$$

Ya que la  $dq^i$  son las componentes de un vector contravariante, las derivadas parciales,  $\partial\psi/\partial q^i$  debieran ser un vector covariante, por la regla del cociente. El gradiente de un escalar llega a ser

$$\vec{\nabla}\psi = \frac{\partial \psi}{\partial q^i} \vec{\varepsilon}^i. \quad (1.119)$$

Deberíamos notar que  $\partial\psi/\partial q^i$  no son las componentes del gradiente de la sección 1.2 ya que  $\vec{\varepsilon}^i \neq \hat{e}_i$  de la sección 1.2.

Continuando con las derivadas de un vector, encontramos que la situación es mucho más complicada por que los vectores bases  $\vec{\varepsilon}_i$  en general no son constantes. Recordemos que no estamos más restringidos a las coordenadas cartesianas con sus conveniente  $\hat{x}$ ,  $\hat{y}$ ,  $\hat{z}$ . La diferenciación directa es

$$\frac{\partial \vec{V}}{\partial q^j} = \frac{\partial V^i}{\partial q^j} \vec{\varepsilon}_i + V^i \frac{\partial \vec{\varepsilon}_i}{\partial q^j}. \quad (1.120)$$

Ahora  $\partial\vec{\varepsilon}_i/\partial q^j$  será alguna combinación lineal de  $\vec{\varepsilon}_k$  con coeficientes que dependen de los índices  $i$  y  $j$  de las derivadas parciales y el índice  $k$  del vector base. Escribimos

$$\frac{\partial\vec{\varepsilon}_i}{\partial q^j} = \Gamma_{ij}^k \vec{\varepsilon}_k . \quad (1.121)$$

Multiplicando por  $\vec{\varepsilon}^m$  y usando  $\vec{\varepsilon}^m \cdot \vec{\varepsilon}_k = \delta_k^m$  (pruébelo).

$$\Gamma_{ij}^m = \vec{\varepsilon}^m \cdot \frac{\partial\vec{\varepsilon}_i}{\partial q^j} . \quad (1.122)$$

El  $\Gamma_{ij}^k$  es un símbolo de Christoffel (del segundo tipo). También se le conoce como “coeficiente de conexión”. Estos  $\Gamma_{ij}^k$  no son tensores de rango tres y el  $\partial V^i/\partial q^j$  de la ecuación (1.120) no son tensores de segundo rango. En coordenadas cartesianas,  $\Gamma_{ij}^k = 0$  para todos los valores de los índices  $i, j, k$ .

Usando la ecuación (1.108), obtenemos

$$\frac{\partial\vec{\varepsilon}_i}{\partial q^j} = \frac{\partial^2\vec{r}}{\partial q^j \partial q^i} = \frac{\partial\vec{\varepsilon}_j}{\partial q^i} = \Gamma_{ji}^k \vec{\varepsilon}_k . \quad (1.123)$$

Luego estos símbolos de Christoffel son simétricos en los dos índices inferiores:

$$\Gamma_{ij}^k = \Gamma_{ji}^k . \quad (1.124)$$

### Derivada covariante.

Con los símbolos de Christoffel, la ecuación (1.120) puede ser reescrita

$$\frac{\partial\vec{V}}{\partial q^j} = \frac{\partial V^i}{\partial q^j} \vec{\varepsilon}_i + V^i \Gamma_{ij}^k \vec{\varepsilon}_k . \quad (1.125)$$

Ahora  $i$  y  $k$  en el último término son índices mudos. Intercambiando  $i$  y  $k$ , tenemos

$$\frac{\partial\vec{V}}{\partial q^j} = \left( \frac{\partial V^i}{\partial q^j} + V^k \Gamma_{kj}^i \right) \vec{\varepsilon}_i . \quad (1.126)$$

La cantidad entre paréntesis es denominada **derivada covariante**,  $V_{;j}^i$ . Tenemos

$$V_{;j}^i \equiv \frac{\partial V^i}{\partial q^j} + V^k \Gamma_{kj}^i . \quad (1.127)$$

Los subíndices  $_{;j}$  indican que la diferenciación es con respecto a  $q^j$ . La diferencial  $d\vec{V}$  se convierte en

$$d\vec{V} = \frac{\partial\vec{V}}{\partial q^j} dq^j = [V_{;j}^i dq^j] \vec{\varepsilon}_i . \quad (1.128)$$

Una comparación con las ecuaciones (1.107) o (1.116) muestran que la cantidad entre paréntesis cuadrados es la  $i$ -ésima componente de un vector contravariante. Ya que  $dq^j$  es la  $j$ -ésima

componente de un vector contravariante,  $V^i_{;j}$  debe ser la componente  $ij$ -ésima de un tensor mixto de segundo rango (regla del cociente). Las derivadas covariantes de las componentes contravariantes de un vector forma un tensor mixto de segundo rango,  $V^i_{;j}$ .

Ya que los símbolos de Christoffel desaparecen en coordenadas cartesianas, la derivada covariante y la derivada parcial ordinaria coinciden

$$\frac{\partial V^i}{\partial q^j} = V^i_{;j}, \quad (\text{En coordenadas cartesianas}). \quad (1.129)$$

Se puede demostrar que la derivada covariante de un vector covariante  $V_i$  está dada por

$$V_{i;j} = \frac{\partial V_i}{\partial q^j} - V_k \Gamma^k_{ij}. \quad (1.130)$$

Como  $V^i_{;j}$ , el tensor  $V_{i;j}$  es de rango dos.

La importancia física de la derivada covariante es:

*Un reemplazo consistente de las derivadas parciales regulares por derivadas covariantes lleva las leyes de la Física (por componentes) desde un espacio tiempo plano al espacio tiempo curvo (Riemannian) de la relatividad general. Realmente, esta sustitución puede ser tomada como una declaración matemática del principio de equivalencia.<sup>9</sup>*

### Los símbolos de Christoffel como derivadas del tensor métrico.

A menudo es conveniente tener una expresión explícita para los símbolos de Christoffel en términos de derivadas del tensor métrico. Como un paso inicial, definimos el símbolo de Christoffel del primer tipo  $[ij, k]$  por

$$[ij, k] \equiv g_{mk} \Gamma^m_{ij}. \quad (1.131)$$

Este  $[ij, k]$  no es un tensor de rango tres. De la ecuación (1.122)

$$\begin{aligned} [ij, k] &= g_{mk} \vec{\varepsilon}^m \cdot \frac{\partial \vec{\varepsilon}_i}{\partial q^j}, \\ &= \vec{\varepsilon}_k \cdot \frac{\partial \vec{\varepsilon}_i}{\partial q^j}. \end{aligned} \quad (1.132)$$

Ahora diferenciamos  $g_{ij} = \vec{\varepsilon}_i \cdot \vec{\varepsilon}_j$ , ecuación (1.112):

$$\begin{aligned} \frac{\partial g_{ij}}{\partial q^k} &= \frac{\partial \vec{\varepsilon}_i}{\partial q^k} \cdot \vec{\varepsilon}_j + \vec{\varepsilon}_i \cdot \frac{\partial \vec{\varepsilon}_j}{\partial q^k}, \\ &= [ik, j] + [jk, i], \end{aligned} \quad (1.133)$$

por la ecuación (1.132).

<sup>9</sup>C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation*. San Francisco: W. H. Freeman (1973), p 387.

Luego

$$[ij, k] = \frac{1}{2} \left\{ \frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k} \right\}, \quad (1.134)$$

y

$$\begin{aligned} \Gamma_{ij}^s &= g^{ks} [ij, k], \\ &= \frac{1}{2} g^{ks} \left\{ \frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k} \right\}. \end{aligned} \quad (1.135)$$

Estos símbolos de Christoffel y las derivadas covariantes son aplicadas en la próxima sección.

## 1.11 Operadores diferenciales de tensores.

En esta sección la derivada covariante de la sección 1.10 es aplicada para derivar las operaciones diferenciales vectoriales de la sección 1.2 en la forma tensorial general.

### Divergencia.

Reemplazando las derivadas parciales por la derivada covariante, tomamos la divergencia como

$$\vec{\nabla} \cdot \vec{V} = V_{;i}^i = \frac{\partial V^i}{\partial q^i} + V^k \Gamma_{ik}^i. \quad (1.136)$$

Expresando  $\Gamma_{ik}^i$  por la ecuación (1.135), tenemos

$$\Gamma_{ik}^i = \frac{1}{2} g^{im} \left\{ \frac{\partial g_{im}}{\partial q^k} + \frac{\partial g_{km}}{\partial q^i} - \frac{\partial g_{ik}}{\partial q^m} \right\}. \quad (1.137)$$

Cuando contraemos con  $g^{im}$  los dos últimos términos en el paréntesis de llave se cancelan, ya que

$$g^{im} \frac{\partial g_{km}}{\partial q^i} = g^{mi} \frac{\partial g_{ki}}{\partial q^m} = g^{im} \frac{\partial g_{ik}}{\partial q^m}. \quad (1.138)$$

Entonces

$$\Gamma_{ik}^i = \frac{1}{2} g^{im} \frac{\partial g_{im}}{\partial q^k}. \quad (1.139)$$

Por la teoría de determinantes

$$\frac{\partial g}{\partial q^k} = g g^{im} \frac{\partial g_{im}}{\partial q^k}, \quad (1.140)$$

donde  $g$  es el determinante de la métrica,  $g = \det(g_{ij})$ . Sustituyendo este resultado en la ecuación (1.139), obtenemos

$$\Gamma_{ik}^i = \frac{1}{2g} \frac{\partial g}{\partial q^k} = \frac{1}{g^{1/2}} \frac{\partial g^{1/2}}{\partial q^k}. \quad (1.141)$$

Esto da

$$\vec{\nabla} \cdot \vec{V} = V_{;i}^i = \frac{1}{g^{1/2}} \frac{\partial}{\partial q^k} (g^{1/2} V^k) . \quad (1.142)$$

Para comparar este resultado con la ecuación (1.21), note que  $h_1 h_2 h_3 = g^{1/2}$  y  $V^i$  (coeficiente contravariante de  $\vec{\varepsilon}_i$ ) =  $V_i/h_i$  (no hay suma), donde  $V_i$  es el coeficiente de  $\hat{e}_i$ .

### Laplaciano.

En la sección 1.2 reemplazamos el vector  $\vec{V}$  en  $\vec{\nabla} \cdot \vec{V}$  por  $\vec{\nabla} \psi$  para obtener al Laplaciano  $\vec{\nabla} \cdot \vec{\nabla} \psi$ . Aquí tenemos un  $V^i$  contravariante. Usando el tensor métrico para crear un contravariante  $\vec{\nabla} \psi$ , hacemos la sustitución

$$V^i \rightarrow g^{ik} \frac{\partial \psi}{\partial q^k} . \quad (1.143)$$

Entonces el Laplaciano  $\vec{\nabla} \cdot \vec{\nabla} \psi$  se convierte

$$\vec{\nabla} \cdot \vec{\nabla} \psi = \frac{1}{g^{1/2}} \frac{\partial}{\partial q^i} (g^{1/2} g^{ik} \frac{\partial \psi}{\partial q^k}) . \quad (1.144)$$

Para los sistemas *ortogonales* de la sección 1.2 el tensor métrico es diagonal y el contravariante  $g^{ii}$  (no hay suma) llega a ser

$$g^{ii} = (h_i)^{-2} .$$

La ecuación (1.144) se reduce a

$$\vec{\nabla} \cdot \vec{\nabla} \psi = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial q^i} \left( \frac{h_1 h_2 h_3}{h_i^2} \frac{\partial \psi}{\partial q^k} \right) .$$

en concordancia con la ecuación (1.22).

### Rotor.

La diferencia de derivadas que aparece en el rotor (ecuación (1.26)) será escrita

$$\frac{\partial V_i}{\partial q^j} - \frac{\partial V_j}{\partial q^i} .$$

Nuevamente, recordemos que las componentes de  $V_i$  aquí son los coeficientes contravariantes de los vector no unitarios base  $\vec{\varepsilon}^i$ . Los  $V_i$  de la sección 1.2 son coeficientes de los vectores unitarios  $\hat{e}_i$ . Usando

$$\Gamma_{ij}^k = \Gamma_{ji}^k ,$$

obtenemos

$$\begin{aligned} \frac{\partial V_i}{\partial q^j} - \frac{\partial V_j}{\partial q^i} &= \frac{\partial V_i}{\partial q^j} - V_k \Gamma_{ij}^k - \frac{\partial V_j}{\partial q^i} + V_k \Gamma_{ji}^k \\ &= V_{i;j} - V_{j;i} . \end{aligned} \quad (1.145)$$

Las características diferencias de derivadas del rotor llega a ser diferencias en derivadas covariantes y por lo tanto es un tensor de segundo orden (covariante en ambos índices). Como enfatizamos en la sección 1.9, la forma vectorial especial del rotor existe solamente en el espacio tridimensional.

De la ecuación (1.135) es claro que todos los símbolos de Christoffel de tres índices se anulan en el espacio de Minkowski y en el espacio-tiempo real de la relatividad especial con

$$g_{\lambda\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (1.146)$$

Aquí

$$x_0 = ct, \quad x_1 = x, \quad x_2 = y, \quad y \quad x_3 = z.$$

Esto completa el desarrollo de los operadores diferenciales en la forma tensorial general. En adición a los campos elásticos y electromagnéticos, estas formas diferenciales encuentran aplicación en mecánica (mecánica Lagrangiana, Hamiltoniana, y las ecuaciones de Euler para la rotación de cuerpos rígidos); mecánica de fluidos, y quizás mucho más importante que todas, en el espacio-tiempo curvado de la teoría moderna gravitacional.



# Capítulo 2

## Determinantes y matrices.

versión final 1.31-160401<sup>1</sup>

### 2.1 Determinantes.

Comenzamos el estudio de matrices resolviendo ecuaciones lineales las cuales nos llevan a determinantes y matrices. El concepto de determinante y su notación fueron introducidos por Leibniz.

#### Ecuaciones lineales homogéneas.

Una de las mayores aplicaciones de los determinantes está en el establecimiento de una condición para la existencia de una solución no trivial de un conjunto de ecuaciones algebraicas lineales homogéneas. Supongamos que tenemos tres incógnitas  $x_1, x_2, x_3$  (o  $n$  ecuaciones con  $n$  incógnitas).

$$\begin{aligned}a_1x_1 + a_2x_2 + a_3x_3 &= 0, \\b_1x_1 + b_2x_2 + b_3x_3 &= 0, \\c_1x_1 + c_2x_2 + c_3x_3 &= 0.\end{aligned}\tag{2.1}$$

El problema es: ¿en qué condiciones hay alguna solución, aparte de la solución trivial  $x_1 = 0, x_2 = 0, x_3 = 0$ ? Si usamos notación vectorial  $\vec{x} = (x_1, x_2, x_3)$  para la solución y tres filas  $\vec{a} = (a_1, a_2, a_3), \vec{b} = (b_1, b_2, b_3), \vec{c} = (c_1, c_2, c_3)$  para los coeficientes, tenemos que las tres ecuaciones, ecuación (2.1), se convierten en

$$\vec{a} \cdot \vec{x} = 0, \quad \vec{b} \cdot \vec{x} = 0, \quad \vec{c} \cdot \vec{x} = 0.\tag{2.2}$$

Estas tres ecuaciones vectoriales tienen la interpretación geométrica obvia que  $\vec{x}$  es ortogonal a  $\vec{a}, \vec{b}, \vec{c}$ . Si el volumen sustentado por  $\vec{a}, \vec{b}, \vec{c}$  dado por el determinante (o el producto escalar triple)

$$D_3 = (\vec{a} \times \vec{b}) \cdot \vec{c} = \det(\vec{a}, \vec{b}, \vec{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix},\tag{2.3}$$

---

<sup>1</sup>Este capítulo está basado en el tercer capítulo del libro: *Mathematical Methods for Physicists, fourth edition* de George B. Arfken & Hans J. Weber, editorial ACADEMIC PRESS.

no es cero, claramente sólo existe la solución trivial  $\vec{x} = 0$ .

Vice-versa, si el anterior determinante de coeficientes se anula, luego uno de los vectores columna es una combinación lineal de otros dos. Supongamos que  $\vec{c}$  está en el plano que sustenta  $\vec{a}$ ,  $\vec{b}$ , *i.e.*, la tercera ecuación es una combinación lineal de las primeras dos y no es independiente. Luego  $\vec{x}$  es ortogonal a ese plano tal que  $\vec{x} \sim \vec{a} \times \vec{b}$ . Ya que las ecuaciones homogéneas pueden ser multiplicadas por números arbitrarios, solamente las relaciones de  $x_i$  son relevantes, para lo cual obtenemos razones de determinantes de  $2 \times 2$

$$\begin{aligned} \frac{x_1}{x_3} &= \frac{(a_2b_3 - a_3b_2)}{(a_1b_2 - a_2b_1)}, \\ \frac{x_2}{x_3} &= -\frac{(a_1b_3 - a_3b_1)}{(a_1b_2 - a_2b_1)}, \end{aligned} \quad (2.4)$$

a partir de los componentes del producto cruz  $\vec{a} \times \vec{b}$ .

### Ecuaciones lineales no homogéneas.

El caso más simple es de dos ecuaciones con dos incógnitas

$$\begin{aligned} a_1x_1 + a_2x_2 &= a_3, \\ b_1x_1 + b_2x_2 &= b_3, \end{aligned} \quad (2.5)$$

puede ser reducido al caso previo embebiéndolo en un espacio tridimensional con una solución vectorial  $\vec{x} = (x_1, x_2, -1)$  y el vector fila  $\vec{a} = (a_1, a_2, a_3)$ ,  $\vec{b} = (b_1, b_2, b_3)$ . Como antes, ecuación (2.5) en notación vectorial,  $\vec{a} \cdot \vec{x} = 0$  y  $\vec{b} \cdot \vec{x} = 0$ , implica que  $\vec{x} \sim \vec{a} \times \vec{b}$  tal que el análogo de la ecuación (2.4) se mantiene. Para que esto se aplique la tercera componente de  $\vec{a} \times \vec{b}$  debiera ser distinta de cero, *i.e.*,  $a_1b_2 - a_2b_1 \neq 0$ , ya que la tercera componente de  $\vec{x}$  es  $-1 \neq 0$ . Esto produce que los  $x_i$  tengan la forma

$$x_1 = \frac{(a_3b_2 - a_2b_3)}{(a_1b_2 - a_2b_1)} = \frac{\begin{vmatrix} a_3 & a_2 \\ b_3 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}} \quad (2.6a)$$

$$x_2 = \frac{(a_1b_3 - a_3b_1)}{(a_1b_2 - a_2b_1)} = \frac{\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}}. \quad (2.6b)$$

El determinante en el numerador de  $x_1(x_2)$  es obtenido a partir del determinante de los coeficientes  $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$  reemplazando el primer vector columna (segundo) por el vector  $\begin{pmatrix} a_3 \\ b_3 \end{pmatrix}$  del lado inhomogéneo de la ecuación (2.5).

Estas soluciones de ecuación lineal en términos de determinantes pueden ser generalizados

a  $n$  dimensiones. El determinante es un arreglo cuadrado

$$D_n = \begin{vmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \\ c_1 & c_2 & \dots & c_n \\ \cdot & \cdot & \dots & \cdot \end{vmatrix}, \quad (2.7)$$

de números (o funciones), los coeficientes de  $n$  ecuaciones lineales en nuestro caso. El número  $n$  de columnas (y de filas) en el arreglo es llamado algunas veces el orden del determinante. La generalización de la expansión del producto escalar triple (de vectores fila de las tres ecuaciones lineales) tiende al siguiente valor del determinante  $D_n$  en  $n$  dimensiones,

$$D_n = \sum_{i,j,k,\dots} \varepsilon_{ijk\dots} a_i b_j c_k \dots, \quad (2.8)$$

donde  $\varepsilon_{ijk\dots}$ , análogo al símbolo de Levi-Civita de la sección (1.2), es  $+1$  para permutaciones pares ( $ijk\dots$ ) de  $(123\dots n)$ ,  $-1$  para permutaciones impares, y cero si algún índice es repetido.

Específicamente, para el determinante de orden tres  $D_3$  de las ecuaciones (2.3) y (2.8) tenemos

$$D_3 = +a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1. \quad (2.9)$$

El determinante de orden tres, entonces, es esta particular combinación lineal de productos. Cada producto contiene uno y sólo un elemento de cada fila y de cada columna. Cada producto es sumado si las columnas (los índices) representan una permutación par de  $(123)$  y restando si corresponde a una permutación impar. La ecuación (2.3) puede ser considerada en notación abreviada de la ecuación (2.9). El número de términos en la suma (ecuación (2.8)) es 24 para un determinante de cuarto orden, en general  $n!$  para un determinante de orden  $n$ . A causa de la aparición de signos negativos en la ecuación (2.9) pueden haber cancelaciones. Debido a esto es muy posible que un determinante de elementos grandes tenga un valor pequeño.

Algunas propiedades útiles de los determinantes de  $n$ -ésimo orden siguen de la ecuación (2.8). De nuevo, para ser específico, la ecuación (2.9) para determinantes de orden tres es usada para ilustrar estas propiedades.

### Desarrollo laplaciano por las menores.

La ecuación (2.9) puede ser reescrita

$$\begin{aligned} D_3 &= a_1(b_2 c_3 - b_3 c_2) - a_2(b_1 c_3 - b_3 c_1) + a_3(b_1 c_2 - b_2 c_1) \\ &= a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}. \end{aligned} \quad (2.10)$$

En general, el determinante de orden  $n$ -ésimo puede ser expandido como una combinación lineal de productos de elementos de alguna fila (o columna) por determinantes de orden  $(n-1)$  formados suprimiendo la fila y la columna del determinante original en el cual aparece

el elemento. Este arreglo reducido ( $2 \times 2$  en el ejemplo específico) es llamado una menor. Si el elemento está en la  $i$ -ésima fila y en la  $j$ -ésima columna, el signo asociado con el producto es  $(-1)^{i+j}$ . La menor con este signo es llamada el cofactor. Si  $M_{ij}$  es usado para designar la menor formado omitiendo la fila  $i$  y la columna  $j$  y  $c_{ij}$  es el cofactor correspondiente, la ecuación (2.10) se convierte en

$$D_3 = \sum_{j=1}^3 (-1)^{j+1} a_j M_{1j} = \sum_{j=1}^3 a_j c_{1j} . \quad (2.11)$$

En este caso, expandiendo a lo largo de la primera fila, tenemos  $i = 1$  y la suma es sobre  $j$ , las columnas.

Esta expansión de Laplace puede ser usada para sacar ventaja en la evaluación de determinantes de alto orden en el cual muchos de los elementos son nulos. Por ejemplo, para encontrar el valor de el determinante

$$D = \begin{vmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{vmatrix} , \quad (2.12)$$

expandimos a través de la fila superior para obtener

$$D = (-1)^{1+2} \cdot (1) \begin{vmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{vmatrix} . \quad (2.13)$$

Nuevamente, expandimos a través de la fila superior para obtener

$$\begin{aligned} D &= (-1) \cdot (-1)^{1+1} \cdot (-1) \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} \\ &= \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = 1 . \end{aligned} \quad (2.14)$$

Este determinante  $D$  (ecuación (2.12)) está formado de una de las matrices de Dirac que aparecen en la teoría relativista del electrón de Dirac.

### Antisimetría.

El determinante cambia de signo si cualquier par de filas son intercambiadas o si cualquier par de columnas son intercambiadas. Esto deriva del carácter par-impar del Levi-Civita  $\varepsilon$  en la ecuación (2.8) o explícitamente de la forma de las ecuaciones (2.9) y (2.10).

Esta propiedad fue usada en la sección 1.9 para desarrollar una combinación lineal totalmente antisimétrica. Esto es también frecuentemente usado en Mecánica Cuántica en la construcción de una función de onda de muchas partículas que, en concordancia con el principio de exclusión de Pauli, será antisimétrica bajo el intercambio de cualquier par de partículas idénticas con spin  $1/2$  (electrones, protones, neutrones, etc).

Como un caso especial de antisimetría, cualquier determinante con dos filas iguales o dos columnas iguales es nulo.

Si cada elemento en una fila o de una columna es cero el determinante completo es nulo.

Si cada elemento en una fila o de una columna es multiplicado por una constante, el determinante completo es multiplicado por esa constante.

El valor de un determinante es inalterado si un múltiplo de una fila es añadido (columna por columna) a otra fila o si un múltiplo de una columna es añadido (fila por fila) a otra columna. Tenemos

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 + ka_2 & a_2 & a_3 \\ b_1 + kb_2 & b_2 & b_3 \\ c_1 + kc_2 & c_2 & c_3 \end{vmatrix}. \quad (2.15)$$

Usando el desarrollo de Laplace sobre el lado derecho, obtenemos

$$\begin{vmatrix} a_1 + ka_2 & a_2 & a_3 \\ b_1 + kb_2 & b_2 & b_3 \\ c_1 + kc_2 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + k \begin{vmatrix} a_2 & a_2 & a_3 \\ b_2 & b_2 & b_3 \\ c_2 & c_2 & c_3 \end{vmatrix}, \quad (2.16)$$

entonces por la propiedad de antisimetría el segundo determinante del lado derecho se anula, verificando la ecuación (2.15).

Un caso especial, un determinante es igual a cero, si cualquier par de filas o columnas son proporcionales.

Volviendo a las ecuaciones homogéneas (2.1) y multiplicando el determinante de los coeficientes por  $x_1$ , y luego sumando  $x_2$  veces la segunda columna y  $x_3$  veces la tercera columna, podemos establecer directamente la condición para la presencia de una solución no trivial para la ecuación (2.1):

$$x_1 \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} x_1 a_1 & a_2 & a_3 \\ x_1 b_1 & b_2 & b_3 \\ x_1 c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 x_1 + a_2 x_2 + a_3 x_3 & a_2 & a_3 \\ b_1 x_1 + b_2 x_2 + b_3 x_3 & b_2 & b_3 \\ c_1 x_1 + c_2 x_2 + c_3 x_3 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} 0 & a_2 & a_3 \\ 0 & b_2 & b_3 \\ 0 & c_2 & c_3 \end{vmatrix} = 0. \quad (2.17)$$

Por lo tanto  $x_1$  ( $x_2$  y  $x_3$ ) deberían ser cero a menos que el determinante de los coeficientes sea nulo. Podemos mostrar que si el determinante de los coeficientes es nulo, existe realmente una solución no trivial.

Si nuestras ecuaciones lineales son inhomogéneas, esto es, como en la ecuación (2.5) o si los ceros en el lado derecho de la ecuación (2.1) fueran reemplazados por  $a_4$ ,  $b_4$ ,  $c_4$  respectivamente, luego de la ecuación (2.17) obtenemos,

$$x_1 = \frac{\begin{vmatrix} a_4 & a_2 & a_3 \\ b_4 & b_2 & b_3 \\ c_4 & c_2 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}}, \quad (2.18)$$

la cual generaliza la ecuación (2.6a) a la dimensión  $n = 3$ . Si el determinante de los coeficientes se anula, el conjunto de ecuaciones no homogéneas no tiene solución a menos que el numerador también se anule. En este caso las soluciones pueden existir pero ellas no son únicas.

Para el trabajo numérico, esta solución del determinante, ecuación (2.18), es enormemente difícil de manejar. El determinante puede involucrar grandes números con signos alternados, y en la resta de dos números grandes el error relativo podría remontarse al punto que hace que el resultado no tenga valor. También, aunque el método del determinante es ilustrado aquí con tres ecuaciones y tres incógnitas, podríamos fácilmente tener 200 ecuaciones con 200 incógnitas las cuales, involucran sobre 200! términos por determinante, lo que pone un desafío muy alto a la velocidad computacional. Debería haber una mejor manera. En efecto, hay una mejor manera. Una de las mejores es un proceso a menudo llamado eliminación de Gauss. Para ilustrar esta técnica, consideremos el siguiente conjunto de ecuaciones.

Resolvamos

$$\begin{aligned} 3x + 2y + z &= 11 \\ 2x + 3y + z &= 13 \\ x + y + 4z &= 12 . \end{aligned} \tag{2.19}$$

El determinante de la ecuación lineal no homogénea ecuación (2.19) es 18, por lo tanto existe una solución.

Por conveniencia y para una óptima precisión numérica, las ecuaciones son reordenadas tal que los coeficientes mayores corran a lo largo de la diagonal principal (superior izquierda a inferior derecha). Esto ha sido hecho en el conjunto anterior.

La técnica de Gauss es usar la primera ecuación para eliminar la primera incógnita  $x$  de las ecuaciones restantes. Entonces la (nueva) segunda ecuación es usada para eliminar  $y$  de la última ecuación. En general, descendemos poco a poco a través del conjunto de ecuaciones, y luego, con una incógnita determinada, avanzamos gradualmente para resolver cada una de las otras incógnitas en sucesión.

Dividiendo cada fila por su coeficiente inicial, vemos que las ecuaciones (2.19) se convierten en

$$\begin{aligned} x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3} \\ x + \frac{3}{2}y + \frac{1}{2}z &= \frac{13}{2} \\ x + y + 4z &= 12 . \end{aligned} \tag{2.20}$$

Ahora, usando la primera ecuación, eliminamos  $x$  de la segunda y la tercera:

$$\begin{aligned} x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3} \\ \frac{5}{6}y + \frac{1}{6}z &= \frac{17}{6} \\ \frac{1}{3}y + \frac{11}{3}z &= \frac{25}{3} , \end{aligned} \tag{2.21}$$

y

$$\begin{aligned} x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3} \\ y + \frac{1}{5}z &= \frac{17}{5} \\ y + 11z &= 25 . \end{aligned} \tag{2.22}$$

Repitiendo la técnica, usamos la segunda ecuación para eliminar  $y$  a partir de la tercera ecuación:

$$\begin{aligned}x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3} \\y + \frac{1}{5}z &= \frac{17}{5} \\54z &= 108 ,\end{aligned}\tag{2.23}$$

o

$$z = 2 .$$

Finalmente, al reemplazar obtenemos

$$y + \frac{1}{5} \times 2 = \frac{17}{5} ,$$

o

$$y = 3 .$$

Luego con  $z$  e  $y$  determinados,

$$x + \frac{2}{3} \times 3 + \frac{1}{3} \times 2 = \frac{11}{3} ,$$

y

$$x = 1 .$$

La técnica podría parecer no tan elegante como la ecuación (2.17), pero está bien adaptada a los computadores modernos y es más rápida que el tiempo gastado con los determinantes.

Esta técnica de Gauss puede ser usada para convertir un determinante en una forma triangular:

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & 0 & c_3 \end{vmatrix} ,$$

para un determinante de tercer orden cuyos elementos no deben ser confundidos con aquellos en la ecuación (2.3). De esta forma  $D = a_1 b_2 c_3$ . Para un determinante de  $n$ -ésimo orden la evaluación de una forma triangular requiere solamente  $n - 1$  multiplicaciones comparadas con las  $n!$  requeridas para el caso general.

Una variación de esta eliminación progresiva es conocida como eliminación de Gauss-Jordan. Comenzamos como si fuera el procedimiento de Gauss, pero cada nueva ecuación considerada es usada para eliminar una variable de todas las "otras" ecuaciones, no sólo de aquellas bajo ella. Si hemos usado esta eliminación de Gauss-Jordan, la ecuación (2.23) llegaría a ser

$$\begin{aligned}x + \frac{1}{5}z &= \frac{7}{5} \\y + \frac{1}{5}z &= \frac{17}{5} \\z &= 2 ,\end{aligned}\tag{2.24}$$

usando la segunda ecuación de la ecuación (2.22) para eliminar  $y$  de ambas, la primera y tercera ecuaciones. Entonces la tercera ecuación de la ecuación (2.24) es usada para eliminar  $z$  de la primera y segunda ecuaciones, dando

$$\begin{aligned}x &= 1 \\y &= 3 \\z &= 2 ,\end{aligned}\tag{2.25}$$

Volveremos a la técnica de Gauss-Jordan cuando invertamos matrices.

Otra técnica disponible para el uso computacional es la técnica de Gauss-Seidel. Cada técnica tiene sus ventajas y desventajas. Los métodos de Gauss y Gauss-Jordan pueden tener problemas de precisión para un determinante grande. Esto también es un problema para la inversión de matrices. El método de Gauss-Seidel, como un método iterativo, puede tener problemas de convergencia.

## 2.2 Matrices.

El análisis matricial pertenece al álgebra lineal ya que las matrices son operadores o mapas lineales tales como rotaciones. Supongamos, por ejemplo, que rotamos las coordenadas cartesianas de un espacio bidimensional tal que, en notación vectorial,

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} x_1 \cos \varphi & x_2 \sin \varphi \\ -x_2 \sin \varphi & x_1 \cos \varphi \end{pmatrix} = \left( \sum_j a_{ij} x_j \right) .\tag{2.26}$$

Etiquetamos el arreglo de elementos  $a_{ij}$  por la matriz  $\mathbf{A}$  de  $2 \times 2$  consistente de dos filas y dos columnas, además, consideramos los vectores  $x$ ,  $x'$  como matrices de  $2 \times 1$ . Tomemos la suma de productos de la ecuación (2.26) como una definición de la multiplicación matricial que involucra el producto escalar de cada uno de los vectores fila de  $\mathbf{A}$  con el vector columna  $x$ . Así en notación matricial la ecuación (2.26) se convierte en

$$x' = \mathbf{A}x .\tag{2.27}$$

Para extender esta definición de multiplicación de una matriz por un vector columna a el producto de dos matrices de  $2 \times 2$ , consideremos la rotación de coordenada seguida por una segunda rotación dada por la matriz  $\mathbf{B}$  tal que

$$x'' = \mathbf{B}x' .\tag{2.28}$$

Por componentes

$$x''_i = \sum_j b_{ij} x'_j = \sum_j b_{ij} \sum_k a_{jk} x_k = \sum_k \left( \sum_j b_{ij} a_{jk} \right) x_k .\tag{2.29}$$

La suma sobre  $j$  es la multiplicación matricial definiendo una matriz  $\mathbf{C} = \mathbf{B}\mathbf{A}$  tal que

$$x''_i = \sum_k c_{ik} x_k ,\tag{2.30}$$

o  $x'' = Cx$  en notación matricial. Nuevamente, esta definición involucra el producto escalar de vectores filas de  $B$  con vectores columnas de  $A$ . Esta definición de multiplicación matricial se puede generalizar a matrices de  $m \times n$  y es útil, realmente “su utilidad es la justificación de su existencia”. La interpretación física es que el producto matricial de dos matrices,  $BA$ , es la rotación que conduce del sistema sin prima directamente al sistema de coordenadas con doble prima. Antes de pasar a la definición formal, podemos notar que el operador  $A$  está descrito por sus efectos sobre las coordenadas o vectores base. Los elementos de matriz  $a_{ij}$  constituyen una representación del operador, una representación que depende de la elección de una base.

El caso especial donde una matriz tiene una columna y  $n$  filas es llamada un vector columna,  $|x\rangle$ , con componentes  $x_i$ ,  $i = 1, 2, \dots, n$ . Si  $A$  es una matriz de  $n \times n$ ,  $|x\rangle$  es un vector columna de  $n$  componentes,  $A|x\rangle$  está definida como en la ecuación (2.27) y (2.26). Similarmente, si una matriz tiene una fila y  $n$  columnas, es llamada un vector fila,  $\langle x|$  con componentes  $x_i$ ,  $i = 1, 2, \dots, n$ . Claramente,  $\langle x|$  resulta de  $|x\rangle$  por el intercambio de filas y columnas, una operación matricial llamada *transposición*, y para cualquier matriz  $A$ ,  $\tilde{A}$  es llamada <sup>2</sup> la transpuesta de  $A$  con elementos de matriz  $(\tilde{A})_{ik} = A_{ik}$ . Transponiendo un producto de matrices  $AB$  se invierte el orden y da  $BA$ ; similarmente,  $A|x\rangle$  se transpone como  $\langle x|A$ . El producto escalar toma la forma  $\langle x|y\rangle$ .

### Definiciones básicas.

Una matriz puede ser definida como un arreglo cuadrado o rectangular de números o funciones que obedecen ciertas leyes. Esto es una extensión perfectamente lógica de los conceptos matemáticos familiares. En aritmética tratamos con números simples. En la teoría de variable compleja tratamos con pares ordenados de números,  $(1, 2) = 1 + 2i$ , en el cual el orden es importante. Ahora consideremos números (o funciones) ordenados en un arreglo cuadrados o rectangular. Por conveniencia en el trabajo posterior los números son distinguidos por dos subíndices, el primero indica la fila (horizontal) y el segundo indica la columna (vertical) en la cual aparecen los números. Por ejemplo,  $a_{13}$  es el elemento de matriz en la primera fila y tercera columna. De este modo, si  $A$  es una matriz con  $m$  filas y  $n$  columnas,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \quad (2.31)$$

Quizás el hecho más importante a notar es que los elementos  $a_{ij}$  no están combinados unos con otros. Una matriz no es un determinante. Es un arreglo ordenado de números, no un simple número.

La matriz  $A$  hasta ahora de sólo es un arreglo de números que tiene las propiedades que le asignamos. Literalmente, esto significa construir una nueva forma de matemáticas. Postulamos que las matrices  $A$ ,  $B$  y  $C$ , con elementos  $a_{ij}$ ,  $b_{ij}$  y  $c_{ij}$ , respectivamente, combinan de acuerdo a las siguientes reglas.

---

<sup>2</sup>Algunos textos denotan  $A$  transpuesta por  $A^T$ .

**Igualdad.**

Matriz  $A =$  Matriz  $B$  si y sólo si  $a_{ij} = b_{ij}$  para todos los valores de  $i$  y  $j$ . Esto, por su puesto, requiere que  $A$  y  $B$  sean cada uno arreglos de  $m \times n$  ( $m$  filas y  $n$  columnas).

**Suma.**

$A + B = C$  si y sólo si  $a_{ij} + b_{ij} = c_{ij}$  para todos los valores de  $i$  y  $j$ , los elementos se combinan de acuerdo a las leyes del álgebra lineal (o aritmética si hay números simples). Esto significa que  $A + B = B + A$ , la conmutación. También, se satisface la ley de asociatividad  $(A + B) + C = A + (B + C)$ . Si todos los elementos son cero, la matriz es llamada matriz nula y se denota por  $0$ . Para todo  $A$ ,

$$A + 0 = 0 + A = A ,$$

con

$$0 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} . \quad (2.32)$$

Tal que las matrices de  $m \times n$  forman un espacio lineal con respecto a la suma y la resta.

**Multiplicación (por un escalar).**

La multiplicación de la matriz  $A$  por una cantidad escalar  $\alpha$  está definida como

$$\alpha A = (\alpha A) , \quad (2.33)$$

en la cual los elementos de  $\alpha A$  son  $\alpha a_{ij}$ ; esto es, cada elemento de la matriz  $A$  es multiplicado por el factor escalar. Esto contrasta con el comportamiento de los determinantes en el cual el factor  $\alpha$  multiplica solamente una columna o una fila y no cada elemento del determinante. Una consecuencia de esta multiplicación por escalar es que

$$\alpha A = A\alpha , \quad \text{conmutación.} \quad (2.34)$$

**Multiplicación (multiplicación matricial) producto interno.**

$$AB = C \quad \text{si y solo si} \quad c_{ij} = \sum_k a_{ik} b_{kj} . \quad (2.35)$$

Los elementos  $i$  y  $j$  de  $C$  están formados como un producto escalar de la  $i$ -ésima fila de  $A$  con el  $j$ -ésima columna de  $B$  (el cual demanda que  $A$  tenga el mismo número de columnas como  $B$  tiene de filas). El índice mudo  $k$  toma los valores  $1, 2, \dots, n$  en sucesión, esto es,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} , \quad (2.36)$$

para  $n = 3$ . Obviamente, el índice mudo  $k$  puede ser reemplazado por algún otro símbolo que no esté en uso sin alterar la ecuación (2.35). Quizás la situación puede ser aclarada afirmando que la ecuación (2.35) defina el método de combinar ciertas matrices. Este método de combinación, es llamado multiplicación matricial. Para ilustrar, consideremos dos matrices (matrices de Pauli)

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.37)$$

El elemento  $_{11}$  del producto,  $(\sigma_1\sigma_3)_{11}$  está dado por la suma de productos de elementos de la primera fila de  $\sigma_1$  con el correspondiente elemento de la primera columna de  $\sigma_3$ : Aquí

$$(\sigma_1\sigma_3)_{ij} = \sigma_{1i1}\sigma_{31j} + \sigma_{1i2}\sigma_{32j}.$$

Una aplicación directa de la multiplicación de matrices muestra que

$$\sigma_3\sigma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (2.38)$$

y por la ecuación (2.35)

$$\sigma_1\sigma_3 = -\sigma_1\sigma_3. \quad (2.39)$$

Excepto en casos especiales, la multiplicación de matrices no es conmutativa.<sup>3</sup>

$$AB \neq BA. \quad (2.40)$$

Sin embargo, de la definición de multiplicación de matrices podemos mostrar que se mantiene una ley de asociatividad,  $(AB)C = A(BC)$ . También se satisface una ley de distributividad,  $A(B + C) = AB + AC$ . La matriz unidad tiene elementos  $\delta_{ij}$ , la delta de Kronecker, y la propiedad de que  $1A = A1 = A$  para toda  $A$ ,

$$1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \quad (2.41)$$

Notamos que es posible que el producto de dos matrices sea una matriz nula sin ser ninguna de ellas una matriz nula. Por ejemplo, si

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}.$$

$AB = 0$ . Esto difiere de la multiplicación de números reales o complejos los cuales forman un *campo*, mientras que las estructura aditiva y multiplicativa de las matrices es llamada *anillo* por los matemáticos.

---

<sup>3</sup>La pérdida de la propiedad conmutativa es descrita por el conmutador  $[A, B] = AB - BA$ . La no conmutatividad se expresa por  $[A, B] \neq 0$ .

Si  $A$  es una matriz de  $n \times n$  con determinante  $|A| \neq 0$ , luego tiene una única inversa  $A^{-1}$  tal que  $AA^{-1} = A^{-1}A = 1$ . Si  $B$  es también una matriz de  $n \times n$  con inversa  $B^{-1}$ , luego el producto de  $AB$  tiene la inversa

$$(AB)^{-1} = B^{-1}A^{-1}, \quad (2.42)$$

ya que  $ABB^{-1}A^{-1} = 1 = B^{-1}A^{-1}AB$ .

El *teorema del producto* el cual dice que el determinante de un producto,  $|AB|$ , de dos matrices de  $n \times n$   $A$  y  $B$  es igual al producto de los determinantes,  $|A||B|$ , uniendo matrices con determinantes. El anterior teorema puede ser fácilmente probado.

### Producto directo.

Un segundo procedimiento para multiplicar matrices, conocido como el tensor producto directo o de Kronecker. Si  $A$  es una matriz de  $m \times m$  y  $B$  una matriz de  $n \times n$ , luego el producto directo es

$$A \otimes B = C. \quad (2.43)$$

$C$  es una matriz de  $mn \times mn$  con elementos

$$C_{\alpha\beta} = A_{ij}B_{kl}, \quad (2.44)$$

con

$$\alpha = n(i-1) + k, \quad \beta = n(j-1) + l.$$

Por ejemplo, si  $A$  y  $B$  ambas son matrices de  $2 \times 2$ ,

$$\begin{aligned} A \otimes B &= \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}. \end{aligned} \quad (2.45)$$

El producto directo es asociativo pero no conmutativo. Como un ejemplo de producto directo, las matrices de Dirac pueden ser desarrolladas como productos directos de las matrices de Pauli y de la matriz unidad. Otros ejemplos aparecen en la construcción de grupos en teoría de grupos y en espacios de Hilbert en teoría cuántica.

El producto directo definido aquí es algunas veces llamado la forma *standard* y es denotado por  $\otimes$ . Otros tres tipos de producto directo de matrices existe como posibilidades o curiosidades matemáticas pero tienen muy poca o ninguna aplicación en física matemática.

**Matrices diagonales.**

Un tipo especial muy importante de matrices es la matriz cuadrada en la cual todos los elementos no diagonales son cero. Específicamente, si una matriz  $A$  de  $3 \times 3$  es diagonal,

$$A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} .$$

Una interpretación física de tales matrices diagonales y el método de reducir matrices a esta forma diagonal son considerados en la sección 2.5. Aquí nos limitamos a notar la importante propiedad de que la multiplicación de matrices es conmutativa,  $AB = BA$ , si  $A$  y  $B$  son cada una diagonales.

**Traza.**

En cualquiera matriz cuadrada la suma de los elementos diagonales es llamada la *traza*. Claramente la traza es una operación lineal:

$$\text{traza}(A - B) = \text{traza}(A) - \text{traza}(B) .$$

Una de sus interesantes y útiles propiedades es que la traza de un producto de dos matrices  $A$  y  $B$  es independiente del orden de la multiplicación:

$$\begin{aligned} \text{traza}(AB) &= \sum_i (AB)_{ii} = \sum_i \sum_j a_{ij} b_{ji} \\ &= \sum_i \sum_j b_{ji} a_{ij} = \sum_j (BA)_{jj} \\ &= \text{traza}(BA) . \end{aligned} \tag{2.46}$$

Esto se mantiene aún cuando  $AB \neq BA$ . La ecuación (2.46) significa que la traza de cualquier conmutador,  $[A, B] = AB - BA$ , es cero. De la ecuación (2.46) obtenemos

$$\text{traza}(ABC) = \text{traza}(BCA) = \text{traza}(CAB) ,$$

lo cual muestra que la traza es invariante bajo permutaciones cíclicas de la matriz en un producto.

Para una matriz simétrica o una matriz Hermítica compleja la traza es la suma, y el determinante el producto, de sus autovalores, y ambos son coeficientes del polinomio característico. La traza servirá una función similar para las matrices como la ortogonalidad sirve para los vectores y funciones.

En términos de tensores la traza es una contracción y como el tensor de segundo orden contraído es un escalar (invariante).

Las matrices son usadas ampliamente para representar los elementos de grupos. La traza de las matrices representando los elementos de grupo es conocido en teoría de grupos como el *carácter*. La razón de este nombre especial y espacial atención es que mientras las matrices pueden variar la traza o carácter se mantiene invariante.

**Inversión de matriz.**

Al comienzo de esta sección la matriz  $\mathbf{A}$  fue presentada como la representación de un operador que (linealmente) transforma los ejes de coordenadas. Una rotación podría ser un ejemplo de tal transformación lineal. Ahora buscaremos la transformación inversa  $\mathbf{A}^{-1}$  que restablecerá los ejes de coordenadas originales. Esto significa, ya sea como una ecuación matricial o de operador<sup>4</sup>,

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{1} . \quad (2.47)$$

Podemos probar (ejercicio) que

$$a_{ij}^{-1} = \frac{C_{ji}}{|\mathbf{A}|} , \quad (2.48)$$

con la suposición que el determinante de  $\mathbf{A}$  ( $|\mathbf{A}| \neq 0$ ). Si es cero, etiquetaremos a  $\mathbf{A}$  como singular. No existe la inversa. Como fue explicado en la sección 2.1 esta forma con determinante es *totalmente inapropiado para el trabajo numérico* con grandes matrices.

Hay una amplia variedad de técnicas alternativas. Una de las mejores y más comúnmente usada es la técnica de inversión de matrices de Gauss-Jordan. La teoría está basada en los resultados que muestran que existen matrices  $\mathbf{M}_L$  tal que el producto  $\mathbf{M}_L\mathbf{A}$  será  $\mathbf{A}$  pero con

- a. una fila multiplicada por una constante, o
- b. una fila reemplazada por la fila original menos un múltiplo de otra fila, o
- c. filas intercambiadas.

Otras matrices  $\mathbf{M}_R$  operando sobre la derecha de  $(\mathbf{A}\mathbf{M}_R)$  puede llevar a las mismas operaciones sobre las columnas de  $\mathbf{A}$ .

Esto significa que las filas y las columnas de la matriz pueden ser alteradas (por multiplicación de matrices) como si estuviéramos tratando con determinantes, así podemos aplicar las técnicas de eliminación de Gauss-Jordan a los elementos de matriz. Por tanto existe una matriz  $\mathbf{M}_L$  (o  $\mathbf{M}_R$ ) tal que<sup>5</sup>

$$\mathbf{M}_L\mathbf{A} = \mathbf{1} . \quad (2.49)$$

La  $\mathbf{M}_L = \mathbf{A}^{-1}$ . Determinamos  $\mathbf{M}_L$  realizando las operaciones de eliminación idénticas sobre la matriz unidad. Luego

$$\mathbf{M}_L\mathbf{1} = \mathbf{M}_L . \quad (2.50)$$

Para clarificar ésto consideremos un ejemplo específico.

Deseamos invertir la matriz

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix} . \quad (2.51)$$

<sup>4</sup>Aquí y a través de todo el capítulo nuestras matrices tienen rango finito.

<sup>5</sup>Recordemos que  $\det(\mathbf{A}) \neq 0$ .

Por conveniencia escribimos  $A$  y  $I$  lado a lado realizando operaciones idénticas sobre cada una de ellas

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.52)$$

Para ser sistemáticos, multiplicamos cada fila para obtener  $a_{k1} = 1$ ,

$$\begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 1 & \frac{3}{2} & \frac{1}{2} \\ 1 & 1 & 4 \end{pmatrix} \quad \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.53)$$

Restando la primera fila de la segunda y tercera, obtenemos

$$\begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & \frac{1}{3} & \frac{11}{3} \end{pmatrix} \quad \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & \frac{1}{2} & 0 \\ -\frac{1}{3} & 0 & 1 \end{pmatrix}. \quad (2.54)$$

Entonces dividimos la segunda fila (de ambas matrices) por  $5/6$  y sustrayéndola  $2/3$  veces de la primera fila, y  $1/3$  veces de la tercera fila. Los resultados para ambas matrices son

$$\begin{pmatrix} 1 & 0 & \frac{1}{5} \\ 0 & 1 & \frac{1}{5} \\ 0 & 0 & \frac{18}{5} \end{pmatrix} \quad \begin{pmatrix} \frac{3}{5} & -\frac{2}{5} & 0 \\ -\frac{2}{5} & \frac{3}{5} & 0 \\ -\frac{1}{5} & -\frac{1}{5} & 1 \end{pmatrix}. \quad (2.55)$$

Dividimos la tercera fila (de ambas matrices) por  $18/5$ . Luego como último paso  $1/5$  veces la tercera fila es sustraída de cada una de las dos primeras filas (de ambas matrices). Nuestro par final es

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} \frac{11}{8} & -\frac{7}{18} & -\frac{1}{18} \\ -\frac{7}{18} & \frac{11}{18} & -\frac{1}{18} \\ -\frac{1}{18} & -\frac{1}{18} & \frac{5}{18} \end{pmatrix}. \quad (2.56)$$

El chequeo es multiplicar la original  $A$  por la calculada  $A^{-1}$  para ver si realmente obtuvimos la matriz unidad  $I$ .

Como con la solución de Gauss-Jordan de ecuaciones algebraicas simultáneas, esta técnica está bien adaptada para computadores.

## 2.3 Matrices ortogonales.

El espacio de tres dimensiones ordinario puede ser descrito con las coordenadas cartesianas  $(x_1, x_2, x_3)$ . Consideremos un segundo conjunto de coordenadas cartesianas  $(x'_1, x'_2, x'_3)$  cuyo origen y sentido coinciden con el primero pero su orientación es diferente (figura 2.1). Podemos decir que el sistema de ejes prima ha sido rotado respecto al inicial sistema de coordenadas sin prima. Ya que esta rotación es una operación lineal, esperamos una ecuación matricial que relaciones la base con primas con la sin primas.

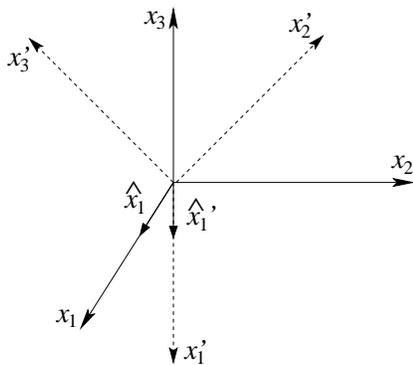


Figura 2.1: Sistemas de coordenadas cartesianos.

**Cosenos directores.**

Un vector unitario a lo largo del eje  $x'_1$  ( $\hat{x}'_1$ ) puede ser resuelto en sus componentes a lo largo de los ejes  $x_1$ ,  $x_2$  y  $x_3$  por las usuales técnicas de proyección.

$$\hat{x}'_1 = \hat{x}_1 \cos(x'_1, x_1) + \hat{x}_2 \cos(x'_1, x_2) + \hat{x}_3 \cos(x'_1, x_3) . \quad (2.57)$$

Por conveniencia estos cosenos, los cuales son los cosenos directores, son etiquetados

$$\begin{aligned} \cos(x'_1, x_1) &= \hat{x}'_1 \cdot \hat{x}_1 = a_{11} , \\ \cos(x'_1, x_2) &= \hat{x}'_1 \cdot \hat{x}_2 = a_{12} , \\ \cos(x'_1, x_3) &= \hat{x}'_1 \cdot \hat{x}_3 = a_{13} . \end{aligned} \quad (2.58)$$

Continuando, tenemos

$$\begin{aligned} \cos(x'_2, x_1) &= \hat{x}'_2 \cdot \hat{x}_1 = a_{21} , & (a_{21} \neq a_{12}) , \\ \cos(x'_2, x_2) &= \hat{x}'_2 \cdot \hat{x}_2 = a_{22} , & \text{y así sucesivamente.} \end{aligned} \quad (2.59)$$

Ahora la ecuación (2.57) puede ser reescrita como

$$\hat{x}'_1 = \hat{x}_1 a_{11} + \hat{x}_2 a_{12} + \hat{x}_3 a_{13}$$

y también

$$\begin{aligned} \hat{x}'_2 &= \hat{x}_1 a_{21} + \hat{x}_2 a_{22} + \hat{x}_3 a_{23} \\ \hat{x}'_3 &= \hat{x}_1 a_{31} + \hat{x}_2 a_{32} + \hat{x}_3 a_{33} . \end{aligned} \quad (2.60)$$

También podemos ir de la otra manera resolviendo  $\hat{x}_1$ ,  $\hat{x}_2$  y  $\hat{x}_3$  en sus componentes en el sistema con primas. Entonces

$$\begin{aligned} \hat{x}_1 &= \hat{x}'_1 a_{11} + \hat{x}'_2 a_{21} + \hat{x}'_3 a_{31} \\ \hat{x}_2 &= \hat{x}'_1 a_{12} + \hat{x}'_2 a_{22} + \hat{x}'_3 a_{32} \\ \hat{x}_3 &= \hat{x}'_1 a_{13} + \hat{x}'_2 a_{23} + \hat{x}'_3 a_{33} . \end{aligned} \quad (2.61)$$

**Aplicaciones a vectores.**

Si consideramos un vector cuyas componentes son funciones de la posición, entonces

$$\begin{aligned}\vec{V}(x_1, x_2, x_3) &= \hat{x}_1 V_1 + \hat{x}_2 V_2 + \hat{x}_3 V_3 \\ &= \vec{V}'(x'_1, x'_2, x'_3) = \hat{x}'_1 V'_1 + \hat{x}'_2 V'_2 + \hat{x}'_3 V'_3 ,\end{aligned}\tag{2.62}$$

ya que el punto puede ser dado en cualquiera de los dos sistemas de coordenadas  $(x_1, x_2, x_3)$  o  $(x'_1, x'_2, x'_3)$ . Notemos que  $\vec{V}$  y  $\vec{V}'$  son geoméricamente el mismo vector (pero con diferentes componentes). Si los ejes de coordenadas son rotados, el vector se mantiene fijo. Usando la ecuación (2.60) para eliminar  $\hat{x}_1, \hat{x}_2, \hat{x}_3$ , podemos separar la ecuación (2.62) en tres ecuaciones escalares

$$\begin{aligned}V'_1 &= a_{11}V_1 + a_{12}V_2 + a_{13}V_3 \\ V'_2 &= a_{21}V_1 + a_{22}V_2 + a_{23}V_3 \\ V'_3 &= a_{31}V_1 + a_{32}V_2 + a_{33}V_3 .\end{aligned}\tag{2.63}$$

En particular, estas relaciones se mantendrán para las coordenadas de un punto  $(x_1, x_2, x_3)$  y  $(x'_1, x'_2, x'_3)$ , dando

$$\begin{aligned}x'_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x'_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 ,\end{aligned}\tag{2.64}$$

y similarmente para las coordenadas primas. En esta notación el conjunto de tres ecuaciones (2.64) pueden ser escritas como

$$x'_i = \sum_{j=1}^3 a_{ij}x_j ,\tag{2.65}$$

donde  $i$  toma los valores 1, 2 y 3 y el resultado son tres ecuaciones separadas.

De la ecuación anterior podemos derivar interesante información sobre los  $a_{ij}$  los cuales describen la orientación del sistema de coordenadas  $(x'_1, x'_2, x'_3)$  relativa al sistema  $(x_1, x_2, x_3)$ . La distancia respecto al origen es la misma en ambos sistemas. Elevando al cuadrado,

$$\begin{aligned}\sum_i x_i^2 &= \sum_i x_i'^2 \\ &= \sum_i \left( \sum_j a_{ij}x_j \right) \left( \sum_k a_{ik}x_k \right) \\ &= \sum_{j,k} x_j x_k \sum_i a_{ij} a_{ik} .\end{aligned}\tag{2.66}$$

Esto sólo puede ser cierto para todos los puntos si y sólo si

$$\sum_i a_{ij} a_{ik} = \delta_{jk} , \quad j, k = 1, 2, 3 .\tag{2.67}$$

La ecuación (2.67) es una consecuencia de requerir que la longitud permanezca constante (invariante) bajo rotaciones del sistema de coordenadas, es llamada la *condición de ortogonalidad*. Los  $a_{ij}$  escritos como una matriz  $A$ , forman una matriz ortogonal. Notemos que la ecuación (2.67) no es una multiplicación matricial.

En notación matricial la ecuación (2.65) llega a ser

$$|x'\rangle = A|x\rangle . \quad (2.68)$$

### Condiciones de ortogonalidad, caso bidimensional.

Podemos ganar un mejor entendimiento de los  $a_{ij}$  y de la condición de ortogonalidad considerando con detalle rotaciones en dos dimensiones. Esto lo podemos pensar como un sistema tridimensional con los ejes  $x_1$  y  $x_2$  rotados respecto a  $x_3$ . De la figura 2.2,

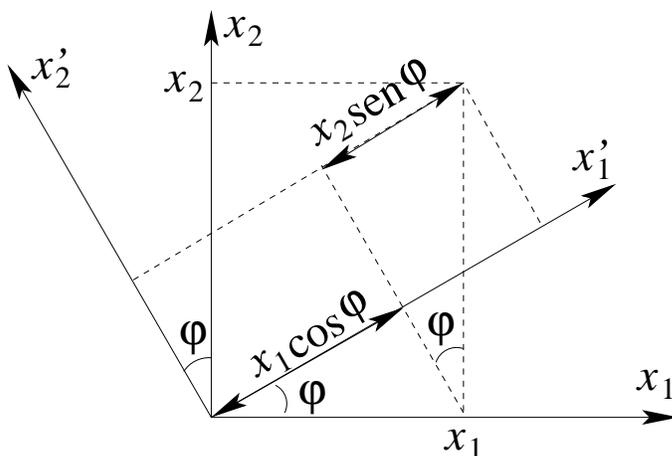


Figura 2.2: Sistemas de coordenadas rotados en dos dimensiones.

$$\begin{aligned} x'_1 &= x_1 \cos \varphi + x_2 \sin \varphi , \\ x'_2 &= -x_1 \sin \varphi + x_2 \cos \varphi . \end{aligned} \quad (2.69)$$

Por lo tanto por la ecuación (2.68)

$$A = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} . \quad (2.70)$$

Notemos que  $A$  se reduce a la matriz unidad para  $\varphi = 0$ . La rotación cero significa que nada ha cambiado. Es claro a partir de la figura 2.2 que

$$\begin{aligned} a_{11} &= \cos \varphi = \cos(x'_1, x_1) , \\ a_{12} &= \sin \varphi = \cos\left(\frac{\pi}{2} - \varphi\right) = \cos(x'_1, x_2) , \quad \text{y así sucesivamente,} \end{aligned} \quad (2.71)$$

de este modo identificamos los elementos de matriz  $a_{ij}$  con los cosenos directores. La ecuación (2.67), la condición de ortogonalidad, llega a ser

$$\begin{aligned} \sin^2 \varphi + \cos^2 \varphi &= 1 , \\ \sin \varphi \cos \varphi - \sin \varphi \cos \varphi &= 0 . \end{aligned} \quad (2.72)$$

la extensión a tres dimensiones ( rotación de las coordenadas a lo largo del eje  $z$  en un ángulo  $\varphi$  en el sentido de los punteros del reloj) es simplemente

$$\mathbf{A} = \begin{pmatrix} \cos \varphi & \operatorname{sen} \varphi & 0 \\ -\operatorname{sen} \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \quad (2.73)$$

El  $a_{33} = 1$  expresa el hecho que  $x'_3 = x_3$ , ya que la rotación ha sido en torno al eje  $x_3$ . Los ceros garantizan que  $x'_1$  y  $x'_2$  no dependen de  $x_3$  y que  $x'_3$  no depende de  $x_1$  y  $x_2$ . En un lenguaje más sofisticado,  $x_1$  y  $x_2$  se extienden sobre un subespacio invariante, mientras que  $x_3$  forma un subespacio invariante por si solo. La forma de  $\mathbf{A}$  es reducible. La ecuación (2.73) da una posible descomposición.

### Matriz inversa $\mathbf{A}^{-1}$ .

Volviendo a la matriz de transformación general  $\mathbf{A}$ , la matriz inversa  $\mathbf{A}^{-1}$  es definida tal que

$$|x\rangle = \mathbf{A}^{-1}|x'\rangle . \quad (2.74)$$

Esto es,  $\mathbf{A}^{-1}$  describe el inverso de la rotación dada por  $\mathbf{A}$  y retorna el sistema de coordenadas a su posición original. Simbólicamente, las ecuaciones (2.68) y (2.74) combinadas dan

$$|x\rangle = \mathbf{A}^{-1}\mathbf{A}|x\rangle , \quad (2.75)$$

y ya que  $|x\rangle$  es arbitrario,

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{1} , \quad (2.76)$$

la matriz unidad, Similarmente,

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{1} . \quad (2.77)$$

usando las ecuaciones (2.68) y (2.74) y eliminando  $|x'\rangle$  en vez de  $|x\rangle$ .

### Matriz transpuesta, $\tilde{\mathbf{A}}$ .

Podemos determinar los elementos de nuestra postulada matriz inversa  $\mathbf{A}^{-1}$  empleando la condición de ortogonalidad. La ecuación (2.67), la condición de ortogonalidad, no está de acuerdo con nuestra definición de multiplicación matricial, pero la podemos definir de acuerdo a una nueva matriz  $\tilde{\mathbf{A}}$  tal que

$$\tilde{a}_{ji} = a_{ij} . \quad (2.78)$$

La ecuación (2.67) llega a ser

$$\tilde{\mathbf{A}}\mathbf{A} = \mathbf{1} . \quad (2.79)$$

Esta es una reformulación de la condición de ortogonalidad y puede ser tomada como una definición de ortogonalidad. Multiplicando (2.79) por  $A^{-1}$  por la derecha y usando la ecuación (2.77), tenemos

$$\tilde{A} = A^{-1} . \quad (2.80)$$

Este importante resultado que la inversa es igual a la transpuesta se mantiene sólo para matrices ortogonales y puede ser tomado como una reformulación de la condición de ortogonalidad.

Multiplicando la ecuación (2.80) por  $A$  por la izquierda, obtenemos

$$A\tilde{A} = 1 , \quad (2.81)$$

o

$$\sum_i a_{ji}a_{ki} = \delta_{jk} , \quad (2.82)$$

lo cual es otra forma más de la condición de ortogonalidad.

Resumiendo, la condición de ortogonalidad puede ser enunciada de varias maneras equivalentes:

$$\sum_i a_{ij}a_{ik} = \delta_{jk} \quad (2.83a)$$

$$\sum_i a_{ji}a_{ki} = \delta_{jk} \quad (2.83b)$$

$$\tilde{A}A = A\tilde{A} = 1 \quad (2.83c)$$

$$\tilde{A} = A^{-1} . \quad (2.83d)$$

Cualquiera de estas relaciones es condición necesaria y suficiente para que  $A$  sea ortogonal.

Es posible ahora ver y entender por qué el nombre *ortogonal* es apropiado para estas matrices. Tenemos la forma general

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} ,$$

de una matriz de cosenos directores en la cual  $a_{ij}$  es el coseno del ángulo entre  $x'_i$  y  $x_j$ . Por lo tanto,  $a_{11}$ ,  $a_{12}$ ,  $a_{13}$  son los cosenos directores de  $x'_1$  relativo a  $x_1$ ,  $x_2$ ,  $x_3$ . Estos tres elementos de  $A$  definen una unidad de longitud a lo largo de  $x'_1$ , esto es, un vector unitario  $\hat{x}'_1$ ,

$$\hat{x}'_1 = \hat{x}_1 a_{11} + \hat{x}_2 a_{12} + \hat{x}_3 a_{13} .$$

La relación de ortogonalidad (ecuación (2.82)) es simplemente una declaración que los vectores unitarios  $\hat{x}'_1$ ,  $\hat{x}'_2$ , y  $\hat{x}'_3$  son mutuamente perpendiculares o ortogonales. Nuestra matriz de transformación ortogonal  $A$  transforma un sistema ortogonal en un segundo sistema ortogonal de coordenadas por rotación y/o reflexión.

### Ángulos de Euler.

Nuestra matriz de transformación  $A$  contiene nueve cosenos directores. Claramente, sólo tres de ellos son independientes, la ecuación (2.67) proveen seis restricciones. De otra manera, uno puede decir que necesita dos parámetros ( $\theta$  y  $\varphi$  en coordenadas polares esféricas) para fijar el eje de rotación, más uno adicional para describir la magnitud de la rotación en torno a ese eje. En la formulación Lagrangiana de la mecánica es necesario describir  $A$  usando algún conjunto de tres parámetros independientes más que los redundantes cosenos directores. La elección usual de estos parámetros es la de los ángulos de Euler<sup>6</sup>

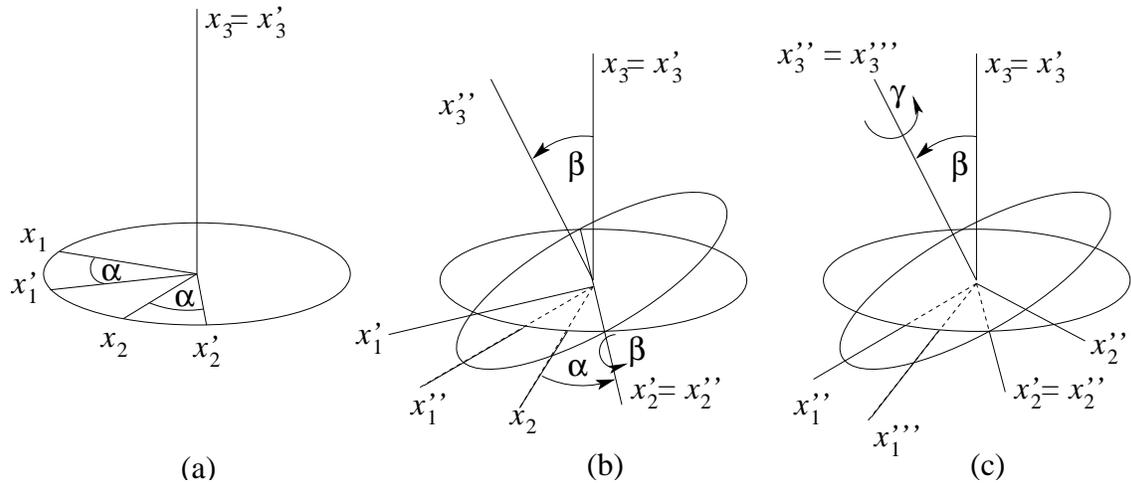


Figura 2.3: (a) Rotación respecto al eje  $x_3$  en un ángulo  $\alpha$ ; (b) Rotación respecto a un eje  $x_2'$  en un ángulo  $\beta$ ; (c) Rotación respecto a un eje  $x_3''$  en un ángulo  $\gamma$ .

El objetivo de describir la orientación de un sistema final rotado  $(x_1''', x_2''', x_3''')$  relativo a algun sistema de coordenadas inicial  $(x_1, x_2, x_3)$ . El sistema final es desarrollado en tres pasos cada paso involucra una rotación descrita por un ángulo de Euler (figura 2.3):

1. Los ejes  $x_1', x_2',$  y  $x_3'$  son rotados respecto al eje  $x_3$  en un ángulo  $\alpha$  en el sentido horario relativo a  $x_1, x_2$  y  $x_3$ . (Los ejes  $x_3$  y  $x_3'$  coinciden.)
2. los ejes  $x_1'', x_2'',$  y  $x_3''$  son rotados respecto al eje  $x_2'$  en un ángulo  $\beta$  en el sentido horario relativo a  $x_1', x_2'$  y  $x_3'$ . (Los ejes  $x_2'$  y  $x_2''$  coinciden.)
3. la tercera y final rotación es en un ángulo  $\gamma$  en sentido horario respecto al eje  $x_3''$  produciendo el sistema  $(x_1''', x_2''', x_3''')$ . (Los ejes  $x_3''$  y  $x_3'''$  coinciden.)

Las tres matrices que describen estas rotaciones son:

$$R_z(\alpha) = \begin{pmatrix} \cos \alpha & \text{sen } \alpha & 0 \\ -\text{sen } \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (2.84)$$

<sup>6</sup>No hay una única manera de definir los ángulos de Euler. Usamos la elección usual en Mecánica Cuántica de momento angular.

exactamente como en la ecuación (2.73),

$$\mathbf{R}_y(\beta) = \begin{pmatrix} \cos \beta & 0 & -\operatorname{sen} \beta \\ 0 & 1 & 0 \\ \operatorname{sen} \beta & 0 & \cos \beta \end{pmatrix} \quad (2.85)$$

y

$$\mathbf{R}_z(\gamma) = \begin{pmatrix} \cos \gamma & \operatorname{sen} \gamma & 0 \\ -\operatorname{sen} \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.86)$$

La rotación total es descrita por el producto matricial triple,

$$\mathbf{A}(\alpha, \beta, \gamma) = \mathbf{R}_z(\gamma)\mathbf{R}_y(\beta)\mathbf{R}_z(\alpha). \quad (2.87)$$

Notemos el orden:  $\mathbf{R}_z(\alpha)$  opera primero, entonces  $\mathbf{R}_y(\beta)$ , y finalmente  $\mathbf{R}_z(\gamma)$ . La multiplicación da

$$\mathbf{A} = \begin{pmatrix} \cos \gamma \cos \beta \cos \alpha - \operatorname{sen} \gamma \operatorname{sen} \alpha & \cos \gamma \cos \beta \operatorname{sen} \alpha - \operatorname{sen} \gamma \cos \alpha & -\cos \gamma \operatorname{sen} \beta \\ -\operatorname{sen} \gamma \cos \beta \cos \alpha - \cos \gamma \operatorname{sen} \alpha & -\operatorname{sen} \gamma \cos \beta \operatorname{sen} \alpha + \cos \gamma \cos \alpha & \operatorname{sen} \gamma \operatorname{sen} \beta \\ \operatorname{sen} \beta \cos \alpha & \operatorname{sen} \beta \operatorname{sen} \alpha & \cos \beta \end{pmatrix}. \quad (2.88)$$

Comparando  $\mathbf{A}(a_{ij})$  con  $\mathbf{A}(\alpha, \beta, \gamma)$  elemento por elemento, nos produce los cosenos directores en términos de los ángulos de Euler.

### Propiedades de simetría.

Nuestra descripción matricial conduce al grupo de rotaciones  $SO(3)$  en el espacio tridimensional  $\mathbb{R}^3$ , y la descripción en términos de ángulos de Euler de las rotaciones forman una base para desarrollar el grupo de rotaciones. Las rotaciones pueden también ser descritas por el grupo unitario  $SU(2)$  en el espacio bidimensional  $\mathbb{C}^2$ .

La matriz transpuesta es útil en la discusión de las propiedades de simetría. Si

$$\mathbf{A} = \tilde{\mathbf{A}}, \quad a_{ij} = a_{ji}, \quad (2.89)$$

la matriz es llamada *simétrica*, mientras que si

$$\mathbf{A} = -\tilde{\mathbf{A}}, \quad a_{ij} = -a_{ji}, \quad (2.90)$$

es llamada *antisimétrica*. Los elementos de la diagonal son nulos. Es fácil mostrar que cualquier matriz cuadrada puede ser escrita como la suma de una matriz simétrica y una antisimétrica. Consideremos la identidad

$$\mathbf{A} = \frac{1}{2} [\mathbf{A} + \tilde{\mathbf{A}}] + \frac{1}{2} [\mathbf{A} - \tilde{\mathbf{A}}]. \quad (2.91)$$

$\mathbf{A} + \tilde{\mathbf{A}}$  es claramente simétrica, mientras que  $\mathbf{A} - \tilde{\mathbf{A}}$  es claramente antisimétrica. Esta es la análoga matricial a la ecuación tensorial (1.68).

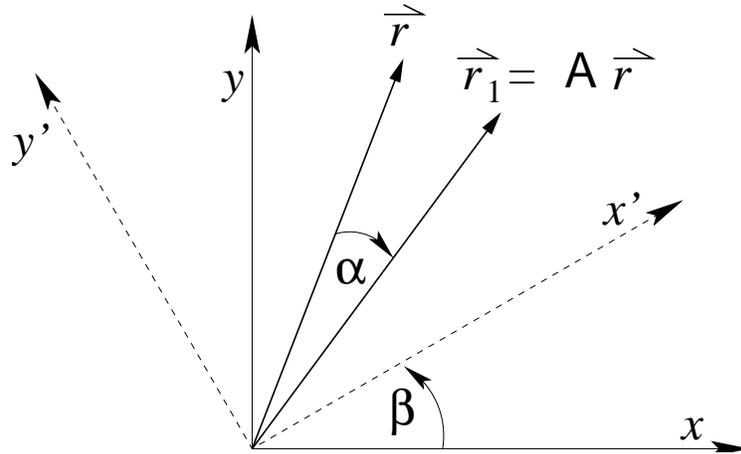


Figura 2.4: Vector fijo con coordenadas rotadas.

Hasta ahora hemos interpretado las matrices ortogonales como rotaciones del sistema de coordenadas. Estas cambian las componentes de un vector fijo. Sin embargo, una matriz ortogonal  $A$  puede ser interpretada igualmente bien como una rotación del vector en la dirección opuesta (figura 2.4).

Estas dos posibilidades, (1) rotar el vector manteniendo la base fija y (2) rotar la base (en el sentido opuesto) manteniendo el vector fijo.

Supongamos que interpretamos la matriz  $A$  como rotar un vector  $\vec{r}$  en una nueva posición  $\vec{r}_1$ , *i.e.*, en un particular sistema de coordenadas tenemos la relación

$$\vec{r}_1 = A\vec{r}. \quad (2.92)$$

Ahora rotemos las coordenadas aplicando una matriz  $B$ , la cual rota  $(x, y, z)$  en  $(x', y', z')$ ,

$$\begin{aligned} \vec{r}'_1 &= B\vec{r}_1 = BA\vec{r} = (A\vec{r})' \\ &= BA(B^{-1}B)\vec{r} \\ &= (BAB^{-1})B\vec{r} = (BAB^{-1})\vec{r}'. \end{aligned} \quad (2.93)$$

$B\vec{r}_1$  es justo  $\vec{r}'_1$  en el nuevo sistema de coordenadas con una interpretación similar se mantiene para  $B\vec{r}$ . Ya que en este nuevo sistema  $(B\vec{r})$  es rotado a la posición  $(B\vec{r}_1)$  por la matriz  $BAB^{-1}$ .

$$\begin{array}{ccc} B\vec{r}_1 & = & (BAB^{-1})B\vec{r} \\ \downarrow & & \downarrow \quad \downarrow \\ \vec{r}'_1 & = & A' \quad \vec{r}'. \end{array}$$

En el nuevo sistema las coordenadas han sido rotadas por la matriz  $B$ ,  $A$  tiene la forma  $A'$ , en la cual

$$A' = BAB^{-1}. \quad (2.94)$$

$A'$  opera en el espacio  $x', y', z'$  como  $A$  opera en el espacio  $x, y, z$ .

La transformación definida por la ecuación (2.94) con  $\mathbf{B}$  cualquier matriz, no necesariamente ortogonal, es conocida como transformación de similaridad. Por componentes la ecuación (2.94) llega a ser

$$a'_{ij} = \sum_{k,l} b_{ik} a_{kl} (\mathbf{B}^{-1})_{lj} . \quad (2.95)$$

Ahora si  $\mathbf{B}$  es ortogonal,

$$(\mathbf{B}^{-1})_{lj} = (\tilde{\mathbf{B}})_{lj} = b_{jl} , \quad (2.96)$$

y tenemos

$$a'_{ij} = \sum_{k,l} b_{ik} b_{jl} a_{kl} . \quad (2.97)$$

La matriz  $\mathbf{A}$  es la representación de un mapeo lineal en un sistema de coordenadas dado o base. Pero hay direcciones asociadas con  $\mathbf{A}$ , ejes cristalinos, ejes de simetría en un sólido rotando y etc. tal que la representación depende de la base. La transformación de similaridad muestran justo como la representación cambia con un cambio de base.

## 2.4 Matrices Hermíticas, matrices unitarias.

### Definiciones.

Hasta aquí hemos generalmente supuesto que nuestro espacio vectorial es un espacio real y que los elementos de las matrices (la representación de los operadores lineales) son reales. Para muchos cálculos en Física Clásica los elementos de matriz reales serán suficientes. Sin embargo, en Mecánica Cuántica las variables complejas son inevitables por la forma de las reglas de conmutación básicas (o la ecuación tiempo dependiente de Schödinger). Con esto en mente, generalizamos al caso de matrices con elementos complejos. Para manejar estos elementos, definamos algunas propiedades.

1. Compleja conjugada,  $\mathbf{A}^*$ , formada por tomar los complejos conjugados ( $i \rightarrow -i$ ) de cada elemento, donde  $i = \sqrt{-1}$ .
2. Adjunta,  $\mathbf{A}^\dagger$ , formada por transponer  $\mathbf{A}^*$ ,

$$\mathbf{A}^\dagger = \widetilde{\mathbf{A}^*} = \tilde{\mathbf{A}}^* . \quad (2.98)$$

3. Matriz hermítica: La matriz es etiquetada como hermítica (o autoadjunta) si

$$\mathbf{A} = \mathbf{A}^\dagger . \quad (2.99)$$

Si  $\mathbf{A}$  es real, entonces  $\mathbf{A}^\dagger = \tilde{\mathbf{A}}$ , y las matrices hermíticas reales son matrices reales y simétricas. En Mecánica Cuántica las matrices son hermíticas o unitarias.

4. Matriz unitaria: La matriz  $U$  es etiquetada como unitaria si

$$U^\dagger = U^{-1} . \quad (2.100)$$

Si  $U$  es real, entonces  $U^{-1} = \tilde{U}$ , tal que las matrices reales unitarias son matrices ortogonales. Este representa una generalización del concepto de matriz ortogonal.

5.  $(AB)^* = B^*A^*$ ,  $(AB)^\dagger = B^\dagger A^\dagger$ .

Si los elementos son complejos, a la Física casi siempre le interesan las matrices adjuntas, hermíticas y unitarias. Las matrices unitarias son especialmente importantes en Mecánica Cuántica porque ellos dejan el largo de un vector (complejo) inalterado, análoga a la operación de una matriz ortogonal sobre un vector real. Una importante excepción a este interés en las matrices unitarias es el grupo de matrices de Lorentz.

En un espacio  $n$ -dimensional complejo el cuadrado del largo de un punto  $\tilde{x} = (x_1, x_2, \dots, x_n)$ , o el cuadrado de su distancia al origen, es definido como  $x^\dagger x = \sum_i x_i^* x_i = \sum_i |x_i|^2$ . Si una transformación de coordenadas  $y = Ux$  deja la distancia inalterada, entonces  $x^\dagger x = y^\dagger y = (Ux)^\dagger Ux = x^\dagger U^\dagger Ux$ . Ya que  $x$  es arbitrario concluimos que  $U^\dagger U = \mathbf{1}_n$ , *i.e.*,  $U$  es una matriz unitaria de  $n \times n$ . Si  $x' = Ax$  es un mapa lineal, entonces su matriz en las nuevas coordenadas llega a ser una transformación unitaria (análogo de una de similaridad)

$$A' = UAU^\dagger ,$$

porque  $Ux' = y' = UAx = UAU^{-1}y = UAU^\dagger y$ .

### Matrices de Pauli y de Dirac.

El conjunto de tres matrices de Pauli de  $2 \times 2$   $\sigma$ ,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \quad (2.101)$$

fueron introducidas por W. Pauli para describir una partícula de spin  $\frac{1}{2}$  en Mecánica Cuántica no relativista. Se puede demostrar que las  $\sigma$  satisfacen

$$\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} \mathbf{1}_2 , \quad \text{anticonmutación} \quad (2.102)$$

$$\sigma_i \sigma_j = i\sigma_k , \quad \text{permutación cíclica de los índices} \quad (2.103)$$

$$(\sigma_i)^2 = \mathbf{1}_2 , \quad (2.104)$$

donde  $\mathbf{1}_2$  es la matriz unidad de  $2 \times 2$ . Así, el vector  $\vec{\sigma}/2$  satisface las mismas reglas de conmutación

$$[\sigma_i, \sigma_j] \equiv \sigma_i \sigma_j - \sigma_j \sigma_i = 2i\varepsilon_{ijk} \sigma_k , \quad (2.105)$$

que el momento angular  $\vec{L}$ .

Las tres matrices de Pauli  $\vec{\sigma}$  y la matriz unitaria forman un conjunto completo tal que cualquier matriz de  $2 \times 2$   $M$  puede ser expandida como

$$M = m_0 \mathbf{1} + m_1 \sigma_1 + m_2 \sigma_2 + m_3 \sigma_3 = m_0 \mathbf{1} + \vec{m} \cdot \vec{\sigma} , \quad (2.106)$$

donde los  $m_i$  son constantes. Usando  $\sigma_i^2 = 1$  y  $\text{tr}(\sigma_i) = 0$  nosotros obtenemos a partir de la ecuación (2.106) los coeficientes  $m_i$  de la expansión formando las trazas,

$$2m_0 = \text{tr}(\mathbf{M}) , \quad 2m_i = \text{tr}(\mathbf{M}\sigma_i) , \quad i = 1, 2, 3 . \quad (2.107)$$

En 1927 P.A.M. Dirac extendió este formalismo para partículas de spin  $\frac{1}{2}$  moviéndose a velocidades cercana a la de la luz tales como electrones Para incluir la relatividad especial su punto de partida es la ecuación de Einstein para la energía  $E^2 = \vec{p}^2 c^2 + m^2 c^4$  en vez de la energía cinética y potencial no relativista  $E = \vec{p}^2/2m + V$ . La clave para la ecuación de Dirac es factorizar

$$E^2 - \vec{p}^2 c^2 = E^2 - (c\vec{\sigma} \cdot \vec{p})^2 = (E - c\vec{\sigma} \cdot \vec{p})(E + c\vec{\sigma} \cdot \vec{p}) = m^2 c^4 , \quad (2.108)$$

usando la identidad matricial en  $2 \times 2$

$$(c\vec{\sigma} \cdot \vec{p})^2 = \vec{p}^2 \mathbf{1}_2 . \quad (2.109)$$

La matriz unidad de  $2 \times 2$   $\mathbf{1}_2$  no es escrita explícitamente en la ecuación (2.108) y (2.109). Podemos presentar las matrices  $\gamma_0$  y  $\gamma$  para factorizar  $E^2 - \vec{p}^2 c^2$  directamente,

$$(\gamma_0 E - \gamma c\vec{\sigma} \cdot \vec{p})^2 = \gamma_0^2 E^2 + \gamma^2 c^2 (\vec{\sigma} \cdot \vec{p})^2 - E c\vec{\sigma} \cdot \vec{p} (\gamma_0 \gamma + \gamma \gamma_0) = E^2 - \vec{p}^2 c^2 = m^2 c^4 . \quad (2.110)$$

Si reconocemos

$$\gamma_0 E - \gamma c\vec{\sigma} \cdot \vec{p} = \gamma \cdot p = (\gamma_0, \gamma \vec{\sigma}) \cdot (E, c\vec{p}) , \quad (2.111)$$

como el producto escalar de dos cuadvectores  $\gamma^\mu$  y  $p^\mu$ , entonces la ecuación (2.110) con  $p^2 = p \cdot p = E^2 - \vec{p}^2 c^2$  puede ser visto como una generalización cuadvectorial de la ecuación (2.109). Claramente, para que la ecuación (2.110) mantenga las condiciones

$$\gamma_0^2 = 1 = -\gamma^2 , \quad \gamma_0 \gamma + \gamma \gamma_0 = 0 , \quad (2.112)$$

debe satisfacerse que las cuatro matrices  $\gamma^\mu$  anticonmuten, justo como las tres matrices de Pauli. Ya que estas últimas son un conjunto completo de matrices anticonmutantes de  $2 \times 2$ , la condición (2.112) no puede ser satisfacerse para matrices de  $2 \times 2$ , pero ella puede ser satisfecha para matrices de  $4 \times 4$

$$\begin{aligned} \gamma_0 = \gamma^0 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \mathbf{1}_2 & 0 \\ 0 & -\mathbf{1}_2 \end{pmatrix} , \\ \gamma &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1}_2 \\ -\mathbf{1}_2 & 0 \end{pmatrix} . \end{aligned} \quad (2.113)$$

Alternativamente, el vector de matrices de  $4 \times 4$

$$\gamma = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} = \gamma \vec{\sigma} = \sigma_1 \times \vec{\sigma} , \quad (2.114)$$

puede obtenerse como el producto directo en el mismo sentido de la sección 2.2 de las matrices de  $2 \times 2$  de Pauli. De la misma manera,  $\gamma_0 = \sigma_3 \times \mathbf{1}_2$  y  $\mathbf{1}_4 = \mathbf{1}_2 \times \mathbf{1}_2$ .

Resumiendo el tratamiento no relativista de una partícula de spin  $\frac{1}{2}$ , produce matrices de  $4 \times 4$ , mientras que las partículas no relativistas de spin  $\frac{1}{2}$  son descritas por las matrices de Pauli  $\sigma$  de  $2 \times 2$ .

## 2.5 Diagonalización de matrices.

### Momento de la matriz de inercia .

En muchos problemas en Física que involucran matrices reales simétricas o complejas hermiticas es deseable llevar a cabo una real transformación de similaridad ortogonal o una transformación unitaria (correspondiente a una rotación del sistema de coordenadas) para reducir la matriz a una forma diagonal, con todos los elementos no diagonales nulos. Un ejemplo particularmente directo de esto es la matriz del momento de inercia  $I$  de un cuerpo rígido. A partir de la definición del momento angular  $\vec{L}$  tenemos

$$\vec{L} = I\vec{\omega} , \quad (2.115)$$

donde  $\vec{\omega}$  viene a ser la velocidad angular. La matriz de inercia  $I$  tiene elementos diagonales

$$I_{xx} = \sum_i m_i (r_i^2 - x_i^2) , \text{ y así sucesivamente,} \quad (2.116)$$

el subíndice  $i$  referencia la masa  $m_i$  localizada en  $\vec{r}_i = (x_i, y_i, z_i)$ . Para las componentes no diagonales tenemos

$$I_{xy} = - \sum_i m_i x_i y_i = I_{yx} . \quad (2.117)$$

Por inspección la matriz  $I$  es simétrica. También, ya que  $I$  aparece en una ecuación física de la forma (2.115), la cual se mantiene para todas las orientaciones del sistema de coordenadas, esta puede ser considerada un tensor (regla del cociente).

La clave ahora es la orientación de los ejes (a lo largo de un cuerpo fijo) tal que  $I_{xy}$  y los otros elementos no diagonales desaparezcan. Como una consecuencia de esta orientación y una indicación de ella, si la velocidad angular está a lo largo de tales realineados ejes, la velocidad angular y el momento angular serán paralelos.

### Autovectores y autovalores (eigenvector y eigenvalues).

Es quizás instructivo considerar un cuadro geométrico asociado a este problema. Si la matriz de inercia  $I$  es multiplicada a cada lado por un vector unitario cuya dirección es variable,  $\hat{n} = (\alpha, \beta, \gamma)$ , entonces en notación de Dirac

$$\langle \hat{n} | I | \hat{n} \rangle = I , \quad (2.118)$$

donde  $I$  es el momento de inercia respecto a la dirección  $\hat{n}$  y es un número positivo (escalar). Llevando a cabo la multiplicación, obtenemos

$$I = I_{xx}\alpha^2 + I_{yy}\beta^2 + I_{zz}\gamma^2 + 2I_{xy}\alpha\beta + 2I_{xz}\alpha\gamma + 2I_{yz}\beta\gamma . \quad (2.119)$$

Si introducimos

$$\vec{n} = \frac{\hat{n}}{\sqrt{I}} = (n_1, n_2, n_3) , \quad (2.120)$$

la cual es variable en dirección y magnitud entonces la ecuación (2.119) llega a ser

$$1 = I_{xx}n_1^2 + I_{yy}n_2^2 + I_{zz}n_3^2 + 2I_{xy}n_1n_2 + 2I_{xz}n_1n_3 + 2I_{yz}n_2n_3 , \quad (2.121)$$

una forma cuadrática positiva la cual debe ser un elipsoide (ver figura 2.5).

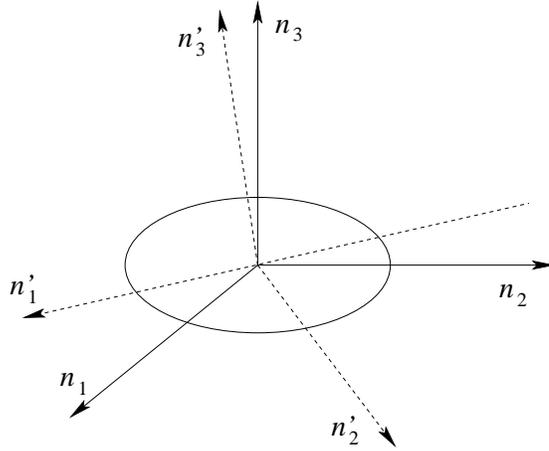


Figura 2.5: Elipsoide del momento de inercia.

A partir de la geometría analítica es sabido que los ejes de coordenadas pueden ser rotados para coincidir con los ejes de nuestro elipsoide. En muchos casos elementales, espacialmente cuando hay simetría, estos nuevos ejes, llamados ejes principales, pueden ser encontrados por inspección. Ahora nosotros procederemos a desarrollar un método general de hallazgo de los elementos diagonales y los ejes principales.

Si  $\mathbf{R}^{-1} = \tilde{\mathbf{R}}$  es la correspondiente matriz ortogonal real tal que  $\vec{n}' = \mathbf{R}\vec{n}$ , o  $|n'\rangle = \mathbf{R}|n\rangle$  en la notación de Dirac, son las nuevas coordenadas, luego obtenemos usando  $\langle n'|\mathbf{R} = \langle n|$  en la ecuación (2.121)

$$\langle n|\mathbf{I}|n\rangle = \langle n'|\mathbf{R}\tilde{\mathbf{R}}\mathbf{I}|n'\rangle = I'_1n_1'^2 + I'_2n_2'^2 + I'_3n_3'^2 , \quad (2.122)$$

donde los  $I'_i > 0$  son los momentos de inercia principales. La matriz de inercia  $I'$  en la ecuación (2.122) es diagonal en las nuevas coordenadas,

$$I' = \mathbf{R}\mathbf{I}\tilde{\mathbf{R}} = \begin{pmatrix} I'_1 & 0 & 0 \\ 0 & I'_2 & 0 \\ 0 & 0 & I'_3 \end{pmatrix} . \quad (2.123)$$

Si reescribimos la ecuación (2.123) usando  $R^{-1} = \tilde{R}$

$$\tilde{R}I' = I\tilde{R} , \quad (2.124)$$

y tomando  $\tilde{R} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$  compuesto de tres vectores columnas, entonces la ecuación (2.124) se separa en tres ecuaciones de autovalores

$$I\vec{v}_i = I'_i\vec{v}_i , \quad i = 1, 2, 3 , \quad (2.125)$$

con autovalores  $I'_i$  y autovectores  $\vec{v}_i$ . Como estas ecuaciones son lineales y homogéneas (para un  $i$  fijo), por la sección 2.1 los determinantes tienen que anularse:

$$\begin{vmatrix} I_{11} - I'_1 & I_{12} & I_{13} \\ I_{21} & I_{22} - I'_2 & I_{23} \\ I_{31} & I_{32} & I_{33} - I'_3 \end{vmatrix} = 0 . \quad (2.126)$$

Reemplazando los autovalores  $I'_i$  por una variable  $\lambda$  veces la matriz unidad  $\mathbf{1}$ , podríamos reescribir la ecuación (2.125) como

$$(\mathbf{I} - \lambda\mathbf{1})|v\rangle = 0 , \quad (2.127)$$

cuyo determinante

$$|\mathbf{I} - \lambda\mathbf{1}| = 0 , \quad (2.128)$$

es un polinomio cúbico en  $\lambda$ ; sus tres raíces, por supuesto, son los  $I'_i$ . Sustituyendo una raíz de regreso en la ecuación (2.125), podemos encontrar los correspondientes autovectores. La ecuación (2.126) (o la (2.128)) es conocida como la *ecuación secular*. El mismo tratamiento se aplica a una matriz simétrica real  $\mathbf{I}$ , excepto que sus autovalores no necesitan ser todos positivos. También, la condición de ortogonalidad en la ecuación (2.83a-2.83d) para  $\mathbf{R}$  dice que, en términos geométricos, los autovectores  $\vec{v}_i$  son vectores mutuamente ortogonales unitarios. Por cierto ellos forman las nuevas coordenadas del sistema. El hecho que cualquier par de autovectores  $\vec{v}_i, \vec{v}_j$  son ortogonales si  $I'_i \neq I'_j$  se deduce de la ecuación (2.125) en conjunción con la simetría de  $\mathbf{I}$  multiplicando con  $\vec{v}_i$  y  $\vec{v}_j$ , respectivamente,

$$\langle v_j | I | v_i \rangle = I'_i \langle v_j | v_i \rangle = \langle v_i | I | v_j \rangle = I'_j \langle v_j | v_i \rangle . \quad (2.129)$$

Ya que  $I'_i \neq I'_j$  y la ecuación (2.129) implica que  $(I'_i - I'_j) \vec{v}_i \cdot \vec{v}_j = 0$ , por lo tanto  $\vec{v}_i \cdot \vec{v}_j = 0$ .

### Matrices hermíticas.

Para espacios vectoriales complejos las matrices unitarias y hermíticas juegan el mismo rol como las matrices ortogonales y simétricas sobre los espacios vectoriales reales, respectivamente. Primero, generalicemos el importante teorema acerca de los elementos diagonales y los ejes principales para la ecuación de autovalores

$$A|r\rangle = \lambda|r\rangle . \quad (2.130)$$

Ahora mostramos que si  $\mathbf{A}$  es una matriz hermítica, sus autovalores son reales y sus autovectores ortogonales.

Sean  $\lambda_i$  y  $\lambda_j$  dos autovalores y  $|r_i\rangle$  y  $|r_j\rangle$ , los correspondientes autovectores de  $\mathbf{A}$ , una matriz hermítica. Entonces

$$\mathbf{A}|r_i\rangle = \lambda_i|r_i\rangle \quad (2.131)$$

$$\mathbf{A}|r_j\rangle = \lambda_j|r_j\rangle . \quad (2.132)$$

La ecuación (2.131) es multiplicada por  $|r_j\rangle$

$$\langle r_j|\mathbf{A}|r_i\rangle = \lambda_i\langle r_j|r_i\rangle . \quad (2.133)$$

La ecuación (2.132) es multiplicada por  $|r_i\rangle$  para dar

$$\langle r_i|\mathbf{A}|r_j\rangle = \lambda_j\langle r_i|r_j\rangle . \quad (2.134)$$

Tomando la adjunta conjugada de esta ecuación, tenemos

$$\langle r_j|\mathbf{A}^\dagger|r_i\rangle = \lambda_j^*\langle r_j|r_i\rangle \quad (2.135)$$

o

$$\langle r_j|\mathbf{A}|r_i\rangle = \lambda_j^*\langle r_j|r_i\rangle , \quad (2.136)$$

ya que  $\mathbf{A}$  es hermítica. Sustrayendo la ecuación (2.136) de la ecuación (2.133), obtenemos

$$(\lambda_i - \lambda_j^*)\langle r_j|r_i\rangle . \quad (2.137)$$

Este es un resultado general para todas las combinaciones posibles de  $i$  y  $j$ . Primero, sea  $j = i$ . Luego la ecuación (2.137) se convierte en

$$(\lambda_i - \lambda_i^*)\langle r_i|r_i\rangle = 0 . \quad (2.138)$$

Ya que  $\langle r_i|r_i\rangle = 0$  sería una solución trivial de la ecuación (2.138), concluimos que

$$\lambda_i = \lambda_i^* , \quad (2.139)$$

es decir,  $\lambda_i$  es real, para todo  $i$ .

Segundo, para  $i \neq j$  y  $\lambda_i \neq \lambda_j$ ,

$$(\lambda_i - \lambda_j)\langle r_i|r_j\rangle = 0 \quad (2.140)$$

o

$$\langle r_i|r_j\rangle = 0 \quad (2.141)$$

lo cual significa que los autovectores de distintos autovalores son ortogonales, la ecuación (2.141) siendo la generalización de ortogonalidad en este espacio complejo.

Si  $\lambda_i = \lambda_j$  (caso degenerado),  $\langle r_i|$  no es automáticamente ortogonal a  $|r_j\rangle$ , pero podría hacerse ortogonal. Consideremos el problema físico de la matriz del momento de inercia

nuevamente. Si  $x_i$  es un eje de simetría rotacional, entonces encontraremos que  $\lambda_2 = \lambda_3$ . Los autovectores  $|r_2\rangle$  y  $|r_3\rangle$  son cada uno perpendiculares al eje de simetría,  $|r_1\rangle$ , pero ellos yacen en alguna parte en el plano perpendicular a  $|r_1\rangle$ ; esto es, alguna combinación lineal de  $|r_2\rangle$  y  $|r_3\rangle$  es también un autovector. Considere  $(a_2|r_2\rangle + a_3|r_3\rangle)$  con  $a_2$  y  $a_3$  constantes. Entonces

$$\begin{aligned} A(a_2|r_2\rangle + a_3|r_3\rangle) &= a_2\lambda_2|r_2\rangle + a_3\lambda_3|r_3\rangle \\ &= \lambda_2(a_2|r_2\rangle + a_3|r_3\rangle) , \end{aligned} \quad (2.142)$$

como es esperado, para  $x_1$  un eje de simetría rotacional. Por lo tanto, si  $|r_1\rangle$  y  $|r_2\rangle$  son fijos,  $|r_3\rangle$ , puede simplemente escogerse yaciendo en el plano perpendicular a  $|r_1\rangle$  y también perpendicular a  $|r_2\rangle$ . Un método general para ortogonalizar soluciones conocido como proceso de Gram-Schmidt, es aplicado a funciones más adelante.

El conjunto de  $n$  autovectores ortogonales de nuestra matriz hermítica de  $n \times n$  forma un conjunto completo, generando el espacio de  $n$  dimensiones complejo. Este hecho es útil en un cálculo variacional de los autovalores. Los autovalores y los autovectores no están limitados a las matrices hermíticas. Todas las matrices tienen autovalores y autovectores. Por ejemplo, la matriz  $T$  de población estocástica satisface una ecuación de autovalores

$$T\vec{P}_{\text{equilibrio}} = \lambda\vec{P}_{\text{equilibrio}} ,$$

con  $\lambda = 1$ . Sin embargo, solamente las matrices hermíticas tienen todos los autovectores ortogonales y todos sus autovalores reales.

### Matrices antihermíticas.

Ocasionalmente, en Mecánica Cuántica encontramos matrices antihermíticas:

$$A^\dagger = -A .$$

Siguiendo el análisis de la primera porción de esta sección, podemos mostrar que

- Los autovalores son imaginarios puros (o cero).
- Los autovectores correspondientes a autovalores distintos son ortogonales.

La matriz  $R$  formada de los autovectores normalizados es unitaria. Esta propiedad antihermítica es preservada bajo transformaciones unitarias.

**Ejemplo:** Autovalores y autovectores de una matriz real simétrica.

Sea

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \quad (2.143)$$

La ecuación secular es

$$\begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = 0 , \quad (2.144)$$

o

$$-\lambda(\lambda^2 - 1) = 0, \quad (2.145)$$

expandiéndolo por las menores. Las raíces son  $\lambda = -1, 0, 1$ . Para encontrar el autovector correspondiente a  $\lambda = -1$ , sustituimos este valor de vuelta en la ecuación de autovalores, ecuación (2.130),

$$\begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (2.146)$$

Con  $\lambda = -1$ , esto produce

$$x + y = 0, \quad z = 0. \quad (2.147)$$

Dentro de un factor de escala arbitrario, y un signo arbitrario (factor de fase),  $\langle r_1 | = (1, -1, 0)$ . Notemos que (para el real  $|r\rangle$  en el espacio ordinario) el autovector define una línea en el espacio. El sentido positivo o negativo no está determinado. Esta indeterminación puede ser entendida si notamos que la ecuación (2.130) es homogénea en  $|r\rangle$ . Por conveniencia requeriremos que los autovectores estén normalizados a la unidad,  $\langle r_1 | r_1 \rangle = 1$ . Con esta elección de signo

$$\langle r_1 | = \vec{r}_1 = \left( \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0 \right), \quad (2.148)$$

está fijo. Para  $\lambda = 0$ , la ecuación (2.130) produce

$$y = 0, \quad x = 0, \quad (2.149)$$

$\langle r_2 |$  o  $\vec{r}_2 = (0, 0, 1)$  es un autovector aceptable. Finalmente, para  $\lambda = 1$ , tenemos

$$-x + y = 0, \quad z = 0, \quad (2.150)$$

o

$$\langle r_3 | = \vec{r}_3 = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right). \quad (2.151)$$

La ortogonalidad de  $\vec{r}_1$ ,  $\vec{r}_2$  y  $\vec{r}_3$ , correspondientes a los tres autovalores distintos, puede ser fácilmente verificada.

**Ejemplo:** Autovalores degenerados.

Consideremos

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \quad (2.152)$$

La ecuación secular es

$$\begin{vmatrix} 1 - \lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = 0 , \quad (2.153)$$

o

$$(1 - \lambda)(\lambda^2 - 1) = 0 , \quad \lambda = -1, 1, 1 , \quad (2.154)$$

un caso degenerado. Si  $\lambda = -1$ , la ecuación de autovalores (2.130) produce

$$2x = 0 , \quad y + z = 0 . \quad (2.155)$$

Un autovector normalizado adecuado es

$$\langle r_1 | = \vec{r}_1 = \left( 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) . \quad (2.156)$$

para  $\lambda = 1$ , tenemos

$$-y + z = 0 . \quad (2.157)$$

Cualquier autovector que satisface la ecuación (2.157) es perpendicular a  $\vec{r}_1$ . Tenemos infinito número de opciones. Tomemos una elección posible tomando

$$\langle r_2 | = \vec{r}_2 = \left( 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) , \quad (2.158)$$

la cual claramente satisface la ecuación (2.157). Entonces  $\vec{r}_3$  debe ser perpendicular a  $\vec{r}_1$  y puede ser escogido perpendicular a  $\vec{r}_2$  por<sup>7</sup>

$$\vec{r}_3 = \vec{r}_1 \times \vec{r}_2 = (1, 0, 0) . \quad (2.159)$$

### Funciones de matrices.

Polinomios con uno o más argumentos matriciales están bien definidos y ocurren a menudo. Series de potencias de una matriz también pueden estar definidas para dar la convergencia de la serie para cada uno de los elementos de matriz. Por ejemplo, si  $A$  es cualquiera matriz de  $n \times n$  entonces la serie de potencia

$$\exp(A) = \sum_{i=0}^{\infty} \frac{A^i}{i!} , \quad (2.160a)$$

$$\text{sen}(A) = \sum_{i=0}^{\infty} (-1)^i \frac{A^{2i+1}}{(2i+1)!} , \quad (2.160b)$$

$$\text{cos}(A) = \sum_{i=0}^{\infty} (-1)^i \frac{A^{2i}}{(2i)!} , \quad (2.160c)$$

---

<sup>7</sup>El uso del producto cruz es limitado a tres dimensiones.

son matrices de  $n \times n$  bien definidas. Para todas las matrices de Pauli  $\sigma_k$  la identidad de Euler para  $\theta$  real y  $k=1, 2$  o  $3$

$$\exp(i\sigma_k\theta) = \mathbf{1}_2 \cos(\theta) + i\sigma_k \text{sen}(\theta) , \quad (2.161)$$

sale a partir de coleccionar las potencias pares e impares en series separadas usando  $\sigma_k^2 = 1$ . Para las matrices de Dirac  $\sigma^{ij}$  de  $4 \times 4$  con  $(\sigma^{ij})^2 = 1$ , si  $j \neq k = 1, 2$  o  $3$ , obtenemos de manera similar (sin escribir las obvias matrices  $\mathbf{1}_4$  nunca más)

$$\exp(i\sigma^{jk}\theta) = \cos(\theta) + i\sigma^{jk} \text{sen}(\theta) , \quad (2.162)$$

mientras

$$\exp(i\sigma^{0k}\zeta) = \cosh(\zeta) + i\sigma^{0k} \text{senh}(\zeta) , \quad (2.163)$$

manteniendo  $\zeta$  real porque  $(i\sigma^{0k})^2 = 1$  para  $k = 1, 2$  o  $3$ .

Para una matriz hermítica  $\mathbf{A}$  hay una matriz unitaria  $\mathbf{U}$  que la diagonaliza, es decir,  $\mathbf{U}\mathbf{A}\mathbf{U}^\dagger = [a_1, a_2, \dots, a_n]$ . Entonces la *fórmula de la traza*

$$\det(\exp(\mathbf{A})) = \exp(\text{tr}(\mathbf{A})) \quad (2.164)$$

Puede ser fácilmente demostrado.

Otra importante relación es la de *fórmula de Baker-Hausdorff*

$$\exp(i\mathbf{G})\mathbf{H}\exp(-i\mathbf{G}) = \mathbf{H} + [i\mathbf{G}, \mathbf{H}] + \frac{[i\mathbf{G}, [i\mathbf{G}, \mathbf{H}]]}{2!} + \dots \quad (2.165)$$

lo cual resulta de multiplicar la serie de potencia para  $\exp(i\mathbf{G})$  y recolectar los términos de la misma potencia en  $i\mathbf{G}$ . Aquí definimos

$$[\mathbf{G}, \mathbf{H}] = \mathbf{GH} - \mathbf{HG}$$

como el conmutador de  $\mathbf{G}$  con  $\mathbf{H}$ .

## 2.6 Matrices normales.

En la sección 2.5 nos concentramos principalmente en matrices hermíticas o reales simétricas y en el proceso de encontrar autovalores y autovectores. En esta sección generalizaremos a matrices normales con matrices hermítica y unitario como casos especiales. Consideramos los casos físicamente importantes como el problema de los modos de vibraciones y el problema numérico importante de matrices patológicas.

Una matriz normal es una matriz que conmuta con su adjunta,

$$[\mathbf{A}, \mathbf{A}^\dagger] = 0 .$$

Ejemplos obvios e importante son las matrices hermíticas y unitarias. Mostraremos que las matrices normales tienen autovectores (ver tabla 2.1)

Matriz	Autovalores	Autovectores (para diferentes autovalores)
Hermítica	Real	Ortogonal
Antihermítica	Imaginaria puro (o cero)	Ortogonal
Unitaria	Magnitud uno	Ortogonal
Normal	Si $A$ tiene autovalor $\lambda$ $A^\dagger$ tiene autovalor $\lambda^*$	Ortogonal $A$ y $A^\dagger$ tienen los mismos autovectores

Tabla 2.1:

I. Sea  $|x\rangle$  un autovector de  $A$  con correspondiente autovalor  $\lambda$ . Entonces

$$A|x\rangle = \lambda|x\rangle \quad (2.166)$$

o

$$(A - \lambda 1)|x\rangle = 0 . \quad (2.167)$$

Por conveniencia la combinación  $A - \lambda 1$  la etiquetamos  $B$ . Tomando la adjunta de la ecuación (2.167), obtenemos

$$\langle x|(A - \lambda 1)^\dagger = 0 = \langle x|B^\dagger . \quad (2.168)$$

Porque

$$[(A - \lambda 1), (A - \lambda 1)^\dagger] = [A, A^\dagger] = 0 ,$$

tenemos

$$[B, B^\dagger] = 0 . \quad (2.169)$$

La matriz  $B$  es también normal.

A partir de las ecuaciones (2.167) y (2.168) formamos

$$\langle x|B^\dagger B|x\rangle = 0 . \quad (2.170)$$

Usando (2.169)

$$\langle x|BB^\dagger|x\rangle = 0 . \quad (2.171)$$

Ahora la ecuación (2.171) puede ser rescrita como

$$(B^\dagger|x\rangle)^\dagger(B^\dagger|x\rangle) = 0 . \quad (2.172)$$

Así

$$B^\dagger|x\rangle = (A^\dagger - \lambda^* 1)|x\rangle = 0 . \quad (2.173)$$

Vemos que para matrices normales,  $\mathbf{A}^\dagger$  tiene los mismos autovectores que  $\mathbf{A}$  pero los autovalores son los complejos conjugados.

II. Ahora, consideremos más que uno autovector-autovalor, tenemos

$$\mathbf{A}|x_i\rangle = \lambda_i|x_i\rangle, \quad (2.174)$$

$$\mathbf{A}|x_j\rangle = \lambda_j|x_j\rangle. \quad (2.175)$$

Multiplicando la ecuación (2.175) por la izquierda por  $\langle x_i|$  produce

$$\langle x_i|\mathbf{A}|x_j\rangle = \lambda_j\langle x_i|x_j\rangle. \quad (2.176)$$

Operando sobre el lado izquierdo de la ecuación (2.176), obtenemos

$$\langle x_i|\mathbf{A} = (\mathbf{A}^\dagger|x_i\rangle)^\dagger. \quad (2.177)$$

A partir de la ecuación (2.173) sabemos que  $\mathbf{A}^\dagger$  tiene los mismos autovectores que  $\mathbf{A}$  pero con los complejos conjugados de los autovalores

$$(\mathbf{A}^\dagger|x_i\rangle)^\dagger = (\lambda_i^*|x_i\rangle)^\dagger = \lambda_i\langle x_i|. \quad (2.178)$$

Sustituyendo en la ecuación (2.176) tenemos

$$\lambda_i\langle x_i|x_j\rangle = \lambda_j\langle x_i|x_j\rangle \quad (2.179)$$

o

$$(\lambda_i - \lambda_j)\langle x_i|x_j\rangle = 0. \quad (2.180)$$

Esta es la misma que la ecuación (2.140).

Para  $\lambda_i \neq \lambda_j$

$$\langle x_i|x_j\rangle = 0.$$

Los autovectores correspondientes a diferentes autovalores de una matriz normal son *ortogonales*. Esto significa que una matriz normal puede ser diagonalizada por una transformación unitaria. La matriz unitaria requerida puede ser construida a partir de los vectores ortonormales como se mostró en la sección anterior.

El converso también es válido. Si  $\mathbf{A}$  puede ser diagonalizada por una transformación unitaria, entonces  $\mathbf{A}$  es normal.

### Modos normales de vibración.

Consideremos las vibraciones de un modelo clásico de la molécula de  $\text{CO}_2$ . Esta es una ilustración de la aplicación de las técnicas matriciales a un problema que no parte como un problema de matrices. También provee un ejemplo de autovalores y autovectores de una matriz real asimétrica.

**Ejemplo:** Modos Normales.

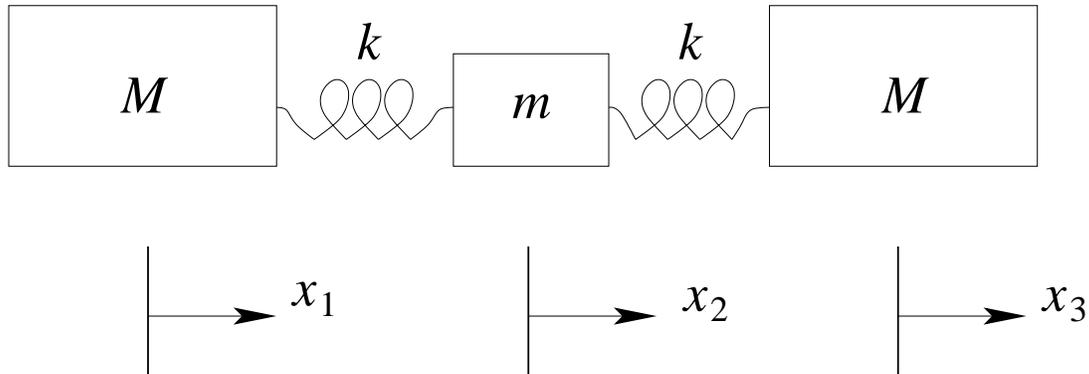


Figura 2.6: Vector fijo con coordenadas rotada.

Consideremos tres masas sobre el eje  $x$  unidas por resortes como muestra la figura 2.6. Las fuerzas de los resortes se suponen lineales (para pequeños desplazamientos, ley de Hooke) y las masas se restringen a mantenerse sobre el eje  $x$ .

Usando una coordenada diferente para cada masa la segunda ley de Newton produce el conjunto de ecuaciones

$$\begin{aligned}\ddot{x}_1 &= -\frac{k}{M}(x_1 - x_2) \\ \ddot{x}_2 &= -\frac{k}{M}(x_2 - x_1) - \frac{k}{m}(x_2 - x_3) \\ \ddot{x}_3 &= -\frac{k}{M}(x_3 - x_2) .\end{aligned}\tag{2.181}$$

El sistema de masa está vibrando. Buscamos las frecuencias comunes,  $\omega$  tal que todas las masas vibren en esta misma frecuencia. Estos son los modos normales. Sea

$$x_i = x_{i0}e^{i\omega t}, \quad i = 1, 2, 3.$$

Substituyendo en la ecuación (2.181), podemos escribir este conjunto como

$$\begin{pmatrix} \frac{k}{M} & -\frac{k}{M} & 0 \\ -\frac{k}{m} & \frac{2k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{M} & \frac{k}{M} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \omega^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{2.182}$$

dividiendo por el factor común  $e^{i\omega t}$ . Tenemos una ecuación matricial de autovalores con la matriz asimétrica. La ecuación secular es

$$\begin{vmatrix} \frac{k}{M} - \omega^2 & -\frac{k}{M} & 0 \\ -\frac{k}{m} & \frac{2k}{m} - \omega^2 & -\frac{k}{m} \\ 0 & -\frac{k}{M} & \frac{k}{M} - \omega^2 \end{vmatrix} = 0 . \tag{2.183}$$

Esto conduce a

$$\omega^2 \left( \frac{k}{M} - \omega^2 \right) \left( \omega^2 - \frac{2k}{m} - \frac{k}{M} \right) = 0$$

Los autovalores son

$$\omega^2 = 0, \quad \frac{k}{M}, \quad \text{y} \quad \frac{k}{M} + \frac{2k}{m},$$

todas reales.

Los correspondientes autovectores son determinados sustituyendo los autovalores de regreso en la ecuación (2.182) un autovalor a la vez. Para  $\omega^2 = 0$ , ecuación (2.182) produce

$$\begin{aligned} x_1 - x_2 &= 0 \\ -x_1 + 2x_2 - x_3 &= 0 \\ -x_2 + x_3 &= 0. \end{aligned}$$

Entonces, tenemos

$$x_1 = x_2 = x_3.$$

Esto describe una translación pura sin movimiento relativo de las masas y sin vibración.

Para  $\omega^2 = \frac{k}{M}$ , la ecuación (2.182) produce

$$x_1 = -x_3, \quad x_2 = 0. \quad (2.184)$$

Las masas exteriores se mueven en direcciones opuestas. El masa del centro está estacionaria.

Para  $\omega^2 = \frac{k}{M} + \frac{2k}{m}$ , las componentes de los autovectores son

$$x_1 = x_3, \quad x_2 = -\frac{2M}{m}x_1.$$

Las dos masas exteriores se están moviendo juntas. La masa del centro se está moviendo opuesta a las otras dos. El momentum neto es cero.

Cualquier desplazamiento de estas tres masas a lo largo del eje  $x$  puede ser descrito como una combinación lineal de estos tres tipos de movimiento: translación más dos formas de vibración.

### Sistemas con condiciones patológicas.

Un sistema lineal de ecuaciones puede ser escrito como

$$\mathbf{A}|x\rangle = |y\rangle \quad \text{o} \quad \mathbf{A}^{-1}|y\rangle = |x\rangle, \quad (2.185)$$

con  $\mathbf{A}$  y  $|y\rangle$  conocido y  $|x\rangle$  desconocido. Podemos encontrar ejemplos en los cuales un pequeño error en  $|y\rangle$  resulta en un gran error en  $|x\rangle$ . En este caso la matriz  $\mathbf{A}$  es llamada de condición

patológica. Si  $|\delta x\rangle$  es el error en  $|x\rangle$  y  $|\delta y\rangle$  es el error en  $|y\rangle$ , entonces los errores relativos pueden ser escritos como

$$\left[ \frac{\langle \delta x | \delta x \rangle}{\langle x | x \rangle} \right]^{1/2} \leq K(\mathbf{A}) \left[ \frac{\langle \delta y | \delta y \rangle}{\langle y | y \rangle} \right]^{1/2}. \quad (2.186)$$

Aquí  $K(\mathbf{A})$ , una propiedad de la matriz  $\mathbf{A}$ , es etiquetado la *condición de número*. Para  $\mathbf{A}$  hermítica una forma de la condición de número es dada por

$$K(\mathbf{A}) = \frac{|\lambda|_{\max}}{|\lambda|_{\min}}. \quad (2.187)$$

Una forma aproximada debido a Turing es

$$K(\mathbf{A}) = n[A_{ij}]_{\max}[A_{ij}^{-1}]_{\max}, \quad (2.188)$$

en la cual  $n$  es el orden de la matriz y  $[A_{ij}]_{\max}$  es el máximo elemento en  $\mathbf{A}$ .

**Ejemplo:** Una matriz patológica.

Un ejemplo común de una matriz con condición patológica es la matriz de Hilbert, la matriz de Hilbert de orden 4 es  $H_{ij} = (i + j - 1)^{-1}$ ,

$$H_4 = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}. \quad (2.189)$$

Los elementos de la matriz inversa (orden  $n$ ) son dados por

$$(H_n^{-1})_{ij} = \frac{(-1)^{i+j}}{i+j-1} \cdot \frac{(n+i-1)!(n+j-1)!}{[(i-1)!(j-1)!]^2(n-i)!(n-j)!}. \quad (2.190)$$

Para  $n = 4$

$$H_4^{-1} = \begin{pmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{pmatrix}. \quad (2.191)$$

A partir de la ecuación (2.188) la estimación de Turing de la condición de número para  $H_4$  llega a ser

$$K_{\text{Turing}} = 4 \times 1 \times 6480 \\ 2.59 \times 10^4.$$

Esto es una advertencia de que un error en la entrada puede ser multiplicado por 26000 en el cálculo del resultado de salida. Esto sentencia que  $H_4$  tiene condición patológica. Si usted encuentra un sistema altamente patológico tiene un par de alternativas (además de abandonar el problema).

- a. Tratar un ataque matemático diferente.
- b. Hacer arreglos para llevar más cifras significativas y a costa de fuerza bruta empujar de principio a fin.

# Capítulo 3

## Teoría de grupo.

versión final 1.2-090502<sup>1</sup>

*Disciplined judgment about what is neat  
and symmetrical and elegant has time and  
time again proved an excellent guide to  
how nature work.*

MURRAY GELL-MANN

### 3.1 Introducción.

En mecánica clásica la *simetría* de un sistema físico conduce a una *ley de conservación*. La conservación del momentum angular es una consecuencia directa de la simetría rotacional, lo cual significa *invariancia* bajo rotaciones espaciales. A principios del siglo pasado, Wigner y otros comprendieron que la invariancia era el concepto clave en el entendimiento de los nuevos fenómenos y en el desarrollo de teorías apropiadas. Así, en mecánica cuántica los conceptos de momento angular y spin han llegado a ser aún más centrales. Sus generalizaciones, el *isospin* en física nuclear y la simetría de *sabor* en física de partículas, son herramientas indispensables en la construcción teórica y en sus soluciones. Las generalizaciones del concepto de invariancia de *gauge* de la electrodinámica clásica para la simetría del isospin conduce a la teoría de gauge electrodébil.

En cada caso el conjunto de estas operaciones de simetría forman un *grupo*. La teoría de grupo es la herramienta matemática para tratar las invariancias y las simetrías. Ella trae consigo unificación y formalización de principios tales como reflexión espacial, o paridad, momento angular, y geometría que son ampliamente usados por los físicos.

En geometría el rol fundamental de la teoría de grupo fue reconocido hace mucho tiempo por los matemáticos. En geometría euclídeana la distancia entre dos puntos, el producto escalar de dos vectores o métrica, no cambia bajo rotaciones o translaciones. Estas simetrías son características de esta geometría. En relatividad especial la métrica, o producto escalar de cuadvectores, difiere del de la geometría euclídeana en que ya no es más positivo definido y es invariante ante transformaciones de Lorentz.

---

<sup>1</sup>Este capítulo está basado en el cuarto capítulo del libro: *Mathematical Methods for Physicists, fourth edition* de George B. Arfken & Hans J. Weber, editorial ACADEMIC PRESS.

Para un cristal el grupo de simetría contiene sólo un número finito de rotaciones en valores discretos del ángulo y reflexiones. La teoría de tales grupos discretos o finitos, desarrollada inicialmente como una rama de las matemáticas pura, ahora es una útil herramienta para el desarrollo de la cristalografía y la física de la materia condensada. Haremos una breve introducción a ellos. Cuando las rotaciones dependen de un ángulo continuo el grupo de rotaciones tiene un número infinito de elementos. Estudiaremos tales grupos continuos o de Lie.

### Definición de grupo.

Un grupo  $G$  puede ser definido como un conjunto de objetos u operaciones, llamados los elementos de  $G$ , que pueden ser combinados o “multiplicados” para formar un producto bien definido en  $G$  el cual satisface las siguientes cuatro condiciones.

1. Si  $a$  y  $b$  son cualquier par de elementos de  $G$ , entonces el producto  $ab$  es también elemento de  $G$ ; o  $(a, b) \rightarrow ab$  mapea  $G \times G$  sobre  $G$ .
2. Esta multiplicación es asociativa,  $(ab)c = a(bc)$ .
3. Hay un elemento unidad o neutro  $I$  en  $G$  tal que  $Ia = aI = a$  para cada elemento  $a$  de  $G$ .<sup>2</sup>
4. Debe haber un inverso o recíproco de cada elemento  $a$  de  $G$ , etiquetado  $a^{-1}$ , tal que  $aa^{-1} = a^{-1}a = I$ .

Un ejemplo para un grupo es el conjunto de rotaciones de coordenadas en el sentido del puntero del reloj,

$$R(\varphi) = \begin{pmatrix} \cos \varphi & \text{sen } \varphi \\ -\text{sen } \varphi & \cos \varphi \end{pmatrix} \quad (3.1)$$

en un ángulo  $\varphi$  del sistema de coordenadas  $xy$  a una nueva orientación. El producto de dos rotaciones  $R(\varphi_1)R(\varphi_2)$  es definida como una rotación primero en un ángulo  $\varphi_2$  y entonces en un ángulo  $\varphi_1$ . De acuerdo a la ecuación (2.29), esto corresponde al producto de las dos matrices ortogonales de  $2 \times 2$

$$\begin{pmatrix} \cos \varphi_1 & \text{sen } \varphi_1 \\ -\text{sen } \varphi_1 & \cos \varphi_1 \end{pmatrix} \begin{pmatrix} \cos \varphi_2 & \text{sen } \varphi_2 \\ -\text{sen } \varphi_2 & \cos \varphi_2 \end{pmatrix} = \begin{pmatrix} \cos(\varphi_1 + \varphi_2) & \text{sen}(\varphi_1 + \varphi_2) \\ -\text{sen}(\varphi_1 + \varphi_2) & \cos(\varphi_1 + \varphi_2) \end{pmatrix}, \quad (3.2)$$

usando las fórmulas de adición de funciones trigonométricas. El producto es claramente una rotación representada por una matriz ortogonal con un ángulo  $\varphi_1 + \varphi_2$ . El producto es la multiplicación asociativa de matrices. Es *conmutativo* o *abeliano* porque el orden en el cual esta rotaciones son realizadas no importa. El inverso de la rotación con ángulo  $\varphi$  es una con ángulo  $-\varphi$ . La unidad o neutro corresponde al ángulo  $\varphi = 0$ . El nombre del grupo es  $SO(2)$ , si el ángulo varía continuamente desde 0 a  $2\pi$ . Claramente,  $SO(2)$  tiene infinitos elementos. La unidad con ángulo  $\varphi = 0$  y la rotación con  $\varphi = \pi$  forman un *subgrupo* finito.

<sup>2</sup>También etiquetan al elemento unidad como  $E$ .

Un subgrupo  $G'$  de un grupo  $G$  consiste de elementos de  $G$  tal que el producto de cualquiera de sus elementos está de nuevo en el subgrupo  $G'$ , *i.e.*,  $G'$  es cerrado bajo la multiplicación de  $G$ . Si  $gg'g^{-1}$  es un elemento de  $G'$  para cualquier  $g$  de  $G$  y  $g'$  de  $G'$ , entonces  $G'$  es llamado un *subgrupo invariante* de  $G$ .

Las matrices ortogonales  $n \times n$  forman el grupo  $O(n)$ , y también  $SO(n)$  si sus determinantes son  $+1$  (S por eSpecial). Si  $\tilde{O}_i = O_i^{-1}$  para  $i = 1$  y  $2$ , entonces el producto

$$\widetilde{O_1 O_2} = \tilde{O}_2 \tilde{O}_1 = O_1^{-1} O_2^{-1} = (O_1 O_2)^{-1}$$

es también una matriz ortogonal en  $SO(n)$ . La inversa es la matriz (ortogonal) transpuesta. La unidad del grupo es  $1_n$ . Una matriz real ortogonal de  $n \times n$  tiene  $n(n-1)/2$  parámetros independientes. Para  $n = 2$  hay sólo un parámetro: un ángulo en la ecuación (3.1). Para  $n = 3$ , hay tres parámetros independientes: los tres ángulos de Euler de la sección 2.3.

De la misma manera, las matrices unitarias de  $n \times n$  forman el grupo  $U(n)$ , y también  $SU(n)$  si sus determinantes son  $+1$ . Si  $U_i^\dagger = U_i^{-1}$ , entonces

$$(U_1 U_2)^\dagger = U_2^\dagger U_1^\dagger = U_2^{-1} U_1^{-1} = (U_1 U_2)^{-1} ,$$

tal que el producto es unitario y un elemento de  $SU(n)$ . Cada matriz unitaria tiene una inversa la cual es también unitaria.

### Homomorfismo, isomorfismo.

Puede haber una correspondencia entre los elementos de dos grupos (o entre dos representaciones), uno a uno, dos a uno, o muchos a uno. Si esta correspondencia preserva la multiplicación del grupo, diremos que los dos grupos son *homomórficos*. Una de las más importantes correspondencias homomórficas entre el grupo de rotaciones  $SO(3)$  y el grupo de matrices unitarias  $SU(2)$  será desarrollado en la próxima sección. Si la correspondencia es uno a uno, y aún preserva la multiplicación del grupo,<sup>3</sup> entonces los grupos son *isomórficos*. Un ejemplo es las rotaciones de coordenadas a través de un ángulo finito  $\varphi$  en el sentido horario respecto al eje  $z$  en el espacio tridimensional descrito por

$$R_z(\varphi) = \begin{pmatrix} \cos \varphi & \text{sen } \varphi & 0 \\ -\text{sen } \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \quad (3.3)$$

El grupo de rotaciones  $R_z$  es isomórfico al grupo de rotaciones en la ecuación (3.1).

### Representaciones matriciales, reducibles e irreducibles.

La representación de los elementos de un grupo por matrices es una técnica muy poderosa y ha sido casi universalmente adoptada por los físicos. El uso de matrices no impone restricciones significativas. Puede mostrarse que los elementos de cualquier grupo finito y de grupos

---

<sup>3</sup>Supongamos que los elementos del primer grupo son etiquetados por  $g_i$  y los elementos del segundo grupo por  $h_i$ . Entonces  $g_i \leftrightarrow h_i$  en una correspondencia uno a uno para todos los valores de  $i$ . Si  $g_i g_j = g_k$  y  $h_i h_j = h_k$ , entonces  $g_k$  y  $h_k$  deben ser los elementos correspondientes del grupo.

continuos pueden ser representados por matrices. Ejemplos son las rotaciones descritas en la ecuación (3.1) y (3.3).

Para ilustrar como las representaciones matriciales surgen a partir de una simetría, consideremos la ecuación estacionaria de Schrödinger (o algún otra ecuación de autovalores tal como  $Iv_i = I_i v_i$  para los momentos principales de inercia de un cuerpo rígido en mecánica clásica)

$$H\psi = E\psi . \quad (3.4)$$

Supongamos que la ecuación (3.4) se mantiene invariante bajo la acción de un grupo  $G$  de transformaciones  $R$  en  $G$  (rotaciones de coordenadas, por ejemplo, para un potencial central  $V(r)$  en el Hamiltoniano  $H$ ), *i.e.*,

$$H_R = RHR^{-1} = H . \quad (3.5)$$

Ahora tomamos una solución  $\psi$  de la ecuación (3.4) y la “rotamos”:  $\psi \rightarrow R\psi$ . Entonces  $R\psi$  tiene la misma energía  $E$  porque multiplicando la ecuación (3.4) por  $R$  y usando (3.5) produce

$$RH\psi = E(R\psi) = (RHR^{-1})R\psi = H(R\psi) . \quad (3.6)$$

En otras palabras, todas las soluciones rotadas  $R\psi$  son *degeneradas* en energía o forman lo que los físicos llaman un *multiplete*. Supongamos que este espacio vectorial  $V_\psi$  de soluciones transformadas tiene una dimensión finita  $n$ . Sean  $\psi_1, \psi_2, \dots, \psi_n$  una base. Ya que  $R\psi_j$  es un miembro del multiplete, podemos expandirlo en términos de esta base

$$R\psi_j = \sum_k r_{jk} \psi_k . \quad (3.7)$$

Así, cada  $R$  en  $G$  puede ser asociado a una matriz  $(r_{jk})$ , y este mapeo  $R \rightarrow (r_{jk})$  es llamada una representación de  $G$ . Si podemos tomar cualquier elemento de  $V_\psi$  y por rotaciones con todos los elementos de  $R$  de  $G$  transforman en todos los otros elementos de  $V_\psi$  entonces la representación es *irreducible*. Si todos los elementos de  $V_\psi$  no son alcanzados, entonces  $V_\psi$  se separa en una suma directa de dos o más subespacios vectoriales,  $V_\psi = V_1 + V_2 + \dots$ , los cuales son mapeados dentro de ellos mismos por rotación de sus elementos. En este caso la representación es llamada *reducible*. Entonces podemos encontrar una base en  $V_\psi$  (*i.e.*, hay una matriz unitaria  $U$ ) tal que

$$U(r_{jk})U^\dagger = \begin{pmatrix} r_1 & 0 & \dots \\ 0 & r_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad (3.8)$$

para todos los  $R$  de  $G$  y todas las matrices  $(r_{jk})$ . Aquí  $r_1, r_2, \dots$ , son matrices de menor dimensiones que  $(r_{jk})$  que están alineadas a lo largo de la diagonal y los  $0$  son matrices de ceros. podemos decir que  $R$  ha sido descompuestas en  $r_1 + r_2 + \dots$  en paralelo con  $V_\psi = V_1 \oplus V_2 \oplus \dots$

Las representaciones irreducibles juegan un rol en teoría de grupo que es aproximadamente análogo a los vectores unitarios en el análisis vectorial. Ellas son las representaciones más simples, toda otra puede ser construida desde ellas.

## 3.2 Generadores de grupos continuos.

Una característica de los grupos continuos conocidos como grupos de Lie es que los parámetros de un producto de elementos son funciones analíticas de los parámetros de los factores. La naturaleza analítica de las funciones nos permite desarrollar el concepto de *generador* y reduce el estudio del grupo completo a un estudio de los elementos del grupo en la vecindad del elemento identidad.

La idea esencial de Lie fue el estudio de elementos  $R$  en un grupo  $G$  que estén infinitesimalmente cercanos a la unidad de  $G$ . Consideremos el grupo  $SO(2)$  como un ejemplo simple. Las matrices de rotación de  $2 \times 2$  en la ecuación (3.1) puede ser escrita en forma exponencial usando la identidad de Euler ecuación (2.161) como

$$R(\varphi) = \begin{pmatrix} \cos \varphi & \text{sen } \varphi \\ -\text{sen } \varphi & \cos \varphi \end{pmatrix} = \mathbf{1}_2 \cos \varphi + i\sigma_2 \text{sen } \varphi = \exp(i\sigma_2 \varphi) . \quad (3.9)$$

A partir de la forma exponencial es obvio que la multiplicación de estas matrices es equivalente a la suma de los argumentos

$$R(\varphi_2)R(\varphi_1) = \exp(i\sigma_2 \varphi_2) \exp(i\sigma_2 \varphi_1) = \exp(i\sigma_2(\varphi_1 + \varphi_2)) = R(\varphi_1 + \varphi_2) .$$

Por supuesto las rotaciones cercanas a  $\mathbf{1}$  tienen un ángulo pequeño  $\varphi \sim 0$ .

Esto sugiere que busquemos una representación exponencial

$$R = \exp(i\varepsilon S) , \quad \varepsilon \rightarrow 0 , \quad (3.10)$$

para elementos del grupo  $R \in G$  cercanos a la  $\mathbf{1}$ . Las transformaciones infinitesimales  $S$  son llamadas los *generadores* de  $G$ . Ellos forman un espacio lineal cuya dimensión es el *orden* de  $G$  porque la multiplicación de los elementos  $R$  del grupo se traduce en la suma de los generadores  $S$ .

Si  $R$  no cambia el elemento de volumen, *i.e.*,  $\det(R) = 1$ , nosotros usamos la ecuación (2.164) para ver que

$$\det(R) = \exp(\text{tr}(\ln R)) = \exp(i\varepsilon \text{tr}(S)) = 1$$

implica que los generadores son de traza nula,

$$\text{tr}(S) = 0 . \quad (3.11)$$

Este es el caso para el grupo de rotaciones  $SO(n)$  y el grupo unitario  $SU(n)$ , como veremos más adelante.

Si  $R$  de  $G$  en la ecuación (3.1) es unitario, entonces  $S^\dagger = S$  es hermítica, lo cual también es el caso para  $SO(n)$  y  $SU(n)$ . Ya que hay un  $i$  extra en la ecuación (3.10).

Expandamos los elementos del grupo

$$\begin{aligned} R_i &= \exp(i\varepsilon_i S_i) = \mathbf{1} + i\varepsilon_i S_i - \frac{1}{2}\varepsilon_i^2 S_i^2 + \dots , \\ R_i^{-1} &= \exp(-i\varepsilon_i S_i) = \mathbf{1} - i\varepsilon_i S_i - \frac{1}{2}\varepsilon_i^2 S_i^2 + \dots , \end{aligned} \quad (3.12)$$

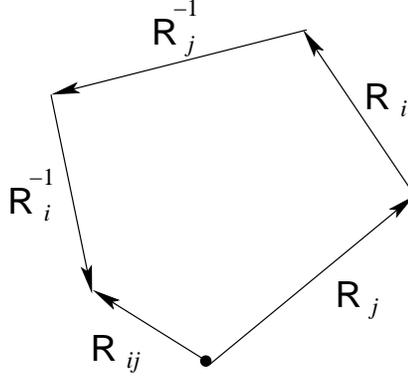


Figura 3.1: Ilustración de la ecuación (3.13).

a segundo orden en el pequeño parámetro del grupo  $\varepsilon_i$  porque los términos lineales y varios términos cuadráticos se cancelan en el producto (figura 3.1)

$$\begin{aligned} R_i^{-1} R_j^{-1} R_i R_j &= 1 + \varepsilon_i \varepsilon_j [S_j, S_i] + \dots, \\ &= 1 + \varepsilon_i \varepsilon_j \sum_k c_{ji}^k S_k + \dots, \end{aligned} \quad (3.13)$$

cuando las ecuaciones (3.12) son sustituidas dentro de la ecuación (3.13). La última línea es debido a que el producto en la ecuación (3.13) es nuevamente un elemento,  $R_{ij}$  cercano a la unidad en el grupo  $G$ . Por tanto su exponente debe ser una combinación lineal de los generadores  $S_k$  y sus parámetros infinitesimales del grupo tienen que ser proporcionales al producto  $\varepsilon_i \varepsilon_j$ . Comparando ambas líneas (3.13) encontramos la relación de *clausura* de los generadores del grupo de Lie  $G$ ,

$$[S_i, S_j] = \sum_k c_{ij}^k S_k \quad (3.14)$$

Los coeficientes  $c_{ij}^k$  son las *constantes de estructura* del grupo  $G$ . Ya que el conmutador en la ecuación (3.14) es antisimétrico en  $i$  y en  $j$ , también lo son las constantes de estructura en los índices inferiores,

$$c_{ij}^k = -c_{ji}^k. \quad (3.15)$$

Si el conmutador en la ecuación (3.14) es tomado como *la regla de multiplicación de los generadores*, vemos que el espacio vectorial de los generadores llega a ser un álgebra, el álgebra de Lie  $G$  del grupo  $G$ . Para  $SU(l+1)$  el álgebra de Lie es llamada  $A_l$ , para  $SO(2l+1)$  es  $B_l$  y para  $SO(2l)$  es  $D_l$ , donde  $l = 1, 2, \dots$  es un entero positivo, esto será llamado el *rango* de grupo de Lie  $G$  o de su álgebra  $G$ .

Finalmente, la *identidad de Jacobi* se satisface para los dobles conmutadores

$$[[S_i, S_j], S_k] + [[S_j, S_k], S_i] + [[S_k, S_i], S_j] = 0, \quad (3.16)$$

lo cual es fácilmente verificable usando la definición de conmutador. Cuando la ecuación (3.14) es substituida en (3.16) encontramos otra restricción sobre las constantes de estructura,

$$\sum_m \{c_{ij}^m [S_m, S_k] + c_{jk}^m [S_m, S_i] + c_{ki}^m [S_m, S_j]\} = 0 . \quad (3.17)$$

Usando de nuevo la ecuación (3.14) en la ecuación (3.17) implica que

$$\sum_{mn} \{c_{ij}^m c_{mk}^n S_n + c_{jk}^m c_{mi}^n S_n + c_{ki}^m c_{mj}^n S_n\} = 0 , \quad (3.18)$$

donde el factor común  $S_n$  (y la suma sobre  $n$ ) pueden eliminarse por que los generadores son linealmente independientes. Por tanto

$$\sum_m \{c_{ij}^m c_{mk}^n + c_{jk}^m c_{mi}^n + c_{ki}^m c_{mj}^n\} = 0 . \quad (3.19)$$

Las relaciones (3.14), (3.15) y (3.19) forman la base de las álgebras de Lie desde la cual los elementos finitos del grupo de Lie cerca de su unidad puede ser reconstruido.

Volviendo a la ecuación (3.5), el inverso de  $R$  es exactamente  $R^{-1} = \exp(-i\varepsilon S)$ . expandimos  $H_R$  de acuerdo a la fórmula de Baker-Haudorff, ecuación (2.17),

$$H = H_R = \exp(i\varepsilon S)H \exp(-i\varepsilon S) = H + i\varepsilon [S, H] - \varepsilon^2 \frac{[S, [iS, H]]}{2!} + \dots . \quad (3.20)$$

Al simplificar  $H$  de la ecuación (3.20), dividiendo por  $\varepsilon$  y haciendo  $\varepsilon \rightarrow 0$ . Entonces la ecuación (3.20) implica que para cualquier rotación cercana a 1 en  $G$  el conmutador

$$[S, H] = 0 . \quad (3.21)$$

Si  $S$  y  $H$  son matrices hermíticas, la ecuación (3.21) dice que  $S$  y  $H$  pueden ser simultaneamente diagonalizados. Si  $S$  y  $H$  son operadores diferenciales como el Hamiltoniano y el momento angular orbital en mecánica cuántica, entoces la ecuación (3.21) dice que  $S$  y  $H$  tienen autofunciones en común y que los autovalores degenerados de  $H$  pueden ser distinguidos por los autovalores de los generadores  $S$ . Esta es con mucho la más importante aplicación de teoría de grupos a mecánica cuántica.

A continuación, estudiaremos los grupos ortogonales y unitarios como ejemplos.

### Grupos de rotaciones $SO(2)$ y $SO(3)$ .

Para  $SO(2)$  definido por la ecuación (3.1) hay sólo un generador linealmente independiente,  $\sigma_2$  y el orden de  $SO(2)$  es 1. Obtenemos  $\sigma_2$  a partir de diferenciar la ecuación (3.9) y evaluarla en cero,

$$\left. -i \frac{dR(\varphi)}{d\varphi} \right|_{\varphi=0} = -i \begin{pmatrix} -\text{sen } \varphi & \text{cos } \varphi \\ -\text{cos } \varphi & -\text{sen } \varphi \end{pmatrix} \Big|_{\varphi=0} = -i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \sigma_2 . \quad (3.22)$$

Para las rotaciones  $R_z(\varphi)$  sobre el eje  $z$  descritas por la ecuación (3.3), el generador es dado por

$$-i \frac{dR(\varphi)}{d\varphi} \Big|_{\varphi=0} = S_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (3.23)$$

donde el factor extra  $i$  es insertado para hacer  $S_z$  hermítica. La rotación  $R_z(\delta\varphi)$  en un ángulo infinitesimal  $\delta\varphi$  puede ser escrita como

$$R_z(\delta\varphi) = \mathbf{1}_3 + i\delta\varphi S_z, \quad (3.24)$$

Una expansión de Maclaurin-Taylor de  $R_z$  cerca de la unidad  $\varphi = 0$  con términos hasta orden  $(\delta\varphi)^2$  y los superiores son despreciados. Una rotación finita puede ser compuesta por sucesivas rotaciones infinitesimales

$$R_z(\delta\varphi_1 + \delta\varphi_2) = (\mathbf{1}_3 + i\delta\varphi_1 S_z)(\mathbf{1}_3 + i\delta\varphi_2 S_z). \quad (3.25)$$

Sea  $\delta\varphi = \varphi/N$  para  $N$  rotaciones, con  $N \rightarrow \infty$ . Entonces,

$$R_z(\varphi) = \lim_{N \rightarrow \infty} \left[ \mathbf{1}_3 + i \frac{\varphi}{N} S_z \right]^N = \exp(iS_z). \quad (3.26)$$

Esta forma identifica  $S_z$  como el generador del grupo  $R_z$ , un subgrupo abeliano de  $SO(3)$ , el grupo de rotaciones en tres dimensiones con determinante  $+1$ . Cada matriz de  $3 \times 3$   $R_z(\varphi)$  es ortogonal, por lo tanto unitaria, y la  $\text{tr}(S_z) = 0$  de acuerdo con la ecuación (3.11).

Por diferenciación de las rotaciones de coordenadas

$$R_x(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & \text{sen } \psi \\ 0 & -\text{sen } \psi & \cos \psi \end{pmatrix}, \quad R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & -\text{sen } \theta \\ 0 & 1 & 0 \\ \text{sen } \theta & 0 & \cos \theta \end{pmatrix}, \quad (3.27)$$

obtenemos los generadores

$$S_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad S_y = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad (3.28)$$

de  $R_x$  y  $R_y$ , los subgrupos de rotaciones en torno a los ejes  $x$  e  $y$  respectivamente.

### Rotaciones de funciones y momento angular orbital.

En la discusión precedente los elementos del grupos son matrices que rotan las coordenadas. Cualquier sistema físico que esta siendo descrito se mantiene fijo. Ahora mantengamos fijas las coordenadas y rotemos una función  $\psi(x, y, z)$  relativa a nuestras coordenadas fijas. Con  $R$  para rotar las coordenadas,

$$\vec{x}' = R\vec{x}, \quad (3.29)$$

definimos  $R$  por

$$R\psi(x, y, z) = \psi'(x, y, z) \rightarrow \psi(\vec{x}') . \quad (3.30)$$

En palabras, la matriz  $R$  opera sobre la función  $\psi$ , creando una *nueva función*  $\psi'$  que es numéricamente igual a  $\psi(\vec{x}')$ , donde  $\vec{x}'$  son las coordenadas rotadas por  $R$ . Si  $R$  rota las coordenadas en el sentido horario, el efecto de la matriz  $R$  es rotar el modelo de la función  $\psi$  en el sentido horario.

Volviendo a las ecuaciones (3.3) y (3.28), consideremos una rotación infinitesimal,  $\varphi \rightarrow \delta\varphi$ . Luego, usando  $R_z$ , ecuación (3.3), obtenemos

$$R_z(\delta\varphi)\psi(x, y, z) = \psi(x + y\delta\varphi, y - x\delta\varphi, z) . \quad (3.31)$$

El lado derecho puede ser expandido como una serie de Taylor de primer orden en  $\delta\varphi$  para dar

$$\begin{aligned} R_z(\delta\varphi)\psi(x, y, z) &= \psi(x, y, z) - \delta\varphi \left\{ x \frac{\partial\psi}{\partial y} - y \frac{\partial\psi}{\partial x} \right\} + \mathcal{O}(\delta\varphi)^2 \\ &= (1 - i\delta\varphi L_z)\psi(x, y, z) , \end{aligned} \quad (3.32)$$

la expresión diferencial en el paréntesis de llave es  $iL_z$ . Ya que una rotación primero en  $\varphi$  y luego en  $\delta\varphi$  alrededor del eje  $z$  está dado por

$$R_z(\varphi + \delta\varphi)\psi(x, y, z) = R_z(\delta\varphi)R_z(\varphi)\psi(x, y, z) = (1 - i\delta\varphi L_z)R_z(\varphi)\psi(x, y, z) , \quad (3.33)$$

tenemos (como una ecuación de operadores)

$$\frac{R_z(\varphi + \delta\varphi) - R_z(\varphi)}{\delta\varphi} = -iL_z R_z(\varphi) . \quad (3.34)$$

El lado izquierdo es justo  $dR_z(\varphi)/d\varphi$  (para  $\delta\varphi \rightarrow 0$ ). En esta forma la ecuación (3.34) se integra inmediatamente a

$$R_z(\varphi) = \exp(-i\varphi L_z) . \quad (3.35)$$

Note cuidadosamente que  $R_z(\varphi)$  rota funciones (en el sentido horario) relativa a las coordenadas fijadas y que  $L_z$  es la componente  $z$  del momento angular orbital  $\vec{L}$ . La constante de integración está fijada por la condición de borde  $R_z(0) = 1$ .

Si reconocemos que los elementos de matriz

$$L_z = (x, y, z)S_z \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} , \quad (3.36)$$

claramente  $L_x, L_y, L_z$  satisface la misma relación de conmutación

$$[L_i, L_j] = i\varepsilon_{ijk}L_k \quad (3.37)$$

que  $S_x, S_y, S_z$  y tienen a la misma constantes de estructura  $i\varepsilon_{ijk}$  de  $SO(3)$ .

**Homomorfismo SU(2)-SO(3).**

El grupo *unitario especial* SU(2) de matrices unitarias de  $2 \times 2$  con determinante +1 tiene las tres matrices de Pauli  $\sigma_i$  como generadores (mientras que las rotaciones de la ecuación (3.3) forman un subgrupo abeliano unidimensional). Por lo tanto SU(2) es de orden 3 y depende de tres parámetros continuos reales  $\xi$ ,  $\eta$  y  $\zeta$  los cuales a menudo son llamados los parámetros de Caley-Klein. Sus elementos generales tienen la forma

$$U_2(\xi, \eta, \zeta) = \begin{pmatrix} e^{i\xi} \cos \eta & e^{i\zeta} \sen \eta \\ -e^{-i\zeta} \sen \eta & e^{-i\xi} \cos \eta \end{pmatrix} = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}. \quad (3.38)$$

Es fácil chequear que el determinante  $\det(U_2) = 1$  y que  $U_2^\dagger U_2 = 1 = U_2 U_2^\dagger$  se mantiene.

Para obtener los generadores diferenciamos

$$\begin{aligned} -i \frac{\partial U_2}{\partial \xi} \Big|_{\xi=0, \eta=0} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_3, \\ \frac{-i}{\sen \eta} \frac{\partial U_2}{\partial \zeta} \Big|_{\zeta=0} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1, \\ -i \frac{\partial U_2}{\partial \eta} \Big|_{\zeta=0, \eta=0} &= \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = \sigma_2. \end{aligned} \quad (3.39)$$

Por supuesto, las matrices de Pauli son todas de traza nula y hermíticas.

Con las matrices de Pauli como generadores de los elementos  $(U_1, U_2, U_3)$  de SU(2) pueden ser generados por

$$U_1 = \exp(ia_1 \sigma_1/2), \quad U_2 = \exp(ia_2 \sigma_2/2), \quad U_3 = \exp(ia_3 \sigma_3/2). \quad (3.40)$$

Los tres parámetros  $a_i$  son reales. El factor extra 1/2 está presente en los exponentes ya que  $s_i = \sigma_i/2$  satisface las mismas relaciones de conmutación <sup>4</sup>

$$[s_i, s_j] = i \varepsilon_{ijk} s_k \quad (3.41)$$

como el momento angular en la ecuación (3.37).

La ecuación (3.3) da un operador de rotación para rotar las coordenadas cartesianas en el espacio tridimensional. Usando la matriz de momento angular  $s_3$ , tenemos el correspondiente operador de rotación en el espacio de dos dimensiones (complejo)  $R_z(\varphi) = \exp(i\varphi \sigma_3/2)$ .

Para rotar una función de onda vectorial de dos componentes (spinor) o una partícula de spin 1/2 relativa a coordenadas fijas, el operador de rotación es  $R_z(\varphi) = \exp(-i\varphi \sigma_3/2)$  de acuerdo a la ecuación (3.35).

Usando la ecuación (3.40) la identidad de Euler, la ecuación (2.161), obtenemos

$$U_j = \cos\left(\frac{a_j}{2}\right) + i\sigma_j \sen\left(\frac{a_j}{2}\right).$$

---

<sup>4</sup>Las constantes de estructuras ( $i\varepsilon_{ijk}$ ) conducen a las representaciones de SU(2) de dimensión  $2J = 1$  para generadores de dimensión  $2j + 1$ , con  $J = 0, 1/2, 1, \dots$ . Los casos con  $J$  entero también conducen a las representaciones de SO(3).

Aquí el parámetro  $a_j$  aparece como un ángulo, el coeficiente de una matriz tipo momento angular  $\varphi$  en la ecuación (3.26). Con esta identificación de los exponenciales, la forma general de la matriz  $SU(2)$  (para rotar funciones más que las coordenadas) podría ser escrita como

$$U(\alpha, \beta, \gamma) = \exp(-i\gamma\sigma_3/2) \exp(-i\beta\sigma_2/2) \exp(-i\alpha\sigma_1/2) .$$

Como vimos, los elementos de  $SU(2)$  describen rotaciones en un espacio complejo bidimensional que deja invariante a  $|z_1|^2 + |z_2|^2$ . El determinante es  $+1$ . Hay tres parámetros independientes. Nuestro grupo ortogonal real  $SO(3)$  de determinante  $+1$ , claramente describe rotaciones comunes en el espacio tridimensional con la importante característica de dejar invariante a  $x^2 + y^2 + z^2$ . También hay tres parámetros independientes. Las interpretaciones de rotación y la igualdad de números de parámetros sugiere la existencia de alguna clase de correspondencia entre los grupos  $SU(2)$  y  $SO(3)$ . Aquí desarrollamos esta correspondencia.

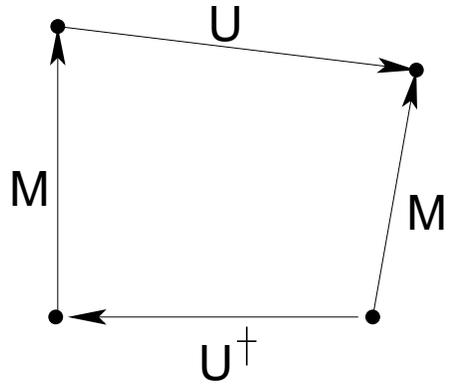


Figura 3.2: Ilustración de  $M' = UMU^\dagger$  ecuación (3.42).

La operación  $SU(2)$  sobre una matriz está dada por una transformación unitaria, la ecuación (3.5), con  $R = U$  y la figura (3.2)

$$M' = UMU^\dagger . \quad (3.42)$$

Tomando  $M$  como una matriz de  $2 \times 2$ , notemos que cualquier matriz de  $2 \times 2$  puede ser escrita como una combinación lineal de la matriz unidad y las tres matrices de Pauli. Sea  $M$  la matriz de traza cero,

$$M = x\sigma_1 + y\sigma_2 + z\sigma_3 = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} , \quad (3.43)$$

la matriz unidad no entra. Ya que la traza es invariante bajo transformaciones unitarias,  $M'$  debería tener la misma forma,

$$M = x'\sigma_1 + y'\sigma_2 + z'\sigma_3 = \begin{pmatrix} z' & x' - iy' \\ x' + iy' & -z' \end{pmatrix} . \quad (3.44)$$

El determinante también es invariante bajo una transformación unitaria. Por lo tanto

$$-(x^2 + y^2 + z^2) = -(x'^2 + y'^2 + z'^2) , \quad (3.45)$$

o  $(x^2 + y^2 + z^2)$  es invariante bajo esta operación de  $SU(2)$ , como con  $SO(3)$ .  $SU(2)$  debe, por lo tanto, describir una rotación. Esto sugiere que  $SU(2)$  y  $SO(3)$  pueden ser isomórficos o homomórficos.

Aproximemos el problema de qué rotación describe  $SU(2)$  considerando casos especiales. Retomando la ecuación (3.38) sea  $a = e^{i\xi}$  y  $b = 0$ , o

$$U_z = \begin{pmatrix} e^{i\xi} & 0 \\ 0 & e^{-i\xi} \end{pmatrix}. \quad (3.46)$$

En anticipación de la ecuación (3.50), esta  $U$  le está dado un subíndice  $z$ .

Realizando una transformación unitaria sobre cada una de las tres matrices de Pauli, tenemos

$$\begin{aligned} U_z \sigma_1 U_z^\dagger &= \begin{pmatrix} e^{i\xi} & 0 \\ 0 & e^{-i\xi} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{-i\xi} & 0 \\ 0 & e^{i\xi} \end{pmatrix} \\ &= \begin{pmatrix} 0 & e^{2i\xi} \\ e^{-2i\xi} & 0 \end{pmatrix}. \end{aligned} \quad (3.47)$$

Reexpresamos este resultado en términos de las matrices de Pauli para obtener

$$U_z x \sigma_1 U_z^\dagger = x \cos 2\xi \sigma_1 - x \sin 2\xi \sigma_2. \quad (3.48)$$

Similarmente,

$$\begin{aligned} U_z y \sigma_2 U_z^\dagger &= y \sin 2\xi \sigma_1 - y \cos 2\xi \sigma_2, \\ U_z z \sigma_3 U_z^\dagger &= z \sigma_3. \end{aligned} \quad (3.49)$$

A partir de esta expresión de doble ángulo vemos que podríamos comenzar con el ángulo medio:  $\xi = \alpha/2$ . Entonces, de las ecuaciones (3.42)–(3.44), (3.48) y (3.49),

$$\begin{aligned} x' &= x \cos \alpha + y \sin \alpha \\ y' &= -x \sin \alpha + y \cos \alpha \\ z' &= z. \end{aligned} \quad (3.50)$$

La transformación unitaria de  $2 \times 2$  usando  $U_z(\alpha/2)$  es equivalente al operador de rotación  $R(\alpha)$  de la ecuación (3.3).

El establecimiento de la correspondencia de

$$U_y(\beta/2) = \begin{pmatrix} \cos \beta/2 & \sin \beta/2 \\ -\sin \beta/2 & \cos \beta/2 \end{pmatrix} \quad (3.51)$$

y  $R_y(\beta)$  y de

$$U_x(\varphi/2) = \begin{pmatrix} \cos \varphi/2 & i \sin \varphi/2 \\ i \sin \varphi/2 & \cos \varphi/2 \end{pmatrix} \quad (3.52)$$

y  $R_x(\varphi)$  pueden calcularse como ejercicio. Podemos notar que  $U_k(\psi/2)$  tiene la forma general

$$U_k(\psi/2) = \mathbf{1} \cos \psi/2 + i \sigma_k \sin \psi/2, \quad (3.53)$$

donde  $k = x, y, z$ .

La correspondencia

$$U_z\left(\frac{\alpha}{2}\right) = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} \leftrightarrow \begin{pmatrix} \cos \alpha & \text{sen } \alpha & 0 \\ -\text{sen } \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = R_z(\alpha), \quad (3.54)$$

no es una simple correspondencia uno a uno. Específicamente, como  $\alpha$  en  $R_z$  recorre desde 0 a  $2\pi$ , el parámetro  $U_z$ ,  $\alpha/2$ , recorre desde 0 a  $\pi$ . Encontramos que

$$\begin{aligned} R_z(\alpha + 2\pi) &= R_z(\alpha) \\ U_z(\alpha/2 + \pi) &= \begin{pmatrix} -e^{i\alpha/2} & 0 \\ 0 & -e^{-i\alpha/2} \end{pmatrix} = -U_z(\alpha/2). \end{aligned} \quad (3.55)$$

Por lo tanto *ambos*  $U_z(\alpha/2)$  y  $U_z(\alpha/2 + \pi) = -U_z(\alpha/2)$  corresponde a  $R_z(\alpha)$ . La correspondencia es de 2 a 1, o SU(2) y SO(3) son homomórficos. Este establecimiento de la correspondencia entre las representaciones de SU(2) y de aquella SO(3) significa que las representaciones conocidas de SU(2) automáticamente nos proporciona de las representaciones de SO(3).

Combinando las rotaciones, encontramos que una transformación unitaria usada

$$U(\alpha, \beta, \gamma) = U_z(\gamma/2)U_y(\beta/2)U_z(\alpha/2), \quad (3.56)$$

corresponde a la rotación general de Euler  $R_z(\gamma)R_y(\beta)R_z(\alpha)$ . Por multiplicación directa,

$$\begin{aligned} U(\alpha, \beta, \gamma) &= \begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{pmatrix} \begin{pmatrix} \cos \beta/2 & \text{sen } \beta/2 \\ -\text{sen } \beta/2 & \cos \beta/2 \end{pmatrix} \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} \\ &= \begin{pmatrix} e^{i(\gamma+\alpha)/2} \cos \beta/2 & e^{i(\gamma-\alpha)/2} \text{sen } \beta/2 \\ -e^{-i(\gamma-\alpha)/2} \text{sen } \beta/2 & e^{-i(\gamma+\alpha)/2} \cos \beta/2 \end{pmatrix}. \end{aligned} \quad (3.57)$$

Esta es nuestra forma general alternativa, la ecuación (3.38), con

$$\varepsilon = \frac{(\gamma + \alpha)}{2}, \quad \eta = \frac{\beta}{2}, \quad \zeta = \frac{(\gamma - \alpha)}{2}. \quad (3.58)$$

De la ecuación (3.57) podemos identificar los parámetros de la ecuación (3.38) como

$$\begin{aligned} a &= e^{i(\gamma+\alpha)/2} \cos \beta/2 \\ b &= e^{i(\gamma-\alpha)/2} \text{sen } \beta/2 \end{aligned} \quad (3.59)$$

### SU(2) isospin y el octeto SU(3).

En el tratamiento de las partículas con interacciones fuertes de Física nuclear y de altas energías que conducen al grupo de SU(2) de isospin y la simetría de sabor SU(3), podríamos mirar el momento angular y el grupo de rotación SO(3) por una analogía. Supongamos que tenemos un electrón en un potencial atractivo esféricamente simétrico de algún núcleo atómico. La función de onda para el electrón puede ser caracterizada por tres números

cuánticos  $n, l, m$ , que están relacionados con los autovalores de los operadores conservados  $H, L^2, L_z$ . La energía,<sup>5</sup> sin embargo, es  $2l + 1$  veces degenerada, dependiendo solamente de  $n$  y  $l$ . La razón para esta degeneración puede ser expresado de dos maneras equivalentes:

1. El potencial es simétricamente esférico, independiente de  $\theta, \varphi$ .
2. El hamiltoniano de Schrodinger  $-(\hbar^2/2m_e)\nabla^2 + V(r)$  es invariante bajo rotaciones espaciales ordinarias  $SO(3)$ .

Como una consecuencia de la simetría esférica del potencial  $V(r)$ , el momento angular orbital  $\vec{L}$  es conservado. En la sección 3.2 las componentes cartesianas de  $\vec{L}$  están indentificadas como los generadores del grupo de rotación  $SO(3)$ . En vez de representar  $L_x, L_y, L_z$  por operadores, usaremos matrices. Las matrices  $L_i$  son matrices  $(2l + 1) \times (2l + 1)$  con la misma dimensión del número de estados degenerados. La dimensión  $2l + 1$  está identificada con los estados degenerados  $2l + 1$ .

Esta degeneración es removida por un campo magnético  $\vec{B}$ , hecho conocido como el efecto Zeeman. Esta interacción magnética añade un término al Hamiltoniano que *no* es invariante bajo  $SO(3)$ . Este es un término quiebra la simetría.

En el caso de partículas con interacción fuerte (protones, neutrones, etc.) no podemos seguir la analogía directamente, ya que todavía no entendemos completamente las interacciones nucleares. La fuerza fuerte está descrita por la teoría gauge de Yang-Mills basada sobre la simetría de color  $SU(3)$  llamada cromodinámica cuántica o abreviada QCD. Sin embargo, QCD es una teoría no lineal y por lo tanto complicada a grandes distancias y baja energía que permanece no resuelta. Por lo tanto, no conocemos el Hamiltoniano, en vez de esto, volveremos a la analogía.

En los años 1930, después del descubrimiento del neutrón, Heisenberg propuso que las fuerzas nucleares eran cargas independientes. Los neutrones difieren en masa de los protones solamente en un 1.6%. Si esta pequeña diferencia es ignorada, el neutrón y el protón podrían ser consideradas como dos estados de cargas (o isospin) de un doblete, llamado nucleón. El isospin  $I$  tiene proyección en el eje  $z$   $I_3 = 1/2$  para el protón y  $I_3 = -1/2$  para el neutrón. El isospin no tiene nada que ver con el spin (el momento angular intrínseco de una partícula) pero las dos componentes del estado de isospin obedece las mismas relaciones matemáticas que el estado de spin  $1/2$ . Para el nucleón,  $I = \tau/2$ , son las matrices usuales de Pauli y los estados de isospin  $(\pm 1/2)$  son autovectores de la matriz de Pauli  $\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . Similarmente, los tres estados de carga del pión  $\pi^+, \pi^0, \pi^-$  forman un triplete. El pión es la partícula más liviana con interacción fuerte y es la mediadora de la fuerza nuclear a distancia, como el fotón es partícula que media la fuerza electromagnética. La interacción fuerte trata igualmente a miembros de esa familia de partículas, o multipletes y conserva el isospin. La simetría es el grupo isospin  $SU(2)$ .

El octuplete mostrado en la tabla 3.1 llama la atención<sup>6</sup>. Los números cuánticos conservados que son análogos y generalizaciones de  $L_z$  y  $L^2$  de  $SO(3)$  son  $I_3$  e  $I^2$  para el isospin, e  $Y$  para *hipercarga*. Las partículas pueden ser agrupadas dentro de multipletes de carga o de

<sup>5</sup>Para un potencial de Coulomb puro la energía depende sólo de  $n$ .

<sup>6</sup>Todas las masas están dadas en unidades de energía.

		Masa [MeV]	$Y$	$I$	$I_3$
$\Xi$	$\Xi^-$	1321.32	-1	$\frac{1}{2}$	$-\frac{1}{2}$
	$\Xi^0$	1314.9			$+\frac{1}{2}$
$\Sigma$	$\Sigma^-$	1197.43	0	1	-1
	$\Sigma^0$	1192.55			0
	$\Sigma^+$	1189.37			+1
$\Lambda$	$\Lambda$	1115.63	0	0	0
$N$	$n$	939.566	1	$\frac{1}{2}$	$-\frac{1}{2}$
	$p$	938.272			$+\frac{1}{2}$

Tabla 3.1: Bariones con spin  $\frac{1}{2}$  y paridad par

isospin. Entonces la hipercarga puede ser tomada como dos veces el promedio de carga del multiplete. Para el nucleón, *i.e.*, el doblete neutrón-protón,  $Y = 2 \cdot \frac{1}{2}(0+1) = 1$ . Los valores de la hipercarga y los del isospin son listados en la tabla 3.1 para bariones como el nucleón y sus compañeros (aproximadamente degenerados). Ellos forman un octeto como lo muestra la figura 3.3. En 1961 Gell-Mann, e independientemente Ne'eman, sugirieron que la interacción fuerte debe ser (aproximadamente) invariante bajo un grupo espacial tridimensional unitario, SU(3), esto es, tienen *simetría de sabor* SU(3).

La elección de SU(3) estuvo basada primero sobre los dos números cuánticos conservados e independientes  $H_1 = I_3$  y  $H_2 = Y$  (*i.e.*, generados con  $[I_3, Y] = 0$ ), que llaman para un grupo de rango 2. Segundo, el grupo ha tenido una representación de ocho dimensiones para tener en cuenta los cercanamente degenerados bariones y cuatro octetos similares para los mesones. En un sentido SU(3) es la generalización más simple del isospin SU(2). Tres de sus generadores son matrices hermíticas de  $3 \times 3$  de traza nula que contienen las matrices de Pauli de  $2 \times 2$  para los isospin  $\tau_i$  en la esquina superior izquierda.

$$\lambda_i = \begin{pmatrix} \tau_i & 0 \\ 0 & 0 \end{pmatrix}, \quad i = 1, 2, 3. \quad (3.60)$$

De este modo, el grupo del isospin SU(2) es un subgrupo de SU(3) con  $I_3 = \lambda_3/2$ . Otros cuatro generadores tienen los no diagonales 1 de  $\tau_1$  e  $i$ ,  $-i$  de  $\tau_2$  en todas las otras posibles

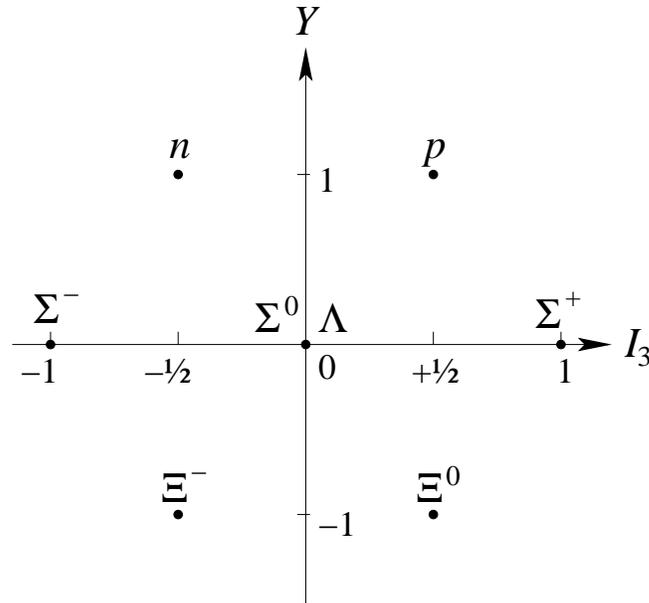


Figura 3.3: Octeto bariónico diagrama de peso para  $SU(3)$ .

ubicaciones para formar las matrices hermíticas  $3 \times 3$  de traza nula.

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad (3.61)$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

El segundo generador diagonal tiene la matriz unidad bidimensional  $\mathbf{1}_2$  en la esquina superior izquierda, la cual la hace claramente independiente del subgrupo  $SU(2)$  isospin ya que su traza no nula en el subespacio, y -2 en el lugar de la tercera diagonal la hace traza nula,

$$\lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad (3.62)$$

Generalmente hay  $3^2 - 1 = 8$  generadores para  $SU(3)$  el cual tiene orden 8. De los conmutadores de esos generadores pueden obtenerse fácilmente las constantes de estructura de  $SU(3)$ .

Volviendo a la simetría de sabor  $SU(3)$  imaginemos que el Hamiltoniano para nuestro octeto de bariones están compuesto de tres partes

$$H = H_{\text{fuerte}} + H_{\text{medio}} + H_{\text{electromagnético}}. \quad (3.63)$$

La primera parte,  $H_{\text{fuerte}}$ , tiene la simetría  $SU(3)$  y conduce a la degeneración ocho. La introducción del término de quiebre de simetría,  $H_{\text{medio}}$ , remueve parte de la degeneración

dando los cuatro multipletes del isospin  $(\Xi^-, \Xi^0)$ ,  $(\Sigma^-, \Sigma^0, \Sigma^+)$ ,  $\Lambda$ , y  $N = (p, n)$  con diferentes masas. Estos aún son multipletes ya que  $H_{\text{medio}}$  tiene la simetría del isospin SU(2). Finalmente, la presencia de fuerzas dependientes de la carga separan los multipletes de isospin y remueve la última degeneración. Esta secuencia se muestra en la figura 3.4

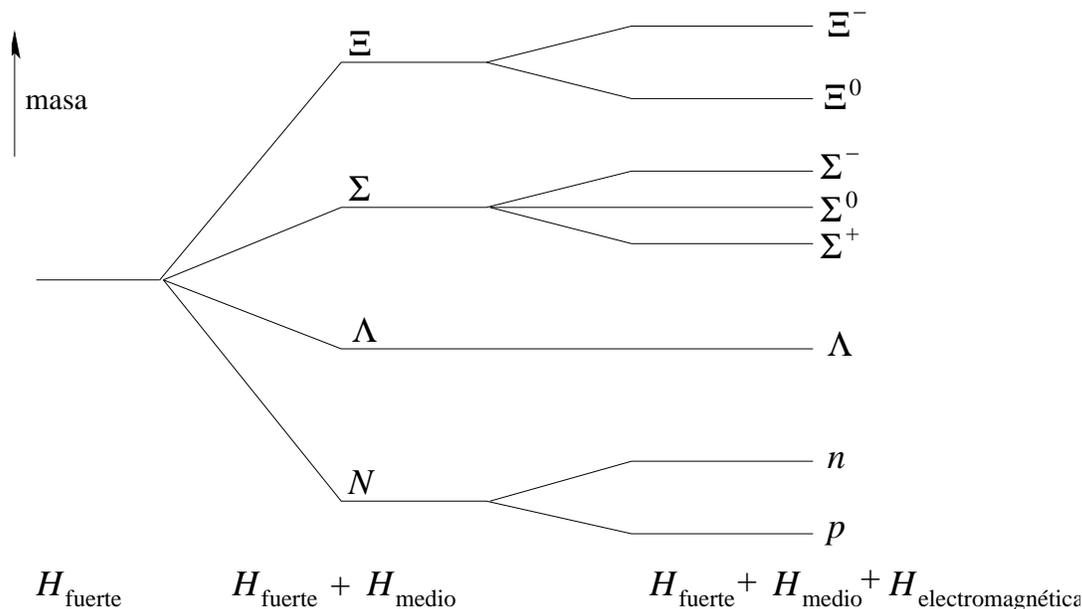


Figura 3.4: Separación de masa bariónica.

Aplicando teoría de perturbación de primer orden de Mecánica Cuántica, relaciones simples de masas de bariónicas pueden ser calculadas. Quizás el suceso más espectacular de este modelo SU(3) ha sido su predicción de nuevas partículas. En 1961 cuatro mesones  $K$  y tres  $\pi$  (todos pseudoescalares; spin 0, paridad impar) sugieren otro octeto, similar al del octeto bariónico. SU(3) predice un octavo mesón  $\eta$ , de masa 563 MeV. El mesón  $\eta$  con una masa determinada experimentalmente de 548 MeV fue encontrado poco después. Agrupamientos de nueve de los bariones más pesados (todos con spin 3/2, con paridad par) sugirió un multiplete de 10 miembros o un decaplete de SU(3). El décimo barión faltante fue predicho con una masa cercana a 1680 MeV y una carga negativa. En 1964  $\Omega^-$  cargada negativamente con masa  $(1675 \pm 12)$  MeV fue descubierta.

La representación de octeto no es la más simple para SU(3). La representación más simple son triangulares como se muestra en la figura 3.5 a partir de las cuales todas las otras pueden ser generadas por acoplamiento del momento angular generalizado. La representación fundamental en la figura 3.5 (a) contiene los quark u (arriba) y d (abajo) y el s (extrañeza), y figura 3.5 (b) los correspondientes antiquarks. Ya que los octetos de mesones pueden ser obtenidos a partir de la representación de quark como pares  $q\bar{q}$ ,  $3^2 = 8 + 1$ , esto sugiere que los mesones contienen quarks (y antiquarks) como sus constituyentes. El modelo de quarks resultante dan una exitosa descripción de la espectroscopía hadrónica. La solución de sus problemas con el principio de exclusión de Pauli eventualmente conduce a la teoría de gauge de SU(3)-color de las interacciones fuertes llamada cromodinámica cuántica o QCD.

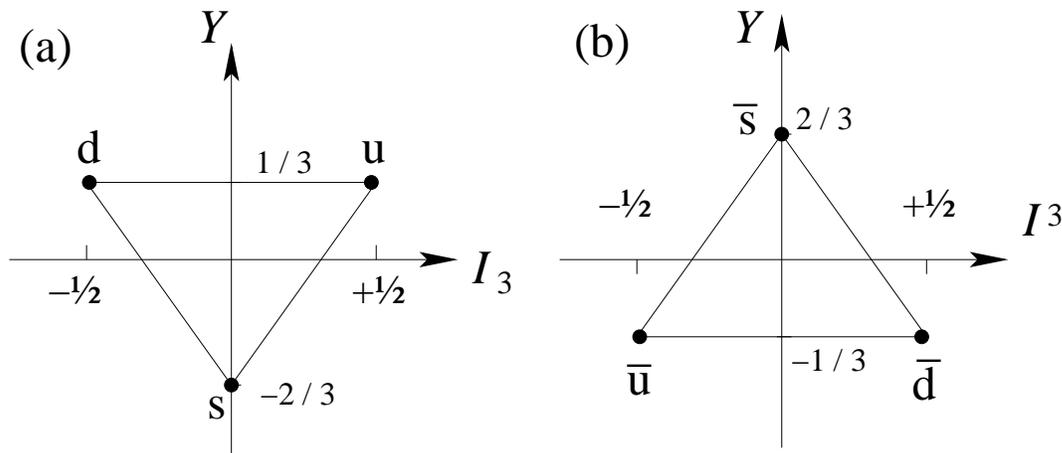


Figura 3.5: Separación de masa bariónica.

Para mantener la teoría de grupo en su real perspectiva, podríamos enfatizar que la teoría de grupo identifica y formaliza las simetrías. Ella clasifica partículas (y algunas veces predice). Pero a parte de decir que una parte del hamiltoniano tiene simetría  $SU(2)$  y otra parte tiene simetría  $SU(3)$ , la teoría de grupo no dice nada a cerca de la interacción de las partículas. Recuerde que la afirmación de que el potencial atómico es esféricamente simétrico no nos dice nada a cerca de la dependencia radial del potencial o de su función de onda. En contraste, en una teoría de gauge la interacción es mediada por bosones vectoriales (como el fotón media en la electrodinámica cuántica) y determinado únicamente por la derivada covariante de gauge.

### 3.3 Momento angular orbital.

El concepto clásico de momento angular  $\vec{L}_{\text{clásico}} = \vec{r} \times \vec{p}$  es mostrado en el capítulo de vectores para presentar el producto cruz. Siguiendo con la representación usual de Schrödinger de la Mecánica Cuántica, el momento lineal clásico  $\vec{p}$  es reemplazado por el operador  $-i\vec{\nabla}$ . El operador de momento angular en la mecánica cuántica se convierte en <sup>7</sup>

$$\vec{L}_{QM} = -i\vec{r} \times \vec{\nabla}. \quad (3.64)$$

Las componentes del momento angular satisfacen las relaciones de conmutación

$$[L_i, L_j] = i\varepsilon_{ijk}L_k. \quad (3.65)$$

El  $\varepsilon_{ijk}$  es el símbolo de Levi-Civita. Una suma sobre el índice  $k$  es sobreentendida.

El operador diferencial correspondiente al cuadrado del momento angular

$$\vec{L}^2 = \vec{L} \cdot \vec{L} = L_x^2 + L_y^2 + L_z^2, \quad (3.66)$$

<sup>7</sup>Por simplicidad,  $\hbar = 1$ . Esto significa que el momento angular es medido en unidades de  $\hbar$ .

puede ser determinado a partir de

$$\vec{L} \cdot \vec{L} = (\vec{r} \times \vec{p}) \cdot (\vec{r} \times \vec{p}) , \quad (3.67)$$

la cual puede verificarse como ejercicio. Ya que  $\vec{L}^2$  es un escalar rotacional,  $[\vec{L}^2, L_i] = 0$ , el cual también puede ser verificado directamente.

La ecuación (3.65) presenta las relaciones de conmutación básicas de los componentes del momento angular en Mecánica Cuántica. Por cierto, dentro del marco de la Mecánica Cuántica y la teoría de grupo, estas relaciones de conmutación definen un operador de momento angular.

### Acercamiento a los operadores de subida y bajada.

Comencemos con una aproximación más general, donde el momento angular  $\vec{J}$  lo consideramos que puede representar un momento angular orbital  $\vec{L}$ , un spin  $\sigma/2$ , o un momento angular total  $\vec{L} + \sigma/2$ , etc. Supongamos que

1.  $J$  es un operador hermítico cuyas componentes satisfacen las relaciones de conmutación

$$[J_i, J_j] = i\varepsilon_{ijk}J_k , \quad [\vec{J}^2, J_i] = 0 . \quad (3.68)$$

2.  $|\lambda M\rangle$  es simultáneamente una autofunción normalizada (o autovector) de  $J_z$  con autovalor  $M$  y una autofunción de  $\vec{J}^2$ ,

$$J_z|\lambda M\rangle = M|\lambda M\rangle , \quad \vec{J}^2|\lambda M\rangle = \lambda|\lambda M\rangle . \quad (3.69)$$

Mostraremos ahora que  $\lambda = J(J+1)$ . Este tratamiento ilustrará la generalidad y potencia de las técnicas de operadores particularmente el uso de operadores de subida y bajada.

Un operador de subida o bajada se defina como

$$J_+ = J_x + iJ_y , \quad J_- = J_x - iJ_y . \quad (3.70)$$

En términos de ese operador  $\vec{J}^2$  puede ser reescrito como

$$\vec{J}^2 = \frac{1}{2}(J_+J_- + J_-J_+) + J_z^2 . \quad (3.71)$$

A partir de las relaciones de conmutación, encontramos que

$$[J_z, J_+] = +J_+ , \quad [J_z, J_-] = -J_- , \quad [J_+, J_-] = 2J_z . \quad (3.72)$$

Ya que  $J_+$  conmuta con  $\vec{J}^2$ , hagalo como ejercicio

$$\vec{J}^2(J_+|\lambda M\rangle) = J_+(\vec{J}^2|\lambda M\rangle) = \lambda(J_+|\lambda M\rangle) . \quad (3.73)$$

Por lo tanto,  $J_+|\lambda M\rangle$  todavía es una autofunción de  $\vec{J}^2$  con autovalores  $\lambda$ , y similarmente para  $J_-|\lambda M\rangle$ . Pero de la ecuación (3.72)

$$J_zJ_+ = J_+(J_z + 1) , \quad (3.74)$$

o

$$J_z(J_+|\lambda M\rangle) = J_+(J_z + 1)|\lambda M\rangle = (M + 1)(J_+|\lambda M\rangle) . \quad (3.75)$$

Por lo tanto,  $J_+|\lambda M\rangle$  todavía es una autofunción de  $J_z$  con autovalores  $M + 1$ .  $J_+$  ha elevado el autovalor en 1 y por eso es llamado *operador de subida*. Similarmente,  $J_-$  baja los autovalores en 1 y a menudo es llamado *operador de bajada*.

Tomando los valores esperados y usando  $J_x^\dagger = J_x$ ,  $J_y^\dagger = J_y$ ,

$$\langle \lambda M | \vec{J}^2 - J_z^2 | \lambda M \rangle = \langle \lambda M | J_x^2 + J_y^2 | \lambda M \rangle = |J_x|\lambda M\rangle|^2 + |J_y|\lambda M\rangle|^2 ,$$

vemos que  $\lambda - M^2 \geq 0$ , tal que  $M$  es ligado. Sea  $J$  el más grande valor de  $M$ . Luego  $J_+|\lambda J\rangle = 0$ , lo cual implica que  $J_-J_+|\lambda J\rangle = 0$ . Luego combinando las ecuaciones (3.71) y (3.72) obtenemos

$$\vec{J}^2 = J_-J_+ + J_z(J_z + 1) , \quad (3.76)$$

encontramos que a partir de la ecuación (3.76)

$$0 = J_-J_+|\lambda M = J\rangle = (\vec{J}^2 - J_z^2 - J_z)|\lambda M = J\rangle = (\lambda - J^2 - J)|\lambda M = J\rangle .$$

Por lo tanto

$$\lambda = J(J + 1) \geq 0; \quad (3.77)$$

con un  $J$  no negativo. Ahora reetiquetaremos los estados  $|\lambda M\rangle = |JM\rangle$ . Similarmente, sea  $J'$  el más pequeño de los  $M$ . Entonces  $J_-|JJ'\rangle = 0$ . A partir de

$$\vec{J}^2 = J_+J_- - J_z(J_z + 1) , \quad (3.78)$$

vemos que

$$0 = J_+J_-|JJ'\rangle = (\vec{J}^2 + J_z - J_z^2)|JJ'\rangle = (\lambda + J' - J'^2)|JJ'\rangle . \quad (3.79)$$

De manera que

$$\lambda = J(J + 1) = J'(J' - 1) = (-J)(-J - 1) .$$

Así  $J' = -J$ , y  $M$  corre en pasos enteros desde  $-J$  a  $+J$ ,

$$-J \leq M \leq +J . \quad (3.80)$$

Comenzando desde  $|JJ\rangle$  y aplicando  $J_-$  repetidas veces, alcanzaremos todos los otros estados  $|JM\rangle$ . De manera que  $|JM\rangle$  forma una representación irreducible;  $M$  varía y  $J$  está fijo.

Entonces usando las ecuaciones (3.68), (3.76) y (3.78) obtenemos

$$\begin{aligned} J_-J_+|JM\rangle &= [J(J + 1) - M(M + 1)]|JM\rangle = (J - M)(J + M + 1)|JM\rangle , \\ J_+J_-|JM\rangle &= [J(J + 1) - M(M - 1)]|JM\rangle = (J + M)(J - M + 1)|JM\rangle . \end{aligned} \quad (3.81)$$

Como  $J_+$  y  $J_-$  son hermíticos conjugados,

$$J_+^\dagger = J_- \quad , \quad J_-^\dagger = J_+ \quad , \quad (3.82)$$

los autovalores o valores esperados en la ecuación (3.81) deberían ser positivos o cero.

Ya que  $J_+$  aumenta el autovalor de  $M$  a  $M + 1$ , reetiquetaremos la autofunción resultante  $|JM + 1\rangle$ . La normalización está dada por la ecuación (3.81) como

$$J_+|JM\rangle = \sqrt{(J - M)(J + M + 1)}|JM + 1\rangle \quad , \quad (3.83)$$

tomando la raíz cuadrada positiva y no introduciendo ningún factor de fase. Por los mismos argumentos

$$J_-|JM\rangle = \sqrt{(J + M)(J - M + 1)}|JM - 1\rangle \quad . \quad (3.84)$$

Finalmente, ya que  $M$  va desde  $-J$  a  $+J$  en pasos unitarios,  $2J$  debería ser un número entero.  $J$  es por lo tanto un entero o la mitad de un entero impar. Como hemos visto, el momento angular orbital está descrito con  $J$  entero. A partir de los spin de algunas partículas fundamentales y de algunos núcleos, tenemos que  $J = 1/2, 3/2, 5/2, \dots$ . Nuestro momento angular está cuantizado esencialmente como un resultado de relaciones de conmutaciones. En coordenadas polares esféricas  $\theta, \varphi$  las funciones  $\langle \theta, \varphi | lm \rangle = Y_l^m(\theta, \varphi)$  son armónicos esféricos.

### Resumen de grupos y álgebras de Lie.

Las relaciones de conmutaciones generales, ecuación (3.14) en la sección 3.2, para un grupo de Lie clásico [SO(n) y SU(n) en particular] pueden ser simplificadas para verse más como la ecuación (3.72) para SO(3) y SU(2) en la sección 3.3.

Primero escogemos generadores  $H_i$  que sean linealmente independientes y que conmuten entre sí, estos son generalizaciones de  $J_z$  de SO(3) y SU(2). Sea  $l$  el número máximo de tales  $H_i$  con

$$[H_i, H_k] = 0 \quad . \quad (3.85)$$

Entonces  $l$  es llamado el *rango* del grupo Lie G o de su álgebra. El rango y la dimensión u orden de algunos grupos Lie son dados en la tabla 3.2. Todos los otros generadores  $E_\alpha$  puede mostrarse que son operadores de subida y bajada con respecto a todos los  $H_i$ , tal que

$$[H_i, E_\alpha] = \alpha_i E_\alpha \quad . \quad (3.86)$$

El conjunto de los  $(\alpha_1, \alpha_2, \dots, \alpha_l)$  son llamados los *vectores raíces*.

Ya que los  $H_i$  conmutan entre sí, ellos pueden ser diagonalizados simultáneamente. Ellos nos dan un conjunto de autovalores  $m_1, m_2, \dots, m_l$ . Al conjunto  $(m_1, m_2, \dots, m_l)$  se les llama vectores de peso de una representación irreducible. Hay  $l$  operadores invariantes  $C_i$ , llamados *operadores de Casimir*, los cuales conmutan con todos los generadores y son generalizaciones de  $J^2$ ,

$$[C_i, H_j] = 0 \quad , \quad [C_i, E_\alpha] = 0 \quad . \quad (3.87)$$

Álgebra de Lie	$A_l$	$B_l$	$D_l$
Grupo de Lie	SU(1+1)	SO(2l+1)	SO(2l)
rango	1	1	1
orden	l(1+2)	l(2l+1)	l(2l-1)

Tabla 3.2: Rango y orden de grupos rotacionales y unitarios.

El primero,  $C_1$ , es una función cuadrática de los generadores, los otros son más complicados. Ya que  $C_i$  conmuta con todos los  $H_j$ , ellos pueden ser diagonalizados simultáneamente con los  $H_j$ . Sus autovalores  $c_1, c_2, \dots, c_l$  caracterizan las representaciones irreducibles y permanecen constantes mientras los vectores de peso varían sobre una representación irreducible particular. Por lo tanto, la autofunción general puede ser escrita como

$$|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle, \quad (3.88)$$

generalizando  $|JM\rangle$  de SO(3) y SU(2). Sus ecuaciones de autovalores son

$$H_i|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle = m_i|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle, \quad (3.89a)$$

$$C_i|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle = c_i|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle. \quad (3.89b)$$

Ahora podemos mostrar que  $E_\alpha|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle$  tiene el vector peso  $(m_1 + \alpha_1, m_2 + \alpha_2, \dots, m_l + \alpha_l)$  usando las relaciones de conmutación, la ecuación (3.86), en conjunto con las ecuaciones (3.89a) y (3.89b),

$$\begin{aligned} H_i E_\alpha|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle &= (E_\alpha H_i + [H_i, E_\alpha])|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle \\ &= (m_i + \alpha_i)E_\alpha|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle. \end{aligned} \quad (3.90)$$

Por lo tanto

$$E_\alpha|(c_1, c_2, \dots, c_l)m_1, m_2, \dots, m_l\rangle \sim |(c_1, c_2, \dots, c_l)m_1 + \alpha_1, m_2 + \alpha_2, \dots, m_l + \alpha_l\rangle$$

estas son las generalizaciones de las ecuaciones (3.83) y (3.84) a partir de SO(3). Esos cambios de autovalores por el operador  $E_\alpha$  son llamados sus reglas de selección en mecánica cuántica.

### 3.4 Grupo homogéneo de Lorentz.

En relatividad especial requerimos que nuestras leyes físicas sean covariantes<sup>8</sup> bajo

- traslaciones en el tiempo y en el espacio,
- rotaciones en el espacio real tridimensional, y
- transformaciones de Lorentz.

---

<sup>8</sup>Ser covariante significa que tienen la misma forma en diferentes sistemas de coordenadas tal que no hay un sistema de referencia privilegiado.

El requerimiento para la covarianza bajo traslaciones está basada en la homogeneidad del espacio y el tiempo. Covarianza bajo rotaciones es una afirmación de la isotropía del espacio. El requerimiento de la covarianza de Lorentz viene de la relatividad especial. Todas estas tres transformaciones en conjunto forman el grupo inhomogeneo de Lorentz o el grupo de Poincaré. Aquí excluimos las traslaciones. Las rotaciones espaciales y las transformaciones de Lorentz forman un grupo, el grupo homogéneo ed Lorentz.

Primero generamos un subgrupo, las transformaciones de Lorentz en el cual la velocidad relativa  $\vec{v}$  está a lo largo del eje  $x = x_1$ . El generador puede ser determinad considerando un marco de referencia espacio-temporal moviendose con una velocidad relativa infinitesimal  $\delta v$ . Las relaciones son similares a aquellas para rotaciones en el espacio real, excepto que aquí el ángulo de rotación es imaginario puro.

Las transformaciones de Lorentz son lineales no sólo en el espacio de ccordenadas  $x_i$  si no que también en el tiempo  $t$ . Ellas se originan a partir de las ecuaciones de Maxwell de la electrodinámica las cuales son invariantes bajo la transformaciones de Lorentz, como veremos luego. Las transformaciones de Lorentz dejan invariante la forma cuadrática siguiente  $c^2t^2 - x_1^2 - x_2^2 - x_3^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2$  donde  $x_0 = ct$ . Vemos esto si encendemos una fuente de luz en el origen del sistema de coordenadas. En tiempo  $t$  la luz ha viajado una distancia  $ct = \sqrt{\sum x_i^2}$ , tal que  $c^2t^2 - x_1^2 - x_2^2 - x_3^2 = 0$ . La relatividad especial requiere esto en todos los sistemas (inercial) que se mueven con velocidad  $v \leq c$  en cualquier dirección relativa al sistema  $x_i$  y que tengan el mismo origen a tiempo  $t = 0$ , se mantenga también que  $c^2t'^2 - x_1'^2 - x_2'^2 - x_3'^2 = 0$ . El espacio cuadridimensional con la métrica  $x_0^2 - x_1^2 - x_2^2 - x_3^2$  es llamado espacio de Minkowski con el producto escalar de dos cuadvectores definido como  $a \cdot b = a_0b_0 - \vec{a} \cdot \vec{b}$ . Usando el tensor métrico

$$(g_{\mu\nu}) = (g^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (3.91)$$

podemos subir y bajar índices de un cuadvector tal como de las coordenadas  $x^\mu = (x_0, \vec{x})$  es decir  $x_\mu = g_{\mu\nu}x^\nu = (x_0, -\vec{x})$  y  $x^\mu g_{\mu\nu}x^\nu = x_0^2 - \vec{x}^2$ , la convección de suma de Einstein se dan por entendida. Para el gradiente  $\partial^\mu = (\partial/\partial x_0, -\vec{\nabla}) = \partial/\partial x_\mu$  y  $\partial_\mu = (\partial/\partial x_0, \vec{\nabla})$  tal que  $\partial^2 = \partial^2/\partial x_0^2 - \nabla^2$  es un escalar de Lorentz, al igual que la métrica  $x_0^2 - \vec{x}^2$ .

Para  $v \ll c$ , en el límite no relativista, las transformaciones de Lorentz deben ser transformaciones de Galileo. Por lo tanto, para derivar la forma de una transformación de Lorentz a lo largo del eje  $x_1$ , partimos con una transformación Galileana para una velocidad relativa infinitesimal  $\delta v$ :

$$x'_1 = x_1 - \delta vt = x_1 - x_0 \delta \beta. \quad (3.92)$$

Como es usual  $\beta = \frac{v}{c}$ . Por simetría también podemos escribir

$$x'_0 = x_0 - ax_1 \delta \beta, \quad (3.93)$$

donde  $a$  es un parámetro a fijar una vez que se imponga que  $x_0^2 - x_1^2$  deba ser invariante,

$$x_0'^2 - x_1'^2 = x_0^2 - x_1^2. \quad (3.94)$$

Recordemos que  $x = (x_0; x_1, x_2, x_3)$  es el prototipo de cuadrivector en el espacio de Minkowski. Así la ecuación (3.94) es simplemente una afirmación de la invariancia del cuadrado de la magnitud del vector distancia bajo rotaciones en el espacio de Minkowski. Aquí es donde la relatividad especial compromete nuestra transformación. Elevando al cuadrado y restando las ecuaciones (3.92) y (3.93) y descartando términos del orden de  $\delta\beta^2$ , encontramos  $a = -1$ . Las ecuaciones (3.92) y (3.93) pueden ser combinadas como una ecuación matricial

$$\begin{pmatrix} x'_0 \\ x'_1 \end{pmatrix} = (1 - \delta\beta\sigma_1) \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \quad (3.95)$$

$\sigma_1$  es la matriz de Pauli, y el parámetro  $\delta\beta$  representa un cambio infinitesimal. Repetimos la transformación  $N$  veces para desarrollar una transformación finita con el parámetro velocidad  $\rho = N\delta\beta$ . entonces

$$\begin{pmatrix} x'_0 \\ x'_1 \end{pmatrix} = \left(1 - \frac{\rho\sigma_1}{N}\right)^N \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}. \quad (3.96)$$

En el límite  $N \rightarrow \infty$

$$\lim_{N \rightarrow \infty} \left(1 - \frac{\rho\sigma_1}{N}\right)^N = \exp(-\rho\sigma_1). \quad (3.97)$$

Interpretamos la exponencial como una serie de Maclaurin

$$\exp(-\rho\sigma_1) = 1 - \rho\sigma_1 + \frac{(\rho\sigma_1)^2}{2!} - \frac{(\rho\sigma_1)^3}{3!} + \dots. \quad (3.98)$$

Notando que  $\sigma^2 = 1$ ,

$$\exp(-\rho\sigma_1) = 1 \cosh \rho + \sigma_1 \sinh \rho. \quad (3.99)$$

Por lo tanto nuestra transformación de Lorentz finita es

$$\begin{pmatrix} x'_0 \\ x'_1 \end{pmatrix} = \begin{pmatrix} \cosh \rho & -\sinh \rho \\ -\sinh \rho & \cosh \rho \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}. \quad (3.100)$$

$\sigma_1$  ha generado las representaciones de esta especial transformación de Lorentz.

El  $\cosh \rho$  y el  $\sinh \rho$  pueden ser identificados considerando el origen del sistema de coordenadas primas,  $x'_1 = 0$ , o  $x_1 = vt$ . Sustituyendo en la ecuación (3.100), tenemos

$$0 = x_1 \cosh \rho - x_0 \sinh \rho. \quad (3.101)$$

Con  $x_1 = vt$  y  $x_0 = ct$ .

$$\tanh \rho = \beta = \frac{v}{c}.$$

Note que la rapidez  $\rho \neq \frac{v}{c}$  excepto en el límite  $v \rightarrow 0$ .

Usando  $1 - \tanh^2 \rho = (\cosh^2 \rho)^{-1}$ ,

$$\cosh \rho = (1 - \beta^2)^{-1/2} \equiv \gamma, \quad \sinh \rho = \beta\gamma. \quad (3.102)$$

El anterior caso especial en que la velocidad es paralela a uno de los ejes espaciales es simple, pero ilustra la velocidad infinitesimal, la técnica de la exponenciación y el generador. Ahora esta técnica puede ser aplicada para derivar las transformaciones de Lorentz para una velocidad relativa  $\vec{v}$  no paralela a ningún eje. Las matrices dadas por la ecuación (3.100) para el caso  $\vec{v} = \hat{x}v_x$  forman un subgrupo. Las matrices en el caso general no lo hacen. El producto de dos matrices de transformaciones de Lorentz,  $L(\vec{v}_1)$  y  $L(\vec{v}_2)$ , producen una tercera matriz de transformación  $L(\vec{v}_3)$ , si las dos velocidades  $\vec{v}_1$  y  $\vec{v}_2$  son paralelas. La velocidad resultante  $v_3$  está relacionada con  $v_1$  y con  $v_2$  mediante la regla de adición de velocidades de Einstein. Si  $\vec{v}_1$  y  $\vec{v}_2$  no son paralelas, no existe entonces una relación simple.

### 3.5 Covarianza de Lorentz de las Ecuaciones de Maxwell.

Si una ley física se mantiene para todas las orientaciones de nuestro (real) espacial sistema de coordenadas (*i.e.* es invariante ante rotaciones), los términos de la ecuación deben ser covariantes bajo rotaciones. Esto significa que escribimos las leyes físicas en la forma matemática escalar=escalar, vector=vector, tensor de segundo rango=tensor de segundo rango, y así sucesivamente. Similarmente, si una ley física se mantiene para todos los sistemas inerciales, los términos de la ecuación deben ser covariantes bajo transformaciones de Lorentz.

Usando el espacio de Minkowski ( $x = x_1, y = x_2, z = x_3, ct = x_0$ ) tenemos un espacio cuadridimensional cartesiano con métrica  $g_{\mu\nu}$ . Las transformaciones de Lorentz son lineales en el espacio y en el tiempo en este espacio real de cuatro dimensiones.

Consideremos las ecuaciones de Maxwell

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad (3.103a)$$

$$\vec{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \rho \vec{v}, \quad (3.103b)$$

$$\vec{\nabla} \cdot \vec{D} = \rho, \quad (3.103c)$$

$$\vec{\nabla} \cdot \vec{B} = 0, \quad (3.103d)$$

y las relaciones

$$\vec{D} = \varepsilon_0 \vec{E}, \quad \vec{B} = \mu_0 \vec{H}. \quad (3.104)$$

Todos los símbolos tienen sus significados usuales y hemos supuesto el vacío por simplicidad.

Supongamos que las ecuaciones de Maxwell se mantienen en todos los sistemas inerciales; esto es, las ecuaciones de Maxwell son consistentes con la relatividad especial. (La covarianza de las ecuaciones de Maxwell bajo transformaciones de Lorentz fue realmente mostrada por Lorentz y Poincaré antes de que Einstein propusiera su teoría de la relatividad especial). Nuestro objetivo inmediato es reescribir las ecuaciones de Maxwell como ecuaciones tensoriales en el espacio de Minkowski. Esto hará la covarianza de Lorentz explícita.

En términos de los potenciales escalar y vectorial, podemos escribir

$$\begin{aligned}\vec{B} &= \vec{\nabla} \times \vec{A} , \\ \vec{E} &= -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla} \varphi .\end{aligned}\tag{3.105}$$

La ecuación anterior especifica el rotor de  $\vec{A}$ ; la divergencia de  $\vec{A}$  no está definida. Podemos, y por futuras conveniencias lo hacemos, imponer la siguiente relación sobre el vector potencial

$$\vec{\nabla} \cdot \vec{A} + \varepsilon_0 \mu_0 \frac{\partial \varphi}{\partial t} = 0 .\tag{3.106}$$

Este es conocido como el gauge de Lorentz. Servirá a nuestros propósitos de desacoplar las ecuaciones diferenciales para  $\vec{A}$  y para  $\varphi$  .

Ahora reescribimos las ecuaciones de Maxwell en términos de los potenciales. A partir de la ecuación (3.103c) para  $\vec{\nabla} \cdot \vec{D}$  y (3.105)

$$\nabla^2 \varphi + \vec{\nabla} \cdot \frac{\partial \vec{A}}{\partial t} = -\frac{\rho}{\varepsilon_0} ,\tag{3.107}$$

considerando que la ecuación (3.103b) para  $\vec{\nabla} \times \vec{H}$  y (3.105) y la identidad vectorial para el rotor del rotor produce

$$\frac{\partial^2 \vec{A}}{\partial t^2} + \vec{\nabla} \frac{\partial \varphi}{\partial t} + \frac{1}{\varepsilon_0 \mu_0} \left[ \vec{\nabla} \vec{\nabla} \cdot \vec{A} - \nabla^2 \vec{A} \right] = \frac{\rho \vec{v}}{\varepsilon_0} .\tag{3.108}$$

Usando el gauge de Lorentz, la ecuación (3.106), y la relación  $\varepsilon_0 \mu_0 = 1/c^2$ , obtenemos

$$\begin{aligned}\left[ \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \vec{A} &= -\mu_0 \rho \vec{v} , \\ \left[ \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \varphi &= -\frac{\rho}{\varepsilon_0} .\end{aligned}\tag{3.109}$$

Ahora el operador diferencial

$$\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} = \partial^2 = -\partial^\mu \partial_\mu ,$$

es un Laplaciano cuatridimensional. Usualmente este operador es llamado el d'Alembertiano y denotado por  $\square^2$ . Puede probarse que es un escalar.

Por conveniencia definimos

$$\begin{aligned}A^1 &\equiv \frac{A_x}{\mu_0 c} = c \varepsilon_0 A_x , & A^3 &\equiv \frac{A_z}{\mu_0 c} = c \varepsilon_0 A_z , \\ A^2 &\equiv \frac{A_y}{\mu_0 c} = c \varepsilon_0 A_y , & A_0 &\equiv \varepsilon_0 \varphi = A^0 .\end{aligned}\tag{3.110}$$

Si ponemos además

$$\frac{\rho v_x}{c} \equiv i^1, \quad \frac{\rho v_y}{c} \equiv i^2, \quad \frac{\rho v_z}{c} \equiv i^3, \quad \rho \equiv i_0 = i^0, \quad (3.111)$$

entonces la ecuación (3.109) puede ser escrita de la forma

$$\partial^2 A^\mu = i^\mu. \quad (3.112)$$

La ecuación anterior parece una ecuación tensorial, pero eso no basta. Para probar que es una ecuación tensorial, partimos investigando las propiedades de transformación de la corriente generalizada  $i^\mu$ .

Ya que un elemento de carga  $de$  es una cantidad invariante, tenemos

$$de = \rho dx_1 dx_2 dx_3, \quad \text{invariante.} \quad (3.113)$$

Vimos que el elemento de volumen cuatridimensional es también un invariante,  $dx_1 dx_2 dx_3 dx_0$ , comparando estos resultados vemos que la densidad de carga  $\rho$  debe transformar de la misma manera que  $x_0$ . Ponemos  $\rho = i^0$  con  $i^0$  establecida como la componente cero de un cuatrivector. Las otras partes de la ecuación (3.111) pueden ser expandidas como

$$\begin{aligned} i^1 &= \frac{\rho v_x}{c} = \frac{\rho dx_1}{c dt} \\ &= i^0 \frac{dx_1}{dt}. \end{aligned} \quad (3.114)$$

Ya que justo mostramos que  $i_0$  transforma como  $dx_0$ , esto significa que  $i_1$  transforma como  $dx_1$ . Con resultados similares para  $i_2$  e  $i_3$ . tenemos que  $i^\lambda$  transforma como  $dx^\lambda$ , probando de esta manera que  $i^\lambda$  es un vector, un vector del espacio cuatridimensional de Minkowski.

La ecuación (3.112), la cual deriva directamente de las ecuaciones de Maxwell, suponemos que se mantiene en todos los sistemas cartesianos. Entonces, por la regla del cociente  $A_\mu$  es también un vector y (3.112) es una legítima ecuación tensorial.

Ahora, devolviendonos, la ecuación (3.105) puede ser escrita

$$\begin{aligned} \varepsilon_0 E_j &= -\frac{\partial A^j}{\partial x_0} + \frac{\partial A^0}{\partial x_j}, \quad j = 1, 2, 3, \\ \frac{1}{\mu c} B_i &= \frac{\partial A^k}{\partial x_j} - \frac{\partial A^j}{\partial x_k}, \quad (i, j, k) = (1, 2, 3), \end{aligned} \quad (3.115)$$

y permutaciones cíclicas.

Definimos un nuevo tensor

$$\partial^\mu A^\nu - \partial^\nu A^\mu = \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu} \equiv F^{\mu\nu} = -F^{\nu\mu} \quad (\mu, \nu = 0, 1, 2, 3)$$

un tensor antisimétrico de segundo rango, ya que  $A^\mu$  es un vector. Escribamoslo explícitamente

$$F_{\mu\nu} = \varepsilon_0 \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix}, \quad F^{\mu\nu} = \varepsilon_0 \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix}. \quad (3.116)$$

Notemos que en nuestro espacio de Minkowski  $\vec{E}$  y  $\vec{B}$  no son más vectores sino que juntos forman un tensor de segundo rango. Con este tensor podemos escribir las dos ecuaciones de Maxwell nohomogéneas (3.103b) y (3.103c) y combinándolas como una ecuación tensorial

$$\frac{\partial F_{\mu\nu}}{\partial x_\nu} = i_\mu . \quad (3.117)$$

El lado izquierdo es una divergencia cuatridimensional de un tensor y por lo tanto un vector. Esto es, por supuesto, equivalente a contraer un tensor de tercer rango  $\frac{\partial F^{\mu\nu}}{\partial x_\lambda}$ . Las ecuaciones de Maxwell (3.103a) para  $\vec{\nabla} \times \vec{E}$  y la ecuación (3.103d) para  $\vec{\nabla} \cdot \vec{B}$  pueden ser expresadas en forma tensorial

$$\frac{\partial F_{23}}{\partial x_1} + \frac{\partial F_{31}}{\partial x_2} + \frac{\partial F_{12}}{\partial x_3} = 0 , \quad (3.118)$$

para (3.103d) y tres ecuaciones de la forma

$$-\frac{\partial F_{30}}{\partial x_2} - \frac{\partial F_{02}}{\partial x_3} \frac{\partial F_{23}}{\partial x_0} = 0 , \quad (3.119)$$

para (3.103a). Una segunda ecuación permutando 120 y una tercera permutando 130.

Ya que

$$\partial^\lambda F^{\mu\nu} = \frac{\partial F^{\mu\nu}}{\partial x_\lambda} \equiv t^{\lambda\mu\nu} ,$$

es un tensor de tercer rango, las ecuaciones (3.117) y (3.119) pueden ser expresadas por la ecuación tensorial

$$t^{\lambda\mu\nu} + t^{\nu\lambda\mu} + t^{\mu\nu\lambda} = 0 . \quad (3.120)$$

En todos los casos anteriores los índices  $\mu$ ,  $\nu$  y  $\lambda$  se suponen diferentes.

### Transformaciones de Lorentz de $\vec{E}$ y $\vec{B}$ .

La construcción de las ecuaciones tensoriales (3.118) y (3.120) completan nuestro objetivo inicial de reescribir las ecuaciones de Maxwell en forma tensorial. Ahora explotamos las propiedades tensoriales de nuestros cuatrivectores y del tensor  $F^{\mu\nu}$ .

Para las transformaciones de Lorentz que corresponden a movimientos a lo largo del eje  $z(x_3)$  con velocidad  $v$ , los “cosenos directores” están dados por

$$\begin{aligned} x'_0 &= \gamma(x_0 - \beta x_3) \\ x'_3 &= \gamma(x_3 - \beta x_0) , \end{aligned} \quad (3.121)$$

donde

$$\beta = \frac{v}{c}$$

y

$$\gamma = (1 - \beta^2)^{-1/2} . \quad (3.122)$$

Usando las propiedades de transformación tensorial, podemos calcular los campos eléctrico y magnético en el sistema en movimiento en términos de los valores en el marco de referencias original. A partir de las ecuaciones (1.59), (3.116) y (3.121) obtenemos

$$\begin{aligned} E'_x &= \frac{1}{\sqrt{1 - \beta^2}} \left( E_x - \frac{v}{c^2} B_y \right) , \\ E'_y &= \frac{1}{\sqrt{1 - \beta^2}} \left( E_y + \frac{v}{c^2} B_x \right) , \\ E'_z &= E_z , \end{aligned} \quad (3.123)$$

y

$$\begin{aligned} B'_x &= \frac{1}{\sqrt{1 - \beta^2}} \left( B_x + \frac{v}{c^2} E_y \right) , \\ B'_y &= \frac{1}{\sqrt{1 - \beta^2}} \left( B_y - \frac{v}{c^2} E_x \right) , \\ B'_z &= B_z . \end{aligned} \quad (3.124)$$

Este acoplamiento de  $\vec{E}$  y  $\vec{B}$  es esperado. Consideremos, por ejemplo, el caso de campo eléctrico nulo en el sistema sin prima

$$E_x = E_y = E_z = 0 .$$

Claramente, no habrá fuerza sobre una partícula carga estacionaria. Cuando la partícula está en movimiento con una velocidad pequeña  $\vec{v}$  a lo largo del eje  $z$  un observador sobre la partícula ve campos (ejerciendo una fuerza sobre la partícula cargada) dados por

$$\begin{aligned} E'_x &= -v B_y , \\ E'_y &= v B_x , \end{aligned}$$

donde  $\vec{B}$  es un campo magnético en el sistema sin primas. Estas ecuaciones pueden ser puestas en forma vectorial

$$\vec{E}' = \vec{v} \times \vec{B} , \quad \text{o bien,} \quad \vec{F} = q\vec{v} \times \vec{B} , \quad (3.125)$$

la cual es usualmente tomada como la definición operacional del campo magnético  $\vec{B}$ .

### Invariantes electromagnéticas.

Finalmente, las propiedades tensoriales (o vectoriales) nos permiten construir una multitud de cantidades invariantes. Una de las importantes es el producto escalar de los cuadvectores  $A_\lambda$  y  $i_\lambda$ . Tenemos

$$\begin{aligned} A^\lambda i_\lambda &= -c\varepsilon_0 A_x \frac{\rho v_x}{c} - c\varepsilon_0 A_y \frac{\rho v_y}{c} - c\varepsilon_0 A_z \frac{\rho v_z}{c} + \varepsilon_0 \varphi \rho \\ &= \varepsilon_0 (\rho \varphi - \vec{A} \cdot \vec{J}) , \quad \text{invariante,} \end{aligned} \quad (3.126)$$

con  $\vec{A}$  el usual potencial vector y  $\vec{J}$  la densidad de corriente ordinaria. El primer término  $\rho\varphi$  es el ordinario acoplamiento electroestático con dimensiones de energía per unidad de volumen. En consecuencia nuestro recién construido invariante escalar es un densidad de energía. La interacción dinámica del campo y corriente es dado por el producto  $\vec{A} \cdot \vec{J}$ . Este invariante  $A^\lambda i_\lambda$  aparece en los Lagrangianos electromagnéticos.

# Capítulo 4

## Series infinitas.

versión final 1.3-180602<sup>1</sup>

### 4.1 Conceptos fundamentales

Las series infinitas, literalmente sumas de un número infinito de términos, ocurre frecuentemente tanto en matemáticas pura como aplicada. Ellas podrían ser usadas por los matemáticos puros para definir funciones como una aproximación fundamental a la teoría de funciones, tan bien como para calcular valores precisos de constantes y funciones trascendentales. En la matemática en ciencias e ingeniería las series infinitas son ubicuas, es por ello que aparecen en la evaluación de integrales, en la solución de ecuaciones diferenciales, en series de Fourier y compite con las representaciones integral para la descripción de funciones especiales. Más adelante veremos la solución en series de Neumann para ecuaciones integrales dan un ejemplo más de la ocurrencia y uso de las series infinitas.

Encaramos el problema que significa la suma de un número infinito de términos. La aproximación usual es por sumas parciales. Si tenemos una sucesión de términos infinitos  $u_1, u_2, u_3, u_4, u_5, \dots$ , definimos la suma parcial  $i$ -ésima como

$$s_i = \sum_{n=1}^i u_n , \quad (4.1)$$

Esta es una suma finita y no ofrece dificultades. Si las sumas parciales  $s_i$  convergen a un límite (finito) cuando  $i \rightarrow \infty$ ,

$$\lim_{i \rightarrow \infty} s_i = S , \quad (4.2)$$

La serie infinita  $\sum_{n=1}^{\infty} u_n$  se dice que es convergente y tiene el valor  $S$ . Note cuidadosamente que nosotros razonablemente y plausiblemente, pero aún arbitrariamente *definimos* que la serie infinita es igual a  $S$ . Podemos notar que una condición necesaria para esta convergencia a un límite es que el  $\lim_{n \rightarrow \infty} u_n = 0$ . Esta condición, sin embargo, no es suficiente para garantizar la convergencia. La ecuación (4.2) usualmente está escrita en notación matemática formal:

---

<sup>1</sup>Este capítulo está basado en el quinto capítulo del libro: *Mathematical Methods for Physicists, fourth edition* de George B. Arfken & Hans J. Weber, editorial ACADEMIC PRESS.

La condición para la existencia de un límite  $S$  es que para cada  $\varepsilon > 0$ , haya un  $N$  fijo tal que

$$|S - s_i| < \varepsilon, \quad \text{para todo } i > N.$$

Esta condición a menudo derivada del *criterio de Cauchy* aplicado a las sumas parciales  $s_i$ . El criterio de Cauchy es:

Una condición necesaria y suficiente para que una sucesión  $(s_i)$  converja es que para cada  $\varepsilon > 0$  exista un número fijo  $N$  tal que

$$|s_j - s_i| < \varepsilon \quad \text{para todos los } i, j > N.$$

Esto significa que las sumas parciales individuales deben mantenerse cercanas cuando nos movemos lejos en la secuencia.

El criterio de Cauchy puede fácilmente extenderse a sucesiones de funciones. La vemos en esta forma en la sección 4.5 en la definición de convergencia uniforme y más adelante en el desarrollo del espacio de Hilbert.

Nuestras sumas parciales  $s_i$  pueden no converger a un límite simple sino que podría oscilar, como en el caso

$$\sum_{n=0}^{\infty} u_n = 1 - 1 + 1 - 1 + 1 - \dots + (-1)^n - \dots .$$

Claramente,  $s_i = 1$  para  $i$  impar pero 0 para  $i$  par. No hay convergencia a un límite, y series tal como una llamadas oscilantes.

Para las series

$$1 + 2 + 3 + \dots + n + \dots$$

tenemos

$$s_n = \frac{n(n+1)}{2}$$

Cuando  $n \rightarrow \infty$ ,

$$\lim_{n \rightarrow \infty} s_n = \infty .$$

Cada vez que las sumas parciales diverjan (tienden a  $\pm\infty$ ), la serie infinita se dice que *diverge*. A menudo el término divergente es extendido para incluir series oscilatorias.

Ya que evaluamos las sumas parciales por aritmética ordinaria, la serie convergente, definida en términos del límite de las sumas parciales, asume una posición de importancia suprema. Dos ejemplos pueden clarificar la naturaleza de convergencia o divergencia de una serie y servirá como una base para una investigación más detallada en la próxima sección.

**Ejemplo** Series geométricas.

La sucesión geométrica, comenzando con  $a$  y con una razón  $r$  ( $r \geq 0$ ), está dado por

$$a + ar + ar^2 + ar^3 + \cdots + ar^{n-1} + \cdots .$$

La suma parcial  $n$ -ésima está dada por

$$s_n = a \frac{1 - r^n}{1 - r} \quad (4.3)$$

Tomando el límite cuando  $n \rightarrow \infty$ ,

$$\lim_{n \rightarrow \infty} s_n = \frac{a}{1 - r}, \quad \text{para } r < 1. \quad (4.4)$$

De modo que, por definición, la serie geométrica infinita converge para  $r < 1$  y está dada por

$$\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1 - r}. \quad (4.5)$$

Por otra parte, si  $r \geq 1$ , la condición necesaria  $u_n \rightarrow 0$  no se satisface y la serie infinita diverge.

**Ejemplo** Series armónicas.

Consideremos la serie armónica

$$\sum_{n=1}^{\infty} n^{-1} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} + \cdots . \quad (4.6)$$

Tenemos que el  $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} 1/n = 0$ , pero esto no es suficiente para garantizar la convergencia. Si agrupamos los términos (no cambiando el orden) como

$$1 + \frac{1}{2} + \left( \frac{1}{3} + \frac{1}{4} \right) + \left( \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) + \left( \frac{1}{9} + \cdots + \frac{1}{16} \right) + \cdots , \quad (4.7)$$

se verá que cada par de paréntesis encierra  $p$  términos de la forma

$$\frac{1}{p+1} + \frac{1}{p+2} + \cdots + \frac{1}{p+p} > \frac{p}{2p} = \frac{1}{2}. \quad (4.8)$$

Formando sumas parciales sumando un grupos entre paréntesis por vez, obtenemos

$$\begin{aligned} s_1 &= 1, & s_4 &> \frac{5}{2}, \\ s_2 &= \frac{3}{2}, & s_5 &> \frac{6}{2}, \\ s_3 &> \frac{4}{2}, & s_n &> \frac{n+1}{2}. \end{aligned} \quad (4.9)$$

Las series armónicas consideradas de esta manera ciertamente son divergentes. Una demostración independiente y alternativa de su divergencia aparece en la sección 4.2.

Usando el teorema del binomio, podríamos expandir la función  $(1+x)^{-1}$ :

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-x)^{n-1} + \dots \quad (4.10)$$

Si tomamos  $x \rightarrow 1$ , la serie se convierte

$$1 - 1 + 1 - 1 + 1 - 1 + \dots, \quad (4.11)$$

una serie que etiquetamos como oscilatoria anteriormente. Aunque no converge en el sentido usual, significa que puede ser ligada a su serie. Euler, por ejemplo, asignado un valor de  $1/2$  a esta sucesión oscilatoria sobre la base de la correspondencia entre esta serie y la bien definida función  $(1+x)^{-1}$ . Desafortunadamente, tal correspondencia entre la serie y la función no es única y esta aproximación deberá ser redefinida. Otros métodos de asignar un significado a una serie oscilatoria o divergente, métodos de definir una suma, han sido desarrollados. Otro ejemplo de generalizar la convergencia lo vemos en las serie asintótica o semiconvergente, consideradas más adelante.

## 4.2 Pruebas de Convergencia

Aunque las series no convergentes pueden ser útiles en ciertos casos especiales, usualmente insistimos, como una materia de conveniencia si no de necesidad, que nuestras series sean convergentes. Por lo tanto esto llega a ser una materia de extrema importancia para ser capaz de decir si una serie dada es o no convergente. Desarrollaremos un número de posibles pruebas, comenzando con una prueba simple pero poco sensible y posteriormente trabajar con una más complicada pero muy sensible.

Por ahora consideremos una serie de términos positivos,  $a_n > 0$ , posponiendo los términos negativos hasta la próxima sección.

### 4.2.1 Pruebas de comparación.

Si término a término una serie de términos  $u_n \leq a_n$ , en el cual los  $a_n$  forman una serie convergente, las series  $\sum_n u_n$  también es convergente. Simbólicamente, tenemos

$$\sum_n a_n = a_1 + a_2 + a_3 + \dots, \quad \text{convergente,}$$

$$\sum_n u_n = u_1 + u_2 + u_3 + \dots.$$

Si  $u_n \leq a_n$  para todo  $n$ , luego  $\sum_n u_n \leq \sum_n a_n$  y  $\sum_n u_n$  por lo tanto es convergente.

Si término a término es una serie de términos  $v_n \geq b_n$ , en el cual  $b_n$  forma una serie divergente, las series  $\sum_n v_n$  también es divergente. Note que las comparaciones de  $u_n$  con  $b_n$

o  $v_n$  con  $a_n$  no dan información. Aquí tenemos

$$\sum_n b_n = b_1 + b_2 + b_3 + \cdots, \quad \text{divergente,}$$

$$\sum_n v_n = v_1 + v_2 + v_3 + \cdots.$$

Si  $v_n \leq b_n$  para todo  $n$ , luego  $\sum_n v_n \leq \sum_n b_n$  y  $\sum_n v_n$  por lo tanto es divergente.

Para las series convergente  $a_n$  tenemos las series geométricas, mientras las series armónicas servirán como las series divergentes  $b_n$ . En tanto otras series son identificadas como convergentes o divergentes, ellas pueden ser usadas como las series conocidas en estas pruebas de comparación.

Todos las pruebas desarrolladas en esta sección son esencialmente pruebas de comparación. La figura 4.1 muestra estas pruebas y sus relaciones.

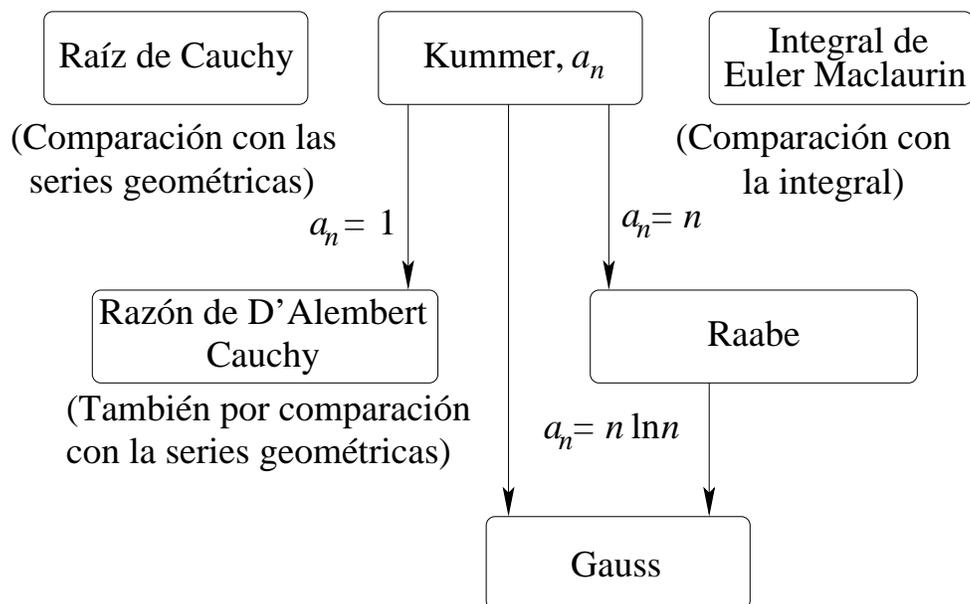


Figura 4.1: Test de comparación.

**Ejemplo** Las series  $p$ .

Probamos  $\sum_n n^{-p}$ ,  $p = 0.999$ , por convergencia. Ya que  $n^{-0.999} > n^{-1}$ , y  $b_n = n^{-1}$  forman la serie armónica divergente, la prueba de comparación muestra que  $\sum_n n^{-0.999}$  es divergente. Generalizando,  $\sum_n n^{-p}$  se ve como divergente para todo  $p \leq 1$ .

### 4.2.2 Prueba de la raíz de Cauchy.

Si  $(a_n)^{1/n} \leq r < 1$  para todo  $n$  suficientemente grande, con  $r$  independiente de  $n$ , entonces  $\sum_n a_n$  es convergente. Si  $(a_n)^{1/n} \geq 1$  para todo  $n$  suficientemente grande, entonces  $\sum_n a_n$  es divergente.

La primera parte de esta prueba se verifica fácilmente elevando  $(a_n)^{1/n} \leq r$  a la  $n$ -ésima potencia. Obtenemos

$$a_n \leq r^n < 1 .$$

Ya que  $r^n$  es sólo el término  $n$ -ésimo en una serie geométrica convergente,  $\sum_n a_n$  es convergente por la prueba de comparación. Conversamente, si  $(a_n)^{1/n} \geq 1$ , entonces  $a_n \geq 1$  y la serie debería diverger. La prueba de la raíz es particularmente útil en establecer las propiedades de la serie de potencias.

### 4.2.3 Prueba de la razón de D' Alembert o Cauchy.

Si  $a_{n+1}/a_n \leq r < 1$  para todo  $n$  suficientemente grande, y  $r$  independiente de  $n$ , entonces  $\sum_n a_n$  es convergente. Si  $a_{n+1}/a_n \geq 1$  para todo  $n$  suficientemente grande, entonces  $\sum_n a_n$  es divergente.

La convergencia está dada por la comparación directa con las series geométricas  $(1 + r + r^2 + \dots)$ . En la segunda parte  $a_{n+1} \geq a_n$  y la divergencia debe ser razonablemente obvia. Aunque la prueba no es tan sensible como la prueba de la raíz de Cauchy, esta prueba de la razón de D' Alembert es una de las más fáciles de aplicar y es ampliamente usada. Una afirmación alternativa de la prueba de la razón está en la forma de un límite: Si

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} < 1 , & \text{ convergencia} \\ & > 1 , \text{ divergencia} \\ & = 1 , \text{ indeterminado.} \end{aligned} \tag{4.12}$$

A causa de la posibilidad de ser indeterminado, la prueba de la razón es probable que falle en puntos cruciales, y se hace necesario una prueba más delicada y sensible.

Podríamos preguntarnos cómo podría levantarse esta indeterminación. Realmente fue disimulado en el primera afirmación  $a_{n+1}/a_n \leq r < 1$ . Podríamos encontrar  $a_{n+1}/a_n < 1$  para todo  $n$  finito pero ser inapropiado escoger un  $r < 1$  e *independiente* de  $n$  tal que  $a_{n+1}/a_n \leq r$  para todo  $n$  suficientemente grande. Un ejemplo está dado por las series armónicas

$$\frac{a_{n+1}}{a_n} = \frac{n}{n+1} < 1 , \tag{4.13}$$

Ya que

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = 1 , \tag{4.14}$$

no existe una razón fija  $r < 1$  y la prueba de la razón falla.

**Ejemplo** Prueba de la razón de D' Alembert.

Probar la convergencia de  $\sum_n \frac{n}{2^n}$

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)/2^{n+1}}{n/2^n} = \frac{1}{2} \frac{n+1}{n} . \tag{4.15}$$

Ya que

$$\frac{a_{n+1}}{a_n} \leq \frac{3}{4} \quad \text{para } n \geq 2, \quad (4.16)$$

tenemos convergencia. Alternativamente,

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \frac{1}{2}, \quad (4.17)$$

y de nuevo converge.

#### 4.2.4 Prueba integral de Cauchy o Maclaurin.

Esta es otra clase de prueba de comparación en la cual comparamos una serie con una integral. Geométricamente, comparamos el área de una serie de un rectángulo de ancho unitario con el área bajo la curva.

Sea  $f(x)$  una función continua, monótonamente decreciente en la cual  $f(n) = a_n$ . Luego  $\sum_n a_n$  converge si  $\int_0^\infty f(x) dx$  es finita y diverge si la integral es infinita. Para la  $i$ -ésima suma parcial

$$s_i = \sum_{n=1}^i a_n = \sum_{n=1}^i f(n). \quad (4.18)$$

Pero

$$s_i > \int_1^{i+1} f(x) dx, \quad (4.19)$$

por la figura 4.2a,  $f(x)$  es monótonamente decreciente. Por otra parte, de la figura 4.2b,

$$s_i - a_1 < \int_1^i f(x) dx, \quad (4.20)$$

en la cual la serie está representada por los rectángulos inscritos. Tomando el límite como  $i \rightarrow \infty$ , tenemos

$$\int_1^\infty f(x) dx < \sum_{n=1}^\infty a_n < \int_1^\infty f(x) dx + a_1. \quad (4.21)$$

De modo que la serie infinita converge o diverge cuando la integral correspondiente converge o diverge respectivamente.

La prueba de la integral es particularmente útil para acotar superior e inferiormente el resto de una serie, después de que algunos números de términos iniciales hayan sido sumados. Esto es,

$$\sum_{n=1}^\infty a_n = \sum_{n=1}^N a_n + \sum_{n=N+1}^\infty a_n,$$

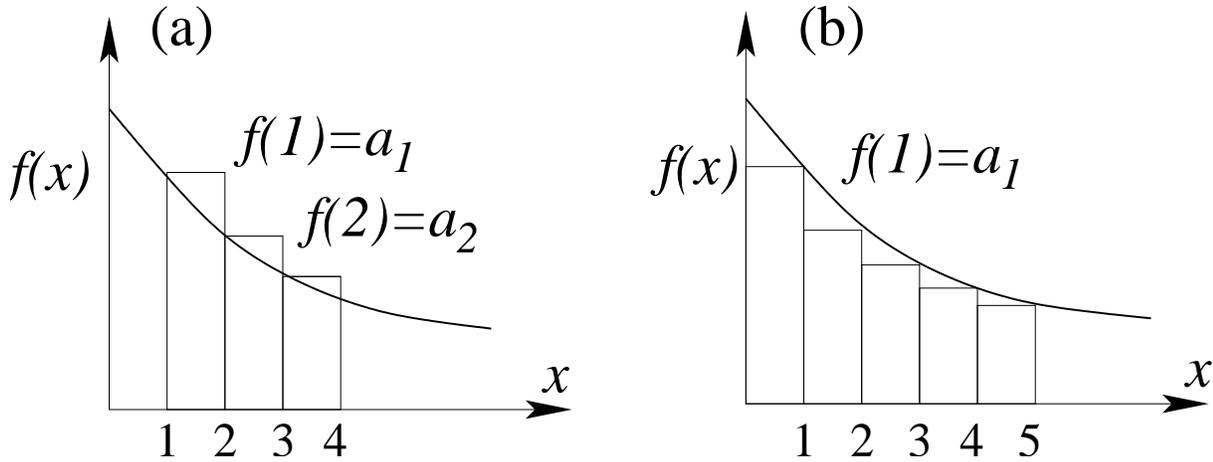


Figura 4.2: (a) Comparación de la integral y la suma de bloques sobresalientes. (b) Comparación de la integral y la suma de bloques envueltos.

donde

$$\int_{N+1}^{\infty} f(x) dx < \sum_{n=N+1}^{\infty} a_n < \int_{N+1}^{\infty} f(x) dx + a_{N+1} .$$

Podemos liberar la prueba de la integral de los requerimientos muy restrictivos de que la función de interpolación  $f(x)$  sea positiva y monótonamente decreciente, basta que la función  $f(x)$  tenga una derivada continua que satisfaga

$$\sum_{n=N_i+1}^{N_f} f(n) = \int_{N_i}^{N_f} f(x) dx + \int_{N_i}^{N_f} (x - [x])f'(x) dx . \quad (4.22)$$

Aquí  $[x]$  denota el entero mayor por debajo de  $x$ , tal que  $x - [x]$  varía como diente de sierra entre 0 y 1.

### Ejemplo Función Zeta de Riemann.

La función zeta de Riemann está definida por

$$\zeta(p) = \sum_{n=1}^{\infty} n^{-p} . \quad (4.23)$$

Podemos tomar  $f(x) = x^{-p}$  y entonces

$$\int_1^{\infty} x^{-p} dx = \begin{cases} \left. \frac{x^{-p+1}}{-p+1} \right|_1^{\infty} , & p \neq 1 \\ \ln x \Big|_1^{\infty} , & p = 1 \end{cases} \quad (4.24)$$

La integral y por lo tanto la serie son divergentes para  $p \leq 1$  y convergente para  $p > 1$ . De modo que la ecuación (4.23) lleva la condición de  $p > 1$ . Esto, incidentalmente, es una prueba independiente de que la serie armónica ( $p = 1$ ) diverge y lo hace en forma logarítmica. La suma del primer millón de términos  $\sum_{n=1}^{1.000.000} n^{-1}$ , es solamente 14.392726 . . . .

Esta comparación con la integral también puede ser usada para dar una cota superior a la constante Euler-Mascheroni definida por

$$\gamma = \lim_{n \rightarrow \infty} \left( \sum_{m=1}^n \frac{1}{m} - \ln n \right) . \quad (4.25)$$

Volviendo a las sumas parciales,

$$s_n = \sum_{m=1}^n m^{-1} - \ln n < \int_1^n \frac{dx}{x} - \ln n + 1 . \quad (4.26)$$

Evaluando la integral del lado derecho,  $s_n < 1$  para todo  $n$  y por lo tanto  $\gamma < 1$ . Realmente la constante de Euler-Mascheroni es 0.57721566 . . . .

### 4.2.5 Prueba de Kummer.

Esta es la primera de tres pruebas que son algo más difíciles para aplicar que las anteriores. Su importancia radica en su poder y sensibilidad. Frecuentemente, al menos una de las tres funcionará cuando las pruebas más fáciles sean indeterminadas. Debe recordarse, sin embargo, que estas pruebas, como aquellas previamente discutidas, están finalmente basadas en comparaciones. Esto significa que todas las pruebas de convergencia dadas aquí, incluyendo la de Kummer, puedan fallar algunas veces.

Consideremos una serie de términos positivos  $u_i$  y una sucesión de constantes positivas finitas  $a_i$ . Si

$$a_n \frac{u_n}{u_{n+1}} - a_{n+1} \geq C > 0 , \quad (4.27)$$

para todo  $n \geq N$ , algún número fijo, entonces  $\sum_{i=1}^{\infty} u_i$  converge. Si

$$a_n \frac{u_n}{u_{n+1}} - a_{n+1} \leq 0 \quad (4.28)$$

y  $\sum_{i=1}^{\infty} a_i^{-1}$  diverge, luego  $\sum_{i=1}^{\infty} u_i$  diverge.

La prueba de este poderoso test es simple y queda como ejercicio.

Si las constantes positivas  $a_n$  de la prueba de Kummer son elegidas como  $a_n = n$ , tenemos la prueba de Raabe.

### 4.2.6 Prueba de Raabe.

Si  $u_n > 0$  y si

$$n \left( \frac{u_n}{u_{n+1}} - 1 \right) \geq P > 1 , \quad (4.29)$$

para todo  $n \geq N$ , donde  $N$  es un entero positivo independiente de  $n$ , entonces  $\sum_i u_i$  converge. Si

$$n \left( \frac{u_n}{u_{n+1}} - 1 \right) \leq 1, \quad (4.30)$$

entonces  $\sum_i u_i$  diverge ( $\sum n^{-1}$  diverge).

La forma en límite en el test de Raabe es

$$\lim_{n \rightarrow \infty} n \left( \frac{u_n}{u_{n+1}} - 1 \right) = P. \quad (4.31)$$

Tenemos convergencia para  $P > 1$ , y divergencia para  $P < 1$ , y no hay prueba para  $P = 1$  exactamente como con el test de Kummer. Esta indeterminancia está expresada en que podemos encontrar ejemplos de una serie convergente y una divergente en que ambas series tienden a  $P = 1$  en la ecuación (4.31).

El test de Raabe es más sensible que la prueba de la razón de D'Alembert ya que  $\sum_{n=1}^{\infty} n^{-1}$  diverge más lentamente que  $\sum_{n=1}^{\infty} 1$ . Obtenemos una prueba aún más sensible (y una relativamente fácil de aplicar) si escogemos  $a_n = n \ln n$ . Esto es la prueba de Gauss.

### 4.2.7 Prueba de Gauss.

Si  $u_n > 0$  para todo  $n$  finito y

$$\frac{u_n}{u_{n+1}} = 1 + \frac{h}{n} + \frac{B(n)}{n^2}, \quad (4.32)$$

en el cual  $B(n)$  es una función acotada de  $n$  para  $n \rightarrow \infty$ , luego  $\sum_i u_i$  converge para  $h > 1$  y diverge para  $h \leq 1$ .

La razón  $u_n/u_{n+1}$  de la ecuación (4.32) a menudo llega a ser como la razón de dos formas cuadráticas:

$$\frac{u_n}{u_{n+1}} = \frac{n^2 + a_1 n + a_0}{n^2 + b_1 n + b_0}. \quad (4.33)$$

Se puede mostrar que tenemos convergencia para  $a_1 > b_1 + 1$  y divergencia para  $a_1 \leq b_1 + 1$ .

El test de Gauss es un test extremadamente sensible para la convergencia de series. Esto funcionará para prácticamente todas las series que encontraremos en Física. Para  $h > 1$  o  $h < 1$  la prueba se deduce directamente del test de Raabe

$$\lim_{n \rightarrow \infty} n \left[ 1 + \frac{h}{n} + \frac{B(n)}{n^2} - 1 \right] = \lim_{n \rightarrow \infty} \left[ h + \frac{B(n)}{n} \right] = h. \quad (4.34)$$

Si  $h = 1$ , falla el test de Raabe. Sin embargo, si volvemos al test de Kummer y usamos  $a_n = n \ln n$ , tenemos

$$\begin{aligned} & \lim_{n \rightarrow \infty} \left\{ n \ln n \left[ 1 + \frac{1}{n} + \frac{B(n)}{n^2} \right] - (n+1) \ln(n+1) \right\} \\ &= \lim_{n \rightarrow \infty} \left[ n \ln n \cdot \frac{(n+1)}{n} - (n+1) \ln(n+1) \right] \\ &= \lim_{n \rightarrow \infty} (n+1) \left[ \ln n - \ln n - \ln \left( 1 + \frac{1}{n} \right) \right]. \end{aligned} \quad (4.35)$$

Pidiendo prestado un resultado de la sección 4.6 (el cual no es dependiente de la prueba de Gauss) tenemos

$$\lim_{n \rightarrow \infty} -(n+1) \ln \left( 1 + \frac{1}{n} \right) = \lim_{n \rightarrow \infty} -(n+1) \left( \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} \dots \right) = -1 < 0. \quad (4.36)$$

De modo que tenemos divergencia para  $h = 1$ . Esto es un ejemplo de una aplicación exitosa del test de Kummer en el cual el test de Raabe falla.

### Ejemplo Series de Legendre.

La relación de recurrencia para la solución en serie de la ecuación de Legendre pueden ser colocadas en la forma

$$\frac{a_{2j+2}}{a_{2j}} = \frac{2j(2j+1) - l(l+1)}{(2j+1)(2j+2)}. \quad (4.37)$$

Esto es equivalente a  $u_{2j+2}/u_{2j}$  para  $x = +1$ . Para  $j \gg l$

$$\frac{a_{2j}}{a_{2j+2}} \rightarrow \frac{(2j+1)(2j+2)}{2j(2j+1)} = \frac{2j+2}{2j} = 1 + \frac{1}{j}. \quad (4.38)$$

Por la ecuación (4.33) la serie es divergente. Más adelante exigiremos que las series de Legendre sean finitas (se corten) para  $x = 1$ . Eliminaremos la divergencia ajustando los parámetros  $n = 2j_0$ , un entero par. Esto truncará la serie, convirtiendo la serie infinita en un polinomio.

### 4.2.8 Mejoramiento de convergencia.

En esta sección no nos preocupará establecer la convergencia como una propiedad matemática abstracta. En la práctica, la razón de convergencia puede ser de considerable importancia. Aquí presentamos un método que mejora la razón de la convergencia de una serie ya convergente.

El principio básico de este método, debido a Kummer, es formar una combinación lineal de nuestra serie lentamente convergente y una o más series cuya suma es conocida. Entre las series conocidas la colección

$$\begin{aligned} \alpha_1 &= \sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1, \\ \alpha_2 &= \sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)} = \frac{1}{4}, \\ \alpha_3 &= \sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)(n+3)} = \frac{1}{18}, \\ &\vdots \\ \alpha_p &= \sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2) \cdots (n+p)} = \frac{1}{p \cdot p!}, \end{aligned}$$

es particularmente útil. Las series están combinadas término a término y los coeficientes en combinación lineal son escogidos para cancelar los términos que convergen lentamente.

**Ejemplo** Función zeta de Riemann,  $\zeta(3)$ .

Sea la serie a ser sumada  $\sum_{n=1}^{\infty} n^{-3}$ . En la sección 4.10 está identificada como una función zeta de Riemann,  $\zeta(3)$ . Formamos una combinación lineal

$$\sum_{n=1}^{\infty} n^{-3} + a_2 \alpha_2 = \sum_{n=1}^{\infty} n^{-3} + \frac{a_2}{4} .$$

$\alpha_1$  no está incluida ya que converge más lentamente que  $\zeta(3)$ . Combinando términos, obtenemos sobre la mano izquierda

$$\sum_{n=1}^{\infty} \left\{ \frac{1}{n^3} + \frac{a_2}{n(n+1)(n+2)} \right\} = \sum_{n=1}^{\infty} \frac{n^2(1+a_2) + 3n+2}{n^3(n+1)(n+2)} .$$

Si escogemos  $a_2 = -1$ , la ecuación precedente tiende a

$$\zeta(3) = \sum_{n=1}^{\infty} n^{-3} = \frac{1}{4} + \sum_{n=1}^{\infty} \frac{3n+2}{n^3(n+1)(n+2)} . \quad (4.39)$$

La serie resultante no es muy bonita pero converge como  $n^{-4}$ , apreciablemente más rápido que  $n^{-3}$ .

El método puede ser extendido incluyendo  $a_3 \alpha_3$  para obtener la convergencia como  $n^{-5}$ ,  $a_4 \alpha_4$  para obtener la convergencia como  $n^{-6}$ , etc. Eventualmente, usted tiene que alcanzar un compromiso entre cuanta álgebra usted hace y cuanta aritmética la computadora hace. Como las computadoras lo hacen más rápido, el balance está seguramente sustituyendo menos algebra hecha por usted por más aritmética realizada por el computador.

## 4.3 Series alternadas.

En la sección 4.2 nos limitamos a series de términos positivos. Ahora, en contraste, consideraremos series infinitas en las cuales los signos se alternan. La cancelación parcial debido a la alternancia de los signos hace la convergencia más rápida y mucho más fácil de identificar. Probaremos que el criterio de Leibniz es una condición general para la convergencia de una serie alternada.

### 4.3.1 Criterio de Leibniz.

Consideremos la serie  $\sum_{n=1}^{\infty} (-1)^{n+1} a_n$  con  $a_n > 0$ . Si  $a_n$  es monótonamente decreciente (para  $N$  suficientemente grande) y el  $\lim_{n \rightarrow \infty} a_n = 0$ , entonces la serie converge.

Para probar esto, examinemos las sumas parciales pares

$$\begin{aligned} s_{2n} &= a_1 - a_2 + a_3 - \dots - a_{2n} , \\ s_{2n+2} &= s_{2n} + (a_{2n+1} - a_{2n+2}) . \end{aligned} \quad (4.40)$$

Ya que  $a_{2n+1} > a_{2n+2}$ , tenemos

$$s_{2n+2} > s_{2n} . \quad (4.41)$$

Por otra parte,

$$s_{2n+2} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \dots - a_{2n+2} . \quad (4.42)$$

De modo que, con cada par de términos  $a_{2p} - a_{2p+1} > 0$ ,

$$s_{2n+2} < a_1 . \quad (4.43)$$

Con las sumas parciales pares acotamos  $s_{2n} < s_{2n+2} < a_1$  y los términos  $a_n$  decrecen monótonamente aproximándose a cero, esta serie alternada converge.

Un resultado más importante puede ser extraído de las sumas parciales. A partir de las diferencias entre el límite de la serie  $S$  y las sumas parciales  $s_n$

$$\begin{aligned} S - s_n &= a_{n+1} - a_{n+2} + a_{n+3} - a_{n+4} + \dots \\ &= a_{n+1} - (a_{n+2} - a_{n+3}) - (a_{n+4} - a_{n+5}) - \dots \end{aligned} \quad (4.44)$$

o

$$S - s_n < a_{n+1} . \quad (4.45)$$

La ecuación (4.45) dice que el error en el corte de una serie alternada después de  $n$  términos es menor que  $a_{n+1}$ , el primer término excluido. Un conocimiento del error obtenido de esta manera puede ser de gran importancia práctica.

### 4.3.2 Convergencia absoluta.

Dada una serie en términos de  $u_n$  en la cual  $u_n$  puede variar en signo, si  $\sum |u_n|$  converge, entonces  $\sum u_n$  se dice que es *absolutamente convergente*. Si  $\sum u_n$  converge pero  $\sum |u_n|$  diverge, la convergencia recibe el nombre de *condicional*.

La serie alternada armónica es un ejemplo simple de esta convergencia condicionada. Tenemos

$$\sum_{n=1}^{\infty} (-1)^{n-1} n^{-1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{1}{n} - \dots \quad (4.46)$$

convergente por el criterio de Leibniz, pero

$$\sum_{n=1}^{\infty} n^{-1} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n} + \dots$$

se ha demostrado que es divergente en la sección 4.1 y 4.2.

Podemos notar que todas las pruebas desarrolladas en la sección 4.2 supone una serie de términos positivos. Por lo tanto, todas las pruebas en esa sección garantizan la convergencia absoluta.

**Ejemplo**

Para  $0 < x < \pi$  la serie de Fourier

$$\sum_{n=1}^{\infty} \frac{\cos(nx)}{n} = -\ln\left(2 \operatorname{sen} \frac{x}{2}\right), \quad (4.47)$$

converge teniendo coeficientes que cambian de signo frecuentemente, pero no tanto para que el criterio de convergencia de Leibniz se aplique fácilmente. Apliquemos el test de la integral de la ecuación (4.22). Usando integración por partes vemos de inmediato que

$$\int_1^{\infty} \frac{\cos(nx)}{n} dn = \left[ \frac{\operatorname{sen}(nx)}{nx} \right]_1^{\infty} - \frac{1}{x} \int_1^{\infty} \frac{\operatorname{sen}(nx)}{n^2} dn$$

converge para  $n \rightarrow \infty$ , y la integral del lado derecho incluso converge absolutamente. El término derivado en la ecuación (4.22) tiene la forma

$$\int_1^{\infty} (n - [n]) \left\{ -\frac{x}{n} \operatorname{sen}(nx) - \frac{\cos(nx)}{n^2} \right\} dn,$$

donde el segundo término converge absolutamente y no necesita ser considerado. Lo próximo es observar que  $g(N) = \int_1^N (n - [n]) \operatorname{sen}(nx) dn$  es acotado para  $N \rightarrow \infty$ , tal como  $\int^N \operatorname{sen}(nx) dn$  es acotado debido a la naturaleza periódica de  $\operatorname{sen}(nx)$  y a su regular cambio de signo. Usando integración por partes nuevamente

$$\int_1^{\infty} \frac{g'(n)}{n} dn = \left[ \frac{g(n)}{n} \right]_1^{\infty} + \int_1^{\infty} \frac{g(n)}{n^2} dn,$$

vemos que el segundo término es absolutamente convergente, y el primero va a cero en el límite superior. Por lo tanto la serie en la ecuación (4.47) converge, lo cual es duro de ver usando otro test de convergencia.

## 4.4 Álgebra de series.

Establecer la convergencia absoluta es importante porque puede probarse que las series absolutamente convergentes pueden ser manipuladas de acuerdo a las reglas familiares del álgebra o aritmética.

1. Si una serie infinita es absolutamente convergente, la suma de la serie es independiente del orden en el cual los términos son añadidos.
2. La serie puede ser multiplicada por otra serie absolutamente convergente. El límite del producto será el producto de los límites de las series individuales. El producto de las series, una doble serie, también será absolutamente convergente.

No hay tales garantías en series condicionalmente convergentes. Nuevamente consideremos la serie armónica alternada. Si escribimos

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots = 1 - \left( \frac{1}{2} - \frac{1}{3} \right) - \left( \frac{1}{4} - \frac{1}{5} \right) - \dots, \quad (4.48)$$

es claro que la suma

$$\sum_{n=1}^{\infty} (-1)^{n-1} n^{-1} < 1. \quad (4.49)$$

Sin embargo, si rearrreglamos los términos sutilmente, podemos hacer que la serie armónica alternada converja a  $3/2$ . Reagrupamos los términos de la ecuación (4.48), tomando

$$\begin{aligned} & \left(1 + \frac{1}{3} + \frac{1}{5}\right) - \frac{1}{2} + \left(\frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15}\right) - \frac{1}{4} \\ & + \left(\frac{1}{17} + \cdots + \frac{1}{25}\right) - \frac{1}{6} + \left(\frac{1}{27} + \cdots + \frac{1}{35}\right) - \frac{1}{8} + \cdots. \end{aligned} \quad (4.50)$$

Tratando los términos agrupados en paréntesis como términos simples por conveniencia, obtenemos las sumas parciales

$$\begin{array}{ll} s_1 = 1.5333 & s_2 = 1.0333 \\ s_3 = 1.5218 & s_4 = 1.2718 \\ s_5 = 1.5143 & s_6 = 1.3476 \\ s_7 = 1.5103 & s_8 = 1.3853 \\ s_9 = 1.5078 & s_{10} = 1.4078 \end{array}$$

A partir de esta tabulación de los  $s_n$  y el gráfico de  $s_n$  versus  $n$  en la figura 4.3 es clara la convergencia a  $3/2$ . Hemos rearrreglado los términos, tomando términos positivos hasta que la suma parcial sea igual o mayor que  $3/2$ , luego sumando los términos negativos hasta que la suma parcial caiga bajo  $3/2$ , etc. Como las series se extienden hasta infinito, todos los términos originales eventualmente aparecerán, pero las sumas parciales de este reordenamiento de esta serie armónica alternada converge a  $3/2$ . Por un reordenamiento de términos una serie condicionalmente convergente podría ser hecha para converger a algún valor deseado o para que diverja. Esta afirmación es dada como el teorema de Riemann. Obviamente, series condicionalmente convergentes deberían ser tratadas con precaución.

#### 4.4.1 Mejoramiento de la convergencia, aproximaciones racionales.

La serie

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}, \quad -1 < x \leq 1, \quad (4.51)$$

converge muy suavemente cuando  $x$  se aproxima a  $+1$ . La razón de convergencia podría ser mejorada sustancialmente multiplicando ambos lados de la ecuación (4.51) por un polinomio y ajustando los coeficientes del polinomio para cancelar las porciones que convergen más lentamente en la serie. Consideremos la posibilidad más simple: Multiplicar  $\ln(1+x)$  por  $1 + a_1x$ .

$$(1 + a_1x) \ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} + a_1 \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^{n+1}}{n}.$$

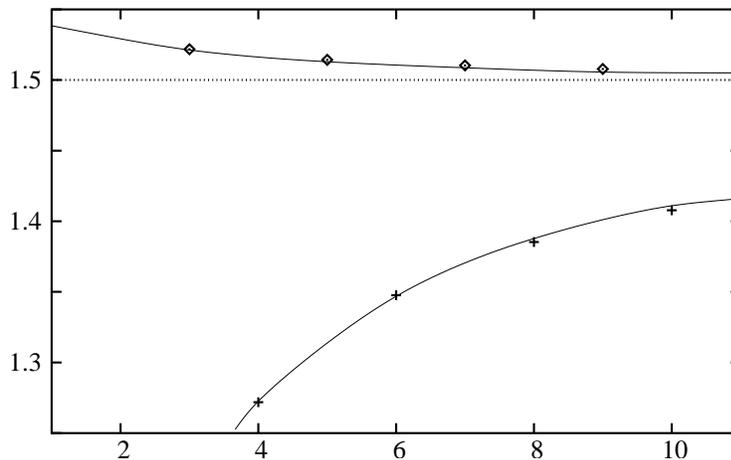


Figura 4.3: Serie armónica alternada, rearreglo de términos para dar convergencia a 1.5.

Combinando las dos series sobre la derecha término a término, obtenemos

$$\begin{aligned} (1 + a_1x) \ln(1 + x) &= x + \sum_{n=2}^{\infty} (-1)^{n-1} \left( \frac{1}{n} - \frac{a_1}{n-1} \right) x^n \\ &= x + \sum_{n=2}^{\infty} (-1)^{n-1} \frac{n(1 - a_1) - 1}{n(n-1)} x^n . \end{aligned}$$

Claramente, si tomamos  $a_1 = 1$ , el  $n$  en el numerador desaparece y nuestra serie combinada converge como  $n^{-2}$ .

Continuando este proceso, encontramos que  $(1 + 2x + x^2) \ln(1 + x)$  se anula como  $n^{-3}$ ,  $(1 + 3x + 3x^2 + x^3) \ln(1 + x)$  se anula cuando  $n^{-4}$ . En efecto estamos desplazándonos desde una expansión de serie simple de la ecuación (4.51) a una representación racional en la cual la función  $\ln(1 + x)$  está representada por la razón de una serie y un polinomio:

$$\ln(1 + x) = \frac{x + \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n(n-1)}}{1 + x} .$$

Tales aproximaciones racionales pueden ser ambas compactas y precisas. Los programas computacionales hacen extensivo el uso de ellas.

#### 4.4.2 Reordenamiento de series dobles.

Otro aspecto del reordenamiento de series aparece en el tratamiento de series dobles (figura 4.4):

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{n,m} .$$

	$m=0$	1	2	3
$n=0$	$a_{00}$	$a_{01}$	$a_{02}$	$a_{03}$
1	$a_{10}$	$a_{11}$	$a_{12}$	$a_{13}$
2	$a_{20}$	$a_{21}$	$a_{22}$	$a_{23}$
3	$a_{30}$	$a_{31}$	$a_{32}$	$a_{33}$

Figura 4.4: Series dobles, la suma sobre  $n$  es indicada por líneas segmentadas verticales.

sustituyamos

$$\begin{aligned} n &= q \geq 0, \\ m &= p - q \geq 0, \\ (q &\leq p). \end{aligned}$$

Esto resulta en la identidad

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{n,m} = \sum_{p=0}^{\infty} \sum_{q=0}^p a_{q,p-q}. \tag{4.52}$$

La suma sobre  $p$  y  $q$  de la ecuación (4.52) está ilustrada en la figura 4.5. La sustitución

	$p=0$	1	2	3
$q=0$	$a_{00}$	$a_{01}$	$a_{02}$	$a_{03}$
1		$a_{10}$	$a_{11}$	$a_{12}$
2			$a_{20}$	$a_{21}$
3				$a_{30}$

Figura 4.5: Series dobles nuevamente, la primera suma es representada por líneas segmentadas verticales pero estas líneas verticales corresponden a las diagonales en la figura 4.4.

$$n = s \geq 0, \quad m = r - 2s \geq 0, \quad \left(s \leq \frac{r}{2}\right)$$

tiende a

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{n,m} = \sum_{r=0}^{\infty} \sum_{s=0}^{[r/2]} a_{s,r-2s}. \quad (4.53)$$

con  $[r/2] = r/2$  para  $r$  par,  $(r-1)/2$  para  $r$  impar. La suma sobre  $r$  y  $s$  de la ecuación (4.53) está mostrada en la figura 4.6. Las ecuaciones (4.52) y (4.53) son claramente reordenamientos del arreglo de coeficientes  $a_{n,m}$ , reordenamientos que son válidos en tanto tengamos convergencia absoluta. La combinación de las ecuaciones (4.52) y (4.53),

	$r=$	0	1	2	3	4
$s=$	0	$a_{00}$	$a_{01}$	$a_{02}$	$a_{03}$	$a_{04}$
				⋮	⋮	⋮
1			$a_{10}$	$a_{11}$	$a_{12}$	
					⋮	
2						$a_{20}$

Figura 4.6: Series dobles. La suma sobre  $s$  corresponde a la suma a lo largo de las líneas segmentadas inclinadas, en la figura 4.4.

$$\sum_{p=0}^{\infty} \sum_{q=0}^p a_{q,p-q} = \sum_{r=0}^{\infty} \sum_{s=0}^{[r/2]} a_{s,r-2s}. \quad (4.54)$$

es usada en la determinación de la forma en serie de los polinomios de Legendre.

## 4.5 Series de funciones.

Extendemos nuestro concepto de series infinitas para incluir la posibilidad que cada término  $u_n$  pueda ser una función de alguna variable,  $u_n = u_n(x)$ . Numerosas ilustraciones de tales series de funciones aparecerán más adelante. Las sumas parciales llegan a ser funciones de la variable  $x$

$$s_n(x) = u_1(x) + u_2(x) + \cdots + u_n(x), \quad (4.55)$$

tal como lo hacemos para la suma de serie, definimos el límite como el límite de las sumas parciales

$$\sum_{n=1}^{\infty} u_n(x) = S(x) = \lim_{n \rightarrow \infty} s_n(x) . \quad (4.56)$$

Hasta ahora nos hemos ocupado del comportamiento de las sumas parciales como una función de  $n$ . Ahora consideremos cómo las cantidades anteriores dependen de  $x$ . Aquí el concepto clave es la convergencia uniforme.

### 4.5.1 Convergencia uniforme.

Si para cualquier  $\varepsilon > 0$  pequeño, existe un número  $N$ , independiente de  $x$  en el intervalo  $[a, b]$  con  $(a \leq x \leq b)$  tal que

$$|S(x) - s_n(x)| < \varepsilon , \forall n \geq N , \quad (4.57)$$

se dice que la serie converge uniformemente en el intervalo  $[a, b]$ . Esto dice que para que nuestra serie sea uniformemente convergente, debe ser posible encontrar un  $N$  finito tal que la cola de la serie infinita,  $|\sum_{i=N+1}^{\infty} u_i(x)|$ , sea menor que un  $\varepsilon$  arbitrariamente pequeño para todo  $x$  en el intervalo dado.

Esta condición, ecuación (4.57), la cual define la convergencia uniforme, es ilustrada en la figura 4.7. El punto es que no importa cuán pequeño sea  $\varepsilon$  podemos siempre tomar un  $n$  suficientemente grande tal que la magnitud absoluta de la diferencia entre  $S(x)$  y  $s_n(x)$  sea menor que  $\varepsilon$  para todo  $x$ ,  $a \leq x \leq b$ . Si esto no puede ser hecho, entonces  $\sum u_n(x)$  no es uniformemente convergente en el intervalo  $[a, b]$ .

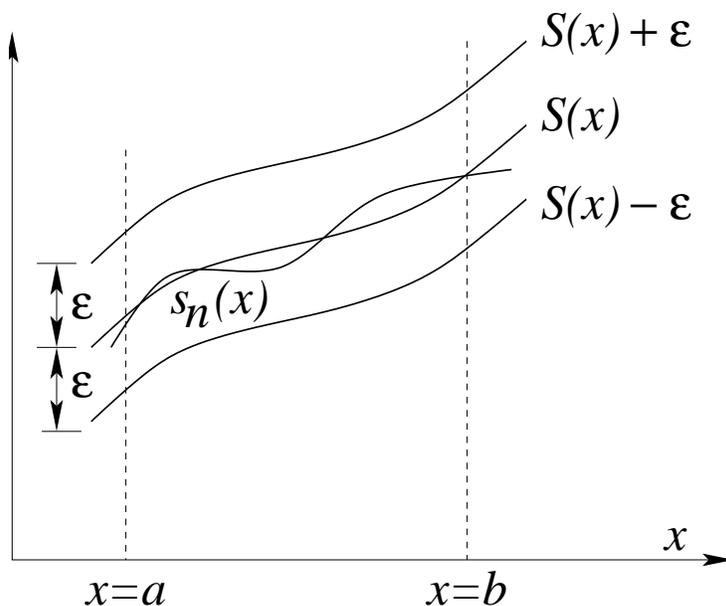


Figura 4.7: Convergencia uniforme.

**Ejemplo**

$$\sum_{n=1}^{\infty} u_n(x) = \sum_{n=1}^{\infty} \frac{x}{[(n-1)x+1][nx+1]} . \quad (4.58)$$

La suma parcial  $s_n(x) = nx(nx+1)^{-1}$  puede ser verificada por inducción matemática. Por inspección esta expresión para  $s_n(x)$  es válida para  $n = 1, 2$ . Suponemos que se mantiene para el término  $n$  y probamos para  $n+1$ .

$$\begin{aligned} s_{n+1} &= s_n + \frac{x}{[nx+1][(n+1)x+1]} \\ &= \frac{nx}{[nx+1]} + \frac{x}{[nx+1][(n+1)x+1]} \\ &= \frac{(n+1)x}{(n+1)x+1} , \end{aligned}$$

completando la prueba.

Tomando  $n \rightarrow \infty$  tenemos

$$\begin{aligned} S(0) &= \lim_{n \rightarrow \infty} s_n(0) = 0 , \\ S(x \neq 0) &= \lim_{n \rightarrow \infty} s_n(x \neq 0) = 1 . \end{aligned}$$

Tenemos una discontinuidad en el límite de la serie en  $x = 0$ . Sin embargo,  $s_n(x)$  es una función continua de  $x$ , en el intervalo  $0 \leq x < 1$ , para todo  $n$  finito. La ecuación (4.57) con  $\varepsilon$  suficientemente pequeño, será violado para todo  $n$  finito. Nuestra serie no converge uniformemente.

**4.5.2 Prueba  $M$  de Weierstrass.**

La prueba más comúnmente usada para la convergencia uniforme es la prueba  $M$  de Weierstrass. Si podemos construir una serie de números  $\sum_1^{\infty} M_i$ , en la cual  $M_i \geq |u_i(x)|$  para todo  $x$  en el intervalo  $[a, b]$  y  $\sum_1^{\infty} M_i$  es convergente, nuestra serie  $\sum_1^{\infty} u_i(x)$  será *uniformemente* convergente en  $[a, b]$ .

La prueba de este test  $M$  de Weierstrass es directa y simple. Ya que  $\sum_i M_i$  converge, existen algunos números  $N$  tal que  $n+1 \geq N$ ,

$$\sum_{i=n+1}^{\infty} M_i < \varepsilon . \quad (4.59)$$

Esto a partir de nuestra definición de convergencia. Entonces, con  $|u_i(x)| \leq M_i$  para todo  $x$  en el intervalo  $a \leq x \leq b$ ,

$$\sum_{i=n+1}^{\infty} |u_i(x)| < \varepsilon . \quad (4.60)$$

De modo que

$$|S(x) - s_n(x)| = \left| \sum_{i=n+1}^{\infty} u_i(x) \right| < \varepsilon , \quad (4.61)$$

y por definición  $\sum_1^{\infty} u_i(x)$  es uniformemente convergente en  $[a, b]$ . Ya que tenemos especificados valores absolutos en el planteamiento de la prueba  $M$  de Weierstrass, la serie  $\sum_1^{\infty} u_i(x)$  también es vista como serie *absolutamente* convergente.

Podemos notar que la convergencia uniforme y convergencia absoluta son propiedades independientes. Una no implica la otra. Para ejemplos específicos,

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n + x^2} , \quad -\infty < x < \infty \quad (4.62)$$

y

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} = \ln(1 + x) , \quad 0 \leq x \leq 1 , \quad (4.63)$$

converge uniformemente en los intervalos indicados pero no converge absolutamente. Por otra parte,

$$\sum_{n=1}^{\infty} (1 - x)x^n = \begin{cases} 1 , & 0 \leq x < 1 \\ 0 , & x = 1 \end{cases} , \quad (4.64)$$

converge absolutamente pero no uniformemente en  $[0, 1]$ .

A partir de la definición de convergencia uniforme podríamos mostrar que cualquier serie

$$f(x) = \sum_{n=1}^{\infty} u_n(x) , \quad (4.65)$$

no puede converger uniformemente en ningún intervalo que incluya una discontinuidad de  $f(x)$ .

Ya que la prueba  $M$  de Weierstrass establece tanto la convergencia uniforme como absoluta, necesariamente falla para series que son uniformes pero condicionalmente convergentes.

### 4.5.3 Prueba de Abel.

Una prueba algo más delicada para la convergencia uniforme ha sido dada por Abel. Si

$$u_n(x) = a_n f_n(x) , \\ \sum a_n = A , \quad \text{convergente,}$$

y las funciones  $f(x)$  son monótonas [ $f_{n+1}(x) \leq f_n(x)$ ] y acotadas,  $0 \leq f_n(x) \leq M$ , para todo  $x$  en  $[a, b]$ , entonces  $\sum u_n(x)$  converge uniformemente en  $[a, b]$ .

Las series uniformemente convergentes tienen tres propiedades particularmente útiles.

1. Si los términos individuales  $u_n(x)$  son continuos, la suma de la serie

$$f(x) = \sum_{n=1}^{\infty} u_n(x) , \quad (4.66)$$

es también continua.

2. Si los términos individuales  $u_n(x)$  son continuos, las series pueden ser integradas término a término. La suma de las integrales es igual a la integral de la suma.

$$\int_a^b f(x) dx = \sum_{n=1}^{\infty} \int_a^b u_n(x) dx . \quad (4.67)$$

3. Las derivadas de la suma de la serie  $f(x)$  es igual a la suma de los términos individuales derivados,

$$\frac{df(x)}{dx} = \sum_{n=1}^{\infty} \frac{du_n(x)}{dx} , \quad (4.68)$$

siempre que las siguientes condiciones sean satisfechas:

$$u_n(x) \text{ y } \frac{du_n(x)}{dx} \text{ son continuas en } [a, b].$$

$$\sum_{n=1}^{\infty} \frac{du_n(x)}{dx} \text{ es uniformemente convergente en } [a, b].$$

La integración término a término de una serie uniformemente convergente<sup>2</sup> requiere sólo continuidad de los términos individuales. Esta condición casi siempre es satisfecha en las aplicaciones físicas. La diferenciación término a término de una serie a menudo no es válida porque deben satisfacer condiciones más restrictivas. Por cierto, encontraremos casos en series de Fourier, en la cual la diferenciación término a término de una serie uniformemente convergente tiende a una serie divergente.

## 4.6 Expansión de Taylor.

Esta es una expansión de una función en una serie infinita o en una serie finita más un término remanente. Los coeficientes de los términos sucesivos de la serie involucra las derivadas sucesivas de la función. Este tipo de expansiones de son ampliamente usadas. Ahora derivaremos la expansión de Taylor.

Supongamos que nuestra función  $f(x)$  tiene una derivada  $n$ -ésima continua en el intervalo  $a \leq x \leq b$ . Entonces, integrando esta  $n$ -ésima derivada  $n$  veces,

$$\int_a^x f^{(n)}(x) dx = f^{(n-1)}(x) \Big|_a^x = f^{(n-1)}(x) - f^{(n-1)}(a)$$

$$\int_a^x \left( \int_a^x f^{(n)}(x) dx \right) dx = \int_a^x [f^{(n-1)}(x) - f^{(n-1)}(a)] dx \quad (4.69)$$

$$= f^{(n-2)}(x) - f^{(n-2)}(a) - (x-a)f^{(n-1)}(a) .$$

<sup>2</sup>La integración término a término también puede ser válida en ausencia de convergencia uniforme.

Continuando, obtenemos

$$\int \int \int_a^x f^{(n)}(x)(dx)^3 = f^{(n-3)}(x) - f^{(n-3)}(a) - (x-a)f^{(n-2)}(a) - \frac{(x-a)^2}{2}f^{(n-1)}(a) . \quad (4.70)$$

Finalmente, integrando por  $n$ -ésima vez,

$$\int_a^x \cdots \int f^{(n)}(x)(dx)^n = f(x) - f(a) - (x-a)f'(a) + \quad (4.71)$$

$$- \frac{(x-a)^2}{2!}f''(a) - \cdots - \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) .$$

Note que esta expresión es exacta. No hay términos que hayan sido excluidos, ni aproximaciones hechas. Ahora, resolviendo para  $f(x)$ , tenemos

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R_n . \quad (4.72)$$

El remanente,  $R_n$ , está dado por la integral  $n$ -dimensional

$$\int_a^x \cdots \int f^{(n)}(x)(dx)^n . \quad (4.73)$$

Este remanente, ecuación (4.73), puede ser puesto en una forma más inteligible usando la forma integral del teorema del valor medio

$$\int_a^x g(x) dx = (x-a)g(\xi) , \quad (4.74)$$

con  $a \leq \xi \leq x$ . Integrando  $n$  veces obtenemos la forma Lagrangiana del remanente:

$$R_n = \frac{(x-a)^n}{n!}f^{(n)}(\xi) . \quad (4.75)$$

Con la expansión de Taylor en esta forma no estamos interesados en cualquier pregunta de convergencia de series infinitas. Esta serie es finita, la sola pregunta que nos importa es la magnitud del remanente.

Cuando la función  $f(x)$  es tal que

$$\lim_{n \rightarrow \infty} R_n = 0 , \quad (4.76)$$

la ecuación (4.72) se convierte en la serie de Taylor

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots \quad (4.77)$$

$$= \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!}f^{(n)}(a) .$$

Nuestra serie de Taylor especifica el valor de una función en un punto,  $x$ , en términos del valor de la función y sus derivadas en un punto de referencia,  $a$ . Esta es una expansión en potencias de un cambio en la variable,  $\Delta x = x - a$  en este caso. La notación puede ser variada según la conveniencia del usuario. Con la sustitución  $x \rightarrow x + h$  y  $a \rightarrow x$  tenemos una forma alterna

$$f(x + h) = \sum_{n=0}^{\infty} \frac{h^n}{n!} f^{(n)}(x) .$$

Cuando usamos el operador  $D = d/dx$  la expansión de Taylor se convierte en

$$f(x + h) = \sum_{n=0}^{\infty} \frac{h^n D^n}{n!} f(x) = e^{hD} f(x) .$$

Una forma en operadores equivalente de la expansión de Taylor. Una derivación de la expansión de Taylor en el contexto de la teoría de variable compleja aparece en el próximo capítulo.

### 4.6.1 Teorema de Maclaurin.

Si expandimos alrededor del origen ( $a = 0$ ), la ecuación (4.77) es conocida como la serie de Maclaurin

$$\begin{aligned} f(x) &= f(0) + x f'(0) + \frac{x^2}{2!} f''(0) + \cdots \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} f^{(n)}(0) . \end{aligned} \tag{4.78}$$

Una aplicación inmediata de la serie de Maclaurin (o serie de Taylor) está en la expansión de varias funciones transcendentales en una serie infinita.

#### Ejemplo

Sea  $f(x) = e^x$ . Diferenciando, tenemos

$$f^{(n)}(0) = 1 , \tag{4.79}$$

para todo  $n$ ,  $n = 1, 2, 3, \dots$ . Entonces, para la ecuación (4.78), tenemos

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!} . \tag{4.80}$$

Esta es la expansión en serie de la función exponencial. Algunos autores usan esta serie para definir la función exponencial.

Aunque esta serie es claramente convergente para todo  $x$ , podríamos chequear el término remanente,  $R_n$ . Por la ecuación (4.75) tenemos

$$\begin{aligned} R_n &= \frac{x^n}{n!} f^{(n)}(\xi) \\ &= \frac{x^n}{n!} e^{\xi} , \quad 0 \leq |\xi| \leq x . \end{aligned} \tag{4.81}$$

Por lo tanto

$$|R_n| \leq \frac{x^n}{n!} e^x \quad (4.82)$$

y

$$\lim_{n \rightarrow \infty} R_n = 0 \quad (4.83)$$

para todo los valores finitos de  $x$ , el cual indica que esta expansión de Maclaurin de  $e^x$  es válida sobre el intervalo  $-\infty < x < \infty$ .

### Ejemplo

Sea  $f(x) = \ln(1+x)$ . Diferenciando, obtenemos

$$\begin{aligned} f'(x) &= \frac{1}{(1+x)}, \\ f^{(n)}(x) &= (-1)^{n-1} (n-1)! \frac{1}{(1+x)^n}. \end{aligned} \quad (4.84)$$

La expansión de Maclaurin produce

$$\begin{aligned} \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + R_n \\ &= \sum_{p=1}^n (-1)^{p-1} \frac{x^p}{p} + R_n. \end{aligned} \quad (4.85)$$

En este caso el remanente está dado por

$$\begin{aligned} R_n &= \frac{x^n}{n!} f^{(n)}(\xi), \quad 0 \leq \xi \leq x \\ &\leq \frac{x^n}{n}, \quad 0 \leq \xi \leq x \leq 1. \end{aligned} \quad (4.86)$$

Ahora el remanente se aproxima a cero cuando  $n$  crece indefinidamente, dado  $0 \leq x \leq 1^3$ . Como una serie infinita

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}, \quad (4.87)$$

la cual converge para  $-1 < x \leq 1$ . El intervalo  $-1 < x < 1$  es fácilmente establecido por la prueba de la razón de D' Alembert. La convergencia en  $x = 1$  se deduce a partir del criterio de Leibniz. En particular, en  $x = 1$ , tenemos

$$\begin{aligned} \ln 2 &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots \\ &= \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}, \end{aligned} \quad (4.88)$$

la serie armónica alterna condicionalmente convergente.

<sup>3</sup>Este intervalo puede ser fácilmente extendido a  $-1 < x \leq 1$  pero no a  $x = -1$ .

### 4.6.2 Teorema Binomial.

Una segunda, aplicación extremadamente importante de las expansiones de Taylor y Maclaurin es la derivación del teorema binomial para potencias negativas y/o no enteras.

Sea  $f(x) = (1+x)^m$ , en la cual  $m$  puede ser negativo y no está limitado a valores integrables. La aplicación directa de la ecuación (4.78) da

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \cdots + R_n . \quad (4.89)$$

Para esta función el remanente es

$$R_n = \frac{x^n}{n!}(1+\xi)^{m-n} \times m(m-1) \cdots (m-n+1) \quad (4.90)$$

y  $\xi$  con  $0 \leq \xi \leq x$ . Ahora, para  $n > m$ ,  $(1+\xi)^{m-n}$  es un máximo para  $\xi = 0$ . Por lo tanto

$$R_n \leq \frac{x^n}{n!} \times m(m-1) \cdots (m-n+1) . \quad (4.91)$$

Note que los factores dependientes de  $m$  no dan un cero a menos que  $m$  sea entero no negativo;  $R_n$  tiende a cero cuando  $n \rightarrow \infty$  si  $x$  está restringido al intervalo  $0 \leq x \leq 1$ . La expansión binomial resulta

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \cdots . \quad (4.92)$$

En otra, notación equivalente

$$\begin{aligned} (1+x)^m &= \sum_{n=0}^{\infty} \frac{m!}{n!(m-n)!} x^n \\ &= \sum_{n=0}^{\infty} \binom{m}{n} x^n . \end{aligned} \quad (4.93)$$

La cantidad  $\binom{m}{n}$ , la cual igual a  $m!/(n!(m-n)!)$  es llamado el *coeficiente binomial*. Aunque hemos mostrado solamente que el remanente se anula,

$$\lim_{n \rightarrow \infty} R_n = 0 ,$$

para  $0 \leq x < 1$ , la serie en la ecuación (4.92) realmente puede mostrarse que converge en el intervalo extendido  $-1 < x < 1$ . Para  $m$  un entero,  $(m-n)! = \pm\infty$  si  $n > m$  y las series automáticamente terminan en  $n = m$ .

#### Ejemplo Energía relativista.

La energía total relativista de una partícula es

$$E = mc^2 \left(1 - \frac{v^2}{c^2}\right)^{-1/2} . \quad (4.94)$$

Comparemos esta ecuación con la energía cinética clásica,  $\frac{1}{2}mv^2$ .

Por la ecuación (4.92) con  $x = -\frac{v^2}{c^2}$  y  $m = -\frac{1}{2}$  tenemos

$$E = mc^2 \left[ 1 - \frac{1}{2} \left( -\frac{v^2}{c^2} \right) + \frac{(-1/2)(-3/2)}{2!} \left( -\frac{v^2}{c^2} \right)^2 + \frac{(-1/2)(-3/2)(-5/2)}{3!} \left( -\frac{v^2}{c^2} \right)^3 + \dots \right].$$

o

$$E = mc^2 + \frac{1}{2}mv^2 + \frac{3}{8}mv^2 \frac{v^2}{c^2} + \frac{5}{16}mv^2 \left( \frac{v^2}{c^2} \right)^2 + \dots \quad (4.95)$$

El primer término,  $mc^2$ , lo identificamos como la masa en reposo. Entonces

$$E_{\text{cinética}} = \frac{1}{2}mv^2 \left[ 1 + \frac{3}{4} \frac{v^2}{c^2} + \frac{5}{8} \left( \frac{v^2}{c^2} \right)^2 + \dots \right]. \quad (4.96)$$

Para la velocidad de la partícula  $v \ll c$ , donde  $c$  es la velocidad de la luz, la expresión en los paréntesis cuadrados se reduce a la unidad y vemos que la porción cinética de la energía relativista total concuerda con el resultado clásico.

Para polinomios podemos generalizar la expansión binomial a

$$(a_1 + a_2 + \dots + a_m)^m = \sum \frac{n!}{n_1!n_2! \dots n_m!} a_1^{n_1} a_2^{n_2} \dots a_m^{n_m},$$

donde la suma incluye todas las combinaciones diferentes de  $n_1, n_2, \dots, n_m$  con  $\sum_{i=1}^m n_i = n$ . Aquí  $n_i$  y  $n$  son enteros. Esta generalización encuentra considerables usos en Mecánica Estadística.

Las series de Maclaurin puede aparecer algunas veces indirectamente más que el uso directo de la ecuación (4.78). Por ejemplo, la manera más conveniente para obtener la expansión en serie

$$\text{sen}^{-1} x = \sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!} \frac{x^{2n+1}}{2n+1} = x + \frac{x^3}{6} + \frac{3x^5}{40} + \dots, \quad (4.97)$$

es hacer uso de la relación

$$\text{sen}^{-1} x = \int_0^x \frac{dt}{(1-t^2)^{1/2}}.$$

Expandimos  $(1-t^2)^{-1/2}$  (teorema binomial) y luego integramos término a término. Esta integración término a término es discutida en la sección 4.7. El resultado es la ecuación (4.97). Finalmente, podemos tomar el límite cuando  $x \rightarrow 1$ . La serie converge por la prueba de Gauss.

### 4.6.3 Expansión de Taylor de más de una variable.

La función  $f$  tiene más de una variable independiente, es decir,  $f = f(x, y)$ , la expansión de Taylor se convierte en

$$\begin{aligned}
 f(x, y) = & f(a, b) + (x - a) \frac{\partial f}{\partial x} + (y - b) \frac{\partial f}{\partial y} + \\
 & + \frac{1}{2!} \left[ (x - a)^2 \frac{\partial^2 f}{\partial x^2} + 2(x - a)(y - b) \frac{\partial^2 f}{\partial x \partial y} + (y - b)^2 \frac{\partial^2 f}{\partial y^2} \right] + \\
 & + \frac{1}{3!} \left[ (x - a)^3 \frac{\partial^3 f}{\partial x^3} + 3(x - a)^2(y - b) \frac{\partial^3 f}{\partial x^2 \partial y} + \right. \\
 & \left. + 3(x - a)(y - b)^2 \frac{\partial^3 f}{\partial x \partial y^2} + (y - b)^3 \frac{\partial^3 f}{\partial y^3} \right] + \dots,
 \end{aligned} \tag{4.98}$$

con todas las derivadas evaluadas en el punto  $(a, b)$ . Usando  $\alpha_j t = x_j - x_{j0}$ , podemos escribir la expansión de Taylor para  $m$  variables independientes en la forma simbólica

$$f(x_j) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left( \sum_{i=1}^m \alpha_i \frac{\partial}{\partial x_i} \right)^n f(x_k) \Big|_{x_k=x_{k0}}. \tag{4.99}$$

Una forma vectorial conveniente es

$$\psi(\vec{r} + \vec{a}) = \sum_{n=0}^{\infty} \frac{1}{n!} (\vec{a} \cdot \vec{\nabla})^n \psi(\vec{r}). \tag{4.100}$$

## 4.7 Series de potencias.

Las series de potencias son un tipo especial y extremadamente útil de series infinitas de la forma

$$\begin{aligned}
 f(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots \\
 &= \sum_{n=0}^{\infty} a_n x^n,
 \end{aligned} \tag{4.101}$$

donde los coeficientes  $a_i$  son constantes e independientes de  $x$ .<sup>4</sup>

### 4.7.1 Convergencia.

La ecuación (4.101) puede testearse rápidamente para la convergencia ya sea por la prueba de la raíz de Cauchy o por la prueba de la razón de D' Alembert. Si

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = R^{-1}, \tag{4.102}$$

<sup>4</sup>La ecuación (4.101) puede ser reescrita con  $z = x + iy$ , reemplazando a  $x$ . Luego todos los resultados de esta sección se aplican a series complejas

la serie converge para  $-R < x < R$ . Este es el intervalo o radio de convergencia. Ya que las pruebas de la raíz y la razón falla cuando el límite es la unidad, el punto final del intervalo requiere atención especial.

Por ejemplo, si  $a_n = n^{-1}$ , entonces  $R = 1$  y, la serie converge para  $x = -1$  pero diverge para  $x = +1$ . Si  $a_n = n!$ , entonces  $R = 0$  y la serie diverge para todo  $x \neq 0$ .

## 4.8 Convergencia uniforme y absoluta.

Supongamos que nuestra serie de potencia sea convergente para  $-R < x < R$ ; entonces será uniforme y absolutamente convergente en cualquier intervalo interior,  $-S \leq x \leq S$ , donde  $0 < S < R$ . Esto podría ser probado directamente por la prueba  $M$  de Weierstrass usando  $M_i = |a_i|S^i$ .

### 4.8.1 Continuidad.

Ya que cada término  $u_n(x) = a_n x^n$  es una función continua de  $x$  y  $f(x) = \sum a_n x^n$  converge uniformemente para  $-S \leq x \leq S$ ,  $f(x)$  debería ser una función continua en el intervalo de convergencia uniforme. Este comportamiento es contradictorio con el comportamiento impresionantemente diferente de las series de Fourier, en el cual las series de Fourier son usadas frecuentemente para representar funciones discontinuas tales como ondas cuadradas y ondas dientes de sierra.

### 4.8.2 Diferenciación e integración.

Con  $u_n(x)$  continua y  $\sum a_n x^n$  uniformemente convergente, encontramos que la serie diferenciada es una serie de potencia con funciones continuas y del mismo radio de convergencia que la serie original. Los nuevos factores introducidos por diferenciación (o integración) no afecta ni a la prueba de la raíz ni a la de la razón. Por lo tanto nuestra serie podría ser diferenciada o integrada tan a menudo como uno deseamos dentro del intervalo de convergencia uniforme.

En vista de las restricciones algo severas puestas en la diferenciación, esto es un resultado valioso y notable.

### 4.8.3 Teorema de la singularidad.

En la sección precedente, usando las series de Maclaurin, expandimos  $e^x$  y  $\ln(1+x)$  en series infinitas. En los capítulos venideros las funciones son frecuentemente representadas e incluso definidas por series infinitas. Ahora estableceremos que la representación de la serie de potencias es única.

Si

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} a_n x^n, & -R_a < x < R_a \\ &= \sum_{n=0}^{\infty} b_n x^n, & -R_b < x < R_b, \end{aligned} \quad (4.103)$$

con intervalos de convergencia sobrepuestos, incluyendo el origen, luego

$$a_n = b_n, \quad (4.104)$$

para todo  $n$ ; esto es, supongamos dos representaciones de serie de potencias (diferentes) y luego procedamos a demostrar que las dos son idénticas.

De la ecuación (4.103)

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n, \quad -R < x < R \quad (4.105)$$

donde  $R$  es el más pequeño entre  $R_a$ ,  $R_b$ . Haciendo  $x = 0$  para eliminar todo salvo el término constante, obtenemos

$$a_0 = b_0. \quad (4.106)$$

Ahora, aprovechandose de la diferenciabilidad de nuestra serie de potencia, diferenciamos la ecuación (4.105), obteniendo

$$\sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=1}^{\infty} n b_n x^{n-1}. \quad (4.107)$$

De nuevo ajustamos  $x = 0$  para aislar el nuevo término constante y encontramos

$$a_1 = b_1. \quad (4.108)$$

Repitiendo este proceso  $n$  veces, obtenemos

$$a_n = b_n, \quad (4.109)$$

lo cual muestra que las dos series coinciden. Por lo tanto nuestra representación en serie de potencia es única.

Esto será un punto crucial cuando usamos una serie de potencia para desarrollar soluciones de ecuaciones diferenciales. Esta unicidad de las series de potencia aparece frecuentemente en física teórica. La teoría de perturbaciones en Mecánica Cuántica es un ejemplo de esto. La representación en serie de potencia de funciones es a menudo útil en formas de evaluación indeterminadas, particularmente cuando la regla de l'Hospital puede ser inconveniente de aplicar.

**Ejemplo**

Evalúemos

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} . \quad (4.110)$$

Remplazando  $\cos x$  por su expansión en serie de Maclaurin, obtenemos

$$\begin{aligned} \frac{1 - \cos x}{x^2} &= \frac{1 - (1 - x^2/2! + x^4/4! - \dots)}{x^2} \\ &= \frac{x^2/2! - x^4/4! + \dots}{x^2} \\ &= \frac{1}{2!} - \frac{x^2}{4!} + \dots . \end{aligned}$$

Tomando  $x \rightarrow 0$ , tenemos

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \frac{1}{2} . \quad (4.111)$$

La unicidad de las series de potencia significa que los coeficientes  $a_n$  pueden ser identificadas con las derivadas en una serie de Maclaurin. A partir de

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) x^n$$

tenemos

$$a_n = \frac{1}{n!} f^{(n)}(0) .$$

**4.8.4 Inversión de series de potencia.**

Supongamos que dada una serie

$$\begin{aligned} y - y_0 &= a_1(x - x_0) + a_2(x - x_0)^2 + \dots \\ &= \sum_{n=1}^{\infty} a_n(x - x_0)^n . \end{aligned} \quad (4.112)$$

Esta dada  $(y - y_0)$  en términos de  $(x - x_0)$ . Sin embargo, podría ser deseable tener una expresión explícita para  $(x - x_0)$  en términos de  $(y - y_0)$ . Podríamos resolver la ecuación (4.112) para  $(x - x_0)$  por inversión de nuestra serie. Supongamos que

$$x - x_0 = \sum_{n=0}^{\infty} b_n (y - y_0)^n , \quad (4.113)$$

con  $b_n$  determinado en términos de la supuestamente conocidos  $a_n$ . Una aproximación a fuerza bruta, la cual es perfectamente adecuada para los primeros pocos coeficientes, es simplemente

sustituir la ecuación (4.112) en la ecuación (4.113). Igualando los coeficientes de  $(x - x_0)^n$  en ambos lados de la ecuación (4.113), ya que la serie de potencia es única, obtenemos

$$\begin{aligned} b_1 &= \frac{1}{a_1}, \\ b_2 &= \frac{a_2}{a_1^3}, \\ b_3 &= \frac{1}{a_1^5}(2a_2^2 - a_1a_3), \\ b_4 &= \frac{1}{a_1^7}(5a_1a_2a_3 - a_1^2a_4 - 5a_2^3), \quad \text{y así sucesivamente.} \end{aligned} \tag{4.114}$$

Los coeficientes mayores son listados en tablas generalmente. Una aproximación más general y mucho más elegante es desarrollada usando variables complejas.

## 4.9 Integrales elípticas.

Las integrales elípticas son incluidas aquí parcialmente como una ilustración del uso de las series de potencias y por su propio interés intrínseco. Este interés incluye la ocurrencia de las integrales elípticas en problemas físicos y aplicaciones en problemas matemáticos.

**Ejemplo** Período de un péndulo simple.

Para pequeñas oscilaciones de amplitud nuestro péndulo, figura 4.8 tiene un movimiento armónico simple con un período  $T = 2\pi(l/g)^{1/2}$ . Para una amplitud grande  $\theta_m$  tal que  $\sin \theta_m \neq \theta_m$ , la segunda ley de movimiento de Newton y las ecuaciones de Lagrange conducen a una ecuación diferencial no lineal ( $\sin \theta$  es una función no lineal de  $\theta$ ), así que tomemos un acercamiento diferente.

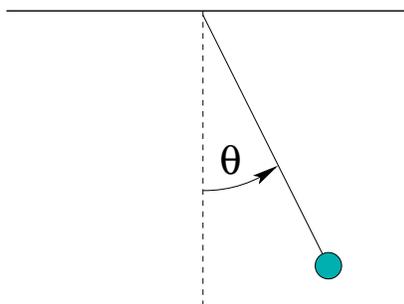


Figura 4.8: Péndulo simple.

La masa oscilante  $m$  tiene una energía cinética de  $\frac{1}{2}ml^2(d\theta/dt)^2$  y una energía potencial de  $-mgl \cos \theta$  ( $\theta = \pi/2$  como la elección del cero de la energía potencial). Ya que  $d\theta/dt = 0$  en  $\theta = \theta_m$ , el principio de la conservación de la energía da

$$\frac{1}{2}ml^2 \left( \frac{d\theta}{dt} \right)^2 - mgl \cos \theta = -mgl \cos \theta_M. \tag{4.115}$$

Resolviendo para  $d\theta/dt$  obtenemos

$$\frac{d\theta}{dt} = \pm \left(\frac{2g}{l}\right)^{1/2} (\cos \theta - \cos \theta_M)^{1/2} \quad (4.116)$$

con la cancelación de la masa  $m$ . Tomando  $t$  como cero cuando  $\theta = 0$  y  $d\theta/dt > 0$ . Una integración desde  $\theta = 0$  a  $\theta = \theta_m$  produce

$$\int_0^{\theta_M} (\cos \theta - \cos \theta_M)^{-1/2} d\theta = \left(\frac{2g}{l}\right)^{1/2} \int_0^t dt = \left(\frac{2g}{l}\right)^{1/2} t. \quad (4.117)$$

Esto es 1/4 del ciclo, y por lo tanto el tiempo  $t$  es 1/4 del período,  $T$ . Notemos que  $\theta \leq \theta_m$ , trataremos la sustitución

$$\text{sen} \left(\frac{\theta}{2}\right) = \text{sen} \left(\frac{\theta_M}{2}\right) \text{sen} \varphi. \quad (4.118)$$

Con esto, la ecuación (4.117) se convierte en

$$T = 4 \left(\frac{l}{g}\right)^{1/2} \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1 - \text{sen}^2 \left(\frac{\theta}{2}\right) \text{sen}^2 \varphi}} \quad (4.119)$$

Aunque no hay un obvio mejoramiento en la ecuación (4.117), la integral ahora define la integral elíptica completa del primer tipo,  $K(\text{sen} \theta_m/2)$ . A partir de la expansión de serie, el período de nuestro péndulo puede ser desarrollado como una serie de potencia en  $\text{sen} \theta_m/2$ :

$$T = 2\pi \left(\frac{l}{g}\right)^{1/2} \left[ 1 + \frac{1}{4} \text{sen}^2 \frac{\theta_M}{2} + \frac{9}{64} \text{sen}^4 \frac{\theta_M}{2} + \dots \right] \quad (4.120)$$

### 4.9.1 Definiciones.

Generalizando el ejemplo anterior para incluir el límite superior como una variable, la *integral elíptica del primere tipo* está definida como

$$F(\varphi \backslash \alpha) = \int_0^\varphi \frac{d\theta}{\sqrt{1 - \text{sen}^2 \alpha \text{sen}^2 \theta}} \quad (4.121)$$

o

$$F(x|m) = \int_0^x \frac{dt}{\sqrt{(1-t^2)(1-mt^2)}}, \quad 0 \leq m < 1. \quad (4.122)$$

Para  $\varphi = \pi/2$ ,  $x = 1$ , tenemos la *integral elíptica completa de primer tipo*,

$$\begin{aligned} K(m) &= \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - m \text{sen}^2 \theta}} \\ &= \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-mt^2)}}, \end{aligned} \quad (4.123)$$

con  $m = \text{sen}^2 \alpha$ ,  $0 \leq m < 1$ .

La *integral elíptica de segundo tipo* está definida por

$$E(\varphi|\alpha) = \int_0^\varphi \sqrt{1 - \text{sen}^2 \alpha \text{sen}^2 \theta} d\theta \quad (4.124)$$

o

$$E(x|m) = \int_0^x \sqrt{\frac{1 - mt^2}{1 - t^2}} dt, \quad 0 \leq m < 1 \quad (4.125)$$

Nuevamente, para el caso  $\varphi = \pi/2$ ,  $x = 1$ , tenemos la *integral elíptica completa de segundo tipo*:

$$\begin{aligned} E(m) &= \int_0^{\pi/2} \sqrt{1 - m \text{sen}^2 \theta} d\theta \\ &= \int_0^1 \sqrt{\frac{1 - mt^2}{1 - t^2}} dt, \quad 0 \leq m < 1. \end{aligned} \quad (4.126)$$

La figura 5.9 muestra el comportamiento de  $K(m)$  y  $E(m)$ . Los valores de ambas funciones pueden encontrarse en tablas o evaluar en software como Mathematica.

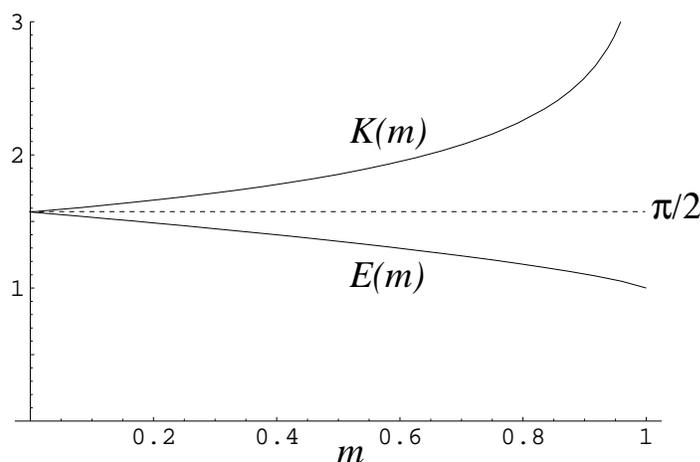


Figura 4.9: Integrales elípticas completas,  $K(m)$ ,  $E(m)$ .

### 4.9.2 Expansión de series.

Para nuestro intervalo  $0 \leq m < 1$ , el denominador de  $K(m)$  puede ser expandido en serie binomial

$$\begin{aligned} (1 - m \text{sen}^2 \theta)^{-1/2} &= 1 + \frac{1}{2} m \text{sen}^2 \theta + \frac{3}{8} m^2 \text{sen}^4 \theta + \dots \\ &= \sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!} m^n \text{sen}^{2n} \theta. \end{aligned} \quad (4.127)$$

Para cualquier intervalo cerrado  $[0, m_{\max}]$ ,  $m_{\max} < 1$  esta serie es uniformemente convergente y puede ser integrada término a término.

$$\int_0^{\pi/2} \text{sen}^{2n} \theta \, d\theta = \frac{(2n-1)!!}{(2n)!!} \frac{\pi}{2}. \quad (4.128)$$

De modo que

$$K(m) = \frac{\pi}{2} \left[ 1 + \left(\frac{1}{2}\right)^2 m + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 m^2 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 m^3 + \dots \right] \quad (4.129)$$

Similarmente,

$$E(m) = \frac{\pi}{2} \left[ 1 - \left(\frac{1}{2}\right)^2 \frac{m}{1} - \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 \frac{m^2}{3} - \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 \frac{m^3}{5} - \dots \right] \quad (4.130)$$

Más adelante estas series son identificadas como funciones hipergeométricas, y tenemos

$$K(m) = \frac{\pi}{2} {}_2F_1 \left( \frac{1}{2}, \frac{1}{2}, 1; m \right) \quad (4.131)$$

$$E(m) = \frac{\pi}{2} {}_2F_1 \left( -\frac{1}{2}, \frac{1}{2}, 1; m \right) \quad (4.132)$$

### 4.9.3 Valores límites.

De las series de las ecuaciones (4.129) y (4.130), o a partir de las integrales definidas,

$$\lim_{m \rightarrow 0} K(m) = \frac{\pi}{2}, \quad (4.133)$$

$$\lim_{m \rightarrow 0} E(m) = \frac{\pi}{2}. \quad (4.134)$$

Para  $m \rightarrow 1$  las expansiones de series son de poco uso. Sin embargo, la integrales tienden

$$\lim_{m \rightarrow 1} K(m) = \infty, \quad (4.135)$$

la integral diverge logarítmicamente, y

$$\lim_{m \rightarrow 1} E(m) = 1. \quad (4.136)$$

Las integrales elípticas han sido usadas ampliamente en el pasado para evaluar integrales. Por ejemplo, integrales de la forma

$$I = \int_0^x R(t, \sqrt{a_4 t^4 + a_3 t^3 + a_2 t^2 + a_1 t + a_0}) \, dt,$$

donde  $R$  es una función rotacional de  $t$  y del radical, puede ser expresado en términos de integrales elípticas. Con los computadores de alta velocidad disponibles para la evaluación numérica directa, el interés en estas técnicas de integrales elípticas ha declinado. Sin embargo, las integrales elípticas todavía mantienen el interés a causa de su apariencia en problemas en Físicos.

## 4.10 Números de Bernoulli.

Los números de Bernoulli fueron introducidos por Jacques Bernoulli. Hay muchas definiciones equivalentes, pero debe tenerse extremo cuidado, porque algunos autores introducen variaciones en la numeración o en signo. Un acercamiento relativamente simple para definir los números de Bernoulli es por la serie<sup>5</sup>

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n x^n}{n!}, \quad (4.137)$$

la cual converge para  $|x| < 2\pi$  usando el test del cociente. Diferenciando esta serie de potencia repetidamente y luego evaluando para  $x = 0$ , obtenemos

$$B_n = \left[ \frac{d^n}{dx^n} \left( \frac{x}{e^x - 1} \right) \right]_{x=0}. \quad (4.138)$$

Específicamente,

$$B_1 = \frac{d}{dx} \left( \frac{x}{e^x - 1} \right) \Big|_{x=0} = \frac{x}{e^x - 1} - \frac{x e^x}{(e^x - 1)^2} \Big|_{x=0} = -\frac{1}{2}, \quad (4.139)$$

como puede ser visto por la expansión en series de los denominadores. Usando  $B_0 = 1$  y  $B_1 = -1/2$ , es fácil verificar que la función

$$\frac{x}{e^x - 1} - 1 + \frac{x}{2} = \sum_{n=0}^{\infty} \frac{B_n x^n}{n!} = -x(e^{-x} - 1) - 1 - \frac{x}{2}, \quad (4.140)$$

es par en  $x$ , tal que todos los  $B_{2n+1} = 0$ .

Para derivar una relación de recurrencia para los números de Bernoulli, multiplicamos

$$\begin{aligned} \frac{e^x - 1}{x} \frac{x}{e^x - 1} &= 1 = \left[ \sum_{m=0}^{\infty} \frac{x^m}{(m+1)!} \right] \left[ 1 - \frac{x}{2} + \sum_{n=1}^{\infty} \frac{B_{2n} x^{2n}}{(2n)!} \right] \\ &= 1 + \sum_{m=1}^{\infty} x^m \left[ \frac{1}{(m+1)!} - \frac{1}{2m!} \right] + \sum_{N=2}^{\infty} x^N \sum_{1 \leq n \leq N/2} \frac{B_{2n}}{[(2n)!(N-2n+1)!]}. \end{aligned} \quad (4.141)$$

La ecuación (4.141) produce

$$\frac{1}{2}(N+1) - 1 = \sum_{1 \leq n \leq N/2} B_{2n} \binom{N+1}{2n} = \frac{1}{2}(N-1), \quad (4.142)$$

la cual es equivalente a

$$\begin{aligned} N - \frac{1}{2} &= \sum_{n=1}^N B_{2n} \binom{2N+1}{2n}, \\ N - 1 &= \sum_{n=1}^{N-1} B_{2n} \binom{2N}{2n}. \end{aligned} \quad (4.143)$$

---

<sup>5</sup>La función  $\frac{x}{e^x - 1}$  puede ser considerada una *función generatriz* ya que genera los números de Bernoulli.

$n$	$B_n$	$B_n$
0	1	1.0000 00000
1	$-\frac{1}{2}$	-0.5000 00000
1	$\frac{1}{6}$	0.1666 66667
1	$-\frac{1}{30}$	-0.0333 33333
1	$\frac{1}{42}$	0.0238 09524
1	$-\frac{1}{30}$	-0.0333 33333
1	$\frac{5}{66}$	0.0757 57576

Tabla 4.1: Números de Bernoulli

A partir de la ecuación (4.143) los números de Bernoulli en la tabla 4.1 se obtienen rápidamente. Si la variable  $x$  en la ecuación (4.137) es remplazada por  $2xi$  (y  $B_1$  elegido igual a  $-1/2$ ), obtenemos una definición alternativa (y equivalente) de  $B_{2n}$ , la expresión

$$x \cot x = \sum_{n=0}^{\infty} (-1)^n B_{2n} \frac{(2x)^{2n}}{(2n)!}, \quad -\pi < x < \pi. \quad (4.144)$$

Usando el método del residuo o trabajando a partir de la representación de producto infinito de  $\operatorname{sen}(x)$ , encontramos que

$$B_{2n} = \frac{(-1)^{n-1} 2(2n)!}{(2\pi)^{2n}} \sum_{p=1}^{\infty} \frac{1}{p^{2n}}, \quad n = 1, 2, 3, \dots. \quad (4.145)$$

Esta representación de los números de Bernoulli fue descubierta por Euler. Es fácil ver a partir de la ecuación (4.145) que  $|B_{2n}|$  aumenta sin límite cuando  $n \rightarrow \infty$ . Ilustrando el comportamiento divergente de los números de Bernoulli, tenemos

$$\begin{aligned} B_{20} &= -5.291 \times 10^2 \\ B_{200} &= -3.647 \times 10^{215}. \end{aligned}$$

Algunos autores prefieren definir los números de Bernoulli con una versión modificada de la ecuación (4.145) usando

$$B_{2n} = \frac{2(2n)!}{(2\pi)^{2n}} \sum_{p=1}^{\infty} \frac{1}{p^{2n}}, \quad (4.146)$$

el subíndice es justo la mitad de nuestro subíndice original y todos los signos son positivos. Nuevamente, se debe chequear cuidadosamente la definición que se está usando de los números de Bernoulli.

Los números de Bernoulli aparecen frecuentemente en teoría de números. El teorema de von Standt-Clausen establece que

$$B_{2n} = A_n - \frac{1}{p_1} - \frac{1}{p_2} - \frac{1}{p_3} - \dots - \frac{1}{p_k}, \quad (4.147)$$

en el cual  $A_n$  es un entero y  $p_1, p_2, \dots, p_k$  son números primos tal que  $p_i - 1$  es un divisor de  $2n$ . Podemos fácilmente verificar que esto se mantiene para

$$\begin{aligned} B_6(A_3 = 1, p = 2, 3, 7) , \\ B_8(A_4 = 1, p = 2, 3, 5) , \\ B_{10}(A_5 = 1, p = 2, 3, 11) , \end{aligned} \quad (4.148)$$

y otros casos especiales.

Los números de Bernoulli aparecen en la suma de potencias enteras de enteros,

$$\sum_{j=1}^N j^p , \quad p \text{ entero.}$$

y en numerosas expansiones de series de las funciones trascendentales, incluyendo  $\tan x$ ,  $\cot x$ ,  $\sin^{-1} x$ ,  $\ln |\sin x|$ ,  $\ln |\cos x|$ ,  $\ln |\tan x|$ ,  $\tanh x$ ,  $\coth x$  y  $\cosh^{-1} x$ . Por ejemplo,

$$\tan(x) = x + \frac{x^3}{3} + \frac{2}{15}x^5 + \dots + \frac{(-1)^{n-1}2^{2n}(2^{2n} - 1)B_{2n}}{(2n)!}x^{2n-1} + \dots . \quad (4.149)$$

Los números de Bernoulli probablemente vengan en tales expansiones en series a causa de las ecuaciones de definición (4.137) y (4.143) y de su relación a la función zeta de Riemann

$$\zeta(2n) = \sum_{p=1}^{\infty} \frac{1}{p^{2n}} . \quad (4.150)$$

### 4.10.1 Funciones de Bernoulli.

Si la ecuación (4.137) puede ser fácilmente generalizada, tenemos

$$\frac{x e^{xs}}{e^x - 1} = \sum_{n=0}^{\infty} B_n(s) \frac{x^n}{n!} . \quad (4.151)$$

definiendo las *funciones de Bernoulli*,  $B_n(s)$ . Las primeras siete funciones de Bernoulli están dadas en la tabla 4.2.

De la función generadora, ecuación (4.151),

$$B_n(0) = B_n , \quad n = 1, 2, \dots . \quad (4.152)$$

la función de Bernoulli evaluadas en cero igual a al correspondiente número de Bernoulli. Dos propiedades particularmente importantes de las funciones de Bernoulli se deducen a partir de la definición: una relación de diferenciación

$$B'_n(s) = n B_{n-1}(s) , \quad n = 1, 2, \dots . \quad (4.153)$$

y una relación e simetría

$$B_n(1) = (-1)^n B_n(0) , \quad n = 1, 2, \dots . \quad (4.154)$$

Estas relaciones son usadas en el desarrollo de la fórmula de integración de Euler-Maclaurin.

---

$B_0$	$=$	$1$
$B_1$	$=$	$x - \frac{1}{2}$
$B_2$	$=$	$x^2 - x + \frac{1}{6}$
$B_3$	$=$	$x^3 - \frac{3}{2}x^2 + \frac{1}{2}x$
$B_4$	$=$	$x^4 - 2x^3 + x^2 - \frac{1}{30}$
$B_5$	$=$	$x^5 - \frac{5}{2}x^4 + \frac{5}{3}x^2 - \frac{1}{6}x$
$B_6$	$=$	$x^6 - 3x^5 + \frac{5}{2}x^4 - \frac{1}{2}x^2 + \frac{1}{42}$

---

Tabla 4.2: Funciones de Bernoulli

### 4.10.2 Fórmula de integración de Euler-Maclaurin.

Uno de los usos de las funciones de Bernoulli es la derivación de la fórmula de integración de Euler-Maclaurin. Esta fórmula es usada en el desarrollo de una expresión asintótica para la función factorial- serie de Stirling. La técnica es integración por partes repetida usando la ecuación (4.153) para crear nuevas derivadas. Comenzamos con

$$\int_0^1 f(x) dx = \int_0^1 f(x)B_0(x) dx . \quad (4.155)$$

A partir de la ecuación (4.153)

$$B_1'(x) = B_0(x) = 1 . \quad (4.156)$$

Sustituyendo  $B_1'(x)$  en la ecuación (4.155) e integrando por partes, obtenemos

$$\begin{aligned} \int_0^1 f(x) dx &= f(1)B_1(1) - f(0)B_1(0) - \int_0^1 f'(x)B_1(x) dx \\ &= \frac{1}{2}[f(1) - f(0)] - \int_0^1 f'(x)B_1(x) dx \end{aligned} \quad (4.157)$$

Nuevamente, usando la ecuación (4.153), tenemos

$$B_1(x) = \frac{1}{2}B_2'(x) , \quad (4.158)$$

e integrando por partes

$$\begin{aligned} \int_0^1 f(x) dx &= \frac{1}{2}[f(1) - f(0)] - \frac{1}{2!}[f'(1)B_2(1) - f'(0)B_2(0)] + \\ &\quad \frac{1}{2!} \int_0^1 f^{(2)}(x)B_2(x) dx . \end{aligned} \quad (4.159)$$

Usando las relaciones,

$$\begin{aligned} B_{2n}(1) &= B_{2n}(0) = B_{2n} , & n &= 0, 1, 2, \dots \\ B_{2n+1}(1) &= B_{2n+1}(0) = 0 , & n &= 1, 2, 3, \dots , \end{aligned} \quad (4.160)$$

y continuando este proceso, tenemos

$$\int_0^1 f(x) dx = \frac{1}{2}[f(1) - f(0)] - \sum_{p=1}^q \frac{1}{(2p)!} B_{2p} [f^{(2p-1)}(1) - f^{(2p-1)}(0)] + \frac{1}{(2q)!} \int_0^1 f^{(2q)}(x) B_{2q}(x) dx . \quad (4.161)$$

Esta es la fórmula de integración de Euler-Maclaurin. Supone que la función  $f(x)$  tiene las derivadas requeridas.

El intervalo de integración en la ecuación (4.161) puede ser trasladado de  $[0, 1]$  a  $[1, 2]$  reemplazando  $f(x)$  por  $f(x + 1)$ . Sumando tales resultados hasta  $[n - 1, n]$ ,

$$\int_0^n f(x) dx = \frac{1}{2}f(0) + f(1) + f(2) + \cdots + f(n-1) + \frac{1}{2}f(n) + - \sum_{p=1}^q \frac{1}{(2p)!} B_{2p} [f^{(2p-1)}(n) - f^{(2p-1)}(0)] + \frac{1}{(2p)!} \int_0^1 B_{2q}(x) \sum_{\nu=0}^{n-1} f^{(2q)}(x + \nu) dx . \quad (4.162)$$

Los términos  $\frac{1}{2}f(0) + f(1) + \cdots + \frac{1}{2}f(n)$  aparecen exactamente como una integración trapezoidal o cuadratura. La suma sobre  $p$  puede ser interpretada como una corrección a la aproximación trapezoidal. La ecuación (4.162) es la forma usada en la derivación de la fórmula de Stirling.

La fórmula de Euler-Maclaurin es a menudo útil para sumar series al convertirlas en integrales.

### 4.10.3 Función zeta de Riemann.

Estas series  $\sum_{p=1}^{\infty} p^{-2n}$  fueron usadas como series de comparación para probar la convergencia y en la ecuación (4.144) como una definición de los números de Bernoulli,  $B_{2n}$ . También sirve para definir la función zeta de Riemann por

$$\zeta(s) \equiv \sum_{n=1}^{\infty} \frac{1}{n^s} , \quad s > 1 . \quad (4.163)$$

La tabla 4.3 muestra los valores de  $\zeta(s)$  para  $s$  entero,  $s = 2, 3, \dots, 10$ . La figura 4.10 es un gráfico de  $\zeta(s) - 1$ . Una expresión integral para esta función zeta de Riemann aparecerá como parte del desarrollo de la función gama.

Otra interesante expresión para la función zeta puede ser derivada como

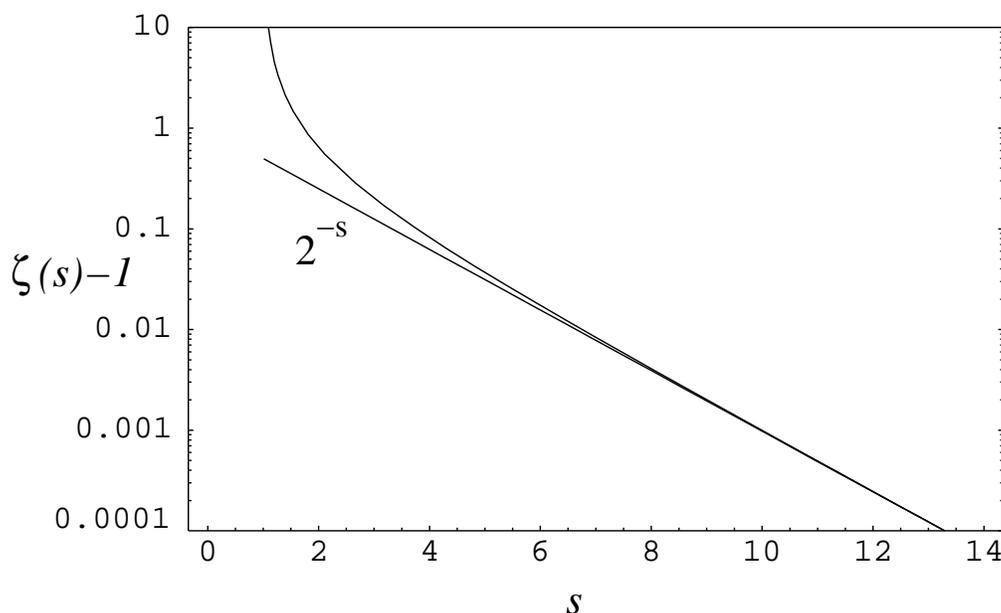
$$\zeta(s)(1 - 2^{-s}) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots - \left( \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \cdots \right) \quad (4.164)$$

eliminando todos los  $n^{-s}$ , donde  $n$  es un múltiplo de 2. Entonces

$$\zeta(s)(1 - 2^{-s})(1 - 3^{-s}) = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{9^s} + \cdots - \left( \frac{1}{3^s} + \frac{1}{9^s} + \frac{1}{15^s} + \cdots \right) , \quad (4.165)$$

$s$	$\zeta(s)$
2	1.64493 40668
3	1.20205 69032
4	1.08232 32337
5	1.03692 77551
6	1.01734 30620
7	1.00834 92774
8	1.00407 73562
9	1.00200 83928
10	1.00099 45751

Tabla 4.3: Función zeta de Riemann.

Figura 4.10: Función zeta de Riemann,  $\zeta(s) - 1$  versus  $s$ .

eliminando todos los términos remanentes en el cual  $n$  es un múltiplo de 3. Continuando, tenemos  $\zeta(s)(1-2^{-s})(1-3^{-s})(1-5^{-s})\dots(1-P^{-s})$ , donde  $P$  es un número primo, y todos los términos  $n^{-s}$ , en el cual  $n$  es un múltiplo entero por sobre  $P$ , son cancelados. Para  $P \rightarrow \infty$ ,

$$\zeta(s)(1-2^{-s})(1-3^{-s})\dots(1-P^{-s}) = \zeta(s) \prod_{P(\text{primo})=2}^{\infty} (1-P^{-s}) = 1. \quad (4.166)$$

Por lo tanto

$$\zeta(s) = \left[ \prod_{P(\text{primo})=2}^{\infty} (1-P^{-s}) \right]^{-1} \quad (4.167)$$

dando  $\zeta(s)$  como un producto infinito.<sup>6</sup>

Este procedimiento de cancelación tiene una clara aplicación en el cálculo numérico. La ecuación (4.164) dará  $\zeta(s)(1 - 2^{-s})$  con la misma precisión como la ecuación (4.163) da  $\zeta(s)$ , pero solamente con la mitad de términos. (En cuyo caso, podría hacerse una corrección para desprestigiar la cola de la serie por la técnica de Maclaurin reemplazando la serie por una integral).

Conjuntamente con la función zeta de Riemann, habitualmente se definen otras tres funciones de sumas de potencia recíprocas:

$$\eta(s) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s} = (1 - 2^{1-s})\zeta(s) ,$$

$$\lambda(s) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)^s} = \left(1 - \frac{1}{2^s}\right)\zeta(s) ,$$

y

$$\beta(s) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)^s} .$$

A partir de los números de Bernoulli o de las series de Fourier podemos determinar algunos valores especiales

$$\zeta(2) = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}$$

$$\zeta(4) = 1 + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = \frac{\pi^4}{90}$$

$$\eta(2) = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \cdots = \frac{\pi^2}{12}$$

$$\eta(4) = 1 - \frac{1}{2^4} + \frac{1}{3^4} - \cdots = \frac{7\pi^4}{720}$$

$$\lambda(2) = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}$$

$$\lambda(4) = 1 + \frac{1}{3^4} + \frac{1}{5^4} + \cdots = \frac{\pi^4}{96}$$

$$\beta(1) = 1 - \frac{1}{3} + \frac{1}{5} - \cdots = \frac{\pi}{4}$$

$$\beta(3) = 1 - \frac{1}{3^3} + \frac{1}{5^3} - \cdots = \frac{\pi^3}{32}$$

La constante de Catalan

$$\beta(2) = 1 - \frac{1}{3^2} + \frac{1}{5^2} - \cdots = 0.9159\ 6559\dots ,$$

---

<sup>6</sup>Este es el punto de partida para la vasta aplicación de la función zeta de Riemann a la teoría de números.

### 4.10.4 Mejoramiento de la convergencia.

Si requerimos sumar una serie convergente  $\sum_{n=1}^{\infty} a_n$  cuyos términos son funciones racionales de  $n$ , la convergencia puede ser mejorada dramáticamente introduciendo la función zeta de Riemann.

**Ejemplo** Mejorando la convergencia.

El problema es evaluar la serie  $\sum_{n=1}^{\infty} \frac{1}{(1+n^2)}$ . Expandiendo  $\frac{1}{(1+n^2)} = \frac{1}{n^2} \frac{1}{\left(1 + \frac{1}{n^2}\right)}$  por

división directa, tenemos

$$\begin{aligned} \frac{1}{1+n^2} &= \frac{1}{n^2} \left( 1 - \frac{1}{n^2} + \frac{1}{n^4} - \frac{n^{-6}}{1+n^{-2}} \right) \\ &= \frac{1}{n^2} - \frac{1}{n^4} + \frac{1}{n^6} - \frac{1}{n^8 + n^6} . \end{aligned}$$

Por lo tanto

$$\sum_{n=1}^{\infty} \frac{1}{1+n^2} = \zeta(2) - \zeta(4) + \zeta(6) - \sum_{n=1}^{\infty} \frac{1}{n^8 + n^6} .$$

Las funciones  $\zeta$  son conocidas y el remanente de la series converge como  $n^{-8}$ . Claramente, el proceso pueden ser continuado hasta cuando uno desee. Usted puede hacer una elección entre cuanta álgebra hará y cuanta aritmética hará el computador.

Otros métodos para mejorar la efectividad computacional están dadas al final de la sección 4.2 y 4.4.

## 4.11 Series asintóticas o semiconvergentes.

Las series asintóticas frecuentemente ocurren en Física. En cálculo numérico ellas son empleadas para la precisión del cálculo de una variedad de funciones. Consideremos aquí dos tipos de integrales que conducen a series asintóticas: primero, una integral de la forma

$$I_1(x) = \int_x^{\infty} e^{-u} f(u) du ,$$

donde la variable  $x$  aparece como el límite inferior de una integral. Segundo, consideremos la forma

$$I_2(x) = \int_0^{\infty} e^{-u} f\left(\frac{u}{x}\right) du ,$$

con la función  $f$  expandible en serie de Taylor. Las series asintóticas a menudo ocurren como solución de ecuaciones diferenciales. Un ejemplo de este tipo de series aparece como una de las soluciones de la ecuación de Bessel.

### 4.11.1 Función gama incompleta.

La naturaleza de una serie asintótica es quizás mejor ilustrada por un ejemplo específico. Supongamos que tenemos una función integral exponencial<sup>7</sup>

$$Ei(x) = \int_{-\infty}^x \frac{e^u}{u} du, \quad (4.168)$$

o

$$-Ei(-x) = \int_x^{\infty} \frac{e^{-u}}{u} du = E_1(x), \quad (4.169)$$

para ser evaluada para grandes valores de  $x$ . Mejor todavía, tomemos una generalización de la función factorial incompleta (función gama incompleta),

$$I(x, p) = \int_x^{\infty} e^{-u} u^{-p} du = \Gamma(1 - p, x), \quad (4.170)$$

en la cual  $x$  y  $p$  son positivas. De nuevo, buscamos evaluarla para valores grandes de  $x$ .

Integrando por partes, obtenemos

$$I(x, p) = \frac{e^{-x}}{x^p} - p \int_x^{\infty} e^{-u} u^{-p-1} du = \frac{e^{-x}}{x^p} - \frac{pe^{-x}}{x^{p+1}} + p(p+1) \int_x^{\infty} e^{-u} u^{-p-2} du \quad (4.171)$$

Continuando para integrar por partes, desarrollamos la serie

$$I(x, p) = e^{-x} \left( \frac{1}{x^p} - \frac{p}{x^{p+1}} + \frac{p(p+1)}{x^{p+2}} - \dots (-1)^{n-1} \frac{(p+n-2)!}{(p-1)!x^{p+n-1}} \right) + (-1)^n \frac{(p+n-1)!}{(p-1)!} \int_x^{\infty} e^{-u} u^{-p-n} du. \quad (4.172)$$

Esta es una serie notable. Chequeando la convergencia por la prueba de D' Alembert, encontramos

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{|u_{n+1}|}{|u_n|} &= \lim_{n \rightarrow \infty} \frac{(p+n)!}{(p+n-1)!} \frac{1}{x} \\ &= \lim_{n \rightarrow \infty} \frac{(p+n)}{x} \\ &= \infty \end{aligned} \quad (4.173)$$

para todos los valores finitos de  $x$ . Por lo tanto nuestras series son series infinitas que divergen en todas partes!. Antes de descartar la ecuación (4.172) como inútil, veamos cuan bien una suma parcial dada se aproxima a la función factorial incompleta,  $I(x, p)$ .

$$= (-1)^{n+1} \frac{(p+n)!}{(p-1)!} \int_x^{\infty} e^{-u} u^{-p-n-1} du = R_n(x, p). \quad (4.174)$$

<sup>7</sup>Esta función ocurre con frecuencia en problemas astrofísicos que involucran gases con una distribución de energía de Maxwell-Boltzmann.

En valor absoluto

$$|I(x, p) - s_n(x, p)| \leq \frac{(p+n)!}{(p-1)!} \int_x^\infty e^{-u} u^{-p-n-1} du .$$

Luego sustituimos  $u = v + x$  la integral se convierte en

$$\begin{aligned} \int_x^\infty e^{-u} u^{-p-n-1} du &= e^{-x} \int_0^\infty e^{-v} (v+x)^{-p-n-1} dv \\ &= \frac{e^{-x}}{x^{p+n+1}} \int_0^\infty e^{-v} \left(1 + \frac{v}{x}\right)^{-p-n-1} dv . \end{aligned}$$

Para  $x$  grande la integral final se aproxima a 1 y

$$|I(x, p) - s_n(x, p)| \approx \frac{(p+n)!}{(p-1)!} \frac{e^{-x}}{x^{p+n+1}} . \quad (4.175)$$

Esto significa que si tomamos un  $x$  suficientemente grande, nuestra suma parcial  $s_n$  es arbitrariamente una buena aproximación a la función deseada  $I(x, p)$ . Nuestra serie divergente por lo tanto es perfectamente buena para cálculos de sumas parciales. Por esta razón algunas veces es llamada serie semiconvergente. Notemos que la potencia de  $x$  en el denominador del remanente  $(p+n+1)$  es más alto que la potencia de  $x$  en último término incluido en  $s_n(x, p)$ ,  $(p+n)$ .

Ya que el remanente  $R_n(x, p)$  alterna en signo, las sucesivas sumas parciales dan alternadamente cotas superiores e inferiores para  $I(x, p)$ . El comportamiento de la serie (con  $p=1$ ) como una función del número de términos incluidos es mostrado en la figura 4.11. Tenemos

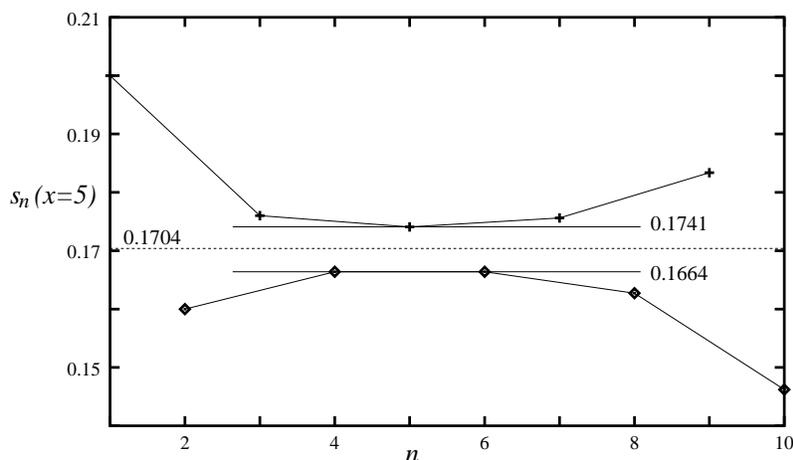


Figura 4.11: Sumas parciales de  $e^x E_1(x) \Big|_{x=5}$ .

$$\begin{aligned} e^x E_1(x) &= e^x \int_x^\infty \frac{e^{-u}}{u} du \\ &= s_n(x) \approx \frac{1}{x} - \frac{1!}{x^2} + \frac{2!}{x^3} - \frac{3!}{x^4} + \cdots + (-1)^n \frac{n!}{x^{n+1}} , \end{aligned} \quad (4.176)$$

la cual es evaluada en  $x = 5$ . Para un valor dado de  $x$  las sucesivas cotas superiores e inferiores dadas por las sumas parciales primero converge y luego diverge. La determinación óptima de  $e^x E_1(x)$  está dado por la aproximación más cercana de las cotas superiores e inferiores, esto es, entre  $s_4 = s_6 = 0.1664$  y  $s_5 = 0.1741$  para  $x = 5$ . Por lo tanto

$$0.1664 \leq e^x E_1(x) \Big|_{x=5} \leq 0.1741 . \quad (4.177)$$

Realmente, a partir de las tablas,

$$e^x E_1(x) \Big|_{x=5} = 0.1704 , \quad (4.178)$$

dentro de los límites establecidos por nuestra expansión asintótica. Note cuidadosamente que la inclusión de términos adicionales en la serie de expansión más allá del punto óptimo literalmente reduce la precisión de la representación.

Cuando aumentamos  $x$ , la diferencia entre la cota superior más baja y la cota inferior más alta disminuirá. Tomando  $x$  suficientemente grande, uno podría calcular  $e^x E_1(x)$  para cualquier grado de precisión deseado.

### 4.11.2 Integrales coseno y seno.

Las series asintóticas también pueden ser desarrolladas a partir de integrales definidas si el integrando tiene el comportamiento requerido. Como un ejemplo, las integrales seno y coseno están definidas por

$$Ci(x) = - \int_x^\infty \frac{\cos t}{t} dt , \quad (4.179)$$

$$si(x) = - \int_x^\infty \frac{\sen t}{t} dt , \quad (4.180)$$

Combinando estas con funciones trigonométricas regulares, podemos definir

$$\begin{aligned} f(x) &= Ci(x) \sen(x) - si(x) \cos(x) = \int_0^\infty \frac{\sen(x)}{y+x} \\ g(x) &= -Ci(x) \cos(x) - si(x) \sen(x) = \int_0^\infty \frac{\cos(x)}{y+x} \end{aligned} \quad (4.181)$$

con la nueva variable  $y = t - x$ . Llevando a variable compleja, tenemos

$$\begin{aligned} g(x) + if(x) &= \int_0^\infty \frac{e^{iy}}{y+x} dy \\ &= \int_0^\infty \frac{ie^{xu}}{1+iu} du \end{aligned} \quad (4.182)$$

en el cual  $u = -iy/x$ . Los límites de integración, 0 a  $\infty$ , a más que de 0 a  $-i\infty$ , puede ser justificado por el teorema de Cauchy. Racionalizando el denominador e igualando las parte reales y las parte imaginarias, obtenemos

$$\begin{aligned} g(x) &= \int_0^\infty \frac{ue^{-xu}}{1+u^2} du, \\ f(x) &= \int_0^\infty \frac{e^{-xu}}{1+u^2} du. \end{aligned} \quad (4.183)$$

La convergencia de las integrales requiere que  $\text{Re}(x) > 0$ .<sup>8</sup>

Ahora, desarrollamos la expansión asintótica, sea  $v = xu$  y expandimos el factor  $[1 + (v/x)^2]^{-1}$  por el teorema del binomio. Tenemos

$$\begin{aligned} f(x) &\approx \frac{1}{x} \int_0^\infty e^{-v} \sum_{0 \leq n \leq N} (-1)^n \frac{v^{2n}}{x^{2n}} dv = \frac{1}{x} \sum_{0 \leq n \leq N} (-1)^n \frac{(2n)!}{x^{2n}} \\ g(x) &\approx \frac{1}{x^2} \int_0^\infty e^{-v} \sum_{0 \leq n \leq N} (-1)^n \frac{v^{2n+1}}{x^{2n}} dv = \frac{1}{x^2} \sum_{0 \leq n \leq N} (-1)^n \frac{(2n+1)!}{x^{2n}}. \end{aligned} \quad (4.184)$$

De las ecuaciones (4.181) y (4.184)

$$\begin{aligned} Ci(x) &\approx \frac{\text{sen}(x)}{x} \sum_{0 \leq n \leq N} (-1)^n \frac{(2n)!}{x^{2n}} - \frac{\cos(x)}{x^2} \sum_{0 \leq n \leq N} (-1)^n \frac{(2n+1)!}{x^{2n}} \\ si(x) &\approx -\frac{\cos(x)}{x} \sum_{0 \leq n \leq N} (-1)^n \frac{(2n)!}{x^{2n}} - \frac{\text{sen}(x)}{x^2} \sum_{0 \leq n \leq N} (-1)^n \frac{(2n+1)!}{x^{2n}}, \end{aligned} \quad (4.185)$$

las expansiones asintóticas deseadas.

La técnica de expandir el integrando de una integral definida e integrar término a término lo volveremos a aplicar para desarrollar una expansión asintótica de la función de Bessel modificada  $K_v$  y también para las expansiones de las dos funciones hipergeométricas confluentes  $M(a, c; x)$  y  $U(a, c; x)$ .

### 4.11.3 Definición de series asintóticas.

El comportamiento de estas series (ecuaciones (4.172) y (4.185)) en consistencia con las propiedades definidas para una serie asintótica<sup>9</sup>. Siguiendo a Poincaré, tomamos

$$x^n R_n(x) = x^n [f(x) - s_n(x)], \quad (4.186)$$

donde

$$s_n(x) = a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \cdots + \frac{a_n}{x^n}. \quad (4.187)$$

<sup>8</sup>La parte real.

<sup>9</sup>No es necesario que las series asintóticas sean series de potencia.

La expansión asintótica de  $f(x)$  tiene las propiedades que

$$\lim_{x \rightarrow \infty} x^n R_n(x) = 0, \quad \text{para } n \text{ fijo}, \quad (4.188)$$

y

$$\lim_{n \rightarrow \infty} x^n R_n(x) = \infty, \quad \text{para } x \text{ fijo}, \quad (4.189)$$

Vemos la ecuaciones (4.172) y (4.173) como un ejemplo de estas propiedades. Para series de potencias, como las supuestas en la forma de  $s_n(x), R_n(x) \sim x^{-n-1}$ . Con condiciones ((4.188)) y ((4.189)) satisfechas, escribimos

$$f(x) \approx \sum_{n=0}^{\infty} a_n \frac{1}{x^n}. \quad (4.190)$$

Notemos el uso de  $\approx$  en lugar de  $=$ . La función  $f(x)$  es igual a la serie solamente en el límite cuando  $x \rightarrow \infty$ .

Las expansiones asintóticas de dos funciones pueden ser multiplicadas entre si y el resultado será una expansión asintótica de un producto de dos funciones.

La expansión asintótica de una función dada  $f(t)$  puede ser integrada término a término (justo como en una serie uniformemente convergente de una función continua) a partir de  $x \leq t < \infty$  y el resultado será una expansión asintótica de  $\int_x^{\infty} f(t) dt$ . Una diferenciación término a término, sin embargo, es válida solamente bajo condiciones muy especiales.

Algunas funciones no poseen una expansión asintótica;  $e^x$  es un ejemplo de tales funciones. Sin embargo, si una función tiene una expansión asintótica, tiene solamente una. La correspondencia no es uno a uno; muchas funciones pueden tener la misma expansión asintótica.

Uno de los métodos más poderoso y útil de generar expansiones asintóticas, es el método de *steepest descents*, será desarrollado más adelante. Las aplicaciones incluyen la derivación de la fórmula de Stirling para la función factorial (completa) y las formas asintóticas de las varias funciones de Bessel. Las series asintóticas ocurren a menudo en física matemática. Una de las aproximaciones más primeras y aún importante de mecánica cuántica, la expansión WKB, es una serie asintótica.

#### 4.11.4 Aplicaciones a cálculo numérico.

Las series asintóticas son usadas frecuentemente en el cálculo de funciones por los computadores. Este es el caso de las funciones de Neumann  $N_0(x)$  y  $N_1(x)$ , y las funciones modificadas de Bessel  $I_n(x)$  y  $K_n(x)$ . Las series asintóticas para integrales del tipo exponencial, ecuación (4.176), para las integrales de Fresnel, y para la función de error de Gauss, son usadas para la evaluación de estas integrales para valores grandes del argumento. Cuan grande debería ser el argumento depende de la precisión requerida.

## 4.12 Productos infinitos.

Consideremos una sucesión de factores positivos  $f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdots f_n (f_i > 0)$ . Usando  $\pi$  mayúscula para indicar el producto, tenemos

$$f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdots f_n = \prod_{i=1}^n f_i . \quad (4.191)$$

Definimos  $p_n$ , como el producto parcial, en analogía con  $s_n$  la suma parcial,

$$p_n = \prod_{i=1}^n f_i , \quad (4.192)$$

y entonces investigamos el límite

$$\lim_{n \rightarrow \infty} p_n = P . \quad (4.193)$$

Si  $P$  es finito (pero no cero), decimos que el producto infinito es convergente. Si  $P$  es infinito o cero, el producto infinito es etiquetado como divergente.

Ya que el producto divergerá a infinito si

$$\lim_{n \rightarrow \infty} f_n > 1 \quad (4.194)$$

o a cero para

$$0 < \lim_{n \rightarrow \infty} f_n < 1 , \quad (4.195)$$

es conveniente escribir nuestro producto como

$$\prod_{n=1}^{\infty} (1 + a_n) .$$

La condición  $a_n \rightarrow 0$  es entonces una condición necesaria (pero no suficiente) para la convergencia.

El producto infinito puede ser relacionado a una serie infinita por el método obvio de tomar el logaritmo

$$\ln \prod_{n=1}^{\infty} (1 + a_n) = \sum_{n=1}^{\infty} \ln(1 + a_n) . \quad (4.196)$$

Una relación más útil es probada por el siguiente teorema.

### 4.12.1 Convergencia de un producto infinito.

Si  $0 \leq a_n < 1$ , el producto infinito  $\prod_{n=1}^{\infty} (1 + a_n)$  y  $\prod_{n=1}^{\infty} (1 - a_n)$  converge si  $\sum_{n=1}^{\infty} a_n$  converge y diverge si  $\sum_{n=1}^{\infty} a_n$  diverge.

Considerando el término  $1 + a_n$ , vemos que de la ecuación (4.80)

$$1 + a_n \leq e^{a_n} . \quad (4.197)$$

Por lo tanto el producto parcial  $p_n$

$$p_n \leq e^{s_n} , \quad (4.198)$$

y haciendo  $n \rightarrow \infty$ ,

$$\prod_{n=1}^{\infty} (1 + a_n) \leq \exp \sum_{n=1}^{\infty} a_n . \quad (4.199)$$

estableciendo una cota superior para el producto infinito.

Para desarrollar una cota más baja, notemos que

$$p_n = 1 + \sum_{i=1}^n a_i + \sum_{i=1}^n \sum_{j=1}^n a_i a_j + \cdots > s_n , \quad (4.200)$$

ya que  $a_i \geq 0$ . De modo que

$$\prod_{n=1}^{\infty} (1 + a_n) \geq \sum_{n=1}^{\infty} a_n \quad (4.201)$$

Si la suma infinita permanece finita, el producto infinito también lo hará. Si la suma infinita diverge, también lo hará el producto infinito.

El caso de  $\prod(1 - a_n)$  es complicado por el signo negativo, pero una prueba de que depende de la prueba anterior puede ser desarrollada notando que para  $a_n < 1/2$  (recuerde que  $a_n \rightarrow 0$  para convergencia)

$$(1 - a_n) \leq \frac{1}{1 + a_n}$$

y

$$(1 - a_n) \geq \frac{1}{1 + 2a_n} . \quad (4.202)$$

### 4.12.2 Funciones seno, coseno y gama.

El lector reconocerá que un polinomio de orden  $n$   $P_n(x)$  con  $n$  raíces reales puede ser escrito como un producto de  $n$  factores:

$$P_n(x) = (x - x_1)(x - x_2) \cdots (x - x_n) = \prod_{i=1}^n (x - x_i) . \quad (4.203)$$

De la misma manera podemos esperar que una función con un número infinito de raíces pueda ser escrito como un producto infinito, un factor para cada raíz. Esto es por cierto el

caso de las funciones trigonométricas. Tenemos dos representaciones muy útiles en productos infinitos,

$$\operatorname{sen}(x) = x \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2\pi^2}\right), \quad (4.204)$$

$$\operatorname{cos}(x) = \prod_{n=1}^{\infty} \left[1 - \frac{4x^2}{(2n-1)^2\pi^2}\right]. \quad (4.205)$$

La más conveniente y quizás la más elegante derivación de estas dos expresiones es usando variable compleja. Por nuestro teorema de convergencia, las ecuaciones (4.204) y (4.205) son convergentes para todos los valores finitos de  $x$ . Específicamente, para el producto infinito para el  $\operatorname{sen}(x)$ ,  $a_n = x^2/n^2\pi^2$ ,

$$\begin{aligned} \sum_{n=1}^{\infty} a_n &= \frac{x^2}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{x^2}{\pi^2} \zeta(2) \\ &= \frac{x^2}{6}. \end{aligned} \quad (4.206)$$

La serie correspondiente a la ecuación (4.205) se comporta en una manera similar.

La ecuación (4.204) conduce a dos resultados interesantes. Primero, si fijamos  $x = \pi/2$ , obtenemos

$$1 = \frac{\pi}{2} \prod_{n=1}^{\infty} \left[1 - \frac{1}{(2n)^2}\right] = \frac{\pi}{2} \prod_{n=1}^{\infty} \left[\frac{(2n)^2 - 1}{(2n)^2}\right]. \quad (4.207)$$

Resolviendo para  $\pi/2$ , obtenemos

$$\frac{\pi}{2} = \prod_{n=1}^{\infty} \left[\frac{(2n)^2}{(2n-1)(2n+1)}\right] = \frac{2 \cdot 2}{1 \cdot 3} \cdot \frac{4 \cdot 4}{3 \cdot 5} \cdot \frac{6 \cdot 6}{5 \cdot 7} \cdots, \quad (4.208)$$

la cual es la famosa fórmula de Wallis para  $\pi/2$ .

El segundo resultado involucra la función factorial o función gama. Una definición de la función gama es

$$\Gamma(x) = \left[ x e^{\gamma x} \prod_{r=1}^{\infty} \left(1 + \frac{x}{r}\right) e^{-\frac{x}{r}} \right]^{-1}, \quad (4.209)$$

donde  $\gamma$  es la constante de Euler-Mascheroni, sección 4.2. Si tomamos el producto de  $\Gamma(x)$  y  $\Gamma(-x)$ , la ecuación (4.209) tiende a

$$\begin{aligned} \Gamma(x)\Gamma(-x) &= - \left[ x e^{\gamma x} \prod_{r=1}^{\infty} \left(1 + \frac{x}{r}\right) e^{-\frac{x}{r}} x e^{-\gamma x} \prod_{r=1}^{\infty} \left(1 - \frac{x}{r}\right) e^{\frac{x}{r}} \right]^{-1} \\ &= - \left[ x^2 \prod_{r=1}^{\infty} \left(1 - \frac{x^2}{r^2}\right) \right]^{-1}. \end{aligned} \quad (4.210)$$

Usando la ecuación (4.204) con  $x$  reemplazado por  $\pi x$ , obtenemos

$$\Gamma(x)\Gamma(-x) = -\frac{\pi}{x \operatorname{sen}(\pi x)}. \quad (4.211)$$

Anticipando una relación de recurrencia desarrollada posteriormente, tenemos que usando  $-x\Gamma(-x) = \Gamma(1-x)$ . La ecuación (4.211) puede ser escrita como

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\operatorname{sen}(\pi x)}. \quad (4.212)$$

Esto será útil cuando tratamos la función gama.

Estrictamente hablando, podríamos chequear el intervalo en  $x$  para el cual la ecuación (4.209) es convergente. Claramente, para  $x = 0, -1, -2, \dots$  los factores individuales se anulan. La prueba que el producto infinito converge para todos los otros valores (finitos) de  $x$  es dejado como ejercicio.

Estos productos infinitos tienen una variedad de usos en matemática analítica. Sin embargo, a causa de su lentitud de convergencia, ellas no son aptas para un trabajo numérico preciso.

# Capítulo 5

## Funciones de una variable compleja I. Propiedades analíticas y Mapeo.

versión final 1.2-260602<sup>1</sup>

Veamos ahora el estudio de funciones de una variable compleja. En esta área desarrollamos alguna de las herramientas más poderosas y útiles de todo el análisis matemático.

1. Para muchos pares de funciones  $u$  y  $v$ , ambas satisfacen la ecuación de Laplace

$$\nabla^2\psi = \frac{\partial^2\psi(x,y)}{\partial x^2} + \frac{\partial^2\psi(x,y)}{\partial y^2} = 0 .$$

De modo que  $u$  o  $v$  pueden ser usados para describir un potencial electrostático bidimensional. La otra función que da una familia de curvas ortogonales a aquella de la primera función, puede ser usada para describir el campo eléctrico  $\vec{E}$ . Una situación similar se mantiene para la hidrodinámica de un fluido ideal en movimiento irrotacional. La función  $u$  podría describir el potencial de velocidades, mientras que la función  $v$  podría entonces ser la función de flujo.

En muchos casos en que las funciones  $u$  y  $v$  son desconocidas, un mapeo conforme o transformación en el plano complejo nos permite crear un sistema de coordenadas hecho a la medida para el problema en particular.

2. Veremos que ecuaciones diferenciales de segundo orden de interés en Física pueden ser resueltas por series de potencia. Las mismas series de potencia pueden ser usadas en el plano complejo reemplazando  $x$  por la variable compleja  $z$ . La dependencia de la solución  $f(z)$  de un  $z_0$  dado sobre el comportamiento de  $f(z)$  en todas partes nos da un mayor discernimiento del comportamiento de nuestra solución y una herramienta poderosa (continuación analítica) para extender la región en la cual la solución es válida.
3. El cambio de un parámetro  $k$  de real a imaginario,  $k \rightarrow ik$ , transforma la ecuación de Helmholtz en la ecuación de difusión. El mismo cambio transforma las soluciones de la ecuación de Helmholtz (funciones de Bessel y funciones esféricas de Bessel) en las soluciones de la ecuación de difusión (funciones de Bessel modificada y funciones esféricas modificadas de Bessel).
4. Las integrales en el plano complejo tienen una amplia variedad de aplicaciones útiles.

---

<sup>1</sup>Este capítulo está basado en el sexto capítulo del libro: *Mathematical Methods for Physicists, fourth edition* de George B. Arfken & Hans J. Weber, editorial ACADEMIC PRESS.

- (a) Evaluación de integrales definidas.
- (b) Inversión de series de potencia.
- (c) Formación de productos infinitos.
- (d) Obtener soluciones de ecuaciones diferenciales para valores grandes de la variable (soluciones asintóticas).
- (e) Investigación de la estabilidad de sistemas potencialmente oscilatorios.
- (f) Inversión de transformadas integrales.

5. Muchas cantidades físicas que originalmente fueron reales se convierten en complejas cuando una teoría física simple se generaliza. Los índices reales de difracción de la luz se convierte en una cantidad compleja cuando la absorción es incluida. La energía real asociada con un nivel de energía se convierte en compleja cuando la vida media finita del nivel es considerado.

## 5.1 Algebra compleja.

Un número complejo es nada más que un par ordenado de dos números reales,  $(a, b)$  o  $a + ib$ , en el cual  $i$  es  $\sqrt{-1}$ . Similarmente, una variable compleja es un par ordenado de dos variables reales,

$$z = (x, y) = x + iy . \quad (5.1)$$

Veremos que el orden es importante, que en general  $a + bi$  no es igual a  $b + ai$  y  $x + iy$  no es igual a  $y + xi$ .<sup>2</sup> Frecuentemente es conveniente emplear una representación gráfica de la variable compleja. Graficando  $x$  la parte real de  $z$  como la abscisa e  $y$  la parte imaginaria de  $z$  como la ordenada, tenemos el plano complejo o plano Argand mostrado en la figura 5.1. Si asignamos valores específicos a  $x$  e  $y$ , entonces  $z$  corresponde a un punto  $(x, y)$  en el plano. En términos del ordenamiento mencionado antes, es obvio que el punto  $(x, y)$  no coincide con el punto  $(y, x)$  excepto para el caso especial de  $x = y$ .

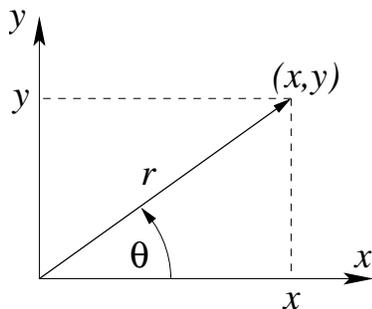


Figura 5.1: Plano complejo, diagrama de Argand.

---

<sup>2</sup>El álgebra de los números complejos,  $a + ib$  es isomórfica con la de las matrices de la forma  $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$

Todo nuestro análisis de variable compleja puede ser desarrollado en términos de pares ordenados de números  $(a, b)$ , variables  $(x, y)$ , y funciones  $(u(x, y), v(x, y))$ . La  $i$  no es necesaria pero es conveniente. Sirve para mantener el par en orden algo como un vector unitario. De modo que la suma y la multiplicación de números complejos puede estar definida en términos de sus componentes cartesianas como

$$\begin{aligned} z_1 + z_2 &= (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \\ &= x_1 + x_2 + i(y_1 + y_2) , \end{aligned} \quad (5.2)$$

$$z_1 z_2 = (x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) . \quad (5.3)$$

De la figura 5.1 podemos escribir

$$\begin{aligned} x &= r \cos(\theta) \\ y &= r \operatorname{sen}(\theta) \end{aligned} \quad (5.4)$$

y

$$z = r(\cos(\theta) + i \operatorname{sen}(\theta)) . \quad (5.5)$$

Usando un resultado que fue sugerido (pero no rigurosamente probado) en la sección 4.6, tenemos la muy útil representación polar

$$z = r e^{i\theta} . \quad (5.6)$$

En esta representación  $r$  es llamado el módulo o magnitud de  $z$  ( $r = |z| = (x^2 + y^2)^{1/2}$ ) y el ángulo  $\theta$  ( $= \tan^{-1}(y/x)$ ) se conoce como argumento  $\arg(z)$  o fase de  $z$ .

La elección de la representación polar, ecuación (5.5), o representación cartesiana, la ecuación (5.1), es un asunto de conveniencia. La suma y la resta de variables complejas son más fáciles en representación cartesiana, ecuación (5.2). La multiplicación, división, potencias, y raíces son más fáciles en la forma polar, ecuación (5.6). Analíticamente o gráficamente, usando la analogía vectorial, podemos mostrar que el módulo de la suma de dos números complejos no es mayor que la suma del módulo y no es menor que la diferencia, ejercicio,

$$|z_1| - |z_2| \leq |z_1 + z_2| \leq |z_1| + |z_2| . \quad (5.7)$$

A causa de la analogía vectorial, estas son llamadas las desigualdades triangulares.

Usando la forma polar, ecuación (5.5), encontramos que la magnitud de un producto es el producto de las magnitudes,

$$|z_1| \cdot |z_2| = |z_1 \cdot z_2| . \quad (5.8)$$

También,

$$\arg(z_1 \cdot z_2) = \arg(z_1) + \arg(z_2) . \quad (5.9)$$

A partir de nuestra variable compleja  $z$  las funciones complejas  $f(z)$  o  $w(z)$  pueden ser construidas. Estas funciones complejas pueden ser resueltas en su parte real y su parte imaginaria

$$w(z) = u(x, y) + iv(x, y) , \quad (5.10)$$

en la cual las funciones separadas  $u(x, y)$  y  $v(x, y)$  son reales puras. Por ejemplo, si  $f(z) = z^2$ , tenemos

$$\begin{aligned} f(z) &= (x + iy)^2 \\ &= (x^2 - y^2) + i2xy . \end{aligned}$$

La parte real de la función  $f(z)$  será etiquetada  $\Re[f(z)]$ , mientras la parte imaginaria será etiquetada  $\Im[f(z)]$ . En la ecuación (5.10)

$$\Re[w(z)] = u(x, y) , \quad \Im[w(z)] = v(x, y) .$$

La relación entre la variable independiente  $z$  y la variable dependiente  $w$  es quizás mejor representada como una operación de mapeo. Dado un  $z = x + iy$ , un punto en el plano  $z$ . El valor complejo  $w(z)$  es entonces un punto en el plano  $w$ . Puntos en el plano  $z$  se mapean en puntos en el plano  $w$  y curvas en el plano  $z$  se mapean en curvas en el plano  $w$  como indica la figura 5.2.

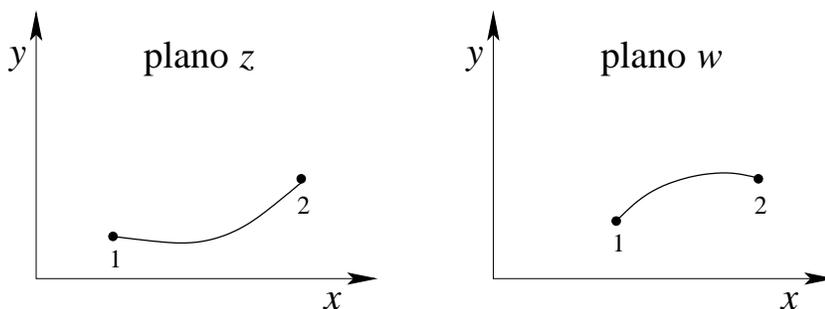


Figura 5.2: La función  $w(z) = u(x, y) + iv(x, y)$  mapea puntos en el plano  $xy$  en puntos en el plano  $w$ .

### 5.1.1 Conjugación compleja.

En todos estos pasos, números complejos, variables, y funciones, la operación de reemplazar  $i$  por  $-i$  es llamada “tomar el complejo conjugado”. El complejo conjugado de  $z$  se denota por  $z^*$ , donde<sup>3</sup>

$$z^* = x - iy . \quad (5.11)$$

<sup>3</sup>El complejo conjugado es a menudo denotado por  $\bar{z}$

La variable compleja  $z$  y su complejo conjugado  $z^*$  es una imagen especular la una de la otra reflejadas por el eje  $x$ , esto es, inversión del eje  $y$  (compare la figura 5.3). El producto  $zz^*$  es

$$zz^* = (x + iy)(x - iy) = x^2 + y^2 = r^2 . \quad (5.12)$$

De modo que

$$(zz^*)^{1/2} = |z|,$$

la magnitud de  $z$

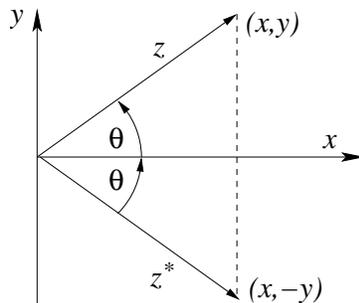


Figura 5.3: Puntos de complejos conjugados.

### 5.1.2 Funciones de una variable compleja.

Todas las funciones elementales de variables reales pueden ser extendidas al plano complejo reemplazando la variable real  $x$  por la variable compleja  $z$ . Esto es un ejemplo de la continuación analítica mencionada en la sección 5.5. La relación extremadamente importante, ecuaciones (5.5) y (5.6), es una ilustración de esto. Moverse en el plano complejo abre nuevas oportunidades para análisis.

**Ejemplo** Fórmula de Moivre.

Si la ecuación (5.6) es elevada a la  $n$ -ésima potencia, tenemos

$$e^{in\theta} = (\cos(\theta) + i \operatorname{sen}(\theta))^n . \quad (5.13)$$

Expandiendo la exponencial ahora con argumento  $n\theta$ , obtenemos

$$\cos(n\theta) + i \operatorname{sen}(n\theta) = (\cos(\theta) + i \operatorname{sen}(\theta))^n . \quad (5.14)$$

Esta es la fórmula de Moivre.

Ahora si el lado derecho de la ecuación (5.14) es expandido usando el teorema del binomio, obtenemos  $\cos(n\theta)$  como una serie de potencias de  $\cos(\theta)$  y  $\operatorname{sen}(\theta)$ . Numerosos de otros ejemplos de relaciones entre funciones exponenciales, hiperbólicas y trigonométricas en el plano complejo podemos encontrarlas como ejercicios. Ocasionalmente hay complicaciones. El logaritmo de una variable compleja puede ser expandida usando la representación polar

$$\begin{aligned} \ln(z) &= \ln(re^{i\theta}) \\ &= \ln(r) + i\theta . \end{aligned} \quad (5.15)$$

Esto no está completo. Al ángulo fase  $\theta$ , podemos añadirle algún múltiplo entero de  $2\pi$  sin cambiar  $z$ . Luego la ecuación (5.15) debería leerse

$$\begin{aligned}\ln z &= \ln (re^{i(\theta+2n\pi)}) \\ &= \ln(r) + i(\theta + 2n\pi) .\end{aligned}\tag{5.16}$$

El parámetro  $n$  puede ser cualquier entero. Esto significa que  $\ln z$  es una función multivaluada teniendo un número infinito de valores para una pareja de valores reales de  $r$  y  $\theta$ . Para evitar esta ambigüedad, usualmente concordamos a tomar  $n = 0$  y limitar la fase a un intervalo de longitud  $2\pi$  tal como  $(-\pi, \pi)$ <sup>4</sup>. La línea en el plano  $z$  que no es cruzada, el eje real negativo en este caso, es conocida como *línea de corte o de ramificación*. El valor de  $\ln z$  con  $n = 0$  es llamado *valor principal* de  $\ln z$ . Más adelante discutiremos estas funciones, incluyendo el logaritmo, aparece en la sección 5.6.

## 5.2 Condiciones de Cauchy-Riemann.

Habiendo establecido las funciones complejas de una variable compleja, ahora procederemos a derivarlas. La derivada de  $f(z)$ , como la de una función real, está definida por

$$\lim_{\delta z \rightarrow 0} \frac{f(z + \delta z) - f(z)}{z + \delta z - z} = \lim_{\delta z \rightarrow 0} \frac{\delta f(z)}{\delta z} = \frac{df}{dz} \quad \text{o} \quad f'(z) ,\tag{5.17}$$

a condición que el límite sea independiente de la forma de aproximación particular al punto  $z$ . Para variables reales requerimos que el límite por la derecha ( $x \rightarrow x_0$  desde arriba) y el límite por la izquierda ( $x \rightarrow x_0$  desde abajo) sea iguales para que la derivada  $df(x)/dx$  exista en  $x = x_0$ . Ahora, con  $z$  (o  $z_0$ ) algún punto en el plano, nuestro requerimiento de que el límite sea independiente de la dirección de aproximación es muy restrictiva.

Consideremos incrementos  $\delta x$  y  $\delta y$  de las variables  $x$  y  $y$ , respectivamente. Entonces

$$\delta z = \delta x + i\delta y .\tag{5.18}$$

También,

$$\delta f = \delta u + i\delta v ,\tag{5.19}$$

tal que

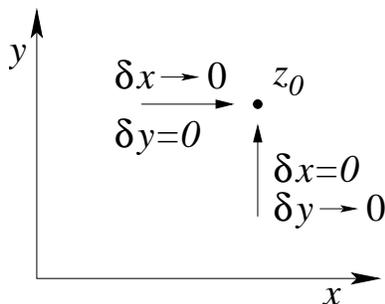
$$\frac{\delta f}{\delta z} = \frac{\delta u + i\delta v}{\delta x + i\delta y} .\tag{5.20}$$

Tomemos el límite indicado en la ecuación (5.17) por dos aproximaciones diferentes como muestra la figura (5.4). Primero, con  $\delta y = 0$ , sea  $\delta x \rightarrow 0$ . La ecuación (5.17) tiende

$$\begin{aligned}\lim_{\delta z \rightarrow 0} \frac{\delta f}{\delta z} &= \lim_{\delta x \rightarrow 0} \left( \frac{\delta u}{\delta x} + i \frac{\delta v}{\delta x} \right) \\ &= \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} ,\end{aligned}\tag{5.21}$$

---

<sup>4</sup>Hay elecciones inusuales de fase. La fase apropiada depende de cada problema.

Figura 5.4: Aproximaciones alternativas a  $z_0$ .

suponiendo que las derivadas parciales existen. Para una segunda aproximación, fijamos  $\delta x = 0$  y entonces hacemos  $\delta y \rightarrow 0$ . Esto tiende a

$$\begin{aligned} \lim_{\delta z \rightarrow 0} \frac{\delta f}{\delta z} &= \lim_{\delta y \rightarrow 0} \left( -i \frac{\delta u}{\delta y} + \frac{\delta v}{\delta y} \right) \\ &= -i \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}. \end{aligned} \quad (5.22)$$

Para tener una derivada  $df/dz$ , las ecuaciones (5.21) y (5.22) deben ser idénticas. Igualando parte real a parte real y parte imaginaria con imaginaria (como los componentes del vector), obtenemos

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}}. \quad (5.23)$$

Estas son las famosas condiciones de Cauchy-Riemann. Ellas fueron descubiertas por Cauchy y usadas ampliamente por Riemann en su teoría de funciones analíticas. Estas condiciones son necesarias para la existencia de una derivada de  $f(z)$ , esto es, si  $df/dz$  existe, la condición de Cauchy-Riemann debe satisfacerse.

Inversamente, si las condiciones de Cauchy-Riemann se satisfacen y la derivada parcial de  $u(x, y)$  y  $v(x, y)$  son continuas, la derivada  $df/dz$  existe. Esto puede ser mostrado escribiendo

$$\delta f = \left( \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) \delta x + \left( \frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y} \right) \delta y. \quad (5.24)$$

La justificación para esta expresión depende de la continuidad de la derivada parcial de  $u$  y  $v$ . Dividiendo por  $\delta z$ , tenemos

$$\begin{aligned} \frac{\delta f}{\delta z} &= \frac{(\partial u/\partial x + i(\partial v/\partial x))\delta x + (\partial u/\partial y + i(\partial v/\partial y))\delta y}{\delta x + i\delta y} \\ &= \frac{(\partial u/\partial x + i(\partial v/\partial x)) + (\partial u/\partial y + i(\partial v/\partial y))\delta y/\delta x}{1 + i(\delta y/\delta x)} \end{aligned} \quad (5.25)$$

Si  $\delta f/\delta z$  tiene un valor único, la dependencia de  $\delta y/\delta x$  debería eliminarse. Aplicando las condiciones de Cauchy-Riemann a las derivadas de  $y$  obtenemos

$$\frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y} = \frac{\partial v}{\partial x} + i \frac{\partial u}{\partial x}. \quad (5.26)$$

Sustituyendo la ecuación (5.26) y (5.25), podemos cancelar la dependencia  $\delta y/\delta x$  y

$$\frac{\delta f}{\delta z} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}, \quad (5.27)$$

la cual muestra que el límite  $\delta f/\delta z$  es independiente de la dirección de aproximación en el plano complejo ya que las derivadas parciales son continuas.

Es digno de atención notar que las condiciones de Cauchy-Riemann garantizan que las curvas  $u = c_1$  serán ortogonales a las curvas  $v = c_2$ . Esto es fundamental en la aplicación de problemas de potencial en una variedad de áreas de la Física. Si  $u = c_1$  es una línea de fuerza eléctrica, entonces  $v = c_2$  es una línea equipotencial (superficie), y vice versa.

### Funciones analíticas.

Finalmente, se  $f(z)$  es diferenciable en  $z = z_0$  y en algunas pequeñas regiones alrededor de  $z_0$  decimos que  $f(z)$  es *analítica*<sup>5</sup> en  $z_0$ . Si  $f(z)$  es analítica en todo el plano complejo (finito), la llamaremos una función *entera*. Nuestra teoría de variables complejas aquí es esencialmente una de funciones analíticas de variables complejas, la cual destaca la importancia crucial de las condiciones de Cauchy-Riemann. El concepto de analiticidad presente en teorías avanzadas de la física moderna, por ejemplo este concepto juega un rol crucial en la teoría de dispersión (de partículas elementales). Si  $f'(z)$  no existe en  $z = z_0$ , entonces  $z_0$  es conocido como punto singular.

Para ilustrar las condiciones de Cauchy-Riemann, consideremos dos ejemplos muy simples.

#### Ejemplo

Sea  $f(z) = z^2$ . Entonces la parte real  $u(x, y) = x^2 - y^2$  y la parte imaginaria  $v(x, y) = 2xy$ . Siguiendo la ecuación (5.23)

$$\frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -2y = -\frac{\partial v}{\partial x}.$$

Vemos que  $f(z) = z^2$  satisface las condiciones de Cauchy-Riemann a través del plano complejo. Ya que las derivadas parciales son claramente continuas, concluimos que  $f(z) = z^2$  es analítica.

#### Ejemplo

Sea  $f(z) = z^*$ . Ahora  $u = x$  y  $v = -y$ . Aplicando las condiciones de Cauchy-Riemann, obtenemos

$$\frac{\partial u}{\partial x} = 1 \neq \frac{\partial v}{\partial y}.$$

Las condiciones de Cauchy-Riemann no son satisfechas y  $f(z) = z^*$  no es una función analítica de  $z$ . Es interesante notar que  $f(z) = z^*$  es continua, aquello provee un ejemplo de una función que es continua en todo lugar pero en ninguno diferenciable.

La derivada de una función real de una variable real es esencialmente una característica local, que provee información a cerca de la función solamente en un vecindario local por

<sup>5</sup>Algunos autores usan el término holomórfica.

ejemplo, como una expansión de Taylor truncada. La existencia de una derivada de una función de una variable compleja tiene muchas más implicancias de conocimiento lejano. Las partes reales e imaginarias de nuestra función analítica debe satisfacer separadamente las ecuaciones de Laplace. Más aún, nuestra función analítica está garantizada para las derivadas de todo orden, sección 4.4. En este sentido las derivadas no solamente gobiernan el comportamiento local de la función compleja, sino que controla el comportamiento distante también.

## 5.3 Teorema integral de Cauchy.

### 5.3.1 Integrales de contorno.

Con la diferenciación bajo control, veamos la integración. La integral de una variable compleja sobre un contorno en el plano complejo puede ser definida en cercana analogía a la integral de (Riemann) de una función real integrada a lo largo del eje  $x$  real.

Dividimos el contorno  $z_0 z_0'$  en  $n$  intervalos tomando  $n - 1$  puntos intermedios  $z_1, z_2, \dots$ , sobre el contorno (figura 5.5). Consideremos la suma

$$S_n = \sum_{j=1}^n f(\zeta_j)(z_j - z_{j-1}) , \quad (5.28)$$

donde  $\zeta_j$  es un punto sobre la curva entre  $z_j$  y  $z_{j-1}$ . Ahora hagamos  $n \rightarrow \infty$  con

$$|z_j - z_{j-1}| \rightarrow 0 ,$$

para todo  $j$ . Si el  $\lim_{n \rightarrow \infty} S_n$  existe y es independiente de los detalles de elección de los puntos  $z_j$  y  $\zeta_j$ , entonces

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n f(\zeta_j)(z_j - z_{j-1}) = \int_{z_0}^{z_0'} f(z) dz . \quad (5.29)$$

El lado derecho de la ecuación es llamado integral de contorno de  $f(z)$  (a lo largo del contorno específico  $C$  desde  $z = z_0$  a  $z = z_0'$ ).

El procedimiento desarrollado para la integral de contorno es cercanamente análogo a la integral de Riemann de una función real de una variable real. Como una alternativa, la integral de contorno puede ser definida por

$$\begin{aligned} \int_{z_0}^{z_0'} f(z) dz &= \int_{x_1, y_1}^{x_2, y_2} [u(x, y) + iv(x, y)][dx + idy] \\ &= \int_{x_1, y_1}^{x_2, y_2} [u(x, y)dx - v(x, y)dy] + i \int_{x_1, y_1}^{x_2, y_2} [v(x, y)dx + u(x, y)dy] , \end{aligned}$$

con el camino que une  $(x_1, y_1)$  y  $(x_2, y_2)$  *especificado*. Esto reduce la integral compleja a la suma compleja de integrales reales. Esto es algo análogo al reemplazo de una integral vectorial por la suma de vectores de integrales escalares.

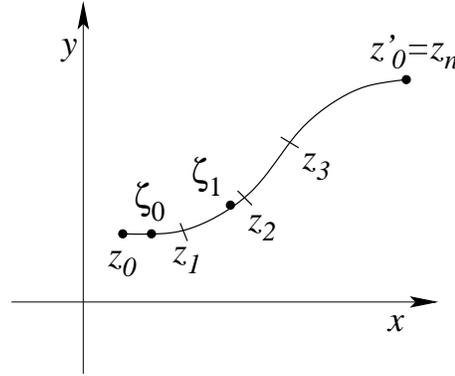


Figura 5.5: Camino de integración.

Un ejemplo importante es la integral de contorno  $\int_C z^n dz$ , donde  $C$  es un círculo de radio  $r > 0$  alrededor del origen  $z = 0$  en el sentido matemático positivo (en el sentido contrario a los punteros del reloj). En coordenadas polares de la ecuación (5.6) parametrizamos  $z = re^{i\theta}$  y  $dz = ire^{i\theta}d\theta$ . Para  $n \neq -1$  obtenemos

$$\begin{aligned} \frac{1}{2\pi i} \int_C z^n dz &= \frac{r^{n+1}}{2\pi} \int_0^{2\pi} \exp[i(n+1)\theta] d\theta \\ &= \frac{r^{n+1}}{2\pi i(n+1)} [e^{i(n+1)\theta}]_0^{2\pi} = 0, \end{aligned} \quad (5.30)$$

ya que  $2\pi$  es un período de  $e^{i(n+1)\theta}$ , mientras que para  $n = -1$

$$\frac{1}{2\pi i} \int_C \frac{dz}{z} = \frac{1}{2\pi} \int_0^{2\pi} d\theta = 1, \quad (5.31)$$

nuevamente independiente de  $r$ . Estas integrales son ejemplos del teorema integral de Cauchy el cual consideraremos en la próxima sección.

### 5.3.2 Prueba del teorema de Stoke.

El teorema de la integral de Cauchy es el primero de dos teoremas básicos en la teoría del comportamiento de funciones de una variable compleja. Primero, una prueba bajo condiciones relativamente restrictivas, condiciones que serían intolerables para matemáticos desarrollando una bella teoría abstracta, pero que son satisfechas usualmente en problemas físicos.

Si una función  $f(z)$  es analítica (por lo tanto simplemente valuada) y sus derivadas parciales son continuas a través de una región simplemente conexa  $R$ ,<sup>6</sup> para cada camino cerrado  $C$  (figura 5.6) en  $R$  la integral de  $f(z)$  alrededor de  $C$  es cero o

$$\int_C f(z) dz = \oint_C f(z) dz = 0. \quad (5.32)$$

<sup>6</sup>Una región o dominio simplemente conexo es aquel en el cual cada contorno cerrado en la región encierra sólo puntos contenidos en ella. Si la región no es simplemente conexa es múltiplemente conexa. Como un ejemplo de una región múltiplemente conexa, consideremos el plano  $z$  con el círculo unitario interior excluido.

El símbolo  $\oint$  es usado para enfatizar que el camino es cerrado. Recordemos que en el capítulo de vectores una función tal como  $f(z)$ , identificada como una fuerza fue etiquetada como conservativa.

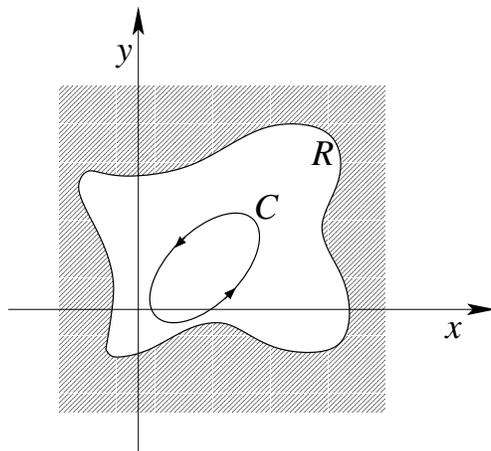


Figura 5.6: Un contorno cerrado  $C$  dentro de una región  $R$  simplemente conexa.

En esta forma el teorema de la integral de Cauchy puede ser dado por aplicación directa del teorema de Stokes. Con  $f(z) = u(x, y) + iv(x, y)$  y  $dz = dx + idy$ ,

$$\begin{aligned} \oint_C f(z)dz &= \oint_C (u + iv)(dx + idy) \\ &= \oint_C (udx - vdy) + i \oint_C (vdx + udy) . \end{aligned} \quad (5.33)$$

Estas dos integrales de línea pueden ser convertidas a integrales de superficie por el teorema de Stokes, un procedimiento que es justificado si las derivadas parciales son continuas dentro de  $C$ . Aplicando el teorema de Stokes podemos notar que las dos últimas integrales de la ecuación (5.33) son completamente reales. Usando

$$\vec{V} = \hat{x}V_x + \hat{y}V_y ,$$

tenemos

$$\oint_C (V_x dx + V_y dy) = \int \left( \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) dxdy . \quad (5.34)$$

Para la primera integral en la última parte de la ecuación (5.33) sea  $u = V_x$  y  $v = -V_y$ .<sup>7</sup>

<sup>7</sup>En la prueba del teorema de Stokes,  $V_x$  y  $V_y$  son cualquier par funciones (con derivadas parciales continuas).

Entonces

$$\begin{aligned}\oint_C (u dx - v dy) &= \oint_C (V_x dx + V_y dy) \\ &= \int \left( \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) dx dy \\ &= - \int \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy .\end{aligned}\tag{5.35}$$

Para la segunda integral del lado derecho de la ecuación (5.33) sea  $u = V_y$  y  $v = V_x$ . Usando nuevamente el teorema de Stokes, obtenemos

$$\oint (v dx + u dy) = \int \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dx dy .\tag{5.36}$$

Aplicando las condiciones de Cauchy-Riemann que se deben satisfacer ya que  $f(z)$  se supone analítica, cada integrando se anula y

$$\begin{aligned}\oint f(z) dz &= - \int \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy + i \int \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dx dy \\ &= 0 .\end{aligned}\tag{5.37}$$

### 5.3.3 Prueba de Cauchy-Goursat.

Esto completa la prueba del teorema integral de Cauchy. Sin embargo, la prueba está estropeada desde el punto de vista teórico por la necesidad de continuidad de las primeras derivadas parciales. Realmente, como se muestra por Goursat, esta condición no es esencial. Un perfil de la prueba de Goursat es como sigue. Subdividimos la región dentro del contorno  $C$  en una red de pequeños cuadrados como está indicado en la figura 5.7. Luego

$$\oint_C f(z) dz = \sum_j \oint_{C_j} f(z) dz ,\tag{5.38}$$

todas las integrales a lo largo de las líneas interiores se cancelan. Para atacar la  $\int_{C_j} f(z) dz$ , construimos la función

$$\delta_j(z, z_j) = \frac{f(z) - f(z_j)}{z - z_j} - \left. \frac{df(z)}{dz} \right|_{z=z_j} ,\tag{5.39}$$

con  $z_j$  un punto interno de la  $j$ -ésima subregión. Note que  $[f(z) - f(z_j)]/(z - z_j)$  es una aproximación a la derivada en  $z = z_j$ . Equivalentemente, podemos notar que si  $f(z)$  tenía una expansión de Taylor (la cual todavía no hemos dado), entonces  $\delta_j(z, z_j)$  sería del orden de  $z - z_j$ , aproximándose a cero cuando la red fuera hecha muy fina. Podemos hacer

$$|\delta_j(z, z_j)| < \varepsilon ,\tag{5.40}$$

donde  $\varepsilon$  es una cantidad arbitrariamente escogida pequeña y positiva.

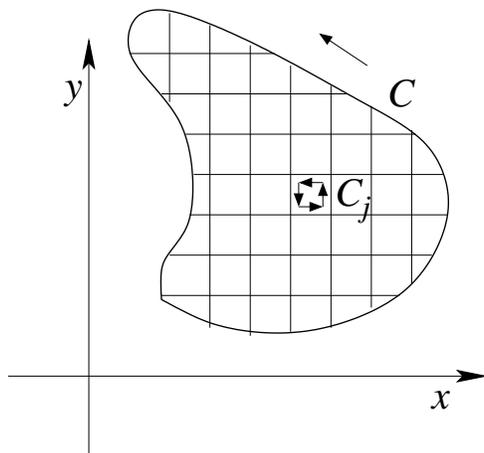


Figura 5.7: Contorno de Cauchy-Goursat.

Resolviendo la ecuación (5.39) para  $f(z)$  e integrando alrededor de  $C_j$ , obtenemos

$$\oint_{C_j} f(z) dz = \oint_{C_j} (z - z_j) \delta_j(z, z_j) dz, \quad (5.41)$$

las integrales de los otros términos se anulan.<sup>8</sup> Cuando las ecuaciones (5.40) y (5.41) están combinadas, uno puede mostrar que

$$\left| \sum_j \oint_{C_j} f(z) dz \right| < A\varepsilon, \quad (5.42)$$

donde  $A$  es un término del orden del área de la región escogida. Ya que  $\varepsilon$  es arbitrario, tomemos  $\varepsilon \rightarrow 0$  y concluimos que si una función  $f(z)$  es analítica sobre y dentro de un camino cerrado  $C$ ,

$$\oint f(z) dz = 0. \quad (5.43)$$

Detalles de la prueba de esta forma significativamente más general y poderosa puede ser encontrada en la literatura de variable compleja (Churchill). Realmente podemos todavía comprobar el teorema para  $f(z)$  analítica dentro del interior de  $C$  y solamente continuas en  $C$ .

La consecuencia del teorema integral de Cauchy es que para funciones analíticas la integral de línea es una función solamente de sus puntos extremos, independiente del camino de integración,

$$\int_{z_1}^{z_2} f(z) dz = F(z_2) - F(z_1) = - \int_{z_2}^{z_1} f(z) dz, \quad (5.44)$$

de nuevo exactamente como el caso de la fuerza conservativa.

<sup>8</sup>  $\oint dz$  y  $\oint z dz = 0$  por la ecuación (5.30).

### 5.3.4 Regiones múltiplemente conexas.

El enunciado original de nuestro teorema demanda una región simplemente conexa. Esta restricción puede ser fácilmente relajada por la creación de una barrera, una línea de contorno. Consideremos la región múltiplemente conexa de la figura 5.8, en la cual  $f(z)$  no está definida para el interior de  $R'$ . El teorema de la integral de Cauchy no es válido para el contorno  $C$ , como lo demostramos, pero podemos construir un contorno  $C'$  para el cual el teorema se mantenga. Dibujemos una línea desde el interior de la región prohibida  $R'$  a la región exterior prohibida a  $R$  y luego corra un nuevo contorno  $C'$ , como se muestra en la figura 5.9.

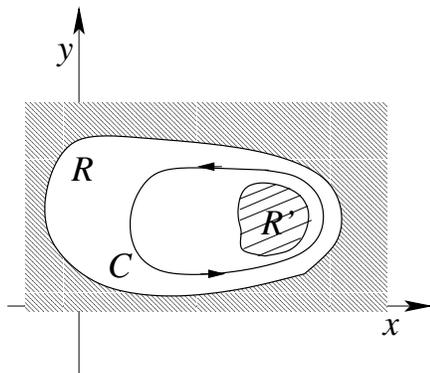


Figura 5.8: Contorno cerrado  $C$  en una región múltiplemente conexa.

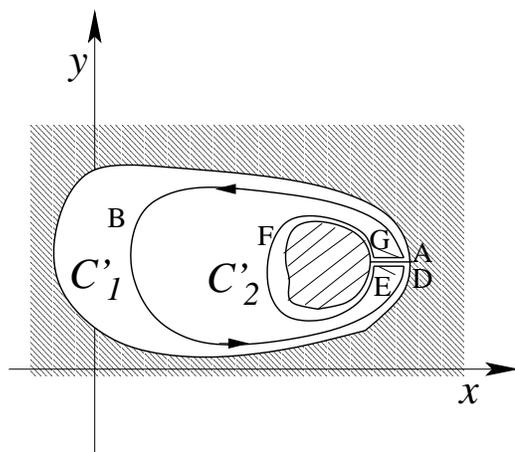


Figura 5.9: Conversión de una región múltiplemente conexa a una región simplemente conexa.

El nuevo contorno  $C'$  a través de  $ABDEFGA$  nunca cruzará la línea de contorno que literalmente convierte a  $R$  en una región simplemente conexa. Por la ecuación (5.44)

$$\int_G^A f(z) dz = - \int_E^D f(z) dz , \quad (5.45)$$

$f(z)$  sido continua a través de la línea de contorno y los segmentos de línea  $DE$  y  $GA$  están arbitrariamente cerca. Entonces

$$\begin{aligned} \oint_{C'} f(z) dz &= \int_{ABD} f(z) dz + \int_{EFG} f(z) dz \\ &= 0 \end{aligned} \quad (5.46)$$

por el teorema integral de Cauchy, con la región  $R$  simplemente conexa. Aplicando la ecuación (5.44) una vez más con  $ABD \rightarrow C'_1$  y  $EFG \rightarrow -C'_2$ , obtenemos

$$\oint_{C'_1} f(z) dz = \oint_{C'_2} f(z) dz, \quad (5.47)$$

en la cual  $C'_1$  y  $C'_2$  ambos son recorridos en la misma dirección (de los punteros del reloj).

Debería enfatizarse que la línea de contorno aquí es un asunto de conveniencia matemática para permitir la aplicación del teorema integral de Cauchy. Ya que  $f(z)$  es analítica en la región anular, es necesariamente univaluada y continua a través de cualquier línea de contorno. Cuando consideremos puntos de ramificación nuestras funciones no serán univaluadas y una línea de corte será requerida para hacerlas univaluadas.

## 5.4 Fórmula integral de Cauchy.

Como en la sección anterior, consideremos una función  $f(z)$  que es analítica sobre un contorno cerrado  $C$  y dentro de la región rodeada por  $C$ . Buscamos probar que

$$\oint_C \frac{f(z)}{z - z_0} dz = 2\pi i f(z_0), \quad (5.48)$$

en la cual  $z_0$  es algún punto en la región interior encerrada por  $C$ . Este es el segundo de los dos teoremas básicos mencionados anteriormente. Note que  $z$  está sobre el contorno de  $C$  mientras  $z_0$  está en el interior, luego  $z - z_0 \neq 0$  y la integral de la ecuación (5.48) está bien definida.

Aunque  $f(z)$  se supone analítica, el integrando es  $f(z)/(z - z_0)$  y no es analítico en  $z = z_0$  a menos que  $f(z_0) = 0$ . Si el contorno está deformado como se muestra en la figura 5.10 se aplica el teorema de la integral de Cauchy. Por la ecuación (5.47)

$$\oint_C \frac{f(z)}{z - z_0} dz - \oint_{C_2} \frac{f(z)}{z - z_0} dz = 0, \quad (5.49)$$

donde  $C$  es el contorno externo original y  $C_2$  es el círculo circundante del punto  $z_0$  recorrido en sentido contrario a los punteros del reloj. Sea  $z = z_0 + re^{i\theta}$ , usando la representación de coordenadas polares a causa de la forma circular del camino en torno a  $z_0$ . Aquí  $r$  es pequeño y eventualmente se aproximará a cero. Tenemos

$$\oint_{C_2} \frac{f(z)}{z - z_0} dz = \oint_{C_2} \frac{f(z_0 + re^{i\theta})}{re^{i\theta}} rie^{i\theta} d\theta.$$

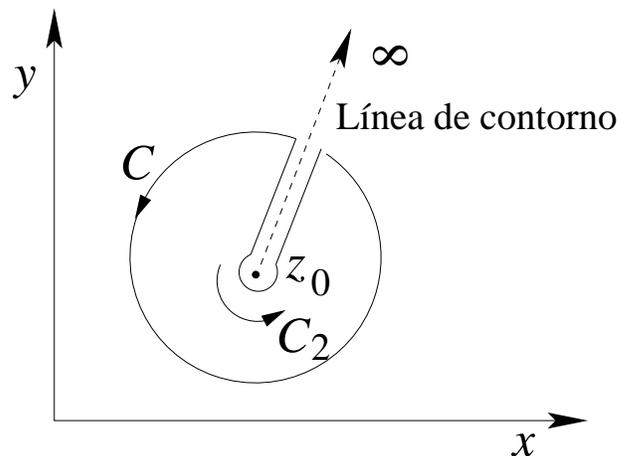


Figura 5.10: Exclusión de un punto singular.

Tomando el límite cuando  $r \rightarrow 0$ , obtenemos

$$\begin{aligned} \oint_{C_2} \frac{f(z)}{z - z_0} dz &= if(z_0) \int_{C_2} d\theta \\ &= 2\pi i f(z_0), \end{aligned} \quad (5.50)$$

ya que  $f(z)$  es analítica y por lo tanto continua en  $z = z_0$ . Esto prueba la fórmula de la integral de Cauchy.

Este es un resultado notable. El valor de una función analítica  $f(z)$  está dado en un punto interior  $z = z_0$  una vez que los valores sobre el contorno  $C$  son especificados. Esto es íntimamente análogo a una forma de la ley de Gauss bidimensional en la cual la magnitud de una carga lineal interna estaría dada en términos de la integral de superficie cilíndrica del campo eléctrico  $\vec{E}$ .

Una analogía adicional es la determinación de una función en el espacio real por una integral de la función y la correspondiente función de Green (y sus derivadas) sobre la superficie de contorno. La teoría de difracción de Kirchoff es un ejemplo de esto.

Ha sido enfatizado que  $z_0$  es un punto interior. ¿Qué ocurre si  $z_0$  es exterior a  $C$ ? En este caso el integrando entero es analítico sobre y dentro de  $C$ . El teorema de la integral de Cauchy, se aplica y la integral se anula. Tenemos

$$\boxed{\frac{1}{2\pi i} \oint_C \frac{f(z) dz}{z - z_0} = \begin{cases} f(z_0), & z_0 \text{ interior} \\ 0, & z_0 \text{ exterior} \end{cases}}$$

### 5.4.1 Derivadas.

La fórmula integral de Cauchy puede ser usada para obtener una expresión para la derivada de  $f(z)$ . A partir de la ecuación (5.48), con  $f(z)$  analítica,

$$\frac{f(z_0 + \delta z_0) - f(z_0)}{\delta z_0} = \frac{1}{2\pi i \delta z_0} \left( \oint \frac{f(z)}{z - z_0 - \delta z_0} dz - \oint \frac{f(z)}{z - z_0} dz \right).$$

Entonces, por definición de la derivada (ecuación (5.17)),

$$\begin{aligned} f'(z_0) &= \lim_{\delta z_0 \rightarrow 0} \frac{1}{2\pi i \delta z_0} \oint \frac{\delta z_0 f(z)}{(z - z_0 - \delta z_0)(z - z_0)} dz \\ &= \frac{1}{2\pi i} \oint \frac{f(z)}{(z - z_0)^2} dz . \end{aligned} \quad (5.51)$$

Podemos ver que este resultado podría haber sido obtenido por diferenciación de la ecuación (5.48) bajo el signo integral con respecto a  $z_0$ .

Esta técnica para construir derivadas puede ser repetida. Escribamos  $f'(z_0 + \delta z_0)$  y  $f'(z_0)$ , usando la ecuación (5.51). Sustrayendo, dividiendo por  $\delta z_0$ , y finalmente tomando el límite cuando  $\delta z_0 \rightarrow 0$ , obtenemos

$$f^{(2)}(z_0) = \frac{2}{2\pi i} \oint \frac{f(z)}{(z - z_0)^3} dz .$$

Notemos que  $f^{(2)}(z_0)$  es independiente de la dirección de  $\delta z_0$  como debe ser. Continuando, obtenemos

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint \frac{f(z)}{(z - z_0)^{n+1}} dz . \quad (5.52)$$

esto es, el requerimiento que  $f(z)$  sea analítica no sólo garantiza una primera derivada sino que las derivadas de todos los ordenes también. Las derivadas de  $f(z)$  son automáticamente analíticas.

### 5.4.2 Teorema de Morera.

Otra aplicación de la fórmula de la integral de Cauchy está en la prueba del teorema de Morera, el cual es la inversión del teorema de la integral de Cauchy. El teorema establece lo siguiente:

*Si una función  $f(z)$  es continua en una región simplemente conexa  $R$  y  $\oint_C f(z) dz = 0$  para cada contorno cerrado  $C$  dentro de  $R$ , entonces  $f(z)$  es analítica a través de  $R$ .*

Integremos  $f(z)$  desde  $z_1$  a  $z_2$ . Ya que cada integral cerrada de camino de  $f(z)$  se anula, la integral es independiente del camino y depende solamente de sus puntos extremos. Etiquetemos el resultado de la integración  $F(z)$ , con

$$F(z_2) - F(z_1) = \int_{z_1}^{z_2} f(z) dz . \quad (5.53)$$

Como una identidad,

$$\frac{F(z_2) - F(z_1)}{z_2 - z_1} - f(z_1) = \frac{\int_{z_1}^{z_2} [f(t) - f(z_1)] dt}{z_2 - z_1} , \quad (5.54)$$

usando  $t$  como otra variable compleja. Ahora tomemos el límite cuando  $z_2 \rightarrow z_1$ .

$$\lim_{z_2 \rightarrow z_1} \frac{\int_{z_1}^{z_2} [f(t) - f(z_1)] dt}{z_2 - z_1} = 0 , \quad (5.55)$$

ya que  $f(t)$  es continua<sup>9</sup>. Por lo tanto

$$\lim_{z_2 \rightarrow z_1} \frac{F(z_2) - F(z_1)}{z_2 - z_1} = F'(z) \Big|_{z=z_1} = f(z_1) \quad (5.56)$$

por definición de derivada (ecuación (5.17)). Hemos probado que  $F'(z)$  en  $z = z_1$  existe y es igual a  $f(z_1)$ . Ya que  $z_1$  es cualquier punto en  $R$ , vemos que  $F(z)$  es analítica. Entonces por la fórmula de la integral de Cauchy (compare la ecuación (5.52))  $F'(z) = f(z)$  es también analítica, probando el teorema de Morera.

Dibujando una vez más en nuestro análogo electrostático, podríamos usar  $f(z)$  para representar el campo electrostático  $\vec{E}$ . Si la carga neta dentro de cada región encerrada en  $R$  es cero (ley de Gauss), la densidad de carga es donde quiera igual a cero en  $R$ . Alternativamente, en términos del análisis vectorial, si  $f(z)$  representa una fuerza conservativa (por definición de conservativa), y entonces encontramos que es siempre posible expresarla como la derivada de una función potencial  $F(z)$ .

Si  $f(z) = \sum a_n z^n$  es analítica y acotada,  $|f(z)| \leq M$  sobre un círculo de radio  $r$  alrededor del origen, entonces

$$|a_n| r^n \leq M \quad (5.57)$$

da cotas superiores para los coeficientes de su expansión de Taylor. Para probar la ecuación (5.57) definamos  $M(r) = \max_{|z|=r} |f(z)|$  y usemos la integral de Cauchy para  $a_n$

$$|a_n| = \frac{1}{2\pi} \left| \int_{|z|=r} \frac{f(z)}{z^{n+1}} dz \right| \leq M(r) \frac{2\pi r}{2\pi r^{n+1}}.$$

Una consecuencia inmediata de la desigualdad (5.57) es el *teorema de Liouville*: Si  $f(z)$  es analítica y acotada en el plano complejo ella es constante. En efecto, si  $|f(z)| \leq M$  para cualquier  $z$ , entonces la desigualdad de Cauchy (5.57) da  $|a_n| \leq M r^{-n} \rightarrow 0$  cuando  $r \rightarrow \infty$  para todo  $n > 0$ . De modo que  $f(z) = a_0$ .

Inversamente, la desviación más leve de una función analítica a partir de un valor constante implica que debe haber al menos una singularidad en alguna parte del plano complejo infinito. A parte de las funciones constantes triviales, entonces, las singularidades son un hecho de la vida, y debemos aprender a vivir con ellas. Pero haremos más que esto. Expandiremos una función en una serie de Laurent en torno a una singularidad, y usaremos singularidades para desarrollar el poderoso y útil cálculo del residuo en el próximo capítulo.

El teorema fundamental del álgebra, el cual dice que cualquier polinomio  $P(z) = \sum_{v=0}^n a_v z^v$  con  $n > 0$  y  $a_n \neq 0$  tiene  $n$  raíces, también se deduce del teorema de Liouville. Suponga que  $P(z)$  no tiene cero. Entonces,  $1/P(z)$  es analítica y acotada cuando  $|z| \rightarrow \infty$ . De modo que  $f(z)$  es una constante por el teorema de Liouville, q.e.a. Por lo tanto  $P(z)$  tiene al menos una raíz por la cual puede ser dividida. Luego repetimos el proceso para el polinomio resultante de grado  $n - 1$ . Esto tiende a la conclusión que  $P(z)$  tiene exactamente  $n$  raíces.

<sup>9</sup>Podemos usar aquí el teorema del valor medio.

## 5.5 Expansión de Laurent.

### 5.5.1 Expansión de Taylor.

La fórmula de la integral de Cauchy de la sección precedente abre el camino para otra derivación de la serie de Taylor, pero esta vez para funciones de una variable compleja. Supongamos que tratamos de expandir  $f(z)$  en torno a  $z = z_0$  y que tenemos  $z = z_1$  como el punto más cercano sobre el plano complejo para el cual  $f(z)$  no es analítica. Construimos un círculo  $C$  centrado en  $z = z_0$  con radio  $|z' - z_0| < |z_1 - z_0|$  (5.11). Ya que supusimos  $z_1$  como el punto más cercano en el cual  $f(z)$  no era analítica,  $f(z)$  es necesariamente analítica sobre y dentro de  $C$ .

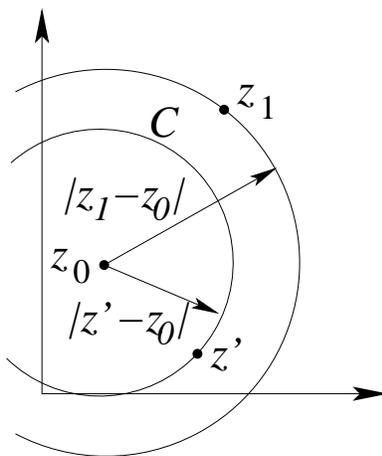


Figura 5.11: Expansión en torno a  $z_0$  con  $z_1$  el primer punto en que  $f(z)$  no es analítica.

A partir de la ecuación (5.48), la fórmula de la integral de Cauchy,

$$\begin{aligned}
 f(z) &= \frac{1}{2\pi i} \oint_C \frac{f(z') dz'}{z' - z} \\
 &= \frac{1}{2\pi i} \oint_C \frac{f(z') dz'}{(z' - z_0) - (z - z_0)} \\
 &= \frac{1}{2\pi i} \oint_C \frac{f(z') dz'}{(z' - z_0)[1 - (z - z_0)/(z' - z_0)]}.
 \end{aligned} \tag{5.58}$$

Aquí  $z'$  es un punto en el contorno de  $C$  y  $z$  es cualquier punto interior a  $C$ . No es rigurosamente legal expandir el denominador del integrando en la ecuación (5.58) por el teorema binomial, porque todavía no hemos probado el teorema binomial para variables complejas. En cambio, notemos la identidad

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + \dots = \sum_{n=0}^{\infty} t^n, \tag{5.59}$$

la cual puede ser fácilmente verificada multiplicando ambos lados por  $1-t$ . La serie infinita, siguiendo el método de la sección 4.2, es convergente para  $|t| < 1$ .

Ahora para el punto  $z$  interior a  $C$ ,  $|z - z_0| < |z' - z_0|$ , y usando la ecuación (5.59), la ecuación (5.58) se convierte en

$$f(z) = \frac{1}{2\pi i} \oint_C \sum_{n=0}^{\infty} \frac{(z - z_0)^n f(z') dz'}{(z' - z_0)^{n+1}} \quad (5.60)$$

Intercambiando el orden de la integración y de la suma (válida ya que la ecuación (5.59) es uniformemente convergente para  $|t| < 1$ ), obtenemos

$$f(z) = \frac{1}{2\pi i} \sum_{n=0}^{\infty} (z - z_0)^n \oint_C \frac{f(z') dz'}{(z' - z_0)^{n+1}} \quad (5.61)$$

Usando la ecuación (5.52), obtenemos

$$f(z) = \sum_{n=0}^{\infty} (z - z_0)^n \frac{f^{(n)}(z_0)}{n!}, \quad (5.62)$$

la cual es nuestra expansión de Taylor deseada. Note que está basada solamente en la suposición de que  $f(z)$  es analítica para  $|z - z_0| < |z_1 - z_0|$ . Al igual que para series de potencia de variable real, esta expansión es única para un  $z_0$  dado.

### 5.5.2 Principio de Reflexión de Schwarz.

A partir de la expansión binomial de  $g(z) = (z - x_0)^n$  para  $n$  entero es fácil ver que el conjugado complejo de una función es la función del complejo conjugado, para  $x_0$  real

$$g^*(z) = ((z - x_0)^n)^* = (z^* - x_0)^n = g(z^*). \quad (5.63)$$

Esto conduce al principio de reflexión de Schwarz: *Si una función  $f(z)$  es (1) analítica sobre alguna región incluyendo el eje real y (2) real cuando  $z$  es real, entonces*

$$f^*(z) = f(z^*). \quad (5.64)$$

Ver figura 5.12.

Expandiendo  $f(z)$  alrededor de algún punto (no singular)  $x_0$  en el eje real,

$$f(z) = \sum_{n=0}^{\infty} (z - x_0)^n \frac{f^{(n)}(x_0)}{n!}, \quad (5.65)$$

por la ecuación (5.61). Ya que  $f(z)$  es analítica en  $z = x_0$ , esta expansión de Taylor existe. Ya que  $f(z)$  es real cuando  $z$  es real,  $f^{(n)}(x_0)$  debe ser real para todo  $n$ . Luego cuando usemos la ecuación (5.63), la ecuación (5.64), el principio de reflexión de Schwarz, sigue inmediatamente.

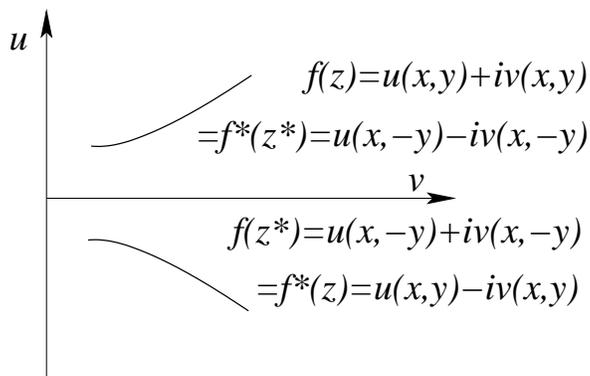


Figura 5.12: Principio de reflexión de Schwarz.

### 5.5.3 Continuación analítica.

Es natural pensar los valores  $f(z)$  de una función analítica  $f$  como una entidad única que usualmente está definida en alguna región restringida  $S_1$  del plano complejo, por ejemplo una serie de Taylor (ver figura 5.13). Entonces  $f$  es analítica dentro del círculo de convergencia  $C_1$ , cuyo radio está dado por la distancia  $r_1$  desde el centro de  $C_1$  a la singularidad más cercana de  $f$  en  $z_1$  (en figura 5.13). Si escogemos un punto dentro de  $C_1$  que este más allá de  $r_1$  desde la singularidad de  $z_1$  y hacemos una expansión de Taylor de  $f$ , luego el círculo de convergencia  $C_2$  usualmente se extenderá más allá del primer círculo  $C_1$ . En la región que se superponen ambos círculos  $C_1, C_2$  la función  $f$  está definida en forma única. En la región del círculo  $C_2$  que se extiende más allá de  $C_1$ ,  $f(z)$  está definida en forma única por la serie de Taylor en torno al centro de  $C_2$  y es analítica allí, aunque la serie de Taylor en torno al centro de  $C_1$  no es más convergente allí. Después Weierstrass este proceso es llamado *continuación analítica*. Esta define la función analítica  $f$  en términos de su definición original (en  $C_1$ ) y todas sus continuaciones.

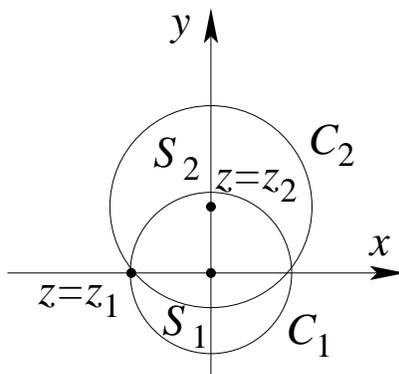


Figura 5.13: Continuación analítica.

Un ejemplo específico es la función *meromórfica*

$$f(z) = \frac{1}{1+z}, \quad (5.66)$$

la cual tiene un polo simple en  $z = -1$  y es analítica donde sea. La expansión de la serie geométrica

$$\frac{1}{1+z} = 1 - z + z^2 - \dots = \sum_{n=0}^{\infty} (-z)^n, \quad (5.67)$$

converge para  $|z| < 1$ , *i.e.*, dentro del círculo  $C_1$  en la figura 5.13. Supongamos que expandimos  $f(z)$  alrededor de  $z = i$ , tal que

$$\begin{aligned} f(z) &= \frac{1}{1+z} = \frac{1}{1+i+(z-i)} = \frac{1}{(1+i)(1+(z-i)/(1+i))} \\ &= \left[ 1 - \frac{z-i}{1+i} + \frac{(z-i)^2}{(1+i)^2} - \dots \right] \frac{1}{1+i}, \end{aligned} \quad (5.68)$$

converge para  $|z-i| < |1+i| = \sqrt{2}$ . Nuestro círculo de convergencia es  $C_2$  en la figura 5.13. Ahora  $f(z)$  está definida por la expansión (5.68) en  $S_2$  la cual se superpone a  $S_1$  y se extiende más allá en el plano complejo<sup>10</sup>. Esta extensión es una continuación analítica, y cuando tenemos solamente puntos singulares aislados con quien enfrentarnos, la función puede ser extendida indefinidamente. Las ecuaciones (5.66), (5.67) y (5.68) son tres representaciones diferentes de la misma función. Cada representación tiene su propio dominio de convergencia. La ecuación (5.67) es una serie de Maclaurin. La ecuación (5.68) es una expansión de Taylor alrededor de  $z = i$  y desde el siguiente párrafo la ecuación (5.66) se ve como una serie de Laurent de un término.

Una continuación analítica puede ser tomada de muchas formas y la serie de expansión ya consideradas no es necesariamente la técnica más conveniente. Como una técnica alternativa usaremos una relación de recurrencia para extender la función factorial alrededor de los puntos singulares aislados,  $z = -n$ ,  $n = 1, 2, 3, \dots$

#### 5.5.4 Permanencia de la forma algebraica.

Todas nuestras funciones elementales,  $e^z$ ,  $\text{sen}(z)$ , etc., pueden ser extendidas al plano complejo. Por ejemplo, ellas pueden ser definidas por expansiones de series de potencias tal como

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \dots = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad (5.69)$$

para la exponencial. Tales definiciones concuerdan con las definiciones de variable real a lo largo del eje  $x$  y literalmente constituyen una continuación analítica de las correspondientes funciones reales en el plano complejo. Este resultado a menudo es llamado *permanencia de la forma algebraica*

<sup>10</sup>Uno de los más poderosos y hermosos resultados de la más abstracta teoría de funciones de variable compleja es que si dos funciones analítica coinciden en una región, tal como la intersección de  $S_1$  y  $S_2$ , o coinciden sobre cualquier segmento de línea, ellas son la misma función en el sentido que ellas coincidirán en cualquier parte siempre que ambas estén bien definidas.

### 5.5.5 Serie de Laurent.

Frecuentemente encontramos funciones que son analíticas en una región anular, es decir, de radio interno  $r$  y radio externo  $R$ , como lo muestra la figura 5.14. Dibujando una línea de

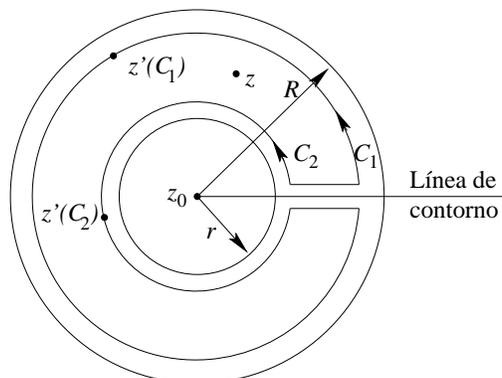


Figura 5.14:  $|z' - z_0|_{C_1} > |z - z_0|$ ;  $|z' - z_0|_{C_2} < |z - z_0|$ .

contorno auxiliar para convertir nuestra región en una región simplemente conexa, aplicamos la fórmula integral de Cauchy, y para los dos círculos,  $C_2$  y  $C_1$ , centrados en  $z = z_0$  y con radios  $r_2$  y  $r_1$ , respectivamente, donde  $r < r_2 < r_1 < R$ , tenemos<sup>11</sup>

$$f(z) = \frac{1}{2\pi i} \oint_{C_1} \frac{f(z') dz'}{z' - z} - \frac{1}{2\pi i} \oint_{C_2} \frac{f(z') dz'}{z' - z}. \quad (5.70)$$

Note cuidadosamente que en la ecuación (5.70) un signo menos explícito ha sido introducido así que el contorno  $C_2$  (como  $C_1$ ) sea recorridos en el sentido positivo (contra los punteros del reloj). El tratamiento de la ecuación (5.70) ahora procede exactamente como la ecuación (5.58) en el desarrollo de la serie de Taylor. Cada denominador escrito como  $(z' - z_0) - (z - z_0)$  y expandido por el teorema del binomio el cual ahora sigue a la serie de Taylor (ecuación (5.62)). Notando que para  $C_1$ ,  $|z' - z_0| > |z - z_0|$  mientras que para  $C_2$ ,  $|z' - z_0| < |z - z_0|$ , encontramos

$$f(z) = \frac{1}{2\pi i} \sum_{n=0}^{\infty} (z - z_0)^n \oint_{C_1} \frac{f(z') dz'}{(z' - z_0)^{n+1}} + \frac{1}{2\pi i} \sum_{n=1}^{\infty} (z - z_0)^{-n} \oint_{C_2} (z' - z_0)^{n-1} f(z') dz'. \quad (5.71)$$

El signo menos de la ecuación (5.70) ha sido absorbido por la expansión binomial. Etiquetando la primera serie  $S_1$  y la segunda  $S_2$ ,

$$S_1 = \frac{1}{2\pi i} \sum_{n=0}^{\infty} (z - z_0)^n \oint_{C_1} \frac{f(z') dz'}{(z' - z_0)^{n+1}}, \quad (5.72)$$

<sup>11</sup>Podemos tomar  $r_2$  arbitrariamente cerca de  $r$  y  $r_1$  arbitrariamente cerca de  $R$ , maximizando el área encerrada por  $C_1$  y  $C_2$ .

la cual es la expansión de Taylor regular, convergente para  $|z - z_0| < |z' - z_0| = r_1$ , esto es, para todo  $z$  interior del círculo,  $C_1$ . Para la segunda serie de la ecuación tenemos

$$S_2 = \frac{1}{2\pi i} \sum_{n=1}^{\infty} (z - z_0)^{-n} \oint_{C_2} (z' - z_0)^{n-1} f(z') dz' \quad (5.73)$$

convergente para  $|z - z_0| > |z' - z_0| = r_2$ , esto es, para todo  $z$  exterior al círculo más pequeño  $C_2$ . Recuerde,  $C_2$  ahora va en contra del sentido de los punteros del reloj.

Estas dos series pueden ser combinadas en una serie<sup>12</sup>( una serie de Laurent) por

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n, \quad (5.74)$$

donde

$$a_n = \frac{1}{2\pi i} \oint_C \frac{f(z') dz'}{(z' - z_0)^{n+1}} \quad (5.75)$$

Ya que, en la ecuación (5.75), la convergencia de la expansión binomial ya no es un problema,  $C$  puede ser cualquier contorno dentro de una región anular  $r < |z - z_0| < R$  encerrando a  $z_0$  una vez en sentido antihorario. Si suponemos que en una región anular la convergencia existe, la ecuación (5.74) es la serie de Laurent o la expansión de Laurent de  $f(z)$ .

El uso de una línea de contorno (figura 5.14) es conveniente para convertir la región anular en una región simplemente conexa. Ya que nuestra función es analítica en esa región anular (y por lo tanto univaluada), la línea de contorno no es esencial y, por cierto, las funciones con puntos de ramificación deben tener líneas de corte. Los coeficientes de las series de Laurent no necesitan venir de la evaluación de las integrales de contorno ( las cuales pueden ser muy difíciles de trabajar). Otras técnicas tales como las expansiones de series ordinarias pueden dar los coeficientes.

Nos limitaremos a un ejemplo sencillo para ilustrar la aplicación de la ecuación (5.74).

### Ejemplo

Sea  $f(z) = [z(z - 1)]^{-1}$ . Si escogemos  $z_0 = 0$ , entonces  $r = 0$  y  $R = 1$ ,  $f(z)$  diverge en  $z = 1$ . A partir de las ecuaciones (5.75) y (5.74)

$$\begin{aligned} a_n &= \frac{1}{2\pi i} \oint \frac{dz'}{(z')^{n+2}(z' - 1)^{n+1}} \\ &= \frac{1}{2\pi i} \oint \sum_{m=0}^{\infty} (z')^m \frac{dz'}{(z')^{n+2}}. \end{aligned} \quad (5.76)$$

De nuevo, intercambiando el orden de la suma y la integración (series uniformemente convergentes), tenemos

$$a_n = \frac{1}{2\pi i} \sum_{m=0}^{\infty} \oint \frac{dz'}{(z')^{n+2-m}}. \quad (5.77)$$

<sup>12</sup>Reemplazando  $n$  por  $-n$  en  $S_2$  y sumando.

Si empleamos la forma polar, como en la ecuación (5.52)

$$\begin{aligned} a_n &= \frac{1}{2\pi i} \sum_{m=0}^{\infty} \oint \frac{r i e^{i\theta} d\theta}{r^{n+2-m} e^{i(n+2-m)\theta}} \\ &= \frac{1}{2\pi i} 2\pi i \sum_{m=0}^{\infty} \delta_{n+2-m,1} . \end{aligned} \quad (5.78)$$

En otras palabras,

$$a_n = \begin{cases} -1 & \text{para } n \geq -1, \\ 0 & \text{para } n < -1. \end{cases} \quad (5.79)$$

La expansión de Laurent (ecuación (5.74)) se convierte en

$$\begin{aligned} \frac{1}{z(z-1)} &= -\frac{1}{z} - 1 - z - z^2 - z^3 - \dots \\ &= -\sum_{n=-1}^{\infty} z^n . \end{aligned} \quad (5.80)$$

Para esta función simple la serie de Laurent puede, por su puesto, ser obtenida por una expansión binomial directa.

La serie de Laurent difiere de la de Taylor por las características obvias de las potencias negativas de  $(z - z_0)$ . Por esta razón la serie de Laurent siempre divergirá al menos en  $z = z_0$  y quizás en cuanto a alguna distancia más allá de  $r$  (figura 5.14).

## 5.6 Mapeo.

En las secciones precedentes hemos definido las funciones analíticas y hemos desarrollado alguna de sus principales características. A partir de estos desarrollo las relaciones integrales del próximo capítulo se derivan directamente. Aquí introducimos algunos de los aspectos más geométricos de las funciones de variables compleja, aspectos que serán útiles en la visualización de las operaciones integrales en el próximo capítulo y que serán valiosas en sí mismas para resolver la ecuación de Laplace en un sistema de dos dimensiones.

En geometría analítica común podemos tomar  $y = f(x)$  y graficar  $y$  versus  $x$ . Nuestro problema aquí es más complicado, porque  $z$  es una función de dos variables  $x$  e  $y$ . Usamos la notación

$$w = f(z) = u(x, y) + iv(x, y) , \quad (5.81)$$

Entonces para un punto en el plano  $z$  (valores específicos de  $x$  e  $y$ ) allí puede corresponder valores específicos para  $u(x, y)$  y  $v(x, y)$  los cuales producen luego un punto en el plano  $w$ . Así los puntos en el plano  $z$  transforman o son mapeados en puntos en el plano  $w$ , líneas o áreas en el plano  $z$  serán mapeados en líneas o áreas en el plano  $w$ . Nuestro propósito inmediato es ver cuantas líneas o áreas mapean desde el plano  $z$  al plano  $w$  para un número de funciones simples.

### 5.6.1 Traslación.

$$w = z + z_0 . \quad (5.82)$$

La función  $w$  es igual a la variable  $z$  más una constante,  $z_0 = x_0 + iy_0$ . Por las ecuaciones (5.1) y (5.81)

$$\begin{aligned} u &= x + x_0 , \\ v &= y + y_0 , \end{aligned} \quad (5.83)$$

representando una traslación pura de los ejes de coordenadas como muestra la figura 5.15.

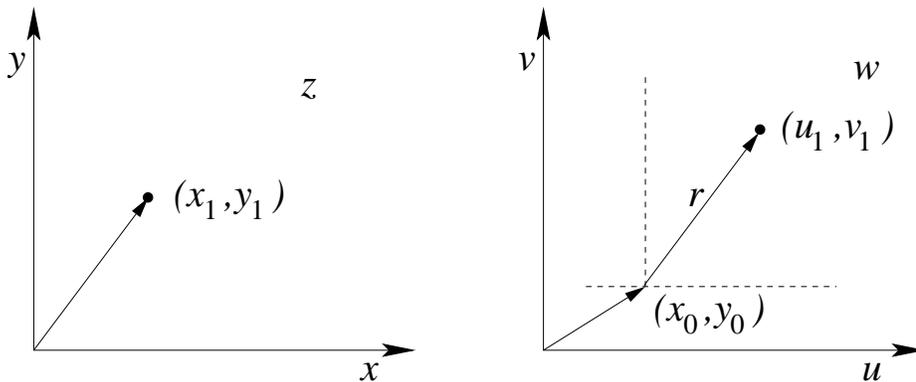


Figura 5.15: Traslación.

### 5.6.2 Rotación.

$$w = z z_0 . \quad (5.84)$$

Aquí es conveniente volver a la representación de coordenadas polares, usando

$$w = \rho e^{i\varphi} , \quad z = r e^{i\theta} \quad \text{y} \quad z_0 = r_0 e^{i\theta_0} , \quad (5.85)$$

entonces

$$\rho e^{i\varphi} = r r_0 e^{i(\theta + \theta_0)} \quad (5.86)$$

o

$$\begin{aligned} \rho &= r r_0 \\ \varphi &= \theta + \theta_0 . \end{aligned} \quad (5.87)$$

Dos cosas han ocurrido. Primero, el módulo  $r$  ha sido modificado, ya sea expandido o contraído, por el factor  $r_0$ . Segundo, el argumento  $\theta$  ha sido aumentado por la constante aditiva  $\theta_0$  (figura 5.16). Esto representa una rotación de la variable compleja en un ángulo  $\theta_0$ . Para el caso especial de  $z_0 = i$ , tenemos una rotación pura en  $\pi/2$  radianes.

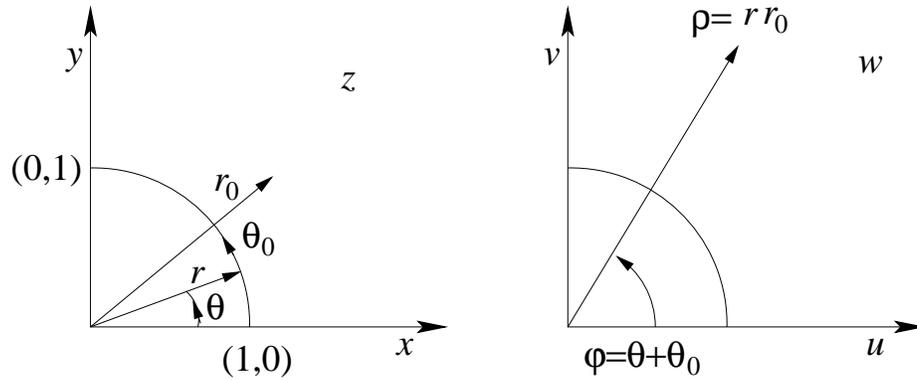


Figura 5.16: Rotación.

### 5.6.3 Inversión.

$$w = \frac{1}{z}. \quad (5.88)$$

Nuevamente, usando la forma polar, tenemos

$$\rho e^{i\varphi} = \frac{1}{r e^{i\theta}} = \frac{1}{r} e^{-i\theta}, \quad (5.89)$$

la cual muestra que

$$\rho = \frac{1}{r}, \quad \varphi = -\theta. \quad (5.90)$$

La primera parte de la ecuación (5.90) muestra claramente una inversión. El interior del círculo unitario es mapeado en el exterior y vice versa (figura 5.17). En suma, la segunda parte de la ecuación (5.90) muestra que el ángulo polar se le invierte el signo. La ecuación (5.88), por lo tanto, también involucra una reflexión del eje  $y$  exactamente como la ecuación del complejo conjugado.

Para ver como las líneas en el plano  $z$  se transforman en el plano  $w$ , simplemente volvamos a la forma cartesiana:

$$u + iv = \frac{1}{x + iy}. \quad (5.91)$$

Racionalizando el lado derecho multiplicando el numerador y el denominador por  $z^*$  y luego igualando las partes reales e imaginarias, tenemos

$$\begin{aligned} u &= \frac{x}{x^2 + y^2}, & x &= \frac{u}{u^2 + v^2}, \\ v &= -\frac{y}{x^2 + y^2}, & y &= -\frac{v}{u^2 + v^2}. \end{aligned} \quad (5.92)$$

Un círculo centrado en el origen en el plano  $z$  tiene la forma

$$x^2 + y^2 = r^2, \quad (5.93)$$

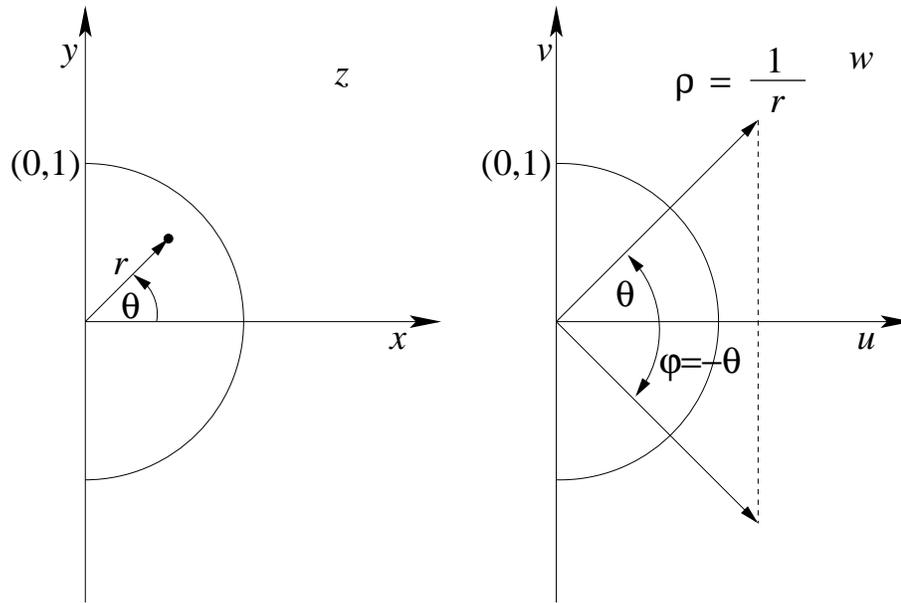


Figura 5.17: Inversión.

y por la ecuación (5.92) se transforma en

$$\frac{u^2}{(u^2 + v^2)^2} + \frac{v^2}{(u^2 + v^2)^2} = r^2. \quad (5.94)$$

Simplificando la ecuación (5.94), obtenemos

$$u^2 + v^2 = \frac{1}{r^2} = \rho^2, \quad (5.95)$$

la cual describe un círculo en el plano  $w$  también centrado en el origen.

La línea horizontal  $y = c_1$  se transforma en

$$\frac{-v}{(u^2 + v^2)^2} = c_1, \quad (5.96)$$

o

$$u^2 + v^2 + \frac{v}{c_1} + \frac{1}{(2c_1)^2} = \frac{1}{(2c_1)^2} \quad (5.97)$$

la cual describe un círculo en el plano  $w$  de radio  $(1/2)c_1$  y centrado en  $u = 0, v = -1/2c_1$  (figura 5.18).

Podemos probar otras posibilidades,  $x = \pm c_1, y = -c_1$ , o rotar los ejes  $xy$ . En general, cualquier línea recta o círculo en el plano  $z$  transformará en línea recta o círculo en el plano  $w$ .

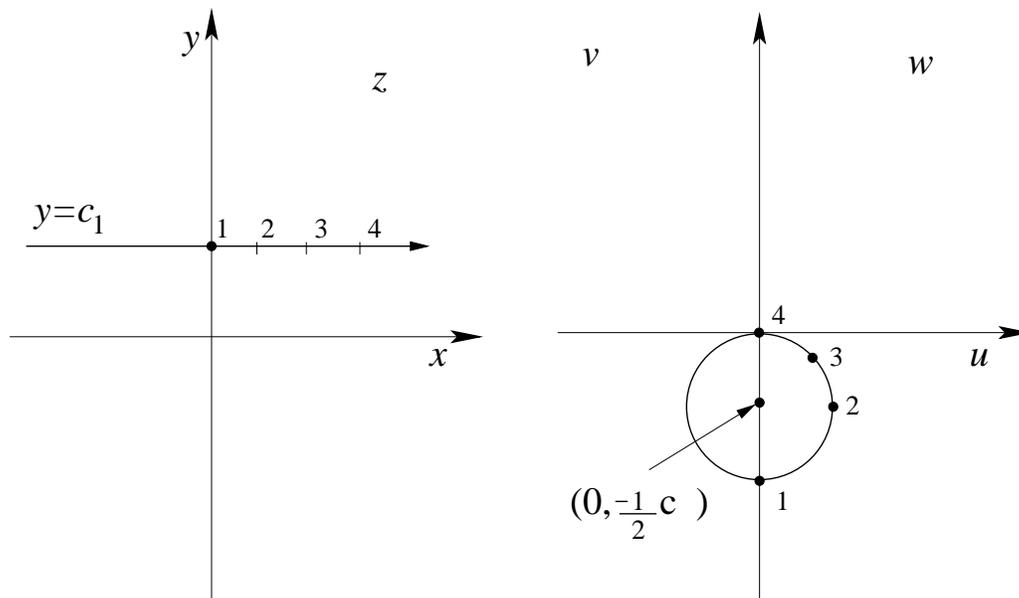


Figura 5.18: Inversión, línea-círculo.

#### 5.6.4 Puntos de ramificación y funciones multivaluadas.

Las tres transformaciones ya discutidas todas han involucrado una correspondencia uno a uno de puntos en el plano  $z$  a puntos en el plano  $w$ . Ahora ilustraremos la variedad de transformaciones que son posibles y los problemas que ellas puedan presentar, introduciremos una correspondencia de dos a uno y luego una correspondencia de muchos a uno. Finalmente, tomaremos los inversos de esas dos transformaciones.

Consideremos primero la transformación

$$w = z^2, \quad (5.98)$$

la cual conduce a

$$\rho = r^2, \quad \varphi = 2\theta. \quad (5.99)$$

Claramente, nuestra transformación es no lineal, para el módulo es cuadrado, pero la característica significativa de la ecuación (5.99) es que el ángulo fase o argumento está doblado. Esto significa que:

primer cuadrante de  $z$ ,  $0 \leq \theta < \pi/2 \rightarrow$  semi plano superior de  $w$ ,  $0 \leq \varphi < \pi$ ,  
 semi plano superior de  $z$ ,  $0 \leq \theta < \pi \rightarrow$  plano completo de  $w$ ,  $0 \leq \theta < 2\pi$ .

El semi plano inferior de  $z$  mapea sobre el ya cubierto el plano completo de  $w$ , cubriendo el plano  $w$  por segunda vez. Esta es nuestra correspondencia dos a uno, dos puntos distintos en el plano  $z$ ,  $z_0$  y  $z_0 e^{i\pi} = -z_0$ , correspondiendo al punto único  $w = z_0^2$ . En representación cartesiana

$$\begin{aligned} u + iv &= (x + iy)^2 \\ &= x^2 - y^2 + i2xy, \end{aligned} \quad (5.100)$$

produciendo

$$\begin{aligned} u &= x^2 - y^2, \\ v &= 2xy. \end{aligned} \quad (5.101)$$

De modo que las líneas  $u = c_1$ ,  $v = c_2$  en el plano  $w$  corresponden a  $x^2 - y^2 = c_1$ ,  $2xy = c_2$ , hipérbolas rectangular (y ortogonal) en el plano  $z$  (figura 5.19). Para cada punto sobre la

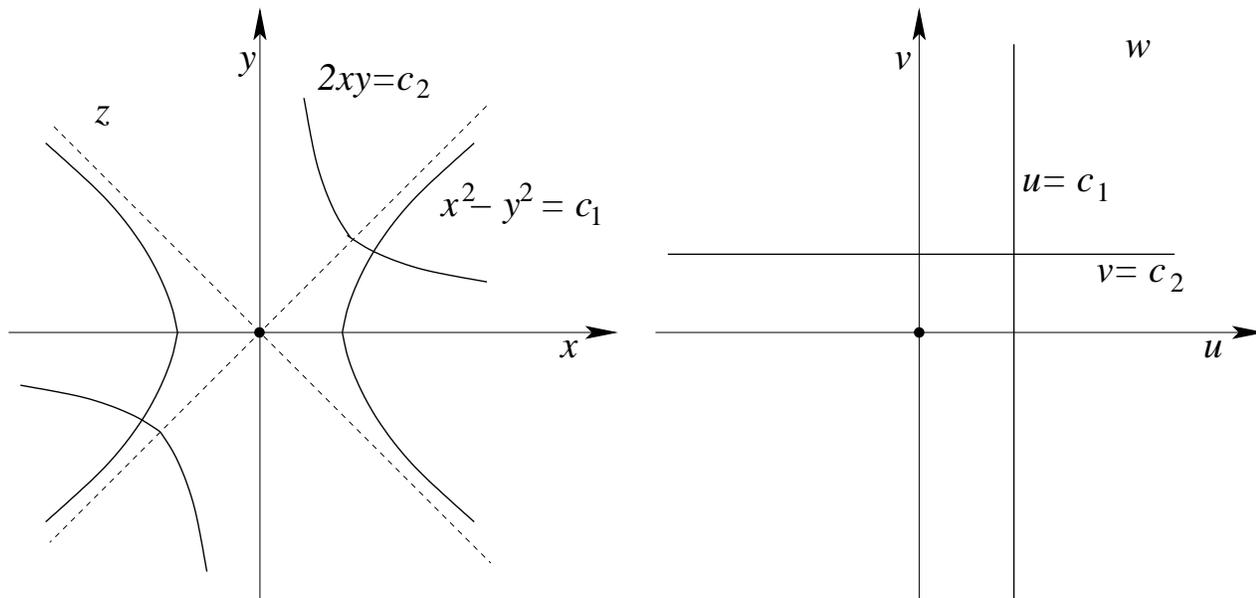


Figura 5.19: Mapeo en coordenadas hiperbólicas.

hipérbola  $x^2 - y^2 = c_1$  en el semi plano derecho,  $x > 0$ , un punto sobre la línea  $u = c_1$  también corresponde a un punto sobre la hipérbola  $x^2 - y^2 = c_1$  y en el semi plano izquierdo,  $x < 0$ , como se explicó.

En la próxima sección mostraremos que si las líneas en el plano  $w$  son ortogonales las correspondientes líneas en el plano  $z$  también son ortogonales, ya que la transformación es analítica. Aunque  $u = c_1$  y  $v = c_2$  son construidas en forma perpendicular una a otra, las correspondientes hipérbolas en el plano  $z$  son ortogonales. Literalmente hemos construido un nuevo sistema ortogonal de líneas hiperbólicas (o superficies si añadimos un eje perpendicular a  $x$  e  $y$ ). Podría notarse que si las líneas hiperbólicas son líneas de fuerza eléctrica o magnética, entonces tenemos un lente cuadrupolar útil en enfocar haces de partículas de alta energía. El inverso de la cuarta transformación (ecuación (5.98)) es

$$w = z^{1/2}. \quad (5.102)$$

De la relación

$$\rho e^{i\varphi} = r^{1/2} e^{i\theta/2}, \quad (5.103)$$

y

$$2\varphi = \theta, \quad (5.104)$$

ahora tenemos dos puntos en el plano  $w$  (argumentos  $\varphi$  y  $\varphi + \pi$ ) corresponden a un punto en el plano  $z$  (excepto para el punto  $z = 0$ ). O, poniendolo de otra manera,  $\theta$  y  $\theta + 2\pi$  corresponden a  $\varphi$  y a  $\varphi + \pi$ , dos puntos distintos en el plano  $w$ . Esto es en variable compleja el análogo de la ecuación en variable real simple  $y^2 = x$ , en la cual dos valores de  $y$ , más y menos, corresponden a cada valor de  $x$ .

El punto importante aquí es que podemos hacer la función  $w$  de la ecuación (5.102) una función univaluada en vez de una función bivaluada si acordamos en restringir  $\theta$  a un rango tal como  $0 \leq \theta < 2\pi$ . Esto puede ser hecho acordando nunca cruzar la línea  $\theta = 0$  en el plano  $z$  (figura 5.20). Cada una de las líneas de demarcación es llamada una línea de corte.

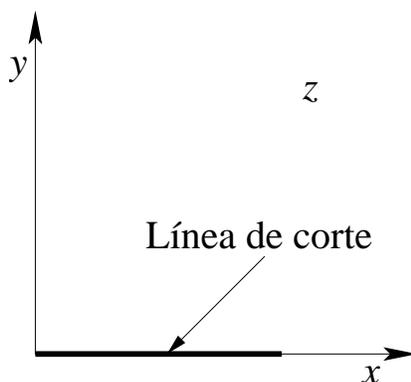


Figura 5.20: Línea de corte.

La línea de corte acopla las dos singularidades de puntos ramificados en  $0$  y  $\infty$ , donde la función es claramente no analítica. Cualquier línea desde  $z = 0$  a infinito serviría igualmente bien. El propósito de la línea de corte es restringir el argumento de  $z$ . Los puntos  $z$  y  $z \exp(2\pi i)$  coincide en el plano  $z$  pero produce diferentes puntos:  $w$  y  $-w = w \exp(\pi i)$ , en el plano  $w$ . De modo que en ausencia de una línea de corte la función  $w = z^{1/2}$  es ambigua. Alternativamente, aunque la función  $w = z^{1/2}$  es bivaluada, también podemos pegar dos hojas del plano complejo  $z$  junto a lo largo de la línea de corte tal que  $\arg(z)$  aumente más allá de  $2\pi$  a lo largo de la línea de corte bajar desde  $4\pi$  sobre la segunda hoja hasta la partida sobre la primera hoja. Esta construcción es llamada superficie de Riemann de  $w = z^{1/2}$ . Encontraremos puntos ramificados y líneas de corte frecuentemente en el próximo capítulo.

La transformación

$$w = e^z, \quad (5.105)$$

produce

$$\rho e^{i\varphi} = e^{x+iy}, \quad (5.106)$$

o

$$\begin{aligned} \rho &= e^x, \\ \varphi &= y. \end{aligned} \quad (5.107)$$

Si  $y$  pertenece a  $0 \leq y < 2\pi$  (o  $-\pi \leq y < \pi$ ), entonces  $\varphi$  cubre el mismo intervalo. Pero esto es el plano  $w$  completo. En otras palabras, una franja horizontal en el plano  $z$  de ancho  $2\pi$  mapea en el plano  $w$  entero. Más amplio, cualquier punto (por la ecuación (5.107)), en el plano  $w$ . Tenemos una correspondencia de muchos a uno (infinitamente muchos). Finalmente, el inverso de la quinta transformación (ecuación (5.105)), tenemos

$$w = \ln z . \quad (5.108)$$

Expandiendola, tenemos

$$\begin{aligned} u + iv &= \ln r e^{i\theta} \\ &= \ln r + i\theta . \end{aligned} \quad (5.109)$$

Para un punto  $z_0$  dado en el plano  $z$  el argumento  $\theta$  es inespecífico dentro de un múltiplo entero de  $2\pi$ . Esto significa que

$$v = \theta + 2n\pi , \quad (5.110)$$

y como en la transformación exponencial, tenemos una correspondencia de muchos a uno.

La ecuación (5.108) una simpática representación física. Si vamos alrededor del círculo unitario en el plano  $z$ ,  $r = 1$  y por la ecuación (5.109),  $u = \ln r = 0$ ; pero  $v = \theta$ , y  $\theta$  está aumentando establemente y continúa aumentando cuando  $\theta$  continúa, pasado  $2\pi$ .

La línea de corte acopla el punto ramificado en el origen con infinito. Cuando  $\theta$  aumenta pasado  $2\pi$  pegamos una nueva hoja del plano complejo  $z$  a lo largo de la línea de corte, etc. Haciendo un recorrido alrededor del círculo unitario en el plano  $z$  es como avanza un desatornillador como si rotara o la subida de una persona ascendiendo una escalera de espiral (figura 5.21), la cual es la superficie de Riemann de  $w = \ln z$ .

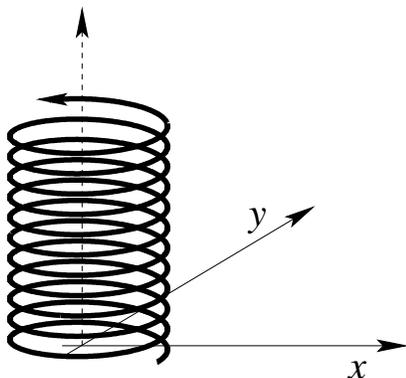


Figura 5.21: Esta es la superficie de Riemann para  $\ln(z)$ , una función multivaluada.

Como en el ejemplo precedente, también podemos hacer la correspondencia única (y ecuación (5.108) inambigua) restringiendo  $\theta$  a un intervalo tal como  $0 \leq \theta < 2\pi$  tomando la línea  $\theta = 0$  (eje real positivo) como una línea de corte. Esto es equivalente a tomar una y solamente una vuelta completa de la escalera de espiral.

Esto a causa de la naturaleza multivaluada de  $\ln z$  que la integral de contorno

$$\oint \frac{dz}{z} = 2\pi i \neq 0 ,$$

integrando alrededor del origen.

El concepto de mapeo es amplio y útil en matemáticas. Nuestro mapeo del plano complejo  $z$  a un plano complejo  $w$  es una generalización simple de una definición de función: un mapeo de  $x$  (a partir de un arreglo) en  $y$  en un segundo arreglo. Una forma más sofisticada de mapeo aparece cuando usamos la función delta de Dirac  $\delta(x - a)$  para mapear una función  $f(x)$  en su valor en el punto  $a$ .

## 5.7 Mapeo conforme

En la sección 5.6 las hipérbolas fueron mapeadas en líneas rectas y líneas rectas mapeadas en círculos. Aún en todas esas transformaciones una característica permaneció constante. Esta constancia fue un resultado del hecho que todas las transformaciones de la sección 5.6 eran analíticas. Ya que  $w = f(z)$  es una función analítica, tenemos

$$\frac{df}{dz} = \frac{dw}{dz} = \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z} \quad (5.111)$$

Suponiendo que esta ecuación está en una forma polar, podemos igualar el módulo al módulo y el argumento al argumento. Para el último (suponiedo que  $df/dz \neq 0$ )

$$\begin{aligned} \arg \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z} &= \lim_{\Delta z \rightarrow 0} \arg \frac{\Delta w}{\Delta z} \\ &= \lim_{\Delta z \rightarrow 0} \arg \Delta w - \lim_{\Delta z \rightarrow 0} \arg \Delta z \\ &= \arg \frac{df}{dz} = \alpha , \end{aligned} \quad (5.112)$$

donde  $\alpha$ , el argumento de la derivada, puede depender de  $z$  pero es una constante para un  $z$  fijo, independiente de la dirección de aproximación. Para ver el significado de esto, consideremos dos curvas,  $C_z$  en el plano  $z$  y la correspondiente curva  $C_w$  en el plano  $w$  (figura 5.22). El incremento  $\Delta z$  está mostrado en un ángulo  $\theta$  relativo al eje real ( $x$ ) mientras el correspondiente incremento  $\Delta w$  forma un ángulo de  $\varphi$  con el eje real ( $u$ ). De la ecuación (5.112)

$$\varphi = \theta + \alpha , \quad (5.113)$$

o cualquier línea en el plano  $z$  es rotado a través de un ángulo  $\alpha$  en el plano  $w$  ya que  $w$  es una transformación analítica y la derivada es distinta de cero<sup>13</sup>. Entonces para el ángulo entre estas dos líneas

$$\varphi_2 - \varphi_1 = (\theta_2 + \alpha) - (\theta_1 + \alpha) = \theta_2 - \theta_1 , \quad (5.114)$$

<sup>13</sup>Si  $df/dz = 0$  las fases quedan indefinidas y la transformación no necesariamente preservará los ángulos.

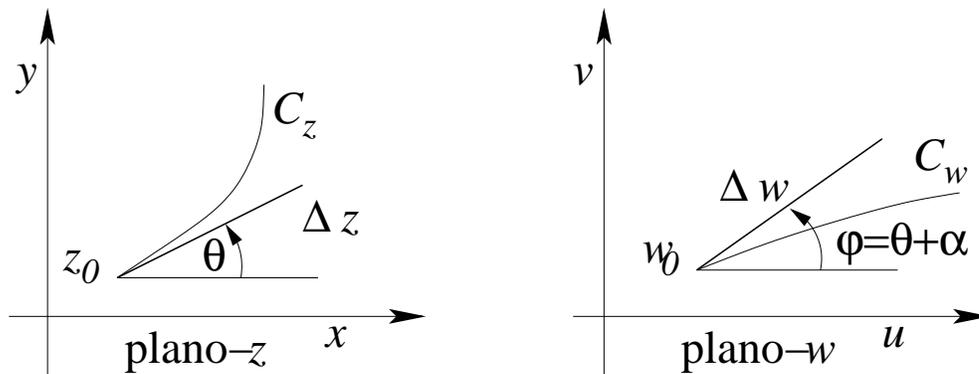


Figura 5.22: Mapeo conforme preservación de los ángulos.

la cual muestra que el ángulo incluido es preservado bajo una transformación analítica. Tal transformación que preserve el ángulo son llamadas *conformes*. El ángulo de rotación  $\alpha$ , en general, dependerá de  $z$ . Adicionalmente  $|f(z)|$  usualmente, será una función de  $z$ .

Históricamente, estas transformaciones conformes han sido de gran importancia para los científicos e ingenieros en resolver la ecuación de Laplace para problemas de electrostática, hidrodinámica, flujo de calor y etc. Desafortunadamente, el acercamiento por transformación conformes, siempre elegante, está limitada a problemas que pueden ser reducidos a dos dimensiones. El método es a menudo bello si hay una alta simetría presente pero a menudo imposible si la simetría está rota o ausente. A causa de estas limitaciones y principalmente a causa de la alta velocidad que los computadores que ofrecen una útil alternativa (soluciones iterativas de la ecuaciones diferenciales parcial), los detalles y aplicaciones del mapeo conforme son omitidas.

# Capítulo 6

## Funciones de una variable compleja II. Cálculo del residuo.

versión final 1.4-120702<sup>1</sup>

### 6.1 Singularidades.

En este capítulo volvemos a la línea de análisis que comenzamos con las condiciones de Cauchy-Riemann en el capítulo anterior y nos condujo a la expansión de Laurent. La expansión de Laurent representa una generalización de las series de Taylor en presencia de singularidades. Definimos el punto  $z_0$  como un punto singular aislado de la función  $f(z)$  si  $f(z)$  no es analítica en  $z = z_0$  pero es analítica en las cercanías del punto. Una función que es analítica en el plano complejo entero finito excepto por polos aislados es llamada *meromórfica*.

#### 6.1.1 Polos.

En la expansión de Laurent de  $f(z)$  alrededor de  $z_0$

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n . \quad (6.1)$$

Si  $a_n = 0$  para todo  $n < -m < 0$  y  $a_{-m} \neq 0$ , decimos que  $z_0$  es un polo de orden  $m$ . Por ejemplo, si  $m = 1$ ; esto es, si  $a_{-1}/(z - z_0)$  es el primer término no nulo en la serie de Laurent, tenemos un polo de orden uno, a menudo llamado un polo simple.

Si, por otra parte, la suma continua en  $n = -\infty$ , entonces  $z_0$  es un polo de orden infinito y es llamado una singularidad esencial. Estas singularidades esenciales tienen muchas características patológicas. Por ejemplo, podemos mostrar que en cualquier pequeña vecindad en torno a una singularidad esencial de  $f(z)$  la función  $f(z)$  llega a ser arbitrariamente cercana a cualquier cantidad compleja seleccionada  $w_0$ . Literalmente, el plano  $w$  complejo es mapeado en el entorno del punto  $z_0$ . Un punto de diferencia fundamental entre un polo de orden finito y una singularidad esencial es que un polo de orden  $m$  puede ser removido multiplicando  $f(z)$  por  $(z - z_0)^m$ . Esto obviamente no puede ser hecho para una singularidad esencial.

---

<sup>1</sup>Este capítulo está basado en el séptimo capítulo del libro: *Mathematical Methods for Physicists, fourth edition* de George B. Arfken & Hans J. Weber, editorial ACADEMIC PRESS.

El comportamiento de  $f(z)$  cuando  $z \rightarrow \infty$  está definido en términos de  $f(1/t)$  cuando  $t \rightarrow 0$ . Consideremos la función

$$\operatorname{sen}(z) = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} . \quad (6.2)$$

Cuando  $z \rightarrow \infty$ , reemplazamos la  $z$  por  $1/t$  para obtener

$$\operatorname{sen}\left(\frac{1}{t}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)! t^{2n+1}} . \quad (6.3)$$

Claramente, a partir de la definición,  $\operatorname{sen}(z)$  tiene una singularidad esencial en infinito. Este resultado podría anticiparse ya que

$$\begin{aligned} \operatorname{sen}(z) &= \operatorname{sen} iy , & \text{cuando } x = 0 \\ &= i \operatorname{senh} y , \end{aligned}$$

el cual tiende a infinito en forma exponencial cuando  $y \rightarrow \infty$ . Por lo tanto, aunque el valor absoluto de  $\operatorname{sen} x$  para  $x$  real es igual a o menor que uno, el valor absoluto de  $\operatorname{sen} z$  no está acotado.

### 6.1.2 Puntos de ramificación.

Hay otra tipo de singularidad que será importante en las últimas secciones de este capítulo. Consideremos

$$f(z) = z^a ,$$

en el cual  $a$  no es un entero<sup>2</sup>. Cuando  $z$  se mueve alrededor del círculo unitario desde  $e^0$  a  $e^{2\pi i}$ ,

$$f(z) \rightarrow e^{2\pi i} \neq e^{0i} ,$$

para  $a$  no entero. Como en el capítulo anterior, tenemos un punto ramificado en el origen y otro en infinito. Los puntos  $e^{0i}$  y  $e^{2\pi i}$  en el plano  $z$  coinciden pero esos puntos coincidentes producen valores diferentes de  $f(z)$ ; esto es,  $f(z)$  es una función multivaluada. El problema es resuelto construyendo una línea de corte uniendo ambos puntos de ramificación tal que  $f(z)$  este únicamente especificada para un punto dado en el plano  $z$ .

Note cuidadosamente que una función con un punto de ramificación y una requerida línea de corte no será continua a través de la línea de corte. En general, será una fase diferente sobre el lado opuesto de esa línea de corte. De modo que las integrales de líneas sobre los lados opuestos de estos puntos ramificados de la línea de corte generalmente no se cancelarán unas con otras.

---

<sup>2</sup> $z = 0$  es técnicamente un punto singular, para  $z^a$  tiene un número finito de derivadas, mientras una función analítica tiene garantizada un número infinito de derivadas. El problema es que  $f(z)$  no es una función monovaluada cuando tomamos un círculo en torno al origen. La fórmula integral de Cauchy no puede ser aplicada.

La línea de contorno usada para convertir una región múltiplemente conexa en una región simplemente conexa es completamente diferente. Nuestra función es continua a través de la línea de contorno, y no existen fases diferentes.

### Ejemplo

Considere la función

$$f(z) = (z^2 - 1)^{1/2} = (z + 1)^{1/2}(z - 1)^{1/2}. \quad (6.4)$$

El primer factor sobre el lado derecho,  $(z - 1)^{1/2}$ , tiene un punto de ramificación en  $z = +1$ . El segundo factor tiene un punto de ramificación  $z = -1$ . En infinito  $f(z)$  tiene un polo simple. La línea de corte conecta ambos puntos ramificados. Consideremos la posibilidad de tomar el segmento de línea que une  $z = +1$  y  $z = -1$  como una línea de corte, sigamos las fases de estos dos factores cuando nos movemos a lo largo del contorno mostrado en la figura 6.1.

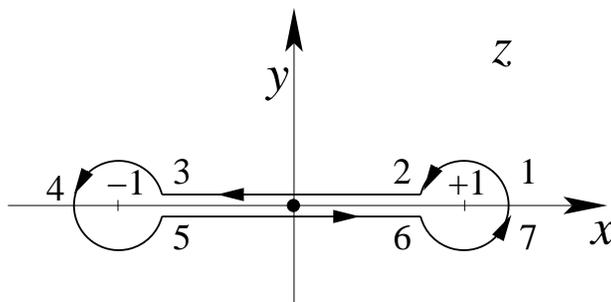


Figura 6.1: Contorno que rodea a un par de puntos de ramificación.

Por conveniencia en los siguientes cambios de fase sea  $z + 1 = re^{i\theta}$  y  $z - 1 = \rho e^{i\varphi}$ . Entonces la fase de  $f(z)$  es  $(\theta + \varphi)/2$ . Comenzamos en el punto 1 donde ambos  $z + 1$  y  $z - 1$  tienen una fase de cero. Moviéndose desde el punto 1 al punto 2,  $\varphi$ , la fase de  $z - 1 = \rho e^{i\varphi}$  aumenta por  $\pi$ . ( $z - 1$  llega a ser negativo).  $\varphi$  luego permanece constante hasta que el círculo es completado, moviéndose desde 6 a 7.  $\theta$ , la fase de  $z + 1 = re^{i\theta}$  muestra un comportamiento similar aumentando por  $2\pi$  cuando nos movemos desde 3 a 5. La fase de la función  $f(z)$  es  $(\theta + \varphi)/2$ . Esto es tabulado en la columna final de la tabla 6.1.

Dos características emergen;

1. La fase en los puntos 5 y 6 no es la misma que la de las fases de los puntos 2 y 3. Este comportamiento puede ser esperado en una línea de corte de punto ramificado.
2. La fase en el punto 7 excede a la del punto 1 por  $2\pi$  y la función  $f(z) = (z^2 - 1)^{1/2}$  es por lo tanto univaluada para el contorno mostrado, circunscribiendo ambos puntos ramificados.

Si tomamos el eje  $x - 1 \leq x \leq 1$  como una línea de corte,  $f(z)$  está únicamente especificada. Alternativamente, el eje  $x$  positivo para  $x > 1$  y el eje  $x$  negativo para  $x < -1$  puede ser tomado como línea de corte. Los puntos ramificados no pueden ser encerrados y la función permanece univaluada.

Punto	$\theta$	$\varphi$	$(\theta + \varphi)/2$
1	0	0	0
2	0	$\pi$	$\pi/2$
3	0	$\pi$	$\pi/2$
4	$\pi$	$\pi$	$\pi$
5	$2\pi$	$\pi$	$3\pi/2$
6	$2\pi$	$\pi$	$3\pi/2$
7	$2\pi$	$2\pi$	$2\pi$

Tabla 6.1:

Generalizando a partir de este ejemplo, tenemos que la fase de una función

$$f(z) = f_1(z) \cdot f_2(z) \cdot f_3(z) \dots$$

es la suma algebraica de la fase de sus factores individuales:

$$\arg[f(z)] = \arg[f_1(z)] + \arg[f_2(z)] + \arg[f_3(z)] + \dots$$

La fase de un factor individual puede ser tomada como el arcotangente de la razón de su parte imaginaria y su parte real,

$$\arg[f_i(z)] = \tan^{-1} \left( \frac{v_i}{u_i} \right).$$

Para el caso de un factor de la forma

$$f_i(z) = (z - z_0)$$

la fase corresponde a la fase angular de un vector bidimensional desde  $+z_0$  a  $z$ , la fase aumenta en  $2\pi$  cuando el punto  $+z_0$  es encerrado. Inversamente, la transversal de cualquier loop cerrado que no encierre a  $z_0$  no cambia la fase de  $z - z_0$ .

## 6.2 Cálculo del residuo.

### 6.2.1 Teorema del residuo.

Si la expansión de Laurent de una función  $f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$  es integrada término a término usando un contorno cerrado que circunscribe un punto singular  $z_0$  aislado en un sentido antihorario, obtenemos

$$\begin{aligned} a_n \oint (z - z_0)^n dz &= a_n \left. \frac{(z - z_0)^{n+1}}{n+1} \right|_{z_1}^{z_2} \\ &= 0 \quad \text{para todo } n \neq -1. \end{aligned} \tag{6.5}$$

Sin embargo, si  $n = -1$ ,

$$a_{-1} \oint (z - z_0)^{-1} dz = a_{-1} \oint \frac{ire^{i\theta} d\theta}{re^{i\theta}} = 2\pi ia_{-1} . \quad (6.6)$$

Resumiendo las ecuaciones (6.5) y (6.6), tenemos

$$\frac{1}{2\pi i} \oint f(z) dz = a_{-1} . \quad (6.7)$$

La constante  $a_{-1}$ , el coeficiente de  $(z - z_0)^{-1}$  en la expansión de Laurent, es llamada el *residuo* de  $f(z)$  en  $z = z_0$ .

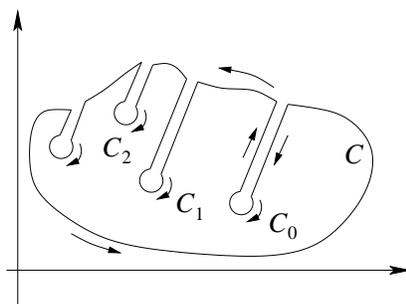


Figura 6.2: Excluyendo singularidades aisladas.

Un conjunto de singularidades aisladas pueden ser manejada adecuadamente deformando nuestro contorno como es mostrado en la figura 6.2. El teorema integral de Cauchy conduce a

$$\oint_C f(z) dz + \oint_{C_0} f(z) dz + \oint_{C_1} f(z) dz + \oint_{C_2} f(z) dz + \dots = 0 . \quad (6.8)$$

La integral cerrada en torno a cualquier punto singular es dada por la ecuación (6.7).

$$\oint_{C_i} f(z) dz = -2\pi ia_{-1, z_i} \quad (6.9)$$

suponiendo una expansión de Laurent en torno del punto singular,  $z = z_i$ . El signo negativo viene de la integración en sentido horario como muestra la figura 6.2. Combinando las ecuaciones (6.8) y (6.9), tenemos

$$\begin{aligned} \oint_C f(z) dz &= 2\pi i(a_{-1, z_0} + a_{-1, z_1} + a_{-1, z_2} + \dots) \\ &= 2\pi i(\text{la suma de los residuos encerrados}) . \end{aligned} \quad (6.10)$$

Este es el *teorema del residuo*. El problema de evaluar una o más integrales de contorno es reemplazado por el problema algebraico del cálculos de residuo en los puntos singulares encerrados.

El primer uso que le daremos al teorema del residuo es desarrollar el concepto del valor principal de Cauchy. Entonces en el remanente de esta sección aplicaremos el teorema del residuo a una amplia variedad de integrales definidas de interés matemático y físico. En la sección 6.3 el concepto del valor principal de Cauchy es usado para obtener las importantes relaciones de dispersión. El teorema del residuo también será necesario más adelante para una variedad de transformadas integrales, particularmente la transformada inversa de Laplace. Usando la transformación  $z = 1/w$  para  $w \rightarrow 0$ , podemos encontrar la naturaleza de la singularidad en  $z \rightarrow \infty$  y el residuo de una función  $f(z)$  con sólo singularidades aisladas y no en puntos ramificados. En tales casos conocemos que

$$\sum \text{residuos en el plano finito } z + \text{residuo en } z \rightarrow \infty = 0 .$$

### 6.2.2 Valor principal de Cauchy.

Ocasionalmente un polo de primer orden aislado estará directamente sobre el contorno de integración. En este caso podemos deformar el contorno para incluir o excluir el residuo deseado incluyendo un desvío semicircular de radio infinitesimal. Esto es mostrado en la figura 6.3. La integración sobre el semicírculo luego da con  $z - x_0 = \delta e^{i\varphi}$ ,  $dz = \delta e^{i\varphi} d\varphi$ ,

$$\int \frac{dz}{z - x_0} = i \int_{\pi}^{2\pi} d\varphi = i\pi , \quad \text{i.e., } \pi i a_{-1} \quad \text{sentido antihorario,}$$

$$\int \frac{dz}{z - x_0} = i \int_{\pi}^0 d\varphi = -i\pi , \quad \text{i.e., } -\pi i a_{-1} \quad \text{sentido horario,}$$

esta contribución, + o -, aparece sobre el lado izquierdo de la ecuación (6.10). Si nuestro desvío es en sentido horario, el residuo no puede ser encerrado y allí no estaría el correspondiente término sobre el lado derecho de la ecuación (6.10). Sin embargo, si nuestro desvío es en sentido antihorario, este residuo podría estar encerrado por el contorno  $C$  y un término  $2\pi i a_{-1}$  aparecería sobre el lado derecho de la ecuación (6.10). El resultado neto del desvío ya sea para el horario o antihorario es que un polo simple sobre el contorno es contabilizado como un medio de lo que podría ser si estuviera dentro del contorno. Esto corresponde a tomar el valor principal de Cauchy.

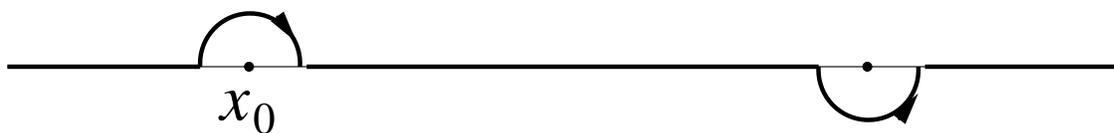


Figura 6.3: Bypass de puntos singulares.

Por ejemplo, supongamos que  $f(z)$  con un polo simple en  $z = x_0$  es integrado sobre el eje entero real. El contorno está encerrado con un semicírculo infinito en la mitad del plano superior (6.4). Entonces

$$\oint f(z) dz = \int_{-\infty}^{x_0-\delta} f(x) dx + \int_{C_{x_0}} f(z) dz + \int_{x_0+\delta}^{\infty} f(x) dx + \int_C \text{semicírculo infinito} \quad (6.11)$$

$$= 2\pi i \sum \text{residuos encerrado} .$$

Si el semicírculo pequeño  $C_{x_0}$  incluye  $x_0$  (moviéndose a lo largo del eje  $x$ , en sentido antihorario),  $x_0$  está encerrado, y su contribución aparece dos veces como  $\pi i a_{-1}$  en  $\int_{C_{x_0}}$  y como  $2\pi i a_{-1}$  en el término  $2\pi i \sum$  residuo encerrado para una contribución neta de  $\pi i a_{-1}$ . Si el semicírculo pequeño superior es elegido,  $x_0$  es excluido. La única contribución es de la integración sobre  $C_{x_0}$  en el sentido antihorario la cual tiende a  $-\pi i a_{-1}$ . Moviendo esto al extremo derecho de la ecuación (6.11), tenemos  $+\pi i a_{-1}$ , como antes.

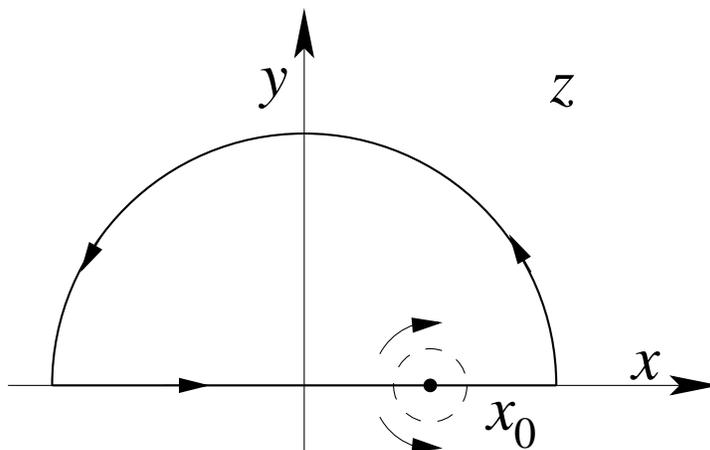


Figura 6.4: Cerrando el contorno con un semicírculo de radio infinito.

Las integrales a lo largo del eje  $x$  pueden ser combinadas y al radio del semicírculo lo hacemos tender a cero. Por lo tanto definimos

$$\lim_{\delta \rightarrow 0} \left\{ \int_{-\infty}^{x_0 - \delta} f(x) dx + \int_{x_0 + \delta}^{\infty} f(x) dx \right\} = \mathcal{P} \int_{-\infty}^{\infty} f(x) dx . \quad (6.12)$$

La  $\mathcal{P}$  indica el valor principal de Cauchy y representa el proceso límite precedente. Note cuidadosamente que el valor principal de Cauchy es un proceso de cancelación o de balance. En la vecindad de nuestra singularidad en  $z = x_0$ ,

$$f(x) \approx \frac{a_{-1}}{x - x_0} . \quad (6.13)$$

Esto es impar, respecto a  $x_0$ . El intervalo simétrico o par (respecto a  $x_0$ ) da la cancelación del área achurada, figura 6.5. La contribución de la singularidad está en la integración alrededor del semicírculo. Algunas veces, esta misma técnica límite es aplicada para los límites de integración  $\pm\infty$ . Podemos definir

$$\mathcal{P} \int_{-\infty}^{\infty} f(x) dx = \lim_{a \rightarrow \infty} \int_{-a}^a f(x) dx . \quad (6.14)$$

Un tratamiento alternativo mueve el polo fuera del contorno y luego considera el comportamiento límite cuando es traído de vuelta.

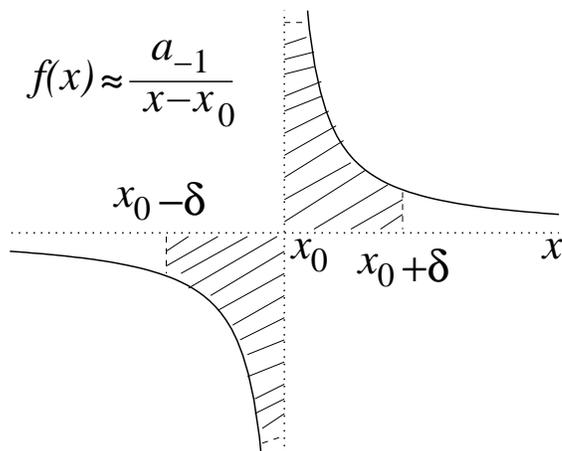


Figura 6.5: Valor principal de Cauchy.

### 6.2.3 Expansión en polos de funciones meromórficas.

Las funciones analíticas  $f(z)$  que tienen solamente polos bien separados como singularidades son llamadas meromórficas. Por simplicidad suponemos esos polos en  $z = a_n$  finito con  $0 < |a_1| < |a_2| < \dots$  son todos simples con residuos  $b_n$ . Entonces una expansión de  $f(z)$  en términos de  $b_n(z - a_n)^{-1}$  depende solamente de las propiedades intrínsecas de  $f(z)$ , en contraste a la expansión de Taylor cerca de un punto analítico arbitrario  $z_0$  de  $f(z)$  o de la expansión de Laurent cerca de algún punto singular de  $f(z)$ .

Consideremos una serie de círculos concéntricos  $C_n$  alrededor del origen tal que  $C_n$  incluya  $a_1, a_2, \dots, a_n$  pero no otros polos, su radio  $R_n \rightarrow \infty$  cuando  $n \rightarrow \infty$ . Para garantizar la convergencia suponemos que  $|f(z)| < \varepsilon R_n$  para cualquier constante positiva pequeña  $\varepsilon$  y todo  $z$  sobre  $C_n$ . Entonces la serie

$$f(z) = f(0) + \sum_{n=1}^{\infty} b_n \left[ \frac{1}{z - a_n} + \frac{1}{a_n} \right], \quad (6.15)$$

converge a  $f(z)$ . Para probar este teorema (debido a Mittag-Leffler) usamos el teorema del residuo para evaluar la integral de contorno para  $z$  dentro de  $C_n$ :

$$\begin{aligned} I_n &= \frac{1}{2\pi i} \int_{C_n} \frac{f(w)}{w(w-z)} dw \\ &= \sum_{m=1}^n \frac{b_m}{a_m(a_m - z)} + \frac{f(z) - f(0)}{z}. \end{aligned} \quad (6.16)$$

En  $C_n$  tenemos para  $n \rightarrow \infty$

$$|I_n| \leq 2\pi R_n \max_{w \text{ sobre } C_n} \frac{|f(w)|}{2\pi R_n(R_n - |z|)} < \frac{\varepsilon R_n}{R_n - |z|} \leq \varepsilon.$$

Usando  $I_n \rightarrow 0$  en la ecuación (6.16) probamos la ecuación (6.15).

Si  $|f(z)| < \varepsilon R_n^{p+1}$ , entonces evaluamos la integral de forma similar

$$I_n = \frac{1}{2\pi i} \int \frac{f(w)}{w^{p+1}(w-z)} dw \rightarrow 0 \quad \text{cuando } n \rightarrow \infty,$$

y obtenemos la expansión en polo análoga

$$f(z) = f(0) + zf'(0) + \dots + \frac{z^p f^{(p)}(0)}{p!} + \sum_{n=1}^{\infty} \frac{b_n z^{p+1}/a_n^{p+1}}{z - a_n}. \quad (6.17)$$

#### 6.2.4 Expansión en producto de funciones enteras.

Una función  $f(z)$  que es analítica para todo  $z$  finito es llamada una función *entera*. La derivada logarítmica  $f'/f$  es una función meromórfica con una expansión en polos.

Si  $f(z)$  tiene un cero simple en  $z = a_n$ , entonces  $f(z) = (z - a_n)g(z)$  con  $g(z)$  analítica y  $g(a_n) \neq 0$ . de modo que la derivada logarítmica

$$\frac{f'(z)}{f(z)} = \frac{1}{z - a_n} + \frac{g'(z)}{g(z)}, \quad (6.18)$$

tiene un polo simple en  $z = a_n$  con residuo 1, y  $g'/g$  es analítica allí. Si  $f'/f$  satisface las condiciones que conducen a la expansión en polos de la ecuación (6.15), entonces

$$\frac{f'(z)}{f(z)} = \frac{f'(0)}{f(0)} + \sum_{n=1}^{\infty} \left[ \frac{1}{a_n} + \frac{1}{z - a_n} \right], \quad (6.19)$$

se mantiene. Integrando la ecuación anterior obtenemos

$$\begin{aligned} \int_0^z \frac{f'(z)}{f(z)} dz &= \ln f(z) - \ln f(0) \\ &= \frac{zf'(0)}{f(0)} + \sum_{n=1}^{\infty} \left\{ \ln(z - a_n) - \ln(-a_n) + \frac{z}{a_n} \right\}, \end{aligned}$$

y expandiendo obtenemos el expansión en producto

$$f(z) = f(0) \exp \frac{zf'(0)}{f(0)} \prod_{n=1}^{\infty} \left( 1 - \frac{z}{a_n} \right) \exp \left( \frac{z}{a_n} \right). \quad (6.20)$$

Ejemplos son las expansiones en productos

$$\begin{aligned} \operatorname{sen}(z) &= z \prod_{n=-\infty}^{\infty} \left( 1 - \frac{z}{n\pi} \right) e^{z/n\pi} = z \prod_{n=1}^{\infty} \left( 1 - \frac{z^2}{n^2\pi^2} \right), \\ \cos(z) &= \prod_{n=1}^{\infty} \left( 1 - \frac{z^2}{(n-1/2)^2\pi^2} \right). \end{aligned} \quad (6.21)$$

Otro ejemplo es la expansión en producto de la función gamma la cual será discutida más adelante.

Como una consecuencia de la ecuación (6.18) la integral de contorno de la derivada logarítmica puede ser usado para contar el número  $N_f$  de ceros (incluyendo sus multiplicidades) de la función  $f(z)$  del contorno  $C$ :

$$\frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)} dz = N_f . \quad (6.22)$$

Por otra parte, usando

$$\int \frac{f'(z)}{f(z)} dz = \ln f(z) = \ln |f(z)| + i \arg[f(z)] , \quad (6.23)$$

vemos que la parte real en la ecuación (6.23) no cambia cuando  $z$  se mueve una vez al rededor del contorno, mientras el correspondiente cambio en  $\arg[f]$  debe ser

$$\Delta_C \arg[f] = 2\pi N_f . \quad (6.24)$$

Esto conduce al *teorema de Rouché*: Si  $f(z)$  y  $g(z)$  son analíticas dentro y sobre un contorno encerrado  $C$ , y  $|g(z)| < |f(z)|$  sobre  $C$ , entonces  $f(z)$  y  $f(z) + g(z)$  tienen el mismo número de ceros dentro de  $C$ .

Para mostrar esto usamos

$$2\pi N_{f+g} = \Delta_C \arg[f + g] = \Delta_c \arg[f] + \Delta_c \arg \left[ 1 + \frac{g}{f} \right] .$$

ya que  $|g| < |f|$  en  $C$ , el punto  $w = 1 + g(z)/f(z)$  es siempre un punto interior del círculo en el plano  $w$  con centro en 1 y radio 1. De modo que  $\arg[1 + g/f]$  debería volver a su valor original cuando  $z$  se mueve al rededor de  $C$ ; no puede disminuir o aumentar por un múltiplo de  $2\pi$  tal que  $\Delta_c \arg[1 + g/f] = 0$ .

El teorema de Rouché puede ser usado como una prueba alternativa del teorema fundamental del álgebra: Un polinomio  $\sum_{m=0}^n a_m z^m$  con  $a_n \neq 0$  tiene  $n$  ceros. Definamos  $f(z) = a_n z^n$ . Entonces  $f$  tiene un cero de orden  $n$  en el origen y no otros ceros. Sea  $g(z) = \sum_{m=0}^{n-1} a_m z^m$ . Apliquemos el teorema de Rouché a un círculo  $C$  con centro en el origen y radio  $R > 1$ . En  $C$ ,  $|f(z)| = |a_n| R^n$  y

$$|g(z)| \leq |a_0| + |a_1| R + \cdots + |a_{n-1}| R^{n-1} \leq \left( \sum_{m=0}^{n-1} |a_m| \right) R^{n-1} .$$

Por lo tanto  $|g(z)| < |f(z)|$  para  $z$  en  $C$  dado  $R > (\sum_{m=0}^{n-1} |a_m|)/|a_n|$ . Para todo círculo  $C$  suficientemente grande, por lo tanto,  $f + g = \sum_{m=0}^n a_m z^m$  tiene  $n$  ceros dentro de  $C$  de acuerdo al teorema de Rouché.

### 6.2.5 Evaluación de integrales definidas.

Las integrales definidas aparecen repetidamente en problemas de física matemática tanto como en matemática pura. Tres técnicas moderadamente generales son útiles para evaluar estas integrales definidas: (1) integrales de contorno, (2) conversión a funciones beta o gamma y (3) cuadratura numérica. Otros acercamientos incluyen expansión en serie con integración término a término y transformadas integrales. Como veremos, el método de integración de contorno es quizás el más versátil de esos métodos, ya que es aplicable a una amplia variedad de integrales.

### 6.2.6 Evaluación de integrales definidas: $\int_0^{2\pi} f(\sin \theta, \cos \theta) d\theta$

El cálculo del residuo es útil evaluando una amplia variedad de integrales finitas tanto en problemas de la física como en matemática pura. Consideremos, primero, integrales de la forma

$$I = \int_0^{2\pi} f(\sin \theta, \cos \theta) d\theta, \quad (6.25)$$

donde  $f$  es finita para todos los valores de  $\theta$ . También requerimos que  $f$  sea una función racional de  $\sin \theta$  y  $\cos \theta$  que será univaluada. Sea

$$z = e^{i\theta}, \quad dz = ie^{i\theta} d\theta.$$

De esto,

$$d\theta = -i \frac{dz}{z}, \quad \sin(\theta) = \frac{z - z^{-1}}{2i}, \quad \cos(\theta) = \frac{z + z^{-1}}{2}. \quad (6.26)$$

Nuestra integral se convierte en

$$I = -i \oint f\left(\frac{z - z^{-1}}{2i}, \frac{z + z^{-1}}{2}\right) \frac{dz}{z}, \quad (6.27)$$

con el camino de integración sobre un círculo unitario. Por el teorema del residuo la ecuación (6.20),

$$I = (-i)2\pi i \sum \text{residuo dentro del círculo unitario.} \quad (6.28)$$

Note que estamos después del residuo de  $f(z)/z$ .

#### Ejemplo

Nuestro problema es evaluar la integral definida

$$I = \int_0^{2\pi} \frac{d\theta}{1 + \varepsilon \cos \theta}, \quad |\varepsilon| < 1.$$

Por la ecuación (6.27) ésta se convierte en

$$\begin{aligned} I &= -i \oint_{\text{círculo unitario}} \frac{dz}{z[1 + (\varepsilon/2)(z + z^{-1})]} \\ &= -i \frac{2}{\varepsilon} \oint \frac{dz}{z^2 + (2/\varepsilon)z + 1}. \end{aligned}$$

El denominador tiene raíces

$$z_- = -\frac{1}{\varepsilon} - \frac{1}{\varepsilon} \sqrt{1 - \varepsilon^2} \quad \text{y} \quad z_+ = -\frac{1}{\varepsilon} + \frac{1}{\varepsilon} \sqrt{1 - \varepsilon^2}$$

$z_+$  está dentro del círculo unitario;  $z_-$  está afuera. Luego por la ecuación (6.28)

$$I = -i \frac{2}{\varepsilon} 2\pi i \frac{1}{z + 1/\varepsilon + (1/\varepsilon)\sqrt{1 - \varepsilon^2}} \Big|_{z=-1/\varepsilon+(1/\varepsilon)\sqrt{1-\varepsilon^2}}.$$

Obtenemos

$$\int_0^{2\pi} \frac{d\theta}{1 + \varepsilon \cos(\theta)} = \frac{2\pi}{\sqrt{1 - \varepsilon^2}}, \quad |\varepsilon| < 1.$$

### 6.2.7 Evaluaciones de integrales definidas: $\int_{-\infty}^{\infty} f(x)dx$ .

Supongamos que nuestra integral definida tiene la forma

$$I = \int_{-\infty}^{\infty} f(x) dx, \quad (6.29)$$

y satisface las dos condiciones:

- $f(z)$  es analítica en el semiplano superior excepto para un número finito de polos. (Suponemos que aquí no hay polos en el eje real. Si se presentan polos en el eje real, ellos podrían ser incluidos o excluidos como se discutió anteriormente).
- $f(z)$  se anula tan fuertemente<sup>3</sup> como  $1/z^2$  para  $|z| \rightarrow \infty$ ,  $0 \leq \arg[z] \leq \pi$ .

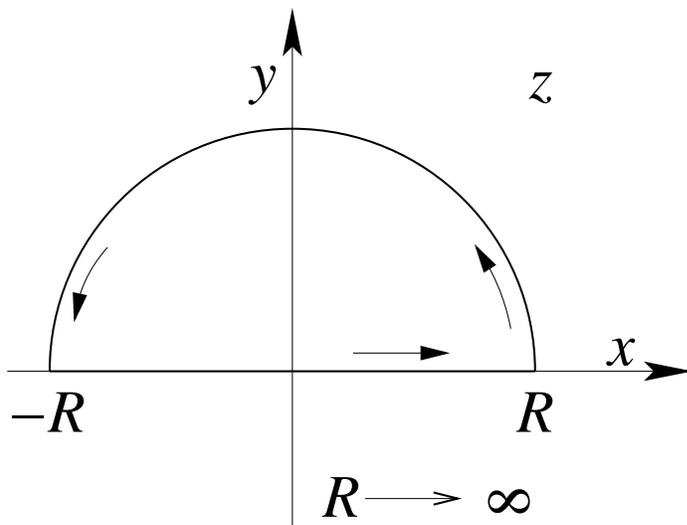


Figura 6.6: Contorno de integración.

Con estas condiciones, podemos tomar como un contorno de integración el eje real y un semicírculo en el plano medio superior como muestra la figura 6.6. Tomemos el radio  $R$  del semicírculo infinitamente grande. Entonces

$$\begin{aligned} \oint f(z) dz &= \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx + \lim_{R \rightarrow \infty} \int_0^{\pi} f(Re^{i\theta}) iRe^{i\theta} d\theta \\ &= 2\pi i \sum \text{residuos (en el semiplano superior)}. \end{aligned} \quad (6.30)$$

De la segunda condición la segunda integral (sobre el semicírculo) se anula y

$$\int_{-\infty}^{\infty} f(x) dx = 2\pi i \sum \text{residuos (en el semiplano superior)}. \quad (6.31)$$

<sup>3</sup>Podríamos usar que  $f(z)$  se anula más rápido que  $1/z$ , pero queremos tener  $f(z)$  monovaluada.

**Ejemplo**

Evalúe

$$I = \int_{-\infty}^{\infty} \frac{dx}{1+x^2}. \quad (6.32)$$

A partir de la ecuación (6.31)

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = 2\pi i \sum \text{residuos (en el semiplano superior)}.$$

Aquí y en cada problema similar tenemos una pregunta: ¿dónde están los polos? Reescribiendo el integrando como

$$\frac{1}{1+z^2} = \frac{1}{z+i} \cdot \frac{1}{z-i}, \quad (6.33)$$

vemos que hay polos simples (orden 1) en  $z = i$  y  $z = -i$ . Un polo simple en  $z = z_0$  indica (y está indicado por) una expansión de Laurent de la forma

$$f(z) = \frac{a_{-1}}{z-z_0} + a_0 + \sum_{n=1}^{\infty} a_n(z-z_0)^n. \quad (6.34)$$

El residuo  $a_{-1}$  es fácilmente aislado como

$$a_{-1} = (z-z_0)f(z)|_{z=z_0}. \quad (6.35)$$

Usando la ecuación (6.35), encontramos que el residuo en  $z = i$  es  $1/(2i)$ , mientras que en  $z = -i$  es  $-1/(2i)$ . Entonces

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = 2\pi i \frac{1}{2i} = \pi. \quad (6.36)$$

Aquí hemos usado  $a_{-1} = 1/(2i)$  para el residuo del único polo incluido en  $z = i$ . Es fácil probar que es posible usar el semicírculo inferior y que esta opción conducirá al mismo resultado,  $I = \pi$ .

**6.2.8 Evaluación de integrales definidas:  $\int_{-\infty}^{\infty} f(x)e^{iax} dx$ .**

Consideremos la integral definida

$$I = \int_{-\infty}^{\infty} f(x)e^{iax} dx, \quad (6.37)$$

con  $a$  real y positivo. Esto es una transformada de Fourier. Suponemos las dos condiciones:

a.  $f(z)$  es analítica en el semiplano superior excepto para un número finito de polos.

b.

$$\lim_{|z| \rightarrow \infty} f(z) = 0, \quad 0 \leq \arg[z] \leq \pi. \quad (6.38)$$

Note que esta es una condición menos restrictiva que la segunda condición impuesta sobre  $f(z)$  para la integral previa  $\int_{-\infty}^{\infty} f(x)dx$ .

Empleemos el contorno mostrado en la figura 6.6. La aplicación del cálculo del residuo es la misma cuando sólo se considera uno, pero aquí tenemos que trabajar un poco más duro para mostrar que la integral sobre el semicírculo (infinito) va a cero. Esta integral se convierte en

$$I_R = \int_0^{\pi} f(Re^{i\theta}) e^{iaR \cos \theta - aR \sin \theta} i R e^{i\theta} d\theta. \quad (6.39)$$

Sea  $R$  tan grande que  $|f(z)| = |f(Re^{i\theta})| < \varepsilon$ . Entonces

$$\begin{aligned} |I_R| &\leq \varepsilon R \int_0^{\pi} e^{-aR \sin \theta} d\theta \\ &= 2\varepsilon R \int_0^{\pi/2} e^{-aR \sin \theta} d\theta. \end{aligned} \quad (6.40)$$

En el rango  $[0, \pi/2]$

$$\frac{2}{\pi} \theta \leq \sin \theta.$$

Por lo tanto (figura 6.7)

$$|I_R| \leq 2\varepsilon R \int_0^{\pi/2} e^{-aR 2\theta/\pi} d\theta. \quad (6.41)$$

Ahora, integrando por inspección, obtenemos

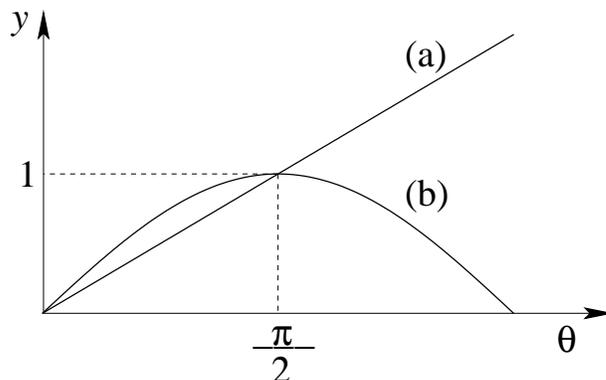


Figura 6.7: (a)  $y = (2/\pi)\theta$ , (b)  $y = \sin \theta$ .

$$|I_R| \leq 2\varepsilon R \frac{1 - e^{-aR}}{aR2/\pi}.$$

Finalmente,

$$\lim_{|I_R|} \leq \frac{\pi}{a} \varepsilon. \quad (6.42)$$

De la ecuación (6.38),  $\varepsilon \rightarrow 0$  cuando  $R \rightarrow \infty$  y

$$\lim_{R \rightarrow \infty} |I_R| = 0. \quad (6.43)$$

Este útil resultado es a veces llamado *lema de Jordan*. Con esto, hemos preparado para atacar las integrales de Fourier de la forma mostrada en la ecuación (6.37).

Usando el contorno mostrado en la figura 6.6, tenemos

$$\int_{-\infty}^{\infty} f(x)e^{iax} dx + \lim_{R \rightarrow \infty} I_R = 2\pi i \sum \text{residuos (en el semiplano superior)}.$$

Ya que la integral sobre el semicírculo superior  $I_R$  se anula cuando  $R \rightarrow \infty$  (lema de Jordan),

$$\int_{-\infty}^{\infty} f(x)e^{iax} dx = 2\pi i \sum \text{residuos (en el semiplano superior)} \quad (a > 0). \quad (6.44)$$

**Ejemplo** Singularidad sobre el contorno de integración.

El problema es evaluar

$$I = \int_0^{\infty} \frac{\text{sen}(x)}{x} dx. \quad (6.45)$$

esto puede ser tomado como la parte imaginaria de<sup>4</sup>

$$I_z = \oint \frac{e^{iz}}{z} dz. \quad (6.46)$$

Ahora el único polo es un polo simple en  $z = 0$  y el residuo dado por la ecuación (6.35) es  $a_{-1} = 1$ . Escogemos el contorno dado por la figura 6.8 (1) para evitar el polo, (2) para incluir el eje real y (3) para procurar que tienda a cero el integrando para  $z = iy$  con  $y \rightarrow \infty$ . Note que en este caso un semicírculo grande (infinito) en el semiplano inferior sería desastroso. Tenemos

$$\oint \frac{e^{iz}}{z} dz = \int_{-R}^{-r} e^{ix} \frac{dx}{x} + \int_{C_1} \frac{e^{iz}}{z} dz + \int_r^R e^{ix} \frac{dx}{x} + \int_{C_2} \frac{e^{iz}}{z} dz = 0, \quad (6.47)$$

<sup>4</sup>Uno puede usar  $\int [(e^{iz} - e^{-iz})/2iz] dz$ , pero entonces dos diferentes contornos serán necesarios para las dos exponenciales.

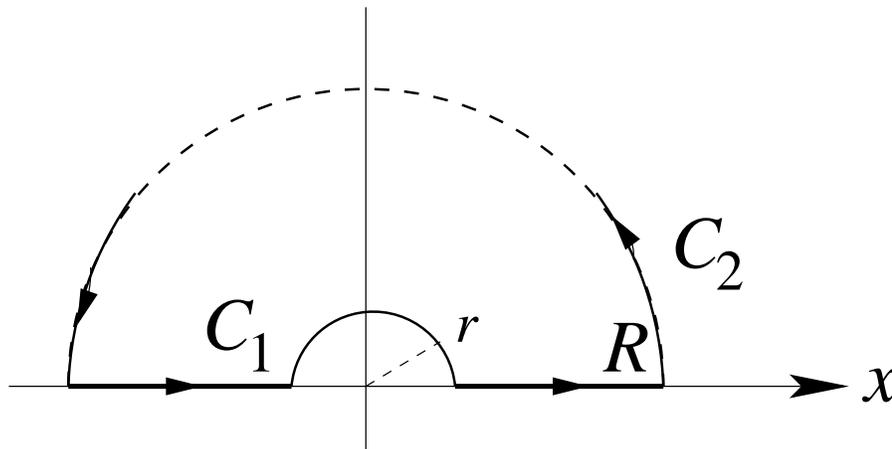


Figura 6.8: Contorno de integración.

el cero final viene del teorema del residuo (ecuación (6.20)). Por el lema de Jordan

$$\int_{C_2} \frac{e^{iz} dz}{z} = 0, \quad (6.48)$$

y

$$\oint \frac{e^{iz} dz}{z} = \int_{C_1} \frac{e^{iz} dz}{z} + \int_{-\infty}^{\infty} \frac{e^{ix} dx}{x} = 0. \quad (6.49)$$

La integral sobre el pequeño semicírculo produce  $(-)\pi i$  veces el residuo de 1, el signo menos es el resultado de ir en sentido horario. Tomando la parte imaginaria, tenemos

$$\int_{-\infty}^{\infty} \frac{\text{sen}(x)}{x} dx = \pi, \quad (6.50)$$

o

$$\int_0^{\infty} \frac{\text{sen}(x)}{x} dx = \frac{\pi}{2}. \quad (6.51)$$

El contorno de la figura 6.8, aunque conveniente, no es único. Encuentre alternativas de contorno para evaluar la ecuación (6.45) como ejercicio.

### Ejemplo *Scattering en mecánica cuántica.*

El análisis mecánico cuántico del *scattering* conduce a la función

$$I(\sigma) = \int_{-\infty}^{\infty} \frac{x \text{sen}(x) dx}{x^2 - \sigma^2}, \quad (6.52)$$

donde  $\sigma$  es real y positivo. A partir de las condiciones físicas del problema hay un requerimiento posterior:  $I(\sigma)$  está para tener la forma  $e^{i\sigma}$  tal que representará una onda saliente (*outgoing scattered*). Usando

$$\text{sen}(z) = \frac{1}{i} \text{senh}(iz) = \frac{1}{2i} e^{iz} - \frac{1}{2i} e^{-iz}, \quad (6.53)$$

escribimos la ecuación (6.52) en el plano complejo como

$$I(\sigma) = I_1 + I_2, \quad (6.54)$$

con

$$\begin{aligned} I_1 &= \frac{1}{2i} \int_{-\infty}^{\infty} \frac{ze^{iz}}{z^2 - \sigma^2} dz, \\ I_2 &= -\frac{1}{2i} \int_{-\infty}^{\infty} \frac{ze^{-iz}}{z^2 - \sigma^2} dz. \end{aligned} \quad (6.55)$$

La integral  $I_1$  es similar al ejemplo anterior y, como en este caso, podríamos completar el contorno por un semicírculo infinito en el semiplano superior como muestra la figura 6.9a. Para  $I_2$  la exponencial es negativa y completamos el contorno con un semicírculo infinito en el semiplano inferior, como muestra la figura 6.9b. Como en el ejemplo anterior, el semicírculo no contribuye a la integral, lema de Jordan.

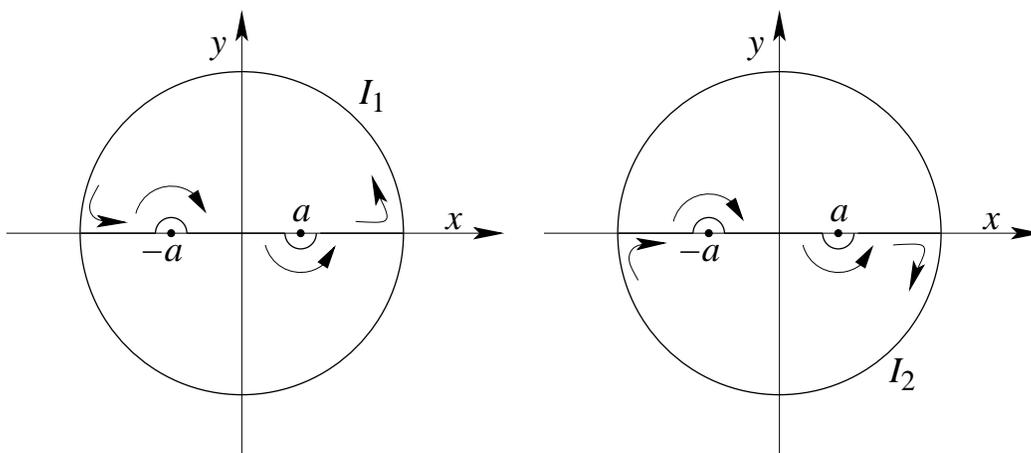


Figura 6.9: Contorno de integración.

Todavía está el problema de localizar los polos y evaluar los residuos. Encontramos polos en  $z = +\sigma$  y  $z = -\sigma$  sobre el contorno de integración. Los residuos son

	$z = \sigma$	$z = -\sigma$
$I_1$	$\frac{e^{i\sigma}}{2}$	$\frac{e^{-i\sigma}}{2}$
$I_2$	$\frac{e^{-i\sigma}}{2}$	$\frac{e^{i\sigma}}{2}$

Rodeando los polos, como se muestra en la figura 6.9 (esto se complica un poco si vamos de arriba a abajo), encontramos que el teorema del residuo conduce a

$$PI_1 - \pi i \left( \frac{1}{2i} \right) \frac{e^{-i\sigma}}{2} + \pi i \left( \frac{1}{2i} \right) \frac{e^{i\sigma}}{2} = 2\pi i \left( \frac{1}{2i} \right) \frac{e^{i\sigma}}{2}, \quad (6.56)$$

por tener encerrada la singularidad en  $z = \sigma$  pero excluida en  $z = -\sigma$ . En un aspecto similar, pero notando que el contorno para  $I_2$  es en sentido horario,

$$PI_2 - \pi i \left( \frac{-1}{2i} \right) \frac{e^{i\sigma}}{2} + \pi i \left( \frac{-1}{2i} \right) \frac{e^{-i\sigma}}{2} = -2\pi i \left( \frac{-1}{2i} \right) \frac{e^{i\sigma}}{2}, \quad (6.57)$$

Sumando las ecuaciones (6.56) y (6.57), tenemos

$$\begin{aligned} PI(\sigma) &= PI_1 + PI_2 = \frac{\pi}{2} (e^{i\sigma} + e^{-i\sigma}) = \pi \cosh i\sigma \\ &= \pi \cos \sigma. \end{aligned} \quad (6.58)$$

Esta es una buena evaluación de la ecuación (6.52), pero desafortunadamente la dependencia del coseno es apropiada para una onda estacionaria y no para una onda saliente como se especificó. Para obtener la forma deseada, tratamos una técnica diferente. En vez de rodear los puntos singulares, los movemos fuera del eje real. específicamente, sea  $\sigma \rightarrow \sigma + i\gamma$ ,  $-\sigma \rightarrow -\sigma - i\gamma$ , donde  $\gamma$  es positivo pero pequeño y eventualmente tenderá a cero, ésto es,

$$I_+(\sigma) = \lim_{\gamma \rightarrow 0} I(\sigma + i\gamma). \quad (6.59)$$

Con esta sustitución simple, la primera integral  $I_1$  se convertirá

$$I_1(\sigma + i\gamma) = 2\pi i \left( \frac{1}{2i} \right) \frac{e^{i(\sigma+i\gamma)}}{2}, \quad (6.60)$$

por directa aplicación del teorema del residuo. También,

$$I_2(\sigma + i\gamma) = -2\pi i \left( \frac{-1}{2i} \right) \frac{e^{i(\sigma+i\gamma)}}{2}. \quad (6.61)$$

Sumando las ecuaciones (6.60) y (6.61) y haciendo luego  $\gamma \rightarrow 0$ , obtenemos

$$\begin{aligned} I_+ &= \lim_{\gamma \rightarrow 0} [I_1(\sigma + i\gamma) + I_2(\sigma + i\gamma)] \\ &= \lim_{\gamma \rightarrow 0} \pi e^{i(\sigma+i\gamma)} = \pi e^{i\sigma}. \end{aligned} \quad (6.62)$$

un resultado que ajustan las condiciones de borde de nuestro problema de scattering.

Es interesante notar que la sustitución  $\sigma \rightarrow \sigma - i\gamma$  tendría que conducir a

$$I_-(\sigma) = \pi e^{-i\sigma}, \quad (6.63)$$

la cual representaría una onda entrante. Nuestros resultados anteriores (ecuación (6.58)) es visto como el promedio aritmético de las ecuaciones (6.62) y (6.63). Este promedio es el valor principal de Cauchy de la integral. Note que tenemos esas posibilidades (ecuaciones (6.58), (6.62) y (6.63)) por que nuestra integral no es únicamente definida hasta que especifiquemos el proceso límite particular (o promedio) a ser usado.

### 6.2.9 Evaluación de integrales definidas: formas exponenciales.

Con funciones exponenciales o hiperbólicas presentes en el integrando, la vida es más complicada que antes. En vez de una prescripción general completa, el contorno debe ser escogido para ajustar la integral específica. Estos casos son oportunidades para ilustrar la versatilidad y poder de la integración de contorno.

Como un ejemplo, consideremos una integral que será muy útil en el desarrollo de una relación entre  $z!$  y  $(-z)!$ . Note cómo la periodicidad a lo largo del eje imaginario es aprovechada.

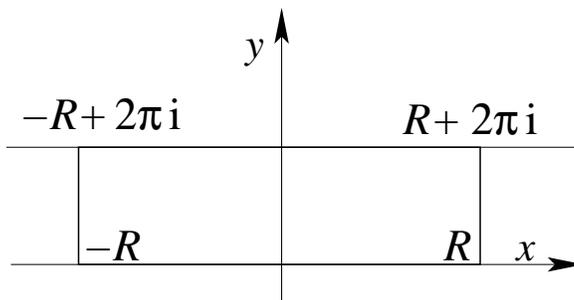


Figura 6.10: Contorno de integración.

#### Ejemplo Función factorial.

Deseamos evaluar

$$I = \int_{-\infty}^{\infty} \frac{e^{ax}}{1 + e^x} dx, \quad 0 < a < 1. \quad (6.64)$$

Los límites sobre  $a$  son necesarios (y suficientes) para prevenir la divergencia de la integral cuando  $x \rightarrow \pm\infty$ . Esta integral (ecuación (6.64)) puede ser manejada reemplazando la variable real  $x$  por la variable compleja  $z$  e integrando alrededor del contorno mostrado en la figura 6.10. Si tomamos el límite cuando  $R \rightarrow \infty$ , el eje real, por supuesto, tiende a la integral que deseamos. El camino de vuelta a lo largo de  $y = 2\pi$  es escogido para dejar el denominador de la integral invariante, al mismo tiempo introducir un factor constante  $e^{i2\pi a}$  en el numerador. Tenemos, en el plano complejo,

$$\begin{aligned} \oint \frac{e^{az}}{1 + e^z} dz &= \lim_{R \rightarrow \infty} \left( \int_{-R}^R \frac{e^{ax}}{1 + e^x} dx - e^{i2\pi a} \int_{-R}^R \frac{e^{ax}}{1 + e^x} dx \right) \\ &= (1 - e^{i2\pi a}) \int_{-\infty}^{\infty} \frac{e^{ax}}{1 + e^x} dx. \end{aligned} \quad (6.65)$$

En suma hay dos secciones verticales ( $0 \leq y \leq 2\pi$ ), la cual se anula (exponencialmente) cuando  $R \rightarrow \infty$ . Ahora ¿dónde están los polos y cuáles son los residuos? Tenemos un polo cuando

$$e^z = e^x e^{iy} = -1. \quad (6.66)$$

La ecuación (6.66) es satisfecha cuando  $z = 0 + i\pi$ . Por una expansión de Laurent<sup>5</sup> en potencias de  $(z - i\pi)$  el polo es un polo simple con un residuo de  $-e^{i\pi a}$ . Luego, aplicando el teorema del residuo,

$$(1 - e^{+i2\pi a}) \int_{-\infty}^{\infty} \frac{e^{ax}}{1 + e^x} dx = 2\pi i(-e^{i\pi a}) . \quad (6.67)$$

Esto rápidamente se reduce a

$$\int_{-\infty}^{\infty} \frac{e^{ax}}{1 + e^x} dx = \frac{\pi a}{\operatorname{sen} \pi a} , \quad 0 < a < 1 . \quad (6.68)$$

Usando la función beta, podemos mostrar que la integral es igual al producto  $(a - 1)!(-a)!$ . Esto resulta en la interesante y útil relación de la función factorial

$$a!(-a)! = \frac{\pi a}{\operatorname{sen} \pi a} . \quad (6.69)$$

Aunque la ecuación (6.68) se mantiene para  $a$  real,  $0 < a < 1$ , la ecuación (6.69) puede ser extendida por una continuación analítica de todos los valores de  $a$ , reales y complejos, excluyendo solamente los valores reales enteros.

Como un ejemplo final de integrales de contorno de funciones exponenciales, consideremos nuevamente los números de Bernoulli

### Ejemplo Números de Bernoulli.

En la sección 5.9 los números de Bernoulli fueron definidos por la expansión

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n . \quad (6.70)$$

Reemplazando  $x$  por  $z$  (continuación analítica), tenemos una serie de Taylor (compare con la ecuación (5.52)) con

$$B_n = \frac{n!}{2\pi i} \oint_{C_0} \frac{z}{e^z - 1} \frac{dz}{z^{n+1}} , \quad (6.71)$$

donde el contorno  $C_0$  está rodeando el origen en sentido antihorario con  $|z| < 2\pi$  para dar los polos en  $\pm 2\pi i$ .

Para  $n = 0$  tenemos un polo simple en  $z = 0$  con un residuo de  $+1$ . De modo que por la ecuación (6.20)

$$B_0 = \frac{0!}{2\pi i} \cdot 2\pi i(1) = 1 . \quad (6.72)$$

Para  $n = 1$  la singularidad en  $z = 0$  llega a ser un polo de segundo orden. El residuo puede mostrarse que es  $-1/2$  por la expansión en serie de la exponencial, seguida por la expansión binomial. Esto resulta en

$$B_1 = \frac{1!}{2\pi i} \cdot 2\pi i \left( -\frac{1}{2} \right) = -\frac{1}{2} . \quad (6.73)$$

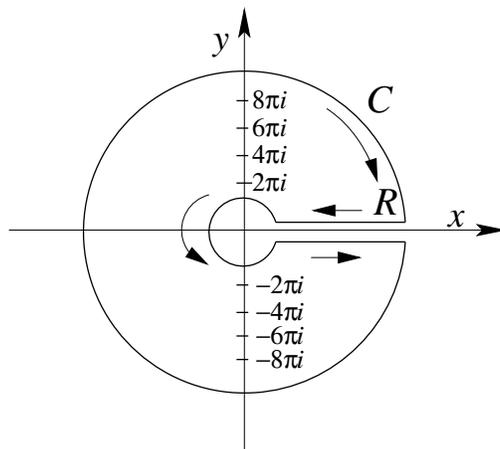


Figura 6.11: Contorno de integración para los números de Bernoulli.

Para  $n \geq 2$  este procedimiento llega a ser algo tedioso, y recurrimos a diferentes maneras de evaluar la ecuación (6.71). El contorno está deformado, como lo muestra la figura 6.11.

El nuevo contorno  $C$  todavía encierra el origen, como se requirió, pero ahora también encierra (en una dirección negativa) series infinitas de puntos singulares a lo largo del eje imaginario en  $z = \pm p2\pi i$ , con  $p = 1, 2, 3, \dots$ . La integración de ida y vuelta a lo largo del eje  $x$  se cancelan, y para  $R \rightarrow \infty$  la integración sobre el círculo infinito tiende a cero. Recuerde que  $n \geq 2$ . Por lo tanto

$$\oint_{C_0} \frac{z}{e^z - 1} \frac{dz}{z^{n+1}} = -2\pi i \sum_{p=1}^{\infty} \text{residuos} \quad (z = \pm p2\pi i). \quad (6.74)$$

En  $z = p2\pi i$  tenemos un polo simple con un residuo  $(p2\pi i)^{-n}$ . Cuando  $n$  es impar, el residuo en  $z = p2\pi i$  se cancela exactamente con el de  $z = -p2\pi i$  y por lo tanto  $B_n = 0$ , con  $n = 3, 5, 7$ , etc. Para  $n$  par los residuos se suman, dando

$$\begin{aligned} B_n &= \frac{n!}{2\pi i} (-2\pi i) 2 \sum_{p=1}^{\infty} \frac{1}{p^n (2\pi i)^n} \\ &= -\frac{(-1)^{n/2} 2n!}{(2\pi)^n} \sum_{p=1}^{\infty} p^{-n} \\ &= -\frac{(-1)^{n/2} 2n!}{(2\pi)^n} \zeta(n) \quad (n \text{ par,}) \end{aligned} \quad (6.75)$$

donde  $\zeta(n)$  es la función zeta de Riemann.

---

<sup>5</sup> $1 + e^z = 1 + e^{z-i\pi} e^{+i\pi} = 1 - e^{z-i\pi} = -(z - i\pi) \left(1 + \frac{z-i\pi}{2!} + \frac{(z-i\pi)^2}{3!} + \dots\right).$

### 6.2.10 Residuos de un polo de orden $m$ .

Si  $f(z)$  tiene un polo de orden  $m$  en  $z = z_0$ , donde  $m$  es un número entero positivo, entonces el residuo  $R$  de  $f(z)$  en  $z = z_0$  es

$$R = \frac{1}{(m-1)!} \lim_{z \rightarrow z_0} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)] , \quad (6.76)$$

donde la derivada de orden cero de la función se entiende la función misma y  $0! = 1$ .

## 6.3 Relaciones de dispersión.

El concepto de relación de dispersión lo introduce en física los trabajos de Kroing y Kramers en óptica. El nombre de dispersión viene de la dispersión óptica, un resultado de la dependencia del índice de refracción con la longitud de onda o frecuencia angular. El índice de refracción  $n$  puede tener una parte real determinada por la velocidad de fase y una (negativa) parte imaginaria determinada por la absorción, ver ecuación (6.91). Kroing y Kramers mostraron en 1926-1927 que la parte real de  $(n^2 - 1)$  podría ser expresada como una integral de la parte imaginaria. Generalizando esto, aplicaremos la etiqueta de relaciones de dispersión a cualquier par de ecuaciones que dan la parte real de una función como una integral de su parte imaginaria y la parte imaginaria como una integral de su parte real ecuaciones (6.82) y (6.83). La existencia de tales relaciones integrales puede entenderse como un análogo integral a las relaciones diferenciales de Cauchy-Riemann, sección 5.2.

Las aplicaciones en física moderna son extensas. Por ejemplo, la parte real de la función podría describir el *scattering* hacia adelante de un rayo gamma en un campo nuclear de Coulomb. Luego la parte imaginaria describiría la producción del par electrón-positrón en este mismo campo de Coulomb. Como será visto más adelante, las relaciones de dispersión pueden ser tomados como una consecuencia de la causalidad y por lo tanto son independientes de los detalles de la interacción particular.

Consideremos una función compleja  $f(z)$  que es analítica en el semiplano superior y sobre el eje real. También requerimos que

$$\lim_{|z| \rightarrow \infty} |f(z)| = 0 , \quad 0 \leq \arg z \leq \pi , \quad (6.77)$$

para que la integral sobre un semicírculo infinito se anule. El punto de esta condición es que podemos expresar  $f(z)$  por la fórmula integral de Cauchy, ecuación (5.48),

$$f(z_0) = \frac{1}{2\pi i} \oint \frac{f(z)}{z - z_0} dz . \quad (6.78)$$

La integral sobre el semicírculo superior<sup>6</sup> se anula y tenemos

$$f(z_0) = \frac{1}{2\pi i} \int \frac{f(x)}{x - z_0} dx . \quad (6.79)$$

La integral sobre el contorno mostrado en la figura 6.12 se ha convertido en una integral a lo largo del eje  $x$ .

<sup>6</sup>El uso de un semicírculo para cerrar el contorno de integración es conveniente, pero no obligatorio. Otros caminos son posibles.

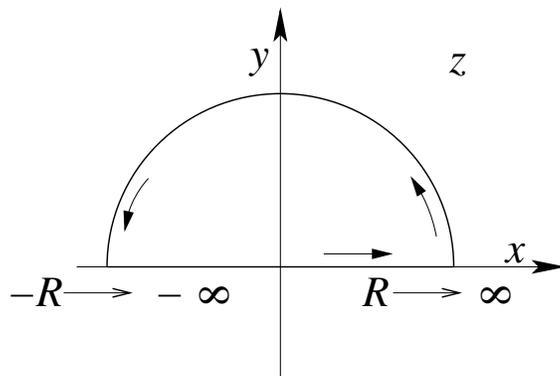


Figura 6.12: Contorno de integración.

La ecuación (6.79) supone que  $z_0$  está en el semiplano superior interior al contorno cerrado. Si  $z_0$  estuviera en el semiplano inferior, la integral produciría cero por el teorema integral de Cauchy. Ahora, ya sea dejando que  $z_0$  se aproxime al eje real desde arriba ( $z_0 \rightarrow x_0$ ), o colocándolo sobre el eje real y tomando un promedio de la ecuación (6.79) y cero, encontramos que la ecuación (6.79) se convierte en

$$f(x_0) = \frac{1}{i\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx, \quad (6.80)$$

donde  $\mathcal{P}$  indica el valor principal de Cauchy.

Separando la ecuación (6.80) en partes real e imaginaria<sup>7</sup> produce

$$\begin{aligned} f(x_0) &= u(x_0) + iv(x_0) \\ &= \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{v(x)}{x - x_0} dx - \frac{i}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{u(x)}{x - x_0} dx. \end{aligned} \quad (6.81)$$

Finalmente, igualando las partes reales y las partes imaginarias obtenemos

$$u(x_0) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{v(x)}{x - x_0} dx, \quad (6.82)$$

$$v(x_0) = -\frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{u(x)}{x - x_0} dx. \quad (6.83)$$

Estas son las relaciones de dispersión. La parte real de nuestra función compleja es expresada como una integral sobre la parte imaginaria. La parte imaginaria está expresada como una integral sobre la parte real. Las partes reales e imaginarias son *transformaciones de Hilbert* una de la otra. Note que estas relaciones son significativas solamente cuando  $f(x)$  es una función compleja de la variable real  $x$ .

Desde un punto de vista físico  $u(x)$  y/o  $v(x)$  representan alguna medida física. Luego  $f(z) = u(z) + iv(z)$  es una continuación analítica sobre el semiplano superior, con el valor en el eje real sirviendo como una condición de borde.

<sup>7</sup>El segundo argumento,  $y = 0$  es omitido.  $u(x_0, 0) \rightarrow u(x_0)$ .

### 6.3.1 Relaciones de Simetría.

En ocasiones  $f(x)$  satisfecerá una relación de simetría y la integral desde  $-\infty$  a  $+\infty$  puede ser reemplazada por una integral solamente sobre valores positivos. Esto es de importancia física considerable ya que la variable  $x$  podría representar una frecuencia y solamente frecuencias  $\geq 0$  son posibles en físicas. Supongamos<sup>8</sup>

$$f(-x) = f^*(x) . \quad (6.84)$$

Entonces

$$u(-x) + iv(-x) = u(x) - iv(x) . \quad (6.85)$$

La parte real de  $f(x)$  es par y la parte imaginaria es impar. En problemas de scattering en mecánica cuántica estas relaciones (ecuación (6.85)) son llamadas *condiciones de cruzamiento*. Para sacar utilidad de estas condiciones, reescribimos la ecuación (6.82) como

$$u(x_0) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^0 \frac{v(x)}{x - x_0} dx + \frac{1}{\pi} \mathcal{P} \int_0^{\infty} \frac{v(x)}{x - x_0} dx . \quad (6.86)$$

Haciendo  $x \rightarrow -x$  en la primera integral al lado derecho de la ecuación (6.86) y sustituyendo  $v(-x) = -v(x)$  de la ecuación (6.85), obtenemos

$$\begin{aligned} u(x_0) &= \frac{1}{\pi} \mathcal{P} \int_0^{\infty} v(x) \left\{ \frac{1}{x + x_0} + \frac{1}{x - x_0} \right\} dx \\ &= \frac{2}{\pi} \mathcal{P} \int_0^{\infty} \frac{xv(x)}{x^2 - x_0^2} dx . \end{aligned} \quad (6.87)$$

Similarmente,

$$v(x_0) = -\frac{2}{\pi} \mathcal{P} \int_0^{\infty} \frac{x_0 u(x)}{x^2 - x_0^2} dx . \quad (6.88)$$

Las originales relaciones de dispersión óptica de Kroing-Kramers eran de esta forma.

### 6.3.2 Dispersión óptica.

La función  $\exp[i(kx - \omega t)]$  describe una onda moviéndose a lo largo del eje  $x$  en la dirección positiva con velocidad  $v = \omega/k$ ,  $\omega$  es la frecuencia angular,  $k$  el número de onda o vector de propagación, y  $n = ck/\omega$  el índice de refracción. A partir de las ecuaciones de Maxwell, la permitividad eléctrica  $\varepsilon$ , y la ley de Ohm con conductividad  $\sigma$  el vector de propagación  $k$  para un dieléctrico llega a ser

$$k^2 = \varepsilon \frac{\omega^2}{c^2} \left( 1 + i \frac{4\pi\sigma}{\omega\varepsilon} \right) \quad (6.89)$$

---

<sup>8</sup>Esto no es una feliz coincidencia. Esto asegura que la transformada de Fourier de  $f(x)$  será real.

(con  $\mu$ , la permeabilidad magnética tomada como uno). La presencia de la conductividad (la cual significa absorción) da un aumento en la parte imaginaria. El vector de propagación  $k$  (y por lo tanto el índice de refracción  $n$ ) se ha convertido en complejo.

Inversamente, la parte imaginaria (positiva) implica absorción. Para conductividades pobres ( $4\pi\sigma/\omega\varepsilon \ll 1$ ) una expansión binomial produce

$$k = \sqrt{\varepsilon} \frac{\omega}{c} + i \frac{2\pi\sigma}{c\sqrt{\varepsilon}}$$

y

$$e^{i(kx-\omega t)} = e^{i\omega(x\sqrt{\varepsilon}/c-t)} e^{-2\pi\sigma x/c\sqrt{\varepsilon}},$$

una onda atenuada.

Volviendo a la expresión general para  $k^2$ , encontramos que la ecuación (6.79) el índice de refracción se convierte en

$$n^2 = \frac{c^2 k^2}{\omega^2} = \varepsilon + i \frac{4\pi\sigma}{\omega}. \quad (6.90)$$

Tomamos  $n^2$  como una función de la variable compleja  $\omega$  (con  $\varepsilon$  y  $\sigma$  dependiendo de  $\omega$ ). Sin embargo,  $n^2$  no se anula cuando  $\omega \rightarrow \infty$  sino realmente se aproxima a la unidad. Tal de satisfacer la condición, ecuación (6.77), uno trabaja con  $f(\omega) = n^2(\omega) - 1$ . Las relaciones originales de dispersión óptica de Kroing-Kramers estaban en la forma de

$$\begin{aligned} \Re[n^2(\omega_0) - 1] &= \frac{2}{\pi} \mathcal{P} \int_0^\infty \frac{\omega \Im[n^2(\omega) - 1]}{\omega^2 - \omega_0^2} d\omega \\ \Im[n^2(\omega_0) - 1] &= -\frac{2}{\pi} \mathcal{P} \int_0^\infty \frac{\omega_0 \Re[n^2(\omega) - 1]}{\omega^2 - \omega_0^2} d\omega. \end{aligned} \quad (6.91)$$

El conocimiento de los coeficientes de absorción en todas las frecuencias especifica la parte real del índice de refracción y vice versa.

### 6.3.3 La relación de Parseval.

Cuando las funciones  $u(x)$  y  $v(x)$  son transformadas de Hilbert una de la otra y cada una es cuadrado integrable<sup>9</sup>, las dos funciones están relacionadas por

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \int_{-\infty}^{\infty} |v(x)|^2 dx. \quad (6.92)$$

Esta es la relación de Parseval.

Para derivar la ecuación (6.92), comenzamos con

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \int_{-\infty}^{\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{v(s) ds}{s-x} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{v(t) dt}{t-x} dx$$

<sup>9</sup>Esto significa que  $\int_{-\infty}^{\infty} |u(x)|^2 dx$  y  $\int_{-\infty}^{\infty} |v(x)|^2 dx$  son finitas.

usando la ecuación (6.82) dos veces.

Integrando primero con respecto a  $x$ , tenemos

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\pi^2} \int_{-\infty}^{\infty} \frac{dx}{(s-x)(t-x)} v(s) ds v(t) dt . \quad (6.93)$$

Parte de la integración en  $x$  produce una función delta:

$$\frac{1}{\pi^2} \int_{-\infty}^{\infty} \frac{dx}{(s-x)(t-x)} = \delta(s-t) .$$

Tenemos

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v(s) \delta(s-t) ds v(t) dt . \quad (6.94)$$

Luego la integración en  $s$  es llevada a cabo por inspección, usando la propiedad que define la función delta.

$$\int_{-\infty}^{\infty} v(s) \delta(s-t) ds = v(t) . \quad (6.95)$$

Sustituyendo la ecuación (6.95) en la (6.94), tenemos la ecuación (6.92), la relación de Parseval.

### 6.3.4 Causalidad.

El significado real de las relaciones de dispersión en física es que ellas son una consecuencia directa de suponer que un sistema físico en particular obedece causalidad. Causalidad es difícil para definir precisamente pero el significado general es que el efecto no puede preceder la causa. Una onda scattered no puede ser emitida por el centro scattering antes de que la onda incidente haya llegado. Para sistemas lineales la relación más general entre una función de entrada  $G$  (la causa) y la función de salida  $H$  (el efecto) puede ser escrita como

$$H(t) = \int_{-\infty}^{\infty} F(t-t')G(t') dt' . \quad (6.96)$$

La causalidad es impuesta requiriendo que

$$F(t-t') = 0 \quad \text{para } t-t' < 0 .$$

La ecuación (6.96) da la dependencia temporal. La dependencia en la frecuencia es obtenida tomando las transformadas de Fourier. Por el teorema de convolución de Fourier,

$$h(\omega) = f(\omega)g(\omega) ,$$

donde  $f(\omega)$  es la transformada de Fourier  $F(t)$ , etc. Inversamente,  $F(t)$  es la transformada de Fourier de  $f(\omega)$ .

La conexión con las relaciones de dispersión están dadas por el teorema de Tichmarsh. Este establece que si  $f(\omega)$  es cuadrado integrable sobre el eje real  $\omega$ , entonces cualquiera de las tres afirmaciones implica las otras dos.

1. La transformada de Fourier para  $f(\omega)$  es cero para  $t < 0$  ecuación (6.96).
2. Remplazando  $\omega$  por  $z$ , la función  $f(z)$  es analítica en el plano complejo  $z$  para  $y > 0$  y se aproxima a  $f(x)$  casi en cualquier lugar cuando  $y \rightarrow 0$ . Además,

$$\int_{-\infty}^{\infty} |f(x + iy)|^2 dx < K \quad \text{para } y > 0 ,$$

esto es, la integral está acotada.

3. Las partes reales e imaginarias de  $f(z)$  son transformaciones de Hilbert la una de la otra: ecuaciones (6.82) y (6.83).

La suposición de esta relación entre la entrada y la salida de nuestro sistema lineal es casual (ecuación (6.96)) significa que la primera afirmación es satisfecha. Si  $f(\omega)$  es cuadrado integrable, entonces el teorema de Tichmarsh tiene la tercera afirmación como una consecuencia y tenemos relaciones de dispersión.

## 6.4 El método de *steepest descents*.

Analizando problemas en física matemática, a menudo uno encuentra deseable conocer el comportamiento de una función para valores grandes de la variable, esto es, el comportamiento asintótico de la función. Ejemplos específicos son proporcionados por la función gamma y los varios tipos de funciones de Bessel. El método de *steepest descents* es un método para determinar tales comportamientos asintóticos cuando la función puede ser expresada como la integral de la forma general

$$I(s) = \int_C g(z) e^{sf(z)} dz . \quad (6.97)$$

Para el momento, tomemos  $s$  real. El contorno  $C$  de integración es entonces escogido tal que la parte real de  $f(z)$  tienda a menos infinito en ambos límites y que el integrando se anulará en ambos límites, o es escogido como un contorno cerrado. Más aún, supongamos que el factor  $g(z)$  en el integrando es dominado por la exponencial en la región de interés.

Si el parámetro  $s$  es grande y positivo, el valor del integrando llegará a ser grande cuando la parte real de  $f(z)$  es grande y pequeño cuando la parte real de  $f(z)$  es pequeña y negativa. En particular, cuando  $s$  se le permite aumentar indefinidamente (dejando la dependencia asintótica), la contribución completa del integrando a la integral vendrá de la región en la cual la parte real de  $f(z)$  tome un valor máximo positivo. Fuera de este valor máximo positivo el integrando llegará a ser en comparación despreciablemente pequeño. Esto se ve expresando  $f(z)$  como

$$f(z) = u(x, y) + iv(x, y) .$$

Entonces la integral puede ser reescrita como

$$I(s) = \int_C g(z) e^{su(x,y)} e^{isv(x,y)} dz . \quad (6.98)$$

Si ahora, adicionalmente, imponemos la condición que la parte imaginaria del exponente,  $iv(x, y)$ , sea constante en la región en la cual la parte real toma su valor máximo, esto es,  $v(x, y) = v(x_0, y_0) = v_0$ , podemos aproximar la integral por

$$I(s) \approx e^{isv_0} \int_C g(z) e^{su(x,y)} dz \quad (6.99)$$

Fuera del máximo de la parte real, la parte imaginaria se le puede permitir oscilar como desee, por el integrando despreciablemente pequeño y el factor de fase variante es por lo tanto irrelevante.

La parte real de  $sf(z)$  es un máximo para un  $s$  dado cuando la parte real de  $f(z)$ ,  $u(x, y)$  es un máximo. Esto implica que

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = 0 ,$$

y por lo tanto, por el uso de las condiciones de Cauchy-Riemann

$$\frac{df(z)}{dz} = 0 , \quad (6.100)$$

procedemos a buscar tales ceros de la derivada.

Es esencial notar que el valor máximo de  $u(x, y)$  es el máximo sólo en un contorno dado. En el plano finito ni la parte real ni la parte imaginaria de nuestra función analítica posee un máximo absoluto. Esto puede ser visto recordando que ambas  $u$  y  $v$  satisfacen la ecuación de Laplace

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 . \quad (6.101)$$

A partir de ésto, si la segunda derivada con respecto a  $x$  es positiva, la segunda derivada con respecto a  $y$  debe ser negativa, y por lo tanto ni  $u$  ni  $v$  pueden poseer un máximo o mínimo absoluto. Ya que la función  $f(z)$  fue tomada como analítica, los puntos singulares claramente están excluidos. La anulación de la derivada, ecuación (6.100), implica que tenemos un punto de ensilladura, un valor estacionario, el cual puede ser un máximo de  $u(x, y)$  para un contorno y un mínimo para otro (figura 6.13).

Nuestro problema, entonces, es escoger el contorno de integración para satisfacer dos condiciones. (1) El contorno debe ser escogido tal que  $u(x, y)$  tenga un máximo en el punto de ensilladura. (2) El contorno debe pasar a través del punto de ensilladura de tal manera que la parte imaginaria,  $v(x, y)$  sea una constante. Esta segunda condición conduce al camino de descenso más abrupto *steepest descent* y da al método su nombre. Del capítulo anterior sabemos que las curvas correspondientes a  $u = \text{constante}$  y a  $v = \text{constante}$  forman un sistema ortogonal. Esto significa que una curva  $v = c_i$ , constante, es tangencial en cualquier parte a la gradiente de  $u$ ,  $\vec{\nabla}u$ . Por lo tanto, la curva  $v = \text{constante}$  es la curva que da la línea de *steepest descent* desde el punto de ensilladura.<sup>10</sup>

<sup>10</sup>La línea de *steepest ascent* está también caracterizada por  $v$  constante. El punto de ensilladura debe ser inspeccionado cuidadosamente para distinguir la línea de *steepest descent* respecto a la línea de *steepest ascent*.

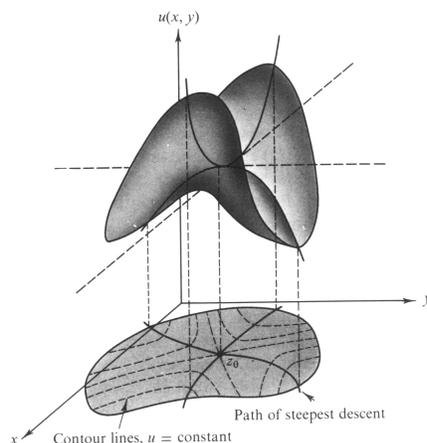


Figura 6.13: Punto de ensilladura.

En el punto de ensilladura la función  $f(z)$  puede ser expandida en serie de Taylor para dar

$$f(z) = f(z_0) + \frac{1}{2}(z - z_0)^2 f''(z_0) + \dots \quad (6.102)$$

La primera derivada está ausente, ya que obviamente la ecuación (6.100) es satisfecha. El primer término de corrección,  $(z - z_0)^2 f''(z_0)/2$ , es real y negativo. Es real, porque hemos especificado que la parte imaginaria es constante a lo largo de nuestro contorno y negativo porque nos estamos moviendo hacia abajo del punto de ensilladura. Entonces, suponiendo que  $f''(z_0) \neq 0$ , tenemos

$$f(z) - f(z_0) \approx \frac{1}{2}(z - z_0)^2 f''(z_0) = -\frac{1}{2s} t^2, \quad (6.103)$$

la cual sirve para definir una nueva variable  $t$ . Si  $(z - z_0)$  es escrito en forma polar

$$(z - z_0) = \delta e^{i\alpha}, \quad (6.104)$$

(con la fase  $\alpha$  mantenida constante), tenemos

$$t^2 = -s f''(z_0) \delta^2 e^{2i\alpha}. \quad (6.105)$$

Ya que  $t$  es real<sup>11</sup>, la ecuación puede ser escrita como

$$t = \pm \delta |s f''(z_0)|^{1/2}. \quad (6.106)$$

Sustituyendo la ecuación (6.103) en (6.97), obtenemos

$$I(s) \approx g(z_0) e^{s f(z_0)} \int_{-\infty}^{\infty} e^{-t^2/2} \frac{dz}{dt} dt. \quad (6.107)$$

<sup>11</sup>La fase del contorno (especificada por  $\alpha$ ) en el punto de ensilladura es elegida tal que  $\Im[f(z) - f(z_0)] = 0$ , lo que implica que  $\frac{1}{2}(z - z_0)^2 f''(z_0)$  debe ser real.

Tenemos

$$\frac{dz}{dt} = \left(\frac{dt}{dz}\right)^{-1} = \left(\frac{dt}{d\delta} \frac{d\delta}{dz}\right)^{-1} = |sf''(z_0)|^{-1/2} e^{i\alpha}, \quad (6.108)$$

a partir de las ecuaciones (6.104) y (6.106). La ecuación (6.107) se transforma en

$$I(s) \approx \frac{g(z_0)e^{sf(z_0)}e^{i\alpha}}{|sf''(z_0)|^{1/2}} \int_{-\infty}^{\infty} e^{-t^2/2} dt. \quad (6.109)$$

Se notará que los límites han sido ajustados como menos infinito a más infinito. Esto es permitido, porque el integrando es esencialmente cero cuando  $t$  se desvía apreciablemente del origen. Note que la integral remanente es una integral de error de Gauss igual a  $\sqrt{2\pi}$ , finalmente obtenemos

$$I(s) \approx \frac{\sqrt{2\pi}g(z_0)e^{sf(z_0)}e^{i\alpha}}{|sf''(z_0)|^{1/2}}. \quad (6.110)$$

La fase  $\alpha$  fue introducida en la ecuación (6.104) como la fase del contorno cuando pasó a través del punto de ensilladura. Esta es escogida tal que las dos condiciones dadas [ $\alpha$ =constante;  $\Re[f(z)]$ =máximo] son satisfechas. Ocurre algunas veces que el contorno pasa a través de dos o más puntos de ensilladura en sucesión. Si este es el caso, necesitamos solamente añadir la contribución hecha por la ecuación (6.110) para cada uno de los puntos de ensilladura para obtener una aproximación a la integral total.

Una nota de advertencia: Supusimos que la única contribución significativa a la integral viene de la vecindad inmediata del o de los puntos de ensilladura  $z = z_0$ , esto es,

$$\Re[f(z)] = u(x, y) \ll u(x_0, y_0),$$

sobre el contorno entero a partir de  $z_0 = x_0 + iy_0$ . Esta condición debe ser chequeada para cada uno de los nuevos problemas.

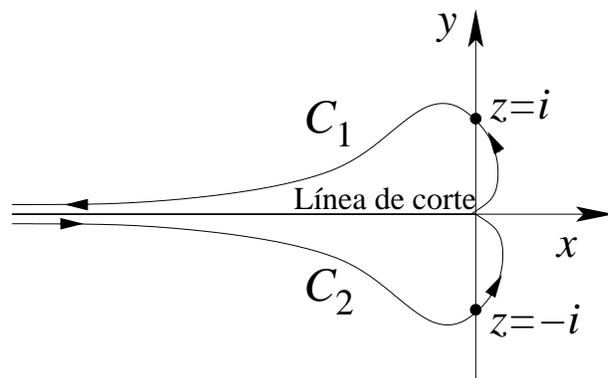


Figura 6.14: Contornos para las funciones de Hankel.

**Ejemplo** Forma asintótica de la función de Hankel,  $H_v^{(1)}(s)$

Las funciones de Hankel satisfacen las ecuaciones de Bessel y pueden ser definidas por

$$H_v^{(1)}(s) = \frac{1}{i\pi} \int_{0C_1}^{\infty} e^{(s/2)(z-1/z)} \frac{dz}{z^{\nu+1}}, \quad (6.111)$$

$$H_v^{(2)}(s) = \frac{1}{i\pi} \int_{-\infty C_2}^{\infty} e^{(s/2)(z-1/z)} \frac{dz}{z^{\nu+1}}. \quad (6.112)$$

El contorno  $C_1$  es la curva en el semiplano superior de la figura 6.14. El contorno  $C_2$  está en el semiplano inferior. Aplicamos el método de *steepest descent* a la primera función de Hankel,  $H_v^{(1)}(s)$ , con  $f(z)$  dado por

$$f(z) = \frac{1}{2} \left( z - \frac{1}{z} \right). \quad (6.113)$$

Diferenciando, obtenemos

$$f'(z) = \frac{1}{2} + \frac{1}{2z^2}. \quad (6.114)$$

Fijando  $f'(z) = 0$  de acuerdo con la ecuación (6.90), obtenemos

$$z = i, -i. \quad (6.115)$$

De modo que hay puntos de ensilladura en  $z = +i$  y  $z = -i$ . La integral para  $H_v^{(1)}(s)$  es escogida tal que comience en el origen, se mueva tangencialmente al eje real positivo, y luego se mueva a través del punto de ensilladura en  $z = +i$  y hacia afuera a menos infinito, asintótica con el eje real negativo. Debemos escoger el contorno a través del punto  $z = +i$  en tal manera que la parte real de  $(z - 1/z)$  sea un máximo y al fase sea constante en la vecindad del punto soporte. Tenemos

$$\Re \left( z - \frac{1}{z} \right) = 0 \quad \text{para } z = i.$$

Requerimos que  $\Re(z - 1/z) < 0$  para el resto de  $C_1(z \neq 1)$ .

En la vecindad del punto de ensilladura en  $z_0 = +i$  tenemos

$$z - i = \delta e^{i\alpha}, \quad (6.116)$$

donde  $\delta$  es un número pequeño. Entonces

$$\begin{aligned} 2f(z) &= z - \frac{1}{z} = \delta e^{i\alpha} + i - \frac{1}{\delta e^{i\alpha} + i} \\ &= \delta \cos \alpha + i(\delta \sin \alpha + 1) - \frac{1}{\delta \cos \alpha + i(\delta \sin \alpha + 1)} \\ &= \delta \cos \alpha + i(\delta \sin \alpha + 1) - \frac{\delta \cos \alpha - i(\delta \sin \alpha + 1)}{1 + 2\delta \sin \alpha + \delta^2}. \end{aligned} \quad (6.117)$$

Por lo tanto nuestra parte real llega a ser

$$\Re\left(z - \frac{1}{z}\right) = \delta \cos \alpha - \delta \cos \alpha (1 + 2\delta \sin \alpha + \delta^2)^{-1}. \quad (6.118)$$

Recordando que  $\delta$  es pequeño, expandimos por el teorema del binomio y despreciando ordenes en  $\delta^3$  y mayores.

$$\Re\left(z - \frac{1}{z}\right) = 2\delta^2 \cos \alpha \sin \alpha + \mathcal{O}(\delta^3) \approx \delta^2 \sin 2\alpha. \quad (6.119)$$

Vemos que la parte real de  $(z - 1/z)$  tomará un valor extremo si  $\sin 2\alpha$  es un extremo, esto es, si  $2\alpha$  es  $\pi/2$  o  $3\pi/2$ . De modo que la fase del contorno  $\alpha$  debe escogerse como  $\pi/4$  o  $3\pi/4$ . Una elección representará el camino de *steepest descent* que deseamos. La otra alternativa representará el camino de *steepest ascent* que debemos evitar. Distinguiamos las dos posibilidades sustituyendo los valores espacíficos de  $\alpha$ . Para  $\alpha = \pi/4$

$$\Re\left(z - \frac{1}{z}\right) = \delta^2. \quad (6.120)$$

Para la elección  $z = i$  es un mínimo.

Para  $\alpha = 3\pi/4$

$$\Re\left(z - \frac{1}{z}\right) = -\delta^2, \quad (6.121)$$

y  $z = i$  es un máximo. Esta es la fase que deseamos.

La sustitución directa en la ecuación (6.110) con  $\alpha = 3\pi/4$  ahora produce

$$\begin{aligned} H_\nu^{(1)}(s) &= \frac{1}{\pi i} \frac{\sqrt{2\pi} i^{-\nu-1} e^{(s/2)(i-1/i)} e^{i(3\pi/4)}}{|(s/2)(-2/i^3)|^{1/2}} \\ &= \sqrt{\frac{2}{\pi s}} e^{(i\pi/2)(-\nu-2)} e^{is} e^{i(3\pi/4)}. \end{aligned} \quad (6.122)$$

Combinando términos, finalmente obtenemos

$$H_\nu^{(1)}(s) = \sqrt{\frac{2}{\pi s}} e^{i(s-\nu(\pi/2)-\pi/4)} \quad (6.123)$$

que conducen a la expansión asintótica de la función de Hankel  $H_\nu^{(1)}(s)$ . Los términos adicionales, si se desean, pueden ser obtenidos suponiendo una serie de potencias descendentes y sustituyendo de regreso en la ecuación de Bessel.

**Ejemplo** Forma asintótica de la función factorial,  $s!$ .

En muchos problemas físicos, particularmente en el campo de la mecánica estadística, es deseable tener una aproximación precisa de la función gamma o función factorial de números muy grandes. Como desarrollaremos más adelante, la función factorial puede ser definida por la integral

$$s! = \int_0^\infty \rho^s e^{-\rho} d\rho = s^{s+1} \int_{0C}^\infty e^{s(\ln z - z)} dz. \quad (6.124)$$

Aquí hemos hecho una sustitución  $\rho = zs$  para que la integral este en la forma requerida por la ecuación (6.97). Como antes, suponemos que  $s$  es real y positiva, de la cual sigue que el integrando se anula en los límites  $0$  y  $\infty$ . Diferenciando con respecto a  $z$  el exponente, obtenemos

$$\frac{df(z)}{dz} = \frac{d}{dz}(\ln z - z) = \frac{1}{z} - 1, \quad (6.125)$$

la cual muestra que el punto  $z = 1$  es un punto de ensilladura. Sea

$$z - 1 = \delta e^{i\alpha}, \quad (6.126)$$

con  $\delta$  pequeño para describir el contorno en la vecindad del punto de ensilladura. Sustituyendo en  $f(z) = \ln z - z$ , desarrollamos una expansión en serie

$$\begin{aligned} f(z) &= \ln(1 + \delta e^{i\alpha}) - (1 + \delta e^{i\alpha}) \\ &= \delta e^{i\alpha} - \frac{1}{2}\delta^2 e^{2i\alpha} + \dots - 1 - \delta e^{i\alpha} \\ &= -1 - \frac{1}{2}\delta^2 e^{2i\alpha} + \dots \end{aligned} \quad (6.127)$$

De esto vemos que el integrando toma un valor máximo ( $e^{-s}$ ) en el punto de ensilladura si escogemos nuestro contorno  $C$  siguiendo el eje real, una conclusión que se puede llegar más o menos intuitivamente.

Una sustitución directa en la ecuación (6.110) con  $\alpha = 0$  ahora da

$$s! \approx \frac{\sqrt{2\pi} s^{s+1} e^{-s}}{|s(-1)^{-2}|^{1/2}}. \quad (6.128)$$

Por lo tanto el primer término en la expansión asintótica de la función factorial es

$$s! \approx \sqrt{2\pi} s^s e^{-s}. \quad (6.129)$$

Este resultado es el primer término en la expansión de Stirling de la función factorial. El método de *steepest descent* es probablemente la más fácil manera de obtener este primer término. Si se desean más términos en la expansión, entonces es preferible otro método, el cual veremos en el próximo curso.

En los ejemplos precedentes el cálculo fue llevado a cabo suponiendo  $s$  como real. Esta suposición no es necesaria. Se puede mostrar que la ecuación (6.129) también se mantiene cuando  $s$  es reemplazada por la variable compleja  $w$ , probando solamente que la parte real de  $w$  se requiere grande y positiva.



# Capítulo 7

## Ecuaciones diferenciales.

versión final 1.2-170702<sup>1</sup>

### 7.1 Ecuaciones diferenciales parciales, características y condiciones de borde.

En Física el conocimiento de la fuerza en una ecuación de movimiento usualmente conduce a una ecuación diferencial. Por lo tanto, casi todas las partes elementales y numerosas partes avanzadas de la Física teórica están formuladas en términos de ecuaciones diferenciales. Algunas veces son ecuaciones diferenciales ordinarias en una variable (ODE). Más a menudo las ecuaciones son ecuaciones diferenciales parciales (PDE) en dos o más variables.

Recordemos del cálculo la operación de tomar una derivada ordinaria o parcial es una *operación lineal*<sup>2</sup> ( $\mathcal{L}$ )

$$\frac{d(a\varphi(x) + b\psi(x))}{dx} = a\frac{d\varphi}{dx} + b\frac{d\psi}{dx} ,$$

para ODE que involucran derivadas en una variable  $x$  solamente y no cuadráticas,  $(d\psi/dx)^2$ , o potencias mayores. Similarmente, para derivadas parciales,

$$\frac{\partial(a\varphi(x, y) + b\psi(x, y))}{\partial x} = a\frac{\partial\varphi(x, y)}{\partial x} + b\frac{\partial\psi(x, y)}{\partial x} ,$$

En general

$$\mathcal{L}(a\varphi + b\psi) = a\mathcal{L}(\varphi) + b\mathcal{L}(\psi) . \quad (7.1)$$

Así, las ODE y las PDE aparecen como ecuaciones de operadores lineales

$$\mathcal{L}(\psi) = F ,$$

---

<sup>1</sup>Este capítulo está basado en el octavo capítulo del libro: *Mathematical Methods for Physicists, fourth edition* de George B. Arfken & Hans J. Weber, editorial ACADEMIC PRESS.

<sup>2</sup>Estamos especialmente interesados en operadores lineales porque en mecánica cuántica las cantidades físicas están representadas por operadores lineales operando en un espacio complejo de Hilbert de dimensión infinita.

donde  $F$  es una función conocida de una (para ODE) o más variables (para PDE),  $\mathcal{L}$  es una combinación lineal de derivadas,  $\psi$  es una función o solución desconocida. Cualquier combinación lineal de soluciones es de nuevo una solución; esto es el *principio de superposición*.

Ya que la dinámica de muchos sistemas físicos involucran sólo dos derivadas, *e.g.*, la aceleración en mecánica clásica y el operador de energía cinética,  $\sim \nabla^2$ , en mecánica cuántica, las ecuaciones diferenciales de segundo orden ocurren más frecuentemente en Física. [Las ecuaciones de Maxwell y de Dirac son de primer orden pero involucran dos funciones desconocidas. Eliminando una incógnita conducen a una ecuación diferencial de segundo orden por la otra.]

### 7.1.1 Ejemplos de PDE.

Entre las PDE más frecuentemente encontradas tenemos:

1. La ecuación de Laplace,  $\nabla^2\psi = 0$ . Esta ecuación muy común y muy importante aparece en el estudio de
  - a. Fenómenos electromagnéticos incluyendo electroestáticos, dieléctricos, corrientes estacionarias y magnetoestática.
  - b. Hidrodinámica (flujo irrotacional de líquidos perfectos y superficies de ondas).
  - c. Flujo de calor.
  - d. Gravitación.
2. La ecuación de Poisson,  $\nabla^2\psi = -4\pi\rho$  En contraste a la ecuación homogénea de Laplace, la ecuación de Poisson es no homogénea con un término de fuente  $-4\pi\rho$ .
3. Las ecuaciones de onda (Helmholtz) y las ecuaciones de difusión tiempo independiente,  $\nabla^2\psi \pm k^2\psi = 0$ . Estas ecuaciones aparecen en fenómenos tan diversos como
  - a. Ondas elásticas en sólidos incluyendo cuerdas vibrantes, barras y membranas.
  - b. En sonido o acústica.
  - c. En ondas electromagnéticas.
  - d. En reactores nucleares.
4. La ecuación de difusión tiempo dependiente

$$\nabla^2\psi = \frac{1}{a^2} \frac{\partial\psi}{\partial t},$$

y su correspondiente forma cuadridimensional que involucra el D'Alembertiano, un análogo cuadridimensional del Laplaciano en el espacio Minkowski,

$$\partial^\mu \partial_\mu = \partial^2 = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2.$$

5. Las ecuaciones de onda tiempo dependiente,  $\partial^2\psi = 0$ .

6. La ecuación del potencial escalar,  $\partial^2\psi = 4\pi\rho$ . Como la ecuación de Poisson esta ecuación es no homogénea con un término de fuente  $4\pi\rho$ .
7. La ecuación de Klein-Gordon,  $\partial^2\psi = -\mu^2\psi$ , y las correspondientes ecuaciones vectoriales en la cual la función escalar  $\psi$  es reemplazada por una función vectorial. Otras formas complicadas son comunes.
8. La ecuación de onda de Schrödinger,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t}$$

y

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi ,$$

para el caso tiempo independiente.

9. Las ecuaciones para ondas elásticas y líquidos viscosos y la ecuación telegráfica.
10. Ecuaciones diferenciales parciales acopladas de Maxwell para los campos eléctricos y magnéticos son aquellas de Dirac para funciones de ondas relativista del electrón.

Algunas técnicas generales para resolver PDE de segundo orden son discutidas en esta sección:

1. Separación de variables, donde el PDE es separada en ODE que están relacionadas por constantes comunes las cuales aparecen como autovalores de operadores lineales,  $\mathcal{L}\psi = l\psi$ , usualmente en una variable. La ecuación de Helmholtz dada como ejemplo 3 anteriormente tiene esta forma, donde el autovalor  $k^2$  puede surgir por la separación del tiempo  $t$  respecto de las variables espaciales. Como en el ejemplo 8, la energía  $E$  es el autovalor que surge en la separación de  $t$  respecto de  $\vec{r}$  en la ecuación de Schrödinger.
2. Conversión de una PDE en una ecuación integral usando funciones de Green que se aplica a PDE no homogéneas tales como los ejemplos 2 y 6 dados más arriba.
3. Otros métodos analíticos tales como el uso de transformadas integrales que serán desarrolladas en el próximo curso.
4. Cálculo numérico. El desarrollo de los computadores ha abierto una abundancia de posibilidades basadas en el cálculo de diferencias finitas. Aquí también tenemos los métodos de relajación. Métodos como Runge-Kutta y predictor-corrector son aplicados a ODEs.

Ocasionalmente, encontramos ecuaciones de orden mayor. En ambos la teoría del movimiento suave de un líquido viscoso y la teoría de un cuerpo elástico encontramos la ecuación

$$(\nabla^2)^2\psi = 0 .$$

Afortunadamente, estas ecuaciones diferenciales de orden más altos son relativamente raras y no son discutidas en una etapa introductoria como esta.

Aunque no son tan frecuentemente encontrados y quizás no son tan importantes como las ecuaciones diferenciales de segundo orden, las ecuaciones diferenciales de primer orden aparecen en Física teórica y algunas veces son pasos intermedios para ecuaciones diferenciales de segundo orden. Las soluciones de algunos de los tipos más importantes de ODE de primer orden son desarrollados en la sección 7.2. Las PDEs de primer orden siempre pueden ser reducidas a ODEs. Este es un proceso directo pero lento e involucra una búsqueda para las características que son presentadas brevemente más adelante.

### 7.1.2 Clases de PDE y característica.

Las PDEs de segundo orden forman tres clases: (i) Las PDEs elípticas que involucran  $\nabla^2$  o  $c^{-2}\partial^2/\partial t^2 + \nabla^2$ ; (ii) Las PDEs parabólica,  $a\partial/\partial t - \nabla^2$ ; (iii) Las PDEs hiperbólica,  $c^{-2}\partial^2/\partial t^2 - \nabla^2$ . Estos operadores canónicos aparecen por un cambio de variables  $\xi = \xi(x, y)$ ,  $\eta = \eta(x, y)$  en un operador lineal (para dos variables sólo por simplicidad)

$$\mathcal{L} = a\frac{\partial^2}{\partial x^2} + 2b\frac{\partial^2}{\partial x\partial y} + c\frac{\partial^2}{\partial y^2} + d\frac{\partial}{\partial x} + e\frac{\partial}{\partial y} + f, \quad (7.2)$$

la cual puede ser reducida a las formas canónicas (i), (ii), (iii) de acuerdo a si el discriminante  $D = ac - b^2 > 0$ ,  $= 0$  o  $< 0$ . Si  $\xi(x, y)$  es determinada a partir de la ecuación de primer orden, pero no lineal, PDE

$$a\left(\frac{\partial\xi}{\partial x}\right)^2 + 2b\left(\frac{\partial\xi}{\partial x}\right)\left(\frac{\partial\xi}{\partial y}\right) + c\left(\frac{\partial\xi}{\partial y}\right)^2 = 0, \quad (7.3)$$

donde los términos de más bajo orden en  $\mathcal{L}$  son ignorados, entonces los coeficientes de  $\partial^2/\partial\xi^2$  en  $\mathcal{L}$  es cero (*i.e.*, ecuación (7.3)). Si  $\eta$  es una solución independiente de la misma ecuación (7.3), entonces el coeficiente de  $\partial^2/\partial\eta^2$  también es cero. El operador remanente  $\partial^2/\partial\xi\partial\eta$  en  $\mathcal{L}$  es característico del caso hiperbólico (iii) con  $D < 0$ , donde la forma cuadrática  $a\lambda^2 + 2b\lambda + c$  es factorizable y, por lo tanto, la ecuación (7.3) tiene dos soluciones independientes  $\xi(x, y)$ ,  $\eta(x, y)$ . En el caso elíptico (i) con  $D > 0$  las dos soluciones  $\xi$ ,  $\eta$  son complejos conjugados los cuales, cuando se sustituyeron en la ecuación (7.2), remueven la derivada de segundo orden mezclada en vez de los otros términos de segundo orden produciendo la forma canónica (i). En el caso parabólico (ii) con  $D = 0$ , solamente  $\partial^2/\partial\xi^2$  permanece en  $\mathcal{L}$ , mientras que los coeficientes de las otras dos derivadas de segundo orden se anulan.

Si los coeficientes  $a, b, c$  en  $\mathcal{L}$  son funciones de las coordenadas, entonces esta clasificación es solamente local, *i.e.*, su tipo podría cambiar cuando las coordenadas varían.

Ilustremos la física implícita en el caso hiperbólico mirando la ecuación de onda (en 1 + 1 dimensiones por simplicidad)

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right)\psi = 0. \quad (7.4)$$

Ya que la ecuación (7.3) se convierte en

$$\left(\frac{\partial\xi}{\partial t}\right)^2 - c^2\left(\frac{\partial\xi}{\partial x}\right)^2 = \left(\frac{\partial\xi}{\partial t} - c\frac{\partial\xi}{\partial x}\right)\left(\frac{\partial\xi}{\partial t} + c\frac{\partial\xi}{\partial x}\right) = 0, \quad (7.5)$$

y es factorizable, determinamos la solución de  $\partial\xi/\partial t - c\partial\xi/\partial x = 0$ . Esta es una función arbitraria  $\xi = F(x + ct)$ , y  $\xi = G(x - ct)$  resuelve  $\partial\xi/\partial t + c\partial\xi/\partial x = 0$ , la cual se verifica rápidamente. Por superposición lineal una solución general de la ecuación (7.4) es  $\psi = F(x + ct) + G(x - ct)$ . Para funciones periódicas  $F, G$  reconocemos los argumentos  $x + ct$  y  $x - ct$  como la fase de la onda plana o frente de ondas, donde las soluciones de la ecuación de onda (7.4) cambia abruptamente (de cero a sus valores actuales) y no están únicamente determinadas. Normal al frente de onda están los rayos de la óptica geométrica. De este modo, las soluciones de la ecuación (7.5) o (7.3) más generalmente, son llamadas *características* o algunas veces *bicaracterísticas* (para PDE de segundo orden) en la literatura matemática corresponde a los frente de ondas de la solución de la óptica geométrica de la ecuación de onda completa.

Para el caso elíptico consideremos la ecuación de Laplace

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = 0, \quad (7.6)$$

para un potencial  $\psi$  de dos variables. Aquí la ecuación característica es

$$\left(\frac{\partial\xi}{\partial x}\right)^2 + \left(\frac{\partial\xi}{\partial y}\right)^2 = \left(\frac{\partial\xi}{\partial x} + i\frac{\partial\xi}{\partial y}\right) \left(\frac{\partial\xi}{\partial x} - i\frac{\partial\xi}{\partial y}\right) = 0 \quad (7.7)$$

tiene soluciones complejas conjugada:  $\xi = F(x + iy)$  para  $\partial\xi/\partial x + i\partial\xi/\partial y = 0$  y  $\xi = G(x - iy)$  para  $\partial\xi/\partial x - i\partial\xi/\partial y = 0$ . Una solución general de la ecuación de potencial (7.6) es por lo tanto  $\psi = F(x + iy) + G(x - iy)$ , tanto la parte real como la imaginaria de  $\psi$ , las cuales son llamadas *funciones armónicas*, mientras que las soluciones polinomiales son llamadas *polinomios armónicos*.

En mecánica cuántica la forma de Wentzel-Kramers-Brillouin (WKB) de  $\psi = \exp(-iS/\hbar)$  para la solución de la ecuación de Schrödinger

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi = i\hbar\frac{\partial\psi}{\partial t}, \quad (7.8)$$

conduce a la ecuación Hamilton-Jacobi de la mecánica clásica,

$$\frac{1}{2m}(\vec{\nabla}S)^2 + V = \frac{\partial S}{\partial t}, \quad (7.9)$$

en el límite  $\hbar \rightarrow 0$ . La acción clásica de  $S$  entonces llega a ser la característica de la ecuación de Schrödinger. Sustituyendo  $\vec{\nabla}\psi = -i\psi\vec{\nabla}S/\hbar$ ,  $\partial\psi/\partial t = -i\psi\partial S/\partial t/\hbar$  en la ecuación (7.8), dejando la totalidad de los factores de  $\psi$  no nulos, y aproximando  $\nabla^2\psi = -i\psi\nabla^2S/\hbar - \psi(\nabla S)^2/\hbar^2 \simeq -\psi(\nabla S)^2/\hbar^2$ , *i.e.*, despreciando  $-i\nabla^2\psi/\hbar$ , realmente obtenemos la ecuación (7.9).

Encontrando las soluciones de PDE por resolver las características es uno de varias técnicas generales. Para más ejemplos y tratamientos detallados de las características, las cuales no perseguimos aquí, nos referimos a H. Bateman, *Partial Differential Equations of Mathematical Physics*. New York: Dover (1994); K.E. Gustafson, *Partial Differential Equations and Hilbert Space Methods*, 2nd ed. New York: Wiley (1987).

### 7.1.3 Las PDE no lineales.

Las ODEs y PDEs no lineales son un campo importante y de rápido crecimiento. Encontramos más arriba la ecuación de onda lineal más simple

$$\frac{\partial \psi}{\partial t} + c \frac{\partial \psi}{\partial x} = 0 ,$$

como la PDE de primer orden a partir de la característica de la ecuación de onda. La ecuación de onda no lineal más simple

$$\frac{\partial \psi}{\partial t} + c(\psi) \frac{\partial \psi}{\partial x} = 0 , \quad (7.10)$$

resulta si la velocidad local de propagación,  $c$ , no es constante sino que depende de la onda  $\psi$ . Cuando una ecuación no lineal tiene una solución de la forma  $\psi(x, t) = A \cos(kx - \omega t)$  donde  $\omega(k)$  varía con  $k$  tal que  $\omega''(k) \neq 0$ , entonces ella es llamada *dispersiva*. Quizás la ecuación dispersiva no lineal más conocida de segundo orden es la ecuación de Korteweg-de Vries

$$\frac{\partial \psi}{\partial t} + \psi \frac{\partial \psi}{\partial x} + \frac{\partial^3 \psi}{\partial x^3} = 0 , \quad (7.11)$$

la cual modela la propagación sin pérdidas de las ondas de agua superficiales y otros fenómenos. Esta es ampliamente conocida por sus soluciones *solitón*. Un solitón es una onda viajera con la propiedad de persistir a través de una interacción con otro solitón: Después de que ellos pasan uno a través del otro, ellos emergen en la misma forma y con la misma velocidad y no adquieren más que un cambio de fase. Sea  $\psi(\xi = x - ct)$  tal onda viajera. Cuando es sustituida en la ecuación (7.11) esta produce la ODE no lineal

$$(\psi - c) \frac{d\psi}{d\xi} + \frac{d^3 \psi}{d\xi^3} = 0 , \quad (7.12)$$

la cual puede ser integrada dando

$$\frac{d^2 \psi}{d\xi^2} = c\psi - \frac{\psi^2}{2} . \quad (7.13)$$

No hay constantes de integración aditivas en la ecuación (7.13) para asegurar que  $d^2 \psi / d\xi^2 \rightarrow 0$  con  $\psi \rightarrow 0$  para  $\xi$  grande, tal que  $\psi$  está localizado en la característica  $\xi = 0$ , o  $x = ct$ . Multiplicando la ecuación (7.13) por  $d\psi/d\xi$  e integrando nuevamente tenemos

$$\left( \frac{d\psi}{d\xi} \right)^2 = c\psi^2 - \frac{\psi^3}{3} , \quad (7.14)$$

donde  $d\psi/d\xi \rightarrow 0$  para  $\xi$  grande. Tomando la raíz de la ecuación (7.14) e integrando una vez más encontramos la solución solitónica

$$\psi(x - ct) = \frac{3c}{\cosh^2 \left( \sqrt{c} \frac{x - ct}{2} \right)} . \quad (7.15)$$

### 7.1.4 Condiciones de borde.

Usualmente, cuando conocemos un sistema físico en algún momento y la ley que rigen ese proceso físico, entonces somos capaces de predecir el desarrollo subsecuente. Tales valores iniciales son las más comunes condiciones de borde asociadas con ODEs y PDEs. Encontrando soluciones que calcen con los puntos, curvas, o superficies dados correspondientes al problema de valores de contorno. Las autofunciones usualmente requieren que satisfagan ciertas condiciones de borde impuestas (*e.g.*, asintóticas). Estas condiciones pueden ser tomadas de tres formas:

1. Condiciones de borde de Cauchy. El valor de una función y su derivada normal especificada en el borde. En electrostática estas significarían  $\varphi$ , el potencial, y  $E_n$  la componente normal del campo eléctrico.
2. Condiciones de borde de Dirichlet. El valor específico en el borde.
3. Condiciones de borde de Neumann. La derivada normal (gradiente normal) de una función específica en el borde. En el caso electrostático este sería  $E_n$  y por lo tanto  $\sigma$ , la densidad de carga superficial.

Un resumen de las relaciones de estos tres tipos de condiciones de borde con los tres tipos de ecuaciones diferenciales parciales bidimensionales están dadas en la tabla 7.1. Para discusiones más extensas de estas ecuaciones diferenciales parciales puede consultar Sommerfeld, capítulo 2, o Morse y Feshbach, capítulo 6.

Partes de la tabla 7.1 son simplemente un asunto de mantener la consistencia interna, o sentido común. Por ejemplo, para la ecuación de Poisson con una superficie cerrada, las condiciones de Dirichlet conducen a una solución única y estable. Las condiciones de Neumann, independiente de las condiciones de Dirichlet, del mismo modo conducen a una solución única y estable independiente de la solución de Dirichlet. Por lo tanto las condiciones de borde de Cauchy (lo que significa la de Dirichlet más la de Neumann) conducen a una inconsistencia.

El término de condiciones de borde incluye como un caso especial el concepto de condiciones iniciales. Por ejemplo, especificando la posición inicial  $x_0$  y la velocidad inicial  $v_0$  en algunos problemas de dinámica correspondería a condiciones de borde de Cauchy. La única diferencia en el presente uso de las condiciones de borde en estos problemas unidimensionales es que estamos aplicando las condiciones en *ambos* extremos del intervalo permitido de la variable.

## 7.2 Ecuaciones diferenciales de primer orden.

La física involucra algunas ecuaciones diferenciales de primer orden, ellas fueron estudiadas en el primer curso. Por completitud parece ser deseable revisarlas brevemente.

Consideremos aquí ecuaciones diferenciales de la forma general

$$\frac{dy}{dx} = f(x, y) = -\frac{P(x, y)}{Q(x, y)}. \quad (7.16)$$

Condiciones de borde	Tipo de ecuación diferencial parcial		
	Elípticas	Hiperbólicas	Parabólicas
	Laplace, Poisson en $(x, y)$	Ecuación de Ondas en $(x, t)$	Ecuación de difusión en $(x, t)$
Cauchy			
Superficie Abierta	Resultados no físicos (inestabilidades)	<i>Solución única y estable</i>	Demasiado restrictivo
Superficie Cerrada	Demasiado restrictivo	Demasiado restrictivo	Demasiado restrictivo
Dirichlet			
Superficie Abierta	Insuficiente	Insuficiente	<i>Solución única y estable en 1 dim</i>
Superficie Cerrada	<i>Solución única y estable</i>	Solución no única	Demasiado restrictivo
Neumann			
Superficie Abierta	Insuficiente	Insuficiente	<i>Solución única y estable en 1 dim</i>
Superficie Cerrada	<i>Solución única y estable</i>	Solución no única	Demasiado restrictivo

Tabla 7.1:

La ecuación (7.16) es claramente una ecuación de primer orden ordinaria. Es de *primer orden* ya que contiene la primera derivada y no mayores. Es *Ordinaria* ya que la derivada  $dy/dx$  es una derivada ordinaria o total. La ecuación (7.16) puede o no puede ser *lineal*, aunque trataremos el caso lineal explícitamente más adelante.

### 7.2.1 Variables separables.

Frecuentemente la ecuación (7.16) tendrá la forma especial

$$\frac{dy}{dx} = f(x, y) = -\frac{P(x)}{Q(y)}. \quad (7.17)$$

Entonces la podemos reescribir como

$$P(x)dx + Q(y)dy = 0.$$

Integrando de  $(x_0, y_0)$  a  $(x, y)$  tiende a

$$\int_{x_0}^x P(x')dx' + \int_{y_0}^y Q(y')dy' = 0. \quad (7.18)$$

Ya que los límites inferiores  $x_0$  e  $y_0$  contribuyen en unas constantes, podríamos ignorar los límites inferiores de integración y simplemente añadir una constante de integración al final. Note que esta técnica de separación de variables *no* requiere que la ecuación diferencial sea lineal.

### Ejemplo Ley de Boyle.

Una forma diferencial de la ley de los gases de Boyle es

$$\frac{dV}{dP} = -\frac{V}{P},$$

para el volumen  $V$  de una cantidad fija de gas a presión  $P$  (y temperatura constante). Separando variables, tenemos

$$\frac{dV}{V} = -\frac{dP}{P}$$

o

$$\ln V = -\ln P + C.$$

Con dos logaritmos presentes, es más conveniente reescribir la constante de integración  $C$  como  $\ln k$ . Entonces

$$\ln V + \ln P = \ln PV = \ln k$$

y

$$PV = k.$$

### 7.2.2 Ecuaciones diferenciales exactas.

Reescribimos la ecuación (7.16) como

$$P(x, y)dx + Q(x, y)dy = 0. \quad (7.19)$$

Esta ecuación se dice que es *exacta* si podemos calzar el lado izquierdo de ella a un diferencial  $d\varphi$ ,

$$d\varphi = \frac{\partial\varphi}{\partial x}dx + \frac{\partial\varphi}{\partial y}dy. \quad (7.20)$$

Ya que la ecuación (7.19) tiene un cero a la derecha, buscamos una función desconocida  $\varphi(x, y) = \text{constante}$  tal que  $d\varphi = 0$ . Tenemos (si tal función  $\varphi(x, y)$  existe)

$$P(x, y)dx + Q(x, y)dy = \frac{\partial\varphi}{\partial x}dx + \frac{\partial\varphi}{\partial y}dy \quad (7.21)$$

y

$$\frac{\partial \varphi}{\partial x} = P(x, y), \quad \frac{\partial \varphi}{\partial y} = Q(x, y). \quad (7.22)$$

La condición necesaria y suficiente para que la ecuación sea exacta es que la segunda derivada parcial mezclada de  $\varphi(x, y)$  (supuesta continua) es independiente del orden de diferenciación:

$$\frac{\partial^2 \varphi}{\partial y \partial x} = \frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x} = \frac{\partial^2 \varphi}{\partial x \partial y}. \quad (7.23)$$

Si la ecuación (7.19) corresponde a un rotor (igual cero), entonces un potencial,  $\varphi(x, y)$ , debiera existir.

Si  $\varphi(x, y)$  existe entonces a partir de las ecuaciones (7.19) y (7.21) nuestra solución es

$$\varphi(x, y) = C. \quad (7.24)$$

Podemos construir  $\varphi(x, y)$  a partir de sus derivadas parciales de la misma manera que construimos un potencial magnético vectorial en el capítulo de vectores a partir de su rotor.

Podemos volver a la ecuación (7.19) y ver que pasa si no es exacta, la ecuación (7.23) no es satisfecha. Sin embargo, siempre existe al menos una o quizás una infinidad de factores de integración,  $\alpha(x, y)$ , tales que

$$\alpha(x, y)P(x, y)dx + \alpha(x, y)Q(x, y)dy = 0$$

es exacta. Desafortunadamente, un factor de integración no siempre es obvio o fácil de encontrar. Diferente es el caso de la ecuación diferencial de primer orden *lineal* considerada a continuación, no hay una manera sistemática de desarrollar un factor de integración para la ecuación (7.19).

Una ecuación diferencial en la cual las variables han sido separadas es automáticamente exacta. Una ecuación diferencial exacta *no* es necesariamente separable.

### 7.2.3 Ecuaciones diferenciales ordinarias de primer orden lineales.

Si  $f(x, y)$  en la ecuación (7.16) tiene la forma  $-p(x)y + q(x)$ , entonces la ecuación (7.16) se convierte en

$$\frac{dy}{dx} + p(x)y = q(x). \quad (7.25)$$

La ecuación (7.25) es la ODE de primer orden *lineal* más general. Si  $q(x) = 0$ , la ecuación (7.25) es homogénea (en  $y$ ). Un  $q(x)$  distinto de cero puede representar una fuente o un término de forzamiento. La ecuación (7.25) es *lineal*; cada término es lineal en  $y$  o  $dy/dx$ . No hay potencias mayores; esto es, no hay  $y^2$ , ni productos,  $y(dy/dx)$ . Note que la linealidad se refiere a el  $y$  y a la  $dy/dx$ ;  $p(x)$  y  $q(x)$  no es necesario que sean lineales en  $x$ . La ecuación (7.25), es la más importante de estas ecuaciones diferenciales de primer orden para los físicos y puede ser resuelta exactamente.

Busquemos un *factor de integración*  $\alpha(x)$  tal que

$$\alpha(x) \frac{dy}{dx} + \alpha(x)p(x)y = \alpha(x)q(x) , \quad (7.26)$$

puede ser reescrito como

$$\frac{d}{dx}[\alpha(x)y] = \alpha(x)q(x) . \quad (7.27)$$

El propósito de esto es hacer el lado izquierdo de la ecuación (7.25) una derivada total que pueda ser integrada por inspección. Esto también, incidentalmente, hace la ecuación (7.25) exacta. Expandiendo la ecuación (7.27), obtenemos

$$\alpha(x) \frac{dy}{dx} + \frac{d\alpha}{dx}y = \alpha(x)q(x) .$$

La comparación con la ecuación (7.26) muestra que debemos requerir que

$$\frac{d\alpha(x)}{dx} = \alpha(x)p(x) . \quad (7.28)$$

Aquí hay una ecuación diferencial para  $\alpha(x)$ , con las variables  $\alpha$  y  $x$  *separables*. Separamos variables, integramos, y obtenemos

$$\alpha(x) = \exp \left[ \int^x p(x') dx' \right] \quad (7.29)$$

como nuestro factor de integración.

Con  $\alpha(x)$  conocida procedemos a integrar la ecuación (7.27). Esto, por supuesto, fue el objetivo de introducir  $\alpha$  en primer lugar. Tenemos

$$\int^x \frac{d}{dx'}[\alpha(x')y] dx' = \int^x \alpha(x')q(x') dx' .$$

Ahora integrando por inspección, tenemos

$$\alpha(x)y = \int^x \alpha(x')q(x') dx' + C .$$

Las constantes a partir del límite inferior de integración constante son reunidas en la constante  $C$ . Dividiendo por  $\alpha(x)$ , obtenemos

$$y(x) = \frac{1}{\alpha(x)} \left\{ \int^x \alpha(x')q(x') dx' + C \right\} .$$

Finalmente, sustituyendo en la ecuación (7.29) por  $\alpha$  conduce

$$y(x) = \exp \left[ - \int^x p(t) dt \right] \left\{ \int^x \exp \left[ \int^s p(t) dt \right] q(s) ds + C \right\} . \quad (7.30)$$

Aquí las variables mudas de integración han sido reescritas para hacerlas inambiguas. La ecuación (7.30) es la solución general completa de la ecuación diferencial lineal, de primer orden, la ecuación (7.25). La porción

$$y_1(x) = C \exp \left[ - \int^x p(t) dt \right] \quad (7.31)$$

corresponde al caso  $q(x) = 0$  y es solución general de la ecuación diferencial homogénea. El otro término en la ecuación (7.30),

$$y(x) = \exp \left[ - \int^x p(t) dt \right] \int^x \exp \left[ \int^s p(t) dt \right] q(s) ds , \quad (7.32)$$

es una solución particular que corresponde al término específico de fuente  $q(x)$ .

Podemos notar que si nuestra ecuación diferencial de primer orden es homogénea ( $q = 0$ ), entonces ella es separable. De lo contrario, salvo casos especiales tal como  $p = \text{constante}$ ,  $q = \text{constante}$ , o  $q(x) = ap(x)$ , la ecuación (7.25) no es separable.

### Ejemplo Circuito RL.

Para un circuito resistencia-inductancia las leyes de Kirchhoff producen a

$$L \frac{dI(t)}{dt} + RI(t) = V(t) ,$$

para la corriente  $I(t)$ , donde  $L$  es la inductancia y  $R$  es la resistencia, ambas constantes.  $V(t)$  es el voltaje aplicado tiempo dependiente.

De la ecuación (7.29) nuestro factor de integración  $\alpha(t)$  es

$$\begin{aligned} \alpha(t) &= \exp \int^t \frac{R}{L} dt \\ &= e^{Rt/L} . \end{aligned}$$

Entonces por la ecuación (7.30)

$$I(t) = e^{-Rt/L} \left[ \int^t e^{Rt/L} \frac{V(t)}{L} dt + C \right] ,$$

con la constante  $C$  es determinada por una condición inicial (una condición de borde).

Para el caso especial  $V(t) = V_0$ , una constante,

$$\begin{aligned} I(t) &= e^{-Rt/L} \left[ \frac{V_0}{L} \frac{L}{R} e^{Rt/L} + C \right] \\ &= \frac{V_0}{R} + C e^{-Rt/L} . \end{aligned}$$

Si la condición inicial es  $I(0) = 0$ , entonces  $C = -V_0/R$  y

$$I(t) = \frac{V_0}{R} [1 - e^{-Rt/L}] .$$

### 7.2.4 Conversión a una ecuación integral.

Nuestra ecuación diferencial de primer orden, ecuación (7.16), puede ser convertida a una ecuación integral por integración directa:

$$y(x) - y(x_0) = \int_{x_0}^x f[x, y(x)] dx . \quad (7.33)$$

Como una ecuación integral hay una posibilidad de una solución en serie de Neumann (se verá en el próximo curso) con la aproximación inicial  $y(x) \approx y(x_0)$ . En la literatura de ecuaciones diferenciales esto es llamado el “método de Picard de aproximaciones sucesivas”.

Ecuaciones diferenciales de primer orden las encontraremos de nuevo en conexión con las transformadas de Laplace y de Fourier.

## 7.3 Separación de variables.

Las ecuaciones de la física matemática listada en la sección 7.1 son todas ecuaciones diferenciales parciales. Nuestra primera técnica para su solución es dividir la ecuación diferencial parcial en  $n$  ecuaciones diferenciales ordinarias de  $n$  variables. Cada separación introduce una constante de separación arbitraria. Si tenemos  $n$  variables, tenemos que introducir  $n - 1$  constantes, determinadas por las condiciones impuestas al resolver el problema.

### 7.3.1 Coordenadas cartesianas.

En coordenadas cartesianas las ecuaciones de Helmholtz llegan a ser

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + k^2 \psi = 0 , \quad (7.34)$$

usando la forma cartesiana para el Laplaciano. Por el momento  $k^2$  será una constante. Quizás la manera más simple de tratar una ecuación diferencial parcial tal como la ecuación (7.34) es dividirla en un conjunto de ecuaciones diferenciales ordinarias. Esto puede ser hecho como sigue. Sea

$$\psi(x, y, z) = X(x)Y(y)Z(z) , \quad (7.35)$$

y sustituir de vuelta en la ecuación (7.34). ¿Cómo sabemos que la ecuación (7.35) es válida?. La respuesta es muy simple. ¡No sabemos si es válida!. Mejor dicho, estamos procediendo en este espíritu y tratando a ver si trabaja. Si nuestro intento es exitoso, entonces la ecuación (7.35) será justificada. Si no es exitoso, lo descubriremos pronto y luego trataremos otro ataque tal como las funciones de Green, transformadas integral, o análisis numérico a la fuerza bruta. Con  $\psi$  supuestamente dada por la ecuación (7.35), la ecuación (7.34) llega a ser

$$YZ \frac{d^2 X}{dx^2} + XZ \frac{d^2 Y}{dy^2} + XY \frac{d^2 Z}{dz^2} + k^2 XYZ = 0 . \quad (7.36)$$

Dividiendo por  $\psi = XYZ$  y reorganizando los términos, obtenemos

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -k^2 - \frac{1}{Y} \frac{d^2 Y}{dy^2} - \frac{1}{Z} \frac{d^2 Z}{dz^2} . \quad (7.37)$$

La ecuación (7.37) exhibe una separación de variables. El lado izquierdo es sólo función de  $x$ , mientras que el lado derecho depende solamente de  $y$  y  $z$ . Así la ecuación (7.37) es una clase de paradoja. Una función de  $x$  es igualada a una función de  $y$  y  $z$ , pero  $x$ ,  $y$  y  $z$  son todas coordenadas independientes. Esta independencia significa que el comportamiento de  $x$  como una variable independiente no está determinada ni por  $y$  ni por  $z$ . La paradoja está resuelta fijando cada lado igual a una constante, una constante de separación. Escogemos<sup>3</sup>

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -l^2 , \quad (7.38)$$

$$-k^2 - \frac{1}{Y} \frac{d^2 Y}{dy^2} - \frac{1}{Z} \frac{d^2 Z}{dz^2} = -l^2 . \quad (7.39)$$

Ahora, volviendo nuestra atención a la ecuación (7.39), obtenemos

$$\frac{1}{Y} \frac{d^2 Y}{dy^2} = -k^2 + l^2 - \frac{1}{Z} \frac{d^2 Z}{dz^2} , \quad (7.40)$$

y una segunda separación ha sido realizada. Aquí tenemos una función de  $y$  igualada a una función de  $z$  y aparece la misma paradoja. La resolvemos como antes igualando cada lado a otra constante de separación,  $-m^2$ ,

$$\frac{1}{Y} \frac{d^2 Y}{dy^2} = -m^2 , \quad (7.41)$$

$$\frac{1}{Z} \frac{d^2 Z}{dz^2} = -k^2 + l^2 + m^2 = -n^2 , \quad (7.42)$$

introduciendo una constante  $n^2$  por  $k^2 = l^2 + m^2 + n^2$  para producir un conjunto simétrico de ecuaciones. Ahora tenemos tres ecuaciones diferenciales ordinarias ((7.38), (7.41), y (7.42)) para reemplazar en la ecuación (7.34). Nuestra suposición (ecuación (7.35)) ha sido exitosa y es por lo tanto justificada.

Nuestra solución sería etiquetada de acuerdo a la elección de nuestras constantes  $l$ ,  $m$ ,  $n$ , esto es,

$$\psi_{lmn}(x, y, z) = X_l(x)Y_m(y)Z_n(z) . \quad (7.43)$$

Sujeto a las condiciones del problema que se resuelve y a la condición  $k^2 = l^2 + m^2 + n^2$ , podemos escoger  $l$ ,  $m$ ,  $n$  como queramos, y la ecuación (7.43) será todavía una solución de la ecuación (7.34), dado que  $X_l(x)$  es una solución de la ecuación (7.38) y así seguimos.

<sup>3</sup>La elección de signo es completamente arbitraria, será fijada en un problema específico por la necesidad de satisfacer las condiciones de borde.

Podemos desarrollar la solución más general de la ecuación (7.34) tomando una combinación lineal de soluciones  $\psi_{lmn}$ ,

$$\Psi = \sum_{l,m,n} a_{lmn} \psi_{lmn} . \quad (7.44)$$

Los coeficientes constantes  $a_{lmn}$  finalmente son escogidos para permitir que  $\Psi$  satisfaga las condiciones de borde del problema.

### 7.3.2 Coordenadas cilíndricas circulares.

Si consideramos que nuestra función desconocida  $\psi$  depende de  $\rho$ ,  $\varphi$ ,  $z$  la ecuación de Helmholtz se convierte en

$$\nabla^2 \psi(\rho, \varphi, z) + k^2 \psi(\rho, \varphi, z) = 0 , \quad (7.45)$$

o

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial \psi}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 \psi}{\partial \varphi^2} + \frac{\partial^2 \psi}{\partial z^2} + k^2 \psi = 0 . \quad (7.46)$$

Como antes, suponemos una forma factorizada para  $\psi$ ,

$$\psi(\rho, \varphi, z) = P(\rho) \Phi(\varphi) Z(z) . \quad (7.47)$$

Sustituyendo en la ecuación (7.46), tenemos

$$\frac{\Phi Z}{\rho} \frac{d}{d\rho} \left( \rho \frac{dP}{d\rho} \right) + \frac{PZ}{\rho^2} \frac{d^2 \Phi}{d\varphi^2} + P\Phi \frac{d^2 Z}{dz^2} + k^2 P\Phi Z = 0 . \quad (7.48)$$

Todas las derivadas parciales han llegado a ser derivadas ordinarias. Dividiendo por  $P\Phi Z$  y moviendo la derivada  $z$  al lado derecho conduce a

$$\frac{1}{P\rho} \frac{d}{d\rho} \left( \rho \frac{dP}{d\rho} \right) + \frac{1}{\rho^2 \Phi} \frac{d^2 \Phi}{d\varphi^2} + k^2 = -\frac{1}{Z} \frac{d^2 Z}{dz^2} . \quad (7.49)$$

De nuevo, tenemos la paradoja. Una función de  $z$  en la derecha aparece dependiendo de una función de  $\rho$  y  $\varphi$  en el lado izquierdo. Resolvemos la paradoja haciendo cada lado de la ecuación (7.49) igual a una constante, la misma constante. Escojamos<sup>4</sup>  $-l^2$ . Entonces

$$\frac{d^2 Z}{dz^2} = l^2 Z , \quad (7.50)$$

y

$$\frac{1}{P\rho} \frac{d}{d\rho} \left( \rho \frac{dP}{d\rho} \right) + \frac{1}{\rho^2 \Phi} \frac{d^2 \Phi}{d\varphi^2} + k^2 = -l^2 . \quad (7.51)$$

---

<sup>4</sup>La elección del signo de la constante de separación es arbitraria. Sin embargo, elegimos un signo menos para la coordenada axial  $z$  en espera de una posible dependencia exponencial en  $z$ . Un signo positivo es elegido para la coordenada azimutal  $\varphi$  en espera de una dependencia periódica en  $\varphi$ .

Ajustando  $k^2 + l^2 = n^2$ , multiplicando por  $\rho^2$ , y reordenando términos, obtenemos

$$\frac{\rho}{P} \frac{d}{d\rho} \left( \rho \frac{dP}{d\rho} \right) + n^2 \rho^2 = -\frac{1}{\Phi} \frac{d^2 \Phi}{d\varphi^2} . \quad (7.52)$$

Podemos ajustar el lado derecho a  $m^2$  y

$$\frac{d^2 \Phi}{d\varphi^2} = -m^2 \Phi \quad (7.53)$$

Finalmente, para la dependencia en  $\rho$  tenemos

$$\rho \frac{d}{d\rho} \left( \rho \frac{dP}{d\rho} \right) + (n^2 \rho^2 - m^2) P = 0 . \quad (7.54)$$

Esta es la ecuación diferencial de Bessel. La solución y sus propiedades serán presentadas en el próximo curso. La separación de variables de la ecuación de Laplace en coordenadas parabólicas también conduce a ecuaciones de Bessel. Puede notarse que la ecuación de Bessel es notable por la variedad de formas que puede asumir.

La ecuación original de Helmholtz, una ecuación diferencial parcial tridimensional, ha sido reemplazada por tres ecuaciones diferenciales ordinarias, las ecuaciones (7.50), (7.53) y (7.54). Una solución de la ecuación de Helmholtz es

$$\psi(\rho, \varphi, z) = P(\rho)\Phi(\varphi)Z(z) . \quad (7.55)$$

Identificando las soluciones específicas  $P$ ,  $\Phi$ ,  $Z$  por subíndices, vemos que la solución más general de la ecuación de Helmholtz es una combinación lineal del producto de soluciones:

$$\psi(\rho, \varphi, z) = \sum_{m,n} a_{mn} P_{mn}(\rho)\Phi_m(\varphi)Z_n(z) . \quad (7.56)$$

### 7.3.3 Coordenadas polares esféricas.

Tratemos de separar la ecuación de Helmholtz, de nuevo con  $k^2$  constante, en coordenadas polares esféricas. Usando la expresión del Laplaciano en estas coordenadas obtenemos

$$\frac{1}{r^2 \sin \theta} \left[ \sin \theta \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin \theta} \frac{\partial^2 \psi}{\partial \varphi^2} \right] = -k^2 \psi . \quad (7.57)$$

Ahora, en analogía con la ecuación (7.35) tratamos

$$\psi(r, \theta, \varphi) = R(r)\Theta(\theta)\Phi(\varphi) . \quad (7.58)$$

Sustituyendo de vuelta en la ecuación (7.57) y dividiendo por  $R\Theta\Phi$ , tenemos

$$\frac{1}{Rr^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \frac{1}{\Theta r^2 \sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{d\Theta}{d\theta} \right) + \frac{1}{\Phi r^2 \sin^2 \theta} \frac{d^2 \Phi}{d\varphi^2} = k^2 . \quad (7.59)$$

Note que todas las derivadas son ahora derivadas ordinarias más que parciales. Multiplicando por  $r^2 \text{sen}^2 \theta$ , podemos aislar  $(1/\Phi)(d^2\Phi/d\varphi^2)$  para obtener<sup>5</sup>

$$\frac{1}{\Phi} \frac{d^2\Phi}{d\varphi^2} = r^2 \text{sen}^2 \theta \left[ -k^2 - \frac{1}{r^2 R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) - \frac{1}{r^2 \text{sen} \theta \Theta} \frac{d}{d\theta} \left( \text{sen} \theta \frac{d\Theta}{d\theta} \right) \right]. \quad (7.60)$$

La ecuación (7.60) relaciona una función únicamente de  $\varphi$  con una función de  $r$  y  $\theta$ . Ya que  $r$ ,  $\theta$ , y  $\varphi$  son variables independientes, igualamos cada lado de la ecuación (7.60) a una constante. Aquí una pequeña consideración puede simplificar el análisis posterior. En casi todos los problemas físicos  $\varphi$  aparecerá como un ángulo azimutal. Esto sugiere una solución periódica más que una exponencial. Con esto en mente, usemos  $-m^2$  como la constante de separación. Cualquier constante lo hará, pero ésta hará la vida un poquito más fácil. Entonces

$$\frac{1}{\Phi} \frac{d^2\Phi}{d\varphi^2} = -m^2 \quad (7.61)$$

y

$$\frac{1}{r^2 R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \frac{1}{r^2 \text{sen} \theta \Theta} \frac{d}{d\theta} \left( \text{sen} \theta \frac{d\Theta}{d\theta} \right) - \frac{m^2}{r^2 \text{sen}^2 \theta} = -k^2. \quad (7.62)$$

Multiplicando la ecuación (7.62) por  $r^2$  y reordenando términos, tenemos

$$\frac{1}{R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + r^2 k^2 = -\frac{1}{\text{sen} \theta \Theta} \frac{d}{d\theta} \left( \text{sen} \theta \frac{d\Theta}{d\theta} \right) + \frac{m^2}{\text{sen}^2 \theta}. \quad (7.63)$$

Nuevamente, las variables son separadas. Igualamos cada lado a una constante  $Q$  y finalmente obtenemos

$$\frac{1}{\text{sen} \theta} \frac{d}{d\theta} \left( \text{sen} \theta \frac{d\Theta}{d\theta} \right) - \frac{m^2}{\text{sen}^2 \theta} \Theta + Q\Theta = 0, \quad (7.64)$$

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + k^2 R - \frac{QR}{r^2} = 0. \quad (7.65)$$

Una vez más hemos reemplazado una ecuación diferencial parcial de tres variables por tres ecuaciones diferenciales ordinarias. Las soluciones de estas tres ecuaciones diferenciales ordinarias son discutidas en el próximo curso. Por ejemplo, la ecuación (7.64) es identificada como la ecuación de asociada de Legendre en la cual la constante  $Q$  llega a ser  $l(l+1)$ ; con  $l$  entero. Si  $k^2$  es una constante (positiva), la ecuación (7.65) llega a ser la ecuación de Bessel esférica.

Nuevamente, nuestra solución más general puede ser escrita

$$\psi_{Qm}(r, \theta, \varphi) = \sum_{q,m} R_Q(r) \Theta_{Qm}(\theta) \Phi_m(\varphi). \quad (7.66)$$

---

<sup>5</sup>El orden en el cual las variables son separadas aquí no es único. Muchos textos de mecánica cuántica separan la dependencia en  $r$  primero.

La restricción que  $k^2$  sea una constante es innecesariamente severa. El proceso de separación será todavía posible para  $k^2$  tan general como

$$k^2 = f(r) + \frac{1}{r^2}g(\theta) + \frac{1}{r^2 \sin^2 \theta}h(\varphi) + k'^2 . \quad (7.67)$$

En el problema del átomo de hidrógeno, uno de los ejemplos más importantes de la ecuación de onda de Schrödinger con una forma cerrada de solución es  $k^2 = f(r)$ . La ecuación (7.65) para el átomo de hidrógeno llega a ser la ecuación asociada de Laguerre.

La gran importancia de esta separación de variables en coordenadas polares esféricas deriva del hecho que el caso  $k^2 = k^2(r)$  cubre una tremenda cantidad de física: las teorías de gravitación, electrostática, física atómica y física nuclear. Y, con  $k^2 = k^2(r)$ , la dependencia angular es aislada en las ecuaciones (7.61) y (7.64), *la cual puede ser resuelta exactamente*.

Finalmente, una ilustración de cómo la constante  $m$  en la ecuación (7.61) es restringida, notamos que  $\varphi$  en coordenadas polares esféricas y cilíndricas es un ángulo azimutal. Si esto es un problema clásico, ciertamente requeriremos que la solución azimutal  $\Phi(\varphi)$  sea univaluada, esto es,

$$\Phi(\varphi + 2\pi) = \Phi(\varphi) . \quad (7.68)$$

Esto es equivalente a requerir que la solución azimutal tenga un período de  $2\pi$  o algún múltiplo entero de él. Por lo tanto  $m$  debe ser un entero. Cuál entero, depende de los detalles del problema. Cada vez que una coordenada corresponda a un eje de translación o a un ángulo azimutal la ecuación separada siempre tendrá la forma

$$\frac{d^2\Phi(\varphi)}{d\varphi^2} = -m^2\Phi(\varphi)$$

para  $\varphi$ , el ángulo azimutal, y

$$\frac{d^2Z}{dz^2} = \pm a^2 Z(z) \quad (7.69)$$

para  $z$ , un eje de translación en un sistema de coordenadas cilíndrico. Las soluciones, por supuesto, son  $\sin az$  y  $\cos az$  para  $-a^2$  y la correspondiente función hiperbólica (o exponencial)  $\sinh az$  y  $\cosh az$  para  $+a^2$ .

Otras ecuaciones diferenciales ordinarias encontradas ocasionalmente incluyen las ecuaciones de Laguerre y la asociada de Laguerre del importante problema del átomo de hidrógeno en mecánica cuántica:

$$x \frac{d^2y}{dx^2} + (1-x) \frac{dy}{dx} + \alpha y = 0 , \quad (7.70)$$

$$x \frac{d^2y}{dx^2} + (1+k-x) \frac{dy}{dx} + \alpha y = 0 . \quad (7.71)$$

De la teoría de la mecánica cuántica del oscilador armónico lineal tenemos la ecuación de Hermite,

$$\frac{d^2y}{dx^2} - 2x \frac{dy}{dx} + 2\alpha y = 0 . \quad (7.72)$$

Finalmente, de vez en vez encontramos la ecuación diferencial de Chebyshev

$$(1 - x^2) \frac{d^2y}{dx^2} - x \frac{dy}{dx} + n^2y = 0 . \tag{7.73}$$

Para una referencia conveniente, las formas de la solución de la ecuación de Laplace, la ecuación de Helmholtz y la ecuación de difusión en coordenadas polares esféricas son resumidas en la tabla 7.2. Las soluciones de la ecuación de Laplace en coordenadas circulares cilíndricas son representadas en la tabla 7.3.

$\psi = \sum_{l,m} a_{lm} \psi_{lm}$		
1.	$\nabla^2 \psi = 0$	$\psi_{lm} = \left\{ \begin{array}{l} r^l \\ r^{-l-1} \end{array} \right\} \left\{ \begin{array}{l} P_l^m(\cos \theta) \\ Q_l^m(\cos \theta) \end{array} \right\} \left\{ \begin{array}{l} \cos m\varphi \\ \text{sen } m\varphi \end{array} \right\}$
2.	$\nabla^2 \psi + k^2 \psi = 0$	$\psi_{lm} = \left\{ \begin{array}{l} j_l(kr) \\ n_l(kr) \end{array} \right\} \left\{ \begin{array}{l} P_l^m(\cos \theta) \\ Q_l^m(\cos \theta) \end{array} \right\} \left\{ \begin{array}{l} \cos m\varphi \\ \text{sen } m\varphi \end{array} \right\}$
3.	$\nabla^2 \psi - k^2 \psi = 0$	$\psi_{lm} = \left\{ \begin{array}{l} i_l(kr) \\ k_l(kr) \end{array} \right\} \left\{ \begin{array}{l} P_l^m(\cos \theta) \\ Q_l^m(\cos \theta) \end{array} \right\} \left\{ \begin{array}{l} \cos m\varphi \\ \text{sen } m\varphi \end{array} \right\}$

Tabla 7.2: Soluciones en coordenadas polares esféricas

$\psi = \sum_{m,\alpha} a_{m\alpha} \psi_{m\alpha} , \quad \nabla^2 \psi = 0$		
a.		$\psi_{m\alpha} = \left\{ \begin{array}{l} J_m(\alpha\rho) \\ N_m(\alpha\rho) \end{array} \right\} \left\{ \begin{array}{l} \cos m\varphi \\ \text{sen } m\varphi \end{array} \right\} \left\{ \begin{array}{l} e^{-\alpha z} \\ e^{\alpha z} \end{array} \right\}$
b.		$\psi_{m\alpha} = \left\{ \begin{array}{l} I_m(\alpha\rho) \\ K_m(\alpha\rho) \end{array} \right\} \left\{ \begin{array}{l} \cos m\varphi \\ \text{sen } m\varphi \end{array} \right\} \left\{ \begin{array}{l} \cos \alpha z \\ \text{sen } \alpha z \end{array} \right\}$
c.	$\alpha = 0$ (no hay dependencia en $z$ )	$\psi_m = \left\{ \begin{array}{l} \rho^m \\ \rho^{-m} \end{array} \right\} \left\{ \begin{array}{l} \cos m\varphi \\ \text{sen } m\varphi \end{array} \right\}$

Tabla 7.3: Soluciones en coordenadas cilíndricas circulares

Para las ecuaciones de Helmholtz y de difusión la constante  $\pm k^2$  se agrega a la constante de separación  $\pm \alpha^2$  para definir un nuevo parámetro  $\gamma^2$  o  $-\gamma^2$ . Para la elección de  $+\gamma^2$  (con  $\gamma^2 > 0$ ) obtenemos  $J_m(\gamma\rho)$  y  $N_m(\gamma\rho)$ . Para la elección  $-\gamma^2$  (con  $\gamma^2 > 0$ ) obtenemos  $I_m(\gamma\rho)$  y  $K_m(\gamma\rho)$  como previamente.

Estas ecuaciones diferenciales ordinarias y sus generalizaciones serán examinadas y sistematizadas en el próximo curso