LARS V. AHLFORS
PROFESOR DE MATEMATICAS DE LA UNIVERSIDAD DE HARVARD

ANALISIS DE VARIABLE COMPLEJA

INTRODUCCION A LA TEORIA DE FUNCIONES ANALITICAS DE UNA VARIABLE COMPLEJA

Traducción del inglés por
A. PARDO FRAILE
Licenciado en Ciencias Matemáticas

AGUILAR

La presente obra, incorporada a este fondo editorial con el asesoramiento de don LUIS BRAVO GALA, *ha sido publicada originalmente en lengua inglesa, por la Casa McGraw-Hill Book Company, Inc., de Nueva York, Toronto y Londres, con el título de*

COMPLEX ANALYSIS

AN INTRODUCTION TO THE THEORY OF ANALYTIC FUNCTIONS OF ONE COMPLEX VARIABLE

y lleva la siguiente dedicatoria:

A LA MEMORIA DE
ERNST LINDELÖF

NÚM. RGTRO.: 1568 - 65.
DEPÓSITO LEGAL. M. 11929.—1966.

Printed in Spain. Impreso en España por Selecciones Gráficas, Avda. del Doctor Federico Rubio y Galí, 184, Madrid. Julio de 1966.

PREFACIO

PREFACIO

Un curso que aproximadamente recoge el material de este libro se da de ordinario en las universidades norteamericanas en el primer año de estudios. La forma de presentar los temas difiere ampliamente de unos centros a otros: en algunos se hace hincapié en la enseñanza de cierta cantidad indispensable de la teoría clásica de funciones; en otros, el curso se utiliza para que el estudiante tome contacto por vez primera con el rigor matemático. En Harvard, también se utiliza tradicionalmente el curso para hacer un repaso de los elementos del análisis infinitesimal, enfocado con completo rigor.

Ha sido ambición del autor escribir un texto que sea a la vez conciso y riguroso, legible y fácil de enseñar. Tal objetivo no es fácil de alcanzar de manera satisfactoria, y el autor no ignora que existen algunos defectos. No se ha hecho intento alguno para que el libro sea suficiente por sí mismo. Por el contrario, se supone al lector en posesión de conocimientos básicos sobre los números reales y el análisis, que incluyen la definición y primeras propiedades de la integral definida. Las cuestiones concernientes a límites y continuidad se repasarán en relación con su aplicación a los números complejos, y se han hecho esfuerzos para no basarse en resultados que en la enseñanza elemental se deducen de ordinario de forma poco rigurosa o insuficiente. Si el profesor decidiese que deberían incluirse en este curso los números reales o la definición de integral, tiene a su disposición una docena de textos de confianza que podría consultar. El autor ha omitido estos temas con objeto de que el libro no resultase demasiado voluminoso.

Aparte de elegir el punto de partida, al autor de un libro de texto se le presenta la difícil tarea de tener que decidir lo que se ha de incluir y lo que ha de omitirse. En el caso presente, se ha deseado proporcionar al lector una base sólida de la teoría clásica de funciones de variable compleja, haciendo hincapié en los resultados generales en que esta se apoya. Se ha creído que una persona a la que se hayan inculcado concienzudamente los métodos fundamentales, no experimentará nuevas dificultades si desea proseguir sus estudios sobre algún punto especializado de la teoría de funciones. Sin em-

bargo, el autor ha tenido que omitir, p. ej., con gran sentimiento por su parte, la teoría de las funciones elípticas. Una de las razones principales para haber procedido así es la de que difícilmente cabe mejorar el magnífico tratamiento de dicha cuestión que aparece en el libro de E. T. Copson "An Introduction to the Theory of Functions of a Complex Variable" *(Londres, 1935).*

Por el contrario, se han incluido algunos temas que normalmente no se consideran como parte de la teoría elemental de funciones. Tal es el caso de la teoría de las funciones subarmónicas y el método de Perron para resolver el problema de Dirichlet, que, ciertamente, son tan elementales como importantes.

El libro empieza con un análisis elemental de los números complejos y acaba con una breve nota sobre la teoría de las funciones analíticas multiformes. El desarrollo de la exposición se hace de modo gradual. Siguiendo a su venerado maestro, Ernst Lindelöf, el autor pospone el uso de la integración compleja hasta que el estudiante se halla completamente familiarizado con las propiedades de las aplicaciones a que dan lugar las funciones analíticas. La intuición geométrica es una fuente de conocimiento tanto como un instrumento didáctico, cuyo valor no puede ser discutido.

Deben hacerse constar otros antecedentes. Así, p. ej., la aparición del libro Funktionentheorie, *de Carathéodory, no ha dejado, por supuesto, de tener influjo sobre la forma definitiva de este manuscrito, prácticamente acabado en aquellas fechas. Sobre todo, el autor ha adoptado sin cambios importantes la espléndida idea de E. Artin de basar la teoría de la homología sobre el concepto del número de vueltas.*

Este procedimiento hace posible dar una demostración completa y rigurosa del teorema de Cauchy, y de todas sus aplicaciones inmediatas, con un mínimo de topología. Naturalmente, prescindir de la topología no es en sí un mérito, pero es muy deseable que un libro sobre teoría de funciones se concentre en aquella parte de la topología que es básica en el estudio de las funciones analíticas. Por idéntica razón, no se ha incluido la demostración del teorema de la curva de Jordan, que, en opinión del autor, jamás se necesita en la teoría de funciones.

Los problemas propuestos en el libro tratan de servir solo como acicate al profesor en su tarea ineludible de idear más y mejores ejercicios; es él quien ha de decidir los métodos en que debe insistirse, las demostraciones que conviene repetir al alumno y la iniciativa que este ha de poner en juego en ciertos casos. Es de esperar

que ningún profesor seguirá este libro página a página, ya que nada sería más peligroso. Un texto es una guía didáctica, que le ahorra la necesidad de elaborar por adelantado un plan detallado; pero el contacto continuo con sus alumnos hace que el profesor sea la autoridad competente para decidir sobre las ampliaciones y omisiones que juzgue necesarias.

LARS V. AHLFORS.

INDICE GENERAL

ANALISIS
DE
VARIABLE COMPLEJA

CAPITULO I

NUMEROS COMPLEJOS

1-1. El álgebra de los números complejos.—Es fundamental que los números reales y complejos obedezcan las mismas leyes básicas de la aritmética. Comenzamos nuestro estudio de la teoría de funciones de variable compleja poniendo de manifiesto y completando esta analogía.

1. *Operaciones aritméticas.*—El lector posee ya, del álgebra elemental, el conocimiento de la *unidad imaginaria* i con la propiedad de que $i^2 = -1$. Si se combina la unidad imaginaria con dos números reales α, β, por el proceso de adición y multiplicación obtenemos un *número complejo* $\alpha + i\beta$, donde α y β son las *partes real* e *imaginaria,* respectivamente, del número complejo. Si $\alpha = 0$, se dice que el número es *imaginario puro;* si $\beta = 0$, dicho número es, naturalmente, real. El cero es el único número que es a la vez real e imaginario puro. Dos números complejos son iguales si (y solo si) tienen la misma parte real y la misma parte imaginaria.

La adición y la multiplicación dan siempre resultados pertenecientes al sistema de los números complejos. De hecho, suponiendo que las reglas aritméticas ordinarias son de aplicación a los números complejos, obtenemos

$$(\alpha + i\beta) + (\gamma + i\delta) = (\alpha + \gamma) + i(\beta + \delta) \qquad [1\text{-}1]$$

y

$$(\alpha + i\beta)(\gamma + i\delta) = (\alpha\gamma - \beta\delta) + i(\alpha\delta + \beta\gamma). \qquad [1\text{-}2]$$

En la segunda identidad hemos hecho uso de la relación $i^2 = -1$.

Es menos obvio el hecho de que la división sea también posible. Queremos probar que $(\alpha + i\beta)/(\gamma + i\delta)$ es un número complejo, siempre y cuando $\gamma + i\delta \neq 0$. Si se denota el cociente por $x + iy$, se tendrá

$$\alpha + i\beta = (\gamma + i\delta)(x + iy).$$

Por [1-2] esta condición puede escribirse

$$\alpha + i\beta = (\gamma x - \delta y) + i(\delta x + \gamma y),$$

y obtenemos las dos ecuaciones

$$\alpha = \gamma x - \delta y$$
$$\beta = \delta x + \gamma y.$$

Este sistema de ecuaciones lineales posee la solución única

$$x = \frac{\alpha\gamma + \beta\delta}{\gamma^2 + \delta^2}$$
$$y = \frac{\beta\gamma - \alpha\delta}{\gamma^2 + \delta^2},$$

ya que, como sabemos, $\gamma^2 + \delta^2$ no es nulo. Tenemos así el resultado

$$\frac{\alpha + i\beta}{\gamma + i\delta} = \frac{\alpha\gamma + \beta\delta}{\gamma^2 + \delta^2} + i\,\frac{\beta\gamma - \alpha\delta}{\gamma^2 + \delta^2} \qquad [1\text{-}3]$$

Una vez que ha sido probada la existencia del cociente, su valor puede hallarse de forma más sencilla. Si se multiplican numerador y denominador por $\gamma - i\delta$, se obtiene inmediatamente

$$\frac{\alpha + i\beta}{\gamma + i\delta} = \frac{(\alpha + i\beta)(\gamma - i\delta)}{(\gamma + i\delta)(\gamma - i\delta)} = \frac{(\alpha\gamma + \beta\delta) + i(\beta\gamma - \alpha\delta)}{\gamma^2 + \delta^2}$$

Como caso especial, el inverso de un número complejo ($\neq 0$) está dado por

$$\frac{1}{\alpha + i\beta} = \frac{\alpha - i\beta}{\alpha^2 + \beta^2}$$

Obsérvese que i^n tiene únicamente cuatro valores posibles: 1, i, -1, $-i$. Corresponden a los valores de n, que divididos por 4 dan de restos 0, 1, 2, 3, respectivamente.

EJERCICIOS

1. Hállense los valores de

$$(1+2i)^3, \qquad \frac{5}{-3+4i}, \qquad \left(\frac{2+i}{3-2i}\right)^2, \qquad (1+i)^n + (1-i)^n.$$

2. Si $z = x + iy$ (x e y reales), hállense las partes real e imaginaria de

$$z^4, \qquad \frac{1}{z}, \qquad \frac{z-1}{z+1}, \qquad \frac{1}{z^2}$$

3. Demuéstrese que

$$\left(\frac{-1 \pm i\sqrt{3}}{2}\right)^3 = 1 \quad \text{y} \quad \left(\frac{\pm 1 \pm i\sqrt{3}}{2}\right)^6 = 1$$

para todas las combinaciones posibles de signos.

2. *Raíces cuadradas.*—Probaremos ahora que la raíz cuadrada de un número complejo puede hallarse de manera explícita. Si el número dado es $\alpha + i\beta$, buscamos un número $x + iy$ tal que

$$(x + iy)^2 = \alpha + i\beta.$$

Esto es equivalente al sistema de ecuaciones

$$x^2 - y^2 = \alpha$$
$$2xy = \beta, \qquad [1\text{-}4]$$

de las cuales obtenemos

$$(x^2 + y^2)^2 = (x^2 - y^2)^2 + 4x^2y^2 = \alpha^2 + \beta^2.$$

Por consiguiente,

$$x^2 + y^2 = \sqrt{\alpha^2 + \beta^2},$$

siendo la raíz cuadrada positiva o cero. Esta, junto con la primera ecuación [1-4], nos da

$$x^2 = \tfrac{1}{2}(\alpha + \sqrt{\alpha^2 + \beta^2})$$
$$y^2 = \tfrac{1}{2}(-\alpha + \sqrt{\alpha^2 + \beta^2}). \qquad [1\text{-}5]$$

Obsérvese que estas cantidades son positivas o nulas independientemente del signo de α.

Las ecuaciones [1-5] dan, en general, dos valores opuestos para x y dos para y. Pero estos valores no pueden combinarse arbitrariamente, pues la segunda ecuación [1-4] no es consecuencia de [1-5]. Debemos, por tanto, elegir cuidadosamente x e y, de forma que su producto tenga el signo de β. Esto conduce a la solución general

$$\sqrt{\alpha + i\beta} = \pm\left(\sqrt{\frac{\alpha + \sqrt{\alpha^2 + \beta^2}}{2}} + (\text{signo } \beta)\, i \sqrt{\frac{-\alpha + \sqrt{\alpha^2 + \beta^2}}{2}}\right), \qquad [1\text{-}6]$$

donde signo $\beta = \pm 1$ según que $\beta > 0$ o $\beta < 0$. Para $\beta = 0$, la fórmula es correcta eligiendo uno u otro de los signos para β. Se sobrentiende que todas las raíces cuadradas de números positivos llevan siempre el signo positivo.

Hemos hallado que la raíz cuadrada de cualquier número complejo existe y tiene dos valores opuestos. Estos coinciden únicamente si $\alpha + i\beta = 0$. Son reales si $\beta = 0$, $\alpha \geqq 0$, e imaginarios puros si $\beta = 0$, $\alpha \leqq 0$. En otras palabras, con excepción del cero, solo los números positivos tienen raíces cuadradas reales, y únicamente los negativos las poseen imaginarias puras.

Puesto que ambas raíces cuadradas son en general complejas, no es posible distinguir entre la raíz cuadrada positiva y la negativa de un número complejo. Naturalmente, podríamos distinguir entre los signos superior e inferior en [1-6], pero esta distinción sería artificiosa y debe evitarse. Lo correcto es tratar ambas raíces cuadradas de manera simétrica.

EJERCICIOS

1. Calcúlense

$$\sqrt{i}, \qquad \sqrt{-i}, \qquad \sqrt{1+i}, \qquad \sqrt{\frac{1-i\sqrt{3}}{2}}$$

2. Hállense los cuatro valores de $\sqrt[4]{-1}$.
3. Calcúlense $\sqrt[4]{i}$ y $\sqrt[4]{-i}$.
4. Resuélvase la ecuación cuadrática

$$z^2 + (\alpha + i\beta)z + \gamma + i\delta = 0.$$

3. *Justificación.*—Hasta ahora nuestro tratamiento de los números complejos ha carecido por completo de rigor. No hemos puesto en duda la existencia de un sistema de números en el que la ecuación $x^2 + 1 = 0$ admite solución, al tiempo que conservan su validez todas las reglas aritméticas.

Empecemos por recordar las propiedades características del sistema de los números reales, al que denotaremos por $\mathfrak{R}$. En primer lugar, $\mathfrak{R}$ es un *cuerpo*. Esto significa que la adición y la multiplicación están definidas y satisfacen las *leyes asociativa, conmutativa* y *distributiva*. Los números 0 y 1 son elementos neutros para la adición y la multiplicación, respectivamente: $\alpha + 0 = \alpha$, $\alpha \cdot 1 = \alpha$ para

todo α. Además, la ecuación $\beta + x = \alpha$ posee siempre una solución, y la $\beta x = \alpha$ la tiene siempre y cuando $\beta \neq 0$[1].

Se demuestra mediante razonamientos elementales que los elementos neutros y los resultados de la sustracción y de la división son únicos. Se prueba también que todo cuerpo es un *dominio de integridad:* $\alpha\beta = 0$ si (y solo si) $\alpha = 0$ o $\beta = 0$.

Estas propiedades son comunes a todos los cuerpos. Además, el cuerpo $\mathfrak{R}$ posee una *relación de orden* $\alpha < \beta$ (o $\beta > \alpha$), la cual se define fácilmente mediante el conjunto $\mathfrak{R}^+$ de los números reales *positivos:* $\alpha < \beta$ si (y solo si) $\beta - \alpha \in \mathfrak{R}^+$. El conjunto $\mathfrak{R}^+$ está caracterizado por las siguientes propiedades: 1) cero no es un número positivo; 2) si $\alpha \neq 0$, o bien α es positivo o lo es $-\alpha$; 3) la suma y el producto de dos números positivos son positivos. A partir de estas condiciones se deducen todas las reglas usuales para el manejo de desigualdades. En particular, se demuestra que todo cuadrado α^2 es o positivo o nulo; por tanto, $1 = 1^2$ es un número positivo.

En virtud de la relación de orden, las sumas $1, 1+1, 1+1+1, \ldots$ son todas diferentes; consiguientemente, $\mathfrak{R}$ contiene a los números naturales, y puesto que es un cuerpo, debe contener al subcuerpo constituido por todos los números racionales.

Por último, $\mathfrak{R}$ satisface la siguiente *condición de complitud:* toda sucesión creciente y acotada de números reales tiene límite. Sea $\alpha_1 < \alpha_2 < \alpha_3 < \ldots < \alpha_n < \ldots$, y supongamos la existencia de un número real B tal que $\alpha_n < B$ para todo n. En tal caso, la condición de complitud exige la existencia de un número $A = \lim_{n\to\infty} \alpha_n$ con la propiedad siguiente: dado cualquier $\epsilon > 0$, existe un número natural n_0 tal que $A - \epsilon < \alpha_n < A$ para todo $n > n_0$.

Nuestro estudio del sistema de los números reales es incompleto, puesto que no hemos demostrado la existencia y unicidad (salvo isomorfismos) de un sistema $\mathfrak{R}$ con las propiedades postuladas[2]. El lector que no esté completamente familiarizado con alguno de los procesos constructivos mediante los cuales se pueden introducir

[1] Suponemos al lector con un conocimiento aceptable de álgebra elemental. Aunque la caracterización de un cuerpo anteriormente dada es completa, resulta obvio que sugerirá muy poco al estudiante que no esté familiarizado, al menos vagamente, con dicho concepto.

[2] Un *isomorfismo* entre dos cuerpos es una correspondencia biunívoca que preserva las sumas y los productos. Se emplea el término con toda generalidad para indicar una correspondencia biunívoca que conserva las relaciones que se consideren importantes en una estructura dada.

los números reales debe llenar esa laguna consultando cualquier texto donde se exponga un tratamiento axiomático completo de los números reales.

La ecuación $x^2+1=0$ no tiene solución en $\mathfrak{R}$, pues α^2+1 es siempre positivo. Supongamos ahora que es posible encontrar un cuerpo $\mathfrak{F}$ que contiene a $\mathfrak{R}$ como subcuerpo y en el cual pueda resolverse la ecuación $x^2+1=0$. Denotemos una solución por i. Entonces $x^2+1=(x+i)(x-i)$, y la ecuación $x^2+1=0$ posee exactamente dos raíces en $\mathfrak{F}$, i y $-i$. Sea $\mathfrak{C}$ el subconjunto de $\mathfrak{F}$ constituido por todos los elementos que pueden expresarse en la forma $\alpha+i\beta$ con α y β reales. Esta representación es única, pues $\alpha+i\beta=\alpha'+i\beta'$ implica $\alpha-\alpha'=-i(\beta-\beta')$; por tanto, $(\alpha-\alpha')^2=-(\beta-\beta')^2$, lo cual es posible únicamente si $\alpha=\alpha'$ y $\beta=\beta'$.

El subconjunto $\mathfrak{C}$ es un subcuerpo de $\mathfrak{F}$. De hecho, excepto por lo que se refiere a comprobaciones triviales, que el lector debería efectuar, esto es exactamente lo que se demostró en la sección 1-1, 1. Lo que es más, la estructura de $\mathfrak{C}$ es independiente de $\mathfrak{F}$. Pues si $\mathfrak{F}'$ es otro cuerpo que contiene a $\mathfrak{R}$ e i' una raíz de la ecuación $x^2+1=0$, el correspondiente subconjunto $\mathfrak{C}'$ está formado por todos los elementos $\alpha+i'\beta$. Existe una correspondencia biunívoca entre $\mathfrak{C}$ y $\mathfrak{C}'$, que hace corresponder $\alpha+i'\beta$ a $\alpha+i\beta$, y evidentemente esta correspondencia es un isomorfismo entre ambos cuerpos. Queda así demostrado que $\mathfrak{C}$ y $\mathfrak{C}'$ son isomorfos.

Definimos ahora el cuerpo de los *números complejos* como el subcuerpo $\mathfrak{C}$ de un cuerpo arbitrariamente dado $\mathfrak{F}$. Acabamos de ver que la elección de $\mathfrak{F}$ es indiferente, pero no hemos probado todavía que exista un cuerpo $\mathfrak{F}$ con las propiedades requeridas. Con objeto de que nuestra definición tenga sentido queda por ver que existe efectivamente un cuerpo $\mathfrak{F}$ que contiene a $\mathfrak{R}$ (o a un subcuerpo isomorfo con $\mathfrak{R}$) y en el cual admite solución la ecuación $x^2+1=0$.

Existen muchos métodos para construir tal cuerpo. El más sencillo y directo es el siguiente: Consideremos todas las expresiones de la forma $\alpha+i\beta$, donde α, β son números reales, mientras que $+$ e i son puramente símbolos ($+$ *no* indica adición e i *no* es un elemento de un cuerpo). Estas expresiones son elementos de un cuerpo $\mathfrak{F}$, en el que la adición y la multiplicación se definen por [1-1] y [1-2] (obsérvense los dos significados distintos del signo $+$). Se comprueba que los elementos de la forma particular $\alpha+i0$ constituyen un subcuerpo isomorfo con $\mathfrak{R}$, y que el elemento $0+i1$ satisface la ecuación $x^2+1=0$; en efecto, se obtiene $(0+i1)^2=-(1+i0)$.

Así, pues, el cuerpo $\mathfrak{F}$ tiene las propiedades requeridas; además es idéntico al correspondiente subcuerpo $\mathfrak{C}$, ya que podemos escribir

$$\alpha+i\beta=(\alpha+i0)+\beta(0+i1).$$

Queda así demostrada la existencia del cuerpo de los números complejos, y podemos volver a la notación más sencilla $\alpha+i\beta$, donde $+$ indica adición en $\mathfrak{C}$ e i es una raíz de la ecuación $x^2+1=0$.

EJERCICIOS

(Para lectores con una base de álgebra)

1. Demuéstrese que el conjunto de todas las matrices de la forma especial

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix},$$

combinadas mediante la adición y la multiplicación de matrices, es isomorfo al cuerpo de los números complejos.

2. Demuéstrese que cabe considerar al sistema de los números complejos como el cuerpo de todos los polinomios con coeficientes reales módulo el polinomio irreducible x^2+1.

4. *Conjugación, valor absoluto.*—Se puede denotar un número complejo, bien por una sola letra a, que representa un elemento del cuerpo $\mathfrak{C}$, o bien en la forma $\alpha+i\beta$, con α y β reales. Otras notaciones típicas son $z=x+iy$, $\zeta=\xi+i\eta$, $w=u+iv$, entendiéndose al utilizar estas formas que x, y, ξ, η, u, v son números reales. Las partes real e imaginaria de un número complejo a se denotarán también por $\operatorname{Re} a$, $\operatorname{Im} a$, respectivamente.

Al deducir las reglas de la adición y de la multiplicación complejas hemos utilizado únicamente el hecho de que $i^2=-1$. Puesto que $-i$ tiene la misma propiedad, todas las reglas continuarán siendo válidas si se reemplaza i por $-i$. Por comprobación directa se ve que esto es realmente cierto. La transformación que sustituye $\alpha+i\beta$ por $\alpha-i\beta$ se llama *conjugación compleja*, y $\alpha-i\beta$ es el *conjugado* de $\alpha+i\beta$. El conjugado de a se denota por $\bar{a}$. Un número es real si (y solo si) es igual a su conjugado. La conjugación es una transformación *involutiva;* esto significa que $\bar{\bar{a}}=a$.

Las fórmulas

$$\operatorname{Re} a=\frac{a+\bar{a}}{2}, \qquad \operatorname{Im} a=\frac{a-\bar{a}}{2i}$$

expresan las partes real e imaginaria de un número complejo en función de este y de su conjugado. Por consiguiente, haciendo uso sistemático de las notaciones a y $\bar{a}$ será posible evitar la utilización de letras distintas para designar las partes real e imaginaria. Sin embargo, resulta más conveniente usar indistintamente ambas notaciones.

La propiedad fundamental de la conjugación es la mencionada, a saber:

$$\overline{a+b}=\bar{a}+\bar{b}$$

$$\overline{ab}=\bar{a}\cdot\bar{b}.$$

La propiedad correspondiente para cocientes es consecuencia de lo anterior: si $ax=b$, $\bar{a}\bar{x}=\bar{b}$, y de aquí $\overline{(b/a)}=\bar{b}/\bar{a}$. Con mayor generalidad, si $R(a, b, c, \ldots)$ indica una operación racional aplicada a los números complejos $a, b, c, \ldots$, se tendrá

$$\overline{R(a, b, c, \ldots)}=R(\bar{a}, \bar{b}, \bar{c}, \ldots).$$

Como aplicación, consideremos la ecuación

$$c_0z^n+c_1z^{n-1}+\ldots+c_{n-1}z+c_n=0.$$

Si ζ es una raíz de esta ecuación, $\bar{\zeta}$ será raíz de la ecuación

$$\bar{c}_0z^n+\bar{c}_1z^{n-1}+\ldots+\bar{c}_{n-1}z+\bar{c}_n=0.$$

En particular, si los coeficientes son *reales*, ζ y $\bar{\zeta}$ son raíces de la misma ecuación, y obtenemos el conocido teorema de que las raíces no reales de una ecuación de coeficientes reales aparecen por parejas de raíces conjugadas.

El producto $a\bar{a}=\alpha^2+\beta^2$ es siempre positivo o cero. Su raíz cuadrada no negativa se llama *módulo* o *valor absoluto* del número complejo a, y se denota por $|a|$. La terminología y la notación quedan justificadas por el hecho de que el módulo de un número real coincide con su valor numérico tomado con signo positivo.

Repetimos la definición

$$a\bar{a}=|a|^2,$$

siendo $|a|\geqq 0$, y obsérvese que $|\bar{a}|=|a|$. Para el valor absoluto de un producto, se obtiene

$$|ab|^2=ab\cdot\overline{ab}=ab\bar{a}\bar{b}=a\bar{a}b\bar{b}=|a|^2|b|^2,$$

y de aquí

$$|ab|=|a|\cdot|b|,$$

puesto que ambos son $\geqq 0$. Expresado verbalmente:

El valor absoluto de un producto es igual al producto de los valores absolutos de los factores.

Es claro que esta propiedad se extiende a productos finitos arbitrarios:

$$|a_1a_2\ldots a_n|=|a_1|\cdot|a_2|\cdots|a_n|.$$

El cociente a/b, $b\neq 0$, verifica $b(a/b)=a$; por tanto, tenemos también $|b|\cdot|a/b|=|a|$, o sea

$$\left|\frac{a}{b}\right|=\frac{|a|}{|b|}$$

La fórmula para el valor absoluto de una suma no es tan sencilla. Se encuentra que

$$|a+b|^2=(a+b)(\bar{a}+\bar{b})=a\bar{a}+(a\bar{b}+b\bar{a})+b\bar{b}$$

o

$$|a+b|^2=|a|^2+|b|^2+2\ \mathrm{Re}\ a\bar{b}. \qquad [1\text{-}7]$$

La fórmula correspondiente en el caso de la sustracción es

$$|a-b|^2=|a|^2+|b|^2-2\ \mathrm{Re}\ a\bar{b}, \qquad [1\text{-}7']$$

y sumando obtenemos la identidad

$$|a+b|^2+|a-b|^2=2(|a|^2+|b|^2). \qquad [1\text{-}8]$$

EJERCICIOS

1. Compruébese mediante el correspondiente cálculo que los valores de

$$\frac{z}{z^2+1}$$

para $z=x+iy$ y $z=x-iy$ son conjugados.

2. Hállense los valores absolutos de

$$-2i(3+i)(2+4i)(1+i) \quad \text{y} \quad \frac{(3+4i)(-1+2i)}{(-1-i)(3-i)}$$

3. Demuéstrese que

$$\left|\frac{a-b}{1-\bar{a}b}\right|=1$$

si $|a|=1$ o $|b|=1$. ¿Qué excepción ha de hacerse en el caso de que $|a|=|b|=1$?

4. Pruébese la identidad de Lagrange en forma compleja:

$$\left|\sum_{i=1}^{n} a_i b_i\right|^2=\sum_{i=1}^{n}|a_i|^2\sum_{i=1}^{n}|b_i|^2-\sum_{1\leqq i<j\leqq n}|a_i\bar{b}_j-a_j\bar{b}_i|^2.$$

5. *Desigualdades.*—Demostraremos ahora algunas desigualdades importantes que serán de uso constante. Quizá sea conveniente indicar que no existe relación de orden en el sistema de los números complejos, y, por tanto, todas las desigualdades deben verificarse entre números reales.

Partiendo de la definición de valor absoluto, se deducen las desigualdades

$$\begin{aligned}-|a|&\leqq \operatorname{Re} a\leqq|a|\\ -|a|&\leqq \operatorname{Im} a\leqq|a|.\end{aligned} \qquad [1\text{-}9]$$

La igualdad $\operatorname{Re} a=|a|$ se verifica si (y solo si) a es real y mayor o igual que cero.

Si se aplica [1-9] a [1-7], se obtiene

$$|a+b|^2\leqq(|a|+|b|)^2$$

y de aquí

$$|a+b|\leqq|a|+|b|. \qquad [1\text{-}10]$$

Esta es la *desigualdad triangular,* así denominada por razones que se pondrán de manifiesto más adelante. Por inducción, se puede extender a sumas cualesquiera:

$$|a_1+a_2+\ldots+a_n|\leqq|a_1|+|a_2|+\ldots+|a_n|. \qquad [1\text{-}11]$$

El valor absoluto de una suma es a lo sumo igual a la suma de los valores absolutos de los sumandos.

El lector conoce perfectamente la importancia de la fórmula [1-11] en el caso real, y comprobará que no es menos importante en la teoría de los números complejos.

Determinemos todos los casos de igualdad en [1-11]. Se verifica el signo de igualdad en [1-10] si (y solo si) $a\bar{b}\geqq 0$ (es conveniente indicar que c es *real* y *positivo* poniendo $c>0$). Si $b\neq 0$, esta condición puede escribirse en la forma $|b|^2(a/b)\geqq 0$, y es, por tanto, equivalente a $a/b\geqq 0$. En el caso general, se procede de la forma siguiente: Supongamos que se verifica la igualdad en [1-11]; entonces

$$\begin{aligned}|a_1|+|a_2|+\ldots+|a_n|&=|(a_1+a_2)+a_3+\ldots+a_n|\\ &\leqq|a_1+a_2|+|a_3|+\ldots+|a_n|\leqq|a_1|+|a_2|+\ldots+|a_n|.\end{aligned}$$

Por consiguiente, $|a_1+a_2|=|a_1|+|a_2|$, y si $a_2\neq 0$, se obtiene $a_1/a_2\geqq 0$. Pero la numeración de los términos es arbitraria; luego el cociente de dos términos cualesquiera no nulos debe ser positivo. Recíprocamente, imaginemos que se cumple esta condición. Suponiendo que $a_1\neq 0$, obtenemos

$$\begin{aligned}|a_1+a_2+\ldots+a_n|&=|a_1|\cdot\left|1+\frac{a_2}{a_1}+\ldots+\frac{a_n}{a_1}\right|\\ &=|a_1|\left(1+\frac{a_2}{a_1}+\ldots+\frac{a_n}{a_1}\right)=|a_1|\left(1+\frac{|a_2|}{|a_1|}+\ldots+\frac{|a_n|}{|a_1|}\right)\\ &=|a_1|+|a_2|+\ldots+|a_n|.\end{aligned}$$

En resumen: *en [1-11] es válido el signo de igualdad si (y solo si) el cociente de dos términos no nulos cualesquiera es positivo.*

Por [1-10] se tiene también

$$|a|=|(a-b)+b|\leqq|a-b|+|b|$$

o

$$|a|-|b|\leqq|a-b|.$$

Por la misma razón, $|b|-|a|\leqq|a-b|$, y estas desigualdades pueden combinarse para dar

$$|a-b|\geqq||a|-|b||. \qquad [1\text{-}12]$$

Naturalmente se puede aplicar la misma acotación a $|a+b|$.

Un caso especial de [1-10] es la desigualdad

$$|\alpha+i\beta|\leqq|\alpha|+|\beta|, \qquad [1\text{-}13]$$

que expresa que el valor absoluto de un número complejo es, a lo sumo, igual a la suma de los valores absolutos de las partes real e imaginaria.

Muchas otras desigualdades cuya demostración es menos inmediata son también de uso frecuente. La más importante es la *desigualdad de Cauchy,* que afirma que

$$|a_1b_1+\ldots+a_nb_n|^2\leqq(|a_1|^2+\ldots+|a_n|^2)(|b_1|^2+\ldots+|b_n|^2)$$

o, con notación más abreviada,

$$\left|\sum_{i=1}^{n}a_ib_i\right|^2\leqq\sum_{i=1}^{n}|a_i|^2\sum_{i=1}^{n}|b_i|^2\,{}^*. \qquad [1\text{-}14]$$

Para demostrarla, sea λ un número complejo arbitrario. Por [1-7] se obtiene

$$\sum_{i=1}^{n}|a_i-\lambda\bar{b}_i|^2=\sum_{i=1}^{n}|a_i|^2+|\lambda|^2\sum_{i=1}^{n}|b_i|^2-2\,\mathrm{Re}\,\bar{\lambda}\sum_{i=1}^{n}a_ib_i. \qquad [1\text{-}15]$$

Esta expresión es $\geqq 0$ para todo λ. Podemos elegir

$$\lambda=\frac{\sum_1^n a_ib_i}{\sum_1^n|b_i|^2}$$

pues si el denominador se anulase no habría nada que demostrar. Esta elección no es arbitraria, pues está dictada por el deseo de hacer la expresión [1-15] tan pequeña como sea posible. Sustituyendo

* i es un índice de sumación conveniente, y, utilizado como subíndice, no puede confundirse con la unidad imaginaria. Parece innecesario proscribir su utilización.

en [1-15], y después de simplificar, se obtiene

$$\sum_1^n|a_i|^2-\frac{\left|\sum_1^n a_ib_i\right|^2}{\sum_1^n|b_i|^2}\geqq 0,$$

que es equivalente a [1-14].

Además, de [1-15] se deduce que el signo de desigualdad es válido en [1-14] si (y solo si) las a_i son proporcionales a las $\bar{b}_i$.

La desigualdad de Cauchy se puede demostrar también mediante la identidad de Lagrange (Sec. 1-1, 4; Ej. 4).

EJERCICIOS

1. Demuéstrese que

$$\left|\frac{a-b}{1-\bar{a}b}\right|<1$$

si $|a|<1$ y $|b|<1$.

2. Demuéstrese por inducción la desigualdad de Cauchy.

3. Si $|a_i|<1$, $\lambda_i\geqq 0$ para $i=1, \ldots, n$ y $\lambda_1+\lambda_2+\ldots+\lambda_n=1$, pruébese que

$$|\lambda_1a_1+\lambda_2a_2+\ldots+\lambda_na_n|<1.$$

4. Determínese el menor valor de $|z-a|^2+|z-b|^2$ cuando se dan a y b.

1-2. Representación geométrica de los números complejos.

El número complejo $a=\alpha+i\beta$ puede representarse, con respecto a un sistema de coordenadas rectangulares en el plano, por el punto de coordenadas (α, β). Se utiliza constantemente esta representación, y a menudo hablaremos del *punto a* como sinónimo del *número a.* El primer eje de coordenadas (eje de las x) toma el nombre de *eje real,* y el segundo eje de coordenadas (eje de las y) se llama *eje imaginario.* Al plano se le denomina *plano complejo.*

La utilidad de la representación geométrica se deriva de lo intuitivas que resultan las imágenes mentales asociadas al lenguaje geométrico. Si los teoremas de geometría sintética estuvieran firmemente establecidos, podríamos utilizarlos para deducir resultados relativos a los números complejos. Desgraciadamente, hay que admitir que en los cursos elementales la geometría no se desarrolla con completo rigor. Por esta razón no consideraremos como concluyente ninguna demostración geométrica. Esta limitación no nos impide utilizar el lenguaje geométrico, siempre y cuando no olvidemos que en definitiva todas las demostraciones han de reducirse a términos analíticos. Siendo esta nuestra postura, estamos, por otra parte, relevados de las exigencias del rigor en relación con consideraciones geométricas de carácter puramente descriptivo.

1. *Representación geométrica de la adición y de la multiplicación.*—Se puede visualizar la adición de números complejos como *adición vectorial.* Con este objeto representaremos un número complejo no solo mediante un punto, sino también mediante un vector con origen en el de coordenadas y extremo el punto. El número, el punto y el vector se denotarán por la misma letra a. Como de costumbre, identificaremos todos los vectores que se pueden obtener unos de otros por desplazamientos paralelos.

Colóquese un segundo vector b de forma que su origen coincida con el extremo de a. Entonces $a+b$ está representado por el vector cuyo origen es el de a y cuyo extremo es el de b. Para construir la diferencia $b-a$ se dibujan ambos vectores a y b a partir del mismo punto inicial; entonces $b-a$ tiene su origen en el extremo de a y su extremo en el de b. Obsérvese que $a+b$ y $a-b$ son las diagonales de un paralelogramo de lados a y b. (Fig. 1-1).

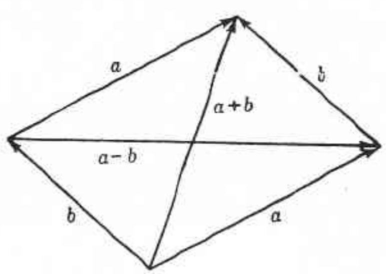

FIG. 1-1.—Adición vectorial.

Una ventaja adicional de la representación vectorial es que la longitud del vector a es igual a $|a|$. Luego la distancia entre los puntos a y b es $|a-b|$. Con esta interpretación, la desigualdad triangular $|a+b| \leqq |a|+|b|$ y la identidad $|a+b|^2+|a-b|^2=2(|a|^2+|b|^2)$ se convierten en teoremas geométricos familiares.

El punto a y su conjugado $\bar{a}$ están colocados simétricamente con respecto al eje real. El punto simétrico de a con respecto al eje imaginario es $-\bar{a}$. Los cuatro puntos a, $-\bar{a}$, $-a$, $\bar{a}$ son los vértices de un rectángulo, que es simétrico respecto a ambos ejes.

Para obtener una interpretación geométrica del producto de dos números complejos introduciremos coordenadas polares. Si las coordenadas polares del punto (α, β) son (r, φ), sabemos que

$$\alpha = r \cos \varphi$$
$$\beta = r \operatorname{sen} \varphi.$$

Luego, podemos escribir $a=\alpha+i\beta=r(\cos\varphi+i\operatorname{sen}\varphi)$. En esta *forma trigonométrica* de un número complejo, r es siempre $\geqq 0$ e igual al módulo $|a|$. Al ángulo polar φ se le llama *argumento* o *amplitud* del número complejo, y se le denota por $\arg a$.

Consideremos dos números complejos $a_1=r_1(\cos\varphi_1+i\operatorname{sen}\varphi_1)$ y $a_2=r_2(\cos\varphi_2+i\operatorname{sen}\varphi_2)$. Su producto puede escribirse en la forma

$$a_1a_2=r_1r_2[(\cos\varphi_1\cos\varphi_2-\operatorname{sen}\varphi_1\operatorname{sen}\varphi_2)+i(\operatorname{sen}\varphi_1\cos\varphi_2+\cos\varphi_1\operatorname{sen}\varphi_2)].$$

Mediante los teoremas de adición del coseno y del seno, esta expresión se reduce a

$$a_1a_2=r_1r_2[\cos(\varphi_1+\varphi_2)+i\operatorname{sen}(\varphi_1+\varphi_2)]. \qquad [1\text{-}16]$$

Se puede reconocer que el módulo del producto es r_1r_2 y su argumento es $\varphi_1+\varphi_2$. El último resultado aparece por primera vez, y lo expresamos mediante la ecuación

$$\arg(a_1a_2)=\arg a_1+\arg a_2. \qquad [1\text{-}17]$$

Naturalmente, esta fórmula puede extenderse a productos arbitrarios; por tanto, puede enunciarse:

El argumento de un producto es igual a la suma de los argumentos de los factores.

Esto es fundamental. La regla que acabamos de formular da una justificación profunda e inesperada a la representación geométrica de los números complejos. Debemos, sin embargo, estar plenamente conscientes de que la forma en que hemos llegado a la fórmula [1-17] vulnera nuestros principios. En primer lugar, la ecuación [1-17] relaciona *ángulos* más bien que números, y en segundo lugar, en su demostración se utiliza la trigonometría. Así, pues, queda por definir el argumento en términos analíticos y demostrar [1-17] mediante procedimientos puramente analíticos. De momento aplazamos esta demostración y nos contentamos con discutir las consecuencias de [1-17] desde un punto de vista menos riguroso.

Observemos, en primer lugar, que el argumento de 0 no está definido, y, por tanto, [1-17] tiene sentido si (y solo si) a_1 y a_2

son $\neq 0$. En segundo lugar, el ángulo polar está definido salvo múltiplos de 360°. Por esta razón, si queremos interpretar [1-17] numéricamente, tenemos que convenir en que los múltiplos de 360° no cuentan.

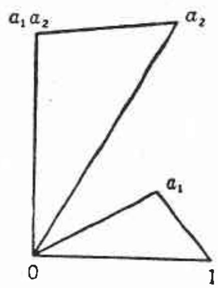

FIG. 1-2.—Multiplicación vectorial.

Mediante [1-17] se puede obtener una sencilla construcción del producto a_1a_2. En efecto, se deduce que el triángulo de vértices 0, 1, a_1 es semejante al triángulo cuyos vértices son 0, a_2, a_1a_2. Dados los puntos 0, 1, a_1 y a_2, esta semejanza determina el punto a_1a_2 (Fig. 1-2).

En el caso de la división, se reemplaza [1-17] por

$$\arg \frac{a_2}{a_1} = \arg a_2 - \arg a_1. \qquad [1\text{-}18]$$

La construcción geométrica es la misma, excepto que los triángulos semejantes son ahora 0, 1, a_1 y 0, a_2/a_1, a_2.

EJERCICIOS

1. Hállense los puntos simétricos de a respecto a las rectas que bisecan los ángulos determinados por los ejes de coordenadas.

2. Pruébese que los puntos a_1, a_2, a_3 son los vértices de un triángulo equilátero si (y solo si) $a_1^2+a_2^2+a_3^2=a_1a_2+a_2a_3+a_3a_1$.

3. Se supone que a y b son dos vértices de un cuadrado. Hállense los otros dos vértices en todos los casos posibles.

4. Hállense el centro y el radio del círculo circunscrito al triángulo de vértices a_1, a_2, a_3. Exprésese el resultado en forma simétrica.

2. *La ecuación binómica.*—De los resultados precedentes, que seguimos admitiendo sin demostración suficiente, se obtiene que las potencias de $a=r(\cos\varphi+i\operatorname{sen}\varphi)$ están dadas por

$$a^n = r^n(\cos n\varphi + i \operatorname{sen} n\varphi). \qquad [1\text{-}19]$$

Esta fórmula es trivialmente válida para $n=0$, y puesto que

$$a^{-1}=r^{-1}(\cos\varphi - i\operatorname{sen}\varphi) = r^{-1}[\cos(-\varphi)+i\operatorname{sen}(-\varphi)],$$

también se verifica cuando n es un entero negativo.

Para $r=1$ obtenemos la *fórmula de De Moivre:*

$$(\cos\varphi + i\operatorname{sen}\varphi)^n = \cos n\varphi + i\operatorname{sen} n\varphi, \qquad [1\text{-}20]$$

que proporciona un procedimiento extraordinariamente sencillo para expresar $\cos n\varphi$ y $\operatorname{sen} n\varphi$ en términos de $\cos\varphi$ y de $\operatorname{sen}\varphi$.

Para hallar la raíz n-ésima de un número complejo a tenemos que resolver la ecuación

$$z^n = a. \qquad [1\text{-}21]$$

En el supuesto de que $a\neq 0$, escribimos $a=r(\cos\varphi+i\operatorname{sen}\varphi)$ y

$$z=\rho(\cos\theta + i\operatorname{sen}\theta).$$

Entonces [1-21] adopta la forma

$$\rho^n(\cos n\theta + i\operatorname{sen} n\theta) = r(\cos\varphi + i\operatorname{sen}\varphi). \qquad [1\text{-}22]$$

Esta ecuación se verifica ciertamente si $\rho^n=r$ y $n\theta=\varphi$. De aquí obtenemos la raíz

$$z=\sqrt[n]{r}\left(\cos\frac{\varphi}{n} + i\operatorname{sen}\frac{\varphi}{n}\right),$$

donde $\sqrt[n]{r}$ denota la raíz n-ésima positiva del número positivo r.

Pero esta no es la única solución. En efecto, también se verifica [1-22] si $n\theta$ difiere de φ en un múltiplo del ángulo total. Si se expresan los ángulos en radianes, el ángulo total es 2π, y tenemos que se verifica [1-22] si (y solo si)

$$\theta = \frac{\varphi}{n} + k\cdot\frac{2\pi}{n},$$

siendo k un entero cualquiera. Sin embargo, solo dan valores de z diferentes los valores 0, 1, ..., $n-1$ de k. Por tanto, la solución completa de la ecuación [1-21] viene dada por

$$z=\sqrt[n]{r}\left[\cos\left(\frac{\varphi}{n}+k\frac{2\pi}{n}\right) + i\operatorname{sen}\left(\frac{\varphi}{n}+k\frac{2\pi}{n}\right)\right], \quad k=0,\ 1,\ \ldots,\ n-1.$$

Existen n raíces n-ésimas de cualquier número complejo $\neq 0$. Todas tienen el mismo módulo y sus argumentos difieren en múltiplos de $\frac{2\pi}{n}$.

Geométricamente, las raíces n-ésimas son los vértices de un polígono regular de n lados.

Es particularmente importante el caso $a=1$. Las raíces de la ecuación $z^n=1$ se llaman raíces n-ésimas de la unidad, y si ponemos

$$\omega=\cos\frac{2\pi}{n}+i\,\mathrm{sen}\,\frac{2\pi}{n} \qquad [1\text{-}23]$$

todas las raíces pueden expresarse mediante $1, \omega, \omega^2, \ldots, \omega^{n-1}$. Resulta evidente que si $\sqrt[n]{a}$ denota cualquier raíz n-ésima de a, entonces todas las raíces n-ésimas pueden expresarse en la forma $\omega^k \cdot \sqrt[n]{a}$, $k=0, 1, \ldots, n-1$.

EJERCICIOS

1. Exprésense $\cos 3\varphi$, $\cos 4\varphi$ y $\mathrm{sen}\, 5\varphi$ en términos de $\cos\varphi$ y de $\mathrm{sen}\,\varphi$.

2. Simplifíquense las expresiones $1+\cos\varphi+\cos 2\varphi+\ldots+\cos n\varphi$ y $\mathrm{sen}\,\varphi+\mathrm{sen}\,2\varphi+\ldots+\mathrm{sen}\,n\varphi$.

3. Exprésense en forma algebraica las raíces quintas y décimas de la unidad.

4. Si se da ω mediante [1-23], demuéstrese que se verifica

$$1+\omega^h+\omega^{2h}+\ldots+\omega^{(n-1)h}=0$$

para cualquier entero h no múltiplo de n.

5. ¿Cuál es el valor de la expresión

$$1-\omega^h+\omega^{2h}-\ldots+(-1)^{n-1}\omega^{(n-1)h}?$$

3. *Definición de argumento.*—Debemos ahora dar una definición de argumento y demostrar la relación fundamental [1-17] sin recurrir a la intuición geométrica. Con este objeto daremos definiciones puramente analíticas del seno y del coseno. Para estas definiciones haremos uso de propiedades sencillas de las integrales definidas. Existen otros métodos más satisfactorios en otros aspectos, pero creemos que una referencia directa a algunos resultados bien establecidos del cálculo elemental facilitará el estudio a la mayoría de los lectores.

En primer lugar *definimos* el número π mediante la ecuación

$$\int_0^1 \frac{dt}{\sqrt{1-t^2}}=\frac{\pi}{2}$$

La integral

$$\int_0^y \frac{dt}{\sqrt{1-t^2}}$$

es una función continua y creciente del límite superior y en el intervalo $0\leqq y\leqq 1$. Por tanto, si $0\leqq\theta\leqq\pi/2$, existirá un valor de este intervalo (y solo uno), para el que

$$\int_0^y \frac{dt}{\sqrt{1-t^2}}=\theta. \qquad [1\text{-}24]$$

Este valor de y es, por definición, $\mathrm{sen}\,\theta$. Para los mismos valores de θ se define el coseno por

$$\cos\theta=\mathrm{sen}\left(\frac{\pi}{2}-\theta\right). \qquad [1\text{-}25]$$

Nuestro procedimiento se inicia con la definición de la función inversa del seno, por la sencilla razón de que esta función puede representarse mediante una integral. De tal representación mediante una integral se sigue que la derivada de θ con respecto a y es $1/\sqrt{1-y^2}$, para $0\leqq y<1$. Mediante la regla de formación de la derivada de la función inversa, obtenemos

$$D\,\mathrm{sen}\,\theta=\sqrt{1-\mathrm{sen}^2\theta}. \qquad [1\text{-}26]$$

Se recuerda al lector que la raíz cuadrada es positiva. El valor $\theta=\pi/2$ ha sido excluido; pero una aplicación familiar y sencilla del teorema del valor medio muestra que [1-26] sigue siendo válida para este valor de θ[1].

Derivando una vez más, se obtiene

$$D^2\,\mathrm{sen}\,\theta=-\frac{\mathrm{sen}\,\theta}{\sqrt{1-\mathrm{sen}^2\theta}}\cdot\sqrt{1-\mathrm{sen}^2\theta}=-\mathrm{sen}\,\theta;$$

[1] La proposición a que nos referimos es la siguiente: supongamos que $f(x)$ es continua para $a\leqq x\leqq b$ y tiene una derivada para $a<x<b$; si $f'(x)$ tiene un límite B cuando x tiende a b, entonces existe la derivada por la izquierda de $f(x)$ en b y es igual a B. Se deduce la demostración a partir de la igualdad $f(b)-f(x)=f'(\xi)(b-x)$ con $x<\xi<b$.

para $\theta=\pi/2$ es preciso usar nuevamente el teorema del valor medio. Mediante su aplicación a [1-25], se obtiene

$$D^{(2)} \cos\theta = -\cos\theta.$$

Es ahora conveniente introducir la función compleja

$$e(\theta) = \cos\theta + i \operatorname{sen}\theta. \qquad [1\text{-}27]$$

El concepto de derivada puede extenderse a funciones complejas de una variable real por el sencillo procedimiento de definir la derivada de $a(t)=\alpha(t)+i\beta(t)$ como $a'(t)=\alpha'(t)+i\beta'(t)$. Puede comprobarse que las reglas de derivación para una suma y para un producto siguen siendo válidas.

Los resultados anteriores muestran que

$$e''(\theta) = -e(\theta).$$

Multiplicando ambos miembros de esta ecuación diferencial por $e'(\theta)$, se llega a la conclusión de que la expresión $e'(\theta)^2+e(\theta)^2$ tiene derivada nula. Por consiguiente, es constante, y, puesto que $e(0)=1$, $e'(0)=i$, se tiene que $e'(\theta)^2+e(\theta)^2=0$, o sea $e'(\theta)=\pm ie(\theta)$. El signo inferior es incompatible con [1-26]; por tanto, se tiene

$$e'(\theta) = ie(\theta). \qquad [1\text{-}28]$$

Esta relación es equivalente a

$$\begin{aligned} D \operatorname{sen}\theta &= \cos\theta \\ D \cos\theta &= -\operatorname{sen}\theta, \end{aligned} \qquad [1\text{-}29]$$

y por comparación con [1-26], se obtiene la siguiente identidad:

$$\operatorname{sen}^2\theta + \cos^2\theta = 1, \qquad [1\text{-}30]$$

de donde $|e(\theta)|=1$.

Esto completa la discusión en lo que se refiere al intervalo $(0, \pi/2)$. Ampliamos ahora el dominio de definición de $e(\theta)$ a valores reales de θ arbitrarios mediante la condición

$$e\left(\theta+\frac{\pi}{2}\right) = ie(\theta).$$

Esto es legítimo, puesto que $e(\pi/2)=ie(0)$, y $e(\theta)$ quedará determinada de manera única como una función periódica de período 2π. Se sigue además que $e'(\theta+\pi/2)=ie'(\theta)$, y, por consiguiente, [1-28] sigue verificándose. Cuando θ es un múltiplo entero de $\pi/2$, se obtiene la demostración considerando por separado las derivadas por la derecha y por la izquierda.

Se amplían las definiciones de las funciones trigonométricas poniendo $e(\theta)=\cos\theta+i\operatorname{sen}\theta$ para todo valor de θ. Es evidente que las relaciones [1-29] y [1-30] siguen siendo válidas.

Elijamos ahora un θ_0 fijo y consideremos la función producto $e(\theta_0+\theta)e(-\theta)$. Su derivada respecto a θ es

$$\begin{aligned} &e'(\theta_0+\theta)e(-\theta) - e(\theta_0+\theta)e'(-\theta) = \\ &\qquad = ie(\theta_0+\theta)e(-\theta) - ie(\theta_0+\theta)e(-\theta) = 0. \end{aligned}$$

Por tanto, la función se reduce a una constante cuyo valor se determina poniendo $\theta=0$. Se obtiene

$$e(\theta_0+\theta)e(-\theta) = e(\theta_0)$$

para θ_0 y θ cualesquiera. Resulta una forma más simétrica si se escribe $\theta_0=\theta_1+\theta_2$ y $\theta=-\theta_2$. Entonces la relación es

$$e(\theta_1+\theta_2) = e(\theta_1)e(\theta_2). \qquad [1\text{-}31]$$

Este es el teorema de adición para la función $e(\theta)$. Si se separan las partes reales y las imaginarias, se obtienen las relaciones

$$\begin{aligned} \cos(\theta_1+\theta_2) &= \cos\theta_1\cos\theta_2 - \operatorname{sen}\theta_1\operatorname{sen}\theta_2 \\ \operatorname{sen}(\theta_1+\theta_2) &= \operatorname{sen}\theta_1\cos\theta_2 + \cos\theta_1\operatorname{sen}\theta_2. \end{aligned}$$

En otras palabras: hemos *demostrado* los teoremas trigonométricos de adición de argumentos.

¿Cuándo es $e(\theta_1)=e(\theta_2)$? Por [1-31] esta ecuación es equivalente a $e(\theta_1-\theta_2)=1$. La ecuación $e(\theta)=1$ se verifica para $\theta=n\cdot 2\pi$, siendo n un número entero. Para ver que estas son las únicas soluciones basta observar que el seno es positivo para $0<\theta\leqq\pi/2$; el coseno, negativo para $\pi/2<\theta<3\pi/2$, y el seno negativo para $3\pi/2\leqq\theta<2\pi$; esto excluye todos los valores entre 0 y 2π, y a causa de la periodicidad, estos son los únicos valores que es necesario investigar. Se concluye que $e(\theta_1)=e(\theta_2)$ implica $\theta_1=\theta_2+n\cdot 2\pi$.

Ahora puede definirse el argumento. Si $\zeta=\xi+i\eta$ tiene de módulo 1, afirmamos que la ecuación $e(\theta)=\zeta$ tiene solución. En primer lugar, si $\xi\geqq 0$, $\eta\geqq 0$, una solución viene dada por

$$\theta=\int_0^\eta \frac{dt}{\sqrt{1-t^2}},$$

pues entonces sen $\theta=\eta$ y $\cos\theta=\sqrt{1-\eta^2}=\xi$. Además, si $\xi\geqq 0$, $\eta\leqq 0$ se puede resolver la ecuación $e(\theta)=i\zeta$, y de este modo la ecuación original tiene la solución $\theta-\pi/2$. Los casos $\xi\leqq 0$, $\eta\geqq 0$ y $\xi\leqq 0$, $\eta\leqq 0$ pueden tratarse de igual manera. A partir de una solución, obtenemos todas sumando múltiplos enteros de 2π.

Adoptamos la definición siguiente:

DEFINICIÓN 1. *Se entiende por argumento de un número complejo* $z\neq 0$ *cualquier solución de la ecuación* $e(\theta)=z/|z|$.

De acuerdo con esta definición, un número complejo $\neq 0$ tiene infinitos argumentos, los cuales difieren unos de otros en múltiplos de 2π. De la definición y de la ecuación funcional [1-31] se obtiene

$$\arg(z_1z_2)=\arg z_1+\arg z_2+n\cdot 2\pi, \qquad [1\text{-}32]$$

donde no puede omitirse el múltiplo aditivo de 2π.

En muchos casos resulta molestísimo tener que manejar infinitos valores del argumento. Existen dos procedimientos para obviar esta dificultad. Primero, puede observarse que existe exactamente un valor del argumento en cualquier intervalo de longitud 2π con tal que se incluya uno (y solo uno) de los extremos. Así, p. ej., podemos escoger el valor que está determinado por la siguiente condición: $-\pi<\arg z\leqq\pi$; se suele hacer referencia al mismo como *valor principal* del argumento. El método tiene el inconveniente de que este argumento no varía en forma continua con z; de hecho, el valor principal salta de $-\pi$ a π cuando z cruza el eje real negativo.

El segundo método es más refinado. Se conviene en identificar cualquier par de números reales cuya diferencia es un múltiplo de 2π, y en llamar cada clase de números identificados un *número real módulo* 2π. Resulta inmediato cómo han de sumarse y restarse tales números. El convenio nos permite interpretar $\arg z$ como un único número real módulo 2π; además, [1-31] tomará entonces la

forma sencilla y deseada

$$\arg(z_1z_2)=\arg z_1+\arg z_2. \qquad [1\text{-}33]$$

La desventaja de este método reside en el hecho de que $\arg z$ ya no es un número real en el sentido ordinario. Por tanto, no se puede sustituir $\arg z$, sin más justificación, en una función como valor de una variable real.

Naturalmente, estas interpretaciones del argumento son fundamentalmente equivalentes. Es una cuestión de conveniencia cuál de las dos interpretaciones elegir en cada caso, y no sería acertado que nos inclinásemos definitivamente por una u otra.

4. *Rectas, semiplanos y ángulos.*—Una *recta* se puede dar en el plano complejo mediante una ecuación paramétrica $z=a+bt$, donde a y b son números complejos, y $b\neq 0$; el parámetro t toma todos los valores reales. Dos ecuaciones $z=a+bt$ y $z=a'+b't$ representan la misma recta si (y solo si) $a'-a$ y b' son múltiplos reales de b. Las rectas son *paralelas* siempre y cuando b' sea un múltiplo real de b, y están *igualmente orientadas* si b' es un múltiplo positivo de b. La última distinción hace posible, evidentemente, el considerar *rectas orientadas*. La *dirección* de una recta orientada puede identificarse con $\arg b$.

Los *semiplanos superior* e *inferior* están caracterizados por las condiciones $\operatorname{Im} z>0$ e $\operatorname{Im} z<0$, respectivamente. Análogamente, los semiplanos de la derecha y de la izquierda están determinados por $\operatorname{Re} z>0$ y $\operatorname{Re} z<0$. Los puntos de la recta $z=a+bt$ satisfacen la ecuación $\operatorname{Im}(z-a)/b=0$. Parece, pues, natural decir que los puntos con $\operatorname{Im}(z-a)/b>0$ y los puntos con $\operatorname{Im}(z-a)/b<0$ están en diferentes semiplanos de los determinados por la recta. Si la recta está orientada, se conviene en llamar al semiplano con $\operatorname{Im}(z-a)/b>0$ *semiplano izquierdo*, y al otro, *semiplano derecho*. Es importante probar que esta distinción es independiente de la representación paramétrica. Supongamos que $z=a+bt$ y $z=a'+b't$ representan la misma recta orientada; entonces $(a-a')/b$ es real y b'/b es positivo. De la primera condición se sigue que $\operatorname{Im}(z-a)/b=\operatorname{Im}(z-a')/b$, y por la segunda $\operatorname{Im}(z-a')/b'$ tiene igual signo que $\operatorname{Im}(z-a')/b$. Se concluye que los semiplanos derecho e izquierdo están determinados de manera única.

El *segmento rectilíneo* que une dos puntos z_1 y z_2 es el conjunto de puntos $z=tz_1+(1-t)z_2$ con $0\leqq t\leqq 1$. No es difícil probar que el segmento rectilíneo que une dos puntos en el mismo semiplano con respecto a una recta está enteramente contenido en dicho semiplano, en tanto que un segmento rectilíneo que une puntos en semiplanos distintos debe cortar a la recta dada.

La palabra *ángulo* tiene al menos dos significados. Primero, se puede hablar de ángulo entre dos rectas orientadas. Si las rectas están dadas por $z=a_1+b_1t$ y $z=a_2+b_2t$, se define el ángulo entre ellas como $\arg b_2/b_1$. Observemos que el ángulo depende del orden en que se mencionan las rectas. Además, o el ángulo tiene infinitos valores, o ha de interpretarse como un número real módulo 2π.

En una segunda interpretación, ángulo significa la porción del plano comprendida entre dos semirrectas con el origen común. Sea a un punto y φ_1 y φ_2 dos números reales que satisfacen la condición $0<\varphi_2-\varphi_1\leqq 2\pi$. Los puntos $z\neq a$, para los que un valor de $\arg(z-a)$ satisface la desigualdad $\varphi_1<\arg(z-a)<\varphi_2$, forman un *sector angular*, que denotamos por $S_a(\varphi_1,\varphi_2)$. Se dice que está contenido entre las semirrectas orientadas de direcciones φ_1 y φ_2, nombradas en este orden, y su *medida angular* es $\varphi_2-\varphi_1$.

Recíprocamente, cualquier par de semirrectas distintas con origen en un mismo punto determinan dos sectores. Si las rectas son $z=a+b_1t$, $z=a+b_2t$, podemos elegir como φ_1 un valor arbitrario de $\arg b_1$; como valor de $\varphi_2=\arg b_2$ elegimos el valor que está contenido en el intervalo $(\varphi_1,\varphi_1+2\pi)$. Los sectores entre las semirrectas son entonces $S_a(\varphi_1,\varphi_2)$ y $S_a(\varphi_2,\varphi_1+2\pi)$, siendo sus medidas angulares $\varphi_2-\varphi_1$ y $2\pi-(\varphi_2-\varphi_1)$. A no ser que las semirrectas tengan direcciones opuestas, una (y solo una) de las medidas angulares es $<\pi$; el sector correspondiente es, por definición, el *ángulo convexo* entre las semirrectas. Es evidente que esta definición no depende del orden de las semirrectas.

Como aplicación probaremos que la suma de los ángulos de un triángulo es π. Obviamente, los ángulos tienen que interpretarse como las medidas angulares de los ángulos convexos entre los lados. Sean los vértices z_1, z_2, z_3; se supone que no están alineados. Es conveniente introducir la notación

$$\lambda_1=\frac{z_2-z_1}{z_3-z_1},\qquad \lambda_2=\frac{z_3-z_2}{z_1-z_2},\qquad \lambda_3=\frac{z_1-z_3}{z_2-z_3}$$

El ángulo en z_i $(i=1,2,3)$ es un valor de $\pm\arg\lambda_i$. Utilizando la identidad $(1/\lambda_1)+\lambda_2=1$, se obtiene que $\operatorname{Im} 1/\lambda_1=-\operatorname{Im}\lambda_2$; por consiguiente, $\operatorname{Im}\lambda_1$ e $\operatorname{Im}\lambda_2$ tienen el mismo signo. Por permutación cíclica, $\operatorname{Im}\lambda_3$ debe tener el mismo signo, y se sigue que el signo en $\pm\arg\lambda_i$ debe ser el mismo para todo i. De donde la suma de los ángulos es $\pm\arg(\lambda_1\lambda_2\lambda_3)$, salvo un múltiplo de 2π. Pero $\lambda_1\lambda_2\lambda_3=-1$; se concluye que la suma de los ángulos es $\pm\pi+n\cdot 2\pi$. Por otra parte, cada ángulo está comprendido entre 0 y π, luego la suma debe estar comprendida entre 0 y 3π. El único múltiplo impar de π entre estos límites es π, con lo cual queda demostrado que la suma de los ángulos es π.

EJERCICIOS

1. Pruébese que los lados y ángulos opuestos de un paralelogramo son respectivamente iguales y que las diagonales se bisecan entre sí.
2. Demuéstrese que los ángulos en la base de un triángulo isósceles son iguales.
3. Pruébese que los n valores de $\sqrt[n]{a}$ son los vértices de un polígono regular (lados y ángulos iguales).

5. *Representación esférica.*—Para muchos propósitos es útil ampliar el sistema $\mathbb{C}$ de los números complejos mediante la introducción de un símbolo ∞ que represente infinito. Su relación con los números finitos se establece poniendo $a+\infty=\infty+a=\infty$ para todo a finito, y $b\cdot\infty=\infty\cdot b=\infty$ para todo $b\neq 0$, incluyendo $b=\infty$. Es imposible, sin embargo, definir $\infty+\infty$ y $0\cdot\infty$ sin violar las reglas aritméticas. No obstante, escribiremos, por convenio especial, $a/0=\infty$ para $a\neq 0$ y $b/\infty=0$ para $b\neq\infty$.

En el plano no hay lugar para un punto correspondiente a ∞, pero podemos, naturalmente, introducir un punto "ideal", que llamaremos *punto del infinito.* Los puntos del plano, junto con el punto del infinito, constituyen el *plano complejo ampliado* [1]. Convenimos en que toda recta pasa por el punto del infinito. En contraste con esto, ningún semiplano contendrá el punto ideal.

Es deseable introducir un modelo geométrico en el cual todos los puntos del plano ampliado tengan un representante concreto. Con este objeto consideremos la esfera unidad S, cuya ecuación en el espacio tridimensional es $x_1^2+x_2^2+x_3^2=1$. A cada punto de S, ex-

[1] A lo largo del texto utilizaremos indistintamente las expresiones *plano ampliado* y *plano completo,* que son equivalentes. *(N. del T.)*

cepto el $(0, 0, 1)$, podemos asociar un número complejo

$$z = \frac{x_1 + ix_2}{1 - x_3}, \qquad [1\text{-}34]$$

y esta correspondencia es biunívoca. Realmente, de [1-34] obtenemos

$$|z|^2 = \frac{x_1^2 + x_2^2}{(1 - x_3)^2} = \frac{1 + x_3}{1 - x_3},$$

y, por tanto,

$$x_3 = \frac{|z|^2 - 1}{|z|^2 + 1} \qquad [1\text{-}35]$$

Cálculos posteriores nos dan

$$x_1 = \frac{z + \bar{z}}{1 + |z|^2}$$
$$x_2 = \frac{z - \bar{z}}{i(1 + |z|^2)} \qquad [1\text{-}36]$$

La correspondencia puede completarse haciendo corresponder al punto del infinito el $(0, 0, 1)$, siendo posible de este modo considerar la esfera como una representación del plano ampliado o del sistema ampliado de números. Observemos que la semiesfera $x_3 < 0$ corresponde al disco $|z| < 1$, y la semiesfera $x_3 > 0$ a su exterior $|z| > 1$. En la teoría de funciones la esfera S se denomina *esfera de Riemann.*

Si se identifica el plano complejo con el plano (x_1, x_2), con los ejes x_1 y x_2 correspondiendo a los ejes real e imaginario, respectivamente, la transformación [1-34] tiene un sencillo significado geométrico. Escribiendo $z = x + iy$, se comprueba que

$$x : y : -1 = x_1 : x_2 : x_3 - 1, \qquad [1\text{-}37]$$

y esto significa que los puntos $(x, y, 0)$, (x_1, x_2, x_3) y $(0, 0, 1)$ están alineados. Por consiguiente, la correspondencia es una proyección central desde el centro $(0, 0, 1)$ como se muestra en la figura 1-3. Se le llama *proyección estereográfica.*

No existe interpretación sencilla de la adición y de la multipli-

cación en la representación esférica. Sus ventajas residen en el hecho de que el punto del infinito ya no es un punto distinguido.

Es geométricamente evidente que la proyección estereográfica transforma toda recta del plano z en una circunferencia en S que pasa por el polo $(0, 0, 1)$, siendo también cierta la recíproca. Con más generalidad, cualquier circunferencia sobre la esfera correspon-

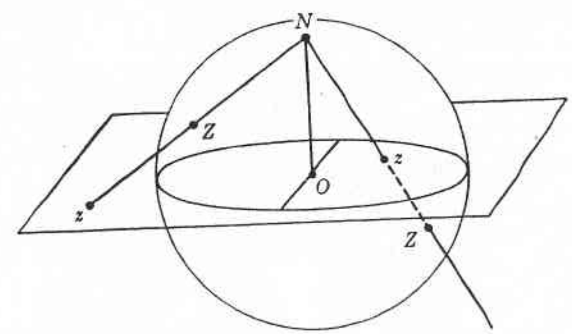

FIG. 1-3.—Proyección estereográfica.

de a una circunferencia o a una recta del plano de las z. Para demostrarlo, observemos que una circunferencia sobre la esfera está en un plano $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \alpha_0$, donde cabe suponer que $\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1$ y $0 \leqq \alpha_0 < 1$. En función de z y $\bar{z}$, esta ecuación toma la forma

$$\alpha_1(z + \bar{z}) - \alpha_2 i(z - \bar{z}) + \alpha_3(|z|^2 - 1) = \alpha_0(|z|^2 + 1)$$

o

$$(\alpha_0 - \alpha_3)(x^2 + y^2) - 2\alpha_1 x - 2\alpha_2 y + \alpha_0 + \alpha_3 = 0.$$

Para $\alpha_0 \neq \alpha_3$, esta es la ecuación de una circunferencia, representando una recta para $\alpha_0 = \alpha_3$. Recíprocamente, la ecuación de cualquier circunferencia o recta puede escribirse en esta forma. La correspondencia es, por tanto, biunívoca.

Es fácil calcular la distancia $d(z, z')$ entre las proyecciones estereográficas de z y z'. Si se denotan los puntos de la esfera por (x_1, x_2, x_3) y (x'_1, x'_2, x'_3), tenemos, en primer lugar,

$$(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + (x_3 - x'_3)^2 = 2 - 2(x_1 x'_1 + x_2 x'_2 + x_3 x'_3).$$

A partir de [1-35] y [1-36] se obtiene mediante un sencillo cálculo

$$x_1x_1' + x_2x_2' + x_3x_3' =$$

$$= \frac{(z+\bar{z})(z'+\bar{z}') - (z-\bar{z})(z'-\bar{z}') + (|z|^2-1)(|z'|^2-1)}{(1+|z|^2)(1+|z'|^2)} =$$

$$= \frac{(1+|z|^2)(1+|z'|^2) - 2|z-z'|^2}{(1+|z|^2)(1+|z'|^2)}$$

Como resultado se deduce que

$$d(z,z') = \frac{2|z-z'|}{\sqrt{(1+|z|^2)(1+|z'|^2)}} \qquad [1\text{-}38]$$

Para $z' = \infty$, la fórmula correspondiente es

$$d(z,\infty) = \frac{2}{\sqrt{1+|z|^2}}$$

EJERCICIOS

1. Pruébese que z y z' corresponden a puntos diametralmente opuestos sobre la esfera de Riemann si (y solo si) $z\bar{z}' = -1$.

2. Un cubo tiene sus vértices sobre la esfera S, y sus lados, paralelos a los ejes coordenados. Hállense las proyecciones estereográficas de los vértices.

3. Problema análogo para un tetraedro regular en posición general.

4. Sean Z y Z' las proyecciones estereográficas de z y z', y N el polo norte. Pruébese que los triángulos NZZ' y Nzz' son semejantes y utilícese este hecho para deducir [1-38].

5. Hállese el radio de la imagen esférica de la circunferencia del plano de centro a y radio R.

1-3. Transformaciones lineales.—Una *transformación* establece una correspondencia entre números o puntos. No difiere en absoluto de una función, pero se utiliza el término transformación cuando se desea dar la idea del reemplazamiento real de un punto por otro. Un ejemplo sencillo de transformación es un desplazamiento paralelo del plano complejo que reemplaza el punto z por el $z+a$. Se pueden imaginar los puntos z y $z+a$ situados en el mismo o en diferentes planos. En el primer caso, una notación adecuada sería $z \to z+a$; en el segundo, deberíamos utilizar la notación funcional $w = z+a$. Con mayor generalidad, si se denota una transformación por T, escribiremos $z \to Tz$ o $w = Tz$.

El lector observará que en esta sección no se hace ninguna consideración relativa al concepto de continuidad.

1. *El grupo lineal.*—Consideremos la transformación

$$w = \frac{az+b}{cz+d}, \qquad [1\text{-}39]$$

cuyos coeficientes a, b, c, d son números complejos. No deseamos que w sea independiente de z, imponiendo por esta razón la condición básica de que $ad - bc \neq 0$. Esta hipótesis evita también que el denominador sea idénticamente nulo, y w está bien definido excepto para el único valor $z = -d/c$ en el caso de que $c \neq 0$.

Con objeto de establecer una correspondencia entre los planos ampliados, añadimos los siguientes valores convencionales a los definidos por [1-39]: 1) si $c \neq 0$, $w = \infty$ para $z = -d/c$ y $w = a/c$ para $z = \infty$; 2) si $c = 0$, $w = \infty$ para $z = \infty$. La transformación ampliada $w = Tz$ se llama una *transformación lineal.*

La ecuación [1-39] puede resolverse con respecto a z dando

$$z = \frac{dw-b}{-cw+a} \qquad [1\text{-}40]$$

Esta transformación puede ampliarse de la misma forma; la transformación lineal resultante es inversa de la T, y se denota, por consiguiente, por T^{-1}. La existencia de una inversa muestra que la correspondencia definida por T es biunívoca.

T está determinada por una matriz de segundo orden

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

cuyo determinante $ad - bc \neq 0$. Está también determinada por cualquier múltiplo no nulo

$$\begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}, \qquad \lambda \neq 0$$

de la misma matriz.

La notación matricial resulta conveniente principalmente porque conduce a la determinación sencilla de una transformación compuesta $w=T_1T_2z$. Si utilizamos subíndices para distinguir entre las matrices correspondientes a T_1 y T_2, es fácil comprobar que T_1T_2 está determinada por la matriz producto

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2+b_1c_2 & a_1b_2+b_1d_2 \\ c_1a_2+d_1c_2 & c_1b_2+d_1d_2 \end{pmatrix}.$$

La comprobación se hace trivial si se pone $z=z_1/z_2$, $w=w_1/w_2$ y se escribe [1-39] en la forma

$$w_1=az_1+bz_2$$
$$w_2=cz_1+dz_2.$$

Esta notación homogénea no se utilizará en lo que sigue.

Todas las transformaciones lineales forman grupo. En efecto, la propiedad asociativa $(T_1T_2)T_3=T_1(T_2T_3)$ se verifica para transformaciones arbitrarias, la identidad $w=z$ es una transformación lineal y la inversa de una transformación lineal es lineal.

Las transformaciones lineales más sencillas están dadas por matrices de la forma

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

La primera de ellas, $w=z+\alpha$, se llama *traslación paralela*. La segunda, $w=kz$, es una *rotación* si $|k|=1$ y una *homotecia* si $k>0$. Para un complejo cualquiera $k\neq 0$ podemos poner $k=|k|\cdot k/|k|$ y, por consiguiente, $w=kz$ puede representarse como el resultado de una homotecia seguida de una rotación. La tercera transformación, $w=1/z$, se llama una *inversión*.

Si $c\neq 0$, podemos escribir

$$\frac{az+b}{cz+d}=\frac{bc-ad}{c^2(z+d/c)}+\frac{a}{c},$$

y esta descomposición muestra que la transformación lineal más general se compone de una traslación, una inversión, una rotación y una homotecia seguida de otra traslación. Si $c=0$, desaparece la inversión y no es necesaria la última traslación.

EJERCICIOS

1. Demuéstrese que la simetría $z\to\bar{z}$ no es una transformación lineal.

2. Si

$$T_1z=\frac{z+2}{z+3}, \qquad T_2z=\frac{z}{z+1},$$

hállense T_1T_2z, T_2T_1z y $T_1^{-1}T_2z$.

3. Pruébese que la transformación más general que deja el origen fijo y conserva todas las distancias es una rotación o una rotación seguida de una simetría respecto al eje real.

4. Pruébese que cualquier transformación lineal que transforma el eje real en sí mismo puede escribirse con coeficientes reales.

2. *La razón doble.*—Dados tres puntos distintos z_2, z_3, z_4 en el plano ampliado, existe una transformación lineal T que transforma estos puntos en los 0, 1, ∞. Si ninguno de los puntos dados es ∞, T podrá escribirse

$$Tz=\frac{z-z_2}{z-z_4}:\frac{z_3-z_2}{z_3-z_4} \qquad [1\text{-}41]$$

Si z_2, z_3 o $z_4=\infty$, la transformación se reduce, respectivamente, a

$$\frac{z_3-z_4}{z-z_4}, \quad \frac{z-z_2}{z-z_4}, \quad \frac{z-z_2}{z_3-z_2}$$

Si S fuera otra transformación con la misma propiedad, entonces ST^{-1} dejaría los puntos 0, 1, ∞ invariantes. Para

$$ST^{-1}=\frac{az+b}{cz+d}$$

estas condiciones implican $b=c=0$, $a=d$. Por tanto, ST^{-1} se reduciría a la identidad, y tendríamos $S=T$. Se concluye que T está determinada de manera única.

DEFINICIÓN 2. *La razón doble (z_1, z_2, z_3, z_4) es la imagen de z_1 mediante la transformación lineal que transforma z_2, z_3, z_4 en* 0, 1, ∞.

La definición tiene sentido únicamente si z_2, z_3 y z_4 son distintos. Se puede definir un valor convencional de [1-42] cuando tres cualesquiera de los cuatro puntos son distintos; pero esto carece de importancia para nuestros propósitos.

La razón doble es invariante con respecto a las transformaciones lineales. Con una formulación más precisa:

TEOREMA 1. *Si z_1, z_2, z_3, z_4 son puntos distintos en el plano ampliado y S una transformación lineal cualquiera, se tiene que* $(Sz_1, Sz_2, Sz_3, Sz_4) = (z_1, z_2, z_3, z_4)$.

La demostración es inmediata, pues si $Tz = (z, z_2, z_3, z_4)$, entonces TS^{-1} transforma Sz_2, Sz_3, Sz_4 en 0, 1, ∞. Por consiguiente, tenemos por la definición

$$(Sz_1, Sz_2, Sz_3, Sz_4) = TS^{-1}(Sz_1) = Tz_1 = (z_1, z_2, z_3, z_4).$$

Con la ayuda de esta propiedad podemos escribir inmediatamente la transformación lineal que transforma tres puntos dados z_1, z_2, z_3 en otros w_1, w_2, w_3 prescritos. La correspondencia viene realmente dada por

$$(w, w_1, w_2, w_3) = (z, z_1, z_2, z_3). \qquad [1\text{-}42]$$

En general, es naturalmente necesario resolver esta ecuación con respecto a w.

TEOREMA 2. *La razón doble (z_1, z_2, z_3, z_4) es real si (y solo si) los cuatro puntos son concíclicos o están alineados.*

Esto es evidente por geometría elemental, pues se obtiene

$$\arg\,(z_1, z_2, z_3, z_4) = \arg \frac{z_1 - z_3}{z_1 - z_4} - \arg \frac{z_3 - z_2}{z_3 - z_4} = 0 \text{ o } \pi,$$

dependiendo de la posición relativa de los puntos.

Para una demostración analítica necesitamos únicamente probar que la imagen del eje real respecto a cualquier transformación lineal es o una circunferencia o una recta. En efecto, $Tz = (z, z_2, z_3, z_4)$ es real cuando está sobre la imagen del eje real bajo la transformación T^{-1} y en ninguna otra parte.

Los valores de $w = T^{-1}z$ para z real satisfacen la ecuación $Tw = \overline{Tw}$. De manera explícita, esta condición es de la forma

$$\frac{aw + b}{cw + d} = \frac{\bar{a}\bar{w} + \bar{b}}{\bar{c}\bar{w} + \bar{d}}$$

Multiplicando en cruz, se obtiene

$$(a\bar{c} - c\bar{a})|w|^2 + (a\bar{d} - c\bar{b})w + (b\bar{c} - d\bar{a})\bar{w} + b\bar{d} - d\bar{b} = 0.$$

Si $a\bar{c} - c\bar{a} = 0$, esta es la ecuación de una recta, pues con esta condición el coeficiente $a\bar{d} - c\bar{b}$ no puede anularse también. Si $a\bar{c} - c\bar{a} \neq 0$, podemos dividir por este coeficiente y completar el cuadrado. Después de un sencillo cálculo obtenemos

$$\left| w - \frac{\bar{a}d - \bar{c}b}{\bar{a}c - \bar{c}a} \right| = \left| \frac{ad - bc}{\bar{a}c - \bar{c}a} \right|,$$

que es la ecuación de una circunferencia.

El último resultado pone de manifiesto que no se debe, en la teoría de transformaciones lineales, distinguir entre circunferencias y rectas. Una justificación posterior está en el hecho de que ambas líneas corresponden a circunferencias en la esfera de Riemann. De acuerdo con esto, convendremos en utilizar la palabra circunferencia en este sentido más amplio [1].

El siguiente es un corolario inmediato de los teoremas 1 y 2:

TEOREMA 3. *Una transformación lineal transforma circunferencias en circunferencias.*

EJERCICIOS

1. Hállese la transformación lineal que transforma 0, i, $-i$ en 1, -1, 0.

2. Exprésense las razones dobles correspondientes a las 24 permutaciones de cuatro puntos, en función de $\lambda = (z_1, z_2, z_3, z_4)$.

3. Si los vértices consecutivos z_1, z_2, z_3, z_4 de un cuadrilátero están situados sobre una circunferencia, demuéstrese que

$$|z_1 - z_3| \cdot |z_2 - z_4| = |z_1 - z_2| \cdot |z_3 - z_4| + |z_2 - z_3| \cdot |z_1 - z_4|$$

e interprétese el resultado geométricamente.

4. Pruébese que cuatro puntos distintos cualesquiera pueden transformarse mediante una transformación lineal en los 1, -1, k, $-k$, dependiendo k de los puntos. ¿Cuántas soluciones hay y cómo están relacionadas?

[1] Este convenio será únicamente válido cuando se trate de transformaciones lineales.

3. *Simetría.*—Los puntos z y $\bar{z}$ son simétricos con relación al eje real. Una transformación lineal con coeficientes reales transforma el eje real en sí mismo, y z, $\bar{z}$, en puntos que son nuevamente simétricos. Más generalmente, si una transformación lineal T transforma el eje real en una circunferencia C, diremos que los puntos $w=Tz$ y $w^*=T\bar{z}$ son *simétricos con respecto a* C. Esta es una relación entre w, w^* y C, que no depende de T. Pues si S es otra transformación que transforma el eje real en C, entonces $S^{-1}T$ es una transformación real y, por consiguiente $S^{-1}w=S^{-1}Tz$ y $S^{-1}w^*=S^{-1}T\bar{z}$ son también conjugados. Se puede entonces definir la simetría en los siguientes términos:

DEFINICIÓN 3. *Los puntos z y z^* se dicen simétricos con respecto a la circunferencia C, que pasa por los puntos z_1, z_2, z_3, si (y solo si)* $(z^*, z_1, z_2, z_3) = \overline{(z, z_1, z_2, z_3)}$.

Los puntos de C, y solo ellos, son simétricos de sí mismos. La aplicación que transforma z en z^* es una correspondencia biunívoca, y se llama *simetría* o *reflexión* respecto a C. Evidentemente, el producto de dos simetrías es una transformación lineal.

Deseamos investigar el significado geométrico de la simetría. Supongamos, en primer lugar, que C es una recta. Entonces podemos elegir $z_3=\infty$, y la condición de simetría se convierte en

$$\frac{z^*-z_1}{z_2-z_1}=\frac{\bar{z}-\bar{z}_1}{\bar{z}_2-\bar{z}_1} \qquad [1\text{-}43]$$

Tomando valores absolutos obtenemos $|z^*-z_1|=|z-z_1|$. Aquí z_1 puede ser cualquier punto finito de C, y concluimos que z y z^* equidistan de todos los puntos de C. Por [1-43] tenemos además

$$\operatorname{Im}\frac{z^*-z_1}{z_2-z_1}=-\operatorname{Im}\frac{z-z_1}{z_2-z_1},$$

y, por tanto, z y z^* están en semiplanos distintos con respecto a C [1]. Dejamos al lector la demostración de que C es la mediatriz del segmento determinado por z y z^*.

Consideremos ahora el caso de una circunferencia finita C de centro a y radio R. El uso sistemático de la invariancia de la razón

[1] A no ser que coincidan y pertenezcan a C.

doble nos permite llegar a

$$\overline{(z, z_1, z_2, z_3)} = \overline{(z-a, z_1-a, z_2-a, z_3-a)} =$$
$$=\left(\bar{z}-\bar{a}, \frac{R^2}{z_1-a}, \frac{R^2}{z_2-a}, \frac{R^2}{z_3-a}\right)=$$
$$=\left(\frac{R^2}{\bar{z}-\bar{a}}, z_1-a, z_2-a, z_3-a\right)=\left(\frac{R^2}{\bar{z}-\bar{a}}+a, z_1, z_2, z_3\right).$$

Esta ecuación muestra, evidentemente, que el punto simétrico de z es $z^*=R^2/(\bar{z}-\bar{a})+a$, o que z y z^* satisfacen la relación

$$(z^*-a)(\bar{z}-\bar{a})=R^2. \qquad [1\text{-}44]$$

El producto $|z^*-a|\cdot|z-a|$ de las distancias al centro es, por tanto, R^2. Resulta, además, que la razón $(z^*-a)/(z-a)$ es positiva, lo que significa que z y z^* están en la misma semirrecta a partir de a. Existe una sencilla construcción geométrica del punto simétrico de z (Fig. 1-4). Obsérvese que el punto simétrico de a es ∞.

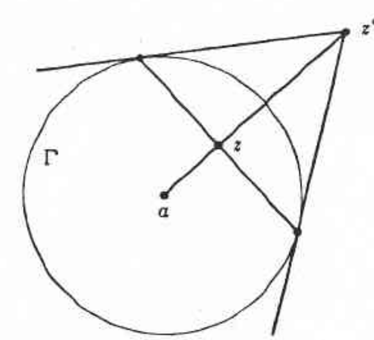

FIG. 1-4.—Simetría respecto a una circunferencia.

TEOREMA 4. *(Principio de simetría.) Si una transformación lineal transforma la circunferencia C_1 en la circunferencia C_2, entonces transforma cualquier par de puntos simétricos con respecto a C_1 en un par de puntos simétricos con respecto a C_2.*

Brevemente, las transformaciones lineales conservan la simetría. Si C_1 o C_2 es el eje real, el principio se sigue de la definición de simetría. En el caso general, la afirmación se prueba utilizando una transformación intermedia que transforme C_1 en el eje real.

Hay dos formas en las que se puede utilizar el principio de simetría. Si se conocen las imágenes de z y C bajo una cierta transformación lineal, entonces el principio nos permite hallar la imagen de z^*. Si, por otra parte, se conocen las imágenes de z y z^*, concluimos que estas imágenes deben ser simétricas respecto a la imagen de C. Aunque esto no es suficiente para determinar la imagen de C, la información obtenida es, sin embargo, valiosa.

Se pone en práctica el principio de simetría en el problema de hallar la transformación lineal que transforma la circunferencia C en la circunferencia C'. Siempre podremos determinar la transformación exigiendo que tres puntos z_1, z_2, z_3 de C se transformen en tres puntos w_1, w_2, w_3 de C'; la transformación es entonces $(w, w_1, w_2, w_3) = (z, z_1, z_2, z_3)$. Pero la transformación está también determinada si imponemos que a un punto z_1 de C corresponda un punto w_1 de C' y que un punto z_2 que *no esté* en C se transforme en un punto w_2 que *no esté* en C'. Sabemos que entonces z_2^* (el punto simétrico de z_2 con respecto a C) debe corresponder a w_2^* (el punto simétrico de w_2 con respecto a C'). Luego la transformación se obtendrá mediante la relación $(w, w_1, w_2, w_2^*) = (z, z_1, z_2, z_2^*)$.

EJERCICIOS

1. Demuéstrese que toda simetría transforma circunferencias en circunferencias.

2. Hállese la simétrica del eje imaginario, de la recta $x=y$ y de la circunferencia $|z|=1$ respecto a la circunferencia $|z-2|=1$.

3. Efectúense las simetrías del ejercicio precedente mediante construcciones geométricas.

4. Hállese la transformación lineal que transforma la circunferencia $|z|=2$ en la $|z+1|=1$, el punto -2 en el origen y el origen en i.

5. Hállese la transformación lineal más general que transforma la circunferencia $|z|=R$ en sí misma.

6. Una transformación lineal transforma un par de circunferencias concéntricas en otro par de circunferencias concéntricas. Demuéstrese que las razones de los radios deben ser iguales.

7. Hállese una transformación lineal que transforme $|z|=1$ y $|z-\frac{1}{4}|=\frac{1}{4}$ en circunferencias concéntricas. ¿Cuál es la razón de los radios?

8. El mismo problema para $|z|=1$ y $x=2$.

4. *Tangentes, orientación y ángulos.*—Se dice que dos circunferencias son *tangentes entre sí* si tienen un único punto común. Si mediante una transformación lineal se lleva el punto común al infinito, las circunferencias se convierten en rectas paralelas. Por la transformación inversa, la familia de rectas paralelas se transforma en una familia de circunferencias tangentes dos a dos.

Deseamos trazar, p. ej., la tangente a una circunferencia en uno cualquiera de sus puntos. Sea la circunferencia C y consideremos un punto z_0 en C. La transformación $z \to 1/(z-z_0)$ transforma C en una recta L e ∞ en 0. Trazamos una recta L' por 0 y paralela a L. La transformación inversa transforma L' en una circunferencia que es tangente a C en z_0 y pasa por ∞. En otras palabras, la circunferencia es tangente a la recta en z_0.

Vamos a definir el ángulo formado por dos circunferencias que se cortan como el ángulo formado por sus tangentes. Con este objeto es necesario orientar las tangentes, y la orientación de una tangente debe depender evidentemente de la orientación de la circunferencia.

Una orientación de la circunferencia C está determinada por una terna ordenada de puntos z_1, z_2, z_3 de C. Con respecto a esta orientación, se dice que un punto no situado sobre C está a la *izquierda* de C si $\operatorname{Im}(z, z_1, z_2, z_3) > 0$, y a la *derecha* de C, si $\operatorname{Im}(z, z_1, z_2, z_3) < 0$.

Es esencial demostrar que solo hay dos orientaciones diferentes. Con esto queremos decir que la distinción entre derecha e izquierda es la misma para todas las ternas, en tanto que los significados pueden invertirse. A causa de la invarianza de la razón doble, es suficiente considerar el caso en que C es el eje real. Tenemos entonces que examinar $\operatorname{Im}(z, z_1, z_2, z_3)$. Escribiendo

$$(z, z_1, z_2, z_3) = \frac{az+b}{cz+d}$$

con coeficientes reales, mediante un sencillo cálculo, obtenemos

$$\operatorname{Im}(z, z_1, z_2, z_3) = \frac{ad-bc}{|cz+d|^2} \cdot \operatorname{Im} z.$$

Luego la distinción entre derecha e izquierda es idéntica a la distinción entre semiplanos superior e inferior.

Una transformación lineal T transforma la circunferencia orientada C en una circunferencia que orientamos mediante la terna Tz_1, Tz_2, Tz_3. De la invarianza de la razón doble se sigue que la izquierda y la derecha de C corresponden a la izquierda y a la derecha de la circunferencia imagen.

Si C es una recta, puede orientarse por medio de la terna z_1, z_2, ∞. Los puntos z a la izquierda están caracterizados por la desigualdad

$$\operatorname{Im} \frac{z-z_1}{z_2-z_1} > 0.$$

Relacionando esto con nuestra definición primitiva, encontramos que estos puntos se encuentran en el semiplano izquierdo, determinado por la recta orientada $z=z_1+t(z_2-z_1)$. Una recta orientada según nuestra primera definición puede considerarse como un caso especial de circunferencia orientada.

En general, no hay procedimiento ni existe razón alguna para comparar las orientaciones de dos circunferencias. Un hecho excepcional ocurre si las circunferencias son tangentes entre sí. En este caso pueden transformarse en rectas paralelas, y se dice que las circunferencias están igualmente orientadas si corresponden a rectas con la misma orientación. En este sentido una circunferencia orientada induce una orientación en sus tangentes, lo cual nos permite definir el ángulo entre dos circunferencias orientadas en el punto de intersección como el ángulo que forman las tangentes orientadas.

Teorema 5. *Una transformación lineal conserva los ángulos entre circunferencias orientadas.*

Sean z_1 y z_2 las intersecciones de C y C'. Elegimos z_0 en C y z'_0 en C' y fijamos las orientaciones mediante las ternas z_1, z_0, z_2 y z_1, z'_0, z_2. El teorema quedará demostrado si probamos que el ángulo en z_1 entre C y C', tomadas en este orden, viene dado por $\arg(z'_0, z_1, z_0, z_2)$.

La afirmación es fácil de comprobar si z_1 o z_2 es igual a ∞. Entonces, C y C' son rectas, la razón doble se reduce a $(z'_0-z_2)/(z_0-z_2)$ o bien a $(z'_0-z_1)/(z_0-z_1)$ y el argumento de uno u otro cociente es, por definición, igual al ángulo entre las rectas orientadas. En el caso general, sean L y L' las tangentes orientadas de C y C' en z_1. Sabemos que una transformación lineal que transforma z_1 en ∞ transforma L, C y L', C' en pares de rectas paralelas e igualmente orientadas. Por la invariancia de la razón doble y el resultado precedente, cuando $z_1=\infty$ se concluye que $\arg(z'_0, z_1, z_0, z_2)$ es igual a la expresión correspondiente para L y L'. Por otra parte, nuestra afirmación ha sido demostrada para L y L', y por definición el ángulo entre C y C' es el mismo que el ángulo entre L y L'. Por consiguiente, el ángulo entre C y C' es realmente igual a $\arg(z'_0, z_1, z_0, z_2)$.

La conservación de la razón doble, simetría y ángulos son las propiedades fundamentales de las transformaciones lineales. Las dos últimas propiedades se ha visto que son consecuencia de la primera.

En la representación geométrica, la orientación z_1, z_2, z_3 puede indicarse por una flecha que vaya de z_1 sobre z_2 a z_3. Con la elección ordinaria del sistema de coordenadas, derecha e izquierda tendrán el significado usual con relación a esta flecha. Cuando se considera el plano complejo sin ampliar como parte del plano ampliado, el punto del infinito es distinguido. Por tanto, podemos definir una orientación positiva absoluta de todas las circunferencias finitas mediante el requerimiento de que ∞ esté situado a la derecha de las circunferencias orientadas. Se dice entonces que los puntos a la izquierda son *interiores* a la circunferencia, mientras que los puntos a la derecha son *exteriores*.

Sobre una esfera de Riemann no hay razón para distinguir entre interior y exterior de una circunferencia.

EJERCICIOS

1. Si z_1, z_2, z_3, z_4 son puntos sobre una circunferencia, demuéstrese que z_1, z_2, z_4 y z_3, z_2, z_4 determinan la misma orientación si (y solo si) $(z_1, z_2, z_3, z_4) > 0$.

2. Hágase la demostración de que dos circunferencias orientadas secantes forman ángulos opuestos en los dos puntos de intersección.

3. Demuéstrese que los puntos interiores a la circunferencia $|z-a|=R$ son todos los que satisfacen la desigualdad $|z-a|<R$.

4. Pruébese que una tangente a una circunferencia es perpendicular al radio en el punto de contacto.

5. *Familias de circunferencias.*—En cuanto a la visualización de una transformación lineal, se puede conseguir mucho mediante la introducción de ciertas familias de circunferencias, que se pueden considerar como curvas coordenadas en un sistema de coordenadas circular.

Consideremos una transformación lineal de la forma

$$w=k\cdot\frac{z-a}{z-b}$$

Aquí $z=a$ corresponde a $w=0$, y $z=b$, a $w=\infty$. Se sigue que las rectas que pasan por el origen del plano de las w son imágenes de las circunferencias que pasan por a y b. Por otra parte, las circunferencias concéntricas con centro en el origen, $|w|=\rho$, corresponden a circunferencias de ecuación

$$\left|\frac{z-a}{z-b}\right|=\rho.$$

Estas son las *circunferencias de Apolonio*, de puntos límites a y b. Por su ecuación, son lugares de puntos cuyas distancias a a y b tienen una razón constante.

Denotemos por C_1 las circunferencias que pasan por a y b, y por C_2, las circunferencias de Apolonio con estos puntos límites. La configuración (Fig. 1-5) formada por todas las circunferencias C_1 y C_2

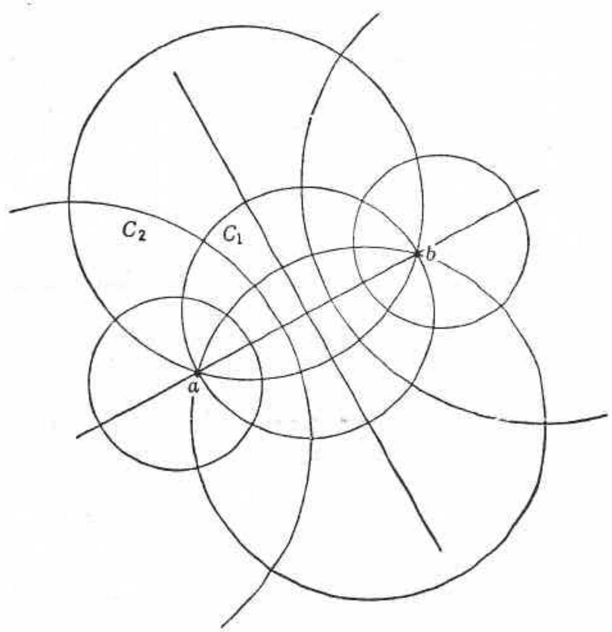

FIG. 1-5.—Circunferencias de Steiner.

se denominará *red circular* o *circunferencias de Steiner*, determinada por a y b. Tiene muchas propiedades interesantes, de las que relacionaremos algunas:

1) Hay exactamente una C_1 y una C_2 que pasan por cada punto del plano, con la excepción de los puntos límites.

2) Toda C_1 corta a toda C_2 en ángulo recto.

3) La simetría respecto a una C_1 transforma toda C_2 en sí misma y toda C_1 en otra C_1. La simetría respecto a una C_2 transforma toda C_1 en sí misma y toda C_2 en otra C_2.

4) Los puntos límites son simétricos con respecto a cada C_2, pero no con respecto a cualquier otra circunferencia.

Estas propiedades son todas triviales cuando los puntos límites son 0 e ∞, es decir, cuando las C_1 son rectas que pasan por el origen y las C_2 circunferencias concéntricas. Puesto que las propiedades son invariantes respecto de las transformaciones lineales, deben seguir verificándose en el caso general.

Si una transformación $w=Tz$ transforma a, b en a', b', puede escribirse en la forma

$$\frac{w-a'}{w-b'}=k\cdot\frac{z-a}{z-b} \qquad [1\text{-}45]$$

Está claro que T transforma las circunferencias C_1 y C_2 en circunferencias C'_1 y C'_2 con puntos límites a' y b'.

La situación es particularmente simple si $a'=a$, $b'=b$. Entonces se dice que a, b son los *puntos fijos* de T, y es conveniente representar z y Tz en el mismo plano. En estas circunstancias toda la red de circunferencias se aplica sobre sí misma. El valor de k sirve para identificar las circunferencias imágenes C'_1 y C'_2. Realmente, con orientaciones apropiadas, C_1 forma el ángulo $\arg k$ con su imagen C'_1, y el cociente de las razones constantes $|a-z|/|z-b|$ en C'_2 y C_2 es $|k|$.

Son particularmente importantes los casos especiales en los que todas las C_1 o todas las C_2 se aplican sobre sí mismas. Se tiene que $C'_1=C_1$ para todo C_1 si $k>0$ (si $k<0$, las circunferencias siguen siendo las mismas, pero se ha invertido la orientación). Se dice entonces que la transformación es *hiperbólica*. Cuando k aumenta, los puntos Tz, $z\neq a, b$ circulan a lo largo de las circunferencias C_1 hacia b. La consideración de esta circulación proporciona una imagen bastante clara de la transformación hiperbólica.

El caso $C'_2=C_2$ tiene lugar cuando $|k|=1$. A las transformaciones con esta propiedad se les llama *elípticas*. Al variar $\arg k$, los puntos Tz se mueven a lo largo de las circunferencias C_2. La circulación correspondiente se verifica en sentidos distintos respecto de a y b.

La transformación lineal general con dos puntos fijos es el producto de una transformación hiperbólica y de otra elíptica con los mismos puntos fijos.

Se hallan los puntos fijos de una transformación lineal resolviendo la ecuación

$$z=\frac{\alpha z+\beta}{\gamma z+\delta} \qquad [1\text{-}46]$$

En general, es esta una ecuación cuadrática con dos raíces; si $\gamma=0$, uno de los puntos fijos es ∞. Puede ocurrir, sin embargo, que las raíces coincidan. Una transformación lineal con puntos fijos coincidentes se llama *parabólica*. La condición correspondiente a este caso es $(\alpha-\delta)^2=4\beta\gamma$.

Si resulta que la ecuación [1-46] tiene dos raíces distintas a y b, la transformación se puede escribir en la forma

$$\frac{w-a}{w-b}=k\,\frac{z-a}{z-b}$$

Podemos entonces utilizar las circunferencias de Steiner determinadas por a y b para discutir la naturaleza de la transformación. Es importante, sin embargo, observar que el método no se limita solo a este caso. Podemos escribir cualquier transformación lineal en la forma [1-45] con a, b arbitrarios y utilizar con gran ventaja las dos redes circulares.

Para la discusión de las transformaciones parabólicas es deseable introducir otro nuevo tipo de red circular. Consideremos la transformación

$$w=\frac{\omega}{z-a}+c.$$

Es evidente que a circunferencias que pasan por a corresponden rectas en el plano w; además, a circunferencias tangentes corresponden rectas paralelas. En particular, si $w=u+iv$, las rectas $u=$constante y $v=$constante corresponden a dos familias de circunferencias tangentes entre sí que se cortan en ángulo recto (Fig. 1-6). Esta configuración puede considerarse como un conjunto degenerado de circunferencias de Steiner. Está determinada por el punto a y la tangente a una de las familias de circunferencias. Denotaremos las imágenes de las rectas $v=$constante por C_1, y las circunferencias de la otra familia, por C_2. Obviamente, la recta $v=\operatorname{Im} c$ corresponde a la tangente de las circunferencias C_1, y su dirección viene dada por $\arg \omega$.

Cualquier transformación que transforme a en a' puede escribirse en la forma

$$\frac{\omega'}{w-a'}=\frac{\omega}{z-a}+c.$$

Está claro que las circunferencias C_1 y C_2 se transforman en circunferencias C'_1 y C'_2 determinadas por a' y ω'. Suponemos ahora que $a=a'$ es el único punto fijo. Entonces $\omega=\omega'$, y podemos escribir

$$\frac{\omega}{w-a}=\frac{\omega}{z-a}+c. \qquad [1\text{-}47]$$

Mediante esta transformación, la configuración constituida por las circunferencias C_1 y C_2 se aplica sobre sí misma. En [1-47] hay factor

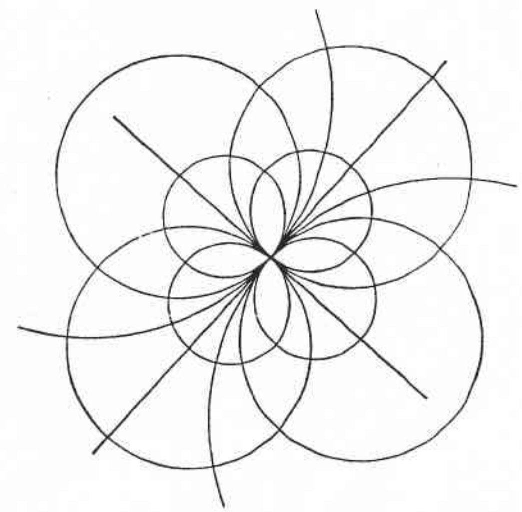

FIG. 1-6.—Circunferencias de Steiner degeneradas.

multiplicativo arbitrario, y podemos suponer, por tanto, que c es real. Entonces cada C_1 se aplica sobre sí misma, y la transformación parabólica puede considerarse como una circulación a lo largo de las circunferencias C_2.

EJERCICIOS

1. Hállense los puntos fijos de las transformaciones lineales

$$w=\frac{z}{2z-1},\qquad w=\frac{2z}{3z-1},\qquad w=\frac{3z-4}{z-1},\qquad w=\frac{z}{2-z}$$

¿Cuál de estas transformaciones es elíptica, hiperbólica o parabólica?

2. Supongamos que los coeficientes de la transformación

$$w=\frac{az+b}{cz+d}$$

están normalizados mediante la condición $ad-bc=1$. ¿Cuál es la condición para que la transformación sea elíptica, hiperbólica o parabólica?

3. Pruébese que toda transformación lineal involutiva es elíptica.

4. Hállense todas las transformaciones lineales que representan las rotaciones de la esfera de Riemann.

5. Hállense todas las circunferencias ortogonales a $|z|=1$ y a $|z-1|=4$.

CAPITULO II

FUNCIONES COMPLEJAS

2-1. Funciones elementales.—En la teoría de variables complejas se consideran funciones de cuatro tipos distintos: funciones reales de una variable real, funciones reales de una variable compleja, funciones complejas de una variable real y funciones complejas de una variable compleja.

Adoptaremos el convenio de que las letras z y w denoten siempre variables complejas; así, para indicar una función compleja de una variable compleja escribiremos $w=f(z)$. Se utilizará la notación $y=f(x)$ de forma neutra, es decir, las letras x e y simbolizarán indistintamente variables reales o complejas. Cuando queramos indicar que una variable toma únicamente valores reales, la denotaremos en general mediante una t. Con todo esto no pretendemos revocar nuestro acuerdo anterior de que x e y representan siempre variables reales en la notación $z=x+iy$.

Es esencial que la ley mediante la que se define una función se formule de manera clara y carente de ambigüedad. En otras palabras: todas las funciones deben estar *bien definidas*, y en tanto no se diga nada en contrario, serán *uniformes*.

No es necesario que una función esté definida para todos los valores de la variable independiente. En esta primera sección no consideraremos la posibilidad de que una función pueda dejar de estar definida para determinados valores. Por tanto, desde un punto de vista formal, estudiaremos solo funciones definidas en toda la recta real o en todo el plano complejo. Por otra parte, podrán omitirse tales restricciones sin originar cambio sustancial.

1. *Límites y continuidad.*—Se adoptarán las siguientes definiciones básicas:

DEFINICIÓN 1. *Se dice que la función $f(x)$ tiene el límite A cuando x tiende hacia a,*

$$\lim_{x\to a} f(x)=A, \qquad [2\text{-}1]$$

si (y solo si) se verifica:

Para todo $\epsilon > 0$ *existe un número* $\delta > 0$ *tal que* $|f(x) - A| < \epsilon$ *para todo* x *que verifique* $|x - a| < \delta$ *y* $x \neq a$.

En esta definición se hace uso del concepto de valor absoluto de manera decisiva. Puesto que este concepto tiene sentido tanto para los números complejos como para los reales, podemos utilizar esta definición tanto si la variable x y la función $f(x)$ son reales como si son complejas.

Otra notación más sencilla que a veces usaremos es $f(x) \to A$ para $x \to a$.

Existen algunas variantes conocidas de la definición, que corresponden a los casos en que a o A son infinitos. En el caso real se puede distinguir entre los límites $+\infty$ y $-\infty$, pero en el complejo hay un único límite infinito. Dejamos al lector la formulación correcta de definiciones que cubran todos los casos.

Los bien conocidos resultados referentes al límite de una suma, un producto o un cociente siguen siendo válidos en el caso complejo. En realidad, las demostraciones están basadas únicamente en las propiedades del valor absoluto, expresadas por

$$|ab| = |a| \cdot |b| \quad \text{y} \quad |a+b| \leqq |a| + |b|.$$

La condición [2-1] equivale evidentemente a

$$\lim_{x \to a} \overline{f(x)} = \overline{A}. \qquad [2\text{-}2]$$

A partir de [2-1] y [2-2] se obtiene

$$\lim_{x \to a} \operatorname{Re} f(x) = \operatorname{Re} A$$
$$\lim_{x \to a} \operatorname{Im} f(x) = \operatorname{Im} A. \qquad [2\text{-}3]$$

Recíprocamente, [2-1] es consecuencia de [2-3].

Se dice que la función $f(x)$ es *continua* en a si (y solo si) $\lim_{x \to a} f(x) = f(a)$. Una *función continua* es la que lo es en todos los puntos.

La suma $f(x) + g(x)$ y el producto $f(x)g(x)$ de dos funciones continuas es una función continua; el cociente $f(x)/g(x)$ está definido y es continuo en a si (y solo si) $g(a) \neq 0$. Si $f(x)$ es continua, también lo son $\operatorname{Re} f(x)$, $\operatorname{Im} f(x)$ y $|f(x)|$.

La derivada de una función se define como un límite particular, pudiendo considerarse indistintamente para variables reales o complejas. La definición formal es

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \qquad [2\text{-}4]$$

Son válidas las reglas usuales para formar la derivada de una suma, un producto o un cociente. La derivada de una función compuesta se determina por la regla de derivación en cadena.

Sin embargo, existe una diferencia fundamental entre los casos de una variable independiente real y una variable independiente compleja. Como ilustración de esto, sea $f(z)$ una función *real* de una variable *compleja* cuya derivada para $z = a$ existe. Entonces $f'(a)$ es, por una parte, real, puesto que es el límite del cociente

$$\frac{f(a+h) - f(a)}{h}$$

cuando h tiende a cero tomando valores reales. Por otra parte es también el límite del cociente

$$\frac{f(a+ih) - f(a)}{ih}$$

y como tal es imaginaria pura. Por tanto, $f'(a)$ tiene que ser cero. Así, pues, una función real de una variable compleja o tiene derivada cero o dicha derivada no existe. Si esta existe en todo punto, la función se reduce a una constante.

El caso de una función compleja de una variable real puede reducirse al caso real. Si se escribe $z(t) = x(t) + iy(t)$ se obtiene

$$z'(t) = x'(t) + iy'(t),$$

y la existencia de $z'(t)$ es equivalente a la existencia simultánea de $x'(t)$ e $y'(t)$. La notación compleja posee, no obstante, ciertas ventajas formales que no conviene despreciar.

En contraste con todo lo anterior, la existencia de la derivada de una función compleja de una variable compleja tiene consecuencias muy importantes en lo que se refiere a las propiedades estructurales de la función. La investigación de estas consecuencias es el tema central de la teoría de funciones de variable compleja.

2. *Funciones analíticas.*—La clase de las *funciones analíticas* está constituida por las funciones complejas de una variable compleja que poseen derivada en cada punto. Por el momento nos limitaremos al estudio de funciones $w=f(z)$ definidas en todo el plano complejo. Como consecuencia de esta limitación, resulta prematuro formular una definición precisa de función analítica, y la presente utilización del término será sometida a revisión posterior.

La suma y el producto de dos funciones analíticas es también una función analítica. Lo mismo es cierto para el cociente $f(z)/g(z)$ de dos funciones analíticas, con tal que $g(z)$ no se anule. En el caso general, es necesario excluir los puntos en los que $g(z)=0$. Este caso típico no estará, por tanto, incluido en nuestras consideraciones; pero está claro que los resultados siguen siendo válidos, salvo por modificaciones obvias.

La definición de derivada puede escribirse también en la forma

$$f'(z)=\lim_{h\to 0}\frac{f(z+h)-f(z)}{h}$$

Como una primera consecuencia, $f(z)$ es necesariamente continua. En efecto, de $f(z+h)-f(z)=h\cdot[f(z+h)-f(z)]/h$ se obtiene

$$\lim_{h\to 0}[f(z+h)-f(z)]=0\cdot f'(z)=0.$$

Si se escribe $f(z)=u(z)+iv(z)$, se deduce además que $u(z)$ y $v(z)$ son ambas continuas.

El límite del cociente incremental debe ser el mismo cualquiera que sea el camino a través del cual h tiende a cero. Si elegimos para h valores reales, entonces la parte imaginaria y se mantiene constante y la derivada se convierte en una derivada parcial con respecto a x. Tenemos así

$$f'(z)=\frac{\partial f}{\partial x}=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}$$

Análogamente, si ponemos valores imaginarios puros ik en lugar de h, obtenemos

$$f'(z)=\lim_{k\to 0}\frac{f(z+ik)-f(z)}{ik}=-i\frac{\partial f}{\partial y}=-i\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y}$$

Se sigue que $f(z)$ ha de satisfacer la ecuación en derivadas parciales

$$\frac{\partial f}{\partial x}=-i\frac{\partial f}{\partial y} \qquad [2\text{-}5]$$

que se desdobla en las ecuaciones reales

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x} \qquad [2\text{-}6]$$

Estas son las ecuaciones diferenciales de *Cauchy-Riemann*, que tienen que ser verificadas por las partes real e imaginaria de toda función analítica[1].

Obsérvese que la existencia de las cuatro derivadas parciales de [2-6] está garantizada por la existencia de $f'(z)$. Utilizando [2-6] podemos escribir cuatro expresiones formalmente diferentes para $f'(z)$; la más sencilla es

$$f'(z)=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}$$

Para $|f'(z)|^2$ tenemos, p. ej.,

$$|f'(z)|^2=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial v}{\partial x}\right)^2=\frac{\partial u}{\partial x}\frac{\partial v}{\partial y}-\frac{\partial u}{\partial y}\frac{\partial v}{\partial x}$$

La última expresión muestra que $|f'(z)|^2$ es el jacobiano de u y v con respecto a x e y.

Demostraremos más adelante que la derivada de una función analítica es también analítica. Como consecuencia de esto, u y v tendrán derivadas parciales continuas de todos los órdenes y, en particular, las derivadas cruzadas serán iguales. De lo anterior y de [2-6] se obtiene

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$$

$$\Delta v=\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}=0.$$

[1] Agustín Cauchy (1789-1857) y Bernhard Riemann (1826-1866) son considerados como los fundadores de la teoría de funciones de variable compleja. El trabajo de Riemann hizo resaltar el aspecto geométrico en contraste con el estudio puramente analítico de Cauchy.

Una función u que satisface la *ecuación de Laplace* $\Delta u=0$ se llama *armónica*. Por tanto, las partes real e imaginaria de una función analítica son armónicas. Si dos funciones armónicas u y v satisfacen las ecuaciones de Cauchy-Riemann [2-6], entonces se dice que v es la *función conjugada armónica* de u. En las mismas circunstancias, u es evidentemente la función conjugada armónica de $-v$.

No es cuestión aquí de estudiar las condiciones menos restrictivas que pueden imponerse, en lo que a regularidad se refiere, a las funciones armónicas. Deseamos demostrar, sin embargo, que la función $u+iv$, determinada por un par de funciones armónicas conjugadas, es siempre analítica, y con este objeto suponemos explícitamente que u y v tienen derivadas parciales de primer orden continuas. Se demuestra en análisis, exactamente bajo las mismas condiciones de regularidad, que las siguientes igualdades son ciertas:

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

$$v(x+h,\, y+k)-v(x,y)=\frac{\partial v}{\partial x}h+\frac{\partial v}{\partial y}k+\epsilon_2,$$

donde los restos ϵ_1, ϵ_2 tienden a cero más rápidamente que $h+ik$, en el sentido de que $\epsilon_1/(h+ik)\to 0$ y $\epsilon_2/(h+ik)\to 0$ para $h+ik\to 0$. Con la notación $f(z)=u(x,y)+iv(x,y)$ obtenemos, en virtud de las relaciones [2-6],

$$f(z+h+ik)-f(z)=\left(\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}\right)(h+ik)+\epsilon_1+i\epsilon_2$$

y, por consiguiente,

$$\lim_{h+ik\to 0}\frac{f(z+h+ik)-f(z)}{h+ik}=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}$$

Concluimos que $f(z)$ es analítica.

Las partes real e imaginaria de una función analítica son funciones armónicas que satisfacen las ecuaciones diferenciales de Cauchy-Riemann.

Recíprocamente, si las funciones armónicas u y v satisfacen dichas ecuaciones, entonces $u+iv$ es una función analítica.

La conjugada de una función armónica puede hallarse por integración, y en los casos más sencillos el cálculo puede efectuarse explícitamente. Así, p. ej., $u=x^2-y^2$ es armónica y $\partial u/\partial x=2x$, $\partial u/\partial y=-2y$. Por tanto, la función conjugada debe satisfacer las condiciones

$$\frac{\partial v}{\partial x}=2y,\qquad \frac{\partial v}{\partial y}=2x.$$

De la primera ecuación $v=2xy+\varphi(y)$, donde $\varphi(y)$ es una función de y únicamente. Por sustitución en la segunda ecuación, se obtiene $\varphi'(y)=0$. Por tanto, $\varphi(y)$ es una constante, y la función conjugada de x^2-y^2 más general es $2xy+c$, donde c es una constante. Obsérvese que $x^2-y^2+2ixy=z^2$. La función analítica cuya parte real es x^2-y^2 resulta, por consiguiente, z^2+ic.

Existe un procedimiento formal muy interesante que pone de manifiesto con bastante claridad la naturaleza de las funciones analíticas. Presentamos este procedimiento advirtiendo al lector de manera explícita que su naturaleza es puramente formal y no constituye una demostración.

Consideremos una función compleja de dos variables reales $f(x,y)$. Introduciendo la variable compleja $z=x+iy$ y su conjugada $\bar z=x-iy$, podemos escribir $x=\frac{1}{2}(z+\bar z)$, $y=-\frac{1}{2}i(z-\bar z)$. Con este cambio de variable podemos considerar a $f(x,y)$ como una función de z y $\bar z$, a las que trataremos como variables independientes (olvidando que son realmente conjugada una de otra). Si fueran aplicables las reglas del cálculo, obtendríamos las siguientes expresiones:

$$\frac{\partial f}{\partial z}=\frac{1}{2}\left(\frac{\partial f}{\partial x}-i\frac{\partial f}{\partial y}\right),\qquad \frac{\partial f}{\partial \bar z}=\frac{1}{2}\left(\frac{\partial f}{\partial x}+i\frac{\partial f}{\partial y}\right)$$

Estas expresiones no tienen definiciones convenientes en términos de límites, pero podemos, sin embargo, introducirlas como derivadas simbólicas con respecto a z y a $\bar z$. Comparándolas con [2-5], se ve que las funciones analíticas están caracterizadas por la condición $\partial f/\partial \bar z=0$. Así, caemos en la tentación de decir que una función analítica es independiente de $\bar z$, y función de z únicamente.

Este razonamiento formal apoya el punto de vista de que las funciones analíticas son verdaderas funciones de una variable compleja en oposición a las funciones que se pueden describir más adecuadamente como funciones complejas de dos variables reales.

Mediante un razonamiento formal análogo, podemos deducir un método muy sencillo que nos permite calcular, sin utilizar integrales, la función analítica $f(z)$, cuya parte real es una función armónica dada $u(x, y)$. Observemos primero que la función conjugada $\overline{f(z)}$ tiene derivada nula con respecto a z, pudiendo, por tanto, considerarse como función de $\bar{z}$; denotamos esta función por $\bar{f}(\bar{z})$. Con esta notación podemos escribir la identidad

$$u(x, y) = \tfrac{1}{2}[f(x+iy) + \bar{f}(x-iy)].$$

Es razonable esperar que la misma sea una identidad formal, y entonces se verifique incluso cuando x e y son complejos. Si sustituimos $x = z/2$, $y = z/2i$, obtenemos

$$u(z/2,\ z/2i) = \tfrac{1}{2}[f(z) + \bar{f}(0)].$$

Puesto que $f(z)$ está determinada únicamente salvo una constante imaginaria pura, cabe también suponer que $f(0)$ es real, lo que implica que $\bar{f}(0) = u(0, 0)$. La función $f(z)$ puede así calcularse mediante la fórmula

$$f(z) = 2u(z/2,\ z/2i) - u(0, 0).$$

Se puede añadir una constante aditiva imaginaria pura cualquiera.

En esta forma el método está definitivamente limitado a funciones $u(x, y)$, que son racionales en x e y, pues la función debe tener significado para valores complejos del argumento. Añadamos que el método puede extenderse al caso general y que es posible dar una justificación completa.

EJERCICIOS

1. Si $g(w)$ y $f(z)$ son funciones analíticas, demuéstrese que $g[f(z)]$ es también analítica.

2. Compruébese que las funciones z^2 y z^3 verifican las ecuaciones de Cauchy-Riemann.

3. Hállese el polinomio armónico de la forma $ax^3+bx^2y+cxy^2+dy^3$ más general. Determínese la función armónica conjugada y la correspondiente función analítica mediante integración, y también utilizando el método formal.

4. Pruébese que una función analítica no puede tener valor absoluto constante sin reducirse a una constante.

5. Demuéstrese con todo rigor que las funciones $f(z)$ y $\overline{f(\bar{z})}$ son analíticas simultáneamente.

6. Demuéstrese que las funciones $u(z)$ y $u(\bar{z})$ son armónicas simultáneamente.

7. Pruébese que una función armónica satisface la siguiente ecuación diferencial formal:

$$\frac{\partial^2 u}{\partial z\, \partial \bar{z}} = 0.$$

3. *Funciones racionales.*—Toda constante es una función analítica con derivada cero. La función analítica más sencilla no constante es z, y su derivada vale uno. Puesto que la suma y el producto de dos funciones analíticas es también analítica, se sigue que todo polinomio

$$P(z) = a_0 + a_1 z + \ldots + a_n z^n \qquad [2\text{-}7]$$

es una función analítica. Su derivada es

$$P'(z) = a_1 + 2a_2 z + \ldots + n a_n z^{n-1}.$$

La notación [2-7] implicará que $a_n \neq 0$, y se dice entonces que el polinomio es de grado n. La constante 0, considerada como un polinomio, es en muchos aspectos excepcional y será excluida de nuestra atención en lo que sigue[1].

Para $n > 0$, la ecuación $P(z) = 0$ tiene al menos una raíz. Este es el llamado teorema fundamental del álgebra, el cual será demostrado más adelante. Si $P(\alpha_1) = 0$, se demuestra en álgebra elemental que $P(z) = (z - \alpha_1)P_1(z)$, siendo $P_1(z)$ un polinomio de grado $n-1$. Repitiendo sucesivamente este proceso, se llega, por último, a la siguiente factorización completa:

$$P(z) = a_n(z-\alpha_1)(z-\alpha_2)\ \ldots\ (z-\alpha_n),$$

donde las $\alpha_1, \alpha_2, \ldots, \alpha_n$ no son necesariamente distintas. De esta factorización se deduce que $P(z)$ no se anula para cualquier valor de z distinto de $\alpha_1, \alpha_2, \ldots, \alpha_n$. Además, la factorización está determinada de manera única salvo el orden de los factores.

Si coinciden exactamente h de las α_j, a este valor común se le llama un *cero* de $P(z)$ de *orden* h. Evidentemente, la suma de los órdenes de los ceros de un polinomio es igual a su grado. Más sencillamente, si se cuenta cada cero tantas veces como indica su orden, un polinomio de grado n tiene exactamente n ceros.

El orden de un cero α puede determinarse también consideran-

[1] Por razones de tipo formal, si se considera a la constante 0 como un polinomio, se le atribuye grado $-\infty$.

do las derivadas sucesivas de $P(z)$ para $z=\alpha$. Supongamos que α es un cero de orden h. Entonces podemos escribir $P(z)=(z-\alpha)^h P_h(z)$ con $P_h(\alpha)\neq 0$. Mediante derivaciones sucesivas se obtiene $P(\alpha)=$ $=P'(\alpha)=\ldots=P^{(h-1)}(\alpha)=0$, mientras que $P^{(h)}(\alpha)\neq 0$. En otras palabras: el orden de un cero es igual al orden de la primera derivada que no se anula. Un cero de orden 1 se llama un cero simple y está caracterizado por las condiciones $P(\alpha)=0$, $P'(\alpha)\neq 0$.

Volvamos al caso de una función racional

$$R(z)=\frac{P(z)}{Q(z)} \qquad [2\text{-}8]$$

dada como cociente de dos polinomios. Suponemos, y esto es esencial, que $P(z)$ y $Q(z)$ no tienen factores comunes y, por tanto, carecen de ceros comunes. Se le dará a $R(z)$ el valor ∞ en los ceros de $Q(z)$. Puede, pues, considerarse como una función con valores en el plano ampliado, y como tal es continua. Los ceros de $Q(z)$ se llaman *polos* de $R(z)$, siendo el orden de un polo, por definición, igual al orden del cero correspondiente de $Q(z)$.

La derivada

$$R'(z)=\frac{P'(z)Q(z)-Q'(z)P(z)}{Q(z)^2} \qquad [2\text{-}9]$$

existe únicamente cuando $Q(z)\neq 0$. Sin embargo, como función racional definida mediante el segundo miembro de [2-9], $R'(z)$ tiene los mismos polos que $R(z)$, estando el orden de cada polo incrementado en una unidad. En el caso de que $Q(z)$ tenga ceros múltiples, obsérvese que la expresión [2-9] no aparece en su forma reducida.

Se obtiene una mayor uniformidad si se supone que, al igual que la función $R(z)$, la variable independiente z toma valores en todo el plano completo. Podemos definir $R(\infty)$ como el límite de $R(z)$ cuando $z\to\infty$, pero esta definición no determinaría el orden de un cero o de un polo en ∞. Es preferible, por tanto, considerar la función $R(1/z)$, que puede escribirse nuevamente como una función racional $R_1(z)$, y poner

$$R(\infty)=R_1(0).$$

Si $R_1(0)=0$ o ∞, el orden del cero o del polo en ∞ se define como el orden del cero o del polo de $R_1(z)$ en el origen.

Con la notación

$$R(z)=\frac{a_0+a_1z+\ldots+a_nz^n}{b_0+b_1z+\ldots+b_mz^m}$$

obtenemos

$$R_1(z)=z^{m-n}\,\frac{a_0z^n+a_1z^{n-1}+\ldots+a_n}{b_0z^m+b_1z^{m-1}+\ldots+b_m}$$

donde la potencia z^{m-n} pertenece al numerador o al denominador. Por consiguiente, si $m>n$, $R(z)$ tiene un cero de orden $m-n$ en ∞; si $m<n$, el punto ∞ es un polo de orden $n-m$, y si $m=n$,

$$R(\infty)=a_n/b_m\neq 0,\ \infty.$$

Podemos contar ahora el número total de ceros y polos en el plano completo. El resultado muestra que el número de ceros, incluyendo los ceros en ∞, es igual al mayor de los números m y n. El número de polos es el mismo. Este número común de ceros y polos se llama *orden* de la función racional.

Si a es una constante cualquiera, la función $R(z)-a$ tiene los mismos polos que $R(z)$ y, por tanto, el mismo orden. Los ceros de $R(z)-a$ son las raíces de la ecuación $R(z)=a$, y si las raíces se cuentan tantas veces como indica el orden del cero, podemos enunciar el siguiente resultado:

Una función racional $R(z)$ de orden p tiene p ceros y p polos, y toda ecuación $R(z)=a$ tiene exactamente p raíces.

Una fracción lineal $(\alpha z+\beta)/(\gamma z+\delta)$ con $\alpha\delta-\beta\gamma\neq 0$ es una función racional de orden uno. El hecho de que una transformación lineal posea una inversa es un caso especial del resultado anterior.

Toda función racional admite una representación mediante *fracciones simples*. Con objeto de deducir esta representación supondremos en primer lugar que $R(z)$ tiene un polo en ∞. Efectuamos la división de $P(z)$ por $Q(z)$ hasta que el grado del resto sea a lo sumo igual al del denominador. El resultado puede entonces escribirse en la forma

$$R(z)=G(z)+H(z), \qquad [2\text{-}10]$$

donde $G(z)$ es un polinomio sin término constante, y $H(z)$, finito en ∞. El grado de $G(z)$ es el orden del polo en ∞, denominándose el polinomio $G(z)$ *parte singular* de $R(z)$ en ∞.

Denotemos los polos finitos distintos de $R(z)$ por $\beta_1, \beta_2, \ldots, \beta_q$. La función $R\left(\beta_j+\frac{1}{\zeta}\right)$ es una función racional de ζ con un polo en $\zeta=\infty$. Utilizando la descomposición [2-10], podemos escribir

$$R\left(\beta_j+\frac{1}{\zeta}\right)=G_j(\zeta)+H_j(\zeta),$$

o mediante un cambio de variable,

$$R(z)=G_j\left(\frac{1}{z-\beta_j}\right)+H_j\left(\frac{1}{z-\beta_j}\right)$$

Aquí $G_j\left(\frac{1}{z-\beta_j}\right)$ es un polinomio en $\frac{1}{z-\beta_j}$ sin término constante, que se llama parte singular de $R(z)$ en β_j. La función $H_j\left(\frac{1}{z-\beta_j}\right)$ es finita para $z=\beta_j$.

Consideremos ahora la expresión

$$R(z)-G(z)-\sum_{j=1}^{q} G_j\left(\frac{1}{z-\beta_j}\right) \qquad [2\text{-}11]$$

Esta es una función racional que no puede tener otros polos que $\beta_1, \beta_2, \ldots, \beta_q$ e ∞. En $z=\beta_j$ hallamos que los dos términos que se hacen infinitos tienen una diferencia $H_j\left(\frac{1}{z-\beta_j}\right)$ con un límite finito, siendo cierto lo mismo en ∞. Por tanto, [2-11] no tiene ni polos finitos ni polo en ∞. Una función racional sin polos ha de reducirse a una constante, y si se engloba esa constante en $G(z)$ obtenemos la expresión

$$R(z)=G(z)+\sum_{j=1}^{q} G_j\left(\frac{1}{z-\beta_j}\right) \qquad [2\text{-}12]$$

Esta representación nos es bien conocida del análisis, donde se utiliza como artificio técnico en la teoría de la integración. Sin embargo, su completo aprovechamiento no se consigue hasta la introducción de los números complejos.

Haremos una pequeña aplicación de [2-12] para la determinación de todas las funciones racionales de orden 2. Con más precisión,

determinaremos las formas más sencillas que se pueden obtener mediante transformaciones lineales de las variables dependiente e independiente.

Una función racional de segundo orden tiene o bien un polo doble o dos simples. En el primer caso podemos llevar el polo a ∞ mediante una transformación lineal previa de z. La función tendrá entonces la forma

$$w=az^2+bz+c=a\left(z+\frac{b}{2a}\right)^2+c-\frac{b^2}{4a}$$

y mediante una nueva transformación lineal puede reducirse a la forma normal $w=z^2$. En el caso de dos polos simples podemos hacer que estos polos estén en $z=0$ y $z=\infty$. La representación [2-12] tomará entonces la forma

$$w=Az+B+\frac{C}{z} \qquad [2\text{-}13]$$

Si reemplazamos z por $z'=z\sqrt{A/C}$, los coeficientes de z' y $1/z'$ son iguales, y un nuevo cambio lineal de w reducirá [2-13] a la forma

$$w=\frac{1}{2}\left(z+\frac{1}{z}\right)$$

que elegimos como forma normal de una función racional de segundo orden con polos distintos.

EJERCICIOS

1. Utilícese el método del texto para desarrollar

$$\frac{z^4}{z^3-1} \quad \text{y} \quad \frac{1}{z(z+1)^2(z+2)^3}$$

en fracciones simples.

2. ¿Cuál es la relación, si es que la hay, entre el orden de una función racional y el orden de su derivada?

3. Determínense todas las funciones racionales de tercer orden y póngase en su forma más sencilla mediante transformaciones lineales.

4. ¿Cuál es la forma general de una función racional que transforma la circunferencia unidad $|z|=1$ en sí misma?

4. *La función exponencial.*—La función

$$\int_1^y \frac{dt}{t}$$

está definida, es continua y creciente para $y > 0$. Tiende a $-\infty$ para $y \to 0$ y a $+\infty$ para $y \to +\infty$. Por consiguiente, la ecuación

$$x = \int_1^y \frac{dt}{t} \qquad [2\text{-}14]$$

determina de manera única un valor positivo de y para todo x real; denotamos esta solución por e^x. Evidentemente, $e^0 = 1$, y la base $e = e^1$ está definida por

$$1 = \int_1^e \frac{dt}{t}$$

Se sigue de [2-14] que $dx/dy = 1/y$; por consiguiente, la derivada de e^x es e^x.

La propiedad fundamental de la función exponencial de una variable real es el *teorema de adición:*

$$e^{x_1+x_2} = e^{x_1} \cdot e^{x_2}. \qquad [2\text{-}15]$$

Se demuestra observando que la derivada de $e^{-x} \cdot e^{x+y}$ con respecto a x es cero. Como tal función de x, esta expresión es, por tanto, constantemente igual a su valor para $x = 0$, y se obtiene $e^{-x} \cdot e^{x+y} = e^y$. Eligiendo $x = -x_1$, $y = x_1 + x_2$, obtendremos [2-15].

La función inversa de e^x es el logaritmo natural, $\log x$. Tenemos, pues, por definición

$$\log y = \int_1^y \frac{dt}{t}$$

para $y > 0$, siendo [2-15] equivalente a la expresión

$$\log (y_1 y_2) = \log y_1 + \log y_2. \qquad [2\text{-}16]$$

Deseamos prolongar la definición de la función exponencial para valores complejos de la variable independiente. Es natural exigir que e^z se reduzca a la función exponencial real cuando z sea real. En se-

gundo lugar exigiremos que el teorema de adición de exponentes [2-15] se verifique para exponentes complejos cualesquiera. Para $z = x + iy$ deberemos tener, por tanto, $e^z = e^x \cdot e^{iy}$, y queda solo por definir e^{iy}.

En el capítulo I se demostró que la función $\cos y + i \operatorname{sen} y$, denotada por $e(y)$, satisface el teorema de adición. Nuestros requisitos quedarán, pues, satisfechos si escribimos

$$e^{iy} = \cos y + i \operatorname{sen} y \qquad [2\text{-}17]$$

y

$$e^z = e^x(\cos y + i \operatorname{sen} y). \qquad [2\text{-}18]$$

Se elige esta fórmula como una definición de la función exponencial compleja. Equivale a las relaciones

$$|e^z| = e^x, \qquad \arg e^z = y. \qquad [2\text{-}19]$$

Los motivos que han conducido a esta definición no son de carácter obligatorio. De hecho, la función $e(ky)$ satisface el teorema de adición para todo valor real de k, y podríamos haber decidido poner $e^z = e^x(\cos ky + i \operatorname{sen} ky)$. ¿Por qué se da preferencia al valor $k = 1$? La respuesta es que esta es la única elección que hace que e^z sea una función analítica. Escribiendo $u = e^x \cos ky$, $v = e^x \operatorname{sen} ky$ se obtiene

$$\frac{\partial u}{\partial x} = e^x \cos ky, \qquad \frac{\partial u}{\partial y} = -ke^x \operatorname{sen} ky$$

$$\frac{\partial v}{\partial x} = e^x \operatorname{sen} ky, \qquad \frac{\partial v}{\partial y} = ke^x \cos ky.$$

Por consiguiente, las ecuaciones de Cauchy-Riemann $\partial u/\partial x = \partial v/\partial y$, $\partial u/\partial y = -\partial v/\partial x$ se verifican si (y solo si) $k = 1$. Con esta elección obtenemos que e^z tiene por derivada e^z; así, la función exponencial compleja posee la misma propiedad de reproducirse por derivación que tenía la función exponencial real.

La definición [2-18] conduce a un buen número de consecuencias notables. En primer lugar observemos las curiosas relaciones

$$e^{\pi i} = -1, \qquad e^{2\pi i} = 1, \qquad [2\text{-}20]$$

las cuales no deberían considerarse como genuinas relaciones numéricas. Está claro que las bases e y π han sido introducidas con objeto de normalizar las funciones exponencial y trigonométricas, y [2-20] expresa únicamente la relación entre estas bases a la que una correcta definición de la función exponencial compleja debe conducir.

De la segunda ecuación [2-20] se sigue que la función exponencial tiene *período* $2\pi i$. Tenemos en efecto que $e^{z+2\pi i}=e^z$. No existen otros períodos aparte de los múltiplos de $2\pi i$.

La función exponencial no se anula nunca. Este hecho es consecuencia directa de la primera ecuación [2-19] y de que la función exponencial real es por definición positiva.

La ecuación [2-17] lleva el nombre de *Euler*. De ella deducimos las siguientes relaciones:

$$\cos y=\frac{e^{iy}+e^{-iy}}{2}$$
$$\operatorname{sen} y=\frac{e^{iy}-e^{-iy}}{2i} \qquad [2\text{-}21]$$

Desde un punto de vista práctico, [2-17] nos da una notación sencilla para un número complejo dado mediante sus coordenadas polares r, θ. Podemos escribir

$$z=r(\cos\theta+i\operatorname{sen}\theta)=re^{i\theta},$$

y esta notación es tan útil que se maneja continuamente, aun cuando en la discusión en cuestión no aparezca la función exponencial.

La función e^z no puede definirse en el plano completo, ya que no tiene límite cuando z tiende a ∞.

Junto con la función exponencial debemos también estudiar su función inversa, *el logaritmo*. Por definición, $z=\log w$ es una raíz de la ecuación $e^z=w$. De acuerdo con [2-19], esto equivale a

$$e^x=|w|$$
$$y=\arg w.$$

Para $w\neq 0$, la primera ecuación tiene una solución única $x=\log|w|$, definida por la integral

$$x=\int_0^{|w|}\frac{dt}{t}$$

Por consiguiente, el valor del logaritmo complejo es

$$\log w=\log|w|+i\arg w. \qquad [2\text{-}22]$$

Debe observarse que todos los valores del argumento dan valores posibles del logaritmo. Así, pues, el logaritmo toma infinitos valores, que difieren en múltiplos de $2\pi i$. Puesto que el logaritmo no es una función uniforme, se debe prestar mucha atención cada vez que se utiliza un logaritmo.

Si ρ es un número positivo, ha de tenerse buen cuidado para distinguir entre el *logaritmo real* con un único valor $\log\rho$ y el logaritmo complejo con infinitos valores $\log\rho+n\cdot 2\pi i$. Se conviene en que $\log\rho$ denote el logaritmo real, a no ser que se diga lo contrario o esté claramente implícito.

La fórmula [2-22] no tiene por qué ser retenida en la memoria, ya que se deduce directamente de la notación $w=\rho e^{i\varphi}$. Hacemos resaltar que $w=0$ no tiene logaritmo, como consecuencia de que e^z nunca es cero.

El símbolo a^b, donde a y b son números complejos arbitrarios, excepto por la condición de que $a\neq 0$, se utiliza siempre como una notación equivalente y más sencilla de $e^{b\log a}$. Si a se restringe a valores positivos, $\log a$ será el logaritmo real, y en este caso a^b toma un solo valor. Si a no está restringido de la forma anterior, $\log a$ es el logaritmo complejo y a^b tiene en general infinitos valores. Existirá un valor único si (y solo si) b es un entero n, y entonces a^n puede calcularse como una potencia (producto iterado) de a o de a^{-1}. Si b es un número racional con forma reducida p/q, entonces $a^{p/q}$ tiene exactamente q valores, y puede representarse mediante $\sqrt[q]{a^p}$.

EJERCICIOS

1. Hállese el valor de e^z para

$$z=-\frac{\pi i}{2},\quad \frac{3}{4}\pi i,\quad \frac{2}{3}\pi i.$$

2. ¿Para qué valores de z es e^z igual a 2, -1, i, $-i/2$, $-1-i$, $1+2i$?
3. Hállense las partes real e imaginaria de e^{e^z}.
4. Determínense todos los valores de 2^i, i^i, $(-1)^{2i}$.
5. ¿En qué sentido es cierto que $(a^b)^c=a^{bc}$?
6. Determínense las partes real e imaginaria de z^z.
7. Discútase el comportamiento de e^z cuando z tiende a ∞ a lo largo de una recta.

5. *Las funciones trigonométricas.*—Las relaciones [2-21] son válidas para todos los valores reales de y. Ampliamos ahora las definiciones del seno y del coseno para valores complejos de la variable, por medio de las siguientes relaciones:

$$\cos z = \frac{e^{iz}+e^{-iz}}{2}, \qquad \operatorname{sen} z = \frac{e^{iz}-e^{-iz}}{2i} \qquad [2\text{-}23]$$

Recíprocamente, estas definiciones dan validez universal a la fórmula

$$e^{iz} = \cos z + i \operatorname{sen} z. \qquad [2\text{-}24]$$

Las definiciones muestran que $\cos z$ y $\operatorname{sen} z$ son funciones analíticas con derivadas

$$D \cos z = -\operatorname{sen} z$$
$$D \operatorname{sen} z = \cos z.$$

Ambas funciones son periódicas con período 2π, y están ligadas por las relaciones $\cos z = \operatorname{sen}(\pi/2 - z)$ y $\cos^2 z + \operatorname{sen}^2 z = 1$. Todas las fórmulas trigonométricas siguen siendo válidas como consecuencia de la definición y del teorema de adición de la función exponencial. Observemos que $\cos z$ es una función par, y $\operatorname{sen} z$, impar.

Para z imaginario puro, $z = iy$, las fórmulas dan

$$\cos iy = \frac{e^y + e^{-y}}{2} = \cosh y$$
$$\operatorname{sen} iy = \frac{e^{-y} - e^y}{2i} = i \operatorname{senh} y,$$

donde hemos utilizado la notación típica para el seno y coseno hiperbólicos. Con ayuda de los teoremas de adición resulta ahora fácil determinar las partes real e imaginaria de $\cos z$ y de $\operatorname{sen} z$. Se obtienen las siguientes expresiones:

$$\cos(x+iy) = \cos x \cosh y - i \operatorname{sen} x \operatorname{senh} y$$
$$\operatorname{sen}(x+iy) = \operatorname{sen} x \cosh y + i \cos x \operatorname{senh} y.$$

A fin de obtener fórmulas más sencillas para los valores absolutos es preferible utilizar las fórmulas trigonométricas

$$\cos\alpha \cos\beta = \tfrac{1}{2}[\cos(\alpha+\beta) + \cos(\alpha-\beta)],$$
$$\operatorname{sen}\alpha \operatorname{sen}\beta = \tfrac{1}{2}[\cos(\alpha-\beta) - \cos(\alpha+\beta)].$$

El cálculo da

$$|\cos z|^2 = \cos(x+iy)\cos(x-iy) = \tfrac{1}{2}(\cos 2x + \cos 2iy) =$$
$$= \tfrac{1}{2}(\cosh 2y + \cos 2x)$$

y análogamente,

$$|\operatorname{sen} z|^2 = \tfrac{1}{2}(\cosh 2y - \cos 2x).$$

Es interesante observar que $|\cos z|^2$ y $|\operatorname{sen} z|^2$ difieren de $\frac{1}{2}\cosh 2y$ únicamente en términos acotados, y este último valor es aproximadamente igual a $\frac{1}{4}e^{2|y|}$.

La tangente y cotangente se introducen, naturalmente, como cocientes del seno y del coseno. De manera explícita, tenemos

$$\operatorname{tg} z = \frac{1}{i}\,\frac{e^{iz}-e^{-iz}}{e^{iz}+e^{-iz}}$$

con el convenio de que $\operatorname{tg} z = \infty$ cuando $\cos z = 0$.

La inversa del coseno se obtiene como solución de la ecuación

$$\cos z = \tfrac{1}{2}(e^{iz}+e^{-iz}) = w.$$

Esta es una ecuación de segundo grado en e^{iz}, cuyas raíces son

$$e^{iz} = w \pm \sqrt{w^2-1},$$

y, por tanto,

$$z = \operatorname{arc\,cos} w = -i \log(w \pm \sqrt{w^2-1}).$$

Estos valores se pueden también escribir en la forma

$$\operatorname{arc\,cos} w = \pm i \log(w + \sqrt{w^2-1}),$$

pues $w + \sqrt{w^2-1}$ y $w - \sqrt{w^2-1}$ son números recíprocos. Los infinitos valores de $\operatorname{arc\,cos} w$ reflejan la paridad y periodicidad de $\cos z$.

La inversa del seno se determina con toda facilidad mediante

$$\operatorname{arc\,sen} w = \frac{\pi}{2} - \operatorname{arc\,cos} w.$$

Es conveniente hacer observar que en la teoría de funciones complejas todas las funciones trascendentes elementales pueden expresarse mediante e^z y su función inversa $\log z$. En otras palabras: esencialmente no existe más que una función trascendente elemental.

EJERCICIOS

1. Hállense los valores de $\operatorname{sen} i$, $\cos i$, $\cos\left(\frac{\pi}{4}-i\right)$, $\operatorname{tg}(1+i)$.

2. Averígüense todas las raíces de las ecuaciones $\operatorname{sen} z=i$, $\cos z=2$, $\operatorname{cotg} z=1+i$.

3. Determínense las partes real e imaginaria de $\operatorname{tg}(x+iy)$.

4. Discútase el comportamiento de $\operatorname{sen} z$, $\cos z$ y $\operatorname{tg} z$ cuando z tiende a ∞.

5. Dedúzcanse los teoremas de adición para la tangente y la cotangente.

6. Pruébese que todos los períodos de $\operatorname{tg} z$ son múltiplos de π.

7. Exprésese $\operatorname{arc\,tg} w$ por medio del logaritmo.

2-2. Conceptos topológicos.—La rama de las matemáticas conocida con el nombre de *topología* trata de todas las cuestiones directa e indirectamente relacionadas con la continuidad. El término es utilizado tradicionalmente en un sentido muy amplio y sin límites estrictos. Las consideraciones topológicas son de gran importancia para el fundamento de la teoría de funciones, y el primer estudio sistemático de la topología tuvo por motivo esta necesidad.

En la presente sección estamos interesados, en primer lugar, en las propiedades topológicas de conjuntos de puntos. Debe quedar bien entendido que las propiedades más generales son las más sencillas de manejar, ya que pueden expresarse en los términos lógicos más simples. Desde el punto de vista lógico, muchas situaciones que parecen intuitivamente sencillas son bastante complicadas, por lo que deben soslayarse. A veces es incluso necesario utilizar definiciones cuyo contenido intuitivo no es inmediatamente claro. Recomendamos al lector que, en tales casos, adopte un punto de vista puramente formal y se concentre en la lógica del razonamiento.

1. *Conjuntos de puntos.*—Para determinar un *conjunto de puntos* tiene que darse una ley mediante la cual sea posible decidir si un punto pertenece al conjunto dado o no. Si se representan los puntos mediante números reales o complejos, muchos conjuntos pueden definirse convenientemente en términos de desigualdades o de ecuaciones. No sería satisfactorio, sin embargo, restringir la atención a conjuntos definidos mediante una ley de cierto tipo prefijado. Realmente, nuestro objeto es deducir proposiciones de una gran generalidad con un correspondiente campo de aplicaciones muy amplio.

Los puntos o números que pertenecen a un conjunto se llaman sus *elementos*. La notación $a \in A$ se utiliza para indicar que el punto a es un elemento del conjunto A. Dos conjuntos son idénticos si (y solo si) tienen los mismos elementos; A es un *subconjunto* de B si todo elemento de A es también un elemento de B, y esta relación se indica mediante la notación $A \subset B$ o $B \supset A$. *No* está excluida la posibilidad de que $A=B$.

Un conjunto es *finito* si tiene solo un número finito de elementos. Admitiremos de manera específica el caso de que un conjunto no tenga elementos en absoluto. Se dice entonces que es vacío, y puesto que todos los conjuntos vacíos son idénticos, podemos hablar del *conjunto vacío*, el cual denotaremos por 0.

La *intersección* de dos conjuntos A y B, denotada por AB o por $A \cap B$, está formada por todos los puntos que son elementos de A y de B. Si $AB=0$, los conjuntos se dicen disjuntos. La *unión* de dos conjuntos se denota por $A+B$ o por $A \cup B$, y se compone de todos los puntos que son elementos de A o de B, incluyendo aquellos que son elementos de ambos. Podemos también considerar la intersección o la unión de una colección arbitraria de conjuntos. Las definiciones son inmediatas.

El *complemento* de un conjunto A es el conjunto de todos los puntos que no pertenecen a A; será denotado por $C(A)$. Obsérvese que el complemento depende de la totalidad de los puntos bajo consideración. Así, p. ej., un conjunto de números reales tiene un complemento con respecto a la recta real y otro con respecto al plano complejo. Más generalmente, si $A \subset B$, podemos considerar el complemento relativo $B-A$ constituido por todos los puntos de B que no están en A.

Es útil recordar las siguientes identidades puramente lógicas:

$$C(AB)=C(A)+C(B)$$
$$C(A+B)=C(A)C(B),$$

y su generalización a colecciones arbitrarias de conjuntos.

La recta real y el plano complejo tienen una estructura especial, cuya característica más importante es la existencia de una distancia $|a-b|$. En términos de distancia podemos dar un significado preciso a los conceptos "suficientemente próximo" y "arbitrariamente próximo", que son fundamentales en todas las cuestiones referentes a límites y continuidad.

Un *entorno* de un punto a es el conjunto de los puntos x que satisfacen una desigualdad de la forma $|x-a| < \epsilon$, donde $\epsilon > 0$. En la recta real, por tanto, un entorno es un segmento sin sus extremos (un intervalo abierto); en el plano complejo es el interior de un círculo (un disco) [1].

Nunca estaremos interesados en entornos individuales. La cuestión que importa es siempre si una determinada propiedad es cierta para *todos* los entornos de un punto a. Sea A una propiedad cualquiera que un punto x puede tener o no. Entonces se dice que A se verifica para puntos *arbitrariamente próximos* a a si todo entorno de a contiene al menos un punto con la propiedad A. Puesto que un entorno "grande" contiene a todos los entornos "pequeños", bastaría exigir que todos los entornos "pequeños" contengan un punto con la propiedad A, pero es lógicamente más sencillo considerar todos los entornos.

Análogamente, se dice que A se verifica para todos los puntos *suficientemente próximos* a a si no es cierto que se verifica la propiedad opuesta no-A para puntos arbitrariamente próximos a a. Formulándolo más positivamente, se verifica A para todos los puntos suficientemente próximos a a si existe un entorno de a todos cuyos puntos tienen la propiedad A.

Todos los puntos con la propiedad A forman un conjunto que podemos denotar por la misma letra A. Recíprocamente, todo conjunto A define una propiedad, la de pertenecer a A. Por tanto, necesitamos solo averiguar si un conjunto A contiene o no puntos arbitrariamente próximos a un punto dado a. Esto nos conduce al concepto de *cierre:*

DEFINICIÓN 2. *El cierre $\bar{A}$ de un conjunto A es el conjunto formado por todos los puntos a tales que A contiene puntos arbitrariamente próximos al a.*

Evidentemente, todo conjunto está contenido en su cierre: $A \subset \bar{A}$. En efecto, todo entorno de $a \in A$ contiene a a y, por tanto, a un punto de A.

Intentamos basar todas las definiciones topológicas en la noción

[1] Se llama *disco* al conjunto de los puntos interiores a una circunferencia (a veces, para mayor claridad, se dirá *disco abierto*). Un *disco cerrado* es el conjunto formado por los puntos interiores y sobre una circunferencia.

de cierre. En interés del lector daremos un mínimo de terminología, pero los términos siguientes son indispensables:

Un conjunto es *cerrado* si es idéntico a su cierre.

Un conjunto es *abierto* si su complemento es cerrado.

El *exterior* de un conjunto es el complemento de su cierre.

El *interior* de un conjunto es el exterior de su complemento.

La *frontera* de un conjunto es la intersección de su cierre con el cierre de su complemento.

Por supuesto que no es fácil dominar la conexión lógica existente entre estas definiciones; por consiguiente, trataremos de ampliarlas. Conjuntos cerrados y abiertos son complementarios, pero en muchos respectos es más sencilla la caracterización directa de los conjuntos abiertos que la de los cerrados. Por definición, A es abierto si ningún punto $a \in A$ pertenece al cierre de $C(A)$. Esto significa que existirá un entorno de a sin parte común con $C(A)$. Luego este entorno está contenido en A, obteniéndose así la siguiente caracterización positiva de los conjuntos abiertos:

Un conjunto A es abierto si todo $a \in A$ posee un entorno contenido en A, y solo en este caso.

Por razones tipográficas, denotaremos de momento el cierre por A^-, y el complemento por A'. Entonces el exterior de A es $A^{-\prime}$, y el interior $A'^{-\prime}$. De $A \subset A^-$ y $A' \subset A'^-$ se sigue que $A^{-\prime} \subset A'$ y $A'^{-\prime} \subset A$; es decir, el interior está contenido en A y el exterior está fuera de A. La frontera es $A^-A'^-$. El complemento de la frontera es, por consiguiente, $(A^-A'^-)' = A^{-\prime} + A'^{-\prime}$. La frontera es, pues, el complementario de la unión del interior y del exterior, así que cada punto puede clasificarse de manera única como un punto interior, un punto exterior o un punto frontera. Un punto interior de A posee un entorno contenido en A, un punto exterior tiene un entorno sin parte común con A y un punto frontera se caracteriza porque todo entorno de él tiene parte común tanto con A como con su complemento A'.

El cierre de todo conjunto es cerrado. Basta probar que $A^{--} \subset A^-$, pues sabemos ya que $A^- \subset A^{--}$. Supongamos que $a \in A^{--}$. Dado cualquier $\epsilon > 0$, existe un punto $b \in A^-$ con $|b-a| < \epsilon$. Por otra parte, puesto que $b \in A^-$ existe un $c \in A$ con $|c-b| < \epsilon - |b-a|$. La desigualdad triangular implica $|c-a| \leqq |c-b| + |b-a| < \epsilon$. Hemos demostrado así que $a \in A^-$, y puesto que a era arbitrario, obtenemos $A^{--} \subset A^-$.

COROLARIO. El interior y el exterior de cualquier conjunto son conjuntos abiertos.

Enunciamos a continuación cuatro proposiciones sencillas, de carácter simétrico, que se utilizan constantemente:

La intersección de cualquier colección de conjuntos cerrados es un conjunto cerrado.

La unión de un número finito de conjuntos cerrados es un conjunto cerrado.

La unión de una colección cualquiera de conjuntos abiertos es un conjunto abierto.

La intersección de un número finito de conjuntos abiertos es un conjunto abierto.

La primera proposición es consecuencia inmediata de la definición; dejamos su demostración a cargo del lector. Demostraremos la segunda para dos conjuntos; el caso general se sigue por inducción. Supongamos que a pertenece al cierre de $A+B$, siendo A y B cerrados. Si a no pertenece a A, posee un entorno $|x-a|<\epsilon_1$ sin parte común con A, y si a no está en B, hay un entorno $|x-a|<\epsilon_2$ disjunto con B. Si a no estuviera ni en A ni en B, el menor de estos entornos no tendría parte común con $A+B$. Esto contradice nuestra hipótesis de que a pertenece al cierre de $A+B$. Por tanto, a tiene que pertenecer a $A+B$, siendo, por tanto, $A+B$ cerrado.

Las proposiciones tercera y cuarta se deducen de la primera y de la segunda mediante consideración de los complementos.

Como consecuencia de la proposición primera obtenemos que la frontera de cualquier conjunto es cerrada, pues, por definición, resulta que es intersección de dos conjuntos cerrados.

A fin de evitar falsas interpretaciones es conveniente hacer resaltar que los atributos "abierto" y "cerrado" no son contradictorios. El conjunto vacío es simultáneamente abierto y cerrado, y lo mismo le ocurre a su complemento, el conjunto de todos los puntos.

Es fácil construir ejemplos que ilustren la naturaleza de los conjuntos abiertos y de los cerrados, lo mismo que los conceptos de interior, exterior y frontera; algunos de estos ejemplos aparecen en los ejercicios que siguen. En el texto demostraremos únicamente que un intervalo abierto o un disco abierto $|x-a|<\rho$ son conjuntos abiertos, mientras que un intervalo cerrado o un disco cerrado $|x-a|\leqq\rho$ son conjuntos cerrados. Estas propiedades derivan directamente de la propiedad triangular. El conjunto $|x-a|<\rho$ es abierto, pues si $|b-a|<\rho$, todos los puntos x pertenecientes al entorno

$|x-b|<\rho-|b-a|$ satisfacen a $|x-a|\leqq|x-b|+|b-a|<\rho$. El mismo razonamiento muestra que el conjunto $|x-a|>\rho$ es abierto, siendo cerrado, por consiguiente, su complemento $|x-a|\leqq\rho$. La frontera de ambos conjuntos es el conjunto $|x-a|=\rho$, que representa los extremos del intervalo o la circunferencia del disco, respectivamente.

Creemos conveniente introducir también los siguientes conceptos de *punto aislado* y *punto de acumulación* de un conjunto. Se dice que el punto $a\in A$ es un punto aislado de A si existe un entorno de a cuya intersección con A se reduce a dicho punto a. Un punto de acumulación es un punto del cierre que no es aislado. Evidentemente, todo entorno de un punto de acumulación de A contiene infinitos puntos de A, y esta propiedad caracteriza a los puntos de acumulación.

Conviene distinguir entre una sucesión de infinitos puntos x_1, x_2, ..., x_n, ... y el conjunto de todos los puntos x_n. Esto es particularmente importante en el caso de que infinitos x_n coincidan. Llamaremos a y *punto límite* de la sucesión si todo entorno de y contiene términos x_n para infinitos valores de n. Una sucesión convergente tiene exactamente un punto límite, su *límite*.

Con tal que definamos un entorno de $z=\infty$ mediante una desigualdad de la forma $|z|>\rho$, todas las consideraciones anteriores pueden aplicarse al plano ampliado.

EJERCICIOS

1. Demuéstrese, aplicando solo las definiciones, que el semiplano $x>0$ del plano complejo z es un conjunto abierto, en tanto que la desigualdad $x\geqq0$ determina un conjunto cerrado.

2. Determínese el interior, el exterior, la frontera y el cierre del conjunto formado por todos los números racionales.

3. Pruébese que los números reales cuya expresión decimal puede escribirse sin utilizar el dígito 7 forman un conjunto cerrado sin puntos aislados. (Obsérvese que $0{,}7=0{,}6999\ldots$ pertenece a este conjunto.)

4. Demuéstrese que $A^{-\prime-\prime-\prime-\prime}=A^{-\prime-\prime}$.

5. Pruébese que el cierre de un conjunto es el menor de los conjuntos cerrados que contiene al conjunto dado, y que el interior es el mayor de los conjuntos abiertos contenidos en dicho conjunto.

6. Demuéstrese que los puntos de acumulación de un conjunto cualquiera forman un conjunto cerrado.

7. Pruébese que los puntos límites de uña sucesión forman un conjunto cerrado.

2. *Conjuntos conexos.*—Las estructuras más sencillas corresponden a los conjuntos abiertos y a los cerrados, y conjuntos más generales se utilizan muy raramente en la teoría elemental de funciones. Cuando se consideren conjuntos más generales se tratarán con relación a los conjuntos abiertos y cerrados.

Intuitivamente, un conjunto *conexo* es aquel que es de una sola pieza. La formalización de esta noción conduce a una definición que parece muy alejada de esta idea intuitiva. Tiene, sin embargo, la ventaja de ser lógicamente muy sencilla. Formularemos primero la definición para el caso de los conjuntos abiertos:

DEFINICIÓN 3. *Un conjunto abierto es conexo si no puede representarse como unión de dos conjuntos abiertos disjuntos no vacíos.*

En otras palabras: si A es un conjunto abierto conexo, y si hemos hallado una descomposición $A=B+C$, donde B y C son abiertos y $BC=0$, entonces o B es vacío o lo es C.

Para conjuntos arbitrarios adoptamos la siguiente definición:

DEFINICIÓN 4. *Un conjunto cualquiera A es conexo si no puede recubrirse por dos conjuntos abiertos cuyas intersecciones con A sean disjuntas y no vacias.*

Si B y C son conjuntos abiertos, el enunciado de la definición significa que $A \subset B+C$, $ABC=0$ implicaría $AB=0$ o $AC=0$. Las dos definiciones son consistentes, pues si A es abierto obtenemos una descomposición $A=AB+AC$ en conjuntos abiertos.

En la definición 4 podríamos haber requerido igualmente que los conjuntos recubridores fueran cerrados. En efecto, si B y C son abiertos, sus complementos B' y C' son cerrados. La hipótesis $A \subset B+C$, $ABC=0$ es equivalente a $AB'C'=0$, $A \subset B'+C'$. En estas condiciones tenemos, además, $AB=AC'$ y $AC=AB'$. Por tanto, la conclusión de que AB o AC es vacío equivale a la conclusión de que AB' o AC' es vacío.

Se sigue que la definición 3 tiene una análoga para conjuntos cerrados:

Un conjunto cerrado es conexo si no puede representarse como unión de dos conjuntos cerrados disjuntos no vacíos.

Conjuntos conexos triviales son el conjunto vacío y el consistente en un solo punto.

En el caso de la recta real, es posible citar todos los conjuntos conexos. El resultado más importante es que toda la recta es conexa

y esta es, indudablemente, una de las propiedades fundamentales del sistema de los números reales.

Se define un *intervalo* por una desigualdad de uno de los cuatro tipos siguientes: $a<x<b$, $a \leqq x<b$, $a<x \leqq b$, $a \leqq x \leqq b$ [1]. Para $a=-\infty$ o $b=+\infty$ quedan incluidos los intervalos semiinfinitos y la recta total.

TEOREMA 1. *Los subconjuntos conexos no vacíos de la recta real son los intervalos.*

Reproducimos una de las demostraciones clásicas, basada en el hecho de que cualquier sucesión monótona tiene límite finito o infinito.

Supongamos que la recta real $\mathfrak{R}$ se representa como la unión $\mathfrak{R}=A+B$ de dos conjuntos cerrados disjuntos. Si ninguno de ellos es vacío, podemos encontrar un $a_1 \in A$ y un $b_1 \in B$; nada impide suponer que $a_1<b_1$. Bisecamos el intervalo (a_1, b_1), y observamos que una de las dos mitades tiene su extremo izquierdo en A y su extremo derecho en B. Denotamos este intervalo por (a_2, b_2) y continuamos este proceso indefinidamente. De esta manera obtenemos una sucesión de intervalos encajados (a_n, b_n) con $a_n \in A$, $b_n \in B$. Las sucesiones $\{a_n\}$ y $\{b_n\}$ tienen un límite común c. Puesto que A y B son cerrados, c debería ser un punto común de A y B. Esta contradicción muestra que o bien A es vacío o lo es B y, por tanto, que $\mathfrak{R}$ es conexo.

La misma demostración, con ligeras modificaciones, es de aplicación a cualquier intervalo.

Antes de demostrar el recíproco haremos una indicación importante. Sea E un subconjunto cualquiera de $\mathfrak{R}$; se dice que α es una *cota inferior* de E si $\alpha \leqq x$ para todo $x \in E$. Consideremos el conjunto A de todas las cotas inferiores. Es evidente que el complemento de A es abierto. En cuanto a A, se ve fácilmente que es abierto, siempre y cuando no contiene un número mayor que todos. Puesto que la recta es conexa, A y su complemento no pueden ser ambos abiertos, a no ser que uno de ellos sea vacío. Existen, pues, tres posibilidades: A es vacío, A contiene un número mayor que todos o A es toda la recta. El mayor número a perteneciente a A, si es que existe, se llama *extremo inferior* de E; se denota corrientemente por

[1] Son de uso corriente las notaciones (a, b), $[a, b)$, $(a, b]$, $[a, b]$. En este libro utilizaremos la notación (a, b) para todos los tipos de intervalos, indicándose en el texto la naturaleza del intervalo.

$\overline{\text{extr}}\, E$. Si A es vacío, se conviene en poner $a = +\infty$, si es toda la recta, ponemos $a = -\infty$. Con este convenio todo conjunto de números reales tiene un extremo inferior determinado de manera única; evidentemente, $a = +\infty$ si (y solo si) el conjunto E es vacío. De manera correspondiente se define el *extremo superior* b, denotado por $\overline{\text{ext}}\, E$.

Volviendo a nuestra demostración, suponemos que E es un conjunto conexo cuyo extremo inferior es a y cuyo extremo superior es b. Todos los puntos de E se encuentran entre a y b, límites incluidos. Supongamos que un punto ξ del intervalo $a < \xi < b$ no pertenece a E. Entonces los conjuntos abiertos $x < \xi$ y $x > \xi$ recubrirían E, y puesto que E es conexo, uno de ellos no tendría parte común con E. Supongamos, p. ej., que ningún punto de E está a la izquierda de ξ. En este caso ξ sería una cota inferior, en contradicción con el hecho de que a es el extremo inferior. La suposición opuesta nos conduciría a una contradicción análoga, por lo que deducimos que ξ debe pertenecer a E. Por consiguiente, E es idéntico a uno de los cuatro intervalos (a, b), con lo que queda terminada la demostración.

En el curso de la demostración hemos introducido las nociones de extremo inferior y de extremo superior. Si el conjunto es cerrado y si los extremos son finitos, tienen que pertenecer al conjunto, en cuyo caso se llaman el mínimo y el máximo del mismo. Para poder estar seguros de que los extremos son finitos debemos saber que el conjunto no es vacío y de que existen ciertas cotas inferior y superior. En otras palabras: el conjunto debe estar contenido en un intervalo finito; a un conjunto como este se le llama *acotado*. Hemos demostrado:

TEOREMA 2. *Cualquier conjunto no vacío cerrado y acotado de números reales tiene un mínimo y un máximo.*

La estructura de los conjuntos conexos en el plano es bastante más complicada que en el caso de la recta, pero la siguiente caracterización de los conjuntos conexos abiertos contiene esencialmente toda la información que necesitaremos.

TEOREMA 3. *Un conjunto abierto no vacío del plano es conexo si (y solo si) dos cualesquiera de sus puntos pueden unirse mediante una línea poligonal contenida en el conjunto.*

La noción de línea poligonal es tan sencilla que no necesitamos dar una definición formal.

Demostraremos primero que la condición es necesaria. Sea A un conjunto abierto conexo y elijamos un punto $a \in A$. Denotamos por A_1 el subconjunto de A cuyos puntos pueden unirse con a mediante líneas poligonales contenidas en A, y por A_2 el conjunto de puntos que no pueden unirse con a de la manera indicada. Probemos que A_1 y A_2 son ambos abiertos. En primer lugar, si $a_1 \in A_1$, existe un entorno $|z - a_1| < \epsilon$ contenido en A. Todos los puntos en este entorno pueden unirse con a_1 mediante un segmento, y de aquí a a, mediante una línea poligonal; luego todo el entorno está contenido en A_1 y A_1 es abierto. En segundo lugar, si $a_2 \in A_2$, sea $|z - a_2| < \epsilon$ un entorno contenido en A. Si un punto de este entorno pudiera unirse a a mediante una línea poligonal, entonces a_2 podría unirse a este punto por un segmento, y de este a a. Esto contradice a la definición de A_2, por lo que A_2 es abierto. Puesto que A era conexo, o A_1 o A_2 tienen que ser vacíos. Pero A_1 contiene al punto a; por consiguiente, A_2 es vacío y todos los puntos pueden unirse a a. Por último, cualquier par de puntos de A puede unirse a través de a, con lo que queda demostrado que la condición es necesaria.

Para un futuro uso, hacemos observar que es incluso posible unir cualquier par de puntos mediante una poligonal cuyos lados sean paralelos a los ejes coordenados. La demostración es la misma.

Con objeto de demostrar la suficiencia, supongamos que A admite la descomposición $A = A_1 + A_2$ en dos conjuntos abiertos disjuntos. Elijamos $a_1 \in A_1$, $a_2 \in A_2$ y supongamos que estos puntos pueden unirse por una línea poligonal contenida en A. Uno de los lados de dicha poligonal debe entonces unir un punto de A_1 con un punto de A_2, bastando considerar por esta razón el caso en que a_1 y a_2 están unidos por un segmento. Este segmento tiene una representación paramétrica $z = a_1 + t(a_2 - a_1)$, donde t toma valores en el intervalo $0 \leqq t \leqq 1$. Los subconjuntos del intervalo $0 < t < 1$, que corresponden a los puntos en A_1 y A_2, respectivamente, son de forma evidente abiertos, disjuntos y no vacíos. Esto contradice el hecho de que el intervalo sea conexo, con lo que queda demostrado que la condición del teorema es suficiente.

DEFINICIÓN 5. *Un subconjunto conexo abierto no vacío del plano se llama una región.*

Del teorema 3 deducimos, pues, que todo el plano, un disco abierto $|z - a| < \rho$ y un semiplano son regiones. Una región es lo análogo en dos dimensiones de un intervalo abierto. Veremos que las regiones son los conjuntos más importantes en la teoría de funciones. El cierre de una región se llama una *región cerrada.* Debe observarse que regiones diferentes pueden tener el mismo cierre.

Ocurre con frecuencia que tenemos que analizar la estructura de conjuntos que se definen implícitamente, p. ej., en el curso de una demostración. En tales casos, el primer paso es descomponer el conjunto en sus *componentes* conexas máximas. Como el nombre indica, una componente es un subconjunto conexo que no está contenido en otro subconjunto conexo mayor. Así, p. ej., si se da un conjunto como unión de dos regiones disjuntas, cabe esperar que esas regiones sean sus componentes. Es importante demostrar esto y probar que en un conjunto arbitrario cada punto pertenece a una componente bien definida.

TEOREMA 4. *Todo conjunto tiene una única descomposición en componentes. Las componentes de un conjunto abierto son regiones.*

Si es E el conjunto dado, consideremos un punto $a \in E$ y sea A la unión de todos los subconjuntos conexos de E que contienen a a. Entonces A contiene seguramente a a, ya que el conjunto que contiene un solo punto a es conexo. Si podemos demostrar que A es conexo, entonces, evidentemente, es un conjunto conexo máximo que contiene a a, siendo el único en estas condiciones. En otras palabras: A será la componente que contiene a a, y dos componentes distintas cualesquiera son disjuntas. Esto es lo que deseamos demostrar.

Supongamos que A no fuera conexo. Por definición podríamos hallar conjuntos abiertos B y C, tales que $A \subset B+C$ y $ABC=0$, mientras que ni AB ni AC son vacíos. Podemos suponer que B contiene el punto a; lo anterior implica que AC contiene un punto c. Puesto que $c \in A$, existe un conjunto conexo $A_0 \subset E$ que contiene a y c. Pero esto es imposible, pues $A_0 \subset B+C$ y $A_0BC=0$, A_0B contiene a a, y A_0C contiene a c, lo que contradice el hecho de ser A_0 conexo. En definitiva, A es conexo.

La última parte del teorema se deduce fácilmente al considerar que cada entorno es conexo. En la recta real cada componente de un conjunto abierto es un intervalo abierto. Por tanto, el conjunto abierto más general sobre la recta es una unión de intervalos abiertos.

EJERCICIOS

1. Pruébese que la unión de dos discos abiertos es conexo si (y solo si) la distancia entre los centros es menor que la suma de los radios. ¿Cuál es el resultado si uno o los dos discos son cerrados?

2. Demuéstrese que una región sigue siendo una región si se suprime un número finito de puntos.

3. Pruébese que el cierre de un conjunto conexo es conexo.

4. Hágase la prueba de que las componentes de un conjunto cerrado son cerradas.

3. *Conjuntos compactos.*—Recordemos que un subconjunto de la recta real se dice acotado si está contenido en un intervalo finito. Análogamente, un conjunto en el plano complejo está acotado si se halla contenido en un disco. En particular, un conjunto está acotado si (y solo si) se halla contenido en un disco cerrado $|z| \leqq M$ con centro en el origen.

Todo conjunto cerrado y acotado de la recta real o del plano complejo se llama *compacto.* En el plano completo, a cualquier conjunto cerrado se le llama compacto.

Para nuestros propósitos, pues, la palabra "compacto" es meramente una abreviatura conveniente. Sin embargo, es importante que demostremos ciertas propiedades características de los conjuntos compactos.

En análisis, muchos razonamientos se simplifican notablemente si están basados en un lema, conocido indistintamente como el lema de Borel, el lema de Heine-Borel o el lema de Heine-Borel-Lebesgue.

TEOREMA 5. *(Lema de Heine-Borel.) Supongamos que un conjunto compacto A está contenido en la unión de una colección de conjuntos abiertos. Entonces A está también contenido en la unión de un número finito de estos conjuntos abiertos.*

A la colección de conjuntos abiertos se le llama *recubridor abierto* de A, y el teorema afirma que es posible extraer un subrecubridor finito. Nos referiremos a esta propiedad como la propiedad de Heine-Borel.

Serán de utilidad unas cuantas observaciones previas. Si A tiene la propiedad de Heine-Borel y si A_0 es un subconjunto cerrado de A, entonces A_0 tiene la misma propiedad. En efecto, supongamos que A_0 está cubierto por una colección $\{U\}$ de conjuntos abiertos. Entonces A está cubierto por la colección que consiste en $\{U\}$ y el complemento de A_0. De acuerdo con nuestra hipótesis, podemos recubrir A con una subcolección finita, que puede incluir el complemento $C(A_0)$. La misma subcolección, sin $C(A_0)$, recubre a A_0.

Como resultado de la anterior observación bastará probar el teorema para un intervalo cerrado en el caso de la recta y para un cuadrado $|x| \leqq M$, $|y| \leqq M$ en el caso del plano. El caso del plano ampliado se reduce fácilmente al del plano. En efecto, si el conjunto cerrado A no incluye ∞, está acotado; si lo incluye, uno de los conjuntos abiertos recubridores contiene a ∞, y bastará considerar el subconjunto cerrado acotado de A exterior a este conjunto abierto particular.

Demostraremos ahora que un intervalo finito cerrado I_1 tiene la propiedad de Heine-Borel. La demostración es indirecta y por el método de bisección. Consideremos un recubridor de I_1 constituido por una colección $\{U\}$ de conjuntos abiertos y supongamos que I_1 no tiene un subrecubridor finito. Si bisecamos I_1, al menos uno de los subintervalos no posee subrecubridor finito. Este subintervalo se denota por I_2; si se ha de elegir entre los dos subintervalos, podemos hacer que la selección quede bien definida tomando siempre el subintervalo izquierdo. Repitiendo el proceso obtenemos una sucesión de intervalos encajados I_n sin subrecubridores finitos. Los extremos izquierdo y derecho de I_n convergen hacia un punto común $\xi \in I_1$. Por hipótesis, ξ pertenece a un conjunto U, y puesto que U es un entorno abierto de ξ, está contenido en U. Tenemos entonces que para n suficientemente grande $I_n \subset U$, lo que contradice el hecho de que I_n no puede recubrirse por un número finito de conjuntos de $\{U\}$. La contradicción prueba el teorema para conjuntos compactos de la recta real.

En el caso del plano complejo puede utilizarse la misma demostración; I_1 es un cuadrado, y en cada paso dividimos el cuadrado dado en cuatro cuadrados más pequeños con lado mitad. La sucesión decreciente de cuadrados I_n se determina como en el caso anterior, llegándose a la misma contradicción.

Como aplicación típica que ilustra la utilización del lema de Heine-Borel probaremos el siguiente teorema:

TEOREMA 6. *(Bolzano-Weierstrass.) Si A es un conjunto compacto, toda sucesión infinita de puntos $a_n \in A$ tiene al menos un punto límite en A.*

Consideremos la colección formada por todos los conjuntos abiertos U, tales que $a_n \in U$, como máximo para un número finito de subíndices n. Todo punto que no es punto límite de la sucesión $\{a_n\}$ tiene un entorno con esta propiedad. Por consiguiente, si no hay punto límite en A, los conjuntos U forman un recubridor abierto de A. Podríamos entonces recubrir A con un número finito de conjuntos U. Se seguiría de aquí que $a_n \in A$ únicamente para un número finito de subíndices, lo que contradice la hipótesis. Por tanto, debe haber al menos un punto límite en A.

No es preciso decir que el teorema de Bolzano-Weierstrass podría haberse demostrado directamente por el método de bisección. Lo importante es el hecho de que el lema de Heine-Borel hace posible evitar repeticiones de lo que esencialmente es la misma demostración. Por otra parte, las demostraciones por bisección son a menudo más elementales, y no dudaremos en utilizarlas cuando tiendan a facilitar la comprensión.

TEOREMA 7. *Todo conjunto con la propiedad de Heine-Borel es compacto.*

Evidentemente, basta probar este recíproco del lema de Heine-Borel para el caso del plano ampliado. En efecto, si un conjunto es cerrado con respecto al plano ampliado y no contiene el punto del infinito, es *eo ipso* acotado.

Sea A un conjunto con la propiedad de Heine-Borel y consideremos un punto $a \in C(A)$. Denotemos por $\{U\}$ la colección de todos los conjuntos abiertos U cuyo cierre $\bar{U}$ no contiene a a. Es evidente que todo punto A tiene un entorno con esta propiedad. Por consiguiente, la colección $\{U\}$ es un recubridor abierto de A, pudiéndose extraer un subrecubridor finito $U_1, U_2, \ldots, U_N$. Entonces a está contenido en la intersección de $C(\bar{U}_1), C(\bar{U}_2), \ldots, C(\bar{U}_N)$, y esta intersección es un subconjunto abierto de $C(A)$. Puesto que a era un punto arbitrario de $C(A)$, se tiene que $C(A)$ es abierto y, por tanto, A es cerrado.

EJERCICIOS

1. Aplíquese el lema de Heine-Borel para demostrar que un conjunto cerrado y acotado de números reales tiene un máximo.

2. Pruébese que una sucesión decreciente de conjuntos compactos $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ tiene una intersección no vacía (lema de Cantor).

4. *Funciones continuas y aplicaciones.*—Deliberadamente, en la sección 2-1 hemos dejado de lado aquellas funciones que no estaban definidas para todos los valores de la variable independiente. Adoptaremos ahora un punto de vista más general y admitiremos funciones $f(x)$ definidas únicamente cuando x pertenece a un cierto conjunto A. Se ha de entender que el citar el conjunto A es una

parte esencial de la definición de la función. Si $A_0 \subset A$, podemos naturalmente definir una función con los mismos valores $f(x)$ para $x \in A_0$. Se le llama la *restricción* de $f(x)$ al conjunto A_0, y ordinariamente se puede utilizar la misma notación para una función y para sus restricciones. Recíprocamente, la función $f(x)$ sobre A es una *prolongación* de la función $f(x)$ en A_0.

No hacemos distinción entre una función f y la *aplicación* que define, excepto que el uso de la palabra "aplicación" hace resaltar la correspondencia entre puntos como opuesta a la correspondencia entre números. El conjunto de todos los puntos $f(x)$ para $x \in A$ se llama la *imagen* de A con respecto a la aplicación f; conviene denotar la imagen mediante $f(A)$. Es habitual decir que f aplica A *en* un conjunto A^* si $f(A) \subset A^*$ y *sobre* el conjunto A^* si $f(A)=A^*$. Para un conjunto arbitrario A^*, se denota mediante $f^{-1}(A^*)$ el conjunto de todos los puntos $x \in A$ tales que $f(x) \in A^*$, y se llama a $f^{-1}(A)^*$ la *imagen inversa* de A^*. La aplicación de A sobre $f(A)$ es *biunívoca* si $f(x_1)=f(x_2)$ implica $x_1=x_2$. En este caso, $y=f(x)$ posee una *función inversa* $x=f^{-1}(y)$ definida en $f(A)$.

La definición de una función continua necesita muy poca modificación. Decimos que $f(x)$ es continua para $x=a$ si, dado $\epsilon>0$, existe un $\delta>0$ tal que $|f(x)-f(a)|<\epsilon$ para todo $x \in A$ que verifique $|x-a|<\delta$. La generalización al caso en que x o $f(x)$ varíen sobre todo el plano completo es obvia.

Si A es o bien un conjunto abierto o cerrado, se pueden dar las siguientes caracterizaciones útiles de las funciones continuas en A.

Una función $f(x)$, definida en un conjunto abierto, es continua si (y solo si) la imagen inversa de todo conjunto abierto es abierta.

Una función $f(x)$, definida en un conjunto cerrado, es continua si (y solo si) la imagen inversa de todo conjunto cerrado es cerrada.

La primera condición es una traducción directa de la definición. La segunda se deduce considerando los complementos, detalles que se dejan al lector.

Es muy importante percatarse que estas propiedades se verifican solo para la imagen inversa y no para la directa. La imagen de un conjunto abierto respecto a una aplicación continua no es siempre un conjunto abierto, ni la de uno cerrado es siempre un conjunto cerrado. La última conclusión es, sin embargo, correcta en el caso siguiente:

TEOREMA 8. *Bajo una aplicación continua, la imagen de todo conjunto compacto es un conjunto compacto.*

Sea $f(x)$ continua en el conjunto compacto A. Demostraremos que $f(A)$ posee la propiedad de Heine-Borel; por el teorema 7 será entonces compacto. Sea $\{U\}$ un recubridor abierto cualquiera de $f(A)$. Consideremos la colección $\{V\}$ de todos los conjuntos abiertos V tales que $f(AV)$ está contenido en un conjunto U. Si $a \in A$, sabemos que $f(a)$ pertenece a algún U, y por la continuidad existe un entorno de a que es un conjunto V. Por consiguiente, $\{V\}$ es un recubridor abierto de A, y podemos seleccionar un subrecubridor finito $\{V_n\}$, $n=1, 2, \ldots, N$. Si $f(AV_n) \subset U_n$, es obvio que $\{U_n\}$, $n=1, 2, \ldots, N$, es un subrecubridor finito de $f(A)$.

COROLARIO. *Una función continua real sobre un conjunto compacto posee un mínimo y un máximo.*

La imagen $f(A)$ es un conjunto cerrado y acotado de la recta real. La existencia de un mínimo y de un máximo se deduce del teorema 2.

TEOREMA 9. *Bajo una aplicación continua, la imagen de cualquier conjunto conexo es un conjunto conexo.*

Suponemos que $f(x)$ es continua en el conjunto conexo A. Si A es abierto, la demostración es particularmente sencilla. Sean B y C conjuntos abiertos tales que $f(A) \subset B+C$ y $f(A)BC=0$. Entonces $f^{-1}(B)$ y $f^{-1}(C)$ son también abiertos, de donde $A=f^{-1}(B)+f^{-1}(C)$, $f^{-1}(B), f^{-1}(C)=0$. Dado que A es conexo, $f^{-1}(B)$ o $f^{-1}(C)$ es vacío, y esto es cierto solo si $f(A)B$ o $f(A)C$ es vacío. Por tanto, $f(A)$ es conexo.

En el caso general, $f^{-1}(B)$ y $f^{-1}(C)$ no son abiertos; pero por la continuidad cada punto de $f^{-1}(B)$ tiene un entorno cuya intersección con A está contenida en $f^{-1}(B)$. La unión de todos estos entornos es un conjunto abierto B_0 tal que $B_0A=f^{-1}(B)$; el conjunto C_0 puede definirse de manera análoga. Ahora el mismo razonamiento hecho anteriormente muestra que o bien B_0A o C_0A es vacío, y esto implica que $f(A)B$ o $f(A)C$ es vacío.

Una aplicación biunívoca f de un conjunto A sobre $f(A)$ se llama una *aplicación topológica* o un *homeomorfismo* si f y f^{-1} son ambas continuas. Una propiedad de un conjunto que es compartida por todas las imágenes topológicas del conjunto se denomina *propiedad topológica*. Así, p. ej., hemos demostrado que la compacidad y el hecho de ser un conjunto conexo son propiedades topológicas (teoremas 8 y 9). *No* hemos demostrado que la noción de conjunto abierto sea topológicamente invariante, aunque el hecho es cierto.

El concepto de *continuidad uniforme* será utilizado constantemente. De manera general, se dice que una condición se verifica de forma uniforme con respecto a un parámetro si puede expresarse mediante desigualdades que no implican al parámetro. En el presente caso venimos a parar a la siguiente definición:

DEFINICIÓN 6. *Se dice que una función $f(x)$ es uniformemente continua en un conjunto A si para todo $\epsilon > 0$ existe un $\delta > 0$, tal que $|f(x_1) - f(x_2)| < \epsilon$, para todos los pares x_1, $x_2 \in A$ que satisfagan $|x_1 - x_2| < \delta$.*

La definición se ha formulado para el caso en que x y $f(x)$ toman sus valores en la recta real o en el plano complejo. Con objeto de extenderla al plano ampliado, los valores absolutos deberán ser reemplazados por distancias en la esfera de Riemann.

Una función uniformemente continua lo es en todos los puntos de A en el sentido ordinario, pero la recíproca no es cierta. En efecto, en la definición de la continuidad ordinaria el punto x_1 (o el x_2) se mantiene fijo, y δ depende de x_1.

En vista de esta diferencia básica entre las definiciones, el siguiente resultado es de gran importancia:

TEOREMA 10. *Toda función continua en un conjunto compacto es uniformemente continua.*

Esta es una consecuencia típica de la propiedad de Heine-Borel. Sea $f(x)$ continua en el conjunto compacto A. Entonces todo $a \in A$ tiene un entorno $|x - a| < \rho$, en el que $|f(x) - f(a)| < \frac{1}{2}\epsilon$ (para abreviar suprimimos la condición obvia de que $x \in A$). Los entornos más pequeños $|x - a| < \frac{1}{2}\rho$ forman un recubridor abierto de A, siendo posible extraer un subrecubridor finito compuesto por entornos $|x - a_n| < \frac{1}{2}\rho_n$. Sea δ el menor de los números $\frac{1}{2}\rho_n$. Consideremos un par x_1, $x_2 \in A$ con $|x_1 - x_2| < \delta$. Existe un a_n tal que $|x_1 - a_n| < \frac{1}{2}\rho_n$, y obtenemos

$$|x_2 - a_n| \leqq |x_1 - a_n| + |x_1 - x_2| < \tfrac{1}{2}\rho_n + \delta \leqq \rho_n.$$

En consecuencia, $|f(x_1) - f(a_n)| < \frac{1}{2}\epsilon$ y $|f(x_2) - f(a_n)| < \frac{1}{2}\epsilon$. Por la desigualdad triangular se deduce que $|f(x_1) - f(x_2)| < \epsilon$, con lo que queda demostrada la continuidad uniforme de $f(x)$ en A.

En conjuntos que no son compactos, algunas funciones continuas

son uniformemente continuas, en tanto que otras no lo son. Así, p. ej., la función z es uniformemente continua en todo el plano, pero no lo es la función z^2.

Debe concederse especial atención al hecho de que una función uniformemente continua en un conjunto A lo es también en todo subconjunto de A. Por consiguiente, si $f(x)$ está definida en A_0 y puede prolongarse a una función sobre un conjunto compacto $A \supset A_0$, entonces $f(x)$ es uniformemente continua en A_0.

EJERCICIOS

1. Constrúyase una aplicación topológica de todo el plano sobre el disco $|z| < 1$. Conclusión: el concepto de conjunto cerrado no es topológico.

2. Pruébese que toda aplicación topológica de un conjunto abierto A sobre un conjunto abierto B aplica subconjuntos abiertos de A sobre subconjuntos abiertos de B, y viceversa.

3. Demuéstrese que toda aplicación biunívoca continua de un conjunto compacto es topológica.

4. Pruébese que dos conjuntos cerrados disjuntos, ninguno de ellos vacío y al menos uno compacto, tienen una distancia mínima positiva.

5. ¿Cuáles de las siguientes funciones son uniformemente continuas en toda la recta real: $\operatorname{sen} x$, $x \operatorname{sen} x$, $x \operatorname{sen} (x)^2$, $|x|^{\frac{1}{2}} \operatorname{sen} x$?

6. ¿Es uniformemente continua la función $x \log x$ en el intervalo abierto $(0, 1)$?

5. *Arcos y curvas cerradas.*—Una manera muy conveniente de dar la ecuación de un *arco* γ en el plano es en forma paramétrica $x = x(t)$, $y = y(t)$, donde t varía en un intervalo $\alpha \leqq t \leqq \beta$, siendo $x(t)$ e $y(t)$ funciones continuas. Podemos también utilizar la notación compleja $z = z(t) = x(t) + iy(t)$, que tiene algunas ventajas.

Así, un arco, considerado como un conjunto de puntos, es la imagen de un intervalo finito cerrado respecto a una aplicación continua. Como tal, es compacto y conexo. No obstante, un arco no es simplemente un conjunto de puntos, sino muy esencialmente una sucesión de puntos también, ordenada para valores crecientes del parámetro. Si una función no decreciente $t = \varphi(\tau)$ aplica un intervalo $\alpha' \leqq \tau \leqq \beta'$ sobre $\alpha \leqq t \leqq \beta$, entonces $z = z[\varphi(\tau)]$ define la misma sucesión de puntos que $z = z(t)$. Decimos que la primera ecuación proviene de la segunda por un *cambio de parámetro*. El cambio es *reversible* si (y solo si) $\varphi(\tau)$ es estrictamente creciente. Así, p. ej., la ecuación $z = t^2 + it^4$, $0 \leqq t \leqq 1$ se obtiene por un cambio reversible de parámetro de la ecuación $z=t+it^2$, $0\leqq t\leqq 1$. Siempre se puede obtener un cambio del intervalo paramétrico (α,β) mediante un cambio *lineal* de parámetro; es decir, de la forma $t=a\tau+b$, $a>0$.

Lógicamente, lo más sencillo es considerar como diferentes dos arcos si vienen dados por ecuaciones diferentes, sin tener en cuenta si una ecuación proviene de la otra por un cambio de parámetro. Al adoptar este punto de vista, como haremos nosotros, es importante mostrar que ciertas propiedades de los arcos son invariantes frente a los cambios de parámetro; p. ej., los puntos *origen* y *extremo* de un arco permanecen los mismos después de un cambio de parámetro.

Si existe la derivada $z'(t)=x'(t)+iy'(t)$ y es distinta de cero, el arco γ tiene una *tangente,* cuya dirección viene dada por $\arg z'(t)$. Diremos que el arco es *diferenciable* si existe $z'(t)$ y es continua (el término continuamente diferenciable es demasiado largo); si, además, $z'(t)\neq 0$, se dice que el arco es *regular.* Un arco es *diferenciable a trozos* o *regular a trozos* si se verifican las mismas condiciones, excepto para un número finito de valores de t; en estos puntos, $z(t)$ será todavía continua con derivadas por la izquierda y por la derecha, las cuales son iguales a los límites por la izquierda y por la derecha de $z'(t)$, y en el caso de un arco regular a trozos, $\neq 0$.

El carácter diferenciable o regular de un arco es invariante frente al cambio de parámetro $t=\varphi(\tau)$ siempre y cuando $\varphi'(\tau)$ sea continua y, para la regularidad, $\neq 0$. Cuando este es el caso, decimos que el cambio de parámetro es diferenciable o regular.

Un arco es *simple,* o un *arco de Jordan,* si $z(t_1)=z(t_2)$ únicamente para $t_1=t_2$. Un arco es una *curva cerrada* si los extremos coinciden: $z(\alpha)=z(\beta)$. Para curvas cerradas un *desplazamiento* del parámetro está definido de la siguiente forma: Si la ecuación original es $z=z(t)$, $\alpha\leqq t\leqq\beta$, elegimos un punto t_0 del intervalo (α,β) y definimos una nueva curva cerrada, cuya ecuación es $z=z(t)$ para $t_0\leqq t\leqq\beta$, y $z=z(t-\beta+\alpha)$ para $\beta\leqq t\leqq t_0+\beta-\alpha$. El propósito del desplazamiento es eliminar la posición distinguida del origen. Las definiciones correctas de una curva cerrada diferenciable o regular y de *curva cerrada simple* (o *curva de Jordan)* son obvias.

El *arco opuesto* del $z=z(t)$, $\alpha\leqq t\leqq\beta$, es el arco $z=z(-t)$, $-\beta\leqq t\leqq-\alpha$. Arcos opuestos se denotan por γ y $-\gamma$, a veces por γ y γ^{-1}, según los casos. Una función constante $z(t)$ define una *curva punto.*

Una circunferencia C, definida originalmente como un lugar geométrico $|z-a|=r$, puede considerarse como una curva cerrada de ecuación $z=a+re^{it}$, $0\leqq t\leqq 2\pi$. Utilizaremos esta parametrización típica siempre que manejemos una circunferencia finita. Este convenio nos ahorra el escribir la ecuación cada vez que se necesite; también y este es su más importante propósito, sirve como una regla perfectamente definida para distinguir entre C y $-C$.

2-3. Funciones analíticas en una región.—Las consideraciones anteriores sobre conjuntos de puntos han preparado el camino para llegar a una definición precisa de las funciones analíticas. Podríamos considerar ahora una función $f(z)$ definida en un conjunto de puntos A arbitrario y exigir la existencia del límite

$$\lim_{h\to 0}\frac{f(z+h)-f(z)}{h}$$

para todo $z\in A$ cuando se restringe h a valores para los cuales $z+h\in A$.

Esta generalidad es decididamente excesiva. Su mayor desventaja estriba en el hecho de que la existencia de la derivada puede ser una condición débil o fuerte, dependiendo de la naturaleza del conjunto A en la inmediata vecindad del punto. La única forma sencilla mediante la cual puede evitarse esto es suponer que cada punto tiene un entorno contenido en A. Llegamos así a tener en cuenta solo las funciones analíticas que están definidas en conjuntos abiertos.

Veremos que se obtienen mayores ventajas si se define $f(z)$ en un conjunto conexo. Para conseguir el apropiado grado de generalidad supondremos, por consiguiente, que toda función analítica está definida en una región.

1. *Definición y consecuencias inmediatas.*—Empezamos con un enunciado formal de la definición:

DEFINICIÓN 7. *Una función compleja $f(z)$ se dice analítica en la región Ω si está definida y tiene derivada en cada punto de Ω.*

De acuerdo con esta definición, una función analítica en Ω es siempre uniforme. Es muy importante que la definición esté localizada en una región fija Ω, y no es lícito hablar de una función analítica sin especificar la región en la cual se considera. A veces esta región aparece claramente implicada en el contexto, pudiendo en tales casos omitirse la referencia explícita.

A fin de disponer de una mayor flexibilidad en el lenguaje es conveniente introducir el complemento siguiente a la definición 7:

DEFINICIÓN 8. *Una función $f(z)$ es analítica en un conjunto de puntos arbitrario A si es analítica en alguna región que contenga a A.*

Es evidente que la última definición es sencillamente un convenio para utilizar una terminología adecuada. Este es un caso en el que la región Ω no necesita ser mencionada explícitamente, pues se comprobará que la elección específica de Ω es indiferente siempre y cuando contenga a A. Una aplicación típica es el uso de la frase: "Sea $f(z)$ analítica en z_0." Ello significa que está definida una función $f(z)$, la cual posee derivada en algún entorno de z_0 que no necesita ser especificado.

Aunque nuestra definición requiere que todas las funciones analíticas sean uniformes, es posible considerar funciones multiformes tales como $\sqrt{z}$, $\log z$ o $\operatorname{arc\,cos} z$, con tal que se restrinjan a una región definida en la que sea posible elegir una rama uniforme y analítica de la función.

Así, p. ej., podemos elegir como Ω el complemento del eje real negativo $z \leqq 0$; indudablemente, este conjunto es abierto y conexo. En Ω, uno (y solo uno) de los valores de $\sqrt{z}$ tiene parte real positiva. Con esta elección $w = \sqrt{z}$ se convierte en una función uniforme en Ω; demostremos que es continua. Elijamos dos puntos $z_1, z_2 \in \Omega$ y denotemos los valores correspondientes de w por $w_1 = u_1 + iv_1$, $w_2 = u_2 + iv_2$, con $u_1, u_2 > 0$. Entonces

$$|z_1 - z_2| = |w_1^2 - w_2^2| = |w_1 - w_2| \cdot |w_1 + w_2|$$

y $|w_1 + w_2| \geqq u_1 + u_2 > u_1$. Luego

$$|w_1 - w_2| < \frac{|z_1 - z_2|}{u_1}$$

y se deduce que $w = \sqrt{z}$ es continua en z_1. Una vez establecida la continuidad, la analiticidad resulta por derivación de la función inversa $z = w^2$. En efecto, con la notación habitual $\Delta z \to 0$ implica $\Delta w \to 0$. Por consiguiente,

$$\lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta w \to 0} \frac{\Delta w}{\Delta z}$$

y obtenemos

$$\frac{dw}{dz} = \frac{1}{\dfrac{dz}{dw}} = \frac{1}{2w} = \frac{1}{2\sqrt{z}}$$

con la misma rama de $\sqrt{z}$.

En el caso de $\log z$ podemos utilizar la misma región Ω, obtenida por la exclusión del eje real negativo, y definir la *rama principal* del logaritmo por la condición $|\operatorname{Im} \log z| < \pi$. Nuevamente, es necesario demostrar la continuidad, pero esta vez no disponemos de una identidad algebraica y estamos forzados a utilizar un razonamiento más general. Denotemos la rama principal por $w = u + iv = \log z$. Para un punto dado $w_1 = u_1 + iv_1$, $|v_1| < \pi$, y un $\epsilon > 0$ dado, consideremos el conjunto A en el plano w definido por las desigualdades $|w - w_1| \geqq \epsilon$, $|v| \leqq \pi$, $|u - u_1| \leqq \log 2$. Este conjunto es cerrado y acotado, y para ϵ suficientemente pequeño no es vacío. La función continua $|e^w - e^{w_1}|$ tiene, por tanto, un mínimo ρ en A (corolario del teorema 8). Este mínimo es positivo, pues A no contiene ningún punto $w_1 + n \cdot 2\pi i$. Elijamos $\delta = \text{mín}\,(\rho, \frac{1}{2}e^{u_1})$, y supongamos que

$$|z_1 - z_2| = |e^{w_1} - e^{w_2}| < \delta.$$

Entonces w_2 no puede estar en A, pues esto haría $|e^{w_1} - e^{w_2}| \geqq \rho \geqq \delta$. Tampoco es posible que $u_2 < u_1 - \log 2$ o $u_2 > u_1 + \log 2$; en el primer caso obtendríamos $|e^{w_1} - e^{w_2}| \geqq e^{u_1} - e^{u_2} > \frac{1}{2}e^{u_1} \geqq \delta$, y en el último $|e^{w_1} - e^{w_2}| \geqq e^{u_2} - e^{u_1} > e^{u_1} > \delta$. Por consiguiente, w_2 debe pertenecer al disco $|w - w_1| < \epsilon$, y hemos demostrado que w es una función continua de z. De la continuidad deducimos, como anteriormente, que la derivada existe y es igual a $1/z$.

Los infinitos valores de $\operatorname{arc\,cos} z$ son los mismos que los valores de $i \log (z + \sqrt{z^2 - 1})$. En este caso restringimos z al complemento Ω' de las semirrectas $x \leqq 0$, $y = 0$ y $x \geqq 1$, $y = 0$. Puesto que $1 - z^2$ nunca es real y $\leqq 0$ en Ω', podemos definir $\sqrt{1 - z^2}$ como en el primer ejemplo y poner entonces $\sqrt{z^2 - 1} = i\sqrt{1 - z^2}$. Además, $z + \sqrt{z^2 - 1}$ no puede ser negativo o nulo en Ω'; en efecto, puesto que $z + \sqrt{z^2 - 1}$ y $z - \sqrt{z^2 - 1}$ son recíprocos, $z + \sqrt{z^2 - 1} < 0$ implicaría $z - \sqrt{z^2 - 1} < 0$, y de aquí, $2z < 0$. Resulta así posible definir una rama analítica de $\log (z + \sqrt{z^2 - 1})$, cuya parte imaginaria quede comprendida entre $-\pi$ y π. De esta forma obtenemos una función analítica uniforme

$$\operatorname{arc\,cos} z = i \log (z + \sqrt{z^2 - 1})$$

en Ω' cuya derivada es

$$D \operatorname{arc\,cos} z = i \frac{1}{z + \sqrt{z^2 - 1}} \left(1 + \frac{z}{\sqrt{z^2 - 1}} \right) = \frac{1}{\sqrt{1 - z^2}}$$

donde $\sqrt{1 - z^2}$ tiene una parte real positiva.

No existe unicidad en la elección de la región y de la rama uniforme en los ejemplos anteriores. Por consiguiente, cada vez que se considere una función tal como $\log z$, se habrá de especificar la elección de la rama. Es un hecho fundamental la *imposibilidad* de definir una rama analítica y uniforme de $\log z$ en ciertas regiones. Esto se demostrará en el capítulo sobre integración.

Todos los resultados de la sección 1, párrafo 2, permanecen válidos para funciones que son analíticas en una región. En particular, las partes real e imaginaria de una función analítica en Ω satisfacen las ecuaciones de Cauchy-Riemann:

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

Recíprocamente, si u y v satisfacen estas ecuaciones en Ω y si las derivadas parciales son continuas, entonces $u+iv$ es una función analítica en Ω.

Una función analítica en Ω *degenera* si se reduce a una constante. En el teorema siguiente relacionamos algunas sencillas condiciones que entrañan esta consecuencia:

TEOREMA 11. *Una función analítica en una región Ω, cuya derivada se anula idénticamente, debe reducirse a una constante. Lo mismo es cierto si la parte real, la parte imaginaria, el módulo o el argumento son constantes.*

La anulación de la derivada implica que $\partial u/\partial x$, $\partial u/\partial y$, $\partial v/\partial x$, $\partial v/\partial y$ son todas cero. Se sigue que u y v son constantes en cualquier segmento de Ω que sea paralelo a uno de los ejes coordenados. En la página 75 hicimos observar, en relación con el teorema 3, que cualquier par de puntos de una región pueden unirse dentro de la región por una línea poligonal de lados paralelos a los ejes. En conclusión, $u+iv$ es constante.

Si u o v son constantes, se tiene que

$$f'(z)=\frac{\partial u}{\partial x}-i\frac{\partial u}{\partial y}=\frac{\partial v}{\partial y}+i\frac{\partial v}{\partial x}=0,$$

y, por tanto, $f(z)$ tiene que ser constante. Si u^2+v^2 es constante, se obtiene

$$u\frac{\partial u}{\partial x}+v\frac{\partial v}{\partial x}=0$$

y

$$u\frac{\partial u}{\partial y}+v\frac{\partial v}{\partial y}=-u\frac{\partial v}{\partial x}+v\frac{\partial u}{\partial x}=0.$$

Estas ecuaciones nos llevan a la conclusión de que $\partial u/\partial x=\partial v/\partial x=0$, a no ser que el determinante u^2+v^2 se anule. Pero si $u^2+v^2=0$ en un único punto, será constantemente cero y $f(z)$ se anula idénticamente. Luego $f(z)$ es en cualquiera de los casos una constante.

Por último, si $\arg f(z)$ es constante, podemos poner $u=kv$, siendo k una constante (a no ser que sea v idénticamente nula). Pero $u-kv$ es la parte real de $(1+ik)f$, y se obtiene nuevamente que f se reduce a una constante.

EJERCICIOS

1. Defínase una rama analítica uniforme de $\sqrt{1+z}+\sqrt{1-z}$ en una región adecuada.
2. El mismo problema para $\log\log z$.
3. Pruébese que una función analítica en una región cuyas partes real e imaginaria satisfacen la ecuación $v=u^2$ tiene que reducirse a una constante.

2. *Representación conforme.*—Supongamos que un arco γ de ecuación $z=z(t)$, $\alpha\leqq t\leqq\beta$, está contenido en una región Ω, y sea $f(z)$ una función definida y continua en Ω. Entonces la ecuación $w=w(t)=f[z(t)]$ define un arco γ' en el plano w, al que llamaremos la *imagen* de γ.

Consideremos el caso en que $f(z)$ es analítica en Ω. Si $z'(t)$ existe, tenemos que $w'(t)$ también existe y está determinada por

$$w'(t)=f'[z(t)]z'(t). \qquad [2\text{-}25]$$

Vamos a investigar el significado de esta ecuación en un punto $z_0=z(t_0)$ con $z'(t_0)\neq 0$ y $f'(z_0)\neq 0$.

La primera conclusión es que $w'(t_0)\neq 0$. Por consiguiente, γ' tiene una tangente en $w_0=f(z_0)$, y su dirección está determinada por

$$\arg w'(t_0)=\arg f'(z_0)+\arg z'(t_0). \qquad [2\text{-}26]$$

Esta relación afirma que el ángulo entre las tangentes orientadas a γ en z_0 y a γ' en w_0 es igual a $\arg f'(z_0)$. Por tanto, es independiente de la curva γ. Por esta razón, curvas tangentes en el punto z_0 se transforman en curvas con tangente común en w_0. Más aún, dos curvas que forman un determinado ángulo en z_0 se transforman en curvas que se cortan bajo el mismo ángulo, tanto en sentido como en magnitud. En razón de verificarse esta propiedad, la aplicación efectuada por $w=f(z)$ se dice que es *conforme* en todos los puntos en que $f'(z) \neq 0$.

Una propiedad relacionada con la anterior se deduce al considerar el módulo $|f'(z_0)|$. Se tiene

$$\lim_{z \to z_0} \frac{|f(z)-f(z_0)|}{|z-z_0|} = |f'(z_0)|,$$

y esto significa que cualquier segmento pequeño con origen en z_0, en el límite, es ampliado o reducido en la razón $|f'(z_0)|$. En otras palabras: el cambio lineal de escala en z_0, efectuado por la transformación $w=f(z)$, es independiente de la dirección. En general, este cambio de escala variará de un punto a otro.

Recíprocamente, está claro que ambas propiedades juntas implican la existencia de $f'(z_0)$. Es menos obvio que cada una de ellas por separado implica el mismo resultado, al menos bajo ciertas hipótesis adicionales de regularidad.

Para obtener mayor precisión, supongamos que las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ son continuas. En estas condiciones, la derivada de $w(t)=f[z(t)]$ puede expresarse en la forma

$$w'(t_0) = \frac{\partial f}{\partial x} x'(t_0) + \frac{\partial f}{\partial y} y'(t_0),$$

donde se han tomado las derivadas parciales en z_0. En términos de $z'(t_0)$, esto puede escribirse nuevamente en la forma

$$w'(t_0) = \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right) z'(t_0) + \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right)\overline{z'(t_0)}$$

Si se conservan los ángulos, $\arg\,[w'(t_0)/z'(t_0)]$ tiene que ser independiente de $\arg z'(t_0)$. La expresión

$$\frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right) + \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right)\frac{\overline{z'(t_0)}}{z'(t_0)} \qquad [2\text{-}27]$$

ha de tener, por tanto, argumento constante. Al variar $\arg z'(t_0)$, el punto representado por [2-27] describe una circunferencia de radio $\frac{1}{2}|(\partial f/\partial x) + i(\partial f/\partial y)|$. El argumento no puede ser constante en esta circunferencia a no ser que se anule el radio; por consiguiente, debemos tener

$$\frac{\partial f}{\partial x} = -i\frac{\partial f}{\partial y} \qquad [2\text{-}28]$$

que es la forma compleja de las ecuaciones de Cauchy-Riemann.

De manera análoga, la condición de que el cambio de escala sea el mismo en todas las direcciones implica que la expresión [2-27] tiene módulo constante. En una circunferencia, el módulo es constante únicamente si el radio es cero o si el centro está en el origen. En el primer caso obtenemos [2-28], y en el segundo,

$$\frac{\partial f}{\partial x} = i\frac{\partial f}{\partial y}$$

La última ecuación expresa el hecho de ser $\overline{f(z)}$ analítica. Una aplicación mediante la conjugada de una función analítica con derivada no nula se dice *indirectamente conforme.* Evidentemente, preserva la magnitud de los ángulos, pero invierte su sentido.

Si la aplicación de Ω mediante $w=f(z)$ es topológica, entonces la función inversa $z=f^{-1}(w)$ es también analítica. Esto se deduce fácilmente si $f'(z) \neq 0$, pues en este caso la derivada de la función inversa debe ser igual a $1/f'(z)$ en el punto $z=f^{-1}(w)$. Demostraremos más adelante que $f'(z)$ no puede anularse nunca en el caso de una aplicación topológica realizada mediante una función analítica.

El conocimiento de que $f'(z_0) \neq 0$ es suficiente para llegar a la conclusión de que la aplicación es topológica si se restringe a un entorno de z_0 suficientemente pequeño. Esto se deduce del teorema de las funciones implícitas ya conocido, pues el jacobiano de las funciones $u=u(x,y)$, $v=v(x,y)$ en el punto z_0 es $|f'(z_0)|^2$ y, por consiguiente, distinto de cero. Posteriormente presentaremos una demostración más sencilla de este importante teorema.

Pero aun cuando sea $f'(z) \neq 0$ en todos los puntos de la región Ω, no se puede afirmar que la aplicación de toda la región es necesariamente topológica. La figura 2-1 indica lo que pudiera ocurrir. Se ve que las aplicaciones de las subregiones Ω_1 y Ω_2 son biunívocas, pero las imágenes se recubren. Es bastante sugeridor pensar en la imagen de toda la región como una película transparente que se recubre parcialmente a sí misma. Esta es la sencilla, a la vez que fecunda, idea utilizada por Riemann cuando introdujo las regiones generalizadas conocidas actualmente como *superficies de Riemann.*

FIG. 2-1.—Región doblemente cubierta.

2-4. Aplicaciones conformes elementales.—La aplicación conforme asociada a una función analítica da una excelente visualización de las propiedades de la última; pudiera compararse a la representación de una función real mediante su gráfica. Por tanto, es natural que todas las cuestiones relativas a aplicaciones conformes hayan recibido una importante atención; los progresos en esta dirección han incrementado considerablemente nuestros conocimientos sobre las funciones analíticas. Además, las aplicaciones conformes aparecen de una manera natural en muchas ramas de la física matemática, dando así lugar a una utilización inmediata de la teoría de funciones de variable compleja.

Uno de los problemas más importantes es la determinación de las aplicaciones conformes de una región en otra. En esta sección consideraremos aquellas aplicaciones que pueden definirse mediante funciones elementales.

1. *Utilización de curvas de nivel.*—Cuando se define una aplicación conforme mediante una función analítica explícita $w=f(z)$, se desea, naturalmente, obtener información sobre las propiedades geométricas específicas de la aplicación. Uno de los métodos más fecundos consiste en el estudio de la correspondencia de curvas inducida por la transformación puntual. Las propiedades especiales de la función $f(z)$ pueden, p. ej., expresarse mediante el hecho de que ciertas curvas sencillas se transforman en curvas de una familia de carácter bien conocido. Toda información de este tipo reforzará nuestra concepción visual de la aplicación.

Tal era el caso para las aplicaciones mediante transformaciones lineales. Demostrábamos en el capítulo I, sección 1-3, que una transformación lineal aplica circunferencias en circunferencias, con tal que se incluya a las rectas como un caso especial de aquellas. Considerando las circunferencias de Steiner, nos era posible obtener una imagen completa de la correspondencia. El hecho de que la aplicación era conforme fue demostrado de manera geométrica; analíticamente se deduce de que

$$D\,\frac{\alpha z+\beta}{\gamma z+\delta}=\frac{\alpha\delta-\beta\gamma}{(\gamma z+\delta)^2}$$

nunca es cero. Los puntos $z=-\delta/\gamma$ y $z=\infty$ se exceptúan, naturalmente, en el tratamiento analítico; pero el que la representación sea conforme aparece patentemente al interpretarla sobre la esfera de Riemann.

También se señaló que una transformación lineal aplica una región circular (el interior de un círculo, el exterior de un círculo o un semiplano) sobre otra. Desde el punto de vista de la aplicación conforme, se tratan todas las regiones circulares como si fueran equivalentes; se puede pasar de una a otra mediante una transformación lineal.

En casos más generales es aconsejable empezar con un estudio de las curvas imágenes de las rectas $x=x_0$ e $y=y_0$. Si ponemos $f(z)=u(x,y)+iv(x,y)$, la imagen de $x=x_0$ viene dada por las ecuaciones paramétricas $u=u(x_0,y)$, $v=v(x_0,y)$, actuando y como parámetro, que puede ser eliminado o retenido según las conveniencias. La imagen de $y=y_0$ se determina de manera análoga. En conjunto, las curvas forman una red ortogonal en el plano w. Análogamente, podríamos considerar las curvas $u(x,y)=u_0$ y $v(x,y)=v_0$ en el plano z. También son ortogonales, y se llaman *curvas de nivel* de u y v.

En otros casos pudiera ser más conveniente utilizar coordenadas polares y estudiar las imágenes de circunferencias concéntricas y de rectas que pasen por el origen.

Entre las aplicaciones más importantes están las obtenidas mediante potencias de z, o sea $w=z^\alpha$. Consideraremos únicamente el caso de ser α real, pudiendo suponerse también que α es positivo. Puesto que

$$|w|=|z|^\alpha$$
$$\arg w=\alpha\arg z,$$

circunferencias con centro en el origen se transforman en circunferencias de la misma familia, y a semirrectas desde el origen corresponden semirrectas análogas. La aplicación es conforme en todos los puntos $z \neq 0$, pero un ángulo θ en el origen se transforma en un ángulo $\alpha\theta$. Para $\alpha \neq 1$ la aplicación de todo el plano no es biunívoca, y si α es fraccionario, z^α no es ni siquiera uniforme. En general, por tanto, únicamente podremos considerar la aplicación de un sector angular sobre otro.

Como en el capítulo I, sección 1-2, 4, el sector $S_0(\varphi_1, \varphi_2)$, donde $0 < \varphi_2 - \varphi_1 \leqq 2\pi$, está formado por todos los puntos $z \neq 0$, tales que un valor de $\arg z$ satisface la desigualdad

$$\varphi_1 < \arg z < \varphi_2. \tag{2-29}$$

Es fácil probar que $S_0(\varphi_1, \varphi_2)$ es una región. En esta región se define un único valor de $w = z^\alpha$ mediante la condición

$$\arg w = \alpha \arg z,$$

donde $\arg z$ denota el valor del argumento indicado por la condición [2-29]. Esta función es analítica con derivada que no se anula:

$$De^{\alpha \log z} = \alpha \frac{w}{z}$$

La aplicación es biunívoca solo si $\alpha(\varphi_2 - \varphi_1) \leqq 2\pi$, y en este caso $S_0(\varphi_1, \varphi_2)$ se aplica sobre el sector $S_0(\alpha\varphi_1, \alpha\varphi_2)$ en el plano w. Debería observarse que $S_0(\varphi_1 + n \cdot 2\pi, \varphi_2 + n \cdot 2\pi)$ es geométricamente idéntico a $S_0(\varphi_1, \varphi_2)$, pero pudiera determinar una rama diferente de z^α.

Consideremos la aplicación $w = z^2$ más detalladamente. Puesto que $u = x^2 - y^2$ y $v = 2xy$, vemos que las curvas de nivel $u = u_0$ y $v = v_0$ son hipérbolas equiláteras, cuyas respectivas asíntotas son las bisectrices de los ángulos que forman los ejes coordenados, y dichos ejes. Naturalmente, son ortogonales entre sí. Por otra parte, la imagen de $x = x_0$ es $v^2 = 4x_0^2(x_0^2 - u)$, y la de $y = y_0$ es $v^2 = 4y_0^2(y_0^2 + u)$. Ambas familias representan parábolas con el foco en el origen, cuyos ejes están dirigidos en las direcciones negativa y positiva, respectivamente, del eje de las u. Su ortogonalidad es bien conocida en geometría analítica. Las familias de curvas de nivel se muestran en las figuras 2-2 y 2-3.

Como una familia diferente de curvas imagen consideremos las circunferencias $|w - 1| = k$ en el plano w. La ecuación de la imagen inversa puede escribirse en la forma

$$(x^2 + y^2)^2 = 2(x^2 - y^2) + k^2 - 1,$$

y representa una familia de lemniscatas con puntos focales ± 1. La familia ortogonal está representada por

$$x^2 - y^2 = 2hxy + 1,$$

y consiste en todas las hipérbolas equiláteras con centro en el origen, que pasan por los puntos ± 1.

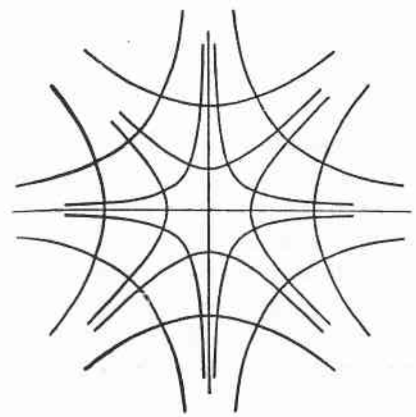

FIG. 2-2.—Plano z.

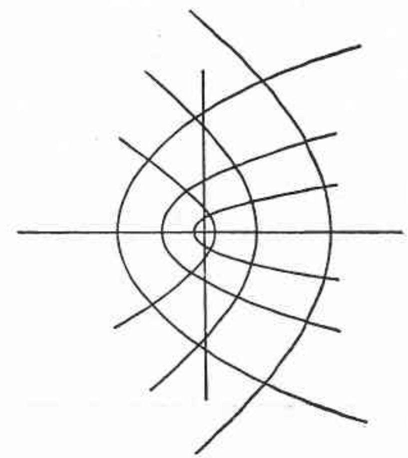

FIG. 2-3.—Plano w.

En el caso del cubo $w = z^3$, las curvas de nivel en ambos planos son curvas cúbicas. No interesa hallar sus ecuaciones, ya que su forma general se puede ver claramente sin necesidad de cálculo. Así, p. ej., las curvas $u = u_0 > 0$ han de tener la forma indicada en la figura 2-4. Análogamente, si seguimos el cambio de $\arg w$ cuando z recorre la curva $x = x_0 > 0$, encontramos que la curva imagen tendrá un lazo (Fig. 2-5). Es, por tanto, un *folium* de Descartes.

La aplicación obtenida por $w = e^z$ es muy sencilla. Las rectas $x = x_0$ e $y = y_0$ se aplican sobre circunferencias con centro en el origen y semirrectas de argumento constante. Cualquier otra recta del plano de las z se aplica sobre una espiral logarítmica. La aplicación es biunívoca en cualquier región que no contenga dos puntos cuya diferencia sea un múltiplo de $2\pi i$. En particular, una banda horizontal $y_1 < y < y_2$, $y_2 - y_1 \leqq 2\pi$, se aplica sobre un sector angular, y si $y_2 - y_1 = \pi$, la imagen es un semiplano. Podemos así aplicar una banda paralela sobre un semiplano y, por consiguiente, sobre una región circular cualquiera. La mitad izquierda de la banda, determinada por el eje imaginario, corresponde a un semicírculo.

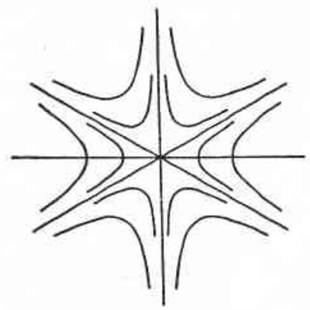

FIG. 2-4.

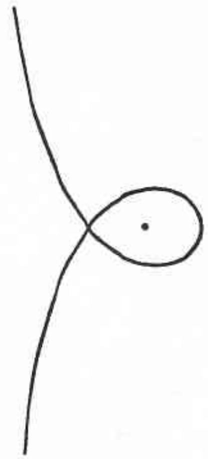

FIG. 2-5.

Convendrá deducir algunas fórmulas explícitas de la aplicación. La función $\zeta = \xi + i\eta = e^z$ aplica la banda $-\pi/2 < y < \pi/2$ sobre el semiplano $\xi > 0$. Por otra parte,

$$w = \frac{\zeta - 1}{\zeta + 1}$$

aplica $\xi > 0$ sobre $|w| < 1$. Por consiguiente,

$$w = \frac{e^z - 1}{e^z + 1} = \operatorname{th} \frac{z}{2}$$

2. *Estudio rápido de aplicaciones elementales.*—Al tratar del problema de aplicar de manera conforme una región Ω_1 sobre otra Ω_2, es generalmente aconsejable proceder en dos etapas: primera, aplicamos Ω_1 sobre una región circular y a continuación aplicamos esta región circular sobre Ω_2. En otras palabras: el problema general de la aplicación conforme puede reducirse al problema de aplicar una región sobre un disco o sobre un semiplano. En el capítulo IV demostraremos que este problema de aplicación tiene una solución para toda región cuya frontera sea una curva cerrada simple.

Las herramientas principales a nuestra disposición son las transformaciones lineales y las transformaciones mediante potencias de z, mediante la función exponencial o mediante la función logarítmica.

Todas estas transformaciones tienen la propiedad característica de que aplican una familia de rectas o de circunferencias en una familia análoga. Por esta razón, su utilización está limitada esencialmente a regiones cuya frontera está constituida por arcos circulares y segmentos de recta. La función potencial sirve para convertir ángulos cualesquiera en rectos, y con la ayuda de la función exponencial es posible incluso transformar ángulos nulos en rectos.

Por estos procedimientos hallaremos, en primer lugar, una aplicación tipo de cualquier región, cuya frontera consiste en dos arcos circulares con extremos comunes. Tal región o es una cuña circular, cuyo ángulo pudiera ser mayor que π, o su complemento. Si los extremos de los arcos son a y b, empezamos con la aplicación previa $z_1 = (z-a)/(z-b)$, que transforma la región dada en un sector angular. Mediante una potencia adecuada $w = z_1^\alpha$ este sector se puede aplicar sobre un semiplano.

Si las circunferencias son tangentes entre sí en el punto a, la transformación $z_1 = 1/(z-a)$ aplicará la región comprendida entre ellas sobre una banda paralela, y una transformación exponencial apropiada aplicará la banda sobre un semiplano.

Con un poco más de generalidad, el mismo método también es de aplicación a un triángulo circular con dos ángulos rectos. En efecto, si el tercer ángulo tiene el vértice en a y si los lados que parten de a se cortan luego en b, la transformación lineal $z_1 = (z-a)/(z-b)$ aplica el triángulo dado sobre un sector circular. Por medio de una potencia este sector puede transformarse en un semicírculo; este es una región en forma de cuña que a su vez puede aplicarse sobre un semiplano.

En relación con todo lo anterior trataremos explícitamente un caso especial que aparece con frecuencia. Supongamos que se pide aplicar el complemento de un segmento sobre el interior o el exterior de un círculo. La región es una cuña con ángulo 2π; sin pérdida de generalidad podemos suponer que los extremos del segmento son ± 1. La transformación previa

$$z_1 = \frac{z+1}{z-1}$$

aplica la cuña en el ángulo total, obtenido por exclusión del eje real negativo. A continuación definimos

$$z_2 = \sqrt{z_1}$$

como la raíz cuadrada cuya parte real es positiva, y obtenemos una aplicación sobre el semiplano derecho. La transformación final

$$w = \frac{z_2 - 1}{z_2 + 1}$$

aplica el semiplano sobre $|w| < 1$.

La eliminación de las variables intermedias nos conduce a la correspondencia

$$z = \frac{1}{2}\left(w + \frac{1}{w}\right) \qquad w = z - \sqrt{z^2 - 1}. \qquad [2\text{-}30]$$

El signo de la raíz cuadrada está determinado de manera única por la condición $|w| < 1$, pues $(z - \sqrt{z^2-1})(z + \sqrt{z^2-1}) = 1$. Si se cambia el signo, obtenemos una aplicación sobre $|w| > 1$.

Para efectuar un estudio más detallado de la aplicación [2-30] ponemos $w = \rho e^{i\theta}$, y obtenemos

$$x = \frac{1}{2}\left(\rho + \frac{1}{\rho}\right)\cos\theta$$

$$y = \frac{1}{2}\left(\rho - \frac{1}{\rho}\right)\operatorname{sen}\theta$$

La eliminación de θ nos da

$$\frac{x^2}{[\frac{1}{2}(\rho + \rho^{-1})]^2} + \frac{y^2}{[\frac{1}{2}(\rho - \rho^{-1})]^2} = 1, \qquad [2\text{-}31]$$

y la de ρ,

$$\frac{x^2}{\cos^2\theta} - \frac{y^2}{\operatorname{sen}^2\theta} = 1 \qquad [2\text{-}32]$$

Por consiguiente, la imagen de una circunferencia $|w| = \rho < 1$ es una elipse con el eje mayor $\rho + \rho^{-1}$ y el menor $\rho^{-1} - \rho$. La imagen de un radio es media rama de una hipérbola. Las elipses [2-31] y las hipérbolas [2-32] tienen los mismos focos. La figura 2-6 ilustra la correspondencia.

Evidentemente, la transformación [2-30] nos permite incluir en nuestra lista de aplicaciones conformes elementales la aplicación del exterior de una elipse o la región entre las ramas de una hipérbola sobre una región circular. Sin embargo, no nos permite aplicar el interior de una elipse o el interior de una rama hiperbólica.

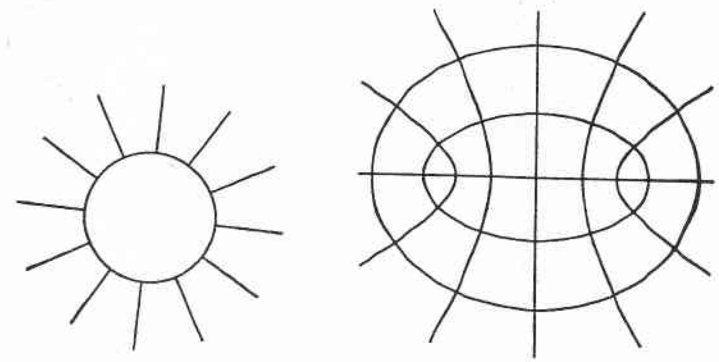

FIG. 2-6.

Como ejemplo final y menos trivial estudiaremos la aplicación definida por un polinomio cúbico $w = a_0z^3 + a_1z^2 + a_2z + a_3$. La conocida transformación $z = z_1 - a_1/3a_0$ nos permite eliminar el término cuadrático, y mediante normalizaciones obvias podemos reducir el polinomio a la forma $w = z^3 - 3z$. Se elige el coeficiente de z de forma que la derivada para $z = \pm 1$ se anule.

Utilizando la transformación [2-30] introducimos una variable auxiliar ζ definida por

$$z = \zeta + \frac{1}{\zeta} \qquad [2\text{-}33]$$

Nuestro polinomio cúbico toma entonces la forma más sencilla

$$w = \zeta^3 + \frac{1}{\zeta^3} \qquad [2\text{-}34]$$

Observemos que cada valor de z determina dos valores ζ, pero ambos son recíprocos y dan el mismo valor de w. Con objeto de obtener un único ζ podríamos imponer la condición $|\zeta| < 1$, pero entonces el segmento $(-2, 2)$ debería ser excluido del plano de las z.

Es fácil visualizar ahora la correspondencia entre los planos z y w. A la circunferencia $|\zeta| = \rho < 1$ corresponde una elipse con semiejes $\rho^{-1} \pm \rho$ en el plano z y una con los semiejes $\rho^{-3} \pm \rho^3$ en el plano w. Análogamente, a ramas hiperbólicas en los planos z y w corresponde una semirrecta $\arg \zeta = \theta$; la del plano z tiene una asíntota que forma ángulo $-\theta$ con el eje real positivo, y en el plano w el ángulo correspondiente es -3θ. El esquema total de elipses e hipérbolas homofocales permanece invariante, pero cuando z describe una elipse, w recorre tres veces la correspondiente elipse mayor. Así, pues, la situación es similar a la del caso de la aplicación más sencilla $w = z^3$. Como orientación, el lector puede utilizar la figura 2-6, excepto que ahora los focos están en ± 2.

Para la región comprendida entre dos ramas hiperbólicas cuyas asíntotas forman ángulo $\leqq 2\pi/3$ la aplicación es biunívoca. Obsérvese en particular que las seis regiones en que la hipérbola $3x^2 - y^2 = 3$ y el eje de las x dividen al plano de las z se aplican sobre semiplanos, tres de ellas sobre el semiplano superior y las otras tres sobre el inferior. El interior de la rama derecha de la hipérbola corresponde a todo el plano w, con una incisión a lo largo del eje real negativo hasta el punto -2.

EJERCICIOS

Todas las aplicaciones habrán de ser conformes.

1. Aplíquese la parte común de los discos $|z| < 1$ y $|z-1| < 1$ en el interior del círculo unidad. Elíjase la aplicación de tal forma que se conserven las dos simetrías.

2. Aplíquese la región entre $|z| = 1$ y $|z - \frac{1}{2}| = \frac{1}{2}$ sobre un semiplano.

3. Aplíquese el complemento del arc $|z| = 1$, $y \geqq 0$ en el exterior del círculo unidad, de tal forma que se correspondan los puntos del infinito.

4. Aplíquese el exterior de la parábola $y^2 = 2px$ sobre el disco $|w| < 1$, de tal forma que $z = 0$ y $z = -p/2$ correspondan a $w = 1$ y $w = 0$ (Lindelöf).

5. Aplíquese el interior de la rama derecha de la hipérbola $x^2 - y^2 = a^2$ en el disco $|w| < 1$, de forma que el foco corresponda a $w = 0$ y el vértice a $w = -1$ (Lindelöf).

6. Aplíquese el interior de la lemniscata $|z^2 - a^2| = \rho^2$ $(\rho > a)$ sobre el disco $|w| < 1$, de forma que se conserven las simetrías (Lindelöf).

7. Aplíquese el exterior de la elipse $(x/a)^2 + (y/b)^2 = 1$ sobre $|w| < 1$, con conservación de simetrías.

8. Aplíquese la parte del plano z a la izquierda de la rama derecha de la hipérbola $x^2 - y^2 = 1$ en un semiplano (Lindelöf).

Sugerencia: Considérese por un lado la aplicación del semiplano superior de la región mediante $w = z^2$, y por otro, la aplicación de un cuadrante por medio de $w = z^3 - 3z$.

3. *Superficies elementales de Riemann.*—La visualización de una función mediante la aplicación correspondiente es completamente clara solo en el caso de que la aplicación sea biunívoca. Si este no es el caso, todavía podremos ayudar a nuestra imaginación mediante la introducción de regiones generalizadas, en las que puntos distintos pueden tener las mismas coordenadas. Para esto es necesario suponer que puntos que ocupan el mismo lugar pueden distinguirse por otras características, p. ej., una señal o un color. Puntos con la misma señal se consideran como pertenecientes a la misma *hoja* o *lámina*.

Esta idea nos conduce al concepto de una *superficie de Riemann*. En relación con este concepto, no es nuestro propósito dar una definición rigurosa del mismo. Es suficiente para nuestros propósitos introducir las superficies de Riemann de forma puramente descriptiva. Podemos hacerlo con toda libertad siempre y cuando utilicemos el concepto para propósitos meramente ilustrativos y nunca en demostraciones lógicas.

La superficie de Riemann más sencilla está relacionada con la aplicación que proporciona la función $w = z^n$, donde $n > 1$ es un entero. Sabemos que existe una correspondencia biunívoca entre cada ángulo $(k-1)(2\pi/n) < \arg z < k(2\pi/n)$, $k = 1, \ldots, n$ y todo el plano w excepto el eje real positivo. La imagen de cada ángulo se obtiene así efectuando un "corte" a lo largo del eje positivo; este corte tiene unos "bordes" superior e inferior. En correspondencia con los n ángulos en el plano z consideramos n copias idénticas del plano w con el corte. Estas serán las "hojas" de la superficie de Riemann, y se distinguirán unas de otras mediante la señal k, que sirve para identificar el ángulo correspondiente. Cuando z se mueve en su plano, el correspondiente punto w ha de estar en libertad para moverse sobre la superficie de Riemann. Por esta razón debemos unir el borde inferior de la primera hoja con el superior de la segunda, el borde inferior de la segunda hoja con el superior de la tercera, y así sucesivamente. En el último paso, el borde inferior de la hoja n-ésima se une al borde superior de la primera, completando el ciclo. En un sentido físico, esto no es posible sin cruzamiento; pero el modelo ideal estará libre de esta discrepancia. El resultado de la construcción es una superficie de Riemann cuyos puntos están en correspondencia biunívoca con los puntos del plano z. Es más, esta correspondencia es continua si se define la continuidad en el sentido sugerido por la construcción.

El corte a lo largo del eje positivo puede reemplazarse por otro a lo largo de cualquier arco simple desde 0 a ∞; la superficie de Riemann obtenida de esta forma deberá considerarse idéntica a la originalmente construida. En otras palabras: los cortes no son de ninguna manera curvas distinguidas de la superficie; pero por motivos descriptivos es necesaria la introducción de cortes específicos.

El punto $w=0$ está en una posición especial. Enlaza todas las hojas, y una curva debe dar n vueltas alrededor del origen antes de cerrarse. Un punto de esta clase se llama *punto de ramificación*. Si consideramos nuestra superficie de Riemann sobre todo el plano ampliado, el punto del ∞ es también un punto de ramificación. En casos más generales un punto de ramificación no enlaza necesariamente todas las hojas; si enlaza h hojas, se dice que es de orden $h-1$.

La superficie de Riemann correspondiente a $w=e^z$ es de naturaleza similar. En este caso, la función aplica cada banda paralela $(k-1)2\pi < y < k\cdot 2\pi$ sobre una hoja con un corte a lo largo del eje positivo. Las hojas están unidas unas a otras de manera que forman una especie de tornillo sin fin. Como consecuencia del hecho de que e^z nunca es cero, el origen *no* es un punto de la superficie de Riemann.

El lector hallará fácil la construcción de otras superficies de Riemann. Ilustraremos el procedimiento considerando la correspondiente a $w=\cos z$. Una región que se aplica biunívocamente sobre todo el plano, excepto por uno o más cortes, se llama *región fundamental*. Como regiones fundamentales de $w=\cos z$ podemos elegir las bandas $(k-1)\pi < x < k\pi$. Cada banda se aplica sobre todo el plano w, con cortes a lo largo del eje real desde $-\infty$ a -1 y de 1 a ∞. La recta $x=k\pi$ corresponde a ambos bordes del corte positivo si k es par, y a los dos bordes del corte negativo si k es impar. Si consideramos las dos bandas adyacentes a lo largo de la recta $x=k\pi$, tenemos que los bordes de los cortes correspondientes tienen que unirse en cruz, de forma que engendren un único punto de ramificación en $w=\pm 1$. La superficie resultante tiene infinitos puntos de ramificación sobre $w=1$ y $w=-1$, que enlazan alternativamente las hojas impares y las pares.

La figura 2-7 muestra un intento para ilustrar el enlace de las hojas. Representa una sección transversal de la superficie en el caso de que los cortes se hayan elegido paralelos entre sí. El lector deberá tener en cuenta que cualquier par de puntos en el mismo nivel puede unirse por un arco, que no corta a ninguno de los cortes.

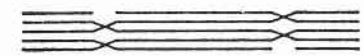

FIG. 2-7.—La superficie de Riemann de $\cos z$.

Cualquiera que sea la ventaja de tales representaciones, quizá la imagen más clara de la superficie de Riemann se obtiene por consideración directa de las regiones fundamentales en el plano z. La interpretación es todavía más sencilla si, como en la figura 2-8, introducimos las subregiones que corresponden a los semiplanos superior e inferior. Las regiones sombreadas son aquellas en que $\cos z$ tiene parte imaginaria positiva. Cada región corresponde a un semiplano, en el que marcamos los puntos frontera 1 y -1. Para dos regiones adyacentes cualesquiera, una sin sombrear y la otra sombreada, los semiplanos deben unirse a través de uno de los tres intervalos $(-\infty, -1)$, $(-1, 1)$ o $(1, \infty)$. La elección de la unión correcta es automática después de echar un vistazo a la situación correspondiente en el plano z.

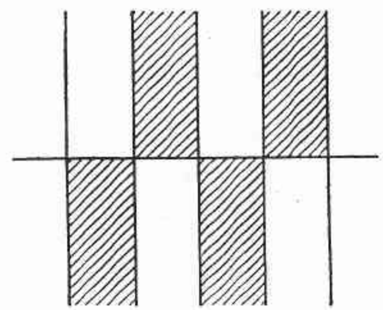

FIG. 2-8.—Regiones fundamentales de $\cos z$.

EJERCICIOS

1. Descríbase la superficie de Riemann que resulta asociada con la función $w=\frac{1}{2}\left(z+\frac{1}{z}\right)$.

2. El mismo problema para $w=(z^2-1)^2$.

3. El mismo problema para $w=z^3-3z$.

4. Descríbase la naturaleza general de la superficie de Riemann asociada a la función e^{e^z}.

5. Pruébese que la ecuación $w^3+z^3-wz=0$ define una correspondencia biunívoca entre dos superficies de Riemann de tres hojas.

CAPITULO III

INTEGRACION EN EL CAMPO COMPLEJO

3-1. Teoremas fundamentales.—Son muchas las propiedades importantes de las funciones analíticas que no pueden demostrarse sin la ayuda de la integración en el campo complejo. Nadie ha conseguido demostrar que la derivada de una función analítica es continua sin utilizar integrales complejas o métodos equivalentes. Aun cuando la continuidad de la derivada forme parte de la definición, no ha sido posible probar la existencia de derivadas de órdenes superiores sin hacer uso de la integración. Este fracaso en el intento de eliminar la integral, en cuestiones que superficialmente conciernen únicamente al cálculo diferencial, señala una diferencia profunda entre las variables real y compleja.

Como en el caso real, distinguimos entre *integrales definidas* e *indefinidas*. Una integral indefinida es una función cuya derivada es igual a una función analítica dada en una región; en muchos casos elementales se pueden hallar integrales indefinidas por inversión de fórmulas de derivación bien conocidas. Las integrales definidas se toman sobre arcos diferenciables o diferenciables a trozos, y no están limitadas a funciones analíticas. Pueden definirse mediante un proceso de paso al límite, que remeda la definición de una integral definida real. En realidad, es preferible la definición de integrales definidas complejas en términos de integrales reales. Esto nos ahorrará la repetición de las demostraciones de existencia, que son esencialmente las mismas que en el caso real. Naturalmente, el lector deberá estar bien familiarizado con la teoría de integrales definidas de funciones continuas reales.

1. *Integrales curvilíneas.*—La generalización más inmediata de una integral real es la integral definida de una función compleja sobre un intervalo real. Si $f(t)=u(t)+iv(t)$ es una función continua, definida en un intervalo (a, b), ponemos por definición

$$\int_a^b f(t)\,dt=\int_a^b u(t)\,dt+i\int_a^b v(t)\,dt. \qquad [3\text{-}1]$$

Esta integral posee la mayoría de las propiedades de la integral real. En particular, si $c=\alpha+i\beta$ es una constante compleja, obtenemos

$$\int_a^b cf(t)\,dt=c\int_a^b f(t)\,dt, \qquad [3\text{-}2]$$

pues ambos miembros son iguales a

$$\int_a^b (\alpha u-\beta v)\,dt+i\int_a^b (\alpha v+\beta u)\,dt.$$

Cuando $a\leqq b$, se verifica la desigualdad fundamental

$$\left|\int_a^b f(t)\,dt\right|\leqq\int_a^b |f(t)|\,dt \qquad [3\text{-}3]$$

para $f(t)$ compleja cualquiera. Para comprobarlo elegimos $c=e^{-i\theta}$ con θ real en [3-2], y tenemos que

$$\operatorname{Re}\left[e^{-i\theta}\int_a^b f(t)\,dt\right]=\int_a^b \operatorname{Re}\,[e^{-i\theta}f(t)]\,dt\leqq\int_a^b |f(t)|\,dt.$$

Para $\theta=\arg\int_a^b f(t)\,dt$, la expresión de la izquierda se reduce al valor absoluto de la integral, y obtenemos [3-3] [1].

Consideremos ahora un arco γ diferenciable a trozos, de ecuación $z=z(t)$, $a\leqq t\leqq b$. Si la función $f(z)$ está definida y es continua sobre γ, entonces $f[z(t)]$ es también continua, y podemos poner

$$\int_\gamma f(z)\,dz=\int_a^b f[z(t)]z'(t)\,dt. \qquad [3\text{-}4]$$

Esta es nuestra *definición* de integral curvilínea compleja de $f(z)$ tomada sobre el arco γ. En el segundo miembro de [3-4], si $z'(t)$ no es continua por doquier, hay que subdividir el intervalo de integración de manera obvia. Siempre que se considere una integral curvilínea sobre un arco γ se supondrá tácitamente que γ es diferenciable a trozos.

---

[1] θ no está definida si $\int_a^b f\,dt=0$, pero entonces no hay nada que demostrar.

La propiedad más importante de la integral [3-4] es su invariancia respecto a un cambio de parámetro, el cual viene determinado por una función creciente $t=t(\tau)$, que aplica un intervalo $\alpha \leqq \tau \leqq \beta$ sobre $a \leqq t \leqq b$; suponemos que $t(\tau)$ es diferenciable a trozos. Mediante la regla correspondiente al cambio de variable de integración, tenemos

$$\int_a^b f[z(t)]z'(t)\,dt = \int_\alpha^\beta f\{z[t(\tau)]\}z'[t(\tau)]t'(\tau)\,d\tau.$$

Pero $z'[t(\tau)]t'(\tau)$ es la derivada de $z[t(\tau)]$ con respecto a τ y, por consiguiente, la integral [3-4] tiene el mismo valor tanto si γ está representado por la ecuación $z=z(t)$ como si lo está por la $z=z[t(\tau)]$.

En el capítulo II, sección 2-2, 5, se definió el arco opuesto $-\gamma$ mediante la ecuación $z=z(-t)$, $-b \leqq t \leqq -a$. Tenemos así

$$\int_{-\gamma} f(z)\,dz = \int_{-b}^{-a} f[z(-t)][-z'(-t)]\,dt,$$

y por un cambio de variable, la integral última puede ponerse en la forma

$$\int_b^a f[z(t)]z'(t)\,dt.$$

En conclusión,

$$\int_{-\gamma} f(z)\,dz = -\int_\gamma f(z)\,dz. \qquad [3\text{-}5]$$

La integral [3-4] posee también una propiedad aditiva completamente obvia. Está bien claro el significado que tiene la subdivisión de un arco γ en un número finito de subarcos. Puede indicarse una subdivisión mediante la ecuación simbólica

$$\gamma = \gamma_1 + \gamma_2 + \ldots + \gamma_n,$$

satisfaciendo las integrales correspondientes la relación

$$\int_{\gamma_1+\gamma_2+\ldots+\gamma_n} f\,dz = \int_{\gamma_1} f\,dz + \int_{\gamma_2} f\,dz + \ldots + \int_{\gamma_n} f\,dz. \qquad [3\text{-}6]$$

Por último, la integral sobre una curva cerrada es también invariante respecto a un desplazamiento del parámetro. Los puntos iniciales original y final determinan dos subarcos γ_1, γ_2, y la invariancia se sigue del hecho de que la integral sobre $\gamma_1+\gamma_2$ es igual a la integral sobre $\gamma_2+\gamma_1$.

Además de integrales de la forma [3-4], podemos también considerar integrales curvilíneas con respecto a $\bar{z}$. La definición más conveniente es mediante doble conjugación:

$$\int_\gamma f\,\overline{dz} = \overline{\int_\gamma \bar{f}\,dz}.$$

Utilizando esta notación se pueden introducir integrales curvilíneas con respecto a x o a y mediante las expresiones

$$\int_\gamma f\,dx = \frac{1}{2}\left(\int_\gamma f\,dz + \int_\gamma f\,\overline{dz}\right)$$

$$\int_\gamma f\,dy = \frac{1}{2i}\left(\int_\gamma f\,dz - \int_\gamma f\,\overline{dz}\right)$$

Con $f=u+iv$ tenemos que la integral [3-4] puede escribirse en la forma

$$\int_\gamma (u\,dx - v\,dy) + i\int_\gamma (u\,dy + v\,dx), \qquad [3\text{-}7]$$

que separa las partes real e imaginaria.

Una integral curvilínea esencialmente diferente se obtiene por integración con respecto a la *longitud del arco.* Hay dos notaciones en uso, siendo la definición

$$\int_\gamma f\,ds = \int_\gamma f|dz| = \int_\gamma f[z(t)]|z'(t)|\,dt. \qquad [3\text{-}8]$$

Esta integral es también independiente de la elección del parámetro. En contraste con [3-5], tenemos ahora

$$\int_{-\gamma} f|dz| = \int_\gamma f|dz|,$$

mientras que [3-6] sigue siendo válida en la misma forma. La desigualdad

$$\left|\int_\gamma f\,dz\right| \leqq \int_\gamma |f|\cdot|dz|. \qquad [3\text{-}9]$$

es una consecuencia de [3-3].

Para $f=1$, la integral [3-8] se reduce $\int_\gamma |dz|$, que es por definición la *longitud* de γ. Como ejemplo calcularemos la longitud de una circunferencia. De la ecuación paramétrica $z=z(t)=a+\rho e^{it}$, $0\leqq t\leqq 2\pi$, de una circunferencia completa obtenemos $z'(t)=i\rho e^{it}$ y, por tanto,

$$\int_0^{2\pi}|z'(t)|\,dt=\int_0^{2\pi}\rho\,dt=2\pi\rho$$

como era de esperar.

Integrales curvilíneas generales de la forma $\int_\gamma p\,dx+q\,dy$ se estudian a menudo como funciones (o *funcionales*) del arco γ. Se supone entonces que p y q están definidas y son continuas en una región Ω y que γ puede variar en Ω. Una clase importante de integrales está caracterizada por la propiedad de que la integral sobre un arco depende únicamente de sus extremos. En otras palabras: si γ_1 y γ_2 tienen los mismos orígenes y extremos, exigimos que $\int_{\gamma_1} p\,dx+q\,dy=$ $=\int_{\gamma_2} p\,dx+q\,dy$. Decir que una integral depende solo de los extremos equivale a decir que la integral sobre cualquier curva cerrada es cero. En efecto, si γ es una curva cerrada, entonces γ y $-\gamma$ tienen los mismos extremos, y si la integral depende únicamente de estos, se obtiene

$$\int_\gamma=\int_{-\gamma}=-\int_\gamma$$

y, por tanto, $\int_\gamma=0$. Recíprocamente, si γ_1 y γ_2 tienen los mismos extremos, entonces $\gamma_1-\gamma_2$ es una curva cerrada, y si la integral sobre cualquier curva cerrada se anula, se sigue que $\int_{\gamma_1}=\int_{\gamma_2}$.

El teorema siguiente da una condición necesaria y suficiente bajo la cual una integral curvilínea depende únicamente de los extremos:

TEOREMA 1. *La integral curvilínea $\int_\gamma p\,dx+q\,dy$, definida en Ω, depende solo de los extremos de γ si (y solo si) existe una función $U(x,y)$ en Ω con derivadas parciales $\partial U/\partial x=p$, $\partial U/\partial y=q$.*

La suficiencia es inmediata, pues si se verifica la condición podemos escribir, con las notaciones habituales,

$$\int_\gamma p\,dx+q\,dy=\int_a^b\left(\frac{\partial U}{\partial x}x'(t)+\frac{\partial U}{\partial y}y'(t)\right)dt=$$
$$=\int_a^b\frac{d}{dt}U[x(t),y(t)]\,dt=U[x(b),y(b)]-U[x(a),y(a)],$$

y el valor de esta diferencia depende únicamente de los puntos extremos. A fin de demostrar la necesidad elegimos un punto fijo $(x_0,y_0)\in\Omega$, lo unimos al (x,y) por una línea poligonal γ, contenida en Ω, cuyos lados sean paralelos a los ejes coordenados (Fig. 3-1) y definimos una función mediante

$$U(x,y)=\int_\gamma p\,dx+q\,dy.$$

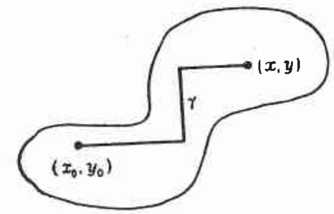

FIG. 3-1.

Puesto que la integral depende solo de los extremos, la función está bien definida. Además, si elegimos horizontal el último segmento de γ, podemos mantener y constante y hacer que varíe x sin cambiar los otros segmentos. En el último elegimos x como parámetro y se obtiene:

$$U(x,y)=\int^x p(x,y)\,dx+\text{cons.},$$

siendo irrelevante el límite inferior de la integral. De esta expresión se deduce inmediatamente que $\partial U/\partial x=p$. De la misma manera, eligiendo vertical el último segmento, podemos demostrar que $\partial U/\partial y=q$.

Se acostumbra escribir $dU=(\partial U/\partial x)dx+(\partial U/\partial y)dy$ y decir que una expresión $p\,dx+q\,dy$ que puede escribirse en esta forma es una *diferencial exacta*. Así, pues, una integral depende únicamente de los extremos si (y solo si) el integrando es una diferencial exacta. Obsérvese que p, q y U pueden ser reales o complejas. La función U, si existe, está determinada de manera única salvo una constante aditiva, pues si dos funciones tienen las mismas derivadas parciales, su diferencia ha de ser constante.

¿Cuándo es $f(z)\,dz = f(z)\,dx + if(z)\,dy$ una diferencial exacta? De acuerdo con la definición, debe existir una función $F(z)$ en Ω con derivadas parciales

$$\frac{\partial F(z)}{\partial x} = f(z)$$

$$\frac{\partial F(z)}{\partial y} = if(z).$$

Si es así, $F(z)$ cumple la ecuación de Cauchy-Riemann:

$$\frac{\partial F}{\partial x} = -i\,\frac{\partial F}{\partial y}$$

puesto que $f(z)$ es por hipótesis continua (de otra forma $\int_\gamma f\,dz$ no estaría definida), $F(z)$ es analítica con derivada $f(z)$ (Cap. II, sección 2-1, 2).

La integral $\int_\gamma f\,dz$, con f continua, depende únicamente de los extremos de γ si (y solo si) f es la derivada de una función analítica en Ω.

En estas circunstancias probaremos más adelante que $f(z)$ es a su vez analítica.

Como aplicación inmediata del resultado anterior tenemos que

$$\int_\gamma (z-a)^n dz = 0 \qquad [3\text{-}10]$$

para todas las curvas cerradas γ, con tal que el entero n sea mayor o igual que cero. En efecto, $(z-a)^n$ es la derivada de la función $(z-a)^{n+1}/(n+1)$, la cual es analítica en todo el plano. Si n es negativo, pero $\neq -1$, se verifica el mismo resultado para todas las curvas cerradas que no pasan por a, pues en la región complementaria del punto a la integral indefinida es todavía analítica y uniforme. Para $n=-1$, [3-10] no siempre se verifica. Consideremos una circunferencia C de centro a, representada por la ecuación $z = a + \rho e^{it}$, $0 \leqq t \leqq 2\pi$. Se obtiene

$$\int_C \frac{dz}{z-a} = \int_0^{2\pi} i\,dt = 2\pi i.$$

Este resultado prueba que es imposible definir una rama uniforme de $\log(z-a)$ en una corona circular $\rho_1 < |z-a| < \rho_2$. Por otra parte, si la curva cerrada γ está contenida en un semiplano que no contiene a a, la integral se anula, pues en tal semiplano se puede definir una rama analítica y uniforme de $\log(z-a)$.

EJERCICIOS

1. Calcúlese

$$\int_\gamma x\,dz$$

siendo γ el segmento orientado de 0 a $1+i$.

2. Hállese

$$\int_{|z|=r} x\,dz,$$

tomando el sentido positivo sobre la circunferencia de dos maneras: primera, tomando un parámetro, y segunda, observando que $x = \frac{1}{2}(z+\bar{z}) = \frac{1}{2}\left(z + \frac{r^2}{z}\right)$ en la circunferencia.

3. Calcúlese

$$\int_{|z|=2} \frac{dz}{z^2-1},$$

tomando en la circunferencia el sentido positivo.

4. Calcúlese

$$\int_{|z|=1} |z-1| \cdot |dz|.$$

5. Sea $f(z)$ analítica en una región Ω. Pruébese que para cualquier curva cerrada γ en Ω

$$\int_\gamma f(z)\overline{f'(z)}\,dz$$

es imaginario puro. [Se supone la continuidad de $f'(z)$.]

6. Supóngase que $f(z)$ es analítica en Ω y que satisface la desigualdad $|f(z)-1| < 1$. Pruébese que

$$\int_\gamma \frac{f'(z)}{f(z)}\,dz = 0$$

para toda curva cerrada en Ω. [Se supone la continuidad de $f'(z)$.]

2. *Teorema de Cauchy para un rectángulo.*—Existen varias formas del teorema de Cauchy que difieren más en su contenido topológico que en el analítico. Es natural comenzar con un caso en el que las consideraciones topológicas son triviales.

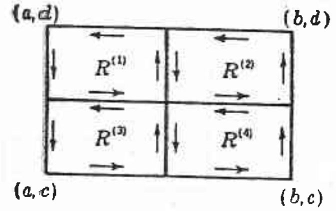

FIG 3-2.—Bisección de un rectángulo.

Consideremos, específicamente, un rectángulo R definido por las desigualdades $a \leqq x \leqq b$, $c \leqq y \leqq d$. Su perímetro puede imaginarse como una curva cerrada simple compuesta por cuatro segmentos, cuya orientación elegimos de forma que R quede a la izquierda de los segmentos orientados. El orden de los vértices es, pues, (a, c), (b, c), (b, d), (a, d). Nos referimos a esta curva cerrada como la *curva frontera* o *contorno* de R, y la denotamos por $\Gamma(R)$[1].

Hacemos hincapié en que se ha elegido a R como un conjunto de puntos cerrado y que, por consiguiente, no es una región. En el teorema que sigue consideraremos una función que es analítica en el rectángulo R. Recordamos al lector que tal función está, por hipótesis, definida y es analítica en una región que contiene a R.

La siguiente es una versión preliminar del *teorema de Cauchy:*

TEOREMA 2. *Si la función $f(z)$ es analítica en R, entonces*

$$\int_{\Gamma(R)} f(z)\,dz = 0. \qquad [3\text{-}11]$$

La demostración se basa en el método de bisección. Introduzcamos la notación

$$\eta(R) = \int_{\Gamma(R)} f(z)\,dz,$$

que también utilizaremos para cualquier rectángulo contenido en el dado. Si se divide R en cuatro rectángulos congruentes $R^{(1)}$, $R^{(2)}$, $R^{(3)}$, $R^{(4)}$, se tiene que

$$\eta(R) = \eta(R^{(1)}) + \eta(R^{(2)}) + \eta(R^{(3)}) + \eta(R^{(4)}), \qquad [3\text{-}12]$$

pues las integrales sobre los lados comunes se anulan unas con otras. Es importante observar que este hecho puede comprobarse explícitamente sin hacer uso ilegítimo de la intuición geométrica. Sin embargo, será útil una referencia a la figura 3-2.

[1] Esta notación se utilizará repetidamente.

Se sigue de [3-12] que al menos uno de los rectángulos $R^{(k)}$, $k = 1, 2, 3, 4$, debe satisfacer la condición

$$|\eta(R^{(k)})| \geqq \tfrac{1}{4} |\eta(R)|.$$

Llamamos a este rectángulo R_1; si varios rectángulos $R^{(k)}$ tienen esta propiedad, se hará la elección de acuerdo con alguna regla bien definida.

Este proceso puede repetirse indefinidamente, obteniéndose una sucesión de rectángulos encajados $R \supset R_1 \supset R_2 \supset \ldots \supset R_n \supset \ldots$ con la propiedad

$$|\eta(R_n)| \geqq \tfrac{1}{4} |\eta(R_{n-1})|$$

y, por tanto,

$$|\eta(R_n)| \geqq 4^{-n} |\eta(R)|. \qquad [3\text{-}13]$$

Los rectángulos R_n convergen a un punto $z^* \in R$ en el sentido de que R_n estará contenido en un entorno prescrito $|z - z^*| < \delta$, siempre y cuando se tome n suficientemente grande. En primer lugar, elegimos δ lo bastante pequeño para que $f(z)$ esté definida y sea analítica en $|z - z^*| < \delta$. En segundo lugar, dado $\epsilon > 0$, podemos elegir δ tal que

$$\left| \frac{f(z) - f(z^*)}{z - z^*} - f'(z^*) \right| < \epsilon$$

o

$$|f(z) - f(z^*) - (z - z^*) f'(z^*)| < \epsilon |z - z^*| \qquad [3\text{-}14]$$

para $|z - z^*| < \delta$. Suponemos que δ satisface ambas condiciones y que R_n está contenido en $|z - z^*| < \delta$.

Hacemos ahora la observación de que

$$\int_{\Gamma(R_n)} dz = 0$$

e

$$\int_{\Gamma(R_n)} z\,dz = 0.$$

Estos casos especiales triviales de nuestro teorema han sido ya demostrados en la sección 3-1, 1. Recordemos que la demostración dependía del hecho de que 1 y z son las derivadas de z y $z^2/2$, respectivamente.

En virtud de estas ecuaciones, podemos escribir

$$\eta(R_n) = \int_{\Gamma(R_n)} [f(z) - f(z^*) - (z-z^*)f'(z^*)]\,dz,$$

y se sigue de [3-14] que

$$|\eta(R_n)| \leqq \epsilon \int_{\Gamma(R_n)} |z-z^*| \cdot |dz|. \qquad [3\text{-}15]$$

En la última integral $|z-z^*|$ es a lo sumo igual a la diagonal d_n de R_n. Si L_n denota la longitud del perímetro de R_n, la integral es, por tanto, $\leqq d_n L_n$. Pero si d y L son las cantidades correspondientes para el rectángulo original R, es claro que $d_n = 2^{-n}d$ y $L_n = 2^{-n}L$. Por [3-15] tenemos entonces

$$|\eta(R_n)| \leqq 4^{-n} dL\,\epsilon,$$

y al comparar con [3-13] obtenemos

$$|\eta(R)| \leqq dL\,\epsilon.$$

Puesto que ϵ es arbitrario, cabe únicamente que sea $\eta(R) = 0$, con lo que queda probado el teorema.

Esta bella demostración, que no podría ser más sencilla, se debe esencialmente a E. Goursat, el cual descubrió que la hipótesis clásica de ser $f'(z)$ continua es redundante. Al mismo tiempo, la demostración es más fácil que las demostraciones clásicas, puesto que no se apoya en integrales dobles ni en diferenciación bajo el signo integral.

La hipótesis en el teorema 2 puede debilitarse considerablemente. Probaremos ahora el siguiente teorema, menos restrictivo, que tendrá una utilización muy importante.

TEOREMA 3. *Sea $f(z)$ analítica en el conjunto R' obtenido del rectángulo R por omisión de un número finito de puntos interiores ζ_j. Si es cierto que*

$$\lim_{z\to\zeta_j} (z-\zeta_j) f(z) = 0$$

para todo j, entonces

$$\int_{\Gamma(R)} f(z)\,dz = 0.$$

Basta considerar el caso de un solo punto excepcional ζ, pues, evidentemente, R puede dividirse en rectángulos más pequeños que contengan a lo sumo un ζ_j.

Dividimos R en nueve rectángulos, según se muestra en la figura 3-3, y aplicamos el teorema 2 a todos menos al rectángulo R_0 del centro. Si se suman las ecuaciones correspondientes [3-11], obtenemos después de eliminar los términos que se anulan:

$$\int_{\Gamma(R)} f\,dz = \int_{\Gamma(R_0)} f\,dz. \qquad [3\text{-}16]$$

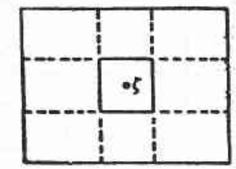

FIG. 3-3.

Si $\epsilon > 0$, podemos elegir el rectángulo R_0 tan pequeño que

$$|f(z)| \leqq \frac{\epsilon}{|z-\zeta|}$$

en $\Gamma(R_0)$. Por [3-16] tenemos así que

$$\left|\int_{\Gamma(R)} f\,dz\right| \leqq \epsilon \int_{\Gamma(R_0)} \frac{|dz|}{|z-\zeta|}$$

Si suponemos, lo que es posible, que R_0 es un cuadrado de centro ζ, un cálculo elemental muestra que

$$\int_{\Gamma(R_0)} \frac{|dz|}{|z-\zeta|} < 8.$$

Obtenemos así

$$\left|\int_{\Gamma(R)} f\,dz\right| < 8\epsilon,$$

y puesto que ϵ es arbitrario, el teorema queda demostrado.

Obsérvese que las hipótesis del teorema se cumplen ciertamente si $f(z)$ es analítica y *acotada* en R'.

3. *Teorema de Cauchy en un disco circular.*—No es cierto que la integral de una función analítica sobre una curva cerrada sea siempre cero. En efecto, hemos hallado que

$$\int_C \frac{dz}{z-a} = 2\pi i$$

cuando C es una circunferencia de centro a. Con objeto de estar seguros de que la integral se anula, es necesaria una hipótesis especial relativa a la región Ω, en la que se sabe que $f(z)$ es analítica y a la que está restringida la curva γ. Aún no estamos en condiciones de formular este supuesto, y por esta razón debemos dedicar nuestra atención a un caso muy especial. En lo que sigue supondremos que Ω es un disco circular abierto $|z-z_0|<\rho$, al que denotaremos por Δ.

TEOREMA 4. *Si $f(z)$ es analítica en un disco abierto Δ, entonces*

$$\int_\gamma f(z)\,dz=0 \qquad [3\text{-}17]$$

para toda curva cerrada γ en Δ.

La demostración es una repetición del razonamiento utilizado al probar la segunda mitad del teorema 1. Definimos una función $F(z)$ mediante

$$F(z)=\int_\sigma f\,dz, \qquad [3\text{-}18]$$

donde σ consiste en el segmento horizontal desde el centro (x_0, y_0) a (x, y_0) y el vertical desde (x, y_0) a (x, y); se ve inmediatamente que $\partial F/\partial y=if(z)$. Por otra parte, por el teorema 2, se puede reemplazar σ por un camino compuesto de un segmento vertical seguido de uno horizontal. Esta elección define la misma función $F(z)$, y se obtiene $\partial F/\partial x=f(z)$. Por tanto, $F(z)$ es analítica en Δ, con derivada $f(z)$, siendo $f(z)\,dz$ una diferencial exacta.

Evidentemente, la misma demostración serviría para cualquier región que contenga al rectángulo de vértices opuestos z_0 y z, siempre y cuando contenga a z. Un rectángulo, un semiplano o el interior de una elipse tienen todos esta propiedad; por tanto, el teorema 4 se verifica para cualquiera de estas regiones. Sin embargo, no es posible alcanzar generalidad completa mediante este método.

Es muy importante para las aplicaciones que la conclusión del teorema 4 siga siendo válida con las condiciones más débiles del teorema 3. Lo enunciaremos como un teorema aparte.

TEOREMA 5. *Sea $f(z)$ analítica en la región Δ' obtenida al omitir un número finito de puntos ζ_j de un disco abierto Δ. Si $f(z)$ satisface la condición $\lim_{z\to\zeta_j}(z-\zeta_j)f(z)=0$ para todo j, entonces se verifica [3-17] para cualquier curva cerrada γ en Δ'.*

La demostración habrá de ser modificada, pues no podemos dejar que σ pase por puntos excepcionales. Supongamos en primer lugar que ningún ζ_j está en las rectas $x=x_0$ e $y=y_0$. Entonces es posible evitar los puntos excepcionales haciendo que σ consista en los tres segmentos indicados (Fig. 3-4). Mediante una aplicación obvia del teorema 3, hallamos que el valor de $F(z)$ en [3-18] es independiente de la elección del segmento intermedio; además, el último segmento puede ser horizontal o vertical. Se determina como antes que $F(z)$ es una integral indefinida de $f(z)$, con lo que se concluye el teorema.

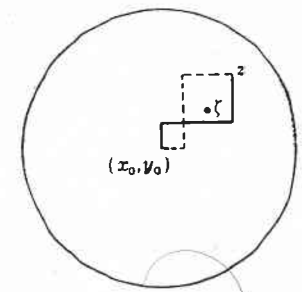

FIG. 3-4.

A fin de liberarnos de las restricciones en relación con el centro, observemos que $|z-z_0|$ tiene un máximo $\rho'<\rho$ en γ. Se sigue que γ está contenida en todo disco $|z-z'_0|\leqq\rho'+|z'_0-z_0|$, y si $|z'_0-z_0|<\frac{1}{2}(\rho-\rho')$, existe un disco ligeramente mayor con centro en z_0 contenido en Δ. Puesto que solo hay un número finito de puntos ζ_j, es evidentemente posible elegir z'_0 de acuerdo con esta condición y de forma que ningún punto ζ_j esté en las rectas vertical u horizontal que pasan por z'_0. El resultado ya probado puede aplicarse ahora al disco excéntrico.

EJERCICIOS

1. Pruébese el teorema análogo al 2 para un triángulo.
2. Demuéstrese el teorema 3 para un sector angular.
3. Pruébese que el teorema 5 sigue verificándose con infinitos puntos excepcionales ζ_j, siempre y cuando no tengan un punto de acumulación en Δ.

3-2. Fórmula de la integral de Cauchy.—Mediante una aplicación muy sencilla del teorema de Cauchy es posible representar una función analítica $f(z)$ como una integral curvilínea, en la que la variable z interviene como parámetro. Esta representación, conocida como la *fórmula de la integral de Cauchy,* tiene numerosas aplicaciones importantes. Sobre todo, nos permite estudiar las propiedades locales de una función analítica con gran detalle.

1. *Indice de un punto con respecto a una curva cerrada.*—Como

preliminar a la fórmula de derivación de Cauchy, debemos definir un concepto que indica de una manera precisa las veces que una curva cerrada rodea a un punto no perteneciente a la misma. Si la curva es diferenciable a trozos, como supondremos sin gran pérdida de generalidad, la definición puede basarse en el lema siguiente:

LEMA 1. *Si la curva cerrada diferenciable a trozos γ no pasa por el punto a, entonces el valor de la integral*

$$\int_\gamma \frac{dz}{z-a}$$

es un múltiplo de $2\pi i$.

Este lema puede parecer trivial, pues podemos escribir

$$\int_\gamma \frac{dz}{z-a} = \int_\gamma d\log(z-a) = \int_\gamma d\log|z-a| + i\int_\gamma d\arg(z-a).$$

Cuando z describe una curva cerrada, $\log|z-a|$ vuelve a su valor inicial y $\arg(z-a)$ aumenta o disminuye en un múltiplo de 2π. Esto podría parecer que implicaba el lema, pero un examen más minucioso nos muestra que el razonamiento carece de valor, a no ser que definamos $\arg(z-a)$ de manera única. En cuanto intentemos hacerlo veremos que una demostración por este método no es precisamente sencilla.

Afortunadamente puede darse una demostración más fácil. Si la ecuación de γ es $z=z(t)$, $\alpha \leqq t \leqq \beta$, consideremos la función

$$h(t) = \int_\alpha^t \frac{z'(t)}{z(t)-a}\,dt.$$

Está definida y es continua en el intervalo cerrado (α, β), siendo su derivada

$$h'(t) = \frac{z'(t)}{z(t)-a}$$

siempre y cuando $z'(t)$ sea continua. De esta ecuación se sigue que la derivada de $e^{-h(t)}[z(t)-a]$ se anula excepto quizá en un número finito de puntos, y puesto que esta función es continua, debe reducirse a una constante. Tenemos así

$$e^{h(t)} = \frac{z(t)-a}{z(\alpha)-a}$$

Puesto que $z(\beta)=z(\alpha)$, obtenemos $e^{h(\beta)}=1$; por tanto, $h(\beta)$ ha de ser un múltiplo de $2\pi i$. Esto prueba el lema.

Podemos ahora definir *el índice del punto a con respecto a la curva γ* mediante la ecuación

$$n(\gamma, a) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a}$$

Con una terminología sugestiva, al índice se le llama también *número de vueltas* de γ con respecto a a.

Evidentemente, $n(-\gamma, a) = -n(\gamma, a)$.

La propiedad siguiente es una consecuencia inmediata del teorema 4:

1) *Si γ está en el interior de un círculo, entonces $n(\gamma, a)=0$ para todos los puntos a exteriores al mismo círculo.*

Como conjunto de puntos, γ es cerrado y acotado. Su complemento es abierto y puede representarse como unión de regiones disjuntas, las componentes del complemento. Diremos brevemente que γ *determina* estas regiones. Si se consideran las regiones complementarias en el plano completo, hay exactamente una que contiene al punto del infinito. En consecuencia, γ determina una (y solo una) región no acotada.

2) *Como función de a, el índice $n(\gamma, a)$ es constante en cada una de las regiones determinadas por γ, y cero en la región no acotada.*

Cualquier par de puntos en la misma región determinada por γ puede unirse mediante una línea poligonal que no corta a γ. Por esta razón basta probar que $n(\gamma, a)=n(\gamma, b)$ si γ no corta al segmento determinado por a y b. Fuera de este segmento, $(z-a)/(z-b)$ nunca es real y $\leqq 0$. Debido a esto la rama principal de la función $\log[(z-a)/(z-b)]$ es analítica en el complemento del segmento. Su derivada es igual a $(z-a)^{-1}-(z-b)^{-1}$, y si γ no corta al segmento, debemos tener

$$\int_\gamma \left(\frac{1}{z-a} - \frac{1}{z-b}\right) dz = 0;$$

por consiguiente, $n(\gamma, a)=n(\gamma, b)$. Si $|a|$ es suficientemente grande, γ está contenido en un disco $|z| < \rho < |a|$, y por 1) se tiene que $n(\gamma, a)=0$. Esto prueba que $n(\gamma, a)=0$ en la región no acotada.

Veremos que el caso $n(\gamma, a)=1$ es particularmente importante, siendo deseable formular una condición geométrica que nos conduzca a esta consecuencia. Por sencillez tomamos $a=0$.

LEMA 2. *Sean z_1 y z_2 dos puntos de una curva cerrada γ que no pasa por el origen. Denotamos el arco parcial desde z_1 a z_2 en la dirección de la curva por γ_1, y el arco parcial desde z_2 a z_1 por γ_2. Supongamos que z_1 está en el semiplano inferior y z_2 en el semiplano superior. Si γ_1 no corta al eje real negativo y γ_2 no corta al eje real positivo, entonces $n(\gamma, 0)=1$.*

Para demostrarlo tracemos las semirrectas L_1 y L_2 por el origen a los puntos z_1 y z_2 (Fig. 3-5). Sean ζ_1 y ζ_2 los puntos en que L_1 y L_2 cortan a la circunferencia C con centro en el origen. Si se describe C en el sentido positivo, el arco C_1 desde ζ_1 a ζ_2 no corta al eje negativo, y el arco C_2 de ζ_2 a ζ_1 no corta al positivo. Denotemos los segmentos orientados de z_1 a ζ_1 y de z_2 a ζ_2 por δ_1 y δ_2. Considerando las curvas cerradas $\sigma_1=\gamma_1+\delta_2-C_1-\delta_1$, $\sigma_2=\gamma_2+\delta_1-C_2-\delta_2$, hallamos que $n(\gamma, 0)=n(C, 0)+n(\sigma_1, 0)+n(\sigma_2, 0)$, después de efectuar simplificaciones. Pero σ_1 no corta al eje negativo; por consiguiente, el origen pertenece a la región no acotada determinada por σ_1, obteniéndose $n(\sigma_1, 0)=0$. Por razones análogas, $n(\sigma_2, 0)=0$, y se tiene que $n(\gamma, 0)=n(C, 0)=1$.

2. *Fórmula de la integral.*—Sea $f(z)$ analítica en un disco circular abierto Δ. Consideremos una curva cerrada γ en Δ y un punto $a \in \Delta$ que no esté en γ. Apliquemos el teorema de Cauchy a la función

$$F(z)=\frac{f(z)-f(a)}{z-a}$$

Esta función es analítica para $z \neq a$. Para $z=a$ no está definida, pero satisface la condición

$$\lim_{z\to a} F(z)(z-a)=\lim_{z\to a}[f(z)-f(a)]=0,$$

que es la condición del teorema 5. Se concluye que

$$\int_\gamma \frac{f(z)-f(a)}{z-a}\,dz=0.$$

Esta ecuación puede escribirse en la forma

$$\int_\gamma \frac{f(z)\,dz}{z-a}=f(a)\int_\gamma \frac{dz}{z-a}$$

y observemos que la integral del segundo miembro es, por definición, $2\pi i \cdot n(\gamma, a)$. Hemos probado así el teorema siguiente:

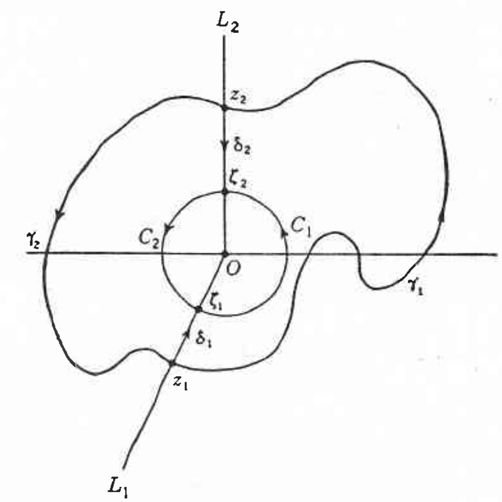

FIG. 3-5.

TEOREMA 6. *Sea $f(z)$ analítica en un disco abierto Δ y sea γ una curva cerrada en Δ. Para cualquier punto a no perteneciente a γ, se tiene*

$$n(\gamma, a)\cdot f(a)=\frac{1}{2\pi i}\int_\gamma \frac{f(z)\,dz}{z-a} \qquad [3\text{-}19]$$

siendo $n(\gamma, a)$ el índice de a con respecto a γ.

En este enunciado hemos suprimido el requisito de que a sea un punto de Δ. Lo hemos hecho de esta forma en vista de la interpretación obvia de la fórmula [3-19] en el caso en que a no pertenece a Δ. Realmente, en este caso, $n(\gamma, a)$ y la integral del segundo miembro son ambos cero, permaneciendo correcta la fórmula cualquiera que sea el valor que deseemos asignar a $f(a)$.

Es evidente que el teorema 6 sigue siendo válido para cualquier región Ω a la que pueda aplicarse el teorema 5. La presencia de puntos excepcionales ζ_j está admitida, con tal que ninguno de ellos coincida con a.

La aplicación más usual es la correspondiente al caso en que $n(\gamma, a) = 1$. Tenemos entonces

$$f(a) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)\,dz}{z-a} \qquad [3\text{-}20]$$

y esta la interpretamos como una *fórmula de representación.* En efecto, nos permite computar $f(a)$ tan pronto como se nos dan los valores de $f(z)$ en γ, junto con el hecho de que $f(z)$ es analítica en Δ. Con tal que el orden de a con respecto a γ siga siendo uno, podríamos dar a a diferentes valores en [3-20]. Podemos, pues, tratar a como si fuera una variable, siendo conveniente cambiar la notación y escribir de nuevo [3-20] en la forma

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(\zeta)\,d\zeta}{\zeta - z} \qquad [3\text{-}21]$$

Esta es la fórmula conocida corrientemente como la *fórmula de la integral de Cauchy.* Debemos recordar que solo es válida cuando $n(\gamma, z) = 1$ y que la hemos demostrado únicamente cuando $f(z)$ es analítica en un disco.

EJERCICIOS

1. Calcúlese

$$\int_{|z|=1} \frac{e^z}{z}\,dz.$$

2. Determínese

$$\int_{|z|=2} \frac{dz}{z^2+1}$$

por descomposición del integrando en fracciones simples.

3. Calcúlese

$$\int_{|z|=\varrho} \frac{|dz|}{|z-a|^2}$$

bajo la condición $|a| \neq \rho$. *Sugerencia:* Utilícese la ecuación $z\bar{z} = \rho^2$ y

$$|dz| = -i\rho \frac{dz}{z}$$

3. *Derivadas de órdenes superiores.*—La fórmula de representación [3-21] nos proporciona una herramienta ideal para el estudio de las propiedades locales de las funciones analíticas. En particular, podemos probar ahora que una función analítica tiene derivadas de todos los órdenes, las cuales son también analíticas.

Consideremos una función $f(z)$ que es analítica en una región arbitraria Ω. Dado un punto $a \in \Omega$, determinamos un entorno Δ contenido en Ω, y en Δ, una circunferencia con centro en a. Se puede aplicar el teorema 6 a $f(z)$ en Δ. Puesto que $n(C, a) = 1$, tenemos que $n(C, z) = 1$ para todos los puntos z interiores a C. Para tales puntos z, de acuerdo con [3-21], obtenemos

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)\,d\zeta}{\zeta - z}$$

Siempre y cuando la integral pueda derivarse bajo el signo de integración, tenemos que

$$f'(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)\,d\zeta}{(\zeta - z)^2} \qquad [3\text{-}22]$$

y

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_C \frac{f(\zeta)\,d\zeta}{(\zeta - z)^{n+1}} \qquad [3\text{-}23]$$

Si las derivaciones pudieran justificarse, habríamos probado la existencia de todas las derivadas en los puntos interiores a C. Puesto que todo punto en Ω es interior a alguno de tales círculos, se demostrará así la existencia en toda la región Ω. Al mismo tiempo habremos obtenido una fórmula de representación conveniente para las derivadas.

En lo que se refiere a la citada justificación, podríamos o referirnos a los teoremas correspondientes en el caso real, o probar un teorema general relativo a integrales curvilíneas, cuyo integrando depende analíticamente de un parámetro. En realidad, únicamente probaremos el lema siguiente, que es todo lo que necesitamos en el presente caso:

LEMA 3. *Sea $\varphi(\zeta)$ continua en el arco γ. Entonces la función*

$$F_n(z) = \int_\gamma \frac{\varphi(\zeta)\,d\zeta}{(\zeta - z)^n}$$

es analítica en cada una de las regiones determinadas por γ, siendo su derivada $F'_n(z) = nF_{n+1}(z)$.

Demostraremos primero que $F_1(z)$ es continua. Sea z_0 un punto no perteneciente a γ; elegimos el entorno $|z-z_0| < \delta$ de forma que sea disjunto con γ. Haciendo que z varíe en el entorno más pequeño $|z-z_0| < \delta/2$, conseguimos que $|\zeta - z| > \delta/2$ para todo $\zeta \in \gamma$. De

$$F_1(z) - F_1(z_0) = (z-z_0)\int_\gamma \frac{\varphi(\zeta)\,d\zeta}{(\zeta-z)(\zeta-z_0)}$$

obtenemos inmediatamente

$$|F_1(z) - F_1(z_0)| < |z-z_0| \cdot \frac{2}{\delta^2}\int_\gamma |\varphi|\,|d\zeta|,$$

y esta desigualdad demuestra la continuidad de $F_1(z)$ en z_0.

Esta parte del lema, aplicada a la función $\varphi(\zeta)/(\zeta - z_0)$, nos da la siguiente expresión para el cociente incremental

$$\frac{F_1(z) - F_1(z_0)}{z-z_0} = \int_\gamma \frac{\varphi(\zeta)\,d\zeta}{(\zeta-z)(\zeta-z_0)}$$

lo que nos dice que dicho cociente tiende al límite $F_2(z_0)$ cuando z tiende a z_0. Queda, por tanto, probado que $F'_1(z) = F_2(z)$.

El caso general se demuestra por inducción. Supongamos que hemos demostrado que $F'_{n-1}(z) = (n-1)F_n(z)$. De la identidad

$$F_n(z) - F_n(z_0) =$$

$$= \left[\int_\gamma \frac{\varphi\,d\zeta}{(\zeta-z)^{n-1}(\zeta-z_0)} - \int_\gamma \frac{\varphi\,d\zeta}{(\zeta-z_0)^n}\right] + (z-z_0)\int_\gamma \frac{\varphi\,d\zeta}{(\zeta-z)^n(\zeta-z_0)}$$

se deduce que $F_n(z)$ es continua. En efecto, de acuerdo con la hipótesis de inducción aplicada a $\varphi(\zeta)/(\zeta-z_0)$, el primer término tiende a cero cuando z tiende a z_0, y en el segundo, el coeficiente de $z-z_0$ está acotado en un entorno de z_0. Ahora bien: si dividimos la identidad por $z-z_0$ y hacemos tender z a z_0, el cociente en el primer miembro tiende a la derivada, que por la hipótesis de inducción es igual a $(n-1)F_{n+1}(z_0)$. El factor restante en el segundo miembro es continuo, por lo que ya ha sido probado, y tiene por límite $F_{n+1}(z_0)$. Por tanto, existe $F'_n(z_0)$ y es igual a $nF_{n+1}(z_0)$.

Evidentemente, el lema 3 es lo que necesitábamos para deducir de manera rigurosa las fórmulas [3-22] y [3-23]. Hemos demostrado

así que una función analítica tiene derivadas de todos los órdenes, las cuales son analíticas y pueden representarse por la fórmula [3-23].

Entre las consecuencias de este resultado nos interesa señalar dos teoremas clásicos. El primero se conoce como *teorema de Morera*, y puede enunciarse en la siguiente forma:

Si $f(z)$ está definida y es continua en la región Ω, y si $\int_\gamma f(z)\,dz = 0$ para todas las curvas cerradas γ en Ω, entonces $f(z)$ es analítica en Ω.

La hipótesis implica, como ya hicimos notar en la sección 3-1, 1, que $f(z)$ es la derivada de una función analítica $F(z)$. Sabemos ahora que $f(z)$ es también analítica.

Un segundo resultado clásico se conoce como *teorema de Liouville:*

Una función analítica y acotada en todo el plano debe reducirse a una constante.

Para demostrarlo haremos uso de una sencilla acotación que se deduce de [3-23]. Sea r el radio de C y supongamos que $|f(\zeta)| \leqq M$ en C. Si aplicamos [3-23] con $z=a$, obtenemos inmediatamente

$$|f^{(n)}(a)| \leqq Mn!\,r^{-n}. \qquad [3\text{-}24]$$

Para el teorema de Liouville necesitamos únicamente el caso $n=1$. La hipótesis significa que $|f(\zeta)| \leqq M$ en todas las circunferencias. Podemos, por tanto, hacer tender r a ∞, y [3-24] nos conduce a $f'(a) = 0$ para todo a. Se concluye que la función es constante.

El teorema de Liouville nos proporciona una demostración casi trivial del *teorema fundamental del álgebra*. Supongamos que $P(z)$ es un polinomio de grado mayor que cero. Si $P(z)$ no fuese nunca cero, la función $1/P(z)$ sería analítica en todo el plano. Sabemos que $P(z)$ tiende a infinito cuando z tiende a infinito, y que, por tanto, $1/P(z)$ tiende a cero. Esto implica que está acotada (el valor absoluto es continuo en la esfera de Riemann, y tiene, por consiguiente, un máximo finito), y de acuerdo con el teorema de Liouville $1/P(z)$ debería ser constante. Puesto que no es así, la ecuación $P(z) = 0$ tiene que tener una raíz.

La desigualdad [3-24] se conoce con el nombre de *desigualdad de Cauchy*. Prueba fundamentalmente que todas las derivadas sucesivas de una función analítica no pueden ser arbitrarias; tienen que existir siempre un r y un M tales que se verifique [3-24]. A fin de conseguir el mejor uso de la desigualdad es importante hacer una juiciosa elección de r con objeto de minimizar la función $M(r)r^{-n}$, siendo $M(r)$ el máximo de $|f|$ en $|\zeta - a| = r$.

EJERCICIOS

1. Pruébese que una función que es analítica en todo el plano y satisface una desigualdad de la forma $|f(z)| < |z|^n$ para algún n y para todo $|z|$ suficientemente grande, tiene que reducirse a un polinomio.

2. Si $f(z)$ es analítica y $|f(z)| \leqq M$ para $|z| \leqq R$, hállese una cota superior de $|f^{(n)}(z)|$ en $|z| \leqq \rho < R$.

3. Si $f(z)$ es analítica para $|z| < 1$ y $|f(z)| \leqq 1/(1-|z|)$, hállese la mejor aproximación de $|f^{(n)}(0)|$ que pueda proporcionar la desigualdad de Cauchy.

4. Pruébese que las derivadas sucesivas en un punto nunca pueden satisfacer las desigualdades $|f^{(n)}(z)| > n!\,n^n$. Formúlese un teorema del mismo tipo más ajustado.

3-3. Propiedades locales de las funciones analíticas.—Hemos demostrado ya que una función analítica tiene derivadas sucesivas de todos los órdenes. En esta sección haremos un estudio más detallado de las propiedades locales. Incluirá una clasificación de las *singularidades aisladas* de las funciones analíticas.

1. *Singularidades evitables. Teorema de Taylor.*—En el teorema 3 se introdujo una condición más débil en sustitución de la analiticidad en un número finito de puntos, sin afectar el resultado final. Además hemos probado, en el teorema 5, que en un disco circular sigue siendo cierto el teorema de Cauchy con estas condiciones más débiles. Esto constituyó un punto esencial en la deducción de la fórmula de la integral de Cauchy, pues necesitábamos aplicar el teorema de Cauchy a una función de la forma $[f(z) - f(a)]/(z-a)$.

Por último, se señaló que la fórmula de la integral de Cauchy seguía siendo válida aun existiendo un número finito de puntos excepcionales que cumpliesen todas las condiciones fundamentales del teorema 3, con tal que ninguno de ellos coincidiera con a. Esta observación es más importante de lo que pudiera parecer en principio. En efecto, la fórmula de Cauchy nos proporciona una representación de $f(z)$ mediante una integral que, al depender de z, tiene el mismo carácter en los puntos excepcionales que en los demás. Se sigue que los puntos excepcionales lo son únicamente por falta de información y no por su naturaleza intrínseca. A los puntos que gozan

de este carácter se les llama *singularidades evitables*. Probaremos el siguiente teorema, más preciso:

TEOREMA 7. *Supongamos que $f(z)$ es analítica en la región Ω', obtenida de la región Ω por omisión de un punto. Una condición necesaria y suficiente para que exista en Ω una función analítica que coincida con $f(z)$ en Ω' es que $\lim_{z \to a} (z-a)f(z) = 0$. La función prolongada está determinada de manera única.*

La necesidad y la unicidad son triviales, puesto que la función prolongada tiene que ser continua en a. A fin de demostrar la suficiencia trazamos una circunferencia C con centro en a, tal que C y sus puntos interiores estén contenidos en Ω. La fórmula de Cauchy es válida, y podemos escribir

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)\,d\zeta}{\zeta - z}$$

para todo $z \neq a$ e interior a C. Pero la integral del segundo miembro representa una función analítica de z en todo el interior de C. Por tanto, la función que es igual a $f(z)$ para $z \neq a$ y que tiene el valor

$$\frac{1}{2\pi i} \int_C \frac{f(\zeta)\,d\zeta}{\zeta - a} \qquad [3\text{-}25]$$

para $z = a$ es analítica en Ω. Es natural que se denote la función prolongada mediante $f(z)$ y el valor [3-25] por $f(a)$.

Apliquemos este resultado a la función

$$F(z) = \frac{f(z) - f(a)}{z - a}$$

utilizada en la demostración de la fórmula de Cauchy. No está definida para $z = a$, pero satisface la condición $\lim_{z \to a} (z-a)F(z) = 0$. El límite de $F(z)$ cuando z tiende a a es $f'(a)$. Existe, por consiguiente, una función analítica igual a $F(z)$ para $z \neq a$, e igual a $f'(a)$ para $z = a$. Denotemos esta función por $f_1(z)$. Repitiendo el mismo proceso podemos definir una función analítica $f_2(z)$ que es igual a $[f_1(z) - f_1(a)]/(z-a)$ para $z \neq a$, y a $f'_1(a)$ para $z = a$, y así sucesivamente.

El esquema recurrente mediante el cual se define $f_n(z)$ puede escribirse en la forma

$$\begin{aligned} f(z) &= f(a)+(z-a)f_1(z)\\ f_1(z) &= f_1(a)+(z-a)f_2(z)\\ &\cdots\cdots\cdots\cdots\\ f_{n-1}(z) &= f_{n-1}(a)+(z-a)f_n(z). \end{aligned}$$

De estas ecuaciones, que son trivialmente válidas también para $z=a$, se obtiene

$$f(z)=f(a)+(z-a)f_1(a)+(z-a)^2f_2(a)+\ldots+(z-a)^{n-1}f_{n-1}(a)+\\ +(z-a)^nf_n(z).$$

Derivando n veces y poniendo $z=a$, se llega a

$$f^{(n)}(a)=n!f_n(a).$$

Con lo que quedan determinados los coeficientes $f_n(a)$ y se obtiene la siguiente forma del *teorema de Taylor:*

TEOREMA 8. *Si $f(z)$ es analítica en una región Ω, que contiene a a, se tiene que*

$$f(z)=f(a)+\frac{f'(a)}{1!}(z-a)+\frac{f''(a)}{2!}(z-a)^2+\ldots+\\ +\frac{f^{(n-1)}(a)}{(n-1)!}(z-a)^{n-1}+f_n(z)(z-a)^n, \qquad [3\text{-}26]$$

donde $f_n(z)$ es analítica en Ω.

Debe distinguirse entre este desarrollo finito y la *serie de Taylor* infinita, que estudiaremos más adelante. Sin embargo, el desarrollo finito [3-26] resulta más útil para el estudio de las propiedades locales de $f(z)$. Su utilidad está acrecentada por el hecho de que $f_n(z)$ tiene una expresión explícita sencilla como integral curvilínea.

Utilizando la misma circunferencia C que antes, tenemos primero

$$f_n(z)=\frac{1}{2\pi i}\int_C\frac{f_n(\zeta)d\zeta}{\zeta-z}$$

En lugar de $f_n(\zeta)$ ponemos la expresión obtenida de [3-26]. Habrá un término principal que contiene a $f(\zeta)$. Los términos restantes

son, salvo factores constantes, de la forma

$$F_\nu(a)=\int_C\frac{d\zeta}{(\zeta-a)^\nu(\zeta-z)}, \qquad \nu\geqq 1.$$

Pero

$$F_1(a)=\frac{1}{z-a}\int_C\left(\frac{1}{\zeta-z}-\frac{1}{\zeta-a}\right)d\zeta=0,$$

idénticamente para todo a interior a C. Por el lema 3 tenemos que $F_{\nu+1}(a)=F_1^{(\nu)}(a)/\nu!$, y así $F_\nu(a)=0$ para todo $\nu\geqq 1$. Por consiguiente, la expresión de $f_n(z)$ se reduce a

$$f_n(z)=\frac{1}{2\pi i}\int_C\frac{f(\zeta)d\zeta}{(\zeta-a)^n(\zeta-z)} \qquad [3\text{-}27]$$

Esta representación es válida en el interior de C.

EJERCICIO

Sea $f(z)$ analítica en el punto a y supongamos que
$f(z)=A_0+A_1(z-a)+A_2(z-a)^2+\ldots+A_{n-1}(z-a)^{n-1}+\varphi_n(z)(z-a)^{n-1}$
siendo $\lim\varphi_n(z)=0$ para $z\to a$. Demuéstrese que $A_k=f^{(k)}(a)/k!$

2. *Ceros y polos.*—Si $f(a)$ y todas sus derivadas $f^{(\nu)}(a)$ se anulan, por [3-26] podemos escribir

$$f(z)=f_n(z)(z-a)^n \qquad [3\text{-}28]$$

para cualquier n. Se puede obtener una acotación de $f_n(z)$ mediante [3-27]. El disco cuya circunferencia es C ha de estar contenido en la región Ω, en la que $f(z)$ está definida y es analítica. El valor absoluto $|f(z)|$ tiene un máximo M en C; si se denota por R al radio de C, resulta que

$$|f_n(z)|\leqq\frac{M}{R^{n-1}(R-|z-a|)}$$

para $|z-a|<R$. Por [3-28] tenemos, pues,

$$|f(z)|\leqq\left(\frac{|z-a|}{R}\right)^n\cdot\frac{MR}{R-|z-a|}$$

Pero $(|z-a|/R)^n \to 0$ para $n \to \infty$, puesto que $|z-a| < R$. Por tanto, $f(z)=0$ en el interior de C.

Probaremos ahora que $f(z)$ es idénticamente cero en todo Ω. Sea E_1 el conjunto en el que $f(z)$ y todas sus derivadas se anulan, y E_2, el conjunto en el que la función o una de las derivadas es distinta de cero. Por el razonamiento anterior E_1 es abierto, siéndolo también E_2 por ser la función y todas sus derivadas continuas. Por consiguiente, o E_1 es vacío o lo es E_2. Si E_2 es vacío, la función es idénticamente cero. Si E_1 es vacío, $f(z)$ nunca puede anularse a la vez que todas sus derivadas.

Supongamos que $f(z)$ no es idénticamente cero. Así, si $f(a)=0$, existirá una primera derivada, $f^{(h)}(a)$, que es diferente de cero. Decimos entonces que *a es un cero de orden h,* y el resultado que acabamos de probar expresa que no existen ceros de orden infinito. A este respecto, una función analítica se comporta localmente como un polinomio, y precisamente como en el caso de los polinomios, vemos que es posible escribir $f(z)=(z-a)^h f_h(z)$, siendo $f_h(z)$ analítica y $f_h(a) \neq 0$.

En la misma situación, $f_h(z) \neq 0$ en un entorno de a, puesto que $f_h(z)$ es continua, y $z=a$ es el único cero de $f(z)$ en este entorno. En otras palabras: los ceros de una función analítica que no se anula idénticamente son puntos *aislados.* Esta propiedad puede también formularse como un teorema de unicidad: *Si $f(z)$ y $g(z)$ son analíticas en Ω, y si $f(z)=g(z)$ en un conjunto que tiene un punto de acumulación en Ω, entonces $f(z)$ es idénticamente igual a $g(z)$.* Esta conclusión se obtiene considerando la diferencia $f(z)-g(z)$.

Ejemplos particulares de este resultado que merezcan enunciarse son los siguientes: Si $f(z)$ es idénticamente cero en una subregión de Ω, entonces es idénticamente cero en Ω, y lo mismo es cierto si $f(z)$ se anula en un arco que no se reduzca a un punto. Podemos también decir que una función analítica está determinada unívocamente por sus valores en cualquier conjunto con un punto de acumulación en la región de analiticidad. Esto no significa que dispongamos de un procedimiento mediante el cual puedan computarse los valores de la función.

Consideremos ahora una función analítica $f(z)$ en un entorno de a, excepto quizá en el mismo punto a. En otras palabras: $f(z)$ será analítica en una región $0<|z-a|<\delta$. Al punto a se le llama *singularidad aislada* de $f(z)$. Hemos tratado ya el caso de una singularidad evitable. Puesto que podemos definir $f(a)$ de forma que $f(z)$ resulte

analítica en el disco $|z-a|<\delta$, no son necesarias más consideraciones[1].

Si $\lim_{z\to a} f(z)=\infty$, se dice que el punto a es un *polo* de $f(z)$, y se pone $f(a)=\infty$. Existe un $\delta' \leqq \delta$ tal que $f(z) \neq 0$ para $0<|z-a|<\delta'$. En esta región la función $g(z)=1/f(z)$ está definida y es analítica. Pero la singularidad de $g(z)$ en a es evitable, teniendo $g(z)$ una prolongación analítica con $g(a)=0$. Puesto que $g(z)$ no se anula idénticamente, el cero en a tiene orden finito, y podemos escribir $g(z)=(z-a)^h g_h(z)$ con $g_h(a) \neq 0$. El número h es el *orden* del polo, y $f(z)$ admite la representación $f(z)=(z-a)^{-h} f_h(z)$, donde $f_h(z) = 1/g_h(z)$ es analítica y distinta de cero en un entorno de a. Así, la naturaleza del polo es exactamente la misma que en el caso de una función racional.

Una función $f(z)$ que es analítica en una región Ω, excepto por polos, se dice *meromorfa* en Ω. Con más precisión, para todo $a \in \Omega$ existirá un entorno $|z-a|<\delta$, contenido en Ω, tal que o $f(z)$ es analítica en todo el entorno, o $f(z)$ es analítica en $0<|z-a|<\delta$, siendo un polo la singularidad aislada. Obsérvese que los polos de una función meromorfa son, *por definición,* aislados. El cociente $f(z)/g(z)$ de dos funciones analíticas en Ω es una función meromorfa en Ω, con tal que $g(z)$ no sea idénticamente cero. Los únicos polos posibles son los ceros de $g(z)$, pero un cero común de $f(z)$ y de $g(z)$ puede ser también una singularidad evitable. Si este es el caso, el valor del cociente debe determinarse por continuidad. Con más generalidad, la suma, el producto y el cociente de dos funciones meromorfas es una función meromorfa. Debe excluirse el caso de denominador que se anula idénticamente, a no ser que deseemos considerar la constante ∞ como una función meromorfa.

Para discutir más detalladamente las singularidades aisladas, consideremos las dos condiciones siguientes: 1) $\lim_{z\to a} |z-a|^\alpha |f(z)|=0$; 2) $\lim_{z\to a} |z-a|^\alpha |f(z)|=\infty$, para valores reales de α. Si se verifica 1) para cierto valor de α, también se verifica para valores mayores que α y, por tanto, para algún entero m. Entonces $(z-a)^m f(z)$ tiene una singularidad evitable y se anula para $z=a$. O bien $f(z)$ es idénticamente cero, en cuyo caso 1) se verifica para todo α, o se tiene que $(z-a)^m f(z)$ tiene un cero de orden finito k. En el último caso se

[1] Cuando a es una singularidad evitable, suele decirse que $f(z)$ es *regular* en a; este término se usa a veces como sinónimo de *analítica.*

deduce inmediatamente que se verifica 1) para todo $\alpha > h = -m-k$, mientras que se verifica 2) para todo $\alpha < h$. Supongamos ahora que se verifica 2) para algún α; entonces se verifica también para algún α menor y, por consiguiente, para cierto entero n. La función $(z-a)^n f(z)$ tiene un polo de orden finito l, y poniendo $h = n + l$ tenemos nuevamente que se verifica 1) para $\alpha > h$ y 2) para $\alpha < h$. La discusión muestra la existencia de tres posibilidades: I) se verifica la condición 1) para todo valor de α, y $f(z)$ se anula idénticamente; II) existe un entero h, tal que se verifica 1) para $\alpha > h$ y 2) para $\alpha < h$; III) no se cumplen 1) y 2) para ningún valor de α.

El caso I) carece de interés. En el caso II) se puede llamar a h el *orden algebraico* de $f(z)$ en a. Es positivo en el caso de un polo, negativo en el de un cero, y cero si $f(z)$ es analítica, pero distinta de cero en a. Lo notable es que el orden es siempre un entero; no existe función analítica uniforme que tienda a cero o a infinito como una potencia fraccionaria de $|z-a|$.

En el caso de un polo de orden h, apliquemos el teorema 8 a la función analítica $(z-a)^h f(z)$. Obtenemos un desarrollo de la forma

$$(z-a)^h f(z) = B_h + B_{h-1}(z-a) + \ldots + B_1(z-a)^{h-1} + \varphi(z)(z-a)^h,$$

siendo $\varphi(z)$ analítica en $z=a$. Para $z \neq a$ podemos dividir por $(z-a)^h$ y obtenemos

$$f(z) = B_h(z-a)^{-h} + B_{h-1}(z-a)^{-h+1} + \ldots + B_1(z-a)^{-1} + \varphi(z).$$

A la parte de este desarrollo que precede a $\varphi(z)$ se le llama *parte singular* de $f(z)$ en $z=a$. Así, pues, un polo no solo tiene un orden, sino también una parte singular perfectamente definida. La diferencia de dos funciones con la misma parte singular es analítica en a.

En el caso III) el punto a es una *singularidad aislada esencial.* En el entorno de una singularidad esencial, $f(z)$ está a la vez sin acotar y tan próxima a cero como se quiera. Como caracterización del complicado comportamiento de una función en un entorno de una singularidad esencial, demostraremos el siguiente teorema clásico debido a Weierstrass:

TEOREMA 9. *Una función analítica toma valores tan próximos como se quiera a un valor complejo dado en todo entorno de una singularidad esencial.*

Si lo enunciado no fuera cierto, podríamos hallar un número

complejo A y un $\delta > 0$ tales que $|f(z) - A| > \delta$ en un entorno de a (excepto para $z = a$). Para cualquier $\alpha < 0$ tendríamos entonces que $\lim_{z \to a} |z-a|^\alpha |f(z) - A| = \infty$. Por tanto, a no sería una singularidad esencial de $f(z) - A$. Así, pues, existe un β con $\lim_{z \to a} |z-a|^\beta |f(z) - A| = 0$, y cabe elegir $\beta > 0$. Puesto que en este caso $\lim_{z \to a} |z-a|^\beta |A| = 0$, se seguiría que $\lim_{z \to a} |z-a|^\beta |f(z)| = 0$, y a no sería una singularidad esencial de $f(z)$. La contradicción prueba el teorema.

El concepto de singularidad aislada se aplica también a funciones que son analíticas en un entorno $|z| > R$ de ∞. Puesto que $f(\infty)$ no está definido, tratamos ∞ como si fuera una singularidad aislada, y tiene, por convenio, el mismo carácter de singularidad evitable, polo o singularidad esencial que la singularidad de $g(z) = f(1/z)$ en $z = 0$. Si la singularidad es no esencial, $f(z)$ tiene un orden algebraico h tal que $\lim_{z \to \infty} z^{-h} f(z)$ no es ni cero ni infinito, y para un polo, la parte singular es un polinomio en z. Si ∞ es una singularidad esencial, la función tiene la propiedad expresada en el teorema 9 en todo entorno de infinito.

EJERCICIOS

1. Si $f(z)$ y $g(z)$ tienen órdenes algebraicos h y k, respectivamente, en $z = a$, pruébese que fg tiene orden $h+k$, f/g tiene orden $h-k$ y $f+g$ un orden que no excede al máx (h, k).

2. Pruébese que una función que es analítica en todo el plano y tiene una singularidad no esencial en infinito se reduce a un polinomio.

3. Demuéstrese que las funciones e^z, $\operatorname{sen} z$ y $\cos z$ tienen singularidades esenciales en infinito.

4. Pruébese que cualquier función que es meromorfa en todo el plano completo es racional.

5. Determínense todas las funciones analíticas $f(z)$ que satisfacen una desigualdad de la forma $|f(z)| < A|z|^\alpha$ para todo valor de z; A y α son constantes positivas.

3. *La aplicación local.*—Empezaremos con la demostración de una fórmula general que nos permita determinar el número de ceros de una función analítica. Consideremos una función $f(z)$ analítica en un disco abierto Δ. Sea γ una curva cerrada en Δ, tal que $f(z) \neq 0$ sobre γ. Para mayor sencillez supongamos primero que $f(z)$ tiene únicamente un número finito de ceros en Δ y convengamos en de-

notarlos por $z_1, z_2, \ldots, z_n$, estando cada cero repetido tantas veces como indique su orden.

Por aplicación reiterada del teorema 8 o, mejor dicho, de su consecuencia [3-28], resulta evidente que podemos escribir $f(z) = (z-z_1)(z-z_2)\ldots(z-z_n)g(z)$, siendo $g(z)$ analítica y distinta de cero en Δ. Formando la derivada logarítmica, obtenemos

$$\frac{f'(z)}{f(z)} = \frac{1}{z-z_1} + \frac{1}{z-z_2} + \ldots + \frac{1}{z-z_n} + \frac{g'(z)}{g(z)}$$

para $z \neq z_j$, y particularmente sobre γ. Puesto que $g(z) \neq 0$ en Δ, el teorema de Cauchy da

$$\int_\gamma \frac{g'(z)}{g(z)}\,dz = 0.$$

Recordando la definición de $n(\gamma, z_j)$, se tiene que

$$n(\gamma, z_1) + n(\gamma, z_2) + \ldots + n(\gamma, z_n) = \frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}\,dz. \qquad [3\text{-}29]$$

Esto sigue siendo cierto si $f(z)$ tiene infinitos ceros en Δ. Evidentemente γ está contenida en un disco concéntrico Δ' más pequeño que Δ. A no ser que $f(z)$ sea idénticamente cero, caso que obviamente debe ser excluido, tiene solo un número finito de ceros en Δ'. Esta es una consecuencia obvia del teorema de Bolzano-Weierstrass, pues si hubiera infinitos ceros, deberían tener un punto de acumulación en el cierre de Δ', y esto es imposible. Podemos ahora aplicar [3-29] al disco Δ'. Los ceros en el exterior de Δ' satisfacen a $n(\gamma, z_j) = 0$ y, por consiguiente, no contribuyen a la suma en [3-29]. Hemos demostrado así:

TEOREMA 10. *Sean z_j los ceros de una función $f(z)$ que es analítica en un disco circular Δ y que no se anula idénticamente, estando contado cada cero tantas veces como indica su orden. Para toda curva cerrada γ en Δ que no pase por un cero se verifica*

$$\sum_j n(\gamma, z_j) = \frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}\,dz, \qquad [3\text{-}30]$$

donde la suma tiene solo un número finito de términos distintos de cero.

La función $w = f(z)$ aplica γ sobre una curva cerrada Γ en el plano w, y tenemos que

$$\int_\Gamma \frac{dw}{w} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$$

Así, pues, la fórmula [3-30] tiene la siguiente interpretación:

$$n(\Gamma, 0) = \sum_j n(\gamma, z_j). \qquad [3\text{-}31]$$

La aplicación más sencilla, y a la vez la más útil, es la correspondiente al caso en que se conoce de antemano que cada $n(\gamma, z_j)$ tiene que ser cero o uno. Entonces [3-30] proporciona una fórmula que da el *número total de ceros* encerrados por γ. Este es, evidentemente, el caso cuando γ es una circunferencia.

Sea a un valor complejo arbitrario; apliquemos el teorema 10 a $f(z) - a$. Los ceros de $f(z) - a$ son las raíces de la ecuación $f(z) = a$, y los denotaremos por $z_j(a)$. En lugar de [3-30] obtenemos la fórmula

$$\sum_j n[\gamma, z_j(a)] = \frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)-a}\,dz$$

y [3-31] toma la forma

$$n(\Gamma, a) = \sum_j n[\gamma, z_j(a)].$$

Es necesario suponer que $f(z) \neq a$ sobre γ.

Si a y b están en la misma región de las determinadas por Γ, sabemos que $n(\Gamma, a) = n(\Gamma, b)$ y, por consiguiente, tenemos también $\sum_j n[\gamma, z_j(a)] = \sum_j n[\gamma, z_j(b)]$. Si γ es una circunferencia, se sigue que $f(z)$ toma los valores a y b el mismo número de veces en el interior de γ. El teorema siguiente sobre correspondencia local es una consecuencia inmediata de este resultado.

TEOREMA 11. *Supongamos que $f(z)$ es analítica en z_0, $f(z_0) = w_0$, y que $f(z) - w_0$ tiene un cero de orden n en z_0. Si $\epsilon > 0$ es suficientemente pequeño, en correspondencia con él existe un $\delta > 0$ tal que* *para todo a con* $|a - w_0| < \delta$ *la ecuación* $f(z) = a$ *tiene exactamente* n *raíces en el disco* $|z - z_0| < \epsilon$.

Podemos elegir ϵ de forma que $f(z)$ esté definida y sea analítica para $|z - z_0| \leqq \epsilon$, y tal que z_0 sea el único cero de $f(z) - w_0$ en este disco. Sea γ la circunferencia $|z - z_0| = \epsilon$ y Γ su imagen por la aplicación $w = f(z)$. Puesto que w_0 pertenece al complemento del conjunto cerrado Γ, existe un entorno $|w - w_0| < \delta$ sin parte común con Γ (Fig. 3-6). Se sigue inmediatamente que todos los valores a en este entorno los toma la función el mismo número de veces en el interior de γ. La ecuación $f(z) = w_0$ tiene exactamente n raíces coincidentes en el interior de γ, y cada valor a, por consiguiente, se alcanza n veces. Se entiende que las raíces múltiples se cuentan de acuerdo con su multiplicidad, pero si ϵ es suficientemente pequeño, podemos afirmar que todas las raíces de la ecuación $f(z) = a$ son simples para $a \neq w_0$. En efecto, basta elegir ϵ de forma que $f'(z)$ no se anule para $0 < |z - z_0| < \epsilon$.

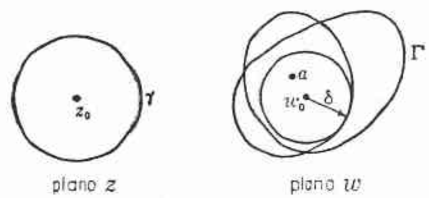

FIG. 3-6.—Correspondencia local.

COROLARIO 1. *Una función analítica no constante aplica conjuntos abiertos sobre conjuntos abiertos.*

Esto es sencillamente otra forma de decir que la imagen de todo disco suficientemente pequeño $|z - z_0| < \epsilon$ contiene un entorno $|w - w_0| < \delta$.

En el caso $n = 1$ existe una correspondencia biunívoca entre el disco $|w - w_0| < \delta$ y un subconjunto abierto Δ de $|z - z_0| < \epsilon$. Puesto que conjuntos abiertos en el plano z corresponden a conjuntos abiertos en el plano w, la función inversa de $f(z)$ es continua, y la aplicación, topológica. La aplicación puede restringirse a un entorno de z_0 contenido en Δ, y podemos entonces enunciar:

COROLARIO 2. *Si* $f(z)$ *es analítica en* z_0 *con* $f'(z_0) \neq 0$, *aplica conforme y topológicamente un entorno de* z_0 *sobre una región.*

De la continuidad de la función inversa se sigue, en la forma usual, que la función inversa es analítica y, por consiguiente, que la aplicación inversa es igualmente conforme. Recíprocamente, si la aplicación local es biunívoca, se verifica el teorema 11 con $n = 1$; por tanto, $f'(z_0)$ tiene que ser diferente de cero.

Para $n > 1$ la correspondencia local puede todavía describirse en términos muy precisos. Con las hipótesis del teorema 11 podemos escribir

$$f(z) - w_0 = (z - z_0)^n g(z),$$

siendo $g(z)$ analítica en z_0 y $g(z_0) \neq 0$. Elijamos $\epsilon > 0$ tal que $|g(z) - g(z_0)| < |g(z_0)|$ para $|z - z_0| < \epsilon$. En este entorno es posible definir una rama analítica uniforme de $\sqrt[n]{g(z)}$, que denotaremos por $h(z)$. Tenemos así

$$f(z) - w_0 = \zeta(z)^n$$
$$\zeta(z) = (z - z_0) h(z).$$

Puesto que $\zeta'(z_0) = h(z_0) \neq 0$, la aplicación $\zeta = \zeta(z)$ es topológica en un entorno de z_0. Por otra parte, la aplicación $w = w_0 + \zeta^n$ es de carácter elemental y determina n valores ζ igualmente espaciados para cada valor de w. Efectuando en dos pasos la aplicación obtenemos una imagen bastante ilustrativa de la correspondencia local. La figura 3-7 muestra la imagen inversa de un pequeño disco circular y los n arcos que se aplican sobre el radio positivo.

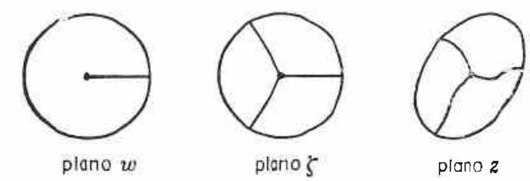

FIG. 3-7.—Punto de ramificación: $n = 3$.

EJERCICIOS

1. Determínese explícitamente el mayor de los discos con centro en el origen, en el que la aplicación $w = z^2 + z$ sea biunívoca.

2. El mismo problema para $w = e^z$.

3. Ilústrese mediante un dibujo la correspondencia local respecto de la aplicación $w = \cos z$ en un entorno del origen.

4. *El principio del módulo máximo.*—El corolario 1 del teorema 11 tiene una consecuencia analítica muy importante, conocida como el principio del módulo máximo para funciones analíticas.

Debido a su formulación explícita y sencilla constituye uno de los teoremas generales más útiles en la teoría de funciones. Por regla general, todas las demostraciones basadas en el principio del módulo máximo son muy directas, por lo que está perfectamente justificado el que sean preferidas a otras.

TEOREMA 12. *(Principio del módulo máximo.) Si $f(z)$ es analítica y no constante en una región Ω, entonces su módulo $|f(z)|$ no alcanza un máximo en Ω.*

La demostración es sencilla. Si $w_0 = f(z_0)$ es un valor cualquiera de los que toma en Ω, existe un entorno $|w - w_0| < \epsilon$ contenido en la imagen de Ω. En este entorno hay puntos de módulo $> |w_0|$ y, por consiguiente, $|f(z_0)|$ no es el máximo de $|f(z)|$.

Con una formulación positiva, el mismo teorema puede esencialmente enunciarse en la forma:

TEOREMA 12'. *Si $f(z)$ es analítica en un conjunto cerrado y acotado E, el máximo de $|f(z)|$ se alcanza en la frontera de E.*

Puesto que E es compacto, $|f(z)|$ tiene un máximo en E. Si $f(z)$ es constante, la afirmación del teorema es trivialmente cierta, pues la frontera de E no es vacía. Supongamos que se alcanza el máximo en un punto interior z_0. Entonces $|f(z_0)|$ sería también el máximo de $|f(z)|$ en un entorno $|z - z_0| < \delta$ contenido en E. Pero esto es imposible, a no ser que $f(z)$ sea constante en este entorno, y entonces $f(z)$ sería constante en toda la región de definición [recordemos que $f(z)$ es analítica en E si está definida y es analítica en una *región* que contenga a E]. Así, pues, el máximo del módulo se alcanza siempre en un punto frontera.

El principio del módulo máximo puede también probarse analíticamente como una consecuencia de la fórmula de la integral de Cauchy. Si se especializa la fórmula [3-21] al caso en que γ es una circunferencia de centro z_0 y radio r, podemos escribir $\zeta = z_0 + re^{i\theta}$, $d\zeta = ire^{i\theta} d\theta$ sobre γ, y obtenemos para $z = z_0$

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta})\, d\theta. \qquad [3\text{-}32]$$

Esta fórmula prueba que el valor de una función analítica en el centro de un círculo es igual a la media aritmética de sus valores sobre la circunferencia del mismo, siempre que el disco cerrado $|z - z_0| \leqq r$ esté contenido en la región de analiticidad.

De [3-32] se deduce la desigualdad

$$|f(z_0)| \leqq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + re^{i\theta})|\, d\theta. \qquad [3\text{-}33]$$

Supongamos que $|f(z_0)|$ sea un máximo. En tal caso tendríamos $|f(z_0 + re^{i\theta})| \leqq |f(z_0)|$, y si se verifica la desigualdad estricta para un solo valor de θ, entonces se verificará, por continuidad, en todo un arco. Pero de este modo el valor medio de $|f(z_0 + re^{i\theta})|$ sería estrictamente menor que $|f(z_0)|$ y [3-33] nos conduciría a la contradicción $|f(z_0)| < |f(z_0)|$. Así, pues, $|f(z)|$ debe ser constantemente igual a $|f(z_0)|$ en todas las circunferencias $|z - z_0| = r$ suficientemente pequeñas y, por consiguiente, en un entorno de z_0. Se sigue fácilmente que $f(z)$ tiene que reducirse a una constante. Este razonamiento proporciona una segunda demostración del principio del módulo máximo. Hemos dado preferencia a la primera demostración, ya que muestra que el principio del módulo máximo es una consecuencia de las propiedades topológicas de la aplicación mediante una función analítica.

Volviendo a la formulación dada en el teorema 12', observamos que la aplicación más natural es la correspondiente al caso en que E es una región cerrada. Resulta así que una función que es analítica en una región cerrada toma su módulo máximo en la frontera de esta región. Para algunas aplicaciones es importante observar que las hipótesis del teorema pueden debilitarse. En efecto, es suficiente suponer que $f(z)$ es continua en la región cerrada y analítica en la abierta. La continuidad en la región acotada y cerrada asegura la existencia del máximo, y la analiticidad en la región abierta implica que el máximo no puede ser alcanzado en un punto interior, a no ser que la función se reduzca a una constante.

Consideremos ahora el caso de una función $f(z)$ que es analítica en el disco abierto $|z| < R$ y continua en el disco cerrado $|z| \leqq R$. Si se sabe que $|f(z)| \leqq M$ en $|z| = R$, entonces $|f(z)| \leqq M$ en todo el disco, de acuerdo con la observación anterior. La igualdad puede únicamente verificarse si $f(z)$ es una constante de valor absoluto M. Por consiguiente, si se sabe que $f(z)$ toma algún valor de módulo menor que M, cabe esperar que pueda darse una acotación mejor. Son muy útiles los teoremas relacionados con esta cuestión. El resultado particular siguiente se conoce con el nombre de *lema de Schwarz:*

TEOREMA 13. *Si $f(z)$ es analítica para $|z| < 1$ y satisface las condiciones $|f(z)| \leqq 1$, $f(0) = 0$, entonces $|f(z)| \leqq |z|$ y $|f'(0)| \leqq 1$. Uni-*

camente se verifica la igualdad si $f(z)=cz$, siendo c una constante de valor absoluto igual a uno.

Aplicamos el principio del módulo máximo a la función $f_1(z)$, que es igual a $f(z)/z$ para $z \neq 0$ y a $f'(0)$ para $z=0$. En la circunferencia $|z|=r<1$ su valor absoluto es $\leqq 1/r$; por consiguiente, $|f_1(z)| \leqq 1/r$ para $|z| \leqq r$. Haciendo que r tienda a 1 tenemos que $|f_1(z)| \leqq 1$ para todo z, y esta es la afirmación del teorema. Si se verifica la igualdad en un solo punto, significa que $|f_1(z)|$ alcanza su máximo y, por tanto, que $f_1(z)$ debe reducirse a una constante.

Las hipótesis algo especializadas del teorema 13 no son esenciales, pero deben considerarse como el resultado de una normalización. Así, p. ej., si se sabe que $f(z)$ satisface las condiciones del teorema en un disco de radio R, puede aplicarse la forma original del teorema a la función $f(Rz)$. Obtenemos como resultado $|f(Rz)| \leqq |z|$, que puede escribirse nuevamente en la forma $|f(z)| \leqq |z|/R$. Análogamente, si la cota superior del módulo es M en lugar de 1, aplicamos el teorema a $f(z)/M$ o, en el caso más general, a $f(Rz)/M$. La desigualdad resultante es $|f(z)| \leqq M|z|/R$.

Con más generalidad todavía, podemos reemplazar la condición $f(0)=0$ por una condición arbitraria $f(z_0)=w_0$, siendo $|z_0|<R$ y $|w_0|<M$. Sea $\zeta = Tz$ una transformación lineal que aplica $|z|<R$ sobre $|\zeta|<1$, siendo el homólogo de z_0 el origen, y sea Sw una transformación lineal con $Sw_0=0$ que aplica $|w|<M$ sobre $|Sw|<1$. Evidentemente, la función $Sf(T^{-1}\zeta)$ satisface las hipótesis del teorema original. Obtenemos, por tanto, $|Sf(T^{-1}\zeta)| \leqq |\zeta|$ o $|Sf(z)| \leqq |Tz|$. Esta desigualdad puede escribirse explícitamente en la forma

$$\left| \frac{M[f(z)-w_0]}{M^2-\bar{w}_0 f(z)} \right| \leqq \left| \frac{R(z-z_0)}{R^2-\bar{z}_0 z} \right|. \qquad [3\text{-}34]$$

EJERCICIOS

1. Pruébese que $|f(z)| \leqq 1$ para $|z|<1$ implica $|f'(0)| \leqq 1$ independientemente del valor de $f(0)$.

2. Hállese la acotación de $|f'(z_0)|$ correspondiente a [3-34].

3. Dedúzcase una desigualdad similar a [3-34] para el caso en que $f(z)$ satisfaga $\operatorname{Im} f(z) \geqq 0$ para $\operatorname{Im} z > 0$.

4. Sea $f(z)$ analítica en la corona circular $r_1 \leqq |z| \leqq r_2$ y supongamos que $|f(z)| \leqq M_1$ en $|z|=r_1$ y $|f(z)| \leqq M_2$ en $|z|=r_2$. Pruébese que, para $r_1 \leqq r \leqq r_2$, el máximo de $|f(z)|$ en $|z|=r$ es a lo sumo igual a

$$M_1^{(\log r_2/r)/(\log r_2/r_1)} \cdot M_2^{(\log r/r_1)/(\log r_2/r_1)}.$$

Sugerencia: Aplíquese el principio del módulo máximo a la función $z^p f(z)^q$ para p y q enteros.

5. Pruébese mediante el lema de Schwarz que toda aplicación conforme biunívoca de un disco circular sobre otro viene dada por una transformación lineal.

3-4. Forma general del teorema de Cauchy.—En nuestro tratamiento previo del teorema de Cauchy y de la fórmula de la integral solo hemos considerado el caso de una región circular. Para efectuar el estudio de las propiedades locales de las funciones analíticas esto resultaba suficiente, pero desde un punto de vista más general no podemos quedar satisfechos con un resultado que obviamente es tan incompleto. La generalización puede encaminarse en dos direcciones. Por una parte, cabe investigar la caracterización de las regiones en las que el teorema de Cauchy tiene validez universal; por otra, podemos considerar una región arbitraria y buscar las curvas γ para las que la afirmación del teorema de Cauchy es cierta.

1. *Cadenas y ciclos.*—Debemos generalizar en primer lugar el concepto de integral curvilínea. Con este objeto examinemos la ecuación

$$\int_{\gamma_1+\gamma_2+\ldots+\gamma_n} f\,dz = \int_{\gamma_1} f\,dz + \int_{\gamma_2} f\,dz + \ldots + \int_{\gamma_n} f\,dz, \qquad [3\text{-}35]$$

que es válida cuando $\gamma_1, \gamma_2, \ldots, \gamma_n$ forman una subdivisión del arco γ. Puesto que el segundo miembro de [3-35] tiene un significado para cualquier colección finita, nada nos impide considerar una suma formal arbitraria $\gamma_1+\gamma_2+\ldots+\gamma_n$, que no sea necesariamente un arco, y definir la correspondiente integral mediante la ecuación [3-35]. Se llama *cadenas* a tales sumas formales de arcos. Está claro que nada se ha perdido y mucho puede ganarse al considerar integrales curvilíneas sobre cadenas arbitrarias.

Al igual que la manera de subdividir un arco no es única, evidentemente diferentes sumas formales pueden representar la misma cadena. El principio que nos ha de guiar es el de que dos cadenas deberían considerarse idénticas si dan las mismas integrales curvilíneas para todas las funciones f. Si se analiza este principio, se halla que las operaciones siguientes no cambian la identidad de una cadena: 1) permutación de dos arcos; 2) subdivisión de un arco; 3) fusión de arcos parciales en un solo arco; 4) nueva parametrización de un arco; 5) cancelación de arcos opuestos. Sobre esta base sería fácil formular una relación de equivalencia lógica que definiese la identidad de cadenas de una manera formal. Puesto que la situación no implica ninguna trampa lógica, omitiremos esta formalización.

Por yuxtaposición, la suma de dos cadenas se define de manera obvia. Evidentemente, la propiedad aditiva [3-35] de las integrales curvilíneas sigue siendo válida para cadenas arbitrarias. Cuando se suman cadenas idénticas es conveniente denotar la suma en forma de múltiplo. Con esta notación toda cadena puede escribirse en la forma

$$\gamma = a_1\gamma_1 + a_2\gamma_2 + \dots + a_n\gamma_n, \qquad [3\text{-}36]$$

donde los a_j son enteros positivos y los γ_j son todos diferentes. Para arcos opuestos podemos escribir $a(-\gamma) = -a\gamma$ y continuar la reducción en [3-36] hasta que no haya dos γ_j opuestos. Los coeficientes son enteros arbitrarios y se pueden añadir, a capricho, términos con coeficientes cero. El último artificio nos permite expresar dos cadenas cualesquiera en términos de los mismos arcos, y su suma se obtiene sumando los coeficientes correspondientes. La cadena cero es, o una suma vacía, o una suma con todos los coeficientes iguales a cero.

Una cadena es un *ciclo* si puede representarse como suma de curvas cerradas. Consideraciones de tipo combinatorio muy sencillas muestran que una cadena es un ciclo si (y solo si) en cualquier representación los puntos origen y extremo de los arcos individuales son idénticos a pares. Por tanto, es posible decidir inmediatamente si una cadena es un ciclo o no.

En las aplicaciones consideraremos cadenas que están contenidas en una región Ω dada. Con esto queremos decir que las cadenas admiten una representación mediante arcos en Ω y que solamente utilizaremos tales representaciones. Está claro que todos los teoremas que hemos formulado hasta ahora, únicamente para curvas cerradas en una región, son realmente válidos para ciclos cualesquiera en una región. En particular, *la integral de una diferencial exacta sobre cualquier ciclo es cero.*

Se define el índice de un punto con respecto a un ciclo exactamente en la misma forma que en el caso de una sola curva cerrada. Tiene las mismas propiedades y además podemos formular la ley aditiva, obvia pero muy importante, expresada por la ecuación $n(\gamma_1+\gamma_2, a) = n(\gamma_1, a) + n(\gamma_2, a)$.

2. *Conexión simple.*—No hay la menor duda que todos los lectores sabrán a qué nos referimos al hablar de una región "sin agujeros". A tales regiones se les llama *simplemente conexas,* y es para regiones simplemente conexas para las que el teorema de Cauchy tiene validez universal. El lenguaje sugerente que hemos empleado no puede sustituir a una definición matemática, pero afortunadamente se necesita muy poco para hacer que el término sea por completo preciso. En efecto, una región sin agujeros es obviamente aquella cuyo complemento consta de una sola pieza. Venimos a parar así a la siguiente definición:

DEFINICIÓN 1. *Una región es simplemente conexa si su complemento con respecto al plano completo es conexo.*

Es fácil ver que un disco circular, un semiplano y una banda paralela son simplemente conexos. El último ejemplo muestra la importancia de tomar el complemento con respecto al plano completo, pues el complemento de la banda en el plano finito evidentemente no es conexo. La definición puede aplicarse a regiones sobre la esfera de Riemann, y esta es evidentemente la situación más simétrica. Sin embargo, para nuestros propósitos es mejor convenir que todas las regiones se encuentran en el plano finito, a no ser que explícitamente se diga lo contrario. De acuerdo con este convenio, el exterior de un círculo no es simplemente conexo, ya que su complemento consiste en un disco cerrado y el punto del infinito.

TEOREMA 14. *Una región Ω es simplemente conexa si (y solo si) $n(\gamma, a) = 0$ para todos los ciclos γ en Ω y todos los puntos a que no pertenezcan a Ω.*

Esta nueva condición es también muy sugerente. Afirma que una curva cerrada en una región simplemente conexa no puede rodear a ningún punto que no pertenezca a la región. Parece bastante evidente que no se cumple esta condición en el caso de una región con un agujero.

La necesidad de la condición es casi trivial. Sea γ un ciclo cualquiera en Ω. Si el complemento de Ω es conexo, debe estar contenido en una de las regiones determinadas por γ, y puesto que ∞ pertenece al complemento, este tiene que ser una región no acotada. Por consiguiente, $n(\gamma, a) = 0$ para todos los puntos finitos del complemento.

Para la demostración precisa de la suficiencia es necesaria una construcción explícita. Suponemos que el complemento de Ω puede representarse como la unión $A+B$ de dos conjuntos cerrados disjuntos. Uno de estos conjuntos contiene a ∞, y el otro es, por tanto, acotado; sea A el conjunto acotado. Los conjuntos A y B tienen una mínima distancia $\delta > 0$. Recubramos todo el plano con una red de cuadrados Q de lado $< \delta/\sqrt{2}$. Tenemos libertad para elegir la red de forma que un cierto punto $a \in A$ esté en el centro de un cuadrado. La curva frontera de Q se denotará por $\Gamma(Q)$; supondremos explícitamente que los cuadrados Q son cerrados y que el interior de Q queda a la izquierda de los segmentos orientados que constituyen $\Gamma(Q)$.

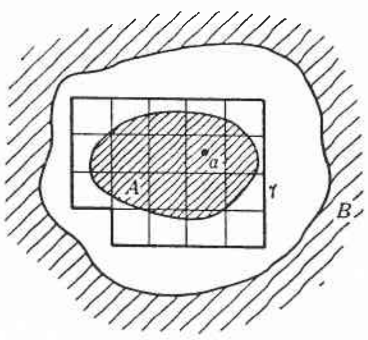

FIG. 3-8.

Consideremos ahora el ciclo

$$\gamma = \sum_j \Gamma(Q_j) \tag{3-37}$$

estando extendidas las sumas a todos los cuadrados Q_j de la red que tienen un punto común con A (Fig. 3-8). Puesto que a está contenido en uno (y solo en uno) de estos cuadrados, es evidente que $n(\gamma, a) = 1$. Es obvio además que γ no tiene parte común con B. Pero si se efectúan las oportunas cancelaciones, es también claro que γ no tiene parte común con A. En efecto, cualquier lado que corte a A es un lado común a dos cuadrados incluidos en la suma [3-37], y puesto que las orientaciones son opuestas, el lado no aparece en la expresión reducida de γ. Por tanto, γ está contenida en Ω, con lo que queda demostrado el teorema.

Observemos ahora que el teorema de Cauchy no es válido ciertamente para regiones que no son simplemente conexas. En efecto, si hay un ciclo γ en Ω, tal que $n(\gamma, a) \neq 0$ para algún a exterior a Ω, entonces $1/(z-a)$ es analítica en Ω, mientras que su integral

$$\int_\gamma \frac{dz}{z-a} = 2\pi i n(\gamma, a) \neq 0.$$

3. *Diferenciales exactas en regiones simplemente conexas.*—Demostraremos ahora que el teorema de Cauchy se verifica en una región simplemente conexa arbitraria. Como consecuencia del teorema 1, únicamente necesitamos probar que $f(z)dz$ es una diferencial exacta. Se sabe que este es el caso cuando se trata de un disco circular contenido en Ω. Por tanto, hemos de probar que una diferencial que es exacta en un entorno de cada punto es también exacta en toda la región Ω, con tal que Ω sea simplemente conexa. En esta forma, la afirmación no está limitada a diferenciales de la forma $f\,dz$, y hemos preferido demostrar la siguiente proposición:

TEOREMA 15. *La diferencial $p\,dx + q\,dy$, cuyos coeficientes están definidos y son continuos en una región simplemente conexa Ω, es exacta en Ω si (y solo si) se verifica*

$$\int_{\Gamma(R)} p\,dx + q\,dy = 0$$

para todo rectángulo R contenido en Ω.

Elegimos un punto $z_0 \in \Omega$ y definimos $U(z)$ mediante

$$U(z) = \int_\sigma p\,dx + q\,dy,$$

donde σ es una línea poligonal de z_0 a z, contenida en Ω y de lados paralelos a los ejes. Si podemos probar que $U(z)$ es independiente de la elección de σ, de la misma forma que en la demostración del teorema 1 se deduce que $\partial U/\partial x = p$, $\partial U/\partial y = q$. La diferencia $\gamma = \sigma_1 - \sigma_2$ de dos líneas poligonales de z_0 a z es una línea poligonal cerrada en Ω de lados verticales y horizontales, y tenemos que probar que la integral de $p\,dx + q\,dy$ sobre γ es cero.

Para la demostración utilizaremos una red rectangular obtenida mediante rectas paralelas a ambos ejes que pasen por todos los vértices de γ (Fig. 3-9). Habrá algunos rectángulos finitos R_i y algunas regiones no acotadas R'_j que pueden considerarse como rectángulos infinitos. Puesto que no es preciso tener en cuenta el caso trivial de que γ sea solo un segmento horizontal o vertical, podemos suponer que hay al menos un rectángulo finito R_i.

Elijamos un punto a_i del interior de cada R_i y formemos el ciclo

$$\gamma_0 = \sum_i n(\gamma, a_i)\Gamma(R_i), \tag{3-38}$$

donde la suma está extendida a todos los rectángulos finitos; los coeficiente $n(\gamma, a_i)$ están perfectamente determinados, pues ningún a_i puede estar en γ. En la discusión que sigue haremos también uso de puntos a'_j elegidos del interior de cada R'_j, aunque no son necesarios para la construcción de γ_0.

La elección de γ_0, definida por [3-38], está dictada por un propósito definido. Es evidente que $n[\Gamma(R_i), a_k]=1$ si $k=i$, y cero si $k \neq i$; análogamente, $n[\Gamma(R_i), a'_j]=0$ para todo j. Teniendo esto en cuenta, en virtud de [3-38] tenemos que $n(\gamma_0, a_i)=n(\gamma, a_i)$ y

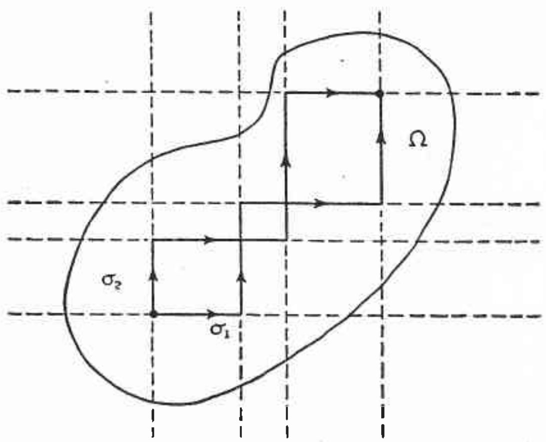

FIG. 3-9.

$n(\gamma_0, a'_j)=0$. Es también cierto que $n(\gamma, a'_j)=0$, pues evidentemente el interior de R'_j debe pertenecer a la región no acotada determinada por γ. Hemos probado así que $n(\gamma-\gamma_0, a)=0$ para todo $a=a_i$ y $a=a'_j$.

De esta propiedad de $\gamma-\gamma_0$ deseamos deducir que γ_0 es idéntico a γ. Sea σ_{ik} el lado común de dos rectángulos adyacentes R_i, R_k; elegimos la orientación de forma que R_i quede a la izquierda de σ_{ik}. Supongamos que la expresión reducida de $\gamma-\gamma_0$ contiene el múltiplo $c\sigma_{ik}$. Entonces el ciclo $\gamma-\gamma_0-c\Gamma(R_i)$ no contiene a σ_{ik}, y se sigue que a_i y a_k han de tener el mismo índice con respecto a este ciclo. Por otra parte, los índices respectivos son, evidentemente, $-c$ y 0; en definitiva, se tiene que $c=0$. Se aplica el mismo razonamiento si σ_{ij} es el lado común de un rectángulo finito R_i y de uno infinito R'_j. Habiendo al menos un rectángulo finito, es claro que dos rectángulos infinitos no pueden tener un lado finito común, y hemos demostrado de forma concluyente que $\gamma-\gamma_0$ debe ser idénticamente cero. Esto significa que γ puede representarse de la siguiente forma:

$$\gamma=\sum_i n(\gamma, a_i)\Gamma(R_i). \qquad [3\text{-}39]$$

Probaremos ahora que la representación [3-39] es una representación *en* Ω; con más precisión, probaremos que los rectángulos R_i, cuyos coeficientes correspondientes $n(\gamma, a_i)$ son diferentes de cero, están contenidos en Ω. Esto es una consecuencia de ser Ω simplemente conexo. En efecto, supongamos que un punto a en el rectángulo cerrado R_i no estuviera en Ω. Entonces $n(\gamma, a)=0$, de acuerdo con el teorema 14. Por otra parte, el segmento determinado por a y a_i no corta a γ; por consiguiente, $n(\gamma, a_i)=n(\gamma, a)$, en contradicción con la hipótesis de que $n(\gamma, a_i)\neq 0$.

Se sigue que la hipótesis del teorema 15 es de aplicación a cada $\Gamma(R_i)$, lo que efectivamente ocurre en [3-39]. Tenemos, por tanto, que

$$\int_\gamma p\,dx+q\,dx=0,$$

con lo que queda probado el teorema.

La consecuencia más importante es el teorema de Cauchy para regiones simplemente conexas:

TEOREMA 16. *Si $f(z)$ es analítica en una región simplemente conexa Ω, se tiene entonces que*

$$\int_\gamma f(z)\,dz=0$$

para todo ciclo γ en Ω.

El corolario siguiente es de frecuente utilización:

COROLARIO. *Si $f(z)$ es analítica y distinta de cero en una región simplemente conexa Ω, se puede definir en Ω una rama analítica uniforme de $\log f(z)$.*

Por el teorema anterior la función $f'(z)/f(z)$ tiene una integral indefinida $F(z)$ en Ω. La función $f(z)e^{-F(z)}$ tiene entonces derivada cero y debe reducirse a una constante. Eligiendo un punto $z_0\in\Omega$ y un valor arbitrario de $\log f(z_0)$, tenemos que

$$e^{F(z)-F(z_0)+\log f(z_0)}=f(z),$$

y, por tanto, es posible escribir $\log f(z)=F(z)-F(z_0)+\log f(z_0)$.

En las mismas circunstancias podemos también determinar una rama uniforme de una potencia arbitraria $f(z)^\mu=e^{\mu\log f(z)}$.

4. *Regiones múltiplemente conexas.*—La demostración del teorema 15 incluye más de lo que ha sido enunciado. En realidad, la

hipótesis de conexión simple se utilizó únicamente para obtener que $n(\gamma, a)=0$ para todos los puntos a no pertenecientes a Ω. Para líneas poligonales particulares esta condición pudiera muy bien verificarse aun cuando la región Ω no fuera simplemente conexa. Hemos demostrado, en efecto, que en una región arbitraria se verifica

$$\int_\gamma p\,dx + q\,dy = 0$$

para cualquier línea poligonal cerrada γ que no rodee puntos exteriores; parece, pues, razonable esperar que este resultado no esté limitado solo a líneas poligonales. Sin embargo, la transición de una línea poligonal de lados verticales y horizontales a un ciclo arbitrario debe basarse en un razonamiento diferente, pues ya no podemos estar seguros de que la diferencial sea exacta.

Antes de seguir con esta demostración es conveniente dar un nombre al tipo de ciclo que vamos a considerar.

DEFINICIÓN 2. *Un ciclo γ en Ω se dice homológico a cero con respecto a Ω si (y solo si) $n(\gamma, a)=0$ para todos los puntos a no pertenecientes a Ω.*

Escribiremos simbólicamente $\gamma \sim 0$ (mod. Ω). Cuando esté claro a qué región nos referimos, no es necesario que se mencione Ω de manera explícita. La notación $\gamma_1 \sim \gamma_2$ es equivalente a la $\gamma_1 - \gamma_2 \sim 0$. Evidentemente, se pueden sumar y restar homologías. Por otra parte, $\gamma \sim 0$ (mod. Ω) implica $\gamma \sim 0$ (mod. Ω') para todo $\Omega' \supset \Omega$.

TEOREMA 17. *Si $p\,dx + q\,dy$ es exacta localmente en Ω, es decir si se verifica*

$$\int_\gamma p\,dx + q\,dy = 0 \qquad [3\text{-}40]$$

para $\gamma = \Gamma(R)$, siendo R cualquier rectángulo contenido en Ω, entonces se verifica [3-40] *para todo ciclo γ que sea homológico a cero en Ω.*

Para demostrarlo basta probar que puede reemplazarse γ por una línea poligonal σ de lados verticales y horizontales, tal que toda diferencial exacta localmente tenga la misma integral sobre γ y sobre σ. Esta propiedad implica, en particular, que $n(\sigma, a) = n(\gamma, a)$ y, por consiguiente, que $\sigma \sim 0$. La demostración del teorema 15 se aplica a σ, y podemos concluir que γ satisface la condición [3-40].

Construimos σ como aproximación de γ. Sea $\geqq \rho > 0$ la distancia de γ al complemento de Ω. Si la ecuación de γ es $z = z(t)$, $a \leqq t \leqq b$, la función $z(t)$ es uniformemente continua en el intervalo cerrado (a, b). Determinamos $\delta > 0$ tal que $|z(t) - z(t')| < \rho$ para $|t - t'| < \delta$ y dividimos (a, b) en subintervalos de longitud $< \delta$. Los arcos parciales correspondientes, que denotaremos por γ_i, tienen la propiedad de que cada γ_i está contenido en un disco de radio ρ totalmente contenido en Ω. Los extremos de γ_i pueden unirse dentro del mismo disco mediante una línea poligonal σ_i, que consiste en segmentos verticales y horizontales. Puesto que nuestra diferencial es exacta en el disco, se tiene que

$$\int_{\sigma_i} p\,dx + q\,dy = \int_{\gamma_i} p\,dx + q\,dy,$$

y poniendo $\sigma = \Sigma\, \sigma_i$, obtenemos

$$\int_\sigma p\,dx + q\,dy = \int_\gamma p\,dx + q\,dy;$$

lo que completa la demostración.

El teorema 17 caracteriza los ciclos que son homológicos a cero. Hemos probado, en efecto, que si la integral sobre γ se anula para todas las diferenciales de la forma especial $dz/(z-a)$ con a exterior a Ω, entonces se anula para todas las diferenciales localmente exactas [1]. En particular se anula para las diferenciales de la forma $f(z)\,dz$, siendo $f(z)$ analítica en Ω. Esta es la forma final y más completa del teorema de Cauchy.

TEOREMA 18. *Si $f(z)$ es analítica en Ω, entonces se verifica*

$$\int_\gamma f(z)\,dz = 0$$

para todo ciclo γ que sea homológico a cero en Ω.

Una región que no es simplemente conexa se llama múltiplemente conexa. Con más precisión: se dice que Ω es de orden de conexión finito n si el complemento de Ω tiene exactamente n componen-

[1] Como en las hipótesis del teorema 17, se dice que una diferencial es *localmente exacta* en una región Ω si es exacta en algún entorno de cada punto de Ω.

tes, y de orden de conexión infinito si el complemento tiene infinitas componentes. En un lenguaje menos preciso, pero más sugerente, una región de orden n de conexión se obtiene haciendo n agujeros en la esfera de Riemann.

En el caso de orden de conexión finito, sean $A_1, A_2, \ldots, A_n$ las componentes del complemento de Ω y supongamos que ∞ pertenece a A_n. Si γ es un ciclo arbitrario en Ω, podemos probar, lo mismo que en el teorema 14, que $n(\gamma, a)$ es constante cuando a varía sobre una cualquiera de las componentes A_i y que $n(\gamma, a)=0$ en A_n. Además, reproduciendo la construcción utilizada en la demostración del mismo teorema, podemos hallar ciclos γ_i, $i=1, \ldots, n-1$, tales que $n(\gamma_i, a)=1$ para $a \in A_i$ y $n(\gamma_i, a)=0$ para todos los demás puntos exteriores a Ω.

Para un ciclo dado γ en Ω, sea c_i el valor constante de $n(\gamma, a)$ para $a \in A_i$. Se tiene que cualquier punto exterior a Ω tiene índice cero con respecto al ciclo $\gamma - c_1\gamma_1 - c_2\gamma_2 - \ldots - c_{n-1}\gamma_{n-1}$. En otras palabras: se verifica

$$\gamma \sim c_1\gamma_1 + c_2\gamma_2 + \ldots + c_{n-1}\gamma_{n-1}.$$

Así, pues, todo ciclo es homológico a una combinación lineal de los ciclos $\gamma_1, \gamma_2, \ldots, \gamma_{n-1}$. Esta combinación lineal está determinada de manera única, pues si dos combinaciones lineales fueran homológicas al mismo ciclo, su diferencia sería una combinación lineal homológica a cero. Pero es claro que el ciclo $c_1\gamma_1 + c_2\gamma_2 + \ldots + c_{n-1}\gamma_{n-1}$ rodea c_i veces a los puntos A_i; por tanto, no puede ser homológico a cero a no ser que todos los c_i se anulen.

En vista de estas circunstancias, los ciclos $\gamma_1, \gamma_2, \ldots, \gamma_{n-1}$ se dice que forman una *base de homología* para la región Ω. No es la única base de homología, pero por un teorema elemental de álgebra lineal se puede concluir que todas las bases de homología tienen el mismo número de elementos. Se tiene que toda región con una base de homología finita tiene conexión finita, y que el número de elementos de la base es uno menos que la conexión.

Por el teorema 18 obtenemos, para cualquier función analítica $f(z)$ en Ω,

$$\int_\gamma f\,dz = c_1 \int_{\gamma_1} f\,dz + c_2 \int_{\gamma_2} f\,dz + \ldots + c_{n-1} \int_{\gamma_{n-1}} f\,dz.$$

Los números

$$P_i = \int_{\gamma_i} f\,dz$$

dependen únicamente de la función y no de γ. Se les llama *módulos de periodicidad* de la diferencial $f\,dz$ o, con menos precisión, los *períodos* de la integral indefinida. Hemos hallado que la integral de $f(z)$ sobre cualquier ciclo es una combinación lineal de los períodos con coeficientes enteros, y que la integral a lo largo de un arco desde z_0 a z está determinada salvo múltiplos aditivos de los períodos. La anulación de los períodos es una condición necesaria y suficiente para la existencia de una integral indefinida uniforme.

Con objeto de ilustrar esto consideremos el caso extremadamente sencillo de una corona circular, definida por $r_1 < |z| < r_2$. El complemento tiene las componentes $|z| \leqq r_1$ y $|z| \geqq r_2$; incluimos los casos degenerados $r_1 = 0$ y $r_2 = \infty$. La corona es doblemente conexa, y una base de homología está formada por cualquier circunferencia $|z| = r$, $r_1 < r < r_2$. Si se denota esta circunferencia por C, cualquier ciclo en la corona satisface la relación $\gamma \sim nC$, siendo $n = n(\gamma, 0)$. La integral de cualquier función analítica sobre un ciclo es un múltiplo del período único

$$P = \int_C f\,dz,$$

cuyo valor es naturalmente independiente del radio r.

EJERCICIOS

1. Pruébese que la región obtenida de una región simplemente conexa quitando m puntos es de orden de conexión $m+1$, y hállese una base de homología.

2. Demuéstrese que las regiones acotadas determinadas por una curva cerrada son simplemente conexas, mientras que la región no acotada es doblemente conexa.

3. Pruébese que pueden definirse ramas analíticas uniformes de $\log z$, z^α y z^z en cualquier región simplemente conexa que no contenga el origen.

4. Demuéstrese que puede definirse una rama analítica uniforme de $\sqrt{1-z^2}$ en cualquier región tal que los puntos ± 1 estén en la misma componente del complemento. ¿Cuáles son los valores posibles de

$$\int \frac{dz}{\sqrt{1-z^2}}$$

sobre una curva cerrada en la región?

3-5. El cálculo de residuos.—Los resultados de la sección precedente han mostrado que la determinación de integrales curvilíneas de funciones analíticas sobre curvas cerradas puede reducirse a la determinación de períodos. Bajo ciertas circunstancias resulta que pueden hallarse los períodos con muy poco o ningún cálculo. Estamos así en posesión de un método que en muchos casos nos permite calcular integrales sin tener que recurrir a cálculos explícitos. Esto es de gran valor tanto en la práctica como para el desarrollo posterior de la teoría.

Con objeto de hacer más sistemático este método, Cauchy, el fundador de la teoría de la integración compleja, introdujo un formalismo sencillo, conocido como el cálculo de residuos. Desde el punto de vista adoptado en este libro, la utilización de los residuos equivale esencialmente a una aplicación de los resultados demostrados en la sección 3-4 en circunstancias particularmente sencillas.

1. *El teorema de los residuos.*—Nuestra tarea inmediata es revisar resultados anteriores a la luz de los teoremas más generales de la sección 3-4. Evidentemente, todos los resultados que se dedujeron como consecuencias del teorema de Cauchy para un disco siguen siendo válidos en regiones arbitrarias para todos los ciclos que sean homológicos a cero. Así, p. ej., y esta aplicación es típica, la fórmula de la integral de Cauchy puede expresarse ahora de la siguiente forma:

Si $f(z)$ es analítica en una región Ω, entonces se tiene

$$n(\gamma, a)f(a) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)dz}{z-a}$$

para todo ciclo γ que sea homológico a cero en Ω.

La demostración es una repetición de la del teorema 6. En relación con ello señalemos que ya no es necesario, naturalmente, dar una demostración separada del teorema 16 cuando existen singularidades evitables. En efecto, nuestra discusión del comportamiento local ha puesto ya en evidencia que pueden ignorarse todas las singularidades evitables.

Volvemos ahora a la discusión de una función $f(z)$, que es analítica en una región Ω, excepto por singularidades aisladas. Como una primera orientación, supongamos que únicamente hay un número finito de puntos singulares, a los que denotaremos por $a_1, a_2, \ldots, a_n$. La región obtenida al excluir los puntos a_j se denotará por Ω'.

Para cada a_j existe un $\delta_j > 0$ tal que la región doblemente conexa $0 < |z - a_j| < \delta_j$ está contenida en Ω'. Trácese una circunferencia C_j de centro a_j y radio menor que δ_j, y sea

$$P_j = \int_{C_j} f(z)dz \qquad [3\text{-}41]$$

el correspondiente período de $f(z)$. La función particular $1/(z-a_j)$ tiene período $2\pi i$. Por tanto, si ponemos $R_j = P_j/2\pi i$, la combinación

$$f(z) - \frac{R_j}{z-a_j}$$

tiene un período nulo. La constante R_j que hace posible este resultado se denomina *residuo* de $f(z)$ en el punto a_j. Repetimos la definición en la forma siguiente:

DEFINICIÓN 3. *El residuo de $f(z)$ en una singularidad aislada a es el único número complejo R que hace que $f(z) - R/(z-a)$ sea la derivada de una función analítica uniforme en una corona circular $0 < |z-a| < \delta$.*

Es cómodo utilizar las notaciones $R = \text{Res}_{z=a} f(z)$, que se explican por sí mismas.

Sea γ un ciclo en Ω' que es homológico a cero con respecto a Ω. Entonces γ verifica la homología siguiente:

$$\gamma \sim \sum_j n(\gamma, a_j)C_j,$$

con respecto a Ω'; en efecto, podemos comprobar fácilmente que los puntos a_j, al igual que todos los puntos exteriores de Ω, tienen el mismo orden con respecto a ambos ciclos. En virtud de la homología obtenemos, con la notación [3-41], la siguiente igualdad:

$$\int_\gamma f\,dz = \sum_j n(\gamma, a_j)P_j,$$

y puesto que $P_j = 2\pi i \cdot R_j$, se tiene por último

$$\frac{1}{2\pi i}\int_\gamma f\,dz = \sum_j n(\gamma, a_j)R_j.$$

Este es el *teorema de los residuos,* excepto por la hipótesis restrictiva de que solo hay un número finito de singularidades. En el caso general necesitamos tan solo probar que $n(\gamma, a_j)=0$, excepto para un número finito de puntos a_j, pues entonces se puede aplicar la misma demostración. La afirmación se deduce mediante un razonamiento rutinario. El conjunto de todos los puntos a con $n(\gamma, a)=0$ es abierto y contiene todos los puntos exteriores a cierto círculo. El complemento es, por tanto, un conjunto compacto, y como tal no puede contener más que un número finito de puntos aislados a_j. Por consiguiente, $n(\gamma, a_j)\neq 0$ únicamente para un número finito de puntos singulares, con lo que hemos probado el teorema siguiente:

TEOREMA 19. *Sea $f(z)$ analítica excepto por singularidades aisladas a_j en una región Ω. Entonces se verifica*

$$\frac{1}{2\pi i}\int_\gamma f(z)\,dz=\sum_j n(\gamma, a_j)\ \mathrm{Res}_{z=a_j} f(z) \qquad [3\text{-}42]$$

para cualquier ciclo γ que sea homológico a cero en Ω y no pase por ninguno de los puntos a_j.

En las aplicaciones se presenta frecuentemente el caso en que cada $n(\gamma, a_j)$ es o cero o uno. Tenemos entonces sencillamente

$$\frac{1}{2\pi i}\int_\gamma f(z)\,dz=\sum_j \mathrm{Res}_{z=a_j} f(z),$$

donde la suma está extendida a todas las singularidades encerradas por γ.

El teorema de los residuos apenas tiene interés práctico si no se dispone de algún procedimiento sencillo para determinar los residuos. Para singularidades esenciales no existe ningún procedimiento de valor práctico, y así no es sorprendente que sea utilizado relativamente poco el teorema de los residuos cuando existen singularidades esenciales. La situación es por completo diferente con relación a los polos. Una sola mirada al desarrollo

$$f(z)=B_h(z-a)^{-h}+\ldots+B_1(z-a)^{-1}+\varphi(z)$$

es suficiente para darse cuenta de que el residuo es igual al coeficiente B_1. En efecto, cuando se omite el término $B_1(z-a)^{-1}$, el resto es evidentemente una derivada. Puesto que la parte principal en un polo siempre se da o puede hallarse fácilmente, tenemos así un método muy sencillo para hallar residuos.

Para polos simples el método es todavía más inmediato, pues entonces el residuo es igual al valor de la función $(z-a)f(z)$ para $z=a$. Así, p. ej., se pide hallar los residuos de la función

$$\frac{e^z}{(z-a)(z-b)}$$

en los polos a y $b\neq a$. El residuo en a es obviamente $e^a/(a-b)$, y el residuo en b es $e^b/(b-a)$. Si $b=a$, la situación es ligeramente más complicada. Debemos entonces desarrollar e^z por el teorema de Taylor en la forma $e^z=e^a+e^a(z-a)+f_2(z)(z-a)^2$. Dividiendo por $(z-a)^2$, hallamos que el residuo de $e^z/(z-a)^2$ en $z=a$ es e^a.

Observación: En la presentación del teorema de Cauchy, la fórmula de la integral y el teorema de los residuos según líneas más clásicas no se hace mención de la homología, ni se utiliza explícitamente el concepto de índice. En su lugar, se supone que la curva γ a la que se aplican los teoremas constituye la frontera completa de una subregión de Ω y se elige la orientación de forma que la subregión quede a la izquierda de Ω. En los textos rigurosos se dedica un gran esfuerzo a la demostración de que estos conceptos intuitivos tienen un significado preciso. La objeción principal a este procedimiento es que precisa conceder tiempo y atención a cuestiones delicadas que son secundarias en comparación con los temas principales.

Con el punto de vista general adoptado por nosotros es todavía posible, y en realidad bastante fácil, aislar el caso clásico. Todo lo que se necesita es aceptar la siguiente definición:

DEFINICIÓN 4. *Se dice que el ciclo γ acota la región Ω si (y solo si) $n(\gamma, a)$ está definido y es igual a 1 para todos los puntos $a\in\Omega$ y no está definido o es igual a cero para todos los puntos a no pertenecientes a Ω.*

Si γ acota a Ω y si $\Omega+\gamma$ está contenido en una región mayor Ω', entonces está claro que γ es homológico a cero con respecto a Ω'. Los enunciados siguientes son, por tanto, consecuencias triviales de los teoremas 18 y 19:

Si γ acota a Ω y $f(z)$ es analítica en el conjunto $\Omega+\gamma$, entonces se verifica

$$\int_\gamma f(z)\,dz=0$$

y

$$f(z)=\frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)\,d\zeta}{\zeta-z}$$

para todo $z\in\Omega$.

Si $f(z)$ es analítica en $\Omega+\gamma$, excepto por singularidades aisladas en Ω, se tiene que

$$\frac{1}{2\pi i}\int_\gamma f(z)\,dz=\sum_j \operatorname{Res}_{z=a_j} f(z)$$

estando extendida la suma a todas las singularidades $a_j\in\Omega$.

Obsérvese que un ciclo γ que acote a Ω debe contener el conjunto de puntos frontera de Ω. En efecto, si z_0 pertenece a la frontera de Ω, entonces todo entorno de z_0 contiene puntos de Ω y puntos no pertenecientes a Ω. Si en tal entorno no hubiera puntos de γ, $n(\gamma, z)$ estaría definido y sería constante en el entorno. Esto contradice a la definición y, por tanto, todo entorno de z_0 debe tener parte común con γ; puesto que γ es cerrado, z_0 debe pertenecer a γ.

El recíproco del enunciado anterior no es cierto, pues un punto de γ pudiera tener un entorno sin parte común con Ω. Normalmente se trataría de elegir γ de forma que fuera idéntico a la frontera de Ω, pero para el teorema de Cauchy y las consideraciones relacionadas con él esto no es necesario.

2. *Principio del argumento.*—La fórmula de la integral de Cauchy puede considerarse como un caso especial del teorema de los residuos. En efecto, la función $f(z)/(z-a)$ tiene un polo simple en a con residuo $f(a)$, y al aplicar [3-42] resulta la fórmula de la integral.

Otra aplicación del teorema de los residuos aparecía en la demostración del teorema 10, que sirve para determinar el número de ceros de una función analítica. Para un cero de orden h podemos escribir $f(z)=(z-a)^h f_h(z)$, con $f_h(a)\neq 0$, y obtener $f'(z)=h(z-a)^{h-1}f_h(z)+(z-a)^h f'_h(z)$. Por tanto, $f'(z)/f(z)=h/(z-a)+f'_h(z)/f_h(z)$, y vemos que f'/f tiene un polo simple con residuo h. En la fórmula [3-30] se tiene en cuenta este residuo mediante una correspondiente repetición de términos.

Podemos ahora generalizar el teorema 10 al caso de una función meromorfa. Si f tiene un polo de orden h, hallamos mediante el mismo cálculo anterior, con $-h$ reemplazando a h, que f'/f tiene el residuo $-h$. Resulta el siguiente teorema:

TEOREMA 20. *Si $f(z)$ es meromorfa en Ω con los ceros a_j y los polos b_k, se tiene entonces que*

$$\frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}\,dz=\sum_j n(\gamma, a_j)-\sum_k n(\gamma, b_k) \qquad [3\text{-}43]$$

para todo ciclo γ que sea homológico a cero en Ω y no pase por ninguno de los ceros ni de los polos.

Se ha de entender que los ceros y los polos múltiples se han de repetir tantas veces como indica su orden; las sumas en [3-43] son finitas.

Al teorema 20 se le suele llamar usualmente el *principio del argumento.* El nombre se refiere a la interpretación del primer miembro de [3-43] como $n(\Gamma, 0)$, siendo Γ el ciclo imagen de γ. Si Γ está en un disco que no contiene al origen, entonces $n(\Gamma, 0)=0$. Esta observación es la base del corolario siguiente, conocido como el *teorema de Rouché:*

COROLARIO. *Sea γ homológico a cero en Ω, y tal que $n(\gamma, z)$ sea cero o uno para cualquier punto no perteneciente a γ. Supongamos que $f(z)$ y $g(z)$ son analíticas en Ω y satisfacen la desigualdad $|f(z)-g(z)|<|f(z)|$ sobre γ. Entonces $f(z)$ y $g(z)$ tienen el mismo número de ceros encerrados por γ.*

El supuesto implica que $f(z)$ y $g(z)$ no tienen ceros sobre γ. Además, satisfacen la desigualdad

$$\left|\frac{g(z)}{f(z)}-1\right|<1$$

sobre γ. Los valores de $F(z)=g(z)/f(z)$ sobre γ están, pues, contenidos en el disco abierto de centro 1 y radio 1. Al aplicar el teorema 20 a $F(z)$, tenemos que $n(\Gamma, 0)=0$, de donde se deduce la afirmación.

Una aplicación típica del teorema de Rouché sería la siguiente. Supongamos que se desea hallar el número de ceros de una función $f(z)$ en el disco $|z|\leqq R$. Utilizando el teorema de Taylor, podemos escribir

$$f(z)=P_{n-1}(z)+z^n f_n(z),$$

siendo P_{n-1} un polinomio de grado $n-1$. Para un n adecuadamente elegido pudiera ocurrir que seamos capaces de demostrar la desigualdad $R^n|f_n(z)| < |P_{n-1}(z)|$ sobre $|z|=R$. Entonces $f(z)$ tiene el mismo número de ceros en $|z| \leqq R$ que $P_{n-1}(z)$, y este número puede determinarse mediante una solución aproximada de la ecuación polinómica $P_{n-1}(z)=0$.

Se puede generalizar el teorema 20 de la siguiente manera. Si $g(z)$ es analítica en Ω, entonces $g(z)\dfrac{f'(z)}{f(z)}$ tiene como residuo $hg(a)$ en un cero a de orden h, y el residuo $-hg(a)$ en un polo. Obtenemos así la siguiente fórmula

$$\frac{1}{2\pi i}\int_\gamma g(z)\frac{f'(z)}{f(z)}\,dz = \sum_j n(\gamma, a_j)g(a_j) - \sum_k n(\gamma, b_k)g(b_k). \qquad [3\text{-}44]$$

Este resultado es importante para el estudio de la función inversa. Con las notaciones del teorema 11 sabemos que la ecuación $f(z)=w$, $|w-w_0|<\delta$ tiene n raíces $z_j(w)$ en el disco $|z-z_0|<\epsilon$. Si aplicamos [3-44] con $g(z)=z$, se obtiene

$$\sum_{j=1}^{n} z_j(w) = \frac{1}{2\pi i}\int_{|z-z_0|=\epsilon}\frac{f'(z)}{f(z)-w}\,z\,dz. \qquad [3\text{-}45]$$

Para $n=1$ la función inversa $f^{-1}(w)$ puede así representarse explícitamente por

$$f^{-1}(w) = \frac{1}{2\pi i}\int_{|z-z_0|=\epsilon}\frac{f'(z)}{f(z)-w}\,z\,dz.$$

Si se aplica [3-44] con $g(z)=z^m$, la ecuación [3-45] es reemplazada por

$$\sum_{j=1}^{n} z_j(w)^m = \frac{1}{2\pi i}\int_{|z-z_0|=\epsilon}\frac{f'(z)}{f(z)-w}\,z^m dz.$$

No es difícil demostrar que el segundo miembro representa una función analítica de w para $|w-w_0|<\delta$. Así, las sumas de las potencias de las raíces $z_j(w)$ son funciones analíticas uniformes de w. Pero es bien conocido que las funciones simétricas elementales pueden expresarse como polinomios en sumas de potencias. Por tanto, también

son analíticas, y tenemos que las $z_j(w)$ son las raíces de una ecuación polinómica

$$z^n + a_1(w)z^{n-1} + \ldots + a_{n-1}(w)z + a_n(w) = 0,$$

cuyos coeficientes son funciones analíticas de w en $|w-w_0|<\delta$.

3. *Cálculo de integrales definidas.*—El cálculo de residuos proporciona una herramienta eficaz para el cálculo de integrales definidas. Es particularmente importante cuando resulta imposible hallar explícitamente la integral indefinida, pero aun cuando los métodos ordinarios del cálculo puedan aplicarse, la utilización de residuos es un artificio que frecuentemente ahorra trabajo. El hecho de que el cálculo de residuos da integrales complejas en lugar de integrales reales no es una desventaja. Pues evidentemente el cálculo de una integral compleja es equivalente al cálculo de dos integrales definidas.

Existen, sin embargo, algunas limitaciones serias, y el método no es precisamente infalible. En primer lugar, el integrando debe estar estrechamente relacionado con alguna función analítica. Este no es un inconveniente muy grave, ya que generalmente solo necesitamos integrar funciones elementales, y todas ellas pueden prolongarse al dominio complejo. Es mucho más serio el hecho de que la técnica de la integración compleja se aplica únicamente a curvas cerradas, mientras que una integral real está siempre extendida a un intervalo. Debe utilizarse un artificio especial con objeto de reducir el problema a otro referente a la integral sobre una curva cerrada. Esto se logra hacer de varias maneras, pero todas se aplican en circunstancias algo especiales. La técnica puede aprenderse siguiendo algunos ejemplos típicos, pero incluso un dominio completo de la misma no garantiza el éxito.

1) Todas las integrales de la forma

$$\int_0^{2\pi} R(\cos\theta, \operatorname{sen}\theta)\,d\theta, \qquad [3\text{-}46]$$

donde el integrando es una función racional de $\cos\theta$ y de $\operatorname{sen}\theta$, pueden calcularse fácilmente por medio de residuos. Naturalmente también es posible calcular estas integrales por integración explícita, pero esta técnica es generalmente muy laboriosa. Es natural efectuar la sustitución $z=e^{i\theta}$, que transforma inmediatamente [3-46] en la integral curvilínea

$$-i\int_{|z|=1} R\left[\frac{1}{2}\left(z+\frac{1}{z}\right),\ \frac{1}{2i}\left(z-\frac{1}{z}\right)\right]\frac{dz}{z}.$$

Queda únicamente la determinación de los residuos correspondientes a los polos del integrando interiores al círculo unidad.

Como un ejemplo calculemos la integral

$$\int_0^{\pi}\frac{d\theta}{a+\cos\theta},\qquad a>1.$$

Esta integral no está extendida al intervalo $(0, 2\pi)$, pero puesto que $\cos\theta$ toma los mismos valores en los intervalos $(0,\pi)$ y $(\pi, 2\pi)$, es claro que la integral desde 0 a π es la mitad de la integral desde 0 a 2π. Teniendo esto en cuenta, hallamos que la integral es igual a

$$-i\int_{|z|=1}\frac{dz}{z^2+2az+1}$$

El denominador puede descomponerse en factores $(z-\alpha)(z-\beta)$, donde tenemos $\alpha=-a+\sqrt{a^2-1}$, $\beta=-a-\sqrt{a^2-1}$. Evidentemente, $|\alpha|<1$, $|\beta|>1$, y el residuo en α es $1/(\alpha-\beta)$. Se halla que el valor de la integral es $\pi/\sqrt{a^2-1}$.

2) Una integral de la forma

$$\int_{-\infty}^{\infty}R(x)\,dx$$

converge si (y solo si) en la función racional $R(x)$ el grado del denominador es por lo menos dos unidades mayor que el grado del numerador y si ningún polo está en el eje real. El procedimiento típico es integrar la función compleja $R(z)$ sobre una curva cerrada consistente en el segmento $(-\rho, \rho)$ y la semicircunferencia desde ρ a $-\rho$, en el semiplano superior. Si ρ es suficientemente grande, esta curva encierra a todos los polos del semiplano superior, siendo la integral correspondiente igual a la suma de los residuos en el semiplano superior multiplicada por $2\pi i$. Cuando $\rho\to\infty$, un cálculo obvio muestra que la integral sobre la semicircunferencia tiende a cero y obtenemos

$$\int_{-\infty}^{\infty}R(x)dx=2\pi i\sum_{y>0}\operatorname{Res}R(z).$$

3) Puede aplicarse el mismo método a una integral de la forma

$$\int_{-\infty}^{\infty}R(x)e^{ix}dx, \qquad [3\text{-}47]$$

cuyas partes real e imaginaria determinan las importantes integrales

$$\int_{-\infty}^{\infty}R(x)\cos x\,dx,\qquad \int_{-\infty}^{\infty}R(x)\operatorname{sen} x\,dx. \qquad [3\text{-}48]$$

Puesto que $|e^{iz}|=e^{-y}$ está acotada en el semiplano superior, podemos concluir nuevamente que la integral sobre la semicircunferencia tiende a cero con tal que la función racional $R(z)$ tenga un cero de al menos orden 2 en el infinito. Se obtiene

$$\int_{-\infty}^{\infty}R(x)e^{ix}dx=2\pi i\sum_{y>0}\operatorname{Res}R(z)e^{iz}.$$

Es menos obvio el que se verifique el mismo resultado cuando $R(z)$ tiene solo un cero simple en el infinito. En este caso no resulta conveniente utilizar semicircunferencias. En primer lugar, no es fácil calcular la integral sobre la semicircunferencia, y en segundo lugar, aun cuando pudiéramos hacerlo habríamos probado únicamente que la integral

$$\int_{-\rho}^{\rho}R(x)e^{ix}dx$$

sobre un intervalo simétrico tiene el límite deseado para $\rho\to\infty$. En realidad lo que deseamos es, naturalmente, probar que

$$\int_{-X_1}^{X_2}R(x)e^{ix}dx$$

tiene límite cuando X_1 y X_2 tienden independientemente a infinito. En los ejemplos anteriores no se suscitaba esta cuestión a causa de que estaba asegurada de antemano la convergencia de la integral.

Para demostrarlo integramos sobre el perímetro de un rectángulo de vértices X_2, X_2+iY, $-X_1+iY$, $-X_1$, siendo $Y>0$. Si X_1, X_2 e Y son suficientemente grandes este rectángulo contiene todos los polos del semiplano superior. De acuerdo con las hipótesis $|zR(z)|$ está acotado. Por tanto, la integral sobre el lado vertical derecho es, salvo un factor constante, menor que

$$\int_0^Y e^{-y}\frac{dy}{|z|}<\frac{1}{X_2}\int_0^Y e^{-y}dy.$$

Se puede calcular explícitamente la última integral y se halla que es menor que uno. Por consiguiente, la integral sobre el lado vertical derecho es menor que una constante multiplicada por $1/X_2$, y se halla un resultado correspondiente para el lado vertical izquierdo. La integral sobre el lado horizontal superior es evidentemente menor que $e^{-Y}(X_1+X_2)/Y$ multiplicado por una constante. Para X_1 y X_2 fijos tiende a cero cuando $Y\to\infty$, y se concluye que

$$\left|\int_{-X_1}^{X_2}R(x)e^{ix}dx-\sum_{y>0}\text{Res}\,R(z)e^{iz}\right|<A\left(\frac{1}{X_1}+\frac{1}{X_2}\right),$$

donde A es una constante. Esta desigualdad prueba que

$$\int_{-\infty}^{\infty}R(x)e^{ix}dx=2\pi i\sum_{y>0}\text{Res}\,R(z)e^{iz}$$

con la única condición de que $R(\infty)=0$.

En la discusión hemos supuesto tácitamente que $R(z)$ no tenía polos en el eje real, pues de otra forma la integral [3-47] no tendría significado. Sin embargo, una de las integrales [3-48] pudiera muy bien existir, a saber: si $R(z)$ tiene polos simples que coincidan con ceros de sen x o cos x. Supongamos, p. ej., que $R(z)$ tiene un polo simple en $z=0$. Entonces la segunda integral de [3-48] tiene sentido y hay que calcularla.

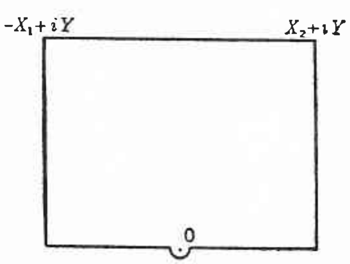

FIG. 3-10.

Utilizamos el mismo método que anteriormente, pero ahora usaremos un camino que evite el origen siguiendo una pequeña semicircunferencia de radio δ en el semiplano inferior (Fig. 3-10). Es fácil ver que esta curva cerrada encierra los polos del semiplano superior, el polo en el origen y ninguno más, siempre y cuando X_1, X_2, Y sean suficientemente grandes y δ suficientemente pequeño. Supongamos que el residuo en cero es B, así que podemos escribir $R(z)e^{iz}=B/z+R_0(z)$, siendo $R_0(z)$ analítica en el origen. La integral del primer término sobre la semicircunferencia es πiB, mientras que la del segundo tiende a cero con δ. Es obvio que se obtiene el resultado

$$\lim_{\delta\to 0}\int_{-\infty}^{-\delta}+\int_{\delta}^{\infty}R(x)e^{ix}dx=2\pi i\left[\sum_{y>0}\text{Res}\,R(z)e^{iz}+\tfrac{1}{2}B\right].$$

El límite del primer miembro se denomina *valor principal de Cauchy* de la integral; existe aunque la integral no tenga significado de por sí. Observemos que en el segundo miembro se ha incluido la mitad del residuo en el punto cero; esto equivale a contar medio polo como perteneciente al semiplano superior.

En el caso general, cuando hay varios polos en el eje real, obtenemos

$$\text{v. pr.}\int_{-\infty}^{\infty}R(x)e^{ix}dx=2\pi i\sum_{y>0}\text{Res}\,R(z)e^{iz}+\pi i\sum_{y=0}\text{Res}\,R(z)e^{iz},$$

donde las notaciones se explican por sí mismas. Es una hipótesis esencial la de que todos los polos sobre el eje real sean simples, y debemos suponer como antes que $R(\infty)=0$.

Como ejemplo más sencillo tenemos que

$$\text{v. pr.}\int_{-\infty}^{\infty}\frac{e^{ix}}{x}dx=\pi i.$$

Separando las partes real e imaginaria observamos que la parte real de la ecuación es trivial como consecuencia del hecho de ser impar el integrando. En la parte imaginaria no es necesario tomar el valor principal, y puesto que el integrando es par hallamos que

$$\int_0^{\infty}\frac{\text{sen}\,x}{x}dx=\frac{\pi}{2}$$

Hacemos constar que las integrales que contienen un factor $\cos^n x$ o $\text{sen}^n x$ pueden calcularse mediante la misma técnica. En efecto, sabemos expresar estos factores como combinaciones lineales de términos $\cos mx$ y $\operatorname{sen} mx$, y las integrales correspondientes pueden reducirse a la forma [3-47] por un cambio de variable:

$$\int_{-\infty}^{\infty} R(x)e^{imx}\,dx = \frac{1}{m}\int_{-\infty}^{\infty} R\left(\frac{x}{m}\right)e^{ix}\,dx.$$

4) La categoría siguiente de integrales tiene la forma

$$\int_0^{\infty} x^{\alpha}R(x)\,dx,$$

donde el exponente α es real y puede suponerse perteneciente al intervalo $0<\alpha<1$. Para la convergencia $R(z)$ debe tener un cero de al menos orden 2 en ∞, y a lo sumo un polo simple en el origen.

El nuevo carácter distintivo es el hecho de que $R(z)z^{\alpha}$ no es uniforme. Esta es, sin embargo, precisamente la circunstancia que hace posible el cálculo de la integral desde 0 a ∞.

El procedimiento más sencillo consiste en empezar con la sustitución $x=t^2$, que transforma la integral en

$$2\int_0^{\infty} t^{2\alpha+1}R(t^2)\,dt.$$

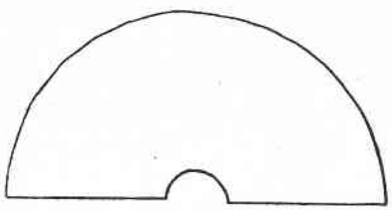

FIG. 3-11.

Para la función $z^{2\alpha}$ podemos elegir la rama cuyo argumento está comprendido entre $-\pi\alpha$ y $3\pi\alpha$; está bien definida y es analítica en la región obtenida al omitir el eje imaginario negativo. Siempre y cuando evitemos el eje imaginario negativo, es posible aplicar el teorema de los residuos a la función $z^{2\alpha+1}R(z^2)$. Utilizaremos una curva cerrada consistente en dos segmentos a lo largo de los ejes positivo y negativo, más dos semicircunferencias en el semiplano superior, una muy grande y otra muy pequeña (Fig. 3-11). Teniendo en cuenta nuestras suposiciones, es fácil probar que las integrales sobre las semicircunferencias tienden a cero. Por tanto, el teorema de los residuos nos da el valor de la integral

$$\int_{-\infty}^{\infty} z^{2\alpha+1}R(z^2)\,dz = \int_0^{\infty} [z^{2\alpha+1} + (-z)^{2\alpha+1}]R(z^2)\,dz.$$

Sin embargo, $(-z)^{2\alpha} = e^{2\pi i\alpha}z^{2\alpha}$, y la integral es igual a

$$(1+e^{2\pi i\alpha})\int_0^{\infty} z^{2\alpha+1}R(z^2)\,dz.$$

Puesto que el factor fuera de la integral es distinto de cero, estamos finalmente en condiciones de hallar el valor de la integral deseada.

El cálculo exige la determinación de los residuos de $z^{2\alpha+1}R(z^2)$ en el semiplano superior. Estos son los mismos que los residuos de $z^{\alpha}R(z)$ en todo el plano. Para fines prácticos pudiera ser preferible no utilizar ninguna sustitución previa e integrar la función $z^{\alpha}R(z)$ sobre la curva cerrada que aparece en la figura 3-12. Debemos entonces utilizar la rama de z^{α}, cuyo argumento está comprendido entre cero y $2\pi\alpha$. Este método necesita alguna justificación, pues no se ajusta a las hipótesis del teorema de los residuos. La justificación es trivial.

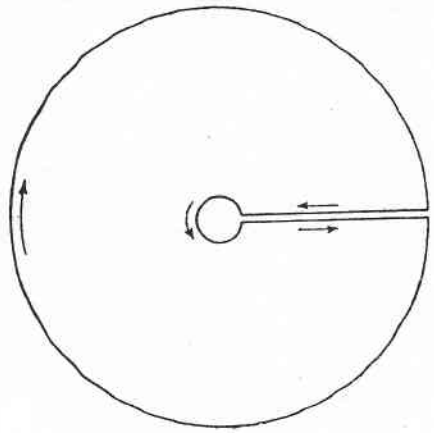

FIG. 3-12.

5) Como ejemplo final calcularemos la integral especial

$$\int_0^{\pi} \log \operatorname{sen}\theta\, d\theta.$$

Consideremos la función $1-e^{2iz} = -2ie^{iz}\operatorname{sen} z$. De la representación $1-e^{2iz} = 1-e^{-2y}(\cos 2x + i\operatorname{sen} 2x)$ hallamos que la función es real y negativa únicamente para $x=n\pi$, $y\leqq 0$. En la región obtenida omitiendo estas semirrectas la rama principal de $\log(1-e^{2iz})$ es, por tanto, analítica y uniforme. Aplicamos el teorema de Cauchy al rectángulo cuyos vértices son 0, π, $\pi+iY$ e iY; sin embargo, los puntos 0 y π han de ser evitados, y lo hacemos siguiendo pequeños cuadrantes circulares de radio δ.

A causa de la periodicidad, las integrales sobre los lados verticales se anulan unas con otras. La integral sobre el lado horizontal superior tiende a cero cuando $Y\to\infty$. Por último, las integrales sobre los cuadrantes puede verse también que tienden a cero cuando $\delta\to 0$. En efecto, ya que la parte imaginaria del logaritmo está acotada, necesitamos únicamente considerar la parte real. Puesto que e^{2iz} tiene como derivada en el origen $2i$, hallamos que $|1-e^{2iz}|/|z| \to 2$ para $z \to 0$. Por consiguiente, solo necesitamos probar que $\delta \log \delta$ tiende a cero con δ. Para no dejar este estudio incompleto incluimos una demostración de este hecho elemental. Estudiando la derivada hallamos que la función $t \log(\epsilon/t)$ es creciente para $t \leqq \epsilon/e$. Para $\delta < \epsilon/e$ tenemos así que $\delta \log(\epsilon/\delta) < \epsilon/e$. Si elegimos $\epsilon < 1$ y hacemos que δ satisfaga la condición adicional $\delta < \epsilon/[e \log(1/\epsilon)]$, tenemos que $\delta \log(1/\delta) < 2\epsilon/e$, y esto prueba que $\lim\limits_{\delta \to 0} \delta \log(1/\delta) = 0$.

La misma demostración se aplica en las proximidades del vértice π, y obtenemos

$$\int_0^\pi \log(-2ie^{ix} \operatorname{sen} x)\, dx = 0.$$

Si elegimos $\log e^{ix} = ix$, la parte imaginaria está comprendida entre 0 y π. Por consiguiente, con objeto de obtener la rama principal con parte imaginaria comprendida entre $-\pi$ y π, tenemos que elegir $\log(-i) = -\pi i/2$. La ecuación puede escribirse ahora en la forma

$$\pi \log 2 - \left(\frac{\pi^2}{2}\right) i + \int_0^\pi \log \operatorname{sen} x\, dx + \left(\frac{\pi^2}{2}\right) i = 0,$$

y se obtiene que

$$\int_0^\pi \log \operatorname{sen} x\, dx = -\pi \log 2.$$

EJERCICIOS

Calcúlense las integrales siguientes por el método de los residuos:

1. $\displaystyle\int_0^{\pi/2} \frac{dx}{a+\operatorname{sen}^2 x}, \qquad a > 0.$
2. $\displaystyle\int_0^\infty \frac{x^2 dx}{x^4+6x^2+13}$
3. $\displaystyle\int_{-\infty}^\infty \frac{x^2-x+2}{x^4+10x^2+9}\, dx.$
4. $\displaystyle\int_0^\infty \frac{x^2 dx}{(x^2+a^2)^3}, \qquad a$ real.

5. $\displaystyle\int_0^\infty \frac{\cos x}{x^2+a^2}\, dx, \qquad a$ real.
6. $\displaystyle\int_0^\infty \frac{x \operatorname{sen} x\, dx}{x^2+a^2}, \qquad a$ real.
7. $\displaystyle\int_0^\infty \frac{\operatorname{sen}^2 kx}{x^2}\, dx, \qquad k > 0.$
8. $\displaystyle\int_0^\infty (1+x^2)^{-2} \log x\, dx.$
9. $\displaystyle\int_0^\infty \log(1+x^2) \frac{dx}{x^{1+\alpha}} \qquad (0 < \alpha < 2).$

CAPITULO IV

SUCESIONES

4-1. Sucesiones convergentes.—Una sucesión de números complejos $a_1, a_2, \ldots, a_n, \ldots$ puede considerarse como una función de una variable independiente n que toma solo valores enteros. El caso de las sucesiones queda así sumergido dentro del caso general de las funciones definidas en un subconjunto del eje real. Sin embargo, no tiene ningún mérito especial el adoptar este punto de vista general. Por el contrario, estamos mucho más interesados en hacer resaltar las características especiales de las sucesiones, que las hacen particularmente sencillas y útiles.

En la teoría de funciones es de primordial importancia la consideración de sucesiones funcionales. Leyes especiales para la formación de sucesiones dan lugar a las series y a los productos infinitos. Uno de los objetivos es el de la representación de funciones analíticas mediante series o productos infinitos. Tales representaciones son bastante explícitas, pudiendo esperarse por esta razón que revelen propiedades importantes de las funciones.

1. *Sucesiones fundamentales.*—La sucesión $\{a_n\}_1^\infty$ tiene el límite A si para todo $\epsilon > 0$ existe un n_0 tal que $|a_n - A| < \epsilon$ para $n > n_0$. Una sucesión con límite finito se dice *convergente*, y cualquier sucesión que no es convergente es *divergente*. Si $\lim_{n\to\infty} a_n = \infty$, se dice que la sucesión *diverge a infinito*.

Solo en casos excepcionales puede probarse la convergencia de una sucesión mostrando su límite; así, pues, es extremadamente importante tener a nuestra disposición un método que nos permita probar la existencia de un límite independientemente de que pueda o no ser determinado de manera explícita. El criterio que sirve para este propósito lleva el nombre de Cauchy. Se dirá que una sucesión es *fundamental*, o *sucesión de Cauchy*, si satisface la condición siguiente: dado un $\epsilon > 0$ cualquiera, existe un n_0 tal que $|a_n - a_m| < \epsilon$, siempre y cuando $n > n_0$ y $m > n_0$. Como consecuencia de esta definición tenemos:

Una sucesión es convergente si (y solo si) es fundamental.

La necesidad de la condición es casi trivial. Si $a_n \to A$, podemos hallar un n_0 tal que $|a_n - A| < \epsilon/2$ para $n > n_0$. Para $m, n > n_0$ se sigue por la desigualdad triangular que $|a_n - a_m| \leq |a_n - A| + |a_m - A| < \epsilon$. La suficiencia está naturalmente relacionada muy estrechamente con la definición de número real; de hecho, una de las formas de introducir los números reales es postulando la suficiencia de la condición de Cauchy. En nuestra exposición, el procedimiento más natural es basar la suficiencia en el teorema de Bolzano-Weierstrass (Cap. II, Sec. 2-2, 3, teorema 6). En primer lugar, la condición de Cauchy implica que los números a_n están acotados. En efecto, si dichos números no estuvieran acotados podríamos hallar, para todo n, un $m > n$ tal que $|a_m| > |a_n| + 1$ y, por tanto, $|a_n - a_m| > 1$; esto contradice a la condición de Cauchy para $\epsilon = 1$. Podemos, pues, suponer que $|a_n| \leq M$ para todo n. Por el teorema de Bolzano-Weierstrass existe un punto límite A. La definición de punto límite implica que $|a_n - A| < \epsilon/2$ para infinitos valores de n. Sea n_0 el que por la condición de Cauchy corresponde al número positivo $\epsilon/2$. Elijamos un $n > n_0$ fijo con $|a_n - A| < \epsilon/2$ y sea m un entero cualquiera $> n_0$. Entonces $|a_m - A| \leq |a_m - a_n| + |a_n - A| < \epsilon$, con lo que hemos demostrado que $\lim_{n\to\infty} a_n = A$.

Una aplicación muy sencilla se refiere a la comparación de dos sucesiones $\{a_n\}$ y $\{b_n\}$. Si es cierto que $|b_m - b_n| \leq |a_m - a_n|$ para todos los pares de subíndices, se puede denominar a la última sucesión una *contracción* de la primera. Satisfecha esta condición, si $\{a_n\}$ es fundamental, es evidente que $\{b_n\}$ también es fundamental. Por tanto, si $\{a_n\}$ converge, lo mismo le ocurre a $\{b_n\}$.

Una serie es una suma infinita formal

$$a_1 + a_2 + \ldots + a_n + \ldots \qquad [4\text{-}1]$$

Asociada a esta serie está la sucesión de sus sumas parciales

$$s_n = a_1 + a_2 + \ldots + a_n.$$

Se dice que la serie es convergente si (y solo si) la sucesión correspondiente es convergente, y en este caso el límite de la sucesión es la *suma* de la serie.

Aplicado a una serie, el criterio de Cauchy da la siguiente condición: la serie [4-1] converge si (y solo si) para todo $\epsilon > 0$ existe un n_0 tal que $|a_n + a_{n+1} + \ldots + a_{n+p}| < \epsilon$ para todo $n > n_0$ y $p \geq 0$. Para

$p=0$ hallamos, en particular, que $|a_n| < \epsilon$. Por consiguiente, el término general de una serie convergente tiende a cero. Esta condición es necesaria; pero, naturalmente, no es suficiente.

Si se omite un número finito de términos de la serie [4-1], la nueva serie converge o diverge junto con [4-1]. En el caso de convergencia, sea R_n la suma de la serie que empieza con el término a_{n+1}. Entonces la suma de toda la serie es $S=s_n+R_n$.

La serie [4-1] puede compararse con la serie

$$|a_1|+|a_2|+\ldots+|a_n|+\ldots, \qquad [4\text{-}2]$$

formada por los valores absolutos de los términos. La sucesión de sumas parciales de [4-1] es una contracción de la correspondiente sucesión de [4-2], pues $|a_n+a_{n+1}+\ldots+a_{n+p}| \leqq |a_n|+|a_{n+1}|+\ldots+|a_{n+p}|$. Por tanto, la convergencia de [4-2] implica la de la serie original [4-1]. Una serie con la propiedad de que la serie formada por los valores absolutos de los términos converge se llama *absolutamente convergente*.

2. *Subsucesiones.*—En el capítulo II, sección 2-2, 1, hemos ya examinado sucesiones desde el punto de vista de sus *puntos límites*. Recordemos que y es un punto límite de la sucesión $\{x_n\}$ si (y solo si) todo entorno de y contiene x_n para infinitos valores de n. Con esta condición existe, en primer lugar, un entero n_1 tal que $|x_{n_1}-y|<1$, y para elegir definitivamente n_1 sea este el menor entero que cumpla esta propiedad. A continuación existe un $n_2>n_1$ tal que $|x_{n_2}-y|<\frac{1}{2}$, y podemos tomar nuevamente como n_2 el menor entero que cumpla las dos propiedades. Es posible ahora continuar la construcción por inducción. Suponiendo que se han hallado ya $n_1<n_2<\ldots<n_{k-1}$, podemos determinar $n_k>n_{k-1}$ tal que $|x_{n_k}-y|<1/k$; para una perfecta definición elegimos n_k tan pequeño como sea posible. La construcción determina una *subsucesión* $\{x_{n_k}\}$ con $\lim\limits_{k\to\infty} x_{n_k}=y$. Así que todo punto límite es el límite de una subsucesión. Recíprocamente, es evidente que el límite de una subsucesión es un punto límite de la sucesión original.

Para simplificar el lenguaje, de ahora en adelante diremos un *límite* en lugar de un punto límite. Esto puede hacerse sin miedo a confusión si distinguimos cuidadosamente entre *el* límite de una sucesión convergente y *un* límite de una sucesión arbitraria.

Tienen particular interés los límites de una sucesión de números reales. Es prácticamente evidente que el conjunto de los límites de

una sucesión $\{x_n\}$ es cerrado (cf. Cap. II, Sec. 2-2, 1, Ej. 7). Si los números x_n están acotados, lo mismo ocurre con el conjunto de los límites. Más aún: por el teorema de Bolzano-Weierstrass (Cap. II, Sec. 2-2, 3, teorema 6), el conjunto de los límites no es vacío. En el caso real podemos, por tanto, aplicar el teorema 2 del capítulo II, sección 2-2, 2, y hallamos que existe un límite mayor que todos A y un límite menor que todos a. Se les llama, respectivamente, *límite superior* y *límite inferior* de la sucesión $\{x_n\}$, y se denotan por

$$\operatorname*{lím\,sup}_{n\to\infty} x_n=A, \qquad \operatorname*{lím\,inf}_{n\to\infty} x_n=a$$

o, mediante esta otra notación,

$$\overline{\lim_{n\to\infty}}\, x_n=A, \qquad \underline{\lim_{n\to\infty}}\, x_n=a.$$

Si la sucesión no está acotada, existe una subsucesión que tiende a $+\infty$ o una que tiende a $-\infty$. Estos límites impropios se comparan con los límites ordinarios, si existen, y el mayor y el menor de los límites toman nuevamente los nombres de límite superior y límite inferior, respectivamente. Si se toman en cuenta los valores $\pm\infty$, se concluye así que toda sucesión tiene un límite superior y otro inferior.

La utilización de estos conceptos tiene sobre todo ventajas de tipo formal. Ocurre frecuentemente que una demostración complicada en términos de épsilon, puede reducirse a una manipulación formal basada en desigualdades del tipo

$$\underline{\lim}\, x_n \leqq \overline{\lim}\, x_n$$

$$\underline{\lim}\, x_n+\underline{\lim}\, y_n \leqq \underline{\lim}\,(x_n+y_n) \leqq \underline{\lim}\, x_n+\overline{\lim}\, y_n$$

$$\underline{\lim}\, x_n+\overline{\lim}\, y_n \leqq \overline{\lim}\,(x_n+y_n) \leqq \overline{\lim}\, x_n+\overline{\lim}\, y_n.$$

EJERCICIOS

1. Demuéstrese la desigualdad $\underline{\lim}\,(x_n+y_n) \leqq \underline{\lim}\, x_n+\overline{\lim}\, y_n$.

2. Pruébese que $\overline{\lim\limits_{n\to\infty}}\, \cos nx=1$ si x/π es irracional.

3. Investíguense los límites inferior y superior de la sucesión a, a^a, $a^{(a^a)}$, ... para diferentes valores positivos de a.

3. *Convergencia uniforme.*—Consideremos una sucesión de funciones $f_n(x)$. Supongamos en primer lugar que todas las funciones $f_n(x)$ están definidas en el mismo conjunto E. Se dice entonces que $f_n(x)$ converge a $f(x)$ en E si $\lim_{n\to\infty} f_n(a)=f(a)$ para todo $a \in E$. Es muy importante a veces considerar una situación más general. Suponiendo que $f_n(x)$ está definida en E_n, consideremos el conjunto E, que consiste en todos los puntos x que pertenecen a todos los E_n salvo a un número finito de ellos. Esto significa que para $x \in E$, $f_n(x)$ está definida para n suficientemente grande; puesto que el límite de una sucesión depende únicamente de los valores para n grande, todavía es lícito hablar del límite de la sucesión $f_n(x)$ en E. Así, p. ej., la función $1/(z-n)$ no está definida para $z=n$. Sin embargo, puede mantenerse que $\lim_{n\to\infty} 1/(z-n)=0$ en todo el plano.

En cuestiones de convergencia el interés está centrado en el subíndice n, mientras que la variable x actúa como parámetro. Con objeto de hacerla compatible con otras utilizaciones de la palabra uniforme, debemos definir la *convergencia uniforme* en los términos siguientes: la sucesión $\{f_n(x)\}$ converge uniformemente hacia $f(x)$ en un conjunto E si para todo $\epsilon > 0$ existe un n_0 tal que $f_n(x)$ está definida y satisface la condición $|f_n(x)-f(x)| < \epsilon$ para todo $n > n_0$ y todo $x \in E$.

En nuestro ejemplo, la convergencia de $1/(z-n)$ a cero no es uniforme en todo el plano, por la sencilla razón de que ninguna de las funciones está siquiera definida en todo el plano. Sin embargo, es fácil ver que la convergencia es uniforme en un conjunto compacto. Es también uniforme en el eje imaginario.

Hay dos consecuencias de la convergencia uniforme que se usan constantemente. En primer lugar podemos afirmar que la función límite de una sucesión uniformemente convergente de funciones continuas es continua. Para demostrarlo, supongamos que las funciones $f_n(x)$ son continuas y tienden uniformemente a $f(x)$ en E. Sea x_0 un punto de E y elijamos un $\epsilon > 0$. Por la convergencia uniforme existe un n tal que $|f_n(x)-f(x)| < \epsilon/3$ en E. Además, puesto que $f_n(x)$ es continua, podemos hallar un $\delta > 0$ tal que $|f_n(x)-f_n(x_0)| < \epsilon/3$, siempre y cuando $x \in E$ y $|x-x_0| < \delta$. Para los mismos puntos tenemos entonces que $|f(x)-f(x_0)| \leqq |f(x)-f_n(x)| + |f_n(x)-f_n(x_0)| + |f_n(x_0)-f(x_0)| < \epsilon$, con lo que hemos demostrado que $f(x)$ es continua en x_0.

La segunda consecuencia se refiere a la integración. Supongamos

que las funciones $f_n(z)$ son continuas y convergen uniformemente a $f(z)$ en un arco diferenciable γ. Entonces es cierto que

$$\lim_{n\to\infty} \int_\gamma f_n(z)\,dz = \int_\gamma f(z)\,dz.$$

La demostración es obvia.

A la vista de estas conclusiones es deseable en grado sumo poder demostrar que una sucesión converge uniformemente. Observemos que la condición necesaria y suficiente de Cauchy tiene un duplicado para la convergencia uniforme. Afirmamos que:

La sucesión $\{f_n(x)\}$ converge uniformemente sobre E si (y solo si) para todo $\epsilon > 0$ existe un n_0 tal que $|f_m(x)-f_n(x)| < \epsilon$ para todo $m > n_0$, $n > n_0$ y todo $x \in E$.

Nuevamente es trivial la necesidad de la condición. Para la suficiencia, observemos que la función límite $f(x)$ existe de acuerdo con la forma ordinaria del criterio de Cauchy. En la desigualdad $|f_m(x) - f_n(x)| < \epsilon$ podemos mantener n fijo y hacer que m tienda a ∞. Se sigue que $|f(x)-f_n(x)| \leqq \epsilon$ para $n > n_0$ y para todo $x \in E$; por tanto, la convergencia es uniforme.

El siguiente criterio es el más aplicable para usos prácticos: Si una sucesión de funciones $\{f_n(x)\}$ es una contracción de una sucesión convergente de constantes $\{a_n\}$, entonces la sucesión $\{f_n(x)\}$ es uniformemente convergente. La hipótesis significa que $|f_m(x)-f_n(x)| \leqq |a_m-a_n|$ sobre E, de donde se sigue inmediatamente la conclusión mediante la condición de Cauchy.

En el caso de series, este criterio, en una forma ligeramente más débil, es particularmente sencillo. Decimos que una serie de términos variables

$$f_1(x) + f_2(x) + \ldots + f_n(x) + \ldots$$

tiene a la serie de términos positivos

$$a_1 + a_2 + \ldots + a_n + \ldots$$

como *mayorante* si se verifica que $|f_n(x)| \leqq Ma_n$ para alguna constante M y para todo n suficientemente grande; recíprocamente, la primera serie es una *minorante* de la segunda. En estas circunstancias, tenemos

$$|f_n(x) + f_{n+1}(x) + \ldots + f_{n+p}(x)| \leqq M(a_n + a_{n+1} + \ldots + a_{n+p}).$$

Por consiguiente, si la mayorante converge, la minorante converge uniformemente. Se conoce esta condición con el nombre de *criterio M de Weierstrass*. Tiene un ligero inconveniente, que solo se aplica a series absolutamente convergentes. El principio general de contracción es más complicado, pero tiene un campo de aplicación más amplio.

EJERCICIOS

1. Pruébese que una sucesión convergente está acotada.
2. Si $\lim_{n\to\infty} z_n = A$, demuéstrese que

$$\lim_{n\to\infty} \frac{1}{n}(z_1+z_2+\ldots+z_n)=A.$$

3. Pruébese que la suma de una serie absolutamente convergente no varía si se cambian de orden los términos.
4. Hágase el análisis completo de la convergencia y convergencia uniforme de la sucesión $\{nz^n\}_1^\infty$.
5. Discútase la convergencia uniforme de la serie

$$\sum_{n=1}^{\infty} \frac{x}{n(1+nx^2)}$$

para valores reales de x.

4. *Límites de funciones analíticas.*—El teorema central relativo a la convergencia de funciones analíticas afirma que el límite de una sucesión uniformemente convergente de funciones analíticas es una función analítica. Deben enunciarse cuidadosamente las hipótesis precisas, que han de ser lo menos restrictivas que sea posible.

Consideremos una sucesión $\{f_n(z)\}$, cada $f_n(z)$, estando definida y siendo analítica en una región Ω_n. La función límite $f(z)$ debe considerarse también en alguna región Ω, y es evidente que si $f(z)$ ha de estar definida en Ω, cada punto de Ω debe pertenecer a todos los Ω_n para n mayor que un cierto n_0. En el caso general n_0 no será el mismo para todos los puntos de Ω, y por esta razón no tendría sentido exigir que la convergencia fuese uniforme en Ω. De hecho, en los casos más típicos las regiones Ω_n forman una sucesión creciente, $\Omega_1 \subset \Omega_2 \subset \ldots \subset \Omega_n \subset \ldots$, y Ω es la unión de las Ω_n. En estas circunstancias no existe una función única $f_n(z)$ que esté definida en todo Ω; sin embargo, puede existir el límite $f(z)$ en todos los puntos de Ω, aunque la convergencia no puede ser uniforme.

Como un ejemplo muy sencillo tomemos $f_n(z)=z/(2z^n+1)$ y sea Ω_n el disco $|z|<2^{-1/n}$. Es prácticamente evidente que $\lim_{n\to\infty} f_n(z)=z$ en el disco $|z|<1$ que elegimos como región Ω. Con objeto de estudiar la convergencia uniforme formemos la diferencia

$$f_n(z)-z=-2z^{n+1}/(2z^n+1).$$

Para cualquier valor dado de z podemos hacer $|z^n|<\epsilon/4$ tomando $n>(\log 4/\epsilon)/(\log 1/|z|)$. Si $\epsilon<1$, se tiene así que $2|z|^{n+1}<\epsilon/2$ y $|1+2z^n|>\frac{1}{2}$, de modo que $|f_n(z)-z|<\epsilon$. Se sigue que la convergencia es uniforme en cualquier disco cerrado $|z|\leqq r<1$ o en cualquier subconjunto de tal disco cerrado.

Con otra formulación, la sucesión del ejemplo precedente $\{f_n(z)\}$ tiende a la función límite $f(z)$ uniformemente en todo subconjunto compacto de la región Ω. En efecto, en un conjunto compacto $|z|$ tiene un máximo $r<1$ y, por tanto, el conjunto está contenido en el disco cerrado $|z|\leqq r$. Esta es la situación típica. Encontraremos que frecuentemente podremos demostrar la convergencia uniforme en todo subconjunto compacto de Ω; por otra parte, esta es la condición natural en el teorema que vamos a demostrar.

TEOREMA 1. *Supongamos que $f_n(z)$ es analítica en la región Ω_n y que la sucesión $\{f_n(z)\}$ converge a una función límite $f(z)$ en una región Ω, uniformemente en todo subconjunto compacto de Ω. Entonces $f(z)$ es analítica en Ω. Además, $f'_n(z)$ converge uniformemente a $f'(z)$ en todo subconjunto compacto de Ω.*

La analiticidad de $f(z)$ se deduce fácilmente mediante el teorema de Morera (Cap. III, Sec. 3-2, 3). Sea $|z-a|\leqq r$ un disco cerrado contenido en Ω; el supuesto implica que este disco está contenido en Ω_n para todo n mayor que un cierto n_0[1]. Si γ es una curva cerrada cualquiera contenida en $|z-a|<r$, por el teorema de Cauchy tenemos que

$$\int_\gamma f_n(z)\,dz=0$$

para $n>n_0$. A causa de la convergencia uniforme sobre γ obtenemos

$$\int_\gamma f(z)\,dz=\lim_{n\to\infty}\int_\gamma f_n(z)\,dz=0,$$

[1] En efecto, las regiones Ω_n forman un recubridor abierto de $|z-a|\leqq r$. El disco es compacto y, por tanto, tiene un subrecubridor finito. Esto significa que está contenido en un Ω_{n_0} fijo.

y por el teorema de Morera se sigue que $f(z)$ es analítica en $|z-a|<r$. En consecuencia, $f(z)$ es analítica en toda la región Ω.

Otra demostración más explícita se basa en la fórmula integral

$$f_n(z)=\frac{1}{2\pi i}\int_C \frac{f_n(\zeta)d\zeta}{\zeta-z},$$

donde C es la circunferencia $|\zeta-a|=r$ y $|z-a|<r$. Haciendo tender n a ∞ obtenemos por la convergencia uniforme

$$f(z)=\frac{1}{2\pi i}\int_C \frac{f(\zeta)d\zeta}{\zeta-z},$$

y esta fórmula prueba que $f(z)$ es analítica en el disco. Partiendo de la fórmula

$$f_n'(z)=\frac{1}{2\pi i}\int_C \frac{f_n(\zeta)d\zeta}{(\zeta-z)^2}$$

el mismo razonamiento da

$$\lim_{n\to\infty} f_n'(z)=\frac{1}{2\pi i}\int_C \frac{f(\zeta)d\zeta}{(\zeta-z)^2}=f'(z),$$

y un sencillo cálculo muestra que la convergencia es uniforme para $|z-a|\leqq\rho<r$. Cualquier subconjunto compacto de Ω puede recubrirse mediante un número finito de tales discos cerrados y, por tanto, la convergencia es uniforme sobre todo subconjunto compacto. El teorema está demostrado, y mediante repetidas aplicaciones se sigue que $f_n^{(k)}(z)$ converge uniformemente a $f^{(k)}(z)$ en todo subconjunto compacto de Ω.

El teorema 1 se debe a Weierstrass, en una formulación equivalente. Es particularmente importante su aplicación a series cuyos términos son funciones analíticas. El teorema puede entonces expresarse en la siguiente forma:

Si una serie de términos analíticos,

$$f(z)=f_1(z)+f_2(z)+\ldots+f_n(z)+\ldots,$$

converge uniformemente en todo subconjunto compacto de una región Ω, entonces la suma $f(z)$ es analítica en Ω y se puede derivar la serie término a término.

La tarea de demostrar la convergencia uniforme en un conjunto compacto de puntos A puede facilitarse mediante el uso del principio del módulo máximo. En efecto, con las notaciones utilizadas en el teorema 1, la diferencia $|f_m(z)-f_n(z)|$ alcanza su máximo en A sobre la frontera de A. Por esta razón, la convergencia uniforme sobre la frontera de A implica la convergencia uniforme en A. Así, p. ej., si las funciones $f_n(z)$ son analíticas en el disco $|z|<1$, y si puede probarse que la sucesión converge uniformemente sobre cada circunferencia $|z|=r_m$, siendo $\lim_{m\to\infty} r_m=1$, resulta aplicable el teorema de Weierstrass, de donde se concluye que la función límite es analítica.

EJERCICIOS

1. Utilizando el teorema de Taylor aplicado a una rama de la función $\log(1+z/n)$, pruébese que

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n=e^z$$

uniformemente en todo conjunto compacto.

2. Pruébese que la serie

$$\zeta(z)=\sum_{n=1}^{\infty} n^{-z}$$

converge para $\mathrm{Re}\, z>1$ y represéntese su derivada en forma de serie.

3. Demuéstrese que

$$(1-2^{1-z})\zeta(z)=1^{-z}-2^{-z}+3^{-z}-\ldots$$

y que la última serie representa una función analítica para $\mathrm{Re}\, z>0$.

4. Si $f(z)$ es analítica para $|z|<1$ y $f(0)=0$, pruébese que la serie

$$f(z)+f(z^2)+f(z^3)+\ldots+f(z^n)+\ldots$$

converge y representa una función analítica para los mismos valores de z.

4-2. Series de potencias.—Una de las propiedades más fundamentales de las funciones analíticas es la de que pueden representarse por medio de series de potencias. Recíprocamente, con excepciones triviales, toda serie de potencias convergente define una función analítica. Las series de potencias son expresiones analíticas muy explícitas, y como tales son extremadamente manejables. No es sorprendente, por tanto, que resulten una herramienta poderosa para el estudio de las funciones analíticas.

1. *El círculo de convergencia.*—Una *serie de potencias* es de la forma

$$a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \ldots + a_n(z - z_0)^n + \ldots, \qquad [4\text{-}3]$$

donde el *centro* z_0 y los *coeficientes* a_n son números complejos arbitrarios. Nuestro problema es discutir la convergencia de la serie y las propiedades de la suma como función de la variable z.

La serie converge trivialmente para $z = z_0$. Este pudiera ser el único valor para el que se verifique la convergencia, o bien la serie converge para todo valor de z. Si no ocurre ninguno de estos dos casos, probaremos que la serie converge dentro de un determinado círculo con centro en z_0, y diverge en el exterior de dicho círculo. Con más precisión, demostraremos el siguiente teorema, debido a Abel:

TEOREMA 2. *Para toda serie de potencias* [4-3] *existe un número* R, $0 \leqq R \leqq \infty$, *llamado radio de convergencia, que goza de las siguientes propiedades:*

I) *La serie converge absolutamente para todo* z *que satisfaga a* $|z - z_0| < R$. *La convergencia es uniforme en todo disco cerrado tal que* $|z - z_0| \leqq \rho < R$.

II) *Si* $|z - z_0| > R$, *los términos de la serie no están acotados y la serie es, por tanto, divergente.*

III) *En* $|z - z_0| < R$, *la suma de la serie es una función analítica. Se puede obtener la derivada derivando término a término, y la serie derivada tiene el mismo radio de convergencia.*

Al círculo $|z - z_0| < R$ se le llama *círculo de convergencia;* no se dice nada respecto a la convergencia *sobre* la circunferencia frontera de dicho círculo.

Para demostrarlo definimos R como el extremo superior de todos los números $r \geqq 0$, con la propiedad de que los números $|a_n|r^n$ estén *acotados*. Sabemos que dicho extremo superior existe y pertenece al intervalo $0 \leqq R \leqq \infty$.

Con esta definición de R, la propiedad II) es trivial. En efecto, para $|z - z_0| > R$ los términos $a_n(z - z_0)^n$ están evidentemente sin acotar. En una serie convergente el término general tiende a cero, y se sigue inmediatamente que los términos están acotados. Se concluye que la serie de potencias diverge para $|z - z_0| > R$.

Para probar I) determinamos un r tal que $|z - z_0| < r < R$. Por hipótesis, $|a_n|r^n \leqq M$ para algún M finito. De esta desigualdad obtenemos

$$|a_n(z - z_0)^n| \leqq M\left(\frac{|z - z_0|}{r}\right)^n.$$

Por tanto, la serie [4-3] tiene como mayorante a la serie geométrica de término general $(|z - z_0|/r)^n$; de la convergencia de esta serie se obtiene que la serie [4-3] es igualmente convergente. Para probar la uniformidad para $|z - z_0| \leqq \rho$ elegimos $\rho < r < R$ y hallamos que la serie geométrica de términos $(\rho/r)^n$ es una mayorante. De acuerdo con el criterio M de Weierstrass, la serie de potencias es, por consiguiente, uniformemente convergente en el disco cerrado.

Por último, III) se deduce por aplicación del teorema 1. Este teorema muestra, en particular, que la serie derivada converge para $|z - z_0| < R$. El hecho de que no pueda converger fuera de la circunferencia $|z - z_0| = R$ se deduce observando que los términos de la serie derivada son en última instancia mayores en valor absoluto que los términos correspondientes de la serie original; por esta razón no pueden estar acotados.

Es posible dar una expresión explícita para R en función de los coeficientes a_n. Si $r < R$, sabemos que $|a_n|r^n \leqq M$ o $r\sqrt[n]{|a_n|} \leqq \sqrt[n]{M}$ para todo n. Haciendo tender n a infinito, obtenemos $\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|} \leqq 1/r$, y puesto que esto es cierto para todo $r < R$, resulta necesariamente que $\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|} \leqq 1/R$. Sea ahora $r > R$ y observemos que $|a_n|r^n \geqq 1$ o $r\sqrt[n]{|a_n|} \geqq 1$ para infinitos valores de r. Esto implica que $\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|} \geqq 1/r$ y, por tanto, que $\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|} \geqq 1/R$. Hemos demostrado así que el radio de convergencia viene dado por la llamada *fórmula de Hadamard*

$$\frac{1}{R} = \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|}.$$

EJERCICIOS

1. Determínense los radios de convergencia de las series de potencias siguientes:

$$\sum n^p z^n, \qquad \sum \frac{z^n}{n!}, \qquad \sum q^{n^2} z^n \qquad (|q| < 1).$$

2. Si $\Sigma a_n z^n$ y $\Sigma b_n z^n$ tienen radios de convergencia R_1 y R_2, respectivamente, pruébese que el radio de convergencia de $\Sigma a_n b_n z^n$ es al menos igual a $R_1 R_2$.

3. Si

$$\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| = R,$$

pruébese que la serie $\Sigma a_n z^n$ tiene radio de convergencia R.

4. ¿Para qué valores de z es convergente la serie

$$\sum_{n=1}^{\infty}\frac{z^n}{1+z^{2n}}?$$

2. *La serie de Taylor.*—Para una serie de potencias con radio de convergencia positivo pongamos

$$f(z) = a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + \ldots + a_n(z-z_0)^n + \ldots \qquad [4\text{-}4]$$

en el interior del círculo de convergencia. La serie puede derivarse cualquier número de veces, obteniéndose

$$f^{(n)}(z) = n!\,a_n + \frac{(n+1)!}{1!}\,a_{n+1}(z-z_0) + \ldots$$

En particular, $f^{(n)}(z_0) = n!\,a_n$, y se puede escribir la serie [4-4] en la forma

$$f(z) = f(z_0) + f'(z_0)(z-z_0) + \ldots + \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n + \ldots \qquad [4\text{-}5]$$

A una serie de esta forma se le llama *serie de Taylor.*

Probaremos ahora que toda función analítica puede desarrollarse en una serie de Taylor convergente. Esto es una consecuencia casi inmediata del desarrollo finito de Taylor dado en el capítulo III, sección 3-3, 1, teorema 8, y de la correspondiente representación del resto. De acuerdo con este teorema, si $f(z)$ es analítica en una región Ω que contenga a z_0, podemos escribir:

$$f(z) = f(z_0) + \frac{f'(z_0)}{1!}(z-z_0) + \ldots + \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n + f_{n+1}(z)(z-z_0)^{n+1},$$

$$f_{n+1}(z) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)\,d\zeta}{(\zeta-z_0)^{n+1}(\zeta-z)}$$

En la última fórmula C es cualquier circunferencia $|z-z_0| = \rho$ tal que el disco cerrado $|z-z_0| \leqq \rho$ esté contenido en Ω.

Si M denota el máximo de $|f(z)|$ sobre C, obtenemos inmediatamente la desigualdad

$$|f_{n+1}(z)(z-z_0)^{n+1}| \leqq \frac{M|z-z_0|^{n+1}}{\rho^n(\rho - |z-z_0|)}$$

Puesto que $|z-z_0| < \rho$, tenemos que la serie de Taylor converge y representa a $f(z)$ en cualquier disco circular de centro z_0 que esté contenido en la región Ω.

TEOREMA 3. *Si $f(z)$ es analítica en la región Ω, que contiene a z_0, entonces el desarrollo*

$$f(z) = f(z_0) + \frac{f'(z_0)}{1!}(z-z_0) + \ldots + \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n + \ldots$$

es válido en el mayor disco abierto de centro z_0 contenido en Ω.

Así, pues, el radio de convergencia de la serie de Taylor es por lo menos igual a la mínima distancia de z_0 a la frontera de Ω. Pudiera muy bien ser mayor, pero si es así no existe garantía de que la serie represente todavía a $f(z)$ en todos los puntos que están simultáneamente en Ω y en el círculo de convergencia.

Con la ayuda de la serie de Taylor es fácil hallar desarrollos en series de potencias para todas las funciones elementales. Así, p. ej., se comprueba inmediatamente que

$$e^z = 1 + z + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \ldots$$

$$\operatorname{sen} z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \ldots$$

donde las series convergen y representan a las funciones indicadas en el primer miembro en todo el plano.

Si queremos representar una potencia fraccionaria de z o $\log z$ mediante una serie de potencias, debemos en primer lugar elegir una rama bien definida, y en segundo lugar tenemos que elegir un centro $z_0 \neq 0$. Esto equivale a desarrollar la función $(1+z)^\mu$ o $\log(1+z)$ alrededor del origen, eligiendo la rama que es igual, respectivamente, a 1 o a 0 en el origen. Puesto que esta rama es uniforme y analítica en $|z|<1$, el radio de convergencia vale por lo menos 1. El cómputo de los coeficientes es elemental, obteniéndose

$$(1+z)^\mu = 1+\mu z+\binom{\mu}{2}z^2+\ldots+\binom{\mu}{n}z^n+\ldots$$

$$\log(1+z) = z-\frac{z^2}{2}+\frac{z^3}{3}-\frac{z^4}{4}+\frac{z^5}{5}-\ldots$$

estando definidos los coeficientes del desarrollo del binomio por

$$\binom{\mu}{n} = \frac{\mu(\mu-1)\ldots(\mu-n+1)}{1\cdot 2\cdots n}$$

Si la serie logarítmica tuviera un radio de convergencia mayor que 1, entonces $\log(1+z)$ estaría acotado para $|z|<1$. Puesto que este no es el caso, el radio de convergencia debe valer exactamente 1. Análogamente, si la serie binómica fuera convergente en un círculo de radio mayor que 1, la función $(1+z)^\mu$ y todas sus derivadas estarían acotadas en $|z|<1$. A no ser que μ sea un entero positivo, una de las derivadas será una potencia negativa de $1+z$ y, por tanto, no estará acotada. Así, pues, el radio de convergencia es precisamente 1, excepto en el caso trivial en que la serie binómica se reduzca a un polinomio.

Los desarrollos en serie de las funciones ciclométricas $\operatorname{arc\,tg} z$ y $\operatorname{arc\,sen} z$ se obtienen con toda facilidad considerando las series derivadas. A partir del desarrollo

$$\frac{1}{1+z^2} = 1-z^2+z^4-z^6+\ldots$$

obtenemos por integración

$$\operatorname{arc\,tg} z = z-\frac{z^3}{3}+\frac{z^5}{5}-\frac{z^7}{7}+\ldots$$

estando la rama determinada de manera única como

$$\operatorname{arc\,tg} z = \int_0^z \frac{dz}{1+z^2}$$

para cualquier camino interior al círculo unidad. Para su justificación podemos, o bien apoyarnos en la convergencia uniforme, o bien aplicar el teorema 2. El radio de convergencia no puede ser mayor que el de la serie derivada, siendo, por tanto, exactamente igual a 1.

Si $\sqrt{1-z^2}$ es la rama con parte real positiva, tenemos

$$\frac{1}{\sqrt{1-z^2}} = 1+\frac{1}{2}z^2+\frac{1\cdot 3}{2\cdot 4}z^4+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}z^6+\ldots$$

para $|z|<1$, y mediante integración obtenemos

$$\operatorname{arc\,sen} z = z+\frac{1}{2}\,\frac{z^3}{3}+\frac{1\cdot 3}{2\cdot 4}\,\frac{z^5}{5}+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\,\frac{z^7}{7}+\ldots$$

La serie representa la rama principal de $\operatorname{arc\,sen} z$ con una parte real entre $-\pi/2$ y $\pi/2$.

Para combinaciones de las funciones elementales, en la mayoría de los casos no es posible encontrar una ley general para los coeficientes. Sin embargo, con objeto de hallar unos cuantos de los primeros coeficientes no necesitamos calcular las derivadas sucesivas. Existen técnicas sencillas que nos permiten computar, con una cantidad de trabajo razonable, todos los coeficientes que podamos necesitar.

Es conveniente introducir la notación $[z^n]$ para cualquier función que sea analítica y tenga un cero de al menos orden n en el origen; con menos precisión, $[z^n]$ denota una función que «contiene el factor z^n». Con esta notación, cualquier función que sea analítica en el origen puede escribirse en la forma

$$f(z) = a_0+a_1z+\ldots+a_nz^n+[z^{n+1}],$$

estando determinados los coeficientes de manera única y siendo iguales a los coeficientes del desarrollo de Taylor de $f(z)$. Así, pues, a fin de hallar los primeros n coeficientes del desarrollo de Taylor, basta con determinar un polinomio $P_n(z)$ tal que $f(z)-P_n(z)$ tenga un cero de al menos orden $n+1$ en el origen. El grado de $P_n(z)$ carece de interés; en cualquier caso, es cierto que los coeficientes de z^m, $m \leqq n$, son los coeficientes del desarrollo de Taylor de $f(z)$.

Así, p. ej., supongamos que

$$f(z) = a_0+a_1z+a_2z^2+\ldots+a_nz^n+\ldots$$
$$g(z) = b_0+b_1z+b_2z^2+\ldots+b_nz^n+\ldots$$

Con una notación abreviada escribiremos $f(z) = P_n(z) + [z^{n+1}]$, $g(z) = Q_n(z) + [z^{n+1}]$. Es claro entonces que $f(z)g(z) = P_n(z)Q_n(z) + [z^{n+1}]$, y los coeficientes de los términos de grado $\leqq n$ en P_nQ_n son los coeficientes del desarrollo de Taylor del producto $f(z)g(z)$. De manera explícita obtenemos

$$f(z)g(z) = a_0b_0 + (a_0b_1 + a_1b_0)z + \ldots + (a_0b_n + a_1b_{n-1} + \ldots + a_nb_0)z^n + \ldots$$

Al deducir tal desarrollo no hemos mencionado siquiera la cuestión de la convergencia, pero dado que esta expresión coincide con el desarrollo de Taylor de $f(z)g(z)$, se deduce por el teorema 3 que el radio de convergencia es al menos igual al menor de los radios de convergencia de las series $f(z)$ y $g(z)$ dadas. En el cálculo práctico de P_nQ_n, naturalmente no es necesario determinar los términos de grado mayor que n.

En el caso de un cociente $f(z)/g(z)$, se puede aplicar el mismo método con tal que $g(0) = b_0 \neq 0$. Mediante el uso de la división ordinaria, continuada hasta que el resto contenga el factor z^{n+1}, podemos determinar un polinomio R_n tal que $P_n = Q_nR_n + [z^{n+1}]$. Entonces $f - R_ng = [z^{n+1}]$, y puesto que $g(0) \neq 0$, tenemos que $f/g = R_n + [z^{n+1}]$. Los coeficientes de R_n son los coeficientes de Taylor de $f(z)/g(z)$. Pueden determinarse explícitamente en forma de determinante, pero las expresiones son demasiado complicadas para que resulten de utilidad esencial.

También es importante conocer cómo formar el desarrollo de una función compuesta $f[g(z)]$. En este caso, si se desarrolla $g(z)$ alrededor de z_0, el desarrollo de $f(w)$ tiene que ser en potencias de $w - g(z_0)$. Para simplificar, supongamos que $z_0 = 0$ y que $g(0) = 0$. Podemos poner entonces

$$f(w) = a_0 + a_1w + \ldots + a_nw^n + \ldots$$

y $g(z) = b_1z + b_2z^2 + \ldots + b_nz^n + \ldots$. Utilizando las mismas notaciones que antes, escribimos $f(w) = P_n(w) + [w^{n+1}]$ y $g(z) = Q_n(z) + [z^{n+1}]$ con $Q_n(0) = 0$. Sustituyendo $w = g(z)$, observemos que

$$P_n(Q_n + [z^{n+1}]) = P_n[Q_n(z)] + [z^{n+1}]$$

y que cualquier expresión de la forma $[w^{n+1}]$ se convierte en un $[z^{n+1}]$. Hemos obtenido así $f[g(z)] = P_n[Q_n(z)] + [z^{n+1}]$, y los coeficientes de Taylor de $f[g(z)]$ son los coeficientes de $P_n[Q_n(z)]$ para las potencias $\leqq n$.

Por último, debemos ser capaces de desarrollar la función inversa de una función analítica $w = g(z)$. Aquí podemos suponer que $g(0) = 0$, y se trata de hallar la rama de la función inversa $z = g^{-1}(w)$ que sea analítica en un entorno del origen y se anule para $w = 0$. Para la existencia de la función inversa es necesario y suficiente que $g'(0) \neq 0$; podemos suponer, por tanto, que

$$g(z) = a_1z + a_2z^2 + \ldots = Q_n(z) + [z^{n+1}]$$

con $a_1 \neq 0$. Nuestro problema es determinar un polinomio $P_n(w)$ tal que $P_n[Q_n(z)] = z + [z^{n+1}]$. En efecto, en el supuesto de que $a_1 \neq 0$, las notaciones $[z^{n+1}]$ y $[w^{n+1}]$ son intercambiables, y de $z = P_n[Q_n(z)] + [z^{n+1}]$ obtenemos $z = P_n\{g(z) + [z^{n+1}]\} + [z^{n+1}] = P_n(w) + [w^{n+1}]$. Por tanto, $P_n(w)$ determina los coeficientes de $g^{-1}(w)$.

Con objeto de probar la existencia de un polinomio P_n, procedemos por inducción. Evidentemente, podemos tomar $P_1(w) = w/a_1$. Si se da P_{n-1}, ponemos $P_n = P_{n-1} + b_nz^n$, y obtenemos

$$\begin{aligned} P_n[Q_n(z)] &= P_{n-1}[Q_n(z)] + b_na_1^nz^n + [z^{n+1}] = \\ &= P_{n-1}[Q_{n-1}(z) + a_nz^n] + b_na_1^nz^n + [z^{n+1}] = \\ &= P_{n-1}[Q_{n-1}(z)] + P'_{n-1}[Q_{n-1}(z)]a_nz^n + b_na_1^nz^n + [z^{n+1}]. \end{aligned}$$

En el último miembro los dos primeros términos forman un polinomio conocido de la forma $z + c_nz^n + [z^{n+1}]$, y únicamente tenemos que tomar $b_n = -c_na_1^{-n}$.

Para fines prácticos se halla el desarrollo de la función inversa mediante sustituciones sucesivas. Para ilustrar el método determinemos el desarrollo de tg w a partir de la serie

$$w = \text{arc tg}\, z = z - \frac{z^3}{3} + \frac{z^5}{5} - \ldots$$

Si queremos que el desarrollo incluya potencias quintas, escribimos

$$z = w + \frac{z^3}{3} - \frac{z^5}{5} + [z^7]$$

y sustituimos esta expresión en los términos del segundo miembro. Con restos apropiados, obtenemos

$$z = w + \frac{1}{3}\left(w + \frac{z^3}{3} + [w^5]\right)^3 - \frac{1}{5}(w + [w^3])^5 + [w^7] =$$

$$= w + \frac{1}{3}w^3 + \frac{1}{3}w^2z^3 - \frac{1}{5}w^5 + [w^7] =$$

$$= w + \frac{1}{3}w^3 + \frac{1}{3}w^2(w + [w^3])^3 - \frac{1}{5}w^5 + [w^7] = w + \frac{1}{3}w^3 + \frac{2}{15}w^5 + [w^7].$$

Así, pues, el desarrolo de tg w empieza con los términos

$$\text{tg } w = w + \frac{1}{3}w^3 + \frac{2}{15}w^5 + \ldots$$

EJERCICIOS

1. Desarróllese $1/(1+z^2)$ en potencias de $z-a$, siendo a un número real. Hállese el coeficiente general, y para $a=1$ redúzcase a la forma más sencilla.

2. Los polinomios de Legendre se definen como los coeficientes $P_n(\alpha)$ del desarrollo

$$(1-2\alpha z+z^2)^{-\frac{1}{2}} = 1 + P_1(\alpha)z + P_2(\alpha)z^2 + \ldots$$

Hállense P_1, P_2, P_3 y P_4.

3. Desarróllese $\log(\text{sen } z/z)$ en potencias de z hasta el término z^6.

4. ¿Cuál es el coeficiente de z^7 en el desarrollo de Taylor de tg z?

5. Compruébese que la inversión de la serie del seno da la serie del arco seno.

6. Se definen los números de Fibonacci mediante $c_0=0$, $c_1=1$, $c_n=c_{n-1}+c_{n-2}$. Pruébese que las c_n son coeficientes de Taylor de una función racional y determínese una expresión cerrada para c_n.

3. *La serie de Laurent.*—Una serie de la forma

$$b_0 + b_1z^{-1} + b_2z^{-2} + \ldots + b_nz^{-n} + \ldots \qquad [4\text{-}6]$$

puede considerarse como una serie de potencias ordinaria en la variable $1/z$. Por tanto, convergerá fuera de alguna circunferencia $|z|=R$, excepto en el caso extremo en que $R=\infty$; la convergencia es uniforme en toda región $z \geqq \rho > R$ y, por consiguiente, la serie representa una función analítica en la región $|z|>R$. Si la serie

[4-6] se combina con una serie de potencias ordinaria, obtenemos una serie más general de la forma

$$\sum_{n=-\infty}^{+\infty} a_nz^n. \qquad [4\text{-}7]$$

Se dirá que es convergente únicamente si las partes que consisten en las potencias no negativas y en las potencias negativas son convergentes por separado. Puesto que la primera parte converge en un disco $|z|<R_2$ y la segunda serie en una región $|z|>R_1$, existe una región común de convergencia solo si $R_1<R_2$, representando [4-7] una función analítica en la corona circular $R_1<|z|<R_2$.

Recíprocamente, podemos partir de una función analítica $f(z)$ cuya región de definición contenga una corona circular $R_1<|z|<R_2$, o con más generalidad, una corona circular $R_1<|z-a|<R_2$. Probaremos que tal función puede siempre desarrollarse en una serie de potencias general de la forma

$$f(z) = \sum_{n=-\infty}^{+\infty} A_n(z-a)^n.$$

La demostración es extremadamente sencilla. Todo lo que hemos de probar es que $f(z)$ puede escribirse como una suma $f_1(z)+f_2(z)$, siendo analíticas $f_1(z)$ para $|z-a|<R_2$ y $f_2(z)$ para $|z-a|>R_1$, con una singularidad evitable en ∞. Bajo tales circunstancias se puede desarrollar $f_1(z)$ en potencias no negativas de $z-a$, y $f_2(z)$ en potencias no negativas de $1/(z-a)$.

Para hallar la representación $f(z)=f_1(z)+f_2(z)$ definimos $f_1(z)$ por

$$f_1(z) = \frac{1}{2\pi i}\int_{|\zeta-a|=r} \frac{f(\zeta)\,d\zeta}{\zeta-z}$$

para $|z-a|<r<R_2$ y $f_2(z)$ por

$$f_2(z) = -\frac{1}{2\pi i}\int_{|\zeta-a|=r} \frac{f(\zeta)\,d\zeta}{\zeta-z}$$

para $R_1<r<|z-a|$. El valor de r en ambas integrales es irrelevante en tanto que se verifique la desigualdad, pues es una consecuencia inmediata del teorema de Cauchy que el valor de la integral no varía con r con tal que la circunferencia no pase por el punto z. Por tal razón $f_1(z)$ y $f_2(z)$ están definidas de manera única y representan funciones analíticas en $|z-a|<R_2$ y $|z-a|>R_1$, respectivamente. Lo que es más, por el teorema de la integral de Cauchy $f(z)=$ $=f_1(z)+f_2(z)$.

El desarrollo de Taylor de $f_1(z)$ es

$$f_1(z)=\sum_{n=0}^{\infty} A_n(z-a)^n$$

con

$$A_n=\frac{1}{2\pi i}\int_{|\zeta-a|=r}\frac{f(\zeta)d\zeta}{(\zeta-a)^{n+1}} \qquad [4\text{-}8]$$

Con objeto de hallar el desarrollo de $f_2(z)$ efectuamos la transformación $\zeta=a+1/\zeta'$, $z=a+1/z'$. Esta transformación aplica $|\zeta-a|=r$ en $|\zeta'|=1/r$ con orientación negativa, y mediante un sencillo cálculo se obtiene

$$f_2\left(a+\frac{1}{z'}\right)=\frac{1}{2\pi i}\int_{|\zeta'|=\frac{1}{r}}\frac{z'}{\zeta'}\,\frac{f\left(a+\frac{1}{\zeta'}\right)d\zeta'}{\zeta'-z'}=\sum_{n=1}^{\infty}B_n z'^n$$

con

$$B_n=\frac{1}{2\pi i}\int_{|\zeta'|=\frac{1}{r}}\frac{f\left(a+\frac{1}{\zeta'}\right)d\zeta'}{\zeta'^{n+1}}=\frac{1}{2\pi i}\int_{|\zeta-a|=r}f(\zeta)(\zeta-a)^{n-1}d\zeta.$$

Esta fórmula prueba que podemos escribir

$$f(z)=\sum_{n=-\infty}^{+\infty}A_n(z-a)^n,$$

donde todos los coeficientes A_n están determinados por [4-8]. Obsérvese que la integral en [4-8] es independiente de r siempre y cuando $R_1<r<R_2$.

EJERCICIOS

1. Pruébese que el desarrollo de Laurent es único.

2. Sea Ω una región doblemente conexa cuyo complemento consiste en las componentes E_1, E_2. Demuéstrese que toda función analítica $f(z)$ en Ω puede escribirse en la forma $f_1(z)+f_2(z)$, siendo $f_1(z)$ analítica fuera de E_1 y $f_2(z)$ analítica fuera de E_2.

3. Pruébese que el desarrollo de Laurent de $(e^z-1)^{-1}$ en el origen es de la forma

$$\frac{1}{z}-\frac{1}{2}+\sum_{1}^{\infty}(-1)^{k-1}\frac{B_k}{(2k)!}z^{2k-1},$$

donde los números B_k, conocidos como números de Bernoulli, son todos positivos. Calcúlense B_1, B_2, B_3.

4. Exprésese el desarrollo de Taylor de tg z y el desarrollo de Laurent de cotg z utilizando los números de Bernoulli.

4-3. Fracciones simples y factorización.

—Una función racional tiene dos representaciones típicas, una mediante fracciones simples y otra mediante factorización del numerador y del denominador. La presente sección está dedicada a representaciones análogas para funciones meromorfas arbitrarias.

1. *Fracciones simples.*—Si la función $f(z)$ es meromorfa en una región Ω, a cada polo b_ν corresponde una parte singular de $f(z)$, que consiste en la parte del desarrollo de Laurent que contiene las potencias negativas de $z-b_\nu$; tal parte se reduce a un polinomio $P_\nu[1/(z-b_\nu)]$. Resulta tentador restar todas las partes singulares, a fin de obtener una representación

$$f(z)=\sum_{\nu}P_\nu\left(\frac{1}{z-b_\nu}\right)+g(z), \qquad [4\text{-}9]$$

donde $g(z)$ sería analítica en Ω. Sin embargo, la suma del segundo miembro es, en general, infinita, no habiendo garantía de que la serie converja. No obstante, existen muchos casos en los que la serie converge, y lo que es más, frecuentemente es posible determinar de manera explícita $g(z)$ mediante consideraciones generales. En tales casos el resultado es muy satisfactorio; se obtiene un desarrollo sencillo que con toda seguridad será muy útil.

Si la serie en [4-9] no converge, es necesario modificar el método. Es evidente que no se pierde nada esencial si quitamos una función analítica $p_\nu(z)$ de cada parte singular P_ν. Mediante una elección juiciosa de las funciones p_ν, la serie $\sum_\nu (P_\nu - p_\nu)$ puede hacerse convergente. Incluso es posible tomar como $p_\nu(z)$ polinomios.

No demostraremos el teorema más general a estos efectos. En el caso en que Ω sea todo el plano, probaremos, no obstante, que toda función meromorfa admite un desarrollo en fracciones simples, y lo que es más, que las partes singulares pueden describirse arbitrariamente. El teorema y su generalización a regiones arbitrarias se debe a Mittag-Leffler.

TEOREMA 4. *Sea $\{b_\nu\}$ una sucesión de números complejos con $\lim_{\nu\to\infty} b_\nu = \infty$, y sean $P_\nu(\zeta)$ polinomios sin término constante. Existen entonces funciones meromorfas en todo el plano con polos en los puntos b_ν y partes singulares correspondientes $P_\nu[1/(z-b_\nu)]$. Además la función meromorfa más general de esta clase puede escribirse en la forma*

$$f(z) = \sum_\nu \left[P_\nu\left(\frac{1}{z-b_\nu}\right) - p_\nu(z) \right] + g(z), \qquad [4\text{-}10]$$

donde los $p_\nu(z)$ son polinomios fijos escogidos adecuadamente y $g(z)$ es analítica en todo el plano.

Puesto que la función $P_\nu[1/(z-b_\nu)]$ es analítica para $|z| < |b_\nu|$, puede desarrollarse en serie de Taylor alrededor del origen[1]. Elegimos como $p_\nu(z)$ una suma parcial de esta serie, terminando, p. ej., en el término de grado n_ν. Puede acotarse la diferencia $P_\nu - p_\nu$ mediante el uso de la expresión explícita para el resto dada en el capítulo III, sección 3-3, 1. Si $|P_\nu| \leqq M_\nu$ para $|z| \leqq |b_\nu|/2$, obtenemos, p. ej.,

$$\left| P_\nu\left(\frac{1}{z-b_\nu}\right) - p_\nu(z) \right| \leqq M_\nu \left(\frac{4|z|}{|b_\nu|}\right)^{n_\nu+1} \qquad [4\text{-}11]$$

para $|z| \leqq |b_\nu|/4$. A causa de esta acotación es evidente que la serie del segundo miembro de [4-10] puede hacerse convergente mediante la elección de un n suficientemente grande. Utilizando la fórmula para el radio de convergencia, hallamos concretamente que la serie de potencias

$$\sum_\nu M_\nu \left(\frac{4z}{b_\nu}\right)^{n_\nu+1}$$

[1] Suponemos, para mayor sencillez, que ningún b_ν es igual a cero.

converge en todo el plano si $\lim_{\nu\to\infty} M_\nu^{1/n_\nu}/|b_\nu| = 0$; esto está asegurado si se elige, p. ej., $n_\nu > \log M_\nu$.

Consideremos ahora un disco cerrado arbitrario $|z| \leqq R$. La serie $\sum_\nu (P_\nu - p_\nu)$ tiene solo un número finito de términos que se hagan infinitos en $|z| \leqq R$, y la desigualdad [4-11] se verificará en todo el disco a partir de un cierto término en adelante. Si los términos con $|b_\nu| \leqq R$ se omiten, se sigue que la serie que queda converge absoluta y uniformemente en $|z| \leqq R$. Puesto que R es arbitrario, la serie converge para todo $z \neq b_\nu$ y representa una función meromorfa en todo el plano. Es obvio que las partes singulares son $P_\nu[1/(z-b_\nu)]$; el resto del teorema se deduce de forma trivial.

Como un primer ejemplo consideremos la función $\pi^2/\mathrm{sen}^2\,\pi z$, que tiene polos dobles en los puntos $z = n$ para n entero. La parte singular en el origen es $1/z^2$, y puesto que $\mathrm{sen}^2\,\pi(z-n) = \mathrm{sen}^2\,\pi z$, la parte singular en $z = n$ es $1/(z-n)^2$. La serie

$$\sum_{n=-\infty}^{+\infty} \frac{1}{(z-n)^2} \qquad [4\text{-}12]$$

es convergente para $z \neq n$, como puede verse por comparación con la serie $\sum_1^\infty 1/n^2$. Es uniformemente convergente en cualquier conjunto compacto después de prescindir de los términos que se hacen infinitos en el conjunto. Por esta razón podemos escribir

$$\frac{\pi^2}{\mathrm{sen}^2\,\pi z} = \sum_{n=-\infty}^{+\infty} \frac{1}{(z-n)^2} + g(z), \qquad [4\text{-}13]$$

siendo $g(z)$ analítica en todo el plano. Afirmamos que $g(z)$ es idénticamente cero.

Para probarlo observemos que la función $\pi^2/\mathrm{sen}^2\,\pi z$ y la serie [4-12] son ambas periódicas con período 1. Por consiguiente, la función $g(z)$ tiene el mismo período. Para $z = x + iy$ tenemos (Cap. II, Sec. 2-1, 5)

$$|\mathrm{sen}\,\pi z|^2 = \tfrac{1}{2}(\cosh 2\pi y - \cos 2\pi x),$$

y, por tanto, $\pi^2/\mathrm{sen}^2\,\pi z$ tiende uniformemente a cero cuando $|y| \to \infty$.

Pero se ve fácilmente que la función [4-12] tiene la misma propiedad. En efecto, la convergencia es uniforme para $|y| \geqq 1$, p. ej., y el límite para $|y| \to \infty$ puede obtenerse así tomando límites en cada término. Se concluye que $g(z)$ tiende uniformemente a cero para $|y| \to \infty$. Esto es suficiente para inferir que $|g(z)|$ está acotada en la banda correspondiente a un período, $0 \leqq x \leqq 1$, con lo que lo estará en todo el plano como consecuencia de la periodicidad. Por el teorema de Liouville, $g(z)$ debe reducirse a una constante, y puesto que el límite es cero, esta constante tiene que anularse. Hemos demostrado así la identidad

$$\frac{\pi^2}{\operatorname{sen}^2 \pi z} = \sum_{-\infty}^{\infty} \frac{1}{(z-n)^2} \tag{4-14}$$

A partir de esta ecuación vamos a deducir mediante integración otra identidad relacionada con ella. El primer miembro es la derivada de $-\pi \operatorname{cotg} \pi z$, siendo los términos de la derecha derivadas de $-1/(z-n)$. La serie de término general $1/(z-n)$ diverge, y debe restarse de todos los términos con $n \neq 0$ una suma parcial de la serie de Taylor. En realidad, es suficiente sustraer los términos constantes, pues la serie

$$\sum_{n \neq 0} \left(\frac{1}{z-n} + \frac{1}{n} \right) = \sum_{n \neq 0} \frac{z}{n(z-n)}$$

es comparable con $\sum_1^{\infty} 1/n^2$ y, por tanto, convergente. La convergencia es uniforme en todo conjunto compacto con tal que se omitan los términos que se hacen infinitos. Por esta razón está permitida la derivación término a término, y se obtiene:

$$\pi \operatorname{cotg} \pi z = \frac{1}{z} + \sum_{n \neq 0} \left(\frac{1}{z-n} + \frac{1}{n} \right) \tag{4-15}$$

excepto por una constante aditiva. Si los términos correspondientes a n y $-n$ se agrupan dentro de un mismo paréntesis, se puede escribir [4-15] en las formas equivalentes

$$\pi \operatorname{cotg} \pi z = \lim \sum_{n=-m}^{m} \frac{1}{z-n} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2} \tag{4-16}$$

Escrito el desarrollo de esta forma, es obvio que ambos miembros de la ecuación son funciones impares de z, por cuya razón debe anularse la constante de integración. Resulta así que las ecuaciones [4-15] y [4-16] están correctamente escritas.

Invirtamos ahora el procedimiento y tratemos de calcular la suma análoga

$$\lim_{m \to \infty} \sum_{-m}^{m} \frac{(-1)^n}{z-n} = \frac{1}{z} + \sum_{1}^{\infty} (-1)^n \frac{2z}{z^2 - n^2} \tag{4-17}$$

que evidentemente representa una función meromorfa. Es perfectamente natural separar los términos pares y los impares y escribir

$$\sum_{-(2k+1)}^{2k+1} \frac{(-1)^n}{z-n} = \sum_{n=-k}^{k} \frac{1}{z-2n} - \sum_{n=-k-1}^{k} \frac{1}{z-1-2n}$$

Por comparación con [4-16], hallamos que el límite es

$$\frac{\pi}{2} \operatorname{cotg} \frac{\pi z}{2} - \frac{\pi}{2} \operatorname{cotg} \frac{\pi(z-1)}{2} = \frac{\pi}{\operatorname{sen} \pi z},$$

con lo que hemos demostrado que

$$\frac{\pi}{\operatorname{sen} \pi z} = \lim_{m \to \infty} \sum_{-m}^{m} (-1)^n \frac{1}{z-n} \tag{4-18}$$

EJERCICIOS

1. Comparando coeficientes en el desarrollo de Laurent de $\operatorname{cotg} \pi z$ y de su expresión como una suma de fracciones simples, hállense los valores de

$$\sum_1^{\infty} \frac{1}{n^2}, \quad \sum_1^{\infty} \frac{1}{n^4}, \quad \sum_1^{\infty} \frac{1}{n^6}$$

Se dará una completa justificación de los pasos necesarios.

2. Exprésese

$$\sum_{-\infty}^{\infty} \frac{1}{z^3 - n^3}$$

en forma cerrada.

2. *Productos infinitos.*—Un producto infinito de números complejos

$$p_1 p_2 \cdots p_n \ldots = \prod_{n=1}^{\infty} p_n \qquad [4\text{-}19]$$

se calcula tomando el límite de los productos parciales $P_n = p_1 p_2 \cdots p_n$. Se dice que converge al valor $P = \lim_{n\to\infty} P_n$ si este límite existe y es diferente de cero. Existen buenas razones para excluir el valor cero. Por una parte, si se admitiese el valor cero para P, cualquier producto infinito con un factor 0 sería convergente, y la convergencia no dependería de toda la sucesión de factores. Por otra, en ciertas situaciones este convenio es demasiado radical. En efecto, se desea expresar una función como producto infinito, y esto tiene que ser posible aun cuando la función tenga ceros. Por esta razón convenimos lo siguiente: Se dice que el producto infinito [4-19] converge si (y solo si) a lo sumo son cero un número finito de factores y si los productos parciales formados por los factores que no se anulan tienden a un límite finito diferente de cero.

En un producto convergente el factor general p_n tiende a 1; esto se ve claramente escribiendo $p_n = P_n/P_{n-1}$, habiendo omitido los factores cero. Como consecuencia de este hecho es preferible escribir todos los productos infinitos en la forma

$$\prod_{n=1}^{\infty} (1 + a_n) \qquad [4\text{-}20]$$

de manera que $a_n \to 0$ sea una condición necesaria para la convergencia.

Si ningún factor es cero, es natural comparar el producto [4-20] con la serie infinita

$$\sum_{n=1}^{\infty} \log (1 + a_n). \qquad [4\text{-}21]$$

Puesto que los a_n son números complejos, debemos adoptar una rama definida para los logaritmos, decidiéndonos por la elección de la rama principal en cada término. Denotemos las sumas parciales de [4-21] por S_n. Entonces $P_n = e^{S_n}$, y si $S_n \to S$, se sigue que P_n tiende al límite $P = e^S$, que es distinto de cero. En otras palabras: la convergencia de [4-21] es una condición suficiente para la convergencia de [4-20].

Con objeto de probar que la condición es también necesaria, supongamos que $P_n \to P \neq 0$ y elijamos un valor fijo de $\log P$, p. ej., el valor de la rama principal. Con el valor correspondiente de $\arg P$ determinemos $\arg P_n$ por la condición $\arg P - \pi < \arg P_n \leq \arg P + \pi$ y hagamos $\log P_n = \log |P_n| + i \arg P_n$. Sabemos, por otra parte, que $S_n = \log P_n + h_n \cdot 2\pi i$, siendo h_n un entero bien determinado. Para dos términos consecutivos se obtiene

$$(h_{n+1} - h_n) 2\pi i = \log (1 + a_{n+1}) + \log P_n - \log P_{n+1}.$$

Nos interesan únicamente las partes imaginarias. Cuando n sea suficientemente grande tendremos, p. ej., que $|\arg (1 + a_{n+1})| < 2\pi/3$, $|\arg P_n - \arg P| < 2\pi/3$ y $|\arg P_{n+1} - \arg P| < 2\pi/3$. Estas desigualdades implican $|h_{n+1} - h_n| < 1$, de donde sigue que $h_{n+1} = h_n$ para todo n suficientemente grande. En definitiva, h_n es, por consiguiente, constantemente igual a un cierto entero h, y tenemos que S_n tiende al límite $S = \log P + h \cdot 2\pi i$. Hemos demostrado así la siguiente proposición:

TEOREMA 5. *El producto infinito* $\prod_1^{\infty} (1 + a_n)$ *con* $1 + a_n \neq 0$ *converge simultáneamente con la serie* $\sum_1^{\infty} \log (1 + a_n)$, *cuyos términos representan los valores de la rama principal del logaritmo.*

La cuestión de la convergencia de un producto puede reducirse, pues, al problema más familiar relativo a la convergencia de una serie. Se puede reducir más observando que la serie [4-21] converge absolutamente a la vez que la serie más sencilla $\Sigma |a_n|$. Esta es una consecuencia inmediata del hecho de que

$$\lim_{z \to 0} \frac{\log (1 + z)}{z} = 1.$$

Si cualquiera de las dos series [4-21] o $\sum_1^{\infty} |a_n|$ converge, tenemos que $a_n \to 0$, y para un $\epsilon > 0$ dado se verifica la doble desigualdad

$$(1 - \epsilon)|a_n| < |\log (1 + a_n)| < (1 + \epsilon)|a_n|$$

para n suficientemente grande. Consecuencia inmediata es que las dos series son, de hecho, absolutamente convergentes a la vez.

Se dice que un producto es absolutamente convergente si (y solo si) la serie [4-21] correspondiente converge absolutamente. Con esta terminología podemos enunciar nuestro resultado en los términos siguientes:

TEOREMA 6. *Una condición necesaria y suficiente para la convergencia absoluta del producto* $\prod_1^\infty (1+a_n)$ *es la convergencia de la serie* $\sum_1^\infty |a_n|$.

En el último teorema se hace hincapié en la convergencia absoluta. Puede probarse con ejemplos sencillos que la convergencia de $\sum_1^\infty a_n$ no es necesaria ni suficiente para la convergencia del producto $\prod_1^\infty (1+a_n)$.

Está claro lo que se entiende por un producto infinito uniformemente convergente cuyos factores son funciones de una variable. La presencia de ceros es posible que origine algunas ligeras dificultades, que generalmente pueden evitarse considerando solo conjuntos sobre los que a lo sumo se puedan anular un número finito de factores. Si se omiten estos factores, es suficiente estudiar la convergencia uniforme del producto restante. Existen proposiciones análogas a los teoremas 5 y 6 para la convergencia uniforme. Si examinamos las demostraciones hallaremos que se pueden uniformar todas las acotaciones, conduciendo las conclusiones a la convergencia uniforme.

EJERCICIOS

1. Pruébese que

$$\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right)=\frac{1}{2}$$

2. Demuéstrese que para $|z|<1$

$$\prod_{n=0}^{\infty}(1+z^{2^n})=\frac{1}{1-z}$$

3. Pruébese que

$$\prod_1^\infty\left(1+\frac{z}{n}\right)e^{-z/n}$$

converge absoluta y uniformemente en todo conjunto compacto.

4. Demuéstrese que el valor de un producto absolutamente convergente no varía si se reordenan los factores.

5. Pruébese que la función

$$\theta(z)=\prod_1^\infty(1+h^{2n-1}e^z)(1+h^{2n-1}e^{-z}),$$

donde $|h|<1$ es analítica en todo el plano y satisface la ecuación funcional

$$\theta(z+2\log h)=h^{-1}e^{-z}\,\theta(z).$$

3. *Productos canónicos.*—Una función que es analítica en todo el plano se dice que es *entera*. Las funciones enteras más sencillas, aparte los polinomios, son e^z, sen z y cos z.

Si $g(z)$ es una función entera, entonces $f(z)=e^{g(z)}$ es entera y distinta de cero. Recíprocamente, si $f(z)$ es una función entera cualquiera que no se anula nunca, probaremos que $f(z)$ es de la forma $e^{g(z)}$. Con este objeto observemos que la función $f'(z)/f(z)$, siendo analítica en todo el plano, es la derivada de una función entera $g(z)$. De este hecho se deduce, mediante cálculo, que $f(z)e^{-g(z)}$ tiene derivada nula y, por consiguiente, $f(z)$ es un múltiplo constante de $e^{g(z)}$; se puede absorber la constante en $g(z)$.

Por este método también podemos hallar la función entera más general con un número finito de ceros. Supongamos que $f(z)$ tiene m ceros en el origen (m pudiera ser cero) y denotemos los otros ceros por $a_1, a_2, \ldots, a_N$, repitiendo los ceros múltiples. Es evidente entonces que es posible escribir

$$f(z)=z^m e^{g(z)}\prod_1^N\left(1-\frac{z}{a_n}\right).$$

Si hay infinitos ceros, podemos tratar de hallar una representación análoga mediante un producto infinito. La generalización obvia sería

$$f(z)=z^m e^{g(z)}\prod_1^\infty\left(1-\frac{z}{a_n}\right). \qquad [4\text{-}22]$$

Esta representación es válida si el producto infinito converge uniformemente en todo conjunto compacto. De hecho, si es así, el producto representa una función entera con ceros en los mismos puntos (excepto en el origen) y con las mismas multiplicidades que $f(z)$. Se sigue que es posible escribir el cociente en la forma $z^m e^{g(z)}$.

El producto que figura en [4-22] converge absolutamente si (y solo si) $\sum_1^\infty 1/|a_n|$ es convergente, en cuyo caso la convergencia también es uniforme en todo disco cerrado $|z| \leqq R$. Unicamente bajo esta condición especial podemos obtener una representación de la forma [4-22].

En el caso general tienen que introducirse factores que produzcan convergencia. Consideremos una sucesión arbitraria de números complejos $a_n \neq 0$ con $\lim_{n\to\infty} a_n = \infty$; probaremos la existencia de polinomios $p_n(z)$ tales que

$$\prod_1^\infty \left(1 - \frac{z}{a_n}\right) e^{p_n(z)} \qquad [4\text{-}23]$$

converge a una función entera. El producto converge al mismo tiempo que la serie de término general

$$r_n(z) = \log\left(1 - \frac{z}{a_n}\right) + p_n(z),$$

donde se elegirá la rama del logaritmo de forma que la parte imaginaria de $r_n(z)$ esté comprendida entre $-\pi$ y π (inclusive).

Para un R dado consideremos únicamente los términos con $|a_n| > R$. En el disco $|z| \leqq R$ se puede desarrollar en serie de Taylor la rama principal de $\log(1 - z/a_n)$

$$\log\left(1 - \frac{z}{a_n}\right) = -\frac{z}{a_n} - \frac{1}{2}\left(\frac{z}{a_n}\right)^2 - \frac{1}{3}\left(\frac{z}{a_n}\right)^3 - \ldots$$

Elegimos para $p_n(z)$ una suma parcial de esta serie que acabe en el término de grado m_n. Entonces $r_n(z)$ admite la representación

$$r_n(z) = \frac{1}{m_n+1}\left(\frac{z}{a_n}\right)^{m_n+1} + \frac{1}{m_n+2}\left(\frac{z}{a_n}\right)^{m_n+2} + \ldots$$

y obtenemos fácilmente la acotación

$$|r_n(z)| \leqq \frac{1}{m_n+1}\left(\frac{R}{|a_n|}\right)^{m_n+1}\left(1 - \frac{R}{|a_n|}\right)^{-1}. \qquad [4\text{-}24]$$

Supongamos ahora que la serie

$$\sum_{n=1}^\infty \frac{1}{m_n+1}\left(\frac{R}{|a_n|}\right)^{m_n+1} \qquad [4\text{-}25]$$

converge. Por la acotación [4-24] se obtiene en primer lugar que $r_n(z) \to 0$, y, por tanto, $r_n(z)$ tiene una parte imaginaria comprendida entre $-\pi$ y π siempre y cuando n sea suficientemente grande. Además, por comparación, se ve que la serie $\Sigma\, r_n(z)$ es absoluta y uniformemente convergente para $|z| \leqq R$, y así el producto [4-23] representa una función analítica en $|z| < R$. Por motivo del razonamiento debemos excluir los valores $|a_n| \leqq R$, pero es evidente que la convergencia uniforme de [4-23] no es afectada al tomar nuevamente en cuenta los factores correspondientes.

Resta únicamente probar que la serie [4-25] puede hacerse convergente para todo R. Pero esto es obvio, pues si tomamos $m_n = n$, [4-25] se convierte en una serie de potencias con radio de convergencia infinito, como puede comprobarse mediante la fórmula del radio de convergencia o considerando una serie mayorante geométrica para cualquier valor fijo de R.

Teorema 7. *Existe una función entera con ceros arbitrariamente prescritos a_n con tal que, en el caso de infinitos ceros, $a_n \to \infty$. Toda función entera con estos únicos ceros puede escribirse en la forma*

$$f(z) = z^m e^{g(z)} \prod_{n=1}^\infty \left(1 - \frac{z}{a_n}\right) e^{\frac{z}{a_n} + \frac{1}{2}\left(\frac{z}{a_n}\right)^2 + \ldots + \frac{1}{m_n}\left(\frac{z}{a_n}\right)^{m_n}} \qquad [4\text{-}26]$$

en donde el producto se extiende a todos los $a_n \neq 0$, los m_n son ciertos enteros y $g(z)$ es una función entera.

Este teorema se debe a Weierstrass. Tiene el siguiente corolario importante:

Corolario. *Toda función que es meromorfa en todo el plano es el cociente de dos funciones enteras.*

En efecto, si $F(z)$ es meromorfa en todo el plano, podemos hallar una función entera $g(z)$ con los polos de $F(z)$ por ceros. El producto $F(z)g(z)$ es en tal caso una función entera $f(z)$, y obtenemos $F(z)=f(z)/g(z)$.

La representación [4-26] resulta considerablemente más interesante si es posible elegir todos los m_n iguales entre sí. La demostración anterior ha probado que el producto

$$\prod_1^\infty \left(1-\frac{z}{a_n}\right) e^{\frac{z}{a_n}+\frac{1}{2}\left(\frac{z}{a_n}\right)^2+\dots+\frac{1}{h}\left(\frac{z}{a_n}\right)^h} \qquad [4\text{-}27]$$

converge y representa una función entera con tal que la serie $\sum_{n=1}^{\infty} (R/|a_n|)^{h+1}/(h+1)$ converja para todo R; es decir, con tal que $\Sigma\, 1/|a_n|^{h+1} < \infty$. Suponiendo que h es el menor entero para el que esta serie converge, la expresión [4-27] recibe el nombre de *producto canónico* asociado a la sucesión $\{a_n\}$, y h es el *género* del producto canónico.

Siempre y cuando sea posible utilizaremos el producto canónico en la representación [4-26], el cual, por otra parte, está determinado de manera única. Si en esta representación $g(z)$ se reduce a un polinomio, se dice que la función $f(z)$ es de género finito, siendo por definición el género de $f(z)$ igual al grado de este polinomio o al género del producto canónico, según cuál de los dos sea el mayor. Así, p. ej., una función entera de género cero es de la forma

$$Cz^m \prod_1^\infty \left(1-\frac{z}{a_n}\right)$$

con $\Sigma\, 1/|a_n| < \infty$. La representación canónica de una función entera de género 1 es de la forma

$$Cz^m e^{\alpha z} \prod_1^\infty \left(1-\frac{z}{a_n}\right) e^{z/a_n}$$

con $\Sigma\, 1/|a_n|^2 < \infty$, $\Sigma\, 1/|a_n| = \infty$, o de la forma

$$Cz^m e^{\alpha z} \prod_1^\infty \left(1-\frac{z}{a_n}\right)$$

con $\Sigma\, 1/|a_n| < \infty$, $\alpha \neq 0$.

Existe una estrecha relación entre el género de una función y la velocidad de crecimiento de $f(z)$ cuando z tiende a infinito. Volveremos sobre este problema en el capítulo V, sección 5-1, 5.

Como aplicación consideremos la representación de $\operatorname{sen} \pi z$ mediante un producto. Los ceros son los enteros $z = \pm n$. Puesto que $\Sigma\, 1/n$ diverge y $\Sigma\, 1/n^2$ converge, debemos tomar $h=1$, obteniendo una representación de la forma

$$\operatorname{sen} \pi z = z e^{g(z)} \prod_{n \neq 0} \left(1-\frac{z}{n}\right) e^{z/n}.$$

Con objeto de determinar $g(z)$ formamos las derivadas logarítmicas de ambos miembros. Hallamos que

$$\pi \operatorname{cotg} \pi z = \frac{1}{z} + g'(z) + \sum_{n \neq 0} \left(\frac{1}{z-n} + \frac{1}{n}\right)$$

siendo fácil justificar el procedimiento basándose en la convergencia uniforme sobre cualquier conjunto compacto que no contenga los puntos $z=n$. Por comparación con la fórmula previa [4-15], se concluye que $g'(z)=0$. Por tanto, $g(z)$ es una constante, y teniendo en cuenta que $\lim_{z\to 0} \operatorname{sen} \pi z/z = \pi$, deberá ser $e^{g(z)} = \pi$, y así

$$\operatorname{sen} \pi z = \pi z \prod_{n \neq 0} \left(1-\frac{z}{n}\right) e^{z/n}. \qquad [4\text{-}28]$$

En esta representación los factores correspondientes a n y $-n$ pueden agruparse dentro de un mismo paréntesis, con lo que se obtiene la forma más sencilla

$$\operatorname{sen} \pi z = \pi z \prod_1^\infty \left(1-\frac{z^2}{n^2}\right). \qquad [4\text{-}29]$$

Se sigue de [4-28] que $\operatorname{sen} \pi z$ es una función entera de género 1.

EJERCICIOS

1. Pruébese que

$$\operatorname{sen} \pi(z+\alpha) = e^{\pi z \operatorname{cotg} \pi \alpha} \operatorname{sen} \pi\alpha \prod_{-\infty}^{\infty} \left(1+\frac{z}{n+\alpha}\right) e^{-z/(n+\alpha)}$$

siempre y cuando α no sea un entero.

2. ¿Cuál es el género de $\cos \sqrt{z}$?

3. Pruébese que una función entera de género h satisface una desigualdad de la forma

$$|f(z)| < Me^{|z|^{h+1+\eta}}$$

para cualquier $\eta > 0$.

4. Si $f(z)$ es de género h, ¿qué valores máximo y mínimo puede tener el género de $f(z^2)$?

5. Pruébese que si $f(z)$ es de género 0 o 1 con ceros reales, y si $f(z)$ es real para z real, entonces todos los ceros de $f'(z)$ son reales.

4. *La función gamma.*—La función $\operatorname{sen} \pi z$ tiene como ceros todos los números enteros, siendo la función más sencilla con esta propiedad. Introduciremos ahora funciones que tengan únicamente como ceros los enteros positivos o los enteros negativos. Así, p. ej., la función más sencilla con los enteros negativos por ceros es el producto canónico correspondiente

$$G(z) = \prod_1^\infty \left(1 + \frac{z}{n}\right) e^{-z/n}. \qquad [4\text{-}30]$$

Es evidente que $G(-z)$ tiene entonces como ceros los enteros positivos, y por comparación con la representación mediante un producto (Ec. [4-28]) de $\operatorname{sen} \pi z$, se tiene inmediatamente que

$$zG(z)G(-z) = \frac{\operatorname{sen} \pi z}{\pi} \qquad [4\text{-}31]$$

Como consecuencia de la forma en que se ha construido $G(z)$, es lógico que tenga otras sencillas propiedades. Observemos que $G(z-1)$ tiene los mismos ceros que $G(z)$ y además un cero en el origen. Es evidente entonces que podemos escribir

$$G(z-1) = ze^{\gamma(z)}G(z),$$

donde $\gamma(z)$ es una función entera. Con objeto de determinar $\gamma(z)$ tomamos en ambos miembros las derivadas logarítmicas. Esto nos da la ecuación

$$\sum_{n=1}^\infty \left(\frac{1}{z-1+n} - \frac{1}{n} \right) = \frac{1}{z} + \gamma'(z) + \sum_{n=1}^\infty \left(\frac{1}{z+n} - \frac{1}{n} \right). \qquad [4\text{-}32]$$

En la serie del primer miembro podemos reemplazar n por $n+1$. Mediante este cambio obtenemos

$$\sum_{n=1}^\infty \left(\frac{1}{z-1+n} - \frac{1}{n} \right) = \frac{1}{z} - 1 + \sum_{n=1}^\infty \left(\frac{1}{z+n} - \frac{1}{n+1} \right) =$$

$$= \frac{1}{z} - 1 + \sum_{n=1}^\infty \left(\frac{1}{z+n} - \frac{1}{n} \right) + \sum_{n=1}^\infty \left(\frac{1}{n} - \frac{1}{n+1} \right).$$

La última serie tiene suma igual a uno y, por tanto, la ecuación [4-32] se reduce a $\gamma'(z) = 0$. Así, pues, $\gamma(z)$ es una constante, que denotaremos por γ, y $G(z)$ tiene la siguiente propiedad de reproducción, $G(z-1) = e^\gamma zG(z)$. Resulta más sencillo considerar la función $H(z) = G(z)e^{\gamma z}$, que evidentemente satisface la ecuación funcional $H(z-1) = zH(z)$.

El valor de γ se determina con facilidad. Tomando $z=1$, tenemos

$$1 = G(0) = e^\gamma G(1),$$

y, por tanto,

$$e^{-\gamma} = \prod_{n=1}^\infty \left(1 + \frac{1}{n}\right) e^{-1/n}.$$

Aquí el n-ésimo producto parcial puede escribirse en la forma

$$(n+1)e^{-(1+\frac{1}{2}+\frac{1}{3}+\dots+1/n)},$$

y obtenemos

$$\gamma = \lim_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \log n \right).$$

A la constante γ se le llama constante de Euler; su valor aproximado es 0,57722.

Si $H(z)$ satisface la ecuación funcional $H(z-1) = zH(z)$, entonces $\Gamma(z) = 1/[zH(z)]$ verifica la relación

$$\Gamma(z-1) = \Gamma(z)/(z-1),$$

o

$$\Gamma(z+1) = z\Gamma(z). \qquad [4\text{-}33]$$

Se comprueba que esta es una relación más útil, y por esta razón se ha convertido en costumbre incluir entre el reducido conjunto de las funciones elementales a $\Gamma(z)$ bajo el nombre de ***función gamma de Euler.***

Nuestra definición conduce a la representación explícita

$$\Gamma(z)=\frac{e^{-\gamma z}}{z}\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)^{-1}e^{z/n}, \qquad [4\text{-}34]$$

y la fórmula [4-31] toma la forma

$$\Gamma(z)\Gamma(1-z)=\frac{\pi}{\operatorname{sen}\pi z} \qquad [4\text{-}35]$$

Observemos que $\Gamma(z)$ es una función meromorfa con polos en $z=0$, -1, -2, ..., pero *sin ceros*.

Tenemos que $\Gamma(1)=1$, y mediante la ecuación funcional hallamos que $\Gamma(2)=1$, $\Gamma(3)=1\cdot 2$, $\Gamma(4)=1\cdot 2\cdot 3$, y en general $\Gamma(n)=(n-1)!$ La función Γ puede así considerarse como una generalización de la factorial. De [4-35] se obtiene que $\Gamma(\frac{1}{2})=\sqrt{\pi}$.

Se obtienen con toda facilidad otras propiedades considerando la derivada segunda de $\log\Gamma(z)$ para la que hallamos, por [4-34], la expresión muy sencilla

$$\frac{d}{dz}\left(\frac{\Gamma'(z)}{\Gamma(z)}\right)=\sum_{n=0}^{\infty}\frac{1}{(z+n)^2} \qquad [4\text{-}36]$$

Así, p. ej., es evidente que $\Gamma(z)\Gamma(z+\frac{1}{2})$ y $\Gamma(2z)$ tienen los mismos polos, y utilizando [4-36] hallamos que

$$\frac{d}{dz}\left(\frac{\Gamma'(z)}{\Gamma(z)}\right)+\frac{d}{dz}\left(\frac{\Gamma'(z+\frac{1}{2})}{\Gamma(z+\frac{1}{2})}\right)=\sum_{n=0}^{\infty}\frac{1}{(z+n)^2}+\sum_{n=0}^{\infty}\frac{1}{(z+n+\frac{1}{2})^2}=$$

$$=4\left[\sum_{n=0}^{\infty}\frac{1}{(2z+2n)^2}+\sum_{n=0}^{\infty}\frac{1}{(2z+2n+1)^2}\right]=4\sum_{m=0}^{\infty}\frac{1}{(2z+m)^2}=$$

$$=2\frac{d}{dz}\left(\frac{\Gamma'(2z)}{\Gamma(2z)}\right).$$

Integrando obtenemos

$$\Gamma(z)\Gamma(z+\tfrac{1}{2})=e^{az+b}\Gamma(2z),$$

donde quedan por determinar las constantes a y b. Haciendo $z=\frac{1}{2}$ y $z=1$, hacemos uso de los valores conocidos $\Gamma(\frac{1}{2})=\sqrt{\pi}$, $\Gamma(1)=1$, $\Gamma(1\frac{1}{2})=\frac{1}{2}\Gamma(\frac{1}{2})=\frac{1}{2}\sqrt{\pi}$, $\Gamma(2)=1$, con lo cual se obtienen las relaciones

$\frac{1}{2}a+b=\frac{1}{2}\log\pi$, $a+b=\frac{1}{2}\log\pi-\log 2$. Se sigue que $a=-2\log 2$, y que

$$b=\tfrac{1}{2}\log\pi+\log 2;$$

el resultado final es, por tanto,

$$\sqrt{\pi}\,\Gamma(2z)=2^{2z-1}\Gamma(z)\Gamma(z+\tfrac{1}{2}),$$

que se conoce con el nombre de fórmula de duplicación de Legendre.

EJERCICIOS

1. Pruébese la fórmula de Gauss:

$$(2\pi)^{\frac{n-1}{2}}\Gamma(z)=n^{z-\frac{1}{2}}\Gamma\left(\frac{z}{n}\right)\Gamma\left(\frac{z+1}{n}\right)\ldots\Gamma\left(\frac{z+n-1}{n}\right).$$

2. Demuéstrese que

$$\Gamma\left(\frac{1}{6}\right)=2^{-\frac{1}{3}}\left(\frac{3}{\pi}\right)^{\frac{1}{2}}\Gamma\left(\frac{1}{3}\right)^2.$$

5. *Fórmula de Stirling*.—En muchas situaciones en que puede aplicarse la función Γ es de la mayor importancia tener alguna información sobre el comportamiento de $\Gamma(z)$ para valores muy grandes de z. Afortunadamente es posible calcular $\Gamma(z)$ con gran precisión y muy poco esfuerzo mediante una fórmula clásica, que se conoce con el nombre de fórmula de Stirling. Existen muchas demostraciones de esta fórmula. Nosotros hemos decidido deducirla utilizando el cálculo de residuos, siguiendo principalmente la presentación de Lindelöf en su libro clásico sobre el cálculo de residuos. Se trata de una demostración muy sencilla y sobre todo muy instructiva, ya que nos da una oportunidad para utilizar los residuos en casos menos triviales que los anteriores.

El punto de partida es la fórmula [4-36] para la derivada segunda de $\log\Gamma(z)$, y nuestra tarea inmediata es la de expresar la suma parcial

$$\frac{1}{z^2}+\frac{1}{(z+1)^2}+\frac{1}{(z+2)^2}+\ldots+\frac{1}{(z+n)^2}$$

como una integral curvilínea conveniente. Con este objeto necesitamos una función con residuos $1/(z+\nu)^2$ en los puntos enteros ν; una buena elección es

$$\Phi(\zeta)=\frac{\pi \operatorname{cotg} \pi\zeta}{(z+\zeta)^2}$$

La variable en dicha función es ζ, mientras que z aparece solo como una parámetro, el cual se mantendrá en un valor fijo $z=x+iy$ con $x>0$ durante la primera parte de la deducción.

Apliquemos la fórmula de los residuos al rectángulo cuyos lados verticales están sobre $\xi=0$ y $\xi=n+\frac{1}{2}$ y los lados horizontales sobre $\eta=\pm Y$, con la intención de hacer tender primero Y y después n a ∞. Este contorno, que denotaremos por K, pasa por el polo en el origen, pero sabemos que la fórmula sigue siendo válida con tal que tomemos el valor principal de la integral e incluyamos un medio del residuo en el origen. Obtenemos, por consiguiente,

$$\text{v. pr. } \frac{1}{2\pi i}\int_K \Phi(\zeta)\,d\zeta=-\frac{1}{2z^2}+\sum_{\nu=0}^{n}\frac{1}{(z+\nu)^2}$$

En los lados horizontales del rectángulo $\operatorname{cotg}\pi\zeta$ tiende uniformemente a $\pm i$ para $Y\to\infty$. Puesto que el factor $1/(z+\zeta)^2$ tiende a cero, las integrales correspondientes tienen límite cero. Nos quedan ahora dos integrales sobre rectas verticales indefinidas; $\operatorname{cotg}\pi\zeta$ está acotada sobre cada recta $\xi=n+\frac{1}{2}$, y como consecuencia de la periodicidad, la cota es independiente de n. Así, pues, la integral sobre la recta $\xi=n+\frac{1}{2}$ es menor que el producto de una constante por la integral

$$\int_{\xi=n+\frac{1}{2}}\frac{d\eta}{|\zeta+z|^2}$$

Esta integral puede calcularse, ya que sobre la recta de integración se verifica

$$\bar{\zeta}=2n+1-\zeta,$$

y por residuos se obtiene

$$\frac{1}{i}\int\frac{d\zeta}{|\zeta-z|^2}=\frac{1}{i}\int\frac{d\zeta}{(\zeta+z)(2n+1-\zeta+\bar{z})}=\frac{2\pi}{2n+1+2x}$$

El límite para $n\to\infty$ es, por tanto, cero.

Por último, el valor principal de la integral sobre el eje imaginario desde $-i\infty$ hasta $+i\infty$ puede escribirse en la forma

$$\frac{1}{2}\int_0^\infty \operatorname{cotg}\pi i\eta\left[\frac{1}{(i\eta-z)^2}-\frac{1}{(i\eta+z)^2}\right]d\eta=$$

$$-\int_0^\infty \coth\pi\eta\cdot\frac{2\eta z}{(\eta^2+z^2)^2}\,d\eta.$$

Ha de cambiarse el signo, con lo que se obtiene la fórmula

$$\frac{d}{dz}\left(\frac{\Gamma'(z)}{\Gamma(z)}\right)=\frac{1}{2z^2}+\int_0^\infty \coth\pi\eta\cdot\frac{2\eta z}{(\eta^2+z^2)^2}\,d\eta. \qquad [4\text{-}37]$$

Es preferible escribir

$$\coth\pi\eta=1+\frac{2}{e^{2\pi\eta}-1}$$

y observar que la integral obtenida del término 1 tiene el valor $1/z$. Así, pues, podemos volver a escribir [4-37] en la forma

$$\frac{d}{dz}\left(\frac{\Gamma'(z)}{\Gamma(z)}\right)=\frac{1}{z}+\frac{1}{2z^2}+\int_0^\infty\frac{4\eta z}{(\eta^2+z^2)^2}\cdot\frac{d\eta}{e^{2\pi\eta}-1} \qquad [4\text{-}38]$$

La integral converge ahora muy rápidamente.

Haciendo que z varíe en el semiplano derecho, se puede integrar esta fórmula. Se obtiene

$$\frac{\Gamma'(z)}{\Gamma(z)}=C+\log z-\frac{1}{2z}-\int_0^\infty\frac{2\eta}{\eta^2+z^2}\cdot\frac{d\eta}{e^{2\pi\eta}-1}, \qquad [4\text{-}39]$$

donde $\log z$ es la rama principal, y C, una constante de integración. La integración del último término exige alguna justificación. Tenemos que estar seguros de que la integral en [4-39] puede derivarse bajo el signo de integral; en efecto es así, ya que la integral converge uniformemente cuando se restringe la variación de z a cualquier conjunto compacto en el semiplano $x>0$.

Deseamos integrar [4-39] nuevamente. Esto introduciría obviamente $\operatorname{arc\,tg}(z/\eta)$ en la integral, y aunque podría definirse una rama uniforme, preferimos evitar el uso de funciones multiformes. Esto es posible si transformamos primero la integral en [4-39] mediante integración parcial. Se obtiene

$$\int_0^\infty \frac{2\eta}{\eta^2+z^2}\cdot\frac{d\eta}{e^{2\pi\eta}-1}=\frac{1}{\pi}\int_0^\infty \frac{z^2-\eta^2}{(\eta^2+z^2)^2}\log\,(1-e^{-2\pi\eta})\,d\eta,$$

donde, naturalmente, el logaritmo es real. Ahora podemos integrar con respecto a z, con lo que se tiene

$$\log\Gamma(z)=C'+Cz+\left(z-\frac{1}{2}\right)\log z+\frac{1}{\pi}\int_0^\infty \frac{z}{\eta^2+z^2}\log\frac{1}{1-e^{-2\pi\eta}}\,d\eta, \tag{4-40}$$

donde C' es una nueva constante de integración, habiéndose reemplazado, por conveniencia, $C-1$ por C. La fórmula significa que existe, en el semiplano derecho, una rama uniforme de $\log\Gamma(z)$ cuyo valor viene dado por el segundo miembro de la ecuación. Mediante una elección adecuada de C' se obtiene la rama de $\log\Gamma(z)$ que es real para z real.

Quedan por determinar las constantes C y C'. Con este objeto debemos estudiar primero el comportamiento de la integral en [4-40], a la que denotaremos por

$$J(z)=\frac{1}{\pi}\int_0^\infty \frac{z}{\eta^2+z^2}\log\frac{1}{1-e^{-2\pi\eta}}\,d\eta. \tag{4-41}$$

Es prácticamente evidente que $J(z)\to 0$ para $z\to\infty$ con tal que z se mantenga alejado del eje imaginario. Supongamos, p. ej., que se restringe z al semiplano $x\geqq c>0$. Separando la integral en dos partes, tenemos

$$J(z)=\int_0^{\frac{|z|}{2}}+\int_{\frac{|z|}{2}}^\infty=J_1+J_2.$$

En la primera integral $|\eta^2+z^2|\geqq|z|^2-|z/2|^2=3|z|^2/4$, y, por tanto,

$$|J_1|\leqq\frac{4}{3\pi|z|}\int_0^\infty\log\frac{1}{1-e^{-2\pi\eta}}\,d\eta.$$

En la segunda integral $|\eta^2+z^2|=|z-i\eta|\cdot|z+i\eta|>c|z|$, y tenemos

$$|J_2|<\frac{1}{\pi c}\int_{\frac{|z|}{2}}^\infty\log\frac{1}{1-e^{-2\pi\eta}}\,d\eta.$$

Puesto que la integral de $\log\,(1-e^{-2\pi\eta})$ es obviamente convergente, se concluye que J_1 y J_2 tienden a cero cuando $z\to\infty$.

Se halla el valor de C sustituyendo [4-40] en la ecuación funcional $\Gamma(z+1)=z\Gamma(z)$ o $\log\Gamma(z+1)=\log z+\log\Gamma(z)$; si se restringe z a valores positivos, no hay duda sobre la rama del logaritmo. La sustitución nos da

$$C'+Cz+C+(z+\tfrac{1}{2})\log\,(z+1)+J(z+1)=$$
$$=C'+Cz+(z+\tfrac{1}{2})\log z+J(z),$$

que se reduce a

$$C=-\left(z+\frac{1}{2}\right)\log\left(1+\frac{1}{z}\right)+J(z)-J(z+1).$$

Haciendo que $z\to\infty$, hallamos que $C=-1$.

A continuación aplicamos [4-40] a la ecuación $\Gamma(z)\Gamma(1-z)=$ $=\pi/\text{sen}\,\pi z$, eligiendo $z=\frac{1}{2}+iy$. Obtenemos

$$2C'-1+iy\log\,(\tfrac{1}{2}+iy)-iy\log\,(\tfrac{1}{2}-iy)+J(\tfrac{1}{2}+iy)+J(\tfrac{1}{2}-iy)=$$
$$=\log\pi-\log\cosh\pi y.$$

Esta ecuación, en la que los logaritmos se toman con sus valores principales, está demostrada hasta ahora salvo un múltiplo constante de $2\pi i$. Pero para $y=0$ la ecuación es correcta tal y como está escrita, ya que [4-40] determina el valor real de $\log\Gamma(\frac{1}{2})$; se verifica, por tanto, para todo y. Haciendo que $y\to\infty$, sabemos que $J(\frac{1}{2}+iy)$ y $J(\frac{1}{2}-iy)$ tienden a cero. Desarrollando el logaritmo del primer miembro en serie de Taylor tenemos que

$$iy\log\frac{\frac{1}{2}+iy}{\frac{1}{2}-iy}=iy\left(\pi i+\log\frac{1+\dfrac{1}{2iy}}{1-\dfrac{1}{2iy}}\right)=-\pi y+1+\epsilon_1(y)$$

mientras que en el segundo miembro,

$$\log\cosh\pi y=\pi y-\log 2+\epsilon_2(y)$$

con $\epsilon_1(y)$ y $\epsilon_2(y)$ tendiendo a cero. Estas consideraciones dan por resultado el valor de $C'=\frac{1}{2}\log 2\pi$. Hemos probado así la fórmula de Stirling en la forma

$$\log \Gamma(z)=\tfrac{1}{2}\log 2\pi - z + (z-\tfrac{1}{2})\log z + J(z) \qquad [4\text{-}42]$$

o, de manera equivalente,

$$\Gamma(z)=\sqrt{2\pi}\, z^{z-\frac{1}{2}}e^{-z}e^{J(z)} \qquad [4\text{-}43]$$

con la representación [4-41] del resto válida en el semiplano derecho. Sabemos que $J(z)$ tiende a cero cuando $z\to\infty$ en un semiplano $x\geqq c>0$.

En la expresión de $J(z)$ podemos desarrollar la integral en serie de potencias de $1/z$, y se obtiene

$$J(z)=\frac{C_1}{z}+\frac{C_2}{z^3}+\ldots+\frac{C_k}{z^{2k-1}}+J_k(z)$$

con

$$C_\nu=(-1)^\nu\frac{1}{\pi}\int^{\infty}\eta^{2\nu-2}\log\frac{1}{1-e^{-2\pi\eta}}\,d\eta \qquad [4\text{-}44]$$

y

$$J_k(z)=\frac{(-1)^k}{z^{2k+1}}\,\frac{1}{\pi}\int^{\infty}\frac{\eta^{2k}}{1+(\eta/z)^2}\log\frac{1}{1-e^{-2\pi\eta}}\,d\eta.$$

Puede demostrarse (p. ej., utilizando residuos) que los coeficientes C_ν están relacionados con los números de Bernoulli (cf. Ej. 4, sección 4-2, 3) por la expresión

$$C_\nu=(-1)^\nu\frac{1}{(2\nu-1)2\nu}B_\nu. \qquad [4\text{-}45]$$

Así, pues, el desarrollo de $J(z)$ toma la forma

$$J(z)=\frac{B_1}{1\cdot 2}\,\frac{1}{z}-\frac{B_2}{3\cdot 4}\cdot\frac{1}{z^3}+\cdots+(-1)^{k-1}\frac{B_k}{(2k-1)2k}\,\frac{1}{z^{2k-1}}+J_k(z). \qquad [4\text{-}46]$$

Se advierte al lector que no debe confundir este desarrollo con un desarrollo de Laurent. La función $J(z)$ no está definida en un entorno de ∞ y, por consiguiente, no posee un desarrollo de Laurent; además, si $k\to\infty$, la serie obtenida de [4-46] no converge. Lo que podemos decir es que para un k fijo la expresión $J_k(z)z^{2k}$ tiende a

cero para $z\to\infty$ (en un semiplano $x\geqq c>0$). Este hecho caracteriza a [4-46] como un *desarrollo asintótico*. Tales desarrollos son de gran valor cuando z es grande en comparación con k, pero para z fijo no tiene ninguna ventaja hacer que k se haga muy grande.

Se puede utilizar la fórmula de Stirling para demostrar que

$$\Gamma(z)=\int_0^{\infty}e^{-t}t^{z-1}dt \qquad [4\text{-}47]$$

siempre y cuando la integral converja; es decir, para $x>0$. Hasta que haya sido demostrada la identidad, denotaremos la integral en [4-47] por $F(z)$. Integrando por partes, se halla inmediatamente que

$$F(z+1)=\int_0^{\infty}e^{-t}t^{z}dt=z\int_0^{\infty}e^{-t}t^{z-1}dt=zF(z).$$

Por tanto, $F(z)$ satisface la misma ecuación funcional que $\Gamma(z)$, y se tiene que $F(z)/\Gamma(z)=F(z+1)/\Gamma(z+1)$. En otras palabras: $F(z)/\Gamma(z)$ es periódica con período 1. Esto prueba, incidentalmente, que $F(z)$ puede definirse en todo el plano, aunque la representación mediante una integral es válida únicamente en un semiplano.

Con objeto de probar que $F(z)/\Gamma(z)$ es constante resulta preciso acotar $|F/\Gamma|$ en la banda correspondiente a un período; p. ej., en la banda $1\leqq x\leqq 2$. Tenemos, en primer lugar, por [4-47],

$$|F(z)|\leqq\int_0^{\infty}e^{-t}t^{x-1}dt=F(x),$$

y, por tanto, $F(z)$ está acotada en la banda. Utilizamos a continuación la fórmula de Stirling para hallar una cota inferior de $|\Gamma(z)|$ para valores grandes de y. De [4-42] se obtiene

$$\log|\Gamma(z)|=\tfrac{1}{2}\log 2\pi - x + (x-\tfrac{1}{2})\log|z| - y\arg z + \operatorname{Re} J(z).$$

Unicamente el término $-y\arg z$ se hace menos infinito, siendo comparable a $-\pi|y|/2$. Así, pues, $|F/\Gamma|$ no crece mucho más rápidamente que $e^{\pi|y|/2}$.

Para una función arbitraria esto no sería suficiente para llegar a la conclusión de que la función es constante; pero para una función de período 1 es más que suficiente. En efecto, es evidente que F/Γ puede expresarse como una función uniforme de la variable $\zeta=e^{2\pi iz}$; a cada valor de $\zeta\neq 0$ corresponden infinitos valores de z, que difieren entre sí en múltiplos de 1, y, por tanto, corresponde un solo valor de F/Γ. La función tiene singularidades aisladas en $\zeta=0$ y $\zeta=\infty$, mostrando nuestro cálculo que $|F/\Gamma|$ crece a lo sumo como $|\zeta|^{-\frac{1}{2}}$ para $\zeta \to 0$ y como $|\zeta|^{\frac{1}{2}}$ para $\zeta \to \infty$. Se sigue que ambas singularidades son evitables y que, por tanto, F/Γ debe reducirse a una constante. Por último, el hecho de que $F(1)=\Gamma(1)=1$ prueba que $F(z)=\Gamma(z)$.

EJERCICIOS

1. Demuéstrese el desarrollo [4-46].
2. Pruébese que

$$\Gamma(x)=\sqrt{2\pi}\, x^{x-\frac{1}{2}}e^{-x}e^{\theta(x)/12x}$$

para $x>0$, con $0<\theta(x)<1$.

4-4. Familias normales.—Al igual que se consideran conjuntos de puntos, también pueden considerarse conjuntos de funciones. Preferimos, con objeto de marcar una clara distinción, hablar de *familias* de funciones, y para mayor conveniencia supondremos siempre que las funciones de una familia están definidas en el mismo conjunto. En la teoría de funciones estamos interesados principalmente en familias de funciones analíticas, definidas en una región fija Ω. Así podremos limitarnos a considerar subfamilias de la familia formadas por todas las funciones analíticas uniformes en Ω. Ejemplos importantes son las familias de funciones analíticas acotadas, funciones que no toman jamás el valor cero o funciones que toman valores prescritos en ciertos puntos dados.

1. *Condiciones de normalidad.*—Denotemos por $\mathfrak{F}$ una familia bien definida de funciones analíticas $f(z)$ en una región Ω. Estaremos interesados en la convergencia de sucesiones $\{f_n\}$ formadas por funciones de $\mathfrak{F}$. No existe ninguna razón particular para suponer que una sucesión $\{f_n\}$ sea convergente; de hecho, quizá es más probable que se dé el caso contrario, el de que $\{f_n\}$ no contenga una sola subsucesión convergente. En muchas situaciones, la última posibilidad es un grave obstáculo. Por tanto, es deseable considerar familias en las que esta forma de comportamiento esté excluida; esto nos conduce a la siguiente definición:

DEFINICIÓN. *Una familia $\mathfrak{F}$ de funciones f, definidas en una región Ω, se dice que es normal si toda sucesión $\{f_n\}$ de funciones*

$f_n \in \mathfrak{F}$ contiene una subsucesión $\{f_{n_k}\}$ que o converge uniformemente o tiende uniformemente a ∞ en todo subconjunto compacto de Ω.

No se ha supuesto explícitamente que las funciones f sean analíticas, pero este es el único caso que nos interesa. Por el teorema 1, la función límite de la sucesión $\{f_{n_k}\}$ es a su vez analítica, a no ser que se reduzca a la constante ∞. *No* se requiere que la función límite pertenezca a $\mathfrak{F}$. La razón de incluir ∞ entre las posibles funciones límite se comprende fácilmente; lo que queremos prevenir es la dispersión de valores, y la convergencia a ∞ sirve ciertamente a este propósito. Es obvio que se obtendría una mayor coherencia si definiésemos familias normales de funciones meromorfas con la ayuda de la convergencia uniforme sobre la esfera de Riemann. Sin embargo, la mayoría de las aplicaciones tratan únicamente con familias de funciones analíticas, por lo que parece aconsejable abstenerse de generalizaciones que no han de ser necesarias.

La siguiente condición es necesaria y suficiente para la normalidad de familias de funciones analíticas:

TEOREMA 8. *Una familia $\mathfrak{F}$ de funciones analíticas en Ω es normal si (y solo si) para todo conjunto compacto $E \subset \Omega$ existe una constante M tal que*

$$|f'(z)| \leqq M(1+|f(z)|^2) \tag{4-48}$$

para todo $z \in E$, $f \in \mathfrak{F}$.

La condición tiene un significado geométrico sencillo. En el capítulo I, sección 1-2, 5, hallamos que las proyecciones estereográficas de dos puntos a, b tienen una distancia cordal

$$d(a,b)=\frac{2|a-b|}{\sqrt{(1+|a|^2)(1+|b|^2)}}$$

Poniendo $a=f(z)$, $b=f(z+h)$ para puntos próximos, vemos que el cambio lineal de escala correspondiente a la aplicación del plano z en la esfera de Riemann sobre el plano $w=f(z)$ viene medida por $2|f'|/(1+|f|^2)$. Se requiere que este cambio de escala esté acotado uniformemente en todo conjunto compacto. Escribiremos

$$\rho(f)=\frac{2|f'|}{1+|f|^2}$$

y observemos que $\rho(1/f)=\rho(f)$.

Probaremos primero que la condición [4-48] es necesaria. Si no se verificara, existiría un conjunto compacto E y una sucesión $\{f_n\}$ tal que el máximo de $\rho(f_n)$ sobre E tendería a ∞. Si una subsucesión $\{f_{n_k}\}$ tiende a una función límite analítica φ, el teorema 1 implica que $\rho(f_{n_k})$ tiende a $\rho(\varphi)$, uniformemente en E. Entonces $\rho(f_{n_k})$ estaría acotado en E, en contra de lo supuesto. Si, por otra parte, $f_{n_k} \to \infty$, ponemos $g_{n_k} = 1/f_{n_k}$, con lo que se tiene que $\rho(g_{n_k}) = \rho(f_{n_k})$ tiende uniformemente a cero en E. Esto es nuevamente una contradicción, luego la familia no puede ser normal.

A fin de probar la suficiencia estableceremos dos propiedades previas. Sea z_0 un punto en Ω y consideremos una sucesión de funciones $f_n \in \mathfrak{F}$.

a) Si los valores $|f_n(z_0)|$ permanecen inferiores a una cota finita, entonces las funciones f_n están acotadas uniformemente en un entorno de z_0.

b) Si $f_n(z_0) \to \infty$, entonces las funciones f_n tienden uniformemente a ∞ en un entorno de z_0.

Supongamos que $|f_n(z_0)| < A$ y que el disco $\Delta: |z - z_0| < \delta$ está contenido en Ω; o es $|f_n(z)| < 2A$ en Δ, o existe un punto $z_1 \in \Delta$, el más próximo a z_0, donde se cumple la condición $|f_n(z_1)| = 2A$. Entonces $|f_n(z)| \leqq 2A$ sobre el segmento de extremos z_0, z_1 y, por tanto, $|f_n'(z)| \leqq M(1 + 4A^2)$ sobre el mismo segmento. Se sigue, pues, que $A \leqq |f_n(z_1) - f_n(z_0)| \leqq M\delta(1 + 4A^2)$. Para A fijo y δ suficientemente pequeño esto constituye una contradicción. Queda así probado *a)*.

Para probar *b)* podemos suponer que $|f_n(z_0)| > 2$. El razonamiento anterior aplicado a $g_n(z) = 1/f_n(z)$ demuestra que $|g_n(z)| < 1$, $|f_n(z)| > 1$ en un entorno $\Delta: |z - z_0| < \delta$. Puesto que $f_n(z) \neq 0$ en Δ, podemos definir una rama uniforme de $\log f_n(z)$. Utilizando esta rama formamos

$$F(z) = \frac{\log f_n(z) - \log f_n(z_0)}{\log f_n(z) + \log f_n(z_0)}$$

La transformación lineal aplicada a $\log f_n(z)$ representa el semiplano derecho sobre el disco unidad. Por tanto, $|F(z)| < 1$ en Δ, mientras que $F(z_0) = 0$. En esta situación el lema de Schwarz (Cap. III, teorema 13) da $|F(z)| < \frac{1}{3}$ para $|z - z_0| < \delta/3$. Por otra parte, $|F(z)| < \frac{1}{3}$ implica que $\log |f_n(z)| > \frac{1}{2}$. Bajo las hipótesis de *b)* concluimos que $f_n(z) \to \infty$ uniformemente en $|z - z_0| < \delta/3$.

Las propiedades *a)* y *b)* se verifican para subsucesiones de $\{f_n(z)\}$ lo mismo que para la sucesión original. Por esta razón, podemos utilizar la propiedad *a)* para concluir que el conjunto de puntos z con $\liminf |f_n(z)| < \infty$ es abierto. Análogamente, el conjunto con $\lim |f_n(z)| = \infty$ es abierto de acuerdo con *b)*. Puesto que los dos conjuntos son complementarios, uno de ellos debe ser vacío: o bien $f_n(z) \to \infty$ en todos los puntos o en ninguno. Si se aplica este resultado a una subsucesión, tenemos que o bien existe una subsucesión $\{f_{n_k}\}$ que tiende a infinito en todo punto, o bien $\limsup |f_n(z)| < \infty$ por doquier.

Si $f_{n_k} \to \infty$, sabemos por *b)* que la convergencia es uniforme en un entorno de cada punto. Un conjunto compacto puede recubrirse por un número finito de entornos con esta propiedad, y se concluye que $f_{n_k} \to \infty$ uniformemente sobre todo conjunto compacto. En este caso hemos alcanzado la conclusión deseada. En el caso contrario, $|f_n(z)|$ está acotado en cada punto z por una cota que pudiera depender de z. Sin embargo, por *a)* las funciones f_n están acotadas uniformemente en un entorno de cada punto, y hallamos exactamente, como antes, que están acotadas uniformemente en todo conjunto compacto. Por consiguiente, de ahora en adelante podemos suponer que las funciones f_n están acotadas uniformemente en cada conjunto compacto. Para el resto de la demostración haremos uso del hecho de que existe una sucesión de puntos ζ_k densa por doquier en Ω; p. ej., el conjunto de puntos con coordenadas racionales. Vamos a extraer de la sucesión $\{f_n\}$ una subsucesión que converja en todos los puntos ζ_k. Esto se efectúa mediante el proceso de la diagonal de Cantor. Para k fijo la sucesión $\{f_n(\zeta_k)\}$ está acotada y tiene, por tanto, un punto límite. Por esta razón es posible disponer los subíndices de la siguiente forma

$$\begin{array}{c} n_{11} < n_{12} < \ldots < n_{1i} < \ldots \\ n_{21} < n_{22} < \ldots < n_{2i} < \ldots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ n_{k1} < n_{k2} < \ldots < n_{ki} < \ldots \\ \cdots\cdots\cdots\cdots\cdots\cdots \end{array} \qquad [4\text{-}49]$$

donde cada fila está contenida en la precedente, y se verifica la propiedad de la existencia de $\lim_{i \to \infty} f_{n_{ki}}(\zeta_k)$. La sucesión diagonal, compuesta por los números $n_i = n_{ii}$, es estrictamente creciente y en definitiva es una subsucesión de cada fila de [4-49]. Por tanto, $\{f_{n_i}\}$ es una verdadera subsucesión de $\{f_n\}$ que converge en todos los puntos ζ_k.

Demostraremos que la sucesión $\{f_{n_i}\}$ converge uniformemente en todo conjunto compacto E; este puede recubrirse con un número finito de entornos Δ cuyo cierre está contenido en Ω. Las funciones f_{n_i} están acotadas uniformemente en la unión de estos entornos; por tanto, por [4-48] también lo están sus derivadas; supongamos que $|f'_{n_i}| < C$. Dado $\epsilon > 0$, consideremos para cada punto en E un entorno Δ' que esté contenido en un Δ y cuyo radio sea $< \epsilon/(6C)$; para puntos cualesquiera z, $\zeta \in \Delta'$, tenemos $|f_{n_i}(z) - f_{n_i}(\zeta)| < \epsilon/3$. Recubrimos E con un número finito de entornos Δ' y elegimos un punto ζ_k en cada Δ'. Puesto que se elige solo un número finito, podemos hallar un entero i_0 tal que $|f_{n_i}(\zeta_k) - f_{n_j}(\zeta_k)| < \epsilon/3$ para todo $i, j > i_0$ y todos estos k. Para un $z \in E$ arbitrario existe un ζ_k contenido en el mismo Δ', con lo que mediante la desigualdad triangular se obtiene

$$|f_{n_i}(z) - f_{n_j}(z)| < |f_{n_i}(z) - f_{n_i}(\zeta_k)| + |f_{n_j}(z) - f_{n_j}(\zeta_k)| + \\ + |f_{n_i}(\zeta_k) - f_{n_j}(\zeta_k)| < \epsilon$$

para todo $i, j > i_0$. Esto prueba que la sucesión $\{f_{n_i}\}$ converge uniformemente en E.

COROLARIO. *Si todas las funciones f de una familia normal $\mathfrak{F}$ son distintas de cero, entonces cada función límite es o distinta de cero o idénticamente cero.*

En efecto, puesto que $\rho(1/f) = \rho(f)$, las funciones $1/f$ forman una familia normal y tienen, por tanto, funciones límites que son o $\neq \infty$ o idénticamente infinitas. El corolario, conocido como teorema de Hurwitz, podría haberse demostrado también utilizando, p. ej., el principio del módulo máximo. Es de aplicación, en particular, a cualquier sucesión de funciones $f_n \neq 0$ que converja uniformemente en todo conjunto compacto, pues las funciones de tal sucesión forman una familia normal. De acuerdo con el corolario, la función límite o no es nunca cero o es idénticamente cero.

En las aplicaciones del teorema 8 son importantes condiciones suficientes para la normalidad más sencillas. El siguiente caso especial es uno de los que se presentan con más frecuencia:

TEOREMA 9. *Una familia $\mathfrak{F}$ de funciones analíticas en una región Ω es normal si las funciones $f \in \mathfrak{F}$ están acotadas uniformemente en todo subconjunto compacto de Ω.*

La demostración es casi trivial si se utiliza la acotación de

Cauchy de la derivada. Recubramos el conjunto compacto E mediante un número finito de discos $|z - a_i| < r_i$, tales que los discos cerrados $|z - a_i| \leqq 2r_i$ de radios dobles estén todavía contenidos en Ω. En los discos últimos f está acotada uniformemente, $|f| \leqq M$. Tenemos entonces, en cada disco más pequeño, por la desigualdad de Cauchy (Cap. III, Sec. 3-2, 3),

$$|f'(z)| \leqq \frac{M}{r_i}$$

Puesto que únicamente hay un número finito de r_i, hallamos inmediatamente que $|f'|/(1 + |f|^2)$ está acotada uniformemente en E, de lo que se sigue el teorema. Bajo las hipótesis del teorema 9, ninguna función límite puede ser infinita, y, por tanto, toda sucesión tiene una subsucesión que converge a una función analítica.

Se dice que una familia normal es *compacta* si toda función límite es un miembro de la familia. Para poner de manifiesto la diferencia entre familias normales compactas y no compactas consideremos la familia $\mathfrak{F}$ de todas las funciones analíticas en Ω que satisfacen la condición $|f(z)| < 1$. Por el teorema 9, esta familia es normal, pero no compacta, pues una sucesión de funciones puede converger a una constante de valor absoluto 1, en tanto que estas constantes no son miembros de $\mathfrak{F}$. En contraste, la familia de funciones que satisfacen la condición $|f| \leqq 1$ es, por supuesto, compacta.

EJERCICIOS

1. Si una sucesión $\{f_n\}$ extraída de una familia que es normal en Ω converge en todos los puntos de Ω, demuéstrese que la convergencia es uniforme en todo subconjunto compacto.

2. En las mismas circunstancias, demuéstrese que la sucesión $\{f_n\}$ convergerá en todos los puntos siempre y cuando converja en un conjunto que tenga un punto de acumulación en Ω.

3. Si $f(z)$ es analítica en todo el plano, pruébese que la familia formada por todas las funciones $f(kz)$, con k constante, es normal en una corona circular $r_1 < |z| < r_2$ si (y solo si) f es un polinomio.

2. *Teorema fundamental de Riemann sobre aplicaciones conformes.*—Como una aplicación de la teoría de las familias normales vamos a probar uno de los teoremas más importantes del análisis de

variable compleja: el hecho de que cualquier región simplemente conexa, que no sea todo el plano, puede aplicarse conformemente sobre un disco. El teorema lo enunció por primera vez Riemann, aunque con una demostración insuficiente. Desde entonces se han dado muchas demostraciones. La que presentamos no es la más satisfactoria en algunos aspectos, pero es la más corta.

TEOREMA 10. *Dada una región simplemente conexa cualquiera* Ω, *que no sea todo el plano, y un punto* $z_0 \in \Omega$, *existe una función analítica única* $f(z)$ *en* Ω, *normalizada por las condiciones* $f(z_0)=0$, $f'(z_0)>0$, *tal que* $f(z)$ *toma todo valor del disco* $|w|<1$ *una (y solo una) vez en* Ω.

La unicidad es trivial, pues si f_1 y f_2 son dos de tales funciones, entonces $f_1[f_2^{-1}(w)]$ define una aplicación biunívoca de $|w|<1$ sobre sí mismo. Sabemos que tal aplicación está dada por una transformación lineal S (Cap. III, Sec. 3-3, 4, Ej. 5). Las condiciones $S(0)=0$, $S'(0)>0$ implican $S(w)=w$; por consiguiente, $f_1=f_2$.

Una función analítica $g(z)$ en Ω se llama *univalente* si $g(z_1)=g(z_2)$ implica $z_1=z_2$; en otras palabras: si la aplicación definida por g es biunívoca[1]. Para demostrar la existencia, consideremos la familia $\mathfrak{F}$ formada por todas las funciones g con las propiedades siguientes: *a)* g es analítica y univalente en Ω; *b)* $|g(z)|\leqq 1$ en Ω; *c)* $g(z_0)=0$ y $g'(z_0)>0$. Afirmamos que f es la función de $\mathfrak{F}$ para la cual la derivada $f'(z_0)$ alcanza un máximo. La demostración consta de tres partes: 1) se demuestra que la familia $\mathfrak{F}$ no es vacía; 2) que existe una f con derivada máxima; 3) que esta f tiene las propiedades deseadas.

Para demostrar que $\mathfrak{F}$ no es vacía observemos que, por hipótesis, existe un punto $a\neq\infty$ no perteneciente a Ω. Puesto que Ω es simplemente conexo, puede definirse en él una rama analítica uniforme de $\log(z-a)$. Esta función no toma el mismo valor dos veces ni toma valores que difieran en un múltiplo de $2\pi i$, pues $\log(z_1-a)=\log(z_2-a)+n\cdot 2\pi i$ implicaría $z_1-a=(z_2-a)e^{n\cdot 2\pi i}=z_2-a$, de donde $z_1=z_2$. Puesto que $\log(z-a)$ toma el valor $w_0=\log(z_0-a)$, no tomará el valor $w_0+2\pi i$. Existe, además, un entorno $|w-w_0|<\rho$, tal que todos sus valores se toman en un entorno de z_0; por tanto, ninguno de los valores contenidos en $|w-w_0-2\pi i|<\rho$ los puede tomar $\log(z-a)$ en Ω; por tanto, $|\log(z-a)-w_0-2\pi i|>\rho$ en Ω.

[1] Es de uso frecuente la palabra alemana *schlicht*, que carece de traducción adecuada.

La función $h(z)=[\log(z-a)-w_0-2\pi i]^{-1}$ es analítica y univalente en Ω y satisface la condición $|h(z)|\leqq 1/\rho$. Se sigue inmediatamente que

$$g(z)=\frac{\rho}{1+\rho|h(z_0)|}\cdot\frac{|h'(z_0)|}{h'(z_0)}\cdot[h(z)-h(z_0)]$$

pertenece a la familia $\mathfrak{F}$.

Las derivadas $g'(z_0)$, $g\in\mathfrak{F}$, tienen un extremo superior B que *a priori* podría ser infinito. Hay una sucesión de funciones $g_n\in\mathfrak{F}$ tal que $\lim_{n\to\infty} g'_n(z_0)=B$. Por el teorema 9 la familia $\mathfrak{F}$ es normal. Existe, por tanto, una subsucesión $\{g_{n_k}\}$ que tiende a una función límite f analítica, uniformemente en conjuntos compactos. Es evidente que $|f|\leqq 1$, $f(z_0)=0$ y $f'(z_0)=B$. Si podemos probar que f es univalente, se seguirá que f pertenece a $\mathfrak{F}$ y tiene una derivada máxima en z_0.

En primer lugar, f no es una constante, pues $f'(z_0)=B>0$. Elijamos un punto $z_1\in\Omega$ y consideremos las funciones $g_1(z)=g(z)-g(z_1)$, $g\in\mathfrak{F}$. Forman una familia normal, pues $|g_1|\leqq 2$ y $g_1(z)\neq 0$ en la región Ω_1 obtenida de Ω por omisión del punto z_1. De acuerdo con el corolario del teorema 8 cada función límite de las funciones g_1 es o diferente de cero o idénticamente cero en Ω_1. Pero $f(z)-f(z_1)$ es una función límite y no es idénticamente cero. Por tanto, se tiene que $f(z)\neq f(z_1)$ para $z\neq z_1$, con lo que hemos demostrado que f es univalente.

Queda por demostrar que f toma todo valor w con $|w|<1$. Sea w_0, $|w_0|<1$, un valor no tomado por f. Entonces, puesto que Ω es simplemente conexa, se puede definir una rama uniforme de

$$F(z)=\log\frac{f(z)-w_0}{1-\bar{w}_0 f(z)}$$

Tiene parte real negativa; por tanto, la función

$$G(z)=\frac{F(z)-F(z_0)}{F(z)+\overline{F(z_0)}}$$

está acotada, $|G(z)|<1$. Mediante el correspondiente cálculo, obtenemos para su derivada en el origen

$$G'(z_0)=-B\cdot\frac{1-|w_0|^2}{2w_0\log|w_0|}$$

La función $g(z)=G(z)|G'(z_0)|/G'(z_0)$ pertenece a $\mathfrak{F}$, y su derivada es

$$g'(z_0)=B\cdot\frac{1-|w_0|^2}{2|w_0|\log 1/|w_0|}$$

Pero esta expresión es mayor que B. En efecto, $2\log(1/t)+t-1/t$ se anula para $t=1$, siendo la derivada de esta función,

$$1-2/t+1/t^2=(1-1/t)^2>0.$$

Por tanto, $2\log(1/t)<1/t-t$ para $t<1$, y con $t=|w_0|$ esta desigualdad implica $g'(z_0)>B$. Hemos llegado así a una contradicción, por lo que concluimos que el teorema tiene que verificarse.

El contenido puramente topológico del teorema 10 es de por sí importante. Sabemos ahora que una región simplemente conexa cualquiera puede aplicarse topológicamente sobre un disco (para todo el plano se puede construir una aplicación trivial), con lo que todas las regiones simplemente conexas son topológicamente equivalentes.

Observaremos finalmente que $|f(z)|$ tiende a 1 cuando z tiende a la frontera de Ω. Esta es realmente una proposición puramente topológica que podemos probar en la formulación siguiente:

TEOREMA 11. *Si la función $f(z)$ aplica una región Ω topológicamente sobre un conjunto Ω', entonces cualquier sucesión $\{z_n\}$ que tienda a la frontera de Ω se transforma en una sucesión $\{f(z_n)\}$ que tiende a la frontera de Ω'.*

La hipótesis significa que todos los puntos límites de la sucesión $\{z_n\}$ pertenecen a la frontera de Ω, la cual pudiera incluir el punto del infinito. Si la afirmación no fuera cierta podríamos hallar una subsucesión $f(z_{n_k})$ que convergiese a un punto $w_\infty\in\Omega'$. Puesto que la aplicación es topológica, la subsucesión $\{z_{n_k}\}$ correspondiente tendería al punto $z_\infty=f^{-1}(w_\infty)\in\Omega$, en contra de lo supuesto. De aquí se deduce el teorema.

Se demuestra en topología que Ω' es siempre una región (invariancia de la región).

EJERCICIO

En el teorema fundamental de Riemann, pruébese que la desigualdad $g'(z_0)>B$ puede deducirse sin cálculo del lema de Schwarz. Dése también otra demostración del teorema que haga uso de la función auxiliar

$$F(z)=\sqrt{\frac{f(z)-w_0}{1-\bar{w}_0f(z)}}$$

CAPITULO V

EL PROBLEMA DE DIRICHLET

5-1. Funciones armónicas.—Las partes real e imaginaria de una función analítica son funciones armónicas conjugadas. Por consiguiente, todos los teoremas referentes a funciones analíticas son también teoremas sobre pares de funciones armónicas conjugadas. No obstante, las funciones armónicas son importantes por derecho propio, y el tratarlas mediante los métodos complejos no siempre simplifica las cosas. Esto es particularmente cierto cuando la función conjugada no es uniforme, así como en todas las cuestiones relativas a *problemas de valores de contorno.*

1. *Definición y propiedades básicas.*—Una función real $u(z)$ o $u(x, y)$, definida y uniforme en una región Ω, se dice *armónica* en Ω, o *función potencial,* si es continua a la vez que sus derivadas parciales de los dos primeros órdenes y satisface a la *ecuación de Laplace:*

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \qquad [5\text{-}1]$$

Veremos más adelante que se pueden debilitar las condiciones de regularidad, pero esto es un punto relativamente de menor importancia.

La suma de dos funciones armónicas y el producto de una constante por una función armónica son nuevamente funciones armónicas; esto es debido al carácter lineal de la ecuación de Laplace. Las funciones armónicas más sencillas son las funciones lineales $ax+by$. En coordenadas polares (r, θ), la ecuación [5-1] toma la forma

$$r\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial \theta^2} = 0 \text{ [1]}.$$

Esto muestra que $\log r$ es una función armónica y que cualquier función armónica que dependa únicamente de r tiene que ser de

[1] Esta forma de la ecuación no puede utilizarse en el origen.

la forma $a\log r + b$. El argumento θ es una función armónica siempre y cuando pueda definirse de manera única.

Si se compone $u(z)$ con una función analítica $z(\zeta)$, la función resultante $u[z(\zeta)]$ es armónica a la vez que $u(z)$. De hecho, hallamos que en general

$$\Delta u[z(\zeta)] = |z'(\zeta)|^2 \Delta u(z).$$

En particular, la clase de las funciones armónicas es invariante respecto a las transformaciones conformes biunívocas de la variable.

Si u es armónica en Ω, entonces se tiene que

$$f(z) = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y} \qquad [5\text{-}2]$$

es analítica, pues poniendo $U = \frac{\partial u}{\partial x}$, $V = -\frac{\partial u}{\partial y}$, tenemos que

$$\frac{\partial U}{\partial x} = \frac{\partial^2 u}{\partial x^2} = -\frac{\partial^2 u}{\partial y^2} = \frac{\partial V}{\partial y}$$

$$\frac{\partial U}{\partial y} = \frac{\partial^2 u}{\partial x\,\partial y} = -\frac{\partial V}{\partial x}$$

Debe recordarse que esta es la forma más natural de pasar de funciones armónicas a funciones analíticas.

De [5-2] se pasa a la diferencial

$$f\,dz = \left(\frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy\right) + i\left(-\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy\right). \qquad [5\text{-}3]$$

En esta expresión la parte real es la diferencial de u,

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy.$$

Si u posee una función armónica conjugada v, entonces la parte imaginaria puede escribirse en la forma

$$dv = \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy = -\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy.$$

En general, sin embargo, no existe función conjugada uniforme, y en estas circunstancias es preferible no utilizar la notación dv. En su lugar escribimos

$$*du = -\frac{\partial u}{\partial y}\,dx + \frac{\partial u}{\partial x}\,dy,$$

y llamamos a $*du$ la *diferencial conjugada* de du. En virtud de [5-3],

$$f\,dz = du + i *du. \qquad [5\text{-}4]$$

Por el teorema de Cauchy la integral de $f\,dz$ se anula a lo largo de cualquier ciclo homológico a cero en Ω. Por otra parte, la integral de la diferencial exacta du se anula a lo largo de todos los ciclos. Se sigue de [5-4] que

$$\int_\gamma *du = \int_\gamma -\frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy = 0 \qquad [5\text{-}5]$$

para todos los ciclos γ que sean homológicos a cero en Ω.

La integral en [5-5] tiene una importante interpretación que no puede dejar de mencionarse. Si γ es una curva regular de ecuación $z = z(t)$, la dirección de la tangente está determinada por el ángulo $\alpha = \arg z'(t)$, y podemos escribir $dx = |dz| \cos\alpha$, $dy = |dz| \operatorname{sen}\alpha$. La normal orientada hacia la derecha de la tangente tiene la dirección $\beta = \alpha - \pi/2$, así que $\cos\alpha = -\operatorname{sen}\beta$, $\operatorname{sen}\alpha = \cos\beta$. La expresión

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}\cos\beta + \frac{\partial u}{\partial y}\operatorname{sen}\beta$$

es una derivada direccional de u, la *derivada normal* por la derecha con respecto a la curva γ. Se obtiene que $*du = (\partial u/\partial n)|dz|$, con lo que podemos escribir [5-5] en la forma

$$\int_\gamma \frac{\partial u}{\partial n}|dz| = 0. \qquad [5\text{-}6]$$

Esta es la notación clásica. Su ventaja principal es la de que $\partial u/\partial n$ representa realmente una velocidad de variación en la dirección perpendicular a γ. Así, p. ej., si γ es la circunferencia $|z| = r$, recorrida en el sentido positivo, se puede reemplazar $\partial u/\partial n$ por la derivada parcial $\partial u/\partial r$. Su inconveniente es que [5-6] no está expresada como una integral curvilínea ordinaria, sino como una integral con respecto a la longitud del arco. Esta es la razón de que la

notación clásica sea menos natural en relación con la teoría de la homología, por lo que preferimos utilizar la notación $*du$.

En una región simplemente conexa, la integral de $*du$ se anula a lo largo de todos los ciclos, y u tiene una función conjugada uniforme v, que está determinada salvo una constante aditiva. En el caso de conexión múltiple, la función conjugada tiene *períodos*

$$\int_{\gamma_i} *du = \int_{\gamma_i} \frac{\partial u}{\partial n}|dz|$$

correspondientes a los ciclos en una base de homología.

Existe una importante generalización de [5-5] que hace intervenir un par de funciones armónicas. Si u_1 y u_2 son armónicas en Ω, afirmamos que

$$\int_\gamma u_1 *du_2 - u_2 *du_1 = 0 \qquad [5\text{-}7]$$

para todo ciclo γ que sea homológico a cero en Ω. De acuerdo con el teorema 17 (Cap. III, Sec. 3-4, 4), es suficiente probar [5-7] para $\gamma = \Gamma(R)$, donde R es un rectángulo contenido en Ω. En R, u_1 y u_2 tienen funciones conjugadas uniformes v_1 y v_2, y podemos escribir

$$u_1 *du_2 - u_2 *du_1 = u_1 dv_2 - u_2 dv_1 = u_1 dv_2 + v_1 du_2 - d(u_2 v_1).$$

Aquí $d(u_2v_1)$ es una diferencial exacta, y $u_1dv_2 + v_1du_2$, la parte imaginaria de

$$(u_1 + iv_1)(du_2 + i\,dv_2).$$

La diferencial última puede escribirse en la forma $F_1 f_2 dz$, siendo $F_1(z)$ y $f_2(z)$ analíticas en R. Por el teorema de Cauchy, la integral de $F_1 f_2 dz$ se anula y lo mismo le ocurre, por tanto, a la integral de su parte imaginaria. Se concluye que [5-7] se verifica para $\gamma = \Gamma(R)$, con lo que hemos demostrado:

TEOREMA 1. *Si u_1 y u_2 son armónicas en una región Ω, entonces se verifica*

$$\int_\gamma u_1 *du_2 - u_2 *du_1 = 0 \qquad [5\text{-}7]$$

para todo ciclo γ que sea homológico a cero en Ω.

Para $u_1=1$, $u_2=u$, la fórmula se reduce a la [5-5]. Con la notación clásica, [5-7] se escribiría en la forma

$$\int_\gamma \left(u_1 \frac{\partial u_2}{\partial n} - u_2 \frac{\partial u_1}{\partial n} \right) |dz| = 0.$$

2. ***Propiedad del valor medio.***—Apliquemos el teorema 1 con $u_1=\log r$ y u_2 igual a una función u, armónica en $|z|<\rho$. Como Ω debemos elegir el disco sin centro $0<|z|<\rho$, y como γ tomamos el ciclo C_1-C_2, siendo C_i la circunferencia $|z|=r_i<\rho$, recorrida en sentido positivo. En una circunferencia $|z|=r$ tenemos que $*du = r(\partial u/\partial r)d\theta$, y, por tanto, [5-7] da

$$\log r_1 \int_{C_1} r_1 \frac{\partial u}{\partial r} d\theta - \int_{C_1} u\, d\theta = \log r_2 \int_{C_2} r_2 \frac{\partial u}{\partial r} d\theta - \int_{C_2} u\, d\theta.$$

En otras palabras: la expresión

$$\int_{|z|=r} u\, d\theta - \log r \int_{|z|=r} r \frac{\partial u}{\partial r} d\theta$$

es constante; lo que es cierto incluso si se conoce únicamente que u es armónica en una corona circular. Mediante [5-5] hallamos de la misma forma que

$$\int_{|z|=r} r \frac{\partial u}{\partial r} d\theta$$

es constante en el caso de una corona circular y cero en el caso de que u sea armónica en todo el disco. Combinando estos resultados obtenemos:

TEOREMA 2. *La media aritmética de una función armónica sobre circunferencias concéntricas $|z|=r$ es una función lineal de $\log r$,*

$$\frac{1}{2\pi} \int_{|z|=r} u\, d\theta = \alpha \log r + \beta, \qquad [5\text{-}8]$$

y si u es armónica en un disco, $\alpha=0$ y la media aritmética es constante.

En el último caso, $\beta=u(0)$, por la continuidad, y mediante un cambio de origen, hallamos que

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{i\theta})\, d\theta. \qquad [5\text{-}9]$$

Evidentemente [5-9] podría también haberse deducido de la fórmula correspondiente para funciones analíticas (Cap. III, Sec. 3-3, 4; [3-32]). Conduce directamente al *principio del máximo* para funciones armónicas:

TEOREMA 3. *Una función armónica no constante no tiene ni máximo ni mínimo en su región de definición. En consecuencia, en un conjunto cerrado y acotado E el máximo y el mínimo se alcanzan en la frontera de E.*

La demostración es la misma que para el principio del módulo máximo para funciones analíticas; por tanto, no la repetiremos. Se aplica también al mínimo como consecuencia de ser $-u$ armónica a la vez que u. En el caso de funciones analíticas, el procedimiento correspondiente hubiese sido aplicar el principio del módulo máximo a $1/f(z)$, lo que no es legítimo a no ser que $f(z) \neq 0$. Obsérvese que el principio del módulo máximo para funciones analíticas se deduce del principio del máximo para funciones armónicas aplicando el último a $\log |f(z)|$, que es armónica cuando $f(z) \neq 0$.

EJERCICIOS

1. Dedúzcase un teorema, análogo al teorema 2, para una familia de elipses homofocales.

2. Si u es armónica y está acotada en $0<|z|<\rho$, pruébese que el origen es una singularidad evitable de u, en el sentido de que u se hace armónica en $|z|<\rho$ cuando se define $u(0)$ de manera apropiada.

3. Si $u(z)$ es armónica en $0<|z|<\rho$ y $\lim_{z\to 0} zu(z)=0$, pruébese que u puede escribirse en la forma $u(z)=\alpha \log|z| + u_0(z)$, siendo α una constante y u_0 armónica en $|z|<\rho$.

3. *Fórmula de Poisson.*—El principio del máximo tiene la importante consecuencia siguiente: Si $u(z)$ es armónica sobre un conjunto acotado y cerrado E, es decir, si está definida y es armónica en una región que contenga a E, entonces está determinada de manera única por sus valores en la frontera de E. En efecto, si u_1 y u_2

son dos funciones armónicas con los mismos valores sobre el contorno, entonces $u_1 - u_2$ es armónica con valor cero sobre el contorno. Utilizando el principio del máximo y mínimo, se deduce que $u_1 - u_2$ debe ser idénticamente cero sobre E.

Usualmente E es una región cerrada, exigiéndose que $u(z)$ sea continua sobre E y armónica en el interior de E. En estas circunstancias todavía es aplicable el principio del máximo, como se indicó en relación con el principio del módulo máximo para funciones analíticas. El problema de determinar $u(z)$ cuando se dan los valores sobre la frontera se conoce con el nombre de *problema de Dirichlet*. No sabemos que exista solución para él, pero si la tiene es necesariamente única.

La fórmula [5-9] nos ayudará a resolver el problema de Dirichlet en un disco circular. Tal y como está, únicamente da el valor de $u(z)$ en el centro; pero mediante una transformación lineal cualquier punto puede ocupar el lugar del centro, con lo que es posible determinar todos los valores de $u(z)$.

Cambiando nuestras notaciones, consideremos el disco cerrado $|z| \leqq R$ y un punto interior a. La transformación lineal

$$z = S(\zeta) = \frac{R(R\zeta + a)}{R + \bar{a}\zeta}$$

aplica $|\zeta| \leqq 1$ sobre $|z| \leqq R$ con $\zeta = 0$ correspondiendo a $z = a$. La función $u[S(\zeta)]$ es armónica en $|\zeta| \leqq 1$, y por [5-9] obtenemos

$$u(a) = \frac{1}{2\pi} \int_{|\zeta|=1} u[S(\zeta)]\, d \arg \zeta.$$

De

$$\zeta = \frac{R(z-a)}{R^2 - \bar{a}z}$$

se deduce que

$$d \arg \zeta = -i \frac{d\zeta}{\zeta} = -i\left(\frac{1}{z-a} + \frac{\bar{a}}{R^2 - \bar{a}z}\right) dz = \left(\frac{z}{z-a} + \frac{\bar{a}z}{R^2 - \bar{a}z}\right) d\theta.$$

Sustituyendo $R^2 = z\bar{z}$, el coeficiente de $d\theta$ de la última expresión puede escribirse nuevamente en la forma

$$\frac{z}{z-a} + \frac{\bar{a}}{\bar{z} - \bar{a}} = \frac{R^2 - |a|^2}{|z-a|^2}$$

o, de manera equivalente, como

$$\frac{1}{2}\left(\frac{z+a}{z-a} + \frac{\bar{z}+\bar{a}}{\bar{z}-\bar{a}}\right) = \operatorname{Re} \frac{z+a}{z-a}$$

Obteniéndose las dos formas

$$u(a) = \frac{1}{2\pi} \int_{|z|=R} \frac{R^2 - |a|^2}{|z-a|^2} u(z)\, d\theta = \frac{1}{2\pi} \int_{|z|=R} \operatorname{Re} \frac{z+a}{z-a} u(z)\, d\theta \qquad [5\text{-}10]$$

de la *fórmula de Poisson*. En coordenadas polares se tiene

$$u(re^{i\varphi}) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2rR \cos(\theta - \varphi) + r^2} u(Re^{i\theta})\, d\theta. \qquad [5\text{-}11]$$

En la deducción de la fórmula de Poisson hemos supuesto que $u(z)$ es armónica en el disco cerrado. A causa de esta restricción la fórmula no prueba todavía la existencia de una solución del problema de Dirichlet. Sin embargo, el segundo miembro de [5-10] o de [5-11] puede formarse con valores arbitrarios sobre la circunferencia $|z| = R$, y se probará que resuelve el problema de Dirichlet.

Eligiendo $R = 1$, definimos, para cualquier función continua a trozos $U(\theta)$ en $0 \leqq \theta \leqq 2\pi$,

$$P_U(z) = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re} \frac{e^{i\theta} + z}{e^{i\theta} - z} U(\theta)\, d\theta. \qquad [5\text{-}12]$$

Como función de una función, $P_U(z)$ se llama un *funcional*. Se dirá que es *lineal* siempre y cuando

$$P_{U+V} = P_U + P_V$$

y

$$P_{cU} = cP_U$$

para c constante. Además, $U \geqq 0$ implica $P_U(z) \geqq 0$; a causa de esta propiedad se dice que P_U es un funcional lineal *positivo*.

Aplicando [5-10] a una función constante, hallamos que $P_c = c$. Lo que junto con el carácter lineal y positivo de P_U nos permite concluir que cualquier desigualdad $m \leqq U \leqq M$ implica $m \leqq P_U \leqq M$.

El teorema fundamental siguiente fue demostrado por vez primera por H. A. Schwarz:

TEOREMA 4. *La función $P_U(z)$ es armónica para $|z|<1$, y*

$$\lim_{z\to e^{i\theta_0}} P_U(z)=U(\theta_0) \qquad [5\text{-}13]$$

con tal que $U(\theta)$ sea continua en θ_0.

El que sea armónica se sigue de la igualdad

$$P_U(z)=\operatorname{Re}\left[\frac{1}{2\pi i}\int_{|\zeta|=1}\frac{\zeta+z}{\zeta-z}U(\theta)\frac{d\zeta}{\zeta}\right]$$

donde se asigna el valor $U(\theta)$ al punto $\zeta=e^{i\theta}$. La expresión entre corchetes es una función analítica de z excepto sobre $|z|=1$, y, por tanto, $P_U(z)$ es armónica para los mismos valores.

Sean C_1 y C_2 arcos complementarios de la circunferencia unidad y denotemos por U_1 la función que coincide con U sobre C_1 y se anula sobre C_2, y por U_2, la correspondiente función para C_2. Puesto que P_{U_1} puede considerarse como una integral curvilínea sobre C_1, por el mismo razonamiento utilizado anteriormente, P_{U_1} es armónica por doquier, excepto sobre el arco cerrado C_1. La expresión

$$\operatorname{Re}\frac{e^{i\theta}+z}{e^{i\theta}-z}=\frac{1-|z|^2}{|e^{i\theta}-z|^2}$$

se anula sobre $|z|=1$ para $z\neq e^{i\theta}$. Se sigue que P_{U_1} se anula en el arco abierto C_2, y como consecuencia de la continuidad satisface la condición $\lim_{z\to e^{i\theta_0}}P_{U_1}(z)=0$ para $e^{i\theta_0}\in C_2$.

Al demostrar [5-13] podemos suponer que $U(\theta_0)=0$, pues si no fuera este el caso necesitaríamos solo reemplazar U por $U-U(\theta_0)$. Dado $\epsilon>0$, podemos hallar C_1 y C_2 tales que $e^{i\theta_0}$ sea un punto interior de C_2 y $|U(\theta)|\leqq\epsilon$ para $e^{i\theta}\in C_2$. Satisfechas estas condiciones, $|U_2(\theta)|\leqq\epsilon$ para todo θ, por lo que $|P_{U_2}(z)|\leqq\epsilon$ para $|z|<1$. Puesto que $P_U=P_{U_1}+P_{U_2}$ y $\lim P_{U_1}=0$ para $z\to e^{i\theta_0}$, tenemos:

$$-\epsilon\leqq\underline{\lim}\,P_U(z)\leqq\overline{\lim}\,P_U(z)\leqq\epsilon\,[1].$$

[1] El *límite inferior* y el *límite superior* se definen para funciones exactamente igual que en el caso de sucesiones. Es decir, $\underline{\lim}_{z\to z_0} f(z)$ y $\overline{\lim}_{z\to z_0} f(z)$ [siendo $f(z)$ una función real] son, respectivamente, el menor y el mayor de los límites de todas las sucesiones convergentes $\{f(z_n)\}$ tales que $z_n\to z_0$.

En esta desigualdad ϵ es un número positivo arbitrario, luego se verifica [5-13].

Supongamos ahora que queremos resolver el problema de Dirichlet para el disco circular $|z-z_0|\leqq\rho$. Dada una función continua $U(\theta)$, con $U(0)=U(2\pi)$, deseamos hallar una función $u(z)$ que sea armónica para $|z-z_0|<\rho$ y continua en el disco cerrado, con los siguientes valores frontera:

$$u(z_0+\rho e^{i\theta})=U(\theta).$$

Por el teorema 4, tal función está dada por $u(z)=P_U[(z-z_0)/\rho]$, y por el principio del máximo, tal solución es única.

Como consecuencia casi inmediata del teorema 4 demostraremos la proposición siguiente.

TEOREMA 5. *Una función $u(z)$ continua en Ω, que en todos los puntos $z_0\in\Omega$ satisface la condición [5-9] (propiedad del valor medio) para todo r suficientemente pequeño, es necesariamente armónica.*

La demostración del principio del máximo se basaba en la propiedad del valor medio, pero un examen más atento de la demostración muestra que esta propiedad necesita postularse únicamente para $r<r_0$, donde r_0 puede depender de z_0. El principio del máximo es, por tanto, válido para $u(z)$ y también para la diferencia entre $u(z)$ y cualquier función armónica. Si $|z-z_0|\leqq\rho$ está contenido en Ω, podemos construir la función armónica $v(z)$ que sea armónica para $|z-z_0|<\rho$ y continua e igual a $u(z)$ para $|z-z_0|=\rho$. Por el principio del máximo y del mínimo, aplicado a $u-v$, se deduce que $u(z)=v(z)$ en todo el disco; por tanto, $u(z)$ es armónica.

El hecho notable del teorema 5 es la carencia absoluta de hipótesis relativas a las derivadas. Un razonamiento análogo muestra que incluso sin la condición [5-9] las hipótesis sobre las derivadas pueden relajarse en grado sumo. Supongamos simplemente que $u(z)$ es continua y que las derivadas $\partial^2u/\partial x^2$, $\partial^2u/\partial y^2$ existen y satisfacen $\Delta u=0$. Con las mismas notaciones que antes probaremos que la función

$$V=u-v+\epsilon(x-x_0)^2,$$

$\epsilon>0$, debe verificar el principio del máximo. En efecto, si V alcanza un máximo, las reglas del cálculo darían $\partial^2V/\partial x^2\leqq0$, $\partial^2V/\partial y^2\leqq0$, y, por tanto, $\Delta V\leqq0$ en dicho punto. Por otra parte,

$$\Delta V=\Delta u-\Delta v+2\epsilon=2\epsilon>0.$$

La contradicción prueba que se verifica el principio del máximo. Podemos concluir de este modo que $u - v + \epsilon(x - x_0)^2 \leqq \epsilon\rho^2$ en el disco $|z - z_0| \leqq \rho$. Haciendo tender ϵ a cero hallamos que $u \leqq v$, y la desigualdad opuesta puede demostrarse de la misma manera. Luego u es armónica.

EJERCICIOS

1. Si C_1 y C_2 son arcos complementarios de la circunferencia unidad, pongamos $U=1$ sobre C_1 y $U=0$ sobre C_2. Hállese explícitamente $P_U(z)$ y pruébese que $2\pi P_U(z)$ es igual a la longitud del arco opuesto a C_1 comprendido entre las rectas que pasan por z y por los extremos de C_1.

2. Demuéstrese que la fórmula [5-9] del valor medio sigue siendo válida para $u = \log|1+z|$, $z_0 = 0$, $r = 1$, y utilícese este hecho para calcular

$$\int_0^{\pi} \log \operatorname{sen} \theta \, d\theta.$$

3. Pruébese que todo polinomio de grado $< n$ tiene la propiedad del valor medio

$$P(z_0) = \frac{1}{n}[P(z_0+a) + P(z_0+a\omega) + \ldots + P(z_0+a\omega^{n-1})],$$

siendo a un número complejo arbitrario y $\omega = e^{2\pi i/n}$.

4. *Principio de Harnack.*—La fórmula de Poisson nos conduce directamente a ciertas acotaciones sencillas, pero muy importantes. Se deducen utilizando la siguiente desigualdad elemental

$$\frac{\rho - r}{\rho + r} \leqq \frac{\rho^2 - r^2}{|\rho e^{i\theta} - z|^2} \leqq \frac{\rho + r}{\rho - r}$$

válida para $|z| = r < \rho$. Con la ayuda de estas desigualdades se obtiene de [5-10]

$$|u(z)| \leqq \frac{1}{2\pi} \frac{\rho + r}{\rho - r} \int_0^{2\pi} |u(\rho e^{i\theta})| d\theta, \qquad [5\text{-}14]$$

y si se sabe que $u(\rho e^{i\theta}) \geqq 0$, se obtiene la siguiente doble desigualdad, que es más potente:

$$\frac{1}{2\pi} \frac{\rho - r}{\rho + r} \int_0^{2\pi} u \, d\theta \leqq u(z) \leqq \frac{1}{2\pi} \frac{\rho + r}{\rho - r} \int_0^{2\pi} u \, d\theta.$$

Puesto que la media aritmética es igual a $u(0)$, se puede escribir la última desigualdad en la forma más sencilla

$$\frac{\rho - r}{\rho + r} u(0) \leqq u(z) \leqq \frac{\rho + r}{\rho - r} u(0). \qquad [5\text{-}15]$$

La aplicación más importante de [5-15] es a series de términos positivos o, de manera equivalente, a sucesiones crecientes de funciones armónicas. Conduce a un teorema sencillo, pero muy poderoso, conocido como el *principio de Harnack.*

TEOREMA 6. *Consideremos una sucesión de funciones $u_n(z)$, cada una de ellas definida y armónica en una cierta región Ω_n. Sea Ω una región tal que todo punto de Ω tiene un entorno contenido en todos los Ω_n salvo en un número finito, y supongamos además que en este entorno $u_n(z) \leqq u_{n+1}(z)$ siempre y cuando n sea suficientemente grande. Hay entonces solo dos posibilidades: o $u_n(z)$ tiende uniformemente a ∞ en todo subconjunto compacto de Ω, o $u_n(z)$ tiende a una función límite armónica $u(z)$ en Ω, uniformemente en conjuntos compactos.*

La situación más sencilla corresponde al caso en que las funciones $u_n(z)$ son armónicas y forman una sucesión no decreciente en Ω. Existen, sin embargo, muchas aplicaciones en las que este caso no es suficientemente general.

Para demostrarlo supongamos en primer lugar que $\lim u_n(z_0) = \infty$, al menos para un punto $z_0 \in \Omega$. Por hipótesis existen un r y un m tales que las funciones $u_n(z)$ son armónicas y forman una sucesión no decreciente para $|z - z_0| < r$ y $n \geqq m$. Si la desigualdad de la izquierda de [5-15] se aplica a las funciones no negativas $u_n - u_m$, se deduce que $u_n(z)$ tiende uniformemente a ∞ en el disco $|z - z_0| \leqq \leqq r/2$. Por otra parte, si $\lim u_n(z_0) < \infty$, la aplicación de la desigualdad de la derecha muestra de la misma manera que $u_n(z)$ está acotada en $|z - z_0| \leqq r/2$. Por consiguiente, los conjuntos en los que $\lim u_n(z)$ es, respectivamente, finito o infinito son ambos abiertos, y puesto que Ω es conexo, uno de los conjuntos tiene que ser vacío. Basta, pues, que el límite sea infinito en un solo punto para que sea idénticamente infinito. La uniformidad se obtiene utilizando el lema de Heine-Borel.

En el caso opuesto, la función límite es finita por doquier, por lo que solo se necesita demostrar que la convergencia es uniforme. Con las mismas notaciones, $u_{n+p}(z) - u_n(z) \leqq 3[u_{n+p}(z_0) - u_n(z_0)]$ para $|z - z_0| \leqq r/2$ y $n \geqq m$. Por tanto, la convergencia en z_0 implica la convergencia uniforme en un entorno, con lo que mediante el lema de Heine-Borel se deduce que la convergencia es uniforme en todo conjunto compacto. El que la función límite es armónica se obtiene del hecho de que $u(z)$ puede representarse por la fórmula de Poisson.

EJERCICIOS

1. Si E es un conjunto compacto contenido en una región Ω, pruébese que existe una constante M, que depende únicamente de E y de Ω, tal que toda función armónica positiva $u(z)$ en Ω satisface la desigualdad $u(z_2) \leqq Mu(z_1)$ para cualquier par de puntos $z_1, z_2 \in E$.

2. Demuéstrese que las funciones analíticas en una región Ω, cuya parte real es positiva, constituyen una familia normal.

5. *Fórmula de Jensen.*—Si $f(z)$ es una función analítica, entonces $\log |f(z)|$ es armónica excepto en los ceros de $f(z)$. Por consiguiente, si $f(z)$ es analítica y no tiene ceros en $|z| \leqq \rho$, se tiene que

$$\log |f(0)| = \frac{1}{2\pi}\int_0^{2\pi} \log |f(\rho e^{i\theta})|\,d\theta, \qquad [5\text{-}16]$$

y se puede expresar $\log |f(z)|$ por la fórmula de Poisson.

La ecuación [5-16] sigue siendo válida si $f(z)$ tiene ceros en la circunferencia $|z| = \rho$. La demostración más sencilla se obtiene dividiendo $f(z)$ por un factor $z - \rho e^{i\theta_0}$ para cada cero. Basta probar que

$$\log \rho = \frac{1}{2\pi}\int_0^{2\pi} \log |\rho e^{i\theta} - \rho e^{i\theta_0}|\,d\theta$$

o

$$\int_0^{2\pi} \log |e^{i\theta} - e^{i\theta_0}|\,d\theta = 0.$$

Esta integral es evidentemente independiente de θ_0, con lo que únicamente tenemos que demostrar que

$$\int_0^{2\pi} \log |1 - e^{i\theta}|\,d\theta = 0.$$

Pero esto es una consecuencia de la fórmula

$$\int_0^{\pi} \log \operatorname{sen} x\,dx = -\pi \log 2$$

demostrada en el capítulo III, sección 3-5, 3 (cf. Sec. 3-1, 3; Ej. 2).

Veamos ahora en qué se convierte [5-16] al existir ceros en el interior de $|z| < \rho$. Denotemos esos ceros por $a_1, a_2, \ldots, a_n$, estando los ceros múltiples repetidos, y supongamos en primer lugar que $z = 0$ no es un cero. Entonces la función

$$F(z) = f(z) \prod_{i=1}^{n} \frac{\rho^2 - \bar{a}_i z}{\rho(z - a_i)}$$

no tiene ceros en el disco y $|F(z)| = |f(z)|$ sobre $|z| = \rho$. Por tanto, se obtiene

$$\log |F(0)| = \frac{1}{2\pi}\int_0^{2\pi} \log |f(\rho e^{i\theta})|\,d\theta$$

y, sustituyendo el valor de $F(0)$, tenemos la igualdad

$$\log |f(0)| = -\sum_{i=1}^{n} \log\left(\frac{\rho}{|a_i|}\right) + \frac{1}{2\pi}\int_0^{2\pi} \log |f(\rho e^{i\theta})|\,d\theta, \qquad [5\text{-}17]$$

conocida como *fórmula de Jensen.* Su importancia reside en el hecho de que relaciona el módulo $|f(z)|$ sobre un círculo con los módulos de los ceros.

Si $f(0) = 0$, la fórmula es algo más complicada. Poniendo $f(z) = cz^h + \ldots$, apliquemos [5-17] a $f(z)(\rho/z)^h$, y obtenemos así que el primer miembro tiene que reemplazarse por $\log |c| + h \log \rho$.

Existe una generalización análoga de la fórmula de Poisson. Todo lo que se necesita es aplicar la fórmula ordinaria de Poisson a $\log |F(z)|$. Se obtiene

$$\log |f(z)| = -\sum_{i=1}^{n} \log \left|\frac{\rho^2 - \bar{a}_i z}{\rho(z - \bar{a}_i)}\right| + \frac{1}{2\pi}\int_0^{2\pi} \operatorname{Re} \frac{\rho e^{i\theta} + z}{\rho e^{i\theta} - z} \log |f(\rho e^{i\theta})|\,d\theta, \qquad [5\text{-}18]$$

siempre y cuando $f(z) \neq 0$. A la ecuación [5-18] se le llama usualmente *fórmula de Poisson-Jensen.*

Hablando estrictamente, la demostración es solo válida si $f \neq 0$ sobre $|z| = \rho$. Pero [5-17] muestra que la integral de la derecha es una función continua de ρ, de donde se deduce con toda facilidad que la integral en [5-18] es igualmente continua. Así, pues, en el caso general se puede deducir [5-18] haciendo que ρ se acerque a un límite.

Las fórmulas de Jensen y de Poisson-Jensen tienen aplicaciones importantes en la teoría de funciones enteras. Las utilizaremos para establecer una cuestión que quedó sin resolver en el capítulo IV, sección 4-3, 3.

Sea $f(z)$ una función entera y denotemos sus ceros por a_n; para mayor sencillez supondremos que $f(0) \neq 0$. Recordemos que el género de una función entera $f(z)$ es el menor entero h tal que $f(z)$ pueda representarse en la forma

$$f(z) = e^{g(z)} \prod_n \left(1 - \frac{z}{a_n}\right) e^{\frac{z}{a_n} + \frac{1}{2}\left(\frac{z}{a_n}\right)^2 + \ldots + \frac{1}{h}\left(\frac{z}{a_n}\right)^h} \qquad [5\text{-}19]$$

siendo $g(z)$ un polinomio de grado $\leqq h$.

Denotemos por $M(r)$ el máximo de $|f(z)|$ sobre $|z| = r$. Se define el *orden* de la función entera $f(z)$ mediante la igualdad

$$\lambda = \overline{\lim_{r \to \infty}} \frac{\log \log M(r)}{\log r}$$

De acuerdo con esta definición, λ es el menor número, tal que

$$M(r) \leqq e^{r^{\lambda+\epsilon}}, \qquad [5\text{-}20]$$

para cualquier $\epsilon > 0$ dado, siempre y cuando r sea suficientemente grande.

El género y el orden están estrechamente relacionados, como puede verse por el siguiente teorema:

TEOREMA 7. *El género y el orden de una función entera satisfacen la doble desigualdad $h \leqq \lambda \leqq h+1$.*

Supongamos primero que $f(z)$ es de género finito h. El factor exponencial en [5-19] es obviamente de orden $\leqq h$, y el orden del producto no puede exceder al orden de ambos factores. Por tanto, es suficiente demostrar que el producto canónico es de orden $\leqq h+1$. La convergencia del producto canónico implica $\sum_n |a_n|^{-h-1} < \infty$; esta es la hipótesis esencial.

Denotemos el producto canónico por $P(z)$ y escribamos los factores individuales como $E_h(z/a_n)$, siendo

$$E_h(u) = (1-u) e^{u + \frac{1}{2}u^2 + \ldots + (1/h)u^h}$$

entendiéndose que $E_0(u) = 1 - u$. Probaremos que

$$\log |E_h(u)| \leqq (2h+1)|u|^{h+1} \qquad [5\text{-}21]$$

para todo u.

Si $|u|$ es menor que 1, tenemos por un desarrollo en serie de potencias

$$\log |E_h(u)| \leqq \frac{|u|^{h+1}}{h+1} + \frac{|u|^{h+2}}{h+2} + \ldots \leqq \frac{1}{h+1} \frac{|u|^{h+1}}{1-|u|}$$

y así

$$(1-|u|) \log |E_h(u)| \leqq |u|^{h+1}. \qquad [5\text{-}22]$$

Para u arbitrario y $h \geqq 1$, se tiene

$$\log |E_h(u)| \leqq \log |E_{h-1}(u)| + |u|^h. \qquad [5\text{-}23]$$

Si se multiplica [5-23] por $|u|$ y se suma a [5-22], se obtiene la desigualdad

$$\log |E_h(u)| \leqq |u| \log |E_{h-1}(u)| + 2|u|^{h+1}, \qquad [5\text{-}24]$$

válida para $|u| < 1$. Pero para $|u| \geqq 1$, [5-24] es una consecuencia de [5-23], con tal que $|E_{h-1}(u)| \geqq 1$. Si no se satisface esta condición, entonces $\log |E_h(u)| \leqq |u|^h < (2h+1)|u|^{h+1}$ en virtud de [5-23], con lo que se verifica [5-21].

Se deduce ahora [5-21] por inducción. Para $h=0$ tenemos únicamente que observar que $\log |1-u| \leqq \log(1+|u|) \leqq |u|$. Supongamos que se verifica [5-21] con $h-1$ en lugar de h. Entonces, exactamente como se demostró, o se verifica [5-21] o se verifica [5-24]. En el último caso, la hipótesis de inducción nos da

$$\log |E_h(u)| \leqq (2h-1)|u|^{h+1} + 2|u|^{h+1} = (2h+1)|u|^{h+1},$$

con lo que queda demostrada [5-21].

La desigualdad [5-21] da inmediatamente

$$\log |P(z)| = \sum_n \log \left| E_h\left(\frac{z}{a_n}\right) \right| \leqq (2h+1)|z|^{h+1} \sum_n |a_n|^{-h-1},$$

de donde se sigue que $P(z)$ es a lo sumo de orden $h+1$.

Para la desigualdad opuesta supongamos que $f(z)$ es de orden finito λ, y sea h el mayor entero $\leqq \lambda$. Entonces $h+1 > \lambda$, y tenemos que demostrar en primer lugar que $\sum_n |a_n|^{-h-1}$ converge. Para esta demostración necesitamos utilizar la fórmula de Jensen.

Denotemos por $\nu(\rho)$ el número de ceros a_n con $|a_n| < \rho$. Con objeto de hallar una cota superior de $\nu(\rho)$ aplicamos [5-17], con 2ρ en lugar de ρ, y omitimos los términos $\log(2\rho/|a_n|)$ con $|a_n| \geqq \rho$. Hallamos que

$$\nu(\rho)\log 2 \leqq \frac{1}{2\pi}\int_0^{2\pi} \log|f(2\rho e^{i\theta})|\,d\theta - \log|f(0)|. \qquad [5\text{-}25]$$

Como consecuencia de [5-20] se deduce que $\lim_{\rho\to\infty} \nu(\rho)\rho^{-\lambda-\epsilon}=0$ para todo $\epsilon > 0$.

Supongamos ahora que los ceros a_n están ordenados de acuerdo con sus valores absolutos: $|a_1| \leqq |a_2| \leqq \ldots \leqq |a_n| \leqq \ldots$ Entonces es evidente que $n \leqq \nu(2|a_n|)$, y desde un cierto n en adelante se debe tener, p. ej.,

$$n \leqq \nu(2|a_n|) < |a_n|^{\lambda+\epsilon}.$$

De acuerdo con esta desigualdad, la serie $\sum_n |a_n|^{-h-1}$ posee la mayorante

$$\sum_n n^{-\frac{h+1}{\lambda+\epsilon}},$$

y si elegimos ϵ de forma que $\lambda+\epsilon < h+1$, la mayorante converge. Hemos demostrado así que $f(z)$ puede escribirse en la forma [5-19], donde de momento solo se sabe que $g(z)$ es entera.

Queda por demostrar que $g(z)$ es un polinomio de grado $\leqq h$. Para este propósito lo más sencillo es utilizar la fórmula de Poisson-Jensen. Si se aplica el operador $(\partial/\partial x) - i(\partial/\partial y)$ a ambos miembros de la identidad [5-18], se obtiene

$$\frac{f'(z)}{f(z)} = \sum_1^{\nu(\varrho)} (z-a_n)^{-1} + \sum_1^{\nu(\varrho)} \bar{a}_n(\rho^2-\bar{a}_n z)^{-1} + \\ + \frac{1}{2\pi}\int_0^{2\pi} 2\rho e^{i\theta}(\rho e^{i\theta}-z)^{-2}\log|f(\rho e^{i\theta})|\,d\theta.$$

Derivando h veces con respecto a z, esto nos da

$$D^{(h)}\frac{f'(z)}{f(z)} = -h!\sum_1^{\nu(\varrho)}(a_n-z)^{-h-1} + h!\sum_1^{\nu(\varrho)}\bar{a}_n^{h+1}(\rho^2-\bar{a}_n z)^{-h-1} + \\ + (h+1)!\frac{1}{2\pi}\int_0^{2\pi} 2\rho e^{i\theta}(\rho e^{i\theta}-z)^{-h-2}\log|f(\rho e^{i\theta})|\,d\theta.$$

Hagamos ahora tender ρ a ∞. La integral tiende a cero en virtud de

$$\frac{1}{2\pi}\int_0^{2\pi} |\log|f(\rho e^{i\theta})||\,d\theta \leqq M(\rho) - \frac{1}{2\pi}\int_0^{2\pi}\log|f(\rho e^{i\theta})|\,d\theta$$

junto con las acotaciones [5-20], [5-25]. En la segunda suma, si $\rho > 2|z|$, tenemos $|a_n|\cdot|\rho^2-\bar{a}_n z|^{-1} < 2/\rho$. Por tanto, la suma es, en valor absoluto, menor que $2^{h+1}\nu(\rho)\rho^{-h-1}$, y hemos demostrado que esta expresión tiende a cero. Por consiguiente, en el límite, se tiene:

$$D^{(h)}\frac{f'(z)}{f(z)} = -h!\sum_{n=1}^{\infty}(a_n-z)^{-h-1}, \qquad [5\text{-}26]$$

donde la convergencia uniforme hace innecesaria la agrupación de términos.

Poniendo $f(z)=e^{g(z)}P(z)$, tenemos

$$g^{(h+1)}(z) = D^{(h)}\frac{f'}{f} - D^{(h)}\frac{P'}{P}$$

Sin embargo, por el teorema de Weierstrass se puede hallar la cantidad $D^{(h)}P'/P$ por derivación separada de cada factor, y de esta forma obtenemos exactamente el segundo miembro de [5-26]. Por tanto, $g^{(h+1)}(z)$ es idénticamente cero y $g(z)$ tiene que reducirse a un polinomio de grado menor o igual que h.

EJERCICIOS

1. Generalícese la fórmula de Jensen al caso de funciones meromorfas.
2. Determínense el género y el orden de $\cos\sqrt{z}$.
3. Constrúyase un ejemplo de función entera con $\lambda=h+1$.

6. *Principio de simetría.*—En relación con las transformaciones lineales ya ha sido discutido un aspecto elemental del principio de simetría (Cap. I, Sec. 1-3, 3; teorema 4). Existen varias formulaciones más generales que en cierto modo proceden de los trabajos de H. A. Schwarz.

El principio de simetría está basado en la observación de que si $u(z)$ es una función armónica, entonces $u(\bar{z})$ es igualmente armónica, y si $f(z)$ es una función analítica, entonces $\overline{f(\bar{z})}$ es también analítica. Con más precisión, si $u(z)$ es armónica y $f(z)$ analítica en una región Ω, resulta que $u(\bar{z})$ es armónica y $\overline{f(\bar{z})}$ analítica como funciones de z en la región $\bar{\Omega}$ obtenida por simetría de Ω respecto al eje real; esto es, $z \in \bar{\Omega}$ si (y solo si) $\bar{z} \in \Omega$. La demostración de estas propiedades se limita a comprobaciones triviales.

Para conclusiones ulteriores consideremos primero el caso de funciones analíticas. Empecemos por suponer que $\Omega = \bar{\Omega}$, en cuyo caso se dice que Ω es simétrica con respecto al eje real; como consecuencia del hecho de ser conexa, dicha región debe cortar al eje real. Sea $f(z)$ analítica en Ω y supongamos que $f(z)$ es real en la intersección de Ω con el eje real. Entonces afirmamos que $f(z)$ satisface la ecuación $f(z) = \overline{f(\bar{z})}$ en todo Ω; esta ecuación significa que $f(z)$ toma valores conjugados en puntos conjugados.

La demostración es extremadamente sencilla. Necesitamos observar únicamente que la función $f(z) - \overline{f(\bar{z})}$ es analítica en Ω y se anula sobre el eje real. Por consiguiente, tiene que anularse idénticamente en Ω, con lo que se obtiene $f(z) = \overline{f(\bar{z})}$. El razonamiento prueba más de lo que se enunció; parece suficiente suponer que $f(z)$ es real en infinitos puntos del eje real con un punto de acumulación en Ω.

Con un poco más de generalidad, suprimamos el supuesto de que Ω es simétrica, pero exijamos que corte al eje real. Supongamos además que todas las componentes de $\Omega\bar{\Omega}$ cortan al eje real, en cuyo caso son regiones simétricas. Sea $f(z)$ analítica en Ω y real sobre el eje real. Entonces $f(z)$ satisface la relación simétrica $f(z) = \overline{f(\bar{z})}$ en $\Omega\bar{\Omega}$. Así, pues, es posible *definir* una función analítica $F(z)$ sobre la región simétrica $\Omega + \bar{\Omega}$ que sea igual a $f(z)$ en Ω e igual a $\overline{f(\bar{z})}$ en $\bar{\Omega}$; de hecho, estas definiciones no son contradictorias sobre $\Omega\bar{\Omega}$. En otras palabras: $f(z)$ tiene una prolongación analítica simétrica en $\Omega + \bar{\Omega}$. Para mayor sencillez, se suele denotar la prolongación nuevamente por $f(z)$, y satisface la relación $\overline{f(z)} = f(\bar{z})$. El resultado sigue siendo válido bajo la hipótesis más débil de que $f(z)$ es real sobre un subconjunto del eje real que tenga al menos un punto de acumulación en cada componente de $\Omega\bar{\Omega}$.

En el caso de funciones armónicas, empezaremos nuevamente con una función $u(z)$ que sea armónica sobre una región simétrica Ω, y esta vez supondremos que $u(z)$ se anula sobre el eje real. Construyamos la función

$$U(z) = u(z) + u(\bar{z}),$$

la cual se sabe que es armónica en Ω. Para z real es evidente que $\partial U/\partial x = 0$, comprobándose fácilmente que $\partial U/\partial y = 0$. Por tanto, la función $(\partial U/\partial x) - i(\partial U/\partial y)$ es analítica sobre el eje real y debe de anularse idénticamente. Se sigue que U debe reducirse a una constante, y esta constante es evidentemente cero. Hemos probado así que $u(\bar{z}) = -u(z)$. El resultado puede generalizarse de la misma manera que se hizo para funciones analíticas.

La misma demostración sirve bajo la hipótesis de que $u(z)$ se anula sobre un segmento abierto del eje real en cada componente de $\Omega\bar{\Omega}$. Si solo se sabe que u se anula sobre un subconjunto del eje real con puntos de acumulación en todas las componentes, es necesario añadir un razonamiento previo. Si x_0 es uno de los puntos de acumulación, consideremos un entorno de x_0 que esté contenido en $\Omega\bar{\Omega}$. En este entorno U tiene una función conjugada V. Aplicando nuestros resultados anteriores sobre funciones analíticas a $V - iU$, hallamos que U se anula idénticamente sobre el eje real en el entorno de x_0.

Resumimos todos estos resultados en la siguiente proposición.

TEOREMA 8. *Supongamos que Ω y todas las componentes de $\Omega\bar{\Omega}$ cortan al eje real. Si $u(z)$ es armónica en Ω y se anula sobre el eje real, o si $f(z)$ es analítica en Ω y real sobre el eje real, entonces $u(z)$ tiene una prolongación armónica en $\Omega + \bar{\Omega}$ que satisface la igualdad $u(\bar{z}) = -u(z)$, y $f(z)$ tiene una prolongación analítica que satisface la relación $f(z) = \overline{f(\bar{z})}$.*

Se verifica la hipótesis en cuanto se tenga que $u(z)$ se anula o que $f(z)$ es real sobre un subconjunto del eje real con un punto de acumulación en cada componente de $\Omega\bar{\Omega}$.

El teorema 8 se aplica, ordinariamente, en circunstancias ligeramente diferentes. Supongamos que Ω está, p. ej., en el semiplano superior, pero que la frontera de Ω contiene un subconjunto abierto E

del eje real. Entonces $\Omega+E+\bar{\Omega}$ es una región simétrica, planteándose la cuestión de si $u(z)$ o $f(z)$ admiten una prolongación armónica o analítica, respectivamente, simétrica en esta región.

En el caso de una función armónica $u(z)$, basta suponer que $u(z)$ tiende a cero cuando z se acerca a E. Necesitamos únicamente probar que la función $u_0(z)$, definida como $u(z)$ en Ω, como cero sobre E y como $-u(\bar{z})$ en $\bar{\Omega}$, es armónica en $\Omega+E+\bar{\Omega}$. Pero esto es evidentemente obvio como consecuencia del teorema 5. En efecto, $u_0(z)$ es desde luego continua, se satisface ciertamente la propiedad del valor medio en $\Omega+\bar{\Omega}$ y para $z_0 \in E$ el valor medio sobre una circunferencia de centro z_0 es cero como consecuencia de la simetría y, por tanto, igual a $u_0(z_0)$.

Para una función analítica $f(z)$, dada en Ω, la suposición apropiada es que $v(z)=\operatorname{Im} f(z)$ tiende a cero cuando z se acerca a E. Sabemos que $v(z)$ tiene una prolongación armónica $v_0(z)$ que satisface la igualdad $v_0(\bar{z})=-v_0(z)$. En un disco circular suficientemente pequeño, de centro $z_0 \in E$, $v_0(z)$ tiene una función armónica conjugada $-u_0(z)$, y mediante una determinación adecuada de la constante aditiva podemos hacer que $u_0(z)=\operatorname{Re} f(z)$ en la mitad superior del disco. Para discos con parte común las funciones $u_0(z)$ coinciden sobre dichas partes comunes; por tanto, $f(z)$ tiene una prolongación analítica en una región que contiene a $\Omega+E$. Puede aplicarse ahora el teorema 8, con lo que se obtiene que $f(z)$ posee una prolongación analítica en toda la región $\Omega+E+\bar{\Omega}$.

TEOREMA 9. *Sea Ω una región en el semiplano superior o en el inferior, cuya frontera contiene un subconjunto abierto E del eje real. Si $v(z)$ es armónica en Ω y tiende a cero cuando z se acerca a E, entonces $v(z)$ tiene una prolongación armónica en $\Omega+E+\bar{\Omega}$ que verifica la relación de simetría $v(\bar{z})=-v(z)$. Si, en la misma situación, $v(z)$ es la parte imaginaria de una función analítica $f(z)$ en Ω, $f(z)$ posee en tal caso una prolongación analítica que satisface la igualdad $f(\bar{z})=\overline{f(z)}$.*

Observemos que en la última parte del teorema la hipótesis no incluye la existencia de un límite de $\operatorname{Re} f(z)$ cuando z se acerca a un punto de E. En muchas aplicaciones, este es un punto muy esencial.

El teorema 9 admite generalizaciones obvias. Puede suponerse que E está en una circunferencia y que los valores de $f(z)$ se aproximan

a una circunferencia arbitraria, en lugar de al eje real, cuando z tiende a E. En tales casos la prolongación satisface una relación de simetría con respecto a las circunferencias, lo mismo que en el caso de las transformaciones lineales.

Sin embargo, la generalización más importante está relacionada con el concepto de *arco analítico* que vamos a definir. La ecuación $z=z(t)$, $0 \leqq t \leqq 1$, se dice que representa un arco analítico γ si para todo t_0 en el intervalo paramétrico, $z(t)$ admite una representación en serie de potencias

$$z(t)=a_0+a_1(t-t_0)+a_2(t-t_0)^2+\dots \qquad [5\text{-}27]$$

con $a_1 \neq 0$, válida en algún intervalo $(t_0-\delta,\ t_0+\delta)$. Un arco analítico de Jordan es, naturalmente, aquel en que $z(t_1) \neq z(t_2)$ para $t_1 \neq t_2$.

Consideremos t como una variable compleja. La serie [5-27] tiene un radio de convergencia $R \geqq \delta$ y representa una función analítica en su círculo de convergencia. En la parte común de dos círculos de convergencia no disjuntos las correspondientes funciones analíticas coinciden sobre el eje real, siendo, por tanto, idénticas. Se deduce que $z(t)$ tiene una prolongación analítica, que seguimos denotando por $z(t)$, en la región formada por la unión de todos los círculos de convergencia. En particular, podemos elegir un $\rho > 0$ tal que $z(t)$ esté definida en la región D_ρ, formada por todos los puntos cuya distancia al segmento $(0,1)$ es $< \rho$.

Si γ es un arco de Jordan, afirmamos que la aplicación de D_ρ mediante $z(t)$ es biunívoca con tal que ρ sea suficientemente pequeño. Si esto no fuera cierto se podrían hallar pares $t'_n \neq t''_n$ que tiendan al segmento $(0,1)$, con $z(t'_n)=z(t''_n)$, y se podrían extraer sendas subsucesiones convergentes $\{t'_{n_k}\}$, $\{t''_{n_k}\}$ con límites t', t'' sobre el segmento; por la continuidad, $z(t')=z(t'')$ y, por tanto, $t'=t''$. Denotando este valor común por t_0, tenemos por hipótesis que $z'(t_0)=a_1 \neq 0$, existiendo un entorno de t_0 en el que la aplicación $z(t)$ es biunívoca. Para k suficientemente grande, los puntos t'_{n_k}, t''_{n_k} están en este entorno, con lo que se llega a una contradicción.

Brevemente, hemos demostrado que todo arco analítico de Jordan γ es la imagen de un segmento respecto de una aplicación conforme de una región simétrica D sobre una región D', situación que viene ilustrada mediante la figura 5-1. Si t y $\bar{t}$ son puntos conjugados de D, los correspondientes puntos $z(t)$ y $z(\bar{t})$ pueden considerarse *simétricos con respecto a* γ. Podemos decir también que $z(t)$ está a la izquierda de γ si t tiene parte imaginaria positiva, y a la derecha de γ, si $\operatorname{Im} t < 0$.

Es importante demostrar que este concepto de simetría depende solo de γ y no de su representación paramétrica. Para probarlo, sean $z = z_1(t_1)$ y $z = z_2(t_2)$ representaciones paramétricas del mismo arco γ, y elijamos regiones correspondientes D_1, D'_1 y D_2, D'_2, como anteriormente se hizo. Consideremos la componente de $D'_1 D'_2$ que contiene

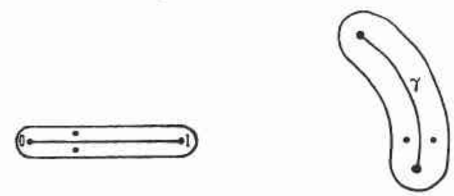

FIG. 5-1.

a γ y sus imágenes inversas en los planos t_1 y t_2. Las aplicaciones inducen una correspondencia analítica biunívoca entre las imágenes inversas, y por el teorema 8. a valores conjugados de t_2 corresponden valores conjugados de t_1. Hemos establecido así la existencia de una región en la que está definida la simetría y tiene el mismo significado con respecto a ambas representaciones. Además, puntos que están en el mismo lado de γ en una representación, también están en el mismo lado en la otra, pues el criterio es que puedan unirse por un arco en el que ningún punto es simétrico de sí mismo. A lo sumo podría existir una inversión de izquierda y derecha, pero puesto que a valores reales crecientes de t_2 corresponden valores crecientes de t_1, del hecho de ser la aplicación conforme se sigue que los conceptos de derecha e izquierda deben realmente conservarse.

Incidentalmente, el mismo razonamiento prueba que dos arcos analíticos γ_1 y γ_2 no pueden cortarse en infinitos puntos sin ser arcos parciales no disjuntos del mismo arco analítico. Con la misma notación de antes, sean t'_1 y t'_2 puntos de acumulación de los valores del parámetro que corresponden a puntos de intersección. En los entornos complejos de t'_1 y t'_2 existe una correspondencia analítica bien definida entre t_1 y t_2, tal que $z_1(t_1) = z_2(t_2)$, y el supuesto significa que t_2 es real para un conjunto de valores reales de t_1 con t'_1 como punto de acumulación. Utilizando la forma más fuerte del teorema 8, se concluye que t_2 es real para todo valor real de t_1 próximo a t'_1, y mediante un razonamiento que utiliza el hecho de ser

conexo en una forma ya familiar, resulta sin dificultad que γ_1 y γ_2 son arcos parciales del mismo arco analítico o curva cerrada analítica.

Consideremos un arco analítico de Jordan γ y una región D' en la que esté definida la simetría con respecto a γ; denotemos por D'_+ y D'_- las subregiones formadas por puntos a la izquierda y a la derecha de γ, respectivamente. Sea Ω una región y E un subconjunto de γ sujeto a la siguiente condición: todo $z_0 \in E$ posee un entorno Δ tal que $\gamma\Delta \subset E$ y $\Delta\Omega = \Delta D'_+$.

TEOREMA 10. *Si $f(z)$ es analítica en Ω y si los límites de $w = f(z)$ cuando z se aproxima a E están todos sobre un arco analítico de Jordan Γ en el plano w, entonces $f(z)$ tiene una prolongación analítica en una región que contiene a $\Omega + E$.*

Se deduce la demostración de la consideración de los parámetros t y τ asociados con γ y Γ, respectivamente. Las hipótesis son tales que puede aplicarse el teorema 9 a la correspondencia analítica que $f(z)$ induce entre los valores complejos t y τ. El lector no encontrará dificultad en suplir los detalles que completan el razonamiento. Hemos suprimido la parte del teorema que afirmaría que a puntos simétricos corresponden puntos simétricos, pues en el caso general esta información carece de valor práctico.

El teorema 10 es particularmente importante cuando Ω es una región simplemente conexa y $f(z)$ aplica Ω conformemente sobre $|w| < 1$. Sabemos entonces por el teorema 11 del capítulo IV, sección 4-4, 2, que los límites de $f(z)$ cuando z se aproxima a E están en la circunferencia $|w| = 1$. Se concluye que $f(z)$ sigue siendo analítica sobre E, y en este caso podemos afirmar incluso que $f(z)$ posee una prolongación analítica en toda la región $\Omega + E + D'_-$ (excepto por un posible polo). Además, $f(z)$ es también univalente en dicha región. En efecto, es evidente que dos puntos de $\Omega + D'_-$ no pueden tener la misma imagen. En cuanto a la posibilidad de que $f(z)$ tomase el mismo valor en dos puntos de E, basta hacer notar que entonces $f(z)$ tendría que aplicar entornos disjuntos sobre conjuntos abiertos que tendrían un punto de $|w| = 1$ en común, pero que carecerían de puntos comunes en el interior del círculo unidad.

EJERCICIOS

1. Si $f(z)$ es analítica en todo el plano, real sobre el eje real e imaginaria pura en el eje imaginario, pruébese que $f(z)$ es impar.

2. Demuéstrese que toda función f que sea analítica en una región

simétrica Ω puede escribirse en la forma f_1+if_2, siendo f_1 y f_2 analíticas en Ω y reales sobre el eje real.

3. Si $f(z)$ es analítica en $|z|\leqq 1$ y satisface la igualdad $|f|=1$ para $|z|=1$, demuéstrese que $f(z)$ es racional. ¿Hasta qué punto pueden debilitarse las hipótesis?

5-2. Funciones subarmónicas.—La ecuación de Laplace en una dimensión tendría la forma $d^2u/dx^2=0$. Las funciones armónicas de una variable serían, pues, las funciones lineales $u=ax+b$. Se dice que una función $v(x)$ es *convexa* si, en cualquier intervalo, es a lo sumo igual a la función lineal $u(x)$ que tiene los mismos valores que $v(x)$ en los extremos del intervalo.

Si se generaliza esta situación a dos dimensiones, obtenemos la clase de las *funciones subarmónicas*. Funciones lineales corresponden a funciones armónicas, intervalos corresponden a regiones y los extremos de un intervalo corresponden a la frontera de la región. De acuerdo con esto, una función $v(z)$ de una variable compleja o de dos variables reales se llamará subarmónica si en cualquier región es menor o igual que la función armónica $u(z)$ que coincide con $v(z)$ sobre la frontera de la región. Puesto que esta formulación entraña que podemos resolver el problema de Dirichlet, resulta preferible reemplazar la condición por el requisito más sencillo de que $v(z)\leqq \leqq u(z)$ sobre la frontera de la región implica $v(z)\leqq u(z)$ en la región.

Las funciones subarmónicas han sido introducidas en fecha bastante reciente. Muchas propiedades importantes de las funciones armónicas son, en realidad, ciertas para la clase más amplia de las funciones subarmónicas y, por tanto, deberían ser demostradas con esta generalidad. Una razón más poderosa para introducir las funciones subarmónicas en un texto elemental es el hecho de que proporcionan la herramienta conocida más sencilla para resolver el problema de Dirichlet.

1. *Definición y propiedades sencillas.*—En lugar de utilizar la definición indicada en el párrafo de introducción elegimos un enunciado equivalente, que en algunos aspectos es más sencillo:

DEFINICIÓN 1. *Una función real continua $v(z)$, definida en una región Ω, se dice que es subarmónica en Ω si para cualquier función armónica $u(z)$ en una región $\Omega' \subset \Omega$ la diferencia $v-u$ satisface el principio del máximo en Ω'.*

La condición significa que $v-u$ no puede tener un máximo en Ω'

sin ser idénticamente constante. En particular, v no puede tener un máximo en Ω. Es importante observar que la definición tiene carácter local: si v es subarmónica en un entorno de cada punto $z \in \Omega$, entonces es subarmónica en Ω. La demostración es inmediata. Se dice que una función es subarmónica en un punto z_0 si lo es en un entorno de z_0. Por tanto, una función es subarmónica en una región si (y solo si) lo es en todos los puntos de la región.

Una función armónica es trivialmente subarmónica.

Una condición suficiente para que una función sea subarmónica es que v tenga su laplaciana positiva. En efecto, si $v-u$ tiene un máximo, se deduce por un cálculo elemental que $\partial^2/\partial x^2(v-u)\leqq 0$, $\partial^2/\partial y^2(v-u)\leqq 0$ en ese punto; esto implicaría que $\Delta v=\Delta(v-u)\leqq 0$. La condición no es necesaria, ya que, de hecho, no es preciso en absoluto que una función subarmónica tenga derivadas parciales. Si la función posee derivadas continuas de primero y segundo orden, puede demostrarse que la condición $\Delta v\geqq 0$ es necesaria y suficiente. Puesto que no necesitaremos esta propiedad, su demostración queda relegada a la sección de ejercicios. La condición da un método sencillo para saber si una función elemental dada de x e y es subarmónica.

Demostraremos ahora que las funciones subarmónicas pueden caracterizarse por una desigualdad que generaliza la propiedad del valor medio de las funciones armónicas:

TEOREMA 11. *Una función continua $v(z)$ es subarmónica en Ω si (y solo si) satisface la desigualdad*

$$v(z_0)\leqq \frac{1}{2\pi}\int_0^{2\pi} v(z_0+re^{i\theta})\,d\theta \qquad [5\text{-}28]$$

para todo disco $|z-z_0|\leqq r$ contenido en Ω.

La suficiencia se deduce del hecho de que [5-28], más que la propiedad del valor medio, es lo que realmente se necesita con objeto de probar que v no puede tener un máximo sin reducirse a una constante. Puesto que $v-u$ satisface la misma desigualdad, se sigue que v es subarmónica.

A fin de demostrar la necesidad, formemos la integral de Poisson $P_v(z)$ en el disco $|z-z_0|<r$ tomando los valores de v sobre la circunferencia $|z-z_0|=r$[1]. Si v es subarmónica, la función $v-P_v$ no puede tener un máximo en el disco a no ser que sea constante. Por

[1] La notación no está completamente de acuerdo con la utilizada en la sección 5-1, 3, pero se entenderá fácilmente.

el teorema 4, $v - P_v$ tiende a cero cuando z se aproxima a un punto sobre la circunferencia. Por tanto, $v - P_v$ tiene un máximo en el disco cerrado. Si el máximo fuera positivo, debería alcanzarse en punto interior, y la función no sería constante. Esto es una contradicción, luego tenemos que $v \leqq P_v$. Para $z = z_0$ obtenemos $v(z_0) \leqq P_v(z_0)$, que es la desigualdad [5-28].

A continuación damos un cierto número de propiedades elementales de las funciones subarmónicas:

a) *Si v es subarmónica, también lo es kv para cualquier constante $k \geqq 0$.*

b) *Si v_1 y v_2 son subarmónicas, también lo es $v_1 + v_2$.*

Estas son consecuencias inmediatas del teorema 11. La propiedad siguiente se deduce con más facilidad de la definición original.

c) *Si v_1 y v_2 son subarmónicas en Ω, entonces $v = \text{máx}\,(v_1, v_2)$ es igualmente subarmónica en Ω.*

Ha de entenderse la notación en el sentido de que $v(z)$ es en cada punto igual al mayor de los valores $v_1(z)$ y $v_2(z)$. La continuidad de v es obvia. Supongamos ahora que $v - u$ tiene un máximo en $z_0 \in \Omega'$, donde u está definida y es armónica en Ω'. Podemos suponer que $v(z_0) = v_1(z_0)$. Tenemos entonces que

$$v_1(z) - u(z) \leqq v(z) - u(z) \leqq v(z_0) - u(z_0) = v_1(z_0) - u(z_0)$$

para $z \in \Omega'$. Por tanto, $v_1 - u$ es constante, y por la misma desigualdad $v - u$ debe ser también constante. Queda probado que v es subarmónica.

Sea Δ un disco cuyo cierre está contenido en Ω, y denotemos por P_v la integral de Poisson formada con los valores de v sobre su circunferencia. Entonces es cierto que

d) *Si v es subarmónica, entonces la función v' (definida como P_v en Δ y como v fuera de Δ) es también subarmónica.*

La continuidad de v' se deduce mediante el teorema de Schwarz (Teorema 4). Hemos probado que $v \leqq P_v$ en Δ y, por tanto, $v \leqq v'$ en todo Ω. Es evidente que v' es subarmónica en el interior y exterior de Δ. Supongamos ahora que $v' - u$ alcanza un máximo en un punto z_0 sobre la circunferencia de Δ. Se sigue inmediatamente que $v - u$ también debería tener un máximo en z_0. Por consiguiente, $v - u$ sería constante, y la desigualdad

$$v - u \leqq v' - u \leqq v'(z_0) - u(z_0) = v(z_0) - u(z_0)$$

muestra que $v' - u$ es igualmente constante. Se tiene, pues, que v' es subarmónica.

EJERCICIOS

1. Demuéstrese que las funciones $|x|$, $|z|^\alpha$ $(\alpha \geqq 0)$ y $\log(1 + |z|^2)$ son subarmónicas.

2. Si $f(z)$ es analítica, pruébese que $|f(z)|^\alpha$ $(\alpha \geqq 0)$ y $\log(1 + |f(z)|^2)$ son subarmónicas.

3. Si v es continua a la vez que sus derivadas parciales hasta el segundo orden, demuéstrese que v es subarmónica si (y sólo si) $\Delta v \geqq 0$. *Sugerencia:* Para la suficiencia, pruébese primero que $v + \epsilon x^2$, $\epsilon > 0$, es subarmónica. Para la necesidad, demuéstrese que si $\Delta v < 0$, el valor medio sobre un círculo sería una función decreciente del radio.

4. Pruébese que una función subarmónica sigue siendo subarmónica si la variable independiente se somete a una aplicación conforme.

5. Enúnciese y pruébese un teorema al efecto de que un límite uniforme de funciones subarmónicas es una función subarmónica.

6. Extiéndase el principio de Harnack a las funciones subarmónicas

2. *Solución del problema de Dirichlet.*—El primero que utilizó las funciones subarmónicas para el estudio del problema de Dirichlet fue O. Perron. Su método se caracteriza por su extrema generalidad y es completamente elemental.

Consideremos una región acotada Ω y una función real $f(\zeta)$ definida sobre su frontera Γ (para mayor claridad, los puntos frontera serán denotados por ζ). Para empezar, no es necesario siquiera que $f(\zeta)$ sea continua, pero por motivos de sencillez supondremos que está acotada, $|f(\zeta)| \leqq M$. Podemos asociar a cada f una función armónica $u(z)$ en Ω, definida mediante un sencillo proceso que detallaremos más adelante. Si f es continua y si Ω satisface ciertas condiciones muy suaves, la correspondiente función u resolverá el problema de Dirichlet en Ω con los valores frontera f.

Definimos la clase $\mathfrak{B}(f)$ de funciones v con las siguientes propiedades:

a) v es subarmónica en Ω;

b) $\overline{\lim}_{z \to \zeta}\, v(z) \leqq f(\zeta)$ para todo $\zeta \in \Gamma$.

El significado preciso de *b)* es este: dado $\epsilon > 0$ y un punto $\zeta \in \Gamma$, existe un entorno Δ de ζ tal que $v(z) < f(\zeta) + \epsilon$ en $\Delta\Omega$. La clase $\mathfrak{B}(f)$ no es vacía, ya que contiene todas las constantes $\leqq -M$. Demostraremos:

LEMA 1. *La función u, definida como $u(z) = \overline{\text{ext}}\, v(z)$ para $v \in \mathfrak{B}(f)$, es armónica en Ω.*

En primer lugar, cada v es $\leqq M$ en Ω. Esta es una consecuencia bastante sencilla del principio del máximo, pero a causa de su importancia deseamos explicar este punto con algún detalle. Para un $\epsilon > 0$ dado, sea E el conjunto de puntos $z \in \Omega$, para los que $v(z) \geqq M + \epsilon$. Los puntos z en el complemento $C(E)$ son de tres clases: 1) puntos en el exterior de Ω; 2) puntos sobre Γ, y 3) puntos en Ω con $v(z) < M + \epsilon$. En el caso 1) z tiene un entorno contenido en el exterior; en el 2) hay un entorno Δ con $v < M + \epsilon$ en $\Delta\Omega$, por la propiedad *b)*, y en el caso 3) existe, por la continuidad, un entorno en Ω con $v < M + \epsilon$. Por tanto, $C(E)$ es abierto y E cerrado. Además, puesto que Ω está acotado, E es compacto. Si E no fuera vacío, v tendría un máximo sobre E, y este debería también ser un máximo en Ω. Esto es imposible, pues a causa de *b)* v no puede ser una constante $> M$. Por consiguiente, E es vacío para todo ϵ, de donde se sigue que $v \leqq M$ en Ω.

Consideremos un disco Δ, cuyo cierre esté contenido en Ω, y un punto $z_0 \in \Delta$. Existe una sucesión de funciones $v_n \in \mathfrak{B}(f)$ tal que $\lim\limits_{n\to\infty} v_n(z_0) = u(z_0)$. Pongamos $V_n = \text{máx}\,(v_1, v_2, \ldots, v_n)$. Entonces las V_n forman una sucesión no decreciente de funciones de $\mathfrak{B}(f)$. Construyamos $V'_n = V_n$ en el exterior de Δ e igual a la integral de Poisson de V_n en Δ. Por la propiedad *d)* de la sección precedente, las V'_n pertenecen todavía a $\mathfrak{B}(f)$. Forman una sucesión no decreciente, y la desigualdad $v_n(z_0) \leqq V_n(z_0) \leqq V'_n(z_0) \leqq u(z_0)$ prueba que $\lim\limits_{n\to\infty} V'_n(z_0) = u(z_0)$. Por el principio de Harnack, la sucesión $\{V'_n\}$ converge a una función límite armónica U en Δ que satisface la desigualdad $U \leqq u$ y para la que $U(z_0) = u(z_0)$.

Supongamos ahora que empezamos el mismo proceso partiendo de otro punto $z_1 \in \Delta$. Elegimos $w_n \in \mathfrak{B}(f)$ tal que $\lim\limits_{n\to\infty} w_n(z_1) = u(z_1)$, pero esta vez, antes de proceder a la construcción, reemplazamos w_n por $\bar{w}_n = \text{máx}\,(v_n, w_n)$. Poniendo $W_n = \text{máx}\,(\bar{w}_1, \ldots, \bar{w}_n)$, construimos la sucesión correspondiente $\{W'_n\}$ con la ayuda de la integral de Poisson, con lo que llegamos a una función límite armónica U_1 que satisface la desigualdad $U \leqq U_1 \leqq u$ y además $U_1(z_1) = u(z_1)$. Se sigue que $U - U_1$ alcanza el máximo cero en z_0. Por tanto, U es idénticamente igual a U_1 y hemos demostrado que $u(z_1) = U(z_1)$ para cualquier $z_1 \in \Delta$. Se deduce que u es armónica en cualquier disco Δ y, en consecuencia, en todo Ω.

Investigaremos ahora las circunstancias bajo las cuales u resuelve el problema de Dirichlet para f continua. Observemos en primer lugar que este problema no siempre tiene solución. Así, p. ej., si Ω es el disco sin centro $0 < |z| < 1$, consideremos los valores frontera $f(0) = 1$ y $f(\zeta) = 0$ para $|\zeta| = 1$. Una función armónica con estos valores frontera debería estar acotada y tendría, por consiguiente, una singularidad evitable en el origen. Pero entonces el principio del máximo implicaría que la función se anula idénticamente, y así no tendría el valor frontera 1 en el origen. Se sigue que no puede existir solución.

También es fácil ver que una solución, si existe, debe ser idéntica a u. En efecto, si U es una solución, ante todo es evidente que $U \in \mathfrak{B}(f)$ y, por tanto, $u \geqq U$. La desigualdad opuesta $u \leqq U$ se sigue del principio del máximo que implica que $v \leqq U$ para todo $v \in \mathfrak{B}(f)$.

Se puede afirmar la existencia de una solución para una amplia clase de regiones. En términos generales, la solución existe siempre que el complemento de Ω no es demasiado "delgado" en el entorno de cualquier punto frontera. Empecemos demostrando un lema que, superficialmente, parece tener poca relación con el concepto de *delgadez*.

LEMA 2. *Supongamos que existe una función armónica $\omega(z)$ en Ω cuyos valores frontera continuos $\omega(\zeta)$ son estrictamente positivos excepto en un punto ζ_0, donde $\omega(\zeta_0) = 0$. Entonces, si $f(\zeta)$ es continua en ζ_0, la función correspondiente u, determinada por el método de Perron, satisface la siguiente igualdad:* $\lim\limits_{z\to\zeta_0} u(z) = f(\zeta_0)$.

Quedará demostrado el lema si probamos que $\overline{\lim\limits_{z\to\zeta_0}}\ u(z) \leqq f(\zeta_0) + \epsilon$ y $\underline{\lim\limits_{z\to\zeta_0}}\ u(z) \geqq f(\zeta_0) - \epsilon$ para todo $\epsilon > 0$. Suponemos todavía que Ω está acotado y que $|f(\zeta)| \leqq M$.

Determinamos un entorno Δ de ζ_0 tal que $|f(\zeta) - f(\zeta_0)| < \epsilon$ para $\zeta \in \Delta$. En el complemento $\Omega - \Delta\Omega$ la función $\omega(z)$ tiene un mínimo positivo ω_0. Consideremos los valores frontera de la función armónica

$$W(z) = f(\zeta_0) + \epsilon + \frac{\omega(z)}{\omega_0}\,[M - f(\zeta_0)].$$

Para $\zeta \in \Delta$ tenemos que $W(\zeta) \geqq f(\zeta_0) + \epsilon > f(\zeta)$, y para ζ en el exterior de Δ obtenemos $W(\zeta) \geqq M + \epsilon > f(\zeta)$. Por el principio del máximo, cualquier función $v \in \mathfrak{B}(f)$ debe, por tanto, satisfacer la desigualdad $v(z) < W(z)$. Se sigue que $u(z) \leqq W(z)$ y, por consiguiente, $\overline{\lim_{z\to\zeta_0}}\, u(z) \leqq W(\zeta_0) = f(\zeta_0) + \epsilon$, que es la primera desigualdad que teníamos que probar.

Para la segunda desigualdad necesitamos únicamente probar que la función

$$V(z) = f(\zeta_0) - \epsilon - \frac{\omega(z)}{\omega_0}\,[M + f(\zeta_0)]$$

pertenece a $\mathfrak{B}(f)$. Para $\zeta \in \Delta$ tenemos que $V(\zeta) \leqq f(\zeta_0) - \epsilon < f(\zeta)$, y en todos los otros puntos frontera, $V(\zeta) \leqq -M - \epsilon < f(\zeta)$. Puesto que V es armónica, pertenece a $\mathfrak{B}(f)$ y resulta $u(z) \geqq V(z)$, $\underline{\lim}_{z\to\zeta_0}\, u(z) \geqq V(\zeta_0) = f(\zeta_0) - \epsilon$. Esto completa la demostración.

A la función $\omega(z)$ del lema 2 se le llama a veces una *barrera* en el punto ζ_0. Evidentemente, podemos decir ahora que el problema de Dirichlet es resoluble con tal que exista una barrera en cada punto frontera. Resta por formular condiciones geométricas que impliquen la existencia de una barrera. En realidad, no se conocen condiciones necesarias y suficientes de carácter puramente geométrico, pero es relativamente fácil encontrar condiciones suficientes con un amplio campo de aplicación.

Para empezar con el caso más sencillo, supongamos que $\Omega + \Gamma$ está contenido en un semiplano abierto, excepto un punto ζ_0, que pertenece a la curva frontera. Si la dirección de esta curva es α (con el semiplano a la izquierda), entonces $\omega(z) = \operatorname{Im} e^{-i\alpha}(z - \zeta_0)$ es una barrera en ζ_0.

Con más generalidad, supongamos que ζ_0 es el extremo de un segmento, todos los puntos del cual, excepto ζ_0, están en el exterior de Ω. Si el otro extremo se denota por ζ_1, sabemos que se puede definir una rama uniforme de

$$\sqrt{\frac{z-\zeta_0}{z-\zeta_1}}$$

fuera del segmento. Con una determinación adecuada del ángulo α se ve fácilmente que la función

$$\operatorname{Im}\left[e^{-i\alpha}\sqrt{\frac{z-\zeta_0}{z-\zeta_1}}\right]$$

es una barrera en ζ_0.

Este no es el resultado más potente que puede obtenerse por estos métodos, pero es suficiente para la mayoría de las aplicaciones. Nos contentaremos, por tanto, con el enunciado siguiente:

TEOREMA 12. *Se puede resolver el problema de Dirichlet para cualquier región Ω tal que cada punto frontera sea el extremo de un segmento cuyos restantes puntos sean exteriores a Ω.*

Se verifica la hipótesis si Ω y su complemento tienen una frontera común consistente en un número finito de curvas cerradas simples con tangente en cada punto. Se admiten también vértices y ciertos tipos de puntos cuspidales [1].

5-3. Aplicaciones canónicas de regiones múltiplemente conexas.—El teorema fundamental de Riemann sobre aplicaciones conformes nos permite concluir que cualquier par de regiones simplemente conexas, con la excepción de todo el plano, pueden aplicarse conformemente una sobre otra, o que son *conformemente equivalentes.* Esto ya no es cierto para regiones múltiplemente conexas del mismo orden de conexión. En su lugar debemos tratar de hallar un sistema de *regiones canónicas,* con la propiedad de que cada región múltiplemente conexa sea conformemente equivalente a una (y solo una) región canónica. La elección de las regiones canónicas es, en cierta forma, arbitraria, existiendo varios tipos con propiedades igualmente sencillas.

Con objeto de mantenernos en un nivel elemental nos limitaremos al estudio de regiones de orden de conexión finito. Hallaremos que la etapa básica que nos conducirá hacia la construcción de aplicaciones canónicas es la introducción de ciertas funciones armónicas con un comportamiento particularmente sencillo sobre la frontera. De estas, las *medidas armónicas* se refieren únicamente a regiones y uno de sus contornos, mientras que la *función de Green* está relacionada con la región y un punto interior.

1. *Medidas armónicas.*—Al estudiar las aplicaciones conformes de una región Ω podemos, naturalmente, reemplazar Ω por cualquier región de la que se sepa que es conformemente equivalente a Ω; es decir, podemos efectuar aplicaciones conformes previas a capricho. Como consecuencia de esta libertad en la elección de la región original, resulta que nunca es necesario enfrentarse con las dificultades a que puede dar lugar una estructura complicada en la frontera.

[1] El mejor resultado que puede demostrarse esencialmente con el mismo método es el siguiente: *El problema de Dirichlet puede resolverse para cualquier región cuyo complemento es tal que ninguna de sus componentes se reduce a un punto.* De esta proposición se puede deducir, con toda facilidad, una demostración independiente del teorema fundamental de Riemann sobre aplicaciones conformes.

En lo que sigue Ω denota una región plana de orden de conexión $n > 1$. Se denotan las componentes del complemento por $E_1, E_2, \ldots, E_n$, y convenimos en que E_n es la componente no acotada. Podemos suponer, sin que haya pérdida de generalidad, que ninguna de las E_k se reduce a un punto, pues es evidente que una componente puntual es una singularidad evitable de cualquier función que pueda dar lugar a la aplicación; por tanto, las aplicaciones siguen siendo las mismas si se añade a la región este punto frontera aislado.

El complemento de E_n es una región simplemente conexa Ω'. Por el teorema de Riemann, Ω' puede aplicarse conformemente sobre el disco $|z| < 1$; bajo esta aplicación Ω se transforma en una nueva región, y las imágenes de $E_1, \ldots, E_{n-1}$ son las componentes acotadas de su complemento. Por motivos de sencillez convenimos en utilizar las mismas notaciones que antes de la aplicación; en particular, E_n es ahora el conjunto $|z| \geqq 1$. La circunferencia unidad $|z| = 1$, trazada en sentido positivo, se denotará por C_n, y se le llama el *contorno exterior* de la nueva región Ω.

Consideremos ahora el complemento de E_1 con respecto al plano ampliado. Esta es nuevamente una región simplemente conexa, que aplicamos sobre el *exterior* del círculo unidad, correspondiéndose ∞ consigo mismo. La imagen de C_n es una curva analítica, cerrada y orientada, que seguiremos llamando C_n, lo mismo que hemos mantenido las otras notaciones. Definimos además el *contorno interior* C_1 como la circunferencia unidad en el nuevo plano, trazada en sentido negativo.

Evidentemente, el proceso puede repetirse hasta conseguir una región Ω acotada por un contorno exterior C_n y $n-1$ contornos interiores $C_1, \ldots, C_{n-1}$ (Fig. 5-2). Es importante observar que el índice de un contorno con respecto a un punto arbitrario del plano puede calcularse rápidamente. Así, p. ej., en la etapa en la que C_k, $k < n$, es la circunferencia unidad, el índice de C_k es -1 con relación a los puntos interiores de E_k y 0 respecto a los restantes puntos que no estén en C_k. Las aplicaciones siguientes no cambiarán este estado de cosas. El hecho es evidente, pudiendo darse fácilmente una demostración formal basada en el principio del argumento. Se demuestra de la misma manera que el contorno exterior C_n tiene índice 0 respecto a los puntos interiores de E_n e índice 1 con relación a todos los demás puntos no pertenecientes a C_n. Se sigue que el ciclo $C = C_1 + C_2 + \ldots + C_n$ acota Ω en el sentido del capítulo III, sección 3-5, 1; definición 4. La distinción entre contornos exterior e interior

es coincidente, pues evidentemente una inversión con respecto a un punto interior de E_k convertirá a C_k en el contorno exterior.

El teorema 12 es claramente de aplicación a Ω. En efecto, es completamente obvia la existencia de una barrera, puesto que cualquier contorno puede transformarse en una circunferencia.

Supongamos ahora que hemos resuelto el problema de Dirichlet en Ω con valores frontera 1 sobre C_k y 0 en los demás contornos. Se

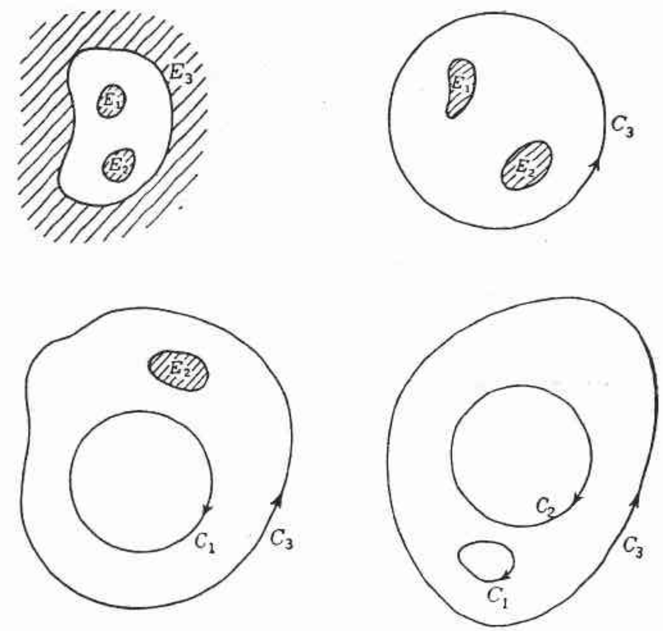

Fig. 5-2.

denota la solución por $\omega_k(z)$, y se le llama *medida armónica* de C_k respecto a la región Ω. Evidentemente, tenemos que $0 < \omega_k(z) < 1$ en Ω y

$$\omega_1(z) + \omega_2(z) + \ldots + \omega_n(z) \equiv 1.$$

Si aplicamos Ω de forma que C_i se convierta en una circunferencia, entonces ω_k puede prolongarse a través de C_i de acuerdo con el principio de simetría. Se obtiene que ω_k es armónica en la región cerrada Ω, en el sentido de que puede prolongarse a una región mayor.

Los contornos $C_1, \ldots, C_{n-1}$ forman una base de homología para los ciclos en Ω, homología que se ha de entender con respecto a una región mayor no especificada. La función armónica conjugada de ω_k es multiforme con período

$$\alpha_{kj} = \int_{C_j} \frac{\partial \omega_k}{\partial n}\, ds = \int_{C_j} {*d\omega_k}$$

a lo largo de C_j. Con mayor generalidad, cabe afirmar que ninguna combinación lineal $\lambda_1\omega_1(z)+\lambda_2\omega_2(z)+\ldots+\lambda_{n-1}\omega_{n-1}(z)$ puede tener una función conjugada uniforme, a no ser que todos los λ_i sean cero. Para verlo, supongamos que esta expresión fuera la parte real de una función analítica uniforme $f(z)$. Por el principio de simetría, $f(z)$ debería tener una prolongación analítica en el cierre de Ω. La parte real de $f(z)$ sería constantemente igual a λ_i sobre C_i, $i=1, \ldots, n-1$, y cero sobre C_n. En consecuencia, cada contorno estaría aplicado sobre un segmento vertical. Si w_0 no pertenece a ninguno de estos segmentos, puede definirse una rama uniforme de arg $[f(z)-w_0]$ sobre cada contorno. Se sigue por el principio del argumento que $f(z)$ no puede tomar al valor w_0 en Ω. Pero entonces $f(z)$ debe reducirse a una constante, puesto que en otro caso la imagen de Ω contendría ciertamente puntos exteriores a los segmentos. En definitiva, todos los λ_i han de ser cero.

El resultado indica que el sistema homogéneo de ecuaciones lineales

$$\lambda_1\alpha_{1j}+\lambda_2\alpha_{2j}+\ldots+\lambda_{n-1}\alpha_{n-1,j}=0 \quad (j=1, \ldots, n-1) \qquad [5\text{-}29]$$

tiene únicamente la solución trivial $\lambda_i=0$. Puesto que $\alpha_{k1}+\alpha_{k2}+\ldots+\alpha_{kn}=0$, la ecuación correspondiente a $j=n$ es consecuencia de [5-29]. Por la teoría de las ecuaciones lineales, cualquier sistema de ecuaciones no homogéneo con los mismos coeficientes que los de [5-29] tiene entonces que tener una solución. En particular, es posible resolver el sistema

$$\begin{aligned}
&\lambda_1\alpha_{11}+\lambda_2\alpha_{21}+\ldots+\lambda_{n-1}\alpha_{n-1,1}=2\pi\\
&\lambda_1\alpha_{12}+\lambda_2\alpha_{22}+\ldots+\lambda_{n-1}\alpha_{n-1,2}=0\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad [5\text{-}30]\\
&\lambda_1\alpha_{1,n-1}+\lambda_2\alpha_{2,n-1}+\ldots+\lambda_{n-1}\alpha_{n-1,n-1}=0\\
&\lambda_1\alpha_{1n}+\lambda_2\alpha_{2n}+\ldots+\lambda_{n-1}\alpha_{n-1,n}=-2\pi,
\end{aligned}$$

donde la última ecuación es consecuencia de las $n-1$ primeras. En otras palabras: podemos hallar una integral multiforme $f(z)$ con

períodos $\pm 2\pi i$ a lo largo de C_1 y de C_n, y todos los demás períodos iguales a cero, cuya parte real es constantemente igual a λ_k sobre C_k (ponemos $\lambda_n=0$). La función $F(z)=e^{f(z)}$ es uniforme. Probaremos el siguiente

TEOREMA 13. *La función $F(z)$ efectúa una aplicación biunívoca conforme de Ω sobre la corona circular $1<|w|<e^{\lambda_1}$ menos $n-2$ arcos concéntricos situados sobre las circunferencias $|w|=e^{\lambda_i}$, $i=2, \ldots, n-1$.*

La aplicación está ilustrada en la figura 5-3. Los contornos C_1 y C_n están en correspondencia biunívoca con las circunferencias completas, mientras que los otros contornos están *aplastados* en hendiduras circulares. Deberá imaginarse cada hendidura con dos bordes, que junto con los extremos forman un contorno cerrado.

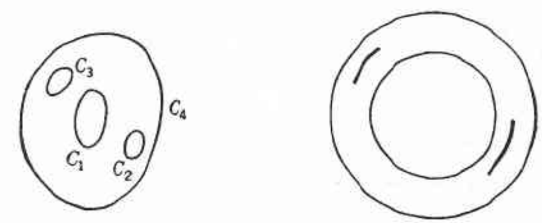

FIG. 5-3.

La demostración utiliza el principio del argumento. Sabemos que $F(z)$ es analítica y que su módulo es constante sobre cada contorno. El número de raíces de la ecuación $F(z)=w_0$ viene dado por

$$\frac{1}{2\pi i}\int_{C_1}\frac{F'(z)\,dz}{F(z)-w_0}+\frac{1}{2\pi i}\int_{C_2}\frac{F'(z)\,dz}{F(z)-w_0}+\ldots+ \int_{C_n}\frac{1}{2\pi i}\,\frac{F'(z)dz}{F(z)-w_0}, \qquad [5\text{-}31]$$

en cualquier caso si w_0 no se toma sobre la frontera. Para $w_0=0$ se conocen los términos de [5-31], siendo sus valores $1, 0, \ldots, 0, -1$, respectivamente. La integral sobre C_1 permanece constantemente igual a 1 para $|w_0|<e^{\lambda_1}$ y se anula para $|w_0|>e^{\lambda_1}$; análogamente, la última integral vale -1 para $|w_0|<1$ y 0 para $|w_0|>1$. Las integrales sobre C_k, $1<k<n$, se anulan para todo w_0 tal que $|w_0|\neq e^{\lambda_k}$. Supongamos ahora que $F(z)$ toma realmente el valor w_0; siempre y cuando

se aplique Ω sobre un conjunto abierto, podemos elegir $|w_0|$ distinto de todos los e^{λ_i}. Para este w_0 la expresión [5-31] tiene que ser positiva. Pero esto es solo posible si $1 < |w_0| < e^{\lambda_1}$. Así, pues, $\lambda_1 > 0$, y por la continuidad, $0 \leqq \lambda_i \leqq \lambda_1$.

A partir de aquí, la demostración puede completarse mediante un razonamiento puramente topológico. Sin embargo, es más instructivo, y de hecho más sencillo, obtener la conclusión a partir del principio del argumento. Cuando existen polos simples en la frontera sigue verificándose el teorema de los residuos, con tal que la integral de contorno se reemplace por su valor principal de Cauchy y la suma de los residuos incluya la mitad de cada residuo sobre la frontera[1]. En la situación presente el segundo convenio significa que un valor tomado sobre la frontera se cuenta con la mitad de su multiplicidad. El cómputo de los valores principales no ofrece dificultad. Si $|w_0| = e^{\lambda_k}$, hallamos que

$$\text{v. pr.} \int_{C_k} \frac{F'(z)\,dz}{F(z) - w_0} = \frac{1}{2} \int_{C_k} \frac{F'(z)\,dz}{F(z)},$$

por geometría elemental (o cálculo directo)

$$d \arg\,[F(z) - w_0] = \tfrac{1}{2}\, d \arg F(z).$$

Por consiguiente, los valores principales en [5-30] son $\frac{1}{2}$ para $k = 1$, 0 para $2 \leqq k \leqq n-1$, $-\frac{1}{2}$ para $k = n$.

Obtenemos en conclusión que cada valor se toma media vez sobre las circunferencias $|w_0| = 1$ o $|w_0| = e^{\lambda_1}$; es decir, una vez sobre la frontera. Esto prueba que C_1 y C_n se aplican biunívocamente una sobre otra y que $0 < \lambda_i < \lambda_1$, $i \neq 1, n$. A continuación se sigue que si $1 < |w_0| < e^{\lambda_1}$, se toma w_0 o una vez en el interior, dos veces en la frontera, o una vez en la frontera con multiplicidad 2. Sobre cada contorno $C_2, \ldots, C_{n-1}$ se puede definir una rama uniforme de $\arg F(z)$, correspondiendo los valores de multiplicidad 2 a los máximos y mínimos relativos de $\arg F(z)$. Existe al menos un máximo y un mínimo, no pudiendo haber más, pues en caso contrario $F(z)$ pasaría más de dos veces por los mismos valores. Además, la diferencia entre el máximo y el mínimo debe ser $< 2\pi$, lo que muestra que cada contorno se aplica sobre un arco propio. Por último, los arcos correspondientes a contornos diferentes deben ser disjuntos.

Hemos probado por completo el teorema 13, habiendo sido capaces de describir la correspondencia entre las fronteras. Lo más importante del teorema es el hecho de que podamos aplicar Ω sobre una región canónica acotada por dos circunferencia y $n-2$ hendiduras circulares concéntricas; por motivos de normalización, el radio del círculo interior se elige igual a 1. Para una elección prescrita de C_1 y C_n, la aplicación canónica está determinada de manera única, salvo una rotación. Esto se sigue del hecho de que el sistema [5-30] tiene una sola solución.

La forma de una región canónica de orden de conexión n depende de $3n-6$ constantes reales. En efecto, la posición y el tamaño de cada hendidura está determinado por tres números, en total $3n-6$; la anchura de la corona circular da otro parámetro adicional, pero hay que descontar un parámetro para permitir la rotación arbitraria.

EJERCICIOS

1. Pruébese directamente que dos coronas circulares son conformemente equivalentes si (y solo si) las razones de sus radios son iguales.
2. Demuéstrese que $\alpha_{ij} = \alpha_{ji}$. *Sugerencia:* Aplíquese el teorema 1.

2. *Función de Green.*—Suponemos nuevamente que Ω es una región de orden de conexión finito, y siempre y cuando sean permisibles aplicaciones conformes previas, podemos suponer que Ω está acotada por contornos analíticos $C_1, \ldots, C_n$; esta vez estará incluido el caso correspondiente a $n = 1$.

Consideremos un punto $z_0 \in \Omega$ y resolvamos el problema de Dirichlet en Ω con valores frontera iguales a $\log |\zeta - z_0|$. Se denota la solución por $G(z)$, pero el interés principal corresponde a la función $g(z) = G(z) - \log |z - z_0|$, conocida como la *función de Green* de Ω con polo en z_0. Cuando se quiere hacer hincapié en su dependencia de z_0, se denota por $g(z, z_0)$.

La función de Green es armónica en Ω salvo en z_0 y se anula sobre la frontera. En un entorno de z_0 difiere de $-\log |z - z_0|$ en una función armónica; $g(z)$ está determinada de manera única por estas propiedades. En efecto, si $g_1(z)$ tiene las mismas propiedades,

[1] En el Cap. III, Sec. 3-5, 3, se introdujo el valor principal de Cauchy en el caso de una integral sobre una recta. En el caso de un arco analítico cualquiera, es más sencillo definir el valor principal mediante una aplicación conforme auxiliar, que transforma un arco parcial en un segmento. La generalización enunciada del teorema de los residuos se deduce con toda facilidad y prueba que el valor principal es independiente de la aplicación conforme auxiliar.

entonces $g-g_1$ es armónica en todo Ω y se anula sobre la frontera. Se sigue del principio del máximo que g_1 es idénticamente igual a g.

Si dos regiones son conformemente equivalentes, entonces las funciones de Green con polos correspondientes son iguales en los puntos homólogos. Para ser más explícitos, sea $z=z(\zeta)$ una función que define una aplicación biunívoca conforme de una región Ω' en el plano ζ sobre una región Ω en el plano z. Elijamos un punto $\zeta_0 \in \Omega'$ y denotemos por $g(z, z_0)$ la función de Green de Ω con polo en $z_0 = z(\zeta_0)$. Afirmamos que $g[z(\zeta), z_0]$ es la función de Green de Ω'. En primer lugar, si ζ tiende a un punto frontera, entonces $z(\zeta)$ se aproxima a la frontera de Ω y, por tanto, $g[z(\zeta), z_0]$ tiene valores frontera cero. En cuanto al comportamiento en ζ_0, sabemos que $g[z(\zeta), z_0]$ difiere de $-\log |z(\zeta)-z(\zeta_0)|$ en una función armónica de $z(\zeta)$ y, consiguientemente, en una función armónica de ζ. Pero la diferencia $\log |z(\zeta)-z(\zeta_0)| - \log |\zeta-\zeta_0|$ es también armónica, y se sigue que $g[z(\zeta), z_0]$ se comporta adecuadamente en ζ_0. Hemos probado que la función de Green es *invariante* respecto a las aplicaciones conformes, y como consecuencia de esta invariancia pueden efectuarse a capricho las aplicaciones conformes previas.

En el caso de una región simplemente conexa existe una relación sencilla entre la función de Green y la función de Riemann del teorema fundamental. Para el disco unidad $|w| < 1$ la función de Green con respecto al origen es evidentemente $-\log |w|$. Por tanto, si $w = f(z)$ aplica Ω sobre el disco unidad llevando z_0 al origen, se tiene por la invariancia que

$$g(z, z_0) = -\log |f(z)|.$$

Recíprocamente, si $g(z, z_0)$ es conocida, se puede determinar la función de Riemann.

La función de Green tiene una importante propiedad de simetría. Dados dos puntos $z_1, z_2 \in \Omega$, escribiremos brevemente $g(z, z_1) = g_1$, $g(z, z_2) = g_2$. Como consecuencia del teorema 1, la diferencial $g_1 *dg_2 - g_2 *dg_1$ es exacta localmente en la región obtenida al omitir los puntos z_1 y z_2 de Ω. Si c_1 y c_2 son pequeñas circunferencias de centros z_1 y z_2, respectivamente, descritas en sentido positivo, entonces el ciclo $C - c_1 - c_2$ es homológico a cero (como antes, $C = C_1 + \ldots + C_n$). Puesto que g_1 y g_2 se anulan sobre C, se concluye que

$$\int_{c_1+c_2} g_1 *dg_2 - g_2 *dg_1 = 0.$$

Poniendo $G_1 = g_1 + \log |z-z_1|$, tenemos que $*dg_1 = *dG_1 - d \arg (z-z_1)$ y hallamos que

$$\int_{c_1} g_1 *dg_2 - g_2 *dg_1 = \int_{c_1} G_1 *dg_2 - g_2 *dG_1 - \int_{c_1} \log |z-z_1| *dg_2 + \\ + \int_{c_1} g_2 d \arg (z-z_1).$$

En el segundo miembro, la primera integral se anula por el teorema 1; la segunda, como consecuencia de ser $|z-z_1|$ constante sobre c_1 y ser $*dg_2$ una diferencial exacta en un entorno de z_1. La última integral, de acuerdo con la propiedad del valor medio para funciones armónicas, es igual a $2\pi g_2(z_1)$. De manera simétrica, la integral sobre c_2 debe ser igual a $-2\pi g_1(z_2)$, con lo que queda demostrado que $g_2(z_1) - g_1(z_2) = 0$ o que

$$g(z_1, z_2) = g(z_2, z_1).$$

Como consecuencia de esta propiedad de simetría, la función de Green $g(z, z_0)$ también es armónica en la segunda variable.

La función conjugada de $g(z, z_0)$, que denotaremos por $h(z, z_0)$, es, naturalmente, multiforme. En primer lugar, tiene el período 2π a lo largo de una pequeña circunferencia c con centro en z_0. Además, tiene los períodos

$$P_k(z_0) = \int_{C_k} dh\,(z, z_0) = \int_{C_k} *dg\,(z, z_0) \qquad (k=1, \ldots, n).$$

Probaremos que el período $P_k(z_0)$ es igual a la medida armónica $\omega_k(z_0)$ multiplicada por 2π.

La demostración es nuevamente una aplicación del teorema 1. Expresemos el hecho de que la integral de $\omega_k *dg - g *d\omega_k$ sobre $C - c$ tiene que anularse. La integral sobre C se reduce a $P_k(z_0)$, y mediante el mismo cálculo utilizado anteriormente obtenemos que la integral sobre c es igual a $2\pi\omega_k(z_0)$. Queda, por tanto, probado que $P_k(z_0) = 2\pi\omega_k(z_0)$.

3. *Regiones paralelas en hendidura.*—Escribiendo $z_0 = x_0 + iy_0$ derivemos la identidad

$$g(z, z_0) = G(z, z_0) - \log |z - z_0|$$

con respecto a x_0, con lo que obtenemos

$$\frac{\partial}{\partial x_0} g(z, z_0) = \frac{\partial}{\partial x_0} G(z, z_0) + \operatorname{Re} \frac{1}{z - z_0}$$

para todo z y z_0 distintos. Es evidente que $(\partial/\partial x_0)G(z, z_0)$ es armónica con respecto a z; en efecto, podemos representar la derivada parcial como el límite uniforme sobre conjuntos compactos de las funciones armónicas $[G(z, z_0+h) - G(z, z_0)]/h$ para h real tendiendo a 0. Además, puesto que $g(z, z_0)$ es idénticamente cero cuando z pertenece a la frontera, la derivada parcial $(\partial/\partial x_0)g(z, z_0)$ también debe anularse sobre la frontera. La función $u_1(z) = (\partial/\partial x_0)g(z, z_0)$ es, por tanto, cero sobre la frontera y difiere de $\operatorname{Re} 1/(z - z_0)$ en una función armónica.

La función conjugada de $u_1(z)$ tiene ciertos períodos A_k a lo largo de los contornos C_k; pero es fácil construir una combinación lineal de $u_1(z)$ y de las medidas armónicas $\omega_i(z)$, cuya función conjugada carece de períodos. En efecto, la función $u_1 + \lambda_1\omega_1 + \ldots + \lambda_{n-1}\omega_{n-1}$ tiene esta propiedad con tal que

$$\lambda_1\alpha_{1k} + \lambda_2\alpha_{2k} + \ldots + \lambda_{n-1}\alpha_{n-1,k} = -A_k \qquad (k = 1, \ldots, n-1).$$

Sabemos ya que este sistema no homogéneo de ecuaciones tiene siempre una solución. Hemos establecido así la existencia de una función $p(z)$, que es uniforme y analítica en Ω, excepto por un polo simple con residuo 1 en z_0, siendo constante sobre cada contorno su parte real; $p(z)$ está determinada de manera única por estos requisitos, salvo una constante aditiva.

Se obtiene un resultado similar derivando con respecto a y_0. De

$$\frac{\partial}{\partial y_0} g(z, z_0) = \frac{\partial}{\partial y_0} G(z, z_0) - \operatorname{Im} \frac{1}{z - z_0}$$

obtenemos que $v_2(z) = -(\partial/\partial y_0)g(z, z_0)$ se anula sobre la frontera y tiene la misma singularidad que $\operatorname{Im} 1/(z - z_0)$. Si se añade una combinación lineal adecuada de medidas armónicas, la función conjugada se hace uniforme. Por tanto, existe una función analítica uniforme $q(z)$ con parte singular $1/(z - z_0)$ cuya parte imaginaria es constante sobre cada contorno.

Las funciones $p(z)$ y $q(z)$ conducen a aplicaciones canónicas sencillas.

TEOREMA 14. *Las aplicaciones determinadas por $p(z)$ y $q(z)$ son biunívocas, y la imagen de Ω es una región en hendidura cuyo complemento consiste en n segmentos verticales u horizontales, respectivamente* [Fig. 5-4, (*a*) y (*b*)].

(*a*) (*b*)

FIG. 5-4.

La demostración es bastante parecida a la del teorema 13. Esta vez la expresión

$$\sum_{k=1}^{n} \frac{1}{2\pi i} \int_{C_k} \frac{p'(z)\,dz}{p(z) - w_0} \qquad [5\text{-}32]$$

representa el número de ceros de $p(z) - w_0$ menos el número de polos. Pero es fácil ver que [5-32] se anula para todo w_0, incluyendo valores frontera. En el último caso debe formarse el valor principal; pero si se toma w_0 sobre C_k, la parte imaginaria de $p'\,dz/(p - w_0)$ se anula a lo largo de C_k, con lo que no existe ninguna dificultad. Puesto que hay exactamente un polo, se concluye que $p(z)$ toma todo valor una vez en el interior de Ω, dos veces sobre la frontera o una vez sobre la frontera con multiplicidad 2. El resto de la demostración es un duplicado exacto del razonamiento anterior. La demostración sigue siendo válida sin cambio alguno para $q(z)$.

Se pueden considerar las regiones paralelas en hendidura como regiones canónicas, pero no son conformemente equivalentes, aun en el caso de que se exija que el punto del infinito se corresponda a sí mismo. Así, p. ej., las aplicaciones mediante $p(z)$ e $iq(z)$ conducen a regiones en hendidura verticales que son diferentes, pero conformemente equivalentes. Las aplicaciones en hendidura están determinadas de manera única, salvo una traslación paralela, solo para aplicaciones con el mismo residuo.

EJERCICIOS

1. Pruébese que una región de orden de conexión finito puede aplicarse sobre un disco circular con hendiduras circulares concéntricas; el centro corresponde a un punto prescrito y las circunferencias a un contorno preasignado.

2. Demuéstrese que la función $e^{i\alpha}(p \cos \alpha - iq \operatorname{sen} \alpha)$ aplica Ω sobre una región limitada por hendiduras oblicuas.

3. Utilizando el ejercicio 2, pruébese que $p+q$ aplica Ω de manera biunívoca sobre una región de contorno convexo.

CAPITULO VI

FUNCIONES MULTIFORMES

6-1. Prolongación analítica.—En los capítulos precedentes hemos hecho hincapié en que todas las funciones deben estar bien definidas y ser, por tanto, uniformes. En el caso de funciones como $\log z$ o $\sqrt{z}$, que no están determinadas de manera única por su expresión analítica, era necesario un esfuerzo especial con objeto de demostrar que, en circunstancias favorables, se podía elegir una rama uniforme. Mientras este punto de vista responde a la necesidad de claridad lógica, no hace justicia al hecho de que la ambigüedad en los casos del logaritmo y de la raíz cuadrada es una característica esencial que no puede ignorarse. Existe, pues, una necesidad evidente de una teoría rigurosa de las funciones multiformes.

Seguiremos aceptando como noción primaria el de función analítica uniforme definida en una región, en términos de la cual se definirán las funciones analíticas multiformes.

1. *Funciones analíticas generales.*—Una función analítica $f(z)$ definida en una región Ω constituirá un *elemento de función,* al que denotaremos por (f, Ω), y una *función analítica general* aparecerá como una colección de elementos de función relacionados unos con otros de una forma prescrita.

Se dice que dos elementos de función (f_1, Ω_1) y (f_2, Ω_2) son *prolongación analítica directa* uno de otro si $\Omega_1\Omega_2$ no es vacío y $f_1(z) = f_2(z)$ en $\Omega_1\Omega_2$. Más específicamente, a (f_2, Ω_2) se le llama prolongación analítica directa de (f_1, Ω_1) en la región Ω_2. No tiene por qué existir una prolongación analítica directa en Ω_2; pero si existe, está determinada de manera única. En efecto, supongamos que (f_2, Ω_2) y (g_2, Ω_2) son dos prolongaciones analíticas directas de (f_1, Ω_1); entonces $f_2 = g_2$ en $\Omega_1\Omega_2$, y esto implica que $f_2 = g_2$ en todo Ω_2. Observemos que si $\Omega_2 \subset \Omega_1$, entonces la prolongación analítica directa de (f_1, Ω_1) es (f_1, Ω_2).

Si (f_1, Ω_1) y (f_2, Ω_2) son prolongaciones analíticas directas una de otra, es evidente que se puede definir una función analítica f en $\Omega_1 + \Omega_2$ haciendo $f = f_1$ en Ω_1 y $f = f_2$ en Ω_2. Puesto que desde el principio se pudo considerar el elemento de función $(f, \Omega_1+\Omega_2)$, parece que no se ha ganado nada. Consideremos, sin embargo, un tercer elemento de función (f_3, Ω_3), el cual suponemos que es una prolongación analítica directa de (f_2, Ω_2). Entonces pudiera muy bien ocurrir que Ω_3 tenga parte común con Ω_1; pero que, sin embargo, (f_3, Ω_3) no sea una prolongación analítica directa de (f_1, Ω_1). En esta situación la colección (f_1, Ω_1), (f_2, Ω_2), (f_3, Ω_3) no puede reemplazarse por un solo elemento de función, pero proporciona una definición satisfactoria de función multiforme.

Con mayor generalidad, se nos ofrece la consideración de cadenas de elementos de función (f_1, Ω_1), (f_2, Ω_2), ..., (f_n, Ω_n), tales que (f_k, Ω_k) es una prolongación analítica directa de (f_{k-1}, Ω_{k-1}). Se dice que los elementos de una cadena como la anterior son *prolongaciones analíticas* uno de otro. Adoptaremos la definición siguiente:

DEFINICIÓN 1. *Una función analítica general es una colección no vacía* **f** *de elementos de función (f, Ω), tal que dos elementos cualesquiera de* **f** *son prolongaciones analíticas uno de otro a través de una cadena cuyos elementos son miembros de* **f**.

Una función analítica completa es una función analítica general que contiene todas las prolongaciones analíticas de cualquiera de sus elementos.

Una función analítica completa es evidentemente *maximal* en el sentido que no puede prolongarse más, y está claro que todo elemento de función pertenece a una sola función analítica completa. Las funciones analíticas generales incompletas son más arbitrarias, y existen muchos casos en los que dos colecciones diferentes de elementos de función deberían considerarse como definiendo la misma función. Así, p. ej., una función uniforme $f(z)$, definida en Ω, puede identificarse o con la colección que consiste en el único elemento de función (f, Ω), o con la colección de todos los (f, Ω') con $\Omega' \subset \Omega$.

Una función analítica general **f** posee una derivada **f**′ determinada de manera única, definida por los elementos de función (f', Ω). En efecto, si (f_1, Ω_1) y (f_2, Ω_2) son prolongaciones analíticas directas uno de otro, también lo son (f'_1, Ω_1) y (f'_2, Ω_2). Se pueden definir las derivadas de órdenes superiores **f**″, **f**‴, ... de la misma manera.

Una relación análoga pudiera existir entre un par cualquiera de funciones analíticas generales **f** y **g**. Supongamos que se da una correspondencia que a todo $(f, \Omega) \in$ **f** asigna un elemento de función único $(g, \Omega) \in$ **g**, de tal forma que a prolongaciones analíticas directas corresponden prolongaciones analíticas directas. En estas circunstancias convenimos en decir que **f** está *subordinada* a **g**, y es posible definir **f** + **g** y **fg** como colecciones que consisten en los elementos $(f+g, \Omega)$, (fg, Ω), que corresponden a los elementos (f, Ω) de **f**. Así, p. ej., **f** está subordinada a cualquier función entera **h**, de donde se sigue que **f** + **h** y **fh** están bien definidas[1].

Podemos ahora formular un principio clásico conocido como el de *permanencia de relaciones funcionales.* Supongamos que se dan ciertas funciones analíticas generales **f**, **g**, ..., y que **f**, p. ej., está subordinada a todas las demás. Supongamos además que se sabe que un conjunto de elementos de función correspondientes (f, Ω), (g, Ω), ... satisfacen una relación de la forma $G(f, g, \ldots)=0$, siendo la expresión G un polinomio de varias variables (la demostración es válida con mayor generalidad). Si (f_1, Ω_1), (g_1, Ω_1), ... es un conjunto de prolongaciones analíticas directas, se deduce inmediatamente que $G(f_1, g_1, \ldots)=0$ en Ω_1, por la sencilla razón de que la función compuesta $G[f_1(z), g_1(z), \ldots]$ es analítica en Ω_1 y se anula en $\Omega\Omega_1$. Podemos, por tanto, llegar a la conclusión de que se verifica la relación $G(f, g, \ldots)=0$ para *todos* los conjuntos de elementos de función correspondientes, hecho que puede también expresarse mediante la ecuación $G(\mathbf{f}, \mathbf{g}, \ldots)=0$.

EJERCICIOS

1. Pruébese que un elemento de función (f, C), donde C es todo el plano, determina una función analítica completa consistente en todos los elementos de función de la forma (f, Ω).

2. Defínase $\sqrt{z}$ como una función analítica general mediante un número finito de elementos de función.

3. Supongamos que (f, Ω) satisface una ecuación diferencial de la forma $P(f, f', f'', \ldots)=0$, siendo P un polinomio cuyos coeficientes son funciones enteras. Pruébese que todos los elementos de función de la función analítica completa determinada por (f, Ω) satisfacen la misma ecuación diferencial.

2. *La superficie de Riemann de una función.*—Con objeto de estudiar la naturaleza multiforme de una función analítica general es conveniente introducir el concepto de *rama.* Se dice que dos elementos de función (f_1, Ω_1) y (f_2, Ω_2) determinan la misma rama en un

[1] Utilizamos solo de momento esta noción de subordinación, que está relacionada, pero no es idéntica, a la que comúnmente se utiliza.

punto $z_0 \in \Omega_1\Omega_2$ siempre y cuando $f_1 = f_2$ en un entorno de z_0. Para que esto ocurra es suficiente (pero no necesario) que los elementos de función sean prolongaciones analíticas directas uno de otro. Son siempre, sin embargo, prolongaciones analíticas uno de otro, pues ambos son prolongaciones analíticas directas de su restricción común a un entorno de z_0. Observemos que dos elementos de función determinan la misma rama en z_0 si (y solo si) tienen el mismo desarrollo de Taylor en un entorno de z_0.

La relación entre elementos de función que acabamos de introducir es evidentemente una relación de equivalencia. Con respecto a esta relación de equivalencia, todos los elementos de función (f, Ω), para los que $z_0 \in \Omega$, se agrupan en clases de equivalencia bien definidas a las que llamaremos *ramas analíticas* en z_0. Es fácil ver que pueden ser identificadas con todas las series de potencias en $z - z_0$ con un radio de convergencia positivo. Denotaremos la rama en z_0 determinada por el elemento de función (f, Ω) mediante (f, z_0).

Para una función analítica general $\mathbf{f}$ seleccionaremos las ramas (f, z_0) determinadas por elementos de función $(f, \Omega) \in \mathbf{f}$ y las llamaremos las ramas de $\mathbf{f}$ en z_0. A cada rama de estas corresponde un único valor funcional $f(z_0)$, así como valores únicos de las derivadas $f'(z_0)$, $f''(z_0)$, ... De manera análoga a la utilizada en la construcción de superficies de Riemann de las funciones multiformes elementales, introduciremos un conjunto $\mathfrak{F}$ (la superficie de Riemann) cuyos elementos $\mathfrak{z}$ (los puntos) son las ramas (f, z) de $\mathbf{f}$. Estamos entonces en situación de considerar $\mathbf{f}$ como una función uniforme $\mathbf{f}(\mathfrak{z})$ sobre $\mathfrak{F}$. La función $z = p(\mathfrak{z})$, que a todo $\mathfrak{z} = (f, z)$ le asigna el valor z determinado de manera única, se denomina la *proyección* de $\mathfrak{F}$ en el plano complejo, siendo z la *traza* de $\mathfrak{z}$.

La consideración de la superficie de Riemann no es de mucha utilidad, a no ser que podamos decir cuándo una función es continua sobre $\mathfrak{F}$. Puesto que se puede expresar la continuidad en términos de entornos, basta con definir los entornos en $\mathfrak{F}$. Dado $\mathfrak{z}_0 = (f_0, z_0)$, determinado por el elemento de función $(f_0, \Omega_0) \in \mathbf{f}$, elijamos un entorno $V \subset \Omega_0$ de z_0 y consideremos el conjunto $\mathfrak{B}$ de todas las ramas (f_0, z) con $z \in V$. Por definición, $\mathfrak{B}$ será un entorno de $\mathfrak{z}_0$. Se deduce con facilidad que $\mathbf{f}(\mathfrak{z})$ y $p(\mathfrak{z})$ son funciones continuas en el sentido de que existe, para un $\epsilon > 0$ cualquiera dado, un entorno $\mathfrak{B}$ con la propiedad de que $|\mathbf{f}(\mathfrak{z}) - \mathbf{f}(\mathfrak{z}_0)| < \epsilon$, $|p(\mathfrak{z}) - p(\mathfrak{z}_0)| < \epsilon$ para todo $\mathfrak{z} \in \mathfrak{B}$.

Mediante la introducción de las superficies de Riemann obtenemos una interpretación muy sencilla de la subordinación. Sean $\mathfrak{F}$ y $\mathfrak{G}$ las superficies de Riemann de $\mathbf{f}$ y $\mathbf{g}$, respectivamente. Entonces $\mathbf{f}$ está subordinada a $\mathbf{g}$ si (y solo si) existe una aplicación continua σ de $\mathfrak{F}$ en $\mathfrak{G}$ tal que $\mathfrak{z}$ y $\sigma(\mathfrak{z})$ tienen la misma proyección z; la demostración es inmediata. Observemos que la aplicación σ no es necesariamente única, lo que significa que $\mathbf{f}$ puede estar subordinada a $\mathbf{g}$ de diferentes formas. En un lenguaje muy sugerente para la imaginación, la existencia de una proyección que conserva las aplicaciones implica que la superficie $\mathfrak{F}$ puede extenderse sobre la $\mathfrak{G}$, o que cabe considerar a $\mathfrak{F}$ como una superficie de Riemann relativamente a $\mathfrak{G}$.

Observemos por último que nuestra definición de superficie de Riemann es provisional, puesto que no incluye todavía el caso de puntos de ramificación.

3. *Prolongación analítica a lo largo de arcos.*—Consideremos una función analítica general $\mathbf{f}$ con superficie de Riemann $\mathfrak{F}$ y un arco γ en el plano complejo, de ecuación $z = z(t)$, $\alpha \leqq t \leqq \beta$. Supongamos que existe en $\mathfrak{F}$ un arco $\bar{\gamma}$ cuya proyección es γ: con esto queremos decir que $\bar{\gamma}$ tiene una ecuación $\mathfrak{z} = \mathfrak{z}(t)$ con $p[\mathfrak{z}(t)] = z(t)$ para todo t. La hipótesis fundamental de que $\bar{\gamma}$ es un arco significa, naturalmente, que $\mathfrak{z}(t)$ es continua con respecto a los entornos introducidos en $\mathfrak{F}$.

Es deseable dar una interpretación paralela que no se refiera de manera explícita a la superficie de Riemann. A cada t corresponde un $\mathfrak{z}$ cuya proyección es $z(t)$ y, por tanto, una rama de la forma $[f, z(t)]$. Para un t_0 dado esta rama está determinada por un elemento de función (f_0, Ω_0) con $z(t_0) \in \Omega_0$. Un entorno consiste en ramas (f_0, z), y la continuidad de $z(t)$ implica evidentemente la existencia de un $\delta > 0$ tal que para $|t - t_0| < \delta$ la rama $[f, z(t)]$ está determinada por el elemento de función (f_0, Ω_0). Cuando es este el caso, decimos que la rama $[f, z(t)]$ y cualquiera de los elementos de función correspondientes han sido obtenidos mediante *prolongación a lo largo del arco* γ. De acuerdo con esta terminología, existe una equivalencia completa entre prolongaciones a lo largo de γ y arcos $\bar{\gamma}$ sobre $\mathfrak{F}$ que se proyectan en γ.

La prolongación a lo largo de un arco corresponde al concepto intuitivo de rama variable con continuidad. No está garantizada la existencia de una prolongación, pero es válido el siguiente importante teorema de unicidad:

TEOREMA 1. *Dos prolongaciones* $[f_1, z(t)]$ *y* $[f_2, z(t)]$ *de una función analítica general* **f** *a lo largo del mismo arco* γ *son o idénticas o diferentes para todo valor de* t.

Consideremos el subconjunto E del intervalo cerrado (α, β), en el que

$$[f_1, z(t)] = [f_2, z(t)].$$

Elijamos $t_0 \in E$ y supongamos que las ramas correspondientes están determinadas por elementos de función (f_1^0, Ω_1), (f_2^0, Ω_2). Por hipótesis, $f_1^0 = f_2^0$ en un entorno de $z(t_0)$. Si t está suficientemente próximo a t_0, el punto $z(t)$ pertenece a este entorno; además podemos elegir $f_1 = f_1^0$, $f_2 = f_2^0$, y se sigue que las ramas $[f_1, z(t)]$, $[f_2, z(t)]$ son idénticas. Este resultado muestra que el complemento de E es cerrado. Supongamos ahora que t_0 no pertenece a E. Con las mismas notaciones, $f_1^0(z)$ y $f_2^0(z)$ no son idénticas en ningún entorno de $z(t_0)$. Por tanto, existe un entorno Δ de $z(t_0)$ en el que $f_1^0(z) \neq f_2^0(z)$, excepto quizá para $z = z(t_0)$. Para t suficientemente próximo a t_0, $z(t) \in \Delta$, y podemos tomar $f_1 = f_1^0$, $f_2 = f_2^0$. Pero si $z(t) \neq z(t_0)$, las ramas $[f_1^0, z(t)]$ y $[f_2^0, z(t)]$ son diferentes, por la sencilla razón de que $f_1^0[z(t)] \neq f_2^0[z(t)]$, y si $z(t) = z(t_0)$, son diferentes por hipótesis. Luego E es cerrado, con lo que queda demostrado el teorema, puesto que el intervalo es conexo.

En virtud de este teorema una prolongación está determinada de manera única; p. ej., por su rama inicial $[f_0, z(\alpha)]$. Podemos, por tanto, hablar de *la* prolongación de **f** a lo largo de γ, a partir de la rama inicial $[f_0, z(\alpha)]$, siempre y cuando tal prolongación exista.

Si γ es un arco arbitrario y **f** una función analítica general, muy bien pudiera ocurrir que **f** no posea ninguna prolongación a lo largo de γ o que exista una prolongación para algunas ramas iniciales, pero no para todas. Investiguemos el caso de una rama inicial $[f_0, z(\alpha)]$ que no pueda prolongarse a lo largo de γ. Si $t_0 > \alpha$ está suficientemente próximo a α, existirá una prolongación de la rama inicial a lo largo del arco parcial correspondiente al intervalo (α, t_0); en efecto, este es trivialmente el caso si el arco parcial está contenido en la región Ω_0 del elemento de función (f_0, Ω_0). El extremo superior de todos estos t_0 es un número τ que satisface la desigualdad $\alpha < \tau < \beta$, y se ve fácilmente que la prolongación será posible para $t_0 < \tau$ e imposible para $t_0 \geqq \tau$. En cierto sentido se puede decir que el arco parcial correspondiente a (α, τ) conduce a un punto en el que **f** deja de estar definida. En particular, si **f** es una función analítica completa, el arco parcial se llama un *camino singular* desde la rama inicial dada; se dice también, con menos precisión, que conduce a un *punto singular* de **f**. Debería utilizarse el término punto singular únicamente cuando estuviese indicado claramente el correspondiente camino.

La relación entre prolongación a lo largo de arcos y prolongación por etapas mediante una cadena de prolongaciones analíticas directas requiere un análisis más preciso. En primer lugar, si (f_1, Ω_1), (f_2, Ω_2), ..., (f_n, Ω_n) forma una cadena de prolongaciones analíticas directas, es siempre posible enlazar un punto $z_1 \in \Omega_1$ a un punto $z_n \in \Omega_n$ mediante un arco γ, tal que **f** tenga una prolongación a lo largo de γ con la rama inicial (f_1, z_1) y rama terminal (f_n, z_n). En efecto, basta con hacer que γ esté compuesta por un arco parcial $\gamma_1 \subset \Omega_1$ desde z_1 a un punto $z_2 \in \Omega_1\Omega_2$, un segundo arco parcial $\gamma_2 \subset \Omega_2$ desde z_2 a $z_3 \in \Omega_2\Omega_3$, y así sucesivamente. La prolongación a lo largo de γ está completamente definida poniendo $\mathfrak{z}(t) = [f_k, z(t)]$ sobre γ_k.

Recíprocamente, si se da una prolongación $\mathfrak{z}(t)$, podemos hallar una cadena de prolongaciones analíticas directas que siguen el arco γ de la misma forma que en la construcción precedente, con tal únicamente que **f** esté definida mediante una colección suficientemente grande de elementos de función. Se prueba, utilizando el lema de Heine-Borel, que se puede subdividir el intervalo paramétrico en (α, t_1), (t_1, t_2), ..., $(t_{n-1}\, \beta)$ de forma que $\mathfrak{z}(t) = [f_k, z(t)]$ en (t_{k-1}, t_k) para elementos de función (f_k, Ω_k) escogidos adecuadamente. Aunque no es necesario que (f_{k-1}, Ω_{k-1}) y (f_k, Ω_k) sean prolongaciones analíticas directas uno de otro, son al menos prolongaciones analíticas directas de su restricción común a un entorno de $z(t_{k-1})$. Si estas restricciones están contenidas en la colección **f**, y este es ciertamente el caso si **f** es una función analítica completa, podemos entonces hallar una cadena de prolongaciones analíticas directas con las propiedades deseadas.

A fin de ilustrar el empleo de la prolongación a lo largo de arcos daremos una definición de $\log z$ como función analítica completa. La definimos como la colección de todos los elementos de función (f, Ω), tales que $e^{f(z)} = z$ en Ω. Hay que demostrar que esta colección es completa.

Hemos de probar que cualquier par de elementos de función (f_1, Ω_1), (f_2, Ω_2) de la colección puede unirse por una cadena de prolongaciones analíticas directas. Como consecuencia de la permanencia de las relaciones funcionales es evidente que los elementos de función intermedios pertenecerán a la misma colección.

Elijamos puntos $z_1 \in \Omega_1$, $z_2 \in \Omega_2$ y unámoslos por un arco diferenciable γ que no pase por el origen. Esto es posible, ya que ni z_1 ni z_2 pueden ser 0. Consideremos la función

$$\varphi(t) = f_1(z_1) + \int_\alpha^t \frac{z'(t)}{z(t)}\, dt.$$

Derivando, tenemos que $z(t)e^{-\varphi(t)}$ es constante; para $t=\alpha$ su valor es 1 y, por tanto, $e^{\varphi(t)} = z(t)$. Para un t dado, existe, en cualquier entorno de $z(t)$ que no incluya al origen, una rama $f(z)$ de $\log z$ determinada de manera única y que toma el valor $\varphi(t)$ para $z=z(t)$. Es evidente que $[f, z(t)]$ define una prolongación a lo largo de γ. La rama terminal quizá no coincida con f_2, pero su valor debe diferir de $f_2(z_2)$ en un múltiplo de $2\pi i$. Con objeto de obtener el valor deseado en z_2, todo lo que resta por hacer es prolongar la rama terminal a lo largo de una curva cerrada que envuelva el origen un número de veces adecuado. Por último, se puede reemplazar la prolongación a lo largo de arcos por una cadena finita de prolongaciones analíticas directas, con lo que queda demostrado que $\log z$ es una función analítica completa.

EJERCICIOS

1. Defínase $\log f(z)$ para una función $f(z)$ distinta de cero y uniforme.

2. Si se define un elemento de función mediante una serie de potencias en el interior de su círculo de convergencia, pruébese que la correspondiente función analítica completa tiene necesariamente un camino singular en el círculo de convergencia, que conduce a un punto sobre la circunferencia. ("Una serie de potencias tiene al menos un punto singular sobre la circunferencia de su círculo de convergencia.")

4. *Curvas homotópicas.*—Debemos ahora estudiar las propiedades topológicas de las curvas cerradas en una región desde un punto de vista que es fundamental para la teoría de prolongaciones analíticas. El problema que nos interesa es el comportamiento de un arco bajo *deformaciones continuas*. Intuitivamente es esta una noción muy sencilla. Si γ_1 y γ_2 son dos arcos que tienen extremos comunes, contenidos en una región Ω, parece natural preguntarse si se puede deformar de manera continua γ_1 en γ_2 cuando los extremos se mantienen fijos y los arcos intermedios quedan siempre dentro de Ω.

Así, p. ej., en la figura 6-1 el arco γ_1 puede deformarse de manera continua en el γ_2, pero no en el γ_3. Dos arcos que pueden deformarse uno en el otro se dice que son *homotópicos* con respecto a Ω. Esta es evidentemente una relación de equivalencia.

Naturalmente, debe darse una definición precisa. Por fortuna, el concepto físico de deformación tiene una interpretación casi inmediata en términos matemáticos. Es realmente evidente que puede

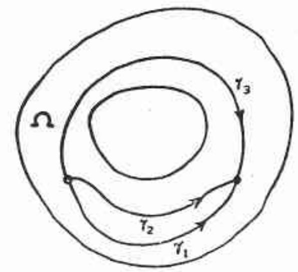

Fig. 6-1.

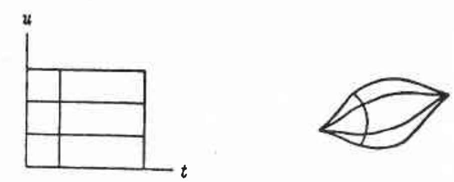

Fig. 6-2.

describirse la deformación de un arco mediante una función continua $z=z(t,u)$ de dos variables, variando el punto (t,u) sobre un rectángulo $\alpha \leqq t \leqq \beta$, $0 \leqq u \leqq 1$. A cada valor fijo $u=u_0$ corresponde un arco $z=z(t,u_0)$, y el efecto de la deformación es el de cambiar el arco inicial $z=z(t,0)$ en el $z=z(t,1)$. La deformación tiene lugar en el interior de Ω si $z(t,u) \in \Omega$ para todo (t,u), y es una deformación con extremos fijos si $z(\alpha,u)$ y $z(\beta,u)$ son constantes. A cada valor fijo de $t=t_0$ corresponde un arco $z=z(t_0,u)$, con u como parámetro, al que podemos llamar *camino de deformación* del punto correspondiente a t_0. La figura 6-2 muestra el efecto de una deformación.

Llegamos así a la siguiente definición de homotopía:

Se dice que dos arcos γ_1 y γ_2 definidos por las ecuaciones $z=z_1(t)$ y $z=z_2(t)$ sobre el mismo intervalo paramétrico $\alpha \leqq t \leqq \beta$ son homotópicos en Ω si existe una función continua $z(t,u)$ de dos variables, definida para $\alpha \leqq t \leqq \beta$, $0 \leqq u \leqq 1$, con las siguientes propiedades:

1) $z(t,u) \in \Omega$ *para todo* (t,u).
2) $z(t,0) = z_1(t)$, $z(t,1) = z_2(t)$ *para todo* t.
3) $z(\alpha,u) = z_1(\alpha) = z_2(\alpha)$, $z(\beta,u) = z_1(\beta) = z_2(\beta)$ *para todo* u.

El tomar los dos intervalos paramétricos iguales se hace solo por razones de comodidad. Si no fuera este el caso, transformaríamos un intervalo en otro mediante un cambio lineal de parámetro, y se conviene en considerar los dos arcos originales como homotópicos si son homotópicos con la nueva parametrización.

El lector puede efectuar con facilidad las sencillas demostraciones formales que muestran que la relación de homotopía, tal como anteriormente se ha definido, es una verdadera relación de equivalencia. Es posible, pues, agrupar todos los arcos en clases de equivalencia, llamadas *clases de homotopía;* los arcos en una clase de homotopía tienen extremos comunes y pueden deformarse unos en otros dentro de Ω. Merece señalarse el hecho de que las diferentes representaciones paramétricas del mismo arco son siempre homotópicas. En efecto, $z=z_1(t)$ es una nueva parametrización de $z=z_2(t)$ si (y solo si) existe una función no decreciente $\tau(t)$ tal que $z_1(t)=z_2[\tau(t)]$. La función

$$z(t, u)=z_1[(1-u)t+u\tau(t)]$$

tiene todos sus valores sobre el arco en consideración y, por tanto, en Ω. Para $u=0$ y $u=1$ obtenemos, respectivamente, $z(t, 0)=z_1(t)$ y

$$z(t, 1)=z_1[\tau(t)]=z_2(t)$$

como se requería, y evidentemente los extremos quedan fijos.

Si se trazan sucesivamente dos arcos γ_1 y γ_2, empezando γ_2 en el punto terminal de γ_1, forman un nuevo arco que denotaremos por $\gamma_1\gamma_2$, en contraste con la notación $\gamma_1+\gamma_2$, preferida en la teoría de la homología. La parametrización de $\gamma_1\gamma_2$ no está determinada de manera única, pero esto carece de importancia en lo que se refiere a la determinación de la clase de homotopía. Sin embargo, un razonamiento muy sencillo muestra que la clase de homotopía de $\gamma_1\gamma_2$ depende solo de las clases de homotopía de γ_1 y de γ_2. En virtud de este hecho fundamental podemos considerar la operación que conduce a la clase de homotopía de $\gamma_1\gamma_2$ como una multiplicación de clases de homotopía. Está definida únicamente cuando coinciden el punto inicial de γ_2 con el punto terminal de γ_1. Si limitamos nuestra atención a las clases de homotopía de curvas cerradas que empiezan y terminan en un punto fijo z_0, el producto está siempre definido y queda representado por una curva de la misma familia. Lo que es más, con esta definición de producto las clases de homotopía de curvas cerradas desde z_0, con respecto a la región Ω, forman un *grupo.* Con objeto de demostrar esta afirmación debemos probar:

1.° La ley asociativa: $(\gamma_1\gamma_2)\gamma_3$ es homotópico a $\gamma_1(\gamma_2\gamma_3)$.

2.° Existencia de una curva *unidad* 1 tal que $\gamma 1$ y 1γ sean homotópicas a γ.

3.° Existencia de una inversa γ^{-1} tal que $\gamma\gamma^{-1}$ y $\gamma^{-1}\gamma$ sean homotópicas a 1.

La ley asociativa es trivial, puesto que $(\gamma_1\gamma_2)\gamma_3$ es a lo sumo una nueva parametrización de $\gamma_1(\gamma_2\gamma_3)$. Como curva unidad podemos elegir la constante $z=z_0$; en realidad, con el símbolo 1 se representa cualquier curva cerrada que se pueda contraer al punto z_0. Por último, la inversa γ^{-1} es la curva γ trazada en sentido opuesto. Si se representa γ por $z=z(t)$, $\alpha\leqq t\leqq\beta$, γ^{-1} puede representarse por $z=z(2\beta-t)$, $\beta\leqq t\leqq 2\beta-\alpha$. La ecuación de $\gamma\gamma^{-1}$ es, pues,

$$\begin{array}{ll} z=z(t) & \text{para } \alpha\leqq t\leqq\beta \\ z=z(2\beta-t) & \text{para } \beta\leqq t\leqq 2\beta-\alpha. \end{array}$$

La curva puede contraerse a un punto mediante la deformación

$$\begin{array}{ll} z(t, u)=z(t) & \text{para } \alpha\leqq t\leqq u\alpha+(1-u)\beta \\ z(t, u)=z[u\alpha+(1-u)\beta] & \text{para } u\alpha+(1-u)\beta\leqq t\leqq u(\beta-\alpha)+\beta \\ z(t, u)=z(2\beta-t) & \text{para } u(\beta-\alpha)+\beta\leqq t\leqq 2\beta-\alpha. \end{array}$$

La interpretación del proceso es obvia: se hace retroceder el punto de giro desde $z(\beta)$ a $z(\alpha)$. Puesto que $z(t, 1)=z(\alpha)=z_0$, hemos demostrado que $\gamma\gamma^{-1}$ es homotópica a 1. La demostración es independiente de la hipótesis de que γ sea una curva cerrada; así, pues, $\gamma\gamma^{-1}$ es homotópica a 1 para cualquier arco γ que parta de z_0.

El grupo que hemos construido se llama *grupo de homotopía* o *grupo fundamental* de la región Ω con respecto al punto z_0. Como grupo abstracto no depende del punto z_0. Si z'_0 es otro punto de Ω, podemos unir z_0 con z'_0 mediante un arco c en Ω. A cada curva cerrada γ' desde z'_0 corresponde una curva cerrada $\gamma=c\gamma'c^{-1}$ desde z_0. Esta correspondencia conserva la homotopía y puede, por tanto, considerarse como una correspondencia entre clases de homotopía. Como tal conserva el producto, pues $(c\gamma'_1c^{-1})(c\gamma'_2c^{-1})$ es homotópico a $c(\gamma'_1\gamma'_2)c^{-1}$, por cancelación de $c^{-1}c$. Por último, la correspondencia es biunívoca, pues si se da γ podemos escoger $\gamma'=c^{-1}\gamma c$ y hallar que la curva correspondiente $c\gamma'c^{-1}=(cc^{-1})\gamma(cc^{-1})$ es homotópica a γ. Queda así demostrado que los grupos de homotopía con respecto a z_0 y a z'_0 son *isomorfos.*

Si γ_1 y γ_2 son dos arcos cualesquiera con punto inicial z_0 y un punto terminal común, entonces γ_1 es homotópico a γ_2 si (y solo si) $\gamma_1\gamma_2^{-1}$ es homotópico a 1. Pues si γ_1 es homotópico a γ_2, entonces $\gamma_1\gamma_2^{-1}$ lo es a $\gamma_2\gamma_2^{-1}$ y, por tanto, a 1. Recíprocamente, si $\gamma_1\gamma^{-1}$ es homotópico a 1,

$$(\gamma_1\gamma_2^{-1})\gamma_2=\gamma_1(\gamma_2^{-1}\gamma_2)$$

es simultáneamente homotópico a γ_1 y γ_2, con tal que γ_1 sea homotópico a γ_2. Por esta razón es suficiente el estudio de la homotopía de las curvas cerradas.

La determinación explícita de los grupos de homotopía se simplifica por el hecho de que el grupo de homotopía es obviamente un invariante topológico. En efecto, cualquier deformación en Ω puede llevarse sobre Ω' mediante una aplicación topológica de Ω sobre Ω', y se comprueba que determina una correspondencia biunívoca entre clases de homotopía que conserva el producto. Regiones topológicamente equivalentes tienen, por consiguiente, grupos de homotopía isomorfos.

El grupo de homotopía de un disco se reduce al elemento unidad; esto significa que dos arcos cualesquiera con extremos comunes son homotópicos. La demostración hace uso de la convexidad del disco: se puede deformar el arco $z=z_1(t)$ en el $z=z_2(t)$ mediante la deformación

$$z(t,u)=(1-u)z_1(t)+uz_2(t),$$

cuyos caminos de deformación son segmentos. La misma demostración sería válida para cualquier región convexa. En particular, todo el plano tiene análogamente un grupo de homotopía que se reduce al elemento unidad.

Se demostró en el capítulo IV, sección 4-4, 2, que cualquier región simplemente conexa que no sea todo el plano puede aplicarse de manera conforme sobre un disco. En lo que ahora nos concierne, el hecho de la conformidad carece de importancia, pero el que la aplicación sea topológica nos permite llegar a la conclusión de que cualquier región simplemente conexa tiene un grupo fundamental que se reduce a su elemento unidad. Hallaremos que la recíproca es también cierta.

5. *Teorema de monodromía.*—Sea Ω una región fija del plano z. Consideraremos el caso de una función analítica general **f** que puede

prolongarse a lo largo de todos los arcos γ contenidos en Ω, empezando con una rama cualquiera definida en el origen de γ. Con más precisión, para cualquier arco $z=z(t)$, $\alpha\leqq t\leqq\beta$, contenido en Ω, y para todo elemento de función $(f_0,\Omega_0)\in\mathbf{f}$ con $z(\alpha)\in\Omega_0$, existirá una prolongación $\mathfrak{z}(t)=[f,z(t)]$ cuya rama inicial es la definida por (f_0,Ω_0).

Si se dan dos arcos γ_1, γ_2 con extremos comunes, nos interesa saber si una rama inicial común, prolongada a lo largo de γ_1 y γ_2 conducirá a la misma rama terminal. El teorema básico, conocido como *teorema de monodromía*, es el siguiente:

TEOREMA 2. *Si los arcos γ_1 y γ_2 son homotópicos con respecto a Ω, y si se puede prolongar una rama inicial de* **f** *a lo largo de todos los arcos contenidos en Ω, entonces la prolongación de esta rama inicial a lo largo de γ_1 y γ_2 debe conducir a la misma rama terminal.*

Para empezar observemos que la prolongación a lo largo de un arco de la forma $\gamma\gamma^{-1}$ nos hará volver evidentemente a la rama inicial. De manera análoga, la prolongación a lo largo de un arco de la forma $\sigma_1(\gamma\gamma^{-1})\sigma_2$ tendrá el mismo efecto que una prolongación a lo largo de $\sigma_1\sigma_2$. Por esta razón, decir que las proiongaciones a lo largo de γ_1 y γ_2 conducen a la misma rama terminal equivale a decir que prolongaciones a lo largo de $\gamma_1\gamma_2^{-1}$ conducen nuevamente al elemento inicial.

De acuerdo con la hipótesis, existe una deformación $z(t,u)$ de γ_1 en γ_2. Todo arco σ en el rectángulo de deformación R se aplica mediante $z(t,u)$ en un arco $\sigma'\in\Omega$, y si σ' comienza en el punto inicial de γ_1 y γ_2, existe una prolongación única de la rama inicial dada a lo largo de σ'. Para mayor sencillez diremos que es una prolongación a lo largo de σ. El teorema afirma que la prolongación a lo largo del perímetro Γ de R hace volver al elemento inicial. Es indiferente el sentido en que se describa Γ, pero debe fijarse de una vez para siempre.

Una demostración sencilla puede basarse en el método de bisección. Empezamos por bisecar R horizontalmente, y denotamos por π_1 el perímetro de la mitad inferior R_1, descrito desde el vértice inferior izquierdo cero y en el sentido que coincide con el sentido de Γ a lo largo del lado común. Asociamos a la mitad superior R_2 una curva π_2 con origen en cero, que sigue verticalmente hasta el vértice inferior izquierdo de R_2, describe el perímetro de R_2 en el sentido que coincide con el de Γ a lo largo del lado común y vuelve verticalmente a cero (Fig. 6-3). Observemos que la curva $\pi_1\pi_2$ difiere de Γ únicamente en un arco intermedio de la forma $\sigma\sigma^{-1}$. Por esta razón, el efecto de una prolongación a lo largo de $\pi_1\pi_2$ es el mismo que el de una prolongación a lo largo de Γ.

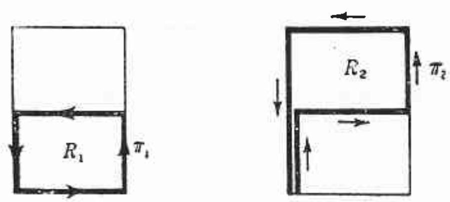

FIG. 6-3.

Por consiguiente, si π_1 y π_2 conducen ambos nuevamente a la rama inicial, lo mismo hace Γ. Hagamos ahora la hipótesis contraria de que Γ no hace volver a la rama inicial. Entonces o π_1 o π_2 tiene la misma propiedad. Se biseca verticalmente el rectángulo correspondiente y se aplica el mismo razonamiento. Por repetición de este proceso obtenemos una sucesión de rectángulos $R \supset R^{(1)} \supset R^{(2)} \supset \ldots \supset R^{(n)} \supset \ldots$ y curvas cerradas correspondientes $\pi^{(n)}$ tales que la prolongación de la rama inicial a lo largo de $\pi^{(n)}$ *no* hace volver a la misma rama. Cada $\pi^{(n)}$ es de la forma $\sigma_n\Gamma_n\sigma_n^{-1}$, siendo σ_n una línea poligonal bien determinada que va desde cero hasta el vértice inferior izquierdo de $R^{(n)}$, y denotando Γ_n el perímetro de $R^{(n)}$; además, σ_n es un arco parcial de σ_{n+1}.

Al tender n a infinito, los rectángulos $R^{(n)}$ convergen a un punto P_∞ y las líneas poligonales σ_n forman, en el límite, una curva continua σ_∞ que acaba en P_∞. Existe una prolongación de la rama inicial a lo largo de σ_∞ que termina en una rama $[f_\infty, z(P_\infty)]$ en el punto correspondiente a P_∞. Para n suficientemente grande, la imagen de Γ_n estará contenida en un entorno Δ de $z(P_\infty)$ y la rama obtenida en el punto terminal de σ_n debe estar determinada por el elemento de función (f_∞, Δ). Cuando este sea el caso, puede utilizarse el elemento (f_∞, Δ) para construir una prolongación a lo largo de $\pi^{(n)}$, que conduce nuevamente a la rama inicial. Esto está en contradicción con la propiedad por la que se definió $\pi^{(n)}$, con lo que hemos demostrado que la prolongación a lo largo de Γ debe acabar en la rama inicial.

El teorema de monodromía implica, sobre todo, que cualquier función analítica general que puede prolongarse a lo largo de todos los arcos en una región simplemente conexa determina una función analítica uniforme para cada elección de la rama inicial. Este hecho puede expresarse también diciendo que una superficie de Riemann (sin puntos de ramificación) sobre una región simplemente conexa debe consistir en una sola hoja.

Podemos sacar además la consecuencia, ya anunciada, de que una región cuyo grupo de homotopía se reduce al elemento unidad es necesario que sea simplemente conexa. En efecto, supongamos que Ω es múltiplemente conexa. Existe entonces una componente acotada E_0 del complemento de Ω, y si $z_0 \in E_0$, sabemos que $\log(z-z_0)$ no es uniforme en Ω. Del teorema de monodromía se deduce que el grupo de homotopía de Ω no puede reducirse al elemento unidad.

Esta es la última etapa de la demostración de la equivalencia de las tres caracterizaciones siguientes de regiones simplemente conexas: 1) Ω es simplemente conexa si su complemento es conexo; 2) Ω es simplemente conexa si es homeomorfa a un disco; 3) Ω es simplemente conexa si su grupo fundamental se reduce al elemento unidad.

6. *Puntos de ramificación.*—Para un estudio más detallado de las singularidades de funciones multiformes es necesario determinar, de manera explícita, el grupo fundamental del disco desprovisto de su centro. Representemos el disco sin su centro por $0 < |z| < 1$ y consideremos un punto fijo; p. ej., el punto $z_0 = r$ sobre el radio positivo. Por una proyección central, dada por

$$z(t, u) = (1-u)z(t) + ur\frac{z(t)}{|z(t)|},$$

cualquier curva cerrada $z = z(t)$ desde z_0 puede deformarse en una curva que esté sobre la circunferencia $|z| = r$. Por tanto, es suficiente considerar curvas γ sobre esta circunferencia. Se escribirá la ecuación de γ nuevamente como $z = z(t)$.

Por continuidad, t_0 posee un entorno en donde $|z(t) - z(t_0)| < r/2$; en tal entorno no es posible que $z(t)$ tome a la vez los valores r y $-r$. Se deduce fácilmente, utilizando el lema de Heine-Borel o el método de bisección, la posibilidad de escribir $\gamma = \gamma_1\gamma_2 \ldots \gamma_n$, donde cada γ_k o no pasa por r o no pasa por $-r$. Para mayor sencillez designaremos los puntos r y $-r$ por las letras P_0 y P'_0 (Fig. 6-4) y denotaremos los extremos de γ_k por P_k y P_{k+1}. Puesto que γ_k está contenida en la región simplemente conexa obtenida por omisión del radio positivo o del radio negativo, podrá deformarse en uno de los dos arcos P_kP_{k+1}. Como consecuencia, puede deformarse γ en un producto de arcos simples con extremos sucesivos $P_0P_1P_2 \ldots P_nP_0$. Este camino puede reemplazarse a su vez por el $P_0P_1P_2P_0P_2P_3P_0 \ldots P_0P_{n-1}P_nP_0$, donde, para mayor precisión, cada arco P_kP_0 y P_0P_k es el que no contiene a P'_0. En efecto, el nuevo camino se obtiene

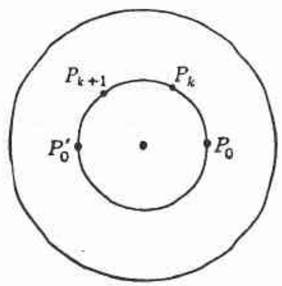

FIG. 6-4.

insertando los arcos trazados dos veces $P_kP_0P_k$, que sabemos que son homotópicos a 1.

Hemos demostrado que cada γ es homotópico a un producto de curvas cerradas de la forma $P_0P_kP_{k+1}P_0$. Si P_kP_{k+1} no contiene a P'_0, esta curva es homotópica a 1. Si, por otra parte, P_kP_{k+1} contiene a P'_0, se comprueba considerando todos los casos posibles que la curva es homotópica a C o a C^{-1}, siendo C la circunferencia completa. Por consiguiente, toda curva cerrada es homotópica a una potencia de C.

Por último, observemos que C^m es homotópica a 1 únicamente si $m=0$. Se ve esto por el hecho de que

$$\int_{C^m} \frac{dz}{z} = m \cdot 2\pi i,$$

en tanto que si la curva fuera homotópica a 1 la integral debería anularse. De estos resultados se concluye que el grupo fundamental del disco sin su centro es isomorfo al *grupo aditivo de los enteros*. Evidentemente, una corona circular cualquiera posee el mismo grupo fundamental.

Consideremos ahora una función analítica general **f**, que puede prolongarse a lo largo de todos los arcos en el disco sin su centro

$0 < |z| < 1$. Elegimos una rama inicial en $z_0 = r$ y la prolongamos a lo largo de las curvas C^m. Una de dos: o la prolongación no vuelve jamás a la rama inicial, o existe un entero positivo h más pequeño que todos, tal que C^h conduce nuevamente a la rama inicial. En el último caso, pongamos $m = qh + r$ con $0 \leqq r < h$. Si C^m conduce nuevamente a la rama inicial, lo mismo ocurre con C^r. Pero puesto que $r < h$, esto es solo posible si $r = 0$, lo que nos dice que C^m conduce nuevamente a la rama inicial si (y solo si) m es un múltiplo de h.

Consideremos la aplicación $z = \zeta^h$ de $0 < |\zeta| < 1$ sobre $0 < |z| < 1$. Afirmamos que se puede expresar **f** como una función analítica uniforme $F(\zeta)$ en el disco sin centro del plano ζ. El sentido preciso de esta afirmación es el de que existe, para todo ζ_1, $0 < |\zeta_1| < 1$, un elemento de función $(f, \Omega) \in \mathbf{f}$ con $\zeta_1^h \in \Omega$, tal que $F(\zeta) = f(\zeta^h)$ en un entorno de ζ_1; se exige, en particular, que el elemento de función que corresponde al punto $\zeta_0 = r^{1/h}$ determine la rama inicial en z_0.

Con objeto de construir $F(\zeta)$, unimos ζ_0 a ζ mediante un arco γ' y prolongamos la rama inicial de **f** a lo largo de la imagen de γ' respecto de la aplicación $z = \zeta^h$; definimos $F(\zeta)$ como el valor terminal obtenido a través de esta prolongación. Hay que demostrar que $F(\zeta)$ está determinada de manera única. Si γ'_1 y γ'_2 son dos caminos desde ζ_0 a ζ, entonces $\gamma'_1\gamma_2'^{-1}$ puede deformarse en una potencia C'^n de la circunferencia que pasa por ζ_0. Por tanto, la curva imagen $\gamma_1\gamma_2^{-1}$ puede deformarse en la imagen de C'^n, que es C^{nh}. Pero C^{nh} conduce nuevamente a la rama inicial y, por tanto, γ_1 y γ_2 determinan la misma rama terminal. Por último, si ζ está en un entorno de ζ_1, podemos seguir primero un arco γ'_1 desde ζ_0 a ζ_1 y después un arco variable γ' desde ζ_1 hasta ζ que permanezca dentro del entorno. Si el entorno está suficientemente restringido, la prolongación a lo largo de la imagen de γ' está determinada por un único elemento de función (f, Ω), y tenemos $F(\zeta) = f(\zeta^h)$ en dicho entorno.

Puesto que $F(\zeta)$ es uniforme y analítica en un entorno reducido del origen, posee un desarrollo de Laurent convergente de la forma

$$F(\zeta) = \sum_{-\infty}^{\infty} A_n \zeta^n.$$

Se dice que la correspondiente función multiforme de z posee el desarrollo

$$f(z) = \sum_{-\infty}^{\infty} A_n z^{n/h}. \qquad [6\text{-}1]$$

Debe observarse que este desarrollo depende de la elección de la rama inicial. Ramas iniciales diferentes pueden dar desarrollos completamente distintos y, en particular, diferentes valores de h. La serie [6-1] da un total de h desarrollos asociados, obtenidos eligiendo diferentes valores iniciales de $z^{1/h}$. Si $\omega = e^{2\pi i/h}$, estos desarrollos están representados por

$$f_\nu(z) = \sum_{-\infty}^{\infty} A_n \omega^{\nu n} z^{n/h} \qquad (\nu = 0, 1, \ldots, h-1). \qquad [6\text{-}2]$$

Cuando se prolonga la rama (f_ν, z_0) a lo largo de C, conduce a la rama $(f_{\nu+1}, z_0)$, en el entendimiento de que hay que identificar el subíndice h con 0.

En casos especiales, el desarrollo de Laurent puede contener únicamente un número finito de potencias negativas. Entonces $F(\zeta)$ tiene o una singularidad evitable o un polo, y la función multiforme $f(z)$ (o, más correctamente, la función analítica general obtenida prolongando la rama inicial dada dentro del disco sin su centro) se dice que tiene una *singularidad algebraica* o *punto de ramificación* en $z=0$, con tal, naturalmente, que $h > 1$. Si $F(\zeta)$ tiene una singularidad evitable, el punto de ramificación es una singularidad algebraica *ordinaria;* en el caso opuesto es un *polo algebraico.* En cada caso, $f(z)$ tiende a un límite definido A_0 o a ∞ al tender z a cero a lo largo de un arco arbitrario.

Evidentemente, podríamos haber estudiado lo mismo una singularidad aislada en un punto cualquiera a o ∞, y el radio del disco sin el centro puede ser tan pequeño como queramos. En el caso de que h sea finita se puede expresar la correspondencia entre $w = f(z)$ y la variable independiente z mediante ecuaciones de la forma

$$w = \sum_{-\infty}^{\infty} A_n \zeta^n$$

$$z = a + \zeta^h \qquad \text{o} \qquad z = \zeta^{-h}.$$

La variable ζ toma el nombre de *variable de uniformización local.*

En el caso de una singularidad algebraica es deseable completar la superficie de Riemann de $\mathbf{f}$ por la inclusión de un punto de ramificación sobre la superficie. La proyección de este punto será a, y el punto en sí no está determinado por una rama de $\mathbf{f}$, sino por desarrollos análogos al [6-2]. Por último, un entorno está formado por las ramas que corresponden a puntos ζ en un entorno de $\zeta=0$. De ahora en adelante se supondrá que la superficie de Riemann de una función analítica general incluye todos estos puntos de ramificación.

6-2. Funciones algebraicas.—Una ecuación del tipo $P(w, z) = 0$, donde P representa un polinomio en dos variables, tiene para cada valor de z un número finito de soluciones $w_1(z), \ldots, w_n(z)$. Deseamos probar que estas raíces pueden interpretarse como valores de una función multiforme $\mathbf{f}(z)$, a la que se llamará *función algebraica.* Recíprocamente, si se da una función analítica general, queremos estar en condiciones de decidir si satisface o no una ecuación polinómica.

1. *La resultante de dos polinomios.*—Un polinomio en dos variables $P(w, z)$ es *irreducible* si no puede expresarse como el producto de dos polinomios, ninguno de los cuales sea constante. Dos polinomios P y Q son *primos entre sí* si no tienen factores comunes salvo constantes.

El siguiente teorema es de carácter algebraico. Sin embargo, a causa de su importancia fundamental en la teoría de funciones algebraicas daremos su demostración.

TEOREMA 3. *Si $P(w, z)$ y $Q(w, z)$ son polinomios primos entre sí, existe únicamente un número finito de valores z_0 para los cuales las ecuaciones $P(w, z_0) = 0$ y $Q(w, z_0) = 0$ tienen una raíz común.*

Supongamos que P y Q están ordenados según las potencias decrecientes de w y pongamos $Q(w, z) = b_0(z)w^m + \ldots + b_m(z)$, donde $b_0(z)$ no es idénticamente nulo. Si se divide P por Q, se obtiene un cociente y un resto que son polinomios en w y funciones racionales en z, con una potencia de $b_0(z)$ como denominador. Por tanto, es posible establecer un algoritmo de Euclides de la forma

$$\begin{aligned} b_0^{r_0} P &= q_0 Q + R_1 \\ b_0^{r_1} Q &= q_1 R_1 + R_2 \\ b_0^{r_2} R_1 &= q_2 R_2 + R_3 \\ &\cdots\cdots\cdots\cdots \\ b_0^{r_{n-1}} R_{n-2} &= q_{n-1} R_{n-1} + R_n \end{aligned} \qquad [6\text{-}3]$$

donde los q_k y R_k son ahora polinomios y $r_0, \ldots, r_{n-1}$ denotan exponentes enteros $\geqq 0$. Los grados en w de los R_k son decrecientes, siendo R_n un polinomio en z solamente. Si $R_n(z)$ fuera idénticamente nulo, el teorema de factorización única implica, como consecuencia de la última relación en [6-3], que R_{n-2} es divisible por cualquier factor irreducible de R_{n-1} que sea de grado positivo en w. El mismo razonamiento, paso a paso, muestra que todos los R_k, lo mismo que Q y P, son divisibles por el mismo factor. Esto es contrario a la hipótesis, pues R_{n-1} es de grado positivo en w y debe tener, por consiguiente, un factor irreducible que contenga a w.

Supongamos ahora que $P(w_0, z_0) = 0$ y $Q(w_0, z_0) = 0$. Sustituyendo estos valores en [6-3] se obtiene $R_1(w_0, z_0) = 0, \ldots, R_{n-1}(w_0, z_0) = 0$, y finalmente, $R_n(z_0) = 0$. Pero puesto que R_n no es idénticamente cero, existe solo un número finito de valores z_0 que satisfacen esta condición, de donde se deduce el teorema.

Al polinomio $R_n(z)$ se le llama la *resultante* de P y Q. Con más precisión, si queremos que la resultante esté determinada de manera única, deberíamos exigir que los exponentes r_k en [6-3] sean lo más pequeños posible. En realidad, no estamos tan interesados en la resultante como en el enunciado del teorema 3. Se aplicará el teorema a un polinomio irreducible $P(w, z)$ y a su derivada parcial $P_w(w, z)$ con respecto a w. Si P tiene grado positivo con respecto a w, estos polinomios serán primos entre sí, y la resultante de P y P_w se denomina el *discriminante* de P. Los ceros del discriminante son los valores z_0 para los que la ecuación $P(w, z_0) = 0$ posee raíces múltiples.

Observemos, por último, que la resultante $R(z)$ de cualquier par de polinomios primos entre sí P y Q puede escribirse en la forma $R = pP + qQ$, siendo p y q polinomios. Esto es consecuencia inmediata de [6-3].

2. *Definición y propiedades de las funciones algebraicas.*—Empezamos formulando una definición precisa:

DEFINICIÓN 2. *Se dice que una función analítica completa* $\mathbf{f}$ *es una función algebraica si todos sus elementos de función* (f, Ω) *satisfacen una relación* $P[f(z), z] = 0$ *en* Ω, *siendo* $P(w, z)$ *un polinomio que no se anula idénticamente.*

Como consecuencia de la permanencia de las relaciones funcionales, es suficiente suponer que un elemento de función satisface la ecuación $P[f(z), z] = 0$. Los otros satisfarán entonces automáticamente la misma relación. Además, puede suponerse que $P(w, z)$ es un polinomio irreducible. En efecto, supongamos que $P(w, z)$ posee la descomposición factorial $P = P_1 P_2 \ldots P_n$ en factores irreduci-

bles. Para cualquier punto fijo $z \in \Omega$ debe verificarse una de las ecuaciones $P_k[f(z), z] = 0$. Si consideramos una sucesión de puntos distintos $z_n \in \Omega$ que tienda a un límite en Ω, entonces una de las relaciones $P_k[f(z_n), z_n] = 0$ debe verificarse infinitas veces. Se sigue que esta relación particular $P_k[f(z), z] = 0$ debe satisfacerse idénticamente en Ω y, por tanto, para todos los elementos de función de $\mathbf{f}$. Podemos, pues, reemplazar P por P_k.

También se ve fácilmente que el polinomio irreducible P determinado por una función algebraica lo está de manera única salvo un factor constante. Si Q es un polinomio irreducible esencialmente diferente, podemos determinar la resultante $R(z) = pP + qQ$. Si $P[f(z), z] = 0$ y $Q[f(z), z] = 0$ para todo $z \in \Omega$, obtendríamos $R(z) = 0$ en Ω, en contradicción con el hecho de que $R(z)$ no es idénticamente nulo. Observemos que P no puede reducirse a un polinomio en z únicamente. Si contiene solo a w, deberá ser de la forma $w - a$, reduciéndose la función $\mathbf{f}$ a la constante a.

Probaremos a continuación que existe una función algebraica en correspondencia con cualquier polinomio irreducible $P(w, z)$ de grado positivo en w. Supongamos que

$$P(w, z) = a_0(z) w^n + a_1(z) w^{n-1} + \ldots + a_n(z).$$

Si z_0 no es un cero ni del polinomio $a_0(z)$ ni del discriminante de P, la ecuación $P(w, z_0) = 0$ tiene exactamente n raíces distintas $w_1, w_2, \ldots, w_n$. Bajo estas condiciones se verifica el siguiente

LEMA 1. *Existe un disco abierto* Δ, *que contiene a* z_0, *y* n *elementos de función* $(f_1, \Delta), (f_2, \Delta), \ldots, (f_n, \Delta)$ *con las propiedades:*

a) $P[f_i(z), z] = 0$ *en* Δ;
b) $f_i(z_0) = w_i$;
c) *si* $P(w, z) = 0$, $z \in \Delta$, *entonces* $w = f_i(z)$ *para algún* i.

El polinomio $P(w, z_0)$ tiene ceros simples en $w = w_i$. Determinemos $\epsilon > 0$ tal que los discos $|w - w_i| \leqq \epsilon$ no tengan parte común y denotemos las circunferencias $|w - w_i| = \epsilon$ por C_i. Entonces $P(w, z_0) \neq 0$ sobre C_i, y por el principio del argumento

$$\frac{1}{2\pi i} \int_{C_i} \frac{P_w(w, z_0)}{P(w, z_0)} \, dw = 1.$$

Si se reemplaza z_0 por z, las integrales se convierten en funciones continuas de z bien definidas en un entorno de z_0. Puesto que solo pueden tomar valores enteros, existe un entorno Δ tal que

$$\frac{1}{2\pi i}\int_{C_i}\frac{P_w(w,z)}{P(w,z)}\,dw=1$$

para todo $z\in\Delta$. Esto significa que la ecuación $P(w,z)=0$ tiene exactamente una raíz en el disco $|w-w_i|<\epsilon$; denotamos esta raíz por $f_i(z)$. Según el cálculo de residuos, su valor viene dado por

$$f_i(z)=\frac{1}{2\pi i}\int_{C_i} w\,\frac{P_w(w,z)}{P(w,z)}\,dw.$$

Esta representación muestra que $f_i(z)$ es analítica. Además, $f_i(z_0)=w_i$, y *c)* se deduce del hecho de que hemos obtenido n raíces de la ecuación $P(w,z)=0$, y esta no puede tener más.

El lema implica inmediatamente la existencia de una función algebraica **f** correspondiente al polinomio P; de hecho, podemos elegir **f** de forma que sea la función analítica completa determinada por el elemento (f_1,Δ) para cualquier z_0 que no coincida con ninguno de los puntos excluidos, en número finito. Demostraremos, además, que todos estos elementos de función pertenecen a la misma función analítica completa; esto probará también que la función **f** que corresponde a P es única. Sea (f,Ω) uno de estos elementos de función. Debe existir un $z_0\in\Omega$ que no sea uno de los puntos excluidos; determinemos un Δ correspondiente. Puesto que $P[f(z),z]=0$ para $z\in\Omega$, se sigue de *c)* que $f(z)$ es igual a alguno de los $f_i(z)$ en cada punto de $\Delta\Omega$. Pero entonces $f(z)$ es igual a la misma $f_i(z)$ en infinitos puntos de cualquier entorno de z_0 y, por tanto, (f,Ω) pertenece a la función analítica completa determinada por (f_i,Δ).

Denotemos los puntos excluidos por $c_1, c_2, \ldots, c_m$. Deseamos probar que un elemento de función (f,Ω) que satisface la ecuación $P[f(z),z]=0$ puede prolongarse a lo largo de cualquier arco que no pase por ningún punto c_k. Si no fuera así, existiría un arco $z=z(t)$, $\alpha\leqq t\leqq\beta$, tal que una rama inicial dada podría prolongarse a lo largo de todos los arcos parciales $\alpha\leqq t\leqq\tau<\beta$, pero no a lo largo del arco total. Pongamos $z_0=z(\beta)$, determinemos Δ de acuerdo con el lema 1 y elijamos τ de forma que $z(t)\in\Delta$ para $\tau\leqq t\leqq\beta$. El mismo razonamiento anterior muestra que la rama $[f,z(\tau)]$ obtenida por la prolongación debe estar determinada por uno de los elementos de función (f_i,Δ). Pero entonces puede prolongarse completamente hasta β, con lo que hemos llegado a una contradicción.

No se ha demostrado todavía el que todos los elementos (f_i,Δ) pertenecen a la misma función analítica. Para esta parte de la demostración es necesario el estudio del comportamiento en los puntos críticos c_k con mayor detalle.

3. *Comportamiento en los puntos críticos.*—Los puntos c_k que hasta ahora fueron excluidos de nuestras consideraciones eran los ceros del primer coeficiente $a_0(z)$ de P y los del discriminante. Elijamos δ de forma que el disco $|z-c_k|\leqq\delta$ no contenga más punto crítico que el c_k. Fijemos un punto $z_0\neq c_k$ en este disco y elijamos una de las ramas $f_i(z)$ en ese punto. Esta rama puede prolongarse a lo largo de todos los arcos en el disco sin centro. Además, si se prolonga a lo largo de la circunferencia C de centro c_k, que pasa por z_0, debemos volver con una rama $f_j(z)$. Puesto que solo hay un número finito de tales ramas, se sigue fácilmente que debe existir un entero positivo $h\leqq n$ más pequeño que todos, con la propiedad de que la prolongación a lo largo de C^h conduce nuevamente a la rama inicial $f_i(z)$. Por el resultado fundamental de la sección 6-1, 6, podemos escribir

$$f_i(z)=\sum_{\nu=-\infty}^{\infty}A_\nu(z-c_k)^{\nu/h}. \qquad [6\text{-}4]$$

Supongamos en primer lugar que c_k no es un cero de $a_0(z)$. Entonces $f_i(z)$ permanece acotada cuando z tiende a c_k. En efecto, en cuanto que $f_i(z)\neq0$, la ecuación $P[f_i(z),z]=0$ puede escribirse en la forma

$$a_0(z)+a_1(z)f_i(z)^{-1}+\ldots+a_n(z)f_i(z)^{-n}=0. \qquad [6\text{-}5]$$

Si $f_i(z)$ no estuviera acotada, existirían puntos $z_n\to c_k$, con $f_i(z_n)\to\infty$. Por sustitución en [6-5] obtendríamos $a_0(z_n)\to0$, en contra de la hipótesis de que $a_0(c_k)\neq0$. Se sigue que el desarrollo [6-4] contiene únicamente potencias positivas, y la rama $f_i(z)$ tiene una singularidad algebraica ordinaria en c_k.

Consideremos ahora el caso en que $a_0(c_k)=0$. Si se denota por m la multiplicidad del cero, sabemos que $\lim\limits_{z\to c_k} a_0(z)(z-c_k)^{-m}\neq0$. Se obtiene de [6-5]

$$a_0(z)(z-c_k)^{-m}+a_1(z)(z-c_k)^{-m}f_i(z)^{-1}+\ldots+$$
$$+a_n(z)(z-c_k)^{-m}f_i(z)^{-n}=0.$$

Si la expresión $f_i(z)(z-c_k)^m$ no estuviera acotada, llegaríamos nuevamente a una contradicción. Al igual que en la sección 6-1, 6, escribimos

$$F(\zeta) = \sum_{-\infty}^{\infty} A_\nu \zeta^\nu$$

y hallamos que $F(\zeta)\zeta^{mh}$ está acotada. En consecuencia, $F(\zeta)$ tiene un polo de a lo sumo orden mh, y la rama $f_i(z)$ tiene un polo algebraico en c_k, o, en casos especiales, una singularidad algebraica ordinaria.

Por último, es también necesario discutir el comportamiento en $z=\infty$. Es evidente que tenemos un desarrollo de la forma

$$f_i(z) = \sum_{-\infty}^{\infty} A_\nu z^{\nu/h},$$

válido en un entorno de ∞. Supongamos que el polinomio $a_i(z)$ es de grado r_i (no se considerarán los coeficientes que se anulen idénticamente). Elijamos un entero m tal que

$$m > \frac{1}{k}(r_k - r_0) \qquad [6\text{-}6]$$

para $k=1, \ldots, n$. Afirmamos que $f_i(z)z^{-m}$ ha de estar acotada cuando $z \to \infty$. Si no fuera así tendríamos $f_i(z)^{-1}z^m \to 0$ para una sucesión que tendiese a ∞. Esto implicaría $f_i(z)^{-k}z^{mk} \to 0$ y, por [6-6], $f_i(z)^{-k}z^{r_k - r_0} \to 0$ para $k \geqq 1$. Si se multiplica [6-5] por z^{-r_0}, se sigue que todos los términos excepto el primero tienden a cero. Esto constituye una contradicción, de donde se concluye que $f_i(z)$ posee a lo sumo un polo algebraico en el infinito.

En resumen, hemos demostrado que una función algebraica tiene a lo sumo singularidades algebraicas en el plano completo. Probaremos ahora un recíproco de este enunciado. Con objeto de obtener tal recíproco es esencial añadir una hipótesis que implica que existe únicamente un número finito de ramas en un punto dado.

Sea **f** una función analítica general. Para cada c supondremos la existencia de un disco Δ, desprovisto de su centro, c, tal que todas las ramas de **f** que estén definidas en un punto $z_0 \in \Delta$ puedan prolongarse a lo largo de todos los arcos en Δ, y probaremos el carácter algebraico en c. La hipótesis se satisfará también para $c=\infty$, en cuyo caso Δ es el exterior de un círculo. Además, para un Δ debe suponerse que el número de ramas distintas en z_0 es finito.

Puesto que el plano completo puede recubrirse por un número finito de discos Δ, el centro incluido, se sigue que únicamente un número finito de puntos c pueden ser singularidades efectivas; denotaremos estos puntos por c_k. Es fácil probar que el número de ramas en cualquier punto $z \neq c_k$ es constante, pues todos estos puntos poseen un entorno en el que todas las ramas de **f** son uniformes y pueden prolongarse a través de este entorno. Se sigue que el conjunto de puntos z con exactamente n ramas es abierto (n puede ser finito o infinito). Puesto que el plano completo menos los puntos c_k es conexo, únicamente uno de estos conjuntos es no vacío. Por tanto, n es constante; por hipótesis, no puede ser infinito, ni tampoco cero, pues en tal caso **f** sería una colección vacía de elementos de función.

Pueden denotarse ahora las ramas en cualquier punto $z \neq c_k$ por $f_1(z), \ldots, f_n(z)$, excepto que el orden permanece indeterminado. Formemos ahora las funciones simétricas elementales de las $f_i(z)$, es decir, los coeficientes del polinomio

$$[w - f_1(z)][w - f_2(z)] \;\ldots\; [w - f_n(z)].$$

Estos coeficientes son funciones bien definidas de z y obviamente analíticas, salvo posibles singularidades aisladas en los puntos c_k. Cuando z se aproxima a c_k sabemos que cada $f_i(z)$ debe crecer sin límite, a lo sumo como una potencia negativa de $|z-c_k|$. Por tanto, lo mismo es cierto para las funciones simétricas elementales. En definitiva, las singularidades aisladas, incluyendo la del infinito, son cuando más polos, y, por tanto, las funciones simétricas elementales son funciones racionales de z. Si se denota el denominador común por $a_0(z)$, hallamos que todas las ramas $f_i(z)$ deben satisfacer una ecuación polinómica

$$a_0(z)w^n + a_1(z)w^{n-1} + \ldots + a_n(z) = 0,$$

y esto prueba que **f** es algebraica.

Es ahora fácil resolver la cuestión que quedó planteada en la sección 6-2, 2. Supongamos que el elemento de función (f, Ω) satisface la ecuación $P[f(z), z] = 0$, donde P es irreducible y de grado n en w. Entonces la función analítica completa correspondiente **f** tiene solo singularidades algebraicas y un número finito de ramas. De acuerdo con lo que acabamos de probar, **f** satisfará una ecuación polinómica cuyo *grado es igual al número de ramas.* Satisfará, por tanto, una ecuación irreducible cuyo grado no es mayor. Pero la única ecuación irreducible que puede satisfacer es $P(w, z) = 0$, siendo su grado n. Por consiguiente, el número de ramas es exactamente n, y hemos demostrado que todas las soluciones de $P(w, z) = 0$ son ramas de esta misma función analítica.

Solo queda reunir los resultados:

TEOREMA 4. *Una función analítica general es una función algebraica si tiene un número finito de ramas y a lo sumo singularidades algebraicas. Toda función algebraica* $w = \mathbf{f}(z)$ *satisface una ecuación irreducible* $P(w, z) = 0$, *única salvo un factor constante, y todas estas ecuaciones determinan de manera única una función algebraica correspondiente.*

Se acostumbra decir que una ecuación irreducible $P(w, z) = 0$ define una *curva algebraica.* La teoría de las curvas algebraicas es una rama bastante desarrollada del álgebra y de la teoría de funciones. Hemos expuesto únicamente la parte más elemental del aspecto correspondiente a la teoría de funciones.

EJERCICIO

Determínese la posición y naturaleza de las singularidades de la función algebraica definida por $w^3 - 3wz + 2z^3 = 0$.

6-3. Ecuaciones diferenciales lineales.

—Con un alto grado de generalidad, la teoría de funciones analíticas multiformes hace posible el estudio de las soluciones complejas de las ecuaciones diferenciales ordinarias. De todas las ecuaciones diferenciales, las más sencillas, y también las más importantes, son las lineales. Una ecuación lineal de orden n tiene la forma

$$a_0(z)\frac{d^n w}{dz^n} + a_1(z)\frac{d^{n-1}w}{dz^{n-1}} + \ldots + a_{n-1}(z)\frac{dw}{dz} + a_n(z)w = b(z), \qquad [6\text{-}7]$$

donde los coeficientes $a_k(z)$ y el segundo miembro $b(z)$ son funciones analíticas uniformes. Con objeto de simplificar el tratamiento, limitaremos nuestra atención al caso en que estas funciones están definidas en todo el plano; se supone así que son funciones enteras.

Una *solución* de [6-7] es una función analítica general **f**, que satisface la identidad

$$a_0\mathbf{f}^{(n)} + a_1\mathbf{f}^{(n-1)} + \ldots + a_{n-1}\mathbf{f}' + a_n\mathbf{f} = b. \qquad [6\text{-}8]$$

Hemos hecho observar ya que esta es una ecuación que tiene sentido y que se verifica en cuanto un elemento de función (f, Ω) de **f** satisfaga la correspondiente ecuación con **f** reemplazada por f. Un elemento de función con esta propiedad se llamará una *solución local.*

El lector que está familiarizado con el caso real esperará que la ecuación [6-8] tenga n soluciones linealmente independientes. Esto es cierto en lo que se refiere a soluciones locales, pero debemos estar preparados para obtener que soluciones locales distintas pueden ser elementos de la misma función analítica general. En otras palabras: en el caso complejo, parte del problema es hallar hasta qué extremo las soluciones locales son prolongaciones analíticas unas de otras.

La ecuación [6-7] es *homogénea* si $b(z)$ es idénticamente cero. Este es, sin duda, el caso más importante, y será el único que tratemos. Podemos suponer, además, que los coeficientes $a_k(z)$ no tienen ceros comunes; en efecto, si z_0 fuera un cero común podríamos dividir todos los coeficientes por $z - z_0$ y las soluciones seguirían siendo las mismas. En realidad, si no tenemos inconveniente en considerar coeficientes meromorfos, podemos dividir [6-7] por $a_0(z)$ desde el principio. Recíprocamente, si se da una ecuación con coeficientes meromorfos, se puede escribir cada coeficiente como cociente de dos funciones enteras, con lo cual se obtiene una ecuación equivalente con coeficientes enteros, después de haber multiplicado por el denominador común. Por tanto, el hecho de que los coeficientes tengan o no polos no modifica el problema.

En el caso correspondiente a $n = 1$, [6-7] posee la solución explícita

$$w = e^{-\int \frac{a_1(z)}{a_0(z)}dz}$$

de donde el único problema es determinar el carácter multiforme de la integral, una cuestión que ya ha sido tratada. Por otra parte, se tiene que el caso correspondiente a $n = 2$ posee todas las características propias del caso general. Por esta razón es suficiente considerar ecuaciones diferenciales lineales homogéneas de segundo orden.

1. *Puntos ordinarios.*—Se dice que un punto z_0 es un *punto ordinario* de la ecuación diferencial

$$a_0(z)w'' + a_1(z)w' + a_2(z)w = 0 \qquad [6\text{-}9]$$

si (y solo si) $a_0(z_0) \neq 0$. El teorema central que hemos de demostrar es el siguiente:

TEOREMA 5. *Si z_0 es un punto ordinario de la ecuación* [6-9], *existe una solución local (f, Ω), $z_0 \in \Omega$, con valores prescritos arbitrarios $f(z_0) = b_0$ y $f'(z_0) = b_1$. La rama (f, z_0) está determinada de manera única.*

Preferimos escribir [6-9] en la forma

$$w'' = p(z)w' + q(z)w, \qquad [6\text{-}10]$$

siendo $p(z) = -a_1/a_0$, $q(z) = -a_2/a_0$. La hipótesis significa que $p(z)$ y $q(z)$ son analíticas en un entorno de z_0; por razones de conveniencia tomamos $z_0 = 0$. Sean

$$\begin{aligned} p(z) &= p_0 + p_1 z + \ldots + p_n z^n + \ldots \\ q(z) &= q_0 + q_1 z + \ldots + q_n z^n + \ldots \end{aligned} \qquad [6\text{-}11]$$

los desarrollos de Taylor de $p(z)$ y de $q(z)$, respectivamente.

Con objeto de resolver [6-10] utilizaremos el método de los coeficientes indeterminados. Si el teorema es cierto, la solución $w = f(z)$ debe tener un desarrollo de Taylor

$$f(z) = b_0 + b_1 z + \ldots + b_n z^n + \ldots \qquad [6\text{-}12]$$

cuyos coeficientes satisfagan las condiciones

$$\begin{aligned} &2b_2 = b_1 p_0 + b_0 q_0 \\ &6b_3 = 2b_2 p_0 + b_1 p_1 + b_1 q_0 + b_0 q_1 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &n(n-1)b_n = (n-1)b_{n-1}p_0 + (n-2)b_{n-2}p_1 + \ldots + b_1 p_{n-2} + \\ &\qquad + b_{n-2}q_0 + b_{n-3}q_1 + \ldots + b_0 q_{n-2} \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \qquad [6\text{-}13]$$

Esto prueba ya la unicidad. Todo lo que queda por demostrar es que las ecuaciones [6-13] conducen a una serie de potencias [6-12]

con radio de convergencia positivo. Se seguirá entonces, mediante operaciones lícitas de derivación término a término, multiplicación y reordenación, que [6-12] es una solución de la ecuación con los valores iniciales de f y f' deseados.

Puesto que las series [6-11] tienen radios de convergencia positivos, existen, por las desigualdades de Cauchy, constantes M_0 y $r_0 > 0$ tales que

$$\begin{aligned} |p_n| &\leqq M_0 r_0^{-n} \\ |q_n| &\leqq M_0 r_0^{-n}. \end{aligned} \qquad [6\text{-}14]$$

A fin de probar que [6-12] tiene igualmente un radio de convergencia positivo basta demostrar que se verifican desigualdades

$$|b_n| \leqq M r^{-n} \qquad [6\text{-}15]$$

análogas para una elección adecuada de M y r.

La idea natural es demostrar [6-15] por inducción con respecto a n. En primer lugar, [6-15] debe verificarse para $n=0$ y $n=1$; esto conduce a las condiciones previas $|b_0| \leqq M$, $|b_1| \leqq M r^{-1}$ que se verifican para un M suficientemente grande y un r suficientemente pequeño. Supongamos que [6-15] es válida para todos los subíndices $< n$. Con objeto de simplificar los cálculos elegimos $r < r_0$; entonces la ecuación general [6-13] conduce inmediatamente a la acotación

$$\begin{aligned} n(n-1)|b_n| &\leqq M M_0[(1+2+\ldots+(n-1))r^{1-n} + (n-1)r^{2-n}] = \\ &= M M_0\left[\frac{n(n-1)}{2} r + (n-1)r^2\right] r^{-n}. \end{aligned}$$

Tenemos así

$$|b_n| \leqq M M_0\left(\frac{r}{2} + \frac{r^2}{n}\right) r^{-n} \leqq M M_0\left(\frac{r}{2} + r^2\right) r^{-n},$$

de donde se sigue [6-15], con tal que $M_0(r/2 + r^2) \leqq 1$. Es evidente que esta condición y las anteriores se cumplen para todo r suficientemente pequeño. Con esto queda completa la demostración.

Existen, en particular, soluciones locales $f_0(z)$ y $f_1(z)$ que satisfacen las condiciones $f_0(z_0) = 1$, $f'_0(z_0) = 0$ y $f_1(z_0) = 0$, $f'_1(z_0) = 1$. Como consecuencia de la unicidad, la solución con valores iniciales b_0 y b_1 debe ser $f(z) = b_0 f_0(z) + b_1 f_1(z)$. Por tanto, toda solución local es una combinación lineal de $f_0(z)$ y $f_1(z)$. Además, las soluciones $f_0(z)$ y $f_1(z)$ son linealmente independientes, pues de la identidad $b_0f_0(z)+b_1f_1(z)=0$ se obtiene primero $b_0=0$, haciendo $z=z_0$, y después $b_1=0$ como consecuencia de que $f_1(z)$ no puede ser idénticamente nula.

EJERCICIOS

1. Hállense los desarrollos en series de potencias en un entorno del origen de dos soluciones linealmente independientes de $w''=zw$.

2. Se definen los polinomios de Hermite mediante $H_n(z)=(-1)^n e^{z^2}(d^n/dz^n)(e^{-z^2})$. Pruébese que $H_n(z)$ es una solución de la ecuación $w''-2zw'+2nw=0$.

2. *Puntos singulares regulares.*—A cualquier punto z_0, tal que $a_0(z_0)=0$, se le llama *punto singular* de la ecuación [6-9]. Si se escribe la ecuación en la forma [6-10], el supuesto significa que o $p(z)$ o $q(z)$ tienen un polo en z_0, pues seguimos excluyendo el caso de ceros comunes a todos los coeficientes de [6-9].

Existen distintas clases de puntos singulares. Empecemos con un estudio previo del caso más sencillo, que ocurre cuando $a_0(z)$ tiene un cero simple. Bajo esta hipótesis, las funciones $p(z)$ y $q(z)$ poseen a lo sumo polos simples, y si elegimos $z_0=0$, los desarrollos de Laurent son de la forma

$$p(z)=\frac{p_{-1}}{z}+p_0+p_1z+\dots$$

$$q(z)=\frac{q_{-1}}{z}+q_0+q_1z+\dots .$$

Si sustituimos esta vez

$$w=b_0+b_1z+b_2z^2+\dots$$

en [6-10], la comparación de coeficientes da

$$\begin{aligned}
&-p_{-1}b_1=b_0q_1\\
&2(1-p_{-1})b_2=b_1p_0+\ +b_1q_{-1}+b_0q_0\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&n(n-1-p_{-1})b_n=(n-1)b_{n-1}p_0+(n-2)b_{n-2}p_1+\dots+\\
&\qquad\qquad +b_1p_{n-2}+b_{n-1}q_{-1}+b_{n-2}q_0+\dots+b_0q_{n-2}\\
&\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}\qquad [6\text{-}16]$$

Este sistema de relaciones es esencialmente diferente del [6-13]. En primer lugar, solo se puede escoger arbitrariamente a b_0 y, por tanto, el método da a lo sumo una solución linealmente independiente. En segundo lugar, si p_{-1} es cero o un entero positivo, el sistema [6-16] o no tiene solución o una de las b_n puede ser elegida arbitrariamente.

Suponiendo que p_{-1} no es ni cero ni un entero positivo, probaremos que la serie de potencias resultante tiene un radio de convergencia positivo. Utilizaremos como antes las acotaciones [6-14], eligiendo $M\geqq|b_0|$, y supondremos que se verifica [6-15] para subíndices $<n$. Bajo la hipótesis auxiliar de que $r\leqq r_0$, se obtiene

$$n|n-1-p_{n-1}|\cdot|b_n|\leqq Mr^{-n}\left\{M_0\left[\frac{n(n-1)}{2}r+(n-1)r^2\right]+|q_{-1}|r\right\}.$$

Se verificará una desigualdad de la forma

$$|b_n|\leqq Mr^{-n}(Ar+Br^2)$$

para todo n, siempre y cuando $(n-1)/|n-1-p_{-1}|$ esté acotado. Para r suficientemente pequeño, esta acotación es más exigente que la [6-15], de donde se sigue la convergencia.

Como ya se indicó, este resultado es de carácter previo. Nuestro verdadero objetivo es el de resolver [6-10] cuando exista una *singularidad regular* en z_0. Se utiliza esta terminología para indicar que $p(z)$ tiene a lo sumo un polo simple en z_0 y $q(z)$ a lo sumo un polo doble en dicho punto.

En estas circunstancias, resulta que existen soluciones de la forma $w=z^\alpha g(z)$, siendo $g(z)$ analítica y $\neq 0$ en $z_0=0$. Hagamos esta sustitución en [6-10] y hallaremos, tras un breve cálculo, que $g(z)$ debe satisfacer la ecuación diferencial

$$g''=\left(\frac{p-2\alpha}{z}\right)g'+\left(q+\frac{\alpha p}{z}-\frac{\alpha(\alpha-1)}{z^2}\right)g. \qquad [6\text{-}17]$$

Para α arbitrario, esta es del mismo tipo que la ecuación original y no se ha ganado nada. Podemos, sin embargo, elegir α de tal forma que el coeficiente de g tenga únicamente un polo simple. Si $q(z)$ tiene el desarrollo

$$q(z)=\frac{q_{-2}}{z^2}+\dots,$$

este será el caso si α satisface la ecuación cuadrática

$$\alpha(\alpha-1)-p_{-1}\alpha-q_{-2}=0, \qquad [6\text{-}18]$$

que se denomina *ecuación indicial.* Para un α tal, nuestro resultado previo muestra que [6-10] tiene una solución de la forma $z^{\alpha}g(z)$, $g(0)\neq 0$, con tal que $p_1-2\alpha$ no sea un entero no negativo.

Denotemos por α_1 y α_2 a las raíces de [6-18]. Entonces $\alpha_1+\alpha_2=$ $=p_{-1}+1$ o $\alpha_2-\alpha_1=p_{-1}-2\alpha_1+1$. Por tanto, α_1 es excepcional si (y solo si) $\alpha_2-\alpha_1$ es un entero positivo. De donde, si las raíces de la ecuación indicial no difieren en un entero, obtenemos dos soluciones: $z^{\alpha_1}g_1(z)$ y $z^{\alpha_2}g_2(z)$, que, como es obvio, son linealmente independientes. Si las raíces son iguales o difieren en un entero, el método da una sola solución.

TEOREMA 6. *Si z_0 es un punto singular regular de la ecuación [6-9], existen soluciones linealmente independientes de la forma $(z-z_0)^{\alpha_1}g_1(z)$ y $(z-z_0)^{\alpha_2}g_2(z)$ con $g_1(0)$, $g_2(0)\neq 0$ correspondientes a las raíces de la ecuación indicial, con tal que $\alpha_2-\alpha_1$ no sea un entero. En el caso de una diferencia entera $\alpha_2-\alpha_1\geqq 0$, todavía puede afirmarse la existencia de una solución correspondiente a α_2.*

Si se conoce una solución, no es difícil hallar otra linealmente independiente de la primera. Los métodos que conducen a una segunda solución son más bien propios de textos sobre ecuaciones diferenciales. Tampoco es posible tratar en este libro el caso de singularidades irregulares.

EJERCICIOS

1. Pruébese que la ecuación $(1-z^2)w''-2zw'+n(n+1)w=0$, donde n es un entero no negativo, admite como soluciones los polinomios de Legendre

$$P_n(z)=\left(\frac{1}{2^n n!}\right)\frac{d^n}{dz^n}(z^2-1)^n.$$

2. Determínense dos soluciones linealmente independientes de la ecuación

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

en las proximidades de 0 y de -1, respectivamente.

3. Demuéstrese que la ecuación de Bessel $zw''+w'+zw=0$ tiene una solución que es una función entera. Determínese su desarrollo en serie de potencias.

3. *Soluciones en el punto del infinito.*—Si $a_0(z)$, $a_1(z)$, $a_2(z)$ son polinomios, parece natural preguntarse cómo se comportarán las soluciones en el entorno de ∞. La forma más conveniente de tratar esta cuestión es la de efectuar el cambio de variables $z=1/Z$. Puesto que

$$\frac{dw}{dz}=-Z^2\frac{dw}{dZ}$$

$$\frac{d^2w}{dz^2}=2Z^3\frac{dw}{dZ}+Z^4\frac{d^2w}{dZ^2},$$

la ecuación [6-10] toma la forma

$$\frac{d^2w}{dZ^2}=-\left[2Z^{-1}+Z^{-2}p\left(\frac{1}{Z}\right)\right]\frac{dw}{dZ}+Z^{-4}q\left(\frac{1}{Z}\right)w. \qquad [6\text{-}19]$$

Decimos, naturalmente, que ∞ es un punto ordinario o una singularidad regular de la ecuación [6-10] si el punto $Z=0$ tiene el mismo carácter para [6-19]. Así, pues, ∞ es un punto ordinario si los coeficientes de [6-10] tienen una singularidad evitable en $Z=0$; esto es lo mismo, por definición, que decir que $-[2z+z^2p(z)]$ y $z^4q(z)$ tienen singularidades evitables en ∞. De manera análoga, ∞ es una singularidad regular si tales funciones tienen, respectivamente, un polo simple y un polo doble, a lo sumo, en ∞.

Es interesante determinar las ecuaciones con el menor número posible de singularidades. Si ∞ ha de ser un punto ordinario, $q(z)$ debe tener al menos cuatro polos, a no ser que se anule idénticamente. En el último caso, $p(z)$ puede tener como mínimo un polo, y si el polo está en el origen, se debe tener $p(z)=-2/z$. La ecuación correspondiente

$$\frac{d^2w}{dz^2}=-\frac{2}{z}\frac{dw}{dz}$$

posee la solución general $w=az^{-1}+b$.

Si $q(z)$ no es idénticamente cero, como mínimo ha de haber dos singularidades regulares. Evidentemente, lo más sencillo es situar las singularidades en 0 e ∞, y por esta razón volvemos inmediatamente al caso en que ∞ es una singularidad regular. Si solo ha de haber una singularidad finita situada en el origen, debemos tener $p(z)=A/z$, $q(z)=B/z^2$. Eligiendo de otra manera las constantes, puede escribirse la ecuación en la forma

$$z^2w''-(\alpha+\beta-1)zw'+\alpha\beta w=0. \qquad [6\text{-}20]$$

Esta tiene las soluciones $w=z^{\alpha}$ y $w=z^{\beta}$, siendo α y β obviamente las raíces de la ecuación indicial. Si $\alpha=\beta$, debe existir otra solución. Para hallarla escribimos [6-20] en la forma simbólica

$$\left(z\frac{d}{dz}-\alpha\right)^2 w=0$$

y hacemos la sustitución $w=z^{\alpha}W$. Se obtiene

$$\left(z\frac{d}{dz}-\alpha\right) z^{\alpha}W=z^{\alpha}\cdot z\frac{dW}{dz}$$

$$\left(z\frac{d}{dz}-\alpha\right)^2 z^{\alpha}W=z^{\alpha}\cdot z\frac{d}{dz}\left(z\frac{dW}{dz}\right).$$

La ecuación $\left(z\dfrac{d}{dz}\right)^2 W=0$ tiene la solución obvia $W=\log z$, por lo que la solución de [6-20] buscada es $w=z^{\alpha}\log z$.

4. *La ecuación diferencial hipergeométrica.*—Acabamos de ver que las ecuaciones diferenciales con una o dos singularidades regulares poseen soluciones triviales. Es al introducir una tercera singularidad cuando obtenemos una nueva e interesante clase de funciones analíticas.

Es completamente evidente que una transformación lineal de la variable transforma una ecuación diferencial lineal de segundo orden en otra del mismo tipo, y que el carácter de las singularidades sigue siendo el mismo. Podemos, por consiguiente, situar las tres singularidades en puntos escogidos de antemano, siendo la elección más sencilla la de 0, 1 e ∞.

Si la ecuación

$$w''=p(z)\,w'+q(z)$$

ha de tener singularidades regulares finitas únicamente en 0 y 1, debemos tener

$$p(z)=\frac{A}{z}+\frac{B}{z-1}+P(z)$$

$$q(z)=\frac{C}{z^2}+\frac{D}{z}+\frac{E}{(z-1)^2}+\frac{F}{z-1}+Q(z),$$

siendo $P(z)$ y $Q(z)$ polinomios. Con objeto de hacer que la singularidad en ∞ sea regular, $2z+z^2p(z)$ debe tener a lo sumo un polo simple en ∞, y $z^4q(z)$, cuando más, uno doble. A la vista de estas condiciones, $P(z)$ y $Q(z)$ tienen que ser idénticamente nulos, debiendo verificarse la relación $D+F=0$. Evidentemente, estas son las únicas condiciones, y podemos escribir nuevamente las expresiones de $p(z)$ y $q(z)$ en la forma

$$p(z)=\frac{A}{z}+\frac{B}{z-1}$$

$$q(z)=\frac{C}{z^2}-\frac{D}{z(z-1)}+\frac{E}{(z-1)^2}$$

La ecuación indicial en el origen se escribe en la forma

$$\alpha(\alpha-1)=A\alpha+C.$$

Así que si denotamos sus raíces por α_1, α_2, obtenemos $A=\alpha_1+\alpha_2-1$, $C=-\alpha_1\alpha_2$. Análogamente, $B=\beta_1+\beta_2-1$ y $E=-\beta_1\beta_2$, siendo β_1, β_2 las raíces de la ecuación indicial en el punto 1. Con objeto de escribir la ecuación indicial en infinito, observemos que los coeficientes directores de $-[2z+z^2p(z)]$ y $z^4q(z)$ son $-(2+A+B)$ y $C-D+E$, respectivamente. Por tanto, las raíces γ_1, γ_2 satisfacen la relación $\gamma_1+\gamma_2=-A-B-1$ y $\gamma_1\gamma_2=-C+D-E$. Obtenemos la relación

$$\alpha_1+\alpha_2+\beta_1+\beta_2+\gamma_1+\gamma_2=1, \qquad [6\text{-}21]$$

y hallamos que la ecuación puede escribirse en la forma

$$w''+\left(\frac{1-\alpha_1-\alpha_2}{z}+\frac{1-\beta_1-\beta_2}{z-1}\right)w'+ \qquad [6\text{-}22]$$

$$+\left(\frac{\alpha_1\alpha_2}{z^2}-\frac{\alpha_1\alpha_2+\beta_1\beta_2-\gamma_1\gamma_2}{z(z-1)}+\frac{\beta_1\beta_2}{(z-1)^2}\right)w=0.$$

Con objeto de evitar los casos excepcionales supondremos ahora que ninguna de las diferencias $\alpha_2-\alpha_1$, $\beta_2-\beta_1$, $\gamma_2-\gamma_1$ es un entero. El paso siguiente consiste en simplificar la ecuación [6-22]. Ya se ha visto en la sección 6-3, 2 que la sustitución $w=z^{\alpha}g(z)$ determina para $g(z)$ una ecuación diferencial análoga, a saber: la ecuación [6-17]. Puesto que la ecuación original posee soluciones de la forma $w=z^{\alpha_1}g_1(z)$ y $w=z^{\alpha_2}g_2(z)$, se concluye que la ecuación transformada [6-17] debe tener soluciones de la forma $g(z)=$

$=z^{\alpha_1-\alpha}g_1(z)$ y $g(z)=z^{\alpha_2-\alpha}g_2(z)$. Por tanto, la ecuación indicial de [6-17] tiene las raíces $\alpha_1-\alpha$, $\alpha_2-\alpha$, como puede comprobarse también mediante cálculo. Las raíces que corresponden a la singularidad en ∞ cambian simultáneamente desde γ_1, γ_2 a $\gamma_1+\alpha$, $\gamma_2+\alpha$. De la misma manera exactamente, podemos separar un factor $(z-1)^\beta$ y hallar que la ecuación resultante tiene exponentes, disminuido en β en el punto 1 y aumentado en β en ∞. La elección natural es tomar $\alpha=\alpha_1$, $\beta=\beta_1$. Entonces en la ecuación final los seis exponentes son 0, $\alpha_2-\alpha_1$, 0, $\beta_2-\beta_1$, $\gamma_1+\alpha_1+\beta_1$, $\gamma_2+\alpha_1+\beta_1$, respectivamente. Con objeto de ajustarse a convenios sancionados por el tiempo, escribiremos $a=\alpha_1+\beta_1+\gamma_1$, $b=\alpha_1+\beta_1+\gamma_2$, $c=1+\alpha_1-\alpha_2$. Como consecuencia de la relación [6-21] obtenemos $c-a-b=\beta_2-\beta_1$. De acuerdo con esto, la nueva ecuación diferencial será de la forma

$$w''+\left(\frac{c}{z}+\frac{1-c+a+b}{z-1}\right)w'+\frac{ab}{z(z-1)}\,w=0$$

o, después de simplificada,

$$z(1-z)w''+[c-(a+b+1)z]w'-abw=0. \qquad [6\text{-}23]$$

A esta ecuación se le llama *ecuación diferencial hipergeométrica*, y hemos probado que las soluciones de [6-22] son iguales a las soluciones de [6-23] multiplicadas por $z^{\alpha_1}(z-1)^{\beta_1}$. Se supone que ninguna de las diferencias $c-1$, $a-b$, $a+b-c$ es un entero.

De acuerdo con la teoría, la ecuación [6-23] tiene una solución de la forma $w=\sum_{n=0}^{\infty}A_nz^n$. Si se sustituye esta serie de potencias en [6-23], se halla con escaso cálculo que los coeficientes deben satisfacer las siguientes relaciones de recurrencia:

$$(n+1)(n+c)A_{n+1}=(n+a)(n+b)A_n.$$

La forma extremadamente sencilla de esta relación hace posible escribir de manera explícita la solución. Eligiendo $A_0=1$, hallamos que la ecuación hipergeométrica se satisface por la función

$$F(a,b,c,z)=1+\frac{a\cdot b}{1\cdot c}z+\frac{a(a+1)\cdot b(b+1)}{1\cdot 2\cdot c(c+1)}z^2+\frac{a(a+1)(a+2)\cdot b(b+1)(b+2)}{1\cdot 2\cdot 3\cdot c(c+1)(c+2)}z^3+\dots,$$

conocida como *función hipergeométrica*. Está definida siempre y cuando c no sea cero o un entero negativo.

Mediante cálculo, puede hallarse con toda facilidad el radio de convergencia de la serie hipergeométrica; pero es más instructivo utilizar el razonamiento puro. En primer lugar, sabemos que $F(a, b, c, z)$ puede prolongarse analíticamente a lo largo de cualquier camino que no pase por el punto 1 y no vuelva al origen. Por tanto, puede definirse una rama uniforme de $F(a,b,c,z)$ en el disco unidad $|z|<1$ (como consecuencia de ser el disco simplemente conexo), y se sigue que el radio de convergencia es al menos igual a 1. Si fuera mayor que 1, $F(a,b,c,z)$ sería una función entera. En las proximidades del infinito debe ser una combinación lineal de las soluciones $z^{-a}g_1(z)$, $z^{-b}g_2(z)$, las cuales sabemos que existen en un entorno de ∞. Pero es evidente que una combinación lineal puede ser uniforme únicamente si a o b es un entero. Por hipótesis, si a es un entero, b no lo es, y $F(a,b,c,z)$ es un múltiplo de $z^{-a}g_1(z)$. Por el teorema de Liouville, si a fuera positivo, $F(a,b,c,z)$ se anularía idénticamente, lo que no ocurre. El único caso en que el radio de convergencia es infinito es, por tanto, cuando a (o b) es un entero negativo o cero, y entonces la serie hipergeométrica se reduce trivialmente a un polinomio.

En un entorno del origen existe también una solución de la forma $z^{1-c}g(z)$. Esta función $g(z)$ satisface una ecuación diferencial hipergeométrica, cuyos seis exponentes son $\alpha_2-\alpha_1$, 0, 0, $\beta_2-\beta_1$, $\gamma_1+\alpha_2+\beta_1$, $\gamma_2+\alpha_2+\beta_1$. Se sigue inmediatamente que podemos poner $g(z)=F(1+a-c,\ 1+b-c,\ 2-c,z)$. Hemos probado así que $F(a,b,c,z)$ y $z^{1-c}F(1+a-c,\ 1+b-c,\ 2-c,\ z)$ son dos soluciones linealmente independientes en las proximidades del origen.

Pueden determinarse las soluciones en las proximidades de 1 exactamente de la misma manera. Sin embargo, resulta más sencillo reemplazar z por $1-z$ e intercambiar las α y las β. Hallamos como resultado que las funciones $F(a,b,1+a+b-c,1-z)$ y $(1-z)^{c-a-b}F(c-b,c-a,1-a-b+c,1-z)$ son soluciones linealmente independientes en un entorno de 1. De manera análoga, se pueden hallar las soluciones en las proximidades de ∞.

Hemos demostrado que la ecuación diferencial lineal de segundo orden con tres singularidades más general puede resolverse de manera explícita mediante la función hipergeométrica. Evidentemente, también es posible, aunque algo más laborioso, determinar la estructura multiforme completa de las soluciones.

EJERCICIOS

1. Pruébese que $(1-z)^{-a}=F(\alpha,\ \beta,\ \beta,\ z)$ y que $\log 1/(1-z)=$ $=zF(1,1,2,z)$.

2. Exprésese la derivada de $F(a,b,c,z)$ como función hipergeométrica.

3. Dedúzcase la siguiente representación integral:

$$F(a,b,c,z)=\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}dt.$$

4. Si w_1 y w_2 son soluciones linealmente independientes de la ecuación diferencial $w''=pw'+qw$, pruébese que el cociente $\eta=w_2/w_1$ satisface la siguiente ecuación diferencial:

$$\frac{d}{dz}\left(\frac{\eta''}{\eta'}\right)-\frac{1}{2}\left(\frac{\eta''}{\eta'}\right)^2=-2q-\frac{1}{2}p^2+p'.$$

5. *El punto de vista de Riemann.*—Riemann fue un mantenedor entusiasta de la idea de que una función analítica, lo mismo que sus propiedades generales, puede definirse mediante sus singularidades tan bien o quizá mejor que por una expresión explícita. Un ejemplo trivial es la determinación de una función racional por las partes singulares relativas a sus polos.

Probaremos, siguiendo las ideas de Riemann, que las soluciones de una ecuación diferencial hipergeométrica pueden caracterizarse por propiedades de esta naturaleza. Consideraremos en lo que sigue una colección **F** de elementos de función (f,Ω) con ciertas propiedades características que enumeramos a continuación:

1.ª La colección **F** es *completa* en el sentido de que contiene todas las prolongaciones analíticas de cualquier $(f,\Omega)\in\mathbf{F}$. *No* se requiere que dos elementos de función de **F** cualesquiera sean prolongaciones analíticas uno de otro, por lo que **F** puede consistir en varias funciones analíticas completas.

2.ª La colección es *lineal*. Esto significa que $(f_1,\Omega)\in\mathbf{F}$ y $(f_2,\Omega)\in\mathbf{F}$ implica que $(c_1f_1+c_2f_2,\Omega)\in\mathbf{F}$ para todo par de constantes c_1 y c_2. Además, cualquier terna de elementos (f_1,Ω), (f_2,Ω), $(f_3,\Omega)\in\mathbf{F}$ con el mismo Ω satisfarán una relación idéntica c_1f_1+ $+c_2f_2+c_3f_3=0$ en Ω con coeficientes constantes, no todos nulos. En otras palabras: **F** será a lo sumo *bidimensional*.

3.ª Las únicas singularidades finitas de las funciones de **F** estarán en los puntos 0 y 1; además, el punto ∞ también se cuenta

como una singularidad. Con más precisión, se exige que cualquier $(f,\Omega)\in\mathbf{F}$ pueda prolongarse a lo largo de todos los arcos en el plano finito que no pasen por los puntos 0 y 1.

4.ª En cuanto al comportamiento de los puntos singulares, suponemos que hay funciones de **F** que se comportan como potencias z^{α_1} y z^{α_2} prescritas en las proximidades de 0; como $(z-1)^{\beta_1}$ y $(z-1)^{\beta_2}$, en las proximidades de 1, y como $z^{-\gamma_1}$ y $z^{-\gamma_2}$, en las proximidades de ∞. En términos precisos, existirán ciertas funciones analíticas $g_1(z)$ y $g_2(z)$ definidas en un entorno Δ de 0 y diferentes de cero en este punto; se pueden definir, para una subregión simplemente conexa Ω de Δ que no contenga el origen, elementos de función $[z^{\alpha_1}g_1(z),\Omega]$, $[z^{\alpha_2}g_2(z),\Omega]$, y se requiere que pertenezcan a **F**. Las suposiciones correspondientes para los puntos 1 e ∞ pueden formularse de manera análoga.

El lector habrá reconocido que las soluciones de la ecuación diferencial [6-22] tienen precisamente estas propiedades, siempre y cuando ninguna de las diferencias $\alpha_2-\alpha_1$, $\beta_2-\beta_1$, $\gamma_2-\gamma_1$ sea un entero. Además se verifica la relación $\alpha_1+\alpha_2+\beta_1+\beta_2+\gamma_1+\gamma_2=1$. Hacemos ambas hipótesis y probaremos, bajo estas restricciones, que existe una (y solo una) colección **F** con las propiedades 1.ª a 4.ª. Por consiguiente, **F** será idéntica a la colección de soluciones locales de la ecuación diferencial [6-22].

Riemann denota cualquier elemento de función de **F** mediante el símbolo

$$P\left\{\begin{matrix}0 & 1 & \infty & \\ \alpha_1 & \beta_1 & \gamma_1, & z\\ \alpha_2 & \beta_2 & \gamma_2 & \end{matrix}\right\}.$$

Así, pues, P no denota una función individual; pero evidentemente esto es de poca importancia. Una vez que se establece la unicidad, se siguen inmediatamente identidades del tipo

$$P\left\{\begin{matrix}0 & 1 & \infty & \\ \alpha_1 & \beta_1 & \gamma_1, & z\\ \alpha_2 & \beta_2 & \gamma_2 & \end{matrix}\right\}=z^{\alpha}(z-1)^{\beta}P\left\{\begin{matrix}0 & 1 & \infty & \\ \alpha_1-\alpha & \beta_1-\beta & \gamma_1+\alpha+\beta, & z\\ \alpha_2-\alpha & \beta_2-\beta & \gamma_2+\alpha+\beta & \end{matrix}\right\}$$

o

$$P\left\{\begin{matrix}0 & 1 & \infty & \\ \alpha_1 & \beta_1 & \gamma_1, & z\\ \alpha_2 & \beta_2 & \gamma_2 & \end{matrix}\right\}=P\left\{\begin{matrix}0 & 1 & \infty & \\ \beta_1 & \alpha_1 & \gamma_1, & 1-z\\ \beta_2 & \alpha_2 & \gamma_2 & \end{matrix}\right\}$$

con tal que se tenga cuidado de interpretarlas debidamente. El hecho de que puedan reconocerse tan fácilmente tales relaciones, algunas de ellas bastante elaboradas, es uno de los motivos que apoyan el punto de vista de Riemann.

Con objeto de probar la unicidad, consideremos dos elementos de función linealmente independientes (f_1, Ω), $(f_2, \Omega) \in \mathbf{F}$, definidos en una región simplemente conexa Ω que no contenga ni a 0 ni a 1. Existen tales elementos de función en cualquier Ω, pues las funciones $z^{\alpha_1}g_1(z)$ y $z^{\alpha_2}g_2(z)$ son linealmente independientes en su región de definición; pueden prolongarse a lo largo de un arco que acabe en Ω y determinan elementos de función linealmente independientes. Si (f, Ω) es un tercer elemento de función de $\mathbf{F}$, las identidades

$$\begin{aligned} cf + c_1f_1 + c_2f_2 &= 0 \\ cf' + c_1f_1' + c_2f_2' &= 0 \\ cf'' + c_1f_1'' + c_2f_2'' &= 0 \end{aligned}$$

implican

$$\begin{vmatrix} f & f_1 & f_2 \\ f' & f_1' & f_2' \\ f'' & f_1'' & f_2'' \end{vmatrix} = 0.$$

Escribimos esta ecuación en la forma

$$f'' = p(z)f' + q(z)f$$

con

$$p(z) = \frac{f_1f_2'' - f_2f_1''}{f_1f_2' - f_2f_1'}, \qquad q(z) = -\frac{f_1'f_2'' - f_2'f_1''}{f_1f_2' - f_2f_1'} \tag{6-24}$$

Aquí el denominador no es idénticamente cero, pues eso significaría que f_1 y f_2 son linealmente dependientes.

Hacemos ahora la observación de que las expresiones [6-24] permanecen invariantes si se someten f_1 y f_2 a una transformación lineal no singular; es decir, si se reemplazan por $c_{11}f_1 + c_{12}f_2$, $c_{21}f_1 + c_{22}f_2$ con $c_{11}c_{22} - c_{12}c_{21} \neq 0$. Esto significa que $p(z)$ y $q(z)$ serán los mismos para cualquier elección de f_1 y f_2; por tanto, son funciones *uniformes* bien determinadas en todo el plano menos en los puntos 0 y 1.

Con objeto de determinar el comportamiento de $p(z)$ y $q(z)$ en las proximidades del origen elegimos $f_1 = z^{\alpha_1}g_1(z)$, $f_2 = z^{\alpha_2}g_2(z)$. Cálculos sencillos nos dan

$$\begin{aligned} f_1f_2' - f_2f_1' &= (\alpha_2 - \alpha_1)z^{\alpha_1+\alpha_2-1}(C + \ldots) \\ f_1f_2'' - f_2f_1'' &= (\alpha_2 - \alpha_1)(\alpha_1 + \alpha_2 - 1)z^{\alpha_1+\alpha_2-2}(C + \ldots) \\ f_1'f_2'' - f_2'f_1'' &= \alpha_1\alpha_2(\alpha_2 - \alpha_1)z^{\alpha_1+\alpha_2-3}(C + \ldots) \end{aligned}$$

donde los paréntesis representan funciones analíticas con el valor común $C = g_1(0)g_2(0)$ en el origen. En conclusión, $p(z)$ tiene un polo simple con residuo $\alpha_1 + \alpha_2 - 1$, mientras que el desarrollo de Laurent de $q(z)$ empieza con el término $-\alpha_1\alpha_2/z^2$. Se obtienen resultados análogos para los puntos 1 e ∞. Se deduce que

$$p(z) = \frac{\alpha_1 + \alpha_2 - 1}{z} + \frac{\beta_1 + \beta_2 - 1}{z - 1} + p_0(z),$$

donde $p_0(z)$ no tiene polos en 0 y 1. Por otra parte, el desarrollo de $p(z)$ en ∞ debe empezar con el término $-(\gamma_1 + \gamma_2 + 1)/z$. De acuerdo con su definición, [6-24], $p(z)$ es una derivada logarítmica. Como tal, tiene en el plano finito únicamente polos simples, con enteros positivos como residuos. En vista de la relación $(\alpha_1 + \alpha_2 - 1) + (\beta_1 + \beta_2 - 1) = -(\gamma_1 + \gamma_2 + 1)$, se sigue que $p_0(z)$ no puede tener polos en absoluto, y debe, de hecho, anularse idénticamente.

Puesto que $f_1f_2' - f_2f_1'$ es $\neq 0$, excepto en 0 y 1, se concluye que $q(z)$ es de la forma

$$q(z) = -\frac{\alpha_1\alpha_2}{z^2} - \frac{\beta_1\beta_2}{(z-1)^2} + \frac{A}{z} + \frac{B}{z-1} + q_0(z),$$

siendo $q_0(z)$ un polinomio. En infinito, el desarrollo debe empezar con $-\gamma_1\gamma_2/z^2$. Hallamos que $q_0(z)$ debe ser idénticamente cero, mientras que

$$A = -B = -(\alpha_1\alpha_2 + \beta_1\beta_2 - \gamma_1\gamma_2).$$

Resumiendo los resultados llegamos a la conclusión de que f satisface la ecuación

$$w'' + \left(\frac{1 - \alpha_1 - \alpha_2}{z} + \frac{1 - \beta_1 - \beta_2}{z - 1}\right)w' + \left(\frac{\alpha_1\alpha_2}{z^2} - \frac{\alpha_1\alpha_2 + \beta_1\beta_2 - \gamma_1\gamma_2}{z(z-1)} + \frac{\gamma_1\gamma_2}{(z-1)^2}\right)w = 0,$$

que es precisamente la ecuación [6-22].

Esto completa las demostraciones de unicidad, pues se deduce ahora que cualquier colección **F** que verifique las propiedades 1.ª a 4.ª debe ser una subcolección de la familia $\mathbf{F}_0$ de soluciones locales de [6-22]. Para cualquier región simplemente conexa Ω que no contenga 0 o 1 sabemos que hay dos elementos de función linealmente independientes (f_1, Ω), (f_2, Ω) de **F**. Todo $(f, \Omega) \in \mathbf{F}_0$ es de la forma $(c_1 f_1 + c_2 f_2, \Omega)$ y, por tanto, está contenido en **F**. Por último, si Ω no es simplemente conexo, entonces $(f, \Omega) \in \mathbf{F}_0$ es la prolongación analítica de una restricción a una subregión simplemente conexa de Ω, y puesto que la restricción pertenece a **F**, lo mismo ocurre con (f, Ω) como consecuencia de la propiedad 1.

INDICE ALFABETICO

INDICE DE AUTORES Y MATERIAS