

A. MARKUSHEVICH

TEORIA

DE LAS

FUNCIONES ANALITICAS

EDITORIAL MIR

MOSCÚ



EDITORIAL MIR

А. И. МАРКУШЕВИЧ
ТЕОРИЯ
АНАЛИТИЧЕСКИХ ФУНКЦИЙ
ТОМ I НАЧАЛА ТЕОРИИ
ИЗДАТЕЛЬСТВО «НАУКА» МОСКВА

A. MARKUSHEVICH

TEORIA
DE LAS
FUNCIONES ANALITICAS

TRADUCIDO DEL RUSO

POR

EMILIANO APARICIO BERNARDO

CANDIDATO A DOCTOR EN CIENCIAS FISICO-MATEMATICAS,

CATEDRATICO DE MATEMATICAS SUPERIORES

EDITORIAL MIR

MOSCU

1970

CDU 517.5 = 60

Impreso en la URSS
Derechos reservados

На испанском языке

PROLOGO
DEL
TRADUCTOR

Es sabido que las funciones analíticas se emplean con frecuencia en distintas disciplinas teóricas y aplicadas: en la teoría analítica de los números, ecuaciones diferenciales, geometría, hidromecánica, mecánica celeste, etc. El conocimiento de los principios fundamentales de la teoría de las funciones analíticas es una premisa indispensable en la educación matemática moderna. Actualmente esta teoría se ha desarrollado tanto que resulta difícil abarcar en un solo libro todas sus múltiples ramificaciones. Entre las diversas obras que se han escrito sobre esta disciplina, la «Teoría de las funciones analíticas» de A. I. Markushévich posee una serie de méritos particulares. En primer lugar, abarca un amplio material, debido a lo cual se edita en dos tomos. Por otra parte, es accesible para aquel que no conozca aún los elementos de la teoría de las funciones de variable compleja y desee emprender un estudio individual de la misma. Además, consideramos que es una de las obras más completas y mejores de la teoría en cuestión. Para poder comprenderla basta con tener los conocimientos que se exigen en los dos primeros cursos de la Facultad de Matemáticas de la Universidad. Puede ser útil también para los que preparan la tesis doctoral y deseen especializarse en teoría de funciones. Posiblemente, para algunos lectores esta obra servirá de punto de partida para ampliar sus investigaciones sobre algún tema.

El autor expone en este curso tanto los elementos de la teoría de las funciones analíticas como sus secciones modernas. Las demostraciones de los teoremas y, en ge-

neral, las cuestiones más difíciles, se analizan con todo detalle y de manera muy razonada.

El libro contiene numerosos ejemplos que, frecuentemente, ilustran los temas que son objeto de estudio. En muchos casos estos ejemplos desempeñan un papel decisivo, comprobando que el teorema estudiado pierde su valor en condiciones más moderadas.

Al final del segundo tomo y como apéndice, se incluye un artículo del mismo autor «Sobre la base en el espacio de las funciones analíticas». Los enunciados y teoremas expuestos en él se deben también al profesor Markushévich.

Al final de cada tomo se inserta una amplia bibliografía sobre los distintos problemas estudiados.

Esperamos que la edición española de esta obra sea bien acogida. Al final del primer tomo de la presente edición española incluimos un apéndice del traductor—«Seudopolinomios ortogonales»—en el que se hace una generalización de los polinomios de Legendre, expuestos en el libro.

Quedaremos muy agradecidos a todos aquellos que deseen darnos su opinión acerca de la traducción.

E. Aparicio Bernardo

PREFACIO
A LA
EDICION
ESPAÑOLA

En el presente curso de «Teoría de las Funciones Analíticas» en dos volúmenes se han recopilado las lecciones dadas a lo largo de varios años por el autor en la facultad Mecánico-Matemática de la Universidad Lomonósov de Moscú. Abarca con suficiente amplitud la teoría de las funciones de una variable compleja, incluyendo las transformaciones conformes, interpolación de las funciones por polinomios, elementos de la teoría de las funciones armónicas y subarmónicas, fundamentos de la teoría de las funciones enteras y meromorfas, el concepto de superficie de Riemann y prolongación analítica.

El lector que desee estudiar los importantes problemas de la teoría moderna de las funciones analíticas de varias variables complejas deberá consultar otros libros.

El autor espera que su obra será útil para los que deseen estudiar una de las partes más interesantes del análisis clásico y de la teoría de las funciones.

Es para mí sumamente grato que a las ediciones en ruso, inglés y chino venga ahora a sumarse esta edición en uno de los idiomas universales, como es el español.

Aprovecho la ocasión para expresar mi sincero agradecimiento al traductor, E. Aparicio, por su esmerada y competente traducción.

A. Markushévich

INDICE

CAPITULO PRIMERO

CONCEPTOS FUNDAMENTALES

§ 1.	El objeto de la teoría	11
§ 2.	Los números complejos	16
§ 3.	Conjuntos y funciones. Teoría de límites. Las funciones continuas	24
§ 4.	Conexividad de los conjuntos. Curvas y recintos	51
§ 5.	El infinito. Proyección estereográfica y plano ampliado	68

CAPITULO SEGUNDO

LA DERIVABILIDAD Y SU SIGNIFICADO GEOMETRICO. LAS FUNCIONES ELEMENTALES

§ 1.	La derivada. Condiciones de D'Alembert — Euler	85
§ 2.	Significado geométrico de la derivada. Transformación conforme	96
§ 3.	Polinomios. Función exponencial. Seno y coseno	103
§ 4.	Funciones racionales. Función homográfica. Geometría de Lobachevski. Funciones trigonométricas	128
§ 5.	Funciones multiformes elementales	175

CAPITULO TERCERO

INTEGRALES Y SERIES DE POTENCIAS

§ 1.	Curvas rectificables. Integrales	206
§ 2.	Teorema integral de Cauchy	216
§ 3.	Integral de Cauchy. Fórmulas de Y. Sojotski	250
§ 4.	Series de funciones y productos infinitos	271
§ 5.	Series de potencias. Relación con las series de Fourier. Desarrollo de una función analítica en serie de potencias	292
§ 6.	Unicidad. A-puntos de una función analítica. Principio del módulo máximo. Puntos singulares del elemento de una función analítica	315
§ 7.	Métodos de desarrollo de las funciones en series de potencias. Comportamiento de la serie de potencias en la frontera del círculo de convergencia	347

CAPITULO CUARTO

DIVERSAS SERIES. RESIDUOS. FUNCIONES INVERSAS
E IMPLICITAS

§ 1. Principio de compacidad	380
§ 2. Serie de Laurent. Series de Dirichlet. Teorema de Runge	392
§ 3. Puntos singulares aislados. Residuos. Principio del argumento	416
§ 4. Aplicación de la teoría de los residuos al desarrollo de las funciones en series. Interpolación	450
§ 5. Funciones inversas e implícitas	476
Apéndice del traductor	499
Bibliografía	502
Indice alfabético	507

CAPITULO
PRIMERO

CONCEPTOS
FUNDAMENTALES

§ 1. EL OBJETO DE LA TEORÍA

1.1. La disciplina matemática que se expone en este libro lleva dos denominaciones: teoría de las funciones analíticas y teoría de las funciones de variable compleja. Cada una de éstas está justificada.

Una función de variable real, definida en cierto intervalo (a, b) se llama analítica en un punto x_0 de este intervalo, si en un entorno de este punto la función se puede expresar en forma de la suma de una serie de potencias convergente, dispuesta según las potencias de $x - x_0$:

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + \dots$$

Una función que es analítica en cada uno de los puntos del intervalo se llama analítica en este intervalo. Todas las funciones elementales que se estudian en el análisis son analíticas en toda la región de su definición, a excepción, posiblemente, de algunos puntos. Por ejemplo, un polinomio $P_m(x) = a_0 + a_1x + \dots + a_mx^m$, la función exponencial e^x , las funciones trigonométricas ($\operatorname{sen} x$ y $\operatorname{cos} x$), son analíticas en todos los puntos, puesto que para cualquier x_0 subsisten los desarrollos:

$$P_m(x) = P_m(x_0) + \frac{P'_m(x_0)}{1!}(x - x_0) + \dots + \frac{P_m^{(m)}(x_0)}{m!}(x - x_0)^m,$$

$$e^x = e^{x_0} + \frac{e^{x_0}}{1!}(x - x_0) + \dots + \frac{e^{x_0}}{n!}(x - x_0)^n + \dots,$$

$$\operatorname{sen} x = \operatorname{sen} x_0 + \frac{\operatorname{cos} x_0}{1!}(x - x_0) + \dots + \frac{\operatorname{sen}\left(x_0 + n \frac{\pi}{2}\right)}{n!}(x - x_0)^n + \dots,$$

$$\operatorname{cos} x = \operatorname{cos} x_0 - \frac{\operatorname{sen} x_0}{1!}(x - x_0) + \dots + \frac{\operatorname{cos}\left(x_0 + n \frac{\pi}{2}\right)}{n!}(x - x_0)^n + \dots;$$

la función $\frac{1}{a-x}$ es analítica para $x \neq a$, puesto que

$$\begin{aligned} \frac{1}{a-x} &= \frac{1}{a-x_0} \cdot \frac{1}{1 - \frac{x-x_0}{a-x_0}} = \\ &= \frac{1}{a-x_0} - \frac{1}{(a-x_0)^2} (x-x_0) + \dots + \frac{1}{(a-x_0)^{n+1}} (x-x_0)^n + \dots \end{aligned}$$

si $x_0 \neq a$ y $\left| \frac{x-x_0}{a-x_0} \right| < 1$; la función $\sqrt[n]{x}$ es analítica para $x \neq 0$, puesto que

$$\begin{aligned} \sqrt[n]{x} &= x_0^{\frac{1}{n}} \left(1 + \frac{x-x_0}{x_0} \right)^{\frac{1}{n}} = \\ &= x_0^{\frac{1}{n}} + \frac{1}{\frac{2}{x_0^{\frac{2}{n}}}} (x-x_0) + \dots + \frac{\frac{1}{n} \left(\frac{1}{n} - 1 \right) \dots \left(\frac{1}{n} - n + 1 \right)}{n! \cdot x_0^{n - \frac{1}{n}}} \times \\ &\quad \times (x-x_0)^n + \dots \end{aligned}$$

si $x_0 \neq 0$ y $\left| \frac{x-x_0}{x_0} \right| < 1$; la función $\ln x$ es analítica en todo el campo de su definición, puesto que

$$\begin{aligned} \ln x &= \ln x_0 + \ln \left(1 + \frac{x-x_0}{x_0} \right) = \ln x_0 + \frac{1}{x_0} (x-x_0) - \\ &\quad - \frac{1}{2x_0^2} (x-x_0)^2 + \dots + (-1)^{n-1} \frac{1}{nx_0^n} (x-x_0)^n + \dots \end{aligned}$$

si $x_0 > 0$ y $\left| \frac{x-x_0}{x_0} \right| < 1$, etc.

Como el resultado de las principales operaciones algebraicas y analíticas (suma, resta, multiplicación, división, derivación e integración), efectuadas con las series de potencias, se expresa también, generalmente ^{*}), por una serie de potencias convergente, es posible hacerse ahora una idea de la amplitud e importancia de la clase de las funciones analíticas. En realidad, su importancia es todavía mayor puesto que son también analíticas, en los correspondientes intervalos, por ejemplo, las funciones analíticas de funciones analíticas, las funciones inversas a las analíticas, las funciones que satisfacen a una ecuación de la forma

$$P_0(x) + P_1(x) f(x) + \dots + P_n(x) [f(x)]^n = 0$$

^{*}) En el caso de la división, son excepcionales solamente los puntos aislados en los que se anula el divisor.

(ecuación algebraica de grado n con respecto a $f(x)$) o a una ecuación de la forma

$$p_0(x)f(x) + p_1(x)\frac{df(x)}{dx} + \dots + p_n(x)\frac{d^n f(x)}{dx^n} = q(x)$$

(ecuación diferencial lineal de orden n con respecto a $f(x)$), donde $p_0(x), \dots, p_n(x), q(x)$ son funciones analíticas.

Por lo tanto, no es de extrañar que todas las clases más importantes de funciones que aparecen en el análisis clásico y en sus aplicaciones a los problemas de mecánica y física, sean analíticas a excepción de algunos puntos singulares de estas funciones.

De aquí se deduce la extraordinaria importancia que tiene un estudio especial de las propiedades generales de las funciones analíticas.

1.2. A pesar de lo amplia que es la clase de funciones analíticas, ésta forma solamente una parte regular de la clase de funciones infinitamente derivables (o sea, de las funciones que poseen derivadas de cualquier orden). Aquí demostraremos la siguiente proposición: *una función $f(x)$, definida en un entorno de un punto x_0 , es analítica en este punto cuando, y sólo cuando, se cumplen las condiciones:*

1) *ésta es indefinidamente derivable en cierto entorno de este punto,*
 2) *existen unos números positivos δ y M tales, que para cualquier x del intervalo $(x_0 - \delta, x_0 + \delta)$ y para cualquier natural k , se verifican las desigualdades:*

$$|f^{(k)}(x)| < M \frac{k!}{\delta^k}. \quad (1.2:1)$$

Demostración. Las condiciones son necesarias. En efecto, si $f(x)$ es analítica en el punto x_0 , en cierto intervalo $(x_0 - \rho, x_0 + \rho)$ se expresa por una serie de potencias

$$f(x) = a_0 + a_1(x - x_0) + \dots + a_n(x - x_0)^n + \dots, \quad (1.2:2)$$

que, como es sabido por el análisis, admite derivación término a término la cantidad de veces que se desee, de modo que existen derivadas de cualquier orden k , expresables también por series de potencias:

$$f^{(k)}(x) = k! a_k + \frac{(k+1)! a_{k+1}}{1!} (x - x_0) + \frac{(k+2)! a_{k+2}}{2!} (x - x_0)^2 + \dots \quad (1.2:3)$$

($|x - x_0| < \rho$). Fijemos δ de manera que sea: $0 < 2\delta < \rho$. Entonces, la serie (1.2:2) convergerá para $x = x_0 + 2\delta$; por consiguiente, $\lim_{n \rightarrow \infty} a_n (2\delta)^n = 0$, de donde se deduce que la sucesión $\{a_n (2\delta)^n\}$ es

acotada, o sea,

$$|a_n (2\delta)^n| < M' \quad (n=0, 1, 2, \dots). \quad (1.2:4)$$

Acotemos ahora $|f^{(h)}(x)|$ en el intervalo $(x_0 - \delta, x_0 + \delta)$; aplicando (1.2:3) y las desigualdades (1.2:4), obtenemos:

$$\begin{aligned} |f^{(h)}(x)| &\leq k! |a_k| + \frac{(k+1)! |a_{k+1}|}{1!} \delta + \frac{(k+2)! |a_{k+2}|}{2!} \delta^2 + \dots \\ &\dots < k! \frac{M'}{(2\delta)^k} + \frac{(k+1)!}{1!} \cdot \frac{M'}{(2\delta)^{k+1}} \delta + \frac{(k+2)!}{2!} \cdot \frac{M'}{(2\delta)^{k+2}} \delta^2 + \dots \\ &\dots = \frac{k! M'}{2^k \delta^k} \left[1 + \frac{k+1}{1!} \left(\frac{1}{2} + \frac{(k+1)(k+2)}{2!} \cdot \frac{1}{2^2} + \dots \right) \right] = \\ &= \frac{k! M'}{2^k \delta^k} \left(1 - \frac{1}{2} \right)^{-(k+1)} = 2M' \cdot \frac{k!}{\delta^k}. \end{aligned}$$

Ahora no queda más que poner $2M' = M$ para obtener las desigualdades (1.2:1).

Demostremos que las condiciones del teorema son suficientes. Sea $f(x)$ una función indefinidamente derivable en el intervalo $(x_0 - \delta, x_0 + \delta)$, y supongamos que se cumplen las desigualdades (1.2:1). Escribamos $f(x)$ según la fórmula de Taylor con el término complementario en forma de Lagrange:

$$\begin{aligned} f(x) &= f(x_0) + \frac{f'(x_0)}{1!} (x-x_0) + \dots + \frac{f^{(n-1)}(x_0)}{(n-1)!} \times \\ &\quad \times (x-x_0)^{n-1} + \frac{f^{(n)}(\xi)}{n!} (x-x_0)^n. \end{aligned} \quad (1.2:5)$$

Se tiene:

$$\left| \frac{f^{(n)}(\xi)}{n!} \right| |x-x_0|^n < \frac{M n!}{\delta^n \cdot n!} |x-x_0|^n = M \left(\frac{|x-x_0|}{\delta} \right)^n,$$

y, por consiguiente, el término complementario tiende a cero cuando $n \rightarrow \infty$, si $|x-x_0| < \delta$. De esto se deduce que en el intervalo $(x_0 - \delta, x_0 + \delta)$ la función $f(x)$ se expresa en serie de potencias:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!} (x-x_0) + \dots + \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k + \dots,$$

es decir, es analítica en el punto x_0 .

El criterio de analiticidad establecido en este teorema es satisfactorio en el sentido de su determinabilidad absoluta y su perfección. Sin embargo, resulta muy incómodo para las aplicaciones, así como para las cuestiones teóricas, puesto que está basado en el conocimiento del comportamiento de las derivadas de cualquier orden en cierto entorno del punto dado (desigualdades (1.2:1)).

1.3. El mérito de Cauchy consiste en haber desarrollado los fundamentos de la teoría general de las funciones analíticas saliendo al campo de la variable compleja. Para las funciones analíticas esta salida se efectúa de un modo simple y natural. No hay más que generalizar los principios básicos de la teoría de límites al campo de los números complejos y observar que la serie

$$a_0 + a_1(x-x_0) + \dots + a_n(x-x_0)^n + \dots,$$

que es convergente para todos los valores de la variable real x que cumplen la condición $|x - x_0| < \rho$, sigue siendo convergente también para todos los valores de la variable compleja $z = x + iy$ que cumplen la condición $|z - x_0| < \rho$ y, por consiguiente, determina una función de la variable compleja z que se puede considerar como la *prolongación* (o generalización) de la función analítica de variable real al campo de los números complejos. Ya hace mucho que se empleaban prolongaciones de este género. Estas son bien conocidas en el álgebra para el caso de polinomios donde, por cierto, no se necesitan estudiar las cuestiones de convergencia, sino que es suficiente establecer las reglas de las operaciones algebraicas sobre los números complejos, y también en el curso de análisis, por ejemplo, para el caso de la función exponencial e^x , donde de un modo semejante se deduce la fórmula clásica de Euler

$$e^{ix} = \cos x + i \operatorname{sen} x.$$

El mérito histórico de Cauchy, mencionado anteriormente, no radica en que él comenzó a reemplazar en las series de potencias la variable real x por la variable compleja z . Esto ya lo hacían antes, en el siglo XVIII. Su mérito consistió en que Cauchy fundó una teoría sistemática de las funciones de variable compleja, semejante a la teoría de las funciones de variable real (aquí tenemos en cuenta el análisis clásico), que incluye la teoría de límites, el concepto de discontinuidad, la derivada, la integral y la teoría de series. Con esto, quedó claro el hecho fundamental de que el concepto de función de variable compleja, analítica en un recinto G , es decir, de función que admite una representación de la forma

$$f(z) = a_0 + a_1(z-z_0) + \dots + a_n(z-z_0)^n + \dots \quad (1.3:1)$$

en cierto entorno de cada punto z_0 del recinto (todos los números que figuran aquí son complejos y, en particular, reales), coincide totalmente con el concepto de función derivable en el mismo recinto. Uno de ellos implica al otro, y la existencia de la derivada primera de la función $f(z)$ da lugar a que la función sea desarrollable en serie de potencias. Ya se vio en el apartado 1.2 lo complicada que es la relación entre la derivabilidad de la función y su analiticidad cuando nos limitamos solamente a valores reales de la variable indepen-

diente. La salida al plano complejo tiene también muchas más ventajas ante el estudio de las funciones en el campo real solamente. En particular, aquí se observa que las propiedades de analiticidad se pueden caracterizar también muy sencillamente considerando la integral de la función compleja $f(z)$, o estudiando sus desarrollos en series de polinomios arbitrarios, etc.

Esta pluralidad de propiedades y relaciones simples entre ellas sirven de verdadero fundamento de la teoría de las funciones analíticas de variable compleja, llamada también frecuentemente teoría de las funciones de variable compleja. Desde luego, las conclusiones obtenidas acerca de las funciones analíticas de variable compleja aclaran también, particularmente, las cuestiones referentes a las funciones analíticas de variable real en las que se convierten las primeras funciones cuando los números $a_0, \dots, a_n, \dots, z_0$, en la expresión (1.3:1) son reales y a la variable z se le atribuyen valores reales.

El presente curso está dedicado a la teoría de las funciones de variable compleja.

§ 2. LOS NUMEROS COMPLEJOS

2.1. Se suponen conocidos por el curso de álgebra *) los números complejos, su representación geométrica y las operaciones con los mismos. Para facilitar la lectura, haremos aquí un resumen de las definiciones y conclusiones fundamentales relativas a los números complejos.

Todo número complejo es de la forma $a + bi$, donde a y b son números reales. El primero de ellos se llama parte real, y el segundo, parte imaginaria del número complejo. Designando $a + bi$ mediante c , escribiremos:

$$a = \text{Re } c \quad \text{y} \quad b = \text{Im } c,$$

donde Re son las primeras letras de la palabra latina *realis* (real) e Im son las primeras letras de la palabra *imaginarius* (imaginario). Dos números complejos se suponen iguales cuando, y sólo cuando, son iguales las partes reales e imaginarias por separado. Un número complejo con la parte imaginaria igual a cero: $c = a + 0 \cdot i$, se escribe así: $c = a$, y se identifica con el número real a . En particular, el número $0 + 0 \cdot i$ se identifica con el cero. Por lo tanto, los números reales representan un caso particular de los complejos. Un número complejo con la parte real igual a cero: $c = 0 + b \cdot i$, se escribe así: $c = bi$. Si también $b = 0$, entonces, c es como antes igual a cero ($a = b = 0$). Si $b \neq 0$, el número c se llama imaginario

*) A. K u r o s c h, *Curso de álgebra superior*, Editorial MIR, Moscú, 1968.

puro. En particular, cuando $b = 1$ resulta la unidad imaginaria: $1 \cdot i = i$. En general, un número complejo $a + bi$ se llama imaginario, si su parte imaginaria es distinta de cero. Por lo tanto, los números imaginarios puros representan un caso particular de los imaginarios (que corresponden al caso, en que la parte real es igual a cero).

Los números complejos forman un campo*). Esto significa que para ellos están definidas las operaciones de sumar y multiplicar, cumpliéndose las leyes conmutativa, asociativa y distributiva

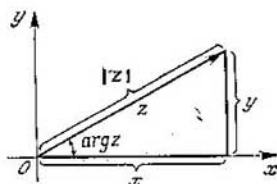


FIG. 1

(para la multiplicación con respecto a la suma), y además, para la suma existe la operación inversa: la resta, y para la multiplicación, la operación inversa: la división. En otras palabras, las ecuaciones $a + x = b$ y $a \cdot x = b$ siempre tienen solución con respecto a la incógnita x (esto último con la condición complementaria de que $a \neq 0$).

Los números 0 y 1 se llaman cero y unidad, respectivamente, del campo de números complejos. Desde el punto de vista de las operaciones en el campo de los números complejos, todo número imaginario puro bi puede ser interpretado como el producto del número real b por la unidad imaginaria i , y todo número complejo $a + bi$, como la suma del número real a y el número imaginario puro bi .

2.2. Lo más sencillo para representar geoméricamente los números complejos son los puntos o los vectores del plano, en el que se ha elegido un sistema cartesiano de coordenadas rectangulares. Por imagen geométrica del número complejo $c = a + bi$ se puede tomar con igual ventaja, tanto el punto de abscisa a y ordenada b (en este caso, el número $a + bi$ se llama afijo de este punto, que proviene de la palabra latina *affixus*, que significa sujeto a algo) como el vector cuya proyección sobre el eje de abscisas es igual a a y su proyección sobre el eje de ordenadas, igual a b ,

*) O sea, un cuerpo conmutativo. (Nota del T.)

Empleando el idioma geométrico se puede hablar entonces del punto o del vector, en lugar de hablar del número complejo correspondiente.

Conservando para la abscisa y ordenada las designaciones ordinarias x e y , escribiremos frecuentemente los números complejos en la forma $z = x + iy$. El plano, cuyos puntos (o vectores) se utilizan para representar geoméricamente a los números complejos, lo llamaremos *plano complejo*, o bien *z-plano*; el eje de abscisas, *eje real*, y el eje de ordenadas, *eje imaginario*. Está claro que los números reales se representan por puntos del eje real (o por vectores paralelos al mismo); los números imaginarios puros, por puntos del eje imaginario (o por vectores paralelos al mismo), y, en general, los números imaginarios, por puntos no situados en el eje real (o por vectores no paralelos al mismo). Si un número complejo es diferente de cero, su punto determinativo es distinto del origen de coordenadas. En este caso, para su determinación, además de las coordenadas cartesianas x e y pueden utilizarse también las polares: el radio vector $r > 0$ y el ángulo polar Φ (que se determina salvo un entero arbitrario, múltiplo de 2π).

Con respecto al afijo del punto $z = x + iy$, estos números se llaman *módulo* y *argumento*, respectivamente, y se designan del siguiente modo:

$$r = |z|, \quad \Phi = \text{Arg } z.$$

Para $z = 0$ el módulo es igual a 0, mientras que el argumento no está definido (carece de sentido).

Como, evidentemente,

$$x = r \cos \Phi \quad \text{e} \quad y = r \sin \Phi,$$

se tiene

$$z = x + iy = r(\cos \Phi + i \sin \Phi).$$

Hemos obtenido la expresión de un número complejo en coordenadas polares (o, como suele decirse, la forma *trigonométrica* del número complejo). Es comprensible que el módulo de un número complejo es al mismo tiempo la longitud del vector que representa a este número, mientras que el argumento, es el ángulo que forma este vector con la dirección positiva del eje real (determinado, como siempre, salvo un entero arbitrario, múltiplo de 2π).

Entre los valores del argumento del número $z \neq 0$, existe uno, y sólo uno, comprendido entre $-\pi$ y $+\pi$ (posiblemente, incluyendo este último valor). Este se denomina *valor principal del argumento* y se designa mediante $\arg z$. Así, pues,

$$-\pi < \arg z \leq \pi$$

y

$$\text{Arg } z = \arg z + 2n\pi,$$

donde n recorre todos los números enteros ($0, \pm 1, \pm 2, \dots$).

Señalemos las relaciones:

$$|z| = \sqrt{x^2 + y^2},$$

$$\text{tg}(\arg z) = \frac{y}{x}.$$

De esta última se puede sacar la conclusión que el $\arg z$ coincide con uno de los valores de $\text{Arctg} \frac{y}{x}$. Designando el valor principal de $\text{Arctg} \frac{y}{x}$, o sea, el valor comprendido entre $-\frac{\pi}{2}$ y $\frac{\pi}{2}$ (incluyendo, posiblemente, este último), mediante $\text{arctg} \frac{y}{x}$, tendremos:

$$\arg z = \text{arctg} \frac{y}{x}, \quad \text{si } x > 0,$$

$$\arg z = \pi + \text{arctg} \frac{y}{x}, \quad \text{si } x < 0, y \geq 0,$$

$$\arg z = -\pi + \text{arctg} \frac{y}{x}, \quad \text{si } x < 0, y < 0,$$

$$\arg z = \frac{\pi}{2}, \quad \text{si } x = 0, y > 0,$$

$$\arg z = -\frac{\pi}{2}, \quad \text{si } x = 0, y < 0.$$

Los números complejos $x + iy$ e $x - iy$ se llaman **conjugados** (entre sí). Estos se representan por puntos simétricos con respecto del eje real, y son iguales entre sí sólo cuando son números reales. Si $x + iy = z$, el número conjugado con éste se designa mediante \bar{z} : $x - iy = \bar{z}$.

Como el conjugado con $x - iy$ es $x + iy$, resulta $\overline{(\bar{z})} = z$.

De la definición misma de números conjugados se deduce que sus módulos son iguales y los valores de sus argumentos se obtienen uno del otro permutando sus signos, o sea, son opuestos. Obsérvese, por otra parte, que si $z = x < 0$, entonces, $\bar{z} = x < 0$, y los valores principales de los argumentos de los números z y \bar{z} son iguales entre sí:

$$\arg z = \arg \bar{z} = \pi.$$

2.3. Las operaciones sobre los números complejos se efectúan según las siguientes reglas:

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2),$$

$$z_1 \cdot z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1),$$

de donde se obtienen las reglas para las operaciones inversas:

$$z_1 - z_2 = (x_1 - x_2) + i(y_1 - y_2),$$

$$\frac{z_1}{z_2} = \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + i \frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2} \quad (x_2^2 + y_2^2 \neq 0).$$

Si los números complejos z_1 y z_2 vienen expresados en coordenadas polares

$$z_1 = r_1 (\cos \Phi_1 + i \operatorname{sen} \Phi_1), \quad z_2 = r_2 (\cos \Phi_2 + i \operatorname{sen} \Phi_2),$$

las reglas de la multiplicación y división proporcionarán los siguientes resultados:

$$z_1 \cdot z_2 = r_1 r_2 [\cos (\Phi_1 + \Phi_2) + i \operatorname{sen} (\Phi_1 + \Phi_2)],$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} [\cos (\Phi_1 - \Phi_2) + i \operatorname{sen} (\Phi_1 - \Phi_2)] \quad (r_2 \neq 0),$$

es decir, al multiplicar dos números complejos se multiplican sus módulos y se suman sus argumentos, mientras que al dividirlos se dividen sus módulos y se restan sus argumentos.

Todas estas operaciones tienen un significado geométrico sencillo. Representando z_1 y z_2 por vectores del plano complejo sacamos la conclusión de que la suma $z_1 + z_2$ se representa por la diagonal del paralelogramo construido sobre los vectores z_1 y z_2 . Para representar geoméricamente la diferencia es conveniente representar los puntos z_1 y z_2 por puntos (o, lo que al fin y al cabo se reduce a lo mismo, por vectores que parten del origen de coordenadas). Entonces, la diferencia $z_1 - z_2$ se representará por un vector cuyo origen coincide con el punto z_2 y su extremo, con el punto z_1 . De aquí se deduce que el módulo de la diferencia $|z_1 - z_2|$ es igual a la distancia entre los puntos z_1 y z_2 ; esta observación resulta muy útil. En particular, la ecuación $|z - z_0| = \rho$ representa una circunferencia con centro en z_0 y de radio ρ ; la desigualdad $|z - z_0| < \rho$ representa el interior de esta circunferencia.

La suma de unos cuantos números complejos se representa por un vector que cierra la poligonal construida con los vectores que representan a los sumandos.

Mediante el significado geométrico indicado de las operaciones de adición y sustracción (o bien, directamente, es decir, de un modo puramente algebraico) se establecen unas desigualdades muy importantes que permiten acotar el módulo de la suma o el módulo de la diferencia de números complejos. Estas son:

$$|z_1 + z_2 + \dots + z_n| \leq |z_1| + |z_2| + \dots + |z_n|,$$

$$|z_1 - z_2| \geq ||z_1| - |z_2||.$$

En cada una de estas desigualdades se alcanza el signo de igualdad cuando, y sólo cuando, los argumentos de los números complejos z_1, z_2, \dots, z_n son iguales entre sí, es decir, cuando los vectores correspondientes son paralelos a una misma recta y llevan una misma dirección.

Obsérvese que, para expresar el módulo de la suma de unos cuantos números complejos por estos números, frecuentemente proceden del siguiente modo. De la regla de multiplicación de los números complejos se deduce que $z \cdot \bar{z} = x^2 + y^2 = |z|^2$, es decir, $|z| = \sqrt{x^2 + y^2}$. Haciendo aquí $z = z_1 + z_2 + \dots + z_n$ y observando que entonces $\bar{z} = \bar{z}_1 + \bar{z}_2 + \dots + \bar{z}_n$, obtenemos:

$$|z| = |z_1 + z_2 + \dots + z_n| = \sqrt{(z_1 + z_2 + \dots + z_n)(\bar{z}_1 + \bar{z}_2 + \dots + \bar{z}_n)}.$$

Este es el resultado que se necesita.

El significado geométrico de la multiplicación se observa inmediatamente de la regla enunciada anteriormente. Precisando, el vector que representa al producto $z_1 \cdot z_2$ se obtiene del vector z_1 haciéndole girar un ángulo igual a $\text{Arg } z_2$ (es decir, igual a uno de los valores de $\text{Arg } z_2$), y alargándolo, o sea, cambiando su longitud $|z_2|$ veces. En particular, a la multiplicación por un número complejo z_2 , cuyo módulo es igual a la unidad:

$$z_2 = \cos \Phi_2 + i \sin \Phi_2,$$

corresponde solamente una rotación del vector z_1 alrededor del origen de coordenadas en el ángulo Φ_2 .

De un modo análogo se interpreta el significado geométrico de la división. De las reglas de sustracción y división se deduce inmediatamente que el ángulo bajo el que se ve desde el punto z_0 el par de puntos z_1 y z_2 , es igual (salvo un entero múltiplo de 2π) al argumento del cociente de las diferencias $z_1 - z_0$ y $z_2 - z_0$

$$\text{Arg} \frac{z_1 - z_0}{z_2 - z_0}.$$

Aquí hemos tomado $z_1 - z_0$ por dividendo y $z_2 - z_0$ por divisor. Esto corresponde a suponer que el ángulo con el vértice en el punto z_0 se calcula desde el vector $z_2 - z_0$ hasta el vector $z_1 - z_0$ en dirección contraria a la del movimiento de las agujas del reloj.

Detengámonos, finalmente, en las operaciones de elevación a potencia y extracción de raíz. Como ordinariamente, por potencia de un número complejo z , siendo natural el exponente n , se entiende el producto de n factores, cada uno de los cuales es igual a z .

*) Mediante $\sqrt[n]{r}$ designamos el valor positivo de la raíz de grado n del número positivo r , es decir, el valor aritmético de la raíz.

Si

$$z = r(\cos \Phi + i \operatorname{sen} \Phi),$$

de aquí se deduce que

$$z^n = r^n (\cos n\Phi + i \operatorname{sen} n\Phi);$$

para $r = 1$, obtenemos:

$$(\cos \Phi + i \operatorname{sen} \Phi)^n = \cos n\Phi + i \operatorname{sen} n\Phi$$

que es la fórmula de Moivre.

El lector puede convencerse fácilmente de que la regla de elevación a una potencia (en particular, la fórmula de Moivre) conserva su valor también para cualquier exponente entero m , si se hace

$$z^0 = 1 \text{ y } z^{-m} = \frac{1}{z^m}.$$

Se llama raíz $\sqrt[n]{z}$ de grado n (n es natural) del número complejo z cualquier número complejo ζ que satisfaga a la ecuación

$$\zeta^n = z.$$

Si $z = 0$, se tiene $\zeta = 0$. Si $z \neq 0$, haciendo las notaciones $z = r(\cos \Phi + i \operatorname{sen} \Phi)$ y $\zeta = \rho(\cos A + i \operatorname{sen} A)$, obtenemos:

$$\rho^n (\cos nA + i \operatorname{sen} nA) = r(\cos \Phi + i \operatorname{sen} \Phi),$$

de donde

$$\rho^n = r, \quad \rho = \sqrt[n]{r}$$

y

$$nA = \Phi + 2k\pi = \operatorname{Arg} z, \quad A = \frac{\operatorname{Arg} z}{n}.$$

Por consiguiente,

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

Para unos mismos valores de z y n se pueden obtener distintos valores de la raíz, tomando valores del $\operatorname{Arg} z$ que se diferencien entre sí en $2k\pi$, donde k no es divisible por n . Por ejemplo, considerando los siguientes valores del $\operatorname{Arg} z$:

$$\operatorname{arg} z, \operatorname{arg} z + 2\pi, \dots, \operatorname{arg} z + 2(n-1)\pi,$$

hallamos n valores distintos de la raíz, con los cuales se agotan todos sus valores posibles, puesto que cualquier valor del $\operatorname{Arg} z$ se diferencia de uno de los elegidos previamente en un número de la forma $2mn\pi$, donde m es un número entero. Así pues, la raíz de grado n del número z tiene n valores distintos (siendo $z \neq 0$),

que están contenidos en la fórmula:

$$\sqrt[n]{z} = \sqrt[n]{+|z|} \left(\cos \frac{\text{Arg } z}{n} + i \text{sen } \frac{\text{Arg } z}{n} \right).$$

Al valor de $\sqrt[n]{z}$, igual a

$$\sqrt[n]{+|z|} \left(\cos \frac{\arg z}{n} + i \text{sen } \frac{\arg z}{n} \right),$$

llamaremos valor principal de la raíz y lo designaremos mediante $\sqrt[n]{+z}$ (cuando z es un número real positivo, su valor principal coincide con el valor aritmético).

Como $\text{Arg } z = \arg z + 2m\pi = \arg z + \text{Arg } 1$, se tiene:

$$\begin{aligned} \sqrt[n]{z} &= \sqrt[n]{+|z|} \left(\cos \frac{\arg z + \text{Arg } 1}{n} + i \text{sen } \frac{\arg z + \text{Arg } 1}{n} \right) = \\ &= \left[\sqrt[n]{+|z|} \left(\cos \frac{\arg z}{n} + i \text{sen } \frac{\arg z}{n} \right) \right] \times \\ &\times \left(\cos \frac{\text{Arg } 1}{n} + i \text{sen } \frac{\text{Arg } 1}{n} \right) = \sqrt[n]{+z} \cdot \sqrt[n]{1}, \end{aligned}$$

o sea, todos los valores de $\sqrt[n]{z}$ se pueden obtener del principal multiplicando este último por los valores distintos de la raíz del mismo grado de la unidad.

Finalmente, introduzcamos la potencia con un exponente racional arbitrario haciendo por definición:

$$z^{\frac{m}{n}} = (\sqrt[n]{z})^m$$

(m es un número entero, n es natural, m y n son primos entre sí). Entonces, tendremos:

$$\begin{aligned} z^{\frac{m}{n}} &= \left\{ \sqrt[n]{+|z|} \left[\cos \frac{\text{Arg } z}{n} + i \text{sen } \frac{\text{Arg } z}{n} \right] \right\}^m = \\ &= (\sqrt[n]{+|z|})^m \cdot \left(\cos \frac{m}{n} \text{Arg } z + i \text{sen } \frac{m}{n} \text{Arg } z \right). \end{aligned}$$

Si convenimos en entender por $|z|_+^{\frac{m}{n}}$ el número positivo $\sqrt[n]{+|z|^m}$, la última relación se escribe así:

$$z^{\frac{m}{n}} = |z|_+^{\frac{m}{n}} \left(\cos \frac{m}{n} \text{Arg } z + i \text{sen } \frac{m}{n} \text{Arg } z \right).$$

Proponemos al lector comprobar que la definición de potencia con exponente racional mediante la igualdad

$$z^{\frac{m}{n}} = \sqrt[n]{z^{\frac{m}{n}}}$$

es equivalente a la anterior.

§ 3. CONJUNTOS Y FUNCIONES. TEORÍA DE LÍMITES. LAS FUNCIONES CONTINUAS

3.1. A continuación, hablando de números complejos, empleamos el idioma geométrico. Por lo tanto, el estudio de diversos conjuntos de números complejos se reduce al estudio de distintos conjuntos de puntos en el plano.

Sea E un conjunto de puntos del plano z . Suponemos que éste no es vacío, es decir, que contiene al menos un punto y que, en general, es un conjunto infinito de puntos. Si a cada punto $z \in E^*$ se ha puesto en correspondencia algún conjunto de puntos no vacío \mathcal{E}_z , se dice que en E está definida (o determinada) una función para la cual los puntos $z \in E$ representan los valores de la variable independiente (o del argumento) y los puntos de \mathcal{E}_z , los valores de la función. Por ejemplo, $\text{Arg } z$ es una función definida en el conjunto de todos los puntos del plano, distintos de cero, y el conjunto \mathcal{E}_z está formado aquí por los números $\arg z + 2k\pi$ ($k = 0, \pm 1, \pm 2, \dots$).

En nuestra definición se supone que, en general, las funciones, son multiformes. Si cada \mathcal{E}_z contiene solamente un punto, obtenemos una función uniforme.

Designemos los valores de la función mediante $w = u + iv$. Entonces, las funciones definidas en E se pueden representar en la forma $w = f(z)$, $w = F(z)$, $w = \varphi(z)$, \dots ($z \in E$); por cierto, la indicación del conjunto E se puede omitir siempre que esto no dé lugar a incompreensiones. Como z toma valores complejos que, generalmente, son imaginarios, se dice que se trata de funciones de variable compleja.

Particularmente, todos los valores de la función pueden ser reales: $w = u$ ($v = 0$). En este caso, la función de variable compleja puede considerarse como una función de dos variables reales, x e y , que toma valores reales u . En efecto, si al complejo $z = x + iy$ se ha puesto en correspondencia el número (o los números) u , esto significa que al par de números reales x e y se ha puesto en correspondencia el número real u .

Veamos el caso general. Como todo número complejo w se determina por su parte real u y su parte imaginaria v , definir una función $w = f(z)$ en el conjunto E significa definir en este mismo conjunto dos funciones de dos variables reales x e y : $u = \varphi(x, y)$, $v = \psi(x, y)$. Recíprocamente, si en el conjunto E están definidas dos funciones $u = \varphi(x, y)$ y $v = \psi(x, y)$, que toman valores reales una independientemente de la otra, entonces, queda definida también una

*) Escribimos $z \in E$ para decir que z es un punto del conjunto E (z pertenece a E).

función compleja:

$$w = u + iv = \varphi(x, y) + i\psi(x, y) = f(z).$$

Por ejemplo, teniendo la función de variable compleja

$$w = z^2 = (x + iy)^2 = x^2 - y^2 + 2xiy,$$

tenemos a la vez también dos funciones reales de x e y :

$$u = x^2 - y^2 \text{ y } v = 2xy.$$

A las dos funciones reales: $u = x^2 - y^2$, $v = e^{x^2+3y^2}$ corresponde la función de la variable compleja $z = x + iy$:

$$w = u + iv = (x^2 - y^2) + ie^{x^2+3y^2} = f(z).$$

Estas observaciones muestran que toda la teoría de la función de la variable compleja z se podría interpretar como la teoría de los pares de funciones de dos variables reales x e y . A veces, utilizaremos esta interpretación.

De lo anterior se deduce que el concepto de función real de dos variables reales es un caso particular del concepto de función de variable compleja. Del mismo modo, el concepto de función real de una variable está contenido en el concepto de función de variable compleja como un caso particular. Para obtenerlo, es suficiente suponer que el conjunto E está situado en el eje real (entonces, $z = x$) y que los valores de la función son números reales ($w = u$).

A continuación, hablando de funciones tendremos en cuenta las funciones de variable compleja.

3.2. Supongamos que E es el conjunto de todos los números naturales: $1, 2, 3, 4, \dots, n, \dots$. Toda función uniforme definida en E se llama *sucesión*, y sus valores, términos de la sucesión. Designándolos mediante $w_1, w_2, \dots, w_n, \dots$, de modo que w_n corresponde al valor $z = n$, denotaremos la sucesión mediante el símbolo $\{w_n\}$. Si k_1, k_2, k_3, \dots es algún conjunto infinito de números naturales (distintos entre sí), los términos correspondientes de la sucesión $\{w_n\}$: $w_{k_1}, w_{k_2}, w_{k_3}, \dots, w_{k_n}, \dots$, forman una nueva sucesión $\{w_{k_n}\}$ que, con respecto a la sucesión $\{w_n\}$ se denomina *sucesión parcial* (o contenida en $\{w_n\}$). Así, por ejemplo, $w_1, w_3, \dots, w_{2h-1}, \dots$; $w_2, w_4, w_6, w_8, \dots, w_{2h}, \dots$; $w_1, w_4, w_9, \dots, w_{h^2}, \dots$, son distintas sucesiones parciales de la sucesión $\{w_n\}$.

Sea z_0 algún punto del plano. Cualquier círculo que contenga en su interior a este punto, se llama *entorno* del mismo. En particular, todo círculo con centro en z_0 , se denomina *entorno del punto* z_0 : $|z - z_0| < \rho$ ($\rho > 0$).

El punto z_0 se llama *punto de acumulación* de la sucesión $\{w_n\}$, si para cualquier entorno del punto z_0

existe una sucesión parcial $\{w_{k_n}\}$ cuyos términos todos pertenecen a este entorno. Por ejemplo, el punto 0 es un punto de acumulación para la sucesión $1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$; los puntos 0 y 1 son puntos de acumulación para la sucesión $0, 1, 0, 1, 0, 1, \dots$ (para el punto 0 la sucesión parcial que figura en la definición de punto de acumulación está formada por los términos de índice impar: $0, 0, 0, 0, \dots$; para el punto 1, por los términos de índice par: $1, 1, 1, 1, \dots$).

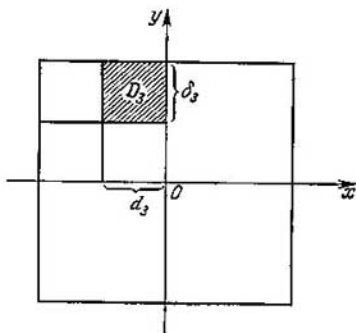


FIG. 2

Obsérvese que, si $\{w_{l_n}\}$ es una sucesión parcial de la sucesión $\{w_{k_n}\}$, y $\{w_{k_n}\}$ es una sucesión parcial de la sucesión $\{w_n\}$, entonces, $\{w_{l_n}\}$ es una sucesión parcial de la sucesión $\{w_n\}$. De aquí se deduce que todo punto de acumulación de una sucesión parcial es también punto de acumulación de la sucesión misma. Lo recíproco puede no ser cierto, como muestra el ejemplo $0, 1, 0, 1, \dots$, donde la sucesión parcial $0, 0, 0, \dots$ tiene solamente un punto de acumulación, mientras que la sucesión misma tiene dos puntos de acumulación.

Se dice que una sucesión $\{w_n\}$ es *acotada*, si existe un entorno del origen de coordenadas que contiene todos los términos de la sucesión, es decir, si existe un $\rho > 0$ tal, que $|w_n| < \rho$ ($n = 1, 2, 3, \dots$). Demostremos ahora el siguiente teorema:

Teorema 1. (Principio de Bolzano-Weierstrass para las sucesiones). *Toda sucesión acotada $\{w_n\}$ tiene al menos un punto de acumulación.*

Demostración. Sea D_1 un cuadrado con los lados paralelos a los ejes coordenados y con el centro en el origen de coordenadas que contenga a todos los términos de la sucesión. Los ejes coordenados le dividen en cuatro cuadrados, al menos uno de los

cuales, sea éste D_2 , contiene en su interior o en sus lados un conjunto infinito de términos de la sucesión. Dividiendo a éste en cuatro cuadrados iguales obtenemos un nuevo cuadrado D_3 , que contiene un conjunto infinito de términos de la sucesión (fig. 2). Continuando estos razonamientos, hallaremos una sucesión de cuadrados encajados

$$D_1 \supset D_2 \supset D_3 \supset D_4 \supset D_5 \supset D_6 \supset \dots \supset D_n \supset \dots^*),$$

cada uno de los cuales contiene en su interior o en los lados un conjunto infinito de términos de la sucesión.

Sus proyecciones sobre los ejes x e y forman dos sucesiones de segmentos encajados

$$\begin{aligned} d_1 \supset d_2 \supset d_3 \supset \dots \supset d_n \supset \dots, \\ \delta_1 \supset \delta_2 \supset \delta_3 \supset \dots \supset \delta_n \supset \dots, \end{aligned}$$

que se ciñen a los puntos ξ y η , respectivamente. El punto ξ pertenece a cada uno de los segmentos d_n , el punto η , a cada uno de los segmentos δ_n ; por lo tanto, el punto $\zeta = \xi + i\eta$ pertenece a cada uno de los cuadrados D_n . Demostremos que ζ es un punto de acumulación de la sucesión $\{w_n\}$. En efecto, cualquiera que sea el entorno U : $|z - \zeta| < \rho$ del punto ζ , es posible elegir n tan grande que el cuadrado D_n esté completamente contenido en U . Para esto es suficiente exigir que la diagonal del cuadrado sea menor que ρ . Pero su longitud es, evidentemente, igual a $\frac{l}{\sqrt{2}^n}$, donde l es la longitud del cuadrado D_1 , y nuestra condición se cumplirá si $\frac{l}{\sqrt{2}^n} < \rho$.

Como D_n contiene un conjunto infinito de términos de la sucesión $\{w_n\}$, también U contendrá un conjunto infinito de éstos, de donde se deduce que ζ es un punto de acumulación de la sucesión $\{w_n\}$.

El teorema queda demostrado.

Una sucesión acotada $\{w_n\}$, que tiene solamente un punto de acumulación z_0 , se llama **convergente**, y z_0 se llama **límite** de la sucesión.

También se dice que $\{w_n\}$ converge a z_0 (o hacia z_0) y se denota así: $w_n \rightarrow z_0$ para $n \rightarrow \infty$ (o, cuando $n \rightarrow \infty$), o bien, $\lim_{n \rightarrow \infty} w_n = z_0$.

Evidentemente, toda sucesión parcial $\{w_{h_n}\}$ de una sucesión convergente $\{w_n\}$ converge al mismo límite z_0 . En efecto, $\{w_{h_n}\}$ es acotada, puesto que es acotada toda la sucesión $\{w_n\}$; por consiguiente, (según el teorema 1) $\{w_{h_n}\}$ tiene al menos un punto de acumulación. Pero los puntos de acumulación para $\{w_{h_n}\}$ son a la vez puntos de acumulación para $\{w_n\}$; por esto, $\{w_{h_n}\}$ tiene sola-

*) Escribimos: $D_{n-1} \supset D_n$ queriendo decir que D_{n-1} contiene a D_n .

mente un punto de acumulación z_0 , de donde se deduce nuestra afirmación.

Demostremos que $\{w_n\}$ converge a z_0 cuando, y sólo cuando, cualquier entorno del punto z_0 contiene todos los términos de la sucesión $\{w_n\}$ comenzando desde uno de ellos; en otras palabras, si para cualquier $\rho > 0$ se cumplen las desigualdades: $|w_n - z_0| < \rho$ para $n > N(\rho)$.

En efecto, supongamos que $\{w_n\}$ converge a z_0 . Si, para cierto $\rho > 0$, en el exterior de la circunferencia $|z - z_0| = \rho$, o en esta misma circunferencia, hay un conjunto infinito de términos de la sucesión $\{w_n\}$: $w_{n_1}, w_{n_2}, w_{n_3}, \dots, w_{n_h}, \dots$ ($n_1 < n_2 < n_3 < \dots < n_h < \dots$), éstos forman a su vez una sucesión que, al igual que la sucesión dada, será acotada. Por consiguiente, según el teorema 1, ésta tiene que tener al menos un punto de acumulación z_1 . Pero $z_1 \neq z_0$, puesto que el entorno $|z - z_0| < \rho$ del punto z_0 no contiene ningún punto de $\{w_{n_h}\}$; por otra parte, z_1 , siendo un punto de acumulación para $\{w_{n_h}\}$, tiene que ser también un punto de acumulación para toda la sucesión $\{w_n\}$.

Hemos obtenido un punto de acumulación de $\{w_n\}$, distinto de z_0 , lo cual contradice a la hipótesis. De aquí se deduce que, en el exterior de la circunferencia $|z - z_0| = \rho$ y en la circunferencia misma, solamente puede haber un número finito de términos de la sucesión $\{w_n\}$, y que, por consiguiente, todos los términos de esta sucesión, comenzando desde cierto número, están situados en el interior de la circunferencia.

Así, pues, la condición mencionada es necesaria para la convergencia de la sucesión. Pero ésta es también suficiente. En efecto, la sucesión $\{w_n\}$, para la que se cumple esta condición, es acotada, puesto que todos los puntos de $\{w_n\}$, comenzando desde uno de ellos, están situados en el interior del círculo $|z - z_0| < 1$ (aquí $\rho = 1$), y siempre se puede tomar un entorno del punto $z = 0$ de un radio tan grande que incluya también al círculo $|z - z_0| < 1$, así como a los puntos $\{w_n\}$ (que hay una cantidad finita de ellos), que no están situados en el interior del último círculo. En virtud del teorema 1, la sucesión considerada tiene que tener al menos un punto de acumulación. Pero, ningún punto z_1 , distinto de z_0 , puede ser punto de acumulación de esta sucesión. En efecto, tomando $\rho < \frac{1}{2} |z_1 - z_0|$ y observando que sólo un número finito de términos de la sucesión $\{w_n\}$ puede estar situado fuera del círculo $|z - z_0| < \rho$, hacemos la conclusión de que solamente un número finito de ellos está situado en el círculo $|z - z_1| < \rho$. Pero esto significa que z_1 no es un punto de acumulación. Por lo tanto, z_0 es el único punto de acumulación de la sucesión $\{w_n\}$, es decir, $\{w_n\}$ converge hacia z_0 .

Demostremos que cualquier sucesión acotada $\{w_n\}$ posee una sucesión parcial convergente. Sea, en efecto, z_0 un punto de acumulación de la sucesión $\{w_n\}$ (tal punto existe en virtud del teorema 1). Entonces, cualquier entorno del punto z_0 contiene un conjunto infinito de términos w_n . Consideremos el entorno $|z - z_0| < 1$ y sea w_{n_1} un punto cualquiera situado en este entorno. Tomemos luego el entorno $|z - z_0| < \frac{1}{2}$, y del conjunto infinito de términos w_n contenidos en él, tomemos un término w_{n_2} con un subíndice mayor que n_1 . Supongamos que ya hemos tomado los términos $w_{n_1}, w_{n_2}, \dots, w_{n_h}$, situados respectivamente en los entornos: $|z - z_0| < 1, |z - z_0| < \frac{1}{2}, \dots, |z - z_0| < \frac{1}{k}$ y tales que $n_1 < n_2 < \dots < n_h$. Entonces, entre el conjunto infinito de términos w_n situados en el entorno $|z - z_0| < \frac{1}{k+1}$, tomaremos por $w_{n_{h+1}}$ un término cualquiera con el subíndice mayor que n_h . Por lo tanto, existe una sucesión parcial $\{w_{n_k}\}$ de la sucesión $\{w_n\}$ tal, que $|w_{n_k} - z_0| < \frac{1}{k}$.

Evidentemente, todos los términos de esta sucesión, cuyos subíndices n_k cumplen la condición $k > \frac{1}{\rho}$, están situados en el entorno $|z - z_0| < \rho$. De aquí se deduce que $\{w_{n_k}\}$ converge hacia z_0 , o sea, $\{w_n\}$ posee una sucesión parcial convergente.

Veamos un criterio de convergencia que no se basa en el conocimiento del límite de la sucesión (criterio de Cauchy).

Teorema 2. *Para que una sucesión $\{w_n\}$ sea convergente, es necesario y suficiente que, para cualquier $\varepsilon > 0$ exista cierto $N(\varepsilon)$ tal, que se cumpla la desigualdad $|w_n - w_{n+p}| < \varepsilon$ para todos los $n > N(\varepsilon)$ y cualquier número natural p .*

Demostración. La condición es necesaria. En efecto, si $\{w_n\}$ converge hacia un límite z_0 , se tiene

$$|w_n - z_0| < \frac{\varepsilon}{2}$$

para $n > N(\varepsilon)$ y, por consiguiente,

$$|w_{n+p} - z_0| < \frac{\varepsilon}{2}$$

para los mismos valores de n y para cualesquiera números naturales p . De aquí que

$$|w_n - w_{n+p}| = |(w_n - z_0) - (w_{n+p} - z_0)| \leq |w_n - z_0| + |w_{n+p} - z_0| < \varepsilon$$

para $n > N(\varepsilon)$, es decir, se cumple la condición.

Demostremos ahora que esta condición también es suficiente para la convergencia. En efecto, tomando $\varepsilon = 1$, obtenemos

$$|w_n - w_{n+p}| < 1,$$

comenzando desde cierto valor de n . Fijando uno de tales n , por ejemplo $n = n_0$, obtenemos que $|w_{n_0} - w_{n_0+p}| < 1$, es decir, todos los puntos $w_{n_0+1}, w_{n_0+2}, \dots$ están situados en el círculo de radio 1 con centro en el punto w_{n_0} .

Tomando ahora un entorno $|z| < \rho$ del origen de coordenadas que contenga el círculo indicado y también todos los puntos $w_1,$

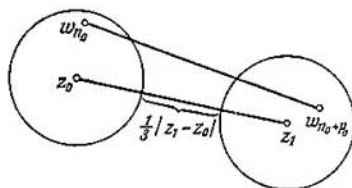


FIG. 3

w_2, \dots, w_{n_0} , obtenemos un entorno que contiene todos los términos de la sucesión, de donde se deduce que $\{w_n\}$ es una sucesión acotada. Por consiguiente, según el teorema 1, ésta tiene al menos un punto de acumulación.

No queda más que demostrar que no pueden existir dos puntos de acumulación distintos. Supongamos, por el contrario, que z_0 y z_1 , $z_0 \neq z_1$, son dos puntos de acumulación de la sucesión $\{w_n\}$. Haciendo $\varepsilon = \frac{1}{3} |z_1 - z_0|$, hallamos que las desigualdades

$$|w_n - w_{n+p}| < \frac{1}{3} |z_1 - z_0| \quad (3.2.1)$$

se cumplen comenzando desde un valor de n suficientemente grande: $n > N$. Por otra parte, cada uno de los entornos

$$|z - z_0| < \frac{1}{3} |z_1 - z_0| \quad \text{y} \quad |z - z_1| < \frac{1}{3} |z_1 - z_0|$$

de los puntos z_0 y z_1 tiene que contener un conjunto infinito de términos de la sucesión $\{w_n\}$. Sea w_{n_0} un término de subíndice mayor que N , perteneciente al primero de estos entornos, y sea $w_{n_0+p_0}$ un término perteneciente al segundo entorno. Entonces, evidentemente, (fig. 3),

$$|w_{n_0} - w_{n_0+p_0}| > \frac{1}{3} |z_1 - z_0|,$$

lo cual, sin embargo, contradice a la desigualdad (3.2.1). De esta contradicción se deduce que $\{w_n\}$ posee solamente un punto de acumulación, es decir, es convergente. El teorema queda demostrado.

3.3. Hagamos $w_n = u_n + iv_n$, donde u_n y v_n son las partes real e imaginaria de w_n . Entonces, junto con la sucesión de números complejos $\{w_n\}$, obtenemos dos sucesiones de números reales $\{u_n\}$ y $\{v_n\}$.

Demostremos la siguiente proposición.

T e o r e m a. Una sucesión $\{w_n = u_n + iv_n\}$ converge hacia el límite $z_0 = x_0 + iy_0$ cuando, y sólo cuando, las sucesiones de números reales $\{u_n\}$ y $\{v_n\}$ convergen hacia los límites x_0 e y_0 , respectivamente.

D e m o s t r a c i ó n. Supongamos que $\lim_{n \rightarrow \infty} w_n = z_0$; entonces, para cualquier $\varepsilon > 0$, la desigualdad $|w_n - z_0| < \varepsilon$ se cumplirá para $n > N(\varepsilon)$. Pero

$$|u_n - x_0| = |\operatorname{Re}(w_n - z_0)| \leq |w_n - z_0|$$

y

$$|v_n - y_0| = |\operatorname{Im}(w_n - z_0)| \leq |w_n - z_0|;$$

por lo tanto, para $n > N(\varepsilon)$ se cumplen las desigualdades

$$|u_n - x_0| < \varepsilon \text{ y } |v_n - y_0| < \varepsilon,$$

de donde se deduce que

$$\lim_{n \rightarrow \infty} u_n = x_0 \text{ y } \lim_{n \rightarrow \infty} v_n = y_0.$$

Recíprocamente, si se sabe que

$$\lim_{n \rightarrow \infty} u_n = x_0 \text{ y } \lim_{n \rightarrow \infty} v_n = y_0,$$

entonces, para cualquier $\varepsilon > 0$ se cumplen las desigualdades

$$|u_n - x_0| < \frac{\varepsilon}{2} \text{ y } |v_n - y_0| < \frac{\varepsilon}{2}$$

para $n > N(\varepsilon)$. Pero

$$|w_n - z_0| = |(u_n - x_0) + i(v_n - y_0)| \leq |u_n - x_0| + |v_n - y_0|;$$

por consiguiente,

$$|w_n - z_0| < \varepsilon$$

para $n > N(\varepsilon)$, de donde se deduce que

$$\lim_{n \rightarrow \infty} w_n = z_0.$$

El teorema queda demostrado.

En virtud de este teorema, cualquier cuestión acerca de la convergencia de una sucesión de números complejos es equivalente al

problema de la convergencia de dos sucesiones de números reales. Por ello, por ejemplo, las proposiciones conocidas sobre el límite de la suma, diferencia, producto o cociente de dos sucesiones convergentes se extienden sin cambio alguno al caso de sucesiones de términos complejos. Precisando, si las sucesiones $\{w'_n\}$ y $\{w''_n\}$ convergen hacia los límites z'_0 y z''_0 , las sucesiones $\{w'_n + w''_n\}$, $\{w'_n - w''_n\}$, $\{w'_n \cdot w''_n\}$ también convergen hacia los límites $z'_0 + z''_0$, $z'_0 - z''_0$, $z'_0 \cdot z''_0$, respectivamente. Si $w''_n \neq 0$ ($n = 1, 2, \dots$) y $z''_0 \neq 0$, la sucesión $\left\{ \frac{w'_n}{w''_n} \right\}$ también es convergente y su límite es igual a $\frac{z'_0}{z''_0}$.

Obsérvese que la proposición expresada por el último teorema se podría tomar por definición de convergencia de una sucesión de términos complejos. Entonces, todos los teoremas del precedente apartado se podrían deducir de los teoremas correspondientes sobre las sucesiones de números reales.

Para resolver el problema de la convergencia de la sucesión $\{w_n\}$, en lugar de considerar las sucesiones de las partes real e imaginaria $\{u_n\}$ y $\{v_n\}$, se pueden considerar las sucesiones de los módulos y de los valores principales de los argumentos: $\{|w_n|\}$ y $\{\arg w_n\}$. Por ejemplo, esto resulta conveniente al estudiar las sucesiones que convergen hacia cero, puesto que, para que una sucesión de números complejos $\{w_n\}$ tienda a cero, es necesario y suficiente que tienda a cero la sucesión de los módulos $\{|w_n|\}$. (En este caso no hace falta examinar la sucesión $\{\arg w_n\}$; ésta puede ser incluso divergente.) Para convencerse que es cierto lo dicho, es suficiente observar que el cumplimiento de la desigualdad $|w_n - 0| = |w_n| < \epsilon$ para $n > N$ (ϵ significa, simultáneamente, la convergencia a cero, tanto de la sucesión de los números complejos $\{w_n\}$ como de la sucesión de los números reales $\{|w_n|\}$).

En el caso general, la convergencia simultánea de las sucesiones $\{|w_n|\}$ y $\{\arg w_n\}$ es suficiente para la convergencia de la sucesión $\{w_n\}$; además, si $\lim_{n \rightarrow \infty} |w_n| = r$ y $\lim_{n \rightarrow \infty} \arg w_n = \varphi$, entonces, $\lim_{n \rightarrow \infty} w_n = r (\cos \varphi + i \sin \varphi)$. En efecto, tenemos: $u_n = \operatorname{Re}(w_n) = |w_n| \cos(\arg w_n)$, $v_n = \operatorname{Im}(w_n) = |w_n| \sin(\arg w_n)$; por lo tanto, $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} [|w_n| \cos(\arg w_n)] = r \cos \varphi$ y $\lim_{n \rightarrow \infty} v_n = \lim_{n \rightarrow \infty} [|w_n| \sin(\arg w_n)] = r \sin \varphi$, de donde se deduce la afirmación pedida.

Recíprocamente, si la sucesión $\{w_n\}$ es convergente, también lo es la sucesión de los módulos: $\{|w_n|\}$. En efecto, $|w_n| = \sqrt{u_n^2 + v_n^2}$, y como las sucesiones $\{u_n\}$ y $\{v_n\}$ son convergentes: $\lim_{n \rightarrow \infty} u_n = u$, $\lim_{n \rightarrow \infty} v_n = v$, existe también el límite: $\lim_{n \rightarrow \infty} |w_n| = \sqrt{u^2 + v^2}$.

igual a $\sqrt{u^2 + v^2}$. Obsérvese que la sucesión de los argumentos $\{\arg w_n\}$ de una sucesión convergente $\{w_n\}$ puede ser divergente incluso cuando $\lim_{n \rightarrow \infty} w_n \neq 0$. Supongamos, por ejemplo, que $w_n = -1 + (-1)^n \frac{i}{n}$. En este caso,

$$\lim_{n \rightarrow \infty} w_n = -1,$$

$$\arg w_{2k} = \pi - \operatorname{arctg} \frac{1}{2k} \text{ y } \arg w_{2k+1} = -\pi + \operatorname{arctg} \frac{1}{2k+1};$$

está claro que la sucesión $\{\arg w_n\}$ es divergente. No obstante, se puede hallar una sucesión convergente de valores $\operatorname{Arg} w_n$. Designemos, para esto, mediante φ_n el valor $\operatorname{Arg} w_n$ comprendido entre 0

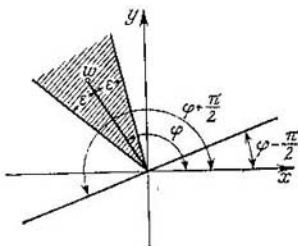


FIG. 4

y 2π : $0 \leq \varphi_n < 2\pi$. Entonces, evidentemente, tendremos que $\varphi_{2k} = \pi - \operatorname{arctg} \frac{1}{2k}$ y $\varphi_{2k+1} = \pi + \operatorname{arctg} \frac{1}{2k+1}$. La sucesión $\{\varphi_n\}$ converge hacia el límite π .

Si, en general, $\lim_{n \rightarrow \infty} w_n = w \neq 0$ y φ es alguno de los valores del $\operatorname{Arg} w$, entonces, comenzando desde cierto valor de $n = N$, todos los puntos de la sucesión $\{w_n\}$ estarán situados dentro del ángulo formado por los rayos ángulos de inclinación con el eje x son $\varphi - \frac{\pi}{2}$ y $\varphi + \frac{\pi}{2}$, y que contiene al punto w (fig. 4). Por lo tanto, para los argumentos $\operatorname{Arg} w_{N+1}, \operatorname{Arg} w_{N+2}, \dots$, se pueden tomar los valores $\varphi_{N+1}, \varphi_{N+2}, \dots$ que cumplen las desigualdades $|\varphi_{N+n} - \varphi| < \frac{\pi}{2}$. Tomando para $\operatorname{Arg} w_1, \dots, \operatorname{Arg} w_N, \dots$ sus valores $\varphi_1, \dots, \varphi_N, \dots$, se puede afirmar que la sucesión $\{\varphi_n\}$ es convergente y su límite es igual a φ . En efecto, para cualquier ε , $0 < \varepsilon < \frac{\pi}{2}$, se puede indicar un $N_1(\varepsilon) \geq N$ tal, que los puntos de

la sucesión $\{w_n\}$ cuyos subíndices sean mayores que $N_1(\varepsilon)$ estarán situados dentro del ángulo formado por los rayos cuyos ángulos de inclinación con el eje real son iguales a $\varphi - \varepsilon$ y $\varphi + \varepsilon$, y que contiene al punto w . Para ellos, los valores φ_n están comprendidos entre los límites $\varphi - \varepsilon < \varphi_n < \varphi + \varepsilon$, es decir, para $n > N_1(\varepsilon)$ se cumple la desigualdad $|\varphi_n - \varphi| < \varepsilon$, de donde se deduce que

$$\lim_{n \rightarrow \infty} \varphi_n = \varphi.$$

Obsérvese que, cuando $w \neq 0$ no es un número real negativo, se pueden tomar por φ y φ_n los valores principales del argumento. Entonces, tendremos:

$$\lim_{n \rightarrow \infty} \arg w_n = \arg w.$$

Así pues, si la sucesión $\{w_n\}$ es convergente y su límite $w \neq 0$, entonces, para cualquier valor $\varphi = \text{Arg } w$ existe una sucesión de valores $\varphi_n = \text{Arg } w_n$ que converge hacia $\text{Arg } w$. Precisamente en este sentido entenderemos a continuación la expresión $\lim_{n \rightarrow \infty} \text{Arg } w_n = \text{Arg } w$. Cuando $w \neq 0$ no es un número negativo, en particular, tendremos:

$$\lim_{n \rightarrow \infty} \arg w_n = \arg w.$$

Proponemos al lector demostrar como ejercicio que, si $\lim_{n \rightarrow \infty} w_n \neq 0$, entonces, cualquier sucesión de valores $\psi_n = \text{Arg } w_n$ que cumpla la condición $|\psi_{n+1} - \psi_n| \leq \pi$ desde cierto valor de $n = N$, será convergente (claro, hacia uno de los valores $\psi = \text{Arg } w$).

3.4. Apliquemos los resultados obtenidos al problema de las series de términos complejos.

Sea

$$w_1 + w_2 + \dots + w_n + \dots \quad (3.4:1)$$

una serie de términos complejos y

$$\{s_n = w_1 + \dots + w_n\} \quad (3.4:2)$$

la sucesión de sus sumas parciales. Según la definición, la serie se llama **convergente** si es convergente la sucesión de sus sumas parciales. El límite de esta sucesión se llama **suma** de la serie. Si $s = \lim_{n \rightarrow \infty} s_n$, escribimos:

$$w_1 + w_2 + \dots + w_n + \dots = s.$$

Una serie que no es convergente, se llama **divergente**. Haciendo $w_n = u_n + iv_n$ y $s_n = \sigma_n + i\tau_n = (u_1 + \dots + u_n) + i(\tau_1 + \dots + \tau_n)$

+ $i(v_1 + \dots + v_n)$, según lo anterior, obtendremos que la serie (3.4:1) es convergente cuando, y sólo cuando, son convergentes las dos sucesiones:

$$\{u_1 + u_2 + \dots + u_n\}, \quad \{v_1 + v_2 + \dots + v_n\}.$$

Pero estas últimas son las sucesiones de las sumas parciales de las dos series de términos reales:

$$u_1 + u_2 + \dots + u_n + \dots, \quad (3.4:3)$$

$$v_1 + v_2 + \dots + v_n + \dots \quad (3.4:4)$$

De este modo, una serie de términos complejos es convergente cuando, y sólo cuando, son convergentes las series formadas por las partes reales e imaginarias de los términos de la serie dada.

Además, si las sumas de las series (3.4:3) y (3.4:4) son iguales a σ y τ , respectivamente, es decir,

$$\lim_{n \rightarrow \infty} (u_1 + \dots + u_n) = \sigma \quad \text{y} \quad \lim_{n \rightarrow \infty} (v_1 + \dots + v_n) = \tau,$$

se tiene,

$$s = \lim_{n \rightarrow \infty} (w_1 + \dots + w_n) = \sigma + i\tau,$$

o sea, las sumas de las series (3.4:3) y (3.4:4) son, respectivamente, las partes real e imaginaria de la suma de la serie (3.4:1).

Aplicando el criterio de Cauchy a la sucesión de las sumas parciales (3.4:2) y observando que

$$s_{n+p} - s_n = w_{n+1} + \dots + w_{n+p},$$

obtenemos la siguiente proposición:

La serie (3.4:1) es convergente cuando, y sólo cuando, para cualquier $\varepsilon > 0$ existe un $N(\varepsilon)$ tal, que para cualquier $n > N(\varepsilon)$ y para todos los números naturales p , se cumple la desigualdad

$$|w_{n+1} + \dots + w_{n+p}| < \varepsilon.$$

En particular, de aquí resulta que la condición $\lim_{n \rightarrow \infty} w_n = 0$ es necesaria para la convergencia de la serie.

La serie (3.4:1) se llama **absolutamente convergente** si es convergente la serie de los módulos de sus términos:

$$|w_1| + |w_2| + \dots + |w_n| + \dots \quad (3.4:5)$$

Como

$$|w_{n+1} + \dots + w_{n+p}| \leq |w_{n+1}| + \dots + |w_{n+p}|,$$

de la convergencia absoluta de la serie de términos complejos se deduce la convergencia de la misma serie. Naturalmente, lo recí-

proco no es cierto, como lo muestra el conocido ejemplo de la serie:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

De las desigualdades

$$|u_n| \leq |w_n|, \quad |v_n| \leq |w_n| \quad \text{y} \quad |u_n| \leq |u_n| + |v_n|$$

se deduce que la serie (3.4:1) es absolutamente convergente cuando, y sólo cuando, son absolutamente convergentes las series (3.4:3) y (3.4:4).

Hagamos una permutación arbitraria de los términos en la serie absolutamente convergente (3.4:1). Se obtiene una nueva serie:

$$w_{n_1} + w_{n_2} + \dots + w_{n_k} + \dots, \quad (3.4:6)$$

donde la sucesión $n_1, n_2, \dots, n_k, \dots$ contiene todos los números naturales y cada uno de éstos aparece una sola vez. Las series de las partes reales e imaginarias de los términos de la serie (3.4:6) tienen la forma:

$$u_{n_1} + u_{n_2} + \dots + u_{n_k} + \dots,$$

$$v_{n_1} + v_{n_2} + \dots + v_{n_k} + \dots,$$

y como estas series se obtienen mediante una permutación de los términos en las series absolutamente convergentes (3.4:3) y (3.4:4), ellas también convergerán hacia las sumas anteriores σ y τ . De esto se deduce que la serie (3.4:6) también converge hacia la suma anterior $s = \sigma + i\tau$. Así, pues, en las series absolutamente convergentes de términos complejos es legítima cualquier alteración de los términos.

Como la convergencia absoluta de la serie (3.4:1) significa la convergencia de la serie (3.4:5) de términos reales no negativos, cualquier criterio conocido de convergencia de las series de términos no negativos se puede utilizar como criterio de convergencia absoluta. Señalemos, en particular, los criterios de D'Alembert y Cauchy.

Para la convergencia absoluta de la serie (3.4:1) es suficiente que, comenzando desde cierto valor de n , se cumplan las desigualdades

$$|w_{n+1}| \leq q |w_n| \quad (0 < q < 1)$$

(criterio de D'Alembert) o bien

$$\sqrt[n]{|w_n|} \leq q \quad (0 < q < 1)$$

(criterio de Cauchy).

Este último criterio es más general, en el sentido de que, cumpliéndose el primero se cumple también el segundo, mientras que puede cumplirse el segundo a pesar de que el primero no se cumpla.

La siguiente proposición, que a menudo se emplea, proporciona una condición suficiente de convergencia (por lo general, no absoluta):

Si

$$w_k = a_k b_k,$$

donde

$$\left| \sum_1^n a_k \right| \leq M \quad (n = 1, 2, \dots), \quad \lim_{k \rightarrow \infty} b_k = 0$$

y la serie $\sum_1^\infty |b_k - b_{k+1}|$ es convergente, la serie $\sum_1^\infty w_k$ también lo es.

D e m o s t r a c i ó n. Designemos la suma $\sum_1^n a_k$ mediante α_n ; entonces, tendremos $a_n = \alpha_n - \alpha_{n-1}$, y la suma $\sum_{n+1}^{n+p} w_k$ se escribirá del siguiente modo:

$$\begin{aligned} \sum_{n+1}^{n+p} w_k &= \sum_{n+1}^{n+p} a_k b_k = \sum_{n+1}^{n+p} (\alpha_k - \alpha_{k-1}) b_k = \sum_{n+1}^{n+p} \alpha_k b_k - \sum_{n+1}^{n+p} \alpha_{k-1} b_k = \\ &= \sum_{n+1}^{n+p} \alpha_k b_k - \sum_n^{n+p-1} \alpha_k b_{k+1} = \alpha_{n+p} b_{n+p} - \alpha_n b_{n+1} - \sum_{n+1}^{n+p-1} \alpha_k (b_{k+1} - b_k) *). \end{aligned} \quad (3.4:7)$$

De aquí se deduce que

$$\begin{aligned} \left| \sum_{n+1}^{n+p} w_k \right| &\leq |\alpha_{n+p}| |b_{n+p}| + |\alpha_n| |b_{n+1}| + \sum_{n+1}^{n+p-1} |\alpha_k| |b_k - b_{k+1}| < \\ &< M (|b_{n+p}| + |b_{n+1}| + \sum_{n+1}^{n+p-1} |b_k - b_{k+1}|), \end{aligned}$$

y si se cumplen las desigualdades

$$|b_n| < \frac{\varepsilon}{3M} \quad \text{y} \quad \sum_{n+1}^{n+p} |b_k - b_{k+1}| < \frac{\varepsilon}{3M},$$

*) La transformación (3.4:7), a que hemos sometido la suma $\sum_{n+1}^{n+p} a_k b_k$, se llama transformación de Abel; ésta es completamente análoga a la integración por partes.

para $n > N(\varepsilon)$ y cualquier p natural, entonces, en las mismas condiciones

$$\left| \sum_{n+1}^{n+p} w_k \right| < \varepsilon,$$

de donde se deduce la convergencia de la serie $\sum_1^{\infty} w_k$.

Las proposiciones conocidas referentes a las operaciones con las series de términos reales se generalizan también a las series de términos complejos.

Señalemos las siguientes proposiciones, que fácilmente se demuestran:

1) Para cualquier número natural n , las series

$$w_1 + w_2 + \dots + w_n + w_{n+1} + \dots \quad \text{y} \quad w_{n+1} + \dots + w_{n+p} + \dots$$

son simultáneamente convergentes o divergentes.

2) Si la serie $w_1 + \dots + w_n + \dots$ es convergente y su suma es igual a s , la serie $\lambda w_1 + \lambda w_2 + \dots + \lambda w_n + \dots$ también lo es y su suma es igual a λs .

3) Si la serie $w_1 + w_2 + \dots + w_n + \dots$ es convergente y su suma es igual a s , la serie

$$(w_1 + \dots + w_{n_1}) + (w_{n_1+1} + \dots + w_{n_2-1}) + \dots \\ \dots + (w_{n_k} + \dots + w_{n_{k+1}-1}) + \dots$$

también lo es y su suma también es igual a s ($n_1, n_2, \dots, n_k, \dots$ es una sucesión creciente arbitraria de números naturales).

4) Si

$$w'_1 + \dots + w'_n + \dots = s' \quad \text{y} \quad w''_1 + \dots + w''_n + \dots = s'',$$

entonces,

$$(w'_1 \pm w''_1) + \dots + (w'_n \pm w''_n) + \dots = s' \pm s''.$$

5) Si las series $\sum_1^{\infty} w'_k$ y $\sum_1^{\infty} w''_k$ son absolutamente convergentes

y sus sumas son iguales a s' y s'' , entonces, la serie $\sum_1^{\infty} (w'_1 w''_1 + w'_2 w''_2 + \dots + w'_k w''_k)$ es absolutamente convergente y su suma es igual a $s' s''$.

Para demostrar esta última proposición, examinemos la serie:

$$\begin{aligned} & |w'_1| |w''_1| + |w'_1| |w''_2| + |w'_2| |w''_1| + \dots \\ & \dots + |w'_1| |w''_k| + \dots + |w'_k| |w''_1| + \dots \end{aligned} \quad (3.4:8)$$

Evidentemente, ésta es convergente (como producto de dos series absolutamente convergentes $\sum_1^{\infty} |w_k'|$ y $\sum_1^{\infty} |w_k''|$ de términos reales). Por consiguiente, la serie de términos complejos

$$w_1'w_1'' + w_1'w_2'' + w_2'w_1'' + \dots + w_1'w_h'' + \dots + w_h'w_1'' + \dots \quad (3.4:9)$$

es absolutamente convergente. Realicemos alguna permutación de los términos de la última serie, colocando sus términos del modo siguiente:

$$w_1'w_1'' + w_1'w_2'' + w_2'w_2'' + w_2'w_1'' + w_1'w_3'' + w_2'w_3'' + w_3'w_3'' + w_3'w_2'' + w_3'w_1'' + \dots$$

Con esto no se altera la suma. Pero, las sumas parciales de uno, cuatro, nueve y, en general, de n^2 términos de la última serie son iguales, respectivamente, a

$$\begin{aligned} w_1'w_1'' &= s_1' \cdot s_1'' \\ w_1'w_1'' + w_1'w_2'' + w_2'w_2'' &= w_2'w_1'' = s_2' \cdot s_2'' \\ w_1'w_1'' + w_1'w_2'' + w_2'w_2'' + w_3'w_1'' + \dots + w_3'w_2'' &= s_3' \cdot s_3'' \\ \dots &\dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ w_1'w_1'' + w_1'w_2'' + \dots + w_n'w_1'' &= s_n' \cdot s_n'' \end{aligned}$$

(s_n' y s_n'' son las sumas parciales respectivas de las series $\sum_1^{\infty} w_k'$ y $\sum_1^{\infty} w_k''$). De aquí se deduce que la suma de la serie (3.4:9) es igual a $\lim_{n \rightarrow \infty} s_n' s_n'' = s' s''$. En virtud de la propiedad 3), la serie

$$\begin{aligned} w_1'w_1'' + (w_2'w_2'' + w_2'w_1'') + \dots \\ \dots + (w_1'w_h'' + w_3'w_{h-1}'') + \dots + w_h'w_1'' + \dots \quad (3.4:10) \end{aligned}$$

tiene la misma suma que la serie (3.4:9), es decir, $s' \cdot s''$. Además, ésta es absolutamente convergente, puesto que

$$|w_1'w_h'' + w_2'w_{h-1}'' + \dots + w_h'w_1''| \leq |w_1'| \cdot |w_h''| + \dots + |w_h'| \cdot |w_1''|,$$

y la serie

$$\begin{aligned} |w_1'|(|w_1''| + (|w_1''| + |w_2''| + |w_2''| + |w_1''|) + \dots \\ \dots + (|w_1''| \cdot |w_h''| + \dots + |w_h''| + |w_1''|) + \dots \end{aligned}$$

es convergente (en virtud de la convergencia de la serie (3.4:8)).

Así, pues, hemos demostrado que la serie (3.4:10) es absolutamente convergente y que su suma es igual a $s' s''$, lo cual se quería demostrar.

Demostremos que toda serie absolutamente convergente $\sum_1^{\infty} w_j$ posee las propiedades asociativa y conmutativa en el siguiente

sentido generalizado. Supongamos que la sucesión de todos los números naturales se ha descompuesto de algún modo en un conjunto infinito de sucesiones parciales crecientes:

$$\{n'_k\}, \{n''_k\}, \dots, \{n^{(m)}_k\}, \dots,$$

tales que cada número natural n figura en una de ellas, y sólo en una, y, además, solamente una vez. Examinemos la serie

$$\begin{aligned} & (w_{n'_1} + w_{n'_2} + \dots) + (w_{n''_1} + w_{n''_2} + \dots) + \dots \\ & \dots + (w_{n^{(m)}_1} + w_{n^{(m)}_2} + \dots) + \end{aligned} \quad (3.4:11)$$

Esta serie se obtiene de la dada mediante una alteración del orden de los términos y una consiguiente división en grupos (cada uno contiene un conjunto infinito de términos). Además, cada uno de los términos de la serie inicial figura en uno de los grupos indicados, y sólo en uno. Cerciorémonos de que la serie (3.4:11) es convergente y que su suma coincide con la suma σ de la serie dada $\sum_1^{\infty} w_j$. En esto consiste, precisamente, la propiedad de la serie absolutamente convergente de que se trataba.

Ante todo, demostremos que son convergentes cada una de las series $\sum_{k=1}^{\infty} w_{n_k}(m)$ ($m = 1, 2, \dots$), que representan términos de la serie (3.4:11). En efecto, para la suma de los módulos de cualquier suma parcial de tal serie, se tiene:

$$|w_{n'_1}(m)| + \dots + |w_{n_k}(m)| \leq |w_1| + |w_2| + \dots + |w_j| + |w_{j+1}| + \dots = s,$$

de donde se deduce la convergencia (y además, absoluta) de las series indicadas.

Para demostrar que la serie (3.4:11) también es absolutamente convergente, formemos primero la siguiente suma:

$$\begin{aligned} & (|w_{n'_1}| + |w_{n'_2}| + \dots + |w_{n'_k}|) + (|w_{n''_1}| + |w_{n''_2}| + \dots \\ & \dots + |w_{n''_h}|) + \dots + (|w_{n^{(m)}_1}| + \dots + |w_{n^{(m)}_k}|). \end{aligned}$$

Evidentemente, ésta no es superior al valor de la suma $\sum_1^{\infty} |w_j| = s$.

Con más razón,

$$\begin{aligned} & |w_{n'_1} + w_{n'_2} + \dots + w_{n'_k}| + |w_{n''_1} + w_{n''_2} + \dots + w_{n''_h}| + \dots \\ & \dots + |w_{n^{(m)}_1} + w_{n^{(m)}_2} + \dots + w_{n^{(m)}_k}| \leq s. \end{aligned}$$

Manteniendo aquí m constante y haciendo crecer a k indefinidamente, obtendremos:

$$\left| \sum_1^{\infty} w_{n_k} \right| + \left| \sum_1^{\infty} w_{n_k}'' \right| + \dots + \left| \sum_1^{\infty} w_{n_k}^{(m)} \right| \leq s,$$

es decir, la suma de los módulos de cualquier cantidad m de términos de la serie (3.4:11) no es superior al número fijado s . De aquí se deduce que esta serie es convergente (y además, absolutamente). No queda más que demostrar que la suma de la serie (3.4:11) coincide con la suma de la serie inicial $\sigma = \sum_1^{\infty} w_j$. Formemos la diferencia entre la suma σ y la suma parcial σ_m de la serie (3.4:11). Como todos los términos de la suma parcial indicada son términos de la serie $\sum_1^{\infty} w_j$, al restar se efectuará una simplificación mutua, y en la resta quedarán solamente aquellos términos de la serie $\sum_1^{\infty} w_j$ que no figuran en el sustraendo.

De la forma en que está construida la serie (3.4:11) se deduce que para cualquier número natural N se puede señalar un $M = M(N)$ tal que siendo $m > M$, la suma parcial

$$\sigma_m = \sum_1^{\infty} w_{n_k} + \sum_1^{\infty} w_{n_k}'' + \dots + \sum_1^{\infty} w_{n_k}^{(m)}$$

contendrá todos los términos w_1, w_2, \dots, w_N (y, además, también una infinidad de términos). Entonces, en la diferencia $\sigma - \sigma_m$, que representa una serie infinita absolutamente convergente, estarán contenidos solamente los términos w_j cuyos subíndices sean mayores que N . Por lo tanto, para $m > M(N)$

$$|\sigma - \sigma_m| \leq |w_{N+1}| + |w_{N+2}| + \dots + |w_{N+p}| + \dots$$

Pero la cantidad que figura en el segundo miembro puede hacerse arbitrariamente pequeña para valores de N suficientemente grandes. Por consiguiente, el primer miembro se puede hacer lo pequeño que se quiera para valores de m suficientemente grandes. Así, pues,

$$\lim_{m \rightarrow \infty} \sigma_m = \sigma,$$

o sea,

$$\sum_{m=1}^{\infty} [w_{n_1}^{(m)} + w_{n_2}^{(m)} + \dots + w_{n_k}^{(m)} + \dots] = \sigma = \sum_1^{\infty} w_j,$$

con lo que se termina la demostración.

La serie (3.4:11) puede ser escrita en forma de una serie doble

$$\sum_{m=1}^{\infty} \sum_{k=1}^{\infty} w_{n_h^{(m)}}.$$

La notación del término general de esta serie se puede simplificar suprimiendo la letra n , ya que no figura en la sumación, y poniendo

$$w_{n_h^{(m)}} = w_{km}.$$

Entonces, se obtiene la serie

$$\sum_{m=1}^{\infty} \sum_{k=1}^{\infty} w_{km}.$$

Desde el punto de vista de la proposición demostrada, esta serie doble representa solamente una de las infinitas formas posibles de expresión de la serie absolutamente convergente $\sum_1^{\infty} w_j$.

3.5. Sea E un conjunto arbitrario de puntos del plano complejo. Se dice que z_0 es un punto de acumulación para el conjunto E (o del conjunto E), si cualquier entorno del punto z_0 contiene un conjunto infinito de puntos pertenecientes a E .

Es evidente que un conjunto que consta de un número finito de puntos no posee puntos de acumulación. Comparemos entre sí los conceptos de punto de acumulación de una sucesión y de punto de acumulación de un conjunto. Supongamos que los puntos del conjunto E representan los términos de una sucesión $\{w_n\}$. Entonces, cada punto z_0 que sea un punto de acumulación para el conjunto E , será también un punto de acumulación para la sucesión, puesto que cualquier entorno del punto z_0 contiene un conjunto infinito de puntos de E y, por consiguiente, contiene también un conjunto infinito de términos de la sucesión $\{w_n\}$.

Sin embargo, puede ocurrir que un punto z_1 , perteneciente a E , no sea un punto de acumulación para E , a pesar de que represente un conjunto infinito de términos distintos de la sucesión $\{w_n\}$: $w_{k_1}, w_{k_2}, \dots, w_{k_n}, \dots$. Tal punto será un punto de acumulación para la sucesión $\{w_n\}$ y no lo será para el conjunto E . Como ejemplo, es suficiente examinar la sucesión $0, 1, 0, 1, 0, 1, \dots$, que posee dos puntos de acumulación 0 y 1 , mientras que el conjunto E que representa los términos de esta sucesión consta solamente de dos puntos 0 y 1 , por lo cual no tiene ningún punto de acumulación.

Refiriéndose a cualquier conjunto E , demostremos que un punto z_0 es un punto de acumulación para este conjunto si, y sólo si, existe una sucesión $\{w_n\}$ de puntos de E , distintos entre sí, que converge

hacia z_0 :

$$\lim_{n \rightarrow \infty} w_n = z_0.$$

En efecto, sea z_0 un punto de acumulación del conjunto E . Consideremos el entorno $|z - z_0| < 1$ del punto z_0 . En éste está contenido un conjunto infinito de puntos de E . Sea w_1 uno de ellos.

Como en el entorno $|z - z_0| < \frac{1}{2}$ también hay un conjunto infinito de puntos de E , tiene que haber entre éstos un punto $w_2 \neq w_1$. Supongamos que ya se han hallado los puntos w_1, w_2, \dots, w_n , pertenecientes a E y distintos entre sí, tales que w_k ($k = 1, 2, \dots, n$) están contenidos en los entornos $|z - z_0| < \frac{1}{k}$, respectivamente.

Como el entorno $|z - z_0| < \frac{1}{n+1}$ contiene un conjunto infinito de puntos de E , entre éstos existe un punto w_{n+1} perteneciente a E y distinto de los puntos w_1, w_2, \dots, w_n . De aquí se deduce que existe una sucesión $\{w_n\}$ de puntos de E , distintos entre sí, tales que $|w_n - z_0| < \frac{1}{n}$. Por lo tanto, $\lim_{n \rightarrow \infty} w_n = z_0$. Recíprocamente,

si se sabe que existe una sucesión $\{w_n\}_{n \rightarrow \infty}$ de puntos distintos de E , tales que $\lim_{n \rightarrow \infty} w_n = z_0$, entonces, en cualquier entorno del punto z_0 está contenido un conjunto infinito de puntos w_n pertenecientes a E , de donde se deduce que z_0 es un punto de acumulación del conjunto E . Con esto, nuestra proposición queda demostrada. Evidentemente, a los puntos w_n se les podría someter a la condición complementaria: suponer que todos ellos son distintos de z_0 .

Diremos que un conjunto E es *acotado* si existe un círculo $|z| < R$ que contiene a todos los puntos de este conjunto. Demostremos que *todo conjunto infinito acotado posee al menos un punto de acumulación*. En efecto, sea w_1 algún punto de E ; como E es un conjunto infinito, existe un punto w_2 perteneciente a E y distinto de w_1 . Supongamos que ya se han hallado n puntos distintos pertenecientes a E : w_1, w_2, \dots, w_n ; entonces, del conjunto infinito E se puede extraer otro punto más w_{n+1} , distinto de cada uno de estos puntos. Por lo tanto, existe una sucesión $\{w_n\}$ de puntos de E distintos entre sí. Esta sucesión es acotada, al igual que el mismo conjunto E . Por esto, ésta posee al menos un punto de acumulación z_0 . Este último tiene que ser también punto de acumulación para el conjunto E , puesto que cualquier entorno del mismo contiene un conjunto infinito de puntos w_n pertenecientes a E . La afirmación queda demostrada.

De las proposiciones demostradas en este apartado se deduce en particular que, *todo conjunto infinito acotado contiene una sucesión convergente de puntos de este conjunto, distintos entre sí*.

3.6. Sea E un conjunto infinito arbitrario, sea $f(z)$ una función definida en este conjunto y z_0 un punto de acumulación de E .

Supongamos que para cada sucesión de puntos z_n , pertenecientes a E y distintos de z_0 , convergente hacia z_0 (tales sucesiones existen en virtud de lo demostrado en el apartado precedente), la sucesión de los valores correspondientes de la función $\{f(z_n)\}$ es convergente. Entonces, para dos sucesiones distintas $\{z'_n\}$ y $\{z''_n\}$ que satisfacen a las condiciones indicadas, los límites $\lim_{n \rightarrow \infty} f(z'_n)$ y $\lim_{n \rightarrow \infty} f(z''_n)$ tienen que ser iguales entre sí. En efecto, es evidente que la sucesión

$$z'_1, z''_1, z'_2, z''_2, \dots, z'_n, z''_n, \dots,$$

converge hacia el punto z_0 y consta de puntos del conjunto E , distintos de z_0 ; según la hipótesis, también tiene que converger la sucesión de los valores correspondientes de la función

$$f(z'_1), f(z''_1), f(z'_2), f(z''_2), \dots, f(z'_n), f(z''_n), \dots$$

Por lo tanto, cualquiera sucesiones parciales de ésta tienen que tener un mismo límite; en particular:

$$\lim_{n \rightarrow \infty} f(z'_n) = \lim_{n \rightarrow \infty} f(z''_n),$$

que es lo que afirmábamos.

Designando con A el valor común del límite de las sucesiones $\{f(z_n)\}$, para todas las sucesiones posibles $\{z_n\}$ convergentes hacia z_0 (y compuestas de puntos z_n pertenecientes a E y distintos de z_0), escribiremos simplemente:

$$\lim_{z \rightarrow z_0, z \in E} f(z) = A \quad (3.6:1)$$

y diremos que A es el límite de $f(z)$ en el punto z_0 (respecto al conjunto E).

Demostremos que la condición (3.6:1) es equivalente a lo siguiente:

Para $f(z)$, z_0 y cualquier $\varepsilon > 0$ existe un $\delta(\varepsilon) > 0$, tal, que de $z \in E$, $|z - z_0| < \delta(\varepsilon)$ y $z \neq z_0$ se deduce que

$$|f(z) - A| < \varepsilon \quad (3.6:2)$$

Supongamos que se cumple la condición (3.6:2). Entonces, para cualquier $\varepsilon > 0$ y para una sucesión de puntos z_n pertenecientes a E y distintos de z_0 que converge hacia z_0 , se tiene:

$$|z_n - z_0| < \delta(\varepsilon)$$

para $n' > N[\delta(\varepsilon)] = N'(\varepsilon)$. Por lo tanto, también se cumplen las desigualdades

$$|f(z_n) - A| < \varepsilon$$

para $n > N'(\varepsilon)$, lo cual significa que la sucesión $\{f(z_n)\}$ es convergente y que $\lim_{n \rightarrow \infty} f(z_n) = A$.

Como esta conclusión es válida para cualquier sucesión de puntos $\{z_n\}$ tal que $\lim_{n \rightarrow \infty} z_n = z_0$, $z_n \in E$ y $z_n \neq z_0$, se tiene $\lim_{z \rightarrow z_0, z \in E} f(z) = A$. Así, pues, de la condición (3.6:2) se deduce (3.6:1).

Supongamos ahora cumplida la condición (3.6:1) y que no se cumple la condición (3.6:2). Entonces, para cierto $\varepsilon > 0$ no existirá un valor $\delta > 0$ tal, que de $z \in E$, $|z - z_0| < \delta$ y $z \neq z_0$ se deduzca la desigualdad $|f(z) - A| < \varepsilon$. Tomando $\delta = 1, \frac{1}{2}, \dots, \frac{1}{n}, \dots$, obtenemos que en cada entorno $|z - z_0| < \frac{1}{n}$ ($n = 1, 2, \dots$) existe un punto $z_n \in E$, $z_n \neq z_0$, tal que $|f(z_n) - A| \geq \varepsilon$. Es evidente que la sucesión $\{z_n\}$ converge hacia z_0 y, por consiguiente, $\lim_{n \rightarrow \infty} f(z_n) = A$. Pero esto contradice a la desigualdad $|f(z_n) - A| \geq \varepsilon > 0$, con lo que se termina la demostración de la equivalencia de las condiciones (3.6:1) y (3.6:2).

La definición introducida de límite de una función en un punto abarca como casos particulares las definiciones conocidas en el análisis de límite de las funciones de una o dos variables reales, que toman valores reales. En efecto, si, por ejemplo, $f(z) = u(x, y)$ es una función real de dos variables reales x e y ($z = x + iy$), $z_0 = x_0 + iy_0$ y A es un número real que representa el límite de la función $f(z)$ en el punto z_0 :

$$\lim_{z \rightarrow z_0} f(z) = A,$$

entonces, según lo anterior, esto significa que para cualquier $\varepsilon > 0$ existe un $\delta(\varepsilon) > 0$ tal, que

$$|f(z) - A| = |u(x, y) - A| < \varepsilon \text{ para } 0 < |z - z_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta(\varepsilon).$$

Evidentemente, las desigualdades obtenidas coinciden con aquellas, mediante las cuales se introduce el concepto de límite en el curso de análisis.

Volviendo a examinar el caso general de una función que toma valores complejos

$$f(z) = u(x, y) + iv(x, y),$$

demostramos que la relación

$$\lim_{z \rightarrow z_0} f(z) = A, \tag{3.6:1}$$

donde $z = x + iy$, $z_0 = x_0 + iy_0$, $A = a + ib$, es equivalente a dos relaciones en las que figuran solamente las funciones reales

$u(x, y) = \operatorname{Re} f(z)$ y $v(x, y) = \operatorname{Im} f(z)$:

$$\lim_{x \rightarrow x_0, y \rightarrow y_0} u(x, y) = a, \quad \lim_{x \rightarrow x_0, y \rightarrow y_0} v(x, y) = b. \quad (3.6:3)$$

En efecto, supongamos que se cumple (3.6:1). Entonces, para cualquier $\varepsilon > 0$ existe un $\delta(\varepsilon) > 0$ tal, que

$$|f(z) - A| < \varepsilon \text{ para } |z - z_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta(\varepsilon).$$

Pero

$$|u(x, y) - a| = |\operatorname{Re}[f(z) - A]| \leq |f(z) - A| < \varepsilon$$

y

$$|v(x, y) - b| = |\operatorname{Im}[f(z) - A]| \leq |f(z) - A| < \varepsilon,$$

de donde, debido a que $\varepsilon > 0$ es arbitrario, se deducen las igualdades (3.6:3).

Recíprocamente: si se cumplen las últimas relaciones, entonces para cualquier $\varepsilon < 0$, se tiene:

$$|u(x, y) - a| < \frac{\varepsilon}{\sqrt{2}} \text{ y } |v(x, y) - b| < \frac{\varepsilon}{\sqrt{2}},$$

si es que

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = |z - z_0| < \delta'(\varepsilon).$$

Por lo tanto, en las mismas condiciones,

$$|f(z) - A| = \sqrt{(u - a)^2 + (v - b)^2} < \varepsilon,$$

de donde se deduce (3.6:1).

De la equivalencia demostrada de (3.6:1) y (3.6:3) se deduce que las proposiciones elementales referentes a los límites de las funciones, conocidas en el curso de análisis, son válidas sin alteración alguna para el caso de funciones de variable compleja. Precisando, si existen los límites

$$\lim_{z \rightarrow z_0, z \in E} f(z) = A \text{ y } \lim_{z \rightarrow z_0, z \in E} \varphi(z) = B,$$

entonces, también existen los límites

$$\lim_{z \rightarrow z_0, z \in E} (f(z) \pm \varphi(z)) = A \pm B, \quad \lim_{z \rightarrow z_0, z \in E} [f(z) \cdot \varphi(z)] = A \cdot B$$

y

$$\lim_{z \rightarrow z_0, z \in E} \frac{f(z)}{\varphi(z)} = \frac{A}{B}$$

(lo último, suponiendo que $B \neq 0$).

3.7. Examinemos ahora el concepto de continuidad de una función de variable compleja. Si el punto de acumulación z_0 de un

conjunto infinito E pertenece al conjunto mismo y para una función $f(z)$, definida en E , se tiene:

$$\lim_{z \rightarrow z_0} f(z) = f(z_0), \quad (3.7:1)$$

la función $f(z)$ se llama **continua** en el punto z_0 . Una función que es continua en cada punto del conjunto E , se llama **continua en este conjunto**. Como, en el caso particular de las funciones que toman valores reales, el concepto de límite en un punto coincide con el conocido en el análisis, la definición de continuidad que acabamos de introducir coincide, en el mismo caso particular, con la que se conoce por el análisis. Hagamos, $z = x + iy$, $z_0 = x_0 + iy_0$ y $f(z) = u(x, y) + iv(x, y)$. Entonces, según lo expuesto en el apartado anterior, se puede sustituir la relación (3.7:1) por dos equivalentes a ésta:

$$\lim_{x \rightarrow x_0, y \rightarrow y_0} u(x, y) = u(x_0, y_0)$$

y

$$\lim_{x \rightarrow x_0, y \rightarrow y_0} v(x, y) = v(x_0, y_0). \quad (3.7:2)$$

De aquí se deduce que una función compleja $f(z)$ es continua en el punto $z_0 = x_0 + iy_0$ si, y sólo si, las partes real e imaginaria de $f(z)$, consideradas como funciones de dos variables reales x e y , son continuas en el punto (x_0, y_0) .

Señalemos las propiedades elementales de las funciones continuas de variable compleja que se deducen de la definición.

Si $f(z)$ y $\varphi(z)$ son continuas en el punto z_0 , las funciones $f(z) \pm \varphi(z)$, $f(z) \cdot \varphi(z)$ y $\frac{f(z)}{\varphi(z)}$ (ésta última, suponiendo que $\varphi(z) \neq 0$) también son continuas en el mismo punto.

Supongamos ahora que $f(z)$ es continua en el conjunto E en el punto z_0 y que sus valores $w = f(z)$ forman ellos mismos un conjunto infinito \mathcal{G} para el cual $w_0 = f(z_0)$ es un punto de acumulación. Sea, finalmente, $\varphi(w)$ una función definida en \mathcal{G} y continua en el punto w_0 . Demostremos que entonces la función compuesta $\varphi[f(z)] = F(z)$ es también continua en el punto z_0 . En efecto, si $\{z_n\}$ es una sucesión arbitraria de puntos de E , convergente hacia z_0 , la sucesión de puntos $\{w_n = f(z_n)\}$, en virtud de la continuidad de $f(z)$, tiene que converger hacia el límite $w_0 = f(z_0)$. Debido a esto, la sucesión $\{\varphi(w_n) = \varphi[f(z_n)] = F(z_n)\}$, en virtud de la continuidad de $\varphi(w)$, tiene que converger hacia $\varphi(w_0) = F(z_0)$. Así, pues, $\lim_{z \rightarrow z_0, z \in E} F(z) = F(z_0)$, con lo que queda demostrada nuestra afirmación.

Como consecuencia de la relación que hemos establecido entre la continuidad de una función compleja $f(z)$ y la continuidad de las funciones reales $u(x, y) = \operatorname{Re} [f(z)]$ y $v(x, y) = \operatorname{Im} [f(z)]$, para las funciones complejas continuas también son válidas las demás propiedades de las funciones reales continuas.

Antes de ocuparnos de éstas, introduzcamos el concepto de conjunto cerrado.

Un conjunto F se llama **c e r r a d o**, si contiene todos sus puntos de acumulación.

Los conjuntos que carecen de puntos de acumulación (por ejemplo, los conjuntos finitos, el conjunto vacío o el conjunto de los números naturales $1, 2, 3, \dots, n, \dots$, etc.), también se consideran conjuntos cerrados. Para obtener una definición que abarque explícitamente a estos casos particulares, es conveniente enunciarla en forma negativa.

Un conjunto F se llama **c e r r a d o**, si un punto no perteneciente a F no puede ser punto de acumulación para F .

He aquí otros (ejemplos de conjuntos cerrados: el conjunto de todos los puntos del plano, el conjunto de todos los puntos de una recta o de una circunferencia, el conjunto de puntos pertenecientes a un segmento arbitrario de la recta, etc. Para el primer ejemplo (el plano), esto es evidente; en los demás, la prueba es inmediata, puesto que para un punto z no perteneciente a uno cualquiera de estos conjuntos F , se puede trazar un entorno que no contenga ningún punto de F . Por consiguiente, tal punto z no puede ser para F un punto de acumulación, y F es cerrado.

Examinemos una función $f(z)$ que sea continua en un conjunto acotado y cerrado F . Esta posee las siguientes propiedades:

1) **Propiedad de continuidad uniforme.**
Para todo $\varepsilon > 0$ existe un $\delta(\varepsilon) > 0$, tal, que la desigualdad $|f(z') - f(z'')| < \varepsilon$ se cumple para cualquier par de puntos z' y z'' que cumplan la condición $|z' - z''| < \delta(\varepsilon)$.

Lo nuevo que hay aquí, en comparación con la propiedad de continuidad de una función en un punto determinado z_0 , es que $\delta(\varepsilon)$ no depende de cómo se hayan elegido los puntos en el conjunto F .

2) **Acotación del módulo de la función.**
Existe un número real positivo M tal, que en todos los puntos del conjunto F se cumple la desigualdad

$$|f(z)| \leq M.$$

3) **Alcanzamiento del extremo superior (e inferior) del módulo.** *En el conjunto F existe al menos un punto Z_0 (z_0) tal, que en todos los puntos de F se cumple la desigualdad $|f(z)| \leq |f(Z_0)|$ ($|f(z)| \geq |f(z_0)|$).*

Las propiedades análogas para las funciones continuas que toman valores reales se demuestran ordinariamente en los cursos de análisis; demostremos que éstas se generalizan inmediatamente al caso de funciones complejas. En efecto, sea $f(z) = u(x, y) + iv(x, y)$. Si $f(z)$ es continua en todos los puntos del conjunto acotado y cerrado F , también lo serán las funciones $u(x, y)$ y $v(x, y)$ en todos sus puntos (considerando a éstas como funciones reales de dos variables reales). Por lo tanto, éstas también serán uniformemente continuas en este conjunto, es decir, para cualquier $\varepsilon > 0$ existe un $\delta(\varepsilon) > 0$ tal, que para cualquier par de puntos (x', y') y (x'', y'') que cumplan

$$\sqrt{(x' - x'')^2 + (y' - y'')^2} < \delta(\varepsilon)$$

se verifican las desigualdades

$$|u(x', y') - u(x'', y'')| < \frac{\varepsilon}{\sqrt{2}} \text{ y } |v(x', y') - v(x'', y'')| < \frac{\varepsilon}{\sqrt{2}}.$$

Pero entonces, también

$$|f(z') - f(z'')| = \sqrt{|u(x', y') - u(x'', y'')|^2 + |v(x', y') - v(x'', y'')|^2} < \varepsilon$$

para $|z' - z''| < \delta(\varepsilon)$ ($z' = x' + iy'$, $z'' = x'' + iy''$), es decir, también es válida la propiedad de continuidad uniforme en el caso de una función compleja.

Para obtener las otras dos propiedades es suficiente observar que el módulo de una función continua también es una función continua. Esto se deduce inmediatamente de la desigualdad

$$\left| |f(z)| - |f(z_0)| \right| \leq |f(z) - f(z_0)|.$$

3.8. Sea Φ algún conjunto no vacío de puntos del plano, acotado y cerrado, y sea z un punto arbitrario del plano. Llamemos distancia $\rho(z, \Phi)$ del punto z al conjunto Φ al extremo inferior de las distancias desde z hasta los distintos puntos pertenecientes a Φ . Como $z = \zeta$ ($\zeta \in \Phi$) es una función continua de ζ , definida en Φ , también $|z - \zeta|$ será una función continua de ζ en Φ y, por consiguiente, alcanzará su extremo inferior $\rho(z, \Phi)$ en cierto punto $\zeta_0 \in \Phi$: $\rho(z, \Phi) = |z - \zeta_0|$.

Así, pues, la distancia del punto z hasta el conjunto Φ coincide con la distancia del punto z hasta cierto punto del conjunto Φ . Por consiguiente, $\rho(z, \Phi)$ se anula solamente para los puntos que pertenecen a Φ ; si z no pertenece a Φ , se tiene: $\rho(z, \Phi) > 0$.

Demostremos que, considerando a $\rho(z, \Phi)$ como función de z , definida en todo el plano, resulta una función continua. En efecto, sea $\rho(z, \Phi) = |z - \zeta_0|$, $\zeta_0 \in \Phi$. Entonces,

$$|z' - \zeta_0| = |(z' - z) + (z - \zeta_0)| \leq |z' - z| + |z - \zeta_0| = \rho(z, \Phi) + |z' - z|,$$

y, por consiguiente,

$$\rho(z', \Phi) \leq |z' - \zeta_0| \leq \rho(z, \Phi) + |z' - z|,$$

de donde

$$\rho(z', \Phi) - \rho(z, \Phi) \leq |z' - z|.$$

Cambiando de lugar en estos razonamientos z' por z , hallaremos

$$\rho(z, \Phi) - \rho(z', \Phi) \leq |z - z'|.$$

Por lo tanto

$$|\rho(z, \Phi) - \rho(z', \Phi)| \leq |z - z'|,$$

y es suficiente tomar $|z - z'| < \delta(\varepsilon) = \varepsilon$ para tener también

$$|\rho(z, \Phi) - \rho(z', \Phi)| < \varepsilon.$$

De la continuidad demostrada de la función $\rho(z, \Phi)$ se deduce, en particular, que si z pertenece a algún conjunto de puntos F , acotado y cerrado, la función $\rho(z, \Phi)$ también será continua sobre el mismo y, por consiguiente, alcanzará en cierto punto $z_0 \in F$ su extremo inferior

$$\inf_{z \in F} \rho(z, \Phi) = \rho(z_0, \Phi).$$

Este último posee la propiedad de que para cualquier par de puntos $z \in F$ y $\zeta \in \Phi$ se tiene:

$$|z - \zeta| \geq \rho(z, \Phi) \geq \rho(z_0, \Phi).$$

Por otra parte, en virtud de las propiedades que ya conocemos de la distancia $\rho(z_0, \Phi)$, existe un punto $\zeta_0 \in \Phi$ tal, que

$$\rho(z_0, \Phi) = |z_0 - \zeta_0|.$$

El número hallado es, evidentemente, el extremo inferior de las distancias entre los pares de puntos pertenecientes a los conjuntos F y Φ , respectivamente. A la vez, nos hemos convencido de que este extremo inferior es alcanzado por cierto par de puntos $z_0 \in F$ y $\zeta_0 \in \Phi$. Este número se llama distancia entre los dos conjuntos F y Φ , y se designa mediante $\rho(F, \Phi)$:

$$\rho(F, \Phi) = \inf_{z \in F, \zeta \in \Phi} |z - \zeta| = \inf_{z \in F} \rho(z, \Phi).$$

En esta definición los conjuntos F y Φ desempeñan exactamente un mismo papel. Por esto, podríamos cambiarlos de sitio (en la expresión):

$$\rho(F, \Phi) = \rho(\Phi, F) = \inf_{\zeta \in \Phi} \rho(\zeta, F).$$

Está claro que el concepto de distancia de un punto hasta un conjunto es un caso particular del concepto de distancia entre dos conjuntos F y Φ (cuando F consta de un solo punto z).

Como la distancia entre dos conjuntos acotados y cerrados F y Φ es al mismo tiempo la distancia entre los puntos de cierto par $z_0 \in F$ y $\zeta_0 \in \Phi$:

$$\rho(F, \Phi) = |z_0 - \zeta_0|,$$

la magnitud $\rho(F, \Phi)$ se anula cuando, y sólo cuando, los conjuntos F y Φ tienen al menos un punto común. En particular, si F y Φ no tienen puntos comunes, se tiene: $\rho(F, \Phi) > 0$.

Fácilmente se observa que, cuando F o Φ no son cerrados, o bien, son cerrados pero no son acotados, el extremo inferior de las distancias entre los pares de puntos pertenecientes a F y Φ , respectivamente, puede ser igual a cero incluso cuando estos conjuntos carecen de puntos comunes. Como ejemplo es suficiente tomar por F el conjunto de los puntos de acumulación del conjunto Φ no pertenecientes a Φ (si Φ no es cerrado), o tomar por F el conjunto de todos los puntos situados en las asíntotas de una hipérbola, y por Φ , el conjunto de puntos de la hipérbola misma.

No obstante, el lector demostrará sin dificultad alguna que, si F y Φ son conjuntos cerrados no vacíos y solamente uno de ellos no es acotado, existe de nuevo un par de puntos $z_0 \in F$ y $\zeta_0 \in \Phi$ tal, que la distancia entre z_0 y ζ_0 coincide con el extremo inferior de las distancias entre todos los pares posibles de puntos $z \in F$ y $\zeta \in \Phi$. Por consiguiente, en este caso, el extremo inferior de las distancias entre $z \in F$ y $\zeta \in \Phi$ también será distinto de cero, siempre que F y Φ no tengan puntos comunes.

§ 4. CONEXIVIDAD DE LOS CONJUNTOS. CURVAS Y RECINTOS

4.1. Un conjunto E se llama *conexo* si en cualquier división del mismo en dos subconjuntos no vacíos disjuntos (sin puntos comunes) E_1 y E_2 , al menos uno de estos conjuntos contiene un punto de acumulación del otro conjunto.

El conjunto vacío y el conjunto que consta de un solo punto también se consideran conexos. Esto tiene su justificación, si se expresa la definición de conexividad en la siguiente forma negativa: un conjunto E se llama *conexo*, si no existe una división del mismo en dos conjuntos no vacíos disjuntos E_1 y E_2 , ninguno de los cuales contiene puntos de acumulación del otro.

Puede servir de ejemplo de conjunto desconexo cualquier conjunto finito que conste de más de un punto, o más generalmente, el

conjunto formado por los puntos de un número finito de conjuntos cerrados F_1, F_2, \dots, F_n , sin puntos comunes dos a dos.

Forman una clase de conjuntos cerrados y conexos de suma importancia los conjuntos de puntos pertenecientes a las curvas continuas.

Con respecto a cada función compleja $z = f(t)$ de la variable real t , definida y continua en cierto segmento $\alpha \leq t \leq \beta$ (es decir, los pares de funciones reales continuas $x = x(t)$, $y = y(t)$), se dice que ésta define o determina una curva (una línea) continua en el plano z . En este caso, los valores de la función se llaman puntos de la curva; el conjunto de todos los valores de la función se llama conjunto de puntos de la curva (frecuentemente, para abreviar, se dice simplemente curva).

En particular, los puntos $z_0 = f(\alpha)$ y $Z = f(\beta)$ se denominan punto inicial y punto final de la curva. Estos pueden coincidir, en cuyo caso se dice que la curva es cerrada.

Ordinariamente, la variable real t se llama parámetro y la igualdad $z = f(t)$, que liga los valores del parámetro con los puntos de la curva, se llama ecuación paramétrica de la curva o, simplemente, ecuación de la curva.

Dos curvas L y Λ , determinadas por las funciones $z = f(t)$ ($a \leq t \leq b$) y $z = \varphi(\tau)$ ($\alpha \leq \tau \leq \beta$), se consideran idénticas, si la ecuación de una de ellas se puede transformar en la ecuación de la otra mediante una sustitución continua y estrictamente monótona del parámetro, es decir, si existe una función continua y estrictamente monótona $\tau = \lambda(t)$ ($a \leq t \leq b$) tal, que la función $z = \varphi[\lambda(t)]$ coincide con la función $z = f(t)$ en el segmento $[a, b]$ *).

Además, cuando la función $\tau = \lambda(t)$ que realiza la sustitución de una representación paramétrica por la otra es creciente, se dice que en ambas curvas está determinada una dirección (o un sentido) igual; si es decreciente, se dice que las direcciones son contrarias. En el último caso, el punto inicial de L sirve de punto final para Λ y viceversa. La curva Λ , que solamente se diferencia de L en el sentido del recorrido, la designaremos a veces mediante L_- . Como las notaciones L y Λ son equitativas, se puede también designar L mediante Λ_- .

Un mismo punto z que corresponde a distintos valores del parámetro, entre los cuales al menos uno es distinto de sus valores extremos, se llama punto múltiple de la curva. La curva que carece de puntos múltiples se llama curva de Jordan.

Ilustremos con ejemplos todas estas definiciones.

*) Entonces, naturalmente, también $f[\lambda^{-1}(\tau)]$ coincide con $\varphi(\tau)$ en el segmento $[\alpha, \beta]$. Aquí $\lambda^{-1}(\tau)$ es la función inversa con respecto a $\lambda(t)$.

1) Las funciones $z = t$, $z = t^2$, $z = 1 - t$ ($0 \leq t \leq 1$) determinan una misma curva continua, representada por el segmento del eje real comprendido entre los puntos $z = 0$ y $z = 1$. A la primera y segunda expresiones corresponde una misma dirección (o un mismo sentido) en la curva (el punto inicial está situado en el cero y el punto final en la unidad); a la tercera expresión, la dirección contraria (el punto inicial está situado en la unidad y el punto final en el cero). Es obvio que esta curva no tiene puntos múltiples y, por consiguiente, es de Jordan. Como sus puntos inicial y final no coinciden, ésta es una curva de Jordan no cerrada.

2) Las funciones $z = t$ y $z = \operatorname{sen} \pi t$ ($0 \leq t \leq 1$) determinan curvas continuas distintas L y Λ , puesto que no existe una transformación monótona del parámetro que transforme una de ellas en la otra. En efecto, una transformación tal tendría que transformar la función monótona $f(t) = t$ de nuevo en una función monótona, mientras que la función $\varphi(t) = \operatorname{sen} \pi t$ no es monótona en el segmento $[0, 1]$.

Obsérvese que los conjuntos de puntos de las curvas $z = t$ y $z = \operatorname{sen} \pi t$ ($0 \leq t \leq 1$) coinciden con el conjunto de puntos del segmento $[0, 1]$, a pesar de que ambas curvas son distintas. Por consiguiente, la coincidencia de los conjuntos de puntos de dos curvas continuas, siendo condición necesaria para la identidad de estas curvas, no es, sin embargo, condición suficiente.

Obsérvese también que, en este ejemplo, la curva $z = t$ es de Jordan y no es cerrada. No obstante, la curva $z = \operatorname{sen} \pi t$ tiene puntos múltiples ($\operatorname{sen} \pi t = \operatorname{sen} \pi(1 - t)$, $0 < t < \frac{1}{2}$) y, por consiguiente, no es de Jordan. Además, esta última es cerrada, puesto que su punto final coincide con el inicial ($\operatorname{sen}(\pi \cdot 0) = \operatorname{sen}(\pi \cdot 1) = 0$).

3) Sea $\Delta_1, \Delta_2, \dots, \Delta_n$ un sistema de segmentos rectilíneos, orientados de un modo determinado y situados en el plano de tal manera que el extremo del segmento precedente Δ_j ($j = 1, 2, \dots, n - 1$) coincide con el origen del siguiente Δ_{j+1} . Designando mediante a_j el número complejo representado por el vector Δ_j y mediante z_0 el punto inicial del segmento Δ_1 , podemos obtener en el segmento $0 \leq t \leq n$ una función continua elemental que determina la curva representada por el sistema de los segmentos dados:

$$z = z_0 + a_1 + \dots + a_{j-1} + a_j(t - j + 1) \\ (j - 1 \leq t \leq j, j = 1, 2, \dots, n).$$

Esta curva se llama *poligonal*, o *quebrada*, y los segmentos Δ_j ($j = 1, 2, \dots, n$), *lados* de la poligonal. La poligonal puede ser cerrada o no, según que el extremo del segmento Δ_n coincida con el origen del segmento Δ_1 o no. Esta será una curva de Jordan solamente cuando no haya intersecciones consigo mismo,

de modo que dos lados distintos no pueden tener más de un punto común; además, este punto común sólo puede ser el extremo del lado anterior y el origen del siguiente, y cuando la curva poligonal es cerrada, también puede ser el extremo del último lado y el origen del primero.

4) Fácilmente se observa que la circunferencia con centro en el punto z_0 de radio r :

$$z = z_0 + r(\cos t + i \operatorname{sen} t) \quad (0 \leq t \leq 2\pi),$$

es una curva cerrada de Jordan.

Un conjunto no vacío, acotado, cerrado y conexo se llama *continuo**); demostremos que el conjunto de puntos de cualquier curva continua es un continuo.

En efecto, sea L una curva continua, definida por la ecuación $z = f(t)$ ($\alpha \leq t \leq \beta$), y sea ζ un punto de acumulación del conjunto de puntos de la curva. Entonces, según el apartado 3.5, existe una sucesión $\{z_n\}$ de puntos distintos de la curva L que converge hacia ζ . Hagamos $z_n = f(t_n)$; entonces, $\lim_{n \rightarrow \infty} f(t_n) = \zeta$. Extraigamos luego de la sucesión infinita de puntos $\{t_n\}$, pertenecientes al conjunto acotado $\alpha \leq t \leq \beta$, una sucesión parcial convergente $\{t_{n_k}\}$ (ap. 3.2). Entonces, tendremos: $\lim_{k \rightarrow \infty} t_{n_k} = t_0$, donde t_0 , evidentemente, pertenece al segmento $[\alpha, \beta]$. Por consiguiente, como $f'_\lambda(t)$ es continua, se tiene: $\lim_{k \rightarrow \infty} f(t_{n_k}) = f(t_0)$. Por otra parte, $\{f(t_{n_k})\}$, siendo una sucesión parcial contenida en la sucesión $\{f(t_n)\}$, tiene que tener el mismo límite ζ : $\lim_{k \rightarrow \infty} f(t_{n_k}) = \zeta$. Por lo tanto, $\zeta = f(t)$, o sea, ζ pertenece a la curva L . Queda demostrado que el conjunto de puntos de una curva continua es cerrado. Demostremos ahora que este conjunto es conexo. Suponiendo lo contrario, tendremos una división de L en dos conjuntos no vacíos L_1 y L_2 , sin puntos comunes, tales que ninguno de éstos contiene puntos de acumulación del otro. Respectivamente, el segmento $\alpha \leq t \leq \beta$ se dividirá en dos conjuntos no vacíos E_1 y E_2 , sin puntos comunes; uno de ellos estará formado por aquellos valores del parámetro t a los cuales corresponden los puntos de L_1 , y el otro, por aquellos valores de t a los cuales corresponden los puntos de L_2 . Cercioremonos de que ninguno de los conjuntos E_1 , E_2 puede contener puntos de acumulación del otro. En efecto, si, por ejemplo, el punto $t' \in E_1$ fuese para E_2 un punto de acumulación, entonces en E_2 se podría hallar una sucesión de puntos $\{t'_n\}$ que convergería hacia t' (ap. 3.5). Como $f(t)$ es una función continua, de aquí resultaría que $\lim_{n \rightarrow \infty} f(t'_n) =$

*) Según esta definición, cada punto es un continuo.

$= f(t')$, es decir, el punto $f(t') \in L_1$ sería un punto de acumulación para L_2^*), lo cual contradice a la hipótesis relativa a L_1 y L_2 .

Así, pues, la suposición de que la curva L es desconexa implica la suposición análoga relativa al segmento de la recta $[\alpha, \beta]$; no queda más que refutar este último. Obsérvese con este fin que E_1 y E_2 son conjuntos cerrados. En efecto, cada punto de acumulación de E_1 pertenece al segmento $[\alpha, \beta]$, y como dicho punto no puede pertenecer a E_2 tiene que pertenecer por lo tanto a E_1 . Por consiguiente, E_1 es cerrado. Del mismo modo, E_2 también es cerrado. Como E_1 y E_2 son conjuntos cerrados no vacíos sin puntos comunes, la distancia entre ellos $\rho(E_1, E_2)$ es distinta de cero.

Dividamos el segmento $[\alpha, \beta]$ en segmentos iguales de longitud $\delta < \rho(E_1, E_2)$: $\delta_1, \delta_2, \dots, \delta_n$, de modo que el punto final de cada uno de ellos sea el punto inicial del siguiente. Evidentemente, cualquiera de estos segmentos δ_j ($j = 1, 2, \dots, n$) pertenece enteramente a uno de los conjuntos E_1 o E_2 ; en caso contrario llegaríamos a la conclusión absurda de que la distancia entre E_1 y E_2 no superaría a δ , o sea, sería menor que $\rho(E_1, E_2)$. Supongamos, por ejemplo, que δ_1 pertenece a E_1 . Entonces, también el punto inicial del segmento δ_2 (que coincide con el punto final del segmento δ_1) y, por consiguiente, todo el segmento δ_2 pertenece a E_1 . Si se ha establecido que δ_k pertenece a E_1 ($k < n$), entonces el punto inicial del segmento δ_{k+1} , que coincide con el punto final del segmento δ_k , también pertenece a E_1 y, por consiguiente, todo el segmento δ_{k+1} pertenecerá a E_1 . Por lo tanto, todos los segmentos δ_j ($j = 1, 2, \dots, n$) tienen que pertenecer solamente a uno de los conjuntos E_1 o E_2 , de modo que el otro resultará vacío, lo cual contradice a la hipótesis. Con esto, queda terminada la demostración.

Ahora podemos demostrar que, en general, *es conexo todo conjunto E para el cual dos puntos cualesquiera del mismo z' y z'' pueden unirse mediante una curva continua L* (es decir, que se puede construir una curva para la cual uno de los puntos z', z'' es el punto inicial, y el otro, el punto final), *cuyos puntos pertenecen a E* . (Obsérvese que, sin embargo, la condición suficiente de conexión comprendida aquí no es necesaria).

En efecto, supongamos que el conjunto E se ha dividido de algún modo en dos subconjuntos no vacíos E_1 y E_2 que carecen de puntos comunes. Sea z' un punto fijo de E_1 y z'' , de E_2 . Unamos z' y z'' mediante una curva continua L cuyos puntos todos pertenezcan a E , y sean L_1 y L_2 los subconjuntos del conjunto de los puntos de

* En el caso considerado, $f(t')$, siendo un punto de acumulación de la sucesión $\{f(t_n)\}$, será también un punto de acumulación para el conjunto que representa los términos de esta sucesión, puesto que ninguno de los puntos $f(t_n)$ coincide con $f(t')$ (L_1 y L_2 no tienen puntos comunes).

la curva L formados por todos los puntos de L que pertenecen a E_1 o a E_2 , respectivamente. Evidentemente, éstos son conjuntos no vacíos ($z' \in L_1$, $z'' \in L_2$) que no tienen puntos comunes. Como L es conexo, al menos uno de ellos, por ejemplo L_1 , tiene que contener un punto de acumulación del otro. Sea z_1 este punto. Este pertenece a L_1 y, por consiguiente, a E_1 . Además, es un punto de acumulación para L_2 y, por lo tanto, también es un punto de acumulación para $E_2 \supset L_2$.

Así, pues, resulta que cualquiera que sea la división del conjunto E en conjuntos no vacíos sin puntos comunes, al menos uno de ellos tiene que poseer un punto de acumulación del otro. Con esto queda demostrado que el conjunto E es conexo. Como dos puntos cualesquiera del plano pueden unirse, por ejemplo, mediante un segmento recto, de aquí se deduce, en particular, que el plano es un conjunto conexo.

Para un conjunto cerrado y acotado, la condición de conexividad puede expresarse del modo siguiente: *Un conjunto cerrado y acotado F es conexo, es decir, es un continuo, si para cualesquiera dos de sus puntos z_0 y z' y para cualquier $\varepsilon > 0$ es posible señalar un número finito de puntos de este conjunto: $z_0, z_1, \dots, z_n = z'$, tales que $|z_k - z_{k-1}| < \varepsilon$ ($k = 1, \dots, n$).*

En efecto, supongamos que se cumple esta condición y que F no es conexo. Entonces F puede dividirse en dos subconjuntos no vacíos F_1 y F_2 , sin puntos comunes, tales que ninguno de ellos contiene puntos de acumulación del otro. Como todos los puntos de acumulación del subconjunto F_j ($j = 1, 2$) tienen que pertenecer a F , deducimos que éstos pertenecen a F_j , es decir, F_j es un conjunto cerrado. Así, pues, F_1 y F_2 son acotados, cerrados y carecen de puntos comunes. Por lo tanto, la distancia $\rho(F_1, F_2) > 0$; tomando un punto $z_0 \in F_1$ y un punto $z' \in F_2$ no podremos indicar, evidentemente, un número finito de puntos de F : $z_0, z_1, \dots, z_n = z'$, tales que $|z_k - z_{k-1}| < \rho(F_1, F_2)$ ($k = 1, 2, \dots, n$). En efecto, entre éstos tiene que haber dos puntos con subíndices consecutivos: z_{k-1} y z_k , el primero de los cuales pertenece a F_1 y el segundo a F_2 , por lo cual $|z_k - z_{k-1}| \geq \rho(F_1, F_2)$. De esta contradicción se deduce la proposición enunciada.

Apliquemos este criterio de conexión para la demostración del siguiente teorema: *si $\{F_n\}$ es una sucesión de continuos tales que $F_{n+1} \subset F_n$, entonces su intersección F (es decir, el conjunto de puntos comunes a todos los F_n) es también un continuo.*

Consideremos una sucesión de puntos $\{z_n\}$, donde $z_n \in F_n$. En virtud de la condición $F_{n+1} \subset F_n$, los puntos z_n, z_{n+1}, \dots pertenecen a F_n ; de aquí se deduce que todos los puntos de acumulación de la sucesión $\{z_n\}$ pertenecen a F_n ($n = 1, 2, \dots$), y por consiguiente, también a F . Pero el conjunto de los puntos de acumulación

de la sucesión $\{z_n\}$ no es vacío (puesto que la sucesión $\{z_n\}$ es acotada); por esto, F es un conjunto no vacío. Si éste carece de puntos de acumulación, entonces F es cerrado. Supongamos que existen puntos de acumulación de F y que z_0 es uno de ellos. Como $F \subset F_n$, el punto z_0 es para F_n un punto de acumulación y, por consiguiente, pertenece a F_n ($n = 1, 2, \dots$); por lo tanto, $z_0 \in F$. Así pues, F es un conjunto no vacío, cerrado (y, evidentemente, acotado).

Demostremos, finalmente, que F es un conjunto conexo. Suponiendo lo contrario, hallaremos que F puede dividirse en dos conjuntos no vacíos, acotados y cerrados F_0 y F' , sin puntos comunes. Supongamos que $z_0 \in F_0$ y $z' \in F'$; como $z_0 \in F_n$ y $z' \in F_n$ para cualquier n , existen unos puntos $z_0, z_1, \dots, z_\nu = z'$ del conjunto F_n tales que $|z_k - z_{k-1}| < \frac{1}{3} \rho(F_0, F') = \varepsilon$ ($k = 1, \dots, \nu$).

Entre éstos tiene que haber al menos un punto z_h cuya distancia hasta F_0 y hasta F' y, por consiguiente, también hasta F , es superior a ε .

En efecto, si para cada punto z_0, z_1, \dots, z_ν , la distancia hasta uno de los conjuntos F_0 o F' no fuese superior a ε , entonces todos estos puntos podrían dividirse en dos clases no vacías de modo que, para los puntos de una de ellas, las distancias hasta F_0 no superarían a ε , y para la otra, las distancias hasta F' no superarían a ε . Evidentemente, la distancia entre dos puntos cualesquiera de clases distintas

tiene que ser mayor que ε , puesto que $\varepsilon = \frac{1}{3} \rho(F_0, F')$. Pero, por otra parte, si $\mu < \nu$ es el subíndice superior de los puntos de la primera clase, entonces el punto de subíndice $\mu + 1$ tiene que pertenecer a la segunda clase, obteniendo para los puntos z_μ y $z_{\mu+1}$ de distintas clases la desigualdad:

$$|z_{\mu+1} - z_\mu| < \varepsilon.$$

De esta contradicción se deduce la existencia del punto pedido z_h . Este punto lo designaremos mediante ζ_n , para señalar su pertenencia al conjunto F_n .

Así, pues, $\zeta_n \in F_n$ y $\rho(\zeta_n, F) > \varepsilon$ ($n = 1, 2, \dots$).

Para cualquier punto de acumulación ζ_0 de la sucesión $\{\zeta_n\}$ tendremos que $\rho(\zeta_0, F) \geq \varepsilon$. Pero esto es imposible, puesto que todos los puntos de la sucesión $\{\zeta_n\}$ pertenecen a F . Así, pues, F tienen que ser un conjunto conexo. La demostración del teorema queda terminada.

4.2. Un conjunto O se llama a b i e r t o, si para cada uno de sus puntos existe un entorno cuyos puntos todos pertenecen a O . Son ejemplos de conjuntos abiertos: el conjunto de todos los puntos del plano, el conjunto de todos los puntos no pertenecientes a un con-

junto finito dado de rectas o circunferencias, el conjunto de todos los puntos situados en el interior de un círculo dado, etc. El conjunto vacío también se considera abierto.

Un conjunto abierto y conexo se llama *recinto*. En virtud de lo demostrado anteriormente, un conjunto abierto será conexo y, por consiguiente, será un recinto, si dos puntos cualesquiera del mismo pueden unirse mediante una curva continua perteneciente a O . Pero se verifica también lo recíproco: *dos puntos cualesquiera de un recinto arbitrario pueden unirse mediante una curva continua perteneciente al recinto*.

Demostraremos esta proposición por reducción a lo absurdo. Si fuese injusta, en cierto recinto G existirían dos puntos z' y z'' que no podrían unirse mediante una curva continua perteneciente al recinto. Designemos entonces mediante G_1 el conjunto de aquellos puntos de G que pueden unirse con z' mediante una curva continua perteneciente a G , y mediante G_2 , el conjunto de todos los demás puntos. Como z' pertenece a G junto con cierto entorno suyo $|z - z'| < \rho$ (según la definición de conjunto abierto), todos los puntos de este entorno quedarán incluidos en G_1 ; éstos pueden unirse con z' mediante segmentos rectos. Por lo tanto, G_1 y G_2 son conjuntos no vacíos ($z'' \in G_2$) que no tienen puntos comunes. Supongamos que $z_1 \in G_1$; entonces existe una curva continua L_1 que une z' con z_1 y pertenece a G . Si $|z - z_1| < \rho_1$ es un entorno del punto z_1 , perteneciente a G , entonces cualquier punto z de este entorno puede unirse con z_1 mediante un segmento recto Δ_1 , también perteneciente a G . Por consiguiente, tal punto se une con z' mediante una curva continua $L_1 + \Delta_1$ perteneciente a G ; así, pues, $z \in G_1$, es decir, el entorno $|z - z_1| < \rho_1$ del punto z_1 pertenece totalmente a G_1 . Vemos que el punto z_1 no puede ser de acumulación para el conjunto G_2 .

Supongamos, finalmente, que $z_2 \in G_2$; entonces tomamos de nuevo un entorno $|z - z_2| < \rho_2$ del punto z_2 , perteneciente a G . Si algún punto z de este entorno perteneciese a G_1 , entonces este punto podría unirse con el punto z' mediante una curva continua L_2 , perteneciente a G ; pero z puede unirse con z_2 mediante un segmento recto Δ_2 , también perteneciente al recinto G . Por consiguiente, el punto z_2 se uniría con el punto z' mediante una curva continua $L_2 + \Delta_2$, perteneciente a G , por lo cual pertenecería a G_1 en contra de la hipótesis ($z_2 \in G_2$). Así, pues, ningún punto del entorno $|z - z_2| < \rho_2$ puede pertenecer al conjunto G_1 . Por lo tanto, ningún punto del conjunto G_2 puede ser de acumulación para G_1 . Vemos que el conjunto G se divide en dos conjuntos no vacíos G_1 y G_2 sin puntos comunes, ninguno de los cuales contiene puntos de acumulación del otro. Pero esto contradice a la condición de conexividad del conjunto G . La demostración queda terminada.

De lo expuesto se deduce que la definición de recinto puede enunciarse del modo siguiente:

Un conjunto abierto O se denomina recinto, si dos puntos cualesquiera del mismo pueden unirse mediante una curva continua L cuyos puntos todos pertenecen a O .

Sea O un conjunto abierto arbitrario y z_0 , alguno de sus puntos. Consideremos el conjunto de todos los puntos de O que pueden unirse con z_0 mediante curvas continuas, pertenecientes a O . Evidentemente, este conjunto no es vacío (perteneciendo al mismo el punto z_0 junto con su entorno, perteneciente a O), es abierto (si z_1 puede unirse con z_0 mediante una curva continua γ , entonces cualquier punto z del entorno del punto z_1 , perteneciente a O , puede unirse con z_0 mediante una curva continua, constituida por γ y por el segmento rectilíneo con los extremos z_1 y z_2) y, según la definición misma, es conexo. Por lo tanto, el conjunto indicado es un recinto. Este recinto se llama **componente conexa** de O (que contiene al punto z_0).

Se puede caracterizar también la componente conexa de O como el recinto mayor, perteneciente a O y que contiene al punto dado z_0 . Es evidente que dos componentes de O que tengan un punto común tienen que coincidir totalmente. Si O no es un recinto, es decir, si O es desconexo, entonces existen al menos dos componentes de O distintas. El conjunto de componentes de O puede ser infinito (figurémonos, por ejemplo, el conjunto O formado por los círculos de radio $\leq \frac{1}{2}$, con centros en los puntos $0, \pm 1, \pm 2, \dots$; en este caso, los círculos por separado son componentes de O). En todo caso, este conjunto tiene que ser a lo sumo numerable. En efecto, tomemos un punto en cada componente de O , de coordenadas racionales x o y . Entonces, obtenemos una aplicación biyectiva del conjunto de todas las componentes sobre cierto subconjunto de un conjunto numerable, de donde se deduce lo que se afirmaba.

Respecto de cualquier conjunto abierto y, en particular, de un recinto O , todos los puntos del plano se dividen en las siguientes tres categorías:

- 1) puntos pertenecientes a O ; éstos se llaman **puntos interiores** a O ;
- 2) puntos no pertenecientes a O que son para O puntos de acumulación; éstos se llaman **puntos frontera** de O ;
- 3) puntos no pertenecientes a O que no son para O puntos de acumulación; éstos se llaman **puntos exteriores** a O .

El conjunto de los puntos interiores coincide con O y, por consiguiente, no es vacío, siempre que O no sea vacío. El conjunto de los puntos frontera — denominado **frontera del conjunto abierto** (del recinto) — será vacío cuando O sea todo el plano. Si el conjunto O

es distinto de todo el plano, entonces tiene que existir un conjunto de puntos no vacío E , no pertenecientes a O . Este conjunto puede constar solamente de puntos frontera y de puntos exteriores. Si no hay puntos exteriores, entonces E consta de puntos frontera, los cuales tiene que haber, por lo tanto, no menos de uno. Si para E hay puntos exteriores, tiene que haber, evidentemente, un conjunto infinito de éstos (puesto que para un punto exterior existe un entorno cuyos puntos son también todos exteriores). Cerciorémonos que, entonces, el conjunto de puntos frontera también tiene que ser infinito. Precizando, demostremos que en cualquier segmento rectilíneo, uno de cuyos extremos es un punto interior y el otro, exterior a O , tiene que haber al menos un punto frontera. En caso contrario, todo el segmento quedaría dividido en dos conjuntos no vacíos E_1 y E_2 de puntos interiores y exteriores a O . Pero ningún punto del conjunto E_1 (punto interior a O) puede ser punto de acumulación del conjunto E_2 (conjunto de puntos exteriores a O), del mismo modo que ningún punto del conjunto E_2 puede ser punto de acumulación de E_1 . Esto contradice a la propiedad de conexión del segmento, de donde se deduce que el segmento, además de puntos interiores y exteriores, tiene que contener también puntos frontera de O (al menos uno). De lo expuesto se deduce que, si el conjunto abierto O no es todo el plano, el conjunto de sus puntos frontera no puede ser vacío. No obstante, en este caso también puede no haber puntos de tercera categoría: puntos exteriores. En efecto, cada punto exterior se caracteriza, evidentemente, en que existe un entorno del mismo, ninguno de cuyos puntos pertenecen al conjunto dado O . Por consiguiente, si el conjunto de puntos no pertenecientes a O no contiene ningún círculo, entonces no puede contener tampoco ningún punto exterior a O , es decir, todo el conjunto consta de puntos frontera. Esto ocurre, por ejemplo, cuando O está formado por todos los puntos no pertenecientes a una recta (o a una circunferencia). Esta última es la frontera de O ; en este caso no hay puntos exteriores.

4.3. Demostremos que la frontera Γ de cualquier conjunto abierto O es un conjunto cerrado. Para esto es suficiente demostrar que un punto z_0 no perteneciente a Γ , no puede ser para Γ un punto de acumulación. En efecto, tal punto tendría que ser interior o exterior a O . En ambos casos, existe un entorno de z_0 que no contiene ningún punto de Γ . Por consiguiente, z_0 no puede ser para Γ un punto de acumulación, lo que muestra que el conjunto Γ es cerrado.

Agregando a un conjunto abierto O todos sus puntos frontera, obtenemos un conjunto \bar{O} , denominado el *clausura* del conjunto O . Demostremos que \bar{O} es un conjunto cerrado. Para esto es suficiente demostrar que ningún punto z_0 no perteneciente a \bar{O} , puede

ser un punto de acumulación para \bar{O} . En efecto, tal punto solamente puede ser exterior a O . Por lo tanto, existe un entorno del punto z_0 que no contiene ningún punto de O , y que, por consiguiente, tampoco contiene puntos frontera de O . En el entorno indicado no habrá ningún punto de \bar{O} , de donde se deduce que z_0 no es para \bar{O} un punto de acumulación. De aquí que \bar{O} es un conjunto cerrado.

Cuando O es un recinto, la clausura del conjunto O se llama **recinto cerrado** (o dominio). Por ejemplo, el conjunto de todos los puntos situados en el interior de una circunferencia y en ella misma forma un círculo cerrado; el círculo cerrado se distingue del círculo abierto, pues este último es el recinto formado por todos los puntos que están situados en el interior de la circunferencia.

Sea G un recinto acotado, es decir, un recinto cuyos puntos están contenidos todos en cierto círculo con centro en el origen de coordenadas $|z| < R$. Entonces todos los puntos que están situados fuera de este círculo también serán exteriores con respecto a G . De aquí se deduce que los puntos frontera del recinto G tienen que satisfacer a la condición $|z| \leq R$, de modo que la frontera Γ de un recinto acotado G es un conjunto cerrado y acotado.

Llamaremos a un recinto acotado G **simplemente conexo** o **múltiplemente conexo**, según que su frontera sea un conjunto conexo o desconexo. Un ejemplo simple de recinto simplemente conexo es el interior de un círculo, y un ejemplo de recinto múltiplemente conexo es el conjunto de puntos situados entre dos circunferencias concéntricas.

La clase de los recintos múltiplemente conexos admite una subdivisión consiguiente. Los recintos simplemente conexos se caracterizan en que su frontera es un continuo. Si la frontera de un recinto múltiplemente conexo consta de un número finito n ($n > 1$) de continuos, sin puntos comunes dos a dos*), el recinto G se llama de

*) Es fácil comprobar que si un conjunto dado admite alguna división en un número finito de continuos, que carecen de puntos comunes dos a dos, entonces esta división es única. En efecto, sean f_1, \dots, f_n y $\varphi_1, \dots, \varphi_m$, dos divisiones distintas de un mismo conjunto en continuos. Entonces, al menos uno de los continuos de una división (sea éste φ_j) tiene que contener tanto puntos que pertenecen como puntos que no pertenecen a cierto continuo de otra división (sea éste f_k). Designemos mediante φ_{jk} el conjunto de los puntos del continuo φ_j que pertenecen a f_k y mediante φ'_{jk} , el conjunto de todos los demás puntos del continuo φ_j (éstos tienen que pertenecer a alguno de los continuos $f_1, \dots, f_{k-1}, f_{k+1}, \dots, f_n$).

Evidentemente, todo punto de acumulación del conjunto φ_{jk} es también para f_k un punto de acumulación, por lo cual está contenido en f_k y no puede pertenecer a φ'_{jk} . Del mismo modo, un punto de acumulación del conjunto φ'_{jk} tiene que ser de acumulación para alguno de los continuos f_l ($l \neq k$); por lo tanto, este punto está contenido en f_l y no puede pertenecer a φ_{jk} . Hemos obte-

conexión finita o más precisamente, se dice que el orden de conexión es n . Finalmente, si la frontera de un recinto múltiplemente conexo no puede dividirse en un número finito de continuos, sin puntos comunes dos a dos, entonces el recinto se llama de conexión infinita.

Si dentro de algún círculo abierto K se toman $n - 1$ ($n > 1$) círculos cerrados K_1, K_2, \dots, K_{n-1} , sin puntos comunes dos a dos, entonces el conjunto de todos los puntos del círculo K no pertenecientes a ninguno de los círculos K_j ($j = 1, 2, \dots, n - 1$), representa un recinto G de orden de conexión n . Su frontera consta de n circunferencias: C y c_j , que limitan los círculos K y k_j ($j = 1, 2, \dots, n - 1$), respectivamente.

Si, para fijar ideas, tomamos el círculo unidad $|z| < 1$ y excluimos de éste los círculos cerrados $|z - \frac{1}{n}| \leq \frac{1}{2n(n+1)}$ ($n = 2, 3, 4, \dots$), y también su centro $z = 0$, obtenemos un recinto de conexión infinita cuya frontera consta de las circunferencias $|z| = 1$, $|z - \frac{1}{n}| = \frac{1}{2n(n+1)}$ ($n = 2, 3, 4, \dots$) y del punto $z = 0$.

4.4. Consideremos un conjunto cerrado de puntos arbitrario F . El conjunto O de todos los puntos del plano no pertenecientes a F es un conjunto abierto. En efecto, si O no es vacío (en este caso la proposición resulta evidente), sus puntos, al no pertenecer a F , no pueden ser tampoco puntos de acumulación de F . Por lo tanto, para cada uno de ellos existe un entorno que carece de puntos del conjunto F y que, por consiguiente, pertenece totalmente a O .

Supongamos que $z_0 \in O$; designemos con G_0 el conjunto de todos los puntos de O que pueden unirse con z_0 mediante curvas continuas pertenecientes a O . Es obvio que G_0 es un recinto. Este se denomina contiguo con respecto al conjunto cerrado F . Fácilmente se observa que puede haber una infinidad de recintos distintos, contiguos a F .

Como ejemplo, es suficiente figurarse que todo el plano se ha cuadrículado mediante rectas equidistantes, paralelas a los ejes de coordenadas. Aquí, el conjunto F está formado por todos los puntos de estas rectas, mientras que los puntos que están situados dentro de cada cuadro forman los recintos contiguos a F .

Cuando F es una curva de Jordan cerrada γ , existen dos y sólo dos recintos distintos contiguos a γ . En este caso, γ es su frontera común. Esta afirmación, que fácilmente se comprueba en cada

nido que el continuo φ_j admite una división en dos conjuntos no vacíos φ_{jR} y φ_{jL} sin puntos comunes, ninguno de los cuales contiene puntos de acumulación del otro. Pero esto contradice a la definición de continuo; de este modo, nuestra afirmación queda demostrada.

caso concreto (por ejemplo, cuando γ es una circunferencia), forma el contenido del teorema de Jordan. A continuación emplearemos este teorema, recomendando al lector que vea la demostración en algún curso de topología. Obsérvese que uno de los dos recintos siempre es acotado: el interior de la curva de Jordan; el otro es no acotado: el exterior a la curva de Jordan (por ejemplo, las partes interior y exterior de una circunferencia). Mediante estos conceptos es fácil establecer la siguiente propiedad de los recintos simplemente conexos, que es fundamental para la teoría de las funciones: *si γ es una curva de Jordan cerrada, perteneciente a un recinto simplemente conexo G , entonces el interior de γ también pertenece al recinto G .*

En efecto, sea ζ algún punto de γ . Como ζ es un punto del recinto G , existe un entorno de este punto perteneciente a G . En este entorno tiene que haber una infinidad de puntos exteriores a γ , así como una infinidad de puntos interiores a γ (esto es debido a que ζ es un punto frontera para uno y otro recintos). De aquí se deduce que tanto en el exterior como en el interior de γ hay puntos del recinto G . Pero el exterior de γ no puede pertenecer totalmente a G , puesto que en caso contrario G sería no acotado (y esto contradice a la definición admitida anteriormente de recinto simplemente conexo). Por consiguiente, en el exterior de γ tiene que haber tanto puntos pertenecientes al recinto G como puntos que no pertenecen a G . De aquí se deduce que en el exterior de γ tiene que haber puntos frontera de G . En efecto, suponiendo lo contrario haríamos la conclusión de que el exterior de γ se dividiría en dos conjuntos no vacíos sin puntos comunes: uno de puntos exteriores y otro de puntos interiores a G , sin contener ninguno de ellos puntos de acumulación del otro, lo cual contradice a la conexividad del exterior de γ . Ahora es fácil demostrar que el interior de γ pertenece totalmente a G . En efecto, en caso contrario en el interior de γ habría puntos pertenecientes a G y puntos no pertenecientes a G , y aplicando los mismos razonamientos que antes hallaríamos que entre éstos habría puntos frontera del recinto G . Luego obtendríamos de aquí que toda la frontera de G quedaría dividida en dos conjuntos no vacíos sin puntos comunes: uno en el interior y otro en el exterior de γ . Como ninguno de éstos puede contener puntos de acumulación del otro (un punto del interior de γ no puede ser de acumulación para el exterior y viceversa), llegaríamos a la conclusión de que la frontera del recinto G es un conjunto desconexo, lo cual contradice a la hipótesis hecha de que el recinto G es simplemente conexo. Así, pues, el interior de γ pertenece totalmente a G , y nuestra afirmación queda demostrada.

Se puede demostrar, sin dificultad alguna, que es simplemente conexo cualquier recinto acotado que posea la propiedad de que

para cualquier curva de Jordan cerrada γ y perteneciente a G el interior de γ también pertenece a G .

Entre los recintos no acotados, para los cuales no hemos definido por ahora los conceptos de simple o múltiplemente conexos, hay algunos que poseen la propiedad indicada de los recintos simplemente conexos: si una curva de Jordan cerrada pertenece al recinto, entonces su interior también pertenece al mismo. A tal género de recintos también denominaremos simplemente conexos, extendiendo para éstos la definición primaria de recinto simplemente conexo.

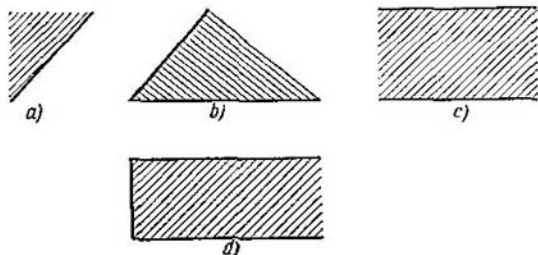


FIG. 5

Por lo tanto, serán simplemente conexos: el plano, un semiplano (fig. 5, a), un ángulo (fig. 5, b), una franja (fig. 5, c), una semifranja (fig. 5, d), etc. Por el contrario, el exterior de un círculo, por ejemplo, no será un recinto simplemente conexo según esta última definición.

4.5. Demostraremos ahora unas cuantas proposiciones auxiliares a las que habrá que hacer referencia a continuación. *)

Hagamos primero la siguiente observación general: si z_0 es un punto del recinto G y Δ es su distancia hasta la frontera de este recinto (cuando la frontera sea un conjunto vacío, supondremos que $\Delta = \infty$, entonces el círculo $|z - z_0| < \Delta$ está contenido en el recinto G . Esta afirmación es evidente cuando la frontera del recinto es un conjunto vacío, puesto que entonces G coincide con todo el plano y el círculo indicado también coincide con todo el plano. Supongamos que la frontera del recinto G no es un conjunto vacío y, por consiguiente, $\Delta < \infty$. El círculo $|z - z_0| < \Delta$ no puede contener puntos frontera de G , puesto que todos ellos están situados fuera de este círculo o en la circunferencia $|z - z_0| = \Delta$ (al menos un punto frontera está situado en esta circunferencia). Pero éste

*) El lector puede omitirlas durante la primera lectura, aludiendo a las proposiciones de este apartado solamente cuando encuentre referencias en la exposición ulterior.

no puede contener tampoco puntos exteriores a G , puesto que en el segmento rectilíneo z_0z_1 , que une el punto interior z_0 con el punto exterior z_1 , tiene que haber al menos un punto frontera. Así, pues, el círculo $|z - z_0| < \Delta$ está contenido en el recinto G .

a) *Dos puntos cualesquiera del recinto pueden unirse mediante una curva de Jordan y, en particular, mediante una línea quebrada.*

Sean z' y z'' dos puntos del recinto G . Según la propiedad conocida del recinto, demostrada en el ap. 4.2, existe una curva continua L : $z = f(t)$ ($\alpha \leq t \leq \beta$), que pertenece a G y une los puntos z' y z'' . Designemos mediante Δ ($\Delta > 0$) la distancia entre la curva L y la frontera del recinto G . Como la función $f(t)$ es continua, el segmento $[\alpha, \beta]$ puede dividirse por los puntos $t_0 = \alpha < t_1 < t_2 < \dots < t_n = \beta$ en partes tan pequeñas, de modo que se cumplan las desigualdades

$$|f(t_{j+1}) - f(t_j)| < \Delta \quad (j = 0, 1, \dots, n-1).$$

Como todo círculo $|z - f(t_j)| < \Delta$ pertenece al recinto G (pues la distancia desde $f(t_j)$ hasta la frontera del recinto G no es menor que Δ) y contiene el punto $f(t_{j+1})$, los segmentos de las rectas Δ_j , que unen $f(t_j)$ con $f(t_{j+1})$ ($j = 0, 1, \dots, n-1$), también pertenecen a G . Por lo tanto, la quebrada Λ con los lados $\Delta_0, \Delta_1, \dots, \Delta_{n-1}$, que une z' con z'' , está situada en el recinto G . Si ésta no se corta consigo mismo, es decir, si carece de puntos múltiples, nuestra afirmación queda demostrada. En caso contrario, existen dos lados Δ_j y Δ_k ($0 \leq j < k$) que tienen un punto común z_0 ; además, cuando los lados son contiguos ($k = j + 1$), se puede suponer que este punto es distinto del punto final del lado Δ_j (que es el punto inicial del lado Δ_{j+1}). Si $k > j + 1$, suprimiendo de Λ la quebrada cerrada que consta: de la parte del lado Δ_j desde el punto z_0 hasta su punto final, de todos los lados de la quebrada Λ cuyos subíndices estén comprendidos entre j y k , y finalmente, de la parte del lado Δ_k desde el punto inicial de este lado hasta el punto z_0 , obtenemos una nueva quebrada Λ' que une como anteriormente z' con z'' , pero que contiene por lo menos un lado menos que Λ . Si $k = j + 1$, entonces uno de los lados Δ_j y Δ_{j+1} , que tienen el punto común z_0 , además de otro punto común, como lo es el punto final de Δ_j y el punto inicial de Δ_{j+1} , tiene que ser parte del otro lado. Por ejemplo, supongamos que Δ_j es una parte de Δ_{j+1} . Entonces, de la quebrada Λ suprimimos la quebrada cerrada que consta de Δ_j y de la parte del lado Δ_{j+1} desde el punto inicial de este lado hasta el punto inicial del lado Δ_j . Resulta una nueva quebrada Λ' que une z' con z'' y que contiene uno o dos lados menos que Λ . Así, pues, en todos los casos, cuando la quebrada Λ se corte consigo misma, es posible sustituir dicha quebrada por otra Λ' o Λ'' , situada de nuevo en el recinto G y que une los mismos puntos z' y z'' , pero

que contiene por lo menos un lado menos que Λ . Repitiendo una cantidad finita de veces esta operación de supresión de quebradas cerradas obtendremos, finalmente, una quebrada λ que carecerá de intersecciones consigo misma (posiblemente, compuesta de un solo lado) y que unirá z' con z'' dentro del recinto G . Con esto, se termina la demostración de nuestra afirmación.

b) Para cada conjunto de puntos cerrado y acotado F , perteneciente a un recinto G , se puede señalar otro conjunto cerrado y acotado E , perteneciente también a G , que contiene a cada punto de F junto con su entorno cerrado de un radio fijado $\rho > 0$.

Esta proposición confirma la posibilidad de inmersión de todo conjunto cerrado F de puntos del recinto G en el interior de otro conjunto cerrado E , perteneciente también a este recinto.

Supongamos primero que G es todo el plano. Si $|z| \leq \rho$ es un círculo que contiene a todos los puntos del conjunto F , se puede tomar por E el círculo cerrado de doble radio: $|z| \leq 2\rho$. Evidentemente, se cumple todo lo que se pide en el teorema.

Supongamos ahora que G no es todo el plano. Entonces, la frontera del recinto G no es un conjunto vacío y se puede hablar de la distancia de F hasta la frontera del recinto G . Designemos esta distancia con d ($d > 0$) y consideremos el conjunto E de aquellos puntos del plano cuyas distancias hasta el conjunto F no son superiores a $\frac{d}{2}$.

Está claro que todo punto del conjunto F es interior con respecto de E , y además, junto con cada punto $z_0 \in F$ pertenece también al conjunto E el círculo $|z - z_0| \leq \frac{d}{2}$. En virtud de la observación hecha al comienzo del presente apartado, el conjunto E mismo pertenece al recinto G .

No queda más que demostrar que E es un conjunto cerrado. Sea z' un punto de acumulación del conjunto E y supongamos que $\{z_n\}$ es una sucesión de puntos de E , convergente hacia z' . Como la distancia del punto al conjunto es continua, la sucesión de números $\rho(z_n, F)$ converge hacia el número $\rho(z', F)$, y como $\rho(z_n, F) \leq \frac{d}{2}$ (debido a la definición del conjunto E), se tiene también: $\rho(z', F) \leq \frac{d}{2}$. De aquí se deduce que z' pertenece a E . Así, pues, el teorema queda demostrado por completo. El número indicado ρ se puede tomar igual a $\frac{d}{2}$.

c) Sea E un subconjunto abierto no vacío del recinto G , que posea la propiedad de que cada uno de sus puntos de acumulación pertenecientes al recinto G , pertenece también a E ; entonces, E necesariamente coincide con todo el recinto G .

D e m o s t r a c i ó n. Supongamos que el teorema no es justo. Entonces en el recinto G existen puntos no pertenecientes a E . Designemos el conjunto de tales puntos mediante E_1 . Como el recinto es conexo, al menos uno de los conjuntos E o E_1 tiene que contener puntos de acumulación del otro. Tal conjunto no puede ser E , puesto que éste es abierto (todos los puntos que figuran en un entorno suficientemente pequeño de un punto de E , también pertenecen a E). Por lo tanto tiene que existir un punto de E_1 que es de acumulación para E . Según la hipótesis del teorema, tal punto tiene que pertenecer a E . Hemos llegado a una contradicción con el hecho de que este punto pertenece a E_1 . De esta contradicción se deduce que el teorema es justo.

Obsérvese que las condiciones del teorema se cumplen particularmente también cuando para cada punto $z_0 \in E$ el entorno del mismo, de radio $d(z_0)$, igual a la distancia desde z_0 hasta la frontera del recinto G , también pertenece a E . En efecto, sea z_1 un punto de acumulación de E y sea $d(z_1)$ su distancia hasta la frontera del recinto G . Entonces, para el punto $z_0 \in E$, situado en el círculo $|z - z_1| < \frac{d(z_1)}{2}$, la distancia $d(z_0)$ hasta la frontera del recinto G será mayor que $\frac{d(z_1)}{2}$, y por consiguiente, el círculo $|z - z_0| < d(z_0)$, que por la hipótesis pertenece a E , contendrá al punto z_1 . Así, pues, todo punto $z_1 \in G$, que es de acumulación para E , pertenece a E . Por lo tanto, se cumplen las condiciones del teorema, y $E = G$.

L e m a d e H e i n e — B o r e l. *Sea F un conjunto cerrado y acotado, y sea $\{K\}$ un sistema de círculos que cubren a F , en el sentido de que cada punto de F es interior al menos para uno de los círculos de K . En estas condiciones, existe un número finito de círculos: K_1, K_2, \dots, K_n ($n \geq 1$), pertenecientes al sistema dado de círculos, que cubren a F .*

Demostremos esta proposición por reducción a lo absurdo. Sea D_1 un cuadrado con los lados paralelos a los ejes coordenados y con el centro en el origen de coordenadas, que contenga al conjunto F . Los ejes coordenados dividen a éste en cuatro cuadrados iguales, que contienen en el interior y en los lados algunos subconjuntos del conjunto F (algunos de estos cuatro subconjuntos pueden ser vacíos). Si el teorema fuese justo para cada uno de estos subconjuntos y cada uno de ellos se cubriese con un número finito de círculos del sistema $\{K\}$, lo mismo sería cierto también para todo el conjunto F . Por esto, negando la justeza del teorema para todo el conjunto F , tenemos que negarla también al menos para uno de los subconjuntos indicados del conjunto F . Sea D_2 el cuadrado que contiene a este subconjunto. Dividiéndolo en cuatro cuadrados iguales obtenemos

un nuevo cuadrado D_3 que contiene algún subconjunto del conjunto F y para el cual no se cumple el teorema. Reiterando indefinidamente estos razonamientos obtendremos una sucesión de cuadrados encajados:

$$D_1 \supset D_2 \supset \dots \supset D_n \supset \dots,$$

las longitudes de cuyos lados, evidentemente, tienden a cero, y cada cuadrado contiene algún subconjunto del conjunto F , al cual no es aplicable el teorema que demostramos.

Designemos con ζ el punto del plano al cual se ciñe la sucesión $\{D_n\}$. En virtud de que cada uno de estos cuadrados D_n contiene un conjunto infinito de puntos de F (si hubiera un número finito de éstos hallaríamos inmediatamente un número finito de círculos del sistema $\{K\}$ que cubriría a este conjunto), ζ es un punto de acumulación de F . En efecto, cualquier entorno del mismo contiene todos los cuadrados, comenzando desde uno de ellos, y por consiguiente, contiene un conjunto infinito de puntos de F . Como F es un conjunto cerrado, resulta que ζ pertenece a F y, por lo tanto, según la hipótesis del teorema, existe un círculo del sistema $\{K\}$ que contiene al punto ζ . Este mismo círculo contiene a todo cuadrado D_n cuyo subíndice n sea suficientemente grande, de donde resulta que contiene al subconjunto del conjunto F que pertenece a D_n . Esto está en contradicción con la hipótesis, según la cual ninguno de los subconjuntos indicados puede cubrirse con un número finito de círculos del sistema $\{K\}$. Con esta contradicción se termina la demostración del teorema.

Obsérvese que en las aplicaciones de este teorema el sistema $\{K\}$ consta generalmente de círculos con centros en puntos del conjunto F , es decir, de una colección de entornos de puntos de este conjunto. Naturalmente, para un sistema finito de entornos que cubren a F , solamente un número finito de puntos del conjunto F estarán situados en los centros de los entornos; los demás puntos de F estarán situados en los círculos K_1, \dots, K_n , sin coincidir con sus centros.

§ 5. EL INFINITO. PROYECCION ESTEREOGRAFICA Y PLANO AMPLIADO

5.1. En la teoría de las funciones de variable compleja desempeña un papel muy importante el infinito (∞), que lo consideraremos como un número complejo impropio. Para definir adecuadamente este concepto introduciremos previamente una interpretación de los números complejos que es muy natural en las cuestiones de carácter analítico. Examinemos, con este fin, todas las sucesiones convergentes posibles de números complejos y distribuyámoslas

en clases, colocando dos sucesiones en una misma clase cuando, y sólo cuando, sus límites son iguales.

Esta clase la designaremos con el símbolo α , donde α es el número complejo hacia el cual convergen todas las sucesiones de la clase, y la llamaremos número complejo propio. En lo que se refiere a cada una de las sucesiones que convergen hacia α , diremos que ella pertenece a α (sobreentendiéndose que α es una clase de sucesiones).

Si α y β son dos clases (distintas o iguales) y $\{u_n\}$ y $\{v_n\}$ son dos sucesiones pertenecientes a α y β , respectivamente, de modo que $\lim_{n \rightarrow \infty} u_n = \alpha$ y $\lim_{n \rightarrow \infty} v_n = \beta$, entonces las sucesiones $\{u_n + v_n\}$, $\{u_n - v_n\}$, $\{u_n \cdot v_n\}$ y $\left\{\frac{u_n}{v_n}\right\}$ (en esta última se supone que $\beta \neq 0$ y que todos los números $v_n \neq 0$) convergen hacia los números $\alpha + \beta$, $\alpha - \beta$, $\alpha \cdot \beta$ y $\frac{\alpha}{\beta}$, respectivamente, es decir, pertenecen a las clases designadas por los mismos símbolos. Tomando en lugar de $\{u_n\}$ y $\{v_n\}$ otras sucesiones $\{u'_n\} \in \alpha$ y $\{v'_n\} \in \beta$, obtenemos de nuevo que

$$\{u'_n \pm v'_n\} \in \alpha \pm \beta, \quad \{u'_n \cdot v'_n\} \in \alpha \cdot \beta \quad \text{y} \quad \left\{\frac{u'_n}{v'_n}\right\} \in \frac{\alpha}{\beta}.$$

O sea, las sucesiones que se obtienen al efectuar una operación algebraica sobre los términos de unas sucesiones arbitrarias de dos clases dadas α y β , pertenecen a una clase determinada, respectivamente: $\alpha + \beta$, $\alpha - \beta$, $\alpha \cdot \beta$ y $\frac{\alpha}{\beta}$. Por esto, resulta natural denominar a estas últimas clases suma, resta, producto y cociente de las clases α y β . Evidentemente, con tal definición, el conjunto de todas las clases representa un campo (o cuerpo conmutativo), que es isomorfo al campo de los números complejos: se obtiene el isomorfismo poniendo en correspondencia a la clase de todas las sucesiones convergentes hacia α , el número mismo α . Hemos obtenido la representación que necesitábamos del campo de los números complejos: *el número complejo α puede interpretarse como la clase de todas las sucesiones que convergen hacia α .*

Para representar geoméricamente a las clases usaremos como antes los puntos del plano. Una sucesión de números complejos $\{u_n\}$ pertenece a la clase α cuando, y sólo cuando, el (único) punto de acumulación de la sucesión de puntos que representan los números u_n , coincide con α . Ahora construiremos el conjunto ampliado de números complejos adjuntando a los números complejos propios (a las clases de sucesiones convergentes) un número impropio: una nueva clase de sucesiones. Precisando, reuniremos en una clase todas aquellas sucesiones $\{u_n\}$ de números complejos, para cada una de las cuales y para cualquier

positivo $M > 0$ se puede señalar un N tal (dependiente de la sucesión y del número M), que $|u_n| > M$ para $n > N$. A esta clase la llamaremos número complejo impropio o infinito y la designaremos con el símbolo ∞ . En lo que se refiere a cada una de las sucesiones $\{u_n\}$ que pertenecen al mismo, se dirá que converge hacia el infinito y escribiremos: $\lim_{n \rightarrow \infty} u_n = \infty$.

Análogamente al modo con que se introdujeron las operaciones sobre los números complejos propios, se pueden establecer también las operaciones para el conjunto ampliado de números complejos. Supongamos de nuevo que α, β son dos clases (propias o impropias) y que $\{u_n\}, \{v_n\}$ son unas sucesiones cualesquiera pertenecientes a ellas. Si la sucesión $\{u_n + v_n\}$ pertenece a cierta clase γ de nuestro conjunto ampliado, y el cambio de $\{u_n\} \in \alpha$ y $\{v_n\} \in \beta$ por cualesquiera otras sucesiones $\{u'_n\} \in \alpha$ y $\{v'_n\} \in \beta$ no influye en el resultado, es decir, $\{u'_n + v'_n\}$ pertenece, igual que anteriormente, a la clase γ , entonces γ se llama suma de α y β : $\gamma = \alpha + \beta$. Pero si no existe en general una clase a la que podría pertenecer la sucesión $\{u_n + v_n\}$, o si esta clase puede variar al sustituir $\{u_n\}$ por $\{u'_n\}$ y $\{v_n\}$ por $\{v'_n\}$, entonces se dirá que α y β no tienen suma, es decir, que la expresión $\alpha + \beta$ carece de sentido. Evidentemente, el concepto de suma introducido de este modo para el conjunto ampliado de números complejos concuerda con el anterior, cuando α y β eran números propios (en este caso, siempre existe la suma). Si al menos una de las clases α y β es impropia, llegamos a los siguientes resultados:

1) $\alpha + \infty = \infty + \alpha = \infty$ (α es un número propio), la expresión $\infty + \infty$ carece de sentido.

Para convencerse de lo último es suficiente tomar, por ejemplo, dos sucesiones de la clase ∞ : 1, 3, 3, 5, 5, 7, 7, 9, 9, . . . y -1, -2, -3, -4, -5, -6, -7, -8, -9, . . . Sumándolas término a término, obtenemos la sucesión 0, 1, 0, 1, 0, 1, 0, 1, . . ., que, evidentemente, no pertenece a ninguna clase propia y tampoco a la clase impropia.

Del mismo modo se consideran las otras operaciones también: la resta, multiplicación y división de las clases α y β . Lo único que en el caso de la división hay que exigir es que sean distintos de cero los términos de la sucesión que figura como divisor.

Se obtienen los siguientes resultados:

2) $\alpha - \infty = \infty - \alpha = \infty$ (α es un número propio), la expresión $\infty - \infty$ carece de sentido;

3) $\alpha \cdot \infty = \infty \cdot \alpha = \infty$ (α es un número propio, distinto de cero), $\infty \cdot \infty = \infty$;

4) $\frac{\alpha}{\infty} = 0$, $\frac{\infty}{\alpha} = \infty$ (α es un número propio), la expresión $\frac{\infty}{\infty}$ carece de sentido;

5) $\frac{\alpha}{0} = \infty$ (α es un número propio, distinto de cero), la expresión $\frac{0}{0}$ carece de sentido.

5.2. Como observó por primera vez Riemann, para la representación geométrica del conjunto ampliado de los números complejos es conveniente emplear la esfera. Consideremos alguna esfera S de radio 1 y un plano Π que pase por su centro O (fig. 6). Tomando en Π un sistema de coordenadas rectangulares xy con el origen en O , se pueden representar todos los números complejos propios en el plano Π . A cada uno de ellos ponemos en correspondencia un punto

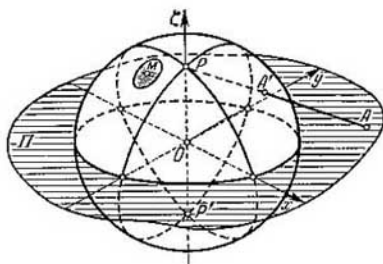


FIG. 6

determinado de la esfera. Con este fin, trazamos por O un diámetro PP' perpendicular al plano Π y unimos uno de sus extremos P con un punto arbitrario $A \in \Pi$ mediante un segmento recto. Este segmento (o su continuación) se cortará con la esfera en cierto punto A' , distinto de P . Al punto A ponemos en correspondencia este punto de la esfera. Evidentemente, de este modo se establece una correspondencia biunívoca entre los puntos del plano y todos los puntos de la esfera (a excepción del punto P). Este método de hacer corresponder a los puntos del plano los de la esfera, se denomina *proyección estereográfica* (de la esfera sobre el plano o del plano sobre la esfera). Si el punto A representaba en el plano un número complejo determinado α , aquí consideraremos a su punto correspondiente A' de la esfera como la imagen del mismo número complejo.

Para mayor facilidad en la expresión emplearemos la terminología geográfica, denominando ecuador de la esfera a la circunferencia obtenida en la intersección de la esfera S con el plano Π ; los puntos P y P' se llamarán polo norte y polo sur, respectivamente; los círculos máximos que pasan por los puntos P y P' se llamarán

meridianos y, en particular, el meridiano situado en el plano POx se llamará meridiano inicial, etc. Entonces, fácilmente se observa que los puntos del plano Π situados dentro de la circunferencia unidad (la cual coincide con el ecuador), se representan en la esfera por puntos del hemisferio sur (el cual contiene a P'), mientras que los puntos situados fuera de la circunferencia unidad se representan por puntos del hemisferio norte (el cual contiene a P). Del mismo modo, los puntos del semiplano superior ($y > 0$) se representan por

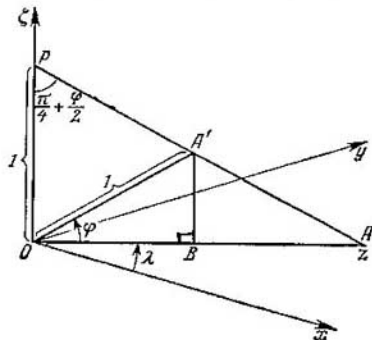


FIG. 7

puntos del hemisferio oriental (con el cual se corta el semieje positivo Oy), mientras que los puntos del semiplano inferior ($y < 0$) se representan por puntos del hemisferio occidental, etc.

Introduzcamos en S las coordenadas esféricas (geográficas): la latitud φ , que se mide desde el ecuador, desde 0 hasta $\frac{\pi}{2}$ en el hemisferio norte y desde 0 hasta $-\frac{\pi}{2}$ en el hemisferio sur, y la longitud λ , que se mide desde el meridiano inicial (o con más precisión, desde su punto de intersección con el semieje positivo x) en dirección positiva desde 0 hasta π (inclusive) y en dirección negativa desde 0 hasta $-\pi$ (exclusive).

De la fig. 7 se deduce que en la proyección estereográfica

$$\arg z = \lambda \text{ y } |z| = \operatorname{tg} \left(\frac{\pi}{4} + \frac{\varphi}{2} \right).$$

Por consiguiente, el punto de la esfera de coordenadas λ y φ representa al número complejo

$$z = \operatorname{tg} \left(\frac{\pi}{4} + \frac{\varphi}{2} \right) (\cos \lambda + i \operatorname{sen} \lambda). \quad (5.2:1)$$

Recíprocamente: el número complejo $z \neq 0$ se representa por el punto de la esfera de coordenadas $\lambda = \arg z$ y $\varphi = -2 \operatorname{arctg} |z| - \frac{\pi}{2}$.

Señalemos también las fórmulas que relacionan las coordenadas cartesianas de los puntos de la esfera A' (ξ , η , ζ) y los puntos del plano A (x , y , 0) en la proyección estereográfica.

Como

$$\xi = \cos \varphi \cdot \cos \lambda, \quad \eta = \cos \varphi \cdot \operatorname{sen} \lambda, \quad \zeta = \operatorname{sen} \varphi$$

y

$$\begin{aligned} \cos \varphi &= \operatorname{sen} \left(\frac{\pi}{2} + \varphi \right) = \frac{2 \operatorname{tg} \left(\frac{\pi}{4} + \frac{\varphi}{2} \right)}{1 + \operatorname{tg}^2 \left(\frac{\pi}{4} + \frac{\varphi}{2} \right)} = \frac{2|z|}{1 + |z|^2}, \\ \operatorname{sen} \varphi &= -\cos \left(\frac{\pi}{2} + \varphi \right) = -\frac{1 - \operatorname{tg}^2 \left(\frac{\pi}{4} + \frac{\varphi}{2} \right)}{1 + \operatorname{tg}^2 \left(\frac{\pi}{4} + \frac{\varphi}{2} \right)} = \frac{|z|^2 - 1}{1 + |z|^2}, \end{aligned}$$

se tiene

$$\left. \begin{aligned} \xi &= \cos \varphi \cos \lambda = \frac{2|z| \cos \lambda}{1 + |z|^2} = \frac{2x}{x^2 + y^2 + 1}, \\ \eta &= \cos \varphi \operatorname{sen} \lambda = \frac{2|z| \operatorname{sen} \lambda}{1 + |z|^2} = \frac{2y}{x^2 + y^2 + 1}, \\ \zeta &= \frac{x^2 + y^2 - 1}{x^2 + y^2 + 1}. \end{aligned} \right\} \quad (5.2:2)$$

Por otra parte,

$$\begin{aligned} z &= \operatorname{tg} \left(\frac{\pi}{4} - \frac{\varphi}{2} \right) (\cos \lambda - i \operatorname{sen} \lambda) = \frac{1 + \operatorname{tg} \frac{\varphi}{2}}{1 - \operatorname{tg} \frac{\varphi}{2}} (\cos \lambda + i \operatorname{sen} \lambda) = \\ &= \frac{\cos \frac{\varphi}{2} - \operatorname{sen} \frac{\varphi}{2}}{\cos \frac{\varphi}{2} + \operatorname{sen} \frac{\varphi}{2}} (\cos \lambda - i \operatorname{sen} \lambda) = \frac{\cos \varphi}{1 - \operatorname{sen} \varphi} (\cos \lambda + i \operatorname{sen} \lambda) = \frac{\xi + i\eta}{1 - \zeta}, \end{aligned}$$

de donde

$$x = \frac{\xi}{1 - \zeta}, \quad y = \frac{\eta}{1 - \zeta}. \quad (5.2:3)$$

Las fórmulas obtenidas resuelven por completo el problema de la proyección estereográfica.

De ellas, en particular, se deduce la propiedad homocíclica de la proyección estereográfica: en esta proyección,

a las circunferencias de la superficie de la esfera corresponden en el plano también circunferencias o rectas (estas últimas corresponden a las circunferencias de la esfera que pasan por su polo norte). En efecto, cualquier circunferencia de la esfera S se obtiene en la intersección de S con cierto plano

$$A\xi + B\eta + C\zeta - D = 0. \quad (5.2:4)$$

De aquí, en virtud de las fórmulas (5.2:2) se deduce que los puntos correspondientes del plano satisfacen a la ecuación

$$Ax + By + \frac{1}{2}(C + D)(x^2 + y^2) + \frac{1}{2}(D - C) = 0. \quad (5.2:5)$$

Esta última es la ecuación de una circunferencia si $C + D \neq 0$ y la ecuación de una recta si $C + D = 0$. Pero $C + D$ es el resultado de sustituir en el segundo miembro de la ecuación (5.2:4) las coordenadas del polo norte $(0, 0, 1)$. Así, pues, en el plano Π se obtiene una recta o una circunferencia, según que pase por el polo norte o no la circunferencia de la esfera.

Señalemos, en particular, que cuando el plano (5.2:4) sea paralelo al plano del ecuador, es decir, cuando la circunferencia de la esfera sea uno de los paralelos, entonces $A = B = 0$ y en el plano resulta una circunferencia con el centro en el origen de coordenadas:

$$(C - D)(x^2 + y^2) = C - D.$$

Si el plano (5.2:4) pasa por el eje PP' , es decir, si la circunferencia de la esfera es uno de los meridianos, entonces $C = D = 0$ y en el plano resulta una recta que pasa por el origen de coordenadas:

$$Ax + By = 0.$$

Claro, estos resultados podrían observarse también partiendo directamente de la definición de proyección estereográfica.

Toda circunferencia γ de la superficie de la esfera, que no pase por un punto dado M' de la esfera, divide la esfera en dos partes, una de las cuales contiene a M' y la otra no lo contiene; a la primera de ellas la llamaremos *entorno* del punto M' .

Una vez definido el concepto de entorno de un punto de la esfera, se puede generalizar inmediatamente el concepto de punto de acumulación, de conjunto cerrado y abierto, de curva continua, de recinto, etc, para los conjuntos de puntos considerados en S .

Por ejemplo: un punto de la esfera se llamará punto de acumulación de un conjunto dado E cuando, y sólo cuando, cualquier entorno del mismo contiene un conjunto infinito de puntos pertenecientes a E . Respecto de cada terna de funciones reales $\xi = \varphi(t)$, $\eta = \psi(t)$, $\zeta = \chi(t)$ del parámetro real t , definidas y continuas en un segmento $a \leq t \leq b$, que cumplen además la condición $[\varphi(t)]^2 +$

$+ |\psi(t)|^2 + |\chi(t)|^2 = 1$, se dirá que ésta determina una curva continua (esférica), etc.

Consideremos una sucesión de números complejos $\{u_n\}$, convergente hacia el número α (propio), el cual es el afijo del punto A . Si los números u_n son los afijos de los puntos U_n , entonces A es un punto de acumulación de la sucesión $\{U_n\}$ (y además el único). Fácilmente se observa que el punto A' , que representa en la esfera el punto α , es también un punto de acumulación (y además, el único) de la sucesión de puntos de la esfera U_n , que representan los números u_n .

Supongamos ahora que la sucesión de números complejos $\{v_n\}$ pertenece a la clase impropia ∞ . Entonces, para cualquier $M > 0$ los puntos V_n , que representan a los números v_n en el plano, comenzando desde cierto subíndice $n > N$, estarán situados fuera del círculo de radio M . Los puntos correspondientes V'_n de la esfera se situarán en el entorno del polo norte, limitado por el paralelo cuya proyección sobre el plano es la circunferencia $|z| = M$. De aquí se deduce que el polo norte será el único punto de acumulación para cualquier sucesión de imágenes de puntos, que pertenezca a la clase impropia.

Recíprocamente: toda sucesión de puntos de la esfera, para la cual el polo norte sea el único punto de acumulación, representa cierta sucesión de la clase impropia. Por cierto, todo esto se deduce de la relación (5.2:1) hallada anteriormente: $|z| = \operatorname{tg}\left(\frac{\pi}{4} + \frac{\varphi}{2}\right)$. En efecto, para la sucesión $\{z_n\}$ la relación

$$\lim_{n \rightarrow \infty} |z_n| = +\infty$$

es equivalente a la relación

$$\lim_{n \rightarrow \infty} \varphi_n = \frac{\pi}{2}.$$

De acuerdo a esto, aquí consideraremos el polo norte de la esfera como la imagen geométrica del número impropio ∞ .

5.3. Ya vimos que mediante la proyección estereográfica se obtiene una representación biunívoca (o aplicación biyectiva) y continua de la esfera, a excepción del polo norte, sobre todo el plano *), de modo que la esfera puede servir para representar los números complejos propios igual que el plano. La existencia en la esfera

*) Es decir, que a cada sucesión de puntos de la esfera, convergente hacia el punto A' , distinto éste del polo norte, corresponde en el plano una sucesión de puntos, convergente hacia el punto A , que es la proyección de A' , y viceversa.

de un punto más, precisamente del polo, que representa naturalmente al número complejo impropio, permite considerar a la esfera como un modelo de plano ampliado, obtenido del plano finito u ordinario mediante la adunción de un punto más: del punto impropio o punto del infinito. A continuación, efectuando la proyección estereográfica, representaremos el plano ampliado (la esfera) en forma de un plano finito, completándolo mentalmente con un punto impropio: con el punto del infinito del plano. En este caso, el entorno del punto del infinito se representará por el exterior de un círculo arbitrario, por ejemplo, de un círculo con centro en el origen de coordenadas.

Obsérvese que no se debe confundir el plano ampliado con el plano proyectivo que se obtiene del plano finito por adunción de un conjunto infinito de puntos impropios que forman una recta impropia. Para construir un modelo de plano proyectivo, es suficiente tomar una esfera que sea tangente al plano en el polo sur. Proyectando sobre el plano, desde el centro de la esfera, los puntos del hemisferio sur, se establece una correspondencia biunívoca y continua entre todos los puntos del plano y los puntos del hemisferio sur (excluyendo de este último los puntos del ecuador). A cada recta del plano le corresponderá en el hemisferio la mitad de un círculo máximo; al conjunto de rectas paralelas entre sí, le corresponderá el conjunto de mitades de los círculos máximos que pasan por un mismo par de puntos diametralmente opuestos del ecuador. Este par de puntos se considera como la imagen de un punto del infinito del plano proyectivo, que es común para todas las rectas que llevan la misma dirección.

A todos los puntos del infinito posibles del plano proyectivo les corresponderá el conjunto de todos los pares de puntos diametralmente opuestos del ecuador. Este conjunto se considera como la imagen de la recta del infinito.

Por lo tanto, el modelo del plano proyectivo se obtiene del hemisferio identificando cada par de puntos diametralmente opuestos del ecuador.

Señalemos que *en el plano ampliado cualquier sucesión de puntos posee al menos un punto de acumulación*. En efecto, si la sucesión está acotada, la existencia de un punto de acumulación se deduce del teorema de Bolzano-Weierstrass. Si no es acotada, el exterior de cualquier círculo, es decir, cualquier entorno del punto del infinito, contiene un conjunto infinito de términos de la sucesión. De aquí se deduce que el punto del infinito es un punto de acumulación para cada sucesión no acotada.

En resumen, en el plano ordinario existen sucesiones que carecen de puntos de acumulación, mientras que en el plano ampliado (en la esfera) cualquier sucesión posee un punto de acumulación. Esta

circunstancia se expresa diciendo que el plano no es compacto y que el plano ampliado es compacto.

Generalizando el concepto de límite, llamaremos convergente a cualquier sucesión de puntos del plano ampliado que contenga solamente un punto de acumulación. Este último se llamará límite de la sucesión.

Si G es un conjunto infinito no acotado del plano ampliado, el punto del infinito será un punto de acumulación para el mismo. Debido a esto, el conjunto G no puede ser cerrado si a él no pertenece el punto ∞ .

Por ejemplo, el conjunto $1, 2, \dots, n, \dots$ no es cerrado en el plano ampliado, mientras que en el plano finito es cerrado.

Si G es un recinto no acotado, el punto ∞ puede ser o punto frontera, o punto interior al mismo. Como ejemplos del primer caso se pueden señalar: un semiplano, un ángulo, cualquiera de los dos recintos limitados por una parábola o por una de las ramas de una hipérbola, una franja o una semifranja, etc.

En el segundo caso tiene que existir un entorno del punto ∞ que pertenezca al recinto y, por consiguiente, que no contenga puntos frontera del recinto G . Por lo tanto, en este caso, la frontera del recinto G es un conjunto acotado. Puede servir de ejemplo el exterior G de cualquier curva de Jordan Γ , por ejemplo de una circunferencia, añadiendo también a G el punto ∞ . Si esto último no se hiciese, resultaría otro recinto G' , cuya frontera constaría de la curva Γ y del punto aislado del infinito.

La definición de orden de conexión de un recinto en el plano ampliado se introduce independientemente de que el recinto sea acotado o no. Así, pues, un recinto G se llama simplemente conexo en el plano ampliado, si su frontera es un continuo. Esta definición concuerda con la anterior (pág. 61) en el caso de un recinto acotado. También concuerda con la definición generalizada dada en la pág. 64 puesto que la frontera de un semiplano, de un ángulo, de una franja, etc. representa un continuo del plano ampliado. No obstante, en el caso del exterior G de una curva de Jordan Γ , la nueva definición da lugar a otro resultado. Precisamente, el exterior mencionado (si se le añade el punto del infinito) es un recinto simplemente conexo del plano ampliado.

Los recintos que no son simplemente conexos los llamaremos múltiplemente conexos. Por ejemplo, excluyendo del exterior G de una curva de Jordan Γ el punto ∞ , obtenemos un recinto múltiplemente conexo G' . Si la frontera del recinto G puede representarse en forma de un número finito p de continuos, que carecen de puntos comunes dos a dos, entonces se dice que G es de conexión finita, precisamente, de conexión de orden p . Por ejemplo, el exterior de un círculo del cual se ha excluido el punto del infinito, es un recinto conexo de

segundo orden. Si la frontera de G no puede agotarse con un número finito de continuos que carezcan de puntos comunes dos a dos, entonces se dice que el recinto G es de conexión infinita.

Se puede demostrar que la definición admitida de recinto simplemente conexo del plano ampliado es equivalente a la siguiente:

Un recinto G se llama simplemente conexo, si para cualquier curva cerrada de Jordan γ , perteneciente a G , el interior o el exterior de γ pertenece a G .

La demostración de la equivalencia de una y otra definición dejamos a cuenta del lector.

El concepto introducido de punto del infinito permite generalizar un poco el concepto de función de variable compleja. Precisamente ahora el punto del infinito puede estar incluido en el conjunto de valores del argumento o en el conjunto de valores de la función, o bien, finalmente, en uno y en otro.

Debido a esto se puede hablar del valor de la función en el punto del infinito (este valor mismo puede ser finito o infinito) o del valor infinito de la función en un punto finito. El motivo de tal generalización del concepto de función se basa en la continuidad. Además, admitimos la siguiente generalización del concepto de límite para los casos en los cuales el punto z_0 o el límite A , o bien, finalmente, z_0 y A , son impropios:

$$\lim_{z \rightarrow z_0} f(z) = A,$$

si para cualquier sucesión $\{z_n\}$, convergente hacia z_0 , se tiene:

$$\lim_{n \rightarrow \infty} f(z_n) = A.$$

La función se llama continua en el punto z_0 (en sentido generalizado) si

$$\lim_{z \rightarrow z_0} f(z) = f(z_0).$$

Hay dos casos fundamentales en los que es conveniente introducir los valores impropios del argumento o de la función.

1) Si la función $f(z)$ está definida en un conjunto E para el cual el punto del infinito es un punto de acumulación, y $f(z)$ en el conjunto E tiende a un límite determinado A (finito o infinito) cuando z tiende al infinito, entonces es conveniente generalizar también la definición de $f(z)$ para el punto del infinito haciendo $f(\infty) = A$.

Así, por ejemplo, si $f(z)$ es una función racional

$$f(z) = \frac{a_0 + a_1 z + \dots + a_n z^n}{b_0 + b_1 z + \dots + b_m z^m} \quad (a_n \neq 0, b_m \neq 0),$$

entonces existe el límite: $\lim_{z \rightarrow \infty} f(z)$, igual a 0, si $m > n$, igual a $\frac{a_m}{b_m}$, si $m = n$, y finalmente, igual a ∞ , si $m < n$. De acuerdo a esto, hacemos:

$$f(\infty) = \begin{cases} 0 \\ \frac{a_m}{b_m} \\ \infty \end{cases}$$

2) Si la función $f(z)$ está definida en un conjunto E , para el cual z_0 es un punto de acumulación, y $f(z)$ tiende hacia ∞ cuando el punto z del conjunto E tiende hacia z_0 , entonces es conveniente generalizar también la definición de $f(z)$ para el punto z_0 , haciendo $f(z_0) = \infty$.

Así, si z_0 es un cero del polinomio $b_0 + b_1z + \dots + b_mz^m$ y no es un cero del polinomio $a_0 + a_1z + \dots + a_nz^n$, entonces $f(z) = \frac{a_0 + a_1z + \dots + a_nz^n}{b_0 + b_1z + \dots + b_mz^m}$ tiende hacia ∞ cuando $z \rightarrow z_0$, lo cual sirve de base para hacer $f(z_0) = \infty$.

En virtud de la definición admitida, la función $f(z)$ es c o n t i n u a en el sentido general en cada uno de estos casos en el punto correspondiente: en el primer caso, en el punto $z = \infty$ y en el segundo, en el punto z_0 .

5.4. Demostremos que en la proyección estereográfica los ángulos entre las curvas de la esfera son iguales a los ángulos entre sus imágenes en el plano. Sea $M'_0(\xi_0, \eta_0, \zeta_0)$ un punto cualquiera de la esfera (distinto del polo norte) y γ' , una curva continua que pase por el punto M'_0 y que tenga tangente en el mismo. Esto significa que entre las distintas representaciones paramétricas de la curva γ' existe una

$$\xi = \varphi(t), \quad \eta = \psi(t), \quad \zeta = \chi(t) \quad (|\varphi(t)|^2 + |\psi(t)|^2 + |\chi(t)|^2 = 1),$$

tal que las funciones φ , ψ y χ son diferenciables para $t = t_0$, y

$$|\varphi'(t)|^2 + |\psi'(t)|^2 + |\chi'(t)|^2 \neq 0.$$

La imagen del punto M'_0 en la proyección estereográfica será el punto del plano $M_0(x_0, y_0, z_0)$ y la imagen de la curva γ' , la curva γ :

$$x = \frac{\xi}{1-\zeta} = \frac{\varphi(t)}{1-\chi(t)}, \quad y = \frac{\eta}{1-\zeta} = \frac{\psi(t)}{1-\chi(t)}.$$

Para los números $x'(t_0)$ y $y'(t_0)$ que caracterizan la dirección de la tangente en el punto M_0 , se tiene:

$$x' = \frac{\xi'(1-\zeta_0) + \xi'\zeta_0}{(1-\zeta_0)^2}, \quad y' = \frac{\eta'(1-\zeta_0) + \eta'\zeta_0}{(1-\zeta_0)^2}$$

$$y' = \frac{(\xi'^2 + \eta'^2)(1 - \zeta_0)^2 + 2(\xi_0 \xi' + \eta_0 \eta') \cdot \zeta' (1 - \zeta_0) + \zeta'^2 (\xi_0^2 + \eta_0^2)}{(1 - \zeta_0)^4}$$

Pero

$$\xi_0^2 + \eta_0^2 = 1 - \zeta_0^2 \quad \text{y} \quad \xi_0 \xi' + \eta_0 \eta' = -\zeta_0 \zeta';$$

por consiguiente,

$$\begin{aligned} x'^2 + y'^2 &= \frac{(\xi'^2 + \eta'^2)(1 - \zeta_0)^2 - \zeta'^2 (2\xi_0 - 2\zeta_0^2) + \zeta'^2 (1 - \zeta_0^2)}{(1 - \zeta_0)^4} = \\ &= \frac{(\xi'^2 + \eta'^2)(1 - \zeta_0)^2 + \zeta'^2 (1 - \zeta_0)^2}{(1 - \zeta_0)^4} = \frac{\xi'^2 + \eta'^2 + \zeta'^2}{(1 - \zeta_0)^2}. \end{aligned}$$

Considerando ahora dos curvas γ' y γ'_1 que pasen por el punto M'_0 de la esfera y sus imágenes γ y γ_1 en el plano, que pasan por el punto M_0 , para el coseno del ángulo α entre las curvas γ y γ_1 en el punto M_0 hallamos:

$$\begin{aligned} \cos \alpha &= \frac{x'x'_1 + y'y'_1}{\sqrt{x'^2 + y'^2} \sqrt{x_1'^2 + y_1'^2}} = \\ &= \frac{[\xi' (1 - \zeta_0) + \zeta' \xi_0] [\xi_1 (1 - \zeta_0) + \zeta_1 \xi_0] + [\eta' (1 - \zeta_0) + \zeta' \eta_0] [\eta_1 (1 - \zeta_0) + \zeta_1 \eta_0]}{(1 - \zeta_0)^2 \sqrt{\xi'^2 + \eta'^2 + \zeta'^2} \sqrt{\xi_1'^2 + \eta_1'^2 + \zeta_1'^2}} = \\ &= \frac{(\xi' \xi_1 + \eta' \eta_1) (1 - \zeta_0)^2 + (\xi_0 \xi' + \eta_0 \eta') \zeta_1 (1 - \zeta_0)}{(1 - \zeta_0)^2 \sqrt{\xi'^2 + \eta'^2 + \zeta'^2} \sqrt{\xi_1'^2 + \eta_1'^2 + \zeta_1'^2}} + \\ &\quad + \frac{(\xi_0 \xi_1 + \eta_0 \eta_1) \zeta' (1 - \zeta_0) + \zeta' \zeta_1 (\xi_0^2 - \eta_0^2)}{(1 - \zeta_0)^2 \sqrt{\xi'^2 + \eta'^2 + \zeta'^2} \sqrt{\xi_1'^2 + \eta_1'^2 + \zeta_1'^2}}. \end{aligned}$$

Pero $\xi_0^2 + \eta_0^2 = 1 - \zeta_0^2$, $\xi_0 \xi' + \eta_0 \eta' = -\zeta_0 \zeta'$ y del mismo modo,

$$\xi_0 \xi_1 + \eta_0 \eta_1 = -\zeta_0 \zeta_1;$$

por lo tanto

$$\begin{aligned} \cos \alpha &= \frac{(\xi' \xi_1 + \eta' \eta_1) (1 - \zeta_0)^2 - \zeta' \zeta_1 (\zeta_0 - \zeta_0^2) - \zeta' \zeta_1 (\zeta_0 - \zeta_0^2) + \zeta' \zeta_1 (1 - \zeta_0^2)}{(1 - \zeta_0)^2 \sqrt{\xi'^2 + \eta'^2 + \zeta'^2} \sqrt{\xi_1'^2 + \eta_1'^2 + \zeta_1'^2}} = \\ &= \frac{\xi' \xi_1 + \eta' \eta_1 + \zeta' \zeta_1}{\sqrt{\xi'^2 + \eta'^2 + \zeta'^2} \sqrt{\xi_1'^2 + \eta_1'^2 + \zeta_1'^2}} = \cos \alpha'. \end{aligned}$$

Aquí, evidentemente, α' es el ángulo entre las curvas esféricas γ_1 y γ'_1 en el punto M'_0 . Con esto queda demostrado lo que se afirmaba.

La representación (o aplicación) de una superficie (o de un recinto sobre una superficie) sobre otra se llama *conforme* si en ella no varían (se conservan) los ángulos entre las curvas. Se puede decir ahora que la *proyección estereográfica de la esfera sobre el plano es una representación (o aplicación) conforme*.

La proposición demostrada permite sustituir la medición de los ángulos entre las curvas de la esfera por la medición de los ángulos entre sus imágenes en el plano, y viceversa. Consideremos, en parti-

cular, los ángulos formados por las curvas que pasan por el punto del infinito. En el plano, tales curvas (curvas continuas del plano ampliado) se representan en forma de proyecciones de las curvas esféricas que pasan por el polo norte.

Si $\xi = \varphi(t)$, $\eta = \psi(t)$, $\zeta = \chi(t)$, $a \leq t \leq b$, son las ecuaciones paramétricas de una curva continua γ de la esfera, la ecuación de su proyección Γ tendrá la forma

$$z = \frac{\xi + i\eta}{1-\zeta} = \frac{\varphi(t) + i\psi(t)}{1-\chi(t)} = f(t).$$

La función $f(t)$ está definida y es continua para todos los valores de t del segmento $[a, b]$, a excepción de aquellos valores $t = \tau$ para los cuales $\chi(t) = 1$ (y, por consiguiente, $[\varphi(t)]^2 + [\psi(t)]^2 = 1 - [\chi(t)]^2 = 0$, o sea, $\varphi(t) = \psi(t) = 0$). Pero, cuando $t \rightarrow \tau$, el punto de la curva γ tiende hacia el polo norte y, por lo tanto, $f(t) \rightarrow \infty$. Debido a esto, se puede definir $f(t)$ en todo el segmento $[a, b]$, conservando la continuidad (en el sentido generalizado), haciendo $f(\tau) = \infty$.

Supongamos que, en general, se tiene alguna función $z = F(t)$, definida y continua en el sentido generalizado en el segmento $[a, b]$ (que ahora puede ser infinito hacia uno o hacia ambos lados). Entonces, las coordenadas (ξ, η, ζ) de los puntos de la esfera que se proyectan sobre $z = F(t)$ serán, evidentemente, funciones de t , definidas y continuas en $[a, b]$ en sentido ordinario (si el segmento es infinito, por ejemplo, $b = \infty$, estas funciones tendrán unos límites determinados cuando $t \rightarrow \infty$; pasando a otro parámetro: $t' = \arctg t$, podemos reducir el problema al caso de un segmento finito). De aquí se deduce que toda función $z = F(t)$, definida y continua en sentido generalizado en un segmento $[a, b]$, determina una curva continua en el plano ampliado.

Puedon servir de ejemplos: la recta $z = (\alpha t + \beta) + i(\gamma t + \delta)$, donde $\alpha^2 + \gamma^2 \neq 0$ y la función z se hace igual a ∞ para $t = \infty$; la parábola, por ejemplo, $z = \alpha t^2 + \beta t + \gamma + i(\delta t + \varepsilon)$, donde $\alpha^2 + \delta^2 \neq 0$ y la función z se hace igual a ∞ para $t = \infty$; la hipérbola (por ejemplo, $z = \frac{a(1+t^2) + 2bit}{1-t^2}$, donde $a \neq 0$, $b \neq 0$, y la función se hace igual a ∞ para $t = \pm 1$), etc.

Dos curvas planas Γ_1 y Γ_2 forman un ángulo en el vértice en el punto del infinito, si las curvas Γ'_1 y Γ'_2 que les corresponden en la esfera (Γ_1 y Γ_2 son las proyecciones estereográficas de Γ'_1 y Γ'_2) forman un ángulo en el polo norte (es decir, poseen tangentes en este punto). Por definición, el valor de este último ángulo se toma por valor del ángulo entre las curvas Γ_1 y Γ_2 en el punto del infinito.

Consideremos, en particular, dos rayos rectilíneos Γ_1 y Γ_2 que partan de un punto M_0 del plano bajo un ángulo α . A ellos, en la

esfera, les corresponderán, dos arcos de circunferencias Γ'_1 y Γ'_2 , que unen el punto M'_0 de la esfera con el polo norte P . Según lo demostrado en el presente apartado, Γ'_1 y Γ'_2 forman en el punto M'_0 un ángulo igual a α ; pero, en virtud de las conocidas propiedades de la esfera, Γ'_1 y Γ'_2 tienen que formar el mismo ángulo en el punto P , es decir, el ángulo α . De aquí se deduce que los rayos Γ_1 y Γ_2 también forman en el punto del infinito un ángulo igual a α .

Tomemos ahora por Γ_1 y Γ_2 dos rectas paralelas; en la esfera a ellas les corresponderán dos circunferencias Γ'_1 y Γ'_2 que pasan por P y que poseen aquí tangente común. De acuerdo a esto, el ángulo entre Γ'_1 y Γ'_2 , y por consiguiente, también entre Γ_1 y Γ_2 será igual a cero.

El lector tiene que acostumbrarse a considerar la parte del plano comprendida entre dos rayos rectilíneos (un ángulo) o entre rectas paralelas (una franja) como un biángulo con dos ángulos iguales. Uno de los vértices de este biángulo (en el caso del ángulo) o ambos (en el caso de la franja) están situados en el punto del infinito.

5.5. A las transformaciones elementales de la esfera en sí misma, mediante la proyección estereográfica, les corresponden ciertas transformaciones del plano en sí mismo. Hagamos girar la esfera en el ángulo π alrededor del diámetro que lleva la dirección del eje z . Entonces, el hemisferio norte pasará al sur y el sur al norte; en particular, se cambiarán de sitio los polos norte y sur, es decir, los puntos que representan al ∞ y al 0 . Además, se cambiarán de sitio los hemisferios occidental y oriental. El ecuador y el meridiano cero se quedarán en sus sitios. De todo esto se deduce que la transformación considerada de la esfera es equivalente a la siguiente transformación de las coordenadas esféricas:

φ se sustituye por $-\varphi$,

λ se sustituye por $-\lambda$.

Veamos cuál es la transformación del plano que corresponde a esta transformación de la esfera. Sustituyendo en la fórmula (5.2: 1) φ y λ por $-\varphi$ y $-\lambda$, en lugar de z obtenemos el nuevo punto:

$$\zeta = \operatorname{tg} \left(\frac{\pi}{4} - \frac{\varphi}{2} \right) (\cos \lambda - i \operatorname{sen} \lambda) = \frac{1}{\operatorname{tg} \left(\frac{\pi}{4} + \frac{\varphi}{2} \right) (\cos \lambda + i \operatorname{sen} \lambda)} = \frac{1}{z}.$$

Por lo tanto, a la rotación de la esfera en el ángulo π alrededor del eje real le corresponde la transformación del plano expresada por la ecuación

$$\zeta = \frac{1}{z}.$$

En este caso, el exterior de la circunferencia unidad (correspondiente al hemisferio norte) se transformará en el interior de la misma (correspondiente al hemisferio sur). Los semiplanos superior e inferior (correspondientes a los hemisferios occidental y oriental) también se cambian de sitio. La circunferencia unidad (imagen del ecuador) y la parte positiva del eje real (imagen del meridiano cero) se transforman en sí mismos.

Es de particular importancia el hecho de que en la transformación indicada se cambian de lugar el origen de coordenadas y el punto del infinito, que representan las imágenes de los polos sur y norte. También se cambian de lugar los entornos de estos puntos. Lo expuesto concuerda con las conocidas igualdades, referentes al número complejo impropio ∞ :

$$\infty = \frac{1}{0}, \quad 0 = \frac{1}{\infty}.$$

Estas expresan la correspondencia existente entre los puntos 0 y ∞ del plano ampliado, establecida mediante la transformación $\zeta = \frac{1}{z}$.

Esta última transformación se emplea constantemente en todas las cuestiones en las que figura el punto del infinito. Ella reduce el estudio de este último punto y de cualquier entorno del mismo al estudio del punto cero y de su entorno correspondiente.

Obsérvese que como la proyección estereográfica es conforme y posee la propiedad homocíclica, la transformación $\zeta = \frac{1}{z}$ también es conforme y posee también la propiedad homocíclica. Comprobemos la primera afirmación.

Si γ es una circunferencia cualquiera o una recta del plano z , la circunferencia de la esfera que le corresponde en la proyección estereográfica, después de haber realizado la rotación considerada de la esfera, se transformará también en una circunferencia. Pero a esta última en la proyección estereográfica le corresponde en el plano una circunferencia o una recta γ' . Esta es precisamente la imagen de la línea γ en la transformación $\zeta = \frac{1}{z}$. Por cierto, la propiedad homocíclica de la última transformación se expresa en que las imágenes de las rectas y circunferencias en esta transformación son también rectas o circunferencias. De un modo semejante se comprueba también que la transformación $\zeta = \frac{1}{z}$ es conforme. Señalemos que ambas propiedades indicadas se establecerán para unas transformaciones más generales en los apartados 2.3—2.4 y 4.3 del segundo capítulo, sin aplicar la proyección estereográfica.

5.6. Consideremos de nuevo la definición de ángulo con el vértice en el punto del infinito. En lugar de la definición formulada al final del apartado 5.4, se puede dar otra equivalente.

Sean Γ_1 y Γ_2 dos curvas continuas del plano ampliado que pasen por el punto del infinito y que formen en el mismo un ángulo α . Según el apartado 5.4, esto significa que Γ_1 y Γ_2 son las proyecciones estereográficas de unas curvas continuas γ_1 y γ_2 de la esfera, que pasan por el polo norte y forman en el mismo el ángulo α .

Haciendo girar la esfera en el ángulo π alrededor del eje x , las curvas γ_1 y γ_2 se transforman en unas curvas nuevas γ'_1 y γ'_2 que pasan por el polo sur de la esfera y que forman en él también el ángulo α . Sus proyecciones estereográficas sobre el plano serán unas curvas continuas Γ'_1 y Γ'_2 que forman en el origen de coordenadas (la proyección del polo sur) el ángulo α . Pero, por otra parte, según lo expuesto anteriormente, las curvas Γ'_1 y Γ'_2 pueden obtenerse de las curvas Γ_1 y Γ_2 mediante la transformación $\zeta = \frac{1}{z}$. De aquí se deduce que, si dos curvas Γ_1 y Γ_2 del plano ampliado forman un ángulo α en el punto del infinito, entonces, sus imágenes Γ'_1 y Γ'_2 , obtenidas mediante la transformación $\zeta = \frac{1}{z}$, forman entre sí en el origen de coordenadas el mismo ángulo α .

Siguiendo nuestros razonamientos en orden inverso nos convencemos de que si se sabe que Γ'_1 y Γ'_2 forman en el origen de coordenadas un ángulo α , de aquí se deduce que sus imágenes Γ_1 y Γ_2 en la transformación $\zeta = \frac{1}{z}$ forman también en el punto del infinito el ángulo α . Por lo tanto, se puede sustituir la definición inicial de ángulo con el vértice en el punto del infinito por la siguiente definición equivalente: se dice que las curvas continuas Γ_1 y Γ_2 del plano ampliado, que pasan por el punto del infinito, forman en este punto un ángulo α cuando, y sólo cuando, las imágenes de estas curvas, obtenidas como resultado de la transformación $\zeta = \frac{1}{z}$, forman en el origen de coordenadas el ángulo α .

CAPITULO
SEGUNDO

LA DERIVABILIDAD Y SU SIGNIFICADO
GEOMETRICO.
LAS FUNCIONES ELEMENTALES

§ 1. LA DERIVADA. CONDICIONES
DE D'ALEMBERT — EULER

1.1. Sea $f(z)$ una función de variable compleja, definida y uniforme en cierto conjunto E , y sea z_0 algún punto de acumulación de este conjunto, perteneciente al mismo. Formemos la razón de diferencias

$$\frac{f(z) - f(z_0)}{z - z_0}.$$

Evidentemente, ésta representa una función de z , definida para todos los puntos del conjunto E , distintos de z_0 . Si existe el límite

$$\lim_{z \rightarrow z_0, z \in E} \frac{f(z) - f(z_0)}{z - z_0},$$

éste se llama derivada de la función $f(z)$ sobre el conjunto E en el punto z_0 y se designa mediante $f'_E(z_0)$ o, más abreviadamente, $f'(z_0)$. La función $f(z)$ que posee derivada se llama derivable (o diferenciable) o monógena sobre el conjunto E en el punto z_0 .

En el caso particular, cuando E es un intervalo del eje real (finito o infinito), $f(z)$ es una función de la variable real $z = x$, que, en el caso general, toma valores complejos: $f(z) = f(x) = \varphi(x) + i\psi(x)$. Si $\psi(x) \equiv 0$, es decir, si los valores de $f(z)$ también son reales, las definiciones dadas de derivada y derivabilidad coinciden, evidentemente, con las definiciones ordinarias del cálculo diferencial. Si $\psi(x) \not\equiv 0$, escribiendo $\frac{f(x) - f(x_0)}{x - x_0}$ en la forma $\frac{\varphi(x) - \varphi(x_0)}{x - x_0} + i \frac{\psi(x) - \psi(x_0)}{x - x_0}$, en virtud del ap. 3.6 del primer capítulo, sacamos la conclusión de que la derivada $f'(x)$ existe cuando, y sólo cuando, existen las derivadas $\varphi'(x)$ y $\psi'(x)$; en este caso, $f'(x) = \varphi'(x) +$

$\pm i\psi'(x)$. Así, por ejemplo, si $f(x) = a \cos x + ib \sin x$, se tiene, $f'(x) = -a \sin x + ib \cos x$.

Designando $f(z) - f(z_0)$ mediante $\Delta f(z)$ (el incremento de la función) y $z - z_0$ mediante Δz (el incremento de la variable independiente), escribimos la condición de derivabilidad así:

$$\frac{\Delta f(z)}{\Delta z} = f'(z_0) + \varepsilon(z_0, \Delta z),$$

donde $\varepsilon(z_0, \Delta z) \rightarrow 0$ para $\Delta z \rightarrow 0$ ($z \in E$). De aquí se deduce que el incremento de una función derivable puede representarse en la forma:

$$\Delta f(z) = A \cdot \Delta z + \varepsilon(z_0, \Delta z) \Delta z \quad (A = f'(z_0)) \quad (1.1:1)$$

donde A no depende de Δz y ε tiende a cero junto con Δz .

Recíprocamente: toda función, para la cual el incremento puede representarse en la forma (1.1:1) con las mismas condiciones respecto de A y $\varepsilon(z, \Delta z)$, es una función derivable y su derivada es igual a A . En efecto, de (1.1:1) se deduce que el límite

$$\lim_{\Delta z \rightarrow 0} \frac{\Delta f(z)}{\Delta z} \quad (\Delta z = z - z_0; \quad z \in E)$$

existe y es igual a A . Por lo tanto, la representación del incremento de la función en la forma (1.1:1), donde A no depende de Δz y ε tiende a 0 junto con Δz , es la condición necesaria y suficiente para la derivabilidad de la función. Obsérvese que de (1.1:1) se deduce inmediatamente que una función derivable en un punto $z_0 \in E$ es continua en este punto (sobre este conjunto).

Designando Δz mediante dz (diferencial de la variable independiente) y $A \cdot \Delta z = f'_E(z_0) dz$ mediante $df(z) = d_E f(z)$ (diferencial de la función), obtenemos para la derivada la siguiente expresión mediante diferenciales:

$$f'_E(z_0) = \frac{d_E f(z)}{dz}.$$

Aclaremos con un ejemplo el papel que desempeña en la definición del concepto de derivabilidad el conjunto E sobre el cual se toma la derivada. Supongamos primero que E es el eje real y $f(z) = f(x) = x$. Entonces, la derivada $f'_E(x)$ existe para cualquier $x \in E$ y es igual a uno, es decir, la función es derivable en todo el conjunto E . Prolonguemos ahora la función $f(x)$ a todo el plano complejo E_1 , poniendo como antes $f(z) = x$. Evidentemente, esta función es continua para cualquier z y coincide con la función inicial cuando $z \in E$ (es decir, cuando $y = 0$). La razón de diferencias es aquí igual a:

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{(x - x_0)}{(x - x_0) + i(y - y_0)}.$$

Esta no tiene límite para $z \rightarrow z_0$ (z_0 es un punto arbitrario del plano), puesto que es igual a 0 cuando $x = x_0$ y $y \neq y_0$, y es igual a 1 cuando $x \neq x_0$ y $y = y_0$. Resumiendo, la función $f(z) = x$ no es derivable sobre el plano en ningún punto.

1.2. De la definición de derivada y de las propiedades de los límites de las funciones de variable compleja se deduce que las reglas fundamentales, conocidas en el cálculo diferencial, son válidas también para las derivadas de las funciones de variables complejas sobre un conjunto.

He aquí estas reglas:

$$1. \text{ Si } f(z) \equiv c, \text{ entonces } \frac{df(z)}{dz} = 0.$$

$$2. \frac{d[cf(z)]}{dz} = c \frac{df(z)}{dz}.$$

$$3. \frac{dz}{dz} = 1.$$

$$4. \frac{d}{dz} [f_1(z) + f_2(z) + \dots + f_n(z)] = \frac{df_1(z)}{dz} + \frac{df_2(z)}{dz} + \dots + \frac{df_n(z)}{dz}.$$

$$5. \frac{d}{dz} [f_1(z)f_2(z)\dots f_n(z)] = f_2(z)f_3(z)\dots f_n(z) \frac{df_1(z)}{dz} + f_1(z)f_3(z)\dots f_n(z) \frac{df_2(z)}{dz} + \dots + f_1(z)f_2(z)\dots f_{n-1}(z) \frac{df_n(z)}{dz}.$$

$$6. \frac{d}{dz} [f(z)]^n = n [f(z)]^{n-1} f'(z). \quad 6'. \frac{d}{dz} (z^n) = nz^{n-1}.$$

$$7. \frac{d}{dz} (a_0 + a_1z + \dots + a_nz^n) = a_1 + 2a_2z + \dots + na_nz^{n-1}.$$

$$8. \frac{d}{dz} \left[\frac{f_1(z)}{f_2(z)} \right] = \frac{f_2(z) \frac{df_1(z)}{dz} - f_1(z) \frac{df_2(z)}{dz}}{[f_2(z)]^2}.$$

Aquí se supone que todas las funciones $f(z)$, $f_1(z)$, $f_2(z)$, ... son derivables en el punto dado z del conjunto E . En la regla 8, se exige, además, que $f_2(z)$ sea distinta de cero.

9. Regla de derivación de las funciones compuestas. Supongamos que la función $w = f(z)$ es derivable en un punto $z_0 \in E$ y que $w_0 = f(z_0)$ es un punto de acumulación para el conjunto F de los valores que toma esta función en E . Consideremos una función $Z = \varphi(w)$, definida en F y derivable en el punto w_0 sobre este conjunto. Entonces, la función compuesta $Z = \varphi[f(z)]$ es derivable en el punto z_0 sobre el conjunto E , y además,

$$\frac{d_Z \varphi[f(z)]}{dz} = \frac{d_Z \varphi(w)}{dw} \frac{d_E f(z)}{dz}. \quad (1.2:1)$$

En efecto, supongamos primero que en cualquier entorno del punto z_0 existen puntos $z \in E$, $z \neq z_0$, tales que $f(z) = f(z_0) = w_0$. Entonces existe una sucesión de puntos $\{z_n\}$ del conjunto E , convergente hacia z_0 y tal que $f(z_n) = f(z_0)$ ($n = 1, 2, \dots$). Para esta sucesión $\lim_{z_n \rightarrow z_0} \frac{f(z_n) - f(z_0)}{z_n - z_0} = 0$ y como, por la hipótesis, existe la

derivada $\frac{d_E f(z_0)}{dz} = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$, ésta tiene que ser igual a cero.

Por lo tanto, el segundo miembro de la relación (1.2:1) es igual a cero, y es suficiente demostrar que también es igual a cero el primer miembro.

Examinemos la relación de diferencias

$$\frac{\varphi[f(z)] - \varphi[f(z_0)]}{z - z_0};$$

ésta se anula para cada $z \neq z_0$ tal que $f(z) = f(z_0)$. Por consiguiente, es suficiente establecer que esta relación posee un límite igual a cero para cualquier sucesión $\{z'_n\} \rightarrow z_0$ tal que $f(z'_n) = w'_n \neq w_0$. Pero entonces,

$$\begin{aligned} \frac{\varphi[f(z'_n)] - \varphi[f(z_0)]}{z'_n - z_0} &= \frac{\varphi(w'_n) - \varphi(w_0)}{w'_n - w_0} \frac{w'_n - w_0}{z'_n - z_0} = \\ &= \frac{\varphi(w'_n) - \varphi(w_0)}{w'_n - w_0} \frac{f(z'_n) - f(z_0)}{z'_n - z_0} \rightarrow \frac{d_F \varphi(w)}{dw} \frac{d_E f(z)}{dz} = 0 \text{ para } z \rightarrow z_0. \end{aligned}$$

Así, pues, bajo la hipótesis hecha, se cumple la relación (1.2:1).

Supongamos ahora que existe un entorno del punto z_0 , en el cual $f(z) = w \neq w_0$ para todos los puntos z pertenecientes a E . Entonces, tendremos:

$$\begin{aligned} \frac{\varphi[f(z)] - \varphi[f(z_0)]}{z - z_0} &= \frac{\varphi(w) - \varphi(w_0)}{w - w_0} \frac{w - w_0}{z - z_0} = \\ &= \frac{\varphi(w) - \varphi(w_0)}{w - w_0} \frac{f(z) - f(z_0)}{z - z_0} \rightarrow \frac{d_F \varphi(w)}{dw} \frac{d_E f(z)}{dz} \text{ para } z \rightarrow z_0, \end{aligned}$$

es decir, de nuevo obtenemos la relación (1.2:1).

10. Regla de derivación de las funciones inversas. Supongamos que la función $w = f(z)$ establece una correspondencia biunívoca entre los puntos de dos conjuntos E y F , y que la función inversa $z = \varphi(w)$ es continua en F . Entonces, si $f(z)$ es derivable en el punto $z_0 \in E$ y $f'_E(z_0) \neq 0$, la función inversa $z = \varphi(w)$ también será derivable en el punto $w_0 = f(z_0) \in F$ y

$$\varphi'_F(w_0) = \frac{1}{f'_E(z_0)}.$$

En efecto, como la aplicación $w = f(z)$ es biyectiva, se tiene, $z \neq z_0$ para $w \neq w_0$, por lo cual la relación de diferencias para la

función $\varphi(w)$ se puede representar en la forma

$$\frac{\varphi(w) - \varphi(w_0)}{w - w_0} = \frac{z - z_0}{w - w_0} = \frac{1}{\frac{w - w_0}{z - z_0}},$$

y como para $w \rightarrow w_0$ también $z = \varphi(w) \rightarrow z_0 = \varphi(w_0)$, se tiene:

$$\lim_{w \rightarrow w_0} \frac{\varphi(w) - \varphi(w_0)}{w - w_0} = \frac{1}{\lim_{z \rightarrow z_0} \frac{w - w_0}{z - z_0}} = \frac{1}{\lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}} = \frac{1}{f'(z_0)},$$

que es lo que se quería demostrar.

1.3. Aquí, fundamentalmente, estudiaremos las funciones definidas en un recinto $E = G$ y, en este caso, en lugar de $f'_E(z)$ o de $\frac{d_E f(z)}{dz}$ escribiremos abreviadamente: $f'(z)$ o $\frac{df(z)}{dz}$.

Sea $f(z) = u(x, y) + iv(x, y)$; recordemos que una función de dos variables reales $u(x, y)$ se llama diferenciable en el punto (x_0, y_0) del recinto de su definición, si se cumple la relación:

$$u(x, y) - u(x_0, y_0) = A(x_0, y_0)(x - x_0) + B(x_0, y_0)(y - y_0) + \varepsilon_1(x, y; x_0, y_0)(x - x_0) + \varepsilon_2(x, y; x_0, y_0)(y - y_0),$$

donde

$$\lim_{x \rightarrow x_0, y \rightarrow y_0} \varepsilon_1(x, y; x_0, y_0) = \lim_{x \rightarrow x_0, y \rightarrow y_0} \varepsilon_2(x, y; x_0, y_0) = 0.$$

Los coeficientes $A(x_0, y_0)$ y $B(x_0, y_0)$ del segundo miembro de la igualdad, son las derivadas parciales de la función $u(x, y)$:

$$A(x_0, y_0) = \left. \frac{\partial u(x, y)}{\partial x} \right|_{\substack{x=x_0 \\ y=y_0}}, \quad B(x_0, y_0) = \left. \frac{\partial u(x, y)}{\partial y} \right|_{\substack{x=x_0 \\ y=y_0}}.$$

Demostremos la siguiente proposición importante:

T e o r e m a. *Para que una función $f(z) = u(x, y) + iv(x, y)$, definida en un recinto G , sea derivable en un punto z de este recinto como función de variable compleja, es necesario y suficiente que las funciones $u(x, y)$ y $v(x, y)$ sean diferenciables en este mismo punto (como funciones de dos variables reales) y que, además, se cumplan en este punto las condiciones:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \quad (1.3:1)$$

Si se cumplen todas las condiciones del teorema, la derivada $f'(z)$ puede expresarse en una de las siguientes formas:

$$f'(z) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} = \frac{\partial u}{\partial x} - i \frac{\partial u}{\partial y} = \frac{\partial v}{\partial y} + i \frac{\partial v}{\partial x}. \quad (1.3:2)$$

Las condiciones (1.3:1) son de fundamental importancia en la teoría de las funciones analíticas y en las aplicaciones de esta teoría a los problemas de la mecánica y física. Estas se llaman **condiciones (o ecuaciones) de Cauchy — Riemann**.

Es necesario señalar que esta denominación, universalmente admitida en los libros de enseñanza y científicos, está injustificada desde el punto de vista histórico, puesto que las condiciones (1.3:1) se estudiaron ya en el siglo XVIII por D'Alembert y, fundamentalmente, por Euler, en los trabajos dedicados a la aplicación de las funciones de variable compleja a la hidromecánica (D'Alembert y Euler), cartografía y cálculo integral (Euler). Por lo tanto, sería justo cambiar la terminología corriente y llamar a las ecuaciones (1.3:1) **ecuaciones de D'Alembert — Euler**.

Veamos la demostración del teorema y demostremos primero que las condiciones del mismo son necesarias para la derivabilidad de la función $f(z)$.

En efecto, si $f(z)$ es derivable en el punto z del recinto G , se tiene¹

$$\Delta f(z) = f'(z) \Delta z + \varepsilon \Delta z, \quad (1.3:3)$$

donde

$$\begin{aligned} \Delta z &= z_1 - z = (x_1 - x) + i(y_1 - y) = \Delta x + i\Delta y, \\ \Delta f(z) &= f(z_1) - f(z) = \\ &= [u(x_1, y_1) - u(x, y)] + i[v(x_1, y_1) - v(x, y)] = \Delta u + i\Delta v, \\ f'(z) &= a + ib, \quad \varepsilon = \varepsilon_1 + i\varepsilon_2, \end{aligned}$$

y ε_1 y ε_2 tienden a cero cuando Δx y Δy tienden simultáneamente a cero.

Separando en la relación (1.3:3) las partes real e imaginaria, tendremos:

$$\begin{aligned} \Delta u &= a\Delta x - b\Delta y + \varepsilon_1\Delta x - \varepsilon_2\Delta y, \\ \Delta v &= b\Delta x + a\Delta y + \varepsilon_2\Delta x - \varepsilon_1\Delta y. \end{aligned}$$

Como $\lim_{\Delta x, \Delta y \rightarrow 0} \varepsilon_1 = \lim_{\Delta x, \Delta y \rightarrow 0} \varepsilon_2 = 0$, de aquí resulta que:

- 1) las funciones $u(x, y)$ y $v(x, y)$ de dos variables reales x e y , son diferenciables en el punto (x, y) ;
- 2) sus derivadas parciales en este punto son:

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

y, por consiguiente, satisfacen a las condiciones:

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

Finalmente, para $f'(z)$ obtenemos:

$$f'(z) = a + ib = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} = \frac{\partial u}{\partial x} - i \frac{\partial u}{\partial y} = \frac{\partial v}{\partial y} + i \frac{\partial v}{\partial x}.$$

Así, pues, queda demostrado que las condiciones del teorema son necesarias.

Demostremos que las condiciones del teorema son suficientes. Supongamos éstas cumplidas. Entonces

$$\left. \begin{aligned} \Delta u &= \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y + \alpha_1 \Delta x + \alpha_2 \Delta y, \\ \Delta v &= \frac{\partial v}{\partial x} \Delta x + \frac{\partial v}{\partial y} \Delta y + \beta_1 \Delta x + \beta_2 \Delta y, \end{aligned} \right\} \quad (1.3:4)$$

donde α_1 , α_2 , β_1 y β_2 tienden a cero cuando Δx y Δy tienden a cero. Además,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = a, \quad -\frac{\partial u}{\partial y} = \frac{\partial v}{\partial x} = b. \quad (1.3:5)$$

Por consiguiente,

$$\begin{aligned} \Delta u &= a \Delta x - b \Delta y + \alpha_1 \Delta x + \alpha_2 \Delta y, \\ \Delta v &= b \Delta x + a \Delta y + \beta_1 \Delta x + \beta_2 \Delta y \end{aligned}$$

y

$$\begin{aligned} \Delta f(z) &= \Delta u + i \Delta v = \\ &= a(\Delta x + i \Delta y) + ib(\Delta x + i \Delta y) + (\alpha_1 + i \beta_1) \Delta x + (\alpha_2 + i \beta_2) \Delta y = \\ &= (a + ib) \Delta z + \left[(\alpha_1 + i \beta_1) \frac{\Delta x}{\Delta z} + (\alpha_2 + i \beta_2) \frac{\Delta y}{\Delta z} \right] \Delta z = A \Delta z + \varepsilon \Delta z. \end{aligned} \quad (1.3:6)$$

Como

$$\begin{aligned} |\varepsilon| &= \left| (\alpha_1 + i \beta_1) \frac{\Delta x}{\Delta z} + (\alpha_2 + i \beta_2) \frac{\Delta y}{\Delta z} \right| \leq \\ &\leq \left| \alpha_1 + i \beta_1 \right| \left| \frac{\Delta x}{\Delta z} \right| + \left| \alpha_2 + i \beta_2 \right| \left| \frac{\Delta y}{\Delta z} \right| \leq \\ &\leq |\alpha_1 + i \beta_1| + |\alpha_2 + i \beta_2| \leq |\alpha_1| + |\beta_1| + |\alpha_2| + |\beta_2|, \end{aligned}$$

tenemos que ε , junto con α_1 , β_1 , α_2 , β_2 , tiende a cero cuando $\Delta z = \Delta x + i \Delta y$ tiende a cero. De aquí y de las relaciones (1.3:6), se deduce que la función $f(z)$ es derivable y su derivada $f'(z)$ es igual a A :

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

Con esto se termina la demostración.

Por el curso general del análisis se sabe que para que las funciones $u(x, y)$ y $v(x, y)$ sean diferenciables, es suficiente que existan y sean continuas sus derivadas parciales: $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$. Por esto, para que la función $f(z) = u + iv$ sea derivable, es suficiente que existan las derivadas parciales $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$, que éstas sean continuas y que satisfagan a las ecuaciones (1.3:1).

Una función $f(z)$, derivable en cada punto del recinto G , se llama derivable (diferenciable) en este recinto, o también holomorfa, o analítica (a veces, regular). La denominación holomorfa (semejante a entera, de las palabras griegas $\delta\lambda\omicron\zeta$ total, entero, y $\mu\omicron\rho\phi\acute{\eta}$ — forma) fue introducida por los alumnos de Cauchy: Briot y Bouquet. «Con esta denominación señalamos —decían ellos— que ella (es decir, la función holomorfa — autor) es semejante a las funciones enteras (es decir, a los polinomios. — autor), que poseen las mismas propiedades en todo el plano». El significado del vocablo «analítica», empleado anteriormente por Lagrange y más tarde por Weierstrass, y universalmente admitido en la actualidad, se explicó en el § 1 del primer capítulo; su aplicación a las funciones de variable compleja, que son derivables en cierto recinto, será justificada en la exposición ulterior, cuando demostremos que una función tal puede expresarse en un entorno de cualquier punto del recinto en forma de suma de una serie de potencias convergente. Por ahora emplearemos la denominación de «función analítica» como sinónimo de la denominación «función de variable compleja, derivable (o diferenciable) en un recinto dado».

Como ejemplo, examinemos la función $f(z) = e^x (\cos y + i \sin y)$, definida en todo el plano. En este caso,

$$u = e^x \cos y, \quad v = e^x \sin y,$$

$$\frac{\partial u}{\partial x} = e^x \cos y = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -e^x \sin y = -\frac{\partial v}{\partial x}.$$

Por lo tanto, se cumplen las condiciones (1.3:1) y la función $f(z)$ es analítica en todo el plano. Para su derivada, se tiene:

$$f'(z) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = e^x \cos y + i e^x \sin y = f(z).$$

En el ejemplo examinado al final del ap. 1.1, $f(z) = x$, $u = x$, $v = 0$, $\frac{\partial u}{\partial x} = 1$, $\frac{\partial u}{\partial y} = \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} = 0$ y no se cumplen las condiciones de D'Alembert — Euler: $\frac{\partial u}{\partial x} \neq \frac{\partial v}{\partial y}$. Ya se vio que esta función no es derivable en ningún punto (sobre el plano).

Las condiciones (1.3:1) pueden expresarse en una forma más compacta aplicando el concepto de las llamadas derivadas formales, que resulta útil en diversas cuestiones de la teoría de las funciones de variable compleja. Examinemos en un recinto una función compleja $f(x, y) = u(x, y) + iv(x, y)$, diferenciable en el mismo respecto de x e y . Efectuando un cambio de variables:

$$x = \alpha\xi + \beta\eta, \quad y = \gamma\xi + \delta\eta,$$

donde $\alpha\delta - \beta\gamma \neq 0$,

se obtiene una función diferenciable de ξ y η , para la cual

$$\frac{\partial f}{\partial \xi} = \alpha \frac{\partial f}{\partial x} + \gamma \frac{\partial f}{\partial y}, \quad \frac{\partial f}{\partial \eta} = \beta \frac{\partial f}{\partial x} + \delta \frac{\partial f}{\partial y}.$$

Desde luego en estas fórmulas, α , β , γ y δ son números reales. Sin embargo podemos asignar a éstos valores imaginarios, conviniendo en que los segundos miembros de las fórmulas servirán entonces de definición de los primeros. Precisamente haremos:

$$\alpha = \beta = \frac{1}{2}, \quad \gamma = -\frac{i}{2}, \quad \delta = \frac{i}{2};$$

en este caso,

$$\xi = x + iy = z, \quad \eta = x - iy = \bar{z}$$

y nuestras fórmulas proporcionan la definición de las derivadas formales:

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

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

Del teorema demostrado en el presente apartado se deduce que, una función $f = u + iv$, diferenciable en el recinto G con respecto de x e y , será analítica en este recinto cuando, y sólo cuando, en todos los puntos del recinto se cumple la condición

$$\frac{\partial f}{\partial \bar{z}} = 0;$$

en este caso,

$$\frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = f'(z).$$

En muchos casos es importante que las condiciones para que una función de variable compleja $f(z) = u + iv$ sea diferenciable en un punto $z \neq 0$ vengan expresadas en coordenadas polares:

$|z| = r$ y $\text{Arg } z = \Phi$. Estas condiciones (necesaria y suficientes) son:

- 1') u y v son funciones diferenciables de r y Φ ;
- 2') sus derivadas parciales están ligadas por las relaciones

$$\frac{\partial u}{\partial r} = \frac{1}{r} \frac{\partial v}{\partial \Phi}, \quad \frac{\partial v}{\partial r} = -\frac{1}{r} \frac{\partial u}{\partial \Phi}. \quad (1.3.7)$$

Para convencerse de esto, es suficiente demostrar que u y v son diferenciables como funciones de r y Φ ($r \neq 0$) si, y sólo si, son diferenciables como funciones de x e y , y que en estas condiciones las ecuaciones (1.3.7) son equivalentes a las ecuaciones (1.3.1). Pero el cumplimiento de lo primero que se pide es consecuencia del hecho, conocido en el curso general de análisis, de que una función diferenciable (por ejemplo, $u = u(x, y)$) de funciones diferenciables (por ejemplo, de $x = r \cos \Phi$ e $y = r \sin \Phi$) también es diferenciable (respecto de las variables r y Φ). La segunda afirmación se comprueba inmediatamente. Por ejemplo, si se cumple la condición 1') y, además, se cumplen las condiciones (1.3.1), se tiene:

$$\left. \begin{aligned} \frac{\partial u}{\partial r} &= \frac{\partial u}{\partial x} \cos \Phi + \frac{\partial u}{\partial y} \sin \Phi = \frac{\partial v}{\partial y} \cos \Phi - \frac{\partial v}{\partial x} \sin \Phi = \frac{1}{r} \frac{\partial v}{\partial \Phi}, \\ \frac{\partial v}{\partial r} &= \frac{\partial v}{\partial x} \cos \Phi + \frac{\partial v}{\partial y} \sin \Phi = -\frac{\partial u}{\partial y} \cos \Phi + \frac{\partial u}{\partial x} \sin \Phi = -\frac{1}{r} \frac{\partial u}{\partial \Phi}. \end{aligned} \right\} \quad (1.3.8)$$

El lector fácilmente realizará también el paso inverso de las condiciones (1.3.7) a las condiciones (1.3.1).

Escribiendo las ecuaciones (1.3.8) de la forma

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x} \cos \Phi - \frac{\partial v}{\partial x} \sin \Phi, \quad \frac{\partial v}{\partial r} = \frac{\partial u}{\partial x} \sin \Phi + \frac{\partial v}{\partial x} \cos \Phi,$$

de ellas obtenemos

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r} \cos \Phi + \frac{\partial v}{\partial r} \sin \Phi, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial r} \sin \Phi + \frac{\partial v}{\partial r} \cos \Phi,$$

y, por consiguiente,

$$\begin{aligned} f'(z) &= \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial u}{\partial r} (\cos \Phi - i \sin \Phi) + i \frac{\partial v}{\partial r} (\cos \Phi - i \sin \Phi) = \\ &= \left(\frac{\partial u}{\partial r} + i \frac{\partial v}{\partial r} \right) (\cos \Phi - i \sin \Phi) = \frac{r}{z} \left(\frac{\partial u}{\partial r} + i \frac{\partial v}{\partial r} \right). \end{aligned} \quad (1.3.9)$$

Esta fórmula es útil para calcular $f'(z)$ mediante las coordenadas polares. Las ecuaciones (1.3.7) permiten también expresar $f'(z)$ en la forma:

$$f'(z) = \frac{1}{z} \left(\frac{\partial v}{\partial \Phi} - i \frac{\partial u}{\partial \Phi} \right). \quad (1.3.10)$$

Como ejemplo, examinemos la función potencial:

$$z^{\frac{m}{n}} = |z|^{\frac{m}{n}} \left(\cos \frac{m \operatorname{Arg} z}{n} + i \operatorname{sen} \frac{m \operatorname{Arg} z}{n} \right) = r^{\frac{m}{n}} \left(\cos \frac{m\Phi}{n} + i \operatorname{sen} \frac{m\Phi}{n} \right),$$

donde m es un número entero y n es natural. Esta función está definida en el recinto $G: z \neq 0$, y es multiforme si el número racional $\frac{m}{n}$ no es entero (véase el cap. 1, ap. 2.3). Esto último es debido a que el argumento Φ también es multiforme. Para tener la posibilidad

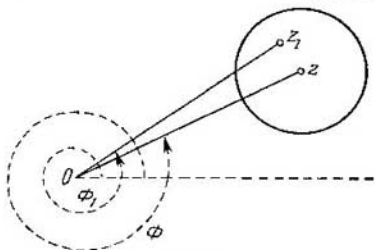


FIG. 8

de hablar de la derivada de esta función multiforme en cierto punto z del recinto G , tomemos en este recinto algún entorno del punto z que no contenga al origen de coordenadas y, fijando uno de los valores que toma Φ en el punto z , tomemos en todos los demás puntos z_1 del mismo entorno los valores Φ_1 que cumplen la condición $|\Phi_1 - \Phi| < \frac{\pi}{2}$ (fig. 8). Entonces, obtendremos en el entorno considerado una rama uniforme y continua*) de la función $z^{\frac{m}{n}}$. Esta función uniforme se designará con la misma notación: $f(z) = z^{\frac{m}{n}}$.

Evidentemente, en este caso

$$u = r^{\frac{m}{n}} \cos \frac{m\Phi}{n}, \quad v = r^{\frac{m}{n}} \operatorname{sen} \frac{m\Phi}{n},$$

$$\frac{\partial u}{\partial r} = \frac{m}{n} r^{\frac{m}{n}-1} \cos \frac{m\Phi}{n} = \frac{1}{r} \frac{\partial u}{\partial \Phi},$$

$$\frac{\partial v}{\partial r} = \frac{m}{n} r^{\frac{m}{n}-1} \operatorname{sen} \frac{m\Phi}{n} = -\frac{1}{r} \frac{\partial v}{\partial \Phi},$$

*) Sobre las ramas uniformes de las funciones multiformes hablaremos más detalladamente en adelante, en el § 5 del presente capítulo.

y, por consiguiente, $f(z)$ es una función diferenciable. En virtud de la fórmula (1.3.9), para su derivada, obtenemos:

$$\begin{aligned} f'(z) &= \frac{r}{z} \left(\frac{m}{n} r^{\frac{m}{n}-1} \cos \frac{m\Phi}{n} + \frac{m}{n} r^{\frac{m}{n}-1} i \operatorname{sen} \frac{m\Phi}{n} \right) = \\ &= \frac{m}{n} r^{\frac{m}{n}-1} \left(\cos \frac{m\Phi}{n} + i \operatorname{sen} \frac{m\Phi}{n} \right) \frac{1}{z} = \frac{m}{n} \frac{f(z)}{z}. \end{aligned}$$

Por lo tanto, el lector puede observar que la regla de derivación de la «potencia fraccionaria» $z^{\frac{m}{n}}$ se conserva formalmente igual que para la función correspondiente de variable real $x^{\frac{m}{n}}$. Solamente hay que tener en cuenta que nuestro cálculo se efectuaba con la condición de que $z \neq 0$, condición que puede omitirse solamente cuando $\frac{m}{n}$ sea un número entero no negativo.

Como ejercicio, proponemos al lector convencerse de que la función $f(z) = \ln r + i\Phi$, definida en el mismo recinto G , es diferenciable y su derivada es igual a $\frac{1}{z}$ (aquí también hay que separar las ramas uniformes y continuas de la función).

§ 2. SIGNIFICADO GEOMETRICO DE LA DERIVADA. TRANSFORMACION CONFORME

2.1. Consideremos primero una función compleja $z = \lambda(t)$ de variable real t , definida y continua en un segmento $E: [\alpha, \beta]$ del eje real. Como se señaló en el ap. 4.1 del capítulo primero, tal función determina una curva continua L . Supongamos que en cierto punto del segmento $[\alpha, \beta]$ existe la derivada (sobre el conjunto E) $\lambda'(t) \neq 0$. Demostremos que entonces existe en el punto correspondiente $z_0 = \lambda(t_0)$ de la curva L la tangente T de ésta (entendida como la posición límite de la secante que pasa por el punto z_0) y que el ángulo entre T y el eje real coincide con $\operatorname{Arg} \lambda'(t_0)$.

En efecto, tracemos una secante por los puntos $z_0 = \lambda(t_0)$ y $z_1 = \lambda(t_1)$ de la curva L . Se puede suponer que estos puntos no coinciden para todos los valores t_1 , distintos de t_0 y suficientemente próximos a t_0 (en caso contrario, existe una sucesión $\{t_{1n}\} \rightarrow t_0$, tal que

$$\lambda(t_{1n}) - \lambda(t_0) = 0$$

para todos n , y por consiguiente,

$$\lambda'(t_0) = \lim_{t_{1n} \rightarrow t_0} \frac{\lambda(t_{1n}) - \lambda(t_0)}{t_{1n} - t_0} = 0).$$

Observando que la dirección de la secante coincide con la dirección del vector $\frac{z_1 - z_0}{t_1 - t_0}$, sacamos la conclusión de que la secante tendrá una posición límite para $t_1 \rightarrow t_0$ ($z_1 \rightarrow z_0$), solamente si el ángulo entre el último vector y el eje real, igual a $\text{Arg} \frac{z_1 - z_0}{t_1 - t_0}$, tiene límite cuando $t_1 \rightarrow t_0$. Pero, según la condición, existe el límite

$$\lim_{t_1 \rightarrow t_0} \frac{z_1 - z_0}{t_1 - t_0} = \lambda'(t_0) \neq 0;$$

por lo cual también existe el límite

$$\lim_{t_1 \rightarrow t_0} \text{Arg} \frac{z_1 - z_0}{t_1 - t_0} = \text{Arg} \lambda'(t_0) *),$$

con lo cual se termina la demostración.

Resumiendo, *para una función compleja de variable real, la existencia de derivada distinta de cero significa la existencia de tangente a la curva correspondiente; el ángulo de inclinación de la tangente al eje real coincide con el argumento de la derivada.*

Examinemos ahora una función de variable compleja $w = f(z)$, definida y continua en un recinto G , y supongamos que en un punto $z_0 \in G$ existe la derivada $f'(z_0) \neq 0$. Trazemos por el punto z_0 alguna curva $L: z = \lambda(t)$, ($\alpha \leq t \leq \beta$, $\lambda(\alpha) = z_0$), para la cual exista la derivada $\lambda'(t_0) \neq 0$; según lo expuesto anteriormente, la curva L posee tangente en el punto $z_0 = \lambda(\alpha)$ con el ángulo de inclinación igual a $\text{Arg} \lambda'(t_0)$. Mediante la transformación $w = f(z)$ esta curva se transforma en una curva Λ , situada en el plano $w: w = f[\lambda(t)] = \mu(t)$ ($\alpha \leq t \leq \beta$, $\mu(t_0) = f(z_0) = w_0$). Según la regla de derivación de las funciones compuestas (ap. 1.1), la función $\mu(t)$ es derivable en el punto $t = t_0$ y $\mu'(t_0) = f'(z_0) \lambda'(t_0) \neq 0$, por lo cual la curva Λ posee tangente en el punto $w_0 = f(z_0)$, y el ángulo entre la tangente y el eje real es igual a

$$\text{Arg} \mu'(t_0) = \text{Arg} [\lambda'(t_0) f'(z_0)] = \text{Arg} \lambda'(t_0) + \text{Arg} f'(z_0).$$

De aquí se deduce que, al pasar de la curva L a su imagen Λ , el ángulo de inclinación de la tangente en el punto inicial de la curva varía en la magnitud

$$\text{Arg} \mu'(t_0) - \text{Arg} \lambda'(t_0) = \text{Arg} f'(z_0),$$

que no depende de la curva. Si del punto z_0 parten dos curvas cualesquiera L_1 y L_2 , que posean tangentes t_1 y t_2 en el punto z_0 , entonces las tangentes τ_1 y τ_2 a sus imágenes Λ_1 y Λ_2 en el punto $w_0 = f(z_0)$ se obtendrán de t_1 y t_2 mediante un giro en un mismo ángulo $\text{Arg} f'(z_0)$,

*) Véase el final del ap. 3.3, cap. 1.

y, por consiguiente, el ángulo entre las curvas L_1 y L_2 será igual (en valor y en el sentido de la dirección) al ángulo entre Λ_1 y Λ_2 . Por lo tanto, al hacer una transformación mediante una función continua $w = f(z)$, que posea derivada $f'(z_0)$ distinta de cero, todas las curvas del plano z que pasan por el punto z_0 y que posean tangente en este punto se transforman en curvas del plano w que pasan por el punto $w_0 = f(z_0)$ y que también poseen tangente en este punto; en esta transformación se conservan los ángulos entre las curvas. La transformación mediante una función continua, que conserva los ángulos entre las curvas que pasan por el punto dado, se llama **conforme en este punto**.

Si, además, se conservan no solamente los valores de los ángulos, sino también los sentidos de sus direcciones, se dice que la transformación es **conforme de primera especie** *), si los sentidos de las direcciones de los ángulos se cambian por los contrarios, se dice que la transformación es **conforme de segunda especie** **).

Resumiendo, *una transformación mediante una función de variable compleja, analítica en un recinto G , es una transformación conforme de primera especie en todos los puntos en los que la derivada es distinta de cero*. Si la transformación es conforme en todos los puntos del recinto G sin excepción, entonces se llama **transformación conforme del recinto G** .

Puede servir de ejemplo de transformación conforme de segunda especie la **simetría** con respecto del eje real: $w = \bar{z}$. Sirven de ejemplos más generales las transformaciones que se realizan mediante las funciones conjugadas con las analíticas: $w = \overline{f(z)}$ (se supone que $f'(z) \neq 0$).

Proponemos demostrar al lector que si la derivada es igual a cero en cierto punto, los ángulos pueden conservarse o variar (examinense las transformaciones

$$f_1(z) = r^2 (\cos \Phi + i \operatorname{sen} \Phi)$$

$$f_2(z) = r^2 (\cos 2\Phi + i \operatorname{sen} 2\Phi) = z^2$$

en el punto $z=0$).

2.2. En el apartado anterior se demostró que $\operatorname{Arg} f'(z_0)$ representa el ángulo de rotación de la tangente a la curva L en el punto z_0 de la misma al pasar a su imagen Λ y al punto $w_0 = f(z_0)$. En particular, si $f'(z_0)$ es un número real positivo, entonces los vectores tangentes a L en z_0 y a Λ en $f(z_0)$ son paralelos y tienen una misma dirección.

*) O conforme directa. (N. del T.)

***) O conforme inversa. (N. del T.)

Veamos ahora el significado geométrico que tiene el módulo de la derivada $|f'(z_0)|$. Con este fin, obsérvese que

$$|f'(z_0)| = \lim_{z \rightarrow z_0} \frac{|f(z) - f(z_0)|}{|z - z_0|}$$

y que los números $|z - z_0|$ y $|f(z) - f(z_0)|$ expresan las distancias entre los puntos z y z_0 del plano z y entre sus imágenes $f(z)$ y $f(z_0)$ del plano w , respectivamente. Si la razón $\frac{|f(z) - f(z_0)|}{|z - z_0|}$ se puede considerar como la dilatación del vector $z - z_0$, obtenida al hacer la transformación mediante la función $w = f(z)$ (esta dilatación puede ser menor que la unidad, igual a la unidad o mayor que la unidad), *el módulo de la derivada $|f'(z_0)|$ se puede considerar como la dilatación en el punto z_0 en la transformación mediante la función $w = f(z)$* . De lo que acabamos de exponer se deduce que la magnitud de dilatación en el punto z_0 no depende del vector $z - z_0$ que se tome con el origen en este punto; no obstante, ésta no coincide con la dilatación del vector $z - z_0$ sino que representa el límite de esta dilatación cuando z tiende a z_0 .

2.3. Como ilustración, estudiemos la función homográfica *) $L(z) = \frac{az+b}{cz+d}$ (al menos uno de los números c o d es distinto de cero). Supongamos primero que $c = 0$. Entonces, $L(z)$ puede expresarse en la forma $L(z) = \alpha z + \beta$ ($\alpha = \frac{a}{d}$, $\beta = \frac{b}{d}$); ésta es una función lineal entera. Está definida para todos los valores de z y posee derivada $L'(z) = \alpha$, que conserva un valor constante y distinto de cero, si $\alpha \neq 0$. Por consiguiente, la función $L(z)$ realiza una transformación conforme de todo el plano de la variable compleja z . En esta transformación, las tangentes a todas las curvas del plano z giran un mismo ángulo, igual a $\text{Arg } \alpha$, y la dilatación en todos los puntos es igual a $|\alpha|$. Si $\alpha = 1$, se tiene, $\text{Arg } \alpha = 2k\pi$, $|\alpha| = 1$, y realmente no existe ni rotación ni dilatación. Como, en este caso, la transformación toma la forma $w = z + \beta$, ésta, evidentemente, se reduce a una traslación del plano como un todo en el vector β . Si $\alpha \neq 1$ (y $\alpha \neq 0$), la transformación puede expresarse en la forma $w - \gamma = \alpha(z - \gamma)$, donde γ se determina por la ecuación $\gamma = \alpha\gamma + \beta$. De aquí se deduce que en la transformación cada vector $z - \gamma$ que parta del punto γ , gira un ángulo igual a $\text{Arg } \alpha$ y se dilata en $|\alpha|$ veces, convirtiéndose en el vector $w - \gamma$ que parte del mismo punto γ . Esto significa que la transformación $L(z) = \alpha z + \beta$, siendo $\alpha \neq 1$ (y $\alpha \neq 0$), se reduce a una rotación

*) La denominación de esta función dada por el autor es: función fraccionaria lineal. (N. del T.)

del plano como un todo, alrededor del punto $\gamma = \frac{\beta}{1-\alpha}$ en el ángulo $\text{Arg } \alpha$, y a una dilatación de razón $|\alpha|$ con respecto de este punto. Evidentemente, ésta es una transformación homotética con centro en el punto $\gamma = \frac{\beta}{1-\alpha}$ de razón (o coeficiente de semejanza) $|\alpha|$, seguida de una rotación en el ángulo $\text{Arg } \alpha$ alrededor del mismo punto. Tal es la transformación conforme en el caso más simple.

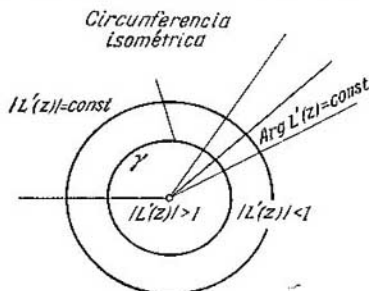


FIG. 9

Supongamos ahora que $c \neq 0$. Entonces, para $z \neq \delta = -\frac{d}{c}$ existe la derivada

$$L'(z) = \frac{ad-bc}{(cz+d)^2} = \frac{ad-bc}{c^2} \cdot \frac{1}{(z-\delta)^2}.$$

Si el determinante $ad-bc \neq 0$ (la igualdad a cero de la expresión $ad-bc$ significa el cumplimiento de la proporción $\frac{a}{c} = \frac{b}{d} = \lambda$, de donde $a = c\lambda$, $b = d\lambda$ y $L(z) = \frac{az+b}{cz+d} = \frac{c\lambda z + d\lambda}{cz+d} \equiv \lambda$), se tiene, $L'(z) \neq 0$ para todos los puntos $z \neq \delta$. Por consiguiente, la transformación $w = L(z)$ es conforme en todos los puntos finitos, distintos de δ . En esta transformación, las tangentes a las curvas que pasan por un punto arbitrario $z \neq \delta$, giran un ángulo igual a

$$\text{Arg } L'(z) = \text{Arg } \frac{ad-bc}{c^2} - 2 \text{Arg } (z-\delta).$$

Evidentemente, el ángulo de rotación de la tangente varía al pasar de un punto a otro, conservando el mismo valor en los puntos, para los cuales $\text{Arg } (z-\delta)$ conserva un mismo valor, es decir, en los puntos

de cada uno de los rayos rectilíneos que parten del punto δ . La dilatación de la longitud en el punto z en esta transformación es igual a $|L'(z)| = \left| \frac{ad-bc}{c^2} \right| : |z - \delta|^2$ y también varía al pasar de un punto a otro. Esta conserva el mismo valor en los puntos para los cuales la magnitud $|z - \delta|$ es la misma, es decir, en los puntos de cada circunferencia con el centro en el punto δ . En particular, esta dilatación es igual a la unidad en cada punto de la circunferencia $\gamma : |z - \delta| = \frac{1}{|c|} \sqrt{|ad - bc|}$ (circunferencia isométrica de la transformación homográfica); es mayor que la unidad en el interior de γ , tendiendo al infinito cuando z tiende a δ , y es menor que la unidad en el exterior de γ , tendiendo a cero cuando z tiende al infinito (fig. 9).

2.4. Supongamos, como anteriormente, que $c \neq 0$ y $ad - bc \neq 0$. Entonces, evidentemente,

$$\lim_{z \rightarrow \delta = -\frac{d}{c}} \frac{az+b}{cz+d} = \infty \quad \text{y} \quad \lim_{z \rightarrow \infty} \frac{az+b}{cz+d} = \frac{a}{c} = \alpha.$$

En correspondencia con esto, obtenemos:

$$L(\delta) = \infty \quad \text{y} \quad L(\infty) = \alpha.$$

Por lo tanto, el punto finito δ se transforma mediante la función $w = L(z)$ en el punto del infinito, y el punto del infinito, en el punto finito α . Demostremos que la transformación también es conforme en estos puntos. En efecto, sean γ_1 y γ_2 dos curvas que pasan por el punto δ y forman en el mismo un ángulo θ , y sean Γ_1 y Γ_2 sus imágenes en el plano w . Queremos demostrar que Γ_1 y Γ_2 forman también el ángulo θ en el punto del infinito. Hagamos con este fin la transformación del plano w : $\xi = \frac{1}{w}$. Entonces, las curvas Γ_1 y Γ_2 se transformarán en las curvas Γ'_1 y Γ'_2 , y el punto del infinito, en el origen de coordenadas (fig. 10). Está claro que el paso de γ_1 y γ_2 en el plano z a Γ'_1 y Γ'_2 en el plano ξ se efectúa mediante la transformación homográfica

$$\xi = \frac{1}{w} = \frac{cz+d}{az+b},$$

que es conforme en el punto $z = \delta = -\frac{d}{c}$. Por consiguiente, Γ'_1 y Γ'_2 también forman el ángulo θ en el origen de coordenadas. De aquí que, en virtud de la definición del ap. 5.6 del capítulo primero, las curvas Γ_1 y Γ_2 forman también entre sí el ángulo θ en el punto del infinito. Así, pues, queda demostrado que la transformación $w = L(z)$ es conforme en el punto $z = \delta$.

De un modo análogo se demuestra que también es conforme en el punto del infinito. Precisando, si las curvas γ_1 y γ_2 pasan por el punto del infinito en el plano z , sus imágenes Γ_1 y Γ_2 en el plano w pasan por el punto α . Supongamos que γ_1 y γ_2 forman entre sí un ángulo θ en el punto del infinito. Esto significa que sus imágenes γ'_1 y γ'_2 , obtenidas como resultado de la transformación $\xi = \frac{1}{z}$,

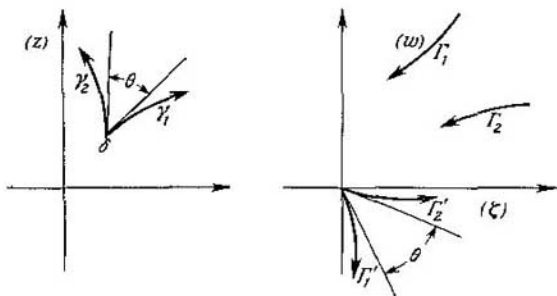


FIG. 10

forman el ángulo θ en el origen de coordenadas. Pero, evidentemente, se puede pasar de γ'_1 y γ'_2 a Γ_1 y Γ_2 mediante la transformación

$$w = \frac{az + b}{cz + d} = \frac{a \frac{1}{\xi} + b}{c \frac{1}{\xi} + d} = \frac{a + b\xi}{c + d\xi}.$$

Esta transformación es homográfica y, por lo tanto, conforme en el punto $\xi = 0$. De aquí se deduce que Γ_1 y Γ_2 forman también entre sí el ángulo θ en el punto $\alpha = \frac{a}{c}$, con lo que se termina la demostración de que la transformación $w = L(z)$ es conforme en el punto del infinito.

Cuando $c = 0$, resulta la función lineal entera

$$w = L(z) = \alpha z + \beta \quad (\alpha \neq 0).$$

Aquí se supone $L(\infty) = \infty$ y mediante la transformación auxiliar $\xi = \frac{1}{w} = \frac{1}{\alpha z + \beta}$ se reduce este caso al que se acaba de examinar (al punto del infinito del plano z corresponde el origen de coordenadas del plano w).

Proponemos al lector hacer la demostración completa.

Resumiendo, se puede decir que la *función homográfica* $w = L(z)$ realiza una *transformación conforme del plano ampliado sobre sí mismo*.

§ 3. POLINOMIOS. FUNCIÓN EXPONENCIAL. SENO Y COSENO

3.1. Las funciones uniformes y analíticas en todo el plano a excepción del punto del infinito, forman la clase más importante y simple de funciones diferenciables. Tales funciones se llaman *enteras*.

Un *polinomio*:

$$a_0 + a_1z + a_2z^2 + \dots + a_nz^n = P_n(z).$$

es un ejemplo muy particular de función entera. Esto puede reducirse a una constante ($n = 0$). Si $n > 0$ y $a_n \neq 0$, se tiene: $\lim_{z \rightarrow \infty} P_n(z) = \infty$. Por consiguiente, un polinomio de grado superior a cero se hace igual a ∞ en el punto del infinito. Como se sabe por el álgebra, si w es un número complejo arbitrario (propio), la ecuación $P_n(z) = w$ tiene n raíces, algunas de las cuales pueden ser iguales entre sí (raíces múltiples). Por lo tanto, en la transformación $w = P_n(z)$ cada punto del plano w pertenece a la imagen del plano z ; además, este punto tendrá n preimágenes: z_1, z_2, \dots, z_n . Agreguemos a esto que $P_n(\infty) = \infty$ y, por consiguiente, ∞ pertenece a la imagen del plano ampliado. Las preimágenes del punto del infinito $w = \infty$ son las raíces de la ecuación $P_n(z) = \infty$, es decir, es también el punto del infinito. Para simetría, a éste le consideraremos como raíz múltiple de esta ecuación, de orden n . Así, pues, un *polinomio de grado n* ($a_n \neq 0, n > 0$) *transforma el plano ampliado sobre sí mismo, de modo que cada punto de la imagen w tiene n preimágenes: z_1, z_2, \dots, z_n* . Por cierto, como ya se advirtió, para algunos valores excepcionales de w (entre los cuales también está incluido $w = \infty$) el número de preimágenes puede ser también menor que n .

Fácilmente se observa que la cantidad de tales valores excepcionales no es superior a n . En efecto, si $w_0 \neq \infty$ y la ecuación $P_n(z) = w_0$ tiene raíces múltiples, entonces, como se sabe por el álgebra, para cada una de ellas $P'_n(z) = 0$. Pero esta última ecuación tiene $n - 1$ raíces (entre las cuales también puede haber iguales entre sí): $\xi_1, \xi_2, \dots, \xi_{n-1}$. De aquí se deduce que w_0 tiene que tener alguno de los siguientes $n - 1$ valores: $P_n(\xi_1), \dots, P_n(\xi_{n-1})$, y agregando aquí el punto del infinito, obtenemos aquellos n puntos (no más) del plano w que poseen menos que n preimágenes cada uno en el plano z .

3.2. En virtud de la teoría general, la transformación $w = P_n(z)$ es conforme en todos los puntos, a excepción de los puntos

$\xi_1, \xi_2, \dots, \xi_{n-1}$, en los cuales la derivada se anula, y también, posiblemente, a excepción del punto $z = \infty$.

Cuando $n = 1$, el polinomio es una función lineal entera, y en este caso, la transformación es biunívoca y conforme, incluyendo también el infinito (véanse los ap. 2.3—2.4). Cuando $n > 1$, la transformación deja de ser conforme en los puntos indicados.

En efecto, supongamos que $P'_n(z_0) = 0$. Entonces, $z = z_0$ es una raíz múltiple de la ecuación $P'_n(z) - P'_n(z_0) = 0$ y, por consiguiente, $P_n(z) - P_n(z_0)$ puede expresarse en la forma:

$$P_n(z) - P_n(z_0) = (z - z_0)^k Q(z),$$

donde $k \geq 2$ es el orden de multiplicidad de la raíz $z = z_0$ (como es sabido, el número k es una unidad mayor que el orden de multiplicidad de la raíz $z = z_0$ para la ecuación $P'_n(z) = 0$) y el polinomio $Q(z)$ no se anula en el punto $z = z_0$. Poniendo $P_n(z) = w$ y $P_n(z_0) = w_0$, obtenemos de aquí, que

$$\text{Arg}(w - w_0) = \text{Arg}(z - z_0)^k + \text{Arg} Q(z),$$

de donde

$$\lim_{z \rightarrow z_0} \{ \text{Arg}(w - w_0) - \text{Arg}(z - z_0)^k \} = \text{Arg} Q(z_0).$$

Supongamos ahora que $z = \lambda(t)$ es una curva L que pasa por el punto z_0 ($z_0 = \lambda(t_0)$) y que en este punto tiene tangente cuyo ángulo de inclinación al eje real es igual a

$$\text{Arg} \lambda'(t_0) = \lim_{t_1 \rightarrow t_0} \text{Arg} \frac{z_1 - z_0}{t_1 - t_0}$$

(véase el ap. 2.1). La imagen Λ de esta curva en el plano w será: $w = P_n[\lambda(t)] = \mu(t)$. De aquí no podemos sacar inmediatamente la conclusión de que existe tangente a Λ en el punto w_0 ($t = t_0$) (puesto que $\mu'(t_0) = P'_n(z_0) \lambda'(t_0) = 0$). Pero, para el ángulo de inclinación de la secante que pasa por los puntos w_0 y $w \neq w_0$, según la anterior, obtenemos:

$$\text{Arg} \frac{w - w_0}{t - t_0} =$$

$$= \text{Arg} \frac{w - w_0}{(z - z_0)^k} + k \text{Arg} \frac{(z - z_0)}{t - t_0} \rightarrow \text{Arg} Q(z_0) + k \text{Arg} \lambda'(t_0) \text{ cuando } t \rightarrow t_0,$$

de donde se deduce que existe la tangente.

Si L_1 y L_2 son dos curvas: $z = \lambda_1(t)$ y $z = \lambda_2(t)$, que pasan por el punto z_0 y forman en éste un ángulo θ :

$$\theta = \text{Arg} \lambda'_2(t_2) - \text{Arg} \lambda'_1(t_1) \quad (\lambda_1(t_1) = \lambda_2(t_2) = z_0),$$

entonces, las imágenes Λ_1 y Λ_2 de estas curvas pasan por el punto w_0 y forman en éste el ángulo

$$\begin{aligned} [\text{Arg } Q(z_0) + k \text{Arg } \lambda'_2(t_2)] - [\text{Arg } Q(z_0) + k \text{Arg } \lambda'_1(t_1)] &= \\ = k [\text{Arg } \lambda'_2(t_2) - \text{Arg } \lambda'_1(t_1)] &= k\theta. \end{aligned}$$

Resumiendo, en la transformación $w = P_n(z)$ todos los ángulos con los vértices en los puntos, en los cuales se anula la derivada $P'_n(z)$, se alteran; precisando: estos ángulos aumentan k veces si el orden de multiplicidad de la raíz correspondiente de la ecuación $P'_n(z) = 0$ es igual a $k - 1$.

Aplicando la transformación $\zeta = \frac{1}{z}$, el lector se convencerá fácilmente que cuando $n > 1$ la transformación deja de ser conforme también en el punto del infinito. Precisamente, en la transformación $w = P_n(z)$, los ángulos con el vértice en el punto del infinito aumentan n veces.

3.3. Examinemos, en particular, la transformación de la forma $w = (z - a)^n$ ($n > 1$). Esta transforma el plano ampliado sobre sí mismo, de modo que cada punto w posee n preimágenes en el plano z . Son una excepción los puntos $w = 0$ y $w = \infty$, para los cuales, las preimágenes se confunden en un punto: a y ∞ , respectivamente. Las preimágenes $z (\neq a, \neq \infty)$ se determinan de la ecuación

$$w = (z - a)^n,$$

de modo que

$$z = a + \sqrt[n]{w - a} = a + \sqrt[n]{|w - a|} \left(\cos \frac{\text{Arg } w}{n} + i \sin \frac{\text{Arg } w}{n} \right).$$

Evidentemente, estos n puntos están situados en los vértices de un polígono regular de n lados con el centro en a .

La transformación $w = (z - a)^n$ es conforme en todos los puntos, a excepción de los puntos $z = a$ y $z = \infty$. En este caso los ángulos con los vértices en los dos últimos puntos aumentan n veces.

Para obtener una idea más clara de esta transformación, señalemos que

$$|w| = |z - a|^n \quad \text{y} \quad \text{Arg } w = n \text{Arg } (z - a).$$

De aquí se deduce que cada circunferencia de radio r con el centro en el punto $z = a$ se transforma en una circunferencia de radio r^n con el centro en el punto $w = 0$. Si, en este caso, el punto z recorre una vez la circunferencia $|z - a| = r$ en dirección positiva (es decir, que el $\text{Arg } (z - a)$, creciendo continuamente, aumenta en 2π), el punto w recorrerá n veces la circunferencia $|w| = r^n$ en la misma dirección (es decir, que el $\text{Arg } w$, creciendo continuamente, aumentará en $2\pi n$). Hagamos ahora recorrer al punto z a lo largo

del rayo rectilíneo $\text{Arg}(z - a) = \varphi_0 + 2k\pi$, desde el punto a hasta el infinito. Nuestras fórmulas muestran que el punto correspondiente w recorrerá entonces el rayo rectilíneo $\text{Arg } w = n\varphi_0 + 2m\pi$, desde el origen de coordenadas hasta el infinito.

Consideremos ahora la región g , que representa el interior del ángulo de magnitud θ , $0 < \theta \leq \frac{2\pi}{n}$, con el vértice en el punto a . Supongamos que este ángulo está limitado por los rayos rectilíneos:

$$\text{Arg}(z - a) = \varphi_0 + 2k\pi$$

y

$$\begin{aligned} \text{Arg}(z - a) &= \varphi_1 + 2m\pi, \\ (\varphi_1 - \varphi_0 &= \theta). \end{aligned}$$

De lo dicho se deduce que la imagen de la región g en el plano w es una región d que representa un ángulo de magnitud $n\theta$ con el

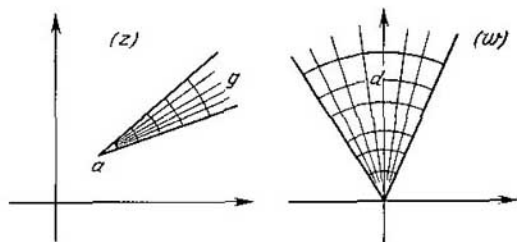


FIG. 11

vértice en el origen de coordenadas, limitado por rayos rectilíneos (fig. 11). La correspondencia entre g y d , establecida mediante la función $w = (z - a)^n$, es biunívoca. En efecto, como la función $w = (z - a)^n$ es uniforme, para comprobar esta afirmación es suficiente establecer que cada punto w de la región d tiene solamente una preimagen en la región g . Obsérvese para esto que todas las n preimágenes del punto w se sitúan en el plano z en los vértices de un polígono regular de n lados con el centro en a , de modo que dos de ellas podrían situarse dentro de un mismo ángulo con el vértice en a solamente cuando la magnitud del ángulo fuese mayor que $\frac{2\pi}{n}$.

Pero la medida del ángulo g no es superior a $\frac{2\pi}{n}$, por consiguiente, al ángulo g solamente pertenece una imagen de cada uno de los puntos de d , con lo cual queda terminada la demostración de nuestra proposición.

Así, pues, la función $w = (z - a)^n$ realiza una transformación biunívoca y conforme del interior de cualquier ángulo de lados rectilíneos, con el vértice en el punto a y de magnitud θ , $0 < \theta \leq \frac{2\pi}{n}$, sobre el interior de un ángulo correspondiente de lados rectilíneos también, con el vértice en el origen de coordenadas y de magnitud $n\theta$.

Por esto mismo, se recurre a la función considerada siempre que sea necesario transformar un ángulo de lados rectilíneos en otro ángulo que sea unas cuantas veces mayor.

Indudablemente, sería erróneo creer que en la transformación $w = (z - a)^n$ ($n > 1$) cualquier recta se transforma en recta y cualquier circunferencia en circunferencia. Supongamos, por ejemplo, que $a = 0$ y $n = 2$. Entonces se obtiene la función $w = z^2$. Veamos en qué se transforman, mediante la función $w = z^2$, las rectas que no pasan por el origen de coordenadas y que son paralelas a uno de los ejes coordenados. Tomemos, por ejemplo, una recta paralela al eje imaginario: $z = c + it$, $c \neq 0$, $-\infty < t < +\infty$. Como imagen obtendremos la línea $w = (c + it)^2$, o bien, poniendo $w = u + iv$ y separando las partes real e imaginaria:

$$u = c^2 - t^2, \quad v = 2ct, \quad -\infty < t < +\infty.$$

Estas son las ecuaciones de la línea transformada, expresadas en coordenadas cartesianas en la forma paramétrica. Excluyendo entre éstas el parámetro t , resulta:

$$v^2 = 4c^2(c^2 - u).$$

Esta es la ecuación de una parábola, cuyo eje va dirigido por el eje real hacia el lado negativo, con el foco en el origen de coordenadas y con el parámetro $p = 2c^2$. Del mismo modo se observa que cada recta paralela al eje real: $z = t + ic'$ se transforma en la parábola

$$v^2 = 4c'^2(u + c'^2)$$

cuyo eje lleva la dirección del eje real hacia el lado positivo, con el foco en el origen de coordenadas y con el parámetro $p' = 2c'^2$. Resumiendo, las dos familias de rectas paralelas a los ejes coordenados se transforman mediante la función $w = z^2$ en dos familias de parábolas con el foco común en el origen y cuyos ejes están situados en el eje real (fig. 12). Como las familias de rectas son ortogonales entre sí y la transformación es conforme, las familias de parábolas obtenidas también serán ortogonales entre sí; esto se comprueba también fácilmente mediante un cálculo inmediato.

Naturalmente, el lector tiene que tener presente que la transformación de todo el plano z mediante la función $w = z^2$ no es biunívoca, puesto que cada punto w , distinto de cero y del infinito,

poseo dos preimágenes. En particular, las preimágenes de la parábola $v^2 = 4c^2(c^2 - u)$ son dos rectas simétricas respecto del eje imaginario: $z = c + it$ y $z = -c + it$; del mismo modo, las preimágenes de la parábola $v^2 = 4c'^2(u + c'^2)$ son dos rectas simétricas respecto del eje real: $z = t + ic'$ y $z = t - ic'$. Pero si se considera solamente la imagen de algún semiplano g , limitado por una recta que pase por el origen de coordenadas (tal semiplano

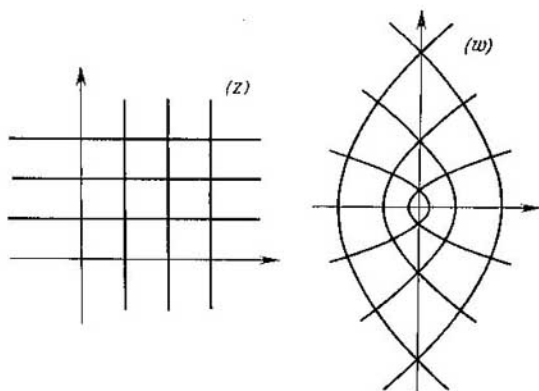


FIG. 12

representa el interior de un ángulo con el vértice en el origen de coordenadas de magnitud π), entonces, según lo expuesto anteriormente, la correspondencia entre g y su imagen d será biunívoca; d representará aquí un ángulo de magnitud 2π con el vértice en el origen de coordenadas; ambos lados de este ángulo se confunden en un rayo rectilíneo que parte del origen de coordenadas.

3.4. Las funciones enteras distintas de los polinomios, se llaman *funciones trascendentes enteras*. La más simple de éstas es la *función exponencial* $\exp z$ o e^z . Esta se obtiene como resultado de la generalización de la función exponencial de variable real e^x al plano complejo. Como es sabido, la función $f(x) = e^x$ se caracteriza completamente por sus propiedades: ésta

- 1) está definida unívocamente para todos los valores reales x , toma valores reales y para $x = 1$ el valor e ;
- 2) satisface al teorema de la suma

$$f(x_1 + x_2) = f(x_1) f(x_2)$$

para cualesquiera x_1 y x_2 ;

3) es continua en el punto $x = 0$.

Aquí construiremos la función exponencial de variable compleja $f(z) = \exp z$, exigiendo el cumplimiento de las siguientes condiciones: ésta

1') está definida unívocamente para todos los valores complejos (finitos) de z , para los valores reales $z = x$ toma también valores reales y para $x = 1$ toma el valor e ;

2') satisface al teorema de la suma

$$f(z_1 + z_2) = f(z_1) f(z_2)$$

para cualesquiera z_1 y z_2 ;

3') es diferenciable en el punto $z = 0$.

De las condiciones 1') y 2') se deduce que $f(0) = 1$. En efecto, $f(0) \cdot f(1) = f(1) \neq 0$ y, por consiguiente, $f(0) = 1$. Luego, sacamos la conclusión de que $f(z) \neq 0$ para todos los valores de z . En efecto, $f(z) \cdot f(-z) = f(0) = 1$, de donde se deduce que $f(z) \neq 0$. Aplicando 2') y 3'), obtenemos para cualquier z :

$$\frac{f(z + \Delta z) - f(z)}{\Delta z} = \frac{f(z) \cdot f(\Delta z) - f(z)}{\Delta z} = \frac{f(\Delta z) - f(0)}{\Delta z} f(z) \rightarrow f'(0) f(z)$$

cuando $\Delta z \rightarrow 0$. Esto significa que $f(z)$ es una función analítica en todo el plano (entera), y satisface a la ecuación diferencial

$$f'(z) = f'(0) f(z). \tag{3.4:1}$$

En particular, $f'(x) = f'(0) f(x)$, de donde $f(x) = e^{f'(0)x + C}$, de las condiciones $f(0) = 1$ y $f(1) = e$ se deduce que $C = 0$ y $f'(0) = 1$. Así, pues, $f(x) = e^x$; por cierto, esta conclusión se deduce inmediatamente de que $f(x)$ satisface a las condiciones 1), 2) y 3).

Poniendo en (3.4:1) el valor $f'(0)$, resulta:

$$f'(z) = f(z). \tag{3.4:2}$$

Expresemos $f(z)$ en la forma

$$f(z) = f(x + iy) = f(x) f(iy) = e^x |\alpha(y) + i\beta(y)|. \tag{3.4:3}$$

Entonces,

$$f'(z) = \frac{\partial |e^x \beta(y)|}{\partial y} - i \frac{\partial |e^x \alpha(y)|}{\partial y} = e^x [\beta'(y) - i\alpha'(y)].$$

De la ecuación (3.4:2), obtenemos:

$$\beta'(y) = \alpha(y), \quad \alpha'(y) = -\beta(y),$$

de donde $\beta''(y) + \beta(y) = 0$ y, por consiguiente:

$$\beta(y) = C_1 \cos y + C_2 \sin y, \quad \alpha(y) = \beta'(y) = -C_1 \sin y + C_2 \cos y.$$

Pero $f(0) = \alpha(0) + i\beta(0) = 1$, es decir, $\alpha(0) = 1$ y $\beta(0) = 0$; por lo tanto, $C_1 = 0$ y $C_2 = 1$. En resumen,

$$\alpha(y) = \cos y, \quad \beta(y) = \operatorname{sen} y$$

y según la fórmula (3.4:3)

$$f(z) = e^x (\cos y + i \operatorname{sen} y).$$

Hemos obtenido la única función que satisface a las condiciones 1') 2') y 3'). Esta se llama función exponencial (de variable compleja) y se designa mediante $\exp z$ o e^z . Así, pues, según la definición

$$\exp z = e^z = e^x (\cos y + i \operatorname{sen} y). \quad (3.4:3)$$

Obsérvese que si la condición 3') se sustituye por otra más general: 3'') la función $f(z)$ es continua en el punto $z = 0$, entonces se puede hallar un conjunto infinito de funciones distintas que satisfacen a las condiciones 1'), 2') y 3''). Todas ellas están comprendidas en la fórmula

$$f(z) = e^{x+\alpha y} (\cos \alpha y + i \operatorname{sen} \alpha y),$$

donde a y α son números reales. Sin embargo, solamente una de ellas, la que se obtiene cuando $a = 0$, $\alpha = 1$, es analítica y además entera.

3.5. De la definición de función exponencial

$$\exp z = e^x (\cos y + i \operatorname{sen} y) \quad (3.5:1)$$

se deduce que ésta no se anula para ningún valor de z y que

$$|\exp z| = e^x \quad \text{y} \quad \operatorname{Arg}(\exp z) = y + 2k\pi.$$

Para $z = iy$ ($x = 0$) obtenemos:

$$\exp(iy) = \cos y + i \operatorname{sen} y.$$

Esta relación permite utilizar, en lugar de la forma trigonométrica del número complejo

$$c = r (\cos \varphi + i \operatorname{sen} \varphi) \quad (r \neq 0),$$

la forma exponencial

$$c = r \exp(i\varphi) = r e^{i\varphi},$$

que es más compacta.

De las fórmulas (3.5:1) vemos que la función exponencial posee período, igual a $2\pi i$ (puesto que, al variar y en 2π , z varía en $2\pi i$, y el valor de la función no se altera):

$$\exp(z + 2\pi i) = \exp z.$$

Demostremos que $2\pi i$ es el período fundamental (p r i m i t i v o) de la función exponencial, es decir, que cualquier otro período de la misma tiene que tener la forma $2k\pi i$, donde k es un número entero.

En efecto, sea $\omega = \alpha + \beta i$ un período de la función exponencial. Entonces

$$\exp(z + \omega) = \exp z$$

para cualquier z y, en particular, para $z = 0$

$$\exp \omega = \exp(\alpha + i\beta) = e^\alpha (\cos \beta + i \operatorname{sen} \beta) = 1.$$

Pero esto significa que $e^\alpha = 1$, es decir, $\alpha = 0$ y $\cos \beta + i \operatorname{sen} \beta = 1$, es decir, $\beta = 2k\pi$. Por consiguiente,

$$\omega = \alpha + i\beta = 2k\pi i,$$

como se quería demostrar.

La expresión $\exp \infty$ carece de sentido, puesto que $\lim e^z$ no existe. Para convencerse de esto es suficiente observar que $e^x \rightarrow \infty$ cuando $x > 0$ y tiende a ∞ , y $e^x \rightarrow 0$ cuando $x < 0$ y tiende a $-\infty$.

De aquí, en particular, se deduce que $\exp z$ no coincide con ninguno de los polinomios, es decir, que verdaderamente es una función trascendente entera. En efecto, todo polinomio, que no se reduce a una constante, tiende al infinito cuando $z \rightarrow \infty$.

En adelante (en el cap. 4 ap. 33), se demostrará que ninguna función trascendente entera puede tener límite en el punto $z = \infty$.

Para la derivada de la función exponencial obtenemos:

$$(\exp z)' = \frac{\partial (e^x \cos y)}{\partial x} + i \frac{\partial (e^x \operatorname{sen} y)}{\partial x} = e^x (\cos y + i \operatorname{sen} y) = \exp z.$$

Por consiguiente, la derivada de la función exponencial no se anula para ningún valor z .

Estudiemos el comportamiento geométrico de la función $w = \exp z$ o, lo que es lo mismo, la transformación que esta función realiza. Ya se observó que esta función no toma el valor $w = 0$ en ningún punto z . Esto significa que el origen de coordenadas del plano w no pertenece a la imagen del plano finito z en la transformación $w = \exp z$. Demostremos que cualquier otro punto finito del plano w pertenece a esta imagen. En efecto, de la ecuación $w = \exp z$, donde $w \neq 0$ está dado mientras que $z = x + iy$ es la incógnita, obtenemos:

$$|w| = e^x, \text{ de donde } x = \ln |w| \text{ y } \operatorname{Arg} w = y + 2\pi k, \text{ o sea } y = \operatorname{Arg} w.$$

Resumiendo, solamente pueden ser preimágenes de los puntos w los puntos de la forma

$$z = \ln |w| + i \operatorname{Arg} w.$$

Evidentemente, hay una infinidad de puntos de éstos, puesto que $\text{Arg } w$ toma un conjunto infinito de valores, que se diferencian dos a dos en múltiplos enteros de 2π . Además, cada uno de los puntos hallados es verdaderamente una preimagen del punto w , ya que

$$\begin{aligned} \exp(\ln |w| + i \text{Arg } w) &= e^{\ln |w|} (\cos \text{Arg } w + i \text{sen } \text{Arg } w) = \\ &= |w| (\cos \text{Arg } w + i \text{sen } \text{Arg } w) = w. \end{aligned}$$

En resumen, el conjunto de todas las raíces de la ecuación $w = e^z$ ($w \neq 0$) se expresa por la fórmula

$$z = \ln |w| + i \text{Arg } w = \ln |w| + i(\arg w + 2k\pi), \quad (3.5:2)$$

donde $k = 0, \pm 1, \pm 2, \dots$

Todos estos puntos están situados en una recta, paralela al eje imaginario, a las distancias entre sí de 2π .

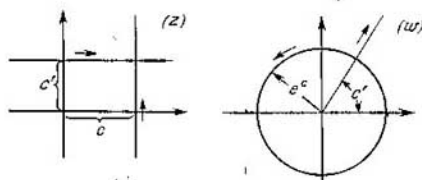


FIG. 13

Vemos, pues, que la función $w = \exp z$ transforma el plano finito z en el recinto que se obtiene del plano finito w excluyendo un punto $w = 0$; esta transformación no es biunívoca, puesto que cada punto $w \neq 0$ posee un conjunto infinito de preimágenes (3.5:2).

Como la derivada de la función exponencial siempre es distinta de cero, esta transformación es conforme en todos los puntos del plano finito z .

Supongamos que z recorre alguna recta paralela a uno de los ejes coordenados (fig. 13). Si ésta es una recta $z = c + it$ paralela al eje imaginario, entonces $w = e^c (\cos t + i \text{sen } t)$, es decir, w estará situado en una circunferencia de radio e^c con el centro en el origen de coordenadas. Además, si el punto z recorre la recta una sola vez, de modo que la ordenada de este punto, igual a t , crece continuamente desde $-\infty$ hasta $+\infty$, entonces w describirá infinitas veces la circunferencia correspondiente en una misma dirección positiva.

Si el punto z recorre una recta $z = t + ic'$ paralela al eje real, entonces $w = e^t (\cos c' + i \operatorname{sen} c')$ recorrerá, evidentemente, un rayo rectilíneo que parte del origen de coordenadas y forma el ángulo c' con la parte positiva del eje real. En este caso, cuando z recorra la recta una sola vez, de modo que la abscisa de este punto, igual a t , crezca continuamente desde $-\infty$ hasta $+\infty$, el punto w describirá una sola vez el rayo correspondiente, de modo que la distancia desde este punto hasta el origen de coordenadas crecerá continuamente desde 0 hasta ∞ (naturalmente, se excluyen tanto un extremo como el otro, puesto que $|w| = e^t$).

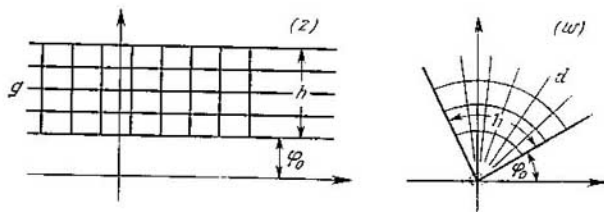


FIG. 14

Resumiendo, en la transformación del plano z mediante la función $w = e^z$, la familia de rectas paralelas al eje imaginario se transforma en la familia de circunferencias con centro en el origen de coordenadas, y la familia de rectas paralelas al eje real, en la familia de rayos rectilíneos que parten del origen de coordenadas.

Consideremos la región g formada por el interior de la franja rectilínea de anchura h , $0 < h \leq 2\pi$, paralela al eje real. Supongamos que esta franja está limitada por las rectas: $y = \varphi_0$ e $y = \varphi_1$ ($\varphi_1 - \varphi_0 = h$). De lo expuesto anteriormente se deduce que la imagen de la región g en el plano w será la región d formada por el ángulo de magnitud h con el vértice en el origen de coordenadas, limitado por los rayos rectilíneos $\operatorname{Arg} w = \varphi_0 + 2k\pi$ y $\operatorname{Arg} w = \varphi_1 + 2l\pi$ (fig. 14). La correspondencia entre las regiones g y d , establecida mediante la función $w = \exp z$, es ahora biunívoca. Para comprobar esto, es suficiente observar que solamente pueden ser preimágenes de algún punto w de la región d los puntos $\ln |w| + i \operatorname{Arg} w$, que se diferencian entre sí por los valores en la parte imaginaria. Dos puntos de éstos están situados en una recta paralela al eje imaginario a una distancia múltiple de 2π . Pero nuestra franja h tiene una anchura no superior a 2π , por lo cual, puede contener en su interior solamente una preimagen del punto w . Por lo tanto, cada punto $z \in g$ posee solamente una imagen y cada punto

$w \in d$ solamente una preimagen en el interior de g , lo cual demuestra que la transformación es biunívoca.

Vemos, pues, que la función exponencial $w = \exp z$ realiza una transformación biunívoca y conforme de una franja de anchura $h \leq 2\pi$, paralela al eje real, en un ángulo de magnitud h con el vértice en el origen de coordenadas.

Por esto mismo se recurre a la función exponencial siempre que se necesita transformar conformemente alguna franja rectilínea en el interior de un ángulo.

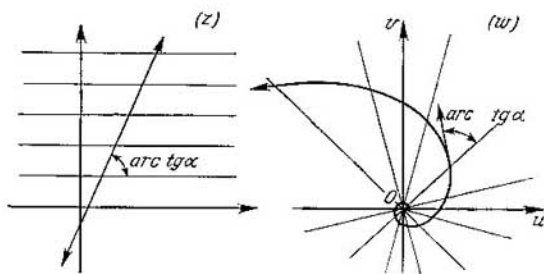


FIG. 15

Si la recta del plano z no es paralela a alguno de los ejes coordenados, su imagen en el plano w no será ya una recta o una circunferencia, sino una espiral logarítmica. En efecto, si esta recta es

$$z = t(1 + i\alpha) + bi, \quad -\infty < t < +\infty$$

(α es el coeficiente angular de la recta y b , la ordenada en el origen), entonces su imagen será la curva

$$w = \exp [t + i(\alpha t + b)] = e^t [\cos (\alpha t + b) + i \operatorname{sen} (\alpha t + b)].$$

Aquí

$$|w| = r = e^t, \quad \varphi = \operatorname{Arg} w = \alpha t + b + 2m\pi,$$

o bien, eliminando el parámetro t : $r = \exp \frac{\varphi - b - 2m\pi}{\alpha}$. Pero $\operatorname{Arg} w$ o el ángulo polar φ se determina salvo un entero múltiplo de 2π . Por lo tanto, designando de nuevo $\varphi - 2m\pi$ mediante φ , obtenemos:

$$r = ce^{\frac{\varphi}{\alpha}}, \quad \text{donde } c = e^{-\frac{b}{\alpha}}.$$

Esta es la ecuación de una espiral logarítmica (fig. 15). Como ésta es la imagen de la recta $z = t(1 + i\alpha) + bi$, la cual

se corta con las rectas paralelas al eje real bajo un ángulo constante igual a $\arctg \alpha$, y como al transformación es conforme, la espiral logarítmica se cortará bajo el mismo ángulo con las imágenes de dichas rectas, es decir, con todos los rayos que partan del origen de coordenadas. Hemos obtenido la propiedad característica de la espiral logarítmica.

Las transformaciones que se realizan mediante las funciones $w = (z - a)^n$ y $w = \exp z$ experimentan cierta semejanza entre sí. Esta semejanza puede aclararse mediante la fórmula

$$\exp z = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n,$$

cuya demostración proponemos hacer al lector como ejercicio.

Examinemos la transformación $w = \left(1 + \frac{z}{n}\right)^n = \frac{1}{n^n} [z - (-n)]^n$ con relación a la cual la transformación $w = \exp z$ es límite. En

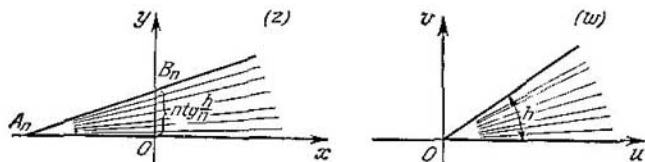


FIG. 16

virtud de lo expuesto en el ap. 3.3, esta función transforma el ángulo de magnitud $\frac{h}{n}$ ($0 < h \leq 2\pi$) con el vértice en el punto $A_n (-n)$, limitado por la parte del eje real $x \geq -n$ ($y = 0$) y el rayo $\text{Arg}(z + n) = \frac{h}{n} + 2k\pi$, en el ángulo de magnitud h con el vértice en el origen de coordenadas, limitado por los rayos $\text{Arg } w = 0$ y $\text{Arg } w = h + 2m\pi$ (fig. 16). Cuando n tiende al infinito, el vértice A_n se aleja al infinito a lo largo de la parte negativa del eje real y la longitud del segmento OB_n tiende al límite $\lim_{n \rightarrow \infty} n \operatorname{tg} \frac{h}{n} = h$, de modo que la posición límite del rayo $A_n B_n$ es la recta $y = h$, la cual, junto con el eje real, limita una franja de anchura h . Además, es evidente que la posición límite para los rayos que parten del vértice del ángulo serán unas rectas paralelas al eje real, y las posiciones límites para los arcos de las circunferencias con centro en el punto A_n serán los segmentos de las rectas perpendiculares al eje real comprendidos en el interior de la franja. Como vemos, el cuadro de la transformación realizada por la función exponencial

puede obtenerse del cuadro correspondiente de la transformación realizada por la función potencial, mediante el paso debido al límite.

3.6. Pasemos ahora a definir el seno y el coseno de un argumento complejo. De las fórmulas

$$\exp(ix) = \cos x + i \operatorname{sen} x \quad \text{y} \quad \exp(-ix) = \cos x - i \operatorname{sen} x$$

obtenemos las fórmulas conocidas de Euler:

$$\cos x = \frac{\exp(ix) + \exp(-ix)}{2}, \quad \operatorname{sen} x = \frac{\exp(ix) - \exp(-ix)}{2i},$$

que, por lo tanto, son válidas para cualquier valor real de x . Como los segundos miembros de estas fórmulas están definidos para cualquier valor complejo de z ($z \neq \infty$) y, evidentemente, son funciones analíticas de z , tenemos aquí dos funciones enteras de z :

$$\frac{\exp(iz) + \exp(-iz)}{2} \quad \text{y} \quad \frac{\exp(iz) - \exp(-iz)}{2i},$$

que para valores reales de $z = x$ toman valores reales que coinciden con $\cos x$ y $\operatorname{sen} x$, respectivamente. Es natural que, por definición, la primera de ellas se denote mediante $\cos z$, la segunda mediante $\operatorname{sen} z$ y que se llamen *funciones trigonométricas principales — coseno y seno de z* :

$$\cos z = \frac{\exp(iz) + \exp(-iz)}{2}, \quad \operatorname{sen} z = \frac{\exp(iz) - \exp(-iz)}{2i}. \quad (3.6:1)$$

Las fórmulas (3.6:1) se denominan *fórmulas de Euler*. También se llama fórmula de Euler la que se obtiene al multiplicar ambos miembros de la segunda fórmula por i y al sumar después el resultado con la primera fórmula

$$\exp(iz) = \cos z + i \operatorname{sen} z. \quad (3.6:2)$$

De las fórmulas (3.6:1) se deduce inmediatamente que $\cos z$ es una función par, y $\operatorname{sen} z$, una función impar:

$$\cos(-z) = \cos z, \quad \operatorname{sen}(-z) = -\operatorname{sen} z. \quad (3.6:3)$$

De las mismas fórmulas (3.6:1) se deduce que $\cos z$ y $\operatorname{sen} z$ poseen período, igual a 2π (puesto que, al variar z en 2π , los argumentos de las funciones exponenciales en los segundos miembros de las fórmulas varían en $\pm 2\pi i$, los cuales son períodos de la función exponencial). Demostremos que 2π es el período primitivo (principal) de las funciones $\cos z$ y $\operatorname{sen} z$. En efecto, si ω es un período de la función $\cos z$, se tiene:

$$\cos(z + \omega) = \cos z,$$

y para $z = \frac{\pi}{2}$ obtenemos:

$$\cos\left(\omega + \frac{\pi}{2}\right) = 0.$$

Pero de aquí se deduce que

$$\exp\left[i\left(\omega + \frac{\pi}{2}\right)\right] + \exp\left[-i\left(\omega + \frac{\pi}{2}\right)\right] = 0,$$

o bien

$$\exp[i(2\omega + \pi)] = -1.$$

Por consiguiente, según la fórmula (3.5:2) $i(2\omega + \pi) = \ln|-1| + i \operatorname{Arg}(-1) = i(\pi + 2k\pi)$, o sea, $\omega = k\pi$, y como $\cos \omega = \cos 0 = 1$, el número k es par y $\omega = 2k\pi$.

De un modo semejante se demuestra que 2π es también el período primitivo de la función $\operatorname{sen} z$.

Dediquémonos ahora a demostrar los teoremas de adición para las funciones $\cos z$ y $\operatorname{sen} z$, es decir, a buscar las relaciones existentes entre $\cos(z_1 + z_2)$ y $\operatorname{sen}(z_1 + z_2)$, de un lado, y $\cos z_1$, $\cos z_2$, $\operatorname{sen} z_1$ y $\operatorname{sen} z_2$, de otro lado (z_1 y z_2 son unos números complejos arbitrarios). Las relaciones pedidas se obtendrán como consecuencia del teorema de adición para la función exponencial.

Sustituyendo en la fórmula (3.6:2) z por $z_1 + z_2$, hallamos:

$$\begin{aligned} \cos(z_1 + z_2) + i \operatorname{sen}(z_1 + z_2) &= \exp[i(z_1 + z_2)] = \\ &= \exp(iz_1) \cdot \exp(iz_2) = (\cos z_1 + i \operatorname{sen} z_1)(\cos z_2 + i \operatorname{sen} z_2) \end{aligned}$$

o bien, efectuando la multiplicación:

$$\begin{aligned} \cos(z_1 + z_2) + i \operatorname{sen}(z_1 + z_2) &= \\ &= (\cos z_1 \cos z_2 - \operatorname{sen} z_1 \operatorname{sen} z_2) + i(\operatorname{sen} z_1 \cos z_2 + \cos z_1 \operatorname{sen} z_2). \end{aligned}$$

Poniendo aquí $-z_1$ y $-z_2$ en lugar de z_1 y z_2 y aplicando las relaciones (3.6:3), resulta:

$$\begin{aligned} \cos(z_1 + z_2) - i \operatorname{sen}(z_1 + z_2) &= \\ &= (\cos z_1 \cos z_2 - \operatorname{sen} z_1 \operatorname{sen} z_2) - i(\operatorname{sen} z_1 \cos z_2 + \cos z_1 \operatorname{sen} z_2). \end{aligned}$$

Sumando y restando estas fórmulas término a término, obtendremos:

$$\left. \begin{aligned} \cos(z_1 + z_2) &= \cos z_1 \cdot \cos z_2 - \operatorname{sen} z_1 \cdot \operatorname{sen} z_2, \\ \operatorname{sen}(z_1 + z_2) &= \operatorname{sen} z_1 \cdot \cos z_2 + \cos z_1 \cdot \operatorname{sen} z_2. \end{aligned} \right\} \quad (3.6:4)$$

Estas fórmulas son fundamentales en la teoría de las funciones trigonométricas. En particular, éstas contienen las denominadas «fórmulas de reducción del argumento». En efecto, poniendo en

las fórmulas (3.6:4) $z_1 = z$ y $z_2 = \frac{\pi}{2}$, obtenemos:

$$\cos\left(z + \frac{\pi}{2}\right) = \cos z \cos \frac{\pi}{2} - \operatorname{sen} z \operatorname{sen} \frac{\pi}{2} = -\operatorname{sen} z,$$

$$\operatorname{sen}\left(z + \frac{\pi}{2}\right) = \operatorname{sen} z \cos \frac{\pi}{2} + \cos z \operatorname{sen} \frac{\pi}{2} = \cos z.$$

Haciendo $z_1 = z$ y $z_2 = \pi$, hallamos otro par de fórmulas de reducción:

$$\cos(z + \pi) = -\cos z,$$

$$\operatorname{sen}(z + \pi) = -\operatorname{sen} z,$$

etc.

Poniendo en la primera de las fórmulas (3.6:4) $z_1 = z$ y $z_2 = -z$, obtenemos la siguiente relación entre $\operatorname{sen} z$ y $\cos z$:

$$1 = \cos^2 z + \operatorname{sen}^2 z. \quad (3.6:5)$$

Ya vemos que todas las relaciones conocidas de la trigonometría entre las funciones trigonométricas de argumento real se conservan también en el campo complejo. No obstante, de las fórmulas (3.6:5) no se puede sacar la conclusión de que $|\cos z| \leq 1$ y $|\operatorname{sen} z| \leq 1$, puesto que, por lo general, $\cos^2 z$ y $\operatorname{sen}^2 z$ no son números reales no negativos.

Con las funciones trigonométricas $\operatorname{sen} z$ y $\cos z$ están estrechamente ligadas las funciones hiperbólicas $\operatorname{ch} z$ y $\operatorname{sh} z$, definidas por las fórmulas

$$\operatorname{ch} z = \frac{\exp z + \exp(-z)}{2}, \quad \operatorname{sh} z = \frac{\exp z - \exp(-z)}{2}. \quad (3.6:6)$$

Cuando $z = x$ es real, estas funciones toman, evidentemente, valores reales y coinciden entonces con las funciones $\operatorname{ch} x$ y $\operatorname{sh} x$, conocidas en el análisis. La primera de éstas (es par) decrece en el semiintervalo $-\infty < x \leq 0$ desde ∞ hasta 1 y después crece desde 1 hasta ∞ en el semiintervalo $0 \leq x < \infty$; la segunda (es impar) crece en todo el intervalo infinito $-\infty < x < +\infty$ desde $-\infty$ hasta $+\infty$, anulándose para $x = 0$.

Comparando las fórmulas (3.6:6) con las fórmulas (3.6:1) se deduce que entre las funciones trigonométricas e hiperbólicas existen las siguientes relaciones:

$$\operatorname{ch} z = \cos(iz), \quad \operatorname{sh} z = -i \operatorname{sen}(iz). \quad (3.6:7)$$

De aquí, en particular, se deduce que

$$\operatorname{ch}^2 z - \operatorname{sh}^2 z = [\cos(iz)]^2 + [\operatorname{sen}(iz)]^2 = 1. \quad (3.6:8)$$

Determinemos las partes reales e imaginarias, y también los módulos de las funciones $\cos z$ y $\operatorname{sen} z$. Poniendo $z = x + iy$, obte-

nemos según las fórmulas (3.6:4) y (3.6:7):

$$\cos(x + iy) = \cos x \cos(iy) - \operatorname{sen} x \operatorname{sen}(iy) = \cos x \operatorname{ch} y - i \operatorname{sen} x \operatorname{sh} y, \quad (3.6:4')$$

$$\operatorname{sen}(x + iy) = \operatorname{sen} x \cos(iy) + \cos x \operatorname{sen}(iy) = \operatorname{sen} x \operatorname{ch} y + i \cos x \operatorname{sh} y.$$

De aquí

$$\left. \begin{aligned} \operatorname{Re}[\cos(x + iy)] &= \cos x \operatorname{ch} y, & \operatorname{Im}[\cos(x + iy)] &= -\operatorname{sen} x \operatorname{sh} y, \\ \operatorname{Re}[\operatorname{sen}(x + iy)] &= \operatorname{sen} x \operatorname{ch} y, & \operatorname{Im}[\operatorname{sen}(x + iy)] &= \cos x \operatorname{sh} y. \end{aligned} \right\} (3.6:9)$$

Para los módulos de las funciones $\cos z$ y $\operatorname{sen} z$ obtenemos las siguientes expresiones:

$$\begin{aligned} |\cos z| &= \sqrt{(\cos x \cdot \operatorname{ch} y)^2 + (\operatorname{sen} x \operatorname{sh} y)^2} = \\ &= \sqrt{\operatorname{ch}^2 y (1 - \operatorname{sen}^2 x) + \operatorname{sen}^2 x \cdot \operatorname{sh}^2 y} = \sqrt{\operatorname{ch}^2 y - \operatorname{sen}^2 x} \end{aligned}$$

$$\text{y análogamente, } |\operatorname{sen} z| = \sqrt{\operatorname{sh}^2 y + \operatorname{sen}^2 x}.$$

Así, pues,

$$|\cos z| = \sqrt{\operatorname{ch}^2 y - \operatorname{sen}^2 x}, \quad |\operatorname{sen} z| = \sqrt{\operatorname{sh}^2 y + \operatorname{sen}^2 x}. \quad (3.6:10)$$

De aquí se deducen las desigualdades:

$$\left. \begin{aligned} \operatorname{ch} y &\geq |\cos z| \geq \sqrt{\operatorname{ch}^2 y - 1} = |\operatorname{sh} y|, \\ \sqrt{\operatorname{sh}^2 y + 1} = \operatorname{ch} y &\geq |\operatorname{sen} z| \geq |\operatorname{sh} y|. \end{aligned} \right\} (3.6:11)$$

Por cierto, estas desigualdades se deducen inmediatamente de las fórmulas (3.6:1). Por ejemplo:

$$\begin{aligned} |\cos z| &\leq \frac{|\exp(iz)| + |\exp(-iz)|}{2} = \frac{\exp(-y) + \exp y}{2} = \operatorname{ch} y, \\ |\cos z| &\geq \left| \frac{|\exp(iz)| - |\exp(-iz)|}{2} \right| = \frac{|\exp(-y) - \exp y|}{2} = |\operatorname{sh} y|. \end{aligned}$$

Vemos que los módulos de las funciones $\cos z$ y $\operatorname{sen} z$ crecen indefinidamente junto con $|y|$ a medida que z se aleja del eje real, y que se cumplen las siguientes fórmulas asintóticas:

$$|\cos z| \approx \frac{1}{2} \exp|y|, \quad |\operatorname{sen} z| \approx \frac{1}{2} \exp|y|.$$

En la fig. 17 está representada la superficie: $u = |\operatorname{sen} z|$, denominada relieve del seno*). Como $\operatorname{sh} y \neq 0$ para $y \neq 0$,

*) El dibujo está adoptado de la «Tabla de funciones» de Jahake y Emde.

de las desigualdades (3.6:11) se deduce luego que $\cos z$ y $\sin z$ no pueden anularse fuera del eje real, es decir, que las ecuaciones $\cos z = 0$ y $\sin z = 0$ no poseen raíces imaginarias. Por consiguiente,

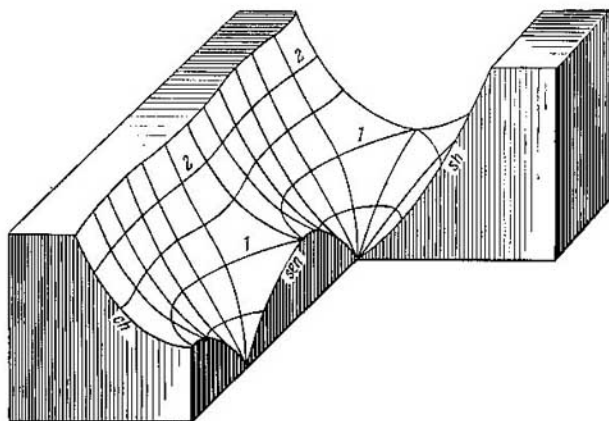


FIG. 17

todas las raíces de estas ecuaciones se reducen a las conocidas en la trigonometría:

$$z = (2k - 1) \frac{\pi}{2} \text{ para la ecuación } \cos z = 0$$

y

$$z = k\pi \text{ para la ecuación } \sin z = 0.$$

Señalemos también las fórmulas para las derivadas de las funciones trigonométricas e hiperbólicas:

$$(\cos z)' = \left[\frac{\exp(iz) + \exp(-iz)}{2} \right]' = i \frac{\exp(iz) - \exp(-iz)}{2} = -\sin z,$$

$$(\sin z)' = \cos z, \quad (\operatorname{ch} z)' = \operatorname{sh} z, \quad (\operatorname{ch} z)' = \operatorname{ch} z.$$

3.7. Ocupémonos del estudio del comportamiento geométrico de las funciones trigonométricas. En este caso podemos limitarnos a estudiar la transformación

$$w = \cos z,$$

puesto que la transformación $w = \sin z$ puede expresarse en la forma

$$w = -\cos \left(z + \frac{\pi}{2} \right)$$

y, por consiguiente, se reduce a la traslación del plano en dirección del eje real $z_1 = z + \frac{\pi}{2}$, a la transformación $z_2 = \cos z_1$ y, finalmente, a la rotación de todo el plano alrededor del origen de coordenadas en el ángulo π : $w = -z_2$.

Examinemos primero las preimágenes del punto w en la transformación $w = \cos z$, es decir, las raíces de la ecuación

$$w = \cos z, \quad (3.7:1)$$

donde w es un número complejo arbitrario, distinto de ∞ . Sustituyendo $\cos z$ según la fórmula de Euler (3.6:1) y poniendo para abreviar

$$\exp(iz) = t, \quad (3.7:2)$$

para determinar t obtenemos la ecuación

$$w = \frac{t + t^{-1}}{2}$$

o bien

$$t^2 - 2wt + 1 = 0, \quad (3.7:3)$$

de donde

$$t_j = w + \sqrt{w^2 - 1} \quad (j = 1, 2) \quad (3.7:4)$$

(ante la raíz cuadrada no ponemos el signo doble, puesto que esta raíz misma posee dos valores). Evidentemente, el producto de los números t_1 y t_2 es igual a 1, por lo cual cada uno de ellos es distinto de cero. Designando uno de estos mediante τ y el otro mediante $\frac{1}{\tau}$, obtenemos de (3.7:2) dos ecuaciones para determinar z :

$$\exp(iz) = \tau (\neq 0) \quad \text{y} \quad \exp(iz) = \frac{1}{\tau} (\neq 0). \quad (3.7:5)$$

Según el ap. 3.5, cada una de estas ecuaciones posee infinitas soluciones, que se expresan según la fórmula (3.5:2)

$$iz' = \ln |\tau| + i \operatorname{Arg} \tau$$

y

$$iz'' = \ln \left| \frac{1}{\tau} \right| - i \operatorname{Arg} \frac{1}{\tau} = -(\ln |\tau| + i \operatorname{Arg} \tau),$$

o bien,

$$z' = \operatorname{Arg} \tau - i \ln |\tau| \quad \text{y} \quad z'' = -(\operatorname{Arg} \tau - i \ln |\tau|). \quad (3.7:6)$$

Hemos obtenido dos conjuntos infinitos de puntos, situados sobre el par de rectas $y = \pm \ln |\tau|$, paralelas al eje real. En cada una de éstas los puntos vecinos z' , respectivamente z'' , están situados

a la distancia 2π unos de otros; además, para cada punto z' , situado en la recta $y = -\ln |\tau|$, existe en la otra recta $y = \ln |\tau|$ un punto z'' simétrico con z' respecto del origen de coordenadas (véase la fig. 18, donde $|\tau| < 1$). Cuando $w = \pm 1$ las raíces τ y $\frac{1}{\tau}$ de la ecuación (3.7:3) son iguales a ± 1 . En este caso, ambas rectas se confunden con el eje real y los dos conjuntos de puntos z' y z'' también coinciden.

Resumiendo, la ecuación (3.7:1) siempre posee soluciones y el conjunto de las soluciones siempre es infinito. De aquí se deduce, en primer lugar, que la función $w = \cos z$ transforma el plano finito

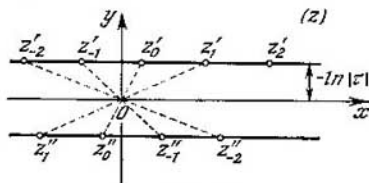


FIG. 18

z sobre todo el plano (finito) w, y , en segundo lugar, que cada punto w posee infinitas preimágenes en el plano z . Esta transformación es conforme en todos los puntos en los cuales $(\cos z)' = -\sin z \neq 0$, es decir, para $z \neq k\pi$ ($k = 0, \pm 1, \pm 2, \dots$).

Supongamos que z recorre una recta cualquiera paralela a uno de los ejes de coordenadas. Si ésta es la recta $z = c + it$, paralela al eje imaginario, entonces la imagen será la curva $L: w = \cos z = \cos c \operatorname{ch} t - i \operatorname{sen} c \operatorname{sh} t$ (véase la primera de las fórmulas (3.6:4')). Si $c = k\pi$, obtenemos $w = \cos k\pi \operatorname{ch} t = (-1)^k \operatorname{ch} t$ ($-\infty < t < +\infty$) o sea, w recorre dos veces la parte $u \geq 1$ del eje real cuando k es par y la parte $u \leq -1$ cuando k es impar. Si $c = (2k-1)\frac{\pi}{2}$ obtenemos $w = (-1)^k i \operatorname{sh} t$, o sea, w describe una vez todo el eje imaginario en dirección del crecimiento de u cuando k es par y en dirección del decrecimiento de u cuando k es impar.

Supongamos ahora que $c \neq m\frac{\pi}{2}$ (para cualquier entero m). Escribamos la ecuación de la curva L en la forma:

$$u = \cos c \cdot \operatorname{ch} t, \quad v = -\operatorname{sen} c \cdot \operatorname{sh} t \quad (-\infty < t < \infty) \quad (3.7:7)$$

o bien, eliminando el parámetro t ($\cos c \neq 0$ y $\operatorname{sen} c \neq 0$):

$$\frac{u^2}{\cos^2 c} - \frac{v^2}{\operatorname{sen}^2 c} = 1. \quad (3.7:8)$$

Hemos obtenido la ecuación de una hipérbola con los semiejes $|\cos c|$ y $|\operatorname{sen} c|$ y con los focos en los puntos ± 1 .

Sin embargo, no hay que creer que la curva L coincide con toda esta hipérbola. De la representación paramétrica (3.7:7) de L se deduce que u conserva todo el tiempo un mismo signo, que coincide con el signo de $\cos c$, mientras que v varía de una manera continua y monótona desde $-\infty$ hasta $+\infty$ (o al revés). De aquí se deduce que la curva L solamente coincide con una de las dos ramas de la hipérbola (3.7:8), precisamente con la rama de la derecha si $\cos c > 0$,

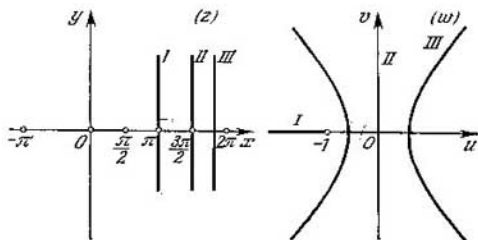


FIG. 19

y con la rama de la izquierda si $\cos c < 0$ (fig. 19, en la cual en la mitad de la derecha están representadas las imágenes de tres rectas z :

$$I (x = \pi), \quad II \left(x = \frac{3\pi}{2}\right) \quad \text{y} \quad III \left(x = c, \text{ donde } \frac{3\pi}{2} < c < 2\pi\right).$$

Además, la transformación de la recta $z = c + it$ en la rama correspondiente es biunívoca y cada una de las dos semirrectas, en las que se divide nuestra recta por el eje real, se transforma biunívocamente en una de las semirramas, en las que se divide la rama de la hipérbola en el vértice.

Supongamos ahora que z describe una recta l' : $z = t + ic'$, paralela al eje real. Su imagen será la curva L' :

$$w = \cos z = \cos t \operatorname{ch} c' - i \operatorname{sen} t \operatorname{sh} c'.$$

Cuando $c' = 0$, l' es el eje real y L' tiene la ecuación $w = \cos t$ ($-\infty < t < +\infty$); por consiguiente, w describe infinitas veces el segmento $-1 \leq u \leq 1$ del eje real, correspondiéndole a cada segmento de la recta l' de longitud 2π un doble recorrido del segmento indicado. Supongamos que $c' \neq 0$; entonces escribimos la ecuación de la curva L' en la forma:

$$u = \cos t \operatorname{ch} c', \quad v = -\operatorname{sen} t \operatorname{sh} c', \quad (-\infty < t < \infty) \quad (3.7:9)$$

y, eliminando el parámetro t ($\operatorname{ch} c' \neq 0$, $\operatorname{sh} c' \neq 0$), obtenemos:

$$\frac{u^2}{\operatorname{ch}^2 c'} + \frac{v^2}{\operatorname{sh}^2 c'} = 1. \quad (3.7:10)$$

Esta es la ecuación de una elipse con los semiejes $|\operatorname{ch} c'|$ y $|\operatorname{sh} c'|$ y con los focos en los puntos ± 1 . De la representación paramétrica (3.7:9) de la curva L' se deduce que el punto w recorre infinitas veces la elipse en una misma dirección, correspondiendo cada recorrido a un desplazamiento del punto z a lo largo de la recta

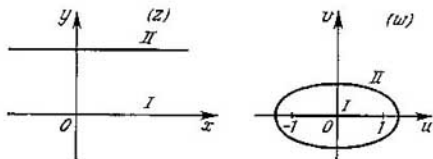


FIG. 20

$z = t + ic'$ a la distancia 2π (fig. 20, donde en la mitad de la derecha están representadas las imágenes de dos rectas del plano z : I ($y = 0$) y II ($y = c \neq 0$)).

Resumiendo, la transformación $w = \cos z$ hace corresponder a la red ortogonal de rectas paralelas a los ejes coordenados, la red de elipses e hipérbolas con los focos comunes ± 1 . Como la transformación es conforme en todos los puntos del plano z , a excepción de los puntos de la forma $z = k\pi$ ($k = 0, \pm 1, \pm 2, \dots$) (las imágenes de los cuales son, precisamente, los focos indicados), la red de elipses e hipérbolas homofocales también tiene que ser ortogonal.

3.8. Tomemos en el plano z una región g que se transforme biunívocamente mediante la función $w = \cos z$ en la región correspondiente del plano w . Esta región se puede elegir de muchos modos. Hay que preocuparse solamente de que a ésta no le pertenezcan dos preimágenes de un mismo punto w . Elijamos por g , por ejemplo, la semifrancia de anchura h ($0 < h \leq 2\pi$), paralela al eje imaginario, con la base en el eje real (fig. 21). Evidentemente, ésta satisface a las condiciones pedidas. En efecto, si para algún punto $z_0 \in g$, $\cos z_0 = w_0$, entonces, como ya se sabe (pág. 121—122), todas las demás preimágenes del punto w_0 en el plano z tienen que estar situadas, en una de sus partes, en la recta paralela al eje real que pasa por el punto z_0 , y en la otra parte, en la recta simétrica a la primera con respecto al eje real. Pero las preimágenes situadas en la primera recta están a unas distancias del punto z_0 que son múltiplos de 2π ; como la anchura de la semifrancia no es superior a 2π , ninguna de

éstas se situará dentro de la semifranja o en su frontera. La segunda recta carece de puntos comunes con la semifranja. En resumen, la función $w = \cos z$ transforma biunívoca y conformemente la región g en cierto conjunto de puntos del plano w .

Para construir este conjunto, hagamos describir al punto z la frontera γ de la región g de modo que éste recorra continua y sucesivamente, primero el lado I de la semifranja, después la base II y, finalmente, el otro lado III de la semifranja. Entonces, el punto

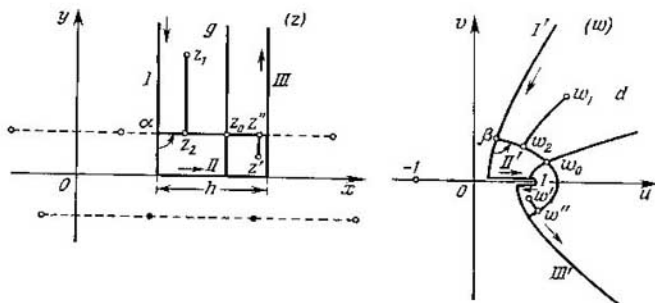


FIG. 21

$w = \cos z$ describirá también continua y sucesivamente la semirrama (I') de una hipérbola, después pasará por la parte (II') de la curva que es la imagen del eje real y se representa por el segmento $-1 \leq u \leq 1, v = 0$ (como la longitud de la base de la semifranja no es superior a 2π , el punto w recorrerá el último segmento no más de dos veces), y, finalmente, pasará una semirrama más (III') de cierta hipérbola.

La imagen completa obtenida en el plano w de la frontera de la región g —designémosla mediante Γ — divide el plano en dos regiones: afirmamos que una de éstas es la imagen buscada d de la región g . Indiquemos dos métodos generales mediante los cuales es posible señalar cuál de las regiones halladas es precisamente la imagen de la región g .

El primer método consiste en que se toma algún punto $z_0 \in g$ y se señala su imagen $w_0 = \cos z_0$. Esta imagen no puede pertenecer al circuito Γ , ya que en caso contrario una de las preimágenes del punto w_0 pertenecería a la región g y la otra, a la frontera γ de esta región, lo cual, como ya se vio, es imposible. Por consiguiente, el punto w_0 se situará en una de las regiones indicadas anteriormente. Esta región será la buscada.

El otro método consiste en que se señala el sentido del recorrido de la frontera γ y de la región g . Esto se puede hacer, por ejemplo, figurándose un observador que se desplaza a lo largo de la frontera de la región g junto con el punto z y que anota hacia qué lado del mismo se encuentra el interior de la región. Durante el recorrido admitido en nuestra figura, la región g quedará, evidentemente, a la izquierda del observador. Obligüemos ahora al observador a desplazarse por Γ junto con el punto $w = \cos z$. Entonces verá la imagen de la región g del mismo lado, es decir, en nuestro ejemplo, a su izquierda.

Expongamos la demostración de todas estas afirmaciones. Supongamos que para el punto $z_0 \in g$ su imagen w_0 pertenece a la región d . Demostremos que, entonces, para cualquier otro punto $z_1 \in g$ su imagen w_1 también pertenece a la misma región d . Tracemos por el punto z_1 una recta paralela al eje imaginario hasta la intersección en el punto z_2 con la recta que pasa por el punto z_0 y es paralela al eje real. Al moverse desde el punto z_0 por el segmento de la última recta hacia z_2 , el punto correspondiente $w = \cos z$ se moverá por el arco de elipse con los focos ± 1 , que pasa por w_0 hacia el punto $w_2 = \cos z_2$. En el camino no se encontrará con ningún punto de la frontera Γ de la región d . En caso contrario, habría un punto w que sería imagen de uno de los puntos situados en γ y también, imagen de algún punto de la región g (del segmento $z_0 z_2$), lo cual, como es sabido, es imposible. Así, pues, todo el arco de elipse $w_0 w_2$ pertenece a la región d . También pertenece a ésta el punto w_2 . Supongamos ahora que z se mueve por el segmento de la recta desde el punto z_2 hasta el punto z_1 . El punto correspondiente $w = \cos z$ se moverá por el arco de una hipérbola con los focos ± 1 , desde el punto w_2 hasta el punto $w_1 = \cos z_1$, y como de nuevo no es posible encontrarse con ningún punto de Γ , todo este arco de hipérbola, inclusive su extremo w_1 , pertenecerá a la región d . Así, pues, la imagen de cualquier punto $z_1 \in g$ pertenece a la misma región d a la que pertenece también la imagen del punto $z_0 \in g$. Por consiguiente, toda la imagen de la región g está contenida en d . No queda más que demostrar que dicha imagen coincide con d , para lo cual hay que verificar que cada punto $w' \in d$ es la imagen de cierto punto $z' \in g$. Tracemos por w' un arco de hipérbola con los focos ± 1 hasta la intersección en el punto w'' con el arco de la elipse que pasa por el punto w_0 . De este modo, obtenemos un arco de elipse $w_0 w''$ y un arco de hipérbola $w'' w'$ pertenecientes por completo a d . Hagamos describir al punto z el segmento de la recta paralela al eje real que pasa por el punto z_0 , desde el punto de intersección con el lado I hasta el punto de intersección con el lado III. El punto correspondiente describirá un arco de elipse, comprendido en la región d , desde el punto de intersección con la semirrama de la hipérbola I'

hasta el punto de intersección con la semirrama III' , y, por consiguiente, pasará por el punto w'' . De aquí se deduce que el segmento indicado de recta contiene la preimagen z'' del punto w'' . Describamos, finalmente, la semirrecta paralela al eje imaginario que pasa por el punto z' . Su imagen será la semirrama de la hipérbola con los focos ± 1 , que pasa por el punto w' . Pero esta semirrama pasa por el punto w' . Por consiguiente, la semirrecta indicada contiene a la preimagen z' del punto w' .

Queda demostrado que w' pertenece a la imagen g , de donde, en virtud de la arbitrariedad del punto w' , se deduce la coincidencia de la imagen de la región g con la región d . Ahora ya es fácil argumentar el método de elección de la región que es la imagen de la región g , basándose en la correspondencia entre los recorridos de los circuitos γ y Γ . Supongamos, por ejemplo, que recorremos el lado I de la semifranja g en la dirección señalada por la flecha, de modo que la región g se mantiene a nuestra izquierda. En algún punto $\alpha \in I$ tracemos hacia dentro de la región un segmento de la normal a la frontera. En este caso, el segmento irá por una recta paralela al eje real. La imagen de este segmento tiene que pertenecer a la imagen de la región g . En el caso considerado ésta representa un arco de la elipse que pasa por un punto $\beta = \cos \alpha$, situado en la semirrama de la hipérbola Γ , y, por consiguiente, por imagen de la región g se debe elegir aquella región (ya se ha verificado que la imagen de la región g es una de las regiones limitadas por el circuito Γ) hacia la cual está dirigida este arco elíptico. Recordemos ahora que en la transformación conforme no sólo se conservan los valores de los ángulos sino también los sentidos de sus direcciones. También tiene que conservarse el sentido de la dirección del ángulo comprendido entre la parte de la frontera γ que parte del punto α en la dirección del recorrido y la normal interior. Pero la indicación de que durante el recorrido del circuito γ la región se mantiene a la izquierda del observador es equivalente al hecho de que el observador, situado en el punto α , tiene que girar un ángulo recto de la derecha hacia la izquierda para ver por la normal el interior de la región g . Debido a esto, el observador que recorre Γ y se encuentra en el punto $\beta = \cos \alpha$ también tiene que girar un ángulo recto de la derecha hacia la izquierda para ver en la dirección de la imagen de la normal (del arco de elipse) el interior de la imagen de la región g . Esto significa que la imagen buscada de la región g estará situada a la izquierda del observador, lo cual se afirmaba.

Señalemos en conclusión que, por lo general, la forma de la región d varía junto con la variación de la disposición y anchura de la semifranja g . En la fig. 22 se ha representado el caso que tiene lugar cuando la base de la semifranja pertenece a uno de los inter-

valos de la forma $(k\pi, (k+1)\pi)$. El caso representado en la fig. 21, que se caracteriza por que la base de la franja, al hacer la transfor-

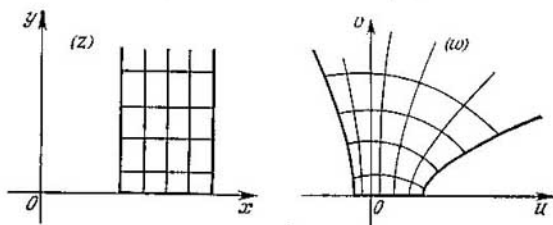


FIG. 22

mación, parece partirse en el punto $w = 1$ o $w = -1$, tiene lugar cuando la base de la semifranja contiene en su interior un punto de la forma $k\pi$.

§ 4. FUNCIONES RACIONALES. FUNCION HOMOGRAFICA. GEOMETRIA DE LOBACHEVSKI. FUNCIONES TRIGONOMETRICAS

4.1. En el párrafo precedente se estudiaron unas cuantas representantes más elementales de la clase de funciones enteras. Siguiendo el orden de aumento de generalidad, después de la clase de funciones enteras viene la clase de funciones meromorfas. Así se llaman las funciones que pueden expresarse en forma de una razón de dos funciones enteras. El vocablo mismo «meromorfa» proviene de las palabras griegas μέρος (parte, fracción) y μορφή (forma) y significa «semejante a una fracción». Está claro que cada función entera $f(z)$ es a la vez meromorfa, puesto que puede expresarse en la forma $\frac{f(z)}{1}$. Naturalmente, lo recíproco no es justo, como muestra el ejemplo de la función $\frac{1}{z}$. Esta función es meromorfa pero no es entera, ya que toma el valor ∞ en el origen de coordenadas.

Las representantes más simples de la clase de las funciones meromorfas en el sentido propio de esta palabra (es decir, que por lo general no se reducen a las enteras), son las funciones racionales. Así se llama toda función que puede expresarse en forma de una razón de dos polinomios:

$$f(z) = \frac{P(z)}{Q(z)} = \frac{a_0 + a_1z + \dots + a_nz^n}{b_0 + b_1z + \dots + b_mz^m} \quad (4.1:1)$$

(el denominador no es idénticamente igual a cero). Supondremos que la fracción $\frac{P(z)}{Q(z)}$ es irreducible, es decir, que las ecuaciones $P(z) = 0$ y $Q(z) = 0$ carecen de raíces comunes. Supongamos, además, que $a_n \neq 0$ y $b_m \neq 0$, es decir, que $P(z)$ y $Q(z)$ son de grados (¡exactos!) n y m , respectivamente. Designemos por $\alpha_1, \dots, \alpha_p$ todas las raíces distintas entre sí de la ecuación $P(z) = 0$, y por k_1, k_2, \dots, k_p , sus órdenes de multiplicidad; del mismo modo, sean β_1, \dots, β_q todas las raíces distintas entre sí de la ecuación $Q(z) = 0$, y l_1, \dots, l_q , sus órdenes de multiplicidad. Entonces (4.1:1) puede escribirse así:

$$f(z) = \frac{P(z)}{Q(z)} = \frac{a_n(z-\alpha_1)^{k_1} \dots (z-\alpha_p)^{k_p}}{b_m(z-\beta_1)^{l_1} \dots (z-\beta_q)^{l_q}}. \quad (4.1:2)$$

Evidentemente, cualquiera de los números $\alpha_1, \dots, \alpha_p$ es distinto de cualquiera de los números β_1, \dots, β_q ; en caso contrario los polinomios $P(z)$ y $Q(z)$ no serían primos entre sí. Además, $k_1 + \dots + k_p = n$ y $l_1 + \dots + l_q = m$. En cada uno de los puntos $z = \alpha_s$ la función $f(z)$ se anula y en cada uno de los puntos $z = \beta_t$ toma el valor ∞ . Los puntos α_s se llaman **ceros** y los puntos β_t , **polos** de la función racional $f(z)$. Los números k_s y l_t que les corresponden, se llaman **órdenes de multiplicidad de los ceros o de los polos**. Si el orden de multiplicidad k_s (o l_t) es igual a uno, el cero α_s (o el polo β_t) se llama **simple**, si el orden de multiplicidad es mayor que uno, se llama **múltiple**.

De esta definición se deduce que los ceros de la función $f(z)$ son polos de la función $\frac{1}{f(z)}$, y los polos de $f(z)$ son ceros de $\frac{1}{f(z)}$ conservándose sus órdenes respectivos de multiplicidad al pasar de $f(z)$ a $\frac{1}{f(z)}$ (es decir, que el cero de un orden determinado se convierte en polo del mismo orden, y viceversa). La función racional está definida en los puntos β_t por las condiciones $f(\beta_t) = \infty$ ($t = 1, 2, \dots, q$); definémosla ahora en el punto del infinito haciendo $f(\infty) = \lim_{z \rightarrow \infty} f(z)$. Obtenemos, evidentemente,

- 1) $f(\infty) = 0$ si $n < m$,
- 2) $f(\infty) = \frac{a_n}{b_n}$ si $n = m$,
- 3) $f(\infty) = \infty$ si $n > m$.

En el caso 1) se dirá que $f(z)$ tiene un cero en el punto del infinito, y en el caso 3), que tiene un polo en el punto del infinito. Para atribuir a este punto un orden de multiplicidad determinado,

hagamos la transformación previa $z = \frac{1}{\zeta}$, que hace corresponder a $z = \infty$ el punto $\zeta = 0$. Entonces resulta:

$$\begin{aligned} f(z) = f\left(\frac{1}{\zeta}\right) &= \frac{a_0 + a_1 \frac{1}{\zeta} + \dots + a_n \frac{1}{\zeta^n}}{b_0 + b_1 \frac{1}{\zeta} + \dots + b_m \frac{1}{\zeta^m}} = \\ &= \frac{\zeta^m}{\zeta^n} \cdot \frac{a_n + a_{n-1}\zeta + \dots + a_0\zeta^n}{b_m + b_{m-1}\zeta + \dots + b_0\zeta^m} = \varphi(\zeta). \end{aligned}$$

Se distinguen los siguientes casos:

1) $n < m$; entonces

$$\varphi(\zeta) = \frac{\zeta^{m-n} (a_n + a_{n-1}\zeta + \dots + a_0\zeta^n)}{b_m + b_{m-1}\zeta + \dots + b_0\zeta^m}.$$

Esta función racional posee un cero de orden $m - n$ en el punto $\zeta = 0$. De acuerdo a esto, se dirá que $f(z)$ posee un cero del mismo orden $m - n$ en el punto $z = \infty$.

2) $n = m$; entonces

$$\varphi(\zeta) = \frac{a_n + a_{n-1}\zeta + \dots + a_0\zeta^n}{b_n + b_{n-1}\zeta + \dots + b_0\zeta^n}.$$

Esta función racional no tiene ni cero ni polo en el punto $\zeta = 0$ (se hace aquí igual a $\frac{a_n}{b_n}$). De acuerdo a esto, $f(z)$ tampoco tiene cero ni polo en el punto $z = \infty$ (se hace aquí igual a $\frac{a_n}{b_n}$).

3) $n > m$; entonces

$$\varphi(\zeta) = \frac{a_n + a_{n-1}\zeta + \dots + a_0\zeta^n}{\zeta^{n-m} (b_m + b_{m-1}\zeta + \dots + b_0\zeta^m)}.$$

Esta función racional tiene un polo de orden $n - m$ en el punto $\zeta = 0$. Respectivamente, se dirá que $f(z)$ tiene un polo del mismo orden $n - m$ en el punto $z = \infty$.

Determinemos el número total de ceros o de polos de la función racional en el plano ampliado, contando cada uno de ellos de acuerdo a su orden de multiplicidad. Ante todo, obtenemos $k_1 + \dots + k_p = n$ ceros finitos. Si $n \geq m$ no hay cero en el infinito. Si $n < m$, obtenemos otro cero en el punto del infinito, de orden $m - n$, y el número total de ceros se hace igual a $n + (m - n) = m$. Si en todos los casos designamos con N el mayor de los números m y n :

$$N = \max(m, n),$$

y lo llamamos orden de la función racional $f(z)$, obtenemos que el número total de ceros de la función $f(z)$ en el plano ampliado es igual al orden de esta función. El número total de polos resulta exactamente igual. Precisando, si $n \leq m$, la función $f(z)$ no tiene polo en el punto del infinito y todos sus polos son finitos. Pero el número de estos últimos es $l_1 + \dots + l_q = m = N$. Si $n > m$, además de los m polos finitos, la función $f(z)$ posee también un polo de orden $n - m$ en el punto del infinito. Por consiguiente, su número total de polos es $m + (n - m) = n = N$.

Sea ahora A un número complejo arbitrario, distinto de 0 y de ∞ . Averiguemos todas las raíces de la ecuación

$$f(z) = \frac{P(z)}{Q(z)} = A \quad (4.1:3)$$

o, lo que es lo mismo, las raíces de la ecuación

$$F(z) = \frac{P(z) - AQ(z)}{Q(z)} = 0. \quad (4.1:4)$$

Si se atribuye a cada raíz de la ecuación (4.1:3) el mismo orden de multiplicidad que tiene esta misma raíz en la ecuación (4.1:4), se puede afirmar, según lo anterior, que el número total de raíces de la ecuación (4.1:3), igual al número de ceros de la función $F(z) = \frac{P(z) - AQ(z)}{Q(z)}$, tiene que coincidir con el orden de esta última.

Pero, si $n > m$, el grado del polinomio $P(z) - AQ(z)$ es igual a n y, por consiguiente, el orden de la función $F(z)$ también es igual a n , es decir, coincide con el orden de la función $f(z)$. Cuando $n = m$, el grado del polinomio $P(z) - AQ(z)$ no es superior a n , y como el grado del polinomio $Q(z)$ es igual a n , el orden de la función $F(z)$ es igual a n , o sea, de nuevo coincide con el orden de la función $f(z)$. Finalmente, si $n < m$, el grado del polinomio $P(z) - AQ(z)$ es igual a m ($A \neq 0$), y como el grado del polinomio $Q(z)$ también es igual a m , el orden de la función $F(z)$ es igual a m . Por consiguiente, en este caso, el orden de la función $F(z)$ también coincide con el orden de la función $f(z)$. Resumiendo, en todos los casos, los órdenes de las funciones $f(z)$ y $F(z)$ son iguales, de donde se deduce que el número total de raíces de la ecuación (4.1:3) en el plano ampliado es igual para todos los valores A y coincide con el orden N de la función $f(z)$.

Queda revelado, pues, que la conocida propiedad de los polinomios de tomar cualquier valor A en un mismo número de puntos del plano, igual al grado del polinomio, se extiende a cualquier función racional sustituyendo el grado por el orden.

Como ejemplo, examinemos la función $f(z) = \frac{z^2+1}{z^2-1}$. Esta tiene dos ceros simples: $\pm i$ y dos polos simples: ± 1 . Para el valor $A = 1$ la ecuación

$$\frac{z^2+1}{z^2-1} = 1 \text{ o sea } \frac{2}{z^2-1} = 0$$

no tiene raíces. Pero tiene una raíz doble en el ∞ , puesto que el grado del denominador de la fracción es dos unidades superior al grado del numerador.

De los resultados obtenidos se deduce que una función racional $w = f(z)$ de orden N transforma el plano ampliado en el plano ampliado de tal modo, que cada punto $w = f(z)$ posee generalmente N preimágenes en el plano z . Para algunos valores w el número de preimágenes distintas entre sí puede resultar menor que N . Así, por ejemplo, si $f(z)$ tiene ceros múltiples o polos múltiples, a los puntos $w = 0$ o $w = \infty$ les corresponderán menos de N distintas preimágenes en el plano z .

Los demás valores w ($w \neq 0$, $w \neq \infty$) de este género se obtienen buscando las condiciones según las cuales la ecuación

$$\frac{P(z)}{Q(z)} = w \text{ o sea } \frac{P(z) - wQ(z)}{Q(z)} = 0$$

tiene raíces múltiples. Todas las raíces múltiples finitas de la última ecuación coinciden con las raíces múltiples de la ecuación

$$P(z) - wQ(z) = 0,$$

y, por consiguiente, satisfacen también a la ecuación

$$P'(z) - wQ'(z) = 0$$

y, finalmente, a la ecuación de grado no superior a $m+n-1$,

$$P(z)Q'(z) - P'(z)Q(z) = 0. \quad (4.1:5)$$

Esta última posee no más de $m+n-1$ raíces distintas $\gamma_1, \gamma_2, \dots, \gamma_r$, a las cuales corresponden no más de $m+n-1$ distintos valores

$$w_j = \frac{P(\gamma_j)}{Q(\gamma_j)} \quad (j = 1, 2, \dots, r),$$

que poseen un número menor que N de preimágenes.

Obsérvese que todos los ceros múltiples finitos de $f(z)$ satisfacen a las ecuaciones

$$P(z) = 0 \text{ y } P'(z) = 0,$$

y los polos múltiples finitos, a las ecuaciones

$$Q(z) = 0 \text{ y } Q'(z) = 0.$$

Por consiguiente, tanto unos como otros satisfacen también a la ecuación (4.1:5), o sea, se encuentran entre sus raíces γ_j . Debido a esto, los números 0 y ∞ también se encontrarán entre los r números w_j (cuando existan ceros múltiples finitos o polos múltiples finitos). Sin embargo, puede ocurrir que la ecuación $f(z) = f(\infty)$ posea raíces múltiples solamente en el punto del infinito; entonces el número $f(\infty)$ (posiblemente, igual a cero o a infinito) no se encontrará entre los números w_j ($j = 1, \dots, r$) y, por consiguiente, habrá que añadirle a éstos. El número de puntos obtenidos no será superior a $m + n$.

Resumiendo, en todos los casos, el número total de aquellos puntos del plano w que poseen, cada uno de ellos, menos de N preimágenes, no es superior a $m + n$. El conjunto de todas las preimágenes de estos puntos consta de los puntos $\gamma_1, \dots, \gamma_r$; a veces, a estos puntos hay que agregar también el punto ∞ .

Si z es distinto de los polos de la función $f(z)$ y del punto del infinito, la función $f(z)$ posee derivada

$$f'(z) = \frac{P'(z)Q(z) - P(z)Q'(z)}{[Q(z)]^2};$$

para que ésta no sea igual a cero, hay que exigir también que el punto z no coincida con ninguno de los puntos γ_j ($j = 1, \dots, r$). De aquí se deduce que la transformación $w = f(z)$ es conforme en todos los sitios, a excepción, posiblemente, de un número finito de puntos.

Dejamos a cuenta del lector la demostración de que la transformación es conforme también en cualquier polo simple de la función $f(z)$, y también en el punto del infinito, suponiendo que $z = \infty$ no es una raíz múltiple de la ecuación $f(z) = f(\infty)$. Del mismo modo se puede demostrar que la transformación no es conforme en cada uno de los puntos γ_j ($j = 1, 2, \dots, r$), a los cuales hay que añadir también $\gamma_0 = \infty$, si $z = \infty$ es una raíz múltiple de la ecuación $f(z) = f(\infty)$. Precisamente, el ángulo con el vértice en el punto γ_j aumenta en la transformación $w = f(z)$ en un número de veces, igual al orden de multiplicidad de la raíz γ_j en la ecuación $f(z) = f(\gamma_j)$ ($j = 0, 1, 2, \dots, r$).

4.2. De los resultados del apartado anterior se deduce que la función racional de primer orden, es decir, la función homográfica $w = L(z) = \frac{az+b}{cz+d}$, es la única de las funciones racionales que realiza una transformación biunívoca del plano ampliado sobre sí mismo. (En este caso hay que exigir además que el determinante de la función $L(z)$ sea distinto de cero: $ad - bc \neq 0$; cuando no se cumple esta condición, $L(z)$ se hace idénticamente igual a una constante, o sea, transforma el plano ampliado en un punto). Ya se vio en el

ap. 2.3 que la transformación $w = L(z)$ es conforme en todos los puntos del plano ampliado. En las cuestiones más diversas de la teoría de las funciones de variable compleja se suele recurrir a esta transformación (a sus distintos casos particulares). Debido a esto, merece un estudio particular. Consideremos aquí sus propiedades principales.

Estudiaremos el conjunto M de todas las transformaciones homográficas con el determinante diferente de cero. Dos transformaciones

$$L_1(z) = \frac{a_1z + b_1}{c_1z + d_1} \quad \text{y} \quad L_2(z) = \frac{a_2z + b_2}{c_2z + d_2}$$

se consideran iguales cuando, y sólo cuando, $L_1(z) = L_2(z)$ para todos los valores de z . Para esto es suficiente que los coeficientes correspondientes sean proporcionales entre sí:

$$a_2 = \lambda a_1, \quad b_2 = \lambda b_1, \quad c_2 = \lambda c_1 \quad \text{y} \quad d_2 = \lambda d_1 \quad (\lambda \neq 0).$$

Estas mismas condiciones son necesarias. En efecto, si $L_1(z) = L_2(z)$, entonces, en particular,

$$L_1(0) = L_2(0), \quad L_1(1) = L_2(1) \quad \text{y} \quad L_1(\infty) = L_2(\infty);$$

lo cual significa que

$$\frac{b_1}{d_1} = \frac{b_2}{d_2} = p, \quad \frac{a_1 + b_1}{c_1 + d_1} = \frac{a_2 + b_2}{c_2 + d_2}, \quad \frac{a_1}{c_1} = \frac{a_2}{c_2} = q.$$

Poniendo en la igualdad del medio

$$b_1 = d_1 p, \quad b_2 = d_2 p, \quad a_1 = c_1 q \quad \text{y} \quad a_2 = c_2 q$$

obtenemos

$$\frac{c_1 q + d_1 p}{c_1 + d_1} = \frac{c_2 q + d_2 p}{c_2 + d_2}, \quad \text{o sea, } (c_1 d_2 - c_2 d_1)(q - p) = 0.$$

Pero $q \neq p$ (en caso contrario sería $\frac{a_1}{c_1} = \frac{b_1}{d_1}$, es decir, $a_1 d_1 - b_1 c_1 = 0$ lo cual contradice a la hipótesis). Por consiguiente,

$$\frac{c_1}{d_1} = \frac{c_2}{d_2}.$$

Las relaciones obtenidas pueden escribirse en la forma

$$\frac{a_2}{a_1} = \frac{b_2}{b_1} = \frac{c_2}{c_1} = \frac{d_2}{d_1} = \lambda,$$

que es lo que se quería demostrar.

De lo expuesto se deduce que el valor del determinante de la transformación homográfica no es, de por sí mismo, característico para esta transformación. En efecto, al pasar de los coeficientes

a_1 , b_1 , c_1 y d_1 a los coeficientes λa_1 , λb_1 , λc_1 y λd_1 ($\lambda \neq 0$) el determinante se multiplica por λ^2 . Pero, en todo caso, este determinante, siendo distinto de cero para algunos valores de los coeficientes, se mantiene siempre distinto de cero.

La transformación

$$U(z) = z,$$

pertenciente, evidentemente, al conjunto M , se llamará transformación idéntica o unidad.

Llamaremos inversa con respecto a una transformación

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

a la transformación según la cual para cada z se toma por imagen su preimagen z_1 en la transformación dada. La transformación inversa es

$$z = \frac{dz_1 - b}{-cz_1 + a}.$$

La transformación inversa con respecto a L se designará mediante L^{-1} .

Si

$$z_1 = L(z) = \frac{az+b}{cz+d} \text{ y } z_2 = L_1(z_1) = \frac{a_1z_1+b_1}{c_1z_1+d_1}$$

son dos transformaciones homográficas arbitrarias (como siempre, con los determinantes distintos de cero), la transformación que se obtiene al realizarlas consecutivamente una tras otra en cierto orden, se llama p r o d u c t o de las transformaciones dadas. Supongamos que se efectúa primero la transformación $z_1 = L(z)$, y después la transformación $z_2 = L_1(z_1)$. Entonces su producto se designa mediante $L_1 L(z)$. Para éste se tiene:

$$\begin{aligned} z_2 = L_1 L(z) &= \left[a_1 \frac{az+b}{cz+d} + b_1 \right] : \left[c_1 \frac{az+b}{cz+d} + d_1 \right] = \\ &= [(aa_1 + cb_1)z + (ba_1 + db_1)] : [(ac_1 + cd_1)z + (bc_1 + dd_1)]. \end{aligned}$$

Por consiguiente, $z_2 = L_1 L(z)$ también es una transformación homográfica; para su determinante obtenemos:

$$\begin{aligned} (aa_1 + cb_1)(bc_1 + dd_1) - (ba_1 + db_1)(ac_1 + cd_1) &= \\ = (ad - bc)(a_1d_1 - b_1c_1) &\neq 0. \end{aligned}$$

Así, pues, la transformación $L_1 L(z)$ pertenece al conjunto M .

Evidentemente,

$$LL^{-1}(z) = L^{-1}L(z) = U(z).$$

Obsérvese que la transformación $z_2 = LL_1(z)$, obtenida al realizar primero la transformación $z_1 = L_1(z)$ y después la transformación $z_2 = L(z_1)$, se diferencia, generalmente, de la transformación $L_1L(z)$. Así, por ejemplo, si

$$L(z) = \frac{z}{z+1} \quad \text{y} \quad L_1(z) = \frac{z+1}{z-1},$$

se tiene

$$L_1L(z) = -2z-1, \quad \text{mientras que} \quad LL_1(z) = \frac{z+1}{2z}.$$

La operación definida de multiplicación de transformaciones es asociativa, es decir, para cualesquiera transformaciones homográficas L , L_1 y L_2 , se cumple la igualdad

$$(LL_1)L_2 = L(L_1L_2).$$

Esta propiedad se comprueba fácilmente. En efecto, supongamos que $L_2(z) = z_2$. Entonces

$$(LL_1)L_2(z) = LL_1[L_2(z)] = LL_1(z_2)$$

y

$$L(L_1L_2)(z) = L[L_1L_2(z)] = LL_1(z_2),$$

así, pues,

$$(LL_1)L_2(z) = L(L_1L_2)(z).$$

La propiedad asociativa se extiende al producto de cualquier número de transformaciones. Esta nos libra de la necesidad de poner los paréntesis en este producto. Así, por ejemplo,

$$L[L_1(L_2L_3)](z) = LL_1(L_2L_3)(z) = L(L_1L_2)L_3(z) = \dots = LL_1L_2L_3(z).$$

Como el conjunto M , junto con cada dos transformaciones L y L_1 , contiene también su producto L_1L (y LL_1), y junto con la transformación L contiene también la inversa L^{-1} , éste forma un grupo de transformaciones*).

4.3. Veamos la demostración de la propiedad homocíclica (o circular) de la transformación homográfica, que se expresa en que la imagen de una recta o una circunferencia en la transformación $w = L(z)$, es una recta o una circunferencia; la imagen de una recta puede ser una recta o una circunferencia y, del mismo modo, la imagen de una circunferencia puede ser una recta o una circunferencia.

Para una función lineal entera $L(z) = az + \beta$ esta propiedad es inmediata, puesto que la transformación $w = L(z)$ se reduce

*) Véase, por ejemplo, A. Kurosch, *Curso de álgebra superior*, Editorial Mir, Moscú, 1968, § 63.

a una traslación (cuando $\alpha = 1$) o a una rotación seguida de una homotecia (o dilatación) (si $\alpha \neq 1$); véase el ap. 2.3.

Consideremos ahora la transformación

$$w = \Lambda(z) = \frac{1}{z},$$

que utilizamos ya varias veces.

Evidentemente, la ecuación de cualquier recta o circunferencia puede expresarse en la forma

$$A(x^2 + y^2) + 2Bx + 2Cy + D = 0.$$

Cuando $A = 0$ y B y C no son simultáneamente iguales a cero, resulta una recta, y cuando $A \neq 0$ y $B^2 + C^2 - AD > 0$, resulta una circunferencia. Sustituyendo aquí $x^2 + y^2$ por $z\bar{z}$, $2x$ por $z + \bar{z}$ y $2y$ por $-i(z - \bar{z})$, representamos esta ecuación en la forma

$$A z \bar{z} + (B - iC)z + (B + iC)\bar{z} + D = 0,$$

o bien

$$A z \bar{z} + \bar{E}z + E\bar{z} + D = 0, \quad (4.3:1)$$

donde $E = B + Ci$. Aquí $A = 0$, y el número complejo E es distinto de cero en el caso de una recta y $A \neq 0$ y $E\bar{E} - AD > 0$ en el caso de una circunferencia. Recíprocamente: cualquier ecuación de esta forma con coeficientes reales A y D y coeficientes complejos conjugados E y \bar{E} será la ecuación de una recta si $A = 0$ y el número E es distinto de cero, y la ecuación de una circunferencia, si $A \neq 0$ y $E\bar{E} - AD > 0$. Para cerciorarse de esto, es suficiente pasar de z a x e y según las fórmulas

$$z\bar{z} = x^2 + y^2, \quad z = x + iy, \quad \bar{z} = x - iy.$$

Queriendo obtener la imagen de la línea (4.3:1) en la transformación $w = \frac{1}{z}$, substituyamos z en la ecuación (4.3:1) por $\frac{1}{w}$. Obtenemos:

$$A \frac{1}{w\bar{w}} + \bar{E} \frac{1}{w} + E \frac{1}{\bar{w}} + D = 0,$$

o bien

$$Dw\bar{w} + Ew + \bar{E}\bar{w} + A = 0. \quad (4.3:2)$$

La ecuación (4.3:2) tiene la misma forma que la ecuación (4.3:1) con la substitución de A por D , D por A y E por \bar{E} . De aquí se deduce que si $D = 0$ ésta es la ecuación de una recta (puesto que entonces será $A = 0$ y $E \neq 0$, o bien $A \neq 0$ y $E\bar{E} - AD = E\bar{E} > 0$, es

decir, de nuevo $E \neq 0$), y si $D \neq 0$, la ecuación de una circunferencia (puesto que cuando $A \neq 0$ la ecuación (4.3:1) representaba una circunferencia y, por consiguiente, se cumplía la condición $E\bar{E} - AD > 0$, mientras que cuando $A = 0$ ésta representaba una recta, por consiguiente, E era distinto de cero, de donde $E\bar{E} - AD = E\bar{E} > 0$). Queda demostrado que *la imagen de una recta o de una circunferencia en la transformación $w = \Lambda(z) = \frac{1}{z}$ es una recta o una circunferencia.*

Considerando una función homográfica arbitraria

$$w = L(z) = \frac{az+b}{cz+d} \quad (c \neq 0),$$

representémosla de la forma

$$w = \frac{a}{c} + \frac{bc-ad}{c(cz+d)}.$$

Pongamos

$$z_1 = L_1(z) = cz + d, \quad z_2 = \Lambda(z_1) = \frac{1}{z_1}$$

y

$$w = L_2(z_2) = \frac{a}{c} + \frac{bc-ad}{c} z_2;$$

entonces $L(z)$ se escribirá en forma del producto de tres transformaciones:

$$L = L_2 \Lambda L_1.$$

Como, en cada una de las transformaciones L_1 , Λ y L_2 , la imagen de una recta o circunferencia es una recta o circunferencia, la transformación L posee la misma propiedad. Queda demostrada por completo la propiedad homocíclica de la transformación homográfica.

Designemos mediante $\delta = -\frac{d}{c}$ el polo de la función homográfica $L(z) = \frac{az+b}{cz+d}$ ($c \neq 0$). La imagen de cada recta o circunferencia que pasa por δ tiene que contener al punto del infinito $L(\delta) = \infty$, y, por consiguiente, no puede ser una circunferencia. En virtud de la propiedad homocíclica, esta imagen es una recta. La imagen de una recta o una circunferencia que no pase por el punto δ , no puede contener al punto del infinito y, por consiguiente, no puede ser una recta. En virtud de la propiedad homocíclica, esta imagen es una circunferencia. Resumiendo, *en la transformación $w = L(z)$ todas las rectas y circunferencias que pasan por el polo δ de esta función se transforman en rectas del plano w , mientras que las rectas y circunferencias que no pasan por δ se transforman en circunferencias del plano w .*

Supongamos que $w = L(z)$ es una función homográfica arbitraria; sea γ una recta o circunferencia del plano z , y sea Γ su imagen en el plano w (es decir, también una recta o circunferencia). Consideremos las regiones g_1 y g_2 limitadas por la línea γ en el plano z ; éstas son dos semiplanos o el interior y el exterior de una circunferencia. Demostremos que la imagen de una de éstas es una de las dos regiones limitadas por la línea Γ en el plano w , mientras que la imagen de la otra es la segunda de dichas regiones. En efecto, sea z_1 un punto de la región g_1 y z_2 un punto de la región g_2 . Como z_1

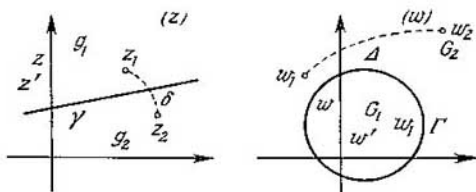


FIG. 23

y z_2 no están situados en γ , las imágenes de estos puntos w_1 y w_2 no pueden estar situados en Γ y, por consiguiente, se sitúan en las regiones en las que Γ divide el plano w . Si éstos estuviesen situados en una misma región, entonces podrían unirse mediante un segmento Δ de recta (o arco de circunferencia) que no tuviese puntos comunes con Γ (fig. 23). La preimagen del segmento Δ en el plano z tiene que ser un segmento de recta (o de arco de circunferencia) δ que una z_1 y z_2 y que no tenga puntos comunes con γ . Pero la existencia de tal segmento contradice a que z_1 y z_2 están situados en distintas regiones g_1 y g_2 . Por lo tanto, si los puntos z_1 y z_2 pertenecen a distintas regiones g_1 y g_2 , sus imágenes w_1 y w_2 también pertenecen a distintas regiones limitadas por la línea Γ . Designemos mediante G_1 la región que contiene a w_1 y mediante G_2 , la que contiene a w_2 , y demostremos que la imagen de la región g_1 es G_1 , y la imagen de la región g_2 es G_2 .

En efecto, si, por ejemplo, z' es algún punto de la región g_1 , entonces el hecho de pertenecer z' y z_2 a distintas regiones g_1 y g_2 implica, según lo demostrado, que sus imágenes w' y w_2 pertenecen a distintas regiones G_1 y G_2 . Pero w_2 pertenece a G_2 , por lo cual w' pertenece a G_1 . Así, pues, la imagen de cada punto de la región g_1 pertenece a G_1 ; del mismo modo, la imagen de cada punto de la región g_2 pertenece a G_2 . Tomemos, finalmente, un punto arbitrario w de la región G_1 . Este tiene que ser la imagen de uno de los puntos z de la región g_1 o g_2 . Pero z no puede ser un punto de la región

g_2 , puesto que en caso contrario w sería un punto de la región G_2 . Por consiguiente, w es la imagen del punto z , perteneciente a g_1 . Resumiendo, G_1 es la imagen de la región g_1 , y G_2 , la imagen de la región g_2 . Por consiguiente, queda demostrado que dos regiones limitadas por la línea γ se transforman en dos regiones limitadas por la línea Γ , además, queda verificado que para saber cuál precisamente de las dos regiones, limitadas por la línea Γ , es la imagen de la región dada g_1 , limitada por la línea γ , es suficiente observar la imagen w_1 de un solo punto $z_1 \in g_1$; la región G_1 a la cual pertenezca w_1 será la imagen de la región g_1 .

4.4. Si la transformación homográfica $w = L(z)$ es distinta de la idéntica, entonces, por lo general, w es distinto de z . No obstante, aquí también existen puntos inmóviles de la transformación, caracterizados por la ecuación

$$z = L(z) = \frac{az + b}{cz + d}.$$

Supongamos primero que $c = 0$ ($d \neq 0$). Entonces $L(z)$ representa una función lineal entera

$$l(z) = \alpha z + \beta \quad \left(\alpha = \frac{a}{d}, \beta = \frac{b}{d} \right).$$

Como $l(\infty) = \infty$, uno de los puntos inmóviles de la transformación lineal entera es el punto del infinito del plano, $z = \infty$. Si $\alpha \neq 1$, existe también otro punto inmóvil, determinado por la ecuación

$$z = \alpha z + \beta.$$

Este es el punto $z = \frac{\beta}{1-\alpha}$. Si $\alpha = 1$ y $\beta \neq 0$, no existe ningún punto finito inmóvil. Pero si $\alpha \neq 1$, $\beta \neq 0$ y $\alpha \rightarrow 1$, entonces el punto finito inmóvil $\frac{\beta}{1-\alpha}$ tiende hacia el punto del infinito. Debido a esto, en el caso de la transformación $l(z) = z + \beta$ ($\beta \neq 0$), el punto del infinito se puede considerar como dos puntos inmóviles confundidos.

Supongamos ahora que $c \neq 0$. Entonces $L(\infty) = \frac{a}{c} \neq \infty$, o sea, el punto $z = \infty$ no es inmóvil. Del mismo modo, no es inmóvil tampoco el punto $-\frac{d}{c}$, pues

$$L\left(-\frac{d}{c}\right) = \infty \neq -\frac{d}{c}.$$

Resolvamos la ecuación

$$z = \frac{az + b}{cz + d},$$

suponiendo que $z \neq \infty$ y $z \neq -\frac{d}{c}$, o sea, la ecuación

$$cz^2 - (a-d)z - b = 0.$$

Ohtenemos:

$$z = \frac{a-d \pm \sqrt{(a-d)^2 + 4bc}}{2c}.$$

Si $(a-d)^2 + 4bc \neq 0$, de aquí se hallan dos puntos inmóviles finitos distintos. Cuando $(a-d)^2 + 4bc = 0$, estos dos puntos se confunden en un solo punto inmóvil finito $\frac{a-d}{2c}$.

En resumen, *la transformación homográfica distinta de la idéntica posee solamente dos puntos inmóviles, que en el caso particular pueden confundirse en uno.* [Las transformaciones lineales enteras se caracterizan por completo en que al menos uno de los puntos inmóviles es el punto del infinito.]

La transformación homográfica que posee más de dos puntos inmóviles puede ser solamente la transformación idéntica $U(z) = z$ (para la cual todos los puntos son inmóviles). De aquí se puede deducir que para la coincidencia de dos transformaciones homográficas $L(z)$ y $\Lambda(z)$ es suficiente que se cumpla la igualdad $L(z) = \Lambda(z)$ para tres puntos distintos z_1, z_2 y z_3 . En efecto, sea $L(z_k) = \Lambda(z_k) = w_k$ ($k = 1, 2, 3$); entonces $\Lambda^{-1}(w_k) = z_k$ ($k = 1, 2, 3$) y, por consiguiente, la transformación $\Lambda^{-1}L(z)$ hace corresponder a los puntos z_k de nuevo los mismos puntos, o sea, $\Lambda^{-1}L(z_k) = z_k$ ($k = 1, 2, 3$) (puesto que $L(z_k) = w_k$ y $\Lambda^{-1}(w_k) = z_k$). Es decir, la transformación $\Lambda^{-1}L$ tiene tres puntos distintos inmóviles: z_1, z_2 y z_3 , o sea, es la transformación idéntica:

$$\Lambda^{-1}L = U.$$

Por consiguiente,

$$\Lambda(\Lambda^{-1}L) = \Lambda U.$$

Pero

$$\Lambda(\Lambda^{-1}L) = (\Lambda\Lambda^{-1})L = UL = L,$$

y

$$\Lambda U = \Lambda.$$

Definitivamente, obtenemos:

$$L = \Lambda.$$

como se quería demostrar.

En resumen, *para determinar una transformación homográfica es suficiente señalar tres puntos distintos: w_1, w_2 y w_3 , correspondientes a tres puntos distintos dados: z_1, z_2 y z_3 .*

Propongámonos hallar la función homográfica que realiza esta transformación, y supongamos primero que z_1, z_2 y z_3 son puntos finitos y que $w_1 = 0, w_2 = \infty$ y $w_3 = 1$.

Para que la función homográfica

$$w = \Lambda(z) = \frac{az + b}{cz + d}$$

tome el valor 0 para $z = z_1$ y el valor ∞ para $z = z_2$, es necesario y suficiente que $z = z_1$ sea un cero del numerador $az + b$, o sea, que el numerador tenga la forma $a(z - z_1)$, y que $z = z_2$ sea un cero del denominador, o sea, que el denominador tenga la forma $c(z - z_2)$. Por lo tanto, la función buscada tiene que poseer la forma

$$w = \frac{a}{c} \frac{z - z_1}{z - z_2}.$$

Pero, cuando $z = z_3$, w tiene que ser igual a 1. De la ecuación

$$1 = \frac{a}{c} \frac{z_3 - z_1}{z_3 - z_2}$$

obtenemos:

$$\frac{a}{c} = 1 : \frac{z_3 - z_1}{z_3 - z_2},$$

de donde

$$w = \Lambda(z) = \frac{z - z_1}{z - z_2} \cdot \frac{z_3 - z_2}{z_3 - z_1}.$$

Esta es la función homográfica buscada, que hace corresponder a los puntos z_1, z_2 y z_3 los puntos 0, ∞ y 1, respectivamente.

Supongamos ahora que w_1, w_2 y w_3 son unos puntos finitos arbitrarios (distintos) y que $w = L(z)$ es la transformación homográfica que satisface a las condiciones:

$$L(z_1) = w_1, \quad L(z_2) = w_2 \quad \text{y} \quad L(z_3) = w_3.$$

Según lo demostrado, la función $\xi = \Lambda_1(w) = \frac{w - w_1}{w - w_2} : \frac{w_3 - w_1}{w_3 - w_2}$ hace corresponder a los puntos w_1, w_2 y w_3 los puntos 0, ∞ y 1. Debido a esto, la función $\Lambda_1 L(z)$ hace corresponder a los puntos z_1, z_2 y z_3 los puntos 0, ∞ y 1, es decir,

$$\Lambda_1 L(z) = \Lambda(z) = \frac{z - z_1}{z - z_2} : \frac{z_3 - z_1}{z_3 - z_2}.$$

De la relación

$$\Lambda_1 L = \Lambda$$

se deduce que

$$\Lambda_1^{-1}(\Lambda_1 L) = \Lambda_1^{-1} \Lambda,$$

o sea,

$$L = \Lambda_1^{-1} \Lambda$$

(puesto que $\Lambda_1^{-1}(\Lambda_1 L) = (\Lambda_1^{-1} \Lambda_1) L = UL = L$).

Esta última relación resuelve el problema, ya que las transformaciones Λ y Λ_1 son conocidas:

$$\Lambda(\zeta) = \frac{\zeta - z_1}{\zeta - z_2} : \frac{z_3 - z_1}{z_3 - z_2}, \quad \Lambda_1(\zeta) = \frac{\zeta - w_1}{\zeta - w_2} : \frac{w_3 - w_1}{w_3 - w_2}.$$

Por cierto, para el estudio de la transformación $w = L(z)$ es mejor utilizar directamente la relación:

$$\Lambda_1 L(z) = \Lambda(z),$$

de donde, después de sustituir $L(z)$ por w , se deduce que:

$$\Lambda_1(w) = \Lambda(z),$$

o bien,

$$\frac{w - w_1}{w - w_2} : \frac{w_3 - w_1}{w_3 - w_2} = \frac{z - z_1}{z - z_2} : \frac{z_3 - z_1}{z_3 - z_2}. \quad (4.4:1)$$

Esta ecuación determina la función homográfica $w = L(z)$ en forma implícita.

Hemos resuelto el problema planteado, bajo la hipótesis de que todos los puntos z_1, z_2, z_3, w_1, w_2 y w_3 sean finitos. Si, por ejemplo, $z_1 = \infty$, la función $\Lambda(z)$, que a los puntos $z_1 = \infty, z_2$ y z_3 hace corresponder los puntos $w_1 = 0, w_2 = \infty$ y $w_3 = 1$, adquiere la forma

$$\Lambda(z) = \frac{1}{z - z_2} : \frac{1}{z_3 - z_2} *).$$

Por lo tanto, la ecuación (4.4:1) se sustituye por la ecuación

$$\frac{w - w_1}{w - w_2} : \frac{w_3 - w_1}{w_3 - w_2} = \frac{1}{z - z_2} : \frac{1}{z_3 - z_2} \quad (4.4:1')$$

(suponiendo que son finitos los puntos w_1, w_2 y w_3). Si $z_2 = \infty$, la función $\Lambda(z)$ que a los puntos $z_1, z_2 = \infty$ y z_3 hace corres-

*) Se puede llegar a esta expresión así: suponiendo z_1 finito, se escribe la relación

$$\frac{z - z_1}{z - z_2} : \frac{z_3 - z_1}{z_3 - z_2}$$

en la forma

$$\frac{z - 1}{z - z_2} : \frac{z_3 - 1}{z_3 - z_2}$$

y se pasa al límite para $z_1 \rightarrow \infty$.

ponder los puntos $w_1=0$, $w_2=\infty$ y $w_3=1$, toma la forma

$$\Lambda(z) = (z - z_1) : (z_3 - z_1)$$

y, por consiguiente, la ecuación (4.4:1) se sustituye por la ecuación

$$\frac{w-w_1}{w-w_2} : \frac{w_3-w_1}{w_3-w_2} = (z-z_1) : (z_3-z_1). \quad (4.4:1'')$$

Si $z_3=\infty$, la función $\Lambda(z)$ que a los puntos z_1 , z_2 y $z_3=\infty$ hace corresponder los puntos $w_1=0$, $w_2=\infty$ y $w_3=1$, toma la forma

$$\Lambda(z) = \frac{z-z_1}{z-z_2}$$

y la ecuación (4.4:1) se sustituye por la ecuación

$$\frac{w-w_1}{w-w_2} : \frac{w_3-w_1}{w_3-w_2} = \frac{z-z_1}{z-z_2}. \quad (4.4:1''')$$

Del mismo modo, el primer miembro de la ecuación (4.4:1) se debe sustituir por

$$\frac{1}{w-w_2} : \frac{1}{w_3-w_2}, \quad (w-w_1) : (w_3-w_1), \quad \frac{w-w_1}{w-w_2}$$

según que sea $w_1 = \infty$, $w_2 = \infty$ o $w_3 = \infty$.

En resumen, llegamos a la siguiente regla mnemotécnica: si $z_k = \infty$ o $w_l = \infty$ ($k = 1, 2, 3$; $l = 1, 2, 3$), entonces en la ecuación (4.4:1) las diferencias en las que figuren z_k o w_l deben sustituirse por 1. El lector fácilmente confirmará la justeza de esta regla mediante el paso al límite en la ecuación (4.4:1) (para $z_k \rightarrow \infty$ o $w_l \rightarrow \infty$).

De la ecuación (4.4:1) se deduce una importante propiedad general de las transformaciones homográficas. Sean a , b , c y d unos números complejos distintos (finitos) arbitrarios. Llamemos a

$$\frac{c-a}{c-b} : \frac{d-a}{d-b}$$

razón doble o anarmónica de cuatro números o de los puntos a , b , c y d . Esta razón la designaremos con la notación (a, b, c, d) :

$$(a, b, c, d) = \frac{c-a}{c-b} : \frac{d-a}{d-b}.$$

La definición de razón doble la extendemos también al caso en que uno de los cuatro puntos a , b , c , d sea el del infinito. Preciando, llamaremos razón doble de cuatro puntos, entre los cuales uno es el del infinito, al límite de la razón doble de cuatro puntos finitos, de los cuales tres coinciden con los puntos dados y el cuarto

tiende hacia el punto del infinito. Siguiendo esta definición, tendremos:

$$\begin{aligned}(\infty, b, c, d) &= \frac{1}{c-b} : \frac{1}{d-b} \\(a, \infty, c, d) &= (c-a) : (d-a), \\(a, b, \infty, d) &= 1 : \frac{d-a}{d-b}, \\(a, b, c, \infty) &= \frac{c-a}{c-b}.\end{aligned}$$

Supongamos ahora que $w = L(z)$ es una transformación homográfica arbitraria. Designemos mediante A, B, C y D los puntos homólogos a los cuatro puntos distintos: a, b, c , y d . Como los tres puntos a, b y d se transforman en los puntos A, B y D , los puntos $w = L(z)$ y z estarán ligados por la relación (4.4:1):

$$\frac{w-A}{w-B} : \frac{D-A}{D-B} = \frac{z-a}{z-b} : \frac{d-a}{d-b},$$

en la cual se deben sustituir por 1 aquellas diferencias en las que figura el punto del infinito. Haciendo $z = c$, tenemos que poner $w = C$ (puesto que en nuestra transformación al punto c corresponde el punto C). Por consiguiente,

$$\frac{C-A}{C-B} : \frac{D-A}{D-B} = \frac{c-a}{c-b} : \frac{d-a}{d-b}$$

(las diferencias en las que figura el punto del infinito se deben sustituir por 1), o

$$(A, B, C, D) = (a, b, c, d). \quad (4.4:2)$$

Resumiendo, en la transformación homográfica la razón doble de cuatro puntos cualesquiera no varía; en otras palabras, la razón doble es un invariante de la transformación homográfica.

4.5. Basándose en la propiedad homocíclica de la transformación homográfica y en la posibilidad de transformar cualquier terna de puntos z_1, z_2, z_3 en otra terna previamente dada w_1, w_2, w_3 , demostremos la siguiente proposición:

Cualesquiera que sean las rectas o circunferencias γ y Γ y dos ternas de puntos z_1, z_2, z_3 y w_1, w_2, w_3 , pertenecientes a γ y Γ , respectivamente, existe una transformación homográfica $w = L(z)$ que transforma γ en Γ y hace corresponder a los puntos z_1, z_2, z_3 los puntos w_1, w_2, w_3 , respectivamente.

En efecto, escribamos la transformación homográfica $w = L(z)$ que satisfice a las condiciones $L(z_j) = w_j$ ($j = 1, 2, 3$). Según lo expuesto anteriormente, tal función existe y es la única que satisfice a estas condiciones. Esta transforma la recta o circunferencia γ

en una recta o circunferencia Γ' . Pero γ pasa por los puntos z_1, z_2 y z_3 ; por lo tanto, Γ' pasa por los puntos w_1, w_2 y w_3 , y como por tres puntos no se pueden trazar dos rectas o circunferencias distintas, Γ' coincide con Γ . Así, pues, $w = L(z)$ satisface a todas las condiciones de la proposición enunciada.

Tomemos de nuevo unas rectas o circunferencias arbitrarias γ y Γ (distintas o coincidentes) y sea g una de las dos regiones limitadas por la línea γ , y G , una de las dos regiones limitadas por la línea Γ . Evidentemente, tanto una como la otra pueden ser un semiplano, el interior de una circunferencia o el exterior de una circunferencia. Elijamos una terna arbitraria de puntos z_1, z_2 y z_3 situados

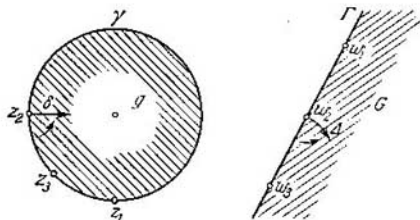


FIG. 24

en γ y supongamos para fijar ideas que al moverse el observador a lo largo de γ , en la dirección del punto z_1 al punto z_3 , pasando por z_2 , la región g se mantiene a su izquierda. Sea w_1, w_2, w_3 una terna de puntos situados en Γ tal que, al moverse el observador a lo largo de Γ en la dirección del punto w_1 al punto w_3 , pasando por el punto w_2 , la región G se mantenga a su izquierda. En lo demás, los puntos w_1, w_2, w_3 son arbitrarios. Formemos, como se mostró anteriormente, una función homográfica $w = L(z)$ que satisfaga a las condiciones $w_j = L(z_j)$ ($j = 1, 2, 3$), la cual, por consiguiente, transformará γ en Γ . Demostremos que esta función transforma también la región g en la región G . En efecto, sea δ el segmento de la normal a la línea γ , trazada por el punto z_2 hacia el interior de la región g , es decir, a la izquierda del observador situado en el punto z_2 , que mira a lo largo de γ en la dirección establecida anteriormente; entonces, como la transformación $w = L(z)$ es conforme, la imagen Δ de este segmento (que será un segmento de recta o de arco de circunferencia) también estará dirigida hacia la izquierda del observador situado en el punto w_2 que mira a lo largo de Γ en la dirección establecida en Γ (fig. 24). Por consiguiente, según la condición, Δ pertenece a G . De este modo, queda ya establecido que la región G contiene las imágenes de ciertos puntos pertenecientes a g (precisamente,

las imágenes de los puntos del segmento δ). Pero, según el ap. 4.3, $L(g)$ es una de las regiones cuya frontera coincide con la imagen de la frontera de la región g , es decir, con $\Gamma = L(\gamma)$. Como solamente hay dos regiones de éstas y una de ellas, G , contiene las imágenes de los puntos de la región g , ésa es la imagen buscada de esta región g : $G = L(g)$.

Aclaremos lo expuesto con un ejemplo. Supongamos que hay que transformar conformemente el semiplano superior $y > 0$ en el interior del círculo unidad.

Para resolver el problema hagamos, por ejemplo, $z_1 = -1$, $z_2 = 0$ y $z_3 = 1$, de modo que el semiplano quede a la izquierda del observador que vaya por el eje real en la dirección de z_1 a z_3 pasando por z_2 , y elegimos también en la circunferencia unidad tres puntos: w_1, w_2 y w_3 , de modo que el interior del círculo quede a la izquierda del observador que vaya por la circunferencia en la dirección de w_1 a w_3 pasando por w_2 . Para mayor sencillez, se pueden tomar: $w_1 = 1$, $w_2 = i$ y $w_3 = -1$. Entonces, la transformación homográfica que cumple las condiciones $w_j = L(z_j)$ ($j = 1, 2, 3$), será la buscada. Esta puede expresarse en la forma

$$\frac{w-1}{w-i} : \frac{-1-1}{-1-i} = \frac{z+1}{z} : \frac{1+1}{1},$$

o bien

$$w = \frac{z-i}{iz-1}.$$

4.6. Sean z_1 y z_2 dos puntos simétricos con respecto a cierta recta γ . Entonces, el centro de una circunferencia arbitraria δ que pase por z_1 y z_2 estará situado en γ y, por consiguiente, δ será ortogonal a γ . Será también ortogonal a γ la recta que pasa por los puntos z_1 y z_2 . Fácilmente se observa que también es justo lo recíproco: si cualquier circunferencia o recta que pase por un par de puntos z_1 y z_2 , es ortogonal a la recta γ , entonces z_1 y z_2 son simétricos respecto de γ .

Hagamos una transformación del plano mediante una función homográfica $w = L(z)$, de modo que la recta γ se transforme en una recta o circunferencia Γ . Entonces, el par de puntos z_1 y z_2 , simétricos respecto de γ , se transformará en cierto par de puntos w_1 y w_2 , y cada circunferencia o recta δ que pase por z_1 y z_2 se transformará en una circunferencia o recta Δ que pasa por w_1 y w_2 , y viceversa: cualquier recta o circunferencia que pase por w_1 y w_2 será la imagen de cierta recta o circunferencia que pasa por z_1 y z_2 . En virtud de la simetría de los puntos z_1 y z_2 respecto de γ y debido a que la transformación $w = L(z)$ es conforme, las rectas y circunferencias que pasan por w_1 y w_2 serán todas ortogonales a Γ . Por consiguiente, si Γ es una recta, los puntos $w_1 = L(z_1)$ y $w_2 = L(z_2)$ serán simétricos

respecto de Γ . Generalizando el concepto de simetría, diremos que dos puntos son simétricos respecto de una circunferencia, si cualquier recta o circunferencia que pase por ellos es ortogonal a la circunferencia dada. Entonces se podrá decir que si la recta γ se transforma en una circunferencia Γ mediante la función homográfica $w = L(z)$, entonces cualquier par de puntos simétricos respecto de la recta se transforma en un par de puntos simétricos respecto de la circunferencia, y recíprocamente. De aquí se deduce que, dada la circunferencia Γ y el punto w_1 , el punto simétrico a w_1 respecto de Γ se determina unívocamente. En efecto, si existiesen dos puntos distintos w_2 y w'_2 que fuesen simétricos a w_1 respecto de Γ , entonces en la transformación homográfica de Γ en la recta γ los puntos w_1 , w_2 y w'_2 se transformarían en los puntos z_1 , z_2 y $z'_2 \neq z_2$, donde los puntos z_1 y z_2 , y también z_1 y z'_2 , serían simétricos respecto de γ , lo cual, evidentemente, es imposible.

Los razonamientos expuestos aquí sobre la transformación de una recta γ en una recta o circunferencia Γ se extienden sin cambios algunos al caso en que una circunferencia se transforma en una circunferencia, obteniendo que en la transformación homográfica cualquier par de puntos simétricos respecto de una circunferencia γ se transforma en un par de puntos simétricos respecto de la circunferencia Γ que es la imagen de γ . Así, pues, hemos obtenido la siguiente propiedad general de conservación de la simetría en las transformaciones homográficas:

¶ Si los puntos z_1 y z_2 son simétricos respecto de una recta o circunferencia γ , entonces en cualquier transformación homográfica $w = L(z)$ sus imágenes w_1 y w_2 serán simétricas respecto de la imagen $\Gamma = L(\gamma)$.

Señalemos un caso particular de esta proposición. Supongamos que γ se transforma en una circunferencia Γ , y que z_1 es el punto que se transforma en el centro w_1 de la circunferencia Γ . Entonces el punto z_2 , simétrico a z_1 , tiene que transformarse en un punto w_2 del plano ampliado, simétrico a w_1 respecto de Γ . Pero tal punto es el del infinito. En efecto, la recta que une w_1 y $w_2 = \infty$, es decir, cualquier recta que pase por el centro de la circunferencia Γ será ortogonal a Γ . En virtud de la unicidad del punto simétrico (al punto dado respecto de la circunferencia dada), el punto ∞ , y sólo éste, será simétrico al centro de la circunferencia Γ respecto de Γ .

Sea γ una recta o circunferencia arbitraria. La transformación del plano ampliado que consiste en que cada punto z se transforma en un punto z^* , simétrico a z respecto de γ , se llama transformación de simetría respecto de γ . Cuando γ es una circunferencia, esta transformación se llama también inversión respecto de γ .

Hallemos las expresiones analíticas para la transformación de simetría. Supongamos primero que γ es una recta. Esta se deter-

mina completamente por uno de sus puntos a y el vector unitario $e^{i\theta} = \cos \theta + i \sin \theta$, que lleva su dirección.

Realicemos la transformación lineal

$$z = a + e^{i\theta}w = l(w),$$

que, evidentemente, transforma el eje real en la recta considerada (es una traslación que lleva el origen de coordenadas al punto a , seguida de una rotación en el ángulo θ alrededor del último punto). Como la transformación $w = l^{-1}(z)$ transforma γ en el eje real, la misma transforma a cada par de puntos z y z^* , simétricos respecto de γ , en un par de puntos w y w^* , simétricos respecto del eje real. Estos últimos se expresan por números complejos conjugados $w = t$ y $w^* = \bar{t}$. Debido a esto, $z - a = e^{i\theta}t$, o bien $\overline{z - a} = e^{i\theta}\bar{t}$ y $z^* - a = e^{i\theta}w^* = e^{i\theta}\bar{t}$. Eliminando t entre las dos últimas igualdades, obtenemos:

$$z^* - a = e^{2i\theta} \overline{(z - a)}. \quad (4.6:1)$$

Esta ecuación muestra que para realizar la transformación de simetría respecto de la recta γ que pasa por el punto a formando con el eje real el ángulo θ , se debe pasar del vector $z - a$ a su simétrico respecto del eje real $\overline{z - a}$ y girar luego este último el ángulo 2θ alrededor del punto a .

Consideremos ahora la transformación de simetría respecto de una circunferencia Γ de radio R ($0 < R < \infty$) con centro en a . Realicemos una transformación homográfica que transforme Γ en el eje real. Lo más sencillo es tomar la transformación

$$z = a + R \frac{1 + iw}{1 - iw} = L(w),$$

la cual a los puntos del eje real $w_1 = -1$, $w_2 = 0$ y $w_3 = 1$ hace corresponder los puntos de la circunferencia Γ : $z_1 = a - iR$, $z_2 = a$ y $z_3 = a + iR$ y, por consiguiente, transforma el eje real en Γ . La transformación inversa $w = L^{-1}(z)$ transforma Γ en el eje real y a cada par de puntos z y z^* , simétricos respecto de Γ , hace corresponder un par de puntos w y w^* , simétricos respecto del eje real. Como w y w^* se expresan por números complejos conjugados:

$w = t$ y $w^* = \bar{t}$, se tiene, $z - a = R \frac{1 + it}{1 - it}$ o bien, $\overline{z - a} = R \frac{1 - i\bar{t}}{1 + i\bar{t}}$

y $z^* - a = R \frac{1 + i\bar{t}}{1 - i\bar{t}}$. Multiplicando término a término estas dos últimas relaciones, obtenemos:

$$\overline{(z - a)} (z^* - a) = R^2$$

o sea

$$z^* - a = \frac{R^2}{z - a}. \quad (4.6:2)$$

De aquí se deduce: primero, que

$$\text{Arg}(z^* - a) = -\text{Arg}(\overline{z - a}) = \text{Arg}(z - a)$$

y, segundo, que

$$|z^* - a| |z - a| = R^2.$$

Por consiguiente, los puntos z^* y z están situados en un mismo rayo que parte del centro de la circunferencia y se encuentran a tales distancias del centro que el producto de estas distancias es igual al cuadrado del radio. Por estas dos condiciones, o lo que es lo mismo, por la fórmula (4.6:2), se determina completamente la posición de uno de los puntos z , z^* siendo dado el otro, es decir, se determina la transformación de inversión respecto de la circunferencia $|z - a| = R$.

De las igualdades (4.6:1) o (4.6:2) se deduce que *la transformación general de simetría se reduce a la realización consecutiva de una transformación lineal (entera o fraccionaria) y una transformación de simetría respecto del eje real.*

Así, por ejemplo, la transformación de simetría respecto de una recta puede expresarse en la forma:

$$z_1 = \bar{a} + e^{-2i\theta}(z - a) \quad \text{y} \quad z^* = \bar{z}_1, \quad (4.6:1')$$

y la transformación de simetría respecto de una circunferencia, en la forma

$$z_1 = \bar{a} + \frac{R^2}{z - a} \quad \text{y} \quad z^* = \bar{z}_1. \quad (4.6:2')$$

Como la transformación lineal es conforme y posee la propiedad homocélica, y la transformación de simetría respecto del eje real posee las mismas propiedades con la única diferencia de que, conservando las magnitudes de los ángulos, se cambian sus sentidos por los opuestos, la transformación de simetría en el caso más general posee también las propiedades indicadas. Precisamente, ésta es una transformación conforme de segundo género y transforma las rectas y circunferencias en rectas o circunferencias.

4.7. Ilustremos con dos ejemplos la aplicación de la propiedad de conservación de la simetría en las transformaciones homográficas.

Ejemplo 1. Transformar conformemente el semiplano superior en el interior del círculo $|w| < R$ de modo que a un punto dado del semiplano le corresponda el centro del círculo: $w = 0$.

La función buscada, si existe, se anula para $z = \alpha$: $L(\alpha) = 0$. Así, pues, ya conocemos el cero $z = \alpha$ de la función $L(z)$. Pero el

punto $\bar{\alpha}$, simétrico a α respecto del eje real, tiene que transformarse en un punto simétrico al centro de la circunferencia respecto de la circunferencia misma, es decir, en el punto del infinito. Por consiguiente, ya conocemos también el polo $z = \bar{\alpha}$ de la transformación homográfica $L(z)$. Por lo tanto, $L(z)$ tiene la forma

$$w = L(z) = \frac{a(z-\alpha)}{c(z-\bar{\alpha})} = \lambda \frac{z-\alpha}{z-\bar{\alpha}}, \quad (4.7:1)$$

donde λ es un número complejo distinto de cero.

Demostremos que la función hallada transforma el semiplano en el círculo $|w| < |\lambda|$ de modo que el punto α se transforma en el centro del círculo $w = 0$. Evidentemente, esta última condición se cumple para la función (4.7:1) con cualquier λ . No queda más que comprobar que el eje real se transforma en la circunferencia de radio $|\lambda|$ con el centro en el origen de coordenadas. En efecto, si $z = x$ es un número real arbitrario, los números $x - \alpha$ y $x - \bar{\alpha}$ son complejos conjugados y, por consiguiente,

$$|w| = |L(x)| = \left| \lambda \frac{x-\alpha}{x-\bar{\alpha}} \right| = |\lambda| \left| \frac{x-\alpha}{x-\bar{\alpha}} \right| = |\lambda|.$$

Hemos obtenido que las imágenes de todos los puntos del eje real están situados en la circunferencia $|w| = |\lambda|$, de donde, en virtud de la propiedad homocíclica, se deduce que la imagen del eje real es esta circunferencia.

Para obtener la transformación del semiplano en un círculo de radio R se debe tomar $|\lambda| = R$. Queda indeterminado todavía el argumento del número λ . El significado geométrico de esta indeterminación es completamente claro. El paso en la fórmula (4.7:1) de un valor λ a otro manteniendo invariable el módulo $|\lambda| = R$ equivale a la variación de los argumentos de todos los puntos en una misma magnitud, es decir, a la rotación del círculo alrededor de su centro $w = 0$. En tal rotación, el círculo se transforma en sí mismo, su centro se mantiene en su sitio, y no se infringen las condiciones del problema.

Si se quiere que el problema planteado posea solución única, es necesario introducir una condición complementaria. Se puede exigir, por ejemplo, que:

a) un punto dado del eje real $x = x_0$ se transforme en el punto $w = R$ de la circunferencia, o

b) la derivada $L'(\alpha)$ sea un número real positivo (geométricamente esto significa que las tangentes a las curvas que pasan por el punto α no tienen que cambiar la inclinación al hacer la transformación).

En efecto, en la condición a), obtenemos de (4.7:1):

$$R = L(x_0) = \lambda \frac{x_0 - \bar{\alpha}}{x_0 - \alpha},$$

de donde

$$\lambda = R \frac{x_0 - \bar{\alpha}}{x_0 - \alpha}$$

y

$$L(z) = R \frac{x_0 - \bar{\alpha}}{x_0 - \alpha} \cdot \frac{z - \alpha}{z - \bar{\alpha}}. \quad (4.7:2)$$

Evidentemente,

$$|\lambda| = R \left| \frac{x_0 - \bar{\alpha}}{x_0 - \alpha} \right| = R.$$

Supongamos ahora que $\alpha = \xi + i\eta$, donde $\eta > 0$. Como

$$L'(\alpha) = \frac{\lambda}{\alpha - \alpha},$$

con la condición b) sacamos la conclusión que $\frac{\lambda}{\alpha - \alpha} = \frac{\lambda}{2i\eta}$, lo cual significa que $\frac{\lambda}{i}$ es un número real positivo. Pero, por otra parte, el módulo $|\lambda|$ tiene que ser igual a R . Por lo tanto, $\lambda = iR$ y

$$L(z) = iR \frac{z - \alpha}{z - \bar{\alpha}}. \quad (4.7:3)$$

Ejemplo 2. Transformar conformemente el círculo $|z| < R$ en sí mismo, de modo que el punto dado $z = \alpha$ de este círculo se transforme en su centro $w = 0$.

La función homográfica $L(z)$ buscada, si existe, se anula para $z = \alpha$: $L(\alpha) = 0$. Así, pues, ya conocemos un cero $z = \alpha$ de la función $L(z)$. Pero el punto α^* , simétrico a α respecto de la circunferencia $|z| = R$, tiene que transformarse en el punto simétrico al centro respecto de la misma circunferencia. Por esto, ya conocemos también el polo $z = \alpha^*$ de la función $L(z)$. Por consiguiente, la función homográfica $L(z)$ tiene que tener la forma

$$w = L(z) = \lambda \frac{z - \alpha}{z - \alpha^*},$$

donde λ es un número complejo distinto de cero. Según la fórmula (4.6:2), el punto α^* , simétrico al punto α respecto de la circunferencia $|z| = R$, es:

$$\alpha^* = \frac{R^2}{\alpha}.$$

Por lo tanto,

$$w = L(z) = -\lambda \bar{\alpha} \frac{z - \alpha}{R^2 - \bar{\alpha}z} = \mu \frac{z - \alpha}{R^2 - \bar{\alpha}z}. \quad (4.7:4)$$

Demostremos que la función buscada transforma el círculo $|z| < R$ en el círculo $|w| < \frac{|\mu|}{R}$, de modo que el punto α se transforma en el centro del círculo 0. Evidentemente, esto último se cumple para la función $w = L(z)$ con cualquier μ . No queda más que demostrar que la circunferencia $|z| = R$ se transforma mediante (4.7:4) en la circunferencia $|w| = \frac{|\mu|}{R}$. Pero, en efecto, sea

$$z = Re^{i\theta} \quad (0 < \theta < 2\pi)$$

un punto arbitrario de la circunferencia $|z| = R$. Entonces, para su imagen

$$w = L(Re^{i\theta}) = \mu \frac{Re^{i\theta} - \alpha}{R^2 - \bar{\alpha}Re^{i\theta}} = \frac{\mu}{Re^{i\theta}} \frac{Re^{i\theta} - \alpha}{Re^{-i\theta} - \bar{\alpha}}$$

se tiene:

$$|w| = \left| \frac{\mu}{Re^{-i\theta}} \right| \left| \frac{Re^{i\theta} - \alpha}{Re^{-i\theta} - \bar{\alpha}} \right| = \frac{|\mu|}{R}$$

(puesto que $|e^{i\theta}| = 1$ y $\left| \frac{Re^{i\theta} - \alpha}{Re^{-i\theta} - \bar{\alpha}} \right| = 1$), de donde se deduce que la imagen de la circunferencia $|z| = R$ es la circunferencia $|w| = \frac{|\mu|}{R}$.

Para obtener la transformación del círculo de radio R en sí mismo es debido, evidentemente, tomar en la fórmula (4.7:4) $|\mu| = R^2$. El argumento del número μ continúa manteniéndose indeterminado. Para que el problema de la transformación tenga solución única se puede imponer una de las siguientes condiciones complementarias a):

a) un punto dado a de la circunferencia $|z| = R$ se transforma en el punto $w = R$ de la misma circunferencia;

b) la derivada $L'(\alpha)$ es un número real positivo.

Dejamos a cuenta del lector la comprobación de que, en el caso b):

$$L(z) = R \frac{R^2 - \bar{\alpha}a}{a - \alpha} \frac{z - \alpha}{R^2 - \bar{\alpha}z}, \quad (4.7:5)$$

y en el caso a):

$$L(z) = \frac{R^2(z - \alpha)}{R^2 - \bar{\alpha}z}. \quad (4.7:6)$$

4.8. Arriba se demostró que toda función de la forma

$$w = \Lambda(z) = \mu \frac{z - \alpha}{R^2 - \bar{\alpha}z} \quad (|\alpha| < R, \quad |\mu| = R^2) \quad (4.8:1)$$

transforma el círculo $K: |z| < R$ en sí mismo. En particular, junto con cada función $\Lambda(z) = \mu \frac{z - \alpha}{R^2 - \bar{\alpha}z}$ posee también la misma propiedad la función $\bar{\mu} \frac{z - \bar{\alpha}}{R^2 - \alpha z}$, que convendremos en designar mediante $\bar{\Lambda}(z)$. Evidentemente

$$\overline{\overline{\Lambda(z)}} = \Lambda(z).$$

Fácilmente se observa que, eligiendo en la fórmula (4.8:1) α y μ de todas las maneras posibles, conservando las condiciones $|\alpha| < R$, $|\mu| = R^2$, obtenemos todas las transformaciones homográficas posibles del círculo K en sí mismo. En efecto, si $w = L(z)$ es una de ellas, entonces ésta transforma algún punto del círculo $z = \alpha$ en su centro y, por consiguiente, según el ap. 4.7 tiene que tener la forma (4.8:1).

Si Λ_1 y Λ_2 son dos transformaciones del círculo K en sí mismo, su producto $\Lambda = \Lambda_1\Lambda_2$ posee la misma propiedad; del mismo modo, la transformación inversa a una cualquiera de las transformaciones (4.8:1) transforma el círculo K en sí mismo, o sea, es una transformación de la forma (4.8:1). De aquí se deduce que todas las transformaciones posibles de la forma (4.8:1) (estando fijado $R > 0$) forman un grupo. Designémoslo mediante Γ ; éste es un subgrupo del grupo de todas las transformaciones homográficas. El grupo Γ admite una interpretación geométrica admirable; precisamente se puede considerar como el grupo de los movimientos del plano en la geometría de Lobachevski.

Al interior del círculo K lo llamaremos plano de Lobachevski, a los puntos del círculo K , puntos de Lobachevski y a los arcos de circunferencias o a los segmentos de rectas, pertenecientes a K y ortogonales a la circunferencia $\Gamma: |z| = R$, rectas de Lobachevski, o bien, abreviadamente: \mathcal{L} -plano, \mathcal{L} -puntos, \mathcal{L} -rectas. Por cierto, a continuación, en lugar de \mathcal{L} -punto diremos siempre simplemente: punto. Estas denominaciones quedan justificadas por el hecho de que entre el interior del círculo, sus puntos y los arcos de circunferencias o segmentos de rectas indicados en la geometría de Euclides se verifican las mismas relaciones que entre el plano, sus puntos y rectas en la geometría de Lobachevski. Ante todo, en el \mathcal{L} -plano se verifican las siguientes proposiciones (axiomas de Hilbert *):

Axiomas de unión:

- I₁. Para cada dos puntos A y B siempre existe una \mathcal{L} -recta que pasa por A y B .
- I₂. Para dos puntos A y B existe no más de una \mathcal{L} -recta que pase por A y B .
- I₃. En cada \mathcal{L} -recta existen al menos dos puntos. Existen al menos tres puntos no situados en una \mathcal{L} -recta.

Axiomas de orden:

- II₁. Si A , B y C son puntos de una \mathcal{L} -recta y B está situado entre A y C , entonces B también está situado entre C y A .
- II₂. Si A y C son puntos de una \mathcal{L} -recta, existe al menos un punto B tal que C está situado entre A y B .

*) Véase D. Hilbert, Fundamentos de la geometría.

113. Entre tres puntos cualesquiera de una \mathcal{L} -recta existe no más de un punto situado entre los otros dos.

114. Sean A , B y C tres puntos no situados en una \mathcal{L} -recta y a , una \mathcal{L} -recta que no pase por ninguno de los puntos A , B , C ; si ésta pasa por un punto del segmento AB , entonces pasa indispensablemente por un punto del segmento AC o por un punto del segmento BC .

Para demostrar que se verifican los axiomas I_1 y I_2 *), supondremos primero que uno de los puntos dados, por ejemplo A , coincide con el centro del círculo K . Entonces el diámetro del círculo que pasa por el punto B será la \mathcal{L} -recta que satisface a la condición I_1 . Supongamos, en contra del axioma I_2 , que existe otra \mathcal{L} -recta más que pasa por los puntos A y B . Esta tendrá que ser un arco de circunferencia ortogonal a Γ y, por consiguiente, los radios de Γ trazados a los extremos de este arco tendrán que ser tangentes a éste. Pero esto es imposible, puesto que cada uno de los radios posee dos puntos comunes distintos con el arco (el punto A y uno de los extremos del arco). Supongamos ahora que ninguno de los puntos A y B coincide con el centro del círculo. Entonces, efectuando una transformación de la forma (4.8:1), donde α es el afijo del punto A , reducimos la cuestión al caso que acabamos de considerar y, después, mediante la transformación inversa, obtenemos la única \mathcal{L} -recta que pasa por los puntos dados (la imagen del diámetro del círculo K en la última transformación). Dejamos a cuenta del lector la comprobación de la justeza de los demás axiomas.

Llamaremos movimiento de Lobachevski (abreviadamente, \mathcal{L} -movimiento) a cualquier transformación homográfica de la forma (4.8:1), es decir, a la transformación homográfica del círculo K en sí mismo. Cada \mathcal{L} -movimiento transforma biunívocamente el \mathcal{L} -plano en sí mismo, de modo que los puntos se transforman en puntos y las \mathcal{L} -rectas en \mathcal{L} -rectas (aquí nos basamos en la propiedad homocíclica y en que la transformación homográfica es conforme).

Consideremos un conjunto E de puntos del \mathcal{L} -plano y el conjunto \bar{E} de los puntos simétricos a éste respecto del eje real. Si F es otro conjunto de puntos del \mathcal{L} -plano, diremos que E es congruente al conjunto F en el sentido de Lobachevski (abreviadamente: \mathcal{L} -congruente) cuando, y sólo cuando, existe un \mathcal{L} -movimiento L que transforma F en E o en \bar{E} , de modo que $L(F) = E$ o $L(F) = \bar{E}$.

De la definición de \mathcal{L} -congruencia se deduce que cada conjunto E es \mathcal{L} -congruente a sí mismo, y también al conjunto \bar{E} simétrico a él con respecto al eje real. Además, de la definición se deduce que si E es \mathcal{L} -congruente a F , entonces F es \mathcal{L} -congruente a E . En efecto, si $L(F) = E$, entonces $F = L^{-1}(E)$, y si $L(F) = \bar{E}$, entonces también $F = L^{-1}(\bar{E})$ y $\bar{F} = L^{-1}(\bar{\bar{E}}) = L^{-1}(E)$. Supongamos, finalmente, que E es \mathcal{L} -congruente a F y F es \mathcal{L} -congruente a G ; demostremos que entonces E es \mathcal{L} -congruente a G . En efecto, supongamos primero que $F = L(G)$ y $E = L(F)$ o $\bar{E} = L(F)$. En el primer caso, tendremos $E = L^2(G)$, y en el segundo caso, $\bar{E} = L(G)$; en uno y otro caso, según la definición, obtenemos que E es \mathcal{L} -congruente a G .

Supongamos ahora que $\bar{F} = L(G)$; entonces, si $E = L(F)$, se tiene $\bar{E} = \bar{L}(F) = \bar{L} \wedge (G)$, y si $\bar{E} = L(F)$, resulta $E = \bar{L}(\bar{F}) = \bar{L} \wedge (G)$. De nuevo obtenemos que E es \mathcal{L} -congruente a G .

*) A nosotros no nos debe desconcertar el hecho de que demostremos los axiomas. Para nosotros éstos son solamente unos teoremas de la geometría de Euclides, referentes al interior del círculo, a sus puntos y a los segmentos de rectas y arcos de circunferencias situados en este círculo; demostrándolos, tenemos el derecho de aplicar todas las proposiciones y conceptos conocidos de la geometría de Euclides.

En resumen, el concepto introducido aquí de \mathcal{L} -congruencia posee todas las propiedades de equivalencia, pues ésta es reflexiva (cada conjunto es \mathcal{L} -congruente a sí mismo), simétrica (si E es \mathcal{L} -congruente a F , entonces F es \mathcal{L} -congruente a E) y transitiva (si E es \mathcal{L} -congruente a F y F es \mathcal{L} -congruente a G , entonces E es \mathcal{L} -congruente a G). Designemos la relación de congruencia entre E y F con el signo \equiv . Entonces tendremos: 1) $E \equiv E$; 2) si $E \equiv F$, entonces $F \equiv E$; 3) si $E \equiv F$ y $F \equiv G$, entonces $E \equiv G$.

Ahora se pueden formular y demostrar las siguientes proposiciones:

Axiomas de congruencia:

III₁. Si A y B son dos puntos de una \mathcal{L} -recta a y A' es un punto de la misma o de otra \mathcal{L} -recta a' , entonces en la recta a' , hacia un lado dado del

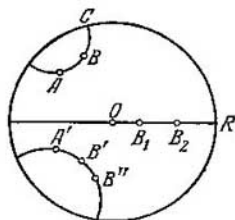


FIG. 25

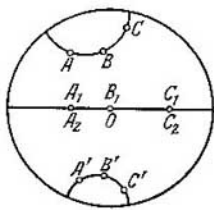


FIG. 26

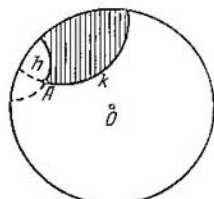


FIG. 27

punto A' , siempre se puede hallar un punto B' , y sólo uno, tal que el segmento $A'B'$ sea \mathcal{L} -congruente al segmento AB .

III₂. Si $AB' \equiv A'B'$ y $AB \equiv A''B''$, entonces $A'B' \equiv A''B''$.

III₃. Sean AB y BC dos segmentos de una \mathcal{L} -recta a sin puntos interiores comunes; sean luego $A'B'$ y $B'C'$ dos segmentos de esta misma o de otra \mathcal{L} -recta a' , también sin puntos interiores comunes. Si en estas condiciones, $AB \equiv A'B'$ y $BC \equiv B'C'$, entonces también $AC \equiv A'C'$.

Antes de enunciar los otros dos axiomas de congruencia, demostremos los axiomas que acabamos de formular.

Aludiendo a la demostración del axioma III₁, designemos mediante C uno de los extremos del arco de circunferencia (o del segmento de recta) que representa la \mathcal{L} -recta a . Según el ap. 4.7 se puede hallar una transformación afín del círculo K en sí mismo de modo que lleve el punto A al centro O del círculo K y el punto C al punto R de la circunferencia Γ de este círculo (fig. 25). Como resultado, obtenemos un \mathcal{L} -movimiento L que transforma a en la \mathcal{L} -recta d , representada por el diámetro del círculo situado en el eje real, y a la \mathcal{L} -semirecta AC , en la \mathcal{L} -semirecta OR .

Si AC contenía al segmento AB , la imagen $B_1 = L(B)$ del punto B estará situada en OR . Señalemos en la \mathcal{L} -recta a' la semirecta $A'C'$ que está situada hacia el lado dado de A' . Sea L' el \mathcal{L} -movimiento que transforma a' en d y $A'C'$ en OR , y sea B' el punto de la \mathcal{L} -recta a' para el cual $L'(B') = B_1$. Entonces, evidentemente, el \mathcal{L} -segmento $A'B'$ estará situado en a' hacia el lado dado de A' y el \mathcal{L} -movimiento $L^{-1}L'$ transformará a' en a y el segmento $A'B'$ en AB .

Hay que demostrar también que B' es el único punto que satisface a las condiciones del axioma. Pero, en efecto, suponiendo que existiese otro punto más B'' , situado en la semirecta $A'C'$ y tal que el segmento $A'B''$ fuese congruente a AB , existiría en la semirecta OR un punto $B_2 \neq B_1$ ($B_2 = L'(B'')$) tal que $OB_2 \equiv OB_1$. De aquí, según la definición, se deduciría que existe un

\mathcal{L} -movimiento Λ tal que $OB_2 = \Lambda(OB_1)$. (La consideración del conjunto $\overline{OB_2}$, simétrico a OB_2 respecto del eje real, es aquí superflua, puesto que $\overline{OB_2}$ coincide con OB_2). Esto significaría que $w = \Lambda(z)$ sería una transformación del círculo K en sí mismo tal, que el segmento rectilíneo OB_1 se transformaría en OB_2 y, por consiguiente, el diámetro situado en el eje real se transformaría en sí mismo. *A priori* existen dos posibilidades:

1) que el punto O se transforme en O y B_1 en B_2 ; entonces [según la fórmula (4.8:1), donde $\alpha = 0$] $\Lambda(z)$ tiene que tener la forma

$$w = \Lambda(z) = \frac{\mu}{R^2} z,$$

y como $|\mu| = R^2$, resulta $|w| = |z|$; pero esto contradice al hecho de que B_1 se transforma en B_2 ($|B_1| \neq |B_2|$);

2) que el punto B_1 se transforme en O y O en B_2 ; entonces

$$w = \Lambda(z) = \mu \cdot \frac{z - B_1}{R^2 - B_1 z} \quad (\alpha = B_1), \quad B_2 = \Lambda(0) = -\frac{\mu B_1}{R^2}.$$

de donde $|B_2| = \frac{|\mu|}{R^2} |B_1| = |B_1|$, lo cual de nuevo contradice a la condición $|B_2| \neq |B_1|$.

Así, pues, la negación de la unicidad del punto B' en las condiciones del axioma III₁ conduce a una contradicción. Queda establecido el contenido fundamental del axioma.

El axioma III₂ es consecuencia inmediata de las propiedades simétrica y transitiva de la \mathcal{L} -congruencia.

Veamos el axioma III₃. Transformemos a en d mediante un \mathcal{L} -movimiento L , de modo que el punto B se transforme en el centro del círculo $B_1 = O$ y el punto C en un punto situado en la parte positiva del eje real. Entonces, en virtud de las hipótesis del axioma, el punto A se transformará en un punto situado en la parte negativa del eje real. Supongamos que A_1 y C_1 son los puntos homólogos de A y C (fig. 26). Si consideramos un \mathcal{L} -movimiento análogo, formado a partir de la \mathcal{L} -recta a' , obtendremos que el punto A_2 estará situado en la parte negativa del eje real, $B_2 = O$ y C_2 estará situado en la parte positiva del eje real.

De las relaciones $AB \equiv A'B'$, $AB \equiv A_1B_1$ y $A'B' \equiv A_2B_2$ se deducirá entonces que $A_1B_1 \equiv A_2B_2$, es decir, $A_1O \equiv A_2O$, y, análogamente, $C_1O \equiv C_2O$. Pero A_1 y A_2 están situados en d hacia un lado de O . Debido a esto, de $A_1O \equiv A_2O$ se deduce que $A_1 = A_2$; análogamente, obtenemos, $C_1 = C_2$ y, por consiguiente, $A_1C_1 \equiv A_2C_2$. De aquí, finalmente, aplicando la propiedad transitiva de la \mathcal{L} -congruencia, obtenemos que $AC \equiv A'C'$, como se quería demostrar.

Introduzcamos el concepto de \mathcal{L} -ángulo (como figura geométrica). Llamaremos \mathcal{L} -ángulo al conjunto de dos \mathcal{L} -semirrectas h y k , que parten de un punto A y pertenecen a distintas \mathcal{L} -rectas. El ángulo se designa mediante $\widehat{h, k}$ o $\widehat{k, h}$; el punto A se llama vértice y h y k , lados del \mathcal{L} -ángulo (fig. 27).

Evidentemente, los lados de un \mathcal{L} -ángulo dividen el \mathcal{L} -plano en dos regiones tales, que cualquier par de puntos pertenecientes a una de las regiones pueden unirse mediante un segmento de \mathcal{L} -recta que no se corta con los lados del ángulo (región convexa), mientras que existen pares de puntos, pertenecientes a la otra región, para los cuales es imposible una unión semejante (región no convexa). Para convencerse de lo dicho, lo más fácil es transformar el punto A mediante un \mathcal{L} -movimiento en O , debido a lo cual las semirrectas h y k se transformarán en radios de la circunferencia. A la primera de las regiones indicadas la llamaremos interior, a la segunda, región exterior del \mathcal{L} -ángulo. Cercioré-

monos de la justeza de los siguientes axiomas planos de congruencia (los axiomas considerados anteriormente se llaman lineales).

III₄. Sean dados un \mathcal{L} -ángulo $\widehat{h, k}$, una \mathcal{L} -recta a' y su \mathcal{L} -semirrecta h' que parte de un punto A de esta \mathcal{L} -recta. Entonces hacia un lado dado de a' existe una, y sólo una, \mathcal{L} -semirrecta k' que parte de A , tal que el ángulo $\widehat{h', k'}$ es congruente al ángulo $\widehat{h, k}$. Cada ángulo es congruente a sí mismo, es decir, $\widehat{h, k} \equiv \widehat{h, k}$ y también $\widehat{h, k} \equiv \widehat{k, h}$.

La formulación de este axioma y la comprobación de su justeza son semejantes al axioma III₁. Lo único es que en la demostración de la unicidad de la \mathcal{L} -semirrecta k' hay que basarse ahora en la conservación de la magnitud (la medida) del ángulo en el \mathcal{L} -movimiento. Según esto, es evidente que dos \mathcal{L} -ángulos $\widehat{h', k'}$ y $\widehat{h'', k''}$ con un vértice común A y un lado común h' , para los cuales k' y k'' están situados hacia un lado de a' y representan distintas \mathcal{L} -semirrectas, no pueden ser \mathcal{L} -congruentes.

Consideremos, finalmente, tres segmentos AB , BC y CA , pertenecientes a distintas \mathcal{L} -rectas. Estos segmentos forman un \mathcal{L} -triángulo, para el cual los puntos A , B y C son los vértices, y AB , BC y CA , los lados. Si h y k son unas semirrectas que parten de A y contienen a los lados AB y AC , entonces el ángulo $(\widehat{h, k})$ se llama ángulo del triángulo ABC c o m p r e n d i d o entre los lados AB y AC u o p u e s t o al lado BC . Dentro del mismo están situados todos los puntos interiores del triángulo ABC ; se designa mediante \widehat{BAC} o \widehat{A} .

III₅. Si en los triángulos ABC y $A'B'C'$: $AB \equiv A'B'$, $AC \equiv A'C'$ y $\widehat{BAC} = \widehat{B'A'C'}$, entonces, $\widehat{ABC} \equiv \widehat{A'B'C'}$ y $\widehat{ACB} \equiv \widehat{A'C'B'}$.

Para comprobar este último axioma de congruencia, traslademos los triángulos ABC y $A'B'C'$, de modo que los vértices A y A' caigan al punto $z = \hat{0}$ y los lados AB y $A'B'$ coincidan con un mismo segmento de la recta OR . Entonces los lados AB , $A'B'$, AC y $A'C'$ se representarán por segmentos de radios que parten del vértice común O (A y A'). Si AC y $A'C'$ se sitúan en este caso hacia un lado del diámetro d de la circunferencia que lleva el segmento $AB = A'B'$, entonces, en virtud de la congruencia de los ángulos \widehat{BAC} y $\widehat{B'A'C'}$, AC y $A'C'$ tienen que estar situados en una misma \mathcal{L} -semirrecta que parte de O . Por esto coinciden los puntos C y C' (puesto que $AC \equiv A'C'$) y, debido a la coincidencia de los puntos A y A' y C y C' , coincidirán los lados AC y $A'C'$ y también BC y $B'C'$, de donde se deduce la congruencia de los ángulos \widehat{ABC} y $\widehat{A'B'C'}$ y \widehat{BCA} y $\widehat{B'C'A'}$ (y junto con ello la congruencia de los triángulos mismos). Si AC y $A'C'$ se sitúan hacia diversos lados del diámetro d , entonces mediante la transformación de simetría respecto de este diámetro, llegamos al caso ya estudiado.

Introduzcamos ahora el concepto de magnitud de un ángulo y longitud de un segmento. Estos conceptos, que denominaremos con el vocablo general de «medida», tienen que satisfacer a las siguientes condiciones:

I. La medida tiene que ser un número no negativo.

II. La medida no tiene que variar en los \mathcal{L} -movimientos, es decir, a los ángulos o segmentos congruentes tienen que corresponder medidas iguales.

III. La medida tiene que poseer la propiedad aditiva, en el sentido de que la medida de un segmento AC tiene que ser igual a la suma de las medidas de AB y BC para cualquier punto interior B de este segmento y la medida del

ángulo $(\widehat{h}, \widehat{k})$ tiene que ser igual a la suma de las medidas de los ángulos $(\widehat{h}, \widehat{l})$ y $(\widehat{l}, \widehat{k})$ para cualquier semirrecta l que parta del vértice del ángulo y esté contenida dentro del ángulo.

El problema de la medida de los ángulos se resuelve con una facilidad particular. Por medida del \mathcal{L} -ángulo $(\widehat{h}, \widehat{k})$ se toma la medida euclídea del ángulo formado por los arcos de las circunferencias que representan h y k . Evidentemente, en este caso se cumplen las condiciones I y III; en lo que se refiere a la condición II, ésta también se cumple debido a la conservación de las magnitudes de los ángulos en los \mathcal{L} -movimientos.

Consideremos ahora la medida de los segmentos. La longitud euclídea del arco de la circunferencia que representa el segmento AB también satisface a las condiciones I y III. Sin embargo, ésta no satisface a la condición II, puesto que los \mathcal{L} -movimientos (a excepción de la rotación alrededor del punto O y de la transformación de simetría respecto del diámetro) hacen variar la longitud euclídea. Para hallar una medida que satisfaga a todas las tres condiciones, apliquemos la invariabilidad en los \mathcal{L} -movimientos de la razón doble de cuatro puntos situados en cualquier recta o circunferencia euclídea.

Sean a y b los afijos de los extremos del segmento AB y sean α y β los afijos de los extremos del arco de circunferencia que lleva este segmento. Elijamos las designaciones para estos dos últimos puntos de modo que los cuatro puntos: α, β, a y b se sitúen en el arco de la circunferencia en el orden α, a, b y β . En el \mathcal{L} -movimiento los puntos α, a, b y β se transformarán en nuevos puntos: α', a', b' y β' , además, los puntos α' y β' se quedarán en la circunferencia y no se alterará el orden relativo de los puntos α', a', b' y β' en el arco, en el que se transformará el arco inicial. No variará tampoco la razón doble $(\alpha, \beta,$

$b, a) = \frac{b-\alpha}{b-\beta} : \frac{a-\alpha}{a-\beta}$. Demostremos que éste es un número positivo mayor

que la unidad. Con este fin, hagamos coincidir, mediante un \mathcal{L} -movimiento, la \mathcal{L} -recta que lleva el segmento AB con el diámetro d situado en el eje real, de modo que el punto a se traslade al punto O . Supongamos que, en este caso, el punto α se transforma en R , entonces β se transformará en $-R$ y el punto b , en un punto b' tal, que $0 > b' > -R$ (esto último es consecuencia de la conservación de la posición relativa de los puntos de la \mathcal{L} -recta en el \mathcal{L} -movimiento).

Obtendremos: $(\alpha, \beta, b, a) = (R, -R, b', 0) = \frac{b'-R}{b'+R} : \frac{0-R}{0+R} = \frac{R-b'}{R+b'} > 1$, como se quería demostrar. Por lo tanto, la razón doble $(\alpha, \beta,$

$b, a)$ como medida del segmento AB posee las propiedades I y II. No obstante, fácilmente se observa que no posee la propiedad aditiva III. En efecto, para un punto C de afijo c , perteneciente a AB , se tiene:

$$(\alpha, \beta, c, a) = \frac{c-\alpha}{c-\beta} : \frac{a-\alpha}{a-\beta}, \quad (\alpha, \beta, b, c) = \frac{b-\alpha}{b-\beta} : \frac{c-\alpha}{c-\beta}$$

y por consiguiente

$$\begin{aligned} (\alpha, \beta, b, a) &= \frac{b-\alpha}{c-\beta} : \frac{a-\alpha}{a-\beta} = \left(\frac{c-\alpha}{c-\beta} : \frac{a-\alpha}{a-\beta} \right) \left(\frac{b-\alpha}{b-\beta} : \frac{c-\alpha}{c-\beta} \right) = \\ &= (\alpha, \beta, c, a) \cdot (\alpha, \beta, b, c), \end{aligned}$$

o sea, la razón doble (α, β, b, a) no es igual a la suma de las razones dobles (α, β, c, a) y (α, β, b, c) , sino a su producto (que, por lo general, no es igual a la suma).

Pero ahora está claro que se satisfacen todas las condiciones que se exigen para la medida si se toma por longitud del segmento el logaritmo de la razón

doble: $\ln(\alpha, \beta, b, a)$. En efecto, como $(\alpha, \beta, b, a) > 1$, se tiene $\ln(\alpha, \beta, b, a) > 0$. Por otra parte, $\ln(\alpha, \beta, b, a)$ no varía en los \mathcal{L} -movimientos, pues esto se cumple para (α, β, b, a) . Finalmente, para cualquier punto C perteneciente al segmento AB , se tiene:

$$\ln(\alpha, \beta, b, a) = \ln(\alpha, \beta, c, a) + \ln(\alpha, \beta, b, c).$$

Obsérvese que en nuestra definición de \mathcal{L} -longitud del segmento AB (o, lo que es lo mismo, de \mathcal{L} -distancia entre los puntos A y B) los puntos de la circunferencia se pueden considerar como puntos del infinito del \mathcal{L} -plano. En efecto, fijando en la \mathcal{L} -recta un punto A (a), hacemos que el punto B (b) aproximarse indefinidamente a β (o fijando B , hacemos que A se aproxime indefinidamente a α). Entonces la razón doble $(\alpha, \beta, b, a) = \frac{b - \alpha}{b - \beta} : \frac{a - \alpha}{a - \beta}$ crecerá indefinidamente y, por consiguiente, la \mathcal{L} -longitud del segmento AB : $\ln(\alpha, \beta, b, a)$ tenderá hacia $+\infty$.

Continuando la verificación de los sistemas de axiomas de Hilbert para nuestro modelo de geometría de Lobachevski, omitimos por ahora el IV grupo de axiomas (el axioma de paralelismo), pasando inmediatamente al último, V grupo de axiomas de continuidad.

V_1 (axioma de Arquímedes). Sea A_1 un punto cualquiera de una \mathcal{L} -recta, situado entre unos puntos A y B arbitrariamente dados; tracemos los puntos A_2, A_3, A_4, \dots de modo que el punto A_1 esté situado entre A y A_2 , A_2 entre A_1 y A_3 , A_3 entre A_2 y A_4 , etc. y además, que los segmentos $AA_1, A_1A_2, A_2A_3, A_3A_4, \dots$ sean congruentes entre sí; entonces en la serie de puntos A_2, A_3, A_4, \dots siempre existe un punto A_n tal, que el punto B está situado entre A y A_n .

Para convencerse de que este axioma es justo es suficiente observar que, en virtud de las propiedades II y III de la medida, la longitud del segmento AA_n es igual al producto de la longitud del segmento AA_1 por n . Por consiguiente, la longitud del segmento AA_n crece indefinidamente junto con n y el punto A_n tiende hacia uno de los extremos del arco que representa la \mathcal{L} -recta. Como el punto B está situado entre A y este extremo, resulta que comenzando desde cierto n en adelante éste quedará situado entre A y A_n .

Como segundo axioma de continuidad, en lugar del axioma de Hilbert de plenitud, tomaremos el axioma de continuidad de Cantor.

V_2 . Si en una \mathcal{L} -recta existen dos sucesiones de puntos $A_1, A_2, \dots, A_n, \dots$ y $B_1, B_2, \dots, B_n, \dots$ tales que, cualesquiera que sean p y q , B_q está situado entre A_p y B_{q-1} , y A_p entre B_q y A_{p-1} , entonces en esta \mathcal{L} -recta existe al menos un punto C situado entre A_p y B_q para cualesquiera p y q .

Considerando esta proposición como un teorema de la geometría euclídea, que expresa una propiedad del arco de circunferencia que representa a nuestra \mathcal{L} -recta, nos convencemos inmediatamente de su justeza.

Hasta ahora observábamos una semejanza completa de las dos geometrías: la geometría de Euclides y la de Lobachevski. Esta semejanza se expresaba en que todos los axiomas de unión (relativos a la geometría plana), orden de congruencia y continuidad se formulaban igualmente en estas geometrías. De aquí se deduce que todos los teoremas de la geometría euclídea son válidos también en la geometría de Lobachevski. La diferencia entre las dos geometrías se manifiesta solamente al considerar los axiomas del IV grupo de la geometría euclídea, precisamente, el axioma de paralelismo. Este no se cumple en la geometría de Lobachevski. Aquí, por cada punto A del \mathcal{L} -plano, no situado en la \mathcal{L} -recta dada a , se puede trazar un conjunto infinito de \mathcal{L} -rectas distintas que no tengan puntos comunes con a . Esta proposición queda clara con solo mirar la fig. 28. Aquí, por el punto A se han trazado dos arcos $A\alpha$ y $A\beta$, ortogonales a la circunferencia Γ , de modo que cada uno de ellos posee un extremo común

con el arco $\alpha\beta$ que representa la \mathcal{L} -recta a . Los arcos $A\alpha$ y $A\beta$ representan \mathcal{L} -rectas que no tienen puntos comunes con la \mathcal{L} -recta a (recordemos al lector que, según la definición, son puntos del \mathcal{L} -plano solamente los que están situados en el interior de la circunferencia Γ). Estas dos \mathcal{L} -rectas se llaman \mathcal{L} -p a r a l e l a s a a trazadas por el punto A . Cada una de éstas tienen con a un punto del infinito común (α y β , respectivamente). Evidentemente, toda \mathcal{L} -recta que pase por A y esté comprendida en el ángulo formado por las \mathcal{L} -paralelas, no tiene puntos comunes con a , no sólo finitos sino tampoco del infinito. Las últimas \mathcal{L} -rectas no se llaman \mathcal{L} -paralelas.

Del hecho indicado se deducen una serie de consecuencias importantes. Demostremos ante todo el siguiente teorema: *la suma de los ángulos de un \mathcal{L} -triángulo es menor que π* . Sea ABC un \mathcal{L} -triángulo con los ángulos α , β y γ .

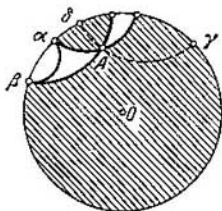


FIG. 28

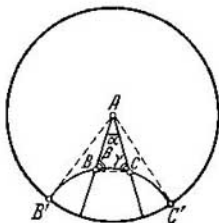


FIG. 29

Apliquémosle un \mathcal{L} -movimiento que lleve el vértice A al centro $o = 0$ del círculo. Entonces este \mathcal{L} -triángulo tomará la forma indicada en la fig. 29. Evidentemente, el arco $B'C'$, que representa la \mathcal{L} -recta en la cual está situada el lado BC del \mathcal{L} -triángulo, tendrá la convexidad dirigida hacia el vértice A . En efecto, este punto es el punto de intersección de las tangentes al arco $B'C'C'$ trazadas en los puntos B' y C' y, por consiguiente, está situado fuera de la circunferencia a la cual pertenece el arco $B'C'C'$. Uniendo B y C con un segmento de recta (euclídea), obtenemos un triángulo euclídeo ABC , cuya suma de ángulos $\hat{A} + \hat{B} + \hat{C}$ es igual a π . Pero $\alpha = \hat{A}$, $\beta < \hat{B}$ y $\gamma < \hat{C}$, por lo tanto, $\alpha + \beta + \gamma < \pi$, como se quería demostrar.

Demostremos otro teorema más que no hay en la geometría euclídea: *dos \mathcal{L} -triángulos con ángulos iguales son \mathcal{L} -congruentes*.

Sean ABC y $A_1B_1C_1$ triángulos con ángulos iguales, es decir, sean $\alpha = \alpha_1$, $\beta = \beta_1$ y $\gamma = \gamma_1$. Sometamos estos triángulos a sendos \mathcal{L} -movimientos, que transformen los puntos A y A_1 en el punto O , y los lados AB y A_1B_1 en segmentos de un mismo radio h de la circunferencia Γ (claro, estos movimientos serán distintos). Si, en este caso, los lados AC y A_1C_1 no caen en un mismo radio de la circunferencia, entonces, debido a la igualdad de los ángulos α y α_1 , se situarán en radios simétricos respecto de h . Por lo tanto, aplicando a uno de estos triángulos la transformación de simetría respecto de h , hacemos coincidir a los radios en los cuales están situados los lados AC y A_1C_1 . Si, entonces, los triángulos no coinciden, nos encontraremos con una de las dos posibilidades señaladas en la fig. 30.

En el caso a) los lados BC y B_1C_1 se cortan en cierto punto D , gracias a lo cual se forman dos \mathcal{L} -triángulos: BDB_1 y CDC_1 (si D coincide con B y B_1 o con C y C_1 , entonces uno de estos últimos degenera en un punto). Calculando los ángulos del \mathcal{L} -triángulo BDB_1 , tendremos, por ejemplo, para el caso repre-

sentado on la fig. a):

$$\hat{B}_1 = \beta_1, \quad \hat{B} = \pi - \beta \quad \text{y} \quad \hat{D} > 0.$$

Por lo tanto

$$\hat{B}_1 + \hat{B} + \hat{D} = \pi + \beta_1 - \beta + \hat{D} = \pi + D \quad (\text{puesto que } \beta_1 = \beta).$$

Así, pues, hemos llegado a la conclusión de que la suma de los ángulos del \mathcal{L} -triángulo BDB_1 es mayor que π , lo cual es imposible.

En el caso b), los ángulos del \mathcal{L} -cuadrilátero BB_1CC_1 son:

$$\hat{B} = \pi - \beta, \quad \hat{B}_1 = \beta_1, \quad \hat{C} = \pi - \gamma \quad \text{y} \quad \hat{C}_1 = \gamma_1.$$

Por lo tanto

$$\hat{B} + \hat{B}_1 + \hat{C} + \hat{C}_1 = 2\pi + \beta_1 - \beta + \gamma_1 - \gamma = 2\pi.$$

Pero la suma de los ángulos de este mismo \mathcal{L} -cuadrilátero se puede calcular dividiéndolo por la \mathcal{L} -diagonal B_1C en dos \mathcal{L} -triángulos BCB_1 y B_1CC_1 y

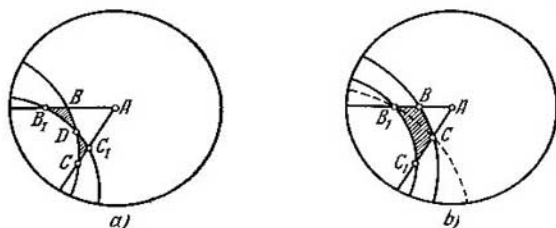


FIG. 30

hallando la suma de todos los ángulos de estos últimos. Como tiene que resultar un número menor que 2π , de aquí obtenemos de nuevo una contradicción.

Resumiendo, como resultado de los \mathcal{L} -movimientos señalados, los \mathcal{L} -triángulos ABC y $A_1B_1C_1$ tienen que coincidir, es decir, son congruentes.

Para la construcción de la geometría de Lobachevski tiene gran importancia el denominado ángulo de paralelismo. Sea a una \mathcal{L} -recta y A un punto fuera de ella. Tracemos por A dos \mathcal{L} -paralelas a a y una \mathcal{L} -perpendicular a a (fig. 31). Las paralelas forman con esta perpendicular dos ángulos en el punto A . Estos ángulos se llaman ángulos de paralelismo y dependen solamente de la \mathcal{L} -distancia del punto A hasta la \mathcal{L} -recta a (esta distancia se mide por la \mathcal{L} -longitud δ de la perpendicular AB bajada de A a a). Para estudiar el ángulo de paralelismo como función de δ —esta función se designa mediante $\Pi(\delta)$ — apliquemos el \mathcal{L} -movimiento que lleva la \mathcal{L} -recta perpendicular a a al diámetro d de modo que el punto A se transforme en el punto $A' = O$ (fig. 32). Entonces las \mathcal{L} -paralelas a a se transformarán en los radios de la circunferencia. De la simetría del dibujo se observa la igualdad de los ángulos entre las \mathcal{L} -paralelas y la perpendicular $A'B'$.

Designemos mediante b' el afijo del punto B' y sea, para fijar ideas, $b' > 0$. Entonces, para la \mathcal{L} -distancia del punto A' hasta la recta a' , tendremos:

$$\delta = \ln(-1, 1, b', 0) = \ln\left(\frac{b'+1}{b'-1} : \frac{0+1}{0-1}\right) = \ln \frac{1+b'}{1-b'},$$

de donde

$$\frac{1+b'}{1-b'} = e^\delta \quad \text{y} \quad b' = \frac{e^\delta - 1}{e^\delta + 1}.$$

Tracemos por el punto β' la tangente a la circunferencia unidad hasta su intersección con el eje real en el punto C' . Evidentemente, C' será el centro de la circunferencia cuyo arco representa la \mathcal{L} -recta a' , y $C'\beta' = \rho$, su radio.

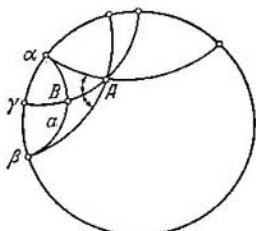


FIG. 31

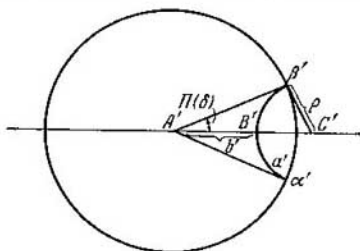


FIG. 32

Como $A'\beta'$ es un segmento tangente a esta circunferencia, resulta: $A'\beta'^2 = A'B'(A'B' + 2\rho)$ o sea,

$$1 = b'(b' + 2\rho),$$

de donde

$$\rho = \frac{1-b'^2}{2b'} = \frac{1 - \left(\frac{e^\delta - 1}{1 + e^\delta}\right)^2}{2 \frac{e^\delta - 1}{1 + e^\delta}} = \frac{2e^\delta}{e^{2\delta} - 1} = \frac{1}{\frac{e^\delta - e^{-\delta}}{2}} = \frac{1}{\text{sh } \delta},$$

donde $\text{sh } \delta$ es el seno hiperbólico de δ . Definitivamente, del triángulo $A'\beta'C'$ hallamos:

$$\Pi(\delta) = \text{arctg} \frac{\beta'C'}{A'\beta'} = \text{arctg} \frac{1}{\text{sh } \delta}.$$

De aquí se deduce que el ángulo de paralelismo $\Pi(\delta)$ está comprendido entre 0 y $\frac{\pi}{2}$:

$$0 < \Pi(\delta) < \frac{\pi}{2},$$

además, $\Pi(\delta) \rightarrow \frac{\pi}{2}$ a medida que $\delta \rightarrow 0$ (puesto que $\text{sh } \delta \rightarrow 0$, $\frac{1}{\text{sh } \delta} \rightarrow +\infty$ y, por consiguiente, $\text{arctg} \frac{1}{\text{sh } \delta} \rightarrow \frac{\pi}{2}$), y $\Pi(\delta) \rightarrow 0$ cuando $\delta \rightarrow \infty$ (puesto que $\text{sh } \delta \rightarrow \infty$, $\frac{1}{\text{sh } \delta} \rightarrow 0$ y $\text{arctg} \frac{1}{\text{sh } \delta} \rightarrow 0$).

La fórmula que acabamos de deducir para $\Pi(\delta)$ es básica para toda la geometría de Lobachevski.

Aquí no vamos a desarrollar más la geometría de Lobachevski, recomendando al lector que lea obras especiales *).

4.9. Entre las funciones racionales de grado superior al primero, en el ap. 3.3 estudiamos la función de la forma $w = (z - a)^n$, donde n es un número natural ($n > 1$).

Detengámonos también en la función $w = \frac{1}{2} \left(z + \frac{1}{z} \right) = \lambda(z)$, que frecuentemente aparece en la resolución de diversos problemas. Por todas las aplicaciones de esta función que N. Joukowski halló para la aeromecánica (véase el cap. V), se llama función de J o u k o w s k i.

Evidentemente, ésta es una función de segundo orden $\left(w = \frac{z^2 + 1}{2z} \right)$, que satisface a la condición $\lambda(z) = \lambda\left(\frac{1}{z}\right)$. De aquí se deduce que cada punto del plano w tiene en la transformación $w = \lambda(z)$ dos (no más de dos) preimágenes z_1 y z_2 , ligadas por la relación $z_1 z_2 = 1$. Si una de éstas pertenece al interior del círculo unidad, la otra pertenece a su exterior y viceversa. Por consiguiente, los conjuntos de valores $w = \lambda(z)$ que se toman en el interior y en el exterior del círculo unidad, tienen que ser iguales. Demostremos que la función $w = \lambda(z)$, siendo continua (en el sentido generalizado) en el dominio $|z| \leq 1$ (o $|z| \geq 1$) y tomando distintos valores en los puntos del recinto $|z| < 1$ (o $|z| > 1$), transforma biunívoca y continuamente el recinto $|z| < 1$ (o $|z| > 1$) en cierto recinto G del plano w . Para hallar la frontera Γ de este recinto hay que hallar la imagen de la circunferencia unidad $\gamma: |z| = 1$. Pero, si

$$z = e^{it}, \quad 0 \leq t \leq 2\pi, \quad \text{entonces } w = \frac{1}{2} (e^{it} + e^{-it}) = \cos t \quad (0 \leq t \leq 2\pi),$$

es decir, la imagen de la circunferencia unidad γ es el segmento del eje real $[-1, 1]$, recorrido dos veces. Por lo tanto, se puede esperar que el recinto G está formado por todos los puntos del plano w , a excepción de aquellos que pertenecen al segmento del eje real $\Gamma: -1 \leq x \leq 1$.

Para demostrar esto, hagamos un estudio más detallado de esta transformación; con este fin, consideremos (fig. 33) las imágenes de las circunferencias $|z| = r$ y los radios $\text{Arg } z = \alpha + 2k\pi$. Podemos limitarnos a considerar, por ejemplo, el interior de la circunferencia unidad $|z| < 1$.

Hagamos

$$z = re^{it}, \quad 0 \leq t \leq 2\pi \quad (0 < r < 1);$$

* N. Lobachevski. *Tres obras de geometría. Estudios geométricos sobre la teoría de las líneas paralelas*. Véase también N. Efimov, *Geometría superior*.

entonces

$$w = \frac{1}{2} \left[r e^{it} + \frac{1}{r} e^{-it} \right] = \frac{1}{2} \left(\frac{1}{r} + r \right) \cos t - i \frac{1}{2} \left(\frac{1}{r} - r \right) \sin t$$

o bien

$$u = \frac{1}{2} \left(\frac{1}{r} + r \right) \cos t, \quad v = -\frac{1}{2} \left(\frac{1}{r} - r \right) \sin t \quad (0 \leq t \leq 2\pi). \quad (4.9:1)$$

De aquí, eliminando el parámetro t , obtenemos:

$$\frac{u^2}{\left[\frac{1}{2} \left(\frac{1}{r} + r \right) \right]^2} + \frac{v^2}{\left[\frac{1}{2} \left(\frac{1}{r} - r \right) \right]^2} = 1. \quad (4.9:2)$$

Esta es la ecuación de una elipse con los semiejes $a = \frac{1}{2} \left(\frac{1}{r} + r \right)$ y $b = \frac{1}{2} \left(\frac{1}{r} - r \right)$ y focos ± 1 . De las fórmulas (4.9:1) se deduce que,

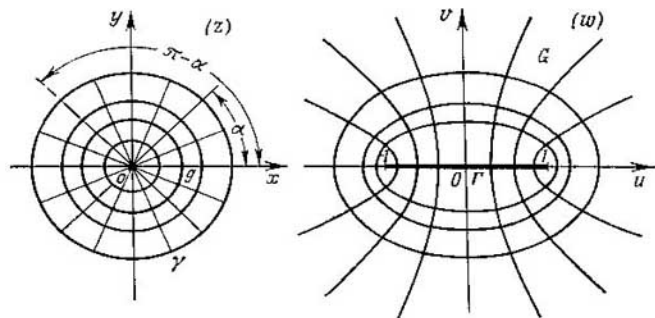


FIG. 33

cuando t crece continuamente desde 0 hasta 2π (es decir, que el punto z describe una sola vez toda la circunferencia $|z| = r$ en dirección positiva), el punto correspondiente describe una sola vez toda la elipse (4.9:2) en dirección negativa. En efecto, cuando $0 \leq t \leq \frac{\pi}{2}$, u es positivo y decrece desde a hasta 0, mientras que v es negativo y decrece desde 0 hasta $-b$; cuando $\frac{\pi}{2} < t < \pi$, u continúa decreciendo desde 0 hasta $-a$, mientras que v crece desde $-b$ hasta 0; cuando $\pi < t < \frac{3\pi}{2}$, u crece desde $-a$ hasta 0, mientras que v crece

desde 0 hasta b ; finalmente, cuando $\frac{3\pi}{2} < t < 2\pi$, u crece desde 0 hasta a , mientras que v decrece desde b hasta 0.

Variando el radio r de la circunferencia $|z| = r$ desde 0 hasta 1, hacemos decrecer a a desde ∞ hasta 1 y b , desde ∞ hasta 0; las elipses correspondientes describirán todo el conjunto de elipses del plano w con los focos ± 1 . De esto ya se deduce que $w = \lambda(z)$ transforma biunívocamente el círculo unidad en el recinto G que representa el exterior del segmento Γ . Además, la imagen del centro del círculo unidad es el punto del infinito, y la imagen de la circunferencia unidad es el segmento Γ (doblemente recorrido).

Para la imagen del radio $z = te^{i\alpha}$, $0 \leq t < 1$, obtenemos primero la ecuación

$$w = \frac{1}{2} \left(\frac{1}{t} + t \right) \cos \alpha - i \frac{1}{2} \left(\frac{1}{t} - t \right) \sin \alpha,$$

o bien

$$u = \frac{1}{2} \left(\frac{1}{t} + t \right) \cos \alpha, \quad v = -\frac{1}{2} \left(\frac{1}{t} - t \right) \sin \alpha \quad (0 \leq t < 1). \quad (4.9:3)$$

De aquí se ve que las imágenes de dos radios, simétricos respecto del eje real (si a uno de ellos le corresponde el ángulo α , entonces al otro le corresponderá el ángulo $-\alpha$), también son simétricos respecto del eje real, mientras que las imágenes de dos radios, simétricos respecto del eje imaginario (si a uno de ellos le corresponde el ángulo α , entonces al otro le corresponderá el ángulo $\pi - \alpha$), son simétricas respecto del eje imaginario. Por lo tanto, es suficiente considerar solamente las imágenes de los radios pertenecientes, por ejemplo, al primer cuadrante: $0 \leq \alpha \leq \frac{\pi}{2}$.

Obsérvese que para $\alpha = 0$ se tiene:

$$u = \frac{1}{2} \left(\frac{1}{t} + t \right), \quad v = 0 \quad (0 \leq t < 1).$$

Esto es un semiintervalo infinito del eje real: $1 < u \leq \infty$. El intervalo simétrico a éste $-\infty \leq u < -1$ es la imagen del radio que corresponde a $\alpha = \pi$.

Para $\alpha = \frac{\pi}{2}$, se tiene:

$$u = 0, \quad v = -\frac{1}{2} \left(\frac{1}{t} - t \right) \quad (0 \leq t < 1).$$

Esto es el semieje imaginario: $-\infty \leq v < 0$. El otro semieje imaginario $0 < v \leq \infty$ es la imagen del radio que corresponde a $\alpha = -\frac{\pi}{2}$.

En resumen, la imagen del diámetro «horizontal» de la circunferencia unidad es el intervalo infinito del eje real que va desde el punto -1 hasta el punto $+1$ pasando por ∞ , mientras que la imagen del diámetro «vertical» es todo el eje imaginario, a excepción del origen de coordenadas (incluyendo el punto del infinito).

Supongamos ahora que $0 < \alpha < \frac{\pi}{2}$. Entonces, eliminando el parámetro t entre las ecuaciones (4.9:3), obtenemos:

$$\frac{u^2}{\cos^2 \alpha} - \frac{v^2}{\sin^2 \alpha} = 1. \quad (4.9:4)$$

Esta es la ecuación de una hipérbola con el semieje real $a = \cos \alpha$; con el semieje imaginario $b = \sin \alpha$ y con los focos ± 1 . Sin embargo, el punto w no describe toda la hipérbola por completo cuando el punto z describe todo el radio $z = te^{i\alpha}$ ($0 \leq t < 1$). En efecto, de las ecuaciones (4.9:3) se deduce que, al crecer t desde 0 hasta 1, u decrece desde ∞ hasta $\cos \alpha$, mientras que v crece desde $-\infty$ hasta 0. Por consiguiente, el punto describe una vez solamente la cuarta parte de toda la hipérbola perteneciente al cuarto cuadrante. En virtud de lo observado anteriormente, la cuarta parte perteneciente al primer cuadrante, es decir, la simétrica a la dada respecto del eje real, será la imagen del radio simétrico al radio dado respecto del eje real, es decir, del radio correspondiente al ángulo $-\alpha$. Pero no sería justo decir que toda la rama de la hipérbola que pasa por el primero y cuarto cuadrantes es la imagen del par de radios indicados. En efecto, el vértice de la hipérbola $u = a, v = 0$ no pertenece a esta imagen (no olvidemos que nuestros radios se toman sin sus puntos extremos y que el vértice de la hipérbola es la imagen de cada uno de estos puntos: $t = 1$).

Luego obtenemos que las imágenes de los radios que corresponden a los ángulos $\pi - \alpha$ y $\alpha + \pi$ (o $\alpha - \pi$), son las cuartas partes de la misma hipérbola, situadas en el tercero y segundo cuadrantes.

La hipérbola completa, a excepción de sus dos vértices, es la imagen de la cuaterna de radios: $\pm \alpha, \pi \pm \alpha$. Obsérvese que la imagen de cada uno de los diámetros formados por estos radios será la parte de la hipérbola formada por los pares de sus cuartas partes que son simétricas respecto del origen de coordenadas y que están ligadas entre sí en el punto del infinito.

Resumiendo, la función $w = \lambda(z) = \frac{1}{2} \left(z + \frac{1}{z} \right)$ transforma biunívocamente tanto el interior como el exterior del círculo unidad en el exterior del segmento $-1 \leq u \leq +1$ (del eje real).

En este caso, las circunferencias $|z| = r$ se transforman en elipses con los focos ± 1 y semiejes: $\frac{1}{2} \left| \frac{1}{r} \pm r \right|$, y los pares de diámetros simé-

tricos respecto de los ejes coordenados (formados por los radios $z = \pm re^{i\alpha}$ ($0 \leq r < 1$) se transforman en hipérbolas con los focos ± 1 y semiejes $|\cos \alpha|$, $|\sin \alpha|$, a excepción de los vértices de estas hipérbolas.

Como la derivada de nuestra función

$$w = \lambda'(z) = \frac{1}{2} \left(1 - \frac{1}{z^2} \right)$$

es diferente de cero para $z \neq \pm 1$, la transformación es conforme en todos los puntos de los recintos considerados (el interior y el exterior del círculo unidad). De aquí se deduce que las hipérbolas se cortan con las elipses bajo ángulos iguales a los formados en las intersecciones de los radios y las circunferencias, es decir, bajo ángulos rectos. Esta deducción ya se hizo antes (ap. 3.7).

Consideremos también las imágenes de las circunferencias que pasan por los puntos ± 1 . De la igualdad

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

obtenemos:

$$w - 1 = \frac{z^2 - 2z + 1}{2z} = \frac{(z-1)^2}{2z}, \quad w + 1 = \frac{z^2 + 2z + 1}{2z} = \frac{(z+1)^2}{2z},$$

de donde

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

Fácilmente se observa que esta ecuación es equivalente a la dada. Haciendo $\frac{z-1}{z+1} = z'$ y $\frac{w-1}{w+1} = w'$, hallaremos que la transformación $w = \lambda(z)$ puede sustituirse por las siguientes:

$$z' = \frac{z-1}{z+1}, \quad w' = z'^2 \quad \text{y} \quad \frac{w-1}{w+1} = w'. \quad (4.9:5)$$

La primera transforma las circunferencias que pasan por los puntos ± 1 en rectas que pasan por el origen de coordenadas. la segunda transforma cada una de estas rectas en un rayo que parte del origen de coordenadas, y, finalmente, la última transforma cada uno de estos rayos en el arco de la circunferencia que une los puntos ± 1 . De las fórmulas (4.9:5) se ve fácilmente que, si el ángulo entre la circunferencia y la dirección positiva del eje real en el punto $z = 1$ era igual a θ , entonces el ángulo en el punto $w = 1$ entre su imagen (el arco de la circunferencia) y la dirección positiva del eje real será igual a 2θ (fig. 34).

Así, pues, la función $w = \lambda(z)$ transforma cada circunferencia γ , que pase por los puntos ± 1 y que forme en el punto 1 el ángulo θ con

la dirección positiva del eje real, en un arco δ de circunferencia que pasa por los puntos ± 1 y forma el ángulo 2θ con la dirección positiva del eje real. Las mismas fórmulas (4.9:5) muestran que en este caso cada uno de los arcos de γ por separado, con los extremos ± 1 , se transforma en el mismo arco δ .

Obsérvese que en la primera de las transformaciones (4.9:5) el exterior de la circunferencia γ se transforma en un semiplano, en la segunda obtenemos un recinto limitado por un rayo rectilíneo que parte del origen de coordenadas, y, finalmente, en la tercera,

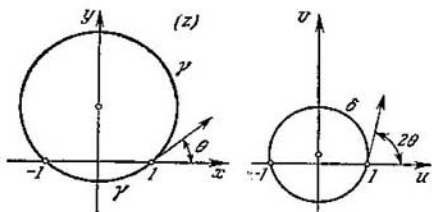


FIG. 34

un recinto cuya frontera es el arco δ . Como todas estas transformaciones son biunívocas en los recintos correspondientes, la función $w = \lambda(z)$ realiza una transformación biunívoca y conforme del exterior de la circunferencia γ (y también del interior de esta circunferencia) en el recinto cuya frontera es el arco de la circunferencia δ que une los puntos ± 1 .

Conviene señalar que la función $w = \lambda(z)$ transforma el semicírculo k_1 : $|z| < 1$, situado en el semiplano superior, en el semiplano inferior w , y el semicírculo k_2 situado en el semiplano inferior, en el semiplano superior w . Pero, en virtud de la relación $\lambda(z) = \lambda\left(\frac{1}{z}\right)$, la función toma en los puntos del semicírculo k_2 los mismos valores que en los puntos del semiplano superior que son exteriores al semicírculo k_1 . Designando el conjunto de estos últimos puntos mediante K_2 (fig. 35), se puede afirmar que la imagen del recinto K_2 también es el semiplano superior. Tengamos en cuenta, finalmente, que la imagen de la semicircunferencia que separa k_1 y K_2 es el intervalo $-1 < u < 1$, recorrido una sola vez. De aquí se deduce que la imagen del semiplano superior en la transformación $w = \lambda(z)$ consta de los semiplanos superior e inferior y del intervalo del eje real: $-1 < u < 1$, es decir, es todo el plano, a excepción del segmento infinito del eje real que une los puntos -1 y $+1$ mediante el punto del infinito.

Como aplicación de las observaciones hechas aquí, proponemos al lector estudiar la transformación de la franja $0 < x < \pi$ mediante la función $w = \cos z$, considerando esta transformación como el resultado de las transformaciones

$$z_1 = iz, \quad z_2 = e^{z_1}, \quad w = z_3 = \frac{1}{2} \left(z_2 + \frac{1}{z_2} \right),$$

efectuadas sucesivamente, una tras otra. El lector tiene que demostrar que $w = \cos z$ transforma biunívoca y conformemente la

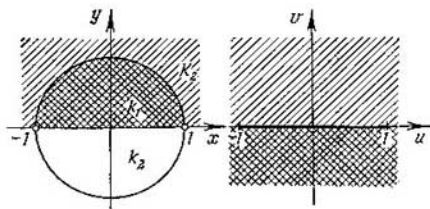


FIG. 35

franja indicada en el recinto del plano w limitado por el segmento infinito del eje real que une los puntos -1 y 1 mediante el punto del infinito.

4.10. Aquí nos detendremos brevemente en las funciones meromorfas trascendentes elementales (es decir, funciones meromorfas no racionales). Entre las más elementales (a excepción de las funciones enteras) figuran, por ejemplo, las funciones

$$\operatorname{tg} z = \frac{\operatorname{sen} z}{\operatorname{cos} z}, \quad \operatorname{ctg} z = \frac{\operatorname{cos} z}{\operatorname{sen} z}, \quad \operatorname{sec} z = \frac{1}{\operatorname{cos} z}, \quad \operatorname{cosec} z = \frac{1}{\operatorname{sen} z}.$$

Que éstas no son racionales, se deduce inmediatamente de que cada una de estas funciones posee un conjunto infinito de polos, es decir, de puntos en los cuales la función se hace infinita, mientras que la función racional posee solamente un número finito de polos.

Las funciones indicadas, así como las funciones enteras $\cos z$ y $\operatorname{sen} z$, figuran en la clase de las funciones trigonométricas. Esta última se define como la clase de funciones meromorfas $f(z)$ que admiten la representación

$$w = f(z) = \frac{P(e^{iz})}{Q(e^{iz})} = \frac{a_0 + a_1 e^{iz} + a_2 e^{2iz} + \dots + a_n e^{niz}}{b_0 + b_1 e^{iz} + b_2 e^{2iz} + \dots + b_m e^{miz}}. \quad (4.10:1)$$

Está claro que obtenemos la misma clase al considerar las funciones expresables en la forma

$$F(z) = \frac{\sum_{j=1}^n a_j e^{ijz}}{\sum_{h=1}^m b_h e^{ikhz}}. \quad (4.10:2)$$

Esta última expresión puede escribirse así:

$$\begin{aligned} F(z) &= \frac{a_0 + \sum_1^n [(a_j - a_{-j}) \cos jz + i(a_j + a_{-j}) \operatorname{sen} jz]}{b_0 + \sum_1^m [(b_h + b_{-h}) \cos kz + i(b_h - b_{-h}) \operatorname{sen} kz]} = \\ &= \frac{a_0 + \sum_1^n (A'_j \cos jz + A''_j \operatorname{sen} jz)}{b_0 + \sum_1^m (B'_h \cos kz + B''_h \operatorname{sen} kz)}. \end{aligned} \quad (4.10:3)$$

En el caso particular, cuando el denominador es constante (entonces se puede suponer que b_0 es igual a 1), obtenemos un polinomio trigonométrico:

$$F(z) = a_0 + \sum_1^n (A'_j \cos jz + A''_j \operatorname{sen} jz).$$

Si al menos uno de los números A'_n o A''_n es diferente de cero, se dice que $F(z)$ es un polinomio trigonométrico de orden n .

Volvamos a considerar de nuevo la fórmula (4.10:1) y hagamos $e^{iz} = t$; entonces obtenemos:

$$w = \frac{P(t)}{Q(t)} = S(t), \quad t = e^{iz},$$

es decir, una función trigonométrica arbitraria es una función racional de la exponencial. Evidentemente, toda función trigonométrica es periódica, de período 2π . Por lo tanto, es suficiente estudiarla en una franja cualquiera g : $x_0 \leq x < x_0 + 2\pi$. En cada una de las franjas

$$g_k: x_0 + 2k\pi \leq x < x_0 + (2k+1)\pi$$

esta función, debido a la periodicidad, tendrá un mismo comportamiento. Supongamos que z describe la franja g (incluyendo en ella la recta $x = x_0$); entonces $z_1 = iz$ describirá una franja $x_0 \leq y_1 < x_0 + 2\pi$ de la misma anchura 2π , paralela al eje real. Por consi-

guiente, $t = e^{iz}$ describirá un ángulo de medida 2π con el vértice en el origen de coordenadas. Los lados de este ángulo se confunden en una semirrecta $\text{Arg } t = x_0 + 2k\pi$, que también se recorrida por el punto t , precisamente cuando z describe la recta $\text{Re } z = x_0$. Al fin y al cabo, cuando z describe la franja g , t describe todo el plano, a excepción de los puntos $t = 0$ y $t = \infty$.

Si N es el orden de la función racional $w = S(t)$, ésta, como ya se sabe por el ap. 4.1, toma cada uno de sus valores, por lo general, en N puntos del plano ampliado t . Como en nuestro caso se han excluido los valores $t = 0$ y $t = \infty$, solamente se puede afirmar que w tomará todos los valores, a excepción, posiblemente, de los valores $S(0)$ y $S(\infty)$, para N valores de t (entre los cuales algunos pueden ser iguales entre sí para valores especiales de w).

Pero la correspondencia entre t y z es biunívoca dentro de los límites de la franja g ; debido a esto la función trigonométrica $w = S(e^{iz})$ toma cada valor complejo (a excepción, posiblemente, de los números $S(0)$ y $S(\infty)$) en N puntos de la franja g y, por consiguiente, de cada franja g_k . Entre estos N puntos puede haber algunos coincidentes; pero esto último es posible solamente para algunos valores de w (la cantidad de éstos no es mayor que $n + m$, donde n y m son los grados de $P(t)$ y $Q(t)$; véase el ap. 4.1).

De lo expuesto se deduce que cada ecuación de la forma

$$f(z) = S(e^{iz}) = A,$$

donde $f(z)$ es una función trigonométrica ($\neq \text{const}$) y A es un número complejo cualquiera, posee un conjunto infinito de raíces (N raíces en cada franja g_k), a excepción, posiblemente, de dos valores especiales de A : $A = S(0)$ y $A = S(\infty)$, para los cuales la ecuación puede no poseer ninguna raíz.

Ejemplos.

$$1) \quad f(z) = e^{inz} = t^n, \quad t = e^{iz}.$$

Aquí $S(t) = t^n$, $S(0) = 0$, $S(\infty) = \infty$ y la función $S(t)$ no toma los valores 0 y ∞ cuando $t \neq 0$, $t \neq \infty$; debido a esto, $f(z)$ no toma los valores 0 y ∞ para ningún valor de z ,

$$2) \quad f(z) = \sec z = \frac{2e^{iz}}{1+e^{2iz}} = \frac{2t}{1+t^2}, \quad t = e^{iz}.$$

Aquí $S(t) = \frac{2t}{1+t^2}$, $S(0) = S(\infty) = 0$ y $S(t)$ no se anula cuando $t \neq 0$, $t \neq \infty$. Por lo tanto, $f(z) = \sec z$ no se anula para ningún valor de z .

$$3) \quad f(z) = \text{tg } z = \frac{1}{i} \frac{e^{2iz} - 1}{e^{2iz} + 1} = \frac{1}{i} \frac{t^2 - 1}{t^2 + 1}, \quad t = e^{iz}.$$

Aquí $S(t) = \frac{1}{t} \frac{t^2 - 1}{t^2 + 1}$, $S(0) = i$, $S(\infty) = -i$ y $S(t)$ no toma los valores i y $-i$ cuando $t \neq 0$, $t \neq \infty$. Por lo tanto, $f(z) = \operatorname{tg} z$ no toma los valores i y $-i$ para ningún valor de z .

$$4) \quad f(z) = \frac{\cos z}{\cos 2z} = \frac{e^{3iz} + e^{iz}}{e^{4iz} + 1} = \frac{t^3 + t}{t^4 + 1}, \quad t = e^{iz}.$$

Aquí $S(t) = \frac{t^3 + t}{t^4 + 1}$, $S(0) = S(\infty) = 0$. Sin embargo, $S(t)$ se anula también en puntos distintos de 0 y ∞ , precisamente, cuando

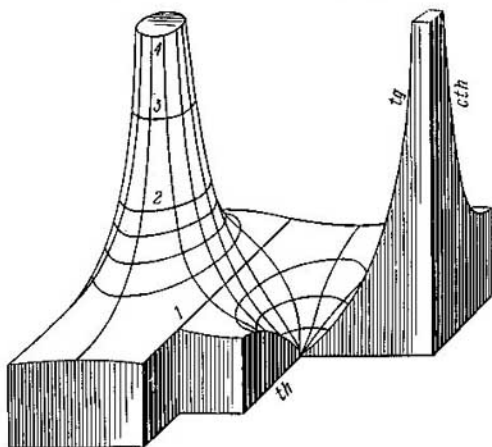


FIG. 36

$t = \pm i$. Por lo tanto, $f(z) = \frac{\cos z}{\cos 2z}$ toma todos los valores complejos sin excepción (y cada uno de ellos en un conjunto infinito de puntos).

Examinemos más detalladamente la función $\operatorname{tg} z = \frac{1}{i} \frac{t^2 - 1}{t^2 + 1}$ ($t = e^{iz}$). En la fig. 36 está representada la superficie $u = |\operatorname{tg} z|$, es decir, el relieve de la tangente*). Para obtener la transformación $w = \operatorname{tg} z$, representémosla en forma de las siguientes transformaciones, realizadas sucesivamente una tras otra:

$$\zeta = iz, \quad t = e^{\zeta}, \quad t_1 = t^2, \quad w = t_2 = \frac{1}{i} \frac{t_1 - 1}{t_1 + 1}.$$

*) El dibujo se ha copiado de las «Tablas de funciones» de Jahnke y Emde.

Si g es la franja $x_0 \leq x \leq x_0 + h$, donde $0 < h \leq \pi$, entonces $\zeta = iz$ transforma a ésta en la franja $x_0 \leq \eta \leq x_0 + h$, $t = e^{\zeta}$ transforma esta última franja en un ángulo de magnitud h con el vértice en el origen de coordenadas y con los lados $\text{Arg } t = x_0 + 2k\pi$ y $\text{Arg } t = x_0 + h + 2n\pi$; por otra parte, la transformación $t_1 = t^2$ nos da un ángulo de magnitud $2h$ con los lados $\text{Arg } t_1 = 2x_0 + 2k\pi$ y $\text{Arg } t_1 = 2x_0 + 2h + 2k\pi$ y, finalmente, $w = \frac{1}{i} \frac{t_1 - 1}{t_1 + 1}$, como transformación homográfica, transforma los

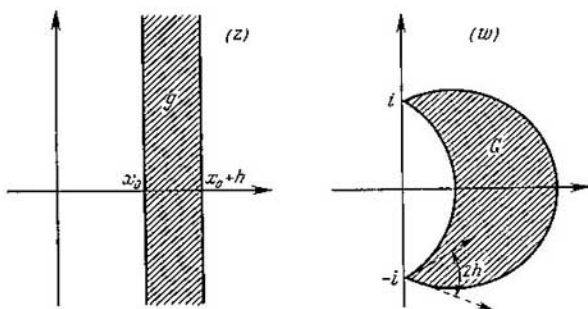


FIG. 37

lados del ángulo en arcos de circunferencias que unen los puntos $w = i$ ($t_1 = 0$) y $w = -i$ ($t_1 = \infty$), y el mismo ángulo, en una lúnula circular (biángulo) con los mismos ángulos $2h$. Como la derivada $\left. \frac{dw}{dt_1} \right|_{t_1=0} = -2i$ y $\text{Arg}(-2i) = -\frac{\pi}{2} + 2k\pi$, las tangentes a las curvas que parten del punto $t_1 = 0$ giran en la transformación $w = \frac{1}{i} \frac{t_1 - 1}{t_1 + 1}$ el ángulo $-\frac{\pi}{2}$; por esto, las tangentes a los arcos de las circunferencias que limitan el biángulo tienen que formar con la dirección positiva del eje real, en el punto $w = i$, los ángulos $2x_0 - \frac{\pi}{2}$ y $2x_0 + 2h - \frac{\pi}{2}$. Con estas condiciones, el biángulo con los vértices i y $-i$ queda completamente determinado (fig. 37).

Cada una de las transformaciones consideradas para separado es biunívoca y conforme. Debido a esto, la transformación resultante $w = \text{tg } z$ transforma biunívoca y conformemente la franja $x_0 \leq x = \text{Re } z \leq x_0 + h$ de anchura h ($h \leq \pi$) en un biángulo circular de ángulos iguales a $2h$, con los vértices en los puntos $-i$ y i .

En particular, si $h = \frac{\pi}{2}$, los ángulos del biángulo se convierten en π , y el biángulo mismo, en un círculo o en un semiplano. De aquí se deduce que la función $w = \operatorname{tg} z$ transforma cada franja $x_0 < x < x_0 + \frac{\pi}{2}$ de anchura $\frac{\pi}{2}$ en un círculo o en un semiplano.

§ 5. FUNCIONES MULTIFORMES ELEMENTALES

5.1. Las funciones enteras y meromorfas $w = f(z)$, estudiadas en los apartados anteriores, toman por lo general un mismo valor w en unos cuantos (dos o más) puntos del plano z . Forma una excepción solamente la función homográfica, que realiza una transformación biunívoca del plano ampliado sobre sí mismo. Dejando de un lado esta excepción, en todos los demás casos la transformación inversa $z = f^{-1}(w)$ no es unívoca. Esto significa que las funciones inversas a las consideradas son multiformes.

Para que sea posible aplicar a las funciones multiformes los conceptos y resultados obtenidos para las funciones uniformes, es necesario saber separar las ramas (o determinaciones) uniformes de estas funciones.

He aquí cómo se consigue esto ordinariamente.

Sea $z = f(w)$ una función definida, uniforme y continua (en sentido generalizado) en un recinto G del plano ampliado. Supongamos que se ha conseguido dividir de algún modo el recinto G en un conjunto finito o numerable de recintos g_1, g_2, \dots , que carecen de puntos comunes dos a dos, de modo que cualquier punto del recinto G sea interior para un solo recinto g_k o sea punto frontera común al menos para dos recintos g_j o g_h , y que en cada uno de estos recintos la transformación $z = f(w)$ sea biunívoca. Entonces la imagen de cada g_k será también un recinto $f(g_k) = G_k$ (en el cap. V esta afirmación fue demostrada con la condición de que $f(w)$ sea analítica) y toda la imagen $f(G)$ se cubrirá por los recintos G_k y también por las imágenes de las partes comunes de las fronteras de los recintos g_k .

Consideremos la función inversa $w = F(z)$ en cada uno de los recintos G_k definiéndola por la condición complementaria de que sus valores pertenezcan a g_k , que es la preimagen del recinto G_k . Entonces la función $F(z)$, generalmente multiforme, se expresará mediante unas cuantas, posiblemente infinitas, funciones $F_k(z)$ uniformes y continuas (en sentido generalizado). A cada una de éstas llamaremos rama (o determinación) uniforme de la función $F(z)$ en el recinto correspondiente G_k . Es importante tener en cuenta en esta definición que el carácter de los recintos G_k , y también el de las ramas uniformes $F_k(z)$ de la función, depende esencialmente de la forma en que se ha dividido el recinto G

en recintos g_h . En los casos más elementales, el recinto G admite una división en recintos g_h tal, que los recintos correspondientes G_h coinciden entre sí. Supongamos, por ejemplo, que G_{h_1}, G_{h_2}, \dots coinciden con un mismo recinto G' . Entonces la función multiforme $w = F(z)$ posee muchas, posiblemente infinitas, ramas uniformes en el recinto G' , precisamente: $F_{h_1}(z), F_{h_2}(z), \dots$.

A todo lo dicho anteriormente hay que añadir que para una función continua arbitraria $z = f(w)$ la división del recinto G en recintos g_h , que satisfagan a las condiciones indicadas anteriormente, es generalmente imposible. Sin embargo, cuando la función $z = f(w)$ es analítica en el recinto G (a excepción de puntos aislados en los cuales ella puede ser igual a ∞), en el ap. 3.4 del octavo capítulo se demostrará que semejante división siempre es posible y además de infinitos modos.

A una función $z = f(w)$, analítica en cierto recinto g (a excepción, posiblemente, de los puntos en los cuales ella es igual a ∞) y que tome en distintos puntos del recinto diferentes valores ($f(w_1) \neq f(w_2)$) si $w_1 \neq w_2$ y $w_1, w_2 \in g$, llamaremos **univalente** en el recinto g . Si existe en el recinto al menos un par de puntos distintos en los cuales $f(w)$ toma un mismo valor: $f(w_1) = f(w_2)$, $w_1 \neq w_2$, diremos que la función es **multivalente** en este recinto.

El hecho al que alegábamos anteriormente, se puede enunciar así: si una función analítica $z = f(w)$ es multivalente en un recinto G , éste puede dividirse en un conjunto finito o numerable de recintos en cada uno de los cuales $f(w)$ es univalente. Los recintos correspondientes g_h se llaman **recintos de univalencia** de la función $f(w)$.

Por lo tanto, el método descrito anteriormente de separación de ramas uniformes siempre es aplicable a las funciones que son inversas respecto de las multivalentes.

En este párrafo se ilustrará el método indicado para el caso de funciones elementales; mas no tendremos que basarnos en el teorema que se ha señalado sin demostración, puesto que la división del recinto G en recintos de univalencia se obtendrá cada vez aplicando las propiedades conocidas de las funciones elementales.

Además de las funciones inversas a las elementales, estudiaremos también aquí otras funciones multiformes, obtenidas como funciones compuestas de la forma $\varphi\psi(z)$ (donde $\varphi(z)$ o $\psi(z)$ son funciones inversas de las elementales), o como combinaciones racionales de tales funciones.

5.2. Consideremos el radical $w = \sqrt[n]{z}$, que representa una función inversa a la función potencial $z = w^n$ (n es un número natural, mayor que la unidad).

Para cada z , distinto de cero e infinito, el radical posee n valores distintos, dados por la fórmula

$$w = \sqrt[n]{+|z|} \left(\cos \frac{\text{Arg } z}{n} + i \text{sen} \frac{\text{Arg } z}{n} \right). \quad (5.2:1)$$

Cuando $z = 0$ o $z = \infty$ se obtiene cada vez un solo valor de la función, $w = 0$ o $w = \infty$, respectivamente. Los n valores (5.2:1), que representan aquellos puntos del plano w en los cuales w^n toma un mismo valor z , se sitúan en los vértices de un polígono regular de n lados inscrito en una circunferencia $|w| = \sqrt[n]{+|z|}$.

Recíprocamente: los vértices de cualquier polígono regular de n lados con el centro en el origen de coordenadas se pueden considerar como n valores de $\sqrt[n]{z}$. Debido a esto, un recinto g del plano w será un recinto de univalencia para $z = w^n$ cuando, y sólo cuando, de los n vértices de cualquier polígono regular con el centro $w = 0$ éste contiene no más de un vértice.

Evidentemente, a esta condición satisface cada ángulo de medida $\frac{2\pi}{n}$ con el vértice en el origen de coordenadas. Tracemos desde el origen de coordenadas n rayos rectilíneos que formen entre sí ángulos iguales. Entonces hallaremos que todo el plano, en el cual está definida la función multivalente $z = w^n$, se dividirá en n recintos de univalencia de esta función: g_1, g_2, \dots, g_n . La imagen de cada uno de éstos será un mismo recinto G' del plano z cuya frontera es un rayo rectilíneo L que parte del origen de coordenadas. Si el recinto g_h está limitado por los rayos que forman los ángulos $\varphi_0 + \frac{2k\pi}{n}$ y $\varphi_0 + \frac{2(k-1)\pi}{n}$ con la parte positiva del eje real, entonces el rayo L formará con la parte positiva del eje real el ángulo $n\varphi_0$.

De acuerdo con lo dicho en el ap. 5.1, obtendremos en el recinto G' n ramas uniformes de la función $\sqrt[n]{z}$. Cada una de ellas: $\sqrt[n]{z}_h$ ($h = 1, 2, \dots, n$) se determina completamente por la condición de que sus valores $w = \sqrt[n]{z}_h$ pertenecen al recinto g_h . Como $z = w^n$ posee derivada diferente de cero en todos los puntos del recinto g_h : $z' = nw^{n-1}$, las ramas $\sqrt[n]{z}_h$ también poseerán derivadas diferentes de cero:

$$\left(\sqrt[n]{z}_h \right)' = \frac{1}{nw^{n-1}} = \frac{1}{n \left(\sqrt[n]{hz} \right)^{n-1}}.$$

Tomemos ahora un sistema de rayos rectilíneos que partan del origen de coordenadas, y que se obtiene del anterior mediante una rotación alrededor del origen de coordenadas en un ángulo α , $0 <$

$< \alpha < \frac{2\pi}{n}$. Entonces el nuevo sistema dividirá el plano w en n recintos d_1, \dots, d_n , cada uno de los cuales d_h tendrá partes comunes con dos recintos vecinos g_h y g_{h+1} (si $k = n$, entonces g_{n+1} debe sustituirse por g_1) (fig. 38).

La imagen de cada uno de los recintos d_h del plano w será un mismo recinto D' , limitado por el rayo rectilíneo M que parte del origen de coordenadas y forma con la parte positiva del eje real

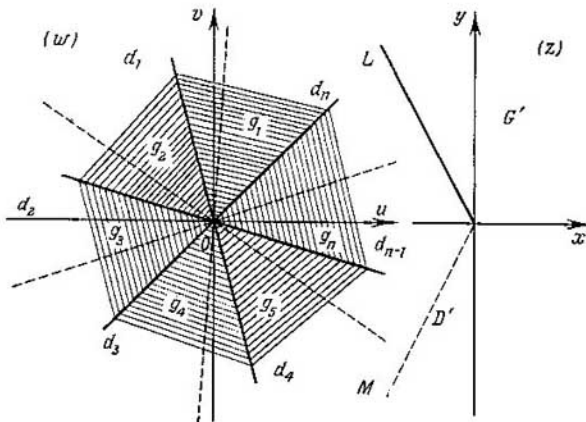


FIG. 38

el ángulo $n\varphi_0 + n\alpha$. En este recinto también obtendremos n ramas uniformes de la función $\sqrt[n]{z}$, cada una de las cuales se determina porque sus valores pertenecen al recinto correspondiente d_h . Designemos estas ramas mediante $(\sqrt[n]{z})_h$. Estas son diferenciables en D' y para sus derivadas se tiene:

$$(\sqrt[n]{z})'_h = 1 : n (\sqrt[n]{z})_h^{n-1}.$$

Comparémoslas con las ramas $\sqrt[n]{z}$; como parte de la preimagen d_h del recinto D' en el plano w pertenece al recinto g_h y parte al recinto g_{h+1} , la rama $(\sqrt[n]{z})_h$ en la parte del recinto D' que representa la imagen de la parte común de los recintos d_h y g_h , coincidirá con $\sqrt[n]{z}$, mientras que en la otra parte del recinto D' (que representa

la imagen de la parte común de los recintos d_h y g_{h+1}) coincidirá con $\sqrt[h+1]{z}$.

Vemos, pues, que al sustituir unos recintos de univalencia por otros cada nueva rama uniforme se obtiene mediante la unión de la parte de definición de una de sus ramas anteriores con la parte de definición de otra rama anterior.

Si el ángulo de rotación $\alpha = 0$, entonces d_h coincide con g_h , D' coincide con G' y cada rama $(\sqrt[h]{z})_k$ coincide con $\sqrt[h]{z}$. Cuando el ángulo α , aumentando continuamente, se aproxima a $\frac{2\pi}{n}$, el recinto d_h se aproxima a g_{h+1} , el recinto correspondiente D' se aproxima a G' y la rama $(\sqrt[h]{z})_k$ en mayor y mayor parte del recinto D' coincide con la rama $\sqrt[h+1]{z}$ (en lugar de g_{h+1} y $\sqrt[h+1]{z}$ se debe tomar g_1 y $\sqrt[1]{z}$).

Cuando $\alpha = \frac{2\pi}{n}$, d_h coincide con g_{h+1} , D' coincide con G' y la rama $(\sqrt[h]{z})_k$ se convierte en la rama $\sqrt[h+1]{z}$.

También se puede observar el paso de una rama $\sqrt[h]{z}$ a otra $\sqrt[h+1]{z}$ haciendo describir al punto z una circunferencia completa con el centro en el origen de coordenadas. Si el valor de $\sqrt[h]{z}$ en el punto z_0 se suponía perteneciente a la rama $\sqrt[h]{z}$ y se representaba por el punto w_0 del recinto g_h :

$$w_0 = \sqrt[h]{|z_0|} \left(\cos \frac{\varphi_0}{n} + i \operatorname{sen} \frac{\varphi_0}{n} \right),$$

al moverse continuamente el punto z por la circunferencia $|z| = |z_0|$ en la dirección positiva, el valor correspondiente del radical

$$w = \sqrt[h]{|z_0|} \left(\cos \frac{\varphi}{n} + i \operatorname{sen} \frac{\varphi}{n} \right)$$

variara continuamente junto con φ , y después de un recorrido completo al volver el punto z a la posición inicial z_0 , el valor del radical se convertirá en

$$w_1 = \sqrt[h]{|z_0|} \left(\cos \frac{\varphi_0 + 2\pi}{n} + i \operatorname{sen} \frac{\varphi_0 + 2\pi}{n} \right).$$

Este último se obtiene de w_0 mediante una rotación alrededor del origen de coordenadas en el ángulo $\frac{2\pi}{n}$; por consiguiente, el

punto w_1 , pertenece al recinto g_{h+1} , vecino de g_h , y w_1 es el valor de la rama $\sqrt[h+1]{z}$ en el punto z_0 .

Esta conclusión se puede aplicar a cualquier punto del recinto G' , de donde se deduce que al recorrer el punto z una circunferencia de cualquier radio con el centro en el origen de coordenadas, los valores de $\sqrt[h]{z}$, variando continuamente, pasan de la rama $\sqrt[h]{z}$ a la rama $\sqrt[h+1]{z}$.

Se necesita un recorrido n -múltiple del punto z alrededor del punto $z = 0$ en la dirección positiva para que las ramas del radical $\sqrt[n]{z}$, sustituyéndose una por otra ($\sqrt[h]{z}$ por $\sqrt[h+1]{z}$, $\sqrt[h+1]{z}$ por $\sqrt[h+2]{z}$,, $\sqrt[n-1]{z}$ por $\sqrt[n]{z}$,, $\sqrt[n]{z}$ por $\sqrt[h]{z}$), vuelvan a las ramas iniciales.

El punto que posee la propiedad de que un recorrido completo (de una vuelta) por una curva cerrada de Jordan alrededor del punto, en cualquier entorno del mismo, sustituye una rama continuamente variable de la función multiforme por otra rama de esta función, se llama punto de ramificación de la función.

El hecho de que después de dar n vueltas en una misma dirección obtenemos la rama inicial, se expresa diciendo que el punto dado de ramificación es de orden finito, precisamente de orden $n - 1$, llamándose punto algebraico de ramificación *).

Resumiendo, el punto $z = 0$ es un punto algebraico de ramificación de orden $n - 1$ para la función $\sqrt[n]{z}$.

Evidentemente, el punto $z = \infty$ también se puede considerar como punto algebraico de ramificación de orden $n - 1$ de la función $\sqrt[n]{z}$, puesto que cada recorrido alrededor del mismo a lo largo de una circunferencia de radio arbitrariamente grande con el centro en el origen de coordenadas es a la vez un recorrido alrededor del origen de coordenadas. Por esto, la función multiforme $w = \sqrt[n]{z}$ posee dos puntos de ramificación en el plano z : $z = 0$ y $z = \infty$, ambos de orden $n - 1$.

Las ramas uniformes de esta función descritas anteriormente fueron construidas para los recintos del tipo G' o D' , cuyas fronteras representaban un rayo rectilíneo que unía ambos puntos de ramificación. Se obtiene un tipo más general de un recinto semejante si en lugar de un rayo rectilíneo se traza una curva de Jordan arbitraria del plano ampliado que una los puntos 0 y ∞ . Sea Γ esta curva

* El último concepto supone también que en el punto dado existe el límite de la función (finito o infinito).

y G , el recinto limitado por ella. Si z describe la curva Γ desde el punto inicial (0) hasta el final (∞), entonces los n puntos $w = \sqrt[n]{z}$ que le corresponden describen n curvas de Jordan γ_k que unen el punto 0 con el punto ∞ . Estas curvas no tienen otros puntos comunes, además de 0 y ∞ , y forman a pares (γ_k con γ_{k+1}) curvas de Jordan cerradas del plano ampliado.

Sea g_k^i aquel de los dos recintos limitados por las curvas γ_k y γ_{k+1} que no contiene a las curvas $\gamma_1, \dots, \gamma_{k-1}, \gamma_{k+1}, \dots, \gamma_n$. Al girar el plano z alrededor del origen de coordenadas el ángulo $\frac{2\pi}{n}$, la curva γ_k , debido a su construcción, pasa a γ_{k+1} , la curva γ_{k+1} a γ_{k+2} y el recinto g_k^i al recinto g_{k+1}^i . Como g_k^i y g_{k+1}^i no tienen puntos comunes, ninguno de estos recintos contendrá un par de puntos que pasen uno a otro como resultado de tal giro. Por esta razón, todos los recintos g_k^i ($k = 1, 2, \dots, n$) son recintos de univalencia para $z = w^n$, y obtenemos n ramas uniformes de la función $\sqrt[n]{z}$ en el recinto G , exigiendo que los valores de cada rama pertenezcan al recinto correspondiente g_k^i . Para fijar una rama es suficiente señalar el valor de $\sqrt[n]{z}$ en un punto cualquiera z_0 del recinto G ; si este valor es w_0 , existirá un recinto único g_k^i que contendrá el punto w_0 , y junto con él una rama única $\sqrt[n]{z}$ en el recinto G , que tomará el valor w_0 en el punto z_0 . Se obra de este modo, precisamente, cuando se desea fijar una rama determinada $\sqrt[n]{z}$ en un recinto del tipo G .

Sean $\sqrt[n]{z}_k$ y $\sqrt[n]{z}_l$ dos ramas de $\sqrt[n]{z}$ en el recinto G y supongamos que sus valores en cierto punto z_0 son iguales a w'_0 y w''_0 , respectivamente. Como

$$w'_0 = \sqrt[n]{kz} = \sqrt[n]{|z|} \left(\cos \frac{\varphi_0 + 2m'\pi}{n} + i \operatorname{sen} \frac{\varphi_0 + 2m'\pi}{n} \right),$$

$$w''_0 = \sqrt[n]{lz} = \sqrt[n]{|z|} \left(\cos \frac{\varphi_0 + 2m''\pi}{n} + i \operatorname{sen} \frac{\varphi_0 + 2m''\pi}{n} \right),$$

donde m' y m'' son números enteros, el valor w''_0 puede obtenerse de w'_0 multiplicándolo por

$$\eta = \cos \frac{2(m'' - m')\pi}{n} + i \operatorname{sen} \frac{2(m'' - m')\pi}{n},$$

es decir, por uno de los valores de $\sqrt[n]{1}$. Pero, al multiplicar la función $\sqrt[n]{z}_k$ por el número η se obtiene, evidentemente, una función $\eta \sqrt[n]{z}_k$, uniforme y continua en el recinto G , cuyos valores representan $\sqrt[n]{z}$ y pertenecen al mismo recinto al cual pertenece el punto $\eta \sqrt[n]{z}_0 = \sqrt[n]{z}_l$. Por consiguiente, $\eta \sqrt[n]{z}_k = \sqrt[n]{z}_l$ en todo el recinto

to G . Vemos, pues, que las dos ramas de $\sqrt[n]{z}$ en un mismo recinto G pueden obtenerse una de otra multiplicando por cierto valor de $\sqrt[n]{1}$.

Todas las conclusiones de este apartado se extienden, con las evidentes variaciones correspondientes, a las funciones de una forma más general:

$$w = \sqrt[n]{z-a} \quad \text{o bien} \quad w = \sqrt[n]{\frac{z-a}{z-b}}.$$

Recomendamos al lector estudiar estos ejemplos, observando que estas funciones son inversas a las funciones $z = a + w^n$ y $z = \frac{bw^n - a}{w^n - 1}$, para las cuales los recintos de univalencia son los mismos que para la función $z = w^n$. Hay que observar en este caso que los puntos de ramificación de la función $w = \sqrt[n]{z-a}$ son a y ∞ , y los puntos de ramificación de la función $w = \sqrt[n]{\frac{z-a}{z-b}}$ son a y b , y que la separación de una rama uniforme de la función es posible en cualquier recinto cuya frontera sea un arco de Jordan que una los puntos de ramificación.

5.3. Para aclarar mejor el concepto de punto de ramificación, consideremos la función multiforme

$$w = f(z) = \sqrt[n]{P(z)}, \quad (5.3:1)$$

donde $P(z)$ es un polinomio arbitrario. Sea N el grado de este polinomio, a_1, a_2, \dots, a_m , todos sus ceros distintos y $\alpha_1, \alpha_2, \dots, \alpha_m$, sus órdenes de multiplicidad ($\alpha_1 + \alpha_2 + \dots + \alpha_m = N$). Entonces $P(z)$ puede expresarse en la forma

$$P(z) = A(z-a_1)^{\alpha_1} \dots (z-a_m)^{\alpha_m},$$

de donde

$$f(z) = \sqrt[n]{A(z-a_1)^{\alpha_1} \dots (z-a_m)^{\alpha_m}}. \quad (5.3:2)$$

Consideremos una curva de Jordan cerrada arbitraria γ (por ejemplo, una circunferencia) que no pase por ninguno de los puntos a_k ($k = 1, \dots, m$). Supongamos que z recorre una vez esta curva en una dirección determinada. Fijemos los valores de los argumentos para $z - a_1, \dots, z - a_m$ en algún punto z_0 de la curva γ . Sean estos valores $\varphi_1^{(0)}, \dots, \varphi_m^{(0)}$. Al recorrer el punto z la curva γ , el ángulo φ_k entre el vector $z - a_k$ y la dirección positiva del eje real variará continuamente, partiendo del valor inicial $\varphi_k^{(0)}$, y como resultado de recorrer una vez la curva γ éste volverá a tomar el valor inicial $\varphi_k^{(0)}$ (si el punto a_k estaba situado en el exterior de γ) o bien adquirirá un incremento $\pm 2\pi$ (si el punto a_k estaba situado en el

interior de γ_h *) (fig. 39). El signo $+o-$ del incremento depende solamente de la dirección del recorrido de la curva γ ; llamaremos positiva la dirección según la cual los ángulos correspondientes adquieren un incremento positivo 2π . Supongamos, para fijar ideas, que el punto z describe γ en la dirección positiva. Si ninguno de los puntos a_k está situado en el interior de γ , entonces todos los ángulos φ_k , después del recorrido, volverán a sus valores iniciales $\varphi_k^{(0)}$, y la función $f(z)$ (5.3:2) volverá a tomar su valor inicial. De aquí se deduce que ninguno de los puntos finitos ζ del plano, distintos de a_k , puede ser punto de ramificación para esta función. En efecto,

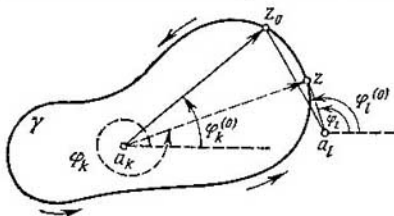


FIG. 39

para tal punto se puede señalar un entorno del mismo que no contenga ningún punto a_k ; entonces el recorrido a lo largo de cualquier curva de Jordan cerrada γ , que pertenezca a este entorno y contenga en su interior al punto ζ , conservará la rama elegida de nuestra función.

Así, pues, *ningún punto finito ζ , distinto de todos los puntos a_k , puede ser punto de ramificación para la función $f(z)$.*

Consideremos ahora un entorno tan pequeño de algún punto a_h que en él no estén contenidos otros puntos: $a_1, \dots, a_{h-1}, a_{h+1}, \dots, a_m$. Entonces, al recorrer una curva γ que pertenezca a este entorno y contenga a_h en su interior, el ángulo φ_h variará en 2π , mientras que los demás ángulos $\varphi_1, \dots, \varphi_{h-1}, \varphi_{h+1}, \dots, \varphi_m$, volverán a tomar sus valores anteriores. De aquí se deduce que, como resultado del recorrido de la curva γ , el argumento de la expresión subradical que figura en la fórmula (5.3:2) variará en $2\pi\alpha_h$, y por consiguiente, el radical (5.3:2) adquirirá el factor $\cos \frac{2\pi\alpha_h}{n} + i \operatorname{sen} \frac{2\pi\alpha_h}{n}$, que, evidentemente, es diferente de la unidad cuando, y sólo cuando, α_h no es múltiplo de n . Así, pues, cada cero a_h del polinomio $P(z)$, cuyo orden de multiplicidad α_h no sea un número entero múltiplo de n , es un punto de ramificación de la

*) Todo esto se comprueba fácilmente en los casos más simples (por ejemplo, cuando γ es una circunferencia, una elipse o un polígono), y puede demostrarse en el caso más general.

función $\sqrt[n]{P(z)}$. Para determinar el orden de este punto, supongamos que δ_k ($\delta_k < n$) es el máximo común divisor de α_k y n . Entonces, haciendo $\alpha_k = \delta_k \alpha'_k$ y $n = \delta_k v_k$ ($v_k > 1$), escribimos el binomio $\cos \frac{2\pi\alpha_k}{n} + i \operatorname{sen} \frac{2\pi\alpha_k}{n}$ en la forma $\cos \frac{2\pi\alpha'_k}{v_k} + i \operatorname{sen} \frac{2\pi\alpha'_k}{v_k}$.

Como resultado de un recorrido p -múltiple de la curva γ en una misma dirección, la función $f(z)$ adquirirá el factor $\cos \frac{2\pi\alpha'_k p}{v_k} + i \operatorname{sen} \frac{2\pi\alpha'_k p}{v_k}$ que, evidentemente, será igual a la unidad cuando, y sólo cuando, p es múltiplo de v_k . El menor valor correspondiente de p es v_k . De aquí que el orden del punto de ramificación a_k es $v_k - 1$.

Consideremos, finalmente, un entorno del punto del infinito que no contenga ningún punto a_k , y en este entorno, una curva de Jordan γ que contenga en su interior todos los puntos a_k . Entonces el exterior de γ contendrá al punto ∞ y no contendrá a ninguno de los puntos a_k . Hagamos un recorrido simple (una vuelta) de la curva γ . Todos los ángulos φ adquirirán el incremento 2π , por consiguiente, el argumento de la expresión subradical en la fórmula (5.3:2) variará en $2\pi(\alpha_1 + \alpha_2 + \dots + \alpha_m)$ y toda la función $f(z)$ adquirirá el factor

$$\cos \frac{2\pi(\alpha_1 + \dots + \alpha_m)}{n} + i \operatorname{sen} \frac{2\pi(\alpha_1 + \dots + \alpha_m)}{n} = \cos \frac{2\pi N}{n} + i \operatorname{sen} \frac{2\pi N}{n}.$$

Este será igual a la unidad o distinto de la unidad según que N sea múltiplo de n o no lo sea. En el primer caso, ∞ no será, y, en el segundo caso, será un punto de ramificación de la función $f(z)$. Si δ es el máximo común divisor de N y n ($\delta < n$) y $n = \delta v$, entonces el orden del punto del infinito, considerado como punto de ramificación, será igual a $v - 1$.

Hemos observado que, cuando α_k es múltiplo de n , el recorrido de la curva de Jordan γ , que contiene en su interior al punto a_k y no contiene a ninguno de los demás puntos a_j , no altera el valor de $f(z)$. Del mismo modo, si N es múltiplo de n , el recorrido de una curva γ que contenga en su interior a todos los puntos a_k , no altera los valores de $f(z)$.

Supongamos que, en general, a_{k_1}, \dots, a_{k_q} , es un grupo tal de puntos de ramificación, para los cuales la suma $\alpha_{k_1} + \dots + \alpha_{k_q}$ es múltiplo de n ; entonces el recorrido de cualquier curva de Jordan cerrada γ , que contenga en su interior a los puntos indicados y no contenga ningún punto a_k distinto de ellos, no puede alterar los valores de $f(z)$. Por esta razón, en cualquier recinto G que contenga solamente a aquellas curvas de Jordan cerradas que no abarcan en su interior a ningún punto de ramificación a_k , o bien, que abarcan

un grupo de puntos de ramificación, para los cuales la suma de sus números respectivos α_h es múltiplo de n , es posible separar ramas uniformes de la función $f(z)$.

Para esto es suficiente fijar un valor w_0 de la función $f(z)$ en uno de los puntos z_0 de este recinto. Entre las n imágenes $f(G)$ del recinto G en el plano w , una contendrá al punto w_0 ; supongamos que esta imagen es g_k . Entonces, la rama uniforme de la función $f(z)$ en el recinto G quedará completamente determinada si se exige que todos sus valores pertenezcan a g_k . El valor de esta rama en cualquier punto z_1 del recinto G se puede obtener también del siguiente modo: Unamos el punto z_0 con el punto z_1 mediante alguna curva continua γ_1 perteneciente al recinto G , y recorramos esta curva desde el punto z_0 hasta el punto z_1 , atendiendo a que el valor correspondiente de $f(z)$ varíe continuamente comenzando desde el valor w_0 . Entonces llegaremos al punto z_1 con uno de los n valores de $f(z)$, que designaremos con w_1 . Este valor depende solamente del valor w_0 elegido en el punto z_0 y del mismo punto z_1 y no depende de cómo se haya elegido el camino que una z_0 con z_1 , y, por consiguiente, representa una función uniforme de z_1 en el recinto G . En efecto, si γ_2 es otra curva que una z_0 y z_1 en el recinto G , entonces al recorrer la curva cerrada γ , formada por γ_1 y γ_2 , obtendremos primero, moviéndonos desde z_0 hasta z_1 a lo largo de γ_1 , el valor w_1 en el punto z_1 , y después, moviéndonos a lo largo de γ_2 desde z_1 hasta z_0 , tendremos que volver al valor inicial w_0 (puesto que, según la condición el recorrido por una curva cerrada en el recinto G no puede conducir a la variación de los valores de la función $f(z)$). De aquí se deduce que, moviéndonos a lo largo de γ_2 desde z_0 hacia z_1 , obtendremos en el punto z_1 al mismo valor w_1 que se obtenía al moverse a lo largo de γ_1 .

Aclaremos lo dicho con ejemplos:

1. $w = \sqrt{(1-z^2)(1-k^2z^2)}$, donde $0 < k < 1$. Esta es una función biforme con cuatro puntos de ramificación: $\pm 1, \pm \frac{1}{k}$. Aquí $N = 4$ es múltiplo de $n = 2$, por lo cual ∞ no es un punto de ramificación.

Como todos los números α_h son iguales a 1 ($\pm 1, \pm \frac{1}{k}$ son ceros simples de la expresión subradical), el recorrido a lo largo de cualquier curva cerrada γ que contenga en su interior solamente a dos puntos de ramificación, no altera los valores de la función. Por esta razón, es posible separar sus ramas uniformes, por ejemplo, en el recinto G cuya frontera consta de los dos segmentos: $-\frac{1}{k} \leq x \leq -1$ y $1 \leq x \leq \frac{1}{k}$, o en el recinto G' cuya frontera consta

de los segmentos: $-1 \leq x \leq 1$ y el segmento infinito del eje real que une los puntos $-\frac{1}{k}$ y $\frac{1}{k}$ mediante el punto del infinito (fig. 40).

En el primero de ellos, las ramas $f_1(z)$ y $f_2(z)$ pueden distinguirse por el valor que toman éstas en el origen de coordenadas. Por ejemplo, $f_1(0) = 1$ y $f_2(0) = -1$.

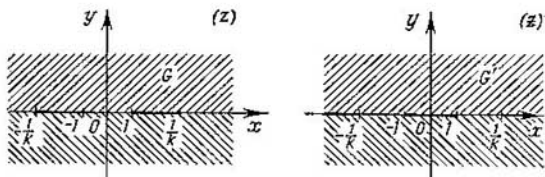


FIG. 40

2. $w = \sqrt{4z^3 - g_2z - g_3}$, donde g_2 y g_3 son unos números complejos que satisfacen a la condición $g_2^3 - 27g_3^2 \neq 0$, lo cual significa que el discriminante del polinomio $4z^3 - g_2z - g_3$ es diferente de cero y, por consiguiente, son distintos los ceros l_1 , l_2 y l_3 de este polinomio. Como en este ejemplo $N = 3$ no es divisible por $n = 2$, el



FIG. 41

punto ∞ también es un punto de ramificación. De nuevo el recorrido a lo largo de una curva de Jordan cerrada en torno de cualquier par de puntos de ramificación no hace variar los valores de la función. Por esto, uniendo con curvas de Jordan γ_1 y γ_2 el punto l_1 con l_2 y l_3 con ∞ , obtenemos un recinto G con la frontera compuesta por γ_1 y γ_2 , en el cual es posible separar las funciones uniformes de la función dada (fig. 41).

3. Consideremos la función inversa respecto de la función $z = \frac{1}{2} \left(w + \frac{1}{w} \right)$, es decir, $w = \varphi(z) = z + \sqrt{z^2 - 1}$. Esta es una función biforme, que posee los mismos puntos de ramificación que la función $\sqrt{z^2 - 1}$, o sea, ± 1 .

Para obtener el recinto G en el cual se pueden separar las ramas uniformes de la función considerada, unamos los puntos -1 y 1 mediante un segmento finito del eje real. Se obtiene un recinto que se transforma biunívocamente mediante la función $w = z + \sqrt{z^2 - 1}$ en cada uno de los dos recintos: el interior del círculo unidad (g_1) y su exterior (g_2) (véase el ap. 4.9). Se puede separar cualquiera de ellas fijando uno de los dos valores de w en un punto cualquiera del recinto G , por ejemplo, en el punto del infinito. Como se ve en la fórmula $z = \frac{1}{2}\left(w + \frac{1}{w}\right)$, z toma el valor ∞ cuando $w = 0$ o cuando $w = \infty$. Por esta razón, una de las ramas de la función $\varphi(z)$ se caracteriza porque, para ella, $\varphi(\infty) = 0$; ésta transforma G en el interior del círculo unidad. La otra rama se caracteriza porque, para ella, $\varphi(\infty) = \infty$; ésta transforma el recinto G en el exterior del círculo unidad.

Podríamos haber tomado en lugar del recinto G , por ejemplo, el recinto G' cuya frontera consta del segmento infinito del eje real que une los puntos -1 y 1 , o los recintos G'' y G''' cuyas fronteras son las semicircunferencias unidad superior e inferior, respectivamente.

Dejamos a cuenta del lector aclarar, basándose en los resultados del ap. 4.9, en qué género de recintos del plano w transforman las ramas uniformes respectivas de la función $\varphi(z)$ los recintos G' , G'' o G''' .

Todo el contenido del presente apartado se refería a las funciones multiformes de la forma $\sqrt[n]{P(z)}$, donde $P(z)$ es un polinomio.

El lector extenderá sin dificultad alguna los resultados obtenidos al caso más general de la función $\sqrt[n]{R(z)}$, donde $R(z)$ es una función racional arbitraria. Para hallar los puntos de ramificación de la función $\sqrt[n]{R(z)}$ habrá que considerar no sólo los ceros sino también los polos de la función racional $R(z)$.

5.4. La función inversa respecto de la función

$$z = e^w = e^u (\cos v + i \operatorname{sen} v),$$

está definida para cualquier z diferente de 0 y ∞ , y se expresa por la fórmula (véase la fórmula (3.5:2)).

$$w = \ln |z| + i \operatorname{Arg} z.$$

Esta función que, evidentemente, es multiforme e incluso de infinitas determinaciones, se llama logaritmo, y se designa por $\operatorname{Ln} z$.

Así, pues, según la definición

$$w = \operatorname{Ln} z = \ln |z| + i \operatorname{Arg} z. \quad (5.4:1)$$

Denominando valor principal (determinación principal) al valor del logaritmo igual a $\ln|z| + i \arg z$, y designándole mediante $\ln z$, para el $\text{Ln } z$ tendremos:

$$\text{Ln } z = \ln z + 2k\pi i, \tag{5.4:2}$$

donde $k = 0, \pm 1, \pm 2, \dots$

De aquí se deduce que cada número complejo, diferente de cero y del infinito, posee un conjunto infinito de logaritmos (es decir, de valores de la función logarítmica); dos valores cualesquiera de éstos se diferencian en un entero múltiplo de $2\pi i$. Si z es un número real positivo, el valor principal del logaritmo coincide con $\ln |z|$ y, por consiguiente, es aquel número real que conocíamos en el curso de análisis matemático cuando considerábamos el logaritmo como una función real de variable real. Así, obtenemos: $\ln 1 = 0$, $\ln e = 1$, $\ln 2 = 0,69314718 \dots$, etc.

Pero, además de estos valores reales, los logaritmos de los números positivos poseen también un conjunto infinito de valores imaginarios que se obtienen por la fórmula (5.4:2). Así, por ejemplo, $\text{Ln } 1 = 2k\pi i$, $\text{Ln } e = 1 + 2k\pi i$, $\text{Ln } 2 = 0,69314718 \dots + 2k\pi i$, etc.

Para los números negativos y para los números imaginarios, el valor principal del logaritmo es un número imaginario

$$\ln |z| + i \arg z \quad (\arg z \neq 0, \quad |\arg z| \leq \pi).$$

Todos los demás valores del logaritmo también son números imaginarios, y se calculan por la fórmula (5.4:2).

Por ejemplo,

$$\text{Ln}(-1) = (2k+1)\pi i, \quad \text{Ln}(-2) = 0,69314718 \dots + (2k+1)\pi i,$$

$$\begin{aligned} \text{Ln}(1-i) &= \ln \sqrt{2} + i \left(-\frac{\pi}{4} + 2k\pi \right) = \\ &= 0,34657359 \dots + (8k-1)\frac{\pi i}{4} \text{ etc.} \end{aligned}$$

Las reglas conocidas del logaritmo del producto y del cociente conservan su valor también para el logaritmo multiforme del número complejo, precisando:

$$\begin{aligned} \text{Ln}(z_1 z_2) &= \ln |z_1 z_2| + i \text{Arg}(z_1 z_2) = \\ &= \ln |z_1| + \ln |z_2| + i (\text{Arg } z_1 + \text{Arg } z_2) = \text{Ln } z_1 + \text{Ln } z_2, \end{aligned} \tag{5.4:3}$$

$$\begin{aligned} \text{Ln} \frac{z_1}{z_2} &= \ln \left| \frac{z_1}{z_2} \right| + i \text{Arg} \frac{z_1}{z_2} = \ln |z_1| - \ln |z_2| + i (\text{Arg } z_1 - \text{Arg } z_2) = \\ &= \text{Ln } z_1 - \text{Ln } z_2. \end{aligned} \tag{5.4:4}$$

Aquí z_1 y z_2 son unos números complejos arbitrarios, diferentes de 0 y ∞ . En cada una de estas igualdades el primero y segundo miem-

bros, para valores dados de z_1 y z_2 , representan conjuntos infinitos de números complejos. Las igualdades deben entenderse en el sentido de que estos conjuntos son iguales, es decir, constan de unos mismos números. El olvido de esta circunstancia puede conducir a errores.

Veamos, por ejemplo, el siguiente sofisma, perteneciente a J. Bernoulli.

Se afirma que $\text{Ln}(-z) = \text{Ln} z$ para cualquier $z \neq 0$.

Para la demostración se considera la siguiente cadena de igualdades:

- 1) $\text{Ln} [(-z)^2] = \text{Ln}(z^2)$, 2) $\text{Ln}(-z) + \text{Ln}(-z) = \text{Ln} z + \text{Ln} z$,
 3) $2 \text{Ln}(-z) = 2 \text{Ln} z$ y 4) $\text{Ln}(-z) = \text{Ln} z$.

Pero esta conclusión es errónea, puesto que

$$\text{Ln} z = \ln |z| + i \text{Arg} z = \ln |z| + i \arg z + 2k\pi i,$$

$\text{Ln}(-z) = \ln |-z| + i \text{Arg}(-z) = \ln |z| + i \arg z + (2k+1)\pi i$ y, evidentemente, ninguno de los números que son valores de $\text{Ln} z$ puede coincidir con alguno de los números que son valores de $\text{Ln}(-z)$.

El error en la demostración expuesta anteriormente fue cometido al pasar de la igualdad 2) a la igualdad 3). Claro, la primera de éstas, obtenida sobre la base de la fórmula (5.4:3), es justa. Pero la suma $\text{Ln}(-z) + \text{Ln}(-z)$ no puede sustituirse por $2 \text{Ln}(-z)$, puesto que la suma indicada se obtiene del conjunto de números $\text{Ln}(-z)$ sumando cualquiera de estos números al mismo o a otro número distinto del mismo conjunto, mientras que el conjunto $2 \text{Ln}(-z)$ se obtiene duplicando cada uno de los números $\text{Ln}(-z)$, es decir, sumando tal número solo consigo mismo. Así, pues, $\text{Ln}(-z) + \text{Ln}(-z) \neq 2 \text{Ln}(-z)$; por la misma razón

$$\text{Ln} z + \text{Ln} z \neq 2 \text{Ln} z.$$

El lector acabará de entender el sentido de esta objeción examinando el sencillo ejemplo. Designemos con A el conjunto que consta de dos números: 0 y 1. Entonces $A + A$ designa el conjunto que consta de tres números: $0 + 0 = 0$, $0 + 1 = 1$ y $1 + 1 = 2$, mientras que el conjunto $2A$ consta solamente de dos números: $2 \cdot 0 = 0$ y $2 \cdot 1 = 2$.

Obsérvese también que, haciendo en la relación (5.4:4) $z_1 = z_2 = z \neq 0$, obtenemos:

$$\text{Ln} 1 = \text{Ln} z - \text{Ln} z.$$

Esta es una relación justa; pero aquí no se puede sustituir el segundo miembro por cero, puesto que se trata del conjunto de todas las

diferencias entre los pares de valores del logaritmo de un mismo número. Este conjunto consta de todos los enteros posibles que son múltiplos del número $2\pi i$, o sea que, de acuerdo a la verdad, se tiene:

$$\operatorname{Ln} 1 = 2k\pi i \quad (k = 0, \pm 1, \pm 2, \dots).$$

Pasando a considerar las ramas uniformes del logaritmo, halleemos primero los recintos de univalencia de la función $z = e^w$, para la cual el logaritmo es la función inversa.

Como todos los valores de w , en los cuales e^w toma un valor dado z ($z \neq 0$ y $z \neq \infty$), vienen dados por la fórmula (3.5:2).

$$w = \ln |z| + i \operatorname{Arg} z$$

y estos valores se obtienen de cualquiera de ellos mediante un traslado en la magnitud $2k\pi i$ ($k = \pm 1, \pm 2, \dots$), el recinto de univalencia de la función exponencial no tiene que contener ningún par de puntos de los cuales uno pueda obtenerse del otro mediante una traslación semejante.

Lo más sencillo para satisfacer a estas condiciones es tomar alguna franja rectilínea g_0 , paralela al eje real, que tenga la anchura 2π : $v_0 < v < v_0 + 2\pi$.

Además de ésta, obtenemos también un conjunto infinito de recintos de univalencia g_k : $v_0 + 2k\pi < v < v_0 + (2k + 2)\pi$ ($k = \pm 1, \pm 2, \dots$).

Evidentemente, cada punto del plano w o bien es interior para uno de los recintos g_k o bien es punto frontera común de dos recintos g_k y g_{k+1} (fig. 42). La imagen de cada franja g_k en el plano z es un mismo recinto G ; es precisamente un ángulo de magnitud 2π con el vértice en el origen de coordenadas. La frontera del recinto G es un rayo rectilíneo que parte del origen de coordenadas formando el ángulo v_0 con el eje real.

En el recinto G obtenemos un conjunto infinito numerable de ramas uniformes distintas de la función $\operatorname{Ln} z$. Cada una de estas ramas $\operatorname{Ln}_k z$ se caracteriza completamente en que sus valores tienen que pertenecer a una franja determinada g_k . Por cierto, es suficiente fijar el valor w_0 de la función $\operatorname{Ln} z$ en un punto z_0 del recinto G , puesto que entre todos los recintos g_k uno, y sólo uno de los recintos g_{k_0} contendrá al punto w_0 .

Consideremos una rama cualquiera del logaritmo:

$$\operatorname{Ln}_k z = \ln |z| + i \operatorname{Arg}_k z,$$

donde $\operatorname{Arg}_k z$ es el valor del argumento que satisface a la condición

$$v_0 + 2k\pi < \operatorname{Arg}_k z < v_0 + (2k + 2)\pi.$$

(Precisamente esta condición significa que los valores $\text{Ln}_h z$ pertenecen a la franja g_k).

Como la función $w = \text{Ln}_h z$ realiza una transformación biunívoca y continua del recinto G en la franja g_k y su función inversa $z = e^w$

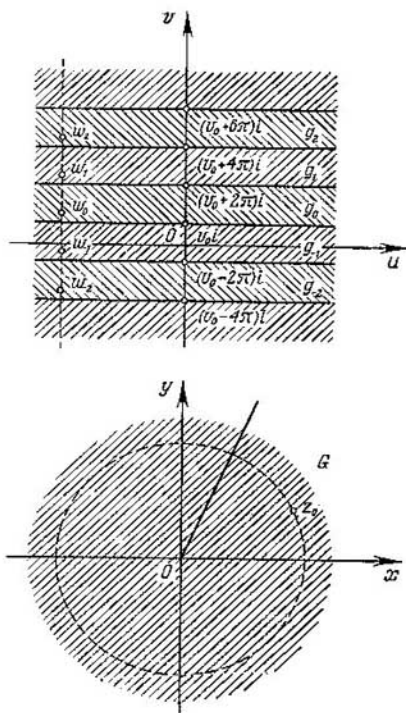


FIG. 42

posee derivada diferente de cero en todos los puntos del recinto g_k , según la regla de derivación de las funciones inversas la función $\text{Ln}_h z$ también posee derivada, la cual se calcula por la fórmula

$$(\text{Ln}_h z)' = \frac{1}{(e^w)'} = \frac{1}{e^w} = \frac{1}{z} .$$

Los puntos de ramificación de la función $\text{Ln } z$ son el cero y el punto del infinito. En efecto, cuando z describe una vez alguna circunferencia con el centro en el origen de coordenadas (de radio arbitrariamente pequeño o arbitrariamente grande), el valor $\text{Arg } z$, variando continuamente, partiendo de cierto valor inicial $\text{Arg}_h z_0$, al volver al punto inicial, obtiene un incremento $\pm 2\pi$ (en dependencia del sentido del recorrido de la circunferencia) y, por consiguiente, la rama

$$\text{Ln}_h z = \ln |z| + i \text{Arg}_h z$$

se convierte en otra rama:

$$\text{Ln}_{h\pm 1} z = \ln |z| + i (\text{Arg}_h z \pm 2\pi) = \ln |z| + i \text{Arg}_{h\pm 1} z.$$

Evidentemente, describiendo la circunferencia cuantas veces se quiera en una misma dirección (por ejemplo, en la dirección positiva), obtendremos cada vez nuevas ramas:

$$\text{Ln}_{h+1} z, \text{Ln}_{h+2} z, \text{Ln}_{h+3} z, \dots$$

y, por consiguiente, nunca volveremos a la rama inicial $\text{Ln}_h z$. Por esta razón, los puntos de ramificación 0 y ∞ se llaman aquí puntos de ramificación de orden infinito, o puntos de ramificación logarítmicos.

Trazando en el plano z , por ejemplo, una curva de Jordan Γ' que una el punto $z = 0$ con el punto $z = \infty$, se obtiene un recinto de un tipo más general que G , en el cual es posible separar las ramas uniformes de $\text{Ln } z$. Las imágenes de esta curva en el plano w , en la transformación $w = \text{Ln } z$, serán las curvas de Jordan γ'_k ($k = 0, \pm 1, \pm 2, \dots$), que dividen el plano w en un conjunto infinito de franjas curvilíneas g'_k cuyas fronteras están formadas por los pares de curvas γ'_k y γ'_{k+1} . . . En el recinto G' cuya frontera es la curva Γ' , obtenemos un conjunto numerable de ramas uniformes de $\text{Ln } z$: $(\text{Ln } z)_h$, cada una de las cuales transforma G' biunívocamente en el recinto correspondiente g'_h .

Cualquiera de las funciones $(\text{Ln } z)_h$ se puede obtener de cualquiera otra de las funciones $(\text{Ln } z)_m$ agregando un entero respectivo, múltiplo de $2\pi i$.

Para la derivada de $(\text{Ln } z)_h$ se tiene la fórmula anterior:

$$(\text{Ln } z)'_h = \frac{1}{z}.$$

La independencia del último resultado de la elección de la rama de $\text{Ln } z$ permite escribir en general:

$$(\text{Ln } z)' = \frac{1}{z},$$

entendiendo el primer miembro como la derivada de una rama uniforme arbitraria de $\text{Ln } z$, elegida en el recinto que contiene al punto dado z .

5.5. En este apartado consideraremos la función potencial general y la función exponencial general, y también el logaritmo de base arbitraria. Previamente tenemos que definir el concepto de potencia de exponente arbitrario.

Sea a un número arbitrario diferente de cero. Si n es un número entero, entonces, como ya sabemos, a^n se define por la relación

$$a^n = |a|^n [\cos (n \text{ Arg } a) + i \text{ sen } (n \text{ Arg } a)].$$

Si n es un número racional arbitrario, igual a $\frac{p}{q}$, donde q es un número natural y la fracción $\frac{p}{q}$ es irreducible, entonces $a^{\frac{p}{q}}$ posee q valores distintos, obtenidos por la fórmula (véase el ap. 2.3 del primer cap.):

$$a^{\frac{p}{q}} = |a|^{\frac{p}{q}} \left[\cos \left(\frac{p}{q} \text{ Arg } a \right) + i \text{ sen } \left(\frac{p}{q} \text{ Arg } a \right) \right].$$

Esta fórmula abarca también el caso de un exponente entero.

Supongamos ahora que ρ es un número real irracional. Fijemos un valor arbitrario $\varphi = \text{Arg } a$ y consideremos una sucesión de números racionales r_n convergente hacia ρ . La sucesión de valores determinados de a^{r_n} :

$$|a|^{r_n} [\cos (r_n \text{ Arg } a) + i \text{ sen } (r_n \text{ Arg } a)],$$

converge, evidentemente, hacia el límite

$$|a|^\rho [\cos (\rho \text{ Arg } a) + i \text{ sen } (\rho \text{ Arg } a)],$$

que tomaremos por uno de los valores de a^ρ . Para obtener todos los valores de la potencia a^ρ de exponente irracional ρ , asignamos a $\text{Arg } a$ en la expresión obtenida todos los valores posibles.

Como dos valores distintos de $\rho \text{ Arg } a$ se diferencian en un número de la forma $2k\rho\pi$, que no puede ser entero múltiplo de 2π (k es un número entero, distinto de cero, y ρ es un número irracional), todos los valores de a^ρ , correspondientes a distintos valores de $\text{Arg } a$, son diferentes entre sí.

Así, pues, hemos definido la potencia a^α para el caso en que α es un número real arbitrario. Todos los valores de la potencia están comprendidos en la fórmula

$$a^\alpha = |a|^\alpha [\cos (\alpha \text{ Arg } a) + i \text{ sen } (\alpha \text{ Arg } a)]. \quad (5.5:1)$$

Si α es un número entero, se obtiene un valor de la potencia; si α es un número racional que se expresa por la fracción irreducible $\frac{p}{q}$,

se obtienen unos cuantos valores, precisamente q valores distintos, y, finalmente, si α es un número irracional, resulta un conjunto infinito (numerable) de valores distintos.

Para definir el concepto de potencia a^α en el caso de un exponente complejo cualquiera α , observemos que la fórmula (5.5:1) puede expresarse en la forma

$$\begin{aligned} a^\alpha &= e^{\alpha \ln |a|} [\cos (\alpha \operatorname{Arg} a) + i \operatorname{sen} (\alpha \operatorname{Arg} a)] = \\ &= \exp (\alpha \ln |a| + i \alpha \operatorname{Arg} a) = \exp (\alpha \operatorname{Ln} a). \end{aligned}$$

El segundo miembro de esta fórmula tiene sentido no sólo cuando α es real, sino que para cualquier α complejo. De acuerdo a esto, hagamos por definición para cualquier α complejo:

$$a^\alpha = \exp (\alpha \operatorname{Ln} a). \quad (5.5:2)$$

Evidentemente, si α es imaginario, todos los valores de a^α que corresponden a distintos valores de $\operatorname{Ln} a$, o lo que es lo mismo, que corresponden a distintos valores de $\operatorname{Arg} a$, también son distintos entre sí. En efecto, dos valores distintos de $\alpha \operatorname{Ln} a$ se diferencian en un número de la forma $2\pi i \alpha$, que siendo α imaginario no puede ser entero múltiplo de $2\pi i$.

Comparando las expresiones

$$\begin{aligned} a^\alpha a^\beta &= \exp (\alpha \operatorname{Ln} a) \cdot \exp (\beta \operatorname{Ln} a) = \exp (\alpha \operatorname{Ln} a + \beta \operatorname{Ln} a) = \\ &= \exp [(\alpha + \beta) \operatorname{Ln} a + 2\pi i (k\alpha + l\beta)] \end{aligned}$$

y

$$a^{\alpha+\beta} = \exp [(\alpha + \beta) \operatorname{Ln} a] = \exp [(\alpha + \beta) \operatorname{Ln} a + 2\pi i m (\alpha + \beta)],$$

donde k , l y m son números enteros arbitrarios, sacamos la conclusión de que entre los valores del producto $a^\alpha a^\beta$ están comprendidos todos los valores de la potencia $a^{\alpha+\beta}$, pero, en el caso general, existen también otros valores. Para que se cumpla la igualdad

$$a^\alpha a^\beta = a^{\alpha+\beta},$$

es necesario y suficiente que para cualesquiera enteros k y l existan unos enteros m y n , tales que

$$k\alpha + l\beta = m(\alpha + \beta) + n.$$

Puede no cumplirse esta condición incluso cuando α y β sean números racionales; por cierto, es suficiente la condición $\alpha = m(\alpha + \beta) + n$ (por ejemplo, $\beta = i$, $\alpha = \frac{n+mi}{1-m}$).

Análogamente, comparando las expresiones

$$\begin{aligned} (a^\alpha)^\beta &= [\exp (\alpha \operatorname{Ln} a)]^\beta = \exp [\beta (\alpha \operatorname{Ln} a + 2l\pi i)] = \\ &= \exp [\alpha\beta \operatorname{Ln} a + 2\pi i\beta (k\alpha + l)] \end{aligned}$$

y

$$a^{\alpha\beta} = \exp(\alpha\beta \operatorname{Ln} a) = \exp(\alpha\beta \ln a + 2m\pi i\alpha\beta),$$

donde k , l y m son números enteros arbitrarios, sacamos la conclusión de que entre los valores $(a^a)^\beta$ están contenidos todos los valores de $a^{\alpha\beta}$, pero en el caso general hay también otros valores. Para que se cumpla la igualdad

$$(a^a)^\beta = a^{\alpha\beta},$$

es necesario y suficiente que para cualesquiera enteros k y l existan unos enteros m y n tales, que

$$k\alpha\beta + l\beta = m\alpha\beta + n.$$

Puede no cumplirse esta condición incluso cuando α y β son números racionales; por cierto, es suficiente la condición $\beta = m\alpha\beta + n$ (por ejemplo, $\beta = i$, $\alpha = \frac{1-ni}{m}$).

Aclaremos con ejemplos la definición de potencia:

- 1) $1^{\sqrt{2}} = \exp(\sqrt{2} \operatorname{Ln} 1) = \exp(2k\pi i \sqrt{2}) =$
 $= \cos(2k\pi \sqrt{2}) + i \operatorname{sen}(2k\pi \sqrt{2}) \quad (k=0, \pm 1, \pm 2, \dots);$
- 2) $e^z = \exp(z \operatorname{Ln} e) = \exp[z(1 + 2k\pi i)] = \exp z \exp(2k\pi iz)$
 $(k=0, \pm 1, \pm 2, \dots).$

Vemos, pues, que solamente uno de los valores de la potencia e^z coincide con $\exp z$. Los otros valores son: $\exp z \exp 2\pi iz$, $\exp z \exp(-2\pi iz)$, etc. En particular, solamente uno de los valores de e^x (x es un número real) coincide con el número real positivo $\exp x$. Los otros valores son: $\exp x \exp 2\pi ix$, $\exp x \exp(-2\pi ix)$, ... Habrá un número finito de valores distintos si x es racional y un conjunto infinito si x es irracional.

A pesar de esto, en nuestro curso utilizamos el símbolo e^z así como estamos acostumbrados a entenderlo en el curso de análisis, donde éste coincide con $\exp z$. Este uso del símbolo multiforme es completamente análogo a como ordinariamente se entiende el símbolo $\sqrt[n]{a}$ en el análisis (a es un número real positivo), es decir, como el único valor positivo («aritmético») del radical.

- 3) $i^i = \exp(i \operatorname{Ln} i) = \exp\left[i\left(\frac{\pi}{2}i - 2k\pi i\right)\right] =$
 $= \exp\left[(4k-1)\frac{\pi}{2}\right] = e^{(4k-1)\frac{\pi}{2}} \quad (k=0, \pm 1, \pm 2, \dots).$

Por lo tanto, todos los valores de la potencia i^i son números reales positivos, entre los cuales hay arbitrariamente pequeños y arbitrariamente grandes.

Basándose en la definición de la potencia, se pueden considerar las dos siguientes funciones multiformes:

$$z^\alpha \text{ y } a^z.$$

La primera de éstas —la potencia de exponente arbitrario— está definida, por lo general, solamente para $z \neq 0$.

Si α es un número real entero, z^α representa una función racional de forma particular. Entonces ésta está definida también para $z = 0$, donde tiene un cero (si $\alpha > 0$) o un polo (si $\alpha < 0$). Cuando α es un número real racional no entero: $\alpha = \frac{p}{q}$ (q es un número natural y la fracción $\frac{p}{q}$ es irreducible), z^α puede expresarse en la forma

$$z^\alpha = \sqrt[q]{z^p}.$$

Esta es una función multiforme; es precisamente q -forme. Para ella, los puntos $z = 0$ y $z = \infty$ son puntos de ramificación de orden $q - 1$. En cualquier recinto G , obtenido del plano ampliado después de haber trazado una curva de Jordan que una los puntos de ramificación, se pueden separar q ramas uniformes diferenciables distintas de la función. Estas ramas se convierten continuamente una en otra al recorrer el punto z por curvas que encierran al origen de coordenadas (o al punto del infinito).

Si, finalmente, α no es un número real racional (o sea, α es real irracional o es un número imaginario), la función z^α es de infinitas determinaciones. Todos sus valores están contenidos en la fórmula

$$z^\alpha = \exp(\alpha \operatorname{Ln} z).$$

Para ésta los puntos $z = 0$ y $z = \infty$ también son puntos de ramificación. Pero ahora estos puntos son de orden infinito.

En efecto, al dar una vuelta en torno del punto $z = 0$, por ejemplo, en la dirección positiva, el valor del $\operatorname{Arg} z$, variando continuamente, aumenta en 2π ; debido a esto, el valor $\alpha \operatorname{Ln} z$ varía en $2\pi i \alpha$, y el valor de la función adquiere un factor $\exp(2\pi i \alpha) \neq 1$.

Vemos ahora la función exponencial general a^z ($a \neq 0$). Esta se define para cualquier valor de z por la fórmula

$$a^z = \exp(z \operatorname{Ln} a).$$

Para obtener una rama uniforme determinada es suficiente fijar uno de los valores $\operatorname{Ln} a = b$.

Suponiendo que esto ya se ha hecho, obtenemos una función $\exp(bz)$, uniforme y diferenciable en cualquier punto. Tomando

todos los valores posibles de $\text{Ln } a$, obtenemos todas las ramas uniformes posibles de la función a^z . Como dos valores de $\text{Ln } a$ se diferencian en un sumando de la forma $2k\pi i$, las dos ramas de la función a^z se diferenciarán en un factor de la forma $\exp(2k\pi iz)$, el cual representa también una función uniforme que es diferenciable en todos los sitios y que toma el valor 1 solamente para valores reales enteros de z . No obstante, en el caso considerado, las ramas de la función uniforme se diferenciarán esencialmente por su carácter de las ramas de todas las funciones multiformes consideradas anteriormente. Precizando, en todos los ejemplos anteriores existían tales puntos del plano ampliado (puntos de ramificación), que moviéndose en torno de los mismos por curvas de Jordan cerradas y haciendo variar continuamente los valores de la función (de una rama determinada), teníamos la posibilidad de convertir continuamente una rama en otra.

Aquí queda excluida tal posibilidad, precisamente porque cada rama representa una función continua y uniforme en todo el plano finito. Independientemente de la curva cerrada por la que nos movamos al volver al punto inicial obtendremos el mismo punto inicial z (no importa que resulte otro valor del argumento), y por consiguiente, el mismo valor de la función $\exp(bz)$ (b es un valor fijado de $\text{Ln } a$).

Por lo tanto, la función multiforme a^z no tiene ningún punto de ramificación y sus ramas uniformes continuas no pueden convertirse continuamente una en otra. Todo esto permite considerar a estas ramas como funciones independientes entre sí que son uniformes y diferenciables en todos los sitios, y por consiguiente, enteras:

$$\exp(z \ln a), \quad \exp\{z(\ln a + 2\pi i)\}, \quad \exp\{z(\ln a - 2\pi i)\}, \dots$$

El hecho de que todas estas funciones enteras distintas puedan expresarse como ramas de una función a^z de infinitas determinaciones, no tiene para nosotros más importancia que, por ejemplo, el hecho de que las funciones $\text{sen } z$ y $-\text{sen } z$ puedan considerarse como ramas de la función biforme $\sqrt{1 - \cos^2 z}$, o que las funciones hiperbólicas $\text{sh } z$ y $\text{ch } z$ se consideren como dos ramas de la función biforme $\frac{1}{2}[\exp z + \sqrt{\exp(-2z)}]$. (Advertimos al lector que las funciones $\sqrt{1 - \cos^2 z}$ y $\frac{1}{2}[\exp z + \sqrt{\exp(-2z)}]$, igual que la función a^z , no poseen puntos de ramificación).

Fijando una de las ramas de la función $z = a^w = \exp(bw)$, donde b es uno de los valores de $\text{Ln } a$, podemos considerar la función inversa respecto de esta rama. Obtenemos, evidentemente:

$$w = \frac{1}{b} \text{Ln } z \quad (b = \ln a + 2k_0\pi i). \quad (5.5:3)$$

Esta función se diferencia de la función $\text{Ln } z$ solamente en el factor constante $\frac{1}{b}$. Como de las relaciones (5.5:3) se deduce que

$$z = \exp(bw) = a^{bw} \quad (\text{uno de los valores de } a^{bw}),$$

se puede considerar que w es el logaritmo de z de base a .

Así, pues, definimos el logaritmo de un número complejo arbitrario respecto de la base a (a es un número complejo, diferente de cero) mediante la fórmula

$$\text{Log}_a z = \frac{\text{Ln } z}{\text{Ln } a}, \quad (5.5:3')$$

donde en el denominador está fijado uno de los infinitos valores de $\text{Ln } a$ (un mismo valor b para todos los z).

Por lo tanto, esta definición no sólo exige que esté indicada la base a del sistema de logaritmos, sino también que esté fijado uno de los valores de $\text{Ln } a$.

Aclaremos lo dicho con ejemplos.

1) $a = e$. Si se fija el valor de $\text{Ln } e$, igual a 1 , se obtiene:

$$\text{Log}_e z = \text{Ln } z.$$

Esta es la definición ordinaria del logaritmo natural. Pero podría tomarse el valor de $\text{Ln } e$, igual, por ejemplo, a $1 + 2\pi i$. Entonces tendríamos:

$$\text{Log}_e z = \frac{\text{Ln } z}{1 + 2\pi i}.$$

El lector comprobará fácilmente que, con tal definición, entre todos los números positivos, solamente los de la forma e^k (k es un número entero) tendrían, cada uno de ellos, un valor real del logaritmo natural.

2) $a = 10$. Tomando el valor de $\text{Ln } 10$ igual a $2,302585 \dots = \frac{1}{0,43429 \dots} = \frac{1}{M}$, tendremos:

$$\text{Log}_{10} z = M \text{Ln } z = 0,43429 \dots \text{Ln } z.$$

Esta definición del logaritmo decimal de un número complejo arbitrario z ($z \neq 0$) concuerda con la definición ordinaria de los logaritmos decimales de los números reales positivos, pues, si $z = x > 0$, tomando los valores principales de los logaritmos, tendremos:

$$\lg x = 0,43429 \dots \text{Ln } x.$$

3) $a = 1$. En este caso, para la definición del valor principal del logaritmo de base 1 no se puede utilizar el valor principal del $\text{Ln } 1$, igual a cero. Tomemos el valor de $\text{Ln } 1$, igual a $2\pi i$. Entonces,

según la definición, tendremos:

$$\text{Log}_1 z = \frac{\text{Ln } z}{2\pi i} = \frac{1}{2\pi} \text{Arg } z - \frac{i}{2\pi} \ln |z|.$$

De aquí se deduce que todos los valores de $\text{Log}_1 z$, son reales si $|z| = 1$, y son imaginarios si $|z| \neq 1$. Por consiguiente, poseen logaritmos reales de base 1 aquellos números, y sólo aquellos, que se representan por puntos de la circunferencia unidad. Para estos números los valores del logaritmo coinciden con los valores de sus argumentos (medidos en fracciones de 2π). En resumen, $\frac{1}{2\pi} \text{Arg } z$ coincide con el logaritmo de z de base 1 para los números complejos cuyos módulos son iguales a uno.

5.6. En este apartado nos detendremos en las funciones trigonométricas inversas $\text{Arc cos } z$ y $\text{Arc tg } z$. La función $w = \text{Arc cos } z$ se define mediante la ecuación

$$z = \cos w. \quad (5.6:1)$$

Sustituyendo $\cos w$ mediante $\frac{\exp(iw) + \exp(-iw)}{2}$ y poniendo, para abreviar, $\exp(iw) = t$, escribimos la ecuación (5.6:1) en la forma

$$z = \frac{t + t^{-1}}{2}, \quad (5.6:2)$$

de donde

$$t^2 - 2zt + 1 = 0 \quad (5.6:3)$$

y

$$t = z + \sqrt{z^2 - 1}. \quad (5.6:4)$$

Como ambas raíces de la ecuación cuadrada (5.6:3) son diferentes de cero (su producto es igual a 1), la ecuación

$$\exp(iw) = t$$

posee raíces (respecto de la incógnita w). Obtenemos:

$$w = \frac{1}{i} \text{Ln } t = \frac{1}{i} \text{Ln } (z + \sqrt{z^2 - 1}). \quad (5.6:5)$$

En resumen, la función multiforme $w = \text{Arc cos } z$ se expresa mediante el logaritmo y la raíz cuadrada

$$w = \text{Arc cos } z = \frac{1}{i} \text{Ln } (z + \sqrt{z^2 - 1}). \quad (5.6:6)$$

Sus puntos de ramificación son, ante todo, $z = \pm 1$. En efecto, al dar una vuelta el punto z a lo largo de alguna curva de Jordan cerrada que encierre en su interior solamente uno de estos puntos, uno

de los valores de $\sqrt{z^2 - 1}$ se sustituye por el otro, el cual se diferencia del primero en el signo. El valor $z + \sqrt{z^2 - 1} = t$, que es la raíz de la ecuación cuadrática (5.6:3), se sustituye por la otra raíz de la misma ecuación, igual a $\frac{1}{t}$ (ya se indicó que el producto de ambas raíces es igual a 1). Por consiguiente, de los valores $w = \frac{1}{i} \text{Ln } t$ pasamos a los valores $\frac{1}{i} \text{Ln } \frac{1}{t}$, que son distintos de los iniciales, si $t \neq \frac{1}{t}$. Pero t puede ser igual a $\frac{1}{t}$ sólo cuando $t = \pm 1$; como se ve en (5.6:3) o en (5.6:4), esto es posible solamente cuando $z = \pm 1$. Como las curvas de Jordan recorridas no pasan por los puntos $z = \pm 1$, el caso indicado no puede aparecer y verdaderamente se puede afirmar que, como resultado de los recorridos señalados varían los valores de $w = \text{Arc cos } z$. Por lo tanto, los puntos ± 1 son puntos de ramificación para $\text{Arc cos } z$. Estos deben su existencia a la presencia del radical cuadrático en la fórmula (5.6:6). Mas la fórmula (5.6:6) posee la forma $w = \frac{1}{i} \text{Ln } t$ ($t = z + \sqrt{z^2 + 1}$), y podemos esperar también puntos de ramificación correspondientes a los dos puntos de ramificación de $\text{Ln } t$: $t = 0$ y $t = \infty$.

Ya se sabe por el ap. 4.9 de este capítulo, que a cada recorrido de una vuelta del punto t , a lo largo de una circunferencia con centro en el origen de coordenadas, corresponde en el plano z un recorrido de una vuelta del punto z a lo largo de una elipse con los focos ± 1 y viceversa.

En resumen, a un recorrido de una vuelta del punto z , a lo largo de una elipse con los focos ± 1 , corresponde una variación del $\text{Arg } t$ en $\pm 2\pi$ y una variación de $\frac{1}{i} \text{Ln } t$ en $\pm 2\pi$. Como tal elipse puede pertenecer a cualquier entorno del punto $z = \infty$ previamente asignado, éste es un punto de ramificación de $\text{Arc cos } z$ de orden infinito.

Claro, el recorrido de cualquier elipse con los focos en los puntos ± 1 se puede considerar también como un recorrido alrededor del punto $z = 0$. Pero ninguna de estas elipses puede estar situada completamente en el entorno $|z| \leq \rho$, si $\rho \leq 1$.

Demostremos que ni el punto $z = 0$ ni, en general, ningún punto z_0 del plano ampliado, a excepción de los indicados anteriormente ($z = \pm 1$ y $z = \infty$), puede ser punto de ramificación para $\text{Arc cos } z$. En efecto, la fórmula (5.6:4) para $z = z_0$ nos da dos valores distintos: t'_0 y t''_0 , diferentes de 0, ± 1 y ∞ , que satisfacen a la condición $t'_0 t''_0 = 1$. Se pueden tomar unos entornos U' y U'' de estos puntos tan pequeños que no contengan a los puntos 0 y ∞ y que ninguno de ellos contengan dos puntos t_1, t_2 que cumplan la condición $t_1 t_2 = 1$.

En efecto, si el punto t_0 no está situado en la circunferencia unidad, por ejemplo $|t_0'| < 1$, entonces t_0'' tampoco estará situado en esta circunferencia y $|t_0''| > 1$. En este caso, es suficiente tomar los entornos de modo que uno de ellos (U') esté situado en el interior de la circunferencia unidad, y el otro (U''), fuera de la circunferencia (fig. 43). Si t_0 está situado en la circunferencia unidad y está, por ejemplo, en el semiplano superior, entonces $t_0'' = \frac{1}{t_0'}$, estará situado

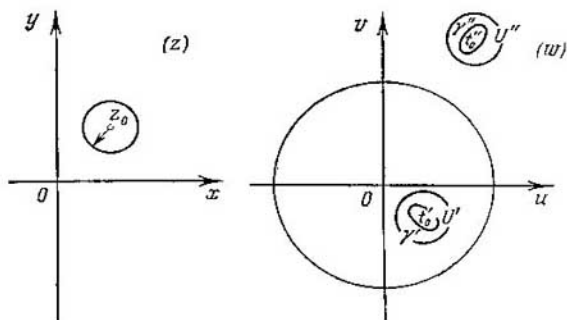


FIG. 43

también en la circunferencia unidad, y además, en el semiplano inferior. En este caso, es suficiente tomar un entorno (U') en el semiplano superior y otro (U''), en el semiplano inferior. En cada uno de los entornos U' y U'' , la función $z = \frac{1}{2} \left(t + \frac{1}{t} \right)$ será univalente (ésta toma un mismo valor solamente en los pares de puntos que están ligados por la relación $t' \cdot t'' = 1$) y, por consiguiente, transformará biunívocamente U' y U'' en unos recintos g' y g'' del plano z que contienen en sus interiores al punto z_0 (la imagen de los centros t_0' y t_0'' de los entornos U' y U'').

Tomemos las circunferencias $|z - z_0| = \rho$ con centro en z_0 tan pequeñas, que estén contenidas tanto en el recinto g' como en el recinto g'' . Evidentemente, todas las circunferencias de radio suficientemente pequeño satisfacen a estas condiciones. Entonces, debido a la transformación (5.6:2), a cada una de estas circunferencias corresponderán unas curvas de Jordan cerradas γ' y γ'' situadas cada una de ellas en uno de los entornos U' y U'' , respectivamente. Ninguna de estas curvas contendrá en su interior al punto O . Por esto, cuando z recorre la circunferencia $|z - z_0| = \rho$, t recorrerá o la curva γ' (que corresponde a una de las ramas de la función (5.6:4)), o la curva γ'' (que corresponde a la otra rama de la función (5.6:4)).

Si antes del recorrido se fija un valor de $\text{Arg } t$ en algún punto de γ' (o de γ''), entonces, después del recorrido éste, variando continuamente, volverá a tomar el valor anterior (debido, precisamente, a que ni γ' ni γ'' contienen en su interior al punto $t = 0$). Por esta razón, como resultado del recorrido volverá también a tomar su valor inicial el valor de $\frac{1}{i} \text{Ln } t = \text{Arccos } z$.

En resumen, un punto z_0 , distinto de ± 1 y ∞ , no puede ser un punto de ramificación para $\text{Arccos } z$.

Para obtener algún recinto del plano z en el cual sea posible separar ramas uniformes continuas de $\text{Arccos } z$, hay que unir entre sí los puntos de ramificación de esta función mediante curvas de Jordan. Tomemos, por ejemplo, el segmento infinito Δ del eje real que une los puntos -1 y $+1$ mediante el punto del infinito. Este segmento Δ es la frontera de un recinto G .

De lo que ya se sabe de la función (5.6:2), se deduce que esta función transforma biunívocamente en el recinto G tanto el semiplano superior t como el inferior. A su vez, la función $w = \frac{1}{i} \text{Ln } t$ transforma cada uno de éstos en franjas del plano w paralelas al eje y , de anchura π ; precisamente, el semiplano superior se transforma en las franjas g_{2k-1} :

$$(2k-1)\pi < u < 2k\pi \quad (k=0, \pm 1, \pm 2, \dots)$$

y el semiplano inferior, en las franjas g_{2k} :

$$2k\pi < u < (2k+1)\pi.$$

En resumen, las imágenes del recinto G en el plano w son las franjas g_n indicadas.

Para fijar alguna rama uniforme de $\text{Arccos } z$ en el recinto G , es suficiente indicar a qué franja g_n pertenecen sus valores. Así, obtenemos las ramas: $\text{Arccos}_0 z$, $\text{Arccos}_1 z$, $\text{Arccos}_{-1} z$, ... Por cierto, es suficiente fijar el valor de $\text{Arccos } z$ en un punto cualquiera del recinto G , por ejemplo, en el origen de coordenadas. Entonces, la franja g_n , donde caiga este valor, determinará toda la rama de $\text{Arccos } z$.

Claro, se pueden definir también las ramas uniformes de $\text{Arccos } z$ en muchos otros recintos del plano z . Señalemos, por ejemplo, el recinto G' cuya frontera consta del segmento finito δ del eje real que une los puntos -1 y $+1$, y de la parte positiva del eje imaginario, o bien, el recinto G'' cuya frontera consta del mismo segmento δ y de la parte negativa del eje imaginario.

Proponemos al lector averiguar a qué recintos del plano w transforman G' y G'' las ramas uniformes de la función $\text{Arccos } z$ que corresponden a estos últimos recintos.

Consideremos ahora la función $w = \text{Arctg } z$, inversa de la función $z = \text{tg } w$. Expresando $\text{tg } w$ mediante $\text{sen } w$ y $\text{cos } w$, y luego mediante la función exponencial, obtenemos:

$$z = \frac{1}{i} \frac{e^{iw} - e^{-iw}}{e^{iw} + e^{-iw}} = \frac{1}{i} \frac{e^{2iw} - 1}{e^{2iw} + 1}, \quad (5.6:7)$$

o bien, haciendo $e^{2iw} = \tau$:

$$z = \frac{1}{i} \frac{\tau - 1}{\tau + 1},$$

de donde

$$\tau = \frac{1 + iz}{1 - iz},$$

y, finalmente,

$$w = \frac{1}{2i} \text{Ln} \frac{1 + iz}{1 - iz}.$$

Así, pues, $\text{Arctg } z$ se expresa mediante el logaritmo de la función homográfica

$$w = \text{Arctg } z = \frac{1}{2i} \text{Ln} \frac{1 + iz}{1 - iz}. \quad (5.6:8)$$

Proponemos al lector verificar que $\text{Arctg } z$ posee solamente dos puntos de ramificación: $\pm i$. Los recintos más simples en los cuales se pueden separar las ramas uniformes de $\text{Arctg } z$, son: el recinto D cuya frontera es el segmento infinito Δ del eje imaginario que une los puntos de ramificación $-i$ y $+i$, y el recinto d cuya frontera es el segmento finito δ que une los mismos puntos.

El lector verificará luego que las ramas uniformes de $\text{Arctg } z$ en el recinto D transforman este recinto biunívoca y conformemente en las franjas de anchura π , paralelas al eje imaginario:

$$k\pi - \frac{\pi}{2} < u < k\pi + \frac{\pi}{2} \quad (k = 0, \pm 1, \dots),$$

y las ramas uniformes de esta función en el recinto d transforman d en las franjas semejantes $k\pi < u < (k + 1)\pi$.

5.7. A continuación necesitaremos la transformación realizada por la función $w = z + \text{Ln } z$, o más exactamente, por su rama uniforme en el semiplano superior

$$w = z + \text{Ln } z. \quad (5.7:1)$$

Haciendo $z = re^{i\theta}$, obtenemos:

$$u = r \cos \theta + \text{Ln } r, \quad v = r \text{sen } \theta + \theta \quad (0 < \theta < \pi). \quad (5.7:2)$$

Consideremos las imágenes de los rayos $\theta = \text{const.}$ Las ecuaciones (5.7:2) determinan estas imágenes, donde r desempeña el papel de parámetro, que varía desde 0 hasta ∞ .

Si $\theta = 0$, entonces $u = r + \ln r$, $v = 0$ y como $r + \ln r$ crece desde $-\infty$ hasta ∞ , cuando r varía desde 0 hasta ∞ , la imagen del rayo $\theta = 0$ es todo el eje real en el plano w . Si $\theta = \pi$, entonces $u = -r + \ln r$, $v = \pi$, y resulta una semirrecta paralela al eje real $-\infty < u \leq -1$, $v = \pi$, doblemente recorrida por el punto w cuando z recorre una sola vez el rayo $\theta = \pi$. Sea, finalmente, $0 < \theta < \pi$. La segunda de las ecuaciones (5.7:2) muestra que v crece

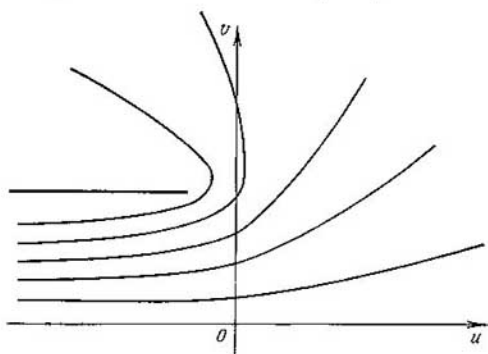


FIG. 44

desde el valor θ hasta ∞ , cuando r varía desde 0 hasta ∞ . Expresando r mediante v y θ de esta ecuación y poniéndolo en la primera de las ecuaciones (5.7:2), obtenemos la ecuación de la imagen del rayo $\theta = \text{const}$ en la forma siguiente:

$$u = (v - \theta) \cotg \theta + \ln \frac{v - \theta}{\text{sen } \theta}, \quad 0 < v < \infty. \quad (5.7:3)$$

El lector se cerciorará fácilmente de que la ecuación (5.7:3) representa una curva, que posee la asíntota $v = \theta$, que toda la curva está situada por encima de la asíntota. Si $0 < \theta \leq \frac{\pi}{2}$, entonces u crece desde $-\infty$ hasta ∞ , cuando v crece desde θ hasta ∞ , y si $\frac{\pi}{2} < \theta < \pi$, entonces u crece desde $-\infty$ hasta el valor máximo de $\ln \left(-\frac{1}{\cos \theta} \right) - 1$, que se alcanza para $v = \theta - \text{tg } \theta$, y después decrece hasta $-\infty$. En todos los casos estas curvas tienen dirigida la convexidad hacia la derecha (fig. 44).

Consideremos a u en la ecuación (5.7:3) como función de θ para un valor fijado de $v = v_0$ ($0 < \theta < \min(v_0, \pi)$). Aplicando la

derivada observamos que u es monótona decreciente *); por esta razón, dos puntos distintos z' y z'' , situados en distintos rayos $\theta = \theta'$ y $\theta = \theta''$, no pueden tener una misma imagen $w_0 = u_0 + iv_0$. Pero dos puntos z' y z'' , situados en un mismo rayo, tampoco pueden tener imágenes iguales; en efecto, como ya se vio, el punto $w = u + iv$ recorre de un modo monótono la curva (5.7:3), que carece de puntos múltiples, cuando $r = |z|$ crece desde 0 hasta ∞ . De aquí se deduce que la función $w = z + \ln z$ es univalente en el semiplano superior; ésta transforma conformemente el semiplano en el recinto D que se obtiene del semiplano superior w excluyendo la semirrecta $-\infty < u \leq -1, v = \pi$.

Haciendo $z = e^t$, se obtiene la función:

$$w = e^t + t, \quad (5.7.4)$$

la cual, como fácilmente verificará el lector, transforma conformemente la franja $0 < \text{Im } t < \pi$ en el mismo recinto D .

Observando que la función (5.7:4) en puntos simétricos respecto del eje real toma valores conjugados, sacamos la conclusión de que ésta transforma la franja $-\pi < \text{Im } t < 0$ en el recinto D^* , simétrico con D respecto del eje real. Como esta función transforma biunívocamente el eje real, la misma realiza una transformación biunívoca de la franja $-\pi < \text{Im } t < \pi$ en el recinto que se obtiene del plano w excluyendo dos rayos: $-\infty < \text{Re } w < -1, \text{Im } w = \pi$ y $-\infty < \text{Re } w < -1, \text{Im } w = -\pi$.

*) $\frac{\partial u}{\partial \theta} = -(v - \theta) - \frac{[1 + (v - \theta) \cotg \theta]^2}{v - \theta} < 0$ para $0 < \theta < \min(v, \pi)$.

Como muestra la fórmula (5.7:3), al crecer θ en el intervalo indicado, u decrece desde $+\infty$ hasta $-\infty$, si $v \neq \pi$, y desde $+\infty$ hasta -1 , si $v = \pi$.