

THE
LIBRARY OF THE
MUSEUM OF MODERN ART
1000 5th Avenue
New York 17, N.Y.

R621.381
P452e
V.1, ej 1

Enciclopedia de la **ELECTRONICA** **INGENIERIA Y TECNICA**

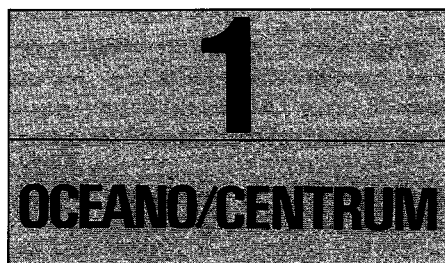
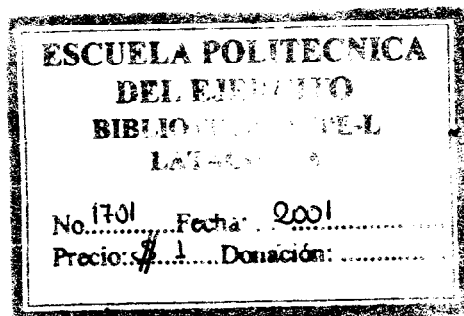
Charles Belove

**Departamento de Ingeniería Eléctrica y Computación
Florida Atlantic University
Boca Raton, Florida**

Director de la versión en español

Francisco Paniagua B., I.M.E.

Universidad Nacional Autónoma de México
Consultor Editorial de Ingeniería Eléctrica
Miembro de la U.S. Metric Association



Es una obra del
GRUPO EDITORIAL OCEANO

Presidente

José Luis Monreal

Director General

José M.^a Martí

Director General de Publicaciones

Carlos Gisbert



Versión española de la edición original de Charles Belove "Handbook of Modern Electronics and Electrical Engineering" publicada por John Wiley & Sons, Nueva York, EE.UU.

VERSIÓN ORIGINAL

Consejo Editorial: Phillip Hopkins

Lockheed Engineering and
Management Services, Co., Inc.
Houston, Texas

Milton Rosenstein

New York Institute of Technology
Old Westbury, Nueva York

Edward Nelson

New York Institute of Technology
Old Westbury, Nueva York

Stanley Shinnars

Jericho, Nueva York

VERSIÓN ESPAÑOLA

Supervisión Editorial: Pedro Basurto Samperio

Edición: Begoña Robles

Dirección Editorial: Marta Bueno

Traducción: Roberto Palacios Martínez

Licenciado en Ciencias
Universidad Autónoma de Baja
California

Juan Carlos Vega Fagoaga

Ingeniería en Sistemas

Hugo Villagómez Velázquez

Licenciado en Física y Matemáticas
Instituto Politécnico Nacional, México
Doctor en Ciencias
Universidad de París, Francia

José Rafael Blengio Pinto

Médico Cirujano
Universidad Nacional Autónoma
de México

M.^a Dolores García Díaz

Traductora Especializada en Ciencias

© MCMLXXXVI Edición Original John Wiley & Sons, Inc.

© MCMXC Edición Española Ediciones Centrum Técnicas y Científicas

Paseo de Gracia, 26 — 08007 Barcelona — España

Tel. (93) 301 01 82 — Télex 51 735 exit e — Fax (93) 317 97 01

Reservados todos los derechos. Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del copyright, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

ISBN Edición española obra completa: 84-7841-016-3

ISBN Edición española volumen 1: 84-7841-017-1

ISBN Edición original: 0-471-09754-3

Impreso en España — Printed in Spain

Depósito legal B: 2007-91 (En)

Imprime: HUROPE, S.A.

Recaredo, 2-4 Barcelona

Presentación

La **Enciclopedia de la Electrónica** representa un hito importante en la literatura científica internacional, pues expone con gran amplitud todos los adelantos técnicos de la electrónica y la electrotecnia modernas.

Escrita por un equipo de destacados especialistas de Estados Unidos de Norteamérica, aporta un completo panorama de los principios y las aplicaciones de la ciencia eléctrica. El objetivo se ha centrado en presentar una introducción práctica y una visión general de cada área.

Aunque está especialmente dirigida a ingenieros y tecnólogos, será de indudable utilidad al personal técnico de formación práctica y a personas no relacionadas directamente con el campo electrónico, como son los ingenieros de diversas especialidades, médicos, tecnólogos biomédicos, personal administrativo y estudiantes.

La **Enciclopedia de la Electrónica** está estructurada en tres grandes núcleos que siguen el flujo natural de los temas en el campo de la ingeniería en sistemas electrónicos y eléctricos.

Dado que los principios y aplicaciones electrónicas se han convertido actualmente en la base de gran parte de la electrotecnia moderna, en las tres primeras divisiones, tituladas **Matemáticas, Propiedades de los materiales y Componentes**, se proporciona todo lo relacionado con los materiales y elementos producidos por la tecnología industrial de la electrónica y la electrofísica.

70-7067 # 147

De forma previa al tratamiento de las diversas materias se ofrecen los fundamentos matemáticos indispensables para el análisis y diseño de sistemas, como son: métodos de cálculo, funciones transformadas, probabilidades y álgebra booleana.

A continuación, se presenta un estudio general de los circuitos en las divisiones tituladas: **Circuitos pasivos, Circuitos activos y Circuitos digitales.**

Por último, un extenso apartado dedicado a los sistemas electrónicos y eléctricos abarca: control automático y robótica; mediciones, modelación y confiabilidad; aplicaciones en medicina; tecnología de los sistemas de audio y vídeo; telecomunicaciones terrestres y espaciales; aplicaciones en la navegación marítima y aérea, y en enlaces vía satélite; sistemas y equipos de computación e ingeniería programática (o de software). Además, se describen la ingeniería eléctrica de potencia y las fuentes alternas de energía, incluyendo la administración energética en la industria y en diversas instalaciones.

En lo que se refiere a las normas metrológicas, en el campo de la electricidad y la electrónica prevalece absolutamente el empleo del Sistema Internacional de Unidades (SI), y sólo en casos muy especiales se utilizan unidades inglesas, que se usarán todavía por algún tiempo, no demasiado prolongado, en Estados Unidos de Norteamérica.

La **Enciclopedia de la Electrónica** supone, por todo lo expuesto, un cabal medio descriptivo de los sistemas basados en los avances de la ciencia electrofísica, en cuya aplicación constituirá con toda seguridad una ayuda inestimable.

Francisco Paniagua B., I.M.E.

Universidad Nacional Autónoma de México
Consultor Editorial de Ingeniería Eléctrica
Miembro de la U.S. Metric Association

Índice de capítulos

PRIMERA PARTE MATEMÁTICAS

1. Unidades y constantes, 3 Shin-R Lin

Unidades SI básicas, 4
Unidades SI derivadas, 5
Unidades SI complementarias, 5
Símbolos de cantidades físicas, 5
Símbolos y valores de constantes físicas, 5
Factores de conversión, 5

2. Matemáticas utilizadas en ingeniería, 11 Paul Baum y Eleanor Baum

Álgebra, 13
Funciones trigonométricas e hiperbólicas y números complejos, 14
Cálculo, 18
Matrices, 29
Vectores, 30
Ecuaciones diferenciales lineales, 32
Ecuaciones en diferencias lineales, 33
Series de potencias, 35
Serie de Fourier, 35

3. Computación, 43 Thomas C. Upson

Herramientas de cómputo disponibles, 44
Fuentes de error en computación, 44

Técnicas numéricas en computación, 45
Integración numérica, 48
Solución numérica de ecuaciones diferenciales ordinarias de primer orden, 50

4. Transformadas, 55 Manu Malek-Zavarei

Introducción, 56
Transformada de Fourier, 56
Transformada discreta de Fourier, 56
Transformada rápida de Fourier, 58
Transformada de Laplace, 60
Transformada z , 65
Funciones de transferencia, 69

5. Probabilidad, 71 Joel Halpert

Definiciones y teoremas de probabilidad, 72
Probabilidad condicional, independencia, teorema de Bayes, 73
Variables aleatorias, distribuciones de probabilidad y densidades, 74
Valor esperado, dispersión y momentos, 75
Algunas densidades y distribuciones útiles, 75
Distribuciones multivariadas, 81
Teorema del límite central, 82
Procesos de Markov, 83

6. Álgebra booleana, 85

Frank T. Boesch y Charles Suffel

Postulados de Huntington para un álgebra booleana, 86

Propiedades de un álgebra booleana, 86

Funciones de conmutación, 86

Simplificación de funciones booleanas, 88

SEGUNDA PARTE

PROPIEDADES DE LOS MATERIALES

7. Propiedades de los materiales, 93

Krishna Seshan y C. F. Bieber

Unidades y clasificación de los materiales, 94

Propiedades resistivas, 94

Capacidad de transporte de corriente de los conductores, 101

Propiedades de conductividad, 101

Propiedades relacionadas con los materiales, 105

Materiales de encapsulamiento microelectrónico, 112

8. Materiales aislantes y dieléctricos, 117

Sorab K. Ghandhi y J. Keith Nelson

Películas aislantes para dispositivos de estado sólido y circuitos integrados, 118

Materiales dieléctricos para almacenamiento de carga, 128

Materiales dieléctricos como aislantes, 133

9. Materiales magnéticos, 141

Alex Goldman

Definiciones, 142

Propiedades eléctricas y magnéticas, 142

Propiedades físicas, 146

Materiales utilizables en la práctica, 147

10. Materiales semiconductores, 153

Edward H. Nicollian

Introducción, 154

Propiedades eléctricas, térmicas y ópticas, 154

Oxidación del silicio, 168

Epitaxia, difusión e implantación de iones, 170

11. Materiales emisivos, ópticos y fotosensibles, 177

Bernard Smith y Mansoor A. Saifi

Materiales emisivos, 178

Materiales ópticos y fotosensibles, 185

**TERCERA PARTE
COMPONENTES**

12. Elementos circuitales concentrados pasivos, 199

Henry Domingos y Joseph Scaturro

Componentes resistivos, 200

Capacitores, 213

Inductores, 228

Transformadores, 232

13. Elementos circuitales concentrados activos, 237

Mark B. Barron, Henry Domingos y George H. Ebel

Diodos, rectificadores, varistores, 238

Transistores bipolares, 251

Transistores de efecto de campo, 261

Circuitos integrados, 273

Circuitos híbridos: de película gruesa y de película delgada, 300

Otros dispositivos electrónicos, 309

14. Componentes ferromagnéticos y piezoeléctricos, 321

Alex Goldman y Don Berlincourt

Componentes ferromagnéticos, 322

Componentes piezoeléctricos, 333

15. Dispositivos ópticos y pantallas, 345

L. S. Watkins, Peter Pleshko y Martin Wolf

Fuentes ópticas, 346

Fibras y conectores ópticos, 356

Pantallas, 364

Detectores ópticos, 388

Convertidores fotovoltaicos de energía solar, 399

16. Transductores eléctricos, 439

Gerald Weiss

Principios, clasificación y terminología, 441
Transductores para movimientos mecánicos de sólidos, 442
Transductores para cantidades de mecánica de fluidos, 454
Transductores para cantidades térmicas, 465
Transductores para cantidades acústicas, 470
Transductores para cantidades ópticas e infrarrojas, 471
Transductores para radiación nuclear, 471
Transductores para campos magnéticos, 474
Transductores para la instrumentación de análisis, 475
Transductores electroquímicos, 479

17. Técnicas de construcción, 485

Gustave R. Gaschnig

Circuitos impresos, 486
Métodos de interconexión, 489
Encapsulado de circuitos, 493
Encapsulado de circuitos integrados, 496
Dispositivos electromecánicos, 496

18. Dispositivos de microondas, 505

Hing-Loi A. Hung

Componentes pasivos, 506
Componentes activos, 523
Circuitos integrados de microondas, 550

CUARTA PARTE CIRCUITOS PASIVOS

19. Circuitos eléctricos lineales, 555

Jeffrey Hantgan, Yacov Shamash y Martin Gazourian

Fundamentos de los circuitos eléctricos, 556
Análisis de circuitos de CD y de CA, 569
Respuestas forzada y transitoria de circuitos, 582

20. Filtros pasivos, 593

Icarius E. Pyros

Introducción, 594
Factores de atenuación y de fase, 594

Secciones T y π , 596

Filtros de K constante, 600

Filtros derivados en M , 601

Filtro pasabajas ideal, 606

Uso de la función característica, 606

Filtros Butterworth, 607

Filtros Chebychev, 609

21. Análisis de sistemas magnéticos, 613

Gordon R. Slemon

Modelos aproximados de características $B-H$, 614

Corrientes parásitas, 623

Circuitos equivalentes de sistemas magnéticos complejos, 625

Análisis de sistemas de imán permanente, 630

22. Análisis de circuitos auxiliado por computadora, 635

Milton Rosenstein

Introducción, 636

Análisis por computadora: redes de CD, 636

Conjunto de programas para la solución de redes, 640

Revisión de la sensibilidad, 643

Otros temas, 644

QUINTA PARTE CIRCUITOS ACTIVOS

23. Amplificadores operacionales, 649

Yusuf Z. Efe

Amplificadores operacionales ideales, 650

Configuraciones básicas de los

amplificadores operacionales, 650

Aplicaciones de las configuraciones básicas, 652

Características de los amplificadores operacionales, 653

Amplificador de instrumentación, 655

Aplicaciones lineales, 655

Aplicaciones no lineales, 659

24. Generadores de ondas de forma y circuitos relacionados, 669

Yusuf Z. Efe

Osciladores, 670

Modulación, 678

Desmoduladores, 688
Convertidores de frecuencias, 693
Detectores, 701

25. Conformación de ondas de forma y circuitos de sincronización, 707

Yusuf Z. Efe

Conformación de ondas de forma, 708
Circuitos de sincronización, 722
Circuito integrado de sincronización 555, 730

26. Filtros activos, 739

Yusuf Z. Efe

Funciones transferencia de filtros, 740
Sensibilidad, 741
Dispositivos activos en filtros activos, 742
Construcción de filtros pasabajas activos, 743
Construcción de filtros pasaaltas activos, 747
Construcción de filtros pasabanda activos, 748
Filtros activos bicuadrados de estado variable, 750
Diseños de filtros activos por simulación de inductores, 752

27. Fuentes de alimentación, 755

Yusuf Z. Efe

Introducción, 756
Fuentes de alimentación de CD no reguladas, 756
Filtros para fuentes de alimentación, 758
Fuentes de alimentación de CD reguladas, 760
Familias de reguladores de circuitos integrados, 762

**SEXTA PARTE
CIRCUITOS DIGITALES**

28. Compuertas, 779

Peter Graham

Introducción, 780
Símbolos de las compuertas, 780
Variaciones de compuertas disponibles, 781

Parámetros de especificación de compuertas, 783
Compuertas de transistores bipolares, 785
Lógica de semiconductores de metal y óxido complementarios (CMOS), 792
Elección de una familia de lógica, 796
Lógica integrada en inyección (IIL), 796
Dispositivos acoplados por la carga (CCD), 797

29. Multivibradores biestables y registros, 799

Thomas W. Parsons

Multivibradores biestables, 800
Registros, 808

30. Dispositivos de funciones digitales y aritméticas, 813

Peter Graham

Introducción, 814
Contadores, 814
Cronómetros (temporizadores), 820
Multicanalizadores, 827
Decodificadores, codificadores y convertidores de código, 830
Dispositivos de aritmética binaria, 833

31. Convertidores de digital a analógico (D/A) y de analógico a digital (A/D), 839

Michael Baltrush

Convertidores D/A de resistores ponderados en escalera ponderada, 840
Especificaciones de convertidores D/A, 841
Convertidores A/D en paralelo y de aproximación sucesiva, 844
Convertidores A/D de conteo (rampa), 845
Especificaciones de los convertidores A/D, 848

32. Filtros digitales, 849

Thomas W. Parsons

Descripción, 850
Estructuras, 853
Representación numérica y efectos de cuantización, 857
Técnicas de diseño de filtros de respuesta de pulso infinita (IIR), 861

Técnicas de diseño de filtros de respuesta de pulso finita (FIR), 865

SÉPTIMA PARTE INGENIERÍA DE SISTEMAS, CONTROL AUTOMÁTICO Y MEDICIONES

33. Ingeniería de sistemas, 875

Richard G. Costello

Definición de sistema, 876
Clasificación de sistemas por función, 876
Sistemas de control, 877
Sistemas de cálculo, 877
Sistemas de comunicación, 877
Sistemas administrativos, 878
Sistemas de control con operador humano, 883
Clasificación de sistemas por modelo matemático, 893
Resumen del proceso de diseño de sistemas, 899

34. Control automático, 903

Richard G. Costello

Introducción, 904
Control automático clásico, 904
Representaciones de los sistemas clásicos de control, 910
Análisis y compensación de sistemas clásicos de control, 918
Control de datos muestreados, 974
Teoría de la transformada z de los sistemas de control, 975
Diagramas de bloques de la transformada z con elementos de retención, 988
Transformada z modificada, $F(z, m)$, 993
Control no lineal, 997
Control moderno, 1025
Técnicas de control óptimo, 1031
Cálculo de variaciones, 1035
Programación dinámica, 1039
Método del multiplicador de Lagrange, 1044

35. Aplicaciones de los sistemas de control, 1049

Roger D. Fruechte y Thomas E. Marlin

Sistemas de control de automóviles, 1050
Sistemas de control de procesos industriales, 1065

36. Robótica, 1085

J. Ronald Bailey

Introducción, 1086
Clasificación, 1086
Software, 1090
Especificaciones, 1094
Aplicaciones, 1094
Función de la robótica en la manufactura integrada con computadora, 1097
Tendencias en la industria y aplicaciones futuras, 1103

37. Mediciones de sistemas, 1105

Richard G. Costello

Introducción a la medición de los estados de un sistema, 1107
Requisitos de las mediciones, 1107
Escalación de mediciones, 1110
Convertidores de analógico a digital y de digital a analógico, 1115
Conversión entre mediciones lineales y angulares, 1117
Dispositivos para la medición analógica de la posición angular o lineal de un sistema, 1123
Resumen de mediciones analógicas de la posición angular o lineal de un sistema, 1135
Transductores de sistemas de control digital, 1136
Medición digital de la posición o el ángulo de un sistema, 1140
Medición analógica de la velocidad de un sistema, 1146
Medición analógica de la aceleración de un sistema, 1154
Medición digital de la aceleración de un sistema, 1159
Medición de peso, fuerza, momento de torsión, presión y temperatura, 1162
Detección de errores, 1171

38. Modelación y simulación, 1177

Richard G. Costello

Simulación por computadora analógica, 1178
Simulación de sistemas por computadora digital, 1186
Algoritmos de integración, 1189
Lenguajes de simulación especializados, 1194
Consideraciones sobre el problema del tiempo real, 1195

39. Confiabilidad, 1197

Amrit L. Goel y Joseph J. Naresky

Introducción, 1198

Conceptos y definiciones básicos, 1198

Distribuciones de probabilidad que se utilizan en los modelos de confiabilidad, 1200

Sistemas no reparables, 1203

Sistemas reparables, 1207

Estimación de la confiabilidad, 1211

Confiabilidad del software, 1213

Ingeniería de confiabilidad, 1214

OCTAVA PARTE

**APLICACIONES MÉDICAS
DE LA ELECTRÓNICA**

40. Introducción a la fisiología humana, 1223

Clifford Bogen

Introducción, 1224

Metabolismo, 1225

Estructura y función celulares, 1227

Bioquímica celular, 1231

Homeostasia, biorretroalimentación y equilibrio dinámico, 1233

41. Instrumentación médica, 1235

Benedict Kingsley, Daniel B. Diner,
Josefa Cubina, Jeff Tosk y Julius
Simon

Ultrasonografía, 1236

Computadoras en medicina: panorama de la tomografía computerizada e imágenes de resonancia magnética nuclear, 1245

Mediciones cardiopulmonares, 1251

Laboratorio médico automatizado, 1268

42. Vigilancia instrumental del enfermo, 1271

Steven G. Epstein, John Dawson y Louis
R. M. del Guercio

Introducción, 1272

Monitorización cardiovascular, 1272

Monitorización de la respiración, 1274

Monitorización de la presión intracraneal, 1276

Monitorización de la temperatura, 1277

Computadoras, 1278

Seguridad, 1278

43. Prótesis, 1281

Harold Z. Haut, Harry Levitt y William
Lembeck

Implantes, 1282

Procesamiento de señales en personas con deficiencias de comunicación, 1308

Prótesis eléctricas (miembro superior), 1318

NOVENA PARTE

**GRABACIÓN Y REPRODUCCIÓN
DE AUDIO Y VÍDEO**

44. Características del sonido, 1333

Robert B. Newman y Stanley A. Gelfand

Acústica, 1334

Psicoacústica, 1354

45. Grabación y reproducción de audio, 1365

Katsuaki Tsurushima, Electronic Defense
Laboratories y Albert B. Grundy

Sistema Sony de disco compacto, 1366

Grabación en cinta magnética, 1376

Altavoces, 1400

46. Grabación y reproducción de vídeo, 1409

RCA Corporation

Fundamentos de la grabación en cinta de vídeo, 1410

Reproductor de discos de vídeo, 1421

**DÉCIMA PARTE
COMUNICACIONES**

47. Conceptos sobre comunicaciones, 1435

Yeng S. Kuo, Kwei Tu, Rodger E. Ziemer
y William H. Tranter

Conceptos sobre comunicaciones modernas, 1436

Teoría del ruido, 1440

Teoría de la información, 1451

48. Técnicas de comunicación, 1461

Jack W. Seyl, George W. Raffoul, Matthew
J. Quinn, Jr., K. K. Chow, Bernard Sklar,
Peter Monsen, E. T. Dickerson

Técnicas de modulación, 1463
 Radiocomunicaciones, 1510
 Comunicaciones alámbricas, 1539
 Sistemas de comunicación por fibras
 ópticas, 1552
 Técnicas avanzadas para las
 comunicaciones digitales, 1571
 Comunicaciones por canales
 con desvanecimiento, 1601
 Análisis comparativo de las técnicas
 de comunicaciones, 1610

49. Sistemas de comunicación, 1621

Ray W. Nettleton, Jefferson F. Lindsey
 III, Surya V. Varanasi, Daniel F.
 DiFonzo

Sistemas terrestres, 1622
 Sistemas de radiodifusión, 1640
 Redes de comunicaciones, 1661
 Sistemas de comunicaciones por satélite, 1672

UNDÉCIMA PARTE SISTEMAS DE TELEMETRÍA, NAVEGACIÓN Y ATERRIZAJE

50. Principios de navegación, 1711

Emil R. Schiesser

Definiciones de funciones, 1712
 Propagación del estado, 1712
 Determinación del estado, 1713
 Observación y medición, 1713
 Sistemas de navegación, 1714
 Resumen, 1715

51. Radar, 1717

Frank R. Castella

El problema de la navegación, 1718
 Principios del radar, 1718
 Tipos de radar, 1719
 Ecuaciones de funcionamiento
 del radar, 1720
 Resolución del radar ideal, 1722
 Limitaciones del sistema en cuanto
 a exactitud de la medición, 1723
 Efectos de la trayectoria
 de propagación, 1725
 Radar Doppler y mediciones
 de velocidad, 1727
 Técnicas avanzadas (ondas milimétricas), 1729

52. Auxiliares de la navegación, 1731

Frederick B. Pogust, Stephen C. Martin
 y Thomas J. Cutler

Normas y convenios internacionales
 de navegación aérea, 1732
 Auxiliares de la navegación aérea, 1737
 Auxiliares de la navegación marítima, 1764

53. Sistemas de aterrizaje de aeronaves, 1777

Stephen C. Martin

Radioayuda para el aterrizaje
 de aeronaves, 1778
 Sistema de aterrizaje por instrumentos
 (ILS), 1778
 Radar de precisión para el aterrizaje
 (PAR), 1784
 Sistema de aterrizaje por microondas
 (MLS), 1784
 Sistemas de aterrizaje militares para fines
 especiales, 1789

54. Navegación por satélite, 1793

John H. Painter

Principios de la radionavegación por
 satélite, 1794
 Descripción de la órbita de un satélite
 (ecuación de Kepler), 1795
 Coordenadas del usuario y geodesia, 1797
 Adquisición de datos de navegación, 1798
 Efectos ionosféricos, 1799
 TRANSIT, sistema Doppler de baja
 dinámica, 1802
 GPS NAVSTAR, sistema telemétrico
 de alta dinámica, 1806

DECIMOSEGUNDA PARTE COMPUTADORAS

55. Sistemas numéricos y representación de datos, 1831

Alan B. Marcovitz

Números binarios, octales
 y hexadecimales, 1832
 Suma binaria de números positivos, 1837
 Números con signo, 1837
 Aritmética binaria, 1839
 Decimal codificado en binario (BCD), 1841
 Aritmética de precisión múltiple, 1844

- Aritmética de punto flotante, 1844
 Detección y corrección de errores, 1847
 Representaciones de datos
 alfanuméricos, 1849
- 56. Organización y arquitectura de la computadora, 1851**
 Edward J. Lancevich
- Organización general (diagrama de bloques), 1852
 Registros de la CPU, 1853
 Tipos de instrucciones, 1854
 Formatos de instrucciones y modos de direccionamiento, 1859
 Modos de operación, 1867
 Diseño de la unidad de control, 1872
 Clasificación de los sistemas de computación, 1876
 Comparación de computadoras, 1880
- 57. Estructuras de datos, 1881**
 James A. M. McHugh
- Introducción, 1882
 Listas lineales enlazadas, 1882
 Líneas de espera y pilas, 1885
 Árboles, 1888
 Tablas de asignación, 1892
- 58. Programación, 1897**
 Peter G. Anderson, John A. Biles, James R. Carbin, Warren R. Carithers, James A. Chmura, Chris Comte, Lawrence A. Coon, Mary Ann Dvovich, Henry A. Etlinger, James Hammerton, Jack Hollingsworth, Guy Johnson, Peter H. Lutz y Rayno D. Niemi
- Programación estructurada, 1898
 Panorama general del proceso de programación, 1898
 Software de sistemas, 1901
 Lenguajes de programación, 1902
 Simulación de sistemas, 1918
 Sistemas de bases de datos, 1920
- 59. Organización del procesador central, 1923**
 Melvyn M. Drossman
- Estructura y función, 1924
 Unidad de control, 1925
- Instrucciones, 1926
 Organización de ductos, 1928
 Operación de la CPU, 1928
 Unidad aritmética y lógica, 1931
 Algoritmos aritméticos, 1934
 Microprogramación, 1944
 Microprocesadores de "rebanadas" de bits, 1951
- 60. Sistemas de memoria, 1953**
 Jay Michlin
- Introducción, 1954
 Definiciones, 1954
 Hardware de memoria, 1956
 Arquitecturas de memoria, 1957
 Memoria virtual, 1961
- 61. Entrada y salida (I/O) de los sistemas de computación, 1965**
 Sam Goldwasser
- Introducción, 1966
 Tipos y ejemplos de dispositivos I/O, 1968
 Consideraciones sobre sistemas generales, 1976
 Estructuras de ductos de sistemas I/O, 1978
 Interfaz de hardware y software, 1982
 I/O controlada por el programa, 1985
 I/O activada por interrupción, 1986
 I/O de acceso directo a la memoria, 1989
 I/O mapeada por la memoria, 1990
 Técnicas de I/O avanzadas, 1993
 Resumen y tendencias en los sistemas I/O, 1995
- 62. Microcomputadoras y microprocesadores, 1997**
 Edward J. Lancevich
- Tipos y tecnologías de procesadores, 1998
 Características arquitectónicas y modos de direccionamiento, 1998
 Conjuntos de instrucciones, 2006
 Ejemplos de programación, 2013
 Principios de vinculación, 2015
 Dispositivos de memoria de microprocesadores, 2022
 Controladores de dispositivos y circuitos de interfaz, 2024

Sistemas de desarrollo de
microprocesadores y auxiliares
de diseño, 2030
Selección del microprocesador, 2034

63. Ingeniería de software, 2037
Frederic L. Swern

Introducción, 2038
Metodología de la programación, 2038
Administración de proyectos
de software, 2041

64. Graficado por computadora, 2047
Guy Johnson

Introducción, 2048
Hardware de graficado, 2048
Software de graficado, 2049
Conclusión, 2056

**65. Redes de comunicación
de computadoras, 2057**
Richard Van Slyke

Elementos de redes, 2058
Topología de redes, 2060
Modos de conmutación, 2060
Proceso de diseño, 2062
Arquitectura y protocolos, 2064

**DECIMOTERCERA PARTE
ENERGÍA DE FUENTES ALTERNATIVAS**

66. Conversión de energía, 2069
Jerald D. Parker

Conversión termoeléctrica, 2070
Conversión termoiónica, 2072
Conversión fotovoltaica, 2073
Conversión por celdas
de combustible, 2079
Conversión magnetohidrodinámica
(MHD), 2082
Conversión de energía térmica solar, 2083
Conversión de energía eólica, 2089

Conversión geotérmica, 2094
Combustibles obtenidos de desechos, 2097
Conversión nucleoelectrica, 2100

67. Sistemas de energía eléctrica, 2105
Daniel D. Lingelbach

Sistemas de distribución, 2106
Fluctuación y regulación del voltaje, 2113
Corrección del factor de potencia
y potencia reactiva (kVAr), 2117
Protección de sistemas, 2122
Dispositivos de protección contra
sobrecorriente, 2128
Dispositivos de protección contra
sobrevoltaje, 2130
Conexión a tierra del equipo y del
sistema, 2132

68. Máquinas eléctricas e iluminación, 2135
Daniel D. Lingelbach

Tipos básicos de máquinas
de CD y CA, 2136
Características de funcionamiento de las
máquinas de CA, 2136
Características de funcionamiento de las
máquinas de CD, 2139
Selección y aplicación de motores de CA
y de CD, 2142
Fundamentos de control de motores, 2144
Iluminación, 2147

69. Administración de la energía, 2159
Wayne C. Turner

¿Qué es administración
de la energía?, 2160
Diseño, inicio y operación de programas
de administración de energía, 2160
Auditorías de energía, 2162
Listas de control para proyectos
de administración de energía, 2165
Oportunidades en la administración
de energía, 2170

Índice de materias, 2175

Autores

Peter G. ANDERSON, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

J. Ronald BAILEY, División Productos de Sistemas, IBM Corporation, Boca Raton, Florida.

Michael BALTRUSH, New Jersey Institute of Technology, Wayne, New Jersey.

Mark B. BARRON, Calma Company, Sunnyvale, California.

Eleanor BAUM, Pratt Institute, Brooklyn, Nueva York.

Paul M. BAUM, Queens College, Queens, Nueva York.

Don BERLINCOURT, Presidente, Channel Products, Inc., Chesterland, Ohio.

C. F. BIEBER, Hughes Aircraft Company, Tucson, Arizona.

John A. BILES, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Frank T. BOESCH, Departamento de Ingeniería Eléctrica, Stevens Institute of Technology, Hoboken, New Jersey.

Clifford BOGEN, Profesor de Ciencias Naturales, New York Institute of Technology, Old Westbury, Nueva York.

James R. CARBIN, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Warren R. CARITHERS, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Frank R. CASTELLA, Laboratorio de Física Aplicada, The Johns Hopkins University, Baltimore, Maryland.

Chris COMTE, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Lawrence A. COON, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Richard G. COSTELLO, Profesor, The Cooper Union for the Advancement of Science and Art, Nueva York.

Josefa CUBINA, Profesora Adjunta de Ciencias Naturales, New York Institute of Technology, Old Westbury, Nueva York.

Thomas J. CUTLER, Capitán de Corbeta del Departamento de Marina, United States Naval Academy, Annapolis, Maryland.

James A. CHMURA, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

K. K. CHOW, Lockheed Missiles and Space Company, Palo Alto Research Laboratory, Palo Alto, California.

John DAWSON, M. D., Jefe de Cirugía del Departamento de Cirugía, New York Medical College, Valhalla, Nueva York.

E. T. DICKERSON, University of Houston-Clear Lake, Houston, Texas.

Daniel F. DiFONZO, COMSAT Laboratories, Clarksburg, Maryland.

Daniel B. DINER, Profesor Adjunto del Departamento de Computación, New York Institute of Technology, Old Westbury, Nueva York.

Henry DOMINGOS, Profesor del Departamento de Ingeniería Eléctrica, Clarkson College of Technology, Potsdam, Nueva York.

Melvyn M. DROSSMAN, New York Institute of Technology, Old Bethpage, Nueva York.

Mary Ann DVONCH, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

George H. EBEL, The Singer Company, Wayne, New Jersey.

Yusuf Z. EFE, Profesor, The Cooper Union for the Advancement of Science and Art, Nueva York.

ELECTRONIC DEFENSE LABORATORIES, Sylvania Electric Products, Inc., Mountainview, California.

Steven G. EPSTEIN, M. D., Cirujano, Departamento de Cirugía, New York Medical College, Valhalla, Nueva York.

Henry A. ETLINGER, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Roger D. FRUECHTE, Jefe Adjunto del Departamento de Electrónica, General Motors Research Laboratories, Warren, Michigan.

Gustave R. GASCHNIG, División de Sistemas de Misiles, Raytheon Company, Bedford, Massachusetts.

Martin GAZOURIAN, Profesor del Departamento de Ingeniería Eléctrica y Computación, Florida Atlantic University, Boca Raton, Florida.

Stanley A. GELFAND, Director del Servicio de Patología Auditiva y del Lenguaje, Veterans Administration Medical Center, East Orange, New Jersey.

Sorab K. GHANDHI, Profesor de Electrofísica del Departamento de Ingeniería Eléctrica y de Sistemas, Rensselaer Polytechnic Institute, Troy, Nueva York.

Amrit L. GOEL, Profesor de Ingeniería Industrial e Investigación de Operaciones, Syracuse University, Syracuse, Nueva York.

Alex GOLDMAN, Magnetics Division of Spang Industries, Butler, Pennsylvania.

Sam GOLDWASSER, The Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pennsylvania.

Peter GRAHAM, Departamento de Ingeniería Eléctrica, Florida Atlantic University, Boca Raton, Florida.

Albert B. GRUNDY, Audio Research Institute, Nueva York.

Louis R. M. del GUERCIO, M. D., Profesor y Presidente del Departamento de Cirugía, New York Medical College, Valhalla, Nueva York.

Joel HALPERT, Grumman, Hauppauge, Nueva York.

James HAMMERTON, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Jeffrey HANTGAN, Profesor Adjunto del Departamento de Ingeniería Eléctrica, State University of New York, Stony Brook, Nueva York.

Harold Z. HAUT, Ingeniero Biomédico, Avery Laboratories, Farmingdale, Nueva York.

Jack HOLLINGSWORTH, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Hing-Loi A. HUNG, Director del Departamento de Técnicas Avanzadas de Microondas, COMSAT Laboratories, Clarksburg, Maryland.

Guy JOHNSON, Profesor, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

Benedict KINGSLEY, National Foundation for Non-Invasive Diagnostics, Princeton, New Jersey.

Yeng S. KUO, Lockheed Engineering and Management Services Company, Inc., Houston, Texas.

Edward J. LANCEVICH, Manhattan College, Brooklyn, Nueva York.

William LEMBECK, Investigador, New York University, Post Graduate Medical School, Nueva York.

Harry LEVITT, Director del Laboratorio de Ciencias de la Comunicación, State University of New York, Nueva York.

Shin-R LIN, New York Institute of Technology, Old Westbury, Nueva York.

Jefferson F. LINDSEY III, Southern Illinois University at Carbondale, Carbondale, Illinois.

Daniel D. LINGELBACH, Profesor, School of Electrical Engineering, Oklahoma State University, Stillwater, Oklahoma.

Peter H. LUTZ, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

James A. M. McHUGH, Departamento de Ciencias de la Información y de la Computación, New Jersey Institute of Technology, Newark, New Jersey.

Manu MALEK-ZAVAREI, Bell Laboratories, Holmdel, New Jersey.

Alan B. MARCOVITZ, Florida Atlantic University, Boca Raton, Florida.

Thomas E. MARLIN, Ingeniero Asociado, Exxon Research and Engineering Company, Florham Park, New Jersey.

Stephen C. MARTIN, Palo Alto Research Laboratory, Lockheed Missiles and Space Company, Palo Alto, California.

Jay MICHLIN, Exxon Research and Engineering Company, Morristown, New Jersey.

Peter MONSEN, SIGNATRON, Inc., Lexington, Massachusetts.

J. Keith NELSON, Profesor, Center for Electric Power Engineering, Rensselaer Polytechnic Institute, Troy, Nueva York.

Ray W. NETTLETON, Departamento de Ingeniería Eléctrica y Ciencia de los Sistemas, Michigan State University, East Lansing, Michigan.

Robert B. NEWMAN, Primer Vicepresidente y Consultor, Bolt Beranek and Newman, Inc., Cambridge, Massachusetts.

Edward H. NICOLLIAN, Profesor, University of North Carolina, Charlotte, North Carolina.

Rayno D. NIEMI, School of Computer Science, Rochester Institute of Technology, Rochester, Nueva York.

John H. PAINTER, Altair Corporation, College Station, Texas.

Jerald D. PARKER, Profesor, School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, Oklahoma.

Thomas W. PARSONS, Hofstra University, Hempstead, Nueva York.

Peter PLESHKO, Director, Plasma Display Technology, IBM Corporation, Kingston, Nueva York.

Frederick B. POGUST, Eaton Corporation, AIL Division, Farmingdale, Nueva York.

Icarus E. PYROS, Profesor Adjunto del Departamento de Ingeniería Eléctrica, New York Institute of Technology, Old Westbury, Nueva York.

Matthew J. QUINN, Jr., College of Technology, University of Houston, Houston, Texas.

George W. RAFFOUL, Lockheed Engineering and Management Services Company, Inc., Houston, Texas.

RCA CORPORATION, Instrucción de Técnica Electrónica para el Consumidor, Indianapolis, Indiana.

Milton ROSENSTEIN, Profesor Adjunto del Departamento de Ingeniería Eléctrica, New York Institute of Technology, Old Westbury, Nueva York.

Mansoor A. SAIFI, Belcom Research, Murray Hill, New Jersey.

Joseph SCATURRO, Profesor Adjunto del Departamento de Ingeniería Eléctrica, Clarkson College of Technology, Potsdam, Nueva York.

Emil R. SCHIESSER, NASA Johnson Space Center, Houston, Texas.

Krishna SESHAN, Departamento de Ingeniería Metalúrgica, University of Arizona, Tucson, Arizona.

Jack W. SEYL, NASA Johnson Space Center, Houston, Texas.

Yacov SHAMASH, Profesor del Departamento de Ingeniería Eléctrica y de la Computación, Florida Atlantic University, Boca Raton, Florida.

Julius SIMON, Laboratorio de Cromatología, Flushing, Nueva York.

Bernard SKLAR, Aerospace Corporation, Los Angeles, California.

Gordon R. SLEMON, University of Toronto, Toronto, Canadá.

Bernard SMITH, Comandante, ERADCOM, U. S. Army Fort Monmouth, New Jersey.

Charles SUFFEL, Departamento de Matemáticas, Stevens Institute of Technology, Hoboken, New Jersey.

Frederick L. SWERN, Bendix, Succasunna, New Jersey.

Jeff TOSK, Departamento de Microbiología, Loma Linda University, Loma Linda, California.

William H. TRANTER, Departamento de Ingeniería Eléctrica, University of Missouri, Rolla, Missouri.

Katsuaki TSURUSHIMA, Director de la División de Sonido de Alta Fidelidad y Vídeo, Sony Corporation of America, Park Ridge, New Jersey.

Kwei TU, Lockheed Engineering and Management Services Company, Inc., Houston, Texas.

Wayne C. TURNER, Profesor, Montana State University, Bozeman, Montana.

Thomas C. UPSON, Departamento de Matemáticas, Rochester Institute of Technology, Rochester, Nueva York.

Richard VAN SLYKE, Departamento de Ingeniería Eléctrica, Stevens Institute of Technology, Hoboken, New Jersey.

Surya V. VARANASI, Lockheed Engineering and Management Services Company, Inc., Houston, Texas.

L. S. WATKINS, Investigador del Centro de Investigaciones de Ingeniería, Western Electric Company, Princeton, New Jersey.

Gerald WEISS, Profesor, Polytechnic Institute of New York, Brooklyn, Nueva York.

Martin WOLF, Profesor del Departamento de Ingeniería Eléctrica y Ciencias, University of Pennsylvania, Philadelphia, Pennsylvania.

Rodger E. ZIEMER, Departamento de Ingeniería Eléctrica, University of Missouri, Rolla, Missouri.

PRIMERA PARTE

Matemáticas

CAPÍTULO I

Unidades y constantes

Shin-R Lin

New York Institute of Technology
Old Westbury, Nueva York

1.1 Unidades SI básicas

1.2 Unidades SI derivadas

1.3 Unidades SI complementarias

1.4 Símbolos de cantidades físicas

1.5 Símbolos y valores de constantes físicas

1.6 Factores de conversión

1.1 UNIDADES SI BÁSICAS

En mediciones físicas o de ingeniería se utilizan muchas cantidades diferentes, y para comparar sus magnitudes es necesario contar con una base de comparación. Para una cantidad dada, esta base se denomina unidad de magnitud.

El Sistema Internacional de Unidades (SI) fue adoptado por la 11.^a Conferencia General de Pesas y Medidas (CGPM, del nombre oficial en francés Conférence Générale des Poids et Mesures) en 1960. Se pretende que el SI sea la base para la estandarización mundial de las unidades de medición, por lo que se está adoptando en casi todos los países del mundo. Representa la forma moderna del sistema métrico de medición.

Las unidades SI se dividen en tres clases: básicas, derivadas y complementarias. En la tabla 1-1 se muestran las unidades básicas.

Tabla 1-1. Unidades SI básicas

Cantidad	Unidades SI	
	Nombre	Símbolo
Longitud	metro	m
Masa	kilogramo	kg
Tiempo	segundo	s
Corriente eléctrica	amperio	A
Temperatura termodinámica	kelvin	K
Cantidad de sustancia	mol	mol
Intensidad luminosa	bujía o candela	cd

A continuación se presentan algunas definiciones de las unidades SI básicas, y entre corchetes se indica la conferencia internacional en que se establecieron:

- **Metro (m).** Unidad de longitud, igual a la distancia recorrida por la luz en el vacío durante la fracción $1/299\,792\,458$ de un segundo [17.^a CGPM (1983)].
- **Kilogramo (kg).** Unidad de masa, igual a la masa del Kilogramo Prototipo Internacional [1.^a CGPM (1889); 3.^a CGPM (1901)].
- **Segundo (s).** Unidad de tiempo, igual a la duración de $9\,192\,631\,770$ ciclos de la radiación correspondiente a la transición entre los dos niveles hiperfinos del estado fundamental del átomo de cesio 133 [13.^a CGPM (1967)].
- **Amperio (A).** Unidad de corriente eléctrica, es la intensidad de corriente constante que, al ser mantenida en dos conductores rectilíneos paralelos de longitud infinita y sección transversal circular despreciable, colocados en el vacío con una separación de 1 metro, produciría entre dichos conductores una fuerza igual a 2×10^{-7} N/m [9.^a CGPM (1948)].
- **Kelvin (K).** Unidad de temperatura termodinámica, es igual a la fracción $1/273.16$ de la temperatura termodinámica del punto triple del agua [13.^a CGPM (1967)].
- **Mol (1).** Unidad de cantidad de sustancia o de entidades elementales, igual a la cantidad de sustancia de un sistema que contiene tantas entidades elementales como átomos haya en 0.012 kg de carbono 12 (esta cantidad es el número de Avogadro, 6.0220×10^{23}).
(2) Cuando se utiliza el mol es necesario especificar las entidades elementales, que pueden ser átomos, moléculas, iones, electrones u otras partículas, o grupos de dichas partículas [14.^a CGPM (1971)].
- **Candela (cd).** Intensidad luminosa, en una dirección dada, de una fuente que emite radiación

Tabla 1-2. Unidades SI derivadas con nombres especiales

Cantidad	Unidades SI		
	Nombre	Símbolo	Expresión en términos de otras unidades
Frecuencia	hertz	Hz	
Fuerza	newton	N	
Presión, esfuerzo, tensión	pascal	Pa	N / m ²
Energía, trabajo, cantidad de calor	joule	J	N·m
Energía, flujo radiante	watt	W	J / s
Carga eléctrica, cantidad de electricidad	coulomb	C	
Potencial eléctrico, diferencia de potencial, fuerza electromotriz	volt	V	W / A
Capacitancia	farad	F	C / V
Resistencia eléctrica	ohm	Ω	
Conductancia eléctrica	siemens	S	A / V
Flujo magnético	weber	Wb	V·s
Densidad de flujo magnético	tesla	T	Wb / m ²
Inductancia	henry	H	Wb / A
Temperatura Celsius	grado Celsius	°C	
Flujo lumínico	lumen	lm	
Emitancia luminosa (iluminancia)	lux	lx	lm / m ²

Tabla 1-3. Unidades SI derivadas con nombres especiales, admitidas por razones de preservación de la salud del ser humano

Cantidad	Unidades SI		
	Nombre	Símbolo	Expresión en términos de otras unidades
Actividad de un radionúclido (radiactividad)	becquerel	Bq	
Radiabsorción, dosis absorbida, energía específica impartida, kerma, índice de dosis absorbida	gray	Gy	J / kg
Dosis equivalente, índice de dosis equivalente	sievert	Sv	J / kg

monocromática cuya frecuencia es de 540×10^{12} hertz y cuya intensidad radiante en dicha dirección es de 1/683 watt por esterradián [16.^a CGPM (1979)].

esterradián deben considerarse unidades derivadas adimensionales que pueden usarse u omitirse en las expresiones de la unidad derivada.

1.2 UNIDADES SI DERIVADAS

Las unidades derivadas se expresan algebraicamente en términos de las unidades básicas por medio de los símbolos matemáticos de la multiplicación y de la división, conforme a las respectivas definiciones o leyes físicas.

Algunas unidades derivadas han recibido nombres y símbolos especiales, que se presentan en la tabla 1-2. Estas unidades pueden usarse, a su vez, para expresar otras unidades derivadas, las cuales se presentan en la tabla 1-3.

1.3 UNIDADES SI COMPLEMENTARIAS

Existen dos unidades SI complementarias: el radián, que es la unidad de ángulo plano, y el esterradián o estereoradián, que es la unidad de ángulo sólido [1.^a CGPM (1960)]. (Véase la tabla 1-4.)

En 1980, el Comité Internacional de Pesas y Medidas (CIPM) estipuló que, en unidades SI, las cantidades ángulo plano y ángulo sólido debían considerarse cantidades derivadas adimensionales. Por consiguiente, las unidades complementarias radián y

1.4 SÍMBOLOS DE CANTIDADES FÍSICAS

En la tabla 1-5 se pueden consultar los símbolos y las unidades de medición que corresponden a las distintas cantidades físicas.

1.5 SÍMBOLOS Y VALORES DE CONSTANTES FÍSICAS

En la tabla 1-6 se muestran los símbolos, valores y unidades de medición que se han establecido para las constantes físicas.

1.6 FACTORES DE CONVERSIÓN

En la tabla 1-7 se presentan los factores de conversión para las cantidades que se utilizan con frecuencia. En Mechtly puede consultarse una lista más completa.

LIBROS Y REVISTAS

- D. Goldman y R. J. Bell, eds., *The International System of Units (SI)*, NBS Special Publication 330.
- IEEE, *Letter Symbols for Quantities Used in Electrical Science and Electrical Engineering* (IEEE, No. 280). ASME, Nueva York.
- E. Mechtly, *The International System of Units*, NASA, sp-7012.
- Particle Data Group, *Review of Particle Properties*, *Rev. Mod. Phys.* **52**(2).
- Informe de CODATA, Boletín n.º 11, *Recommended Consistent Values of the Fundamental Physical Constants 1973*, Pergamon Press, Oxford.

Tabla 1-4. Unidades SI complementarias

Cantidad	Unidades SI	
	Nombre	Símbolo
Ángulo plano	radián	rad
Ángulo sólido	esterradián o estereoradián	sr

Tabla 1-5. Símbolos de cantidades físicas

Cantidad	Símbolo	Nombre de la medición	Unidad de medición
Tiempo y espacio			
Área	A, S	metro cuadrado	m^2
Volumen	V	metro cúbico	m^3
Distancia, desplazamiento	d, r	metro	m
Desplazamiento angular	θ	radián	rad
Velocidad	v	metro por segundo	m/s
Velocidad angular	ω	radián por segundo	rad/s
Aceleración lineal	a	metro por segundo cuadrado	m/s^2
Aceleración angular	α	radián por segundo cuadrado	rad/s^2
Período	T	segundo	s
Frecuencia	f, ν	hertz	Hz
Frecuencia angular	ω	revolución por segundo	r/s
Frecuencia rotacional	n	radián por segundo	rad/s
Longitud de onda	λ	metro	m
Número de onda	$\tilde{\nu}$	metro inverso	m^{-1}
Número de onda angular	k	radián por metro	rad/m
Mecánica			
Densidad, densidad de masa	ρ	kilogramo por metro cúbico	kg/m^3
Cantidad de movimiento (momentum)	p	kilogramo-metro por segundo	$kg \cdot m/s$
Cantidad de movimiento angular	L	kilogramo-metro cuadrado por segundo	$kg \cdot m^2/s$
Momento de inercia	I	kilogramo-metro cuadrado	$kg \cdot m^2$
Fuerza	F	newton	N
Peso	W	newton	N
Densidad de peso	γ	newton por metro cúbico	N/m^3
Momento de fuerza	M	newton-metro	$N \cdot m$
Momento de torsión	T, M	newton-metro	$N \cdot m$
Presión o esfuerzo	p	pascal	Pa
Módulo de elasticidad (o de Young)	E	pascal	Pa
Módulo de esfuerzo cortante	G	pascal	Pa
Módulo volumétrico	K	pascal	Pa
Energía	J	joule	J
Energía cinética	K	joule	J
Energía potencial	U	joule	J
Trabajo	W	joule	J
Potencia o flujo de energía	P	vatio	W
Eficiencia	η	(sólo es numérico)	
Electricidad y magnetismo			
Carga eléctrica	Q, q	culombio	C
Densidad lineal de carga	λ	culombio por metro	C/m
Densidad de carga superficial	σ	culombio por metro cuadrado	C/m^2
Densidad de carga volumétrica	ρ	culombio por metro cúbico	C/m^3
Intensidad de campo eléctrico	E	voltio por metro	V/m
Potencial eléctrico, diferencia de potencial, fuerza electromotriz	V	voltio	V
Flujo eléctrico	Ψ	culombio	C
Densidad de flujo eléctrico	D	culombio por metro cuadrado	C/m^2
Capacitancia	C	faradio	F
Permitividad	ϵ	faradio por metro	F/m
Densidad de corriente	J	amperio por metro cuadrado	A/m^2
Resistencia	R	ohmio	Ω
Resistividad	ρ	ohmio-metro	$\Omega \cdot m$
Conductancia	G	siemens	S
Conductividad	σ	siemens por metro	S/m
Intensidad de campo magnético	H	amperio por metro	A/m
Flujo magnético	Φ	weber	Wb
Acoplamiento de flujo magnético	Λ	weber	Wb
Densidad de flujo magnético	B	tesla	T

Tabla 1-5. Símbolos de cantidades físicas (continuación)

Cantidad	Símbolo	Nombre de la medición	Unidad de medición
Electricidad y magnetismo (cont.)			
Permeabilidad del espacio libre	μ_0	henrio por metro	H / m
Permeabilidad	μ	henrio por metro	H / m
Susceptibilidad magnética	X_m	(sólo en numérico)	
Inductancia	L	henrio	H
Reluctancia	R	henrio inverso	H ⁻¹
Impedancia	Z	ohmio	Ω
Reactancia	X	ohmio	Ω
Reactancia capacitiva	X_c	ohmio	Ω
Potencia	P	watt	W
Frecuencia	f	hercio	Hz
Frecuencia angular	ω	radián por segundo	rad / s
Longitud de onda	λ	metro	m
Calor y termodinámica			
Temperatura termodinámica (absoluta)	T, Θ	kelvin	K
Temperatura Celsius	t, θ	grados Celsius	°C
Cantidad de calor	Q	julio	J
Energía interna	U	julio	J
Energía libre	F	julio	J
Entalpía	H	julio	J
Entropía	S	julio o kelvin	J / K
Capacidad de calor	C	julio o kelvin	J / K
Conductividad térmica	λ, k	watt por metro Kelvin	W / m·K
Capacidad calorífica específica (entropía específica)	c	julio por kilogramo-kelvin	J / kg·K
Luminosidad y radiación			
Intensidad radiante	I	vatio por esterradián	W / sr
Potencia radiante, flujo radiante	P	vatio	W
Energía radiante	W	julio	J
Intensidad específica de radiación (radiancia)	L	vatio por esterradián-metro cuadrado	W / sr·m ²
Excitancia radiante	M	vatio por metro cuadrado	W / m ²
Irradiancia (irradiación)	E	vatio por metro cuadrado	W / m ²
Intensidad luminosa (luminosidad)	I	candela	cd
Flujo luminoso	ϕ	lumen	lm
Cantidad de luz	Q	lumen-segundo	lm·s
Luminancia (brillo o brillancia)	L	candela por metro cuadrado	cd / m ²
Eficacia luminosa	$k(\lambda)$	lumen por vatio	lm / W
Índice refractivo o índice de refracción	n	(sólo numérico)	
Emisividad	$\varepsilon(\lambda)$	(sólo numérico)	
Coefficiente de absorción, absortancia	$\alpha(\lambda)$	(sólo numérico)	
Transmitancia	$\tau(\lambda)$	(sólo numérico)	
Reflectancia	$\rho(\lambda)$	(sólo numérico)	
Sonido y acústica			
Velocidad del sonido	v	metro por segundo	m / s
Flujo de energía sonora	P	vatio	W
Intensidad del sonido	I	vatio por metro	W / m
Impedancia acústica	Z	newton-segundo por metro cúbico	N·s / m ³
Factor de absorción acústica	α	(sólo numérico)	
Factor de reflexión	ρ	(sólo numérico)	
Factor de transmisión	τ	(sólo numérico)	
Factor de disipación	δ	(sólo numérico)	
Nivel relativo	T	fon	

Fuente. "Letter Symbols for Quantities Used in Electrical Science and Electrical Engineering", IEEE Standards 280, reproducido con autorización.

Tabla 1-6. Constantes físicas

Cantidad	Símbolo	Valor	Error (ppm)	Prefijo	Unidad de medición
Velocidad de la luz en el vacío	c	2.997 924 58	0.004	10^8	m / s
Constante de gravitación	G	6.672 000	615	10^{-11}	Nm ² / kg ²
Constante de Avogrado	N_A	6.022 045	5.1	10^{26}	kmol ⁻¹
Constante de Boltzmann	k	1.380 662	32	10^{-23}	J / K
Constante de Planck	h	1.054 588 7	5.4	10^{-34}	J.s
Volumen de los gases ideales a presión y temperatura estándares	V	2.241 383	31	10^1	m ³ / kmol
Carga electrónica	e	1.602 189 2	2.9	10^{-19}	C
Magalectrón-volt	Mev	1.602 189 2	2.9	10^{-13}	J
Unidad de masa atómica	amu	1 / 12 M _{el2} = 1.660 566 = 9.315 016	5.1	10^{-27}	kg
Masa del electrón en reposo	m_e	9.109 534	2.8	10^{-31}	MeV / c ²
Masa del protón en reposo	m_p	0.511 003 4	2.8	10^{-27}	kg
Constante de Stefan-Boltzmann	σ	1.672 648	5.5	10^{-27}	MeV / c ²
Constante de Rydberg	R_∞	9.382 796	2.8	10^2	MeV / c ²
Constante de estructura fina	α	5.670 32	125	10^{-8}	W / m ² K ⁴
Radio de Bohr	a_0	1.097 373 177	0.075	10^7	m ⁻¹
Radio clásico del electrón	r_e	1 / 137.03604	0.82	10^{-11}	m
Longitud de onda del electrón de Compton	λ_c	5.291 770 6	0.82	10^{-15}	m
Magnetón de Bohr	μ_B	2.817 938 0	2.5	10^{-13}	m
Magnetón nuclear	μ_N	3.861 590 5	1.6	10^{-24}	J / T
Momento magnético del electrón	μ_e	9.274 077 5	4.5	10^{-27}	J / T
Momento magnético del protón	μ_p	5.050 823 7	4.6	10^{-27}	J / T
		9.284 851	7.0	10^{-24}	J / T
		1.410 617 0	4.5	10^{-26}	J / T

Fuente. Particle Data Group, Rev. Mod. Phys., 52(2), parte II, y CODATA, Bull. 11.

Tabla 1-7. Factores de conversión para el Sistema Internacional de Unidades

Para convertir de	Multiplicar por	Para obtener
Longitud		
Angstrom	1.000 000 E - 10	metros
Pulgada	2.540 000 E - 02	metros
Pie	3.048 000 E - 01	metros
Año-luz	9.46055 E 15	metros
Micrón	1.000 000 E - 06	metros
Milla (náutica internacional)	1.852 000 E 03	metros
Milla (internacional)	1.609 344 E 03	metros
Área		
Acre (EE.UU.)	4.046 873 E - 03	metros cuadrados
Pulgada cuadrada	6.451 600 E - 04	metros cuadrados
Pie cuadrado	9.290 304 E - 02	metros cuadrados
Milla cuadrada (internacional)	2.589 988 E 06	metros cuadrados
Volumen		
Onza líquida (EE.UU.)	2.957 353 E - 05	metros cúbicos
Galón líquido (EE.UU.)	3.785 412 E - 03	metros cúbicos
Pulgadas cúbicas	1.638 706 E - 05	metros cúbicos
Pinta líquida (EE.UU.)	4.731 765 E - 04	metros cúbicos
Cuarto líquido (EE.UU.)	9.463 529 E - 04	metros cúbicos
Ángulos		
Grados	1.745 329 E - 02	radianes
Minutos	2.908 882 E - 04	radianes
Segundos	4.848 137 E - 06	radianes
Velocidad		
Pies por segundo	3.048 000 E - 01	metros por segundo
Millas por hora	4.470 400 E - 01	metros por segundo
Velocidad		
Millas por hora	1.609 344 E 03	kilómetros por hora
Kilómetros por hora	2.777 778 E - 01	metros por segundo
Nudos (internacionales)	5.144 444 E - 01	metros por segundo
Aceleración, pies por segundo cuadrado	3.048 000 E - 01	metros por segundo cuadrado
Masa		
Onzas (avoirdupois)	2.834 952 E - 02	kilogramos
Onzas (troy)	3.110 348 E - 02	kilogramos
Libras (avoirdupois)	4.535 924 E - 01	kilogramos
Libras (troy)	3.732 417 E - 01	kilogramos
Slugs	1.459 390 E 01	kilogramos
Toneladas (largas, 2 240 lb)	1.016 047 E 03	kilogramos
Toneladas (métricas)	1.000 000 E 03	kilogramos
Toneladas (cortas, 2 000 lb)	9.071 847 E 02	kilogramos
Densidad		
Libras por pie cuadrado	1.601 846 E 01	kilogramos por metro cúbico
Slugs por pie cúbico	5.153 788 E 02	kilogramos por metro cúbico

Tabla 1-7. Factores de conversión para el Sistema Internacional de Unidades (continuación)

Para convertir de	Multiplicar por	Para obtener
Fuerza		
Dinas	1.000 000 E - 05	newtons
Onzas-fuerza	2.780 139 E - 01	newtons
Libras-fuerza	4.448 222 E 00	newtons
Toneladas-fuerza	8.896 444 E 03	newtons
Trabajo-energía		
British thermal units (Btu) (termoquímica)	1.054 350 E 03	julios
Calorías (termoquímica)	4.184 000 E 00	julios
Electrovoltio	1.602 189 E - 19	julios
Ergs	1.000 000 E - 07	julios
Pies-libra-fuerza	1.355 818 E 00	julios
Potencia		
Caballos (550 pies-libra-fuerza por segundo)	7.456 999 E 02	vattios
Pies-libra-fuerza por segundo	1.355 818 E 00	vattios
Presión o esfuerzo (fuerza por unidad de área)		
Atmósferas (estándares)	1.013 250 E 05	pascales
Bars	1.000 000 E 05	pascales
Torr (milímetros de mercurio a 0 °C)	1.333 224 E 02	pascales
Temperatura		
Grados Celsius	$t_k = t^{\circ}C + 273.15$	kelvins
Grados Fahrenheit	$t_k = (t^{\circ}F + 459.67) / 1.8$	kelvins
Grados Rankine	$t_k = t^{\circ}R / 1.8$	kelvins
Kelvins	$t^{\circ}C = t_k - 273.15$	grados Celsius
Luz		
Pie-candela	1.076 391 E - 01	lux
Pie-lambert	3.426 259 E - 02	candelas por metro cuadrado
Lambert	3.183 099 E - 03	candelas por metro cuadrado

CAPÍTULO 2

Matemáticas utilizadas en ingeniería

Paul Baum

Queens College
Flushing, Nueva York

Eleanor Baum

Pratt Institute
Brooklyn, Nueva York

2.1 Álgebra

- 2.1.1 Exponentes, radicales y logaritmos
- 2.1.2 Teorema del binomio
- 2.1.3 Polinomios
- 2.1.4 Ecuación cuadrática o de segundo grado
- 2.1.5 Determinantes
- 2.1.6 Ecuaciones simultáneas

2.2 Funciones trigonométricas e hiperbólicas y números complejos

- 2.2.1 Definiciones trigonométricas
- 2.2.2 Fórmulas de reducción
- 2.2.3 Funciones trigonométricas inversas
- 2.2.4 Identidades trigonométricas
- 2.2.5 Relaciones entre los lados y ángulos de cualquier triángulo plano
- 2.2.6 Funciones exponencial e hiperbólica
- 2.2.7 Números complejos
- 2.2.8 Adición y sustracción
- 2.2.9 Multiplicación
- 2.2.10 Complejos conjugados
- 2.2.11 División

2.3 Cálculo

- 2.3.1 Diferenciación
- 2.3.2 Integración

2.4 Matrices

- 2.4.1 Matrices especiales
- 2.4.2 Operaciones con matrices
- 2.4.3 Otras definiciones
- 2.4.4 Solución matricial de ecuaciones lineales simultáneas

2.5 Vectores

- 2.5.1 Álgebra vectorial
- 2.5.2 Análisis vectorial

2.6 Ecuaciones diferenciales lineales

- 2.6.1 Definiciones
- 2.6.2 Solución de la ecuación diferencial homogénea
- 2.6.3 Solución de ecuaciones diferenciales no homogéneas
- 2.6.4 Solución completa de ecuaciones diferenciales no homogéneas

2.7 Ecuaciones en diferencias lineales

- 2.7.1 Ecuaciones en diferencias

- 2.7.2 Solución de ecuaciones en diferencias homogéneas
- 2.7.3 Solución de ecuaciones en diferencias no homogéneas
- 2.7.4 Solución completa de ecuaciones en diferencias
- 2.8 Series de potencias**
 - 2.8.1 Progresión aritmética
 - 2.8.2 Progresión geométrica
 - 2.8.3 Suma de potencias de números naturales
 - 2.8.4 Series de potencias de funciones comunes
 - 2.8.5 Desarrollo en serie de Taylor
- 2.9 Serie de Fourier**
 - 2.9.1 Serie trigonométrica de Fourier
 - 2.9.2 Serie exponencial de Fourier
 - 2.9.3 Teoremas en los que se utilizan los coeficientes de Fourier
 - 2.9.4 Serie de Fourier de algunas funciones comunes

2.1 ÁLGEBRA

A continuación se presentan algunas fórmulas y conceptos básicos útiles para los ingenieros eléctricos.

2.1.1 Exponentes, radicales y logaritmos

$$x^0 = 1 \quad (x \neq 0) \quad x^{-n} = \frac{1}{x^n} \quad x^n x^m = x^{n+m}$$

$$\frac{x^n}{x^m} = x^{n-m} \quad (x^n)^m = x^{nm} \quad (xy)^n = x^n y^n \quad \left(\frac{x}{y}\right)^n = \frac{x^n}{y^n}$$

$$\sqrt{x} = x^{1/2} \quad \sqrt[n]{x} = x^{1/n} \quad \sqrt[n]{x^m} = x^{m/n} \quad \sqrt[n]{xy} = x^{1/n} y^{1/n}$$

$$\log x^n = n \log x \quad \log xy = \log x + \log y$$

$$\log \frac{x}{y} = \log x - \log y$$

2.1.2 Teorema del binomio

$$(x \pm y)^n = \sum_{k=0}^n \frac{n!}{k!(n-k)!} x^k (\pm y)^{n-k}$$

donde $n! = n$ factorial $= n(n-1)\dots(3)(2)(1)$. Por ejemplo:

$$(x-y)^4 = x^4 - 4yx^3 + 6y^2x^2 - 4y^3x + y^4$$

2.1.3 Polinomios

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

b es un "cero" del polinomio y una raíz de la ecuación $p(x) = 0$ si $p(b) = 0$. De hecho, $(x-b)$ es un factor del polinomio.

Un polinomio de grado n tiene n ceros, que pueden ser números reales o complejos y pueden ser valores repetidos. Sólo los coeficientes (a 's) del polinomio son números reales; entonces las raíces complejas se presentan en pares complejos conjugados.

Algunas veces es útil observar que:

$$\frac{a_{n-1}}{a_n} = -\sum \text{raíces} \quad \frac{a_0}{a_n} = (-1)^n \prod \text{raíces}$$

2.1.4 Ecuación cuadrática o de segundo grado

La solución de $ax^2 + bx + c = 0$ está dada por:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Si a , b y c son reales, entonces las raíces x_1 y x_2 son:

$x_1 \neq x_2$, y x_1 y x_2 son reales si $\sqrt{b^2 - 4ac} > 0$

$$x_1 = x_2 \text{ si } \sqrt{b^2 - 4ac} = 0$$

x_1 y x_2 son complejos conjugados si $\sqrt{b^2 - 4ac} < 0$.

2.5.1 Determinantes

El arreglo cuadrado que se muestra a continuación se conoce como determinante de orden n :

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}$$

Los coeficientes de las líneas horizontales forman filas, y los de las verticales columnas. En este arreglo, los elementos se identifican por su posición mediante un doble subíndice: el primero indica el renglón y el segundo la columna. La diagonal principal consta de los elementos que se encuentran en la línea que va de a_{11} a a_{nn} .

El menor de cualquier elemento a_{jk} es el determinante que queda cuando se eliminan el renglón y la columna que contienen dicho elemento.

Por ejemplo, para el determinante A :

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

el menor para a_{22} es:

$$M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$$

El cofactor de un elemento a_{jk} se define como sigue:

$$\text{cofactor} = (-1)^{j+k} (\text{menor})_{jk}$$

donde j es el renglón del elemento y k la columna del elemento.

El determinante de orden n se evalúa por reducción sucesiva. Es igual a la suma del producto de los elementos de cualquier renglón o columna, multiplicados por el cofactor del elemento correspondiente. Para el determinante de tercer orden dado:

$$A = a_{11}M_{11} - a_{21}M_{21} + a_{31}M_{31}$$

$$A = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

Este proceso continúa:

$$A = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13})$$

2.1.6 Ecuaciones simultáneas

Los determinantes se utilizan para resolver conjuntos de ecuaciones simultáneas de la forma:

$$a_{11}X_1 + a_{12}X_2 + a_{13}X_3 + \dots + a_{1n}X_n = y_1$$

$$a_{n1}X_1 + a_{n2}X_2 + a_{n3}X_3 + \dots + a_{nn}X_n = y_n$$

Mediante la regla de Cramer se obtiene la solución para tal sistema de ecuaciones como:

$$X_1 = \frac{D_1}{\Delta} \quad X_2 = \frac{D_2}{\Delta} \quad \dots \quad X_n = \frac{D_n}{\Delta}$$

Aquí:

$$\Delta \text{ determinante del sistema} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

y D_j es el determinante formado al sustituir la j -ésima columna de coeficientes en Δ por la columna y_1, y_2, \dots, y_n .

Por ejemplo:

$$5x_1 - 2x_2 - 3x_3 = 10$$

$$-2x_1 + 4x_2 - x_3 = 0$$

$$-3x_1 - x_2 + 6x_3 = 0$$

Usando la regla de Cramer:

$$x_1 = \frac{D_1}{\Delta} = \frac{\begin{vmatrix} 10 & -2 & -3 \\ 0 & 4 & -1 \\ 0 & -1 & 6 \end{vmatrix}}{\begin{vmatrix} 5 & -2 & -3 \\ -2 & 4 & -1 \\ -3 & -1 & 6 \end{vmatrix}}$$

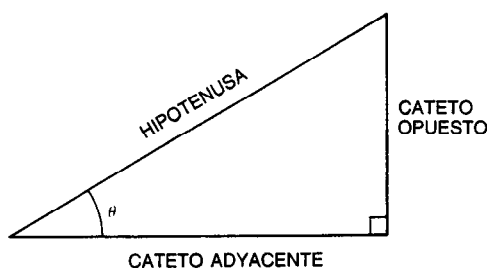


Fig. 2-1. Triángulo rectángulo.

$$X_1 = \frac{10 \begin{vmatrix} 4 & -1 \\ -1 & 6 \end{vmatrix} - 0 \begin{vmatrix} -2 & -3 \\ -1 & 6 \end{vmatrix} + 0 \begin{vmatrix} -2 & -3 \\ 4 & -1 \end{vmatrix}}{5 \begin{vmatrix} 4 & -1 \\ -1 & 6 \end{vmatrix} - (-2) \begin{vmatrix} -2 & -3 \\ -1 & 6 \end{vmatrix} + (-3) \begin{vmatrix} -2 & -3 \\ 4 & -1 \end{vmatrix}}$$

$$X_1 = \frac{10(24 - 1)}{5(24 + 1) + 2(-12 - 3) - 3(2 + 12)} = \frac{230}{43}$$

2.2 FUNCIONES TRIGONÓMICAS E HIPERBÓLICAS Y NÚMEROS COMPLEJOS

2.2.1 Definiciones trigonométricas (fig. 2-1):

$$\text{sen } \theta = \frac{\text{cat. opuesto}}{\text{hipotenusa}} = \frac{1}{\text{csc } \theta}$$

$$\text{cos } \theta = \frac{\text{cat. adyacente}}{\text{hipotenusa}} = \frac{1}{\text{sec } \theta}$$

$$\text{tan } \theta = \frac{\text{cat. opuesto}}{\text{cat. adyacente}} = \frac{\text{sen } \theta}{\text{cos } \theta}$$

$$\text{cot } \theta = \frac{\text{cat. adyacente}}{\text{cat. opuesto}} = \frac{\text{cos } \theta}{\text{sen } \theta}$$

$$\text{sec } \theta = \frac{\text{hipotenusa}}{\text{cat. adyacente}} = \frac{1}{\text{cos } \theta}$$

$$\text{csc } \theta = \frac{\text{hipotenusa}}{\text{cat. opuesto}} = \frac{1}{\text{sen } \theta}$$

Respecto al sistema de coordenadas rectangulares y a la ubicación de los cuadrantes de la figura 2-2, el ángulo θ está en el segundo cuadrante. También debe observarse que un ángulo no cambia si se le suman 360° o 2π radianes. Los signos algebraicos de las funciones son:

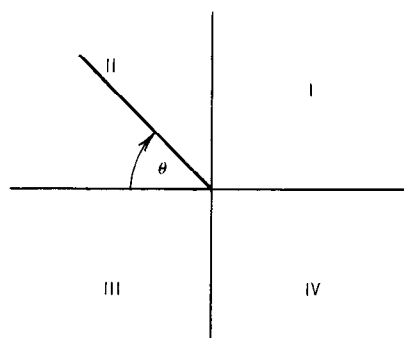


Fig. 2-2. Cuadrantes.

Cuadrante	Senos	Cosenos	Tangente
I	+	+	+
II	+	-	-
III	-	-	+
IV	-	+	-

Algunos valores que se utilizan con frecuencia para las funciones trigonométricas son:

Ángulo en grados y radianes	Sen θ	Cos θ	Tan θ
$0^\circ, 360^\circ, 2\pi$	0	1	0
$30^\circ, \pi/6$	$1/2$	$\sqrt{3}/2$	$1/\sqrt{3}$
$45^\circ, \pi/4$	$1/\sqrt{2}$	$1/\sqrt{2}$	1
$60^\circ, \pi/3$	$\sqrt{3}/2$	$1/2$	$\sqrt{3}$
$90^\circ, \pi/2$	1	0	$\pm \infty$
$180^\circ, \pi$	0	-1	0
$270^\circ, 3\pi/2$	-1	0	$\pm \infty$

2.2.2 Fórmulas de reducción

Ángulo \ Función	Senos	Cosenos	Tangente
$-\theta$	$-\sin \theta$	$\cos \theta$	$-\tan \theta$
$90^\circ - \theta$	$\cos \theta$	$\sin \theta$	$\cot \theta$
$90^\circ + \theta$	$\cos \theta$	$-\sin \theta$	$-\cot \theta$
$180^\circ - \theta$	$\sin \theta$	$-\cos \theta$	$-\tan \theta$
$180^\circ + \theta$	$-\sin \theta$	$-\cos \theta$	$\tan \theta$
$270^\circ - \theta$	$-\cos \theta$	$-\sin \theta$	$\cot \theta$
$270^\circ + \theta$	$-\cos \theta$	$\sin \theta$	$-\cot \theta$
$360^\circ - \theta$	$-\sin \theta$	$\cos \theta$	$-\tan \theta$

2.2.3 Funciones trigonométricas inversas

Dado que $\sin \theta = a$, puede decirse que θ es el ángulo cuyo seno es a . Este hecho se escribe como $\sin^{-1}a = \theta$, es decir, "el seno inverso de a es θ ".

Por ejemplo:

$$\sin^{-1} 0.319 = 18.6^\circ, \text{ o bien } 180^\circ - 18.6^\circ = 161.4^\circ$$

El ángulo puede estar en el primero o en el segundo cuadrantes.

2.2.4 Identidades trigonométricas*

Funciones de una suma o diferencia

(a) $\sin(x + y) = \sin x \cos y + \cos x \sin y.$

(b) $\sin(x - y) = \sin x \cos y - \cos x \sin y.$

(c) $\cos(x + y) = \cos x \cos y - \sin x \sin y.$

(d) $\cos(x - y) = \cos x \cos y + \sin x \sin y.$

(e) $\tan(x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$

(f) $\tan(x - y) = \frac{\tan x - \tan y}{1 + \tan x \tan y}$

Suma o diferencia de funciones como producto de funciones

(a) $\sin x + \sin y = 2 \sin \frac{1}{2}(x + y) \cos \frac{1}{2}(x - y).$

(b) $\sin x - \sin y = 2 \cos \frac{1}{2}(x + y) \sin \frac{1}{2}(x - y).$

(c) $\cos x + \cos y = 2 \cos \frac{1}{2}(x + y) \cos \frac{1}{2}(x - y).$

(d) $\cos x - \cos y = -2 \sin \frac{1}{2}(x + y) \sin \frac{1}{2}(x - y).$

Productos de funciones como sumas o diferencias de funciones

(a) $2 \sin x \cos y = \sin(x + y) + \sin(x - y)$

(b) $2 \cos x \sin y = \sin(x + y) - \sin(x - y)$

(c) $2 \cos x \cos y = \cos(x + y) + \cos(x - y)$

(d) $2 \sin x \sin y = -\cos(x + y) + \cos(x - y)$

Funciones del doble o la mitad de un ángulo

(a) $\sin 2x = 2 \sin x \cos x.$

(b) $\cos 2x = \cos^2 x - \sin^2 x, = 2 \cos^2 x - 1,$
 $= 1 - 2 \sin^2 x.$

(c) $\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$

(d) $\sin \frac{1}{2}x = \pm \sqrt{\frac{1 - \cos x}{2}}$ o bien $2 \sin^2 \theta = 1 - \cos 2\theta$

(e) $\cos \frac{1}{2}x = \pm \sqrt{\frac{1 + \cos x}{2}}$ o bien $2 \cos^2 \theta = 1 + \cos 2\theta$

(f) $\tan \frac{1}{2}x = \pm \sqrt{\frac{1 - \cos x}{1 + \cos x}} = \frac{\sin x}{1 + \cos x} = \frac{1 - \cos x}{\sin x}$

Una transformación importante

(a) $a \cos x + b \sin x = \sqrt{a^2 + b^2} \cos(x - \alpha),$

donde $\cos \alpha = \frac{a}{\sqrt{a^2 + b^2}}, \sin \alpha = \frac{b}{\sqrt{a^2 + b^2}}$

(b) $a \cos x + b \sin x = \sqrt{a^2 + b^2} \sin(x + \beta),$

donde $\sin \beta = \frac{a}{\sqrt{a^2 + b^2}}, \cos \beta = \frac{b}{\sqrt{a^2 + b^2}}$

* Reproducido con autorización de L. Smail, *Calculus*, Appleton-Century-Crofts, Nueva York.

La multiplicación de un número complejo y su conjugado da por resultado:

$$\underline{A} \underline{A}^* = (a + jb)(a - jb) = a^2 + b^2$$

o bien:

$$\underline{A} \underline{A}^* = (Ae^{j\theta})(Ae^{-j\theta}) = A^2$$

2.2.11 División

Si $\underline{A}/\underline{B} = \underline{C}$, entonces:

$$\frac{\underline{A}}{\underline{B}} = \frac{\underline{A}}{\underline{B}} \cdot \frac{\underline{B}^*}{\underline{B}^*} = \frac{\underline{A} \underline{B}^*}{\underline{B}^2}$$

La multiplicación por $\underline{B}^*/\underline{B}^*$ se denomina racionalización y se hace para convertir en real el denominador de la fracción:

$$\frac{\underline{A}}{\underline{B}} = \frac{(a + jb)(c + jd)}{(c + jd)(c - jd)} = \frac{ac + bd}{c^2 + d^2} + j \frac{bc - ad}{c^2 + d^2}$$

En forma exponencial o polar:

$$\frac{\underline{A}}{\underline{B}} = \frac{A/\theta}{B/\phi} = \frac{Ae^{j\theta}}{Be^{j\phi}} = \frac{A}{B} e^{j(\theta-\phi)}$$

o bien:

$$\frac{\underline{A}}{\underline{B}} = \frac{A}{B} \angle (\theta - \phi)$$

Las potencias y raíces de los números complejos se determinan utilizando las leyes de los exponentes. Por tanto:

$$(\underline{A})^N = (Ae^{j\theta})^N = A^N e^{jN\theta} = A^N \angle N\theta$$

También:

$$(\underline{A})^{1/N} = [Ae^{j(\theta+2\pi k)}]^{1/N} = \sqrt[N]{A} \angle \frac{\theta + k2\pi}{N}$$

2.3 CÁLCULO

2.3.1 Diferenciación

Si la función $y = f(t)$ es diferenciable en t , entonces su derivada:

$$\frac{dy}{dt} = f'(t) = \lim_{\Delta t \rightarrow 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

representa:

1. La pendiente de la tangente a la gráfica de $f(t)$ contra el tiempo en un instante t .
2. La razón de cambio instantánea en y con respecto a t .

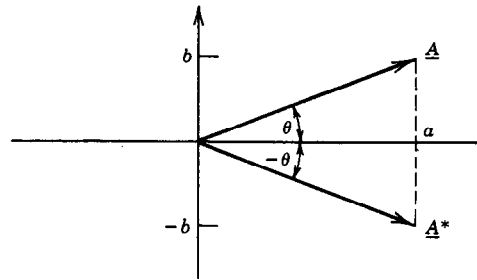


Fig. 2-5. Complejos conjugados.

A continuación se proporcionan las derivadas de funciones que se utilizan con frecuencia.

$f(t)$	$\frac{df}{dt} = f'(t)$
t^n	nt^{n-1}
e^{at}	ae^{at}
a^t	$a^t \ln a$
$\cos at$	$-a \sin at$
$\sin at$	$a \cos at$
$\ln t$	$1/t$
$\tan at$	$a \sec^2 at$
$\cot at$	$-a \csc^2 at$
$\sec at$	$a(\sec at)(\tan at)$
$\csc at$	$-a(\csc at)(\cot at)$
$\sinh at$	$a \cosh at$
$\cosh at$	$a \sinh at$
$\tanh at$	$a \operatorname{sech}^2 at$
$\coth at$	$-a \operatorname{csch}^2 at$
$\operatorname{sech} at$	$-a \operatorname{sech} at (\tanh at)$
$\sin^{-1} t$	$1/\sqrt{1-t^2}$
$\cos^{-1} t$	$-1/\sqrt{1-t^2}$
$\tan^{-1} t$	$1/(1+t^2)$
$\cot^{-1} t$	$-1/(1+t^2)$

Reglas de diferenciación. Supóngase que u y v son funciones de t .

1. Derivada de una constante:

$$\frac{d}{dt} (\text{const}) = 0$$

2. Derivada del producto escalar:

$$\frac{d}{dt} (cu) = c \frac{du}{dt}$$

3. Derivada de la suma:

$$\frac{d}{dt} (u \pm v) = \frac{du}{dt} \pm \frac{dv}{dt}$$

4. Derivada del producto:

$$\frac{d}{dt} (uv) = u \frac{dv}{dt} + v \frac{du}{dt}$$

5. Derivada del cociente:

$$\frac{d}{dt} \left(\frac{u}{v} \right) = \frac{v \frac{du}{dt} - u \frac{dv}{dt}}{v^2}$$

6. Regla de la cadena (y es función de u):

$$\frac{d}{dt} (y) = \frac{dy}{du} \cdot \frac{du}{dt}$$

7. Derivada de una función elevada a una potencia:

$$\frac{d}{dt} (u^n) = nu^{n-1} \frac{du}{dt}$$

$$\frac{d}{dt} (t^n) = nt^{n-1}$$

8. Derivada del valor absoluto de una función:

$$\frac{d}{dt} (|u|) = \frac{u}{|u|} \frac{du}{dt}$$

Máximos y mínimos. Sea x una función de t , por ejemplo $x(t)$. Para encontrar el valor de t que minimice o maximice $x(t)$, sea:

$$\frac{dx(t)}{dt} = 0$$

El valor correspondiente de t , por ejemplo $t = t_m$, que satisface la ecuación anterior será un máximo si se cumple:

$$\frac{d^2x(t)}{dt^2} < 0 \quad \text{en } t = t_m$$

y $t = t_m$ será un mínimo si:

$$\frac{d^2x(t)}{dt^2} > 0 \quad \text{en } t = t_m$$

Los valores extremos de $f(t)$ suelen denominarse como máximo y mínimo.

Diferencial de una función. Si $y = f(t)$, entonces la diferencial de $f(t)$ se denota por $df(t)$ y se define con la siguiente expresión:

$$dy = df(t) = f'(t)dt$$

Ejemplo. Si $y = f(t) = t^2$, entonces la derivada es:

$$f'(t) = \frac{dy}{dt} = 2t$$

y la diferencial es:

$$dy = 2t dt$$

Derivadas parciales. Sea $u = f(t, x)$ una función de dos variables independientes t y x . Si x se mantiene fija, entonces u se vuelve temporalmente una función sólo de la variable t . Desde este punto de vista, la función tiene una derivada con respecto a t , denominada derivada parcial de u con respecto a t y que se denota por:

$$\frac{\partial u}{\partial t} \text{ o bien } u'_t, \text{ o bien } f'_t(t, x)$$

La derivada parcial de u con respecto a x se define de forma parecida. En general, para una función de varias variables, existe una derivada parcial con respecto a cada variable independiente.

Ejemplo. Sea $f(t, x) = t^3x^2$. Entonces:

$$\frac{\partial f}{\partial t} = 3t^2x^2 \quad \frac{\partial f}{\partial x} = 2t^3x \quad \frac{\partial^2 f}{\partial t \partial x} = 6t^2x \quad \frac{\partial^2 f}{\partial x^2} = 2t^3$$

2.3.2 Integración

La integración es la operación contraria a la diferenciación. Por tanto, si:

$$\frac{d}{dt} \left[\int f(t) dt \right] = f(t)$$

entonces:

$$\int f'(t) dt = f(t) + G$$

donde G = cualquier constante.

Para efectuar el proceso de integración es necesario simplemente aprender a utilizar en sentido inverso las fórmulas de diferenciación.

La constante arbitraria G se denomina constante de integración.

Integrales definidas. El teorema fundamental del cálculo establece que si $f(t)$ es una función continua en el intervalo $[a, b]$, entonces:

$$\int_a^b f(t) dt = F(t) \Big|_a^b = F(b) - F(a)$$

donde $F(t)$ es cualquier función tal que $F'(t) = f(t)$ para toda t en $[a, b]$.

Es importante advertir que este teorema describe un método para evaluar una integral definida, en vez de un procedimiento para determinar antiderivadas.

Propiedades de las integrales definidas

$$\int_a^b Kf(t) dt = K \int_a^b f(t) dt \quad \text{donde } K = \text{constante}$$

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt \quad \text{donde } a < c < b$$

$$\int_a^b [f(t) \pm g(t)] dt = \int_a^b f(t) dt \pm \int_a^b g(t) dt$$

Límite integración variable. Si la integral de $f(t)$ se considera sobre un intervalo variable (a, t) , entonces la integral definida es una función del límite superior de integración:

$$F(t) = \int_a^t f(x) dx$$

donde x es una variable muda.

Obsérvese además que la derivada de una integral definida con respecto al límite superior de integración es igual al valor de la integral en dicho límite superior:

$$\frac{d}{dt} \int_0^t f(x) dx = f(t)$$

Tabla de integrales. En la siguiente tabla de fórmulas de integración estándar se emplea la variable u en lugar de t para cubrir un mayor intervalo de aplicación.

Tabla 2-1. Breve tabla de integrales

Algunas formas fundamentales

$$\int du = u + C$$

$$\int c du = c \int du$$

$$\int (f + g + \dots) du = \int f du + \int g du + \dots$$

$$\int u dv = uv - \int v du$$

$$\int u^n du = \frac{u^{n+1}}{n+1} + C \quad (n \neq -1)$$

$$\int \frac{du}{u} = \ln u + C$$

Formas racionales en las que aparece $a + bu$

$$\int \frac{u du}{a + bu} = \frac{1}{b^2} [a + bu - a \ln(a + bu)] + C$$

$$\int \frac{u^2 du}{a + bu} = \frac{1}{b^3} [\frac{1}{2} (a + bu)^2 - 2a(a + bu) + a^2 \ln(a + bu)] + C$$

$$\int \frac{u du}{(a + bu)^2} = \frac{1}{b^2} \left[\frac{a}{a + bu} + \ln(a + bu) \right] + C$$

$$\int \frac{u^2 du}{(a + bu)^2} = \frac{1}{b^3} \left[a + bu - \frac{a^2}{a + bu} - 2a \ln(a + bu) \right] + C$$

$$\int \frac{du}{u(a + bu)} = -\frac{1}{a} \ln \frac{a + bu}{u} + C$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas racionales en las que aparece $a + bu$ (cont.)

$$\int \frac{du}{u^2(a+bu)} = -\frac{1}{au} + \frac{b}{a^2} \ln \frac{a+bu}{u} + C$$

$$\int \frac{du}{u(a+bu)^2} = \frac{1}{a(a+bu)} - \frac{1}{a^2} \ln \frac{a+bu}{u} + C$$

Formas en las que aparece $\sqrt{a+bu}$

$$\int u \sqrt{a+bu} du = \frac{2(3bu-2a)}{15b^2} (a+bu)^{3/2} + C$$

$$\int u^2 \sqrt{a+bu} du = \frac{2(15b^2u^2-12abu+8a^2)}{105b^3} (a+bu)^{3/2} + C$$

$$\int \frac{u du}{\sqrt{a+bu}} = \frac{2(bu-2a)}{3b^3} \sqrt{a+bu} + C$$

$$\int \frac{u^2 du}{\sqrt{a+bu}} = \frac{2(3b^2u^2-4abu+8a^2)}{15b^3} \sqrt{a+bu} + C$$

$$\int \frac{du}{u\sqrt{a+bu}} = \frac{1}{\sqrt{a}} \ln \frac{\sqrt{a+bu}-\sqrt{a}}{\sqrt{a+bu}+\sqrt{a}} + C \quad (a > 0)$$

$$\int \frac{du}{u\sqrt{a+bu}} = \frac{2}{\sqrt{-a}} \arctan \sqrt{\frac{a+bu}{-a}} + C \quad (a < 0)$$

$$\int \frac{du}{u^2\sqrt{a+bu}} = -\frac{\sqrt{a+bu}}{au} - \frac{b}{2a} \int \frac{du}{u\sqrt{a+bu}}$$

$$\int \frac{\sqrt{a+bu}}{u} du = 2\sqrt{a+bu} + a \int \frac{du}{u\sqrt{a+bu}}$$

$$\int \frac{\sqrt{a+bu}}{u^2} du = -\frac{\sqrt{a+bu}}{u} + \frac{b}{2} \int \frac{du}{u\sqrt{a+bu}}$$

Formas en las que aparecen $a^2 \pm u^2$ y $u^2 - a^2$

$$\int \frac{du}{a^2+u^2} = \frac{1}{a} \arctan \frac{u}{a} + C, \quad \text{si } a > 0$$

$$\int \frac{du}{a^2-u^2} = \frac{1}{2a} \ln \frac{a+u}{a-u} + C = \frac{1}{a} \tanh^{-1} \frac{u}{a} + C, \quad \text{si } u^2 < a^2$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas en las que aparecen $a^2 \pm u^2$ y $u^2 - a^2$ (cont.)

$$\int \frac{du}{u^2 - a^2} = \frac{1}{2a} \ln \frac{u - a}{u + a} + C = -\frac{1}{a} \coth^{-1} \frac{u}{a} + C, \text{ si } u^2 > a^2$$

Formas en las que aparece $\sqrt{a^2 - u^2}$

$$\int \frac{du}{\sqrt{a^2 - u^2}} = \arcsen \frac{u}{a} + C, \text{ si } u^2 < a^2, a > 0$$

$$\int \sqrt{a^2 - u^2} du = \frac{u}{2} \sqrt{a^2 - u^2} + \frac{a^2}{2} \arcsen \frac{u}{a} + C$$

$$\int u^2 \sqrt{a^2 - u^2} du = -\frac{u}{4} (a^2 - u^2)^{3/2} + \frac{a^2}{8} u \sqrt{a^2 - u^2} + \frac{a^4}{8} \arcsen \frac{u}{a} + C$$

$$\int \frac{\sqrt{a^2 - u^2}}{u} du = \sqrt{a^2 - u^2} - a \ln \left(\frac{a + \sqrt{a^2 - u^2}}{u} \right) + C$$

$$\int \frac{\sqrt{a^2 - u^2}}{u^2} du = -\frac{\sqrt{a^2 - u^2}}{u} - \arcsen \frac{u}{a} + C$$

$$\int \frac{u^2 du}{\sqrt{a^2 - u^2}} = -\frac{u}{2} \sqrt{a^2 - u^2} + \frac{a^2}{2} \arcsen \frac{u}{a} + C$$

$$\int \frac{du}{u \sqrt{a^2 - u^2}} = -\frac{1}{a} \ln \left(\frac{a + \sqrt{a^2 - u^2}}{u} \right) + C$$

$$\int \frac{du}{u^2 \sqrt{a^2 - u^2}} = -\frac{\sqrt{a^2 - u^2}}{a^2 u} + C$$

$$\int (a^2 - u^2)^{3/2} du = -\frac{u}{8} (2u^2 - 5a^2) \sqrt{a^2 - u^2} + \frac{3a^4}{8} \arcsen \frac{u}{a} + C$$

$$\int \frac{du}{(a^2 - u^2)^{3/2}} = \frac{u}{a^2 \sqrt{a^2 - u^2}} + C$$

Formas en las que aparece $\sqrt{a^2 + u^2}$

$$\int \frac{du}{\sqrt{a^2 + u^2}} = \ln(u + \sqrt{a^2 + u^2}) + C = \sinh^{-1} \frac{u}{a} + C$$

$$\int \sqrt{a^2 + u^2} du = \frac{u}{2} \sqrt{a^2 + u^2} + \frac{a^2}{2} \ln(u + \sqrt{a^2 + u^2}) + C = \frac{u}{2} \sqrt{a^2 + u^2} + \frac{a^2}{2} \sinh^{-1} \frac{u}{a} + C$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas en las que aparece $\sqrt{a^2 + u^2}$ (cont.)

$$\int u^2 \sqrt{a^2 + u^2} du = \frac{u}{8} (2u^2 + a^2) \sqrt{a^2 + u^2} - \frac{a^4}{8} \ln(u + \sqrt{a^2 + u^2}) + C$$

$$= \frac{u}{8} (2u^2 + a^2) \sqrt{a^2 + u^2} - \frac{a^4}{8} \sinh^{-1} \frac{u}{a} + C$$

$$\int \frac{\sqrt{a^2 + u^2}}{u} du = \sqrt{a^2 + u^2} - a \ln \left(\frac{a + \sqrt{a^2 + u^2}}{u} \right) + C$$

$$\int \frac{\sqrt{a^2 + u^2}}{u^2} du = -\frac{\sqrt{a^2 + u^2}}{u} + \ln(u + \sqrt{a^2 + u^2}) + C = -\frac{\sqrt{a^2 + u^2}}{u} + \sinh^{-1} \frac{u}{a} + C$$

$$\int \frac{u^2 du}{\sqrt{a^2 + u^2}} = \frac{u}{2} \sqrt{a^2 + u^2} - \frac{a^2}{2} \ln(u + \sqrt{a^2 + u^2}) + C = \frac{u}{2} \sqrt{a^2 + u^2} - \frac{a^2}{2} \sinh^{-1} \frac{u}{a} + C$$

$$\int \frac{du}{u \sqrt{a^2 + u^2}} = -\frac{1}{a} \ln \left(\frac{\sqrt{a^2 + u^2} + a}{u} \right) + C$$

$$\int \frac{du}{u^2 \sqrt{a^2 + u^2}} = -\frac{\sqrt{a^2 + u^2}}{a^2 u} + C$$

$$\int (a^2 + u^2)^{3/2} du = \frac{u}{8} (2u^2 + 5a^2) \sqrt{a^2 + u^2} + \frac{3a^4}{8} \ln(u + \sqrt{a^2 + u^2}) + C$$

$$= \frac{u}{8} (2u^2 + 5a^2) \sqrt{a^2 + u^2} + \frac{3a^4}{8} \sinh^{-1} \frac{u}{a} + C$$

$$\int \frac{du}{(a^2 + u^2)^{3/2}} = \frac{u}{a^2 \sqrt{a^2 + u^2}} + C$$

Formas en las que aparece $\sqrt{u^2 - a^2}$

$$\int \frac{du}{\sqrt{u^2 - a^2}} = \ln(u + \sqrt{u^2 - a^2}) + C = \cosh^{-1} \frac{u}{a} + C$$

$$\int \sqrt{u^2 - a^2} du = \frac{u}{2} \sqrt{u^2 - a^2} - \frac{a^2}{2} \ln(u + \sqrt{u^2 - a^2}) + C = \frac{u}{2} \sqrt{u^2 - a^2} - \frac{a^2}{2} \cosh^{-1} \frac{u}{a} + C$$

$$\int u^2 \sqrt{u^2 - a^2} du = \frac{u}{8} (2u^2 - a^2) \sqrt{u^2 - a^2} - \frac{a^4}{8} \ln(u + \sqrt{u^2 - a^2}) + C$$

$$= \frac{u}{8} (2u^2 - a^2) \sqrt{u^2 - a^2} - \frac{a^4}{8} \cosh^{-1} \frac{u}{a} + C$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas en las que aparece $\sqrt{u^2 + a^2}$ (cont.)

$$\int \frac{\sqrt{u^2 - a^2}}{u} du = \sqrt{u^2 - a^2} - a \arccos \frac{a}{u} + C = \sqrt{u^2 - a^2} - a \operatorname{arcsec} \frac{u}{a} + C$$

$$\int \frac{\sqrt{u^2 - a^2}}{u^2} du = -\frac{\sqrt{u^2 - a^2}}{u} + \ln(u + \sqrt{u^2 - a^2}) + C = -\frac{\sqrt{u^2 - a^2}}{u} + \cosh^{-1} \frac{u}{a} + C$$

$$\int \frac{u^2 du}{\sqrt{u^2 - a^2}} = \frac{u}{2} \sqrt{u^2 - a^2} + \frac{a^2}{2} \ln(u + \sqrt{u^2 - a^2}) + C = \frac{u}{2} \sqrt{u^2 - a^2} + \frac{a^2}{2} \cosh^{-1} \frac{u}{a} + C$$

$$\int \frac{du}{u\sqrt{u^2 - a^2}} = \frac{1}{a} \arccos \frac{a}{u} + C = \frac{1}{a} \operatorname{arcsec} \frac{u}{a} + C$$

$$\int \frac{du}{u^2 \sqrt{u^2 - a^2}} = \frac{\sqrt{u^2 - a^2}}{a^2 u} + C$$

$$\begin{aligned} \int (u^2 - a^2)^{3/2} du &= \frac{u}{8} (2u^2 - 5a^2) \sqrt{u^2 - a^2} + \frac{3a^4}{8} \ln(u + \sqrt{u^2 - a^2}) + C \\ &= \frac{u}{8} (2u^2 - 5a^2) \sqrt{u^2 - a^2} + \frac{3a^4}{8} \cosh^{-1} \frac{u}{a} + C \end{aligned}$$

$$\int \frac{du}{(u^2 - a^2)^{3/2}} = -\frac{u}{a^2 \sqrt{u^2 - a^2}} + C$$

Formas en las que aparece $\sqrt{2au - u^2}$

$$\int \sqrt{2au - u^2} du = \frac{u - a}{2} \sqrt{2au - u^2} + \frac{a^2}{2} \arccos \left(1 - \frac{u}{a} \right) + C$$

$$\int u \sqrt{2au - u^2} du = \frac{2u^2 - au - 3a^2}{6} \sqrt{2au - u^2} + \frac{a^3}{2} \arccos \left(1 - \frac{u}{a} \right) + C$$

$$\int \frac{\sqrt{2au - u^2}}{u} du = \sqrt{2au - u^2} + a \arccos \left(1 - \frac{u}{a} \right) + C$$

$$\int \frac{\sqrt{2au - u^2}}{u^2} du = -\frac{2\sqrt{2au - u^2}}{u} - \arccos \left(1 - \frac{u}{a} \right) + C$$

$$\int \frac{du}{\sqrt{2au - u^2}} = 2 \arcsin \sqrt{\frac{u}{2a}} + C = \arccos \left(1 - \frac{u}{a} \right) + C$$

$$\int \frac{u du}{\sqrt{2au - u^2}} = -\sqrt{2au - u^2} + a \arccos \left(1 - \frac{u}{a} \right) + C$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas en las que aparece $\sqrt{2au - u^2}$ (cont.)

$$\int \frac{u^2 du}{\sqrt{2au - u^2}} = -\frac{(u + 3a)}{2} \sqrt{2au - u^2} + \frac{3a^2}{2} \arccos \left(1 - \frac{u}{a} \right) + C$$

$$\int \frac{du}{u \sqrt{2au - u^2}} = -\frac{\sqrt{2au - u^2}}{au} + C$$

$$\int \frac{du}{(2au - u^2)^{3/2}} = \frac{u - a}{a^2 \sqrt{2au - u^2}} + C$$

Formas trigonométricas

$$\int \operatorname{sen} u \, du = -\cos u + C$$

$$\int \cos u \, du = \operatorname{sen} u + C$$

$$\int \tan u \, du = -\ln |\cos u| + C = \ln |\sec u| + C$$

$$\int \cot u \, du = \ln |\operatorname{sen} u| + C = -\ln |\csc u| + C$$

$$\int \sec u \, du = \ln |\sec u + \tan u| + C = \ln \left| \tan \left(\frac{u}{2} + \frac{\pi}{4} \right) \right| + C$$

$$\int \csc u \, du = -\ln |\csc u + \cot u| + C = \ln \left| \tan \frac{u}{2} \right| + C$$

$$\int \sec^2 u \, du = \tan u + C$$

$$\int \csc^2 u \, du = -\cot u + C$$

$$\int \sec u \tan u \, du = \sec u + C$$

$$\int \csc u \cot u \, du = -\csc u + C$$

$$\int \operatorname{sen}^2 u \, du = \frac{1}{2}(u - \operatorname{sen} u \cos u) + C = \frac{1}{2}u - \frac{1}{4} \operatorname{sen} 2u + C$$

$$\int \cos^2 u \, du = \frac{1}{2}(u + \operatorname{sen} u \cos u) + C = \frac{1}{2}u + \frac{1}{4} \operatorname{sen} 2u + C$$

$$\int \tan^2 u \, du = \tan u - u + C$$

$$\int \sec^3 u \, du = \frac{1}{2} \sec u \tan u + \frac{1}{2} \ln |\sec u + \tan u| + C$$

$$\int \operatorname{sen} mu \operatorname{sen} nu \, du = \frac{\operatorname{sen}(m-n)u}{2(m-n)} - \frac{\operatorname{sen}(m+n)u}{2(m+n)} + C$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas trigonométricas (cont.)

$$\int \operatorname{sen} mu \cos nu \, du = \frac{\cos(m-n)u}{2(m-n)} - \frac{\cos(m+n)u}{2(m+n)} + C$$

$$\int \cos mu \cos nu \, du = \frac{\operatorname{sen}(m-n)u}{2(m-n)} + \frac{\operatorname{sen}(m+n)u}{2(m+n)} + C$$

$$\int u \operatorname{sen} u \, du = \operatorname{sen} u - u \cos u + C$$

$$\int u \cos u \, du = \cos u + u \operatorname{sen} u + C$$

$$\int u^2 \operatorname{sen} u \, du = (2 - u^2) \cos u + 2u \operatorname{sen} u + C$$

$$\int u^2 \cos u \, du = (u^2 - 2) \operatorname{sen} u + 2u \cos u + C$$

$$\int \operatorname{sen}^m u \cos^n u \, du = -\frac{\operatorname{sen}^{m-1} u \cos^{n+1} u}{m+n} + \frac{m-1}{m+n} \int \operatorname{sen}^{m-2} u \cos^n u \, du$$

$$\int \operatorname{sen}^m u \cos^n u \, du = \frac{\operatorname{sen}^{m+1} u \cos^{n-1} u}{m+n} + \frac{n-1}{m+n} \int \operatorname{sen}^m u \cos^{n-2} u \, du$$

$$\int \frac{du}{a + b \cos u} = \frac{2}{\sqrt{a^2 - b^2}} \operatorname{arc} \tan \left(\frac{\sqrt{a^2 - b^2} \tan \frac{u}{2}}{a + b} \right) + C, \text{ si } a^2 > b^2$$

$$\int \frac{du}{a + b \cos u} = \frac{1}{\sqrt{b^2 - a^2}} \ln \left[\frac{a + b + \sqrt{b^2 - a^2} \tan \frac{u}{2}}{a + b - \sqrt{b^2 - a^2} \tan \frac{u}{2}} \right] + C, \text{ si } b^2 > a^2$$

$$\int \frac{du}{a + b \operatorname{sen} u} = \frac{2}{\sqrt{a^2 - b^2}} \operatorname{arc} \tan \left(\frac{a \tan \frac{u}{2} + b}{\sqrt{a^2 - b^2}} \right) + C, \text{ si } a^2 > b^2$$

$$\int \frac{du}{a + b \operatorname{sen} u} = \frac{1}{\sqrt{b^2 - a^2}} \ln \left[\frac{a \tan \frac{u}{2} + b - \sqrt{b^2 - a^2}}{a \tan \frac{u}{2} + b + \sqrt{b^2 - a^2}} \right] + C, \text{ si } b^2 > a^2$$

Formas trigonométricas inversas

$$\int \operatorname{arc} \operatorname{sen} u \, du = u \operatorname{arc} \operatorname{sen} u + \sqrt{1 - u^2} + C$$

$$\int \operatorname{arc} \cos u \, du = u \operatorname{arc} \cos u - \sqrt{1 - u^2} + C$$

$$\int \operatorname{arc} \tan u \, du = u \operatorname{arc} \tan u - \frac{1}{2} \ln(1 + u^2) + C$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas exponenciales y logarítmicas

$$\int e^u du = e^u + C$$

$$\int a^u du = \frac{a^u}{\ln a} + C$$

$$\int u e^u du = e^u(u - 1) + C$$

$$\int u^n e^u du = u^n e^u - n \int u^{n-1} e^u du$$

$$\int \frac{e^u}{u^n} du = -\frac{e^u}{(n-1)u^{n-1}} + \frac{1}{n-1} \int \frac{e^u}{u^{n-1}} du$$

$$\int \ln u du = u \ln u - u + C$$

$$\int u^n \ln u du = u^{n+1} \left[\frac{\ln u}{n+1} - \frac{1}{(n+1)^2} \right] + C$$

$$\int \frac{du}{u \ln u} = \ln(\ln u) + C$$

$$\int e^{au} \sin nu du = \frac{e^{au}(a \sin nu - n \cos nu)}{a^2 + n^2} + C$$

$$\int e^{au} \cos nu du = \frac{e^{au}(a \cos nu + n \sin nu)}{a^2 + n^2} + C$$

Formas hiperbólicas

$$\int \sinh u du = \cosh u + C$$

$$\int \cosh u du = \sinh u + C$$

$$\int \tanh u du = \ln \cosh u + C$$

$$\int \coth u du = \ln \sinh u + C$$

$$\int \operatorname{sech} u du = \arctan(\sinh u) + C$$

$$\int \operatorname{csch} u du = \ln \tanh \frac{1}{2} u + C$$

$$\int \operatorname{sech}^2 u du = \tanh u + C$$

Tabla 2-1. Breve tabla de integrales (continuación)

Formas hiperbólicas (cont.)

$$\int \operatorname{csch}^2 u \, du = -\coth u + C$$

$$\int \operatorname{sech} u \tanh u \, du = -\operatorname{sech} u + C$$

$$\int \operatorname{csch} u \coth u \, du = -\operatorname{csch} u + C$$

$$\int \sinh^2 u \, du = \frac{1}{4} \sinh 2u - \frac{1}{2}u + C$$

$$\int \cosh^2 u \, du = \frac{1}{4} \sinh 2u + \frac{1}{2}u + C$$

$$\int u \sinh u \, du = u \cosh u - \sinh u + C$$

$$\int u \cosh u \, du = u \sinh u - \cosh u + C$$

$$\int e^{au} \sinh nu \, du = \frac{e^{au}(a \sinh nu - n \cosh nu)}{a^2 - n^2} + C$$

$$\int e^{au} \cosh nu \, du = \frac{e^{au}(a \cosh nu - n \sinh nu)}{a^2 - n^2} + C$$

Fórmulas de Wallis

$$\int_0^{\frac{\pi}{2}} \sin^n u \, du = \int_0^{\frac{\pi}{2}} \cos^n u \, du = \begin{cases} \frac{(n-1)(n-3)\cdots 4\cdot 2}{n(n-2)\cdots 5\cdot 3\cdot 1}, & \text{si } n \text{ es un entero impar } > 1 \\ \frac{(n-1)(n-3)\cdots 3\cdot 1}{n(n-2)\cdots 4\cdot 2} \cdot \frac{\pi}{2}, & \text{si } n \text{ es un entero positivo par} \end{cases}$$

$$\int_0^{\frac{\pi}{2}} \sin^m u \cos^n u \, du = \begin{cases} \frac{(n-1)(n-3)\cdots 4\cdot 2}{(m+n)(m+n-2)\cdots (m+5)(m+3)(m+1)}, & \text{si } n \text{ es un entero impar } > 1 \\ \frac{(m-1)(m-3)\cdots 4\cdot 2}{(n+m)(n+m-2)\cdots (n+5)(n+3)(n+1)}, & \text{si } m \text{ es un entero impar } > 1 \\ \frac{(m-1)(m-3)\cdots 3\cdot 1 \cdot (n-1)(n-3)\cdots 3\cdot 1}{(m+n)(m+n-2)\cdots 4\cdot 2} \cdot \frac{\pi}{2}, & \text{si } m \text{ y } n \text{ son enteros positivos pares} \end{cases}$$

Fuente: L., Smail, *Calculus*, Appleton-Century-Crofts, Nueva York.

2.4 MATRICES

Los arreglos rectangulares de números o funciones como:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

se denominan matrices de orden (m, n) o matrices de $m \times n$. Los números o funciones a_{ij} se denominan elementos de la matriz, donde el primer subíndice indica el número de renglón y el segundo el número de columna. Una matriz de orden $m \times n$ tiene m renglones y n columnas. Dos matrices son iguales si su orden es el mismo y sus elementos correspondientes son iguales. Por ejemplo, si dos matrices A y B del mismo orden son iguales, entonces $a_{ij} = b_{ij}$ para todo i y todo j .

2.4.1 Matrices especiales

Matriz columna. Una matriz de una columna y cualquier número de renglones se conoce como matriz columna o vector.

Matriz cuadrada. Una matriz cuadrada de orden (n, n) es una matriz cuadrada de orden n . La diagonal principal de una matriz cuadrada consta de los elementos $a_{11}, a_{11}, a_{33}, \dots, a_{nn}$. Una matriz cuadrada en la que todos los elementos con excepción de los elementos de la diagonal principal son iguales a cero, se denomina **matriz diagonal**. Además, si todos los elementos de una matriz diagonal son iguales a 1, la matriz se denomina **matriz unitaria o identidad**.

Matriz cero. Si todos los elementos de una matriz son cero, $a_{ij} = 0$, la matriz se conoce como matriz cero.

Matriz simétrica. Si $a_{ij} = a_{ji}$ en una matriz, ésta se denomina matriz simétrica.

2.4.2 Operaciones con matrices

Adición y sustracción. La suma de dos matrices del mismo orden se determina sumando los elementos correspondientes. Si los elementos de A son a_{ij} y los de B son b_{ij} , y si $C = A + B$, entonces $C_{ij} = a_{ij} + b_{ij}$ y evidentemente $A + B = B + A$ para todas las matrices.

Si una matriz se multiplica por una constante α , entonces todo elemento de A se multiplica por α , lo cual se expresa de la siguiente forma:

$$\alpha A = [\alpha a_{ij}]$$

En particular, si $\alpha = -1$, entonces $-A = [-a_{ij}]$. A partir de esto se observa que para restar B de A se multiplican todos los elementos de B por -1 y se suma la matriz resultante a A .

Si $C = A - B$, entonces $C_{ij} = a_{ij} - b_{ij}$.

Multiplicación. La multiplicación de matrices sólo está definida si el número de columnas de A es igual al número de renglones de B .

Si A es una matriz de orden $m \times n$ y B es de orden $n \times p$, entonces el producto AB es una matriz C de orden $m \times p$.

$$C_{m \times p} = A_{m \times n} B_{n \times p}$$

Los elementos de C se determinan a partir de los elementos de A y de B , multiplicando el i -ésimo renglón de A por la j -ésima columna de B y sumando dichos productos para obtener C_{ij} :

$$C_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ip}b_{pj} = \sum_{k=1}^p a_{ik}b_{kj}$$

En general, si AB está definida, BA no necesariamente lo está, o inclusive si BA está definida, $AB \neq BA$.

Por último, si U (matriz identidad) y A son ambas cuadradas y de orden n , $UA = AU = A$.

2.4.3 Otras definiciones

Transpuesta. La *transpuesta* de una matriz A se denota por A^T y se forma intercambiando los renglones y las columnas de A . Por ejemplo, si:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{entonces } A^T = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -1 & 1 \end{bmatrix}$$

Determinante. El *determinante* de una matriz cuadrada tiene por elementos los elementos de dicha matriz:

$$\det A = \det [a_{ij}] = [a_{ij}]$$

Cofactor. El *cofactor* A_{ij} del elemento a_{ij} de una matriz se define para una matriz cuadrada A . Su valor es $(-1)^{i+j}$ veces el determinante formado al eliminar el i -ésimo renglón y la j -ésima columna en el determinante de A .

Matriz adjunta. La *matriz adjunta* de una matriz cuadrada A se forma sustituyendo cada elemento a_{ij} por el cofactor A_{ij} y transponiendo, lo cual se puede expresar como:

$$\text{adjunta de } A = [A_{ij}]^T$$

Matriz inversa. La *matriz inversa* de A , A^{-1} , se define como la matriz adjunta dividida entre el determinante de A :

$$A^{-1} = \frac{\text{adjunta de } A}{\det A}, \quad \det A \neq 0$$

Ejemplo. Dada:

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, \det A = 3$$

entonces:

$$[A_{ij}] = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, [A_{ij}]^T = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$$

$$\text{y } A^{-1} = \frac{1}{3} \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Una propiedad de la matriz inversa A^{-1} es que:

$$AA^{-1} = A^{-1}A = 1$$

En el ejemplo dado:

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1$$

$$\text{y } A^{-1}A = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1$$

2.4.4 Solución matricial de ecuaciones lineales simultáneas

Como ejemplo de un conjunto de ecuaciones lineales simultáneas, considérense las ecuaciones del voltaje de Kirchhoff, que tienen la forma:

$$ZI = V$$

o bien:

$$\sum_{j=1}^n Z_{ij} I_j = V_i$$

En rotación matricial, Z es una matriz cuadrada de orden n , e I y V son matrices columna.

$$\begin{bmatrix} Z_{11} & \dots & Z_{1n} \\ Z_{21} & \dots & Z_{2n} \\ \vdots & \vdots & \vdots \\ Z_{n1} & \dots & Z_{nn} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix}$$

Para resolver estas ecuaciones utilizando matrices, al multiplicar $ZI = V$ por Z^{-1} se obtiene:

$$Z^{-1}ZI = Z^{-1}V$$

pero:

$$Z^{-1}Z = U \text{ y } UI = I$$

de modo que:

$$I = Z^{-1}V = \frac{\text{adjunta de } Z}{\det Z} V$$

que es equivalente a la regla de Cramer.

2.5 VECTORES

2.5.1 Álgebra vectorial

Componentes y vectores unitarios. Un vector \mathbf{A} puede especificarse por medio de sus componentes (proyecciones) a lo largo de cualesquiera tres ejes perpendiculares entre sí, A_x, A_y, A_z . El vector puede expresarse de manera única en términos de sus componentes utilizando los vectores unitarios $\hat{i}, \hat{j}, \hat{k}$, que se definen como vectores de magnitud igual a 1 en las direcciones positivas x, y y z .

$$\mathbf{A} = A_x\hat{i} + A_y\hat{j} + A_z\hat{k}$$

Adición y sustracción. Dados dos vectores:

$$\mathbf{A} = A_x\hat{i} + A_y\hat{j} + A_z\hat{k} \text{ y } \mathbf{B} = B_x\hat{i} + B_y\hat{j} + B_z\hat{k}$$

la suma de \mathbf{A} y \mathbf{B} es:

$$\mathbf{A} + \mathbf{B} = (A_x + B_x)\hat{i} + (A_y + B_y)\hat{j} + (A_z + B_z)\hat{k}$$

Obsérvese que:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

y que:

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

El negativo de un vector \mathbf{B} se define como:

$$-\mathbf{B} = -B_x\hat{i} - B_y\hat{j} - B_z\hat{k}$$

La diferencia de los vectores \mathbf{A} y \mathbf{B} es:

$$\mathbf{A} - \mathbf{B} = (A_x - B_x)\hat{i} + (A_y - B_y)\hat{j} + (A_z - B_z)\hat{k}$$

Multiplicación por un escalar. Si a es un escalar, entonces $a\mathbf{A}$ es un vector en la dirección de \mathbf{A} si $a > 0$ y en la dirección de $-\mathbf{A}$ si $a < 0$.

$$a\mathbf{A} = aA_x\hat{i} + aA_y\hat{j} + aA_z\hat{k}$$

1. Producto escalar (punto) de vectores. Si \mathbf{A} y \mathbf{B} son vectores y θ es el ángulo entre \mathbf{A} y \mathbf{B} , entonces el producto escalar se define como el escalar:

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}| |\mathbf{B}| \cos \theta$$

De la definición de producto escalar:

$$\mathbf{B} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{B}$$

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$$

$$a\mathbf{A} \cdot \mathbf{B} = a(\mathbf{A} \cdot \mathbf{B})$$

$$\hat{i} \cdot \hat{i} = \hat{j} \cdot \hat{j} = \hat{k} \cdot \hat{k} = 1$$

$$\hat{i} \cdot \hat{j} = \hat{j} \cdot \hat{k} = \hat{k} \cdot \hat{i} = 0$$

$$\mathbf{A} \cdot \mathbf{B} = (A_x\hat{i} + A_y\hat{j} + A_z\hat{k}) \cdot (B_x\hat{i} + B_y\hat{j} + B_z\hat{k}) = A_xB_x + A_yB_y + A_zB_z$$

2. **Producto vectorial (aspa) de vectores.** El producto cruz de dos vectores \mathbf{A} y \mathbf{B} se denota por $\mathbf{A} \times \mathbf{B}$ y se define como un vector perpendicular al plano que contiene a \mathbf{A} y \mathbf{B} y que está en la dirección del movimiento axial de un tornillo con rosca a la derecha que gira de \mathbf{A} a \mathbf{B} .

La magnitud de $\mathbf{A} \times \mathbf{B}$ es:

$$|\mathbf{A} \times \mathbf{B}| = |\mathbf{A}| |\mathbf{B}| \sin \theta$$

De la definición del producto vectorial:

$$\mathbf{A} \times \mathbf{B} = -(\mathbf{B} \times \mathbf{A})$$

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) + (\mathbf{A} \times \mathbf{C})$$

$$\hat{i} \times \hat{i} = \hat{j} \times \hat{j} = \hat{k} \times \hat{k} = 0$$

$$\hat{i} \times \hat{j} = \hat{k}, \quad \hat{j} \times \hat{k} = \hat{i}, \quad \hat{k} \times \hat{i} = \hat{j},$$

$$\hat{j} \times \hat{i} = -\hat{k}, \quad \text{etc.}$$

En términos de las componentes:

$$\begin{aligned} \mathbf{A} \times \mathbf{B} &= (A_x\hat{i} + A_y\hat{j} + A_z\hat{k}) \times (B_x\hat{i} + B_y\hat{j} + B_z\hat{k}) \\ &= \hat{i}(A_yB_z - A_zB_y) + \hat{j}(A_zB_x - A_xB_z) + \hat{k}(A_xB_y - A_yB_x) \end{aligned}$$

$$= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$$

3. Triple producto escalar

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{ABC}$$

$$A_y = \begin{vmatrix} A_x & B_x & C_x \\ B_y & C_y & A_y \\ A_z & B_z & C_z \end{vmatrix}$$

y:

$$(\mathbf{ABC})(\mathbf{DEF}) = \begin{vmatrix} \mathbf{A} \cdot \mathbf{D} & \mathbf{A} \cdot \mathbf{E} & \mathbf{A} \cdot \mathbf{F} \\ \mathbf{B} \cdot \mathbf{D} & \mathbf{B} \cdot \mathbf{E} & \mathbf{B} \cdot \mathbf{F} \\ \mathbf{C} \cdot \mathbf{D} & \mathbf{C} \cdot \mathbf{E} & \mathbf{C} \cdot \mathbf{F} \end{vmatrix}$$

Otras identidades útiles son:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$$

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = \mathbf{A} \cdot [\mathbf{B} \times (\mathbf{C} \times \mathbf{D})]$$

$$= (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C})$$

$$(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D}) = [(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{D}]\mathbf{C} -$$

$$- [(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}]\mathbf{D}$$

$$= (\mathbf{ABD})\mathbf{C} - (\mathbf{ABC})\mathbf{D}$$

$$\mathbf{A} \times [\mathbf{B} \times (\mathbf{C} \times \mathbf{D})] = (\mathbf{B} \cdot \mathbf{D})(\mathbf{A} \times \mathbf{C}) -$$

$$- (\mathbf{B} \cdot \mathbf{C})(\mathbf{A} \times \mathbf{D})$$

$$(\mathbf{A} \times \mathbf{B}) \cdot [(\mathbf{B} \times \mathbf{C}) \times (\mathbf{C} \times \mathbf{A})] = [\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})]^2 = (\mathbf{ABC})^2$$

2.5.2 Análisis vectorial

Operadores diferenciales. El gradiente de una función escalar $\Phi(x, y, z)$ se define como:

$$\text{grad } \Phi \equiv \nabla \Phi = \hat{i} \frac{\partial \Phi}{\partial x} + \hat{j} \frac{\partial \Phi}{\partial y} + \hat{k} \frac{\partial \Phi}{\partial z}$$

La divergente de un vector \mathbf{A} se define como:

$$\text{div } \mathbf{A} \equiv \nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

El rotacional de un vector \mathbf{A} se define como:

$$\begin{aligned} \text{rot } \mathbf{A} \equiv \nabla \times \mathbf{A} &= \begin{vmatrix} \hat{i} & \frac{\partial}{\partial x} & A_x \\ \hat{j} & \frac{\partial}{\partial y} & A_y \\ \hat{k} & \frac{\partial}{\partial z} & A_z \end{vmatrix} \\ &= \hat{i} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \hat{j} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \hat{k} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \end{aligned}$$

El laplaciano de una función escalar Φ se define como:

$$\nabla^2 \Phi \equiv \nabla \cdot \nabla \Phi = \text{div grad } \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2}$$

El laplaciano de una función vectorial $A(x, y, z)$ se define como:

$$\nabla^2 \mathbf{A} = (\text{div grad}) \mathbf{A} = \hat{i} \nabla^2 A_x + \hat{j} \nabla^2 A_y + \hat{k} \nabla^2 A_z$$

Algunas operaciones diferenciales de utilidad son:

$$\text{grad} (\Phi + \psi) = \text{grad} \Phi + \text{grad} \psi$$

donde Φ y ψ son funciones escalares.

$$\text{grad} (\Phi \psi) = \Phi \text{grad} \psi + \psi \text{grad} \Phi$$

$$\text{div} (\mathbf{A} + \mathbf{B}) = \text{div} \mathbf{A} + \text{div} \mathbf{B}$$

$$\text{rot} (\mathbf{A} + \mathbf{B}) = \text{rot} \mathbf{A} + \text{rot} \mathbf{B}$$

$$\text{div} (\Phi \mathbf{A}) = \mathbf{A} \cdot \text{grad} \Phi + \Phi \text{div} \mathbf{A}$$

$$\text{rot} (\Phi \mathbf{A}) = \Phi \text{rot} \mathbf{A} - \mathbf{A}_x \text{grad} \Phi$$

$$(\mathbf{A} \cdot \text{grad}) \mathbf{B} = A_x \frac{\partial \mathbf{B}}{\partial x} + A_y \frac{\partial \mathbf{B}}{\partial y} + A_z \frac{\partial \mathbf{B}}{\partial z} =$$

$$\hat{i}(\mathbf{A} \cdot \nabla B_x) + \hat{j}(\mathbf{A} \cdot \nabla B_y) + \hat{k}(\mathbf{A} \cdot \nabla B_z)$$

$$\text{grad} (\mathbf{A} \cdot \mathbf{B}) = (\mathbf{A} \cdot \text{grad}) \mathbf{B} + (\mathbf{B} \cdot \text{grad}) \mathbf{A} + \mathbf{A} \times \text{rot} \mathbf{B} + \mathbf{B} \times \text{rot} \mathbf{A}$$

$$\text{div} (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \text{rot} \mathbf{A} - \mathbf{A} \cdot \text{rot} \mathbf{B}$$

$$\text{rot} (\mathbf{A} \times \mathbf{B}) = \mathbf{A} \text{div} \mathbf{B} - \mathbf{B} \text{div} \mathbf{A} + (\mathbf{B} \cdot \text{grad}) \mathbf{A} - (\mathbf{A} \cdot \text{grad}) \mathbf{B}$$

$$\text{rot rot} \mathbf{A} = \text{grad div} \mathbf{A} - \nabla^2 \mathbf{A}$$

$$\text{rot grad} \Phi = 0$$

$$\text{div rot} \mathbf{A} = 0$$

2.6 ECUACIONES DIFERENCIALES LINEALES

2.6.1 Definiciones

Se considerará la ecuación diferencial lineal con coeficientes constantes, que puede escribirse de la forma:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y =$$

$$= b_m \frac{d^m x}{dt^m} + \dots + b_1 \frac{dx}{dt} + b_0 x$$

donde $x(t)$ es la entrada e $y(t)$ la salida. Dado que $x(t)$ suele conocerse, la ecuación anterior será:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = F(t)$$

Con frecuencia $F(t)$ se denomina función de forzamiento, y cuando ésta es distinta de cero la ecuación diferencial se llama **no homogénea**. Si $F(t)$ es igual a cero, la ecuación diferencial es **homogénea**.

2.6.2 Solución de la ecuación diferencial homogénea

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = 0$$

La ecuación homogénea tiene n soluciones linealmente independientes $y_1(t), y_2(t), \dots, y_n(t)$. La solución general es la combinación lineal de dichas soluciones:

$$y_h(t) = A_1 y_1(t) + A_2 y_2(t) + \dots + A_n y_n(t)$$

donde las A son constantes arbitrarias. La solución debe satisfacer n condiciones iniciales, que suelen conocerse, a saber:

$$y(0), \frac{dy}{dt}(0), \dots, \frac{d^{n-1} y}{dt^{n-1}}(0)$$

Los valores específicos de las A deben encontrarse después, de modo que la solución $y_h(t)$ satisfaga las condiciones iniciales. Se supone que las soluciones de la ecuación homogénea son de la forma:

$$y(t) = e^{pt}$$

Aquí p es una constante por determinar. Sustituyendo esta solución hipotética en la ecuación diferencial, se concluye que la suposición es correcta si p es una raíz de la ecuación algebraica:

$$a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + a_0 = 0$$

Esta ecuación algebraica, denominada **ecuación característica**, produce n valores de p , por ejemplo, p_1, p_2, \dots, p_n .

Supóngase que estos valores son reales y que todos son diferentes entre sí. Entonces, las soluciones de la ecuación diferencial homogénea son:

$$y_1 = e^{p_1 t}, y_2 = e^{p_2 t}, \dots, y_n = e^{p_n t}$$

Y la solución homogénea es:

$$y_h(t) = A_1 e^{p_1 t} + A_2 e^{p_2 t} + \dots + A_n e^{p_n t}$$

Si las raíces de la ecuación característica son complejas, entonces deben presentarse en pares complejos conjugados (si los coeficientes de la ecuación diferencial original son números reales). De este modo, si $p_1 = -\alpha + j\omega$, entonces $p_2 = -\alpha - j\omega$ e $y_1 = e^{-\alpha t} (C_1 \cos \omega t + C_2 \sin \omega t)$.

Si la ecuación característica produce raíces repetidas, por ejemplo que p_1 ocurra q veces, entonces:

$$y_h(t) = (A_1 + A_2 t + \dots + A_q t^{q-1}) e^{p_1 t} + \dots + A_n e^{p_n t}$$

Deben mencionarse otros nombres dados a la solución homogénea, como "solución independiente de la fuente", "solución natural" y "respuesta transitoria".

2.6.3 Solución de ecuaciones diferenciales no homogéneas

La solución de la ecuación no homogénea $F(t) \neq 0$ consta de la suma de dos partes:

$$y(t) = y_h(t) + y_p(t)$$

donde $y_h(t)$ es la solución homogénea, que contiene n constantes arbitrarias, e $y_p(t)$ se conoce como solución particular, solución forzada, solución complementaria o solución debida a la fuente.

La solución completa $y(t)$ debe satisfacer la ecuación diferencial y debe contener n constantes arbitrarias para satisfacer las condiciones iniciales. Debido a que $y_h(t)$ contiene estas n constantes arbitrarias, $y_p(t)$ debe tener determinadas todas sus constantes (es decir, éstas deben tener valores específicos).

Método de los coeficientes indeterminados. El procedimiento que se presenta a continuación se utiliza cuando la función excitadora o de entrada $F(t)$ es alguna de las funciones que suelen encontrarse en la ingeniería eléctrica. El procedimiento consiste en suponer una solución particular cuya forma satisfaga la forma de la solución de una ecuación diferencial no homogénea. Esto suele significar que se supone que $y_p(t)$ es la combinación lineal de las funciones de la forma de $F(t)$ y de todas sus derivadas sucesivas. Si un término de $F(t)$ es alguna suma de la forma de las que se muestran en la lista que sigue, entonces se supone que $y_p(t)$ es la suma de la elección correspondiente (en la que se combinan funciones de tiempo semejantes). Si un término de $F(t)$ ya aparece como un término de la solución homogénea, entonces la forma $y_p(t)$ de la lista debe multiplicarse por t .

$F(t)$	Para $y_p(t)$, suponer
Constante	Constante
At	$K_1t + K_0$
At^2	$K_2t^2 + K_1t + K_0$
At^n	$K_nt^n + K_{n-1}t^{n-1} + \dots + K_1t + K_0$
Ae^{mt}	Ke^{mt}
$At^n e^{mt}$	$e^{mt}(K_nt^n + \dots K_1 + K_0)$
$A \cos wt$	$K_1 \cos wt + K_2 \sin wt$
$A \sin wt$	$K_1 \cos wt + K_2 \sin wt$
$At^n \cos wt$	$(K_nt^n + \dots + K_1t + K_0) \cos wt$ $+ (C_nt^n + \dots + C_1t + C_0) \sin wt$
$Ae^{mt} \cos wt$	$e^{mt}(K_1 \cos wt + K_2 \sin wt)$

Ejemplo. Si:

$$y_h(t) = A + Be^{-t} + Ce^{-2t} \text{ y } F(t) = 6 + 3e^{-t}$$

entonces sea:

$$y_p(t) = K_1t + K_2te^{-t}$$

El procedimiento para determinar las constantes que debe contener $y_p(t)$ es:

1. Suponer una solución para $y_p(t)$ de la lista anterior.
2. Sustituir esta solución en la ecuación diferencial:

$$a_n \frac{d^n y}{dt^n} + \dots + a_1 \frac{dy}{dt} + a_0 y = F(t)$$

3. Igualar los coeficientes de las funciones del tiempo parecidas del miembro izquierdo de la ecuación con las del miembro derecho.
4. Despejar las constantes (K) que aparezcan en la $y_p(t)$ supuesta.

2.6.4 Solución completa de ecuaciones diferenciales no homogéneas

Debe seguirse el siguiente procedimiento:

1. Despejar $y_h(t)$, la solución de la ecuación homogénea.
2. Despejar $y_p(t)$, la solución particular.
3. $y(t) = y_h(t) + y_p(t)$.
4. Dado que la solución completa debe satisfacer las condiciones iniciales, ahora se despejan las constantes arbitrarias, de modo que en efecto se satisfagan dichas condiciones iniciales.

2.7 ECUACIONES EN DIFERENCIAS LINEALES

2.7.1 Ecuaciones en diferencias

Estas ecuaciones se presentan cuando las funciones no varían de forma continua en el tiempo, sino que tienen sentido sólo en intervalos discretos de tiempo. Ciertamente, cualquier sistema que contenga componentes digitales implica el procesamiento de números o señales discretos, conforme a alguna regla específica. Esta regla es una fórmula recurrente o una ecuación en diferencias, que relaciona el valor de una función o señal en algún instante k con su valor en otros instantes específicos. También puede considerarse una relación entre una muestra de una señal o una entrada y otras muestras de la salida.

Una ecuación en diferencias lineal de n -ésimo orden con coeficientes constantes es:

$$a_n y_{k+n} + a_{n-1} y_{k+n-1} + \dots + a_1 y_{k+1} + a_0 y_k = U_k$$

Aquí U_k es la sucesión de entrada y actúa de manera análoga a como lo hace la función de forzamiento en la ecuación diferencial. La similitud entre la solución de las ecuaciones en diferencias y la solución de las ecuaciones diferenciales se utilizará a menudo.

La solución general de la ecuación en diferencias general debe satisfacer n valores muestrales conocidos (o condiciones iniciales). Es posible especificar valores para:

$$y_0, y_1, y_2, \dots, y_{n-1}$$

De nuevo, la solución completa y_k es la suma de la solución homogénea y la solución particular; es decir, $y_k = y_k$ homogénea + y_k particular.

2.7.2 Solución de ecuaciones en diferencias homogéneas

Las ecuaciones, cuando el valor de $u_k = 0$, son homogéneas:

$$a_n y_{k+n} + a_{n-1} y_{k+n-1} + \dots + a_1 y_{k+1} + a_0 y_k = 0$$

De nuevo, supóngase una solución de la forma que se indica a continuación:

$$y_k = e^{bk} = b^k \quad \text{donde } b = e^p$$

Si este resultado se sustituye en la ecuación en diferencias puede concluirse que la hipótesis es correcta para los n valores de b , que son las soluciones de la ecuación algebraica:

$$a_n b^n + a_{n-1} b^{n-1} + \dots + a_1 b + a_0 = 0$$

Si las soluciones de esta "ecuación característica" son valores reales y no repetidos, entonces:

$$y_k = b_1^k, y_k = b_2^k, \dots, y_k = b_n^k$$

La combinación lineal de estas soluciones es la solución homogénea completa:

$$y_k = A_1 b_1^k + A_2 b_2^k + \dots + A_n b_n^k$$

Obsérvese que los valores de b son números. Si la ecuación característica produce un valor de b repetido q veces, entonces:

$$y_k = (A_1 + A_2 k + A_3 k^2 + \dots + A_q k^{q-1}) b_1^k + \dots + A_k b_n^k$$

Si se produce un valor complejo para b y los coeficientes de la ecuación en diferencias son reales, entonces también el conjugado complejo de b será una raíz:

$$\text{Si } b_1 = \alpha + j\beta \text{ entonces } b_2 = \alpha - j\beta$$

$$\text{entonces } y_k = r^k (C_1 \cos \theta k + C_2 \sin \theta k) + \dots + A_k b_n^k$$

$$\text{donde } r^k = \sqrt{\alpha^2 + \beta^2} \text{ y } \theta = \tan^{-1} \frac{\beta}{\alpha}$$

Las raíces complejas múltiples se tratan en la forma acostumbrada.

2.7.3 Solución de ecuaciones en diferencias no homogéneas

$$y_k = y_k \text{ homogénea} + y_k \text{ particular}$$

Para encontrar la y_k particular se utilizará el método de los coeficientes indeterminados.

U_k	Para y_k particular suponer
a^k o e^{rk}	Ca^k
k^p	$C_p k^p + C_{p-1} k^{p-1} + \dots + C_1 k + C_0$
$\cos wk$	$C_1 \cos wk + C_2 \sin wk$
$\sin wk$	$C_1 \cos wk + C_2 \sin wk$
$a^k \cos wk$	$a^k (C_1 \cos wk + C_2 \sin wk)$
$k^p a^k$	$a^k (C_p k^p + \dots + C_1 k + C_0)$
$k^p \cos wk$	$(C_p k^p + C_1 k + C_0) \cos wk$ $+ (D_p k^p + \dots + D_1 k + D_0) \sin wk$

Si U_k contiene una suma de funciones, se supone que la solución es la suma de las hipótesis individuales (combinando los términos parecidos, por supuesto). Si la forma funcional de un término en U_k es igual a la forma funcional de algún término de la solución homogénea, entonces la solución supuesta debida a dicho término se multiplica por k .

En el caso de términos de la forma $\cos wk$ y $\sin wk$ es importante recordar que:

$$\cos w(k+m) = (\cos wm) \cos wk - (\sin wm) \sin wk$$

$$\sin w(k+m) = (\sin wm) \cos wk + (\cos wm) \sin wk$$

El procedimiento para determinar los coeficientes (C) de la y_k particular es:

1. Suponer una solución para la y_k particular.
2. Sustituir esta solución en la ecuación en diferencias:

$$a_n y_{k+n} + a_{n-1} y_{k+n-1} + \dots + a_1 y_{k+1} + a_0 y_k = U_k$$

Obsérvese que esta sustitución implica encontrar los valores $k+1, k+2, \dots, k+n$ de la solución particular de que se trate. Por ejemplo, si se supone que y_k particular = $C_1 3^k$, entonces:

$$y_{k+1} = C_1 3^{k+1} = 3C_1 3^k$$

$$y_{k+2} = C_1 3^{k+2} = 9C_1 3^k$$

$$y_{k+3} = C_1 3^{k+3} = 27C_1 3^k \text{ etc.}$$

3. Igualar los coeficientes de los términos parecidos en k del lado izquierdo de la actuación con los del lado derecho.
4. Despejar las C (constantes) que aparecen en la y_k particular.

2.7.4 Solución completa de ecuaciones en diferencias

Se sigue el procedimiento descrito a continuación:

1. Despejar y_k homogénea.
2. Despejar y_k particular.
3. $y_k = y_k$ homogénea + y_k particular.
4. Ahora, despejar las constantes arbitrarias en y_k homogénea, de modo que se satisfagan los valores muestrales iniciales.

2.8 SERIES DE POTENCIAS

A continuación se presentan las expresiones en forma cerrada de algunas series frecuentes.

2.8.1 Progresión aritmética

$$a + (a+d) + (a+2d) + \dots n \text{ términos} = \frac{n}{2} [2a + (n-1)d]$$

2.8.2 Progresión geométrica

$$a + ar + ar^2 + \dots n \text{ términos} = a \left[\frac{1 - r^n}{1 - r} \right]$$

$$a + ar + ar^2 + \dots \infty \text{ términos} = \frac{a}{1 - r} \text{ donde } r < 1$$

$$a + ax + a^2x^2 + \dots \infty \text{ términos} = \frac{1}{1 - ax}$$

2.8.3 Suma de potencias de números naturales

$$\sum_{1}^{\infty} x^n = \frac{x}{1 - x} \text{ donde } x < 1$$

$$\sum_{1}^{\infty} nx^n = \frac{x}{(1 - x)^2} \text{ donde } x < 1$$

2.8.4 Series de potencias de funciones comunes

$$e^{ax} = 1 + ax + \frac{a^2x^2}{2!} + \frac{a^3x^3}{3!} + \dots$$

$$\text{sen } ax = ax - \frac{a^3x^3}{3!} + \frac{a^5x^5}{5!} - \dots$$

$$\cos ax = 1 - \frac{a^2x^2}{2!} + \frac{a^4x^4}{4!} - \dots$$

$$\frac{1}{x} = 1 - (x-1) + (x-1)^2 - (x-1)^3 + (x-1)^4 - \dots$$

2.8.5 Desarrollo en serie de Taylor

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \dots$$

Esta fórmula se cumple para puntos cercanos a $x = a$ y para funciones cuyas derivadas en el punto $x = a$ existen.

Si $a = 0$, la serie se denomina serie de MacLaurin y adopta la forma:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)x^2}{2!} + \dots + \frac{f^{(n)}(0)}{n!}x^n + \dots$$

Si la serie de Taylor se termina después de n términos, el residuo es:

$$R_n = \frac{f^{(n+1)}(z)}{(n+1)!} (x - c)^{n+1}$$

donde z es algún número entre c y x .

2.9 SERIE DE FOURIER

2.9.1 Serie trigonométrica de Fourier

Las funciones periódicas son aquellas tales que $f(t) = f(t + T)$. Si en un periodo la función es univaluada, tiene un número finito de máximos y mínimos, tiene un número finito de discontinuidades y satisface la expresión:

$$\int_{t_0}^{t_0+T} |x(t)| dt < \infty$$

entonces puede desarrollarse en una serie de Fourier de la forma:

$$f(t) = A_0 + \sum_{k=1}^{\infty} [A_k \cos k\omega_0 t + B_k \text{sen } k\omega_0 t]$$

Aquí $\omega_0 = 2\pi/T = 2\pi f$ se denomina frecuencia fundamental y $k2\pi/T$ es el k -ésimo armónico. Los coeficientes deben evaluarse utilizando:

$$A_0 = \frac{1}{T} \int_{t_0}^{t_0+T} f(t) dt$$

Obsérvese que lo anterior también puede interpretarse como el valor promedio de $f(t)$ en un periodo:

$$A_k = \frac{2}{T} \int_{t_0}^{t_0+T} f(t) \cos k\omega_0 t dt$$

$$B_k = \frac{2}{T} \int_{t_0}^{t_0+T} f(t) \text{sen } k\omega_0 t dt$$

La integración puede empezar en cualquier lugar conveniente, pero debe comprender un periodo. Si $t_0 = 0$, entonces:

$$\int_0^T$$

pero si $t_0 = -T/2$, entonces la forma es:

Tabla 2-2. Coeficientes de Fourier de formas simétricas de onda periódicas

Tipo de simetría	Condiciones	Forma de la serie de Fourier	Fórmulas para los coeficientes
Función impar	$f(t) = -f(-t)$	$f(t) = \sum_{k=1}^{\infty} B_k \sin k \frac{2\pi}{T} t$	$A_k = 0; B_k = \frac{4}{T} \int_0^{T/2} f(t) \sin k \frac{2\pi}{T} t dt$
Función par	$f(t) = f(-t)$	$f(t) = \sum_{k=0}^{\infty} A_k \cos k \frac{2\pi}{T} t$	$B_k = 0; A_k = \frac{4}{T} \int_0^{T/2} f(t) \cos k \frac{2\pi}{T} t dt; A_0 = \frac{2}{T} \int_0^{T/2} f(t) dt$
Semionda	$f(t) = -f\left(t + \frac{T}{2}\right)$	$f(t) = \sum_{i=1}^{\infty} \left[A_{2k-1} \cos(2k-1) \frac{2\pi}{T} t + B_{2k-1} \sin(2k-1) \frac{2\pi}{T} t \right]$	$\begin{aligned} A_{2k-1} \\ B_{2k-1} \end{aligned} \left\} = \frac{4}{T} \int_0^{T/2} f(t) \begin{aligned} \cos \\ \sin \end{aligned} (2k-1) \frac{2\pi}{T} t dt$
Cuarto de onda impar	$f(t) = -f(-t)$ y $f(t) = -f\left(t + \frac{T}{2}\right)$	$f(t) = \sum_{k=1}^{\infty} B_{2k-1} \sin(2k-1) \frac{2\pi}{T} t$	$A_k = 0; B_{2k-1} = \frac{8}{T} \int_0^{T/4} f(t) \sin(2k-1) \frac{2\pi}{T} t dt$
Cuarto de onda par	$f(t) = f(-t)$ y $f(t) = -f\left(t + \frac{T}{2}\right)$	$f(t) = \sum_{k=1}^{\infty} A_{2k-1} \cos(2k-1) \frac{2\pi}{T} t$	$B_k = 0; A_{2k-1} = \frac{8}{T} \int_0^{T/4} f(t) \cos(2k-1) \frac{2\pi}{T} t dt$

Fuente: Javid y Brenner, reproducido con permiso.

Tabla 2-3. Las tres formas de una serie de Fourier

Forma compleja	Forma trigonométrica 1	Forma trigonométrica 2
$f(t) = \sum_{k=-\infty}^{+\infty} \mathbf{D}_k e^{jk(2\pi/T)t}$	$f(t) = C_0 + \sum_{k=1}^{\infty} C_k \cos\left(k \frac{2\pi}{T} t + \varphi_k\right)$ $= F_0 + \sum_{k=1}^{\infty} \sqrt{2} F_k \cos\left(k \frac{2\pi}{T} t + \varphi_k\right)$ $= D_0 + \sum_{k=1}^{\infty} 2 \operatorname{Re}[\mathbf{D}_k e^{jk(2\pi/T)t}]$	$f(t) = A_0 + \sum_{k=1}^{\infty} \left(A_k \cos k \frac{2\pi}{T} t + B_k \sin k \frac{2\pi}{T} t \right)$
<p>Fórmulas de conversión para $k = 1, 2, 3, \dots$</p> $\mathbf{D}_k = \frac{1}{2} C_k < \varphi_k \quad \mathbf{D}_{-k} = \mathbf{D}_k^* \quad \mathbf{D}_k = \frac{1}{2} A_k - j \frac{1}{2} B_k$ $C_k = 2 D_k \quad \varphi_k = \text{ángulo } \mathbf{D}_k \quad C_k = \sqrt{A_k^2 + B_k^2} \quad \varphi_k = -\tan^{-1} \frac{B_k}{A_k} \quad F_k = \frac{C_k}{\sqrt{2}}$ $A_k = 2 \operatorname{Re}[\mathbf{D}_k] \quad B_k = -\operatorname{Im}[\mathbf{D}_k] \quad A_k = C_k \cos \varphi_k \quad B_k = -C_k \sin \varphi_k$ <p>Para $k = 0$,</p> $D_0 = C_0 = A_0 = F_0 \quad B_0 \equiv 0$		

Fuente: Javid y Bremer, reproducido con autorización.

$$\int_{-T/2}^{T/2}$$

En la tabla 2-2 se muestra la forma de estas ecuaciones para situaciones en las que $f(t)$ presenta ciertas propiedades definidas de simetría.

Una forma alternativa de la serie de Fourier es:

$$f(t) = \sum_{k=0}^{\infty} C_k \cos \left(k \frac{2\pi}{T} t + \phi_k \right)$$

2.9.2 Serie exponencial de Fourier

Empleando la ecuación de Euler:

$$e^{j\theta} = \cos \theta + j \sin \theta$$

es posible desarrollar una forma exponencial de la serie de Fourier:

$$f(t) = \sum_{-\infty}^{\infty} D_k e^{jk(2\pi/T)t}$$

donde:

$$D_k = \frac{1}{T} \int_{t_0}^{t_0+T} f(t) e^{-jk(2\pi/T)t} dt$$

En la tabla 2-3 se presentan relaciones entre los coeficientes de las distintas formas de la serie de Fourier.

2.9.3 Teoremas en los que se utilizan los coeficientes de Fourier

Superposición. Si una función periódica puede expresarse como la suma de varias funciones periódicas más sencillas, entonces la serie de Fourier de la suma es igual a la suma de las series de Fourier individuales.

Desplazamiento en el tiempo. Si la serie de Fourier de una función periódica $f(t)$ es:

$$f(t) = \sum_{-\infty}^{\infty} D_k e^{jk(2\pi/T)t}$$

entonces el coeficiente de la función retrasada $f(t - a)$ es:

$$[D_k] \text{ función retrasada} = [D_k] \text{ función original} \cdot e^{-jk(2\pi/T)a}$$

Diferenciación. Si $f(t)$ es diferenciable:

$$f(t) = \sum_{-\infty}^{\infty} D_k e^{jk(2\pi/T)t}$$

entonces:

$$f'(t) = \sum_{-\infty}^{\infty} \left[jk \frac{2\pi}{T} D_k \right] e^{jk(2\pi/T)t}$$

2.9.4 Serie de Fourier de algunas funciones comunes

Ejemplo 1. Forma de onda rectangular (fig. 2-6)

$$f(t) = V \sum_{-\infty}^{\infty} \frac{1 - e^{jk(2\pi/T)T_1}}{jk2\pi} e^{jk(2\pi/T)t}$$

Para $T_1 = T/2$:

$$f(t) = V \left(\frac{1}{2} + \frac{2}{\pi} \sin \frac{2\pi}{T} t + \frac{2}{3\pi} \sin \frac{6\pi}{T} t + \dots \right)$$

Para $T_1 = T/4$:

$$f(t) = V \left(\frac{3}{4} - \frac{1}{\pi} \cos \frac{2\pi}{T} t + \frac{1}{\pi} \sin \frac{4\pi}{T} t + \frac{1}{3\pi} \cos \frac{6\pi}{T} t + \frac{1}{3\pi} \sin \frac{6\pi}{T} t - \frac{1}{5\pi} \cos \frac{10\pi}{T} t + \frac{1}{5\pi} \sin \frac{10\pi}{T} t + \dots \right)$$

Ejemplo 2. Onda cuadrada (fig. 2-7)

$$f(t) = \frac{4V}{\pi} \left(\cos \frac{2\pi}{T} t - \frac{1}{3} \cos \frac{6\pi}{T} t + \frac{1}{5} \cos \frac{10\pi}{T} t - \frac{1}{7} \cos \frac{14\pi}{T} t + \dots \right)$$

Ejemplo 3. Onda cosenoidal rectificada completa (fig. 2-8)

$$f(t) = \frac{4V}{\pi} \left(\frac{1}{2} + \frac{1}{3} \cos \frac{2\pi}{T} t - \frac{1}{15} \cos \frac{4\pi}{T} t + \frac{1}{35} \cos \frac{6\pi}{T} t - \dots \right)$$

Ejemplo 4. Onda senoidal rectificada completa (fig. 2-9)

$$f(t) = \frac{2V}{\pi} \sum_{-\infty}^{\infty} \frac{1}{1 - 4k^2} e^{jk2t}$$

y si $T/2 = \pi$, entonces

$$f(t) = \frac{2V}{\pi} + \frac{4V}{\pi} \sum_{k=1}^{\infty} \frac{1}{1 - 4k^2} \cos 2kt$$

Ejemplo 5. Semionda cosenoidal rectificada (fig. 2-10)

$$f(t) = \frac{V}{\pi} \left(1 + \frac{\pi}{2} \cos \frac{2\pi}{T} t + \frac{2}{3} \cos \frac{4\pi}{T} t - \frac{2}{15} \cos \frac{8\pi}{T} t + \dots \right)$$

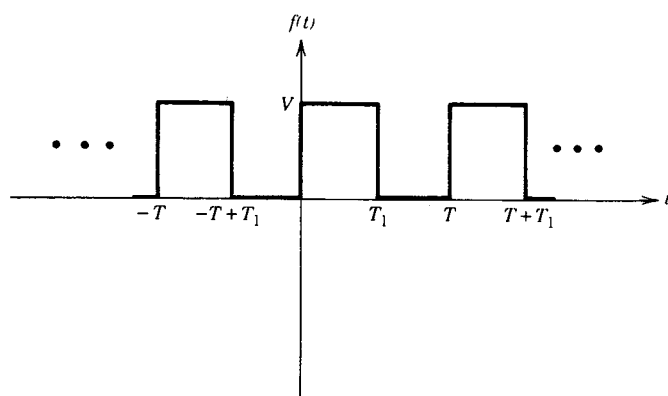


Fig. 2-6. Forma de onda rectangular.

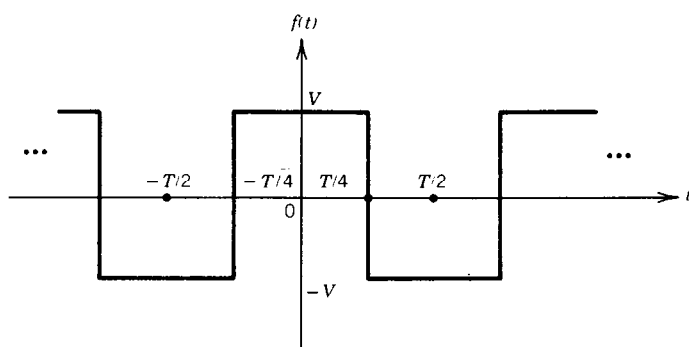


Fig. 2-7. Onda cuadrada.

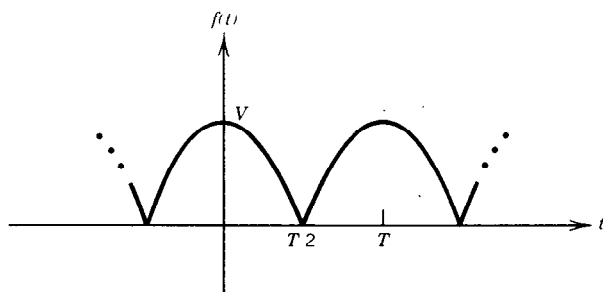


Fig. 2-8. Onda cosenoidal rectificada completa.

Ejemplo 6. Forma de onda en diente de sierra (fig. 2-11)

Ejemplo 7. Forma de onda triangular (fig. 2-12)

$$f(t) = V \left[\frac{1}{2} - \frac{1}{\pi} \left(\sin \frac{2\pi}{T} t + \frac{1}{2} \sin \frac{4\pi}{T} t + \frac{1}{3} \sin \frac{6\pi}{T} t + \dots \right) \right]$$

$$f(t) = \frac{8V}{\pi^2} \left(\sin \frac{2\pi}{T} t - \frac{1}{3^2} \sin \frac{6\pi}{T} t + \frac{1}{5^2} \sin \frac{10\pi}{T} t - \dots \right)$$

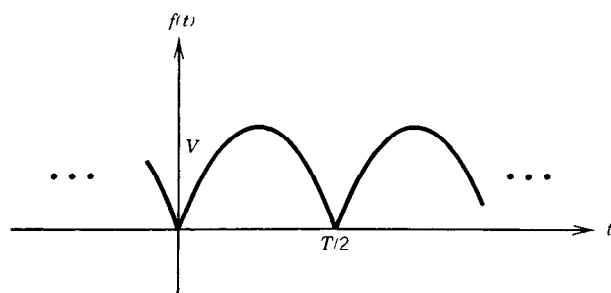


Fig. 2-9. Onda senoidal rectificada completa.

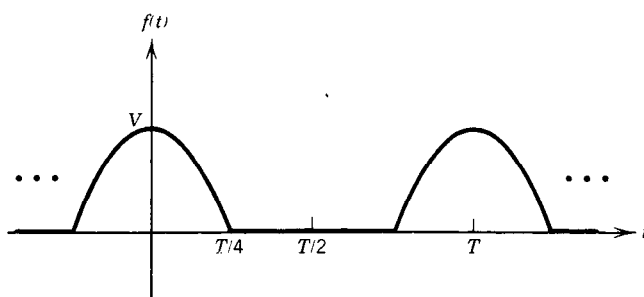


Fig. 2-10. Semionda cosenoidal rectificada.

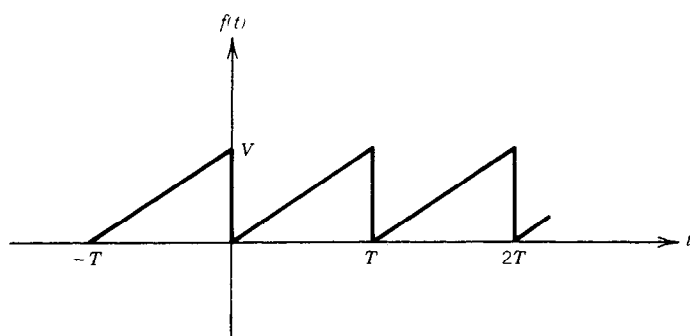


Fig. 2-11. Forma de onda en diente de sierra.

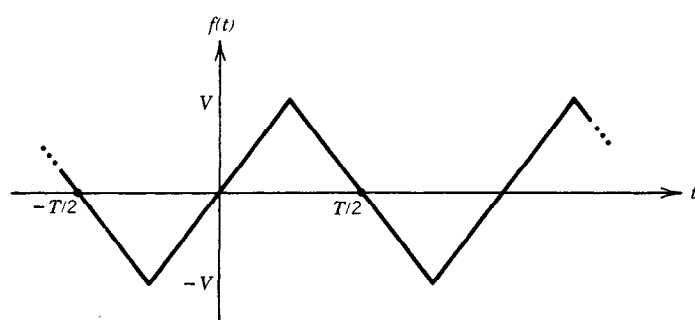


Fig. 2-12. Forma de onda triangular.

LIBROS Y REVISTAS

- F. Ayres, *Matrices*, Schaum Outline Series, McGraw-Hill, Nueva York.
- J. Cadzow, *Discrete Time Systems*, Prentice-Hall, Englewood Cliffs, NJ.
- G. R. Cooper y C. D. McGillem, *Methods of Signal and System Analysis*, Holt, Rinehart & Winston, Nueva York.
- M. Fraser, *College Algebra and Trigonometry*, Benjamin/Cummings, Menlo Park, CA.
- R. A. Gabel y R. A. Roberts, *Signals and Linear Systems*, 2.^a ed., Wiley, Nueva York.
- A. W. Goodman, *Analytic Geometry with Calculus*, MacMillan, Nueva York.
- F. B. Hildebrand, *Methods of Applied Mathematics*, Prentice-Hall, Englewood Cliffs, NJ.
- E. Jahnke y F. Emde, *Table of Functions*, Dover, Nueva York.
- M. Javid y E. Brenner, *Analysis, Transmission and Filtering of Signals*, McGraw-Hill, Nueva York.
- L. B. W. Jolley, *Summation of Series*, Dover, Nueva York.
- R. A. Larson y R. P. Hostetler, *Calculus*, D. C. Heath, Lexington, MA.
- B. P. Lathi, *Signals, Systems and Communications*, Wiley, Nueva York.
- S. Lipschutz, *Finite Mathematics*, Schaum Outline Series, McGraw-Hill, Nueva York.
- J. Marion, *Classical Electromagnetic Radiation*, Academic, Nueva York.
- J. Robson, *Basic Tables in Physics*, McGraw-Hill, Nueva York.
- S. Ross, *Introduction to Ordinary Differential Equations*, Wiley, Nueva York.
- L. Smail, *Calculus*, Appleton-Century-Crofts, Nueva York.
- M. Spiegel, *Vector Analysis*, Schaum Outline Series, McGraw-Hill, Nueva York.
- M. E. Van Valkenberg, *Network Analysis*, 3.^a ed., Prentice-Hall, Englewood Cliffs, NJ.

CAPÍTULO 3

Computación

Thomas C. Upson

Rochester Institute of Technology
Rochester, Nueva York

3.1 Herramientas de cómputo disponibles

3.2 Fuentes de error en computación

3.3 Técnicas numéricas en computación

3.3.1 Método de Newton-Raphson
para encontrar raíces

3.3.2 Ajuste de curvas

3.4. Integración numérica

**3.5 Solución numérica de ecuaciones
diferenciales ordinarias
de primer orden**

3.1 HERRAMIENTAS DE CÓMPUTO DISPONIBLES

En la actualidad, pocas son las operaciones de cómputo que se realizan manualmente. Las máquinas han hecho posible resolver problemas que hasta hace muy poco tiempo permanecían sin solución, debido al tiempo requerido para resolverlos.

Las herramientas de cómputo disponibles van desde máquinas sumadoras de bolsillo hasta grandes sistemas de tiempo compartido. No es oportuno comentar en detalle las herramientas a disposición del ingeniero, ya que se desarrollan tan rápido que cualquier comparación es obsoleta antes de su publicación. Basta decir que el nivel actual de complejidad, incluso del equipo más elemental, permite excluir de los manuales de ingeniería y científicos las tablas matemáticas, que tradicionalmente se habían incluido.

Además de las máquinas calculadoras (hardware), se dispone de diversos "paquetes" de software. De nuevo, el espectro de complejidad y versatilidad es amplio. En la tabla 3-1 se describen algunos de los paquetes de software científico que existen en el mercado, su lenguaje y el número de subrutinas que contienen. Pueden consultarse más detalles en el *Engineering Systems Software Referral Catalog*¹. Por ejemplo, el paquete de la IMSL (Internacional Mathematical and Statistical Library) contiene más de 400 subrutinas programadas en Fortran capaces de invertir matrices, evaluar funciones de Bessel, resolver programas lineales y efectuar otras tareas diversas de análisis de datos. El paquete IMSL es capaz de apoyar muchos ambientes de computación-compilación*.

3.2 FUENTES DE ERROR EN COMPUTACIÓN²

El usuario de cualquier esquema numérico debe estar enterado de las fuentes de error. El error en alguna aproximación puede estar producido por varias componentes:

1. **Exactitud de los datos utilizados.** En un laboratorio o cualquier otro ambiente de prueba en el que se coleccionen datos, las mediciones representan sólo una aproximación al valor real de la variable. Por tanto, el dispositivo de medición introduce por sí mismo un error. El hecho de que el usuario conozca el nivel al que se introduce el error adquiere gran importancia. Por ejemplo, no es posible esperar cinco cifras significativas de exactitud en un esquema de aproximación numérica cuando los datos coleccionados tienen solamente tres cifras significativas.
2. **Instrumento de cálculo.** Cada máquina tiene capacidad para un nivel máximo de exactitud. El

usuario debe conocer dicho nivel y la forma en que afecta los resultados de cálculos extensos.

Por ejemplo, en una calculadora con ocho posiciones de exactitud, los cálculos en los que se utiliza e (la base de los logaritmos naturales) introducen un error, denominado error de redondeo.

Esto se debe a que, como e es un número irracional, su expresión decimal es infinita, y en la calculadora de este ejemplo sólo es posible representar sus ocho primeros dígitos. En cálculos extensos los errores de redondeo se acumulan y afectan el resultado.

3. **Algoritmo.** Cada método de cómputo proporciona un valor numérico aproximado para el valor real deseado. Por tanto, existe un error inherente a la aproximación. Un incremento significativo en la exactitud sólo puede obtenerse a expensas de incrementar la complejidad del algoritmo, el tiempo de computación o ambos. Por ejemplo, las diferentes técnicas numéricas para aproximar $y(x')$, dadas $y' = f(x, y)$, $y(x_0) = y_0$ (que se describen posteriormente en este capítulo), empiezan en el punto (x_0, y_0) y se realizan en n pasos, cada uno de longitud h , para calcular un valor aproximado de la solución en $x' = x_0 + nh$ [$h = (x' - x_0)/n$]. Incluso si los datos iniciales son exactos y no se introduce error de redondeo durante los cálculos, se introduce un error en cada paso, llamado error local. Los errores locales se acumulan de un paso a otro, creando un "error global" en x' .

La diferencia entre los errores global y local puede ser tan grande como un orden de magnitud. Por ejemplo, el error local en el método de Euler (que se presenta más adelante) es directamente proporcional al cuadrado del tamaño del paso, h^2 , mientras que el error global es directamente proporcional a h .

Otra forma de error se introduce en un algoritmo cuando sólo se utiliza un número finito de términos de una serie infinita. Este error se denomina error de truncado. Por ejemplo, en la fórmula para evaluar:

$$\int_a^b f(x) dx$$

por la regla de Simpson se utiliza una serie de Taylor truncada para $f(x)$. Tanto la regla de Simpson como la serie de Taylor se analizan posteriormente en este capítulo.

Sobra decir que no se conoce el error exacto que se introduce con estas formas. Lo más que puede determinarse es el error máximo que puede existir.

Por ejemplo, se sabe que cuando se emplean los primeros tres términos distintos de cero de la serie:

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

* Una lista completa de los ambientes computadora-compilador apoyados por la IMSL puede verse en la IMSL Library, disponible en IMSL Inc., 7500 Bellaire Blvd., 6.º piso, NBC Building, Houston TX 72036. E.U.A.

Tabla 3-1. Paquetes representativos disponibles de software*

Nombre del software	Capacidad del software	Lenguajes	Fuente de mayor información
Biblioteca NAG (Numerical Algorithms Group)	Análisis numérico y estadístico (466 subrutinas)	FORTRAN	Numerical Algorithms Group Inc. Atn: Company Secretary 1250 Grace Ct. Downers Grove, IL 60516 312 971-2337
IMSL (Biblioteca Internacional de Matemáticas y Estadísticas)	Análisis numérico y estadístico (550 subrutinas)	FORTRAN IV, FORTRAN IV-Plus	IMSL, Inc. 7500 Bellaire Blvd. 6th Floor, NBC Building Houston, Texas 72036 713 772-1927
SPSS	Análisis estadístico y administración de datos	FORTRAN	SPSS Inc. Suite 3000 444 North Michigan Ave. Chicago, IL 60611 312 329-2400
Microsolve	Sistema de programación lineal para su uso en microcomputadoras	FORTRAN IV	Advanced Management Techniques Box 1597 Houston, TX 77057 713 373-1905

* Véanse los detalles en el *Engineering Systems Software Referral Catalog*¹.

para representar la función seno, el error no es mayor que el primer término omitido, en este caso:

$$E_5(x) \leq \left| \frac{x^7}{7!} \right|$$

ya que los términos de la serie tienen signos alternados.

Tal vez sea posible para el usuario reducir de alguna manera el error de redondeo tomando en cuenta el orden en que se realizan las operaciones. Por ejemplo, para calcular $(\sin x + \cos x)^2$, sería prudente proceder de la siguiente forma:

$$(\sin x + \cos x)^2 = \sin^2 x + 2 \sin x \cos x + \cos^2 x = 1 + \sin 2x$$

Aprovechar las identidades trigonométricas reduce de tres a uno el número de operaciones que implican redondeo.

3.3 TÉCNICAS NUMÉRICAS EN COMPUTACIÓN

3.3.1 Método de Newton-Raphson para encontrar raíces

El algoritmo se presentará para funciones diferenciables de una variable, $y = f(x)$.

Para resolver una ecuación de la forma $f(x) = 0$ se utiliza el siguiente esquema de iteración (fig. 3-1).

Se elige un valor inicial x_1 , que se supone en la vecindad de la solución real*. A continuación, se calcula el valor x_2 como $x_1 - f(x_1)/f'(x_1)$, que representa la intersección de la tangente a la gráfica de la función en x_1 con el eje x .

De forma parecida, después de que se ha calculado la n -ésima iteración x_n , es posible calcular la $(n + 1)$ -ésima iteración por medio de:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

La convergencia del método de Newton-Raphson a la raíz es rápida, siempre que el cociente de la segunda derivada entre la primera derivada de la función no sea grande.

De hecho, se sabe que el error en la n -ésima iteración está dado por (ref. 2, págs. 43-50):

$$\epsilon_{n+1} = \frac{1}{2} (\epsilon_n)^2 \frac{f''(a)}{f'(a)}$$

donde a es la raíz real de la ecuación; es decir, $f(a) = 0$.

Dado que a es desconocida, sólo es necesario determinar una cota superior para:

$$\left| \frac{f''(a)}{f'(a)} \right|$$

* x_1 puede determinarse evaluando la función hasta que se detecte un cambio de signo. (En la referencia 2 se analizan los casos de raíces múltiples.)

Ejemplo. Determinar una raíz positiva de $x^2 - 7x + 10$ cerca de $x = 1$.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Haciendo:

$$f(x) = x^2 - 7x + 10$$

entonces:

$$f'(x) = 2x - 7$$

de modo que:

$$x_{n+1} = x_n - \frac{x_n^2 - 7x_n + 10}{2x_n - 7}$$

$x_1 = 1$:

$$x_2 = x_1 - \frac{x_1^2 - 7x_1 + 10}{2x_1 - 7}$$

$$x_2 = 1 - \frac{4}{-5} = 1.8$$

$x_2 = 1.8$:

$$x_3 = 1.8 - \frac{0.64}{-3.4} = 1.9882$$

$x_3 = 1.9882$:

$$x_4 = 1.9882 - \frac{0.0355}{-3.0236} = 1.9999$$

$x_4 = 1.9999$.

Para fines de comparación, el valor real es $x = 2$, que se encuentra por factorización:

$$x^2 - 7x + 10 = (x - 2)(x - 5)$$

3.3.2 Ajuste de curvas

Existen muchas técnicas para obtener información de una relación entre variables cuando los datos disponibles son limitados. A continuación se analizan tres métodos:

1. **Método de mínimos cuadrados**³. En el método de mínimos cuadrados que se describe aquí se ajusta una recta $y = a + bx$ a los n puntos de coordenadas:

$$(x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)$$

de tal forma que se minimice la suma de los cuadrados del error, S :

$$S = \sum_{i=1}^n [y_i - (a + bx_i)]^2$$

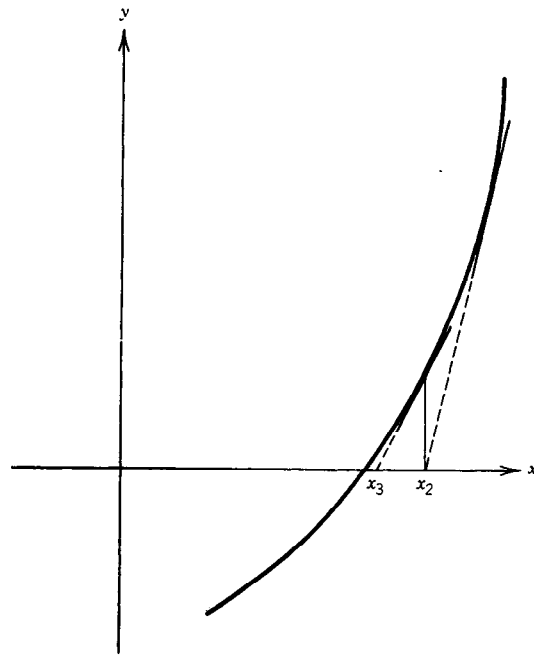


Fig. 3-1. Esquema de iteración en la solución de $f(x) = 0$ utilizando la técnica de Newton-Raphson para encontrar raíces.

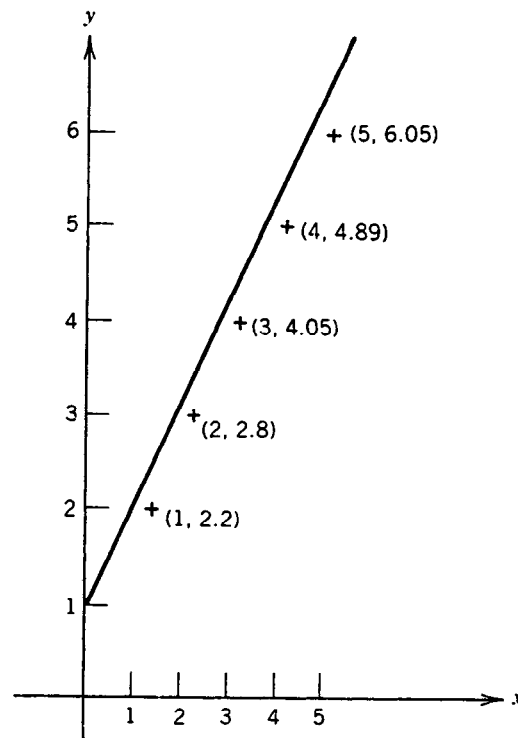


Fig. 3-2. Aproximación lineal por mínimos cuadrados.

Los coeficientes a y b están dados por las expresiones:

$$a = \frac{\left(\sum_{i=1}^n y_i\right)\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n x_i y_i\right)}{n\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2}$$

$$b = \frac{n\left(\sum_{i=1}^n x_i y_i\right) - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2}$$

En la referencia 3 se analizan otros métodos de mínimos cuadrados en los que se utiliza un polinomio de orden superior y datos relacionados de manera exponencial.

Ejemplo. Determinar la aproximación lineal por mínimos cuadrados a los puntos de la figura 3-2.

En la tabla 3-2 se presentan de manera organizada los datos. A partir de ella es posible obtener:

$$a = \frac{\left(\sum_{i=1}^5 y_i\right)\left(\sum_{i=1}^5 x_i^2\right) - \left(\sum_{i=1}^5 x_i\right)\left(\sum_{i=1}^5 x_i y_i\right)}{5\left(\sum_{i=1}^5 x_i^2\right) - \left(\sum_{i=1}^5 x_i\right)^2}$$

$$a = \frac{(19.99)(55) - (15)(21.21)}{(5)(55) - (15)^2}$$

$$a = \frac{1099.45 - 1046.40}{275 - 225} = \frac{53.05}{50} = 1.061$$

$$b = \frac{5\left(\sum_{i=1}^5 x_i y_i\right) - \left(\sum_{i=1}^5 x_i\right)\left(\sum_{i=1}^5 y_i\right)}{50}$$

$$b = \frac{5(71.71) - (15)(19.99)}{50}$$

$$b = \frac{358.55 - 299.85}{50} = 1.174$$

$$y = a + bx.$$

Por tanto, $y = 1.061 + 1.174x$ es la recta que minimiza el error por mínimos cuadrados con los puntos dados.

Tabla 3-2. Datos obtenidos para la aproximación lineal por mínimos cuadrados a los puntos que se muestran en la figura 3-2

i	x_i	y_i	x_i^2	$x_i y_i$
1	1	2.20	1	2.20
2	2	2.80	4	5.60
3	3	4.05	9	12.15
4	4	4.89	16	19.56
5	5	6.05	25	30.25
$\sum_{i=1}^5$	15	19.99	55	69.76

linomio de grado n por los $(n+1)$ puntos dados $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$. La relación entre x e y es:

$$y = \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)} y_0 + \\ + \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} y_1 + \\ + \dots + \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\dots(x_n-x_{n-1})} y_n$$

Que se utiliza a fin de determinar el valor de y para valores intermedios de x^* .

Ejemplo. Utilizar los datos del ejemplo de la aproximación lineal por mínimos cuadrados para ajustar un polinomio de cuarto grado a través de los puntos (1, 2.2), (2, 2.8), (3, 4.05), (4, 4.89) y (5, 6.05):

$$y = \frac{(x-2)(x-3)(x-4)(x-5)}{(1-2)(1-3)(1-4)(1-5)} (2.2) + \\ + \frac{(x-1)(x-3)(x-4)(x-5)}{(2-1)(2-3)(2-4)(2-5)} (2.8) + \\ + \frac{(x-1)(x-2)(x-4)(x-5)}{(3-1)(3-2)(3-4)(3-5)} (4.05) + \\ + \frac{(x-1)(x-2)(x-3)(x-5)}{(4-1)(4-2)(4-3)(4-5)} (4.89) + \\ + \frac{(x-1)(x-2)(x-3)(x-4)}{(5-1)(5-2)(5-3)(5-4)} (6.05)$$

Después de los cálculos necesarios:

2. Interpolación de Lagrange⁴. En el esquema de la interpolación de Lagrange se hace pasar un po-

* En la referencia 4 se presentan otras interpolaciones, polinomios y errores asociados.

$$y = 0.0329x^4 - 0.5058x^3 + 2.5371x^2 - 3.9641x + 4.1$$

3. Serie de Taylor⁵. El polinomio de Taylor aproxima una función conocida con un polinomio de grado n y es de utilidad cuando se busca una buena aproximación local. Esto contrasta con las dos técnicas anteriores, en las que la aproximación proporciona un buen ajuste a lo largo de algún intervalo. El polinomio de Taylor de grado n coincide con la función y sus primeras n derivadas en $x = a$. Si $y = f(x)$ y se conocen sus derivadas en $x = a$, entonces la serie de Taylor para $f(x)$ es:

$$f(x) = f(a) + \sum_{i=1}^{\infty} \frac{f^{(i)}(a)}{i!} (x - a)^i$$

Si la serie se trunca después de n términos o si solamente se conocen las primeras n derivadas, con el residuo $R_n(x)$ se obtiene una aproximación polinomial:

$$R_n(x) = \frac{f^{(n+1)}(t)}{(n+1)!} (t - a)^{n+1}$$

para alguna t :

$$a \leq t \leq x$$

En la tabla 3-3 se presenta la serie de Taylor para funciones que suelen utilizarse de forma habitual.

3.4 INTEGRACIÓN NUMÉRICA

Es posible obtener un valor aproximado de $\int_a^b f(t)dt$ cuando $f(t)$ no tenga antiderivada en forma cerrada o cuando $f(t)$ esté definida por un conjunto de puntos dados.

Aquí se describen tres técnicas. Las dos primeras se utilizan cuando $f(t)$ no tiene antiderivada en forma cerrada, y la tercera cuando sólo se dispone de f y sus primeras n derivadas en a .

1. Regla del trapecio⁶. Con esta regla se utiliza una serie de aproximaciones lineales a la función $y = f(t)$ a fin de determinar un valor aproximado de la integral definida de f para $a \leq x \leq b$, como se muestra a continuación (fig. 3-3).

Se hace una equipartición del intervalo (a, b) mediante los puntos:

$$a = t_0 < t_1 < t_2 < \dots < t_n = b$$

con

$$t_i = a + ih$$

y

$$h = \frac{b - a}{n}$$

Haciendo $y_i = f(t_i)$ se obtiene la siguiente aproximación:

$$\int_a^b f(t)dt = \frac{h}{2} (y_0 + 2y_1 + 2y_2 + \dots + 2y_{n-1} + y_n)$$

Tabla 3-3. Polinomios de Taylor de uso común

$f(x)$	Polinomio de Taylor	Intervalo de convergencia
$\frac{1}{1-x}$	$\sum_{k=0}^n x^k$	$ x < 1$
$\sin x$	$\sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!}$	$ x < \infty$
$\cos x$	$\sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!}$	$ x < \infty$
e^x	$\sum_{k=0}^n \frac{x^k}{k!}$	$ x < \infty$
$\ln(1+x)$	$\sum_{k=0}^n (-1)^k \frac{x^{k+1}}{k+1}$	$ x < 1$
$\sqrt{1+x}$	$1 + \frac{x}{2} - \frac{x^2}{22} + \frac{3}{23}x^3 - \frac{3.5}{24}x^4 + \dots$	$ x < 1$

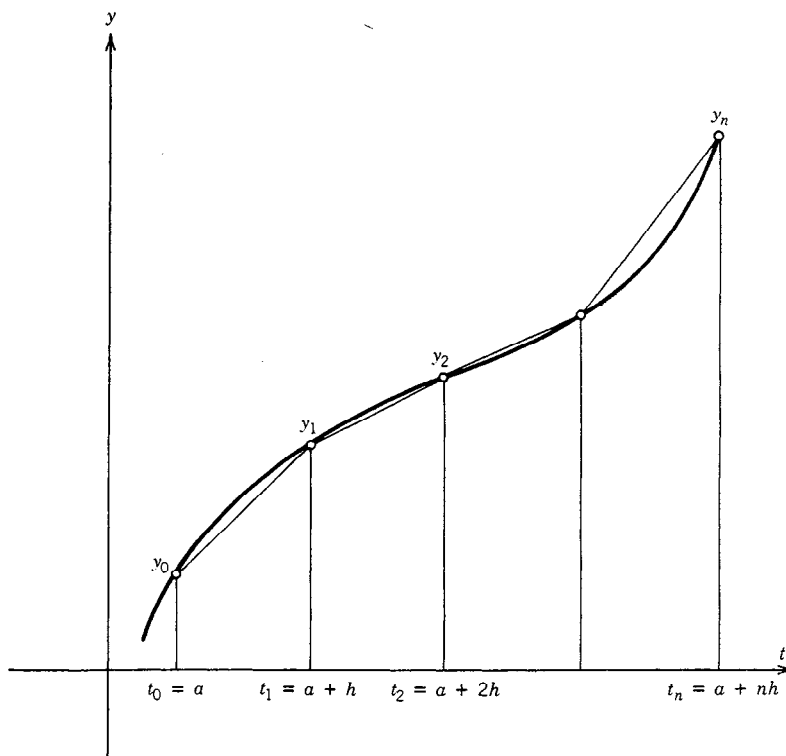


Fig. 3-3. Aproximación trapezoidal a $y = f(t)$ para $a \leq t \leq b$.

Ejemplo. Determinar un valor aproximado para:

$$\int_0^1 \sqrt{1+t^3} dt$$

Primero se elige el número de subintervalos n , por ejemplo, $n = 4$. Después se calcula:

$$h = \frac{b-a}{n} = \frac{1-0}{4} = \frac{1}{4}$$

de forma que $t_0 = 0$, $t_1 = 1/4$, $t_2 = 1/2$, $t_3 = 3/4$, y $t_4 = 1$. Luego se calculan los valores correspondientes de la función para llegar a la última columna de la tabla 3-4. Usando la regla del trapecio:

$$\int_0^1 \sqrt{1+t^3} dt = \frac{1}{2} [1 + 2(1.008) + 2(1.061) + 2(1.192) + 1.414]$$

o bien:

$$\int_0^1 \sqrt{1+t^3} dt = 1.117, \text{ aproximadamente}$$

2. Regla de Simpson⁶. Mientras que en la regla del trapecio se utilizan aproximaciones lineales a las funciones que habrán de integrarse, la regla de

Simpson aproxima la función con funciones cuadráticas diseñadas para coincidir con la función original en tres puntos consecutivos de la partición. Así, se espera una exactitud aún mayor con el mismo número de puntos, o bien la misma exactitud con menos puntos.

El intervalo (a, b) se divide en subintervalos iguales definidos por los puntos:

$$a = t_0 < t_1 < t_2 < \dots < t_n = b$$

$$\text{con } t_i = a + ih$$

$$h = \frac{b-a}{n} \text{ y } n \text{ debe ser par}$$

Haciendo $y_i = f(t_i)$ se obtiene la regla de aproximación integral de Simpson:

$$\int_a^b f(t) dt = \frac{h}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{n-2} + 4y_{n-1} + y_n)$$

Ejemplo. Empleando los datos de la tabla 3-4, con la regla de Simpson se obtiene:

$$\int_0^1 \sqrt{1+t^3} dt = \frac{h}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + y_4)$$

o bien:

Tabla 3-4. Datos calculados para evaluar $\int_0^1 \sqrt{1+t^3} dt$ utilizando la aproximación trapezoidal

i	t_i	$y_i = f(t_i) = \sqrt{1+t_i^3}$
0	0	1
1	$\frac{1}{4}$	1.008
2	$\frac{1}{2}$	1.061
3	$\frac{3}{4}$	1.192
4	1	1.414

$$\int_0^1 \sqrt{1+t^3} dt = \frac{1}{3} [1.000 + (4)(1.008) + (2)(1.061) + 4(1.192) + 1.414]$$

$$\int_0^1 \sqrt{1+t^3} dt = 1.111$$

En contraste, para obtener esta misma exactitud con la regla del trapecio se necesitarían 100 subintervalos.

3. Serie de Taylor. Cuando se conocen el valor de una función y sus primeras n derivadas en un punto a , entonces a partir de:

$$f(t) = f(a) + \sum_{i=1}^n \frac{f^{(i)}(a)}{i!} (t-a)^i$$

es posible aproximar la integral por medio de:

$$\int_a^b f(t) dt = f(a)(b-a) + \sum_{i=1}^n \frac{f^{(i)}(a)(b-a)^{i+1}}{(i+1)!}$$

Ejemplo. Empleando la serie básica de Taylor para:

$$\sqrt{1+t^3} = 1 + \frac{1}{2} t^3 + \frac{t^6}{2^2 \cdot 2!} + \frac{3t^9}{3^3 \cdot 3!} \dots$$

y usando sólo los primeros tres términos diferentes de cero se obtiene:

$$\int_0^1 \sqrt{1+t^3} dt = \int_0^1 \left(1 + \frac{1}{2} t^3 - \frac{t^6}{2^2 \cdot 2!} + \frac{t^9}{2^4} \right) dt = 1.113392$$

3.5 SOLUCIÓN NUMÉRICA DE ECUACIONES DIFERENCIALES ORDINARIAS DE PRIMER ORDEN

Aquí se describen tres métodos para determinar una solución de la ecuación diferencial ordinaria de primer orden $y' = f(x, y)$ que pasa por el punto (x_0, y_0) . Esta ecuación se conoce como problema de valor inicial. Se utiliza el mismo ejemplo para ilustrar los tres métodos.

Es necesario observar que las técnicas más complejas reducen el error a costa de incrementar el tiempo de cómputo.

1. Método de Euler. Con este método se aproxima la solución real al problema de valor inicial mediante rectas paralelas a la tangente a la curva solución (véase en la fig. 3-4 un ejemplo arbitrario). El procedimiento consiste en comenzar en el punto inicial (x_0, y_0) sobre la curva y llegar al punto (x_1, y_1) ($x_1 = x_0 + h$, donde h es el tamaño del paso o incremento) a lo largo de la tangente a la curva en (x_0, y_0) . En el siguiente paso el proceso se inicia en (x_1, y_1) y prosigue a lo largo de una recta paralela a la tangente en el punto (x_2, y_2) ($x_2 = x_1 + h$). El proceso continúa de esta forma hasta que se alcanza el valor deseado de x y se obtiene la solución aproximada. La relación con la que se calcula la aproximación es:

$$y_{i+1} = y_i + hf(x_i, y_i)$$

donde h es el tamaño del paso o incremento, y_{i+1} es el nuevo valor de y , y_i es la aproximación después de i pasos, y $x_i = x_{i-1} + h = x_0 + ih$. En particular:

$$y_1 = y_0 + hf(x_0, y_0)$$

Ejemplo. Con objeto de resolver el problema de valor inicial:

$$y' = x - y, \quad y(0) = 1$$

para $y(1)$, se elige como tamaño del paso $h = 0.25$. Por tanto, $x_0 = 0.00$, $x_1 = 0.25$, $x_2 = 0.50$, $x_3 = 0.75$ y $x_4 = 1.00$. Es conveniente coleccionar los datos como se muestra en la tabla 3-5.

Con fines de comparación, la forma cerrada de la solución es:

$$y = (x - 1) + 2e^{-x}$$

Tabla 3-5. Datos para la aproximación de Euler a la solución de $y' = x - y$, $y(0) = 1$

i	x_i	y_i	$y'_i = x_i - y_i$	y_{i+1} (nueva y_i)
0	0.00	1.00	-1.00	0.75
1	0.25	0.75	-0.50	0.625
2	0.50	0.625	-0.425	0.59375
3	0.75	0.59375	0.15625	0.7890625
4	1.00	0.7890625	—	—

Tabla 3-6. Comparación de la solución real y la aproximación de Euler de la solución para $y' = x - y$, $y(0) = 1$

x	y	Error
0.00	1.0000	0.0000
0.25	0.8076	0.0576
0.50	0.7131	0.1197
0.75	0.6947	0.1010
1.00	0.7358	0.0533

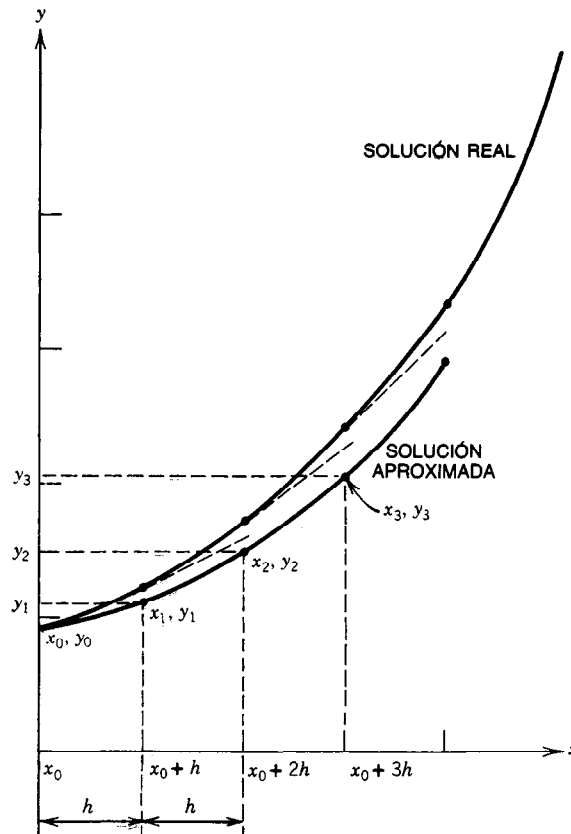


Fig. 3-4. Aproximación de Euler.

Tabla 3-7. Aproximación de Euler mejorada para $y' = x - y$ e $y(0) = 1$

i	x_i	y_i	$f(x_i, y_i)$	$x_i + h = k_i$	$y_i + hf(x_i, y_i) = l_i$	$f(k_i, l_i)$	y_{i+1}
0	0.00	1.00	-1.00	0.25	0.75	-0.50	0.8125
1	0.25	0.8125	-0.5625	0.50	0.6719	-0.1719	0.7207
2	0.50	0.7207	-0.2207	0.75	0.6655	0.0845	0.7036
3	0.75	0.7036	0.0463	1.00	0.7152	0.2848	0.7451
4	1.00	0.7451					

Los valores de esta función, así como el error de la aproximación anterior, se presentan en la tabla 3-6.

2. Método de Euler mejorado. En este método se utiliza un promedio de las pendientes en los dos puntos extremos, en lugar de sólo la pendiente en un punto extremo, para corregir el alto peso asignado al comportamiento en el punto único. La ecuación diferencial de primer orden tiene la siguiente forma:

$$y' = f(x, y) \text{ sujeta a } y(x_0) = y_0$$

La relación entre la aproximación para x_i y para x_{i+1} es:

Tabla 3-8. Solución de $y' = x - y$, $y(0) = 1$, por el método de Runge-Kutta

i	x_i	y_i
0	0.00	1.0000
1	0.25	0.8076
2	0.50	0.7131
3	0.75	0.6948
4	1.00	0.7358

$$y_{i+1} = y_i + h \left\{ \frac{f(x_i, y_i) + f[x_i + h, y_i + hf(x_i, y_i)]}{2} \right\}$$

Este método también se conoce como método del promedio de las pendientes.

Tabla 3-9. Soluciones de $y' = x - y$, $y(0) = 1$, para varios tamaños de incremento utilizando los métodos de cómputo de Euler, de Euler modificado y de Runge-Kutta

Tamaño del incremento	Euler	Euler mejorado	Runge-Kutta	Real
0.25	0.6328	0.7451	0.7358	0.7358
0.2	0.6554	0.7414	0.7358	0.7358
0.1	0.6974	0.7371	0.7358	0.7358
0.05	0.7170	0.7360	0.7358	0.7358
0.025	0.7265	0.7358	0.7358	0.7358
0.01	0.7321	0.7358	0.7358	0.7358

Ejemplo. Resolver $y' = x - y$, sujeta a $y(0) = 1$, utilizando el método de Euler mejorado con $h = 0.25$ a fin de encontrar un valor aproximado de $y(1)$. Para empezar:

$$y_1 = y_0 + (0.25) \left\{ \frac{f(x_0, y_0) + f[0.25, y_0 + 0.25f(x_0, y_0)]}{2} \right\}$$

$$x_0 = 0, \quad y_0 = 1$$

de modo que:

$$y_1 = 1 + (0.25) \left\{ \frac{f(0,1) + f[0.25, 1 + 0.25f(0,1)]}{2} \right\}$$

con:

$$f(x, y) = x - y$$

$$f(0,1) = -1$$

y

$$y_1 = 1 + 0.25 \left[\frac{-1 + f(0.25, 0.75)}{2} \right]$$

$$y_1 = 1 + 0.25 \left[\frac{-1 - 0.50}{2} \right]$$

$$y_1 = 1 + (0.25)(-0.75)$$

$$y_1 = 0.8125$$

Se continúa el algoritmo y los datos se registran en la tabla 3-7, para la cual se utiliza la notación:

$$k_i = x_i + h \quad y \quad l_i = y_i + hf(x_i, y_i)$$

3. Método de Runge-Kutta de cuarto orden. En el método (de cuarto orden) de Runge-Kutta se hace uso de un promedio ponderado de la pendiente en diferentes lugares ubicados a lo largo del intervalo, y se basa en una aproximación en serie de Taylor de cinco términos.

El algoritmo para la solución de $x = x'$ de la ecuación diferencial $y' = f(x, y)$ sujeta a $y(x_0) = y_0$ es:

$$y_{i+1} = y_i + \frac{h}{6}(k_{i1} + 2k_{i2} + 2k_{i3} + k_{i4})$$

$$0 \leq i \leq n$$

donde:

$$k_{i1} = f(x_i, y_i)$$

$$k_{i2} = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_{i1}}{2}\right)$$

$$k_{i3} = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_{i2}}{2}\right)$$

$$k_{i4} = f(x_i + h, y_i + hk_{i3})$$

y

$$h = \frac{x' - x_0}{n}$$

Ejemplo. La solución de $y' = x - y$, sujeta a $y(0) = 1$, utilizando el método de Runge-Kutta para determinar $y(1)$ con $h = 0.25$ arroja los datos que se muestran en la tabla 3-8. En la tabla 3-9 se presentan las soluciones de $y' = x - y$, $y(0) = 1$ en $x = 1$ para varios tamaños de incremento y técnicas de cómputo.

En la tabla 3-10 se comparan los órdenes de magnitud de los errores global y local para las diferentes técnicas de cómputo presentadas.

En caso de que f no dependa de y , la aproximación de Runge-Kutta concuerda con la regla de Simpson para evaluar la integral:

$$y_{n+1} - y_n = \int_{x_n}^{x_{n+h}} f(x) dx$$

Tabla 3-10. Orden de magnitud del error para los métodos de Euler, de Euler mejorado y de Runge-Kutta

Tipo de error	Euler	Euler mejorado	Runge-Kutta
Local	h^2	h^3	h^5
Global	h	h^2	h^4

REFERENCIAS BIBLIOGRÁFICAS

1. *Engineering Systems Software Referral Catalog*, 7.^a ed., Digital Equipment Corp., Marlboro, MA.
2. A. J. Bajpai, I. M. Calus y J. A. Fairley, *Numerical Methods for Scientists and Engineers*, Wiley, Nueva York.
3. R. L. Burden, J. D. Faires y A. C. Reynolds, *Numerical Analysis*, Prindle, Weber & Schmidt, Boston.
4. P. A. Stark, *Introduction to Numerical Methods*, Macmillan, Nueva York.
5. M. Kline, *Calculus/An Intuitive and Physical Approach*, 2.^a ed., Wiley, Nueva York.
6. F. B. Hildebrand, *An Introduction to Numerical Analysis*, 2.^a ed., McGraw-Hill, Nueva York.
7. W. E. Boyce y R. C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, 3.^a ed., Wiley, Nueva York.

CAPÍTULO 4

Transformadas

Manu Malek-Zavarei

Bell Communications Research
Holmdel, New Jersey

4.1 Introducción

4.2 Transformada de Fourier

4.3 Transformada discreta de Fourier

4.4 Transformada rápida de Fourier

4.5 Transformada de Laplace

4.5.1 Determinación de la transformada de Laplace

4.5.2 Inversión de la transformada de Laplace

4.5.3 Tabla de transformadas de Laplace

4.5.4 Solución de ecuaciones diferenciales utilizando la transformada de Laplace

4.6 Transformada z

4.6.1 Determinación de la transformada z

4.6.2 Inversión de la transformada z

4.6.3 Tabla de transformadas z

4.6.4 Solución de ecuaciones en diferencias utilizando la transformada z

4.7 Funciones de transferencia

4.1 INTRODUCCIÓN

Las transformadas lineales, especialmente la de Fourier, la de Laplace y la z , son técnicas para resolver problemas relacionados con sistemas lineales. Se utilizan como herramientas matemáticas que hacen resolubles problemas difíciles. En este capítulo se presenta una introducción a dichas transformadas.

Se empezará con el concepto básico de la transformada de Fourier. Se presentarán la serie de Fourier, la transformada discreta de Fourier y el método de la transformada rápida de Fourier. Después se analizarán definiciones, propiedades y aplicaciones fundamentales de la transformada de Laplace, y un artículo parecido se dedica a la transformada z . Por último, se analizarán los conceptos de función de transferencia en tiempo continuo y de sistemas lineales invariantes en tiempo discreto.

4.2 TRANSFORMADA DE FOURIER

En su estudio sobre cuerdas de instrumentos en vibración, Daniel Bernoulli (1700-1782) fue el primero en utilizar la idea de que cualquier función periódica puede representarse por medio de una serie de componentes senoidales relacionadas armónicamente. Sin embargo, esta técnica de análisis no fue ampliamente aceptada, sino hasta después de que Jean B. J. Fourier publicara en 1822 su estudio sistemático de dichas representaciones¹⁻⁴.

Para comprender esta idea fundamental, primero se considerará una función arbitraria (no necesariamente periódica), $g(t)$. La "transformada de Fourier" $G(f)$, $G_1(\omega)$ o $g(t)$ se define como:

$$G(f) = \int_{-\infty}^{+\infty} g(t) e^{-2\pi j f t} dt \quad (4-1a)$$

$$G_1(\omega) = \int_{-\infty}^{+\infty} g(t) e^{-j\omega t} dt \quad (4-1b)$$

donde $j = \sqrt{-1}$ y $\omega = 2\pi f$. La función $G(f)$ o $G_1(\omega)$ también se conoce como espectro de frecuencia de $g(t)$.

La transformada inversa de Fourier de $G(f)$, o $G_1(\omega)$, está dada por:

$$g(t) = \int_{-\infty}^{+\infty} G(f) e^{2\pi j f t} df = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G_1(\omega) e^{j\omega t} d\omega \quad (4-2)$$

Las ecuaciones 4-1a y 4-1b también se conocen como integral de Fourier. Se dice que una función tiene transformada de Fourier si la integral de Fourier correspondiente converge.

A partir de la definición de integral de Fourier, ecuaciones 4-1a y 4-1b, es posible deducir la correspondencia entre una función y su transformada de Fourier (tabla 4-1).

Los métodos de la transformada de Fourier se utilizan en el procesamiento de señales. En la tabla 4-2 se

Tabla 4-1. Correspondencia entre una función y su transformada de Fourier

Función	Transformada de Fourier
Real y par	Real y par
Real e impar	Imaginaria e impar
Imaginaria y par	Imaginaria y par
Compleja y par	Compleja y par
Compleja e impar	Compleja e impar
Par	Par
Impar	Impar

presentan algunos pares transformados de Fourier comunes, y $g(t)$ está definida en el intervalo $(-\infty, +\infty)$.

4.3 TRANSFORMADA DISCRETA DE FOURIER

Ahora sea $g(t)$ una función periódica continua con período T ; es decir, sea $g(t+T) = g(t)$ para toda t , donde T es el menor número positivo con esta propiedad. La transformada de Fourier $G(f)$ de $g(t)$ es entonces una función discreta dada por:

$$G(f_n) = \int_T g(t) e^{-2\pi j f_n t} dt \quad (4-3)$$

donde:

$$f_n = \frac{n}{T} = n \Delta f \quad (4-4)$$

y la integral se evalúa en un período de $g(t)$. La transformada inversa de Fourier de $G(f_n)$ en la ecuación 4-3 es:

$$g(t) = \sum_{n=-\infty}^{+\infty} G(f_n) e^{2\pi j f_n t} \quad (4-5)$$

y se le conoce a menudo como *serie de Fourier* de $g(t)$. Obsérvese que la ecuación 4-5 corresponde a la versión discreta de la ecuación 4-2 cuando los incrementos Δf son muy pequeños. De hecho, empleando la ecuación 4-4 se tiene:

$$\int_{-\infty}^{+\infty} G(f) e^{2\pi j f t} df = \lim_{\Delta f \rightarrow 0} \sum_{n=-\infty}^{+\infty} G(f_n) e^{2\pi j f_n t} \Delta f = \lim_{\Delta f \rightarrow 0} \sum_{n=-\infty}^{+\infty} G(f_n) e^{2\pi j f_n t} \quad (4-6)$$

Sin embargo, en el par transformado de las ecuaciones 4-1a [o 4-1b] y 4-2 la función $g(t)$ no es periódica, mientras que en el par transformado de las ecuaciones 4-3 y 4-5 sí lo es, con período igual a T .

Considérese el caso en que tanto el tiempo como la frecuencia son variables discretas. En este caso, la función del tiempo y su transformada de Fourier son pec-

Tabla 4-2. Algunas transformadas de Fourier comunes

$g(t)$	$G(f)$	$G_1(\omega)$
1	$\delta(f)$	$2\pi\delta(\omega)$
$\delta(t - T)$	$\cos 2\pi fT - j \operatorname{sen} 2\pi fT$	$e^{-j\omega T}$
$u(t)$ (función escalón unitario)	$\frac{\delta(f)}{2} - \frac{j}{2\pi f}$	$\pi\delta(\omega) - \frac{j}{\omega}$
$\operatorname{sgn}(t)$	$\frac{-j}{\pi f}$	$\frac{-2j}{\omega}$
$ t ^{-1/2}$	$ f ^{-1/2}$	$2\pi \omega ^{-1/2}$
$e^{-a t }$	$\frac{2a}{a^2 + (2\pi f)^2}$	$\frac{2a}{a^2 + \omega^2}$
e^{-at^2}	$\sqrt{\frac{\pi}{a}} e^{-\pi^2 f^2/a}$	$\sqrt{\frac{\pi}{a}} e^{-\omega^2/4a}$
$\operatorname{sen} at$	$\frac{j}{2} \left[\delta\left(f + \frac{a}{2\pi}\right) - \delta\left(f - \frac{a}{2\pi}\right) \right]$	$j\pi[\delta(\omega + a) - \delta(\omega - a)]$
$\cos at$	$\frac{1}{2} \left[\delta\left(f + \frac{a}{2\pi}\right) + \delta\left(f - \frac{a}{2\pi}\right) \right]$	$\pi[\delta(\omega + a) + \delta(\omega - a)]$

riódicas. Sea $g(t_k)$ la señal de tiempo discreto en la que $t_k = k \Delta t$, y sea el incremento de tiempo Δt definido como:

$$\Delta t = \frac{1}{f_s} \quad (4-7)$$

donde f_s es la frecuencia de muestreo a la que se muestrea la función $g(t)$.

Entonces, el par transformado de Fourier correspondiente será:

$$G(f_n) = \Delta t \sum_k g(t_k) e^{-2\pi j f_n t_k} \quad (4-8)$$

$$g(t_k) = \sum_n G(f_n) e^{2\pi j f_n t_k} \quad (4-9)$$

donde la sumatoria de la ecuación 4-8 se evalúa en un período de $g(t_k)$ y la sumatoria de la ecuación 4-9 se evalúa en un período de $G(f_n)$. Las ecuaciones 4-8 y 4-9 son de importancia fundamental en el procesamiento de señales discretas, y se las conoce como *par transformado discreto de Fourier* (TDF). Obsérvese que las ecuaciones 4-8 y 4-9 pueden considerarse las versiones discretas de las ecuaciones 4-1a [o 4-1b] y 4-2. Sin embargo, en el par transformado de las ecuaciones 4-1a [o 4-1b] y 4-2 las funciones $g(t)$ y $G(f)$ no son periódicas, mientras que en el par TDF de las ecuaciones 4-8 y 4-9 dichas funciones son periódicas con periodos T y $2\pi f_s$, respectivamente.

En la figura 4-1 se muestran algunas de las características del par TDF.

Si el intervalo en el dominio del tiempo se desplaza hacia la derecha de modo que un período de $g(t_k)$ empiece en $t_k = 0$, entonces por la ecuación 4-7 obtenemos:

$$K \Delta t = \frac{K}{f_s} = T \quad (4-10)$$

donde K es el número de muestras en el intervalo $0 \leq t \leq T$, y por la ecuación 4-4:

$$K \Delta \omega = K \frac{2\pi}{T} = 2\pi f_s \quad (4-11)$$

Si el intervalo de interés en el dominio de la frecuencia se desplaza hacia la derecha de modo que $n = 0$ corresponda al inicio de un período, entonces de la ecuación 4-11 es obvio que K es también el número de muestras en el intervalo de frecuencia $0 \leq f \leq f_s$. Obsérvese que $n = K$ corresponde al inicio de un nuevo período.

Sean $T = 1$, $W = e^{-j(2\pi / K)}$, $G(n) = G(f_n)$ y $g(k) = g(t_k)$. Las ecuaciones 4-8 y 4-9 pasan a ser:

$$G(n) = \frac{1}{K} \sum_{k=0}^{K-1} g(k) W^{nk}, \quad n = 0, 1, \dots, K-1 \quad (4-12)$$

$$g(k) = \sum_{n=0}^{K-1} G(n) W^{-nk}, \quad k = 0, 1, \dots, K-1 \quad (4-13)$$

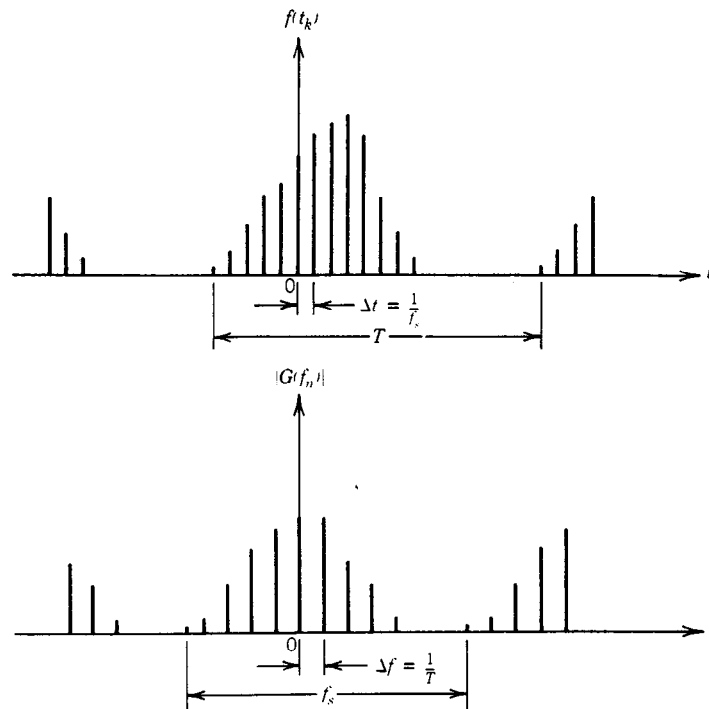


Fig. 4-1. Ilustración de algunas propiedades del par TDF. (Las funciones no necesariamente representan algún par transformado real.)

Las ecuaciones anteriores representan el par TDF comúnmente empleado. En la práctica, K suele tomarse como una potencia de 2 para facilitar el cálculo; es decir:

$$K = 2^\gamma, \quad \gamma \text{ entero} > 1 \quad (4-14)$$

A partir de las ecuaciones 4-12 y 4-13 puede demostrarse con facilidad que el comportamiento de la función $g(k)$ en los intervalos de $K/2$ a K y de $-K/2$ a 0 es idéntico al de la función $G(n)$ en los mismos intervalos. Por tanto, si $g(t_k)$ es una función par, se tiene $g(K-k) = g(k)$ y si es impar se tiene $g(K-k) = -g(k)$ para todo entero k . Propiedades parecidas se cumplen para la función $G(f_n)$. Por tanto, es posible comprobar las propiedades del par TDF en la tabla 4-3 utilizando las ecuaciones 4-12 y 4-13.

Como puede observarse en la tabla 4-3, cuando $g(k)$ es función real del tiempo —lo cual ocurre en la mayor parte de los casos— la magnitud $|G(n)|$ sobre la mitad

del intervalo es idéntica a la de la otra mitad del intervalo. De este modo, la máxima frecuencia es precisamente:

$$f_0 = \frac{f_s}{2} \quad (4-15)$$

que se conoce como *frecuencia de doblaje*. El punto $n = K/2$ corresponde a esta frecuencia (fig. 4-2).

4.4 TRANSFORMADA RÁPIDA DE FOURIER

La transformada rápida de Fourier (TRF) es un algoritmo eficiente que permite calcular con rapidez la TDF en una computadora digital. En el algoritmo, que fue desarrollado por Cooley y Tukey³ en 1965, se aprovechan las propiedades cíclicas de W^{nk} .

Para ilustrar el algoritmo de la TRF, considérese la fórmula de la TDF de la ecuación 4-12. Para $K = 4$ se obtiene:

$$KG(0) = g(0)W^0 + g(1)W^0 + g(2)W^0 + g(3)W^0 \quad (4-16a)$$

$$KG(1) = g(0)W^0 + g(1)W^1 + g(2)W^2 + g(3)W^3 \quad (4-16b)$$

$$KG(2) = g(0)W^0 + g(1)W^2 + g(2)W^4 + g(3)W^6 \quad (4-16c)$$

Tabla 4-3. Propiedades de un par transformado discreto de Fourier

$g(k)$	$G(n)$
Real	Parte real par, parte imaginaria impar
Real y par	Real y par
Real e impar	Imaginaria e impar

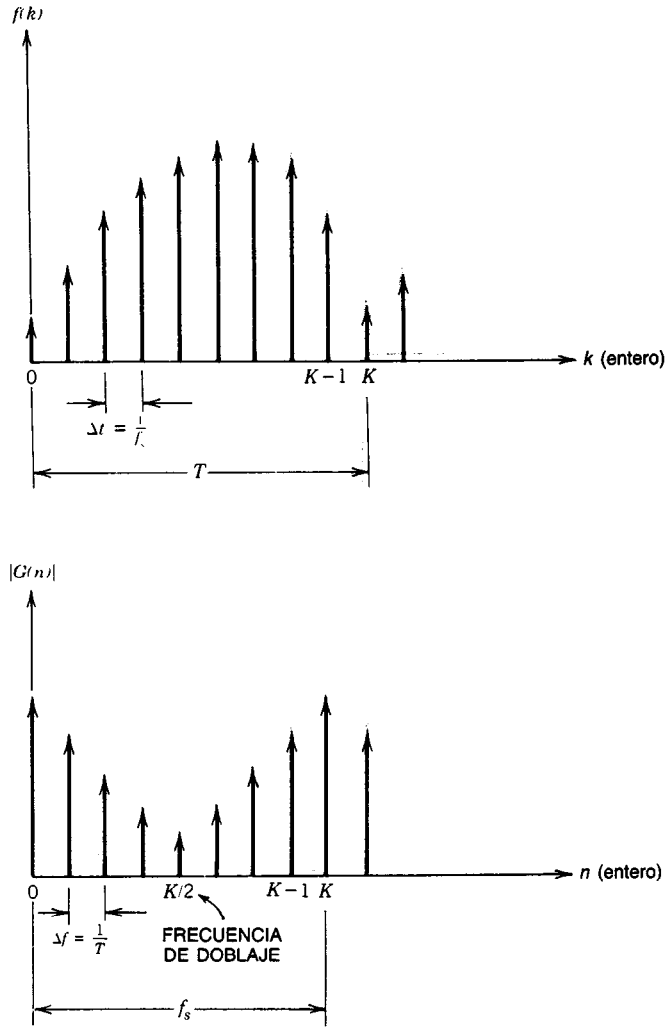


Fig. 4-2. Ilustración de algunas propiedades del par TDF cuando la función del tiempo $f(k)$ es real. (Las funciones no necesariamente representan algún par transformado real.)

$$KG(3) = g(0)W^0 + g(1)W^3 + g(2)W^6 + g(3)W^9 \quad (4-16d)$$

$$KG(1) = g(0)W^0 + g(1)W^1 - g(2)W^0 - g(3)W^1 \quad (4-18b)$$

Obsérvese que para $K = 4$ se tiene $We^{-j(2\pi/4)} = -j$.
Por tanto:

$$W^0 = W^4 = W^8 = 1 \quad (4-17a)$$

$$W^1 = W^5 = W^9 = -j \quad (4-17b)$$

$$W^2 = W^6 = -1 \quad (4-17c)$$

$$W^3 = W^7 = +j \quad (4-17d)$$

$$KG(2) = g(0)W^0 - g(1)W^0 + g(2)W^0 - g(3)W^0 \quad (4-18c)$$

$$KG(3) = g(0)W^0 - g(1)W^1 - g(2)W^0 + g(3)W^1 \quad (4-18d)$$

o bien:

$$KG(0) = P(0)W^0 + P(2)W^0 \quad (4-19a)$$

$$KG(1) = P(1)W^0 + P(3)W^1 \quad (4-19b)$$

$$KG(2) = P(0)W^0 - P(2)W^0 \quad (4-19c)$$

$$KG(3) = P(1)W^0 - P(3)W^1 \quad (4-19d)$$

Obsérvese que W^{nk} es cíclica con periodo 4. Por consiguiente, las ecuaciones 4-16a a 4-16d pasan a tener la forma:

$$KG(0) = g(0)W^0 + g(1)W^0 + g(2)W^0 + g(3)W^0 \quad (4-18a)$$

donde:

$$P(0) = g(0) + g(2) \quad (4-20a)$$

$$P(1) = g(0) - g(2) \quad (4-20b)$$

$$P(2) = g(1) + g(3) \quad (4-20c)$$

$$P(3) = g(1) - g(3) \quad (4-20d)$$

Comparando las ecuaciones 4-16a a 4-16d con las ecuaciones 4-19a a 4-19d se demuestra que el algoritmo de la TRF reduce considerablemente la cantidad necesaria de cálculos para evaluar $G(n)$, $n = 0, 1, 2, 3$. En general, existen K^2 multiplicaciones con forma semejante a la propia de las ecuaciones 4-16a a 4-16d, pero solamente existen $2K$ multiplicaciones por factor parecidas a las ecuaciones 4-19a a 4-19d.

El método de la TRF tiene aplicaciones en el procesamiento de señales, especialmente cuando el procesamiento en tiempo real tiene importancia básica (p. ej., en el procesamiento de señales de radar).

4.5 TRANSFORMADA DE LAPLACE

4.5.1 Determinación de la transformada de Laplace

La transformada de Laplace es de utilidad para transformar funciones en el dominio del tiempo en funciones en el dominio complejo de la frecuencia. Puede utilizarse para transformar una ecuación diferencial lineal con coeficientes constantes en una ecuación algebraica, reduciendo así considerablemente el esfuerzo de cálculo requerido para resolverla.

Sea $f(t)$ una función integrable del tiempo definida en el intervalo $[0, \infty]$. Entonces la transformada (unilateral) de Laplace de $f(t)$, denotada por $L[f(t)] \triangleq F(s)$, se define como la integral:

$$F(s) = \int_{0-}^{\infty} f(t)e^{-st} dt \quad (4-21)$$

para todos los valores de la frecuencia compleja s para los cuales existe la integral. El límite inferior de inte-

gración es 0^- , de modo que si la función del tiempo f incluye un impulso en $t = 0$, también la integral lo incluye.

De la ecuación 4-21 es claro que si $F(s)$ existe para alguna $s = \sigma_c + j\omega_c$, entonces también existe para toda s tal que $\text{Re}(s) > \sigma_c$. La máxima cota inferior en σ_c para la que existe la integral de la ecuación 4-21 se denomina *abscisa de convergencia* para la función $f(t)$. La región del plano complejo que se encuentra a la derecha de la abscisa de convergencia se llama *región o dominio de convergencia* de $F(s)$ (fig. 4-3).

Es posible extender la integral de Laplace (ec. 4-21) a otros valores de $s = \sigma + j\omega$. Es decir, puede considerarse que $F(s)$ es una función bien definida para todos los valores de s , excepto para aquellos que hacen diverger la integral de la ecuación 4-21.

Ejemplo 1. Considérese la función $f(t) = e^{\alpha t}$, donde α es un número real o complejo. Entonces:

$$L[f(t)] \triangleq F(s) = \int_{0-}^{\infty} e^{\alpha t} e^{-st} dt = \frac{e^{-(s-\alpha)t}}{-(s-\alpha)} \Big|_{0-}^{\infty} \quad (4-22)$$

Por tanto, $F(s)$ existe; es decir, es finita cuando $\text{Re}(s - \alpha) > 0$, y en cualquier otro caso es infinita. Esto es:

$$F(s) = L[e^{\alpha t}] = \frac{1}{s - \alpha} \quad \text{para } \text{Re}(s) > \text{Re}(\alpha) \quad (4-23)$$

y $\text{Re}(\alpha)$ es la abscisa de convergencia para $e^{\alpha t}$. Sin embargo, se considerará que la transformada de Laplace de $e^{\alpha t}$ es $1/(s - \alpha)$, que está bien definida para toda s excepto para $s = \alpha$.

La unicidad es una propiedad fundamental de la transformada de Laplace. Es decir, si dos funciones del tiempo $f_1(t)$ y $f_2(t)$ tienen por transformada de Laplace la misma función de la frecuencia compleja s , por ejemplo, $F(s)$, entonces la diferencia entre $f_1(t)$ y $f_2(t)$ es trivial. Por consiguiente, para una función del tiem-

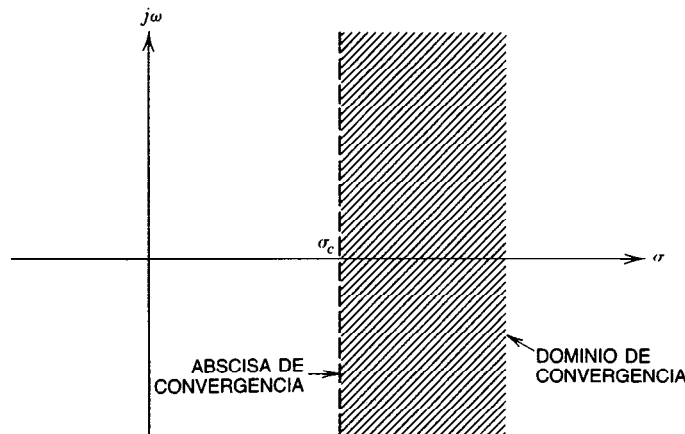


Fig. 4-3. Dominio de convergencia de la transformada de Laplace en el plano s .

po $f(t)$ sólo existe una "única" función $F(s)$ como su transformada de Laplace. Recíprocamente, dada una transformada de Laplace $F(s)$, existe una única función del tiempo $f(t)$ sobre el intervalo $[0, \infty]$ (salvo trivialidades) tal que $F(s) = L[f(t)]$. Lo anterior se expresa como:

$$f(t) = L^{-1}[F(s)] \quad (4-24)$$

lo que significa que $f(t)$ es la transformada inversa de Laplace de $F(s)$.

Otra propiedad importante de la transformada de Laplace es la linealidad. Es decir, si $F_1(s)$ y $F_2(s)$ son las transformadas de Laplace de las funciones del tiempo $f_1(t)$ y $f_2(t)$, respectivamente, entonces:

$$L[\alpha_1 f_1(t) + \alpha_2 f_2(t)] = \alpha_1 F_1(s) + \alpha_2 F_2(s) \quad (4-25)$$

para constantes arbitrarias α_1 y α_2 . La propiedad de linealidad de la transformada de Laplace es una consecuencia inmediata de la linealidad de la integral de Laplace (ec. 4-21).

Ejemplo 2. Considérese:

$$f(t) = \sin at = \frac{1}{2j} (e^{j\alpha t} - e^{-j\alpha t}) \quad (4-26)$$

Utilizando la ecuación 4-23 y la propiedad de linealidad de la transformada de Laplace se obtiene la siguiente expresión:

$$L[\sin at] = \frac{1}{2j} \left[\frac{1}{s - j\alpha} - \frac{1}{s + j\alpha} \right] = \frac{\alpha}{s^2 + \alpha^2} \quad (4-27)$$

Sea $F(s)$ la transformada de Laplace de $f(t)$. Entonces:

$$L \left[\frac{d^n}{dt^n} f(t) \right] = s^n F(s) - s^{n-1} f(0^-) - s^{n-2} f^{(1)}(0^-) - \dots - s f^{(n-2)}(0^-) - f^{(n-1)}(0^-) \quad (4-28)$$

donde $f^{(i)}$ indica la i -ésima derivada de f con respecto a t .

Esta relación se denomina *regla de diferenciación* de la transformada de Laplace, que puede comprobarse evaluando por partes la integral de Laplace (ec. 4-21) para $(d/dt)f(t)$:

$$\begin{aligned} L \left[\frac{d}{dt} f(t) \right] &= f(t)e^{-st} \Big|_{0^-}^{\infty} - \int_{0^-}^{\infty} f(t)(-se^{-st}) dt = \\ &= -f(0^-) + s \int_{0^-}^{\infty} f(t)e^{-st} dt = sF(s) - f(0^-) \end{aligned} \quad (4-29)$$

Aplicando $n-1$ veces la ecuación 4-29 se obtiene la ecuación 4-28.

Ejemplo 3. Sea $(d/dt) \sin at = \alpha \cos at$. Por tanto, utilizando las ecuaciones 4-27 y 4-28 se obtiene:

$$L \left[\frac{d}{dt} \sin at \right] = L[\alpha \cos at] = s \left[\frac{\alpha}{s^2 + \alpha^2} \right] \quad (4-30)$$

y usando la linealidad de la transformada de Laplace:

$$L[\cos at] = \frac{s}{s^2 + \alpha^2} \quad (4-31)$$

Si $L[f(t)] = F(s)$, entonces:

$$L \left[\int_{0^-}^t \int_{0^-}^{\tau_1} \dots \int_{0^-}^{\tau_{n-1}} f(\tau_n) d\tau_n d\tau_{n-1} \dots d\tau_1 \right] = \frac{1}{s^n} F(s) \quad (4-32)$$

La ecuación 4-32 se denomina *regla de integración* de la transformada de Laplace. De nuevo, puede comprobarse utilizando la integración por partes:

$$\begin{aligned} L \left[\int_{0^-}^t f(\tau) d\tau \right] &\triangleq \int_{0^-}^{\infty} \left[\int_{0^-}^t f(\tau) d\tau \right] e^{-st} dt \\ &= \left[\int_{0^-}^t f(\tau) d\tau \right] \frac{e^{-st}}{-s} \Big|_{0^-}^{\infty} - \int_{0^-}^{\infty} f(t) \left(\frac{e^{-st}}{-s} \right) dt \\ &= 0 + \frac{1}{s} \int_{0^-}^{\infty} f(t) e^{-st} dt = \frac{1}{s} F(s) \end{aligned} \quad (4-33)$$

La aplicación repetida de la ecuación 4-33 produce la ecuación 4-32.

La propiedad de "traslación de la función del tiempo" de la transformada de Laplace implica que si $L[f(t)] = F(s)$, entonces para cualquier escalar positivo α :

$$L[f(t - \alpha)] = e^{-s\alpha} F(s) \quad (4-34)$$

Además, el escalar α puede ser negativo si $f(t - \alpha)$ desaparece para $t < 0$. Esta propiedad es una consecuencia directa de la integral que define la transformada de Laplace:

$$\begin{aligned} L[f(t - \alpha)] &\triangleq \int_{0^-}^{\infty} f(t - \alpha) e^{-st} dt = \int_{0^-}^{\infty} f(\tau) e^{-s\tau} d\tau = \int_{0^-}^{\infty} f(\tau) e^{-s(\tau + \alpha)} d\tau \\ &= e^{-s\alpha} \int_{0^-}^{\infty} f(\tau) e^{-s\tau} d\tau = e^{-s\alpha} F(s) \end{aligned} \quad (4-35)$$

donde $\tau = t - \alpha$. Obsérvese que $f(t - \alpha)$ es la función original retrasada por α .

La propiedad de *traslación de la transformada de Laplace* implica que si $L[f(t)] = F(s)$, entonces:

$$L[e^{at} f(t)] = F(s - \alpha) \quad (4-36)$$

donde α es cualquier escalar complejo. Otras dos propiedades de la transformada de Laplace son las que se refieren al "valor inicial" y al "valor final".

La primera propiedad mencionada implica que si $f(t)$ tiende al límite cuando t tiende a cero por la derecha, entonces:

$$f(0+) = \lim_{s \rightarrow \infty} sL[f(t)] \quad (4-37)$$

Obsérvese que el límite del miembro derecho de la ecuación 4-37 puede existir aun cuando $f(0+)$ no exista.

La segunda propiedad implica que si $f(t)$ tiende al límite cuando t tiende a infinito, entonces:

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} sL[f(t)] \quad (4-38)$$

De nuevo, obsérvese que el límite del miembro derecho de la ecuación 4-38 puede existir aunque no exista el del miembro izquierdo.

Ejemplo 4. Considérese $f(t) = e^{-\alpha t} \cos \omega t$, donde α es un número real positivo.

Empleando las ecuaciones 4-31 y 4-36 se tiene la expresión:

$$L[f(t)] \triangleq F(s) = \frac{s + \alpha}{(s + \alpha)^2 + \omega^2} \quad (4-39)$$

Ahora $\lim_{s \rightarrow \infty} sF(s) = 1$, que es igual a $f(0+)$. También $\lim_{s \rightarrow 0} sF(s) = 0$, que es igual a $\lim_{t \rightarrow \infty} f(t)$.

La última propiedad de la transformada de Laplace que se estudiará es la propiedad de "convolución". Si sobre el intervalo $[0, \infty]$ se encuentran definidas funciones del tiempo $f_1(t)$ y $f_2(t)$, la convolución de f_1 y f_2 , denotada $f_1 * f_2$, es una función del tiempo definida por:

$$f_1 * f_2(t) = \int_0^t f_1(\tau) f_2(t - \tau) d\tau \quad (4-40)$$

siempre que la integral exista para toda t . Si $F_1(s)$ y $F_2(s)$ son las transformadas de Laplace de $f_1(t)$ y $f_2(t)$, respectivamente, entonces por la propiedad de convolución de la transformada de Laplace:

$$L[f_1 * f_2(t)] = F_1(s) F_2(s) \quad (4-41)$$

Ésta es una propiedad importante de la transformada de Laplace que se utiliza con frecuencia para el estudio de sistemas lineales invariantes de tiempo continuo.

4.5.2 Inversión de la transformada de Laplace

La importancia de la transformada de Laplace radica en el hecho de que existe una única función del tiempo $f(t)$ para cada transformada de Laplace $F(s)$. Sea $F(s) = L[f(t)]$ y sea σ_c la abscisa de convergencia para la función $f(t)$. Entonces $f(t)$ puede recuperarse como sigue:

$$f(t) \triangleq L^{-1}[F(s)] = \frac{1}{2\pi j} \lim_{\omega \rightarrow \infty} \int_{\sigma - j\omega}^{\sigma + j\omega} F(s) e^{st} ds \quad (4-42)$$

para toda $\sigma > \sigma_c$. La demostración de este hecho se encuentra en la referencia 4. Obsérvese que por hipótesis $f(t) = 0$ para toda $t < 0$; por tanto, la integral (ec. 4-42) produce $f(t)$ para $t > 0$. Además, $f(\cdot)$ puede ser discontinua en t ; de hecho, el miembro izquierdo de la ecuación 4-42 puede sustituirse por: $\frac{1}{2} [f(t+0) + f(t-0)]$.

En general, la integral contenida en la ecuación 4-42 es difícil de calcular y rara vez se utiliza directamente para invertir la transformada de Laplace. En lugar de eso, la transformada inversa de Laplace suele determinarse desarrollando la función en componentes inversibles con facilidad y aplicando la propiedad de linealidad de la transformada de Laplace. Este procedimiento puede formalizarse como sigue. Sea $F(s)$ una función racional; es decir, el cociente de dos polinomios en s . Además, sea $F(s)$ estrictamente propia; es decir, si el grado de su numerador es m y el de su denominador es n , entonces $m < n$. Por tanto, $F(s)$ puede expresarse como:

$$\begin{aligned} F(s) = \frac{Q(s)}{P(s)} &= \frac{q_m s^m + q_{m-1} s^{m-1} + \dots + q_1 s + q_0}{p_n s^n + p_{n-1} s^{n-1} + \dots + p_1 s + p_0} = \\ &= \frac{q_m}{p_n} \frac{\prod_{i=1}^m (s - z_i)}{\prod_{i=1}^n (s - p_i)} \end{aligned} \quad (4-43)$$

donde los coeficientes q_j , $j = 0, 1, \dots, m$ y p_i , $i = 0, 1, \dots, n$ son reales.

Entonces existen números complejos K_{ij} tales que es posible expresar $F(s)$ como:

$$F(s) = \sum_{i=1}^p \sum_{j=1}^{m_i} \frac{K_{ij}}{(s - p_i)^j} \quad (4-44)$$

donde m_i es la multiplicidad de la raíz p_i del polinomio $P(s)$ y p es el número de raíces diferentes de $P(s)$.

Los números complejos K_{ij} se determinan de la manera siguiente:

$$\begin{aligned} K_{i, m_i - j + 1} &= \frac{1}{(j - 1)!} \frac{d^{j-1}}{ds^{j-1}} (s - p_i)^{m_i} F(s) \Big|_{s=p_i}, \\ j &= 1, 2, \dots, m_i; \quad i = 1, 2, \dots, p \end{aligned} \quad (4-45)$$

Entonces la transformada inversa de Laplace $F(s)$ es:

$$f(t) = \sum_{i=1}^p \sum_{j=1}^{m_i} \frac{K_{ij} t^{j-1}}{(j-1)!} e^{\lambda_i t} \quad (4-46)$$

La demostración de este teorema puede verse en la referencia 4.

Las raíces z_j , $j = 1, 2, \dots, m$ del polinomio $Q(s)$ se denominan los *ceros* de $F(s)$, y las raíces de p_i , $i = 1, 2, \dots, n$ del polinomio $P(s)$ se denominan los *polos* de

$F(s)$. Claramente, la abscisa de convergencia para $f(t)$ es el $\max_i \{\operatorname{Re}(p_i)\}$. Los números complejos K_j , $j = 1, 2, \dots, m_i$ se denominan los *residuos* correspondientes al polo p_j . Los polos, ceros y residuos o son reales o se presentan en pares complejos conjugados debido a que se ha supuesto que los coeficientes p_i y q_j son reales. Obsérvese que $\sum_{i=1}^n m_i = n$.

También debe observarse que en el caso en que todos los polos de $F(s)$ son diferentes, es decir, en el que $m_i = 1$, $i = 1, 2, \dots, p$, entonces $p = n$, y la ecuación 4-44 pasa a ser:

$$F(s) = \sum_{i=1}^n \frac{K_i}{s - p_i} \quad (4-47)$$

donde:

$$K_i = (s - p_i)F(s)|_{s=p_i}, \quad i = 1, 2, \dots, n \quad (4-48)$$

y la ecuación 4-46 se convierte en:

$$f(t) = \sum_{i=1}^n K_i e^{p_i t} \quad (4-49)$$

Ejemplo 5. Considérese:

$$F(s) = \frac{s}{(s+1)^3(s+2)} \quad (4-50)$$

que puede expresarse como:

$$F(s) = \frac{K_{11}}{s+1} + \frac{K_{12}}{(s+1)^2} + \frac{K_{13}}{(s+1)^3} + \frac{K_{21}}{s+2} \quad (4-51)$$

donde:

$$K_{13} = (s+1)^3 F(s)|_{s=-1} = -1 \quad (4-52)$$

$$K_{12} = \frac{d}{ds} (s+1)^3 F(s)|_{s=-1} = \frac{2}{(s+2)^2} \Big|_{s=-1} = 2 \quad (4-53)$$

$$K_{11} = \frac{1}{2} \frac{d^2}{ds^2} (s+1)^3 F(s)|_{s=-1} = \frac{-2}{(s+1)^3} \Big|_{s=-1} = -2 \quad (4-54)$$

$$K_{21} = (s+2)F(s)|_{s=-2} = 2 \quad (4-55)$$

Por tanto:

$$F(s) = \frac{-2}{s+1} + \frac{2}{(s+1)^2} - \frac{1}{(s+1)^3} + \frac{2}{s+2} \quad (4-56)$$

y

$$L^{-1}[F(s)] = -2e^{-t} + 2te^{-t} - \frac{t^2}{2}e^{-t} + 2e^{-2t} \quad t \geq 0 \quad (4-57)$$

4.5.3 Tabla de transformadas de Laplace

En la tabla 4-4 se han reunido las transformadas de Laplace de algunas funciones comunes del tiempo, así como transformadas inversas de Laplace de algunas funciones comunes de s . Obsérvese que se supone que la función del tiempo $f(t)$ desaparece para $t < 0$.

4.5.4 Solución de ecuaciones diferenciales utilizando la transformada de Laplace

Una de las aplicaciones importantes de la transformada de Laplace se da en la solución de ecuaciones diferenciales lineales con coeficientes constantes. Determinando la transformada de Laplace de ambos miembros de una de estas ecuaciones, es posible escribir la transformada de Laplace de la función desconocida como una función racional de s . Así, la transformada inversa de Laplace de esta función constituye la solución. Por tanto, la transformada de Laplace convierte la solución de una ecuación diferencial lineal con coeficientes constantes en la solución de una ecuación algebraica. El siguiente ejemplo ilustra el método.

Ejemplo 6. Sea la ecuación diferencial:

$$\ddot{f}(t) + 10\dot{f}(t) + 37f(t) + 52f(t) = -3u(t) + \delta(t), \quad t \geq 0 \quad (4-58)$$

con condiciones iniciales:

$$f(0^-) = \dot{f}(0^-) = 1, \quad \ddot{f}(0^-) = 0 \quad (4-59)$$

Determinando la transformada de Laplace de ambos miembros de la ecuación 4-58 se obtiene:

$$\begin{aligned} & [s^3 F(s) - s^2 f(0^-) - s\dot{f}(0^-) - \ddot{f}(0^-)] + \\ & + 10[s^2 F(s) - s\dot{f}(0^-) - \dot{f}(0^-)] + 37[sF(s) - f(0^-)] + \\ & + 52F(s) = \frac{3}{s} + s \end{aligned} \quad (4-60)$$

o bien:

$$F(s) = \frac{s^3 + 12s^2 + 47s - 3}{s(s^3 + 10s^2 + 37s + 52)} \quad (4-61)$$

que puede expresarse como suma de fracciones parciales de la forma:

$$F(s) = \frac{-3/52}{s} + \frac{63/20}{s+4} - \frac{(136/65)(s+3) + 541/130}{(s+3)^2 + 4} \quad (4-62)$$

Por tanto, la solución es:

$$\begin{aligned} f(t) = L^{-1}[F(s)] = & \frac{3}{52} + \frac{63}{20} e^{-4t} - \\ & - \frac{136}{65} e^{-3t} \cos 2t + \frac{541}{130} e^{-3t} \sin 2t, \quad t \geq 0 \end{aligned} \quad (4-63)$$

Tabla 4-4. Algunas funciones comunes y sus transformadas de Laplace

No.	$f(t)$	$F(s)$
1	$\delta(t)$	1
2	$u(t)$ (función escalón unitario)	$\frac{1}{s}$
3	$u(t - \alpha)$	$\frac{e^{-\alpha s}}{s}$
4	$r(t) = tu(t)$ (rampa unitaria)	$\frac{1}{s^2}$
5	t^n	$\frac{n!}{s^{n+1}}$
6	$t^n e^{\alpha t}$	$\frac{n!}{(s - \alpha)^{n+1}}$
7	$-1 - \alpha t + e^{\alpha t}$	$\frac{\alpha^2}{s^2(s - \alpha)^2}$
8	$1 - e^{\alpha t} + \alpha t e^{\alpha t}$	$\frac{\alpha^2}{s(s - \alpha)^2}$
9	$\frac{\beta}{\alpha^2} + \frac{1 + \alpha t - \beta t}{\alpha} e^{\alpha t}$	$\frac{s - \beta}{s(s - \alpha)^2}$
10	$\text{sen } \omega t$	$\frac{\omega}{s^2 + \omega^2}$
11	$\text{cos } \omega t$	$\frac{s}{s^2 + \omega^2}$
12	$e^{\alpha t} \text{sen } \omega t$	$\frac{\omega}{(s - \alpha)^2 + \omega^2}$
13	$e^{\alpha t} \text{cos } \omega t$	$\frac{s - \alpha}{(s - \alpha)^2 + \omega^2}$
14	$\frac{1}{\omega^2} (1 - \text{cos } \omega t)$	$\frac{1}{s(s^2 + \omega^2)}$
15	$e^{-\xi \omega_n t} \text{sen}(\omega_n \sqrt{1 - \xi^2} t)$	$\frac{\omega_n \sqrt{1 - \xi^2}}{s^2 + 2\xi \omega_n s + \omega_n^2}$
16	$\frac{e^{\alpha t}}{\alpha^2 + \omega^2} + \frac{\text{sen}(\omega t - \phi)}{\alpha \sqrt{\alpha^2 + \omega^2}}$ donde $\phi = -\tan^{-1} \frac{\omega}{\alpha}$	$\frac{1}{(s - \alpha)(s^2 + \omega^2)}$
17	$e^{\alpha t} (A \text{cos } \omega t + B \text{sen } \omega t)$	$\frac{A(s - \alpha) + B\omega}{(s - \alpha)^2 + \omega^2}$
18	$e^{\alpha t} (A \text{cos } \omega t + \frac{B + \alpha A}{\omega} \text{sen } \omega t)$	$\frac{As + B}{(s - \alpha)^2 + \omega^2}$
19	$\frac{\sqrt{(\alpha + \beta)^2 + \omega^2}}{\omega} e^{\alpha t} \text{sen}(\omega t + \delta)$ donde $\delta = \tan^{-1} \frac{\omega}{\alpha + \beta}$	$\frac{s + \beta}{(s - \alpha)^2 + \omega^2}$
20	$2 K e^{\alpha t} \text{cos}(\omega t + \angle K)$	$\frac{K}{s - \alpha - j\omega} + \frac{\bar{K}}{s - \alpha - j\omega}$

Tabla 4-4. Algunas funciones comunes y sus transformadas de Laplace (continuación)

No.	$f(t)$	$F(s)$
21	$\frac{1}{\alpha^2 + \omega^2} + \frac{e^{\alpha t}}{\omega \sqrt{\alpha^2 + \omega^2}} \operatorname{sen}(\omega t - \delta)$ donde $\delta = -\tan^{-1} \frac{\omega}{\alpha}$	$\frac{1}{s[(s - \alpha)^2 + \omega^2]}$
22	$\frac{1}{\omega_n^2} - \frac{e^{\xi \omega_n t}}{\omega_n \sqrt{1 - \xi^2}} \operatorname{sen}(\omega_n \sqrt{1 - \xi^2} t + \delta)$ donde $\delta = -\cos^{-1} \xi$	$\frac{1}{s(s^2 + 2\xi \omega_n s + \omega_n^2)}$
23	$\frac{\beta}{\alpha^2 + \omega^2} + \frac{e^{\alpha t} \sqrt{(\alpha + \beta)^2 + \omega^2}}{\omega \sqrt{\alpha^2 + \omega^2}} \operatorname{sen}(\omega t + \delta)$ donde $\delta = -\tan^{-1} \frac{\omega}{\alpha + \beta} - \tan^{-1} \frac{\omega}{\alpha}$	$\frac{s + \delta}{s[(s - \alpha)^2 + \omega^2]}$

4.6 TRANSFORMADA z

4.6.1 Determinación de la transformada z

El concepto de transformada z está estrechamente relacionado con la idea de muestreo de una función del tiempo. Considérese una función de tiempo continuo $f(t)$, donde $f(t) = 0$ para $t < 0$. Un "muestreador ideal" es aquel que toma muestras de ancho infinitesimal de $f(t)$ en intervalos regulares de tiempo (fig. 4-4). Por tanto, una muestra ideal $f^*(t)$ de $f(t)$ puede considerarse una sucesión de impulsos equiespaciados con magnitudes iguales a los valores de la función $f(\cdot)$ en los instantes discretos correspondientes. Es decir:

$$f^*(t) = \sum_{k=0}^{\infty} f(kT) \delta(t - kT) \quad (4-64)$$

donde T se denomina *periodo de muestreo*. Determinando la transformada de Laplace de ambos miembros de la ecuación 4-64 se obtiene:

$$\begin{aligned} L[f^*(t)] &\triangleq F^*(s) = \sum_{k=0}^{\infty} f(kT) L[\delta(t - kT)] = \\ &= \sum_{k=0}^{\infty} f(kT) e^{-kTs} \end{aligned} \quad (4-65)$$

Ahora, la variable compleja z se define de la siguiente forma:

$$z = e^{Ts} \quad (4-66)$$

Por tanto, la ecuación 4-65 se convierte en:

$$F^*(s) = \sum_{k=0}^{\infty} f(kT) z^{-k} \quad (4-67)$$

que se define como la *transformada z (unilateral)* de $f(t)$ y se denota por $F(z)$.

Es decir:

$$Z[f(k)] \triangleq F(z) = \sum_{k=0}^{\infty} f(kT) z^{-k} \quad (4-68)$$

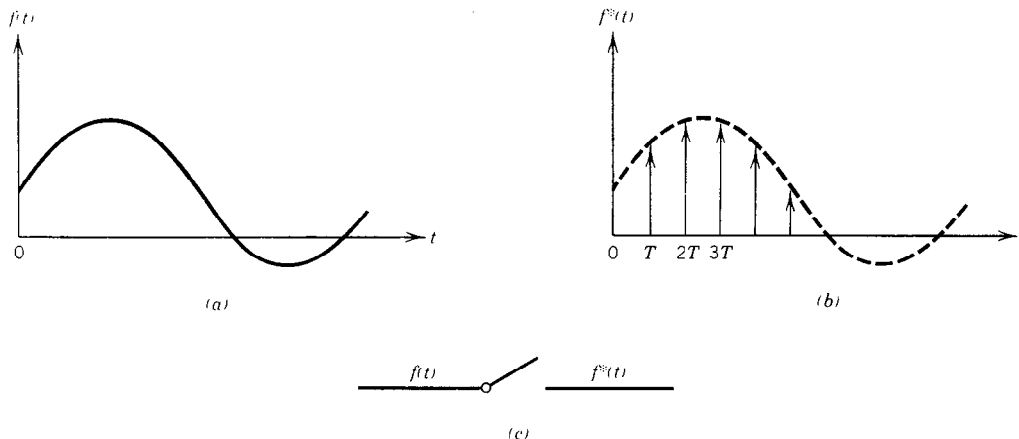


Fig. 4-4. Muestreo ideal: a) función de tiempo continuo; b) sus muestras ideales; c) símbolo del muestreador ideal.

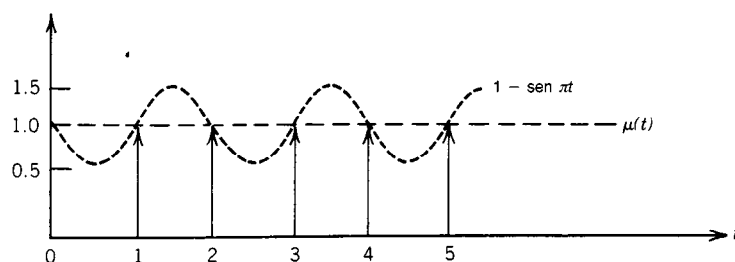


Fig. 4-5. Diferentes funciones del tiempo con las mismas muestras.

La transformada z de una sucesión de escalares $f(kT)$, $k = 0, 1, 2, \dots$ se define como en la ec. 4-68.

La transformada z existe si la suma infinita de la ecuación 4-68 converge. El *radio de convergencia* r_c de la serie infinita de la ecuación 4-68 se define como:

$$r_c = \lim_{k \rightarrow \infty} |f(k)|^{1/k} \quad (4-69)$$

La serie de la ecuación 4-68 es analítica para $|z| > r_c$. Por tanto, converge absolutamente para toda z en el dominio $|z| > r_c$. Por comodidad en la notación, T suele considerarse igual a la unidad.

Ejemplo 1. Considérese la función escalón unitario $u(t)$. Entonces:

$$Z[u(t)] \triangleq U(z) = \sum_{k=0}^{\infty} z^{-k} \quad (4-70)$$

que converge a:

$$\frac{1}{1 - 1/z} = \frac{z}{z - 1} \quad (4-71)$$

si $|z| > 1$. Por tanto, el radio de convergencia de la serie contenida en la ecuación 4-70 es 1.

Obsérvese que la ecuación 4-71 también es la transformada z de la sucesión $u^*(t) = 1, 1, 1, \dots$. Aunque la ecuación 4-71 es la transformada z de esta sucesión única, no corresponde a una función única de tiempo continuo, ya que existen muchas funciones cuyas muestras $f(k)$ en $k = 0, 1, 2, \dots$ son 1 (fig. 4-5).

Ejemplo 2. Considérese la sucesión:

$$f(k) = \alpha^k, \quad k = 0, 1, 2, \dots \quad (4-72)$$

Entonces:

$$\begin{aligned} Z[f(k)] &\triangleq \sum_{k=0}^{\infty} f(k)z^{-k} = \\ &= \sum_{k=0}^{\infty} \alpha^k z^{-k} = \frac{1}{1 - \alpha/z} = \frac{z}{z - \alpha} \end{aligned} \quad (4-73)$$

siempre que $|\alpha/z| < 1$ o que $|z| > |\alpha|$. Por tanto, el radio de convergencia de la serie infinita de la ecuación 4-73 es $|\alpha|$.

Dado que la transformada z está definida a través de la transformada de Laplace, las propiedades de aquella pueden deducirse a partir de las de ésta. La propiedad de unicidad se cumple para la transformada z en el sentido de que una función $F(z)$ dada corresponde a una única "sucesión" $f(k)$. Sin embargo, no corresponde a una única función de tiempo continuo $f(t)$, como se ha analizado.

Así como la de Laplace, la transformada z es lineal; es decir:

$$Z[\alpha_1 f_1(t) + \alpha_2 f_2(t)] = \alpha_1 Z[f_1(t)] + \alpha_2 Z[f_2(t)] \quad (4-74)$$

para constantes arbitrarias α_1 y α_2 .

La propiedad de "avance" de la transformada z es parecida a la regla de diferenciación de la transformada de Laplace:

$$Z[f(k + m)] = z^m F(z) - \sum_{k=0}^{m-1} f(k)z^{m-k} \quad (4-75)$$

para cualquier entero positivo m . Análogamente, la propiedad de "retraso" de la transformada z se parece a la regla de integración de la transformada de Laplace:

$$Z[f(k - m)] = z^{-m} F(z) \quad (4-76)$$

para cualquier entero positivo m .

Otras dos propiedades de la transformada z son las denominadas del "valor inicial" y del "valor final". La primera implica que:

$$f(0+) = \lim_{z \rightarrow \infty} F(z) \quad (4-77)$$

siempre que el límite exista. La segunda implica que si una sucesión $f(k)$ tiende a un límite cuando k tiende a infinito, entonces:

$$\lim_{k \rightarrow \infty} f(k) = \lim_{z \rightarrow 1} (z - 1)F(z) \quad (4-78)$$

si $F(z)$ es analítica para $|z| > 1$. Por último, la propiedad de "convolución" de la transformada z implica que si $F_1(z)$ y $F_2(z)$ son las transformadas z de $f_1(t)$ y $f_2(t)$, respectivamente, entonces:

$$Z\left[\sum_{j=0}^k f_1(j)f_2(k-j)\right] = F_1(z)F_2(z) \quad (4-79)$$

4.6.2 Inversión de la transformada z

Dada una función $F(z)$ de variable compleja z , es posible determinar una función $f(\cdot)$ en el dominio del tiempo de modo que:

$$Z[f(t)] = F(z) \quad (4-80)$$

o, de forma equivalente:

$$f(t) = Z^{-1}[F(z)] \quad (4-81)$$

Sin embargo, obsérvese que a diferencia de la transformada de Laplace, la transformada z inversa $f(t)$ no es única. De hecho, existe una infinidad de funciones $f(t)$ cuya transformada z es $F(z)$. Estas funciones tienen los mismos valores sólo en los instantes de muestreo; es decir, es posible representarlas a todas con la misma muestra ideal $f^*(t)$.

Los métodos para determinar la transformada z inversa pueden agruparse en tres tipos: el de la serie de potencias, el del residuo y el del desarrollo en fracciones parciales. En el método de la serie de potencias se utiliza la serie infinita definida en la ecuación 4-68 para determinar la transformada z inversa de cualquier función $F(z)$ dada. Esto se realiza desarrollando la función racional $F(z)$ en una serie infinita en z^{-k} . Entonces, los coeficientes de z^{-k} serán iguales a $f(k)$. Sin embargo, este procedimiento suele resultar tedioso y, por otra parte, no conduce a una solución en forma cerrada.

Ejemplo 3. Considérese:

$$F(z) = \frac{z}{z-2} \quad (4-82)$$

La división larga produce:

$$F(z) = 1 + 2z^{-1} + 4z^{-2} + 8z^{-3} + \dots \quad (4-83)$$

Por tanto, comparando la expresión precedente con la ecuación 4-68 se tiene:

$$f(0) = 1, \quad f(1) = 2, \quad f(2) = 4, \quad f(3) = 8, \dots \quad (4-84)$$

En el método del residuo se utiliza el hecho de que si $F(z) = Z[f(k)]$, entonces:

$$f(k) = \frac{1}{2\pi j} \oint_{\Gamma} F(z) z^{k-1} dz, \quad k = 0, 1, 2, \dots \quad (4-85)$$

donde la integración se realiza en sentido antihorario y Γ es cualquier curva cerrada que contiene el origen y se encuentra fuera de la circunferencia $|z| > r_c$ [el radio de convergencia de $F(z)$], o de manera equivalente, Γ es cualquier curva cerrada que contenga los polos de $zF(z)$, $k = 0, 1, 2, \dots$. La integral de contorno contenida en la ecuación 4-85 puede evaluarse utilizando el teorema del residuo de Cauchy, el cual establece que si $F(z)$ es una función analítica en una región R excepto

en un número finito de singularidades, y si Γ es una curva cerrada en R que no pasa por dichas singularidades, entonces:

$$\frac{1}{2\pi j} \oint_{\Gamma} F(z) dz = \text{suma de los residuos de } F(z) \quad (4-86)$$

donde la integración se realiza en sentido antihorario, y el residuo de $F(z)$ en un polo $z = p_i$ con multiplicidad m_i se encuentra a partir de la ecuación 4-45. Conforme a este análisis, es obvio que:

$$f(k) = Z^{-1}[F(z)] = \text{suma de los residuos de } F(z) z^{k-1} \quad (4-87)$$

Ejemplo 4. De nuevo, considérese:

$$F(z) = \frac{z}{z-\alpha} \quad (4-88)$$

$F(z)$ tiene un solo polo en $z = \alpha$. Por tanto:

$$f(k) = Z^{-1}[F(z)] = \text{Res} \left[\frac{z}{z-\alpha} z^{k-1} \right] = z^k \Big|_{z=\alpha} = \alpha^k \quad (4-89)$$

El método del desarrollo parcial para determinar la transformada z inversa es parecido al método utilizado para determinar la transformada inversa de Laplace de una función racional. En vez de $F(s)$ en el caso de la transformada de Laplace, aquí se desarrolla $F(z)/z$ en fracciones parciales. La razón de este hecho es que suele ser más conveniente encontrar la transformada z inversa de una fracción con una z libre en su denominador. El siguiente ejemplo ilustra el método.

Ejemplo 5. Considérese:

$$F(z) = \frac{z(-2z+1)}{z^2-3z+2} \quad (4-90)$$

El desarrollo en fracciones parciales de $F(z)/z$ da:

$$\frac{F(z)}{z} = \frac{1}{z-1} - \frac{3}{z-2} \quad (4-91)$$

Por tanto:

$$F(z) = \frac{z}{z-1} - 3 \frac{z}{z-2} \quad (4-92)$$

y utilizando la ecuación 4-89, la transformada z inversa de $F(z)$ es:

$$f(k) = 1 - 3(2^k) \quad (4-93)$$

4.6.3 Tabla de transformadas z

En la tabla 4-5 se presentan las transformadas z de algunas funciones y sucesiones comunes. Obsérvese que $f(k) = 0$ para $k < 0$.

Tabla 4-5. Algunas funciones y sucesiones comunes y sus transformadas z

No.	$f(k)$	$F(z)$
1	$\delta(k)^a$	1
2	1	$\frac{z}{z-1}$
3	α^k	$\frac{z}{z-\alpha}$
4	k	$\frac{z}{(z-1)^2}$
5	k^2	$\frac{z(z+1)}{(z-1)^3}$
6	k^3	$\frac{z(z^2+4z+1)}{(z-1)^4}$
7	k^n	$(-1)^n z^n \frac{d^n}{dz^n} \left(\frac{z}{z-1} \right)$
8	$\frac{1}{k}, \quad k > 0$	$\ln \frac{z}{z-1}$
9	$e^{-\alpha k}$	$\frac{z}{z-e^{-\alpha}}$
10	$ke^{-\alpha k}$	$\frac{ze^{-\alpha}}{(z-e^{-\alpha})^2}$
11	$k^{\alpha k}$	$\frac{\alpha z}{(z-\alpha)^2}$
12	$k^2 \alpha k$	$\frac{\alpha z(z+\alpha)}{(z-\alpha)^3}$
13	$\frac{\alpha^k}{k!}$	$e^{\alpha/z}$
14	$\sin \alpha k$	$\frac{z \sin \alpha}{z^2 - 2z \cos \alpha + 1}$
15	$\cos \alpha k$	$\frac{z(z - \cos \alpha)}{z^2 - 2z \cos \alpha + 1}$
16	$e^{-\alpha k} \sin \beta k$	$\frac{ze^{-\alpha} \sin \beta}{z^2 - 2ze^{-\alpha} \cos \beta + e^{-2\alpha}}$
17	$e^{-\alpha k} \cos \beta k$	$\frac{z(z - e^{-\alpha} \cos \beta)}{z^2 - 2ze^{-\alpha} \cos \beta + e^{-2\alpha}}$
18	$\frac{\alpha^k + (-\alpha^k)}{2\alpha^1}$	$\frac{1}{z^2 - \alpha^2}$
19	$\frac{\alpha^k - \beta^k}{\alpha - \beta}$	$\frac{z}{(z-\alpha)(z-\beta)}$

$^a \delta(k) = \begin{cases} 1 & k=0 \\ 0 & k \neq 0 \end{cases}$ es la delta de Kronecker.

4.6.4 Solución de ecuaciones en diferencias utilizando la transformada z

La transformada z puede utilizarse para resolver ecuaciones en diferencias lineales con coeficientes constantes. Determinando la transformada z de ambos miembros de tal ecuación en diferencias, es posible expresar la transformada z de la sucesión desconocida como una función racional de z . La transformada z inversa de esta función es entonces la solución buscada. Por tanto, la transformada z convierte la solución de una ecuación en diferencias lineal con coeficientes constantes en la de una ecuación algebraica. El siguiente ejemplo ilustra el método.

Ejemplo 6. Considérese la ecuación en diferencias:

$$f(k+2) + 3f(k+1) + 2f(k) = 0, k = 0, 1, 2, \dots \quad (4-94)$$

sujeta a:

$$3f(0) + f(1) = 0 \quad (4-95)$$

Determinando la transformada z de ambos miembros de la ecuación se obtiene:

$$[z^2 F(z) - z^2 f(0) - z f(1)] + 3[z F(z) - z f(0)] + 2 F(z) = 0 \quad (4-96)$$

o, utilizando la ecuación (4-95):

$$F(z) = \frac{z^2 f(0)}{z^2 + 3z + 2} \quad (4-97)$$

que puede expresarse como:

$$\frac{F(z)}{z} = f(0) \left[\frac{-1}{z+1} + \frac{2}{z+2} \right] \quad (4-98)$$

Por tanto, la solución es:

$$f(k) = z^{-1}[F(z)] = f(0)[-(-1)^k + 2(-2)^k] \quad (4-99)$$

4.7 FUNCIONES DE TRANSFERENCIA

Las transformadas de Laplace y z son herramientas fundamentales para el estudio de sistemas lineales invariantes en el tiempo. Uno de los métodos para describir tales sistemas en la forma de entrada-salida es utilizar la función de transferencia, que relaciona la transformada de Laplace o la transformada z de la salida con la transformada de Laplace o la transformada z de la entrada. Para un sistema lineal invariante en el tiempo con r entradas y m salidas, la función de transferencia es una matriz $m \times r$.

Para este análisis, sean $m = r = 1$. Entonces el sistema se denomina sistema *escalar*. Para el caso continuo, la función de transferencia de dicho sistema se define como el cociente de la transformada de Laplace de la salida y la transformada de Laplace de la entrada cuando

todas las condiciones iniciales son cero. Considérese un sistema escalar lineal invariante en el tiempo descrito por la ecuación diferencial con coeficientes constantes:

$$\begin{aligned} p_n y^{(n)}(t) + p_{n-1} y^{(n-1)}(t) + \dots + p_1 y^{(1)}(t) + p_0 y(t) \\ = q_m u^{(m)}(t) + q_{m-1} u^{(m-1)}(t) + \dots + q_1 u^{(1)}(t) + q_0 u(t) \end{aligned} \quad (4-100)$$

donde $u(t)$ indica entrada, $y(t)$ indica salida y el suíndice i indica la derivada de i -ésimo orden con respecto al tiempo. La función de transferencia para este sistema es:

$$H(s) \triangleq \frac{Y(s)}{U(s)} = \frac{q_m s^m + q_{m-1} s^{m-1} + \dots + q_1 s + q_0}{p_n s^n + p_{n-1} s^{n-1} + \dots + p_1 s + p_0} \quad (4-101)$$

Obsérvese que $H(s)$ en la ecuación 4-101 es una función racional de s , de valuación escalar. La ecuación 4-101 también puede expresarse como:

$$H(s) = \frac{q_m(s - z_1)(s - z_2) \dots (s - z_m)}{p_n(s - \lambda_1)(s - \lambda_2) \dots (s - \lambda_n)} \quad (4-102)$$

donde z_i ($i = 1, 2, \dots, m$) se denominan los *ceros* y λ_j ($j = 1, 2, \dots, n$) los *polos* del sistema. Es posible que haya factores comunes en el numerador y en el denominador de la ecuación 4-102. Si dichos factores se cancelan, entonces los factores restantes se denominan polos y ceros de la función de transferencia.

La ecuación 4-101 implica que:

$$Y(s) = H(s)U(s) \quad (4-103)$$

Si la entrada es un impulso unitario, es decir, si $u(t) = \delta(t)$, entonces $U(s) = 1$ e $Y(s) = H(s)$ esto es:

$$y(t) \triangleq h(t) = L^{-1}[H(s)] \quad (4-104)$$

donde $h(t)$ se conoce como la *respuesta de impulso* (o respuesta al impulso) del sistema. Por consiguiente, la función de transferencia de un sistema escalar lineal continuo invariante en el tiempo es la transformada de Laplace de su respuesta de impulso.

Para sistemas lineales invariantes en el tiempo de entradas y salidas múltiples (de varias variables), la transformada de Laplace del vector de salida, $Y(s)$, y la transformada de Laplace del vector de entrada, $U(s)$, están relacionadas por la función de transferencia de la manera siguiente:

$$Y(s) = H(s)U(s) \quad (4-105)$$

que es parecida a la ecuación 4-103. Obsérvese que aquí $H(s)$ es una matriz $m \times r$, donde r es el número de entradas y m es el número de salidas.

El desarrollo de una función de transferencia para sistemas de tiempo discreto es parecido al del caso de los sistemas continuos. La función de transferencia de

un sistema escalar lineal invariante en el tiempo discreto se define como el cociente de la transformada z de la salida y la transformada z de la entrada cuando todas las condiciones iniciales son cero. Considérese el sistema escalar lineal invariante en el tiempo descrito por la ecuación en diferencias con coeficientes constantes:

$$p_n y(k+n) + p_{n-1} y(k+n-1) + \dots + p_1 y(k+1) + p_0 y(k) = q_m u(k+m) + q_{m-1} u(k+m-1) + \dots + q_1 u(k+1) + q_0 u(k) \quad (4-106)$$

donde $u(k)$ e $y(k)$ indican la entrada y la salida, respectivamente, en el instante k . La función de transferencia de este sistema es:

$$H(z) \triangleq \frac{Y(z)}{U(z)} = \frac{q_m z^m + q_{m-1} z^{m-1} + \dots + q_1 z + q_0}{p_n z^n + p_{n-1} z^{n-1} + \dots + p_1 z + p_0} \quad (4-107)$$

que es una función racional de z , de valuación escalar. Los polos y ceros del sistema y de la función de transferencia se definen de manera parecida a los del caso continuo en el tiempo. La ecuación 4-107 implica:

$$Y(z) = H(z)U(z) \quad (4-108)$$

que es la ecuación definitoria de la función de transferencia de los sistemas lineales invariantes en el tiempo descrito para los casos escalar y de varias variables. Se ha demostrado⁵ que:

$$H(z) = Z[h(k-1)] \quad (4-109)$$

donde $h(k)$ es la respuesta de impulso del sistema escalar. La ecuación 4-109 también se cumple para el caso de los sistemas discretos de varias variables. En tal caso $h(k)$ es una matriz $m \times r$, donde r y m son el número de entradas y salidas, respectivamente. Entonces, la i -ésima columna de $h(k)$ es la respuesta de estado cero a un pulso aplicado en la i -ésima entrada cuando todas las demás entradas son cero.

BIBLIOGRAFÍA

REFERENCIAS BIBLIOGRÁFICAS

1. A. Beurling, "Sur les integrales de Fourier absolument convergentes et leur application a une transformation fonctionnelle", 9.º Congreso de los Matemáticos Escandinavos, Helsingfors, Finlandia.
2. S. Bochner y K. Chandrasekharan, *Fourier Transforms*, Princeton University Press, Princeton, NJ.

3. J. W. Cooley y J. W. Tukey, "An algorithm for machine calculation of complex Fourier Serie". *Mathematics of Computation* **19**(90):297-301.
4. W. R. LaPage, *Complex Variables and the Laplace Transform for Engineers*, McGraw-Hill, Nueva York.
5. E. I. Jury, *Theory and Application of the z-Transform Method*, Wiley, Nueva York.

LIBROS Y REVISTAS

- R. E. Bellman, E. Kalaba, y J. A. Lockett, *Numerical Inversion of the Laplace Transform*, American Elsevier, Nueva York.
- R. B. Blackman y J. W. Tukey, *The Measurement of Power Spectra*, Dover, Nueva York.
- R. N. Bracewell, *The Fourier Transform and its Applications*, 2.ª ed., McGraw-Hill, Nueva York.
- G. A. Campbell y R. M. Foster, *Fourier Integrals for Practical Applications*, D. Van Nostrand, N.Y.
- H. S. Carslaw, *Introduction to the Theory of Fourier's Series and Integrals*, Dover, Nueva York.
- G. Doetsch, *Guide to the Applications of the Laplace and z-Transforms*, Van Nostrand Reinhold, Nueva York.
- G. Doetsch, *Introduction to the Theory and Application of the Laplace Transformation*, Springer-Verlag, Nueva York.
- A. Erdelyi, *Tables of Integral Transforms*, Vol. 1, McGraw-Hill, Nueva York.
- H. Freeman y O. Lowenschuss, *I.R.E. Trans. Automatic Control (PGAC)*, pp. 28-30.
- H. A. Helm, *Bell System Technical Journal* **38**(1):177-196.
- E. I. Jury y C. A. Galtieri, *I.R.E. Trans. Circuit Theory* **CT-9**:371-374.
- G. V. Lago *Trans A.I.E.E.* **74**(part II):403-408.
- M. J. Lighthill, *An Introduction to Fourier Analysis and Generalized Functions*, Cambridge University Press, Cambridge, Gran Bretaña.
- H. J. Nussbaumer, *Fast Fourier Transforms and Convolution Algorithms*, Springer-Verlag, Nueva York.
- R. E. Paley y N. Wiener, "Fourier transforms in the complex domain", *American Mathematical Society, Colloquium Publications*, Vol. 19, Nueva York.
- A. Papoulis, *The Fourier Integral and Its Applications*, McGraw-Hill, Nueva York.
- R. Piessens, *J. Comput. Appl. Math.* **1**(2):115-128.
- R. Piessens y N. D. P. Dang, *J. Comput. Appl. Math.* **2**(3):225-230.
- A. Talbot, *J. Inst. Math. Appl.* **23**:97-120.

CAPÍTULO 5

Probabilidad

Joel Halpert

Grumman Aerospace Corp.
Bethpage, Nueva York

5.1 Definiciones y teoremas de probabilidad

5.2 Probabilidad condicional, independencia, teorema de Bayes

5.3 Variables aleatorias, distribuciones de probabilidad y densidades

5.4 Valor esperado, dispersión y momentos

5.5 Algunas densidades y distribuciones útiles

- 5.5.1 Uniforme
- 5.5.2 Binomial
- 5.5.3 Poisson
- 5.5.4 Normal
- 5.5.5 Gamma
- 5.5.6 Beta

5.5.7 Distribución t

5.5.8 Distribución χ^2

5.5.9 Distribución F

5.5.10 Laplace

5.5.11 Cauchy

5.5.12 Rayleigh

5.5.13 Maxwell

5.6 Distribuciones multivariadas

5.6.1 Distribución normal bivariada

5.6.2 Dos funciones de dos variables aleatorias

5.7 Teorema del límite central

5.8 Procesos de Markov

5.1 DEFINICIONES Y TEOREMAS DE PROBABILIDAD¹

La probabilidad es la parte de las matemáticas que estudia el resultado de experimentos en los que hay incertidumbre. Las leyes de probabilidad pueden desarrollarse a partir de conceptos relacionados con lo que se conoce como teoría de conjuntos. A continuación se presentan diez definiciones fundamentales que describen dicha teoría, y cuatro definiciones relativas a resultados experimentales. Por último, se presenta la definición básica de probabilidad junto con algunos teoremas de probabilidad.

- **Conjunto.** Colección bien definida de objetos.
- **Elemento.** Un elemento (o miembro) de un conjunto es un objeto que pertenece al conjunto.
- **Igualdad de conjuntos.** Dos conjuntos son iguales si tienen exactamente los mismos elementos.
- **Conjunto vacío.** El conjunto vacío (o nulo) es un conjunto que no contiene elementos, y suele denotarse por el símbolo \emptyset .

- **Conjunto universal.** Es el conjunto que contiene todos los elementos, y suele denotarse por U .
- **Subconjunto.** Si todo elemento de un conjunto A también es un elemento de un conjunto B , entonces se dice que A es un subconjunto de B .

Un conjunto universal depende de la naturaleza del experimento. Por ejemplo, el conjunto de números que representan el resultado de un lanzamiento de un solo dado es $\{1, 2, 3, 4, 5, 6\}$, que es el conjunto universal de dicho procedimiento. El conjunto $\{2\}$ es un subconjunto de este conjunto universal.

La relación entre conjuntos puede definirse y asociarse con las funciones lógicas Y, O y NO, así como representarse gráficamente con los denominados diagramas de Venn.

- **Intersección.** La intersección de dos conjuntos A y B es el conjunto de elementos comunes a A y a B . Esta es la función lógica Y. La intersección de A y B se representa como $A \cap B$ (fig. 5-1).

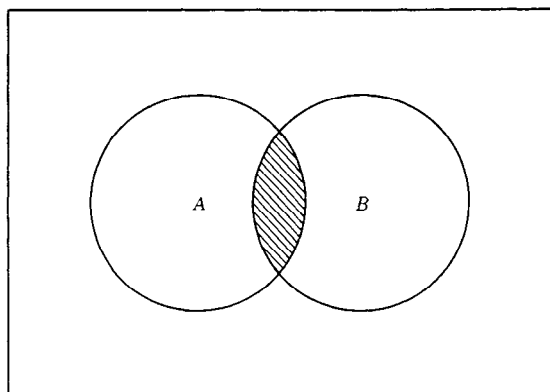


Fig. 5-1. Intersección de dos conjuntos, $A \cap B$, representada por la región sombreada.

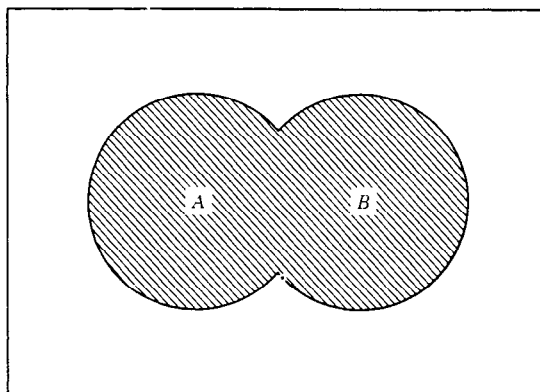


Fig. 5-2. Unión de dos conjuntos, $A \cup B$, representada por la región sombreada.

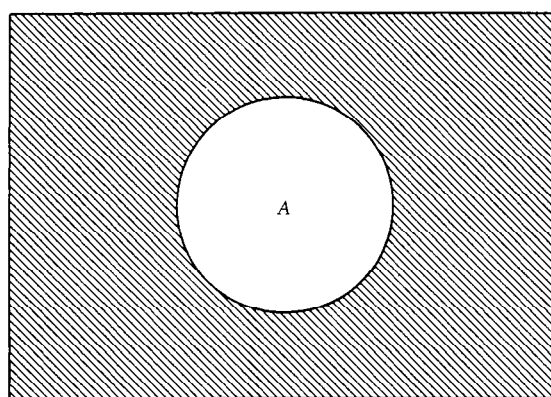


Fig. 5-3. Complemento de un conjunto, \bar{A} (o A'), representado por la región sombreada.

- **Conjuntos ajenos o disjuntos.** $A \cap B = \emptyset$. Los conjuntos A y B no tienen elementos en común.
- **Unión.** La unión de dos conjuntos A y B es el conjunto de elementos que pertenecen a A , a B o a ambos. Ésta es la función lógica o INCLUSIVA. La unión de A y B se representa como $A \cup B$ (fig. 5-2).
- **Complemento.** Si A es un subconjunto del conjunto universal U , entonces el complemento de A con respecto a U es el conjunto de todos los elementos de U que no están en A . Ésta es la función lógica NO. El complemento de A se representa como A' o \bar{A} (fig. 5-3).

Las siguientes definiciones están relacionadas con los resultados de experimentos:

- **Espacio muestral.** Un conjunto cuyos elementos representan todos los posibles resultados de un experimento se denomina espacio muestral (o espacio muestra), que suele representarse con el símbolo S .
- **Punto muestral.** Un elemento de un espacio muestral se denomina punto muestral (o punto muestra).
- **Evento.** Un evento es un subconjunto de un espacio muestral.
- **Eventos mutuamente excluyentes.** Dos eventos A y B son mutuamente excluyentes si $A \cap B = \emptyset$.

Por ejemplo, considérese el experimento de lanzar un solo dado con la intención de observar qué número presenta la cara superior del dado. Así, $S = \{1, 2, 3, 4, 5, 6\}$ es el espacio muestral del experimento, y el número $\{2\}$ es un punto muestral. El evento $\{1, 3, 5\}$ implica que el número de la cara superior sea impar.

Ahora es posible expresar la definición básica de probabilidad en términos de las definiciones anteriores de conjuntos y resultados experimentales:

- **Probabilidad.** La probabilidad de cualquier evento A , denotada por $\Pr(A)$, es la suma de los pesos de todos los puntos muestrales en A :

$$0 \leq \Pr(A) \leq 1$$

$$\Pr(\emptyset) = 0$$

$$\Pr(S) = 1$$

Es decir, la probabilidad de ocurrencia de cualquier evento debe estar en el intervalo cerrado de 0 a 1; la probabilidad de ocurrencia del evento vacío es igual a cero, y la probabilidad de ocurrencia del espacio muestral es igual a 1 (el evento "seguro").

A continuación se presentan algunas leyes básicas de probabilidad:

Teorema 1. La probabilidad de la unión de dos eventos A y B es:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \quad (5-1)$$

La ecuación 5-1 puede demostrarse utilizando un diagrama de Venn. Si A y B son eventos mutuamente excluyentes, es posible obtener un corolario del teorema 1. Entonces:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) \quad (5-2)$$

dado que, por la definición de eventos mutuamente excluyentes, $\Pr(A \cap B) = 0$. Además, si se considera un conjunto de n eventos mutuamente excluyentes $\{A_1, A_2, \dots, A_n\}$, entonces:

$$\Pr(A_1 \cup A_2 \cup \dots \cup A_n) = \Pr(A_1) + \Pr(A_2) + \dots + \Pr(A_n) \quad (5-3)$$

Teorema 2. La probabilidad del complemento de un evento A es:

$$\Pr(\bar{A}) = 1 - \Pr(A) \quad (5-4)$$

dado que A y \bar{A} son mutuamente excluyentes.

5.2 PROBABILIDAD CONDICIONAL, INDEPENDENCIA, TEOREMA DE BAYES¹

La probabilidad condicional es la probabilidad de ocurrencia de un evento cuando se sabe que ha ocurrido otro evento.

Considérense dos eventos, A y B . La probabilidad de ocurrencia de B , dado el hecho de que A ya ha ocurrido, se denota por $\Pr(B/A)$, que se lee "la probabilidad de ocurrencia de B , dado que A ya ha ocurrido" o, más simplemente, "la probabilidad de B dado A ". Esta probabilidad se define como sigue:

- **Probabilidad condicional.** La probabilidad condicional de B dado A es:

$$\Pr(B/A) = \frac{\Pr(A \cap B)}{\Pr(A)} \text{ si } \Pr(A) > 0 \quad (5-5)$$

La ecuación 5-5 puede reordenarse para calcular cualquier cantidad, dadas las otras dos cantidades.

- **Eventos independientes.** Los eventos A y B son independientes si y sólo si:

$$\Pr(A \cap B) = \Pr(A)\Pr(B) \quad (5-6)$$

Obsérvese que si A y B son eventos independientes, entonces la ecuación 5-5 se transforma en:

$$\Pr(B/A) = \Pr(B) \quad (5-7)$$

y que la probabilidad condicional pasa a ser probabilidad simple.

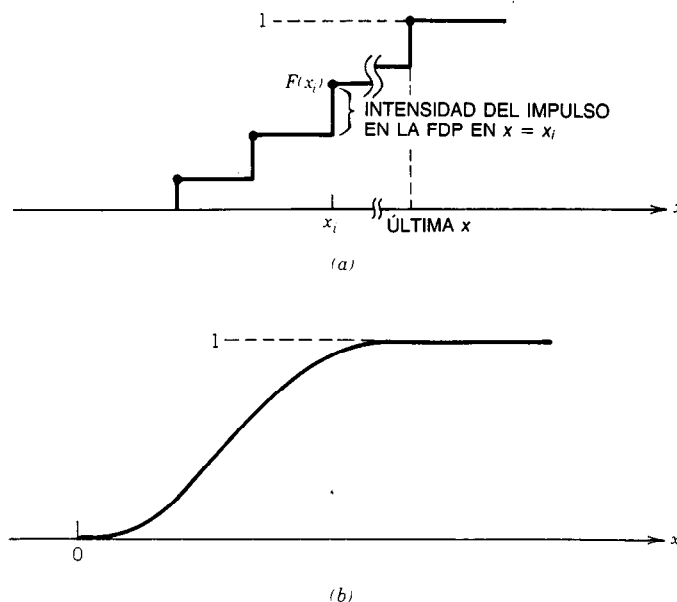


Fig. 5-4. Ejemplos de distribuciones de probabilidad: a) discreta; b) continua.

El teorema de Bayes es un método para calcular la probabilidad condicional de un evento que es un subconjunto de un conjunto de eventos, el cual constituye una partición del espacio muestral.

Teorema 3. Teorema de Bayes. Sea $\{B_1, B_2, \dots, B_n\}$ un conjunto de n eventos que forman una partición del espacio muestral S , con $\Pr(B_i) \neq 0$ para $i = 1, 2, \dots, n$. Sea A cualquier evento de S , con $\Pr(A) \neq 0$. Entonces, para $k = 1, 2, \dots, n$:

$$\Pr(B_k/A) = \frac{\Pr(B_k \cap A)}{\Pr(B_1 \cap A) + \Pr(B_2 \cap A) + \dots + \Pr(B_n \cap A)} \quad (5-8)$$

5.3 VARIABLES ALEATORIAS, DISTRIBUCIONES DE PROBABILIDAD Y DENSIDADES^{1,2}

En este artículo se asigna un número a cada resultado de un experimento, lo que a su vez define una función cuya variable independiente no es un número, sino un elemento del espacio muestral.

Todos los valores y funciones se consideran como reales.

- **Variable aleatoria.** Una variable aleatoria es una función, X , cuyo valor x es un número real determinado por cada elemento del espacio muestral.

Por ejemplo, considérese el evento $X < x$. Su probabilidad, $\Pr(X < x)$, es una función de x . Esta fun-

ción, que se denota por $F(x)$, se denomina función distribución de la variable aleatoria X .

- **Función distribución.** La función distribución de la variable aleatoria X es la función:

$$F(x) = \Pr(X < x) \quad (5-9)$$

definida para cualquier número entre $-\infty$ y $+\infty$. Es una función no decreciente de x ; es decir, $F(x)$ no decrece cuando x aumenta, y tiene $F(-\infty) = 0$ y $F(+\infty) = 1$.

Para definir la función densidad de probabilidad (fdp) suele diferenciarse entre variables aleatorias discretas y variables aleatorias continuas. Las variables aleatorias discretas tienen funciones distribución del tipo "escalonado" (en las que el intervalo entre los valores discretos suele llenarse con una recta horizontal, formando el "escalón").

Las variables aleatorias continuas suelen tener funciones distribución continuas, que se denominan "suaves" (fig. 5-4).

En cualquier caso, la fdp puede definirse de manera rigurosa de la siguiente forma:

- **Función densidad de probabilidad (fdp).** La función densidad $f(x)$, de la variable aleatoria X está dada por:

$$f(x) = \frac{dF(x)}{dx} \quad (5-10)$$

Para variables aleatorias discretas, la fdp consta de funciones impulso en los puntos discretos. Para las va-

riables aleatorias continuas, la función densidad está dada por la ecuación 5-10 en todos los puntos en los que existe la derivada.

Por otra parte, el área total bajo la curva $f(x)$ es igual a 1; es decir:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \quad (5-11)$$

y el área bajo la curva $f(x)$ desde $x = a$ hasta $x = b$ indica la probabilidad de que la variable aleatoria X esté entre a y b ; es decir:

$$\int_a^b f(x)dx = \Pr(a < X < b) \quad (5-12)$$

Para una variable aleatoria continua X , $\Pr(X = \text{cualquier valor entre } a \text{ y } b) = 0$, de ahí la desigualdad estricta en la ecuación 5-12.

5.4 VALOR ESPERADO, DISPERSIÓN Y MOMENTOS²

En esta sección se tratan varios parámetros importantes de las variables aleatorias. Se presentan definiciones para las variables aleatorias continuas y discretas.

- **Valor esperado o media.** El valor esperado de la variable aleatoria X , denotado por $E(X)$, también es el valor medio, denotado por μ , y está dado por:

$$\mu = E(X) = \begin{cases} \int_{-\infty}^{\infty} xf(x)dx & \text{(continua)} \\ \sum_n x_n \Pr(x = x_n) & \text{(discreta)} \end{cases} \quad (5-13)$$

Esto define el centro de gravedad de $f(x)$.

- **Dispersión o varianza.** La varianza de la variable aleatoria X , denotada por σ^2 , está dada por la siguiente expresión:

$$\sigma^2 = E[(X - \mu)^2] = \begin{cases} \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx & \text{(continua)} \\ \sum_n (x_n - \mu)^2 \Pr(x = x_n) & \text{(discreta)} \end{cases} \quad (5-14)$$

La varianza es igual al momento de inercia de las masas de probabilidad y proporciona alguna idea de su concentración en la vecindad de la media. Su raíz cuadrada positiva, σ , se conoce como desviación estándar o típica. Además, puede demostrarse que:

$$\sigma^2 = E(X^2) - \mu^2 \quad (5-15)$$

Es posible una especificación más detallada de los datos estadísticos de la variable aleatoria X si se conocen sus momentos.

- **Momentos.** El k -ésimo momento de la variable aleatoria X está dado por:

$$m_k = E(X^k) = \begin{cases} \int_{-\infty}^{\infty} x^k f(x)dx & \text{(continua)} \\ \sum_n x_n^k \Pr(x = x_n) & \text{(discreta)} \end{cases} \quad (5-16)$$

De la ecuación 5-16, $m_0 = 1$ (área bajo la curva) y $m_1 = \mu$. Además, dado que $m_2 = E(X^2)$, la varianza puede expresarse como:

$$\sigma^2 = m_2 - m_1^2 \quad (5-17)$$

5.5 ALGUNAS DENSIDADES Y DISTRIBUCIONES ÚTILES

En esta sección se presentan algunos de los tipos más comunes de densidades y distribuciones de probabilidad continuas y discretas.

5.5.1 Uniforme^{1, 2}

Una variable aleatoria discreta uniformemente distribuida, X , es aquella en la que la variable aleatoria asume todos sus valores con igual probabilidad. Considérese k de tales valores; es decir, x_1, x_2, \dots, x_k . Entonces, la función distribución puede expresarse como:

$$F(x) = \Pr(X = x) = \frac{1}{k} \quad x = x_1, x_2, \dots, x_k \quad (5-18)$$

Una variable aleatoria continua uniformemente distribuida tiene función densidad parecida a un pulso rectangular. Considérese el intervalo de valores de x_1 a x_2 . La fdp puede expresarse entonces como:

$$f(x) = \begin{cases} \frac{1}{x_2 - x_1} & \text{para } x_1 < x < x_2 \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5-19)$$

5.5.2 Binomial¹

Un experimento binomial consta de un conjunto de ensayos repetidos, cada uno de los cuales tiene dos posibles resultados, que suelen denominarse éxito o fracaso. Un ejemplo sencillo de dicho ensayo es arrojar al aire una moneda, cuyos dos resultados posibles son "cara" o "cruz". Cada resultado puede considerarse el "éxito", y el otro el "fracaso". En general, un experimento binomial tiene las siguientes propiedades:

1. El experimento consta de n ensayos repetidos.
2. Cada ensayo tiene un resultado que puede clasificarse como éxito o fracaso.
3. La probabilidad de un éxito, denotada por p , permanece constante de un ensayo a otro. La probabilidad de fracaso, denotada por q , es por tanto $1 - p$.
4. Los ensayos repetidos son independientes.

Para un experimento binomial es posible definir una variable aleatoria binomial X como el número de éxitos obtenido en n ensayos de un experimento binomial.

La función distribución de esta variable aleatoria se denomina distribución binomial.

Para x éxitos en n intentos independientes, la función distribución binomial está dada por la expresión siguiente:

$$F(x) = \Pr(X = x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} \quad (5-20)$$

$$x = 0, 1, 2, \dots, n$$

En la ecuación 5-20 el símbolo $!$ tiene el significado de factorial.

Para la distribución binomial, la media y la varianza están dadas, respectivamente, por las ecuaciones 5-21 y 5-22:

$$\mu = np \quad (5-21)$$

$$\sigma^2 = npq \quad (5-22)$$

5.5.3 Poisson^{1, 3}

Un experimento de Poisson es aquel que produce valores numéricos de una variable aleatoria, X , durante un intervalo dado de tiempo o en una región específica.

Una variable aleatoria de Poisson X es el número de éxitos en un experimento de Poisson. En general, un experimento de Poisson es aquel que tiene las siguientes propiedades:

1. Se conoce el número promedio de éxitos, μ , que ocurre en el intervalo de tiempo o región específica dados.
2. La probabilidad de ocurrencia de un solo evento durante un intervalo muy corto de tiempo o en una pequeña región es directamente proporcional a la longitud del intervalo de tiempo o al tamaño de la región, y no depende del número de éxitos que ocurren fuera de dicho intervalo de tiempo o región específica.
3. La probabilidad de que ocurra más de un éxito en dicho intervalo corto de tiempo o que se encuentre dentro de tal pequeña región es despreciable.

La función distribución de una variable aleatoria de Poisson se denomina distribución de Poisson, y está dada por:

$$F(x) = \Pr(X = x) = \frac{e^{-\mu} \mu^x}{x!} \quad x = 0, 1, 2, \dots \quad (5-23)$$

En la ecuación 5-23, μ es el número promedio de éxitos que ocurren en el intervalo de tiempo o la región dados. Tanto la media como la varianza tienen el valor μ .

5.5.4 Normal^{1, 2}

Una variable aleatoria continua está distribuida normalmente si su función densidad de probabilidad es una curva gaussiana, llamada así en honor del científico Carl Friedrich Gauss, quien dedujo su ecuación a partir de un estudio de errores en mediciones repetidas de la misma cantidad.

La curva gaussiana (o de Gauss) suele denominarse campana de Gauss, debido a su forma. Esta curva describe tantos conjuntos de datos que suceden en la naturaleza, la industria y la investigación que con toda seguridad es la distribución más importante en estadística. Algunos campos de la estadística en los que se utiliza bastante la distribución normal son la teoría del muestreo (en la cual las distribuciones χ^2 , F y t se desarrollan a partir de las estadísticas de muestras de poblaciones normalmente distribuidas), la teoría de la estimación y la prueba de hipótesis.

Los parámetros de la curva de Gauss son la media, μ , y la desviación estándar, σ (o la varianza σ^2). En la figura 5-5a se presenta la curva, que tiene las siguientes propiedades:

1. El máximo ocurre en la media, donde $x = \mu$.
2. La curva es simétrica respecto al eje vertical que pasa por la media.
3. La curva se acerca asintóticamente al eje horizontal a medida que se aleja de la media en cualquier dirección.
4. Dado que es una función densidad de probabilidad, el área bajo la curva es 1.

Una variable aleatoria X descrita por la curva de Gauss se denomina variable normal aleatoria. Una variable normal aleatoria con media μ y desviación estándar σ tiene una función densidad de probabilidad dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] \quad (5-24)$$

La función distribución correspondiente está dada por:

$$F(x) = \Pr(X < x) = \frac{1}{2} + \operatorname{erf} \left(\frac{x - \mu}{\sigma} \right) \quad (5-25)$$

En la ecuación 5-25, erf es la "función error", definida como:

$$\operatorname{erf} z = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-y^2/2} dy \quad (5-26)$$

Dos propiedades de la función error se expresan mediante las ecuaciones 5-27 y 5-28:

$$\operatorname{erf}(-z) = -\operatorname{erf} z \quad (5-27)$$

y

$$\operatorname{erf}(\infty) = 0.5 \quad (5-28)$$

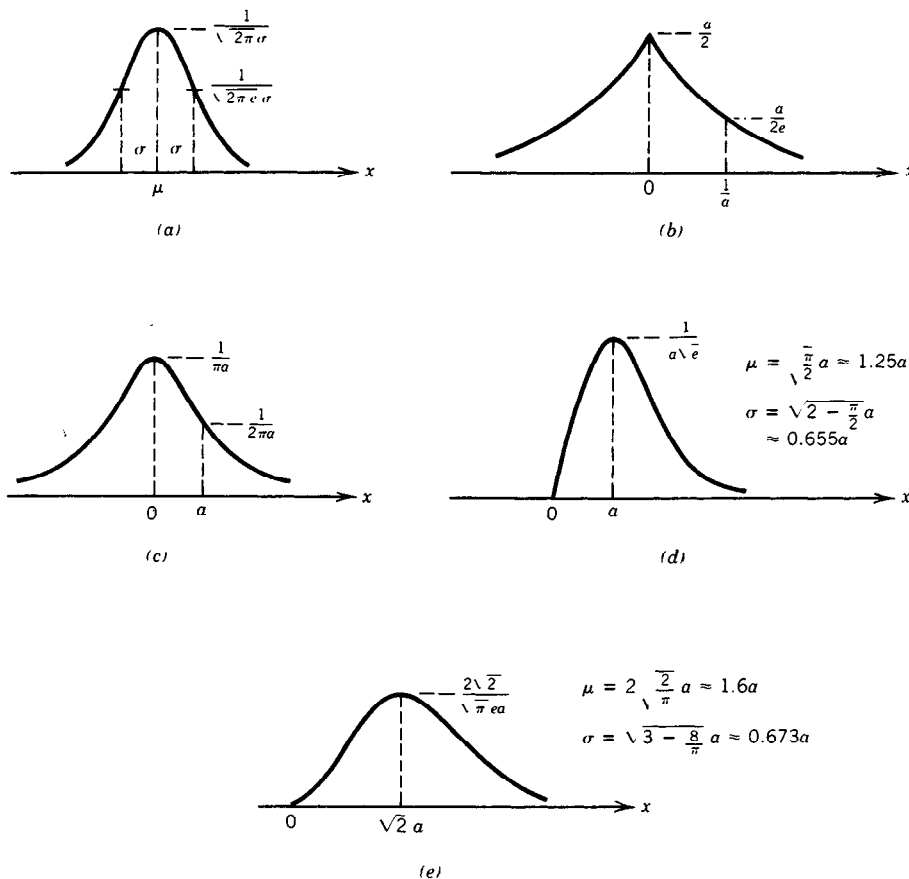


Fig. 5-5. Algunas densidades de probabilidad: a) normal; b) de Laplace; c) de Cauchy; d) de Rayleigh; e) de Maxwell.

Sería bastante difícil hacer tablas de valores para la función distribución de la ecuación 5-25, ya que esta función es paramétrica en μ y σ . En vez de ello, se define una nueva variable normal aleatoria Z con la expresión:

$$Z = \frac{X - \mu}{\sigma} \quad (5-29)$$

Es posible demostrar que la medida de Z es cero y que su varianza es 1. La variable aleatoria Z se denomina variable aleatoria normal estándar. Ahora es posible elaborar una sola tabla para la distribución normal estándar, y los valores correspondientes para las variables normales estándar con valores paramétricos diferentes a $\mu = 0$ y $\sigma = 1$ pueden determinarse a partir de la ecuación 5-29. En la tabla 5-1 se presenta el área bajo la curva normal estándar, correspondiente a $\Pr(Z < z)$, para valores de z de -3.49 a $+3.49$. La ecuación 5-29 se utiliza para convertir la variable aleatoria normal dada en estándar, antes de usar la tabla. La tabla es una evaluación de la ecuación 5-25, con $\mu = 0$ y $\sigma = 1$, y es posible generarla en la mayor parte de las computadoras de alta velocidad que contienen

una subrutina para la función error en su biblioteca de matemáticas. Así es como se generó la tabla 5-1, en la que cada valor tiene cuatro cifras decimales.

Para ilustrar la forma en que se utiliza la tabla, considérese una distribución normal con media igual a 50 y desviación estándar igual a 4. A fin de encontrar la probabilidad de que $x < 45$, se utiliza la ecuación 5-29 para convertir la variable aleatoria normal estándar, de donde se obtiene $z = -1.25$. Así, a fin de determinar $\Pr(X < 45)$, se utiliza la tabla para encontrar $\Pr(Z < -1.25)$, cuyo valor es 0.1056.

5.5.5 Gamma^{2, 3}

La densidad gamma está dada por:

$$f(x) = \frac{c^{b+1}}{\Gamma(b+1)} x^b e^{-cx} \quad x \geq 0, \quad b > 0, \quad c > 0 \quad (5-30)$$

donde:

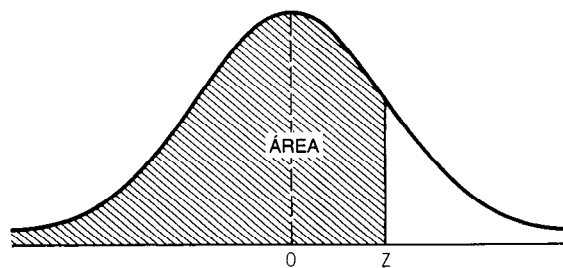
$$\Gamma(b+1) = \int_0^\infty y^b e^{-y} dy \quad (5-31)$$

Tabla 5-1. Áreas bajo la curva normal

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0050	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-.2	.4207	.4158	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5882	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879

Tabla 5-1. Áreas bajo la curva normal (continuación)

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998



Área bajo la curva normal en la región tramada

se denomina función gamma. Un ejemplo de variable aleatoria descrita por la densidad gamma es el tiempo de espera requerido para observar la ocurrencia $b + 1$ de un evento de algún tipo específico, cuando los eventos de este tipo ocurren aleatoriamente a una velocidad c por unidad de tiempo. Las funciones de densidad χ^2 (chi cuadrada) y F contienen funciones gamma en sus descripciones analíticas. Se dice que una variable aleatoria tiene distribución gamma si su densidad está dada por la ecuación 5-30. La media y la varianza están dadas por:

$$\mu = \frac{b + 1}{c} \quad (5-32)$$

$$\sigma^2 = \frac{b + 1}{c^2} \quad (5-33)$$

5.5.6 Beta^{2, 3}

La densidad beta está dada por:

$$f(x) = Ax^b(1-x)^c \quad 0 \leq x \leq 1, \quad b > 0, \quad c > 0 \quad (5-34)$$

donde:

$$A = \frac{\Gamma(b + c + 2)}{\Gamma(b + 1)\Gamma(c + 1)} \quad (5-35)$$

La función beta puede expresarse en términos de la función gamma; la relación es:

$$B(b, c) = \frac{\Gamma(b)\Gamma(c)}{\Gamma(b + c)} \quad (5-36)$$

Por tanto, la constante A de la ecuación 5-34 puede expresarse en términos de la función beta como:

$$A = \frac{1}{B(b + 1, c + 1)} \quad (5-37)$$

Se dice que una variable aleatoria tiene distribución beta si su densidad está dada por la ecuación 5-34. La media y la varianza están dadas por:

$$\mu = \frac{b + 1}{b + c + 2} \quad (5-38)$$

$$\sigma^2 = \frac{(b + 1)(c + 1)}{(b + c + 2)^2(b + c + 3)} \quad (5-39)$$

5.5.7 Distribución t ^{3, 4}

La distribución t está relacionada con las estadísticas de una muestra extraída de una población normal. Específicamente, supóngase que una muestra aleatoria de tamaño n se extrae de una población normal. Se sabe que la media de la población es μ , pero se desconoce la varianza. La media de la muestra es \bar{x} , y su varianza es s^2 ; es decir:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (5-40)$$

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (5-41)$$

Entonces es posible definir un valor t tal que:

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n - 1}} = \sqrt{n - 1} \left(\frac{\bar{x} - \mu}{s} \right) \quad (5-42)$$

t es un valor de una nueva variable aleatoria T que tiene la distribución t con $n - 1$ grados de libertad. Obsérvese que t se ha definido como una función de las estadísticas conocidas de la muestra y la población de la cual se extrajo. Esta distribución se utiliza en la teoría del muestreo como una prueba de la media.

Su función densidad, su media y su varianza están dadas por:

$$f(t) = \frac{1}{\sqrt{n - 1} B\left(\frac{1}{2}, \frac{n - 1}{2}\right)} \left(\frac{1}{1 + \frac{t^2}{n - 1}} \right)^{n/2} \quad -\infty < t < \infty \quad (5-43)$$

$$\mu_{f(t)} = 0 \quad (5-44)$$

$$\sigma_{f(t)}^2 = \begin{cases} \frac{n}{n - 2} & n > 2 \\ \text{indefinida} & n \leq 2 \end{cases} \quad (5-45)$$

En la ecuación 5-43, B es la función beta.

5.5.8 Distribución χ^2

Esta distribución se relaciona con la varianza de una muestra extraída de una población normal. Específicamente, supóngase que de una población normal se ha extraído una muestra aleatoria de tamaño n . La población tiene varianza conocida, σ^2 . Entonces es posible definir un valor χ^2 tal que:

$$\chi^2 = \sum_{i=1}^n x_i^2 \quad (5-46)$$

χ^2 es un valor de una nueva variable aleatoria que tiene la distribución chi cuadrada con n grados de libertad. Esta distribución se utiliza en la teoría del muestreo como una prueba de la varianza. Su función densidad, su media y su varianza están dadas por:

$$f(y) = \frac{1}{2^{n/2} \sigma^n \Gamma\left(\frac{n}{2}\right)} y^{n/2 - 1} e^{-y/2\sigma^2} \quad y > 0 \quad (5-47)$$

$$\mu_{f(y)} = n\sigma^2 \quad (5-48)$$

$$\sigma_{f(y)}^2 = 2n\sigma^4 \quad (5-49)$$

5.5.9 Distribución F^4 , 5

Esencialmente, la función F se halla vinculada con la relación entre dos distribuciones independientes chi cuadrada, cada una dividida entre sus grados de libertad.

Sea y_1 el valor de la variable chi cuadrada Y_1 con n grados de libertad, e y_2 el valor de la variable aleatoria chi cuadrada Y_2 con m grados de libertad, donde Y_1 e Y_2 son independientes. Entonces es posible definir un valor f tal que:

$$f = \frac{y_1/n}{y_2/m} \quad (5-50)$$

f es un valor de una nueva variable aleatoria F que tiene la distribución F con n y m grados de libertad. Existen dos parámetros: los dos diferentes grados de libertad. Si bien es posible definir la estadística f como el cociente de la ecuación 5-50 o su inverso, es importante que la variable chi cuadrada del numerador se especifique de tal modo que la curva f pueda definirse de manera apropiada.

Esto implica que el orden en el que se planteen los grados de libertad es importante; los valores de la distribución F no son los mismos cuando se intercambian los grados de libertad; es decir, en general, $F(n, m) \neq F(m, n)$.

Esencialmente, la distribución F se aplica en el análisis de varianza. Su función de densidad, su media y su varianza están dadas por:

$$g(f) = \frac{\Gamma\left(\frac{n+m}{2}\right) \left(\frac{n}{m}\right)^{n/2}}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{m}{2}\right)} \frac{f^{n/2-1}}{\left(1 + \frac{nf}{m}\right)^{(n+m)/2}} \quad (5-51)$$

$0 < f < \infty$

$$\mu_{g(f)} = \frac{m}{m-2} \quad m > 2 \quad (5-52)$$

$$\sigma_{g(f)}^2 = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)} \quad m > 4 \quad (5-53)$$

$F(n, m)$ no tiene media para $m \leq 2$, y no tiene varianza para $m \leq 4$.

5.5.10 Laplace²

La densidad de Laplace está dada por (fig. 5-5b):

$$f(x) = \frac{a}{2} e^{-a|x|} \quad (5-54)$$

5.5.11 Cauchy²

La densidad de Cauchy está dada por (fig. 5-5c):

$$f(x) = \frac{a/\pi}{a^2 + x^2} \quad (5-55)$$

5.5.12 Rayleigh²

La densidad de Rayleigh está dada por (fig. 5-5d):

$$f(x) = \frac{x}{a^2} e^{-x^2/2a^2} \quad x \geq 0 \quad (5-56)$$

5.5.13 Maxwell²

La densidad de Maxwell está dada por (fig. 5-5e):

$$f(x) = \frac{\sqrt{2}}{(a^3)\sqrt{\pi}} x^2 e^{-x^2/2a^2} \quad x \geq 0 \quad (5-57)$$

5.6 DISTRIBUCIONES MULTIVARIADAS^{2, 5}

Las distribuciones multivariadas están relacionadas con distribuciones de más de una variable aleatoria. Esto implica que la dimensión del espacio muestral es mayor que 1; el espacio muestral unidimensional se describe por medio de una variable aleatoria. En consecuencia, dos diferentes variables aleatorias pueden describirse mediante un espacio muestral bidimensional.

Considérense dos variables aleatorias, X e Y , cuyas medias son μ_1 y μ_2 , respectivamente. Es posible definir un número denominado covarianza de X e Y de forma parecida a la ecuación 5-14:

$$\sigma_{12} = E[(X - \mu_1)(Y - \mu_2)] = E(XY) - \mu_1\mu_2 \quad (5-58)$$

Si X e Y son variables aleatorias independientes, entonces su covarianza es cero. Además, si las variables aleatorias X e Y tienen desviaciones estándar positivas σ_1 y σ_2 , respectivamente, entonces es posible definir un número denominado coeficiente de correlación entre X e Y tal que:

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1\sigma_2} \quad (5-59)$$

De nuevo, si X e Y son variables aleatorias independientes, entonces su coeficiente de correlación es cero.

Una función densidad de probabilidad (fdp) conjunta y una función distribución de probabilidad conjunta pueden definirse para dos variables aleatorias, como se hizo para una variable aleatoria en las ecuaciones 5-9 y 5-10.

Por tanto, la función distribución conjunta para variables aleatorias (continuas) está dada por:

$$F(x, y) = \Pr(X \leq x, Y \leq y) \quad (5-60)$$

y la función densidad conjunta está dada por:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} \quad (5-61)$$

Es posible definir una función denominada distribución marginal como la distribución unidimensional

de cualquier variable aleatoria sola. Puede afirmarse algo parecido para la densidad marginal. Por ejemplo, la distribución marginal para X está dada por:

$$F_X(x) = F(x, \infty) \quad (5-62)$$

y la densidad marginal para X está dada por:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (5-63)$$

5.6.1 Distribución normal bivariada⁵

Considérense las variables aleatorias X e Y con medias μ_1 y μ_2 y desviaciones estándar σ_1 y σ_2 , respectivamente. Se dice que estas variables aleatorias tienen distribución normal bivariada si tienen una fdp conjunta dada por:

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right]\right\} \quad (5-64)$$

En la ecuación 5-64, ρ es el coeficiente de correlación. Los momentos conjuntos de X e Y , con covarianza σ_{12} , están dados por*:

$$m_{kn} = E(X^k Y^n) = kn \int_0^{\sigma_{12}} E(X^{k-1} Y^{n-1}) d\sigma_{12} + E(X^k)E(Y^n) \quad (5-65)$$

5.6.2 Dos funciones de dos variables aleatorias²

Dadas dos funciones $g(x, y)$ y $h(x, y)$ de las variables reales x e y , y dos variables aleatorias X e Y , se forman otras dos variables aleatorias, $V = g(X, Y)$ y $W = h(X, Y)$. Estas variables aleatorias tienen una distribución conjunta $F(v, w)$ y una densidad conjunta $f(v, w)$. Si v y w son dos números reales, R denota la región del plano xy tal que $g(x, y) \leq v$ y $h(x, y) \leq w$. Entonces:

$$F(v, w) = \iint_R f(x, y) dx dy \quad (5-66)$$

es la función distribución conjunta para V y W . Para determinar la función densidad conjunta, es necesario resolver las ecuaciones $v = g(x, y)$ y $w = h(x, y)$ para x e y en términos de v y w .

Supóngase que existen n parejas ordenadas (x, y) que son solución de estas ecuaciones. Dado que el jacobiano de la transformación de variables está dado por la expresión:

$$J(x, y) = \begin{vmatrix} \frac{\partial g(x, y)}{\partial x} & \frac{\partial g(x, y)}{\partial y} \\ \frac{\partial h(x, y)}{\partial x} & \frac{\partial h(x, y)}{\partial y} \end{vmatrix} \quad (5-67)$$

es posible demostrar que:

$$f(v, w) = \frac{f(x_1, y_1)}{J(x_1, y_1)} + \dots + \frac{f(x_n, y_n)}{J(x_n, y_n)} + \dots \quad (5-68)$$

Si para ciertos valores de (v, w) no existen soluciones reales, entonces $f(v, w) = 0$.

5.7 TEOREMA DEL LÍMITE CENTRAL^{2, 5}

Un planteamiento muy general del teorema del límite central establece que, dado un conjunto de variables aleatorias independientes de cualquier tipo de distribución, la distribución de su suma tiende asintóticamente a una distribución normal a medida que aumenta el número de variables.

Este teorema es útil no sólo en el estudio de la probabilidad, sino también en otras áreas; por ejemplo, la salida de un sistema lineal tiende a distribuirse normalmente, aun cuando las entradas aleatorias no lo estén. También puede demostrarse que representa la convolución de un gran número de funciones positivas, por lo que puede aplicarse a sistemas lineales en cascada.

Considérense n variables aleatorias continuas independientes X_1, X_2, \dots, X_n tales que la media y la varianza de cada una sean μ_i y σ_i^2 , $i = 1, 2, \dots, n$ (en general n valores diferentes para cada parámetro). Se define una nueva variable aleatoria igual a la suma, es decir $X = X_1 + X_2 + \dots + X_n$. Entonces la media de X es la suma de las medias de los sumandos, y la varianza de X es la suma de las varianzas de los sumandos. Es posible demostrar que la función densidad de X está dada por:

$$f(x) = f_1(x) * f_2(x) * \dots * f_n(x) \quad (5-69)$$

donde el signo $*$ denota la convolución. Por tanto, la densidad de X es igual a la convolución de las densidades de X_1, X_2, \dots, X_n . El teorema del límite central establece que bajo ciertas condiciones generales, $f(x)$ tiende a una curva normal a medida que crece n .

Un caso especial del teorema del límite central se relaciona con una muestra de tamaño n extraída de cualquier población de la que se conocen la media μ y la varianza σ^2 , ambas finitas. Por tanto, considérese que X_1, X_2, \dots, X_n denotan los objetos de una muestra aleatoria de tamaño n extraída de cualquier distribución que tenga media μ y varianza σ^2 . Se define una variable aleatoria denominada media de la muestra, $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$. Ahora es posible definir una nueva variable aleatoria, Y , como función de \bar{X} y los parámetros de la población μ y σ de forma:

$$Y = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \quad (5-70)$$

*Ecuación 7-121 en la referencia 2.

El teorema del límite central establece que Y tiene distribución normal limitante con media 0 y varianza 1. Por tanto, la distribución de Y es aproximadamente normal con media 0 y varianza 1. Entonces, es posible utilizar esta distribución aproximadamente normal para calcular las probabilidades aproximadas relacionadas con \bar{X} , para encontrar un intervalo aproximado de confianza para μ , y para probar ciertas hipótesis estadísticas, sin conocer aún la función distribución de probabilidad de \bar{X} en cada caso.

5.8 PROCESOS DE MARKOV²

En esta sección se definen los procesos de Markov y se presenta un ejemplo de primer orden.

Antes de definir qué es un proceso de Markov, es necesario definir lo que se entiende por proceso estocástico. Supóngase un experimento especificado por sus resultados ξ , que forman el espacio muestral S , por ciertos subconjuntos de S denominados eventos, y por las probabilidades de estos eventos. A cada evento se asigna, mediante una regla, una función del tiempo $x(t, \xi)$, real o compleja, con lo que se crea una familia de funciones; una para cada ξ . A esta familia se le denomina proceso estocástico, y puede considerarse una función de las dos variables, t y ξ . El dominio de ξ es el conjunto S , y el de t es el conjunto de números reales, que en este caso se supone es el eje real positivo del tiempo. La notación $x(t)$ se utiliza para representar un proceso estocástico, omitiendo la dependencia respecto de ξ . Sin embargo, debe observarse que $x(t)$ representa cuatro cosas diferentes:

1. Una familia de funciones del tiempo (con t y ξ variables).
2. Una simple función del tiempo (con t variable y ξ fija).
3. Una variable aleatoria (con t fija y ξ variable).
4. Un solo número (con t y ξ fijos).

Un proceso estocástico $x(t)$ se denomina proceso de Markov si para toda n y $t_1 < t_2 \dots < t_n$ se tiene:

$$\Pr[x(t_n) \leq x_n / x(t_{n-1}), \dots, x(t_1)] = \Pr[x(t_n) \leq x_n / x(t_{n-1})] \quad (5-71)$$

La ecuación 5-71 es equivalente a:

$$\Pr[x(t_n) \leq x_n / x(t)] \text{ para toda } t \leq t_{n-1} = \Pr[x(t_n) \leq x_n / x(t_{n-1})] \quad (5-72)$$

Algunas de las propiedades de los procesos de Markov son las que se especifican a continuación:

1. Un proceso $x(t)$ es de Markov si el pasado no afecta las estadísticas del futuro a condición de que se conozca el presente.
2. Un proceso de Markov se denomina proceso de Markov hacia atrás si para $t_1 < t_2$, entonces:

$$\Pr[x(t_1) \leq x_1 / x(t)] \text{ para toda } t \geq t_2 = \Pr[x(t_1) \leq x_1 / x(t_2)] \quad (5-73)$$

3. Un proceso $x(t)$ es de Markov si, con $t_1 < t_2$, $x(t_2) - x(t_1)$ es independiente de $x(t)$ para toda $t \leq t_1$.
4. Si $x(t)$ es un proceso normal de Markov con media 0, entonces su autorrelación satisface:

$$R(t_3, t_2)R(t_3, t_1) = R(t_3, t_1)R(t_2, t_2) \quad (5-74)$$

para toda $t_3 > t_2 > t_1$.

5. Un proceso de Markov $x(t)$ puede asociarse con una ecuación diferencial de primer orden:

$$\frac{dx(t)}{dt} - h[x(t), t] = g(t) \quad (5-75)$$

donde $g(t)$ es tal que las variables aleatorias $g(t_1), \dots, g(t_n)$ son independientes para cualquier n, t_1, \dots, t_n . Con:

$$g(t) = \frac{dw(t)}{dt} \quad (5-76)$$

el proceso $w(t)$ tiene incrementos independientes. Un proceso $w(t)$ tiene incrementos independientes si $w(t_i) - w(t_{i+1}), i = \dots, -1, 0, 1, \dots$ es una sucesión de variables aleatorias no correlacionadas donde los intervalos (t_i, t_{i+1}) no se superponen pero son arbitrarios.

Un ejemplo de proceso de Markov de primer orden es la velocidad $v(t)$ de una partícula libre en el movimiento browniano. $v(t)$ satisface:

$$\frac{dv(t)}{dt} + \beta v(t) = \frac{dw(t)}{dt} \quad (5-77)$$

donde $w(t)$ es un proceso con incrementos independientes tal que $E[dw(t)] = 0$ y $E[(dw(t))^2] = \alpha dt$. La solución de la ecuación 5-77 es un proceso de Markov $v(t)$ con estadísticas $\mu(v, t) = -\beta v$ y $\sigma^2(v, t) = \alpha$.

6. Si $x(t)$ es continuo, es posible definir una densidad condicional como:

$$p(x, t; x_0, t_0) = f[x / x(t_0) = x_0] \quad (5-78)$$

suponiendo $x(t_0) = x_0$. Entonces:

$$p(x, t; x_0, t_0) = \int_{-\infty}^{\infty} p(x, t; x_1, t_1) p(x_1, t_1; x_0, t_0) dx_1 \quad (5-79)$$

La ecuación 5-79 es la ecuación de Chapman-Kolmogorov para los procesos de Markov. Si p depende sólo de $t - t_0$, el proceso se denomina homogéneo. Por otra parte, la densidad condicional $p(x, t; x_0, t_0), t > t_0$, también satisface las siguientes ecuaciones:

$$\frac{\partial p}{\partial t_0} + \mu(x_0, t_0) \frac{\partial p}{\partial x_0} + \frac{1}{2} \sigma^2(x_0, t_0) \frac{\partial^2 p}{\partial x_0^2} = 0 \quad (5-80)$$

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial x} [\mu(x, t)p] - \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma^2(x, t)p] = 0 \quad (5-81)$$

Éstas son las ecuaciones de difusión hacia delante y hacia atrás deducidas por Kolmogorov. La ecuación 5-81 se conoce como ecuación de Fokker-Planck. Para el problema de la velocidad de una partícula libre en movimiento browniano, utilizando los valores resultantes de las estadísticas de $v(t)$ en la ecuación 5-81 la densidad condicional p satisface la ecuación:

$$\frac{\partial p}{\partial t} = \beta \frac{\partial(vp)}{\partial v} + \frac{\alpha}{2} \frac{\partial^2 p}{\partial v^2} \quad (5-82)$$

BIBLIOGRAFÍA

REFERENCIAS BIBLIOGRÁFICAS

1. R. E. Walpole, *Introduction to Statistics*, 2.^a ed., Macmillan, Nueva York.
2. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Nueva York.
3. M. Fisz, *Probability and Mathematical Statistics*, S. Chand, Nueva Delhi.
4. F. L. Wolf, *Elements of Probability and Statistics*, McGraw-Hill, Nueva York.
5. R. V. Hogg y A. T. Craig, *Introduction to Mathematical Statistics*, Macmillan, Nueva York.

LIBROS Y REVISTAS

- N. A. Hastings y J. B. Peacock, *Statistical Distributions*, Halsted Press, Wiley, Nueva York.

CAPÍTULO 6

Álgebra booleana

Frank T. Boesch
Charles Suffel

Stevens Institute of Technology
Hoboken, New Jersey

6.1 Postulados de Huntington para un álgebra booleana

6.2 Propiedades de un álgebra booleana

6.3 Funciones de conmutación

6.4 Simplificación de funciones booleanas

6.1 POSTULADOS DE HUNTINGTON PARA UN ÁLGEBRA BOOLEANA¹

El álgebra booleana o de conmutación es la base matemática para el análisis y diseño de los circuitos de conmutación. Existen diversas formas de definir el álgebra booleana. Aquí se presentan los postulados de Huntington, que constituyen un conjunto de axiomas para un álgebra booleana no sólo consistentes (es decir, no contradictorios por sí mismos o entre sí), sino también independientes (es decir, no se deduce ningún axioma del conjunto a partir de los demás).

El conjunto B , asociado con dos operaciones binarias, denotadas por $+$ y \cdot , se denomina un álgebra booleana si posee las propiedades que se exponen a continuación:

- 1. Cerradura.** Para todo par $a, b \in B$, $a + b \in B$ y $a \cdot b \in B$.
- 2. Identidades.** Existen elementos diferentes 0 y 1 tales que, para toda $b \in B$:

$$b + 0 = b \quad y \quad b \cdot 1 = b$$

- 3. Conmutatividad.** Para todo $a, b \in B$:

$$a + b = b + a \quad y \quad a \cdot b = b \cdot a$$

- 4. Distributividad.** Para toda triplete $a, b, c \in B$:

$$a + b \cdot c = (a + b) \cdot (a + c) \quad y$$

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

- 5. Complementos.** Para todo $b \in B$ existe un elemento \bar{b} , denominado complemento de b , tal que:

$$b + \bar{b} = 1 \quad y \quad b \cdot \bar{b} = 0$$

Con fines de ilustración se presentan dos ejemplos, de los cuales el primero es el álgebra booleana más sencilla que existe.

Ejemplo 1. Álgebra booleana binaria. Sea $B = \{0, 1\}$ y defínase $+$ y \cdot como se muestra en la figura 6-1. Se comprueba con facilidad que B , junto con las operaciones $+$ y \cdot , satisfacen los postulados de Huntington¹.

Ejemplo 2. Conjunto potencia¹. Sea U un conjunto y sea $P(U)$ la colección de todos los subconjuntos de U . Se define:

$$X + Y = X \cup Y \quad y \quad X \cdot Y = X \cap Y$$

para cada par $X, Y \in P(U)$. Aquí, 0 es el conjunto vacío y 1 es el conjunto universal U .

Las bien conocidas propiedades de unión e intersección de conjuntos se generalizan a álgebras booleanas arbitrarias. A continuación se presentan algunas de las más importantes.

+	0	1		0	1
0	0	1		0	0
1	1	1		1	1

Fig. 6-1. Un álgebra booleana de dos elementos.

6.2 PROPIEDADES DE UN ÁLGEBRA BOOLEANA

Sea B un álgebra booleana. Entonces, las siguientes propiedades son válidas¹.

- 1. Unicidad de las identidades.** Los elementos 0 y 1 son únicos.
- 2. Cotas universales.** Las identidades 0 y 1 son cotas universales.
Para todo $b \in B$:

$$b \cdot 0 = 0 \quad y \quad b + 1 = 1$$

- 3. Idempotencia.** Para toda $b \in B$:

$$b + b = b \quad y \quad b \cdot b = b$$

- 4. Unicidad del complemento.** Para toda $b \in B$ existe un solo complemento \bar{b} . Además:

$$\bar{\bar{0}} = 1 \quad y \quad \bar{\bar{1}} = 0$$

- 5. Involución.** Para toda $b \in B$:

$$\overline{(\bar{b})} = b$$

- 6. Asociatividad.** Si $a, b, c \in B$, entonces:

$$(a + b) + c = a + (b + c) \quad y \quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

- 7. Leyes de De Morgan.** Si $a, b \in B$, se cumple:

$$\overline{(a + b)} = \bar{a} \cdot \bar{b} \quad y \quad \overline{(a \cdot b)} = \bar{a} + \bar{b}$$

- 8. Absorción.** Si a y $b \in B$, entonces:

$$a + a \cdot b = a \quad y \quad a \cdot (a + b) = a$$

6.3 FUNCIONES DE CONMUTACIÓN

Como preparación para aplicar el álgebra booleana a los circuitos de conmutación, se presentan las definiciones de variable y polinomio de conmutación. Sea $\{0, 1\}$, denotado por Q .

- Variable de conmutación.** Es una variable x sobre el conjunto Q .
- Polinomio de conmutación.** Un polinomio de conmutación $p(x_1, x_2, \dots, x_n)$ en las n variables independientes de conmutación x_1, x_2, \dots, x_n es cualquier expresión obtenida recursivamente

aplicando de manera repetida las operaciones $+$, \cdot y la complementación a 0 y 1 de las variables x_1, x_2, \dots, x_n .

Sustituyendo los valores de las variables de conmutación x_1, x_2, \dots, x_n en $p(x_1, x_2, \dots, x_n)$ se obtiene una función de n variables en Q valuada en Q . Esta función se conoce como función de n variables asociada con $p(x_1, x_2, \dots, x_n)$.

Al aplicarla a la teoría de la conmutación, una variable de conmutación representa una "compuerta"; es decir, un dispositivo bicondicional que permite o impide el flujo de una señal a través de él. Cuando el dispositivo asume el valor 0 la compuerta está abierta, y cuando el valor es 1 se encuentra cerrada. Una interconexión física entre compuertas se denomina circuito de conmutación. El circuito asume un estado abierto o cerrado dependiendo de los estados individuales de la compuerta. Así, asociada con el circuito de conmutación existe una función de n variables en Q , valuada en Q , que se llama función de transmisión del circuito.

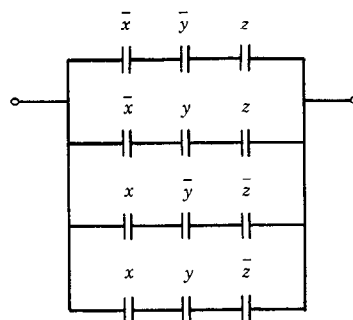
Si x e y representan compuertas, entonces $x + y$ representa la combinación en paralelo de las compuertas, y $x \cdot y$ representa la combinación en serie. La variable \bar{x} representa la compuerta dependiente cuyo estado es el opuesto al de la compuerta representada por x . Por conveniencia, 0 representa una compuerta abierta de un circuito, y 1 representa una compuerta cerrada de un circuito. Así, la función de n variables asociada con un polinomio de transmisión es la función de transmisión de un circuito de conmutación en serie-paralelo. La importancia del siguiente teorema es que cualquier función de transmisión puede sintetizarse en una configuración en serie-paralelo.

Teorema de la síntesis. Cualquier función de n variables (es decir, toda función de transmisión) es la función de n variables asociada con algún polinomio de conmutación.

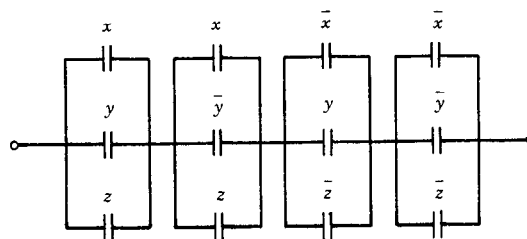
Ahora se presenta un algoritmo para determinar un polinomio de conmutación para una función de transmisión dada. Sea una función de n variables especificada como una tabla de $(n + 1) \times 2^n$, donde las n primeras entradas de una columna contienen una elección específica de valores para las variables de conmutación x_1, x_2, \dots, x_n y la última entrada proporciona el valor correspondiente de la función. Cualquiera de los dos métodos siguientes produce un polinomio de conmutación cuya función asociada de n variables es igual a la función de transmisión dada².

Método 1. Para cada columna con 1 en la última entrada, formar un producto de n variables en el que aparece x_i si la i -ésima entrada es 1, y \bar{x}_i si la i -ésima entrada es 0. Igualar $p(x_1, x_2, \dots, x_n)$ a la suma de estos productos.

Método 2. Para cada columna con 0 en la última entrada, formar una suma en la que aparece x_i si la i -ésima entrada es 0, y \bar{x}_i si la i -ésima entrada es 1. Igualar $p'(x_1, x_2, \dots, x_n)$ al producto de estas sumas.



$p(x, y, z)$



$p'(x, y, z)$

Fig. 6-2. Circuitos de conmutación para la tabla 6-1.

Ejemplo 3. Una función de transmisión. Considérese la función de transmisión definida en la tabla 6-1. Usando el método 1 del algoritmo recién expuesto se obtiene el polinomio de conmutación:

$$p(x, y, z) = \bar{x} \cdot \bar{y} \cdot z + \bar{x} \cdot y \cdot z + x \cdot \bar{y} \cdot \bar{z} + x \cdot y \cdot \bar{z}$$

Si se aplica el método 2 del algoritmo se obtiene:

$$\begin{aligned} p'(x, y, z) &= \\ &= (x + y + z) \cdot (x + \bar{y} + z) \cdot (\bar{x} + y + \bar{z}) \cdot (\bar{x} + \bar{y} + \bar{z}) \end{aligned}$$

En la figura 6-2 se observan representaciones de estos polinomios. Obsérvese que p tiene asociada la misma función de n variables que el polinomio más sencillo:

$$\hat{p}(x, y, z) = \bar{x} \cdot z + x \cdot \bar{z}$$

debido a que para todas las elecciones de valores particulares de x, y y z :

$$p(x, y, z) = \bar{x} \cdot \bar{y} \cdot z + \bar{x} \cdot y \cdot z + x \cdot \bar{y} \cdot \bar{z} + x \cdot y \cdot \bar{z}$$

Tabla 6-1. Una función de transmisión de tres variables

x	0	0	0	0	1	1	1	1
y	0	0	1	1	0	0	1	1
z	0	1	0	1	0	1	0	1
f	0	1	0	1	1	0	1	0

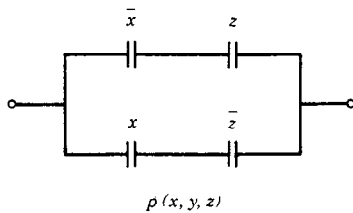


Fig. 6-3. Otra representación de la tabla 6-1.

pero:

$$\bar{x} \cdot \bar{y} \cdot z + \bar{x} \cdot y \cdot z = \bar{x} \cdot (\bar{y} + y) \cdot z = \bar{x} \cdot 1 \cdot z = \bar{x} \cdot z$$

y

$$x \cdot \bar{y} \cdot \bar{z} + x \cdot y \cdot \bar{z} = x \cdot (\bar{y} + y) \cdot \bar{z} = x \cdot \bar{z}$$

Por tanto, la representación física más económica es la que se muestra en la figura 6-3.

6.4 SIMPLIFICACIÓN DE FUNCIONES BOOLEANAS

En el resto de este artículo se describe un método útil para simplificar los polinomios de conmutación similar al método 1 del algoritmo. Sin embargo, primero debe describirse qué se entiende por expresión "más sencilla". Obsérvese que en el ejemplo precedente se requirieron menos bancos paralelos para representar $\bar{p}(x, y, z)$ que para representar $p(x, y, z)$. Al diseñar un circuito de conmutación físico, la medida real de la economía depende del tipo de implementación; por ejemplo, diodos, relevadores o compuertas lógicas. En vez de desarrollar una teoría independiente para cada uno de estos tipos de implementación, se tiene una medida común de minimalidad. Sean $p(x_1, x_2, \dots, x_n)$ y $q(x_1, x_2, \dots, x_n)$ dos polinomios de conmutación asociados con la misma función de n variables, ambos en forma de sumas de productos, en las que puede aparecer cada término de producto (ya sea x_i o \bar{x}_i), pero no ambos. Entonces, p es más sencillo que q si: 1) p tiene menos términos que q , o 2) p y q tienen el mismo número de términos, pero p tiene menos literales que q (donde x y \bar{x} se consideran literales diferentes). Se dice que p es mínimo si no existe un polinomio de conmutación q más sencillo cuya forma sea la de una suma de productos y tenga asociada la misma función de n variables.

Ahora se describe un procedimiento geométrico sistemático para aplicar la regla:

$$x \cdot y + \bar{x} \cdot y = y$$

con la cual se simplifica un polinomio de conmutación en un polinomio mínimo. Este procedimiento es más efectivo para polinomios de conmutación en tres o cuatro variables de conmutación, aunque existen extensiones al caso de cinco o seis variables^{1, 2}.

Para el caso de tres variables, x , y y z , se construye una figura geométrica conocida como mapa de Karnaugh del polinomio de conmutación dado $p(x, y, z)$, como sigue:

1. Primero se forma una matriz 2×4 , en la que los renglones y las columnas se identifican como se muestra en la figura 6-4. Los dos renglones representan los valores posibles de la variable x , y las columnas indican las posibles combinaciones de las variables y y z , respectivamente.
2. Para cada posible combinación de valores de x , y y z , colocar el valor de $p(x, y, z)$ en el lugar apropiado de la matriz.

En la figura 6-5 se presenta el mapa de Karnaugh para el ejemplo 3. Obsérvese que cualesquiera dos celdas adyacentes, como las posiciones 001 y 011 de xyz , representan productos que difieren exactamente en una literal.

Más precisamente, la literal de excepción aparece en forma no complementada en un producto, y en forma complementada en el otro. Por tanto, los términos en una expresión de suma de productos que corresponden a celdas adyacentes que contengan números 1 pueden combinarse por medio de la ley:

$$A \cdot a + A \cdot \bar{a} = A$$

donde A representa alguna combinación de literales y a es cualquier otra literal. También debe observarse qué celdas, como las posiciones 100 y 110, deben considerarse celdas adyacentes. Por consiguiente, podría ser de utilidad considerar que el mapa está enrollado en forma de cilindro. Dos celdas adyacentes que contengan números 1 forman una "celda dos"; un cuadrado con cuatro números 1 en el cilindro forma una "celda cuatro". En este caso los cuatro términos correspondientes en la expresión de la suma de productos pueden combinarse de modo que se eliminen dos literales.

		yz			
		00	01	11	10
x	0				
	1				

Fig. 6-4. Mapa general de Karnaugh para tres variables.

		yz			
		00	01	11	10
x	0	0	1	1	
	1	1	0	0	1

Fig. 6-5. Mapa de Karnaugh correspondiente al ejemplo 6-3.

En este momento ya debe ser claro que la forma de identificar las columnas y renglones permite una visualización geométrica por medio de "k celdas" de todas las simplificaciones posibles que pueden obtenerse utilizando la ley:

$$A \cdot a + A \cdot \bar{a} = A$$

Resta describir un procedimiento algorítmico aplicado al mapa, el cual produce una expresión mínima de suma de productos con la misma función de tres variables que la expresión original utilizada para formar el mapa.

Con este propósito, considérese el mapa de Karnaugh para una expresión de tres variables que no contenga todos los números 1:

1. Determinar todas las celdas uno que no están contenidas en las celdas dos.
2. Encontrar todas las celdas dos que no están contenidas en las celdas cuatro.
3. Encontrar todas las celdas cuatro.
4. Escribir una expresión mínima que conste de los productos correspondientes a las simplificaciones relacionadas con las celdas uno, dos y cuatro ya determinadas^{1, 2}.

Al aplicar este algoritmo al ejemplo se obtiene un par de dos celdas, como se muestra en la figura 6-5, a saber, 001-011 y 100-110; la expresión mínima correspondiente es:

$$\hat{p}(x, y, z) = \bar{x} \cdot z + x \cdot \bar{z}$$

El procedimiento se extiende de forma natural a las expresiones de cuatro variables, como se describe enseguida. Primero se forma la matriz 4×4 , cuyas columnas y renglones se identifican como se observa en la figura 6-6. Como en el caso de tres variables, se escriben en la matriz los valores de la función de cuatro variables asociada con una expresión de suma de productos $p(w, x, y, z)$. Para obtener la visualización geométrica de adyacencia, se supone que la forma del cuadro es un toroide; es decir, que están unidas las partes superior e inferior de la matriz, así como los lados izquierdo y derecho. Empleando las adyacencias resultantes se definen las celdas uno, dos y cuatro como en el caso de tres variables. Además, se introduce la notación de "celda ocho", es decir, un rectángulo de 2×4 en el toroide, que contiene ocho números 1. De

yz \ wx				
	00	01	11	10
00				
01				
11				
10				

Fig. 6-6. Mapa general de Karnaugh para cuatro variables.

yz \ wx				
	00	01	11	10
00		1	1	1
01		1	1	1
11	1	1	1	
10		1	1	

Fig. 6-7. Ejemplo de mapa de Karnaugh para cuatro variables.

nuevo, estas "K celdas" conducen a las simplificaciones que se obtienen aplicando la ley $A \cdot a + A \cdot \bar{a} = A$. El algoritmo para escribir una expresión mínima utilizando el cuadro es la extensión inmediata del caso planteado para tres variables.

Se concluye con la ilustración del algoritmo para el mapa de Karnaugh que se muestra en la figura 6-7. Al aplicar el algoritmo se obtienen las celdas dos, cuatro y ocho que se observan en dicha figura. En la celda dos ahora aparecen \bar{x} y x , por lo que se absorben y eliminan para obtener $\bar{w} \cdot y \cdot z$. En la celda cuatro se eliminan x y z , con lo que se obtiene $w \cdot \bar{y}$. En la celda ocho se eliminan todas las variables, salvo x . Por tanto, la expresión mínima correspondiente es:

$$\hat{p}(w, x, y, z) = \bar{w} \cdot y \cdot z + w \cdot \bar{y} + x$$

REFERENCIAS BIBLIOGRÁFICAS

1. F. J. Hill y G. R. Peterson, *Introduction to Switching Theory and Logical Design*, 3.^a ed., Wiley, Nueva York.
2. Z. Kohavi, *Switching and Finite Automata Theory*, 2.^a ed., McGraw-Hill, Nueva York.

SEGUNDA PARTE

Propiedades de los materiales

CAPÍTULO 7

Propiedades de los materiales

Krishna Seshan

IBM, E. Fishkill
NY, 12533

C. F. Bieber

Hughes Aircraft Company
Tucson, Arizona

7.1 Unidades y clasificación de los materiales

7.1.1 Clasificación de los materiales

7.2 Propiedades resistivas

7.2.1 Origen de la resistividad: regla de Matthiessen

7.2.2 Relación de resistividad residual (RRR)

7.2.3 Resistividad volumétrica

7.2.4 Resistividad de masa

7.2.5 Resistividad laminar

7.2.6 Resistividad superficial

7.2.7 Regla de Nordheim

7.2.8 Resistividad de películas delgadas

7.3 Capacidad de transporte de corriente de los conductores

7.4 Propiedades de conductividad

7.4.1 Conductividad volumétrica

7.4.2 Conductividad de masa

7.5 Propiedades relacionadas con los materiales

7.5.1 Mecánicas

7.5.2 Térmicas

7.5.3 Electroemigración

7.5.4 Difusión

7.6 Materiales de encapsulamiento microelectrónico

7.1 UNIDADES Y CLASIFICACIÓN DE LOS MATERIALES

Aunque las unidades del SI son bastante recomendadas¹, con frecuencia se utilizan varias unidades útiles y prácticas que no pertenecen a ese sistema. En la parte inferior se presenta una lista con dichas unidades y su equivalencia en las unidades SI.

En este capítulo se abordan las propiedades resistivas, de conductividad y de encapsulamiento de los materiales que suelen emplearse en las industrias eléctrica, microelectrónica y de encapsulamiento. Se inicia con las propiedades de los materiales resistivos y después se tratan las de los materiales conductores; con frecuencia en este capítulo también se hacen algunas consideraciones acerca de propiedades mecánicas y térmicas que pueden influir en dichos aspectos. Por consiguiente, también se analizan propiedades como la resistencia mecánica a la tensión y el coeficiente de dilatación.

Las aplicaciones en dispositivos de los materiales conductores, resistivos y de encapsulamiento están limitadas por procesos como flujo plástico, electroemigración y difusión. En este capítulo se incluyen, asimismo, datos acerca de estas propiedades.

También se incluye un análisis de los escasos datos de un campo en desarrollo: el encapsulamiento microelectrónico. Aquí existen serias restricciones en cuanto a la compatibilidad de los materiales. Los materiales que suelen ser aislantes, como las resinas epóxicas o el vidrio, deben hacerse conductores térmicos.

El capítulo termina con una revisión de algunos materiales comunes que se utilizan en aplicaciones de conducción, resistivas y de encapsulamiento, y se presentan algunas de sus limitaciones.

7.1.1 Clasificación de los materiales

Los materiales pueden clasificarse como conductores metálicos² (de 1 a 150 $\mu\Omega\text{-cm}$), semiconductores (de 150 a 1 000 $\mu\Omega\text{-cm}$) o aislantes (más de 10³ $\mu\Omega\text{-cm}$).

Los aisladores y conductores amorfos, así como los conductores vítreos, se producen por medio de técnicas especiales de rápida solidificación y pueden adquirir importancia en un futuro próximo en la tecnología de los semiconductores. Los conductores iónicos, como las sales fundidas, son importantes a temperaturas mayores de 500 °C.

La teoría de banda de los sólidos proporciona una base teórica para esta clasificación, y también explica las diferencias en cuanto a la dependencia respecto a la temperatura de la resistividad presentada por dichos materiales.

Una clase de materiales importantes para el encapsulamiento microelectrónico es la de los materiales conductores y no conductores de unión, entre los cuales se encuentran las resinas epóxicas conductoras y no conductoras, las polímidas, las soldaduras blandas, las soldaduras vítreas y las soldaduras duras. Estos materiales se producen deliberadamente con valores específicos de coeficientes de dilatación y soldabilidad, a fin de hacerlos compatibles con las superficies distintas que unen.

En el artículo 7.6 se presentan las propiedades de dichos materiales.

7.2 PROPIEDADES RESISTIVAS

7.2.1 Origen de la resistividad: regla de Matthiessen

Cualquier perturbación en la regularidad absoluta de la estructura cristalina, o sea, cualquier perturbación de los átomos respecto a su colocación ideal en la estructura reticular, dispersará los electrones, debido a lo cual provocará resistencia eléctrica^{3,4}.

A 0 K, la contribución del fonón (vibración de la malla cristalina) a la resistividad decae a cero. La resistividad volumétrica restante se debe a impurezas, fronteras granulares, dislocaciones, huecos y otras imperfecciones.

En los metales puros y aleaciones diluidas, la resistividad total (ρ) es la suma de dos términos: la componente térmica, ρ_T , debida a vibraciones de la malla, y la resistividad residual, ρ_r , provocada por impurezas e imperfecciones estructurales⁴. Esta última no depende de la temperatura. La resistividad total está dada por:

$$\rho = \rho_T + \rho_r \quad (7-1)$$

La ecuación 7-1 es la regla de Matthiessen².

7.2.2 Relación de resistividad residual (RRR)

La **relación de resistividad residual** es un número utilizado para denotar la pureza de una muestra. A medida que la muestra es más pura, la RRR es más alta. La RRR se define como:

$$RRR = \frac{\rho_{298\text{ K}}}{\rho_{4.2\text{ K}}} \quad (7-2)$$

Por medio de técnicas de refinación por zonas se han obtenido relaciones hasta de 100 000.

	Unidad práctica	Equivalencia en el SI	Tabla
Coeficiente de expansión o dilatación	ppm/°C	10 ⁻⁶ /K	7.1
Peso/1 000 pies	lb/1 000 pies	kg/m	7.7
Resistividad eléctrica	$\mu\Omega\text{-cm}$	S-m	7.7
Conductividad eléctrica	mhos/m	S/m	
Coeficiente de resistencia	ppm/°C	S-m/K	

7.2.3 Resistividad volumétrica

La **resistividad volumétrica** o de volumen, ρ , se define como el cociente del campo eléctrico entre la densidad de corriente resultante:

$$\rho = \frac{E}{J} \quad (7-3)$$

donde ρ = resistividad volumétrica, $\Omega \cdot m$
 E = intensidad del campo eléctrico, V/m
 J = densidad de corriente, A/m²

La resistencia, R , está relacionada con la resistividad por:

$$R = \rho \cdot \frac{L}{A} \quad (7-4)$$

donde A = sección transversal del conductor, m²
 L = longitud del conductor, m
 R = resistencia, Ω

En la tabla 7-1 se presentan las resistividades eléctricas de metales puros, y en la 7-2, las de algunas aleaciones y compuestos seleccionados.

La resistividad volumétrica también puede expresarse como el recíproco de la conductividad volumétrica, σ :

$$\rho = \frac{1}{\sigma} \quad (7-5)$$

en cuyo caso sus unidades son $10^{-8} \Omega \cdot m$ o microohms-cm, $\mu\Omega \cdot cm$. (Obsérvese que ambas unidades son idénticas).

7.2.4 Resistividad de masa

La resistividad de masa, δ , se define como:

$$\delta = \frac{R \cdot m}{L^2} \quad (\Omega kg/m) \quad (7-6)$$

donde R = resistencia, Ω
 m = masa del conductor, kg
 L = longitud del conductor, m

7.2.5 Resistividad laminar

La resistividad laminar, ρ/t , se utiliza cuando el grosor del conductor (t) es mucho más pequeño que su anchura o su longitud. Se define como:

$$\frac{\rho}{t} = R \cdot \frac{W}{L} \frac{\text{ohm}}{\text{cuadrado}} \quad (7-7)$$

donde t = grosor del conductor, m
 L = longitud del conductor, m
 W = ancho del conductor, m

La resistencia de una lámina es el producto de su resistividad laminar por el número de cuadrados L/W . L/W también se conoce como "relación altura/longitud".

7.2.6 Resistividad superficial

A altas frecuencias, la corriente eléctrica se confina a una región cercana a la superficie. A una profundidad pelicular δ , la resistividad superficial, R_s , cae a $1/e$ de su valor en la superficie. La **resistividad superficial**, R_s , es la resistividad laminar a la CD de un conductor cuyo grosor es:

$$R_s = \frac{\rho}{\delta} = \frac{1}{\sigma \delta} \quad (7-8)$$

donde ρ = resistividad eléctrica
 δ = profundidad pelicular, m
 σ = conductividad, S/m

En la tabla 7-3 se presentan las características a alta frecuencia de algunos metales.

7.2.7 Regla de Nordheim

La adición de una impureza eleva la resistividad residual conforme a la relación:

$$\rho_r(X) = A \times (1 - X) \quad (7-9)$$

donde ρ_r = resistividad de la aleación diluida
 X = concentración de la impureza
 A = constante que depende del metal huésped y la impureza; aumenta con la valencia, tamaño atómico y otras diferencias entre el huésped y la impureza.

La ecuación 7-9 se conoce como regla de Nordheim. En la tabla 7-1 se muestra la relación entre la resistividad y el diagrama de fases de un sistema de dos componentes. En la tabla 7-4 se presenta el efecto que tiene sobre la resistividad la adición de aleaciones a metales seleccionados.

7.2.8 Resistividad de películas delgadas

La resistividad de las películas delgadas se incrementa desde el valor volumétrico a medida que el grosor tiende a la trayectoria libre media de los electrones de conducción, que suele ser de 10 a 100 nm a temperatura ambiente^{3,17,18}. El mecanismo dominante es la dispersión de electrones en la superficie.

La resistividad de las películas delgadas tiende a ser mayor que la volumétrica, incluso cuando las películas son más gruesas que la trayectoria libre media de los electrones. Fronteras granulares, óxidos, huecos y otras imperfecciones estructurales suelen ser causas de aumento en la resistividad. Algunos de los procesos que hacen que la resistividad de las películas tienda al valor volumétrico son el recocido, la deposición a mayores temperaturas y el crecimiento epitaxial.

Ocurre otro incremento en la resistividad cuando las películas son en extremo delgadas y físicamente discontinuas. La conducción en tales películas es por emisión termoiónica o por efecto túnel, y la resistencia es no óhmica.

El coeficiente térmico de resistividad (TCR) de películas muy delgadas tiende a ser negativo. Esto ocurre

Tabla 7-1. Propiedades físicas de los metales puros¹⁵

Metal	Punto de fusión (°C)	Densidad a 20 °C (g/cm ³)	Conductividad térmica, 0 a 100 °C (W/m-K)	Calor específico medio, 0 a 100 °C (J/kg-K)	Resistividad a 20 °C (μΩ-cm)	Coefficiente térmico de resistividad, 0 a 100 °C (ppm/K)	Coefficiente de dilatación 0 a 100 °C (10 ⁻⁶ /K)
Aluminio	660.1	2.70	238	917	2.67	4 500	23.5
Antimonio	630.5	6.68	23.8	209	40.1	5 100	8-11
Bario	729	3.5	-	285	60(0 °C)	-	18
Berilio	1 287	1.848	194	2 052	3.3	9 000	12
Bismuto	271	9.80	9	124.8	117	4 600	13.4
Cadmio	320.9	8.64	103	233.2	7.3	4 300	31
Calcio	839	1.54	125	624	3.7	4 570	22
Cerio	798	6.75	11.9	188	85.4	8 700	8
Cesio	28.5	1.87	36.1 ^a	234	20	4 800	97
Cinc	419.5	7.14	119.5	394	5.96	4 200	31
Cobalto	1 492	8.9	96	427	6.34	6 600	12.5
Cobre	1 083.4	8.96	397	386.0	1.694	4 300	17.0
Cromo	1 860	7.1	91.3	461	13.2	2 140	6.5
Estaño	231.9	7.3	73.2	226	12.6	4 600	23.5
Estroncio	770	2.6	-	737	23(0 °C)	-	100
Galio	29.7	5.91	41.0 ^a	377	^b	-	18.3
Germanio	937	5.32	56.4	310	~89×10 ³	-	5.75
Hafnio	2 227	13.1	22.9	147	32.2	4 400	6.0
Hierro	1 536	7.87	78.2	456	10.1	6 500	12.1
Indio	156.4	7.3	80.0	243	8.8	5 200	24.8
Iridio	2 454	22.4	146.5	130.6	5.1	4 500	6.8
Litio	181	0.534	76.1	3 517	9.29	4 350	56
Magnesio	649	1.74	155.5	1 038	4.2	4 250	26.0
Manganeso	1 244	7.4	7.8	486	160(α) ^c	-	23
Mercurio	-38.87	13.546	8.65	138	95.9	1 000	61
Molibdeno	2 615	10.2	137	251	5.7	4 350	5.1
Níquel	1 455	8.9	88.5	452	6.9	6 800	13.3
Niobio	2 467	8.6	54.1	268	16.0	2 600	7.2
Oro	1 063	19.3	315.5	130	2.20	4 000	14.1
Osmio	3 030	22.5	87.5	130	8.8	4 100	4.57
Paladio	1 552	12.0	75.5	247	10.8	4 200	11.0
Plata	960.8	10.5	425	234	1.63	4 100	19.1

Tabla 7-1. Propiedades físicas de los metales puros¹⁵ (continuación)

Metal	Punto de fusión (°C)	Densidad a 20 °C (g/cm ³)	Conductividad térmica, 0 a 100 °C (W/m-K)	Calor específico medio, 0 a 100 °C (J/kg-K)	Resistividad a 20 °C (μΩ-cm)	Coefficiente térmico de resistividad, 0 a 100 °C (ppm/K)	Coefficiente de dilatación 0 a 100 °C (10 ⁻⁶ /K)
Platino	1 769	21.45	71.5	134.4	10.58	3 920	9.0
Plomo	327.4	11.68	34.9	129.8	20.6	4 200	29.0
Potasio	63.2	0.86	104 ^a	754	6.8	5 700	83
Radio	700	5	-	-	-	-	-
Renio	3 180	21.0	47.6	138	18.7	4 500	6.6
Rodio	1 966	12.4	149	243	4.7	4 400	8.5
Rubidio	38.8	1.53	58.3 ^a	356	12.1	4 800	9.0
Rutenio	2 310	12.2	116.3	234	7.7	4 100	9.6
Silicio	1 412	2.34	138.5	729	10 ³ -10 ⁶	-	7.6
Sodio	97.8	0.97	128	1 227	4.7	5 500	71
Talio	304	11.85	45.5	130	16.6	5 200	30
Tantalio	2 980	16.6	57.55	142	13.5	3 500	6.5
Telurio	450	6.24	3.8	134	1.6×10 ⁵ (0 °C)	-	<i>d</i>
Titanio	1 667	4.5	21.6	528	54	3 800	8.9
Torio	1 755	11.5	49.2	100	14	4 000	11.2
Tungsteno	3 400	19.3	174	138	5.4	4 800	4.5
Uranio	1 132	19.05(α) ^c 18.89(β) ^e	28	117	27	3 400	<i>c</i>
Vanadio	1 902	6.1	31.6	498	19.6	3 900	8.3
Zirconio	1 852	6.49	22.6	289	44	4 400	5.9

Fuente. Datos de la referencia bibliográfica 5.

^a Sólido.^b 17.4, paralelo al eje *a*; 8.1, paralelo al eje *b*; 54.3, paralelo al eje *c*.

23 paralelo al eje *a*.
^c α-Uranio - 3.5 paralelo al eje *b*.
 17 paralelo al eje *c*.

^d 1.7, paralelo al eje *c*; 27.5 perpendicular al eje *c*.

4.6 paralelo al eje *c*.
^e β Uranio 23 perpendicular al eje *c*.
 de 20 a 720°C.

Tabla 7-2. Propiedades seleccionadas de algunas aleaciones y compuestos^a

Aleación o composición	Intervalo de fusión (°C)	Densidad (g/cm ³)	Conductividad térmica (W/m-K)	Dilatación (10 ⁻⁶ /K)	Resistividad (μΩ-cm)
Acero al carbono					
AISI 1010	1 510-1 525	7.84	60.2	12.2	16.8
1042	1 480-1 510	7.84	50.7	11.3	17.1
1078	1 450-1 500	7.84	48.2	11.1	18.0
Acero inoxidable					
Tipo 304	1 400-1 450	7.9	16.2	17.2	72
316	1 370-1 400	7.9	16.2	16.0	74
405	1 480-1 530	7.7	27.0	10.8	60
420	1 450-1 500	7.7	24.9	10.3	55
Aluminio					
Tipo 1100	643-657	2.71	222	23.6	2.92
2024	502-638	2.77	190	22.9	3.4
5056	568-638	2.64	120	24.1	5.9
6061	582-652	2.70	180	23.6	3.7
7075-T6	477-635	2.80	130	23.4	5.2
Cobre					
Tipo C10100	1 083	8.94	391	17.0	1.72
OFHC					
C11000	1 065-1 083	8.89	388	16.8	1.72
ETP					
C14500	1 051-1 075	8.94	355	17.1	1.86
0,5% Te					
C17200	865-980	8.25	105-130	16.7	5.7-11.5
2,0% Be					
C21000	1 050-1 065	8.86	234	18.1	3.1
Dorado					
C23000	990-1 025	8.75	159	18.7	4.7
Latón rojo					
C27000	905-932	8.47	116	20.3	6.4
Latón amarillo					
C51000	880-1 050	8.86	69	17.8	11.5
Bronce fosforoso					
Otros					
Hierro gris o de segunda fundición	1175-1300	7.2	50	10.8	70
ASTM-A48					
Hierro colado maleable ASTM A-47	1175-1300	7.36	51	13.5	40

Tabla 7-2. Propiedades seleccionadas de algunas aleaciones y compuestos^a (continuación)

Aleación o composición	Intervalo de fusión (°C)	Densidad (g/cm ³)	Conductividad térmica (W/m-K)	Dilatación (10 ⁻⁶ /K)	Resistividad (μΩ-cm)
Otros (cont.)					
Acero martensítico al alto níquel sin carbono ("maraging")	1 430-1 450	8.03	21	10.1	—
Inconel 601	1 370-1 400	8.1	11	11.5	—
80Ni-20Cr	1 395-1 425	8.4	13	13.2	108
Hastelloy C	1 265-1 345	8.9	11	11.3	130
92.5Ti-5Al-2.5Sn	1 550-1 650	4.48	7.8	9.5	157
Boruros					
CrB ₂	1 900	5.2	—	6.0-6.7	21-56
HfB ₂	3 250	11.2	—	5.3-7.5	10-12
MoB ₂	2 100	7.7	—	7.7	30
NbB ₂	2 980	6.9	17	5.9-8.4	12-65
TaB ₂	3 090	12.6	13	5.2-5.9	14-68
TiB ₂	2 899	4.4	26	4.8-6.6	10-20
VB ₂	2 425	5.0	—	3.8-8.0	16-38
W ₂ B ₅	2 200	13.1	—	—	21-56
ZrB ₂	3 040	6.2	23	5.5-6.2	7-10
Carburos					
B ₄ C	2 450	2.5	121	2.6-3.1	10 ⁶
SiC	2 700	2.6-3.2	95	3.9-4.3	100-200
TaC	3 885	14.4	22	8.3	20-30
TiC	3 127	5.0-7.2	17-20	6.6-7.4	105
ZrC	3 521	6.6-6.8	21	6.6	63
Nitruros					
BN	2 730	1.9-2.2	182	7.5	10 ¹⁹
Si ₃ N ₄	1 900	3.2	14-33	2.4	10 ¹⁹
TaN	3 343	14.4	10	5.5	135
TiN	3 220	5.4	17	6.3	22
VN	2 320	6.1	—	—	86
Óxidos					
Al ₂ O ₃ , alúmina	2 040	3.9	17	7.7	10 ²⁰
3Al ₂ O ₃ -2SiO ₂ , mullita	1 871	2.9-3.3	2.5	4.8-5.4	10 ²¹
BeO, berilia	2 550	2.9	250	7.5	10 ²⁰
CaO, calcia	2 599	3.3-5.5	—	12.6	—
MgO, magnesia	2 799	3.5-3.6	47	12.8-14	—
2mgO-SiO ₂ , forsterita	1 888	2.8-3.2	3.3-4.1	8.5	10 ²⁰
SiO ₂ , sílice	1 670	1.9-2.6	1.4	0.5	—
ZrO ₂ , circonia	2 599	5.5-9.9	12-14	5.6	—

Fuente. Datos de las referencias bibliográficas 5 a 10.

^a Valores aproximados debido a los efectos del método de preparación.

Tabla 7-3. Conductores a altas frecuencias

Metal	Resistividad ($\Omega\text{-m}$) ^{10⁸}	Permeabilidad relativa a 0.002 Wb/m ²	$\delta\sqrt{\nu}$ (δ = profundidad de penetración, m; ν = frecuencia, Hz)	$10^7 R_s/\sqrt{\nu}$ (R_s = resistividad de la superficie; Ω/m^2)
Aluminio	2.828	1	0.085	3.33
Bronce fosforado	7.75	1	0.140	5.54
Cinc	5.38	1	0.117	4.60
Cobre	1.724	1	0.066	2.61
Cromo	2.6	1	0.081	3.21
Estaño	11.5	1	0.171	6.73
Grafito	1 000	1	1.592	62.81
Hierro magnético	10	200	0.011	90.9
Latón (65.8 Cu, 34.2 Zn)	6.29	1	0.126	4.99
Latón (90.9 Cu, 9.1 Zn)	3.65	1	0.096	3.79
Magnesio	4.6	1	0.108	4.26
Mercurio	95.8	1	0.493	19.43
Níquel	7.8	100	0.014	55.71
Oro	2.22	1	0.075	2.96
Plata	1.629	1	0.064	2.55
Platino	9.83	1	0.158	6.22
Plomo	22	1	0.236	9.32
Permalloy (78.5 Ni, 2.5 Fe)	16	8.000	0.0022	727
Supermalloy (5 Mo, 79 Ni, 16 Fe)	60	100.000	0.0012	4.880
Mumetal (75 Ni, 2 Cr, 5 Cu, 18 Fe)	62	20.000	0.0029	2.140

Fuente. Datos de la referencia bibliográfica 11.

Tabla 7-4. Máxima solubilidad de sólidos, $c_{\text{máx}}$, y aumento en la resistividad atómica, A , $\Omega\text{-m}/(\text{en } \%)$ para cobre, plata, oro y aluminio

Soluto	Cobre		Plata		Oro		Aluminio	
	$c_{\text{máx}}$	A	$c_{\text{máx}}$	A	$c_{\text{máx}}$	A	$c_{\text{máx}}$	A
Ag	4.9	1.4	20.34	—	*	3.6	23.8	115
Al	19.6	12.5	20.34	5	6	18.7	—	—
As	—	68	8.8	85	—	80	—	—
Au	*	55	*	36				
Be	16.4	62						
Cd	—	30	42.2	3.8	32.5	6.3	0.11	0.57
Cr	0.8	36	—	—	23	42.5	0.40	77
Cu	—	—	14.1	0.77	*	4.5	24.8	7.85
Fe	4.5	93	—	—	75	79	0.025	53
Ga	19.9	14.2	18.7	23.6	13	22	—	2.5
Ge	11.8	37.9	9.6	55	3.2	52		
In	11	10.6	20	17.8	12.6	13.9		
Mg	7	6.5	29.3	19.5	25	13	16.3	4.9
Mn	*	29	47	16	—	24.1	0.90	59.7
Ni	*	12.5	—	—	*	7.9	0.023	17.7
P	3.5	6.7						
Pd	*	8.9	*	4.4	*	4.1		
Pt	*	21	40.5	16	*	10.1		
Sb	6	54	7.2	72.5	1.1	68		
Si	11.25		39.5				1.59	6.5
Sn	9.1	28.8	11.5	43.6	6.8	33.6		
Zn	38.3	3.2	40.2	6.4	31	9.5	66.4	2.11

Fuente. Datos de las referencias bibliográficas 4 y 6 y de la 12 a la 16.

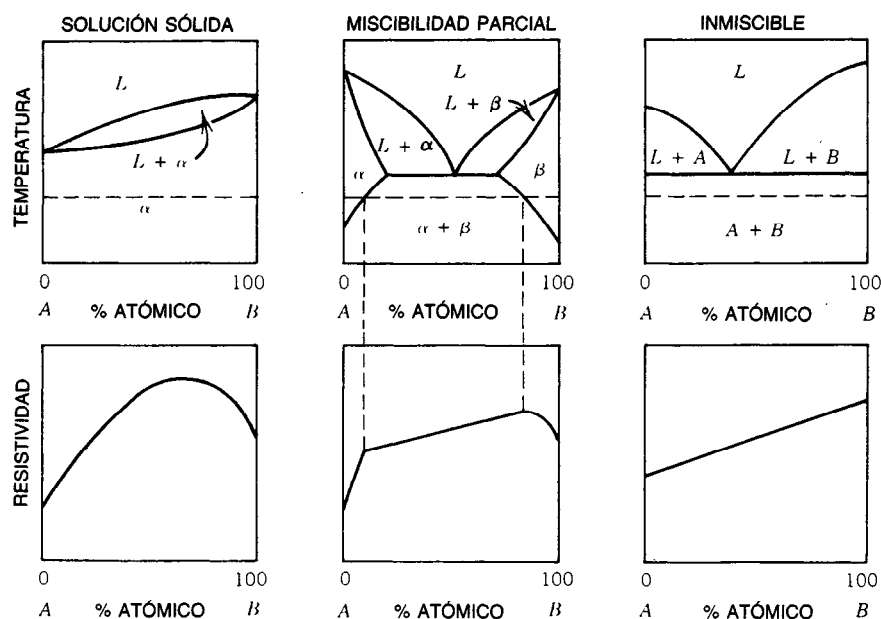


Fig. 7-1. Resistividades típicas de aleaciones binarias como función de la composición^{12,13}. (Se muestra la regla de Nordheim.)

cuando los procesos de emisión termoiónica y de efecto túnel tienen una contribución significativa a la conductividad. Ambos procesos mejoran con el aumento de la temperatura, lo que provoca una disminución en la resistividad⁶.

En la tabla 7-5 se muestran las correlaciones entre el tamaño de grano y la resistividad de películas de aluminio-cobre-silicio formadas por cloruro de plasma, que se utilizan en interconexiones de circuitos integrados¹⁹. En la tabla 7-6 se presentan las propiedades de los siliciuros refractarios, y en las tablas 7-7 y 7-8 las propiedades de materiales resistores y conductores de películas delgadas. En la tabla 7-9 se muestran las propiedades de alambres de cobre y aluminio que se utilizan para interconexiones.

7.3 CAPACIDAD DE TRANSPORTE DE CORRIENTE DE LOS CONDUCTORES

Las capacidades de transporte de corriente de alambres eléctricos, barras colectoras y cables están limitadas por los efectos de calentamiento producidos por la corriente. Un incremento de 30° por encima de una temperatura ambiente de 40° corresponde a una densidad de corriente de alrededor de 10⁶ A/m² en una barra colectoras de cobre.

En las tablas 7-10 y 7-11 se muestran las propiedades y las máximas capacidades de transporte de corriente de alambres.

Las capacidades de transporte de corriente de las películas delgadas están limitadas por la electroemigración; el valor normal de la densidad de corriente en películas delgadas es aproximadamente 100 A/m². La

electroemigración es la emigración de átomos bajo la influencia de diversas corrientes eléctricas^{20,28-31}.

7.4 PROPIEDADES DE CONDUCTIVIDAD

7.4.1 Conductividad volumétrica

La **conductividad eléctrica volumétrica** o de volumen, σ , es la densidad de corriente eléctrica, J , que transporta el conductor por unidad de intensidad de campo eléctrico, E :

$$\sigma = \frac{J}{E} \quad (7-10)$$

donde σ = conductividad eléctrica volumétrica, S/m

J = densidad corriente eléctrica, A/m²

E = intensidad del campo eléctrico, V/m

En un cristal único no cúbico, la conductividad es un tensor con nueve términos. Puede considerarse que los materiales isotrópicos cúbicos tienen una conductividad, y que los materiales policristalinos no cúbicos tienen una conductividad "promedio"²⁰.

La relación entre la conductancia G y la conductividad volumétrica está dada por:

$$G = \sigma \cdot \frac{A}{L} \quad (7-11)$$

donde A = sección transversal del conductor, m²

L = longitud del conductor, m

σ = conductividad, S/m en el SI o bien mho/m en el sistema MKS

Tabla 7-5. Deposición de la película, resistividad y tamaño de grano para películas de aleación de aluminio

Núm.	Película	Composiciones planificadas		Tempe- ratura del sustrato	Velocidad (nm/s)	Grosor (nm)	Resistividad específica		Tamaño de grano			
									Cómo se deposita		Después de recocido en N ₂ durante 15 min. a 450° (μΩ-cm)	Tamaño Intervalo promedio (μm)
		Cu %	Si %				Tamaño Intervalo promedio (μm)	Tamaño Intervalo promedio (μm)	Tamaño Intervalo promedio (μm)			
1 Al	FI-CM	—	—	TA	3.3	838	2.77	2.43 ^a	0.53	0.2–0.7	0.9	0.2–1.2
2 Al	FI-CM	—	—	200 °C	3.6	828	2.85	2.40 ^a	0.96	0.2–1.5	1.4	0.2–1.8
3 Al	M-CM	—	—	TA	2.1	813	2.93	2.68	0.30	0.1–1.0	0.7	0.1–1.8
4 Al+Cu	FI-CM	2	—	TA	2.9	762	4.19	3.28	0.23	0.1–0.5	1.0	0.1–1.8
5 Al+Cu	FI-CM	2	—	200 °C	2.9	762	3.12	3.35	0.67	0.2–1.0	2.3	0.2–2.6
6 Al+Cu	M-CM	2	—	TA	1.8	836	4.10	3.51	0.22	0.1–0.5	1.2	0.2–2.0
7 Al+Cu+Si	FI-CM	3.85	1.0	TA	3.0	850	7.99	4.42	0.55	—	0.2	—
8 Al+Cu+Si	FI-CM	3.85	1.0	200 °C	2.4	813	8.02	4.07	0.07	—	0.2	—
9 Al+Cu+Si	M-CM	2.0	1.0	TA	1.3	938	4.50	3.47 ^b	0.07	—	0.1	—

^a La resistividad de la película de Al recoída a temperatura ambiente (FI=fuente incandescente) es baja (2.43); esto se debe posiblemente a errores en la medición del grosor.

Todos los datos anteriores indican una resistividad específica de 2.71 $\mu\Omega\text{-cm}$.

^b Las películas de Al+Cu+Si clonadas mediante magnetrones de CD presentan menor resistividad respecto a la de las películas formadas a partir de una fuente incandescente. Esta diferencia se atribuye a la diferencia en la concentración de cobre.

FI=Fuente incandescente.

CM=clonado por magnetrones.

Tabla 7-6. Propiedades de los siliciuros refractarios

Siliciuro	Mínima temperatura eutéctica binaria (°C)	Altura de la barrera en el silicio tipo <i>n</i> (eV)	Resistividad ya recocido ($\mu\Omega\text{-cm}$)
CoSi ₂	1 195	0.64	18-20
CrSi ₂	1 300	0.57	~600
FeSi ₂	1 208	—	> 1000
HfSi ₂	1 300	—	45-50
MoSi ₂	1 410	0.55	40-100
NbSi ₂	1 295	—	50
NiSi ₂	966	0.70	~50
Pd ₂ Si	720	0.74	30-35
PtSi	830	0.87	28-35
TaSi ₂	1 385	0.59	35-70
TiSi ₂	1 330	0.60	13-26
VSi ₂	1 385	—	Soldable 50-55
WSi ₂	1 440	0.65	con una capa 30-100
ZrSi ₂	1 355	0.55	subyacente de Ni 35-40

Fuente. Datos de las referencias bibliográficas 20 a 22.

Tabla 7-7. Conductores de película delgada y película gruesa

Material	Grosor (μm)	(ohm/cuad.)	Propiedades
Película gruesa			
80Ag-20Pd	12	0.015-0.020	Soldable; resistencia aceptable al lixiviado
60Ag-40Pd	12	0.040-0.050	Soldable; resistencia moderada al lixiviado
90Ag-10Pt	15	0.015-0.020	Soldable; buena resistencia al lixiviado
Au	12	0.003-0.004	Oro común
Au	7	0.003-0.005	Oro delgado para impresión
Au-Pt	12	0.020-0.050	Soldable; buena resistencia al lixiviado
Cu	15	0.002-0.003	Soldable; excelente resistencia al lixiviado calentado al horno en atmósfera de N ₂
Ni	25	0.060-0.075	Soldable; excelente resistencia al lixiviado calentado al horno en atmósfera de N ₂
Película delgada			
Au	100	0.27-0.30	
Au	500	~0.05	
Au	10 000	~0.004	
Al	100	~0.33	
Al + Cu	5 000-1 000		Interconecta circuitos integrados
Al	1 000	~0.04	

Fuente. Datos de la referencias bibliográficas 7, 14, 23 y 24.

Tabla 7-8. Resistores de película delgada y película gruesa

Sistemas seleccionados de resistores de película gruesa (Coeficiente térmico de resistencia de +25 a -55 °C y de +25 a +125 °C ppm/K)				
om/cuad.	Serie 17 de Dupont ^a	Serie 8000 de Cermalloy ^b	Serie 3900 de ESL ^c	Serie 5500-1 de EMCA ^d
10	0 ± 175	0 ± 100	+ 50 ± 50	—
100	0 ± 50	0 ± 50	0 ± 50	0 ± 50
1 000	0 ± 50	0 ± 50	0 ± 50	0 ± 50
10K	0 ± 100	0 ± 50	0 ± 50	0 ± 50
100K	0 ± 100	0 ± 100	0 ± 100	0 ± 50
1M	0 ± 100	0 ± 100	- 100 ± 100	—

Resistores seleccionados de película delgada Resistividad laminar típica (μΩ-cm)			(ppm/K)
NiCr	50 - 300		0 ± 25
Ta-N	50 - 500		0 a - 200
Cr-SiO	100 - 5 000		0 ± 100
Si difuso	2 - 200		+ 600 a + 2 500

Fuente. Datos de las referencias bibliográficas 7, 9, 15, 24 y 25.

^a Electro-Science Laboratories, Inc., Pennsauken, NJ 08110.

^b Cermalloy, Inc., West Conshohocken, PA 19428.

^c Electro Materials Corp. of America, Mamaroneck, NY 10543.

^d E. I. du Pont de Nemours & Co., Inc., Wilmington, DE 19898.

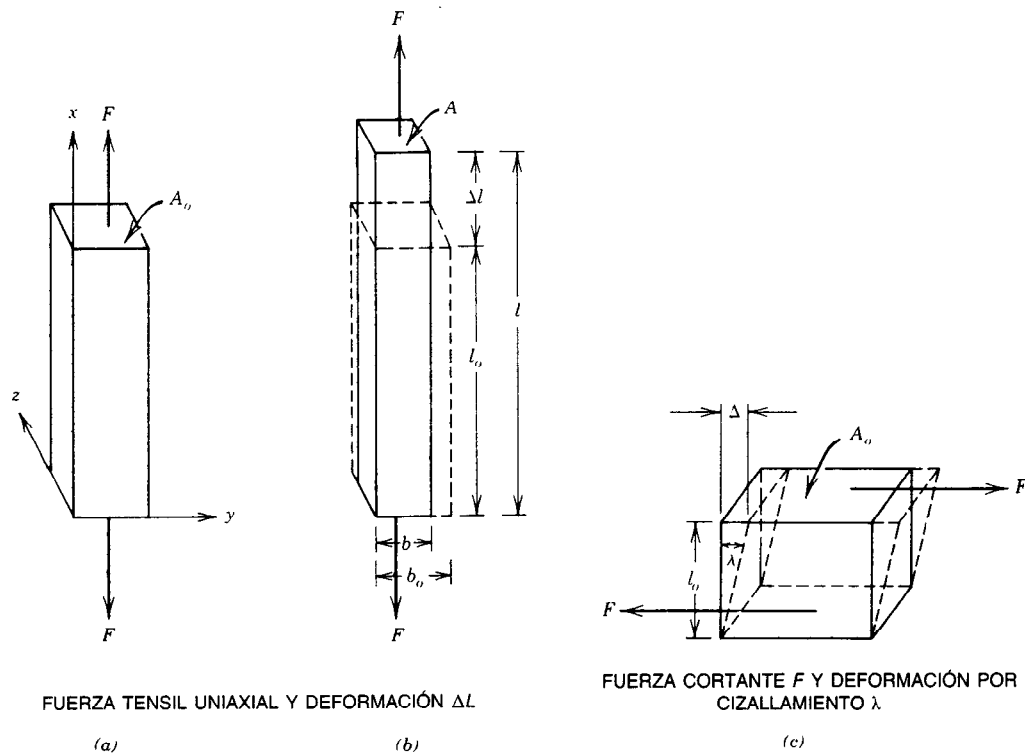


Fig. 7-2. Fuerza uniaxial; a) antes de la deformación; b) después de la deformación, ΔL ; c) fuerza de cizallamiento, F , y deformación por esfuerzo cortante, γ .

7.4.2 Conductividad de masa

La conductividad de masa, σ_m , se define con la expresión:

$$\sigma_m = G \frac{L^2}{m} \quad (7-12)$$

donde σ_m = conductividad de masa, $S \cdot m^2/kg$
 G = conductancia, S
 L = longitud del conductor, m
 m = masa, kg

7.5 PROPIEDADES RELACIONADAS CON LOS MATERIALES

7.5.1 Mecánicas

Esfuerzo de tensión, deformación por tensión. El esfuerzo de tensión o tensil, σ , es la fuerza F por unidad de área no deformada A_0 (fig. 7-2a):

$$\sigma = \frac{F}{A_0} \quad (7-13)$$

donde A_0 = área original sin deformación, m^2

F = fuerza, N

σ = esfuerzo, N/m^2 o Pa

Se recomienda mucho emplear la unidad SI de esfuerzo, Pascal, Pa , y el megaPascal, MPa ¹. Las siguientes son algunas conversiones útiles:

$$1 \frac{lbf}{pulg^2} = 6.8948 \text{ kPa} \quad (7-14)$$

$$1 \frac{kgf}{mm^2} = 9.80665 \text{ MPa} \quad (7-15)$$

La deformación técnica, Σ , es el cambio en longitud por unidad de longitud no distorsionada, l_0 :

$$\Sigma = \frac{\Delta l}{l_0} = \frac{l - l_0}{l_0} \quad (7-16)$$

En la tabla 7-11 se pueden consultar las propiedades tanto elásticas como mecánicas de algunos metales y aleaciones.

Esfuerzo cortante y deformación por esfuerzo cortante. El esfuerzo cortante es la fuerza de cizallamiento F por unidad de área A_0 (fig. 7-2b):

$$\tau = \frac{F}{A_0} \quad (7-17)$$

donde τ = esfuerzo cortante, Pa

F = fuerza de cizallamiento, N

A_0 = área, m^2

La deformación por esfuerzo cortante se define como la cantidad de deformación Δ en un cizallamiento por unidad de longitud (fig. 7-2c):

$$\gamma = \frac{\Delta}{l_0} \quad (7-18)$$

Módulos de Young, de cizallamiento y volumétrico. El módulo de Young, E , es la constante de proporcionalidad entre el esfuerzo σ y la deformación Σ ; lo que se representa:

$$\sigma = E \Sigma \quad (7-19)$$

Esta ecuación es válida sólo para deformaciones elásticas. El módulo de Young también se conoce como ley de Hooke.

El módulo de cizallamiento, G , es la constante de proporcionalidad entre el esfuerzo cortante, τ , y la deformación por esfuerzo cortante, γ :

$$\tau = G\gamma \quad (7-20)$$

Este módulo se conoce también como módulo de corte o coeficiente de rigidez.

El módulo de volumen o volumétrico, K , es el recíproco de la compresibilidad:

$$K = \frac{1}{\beta} \quad (7-21)$$

La compresibilidad β es la constante de proporcionalidad entre el decremento de volumen $\Delta V/V$ y el cambio de presión ΔP :

$$-\frac{\Delta V}{V} = \beta(\Delta P) \quad (7-22)$$

Coefficiente de Poisson. Este coeficiente es la relación entre la contracción lateral ϵ_y y la elongación longitudinal ϵ_x cuando un material se deforma bajo tensión uniaxial.

La deformación lateral ϵ_y se determina a partir del ancho original (b_0) y el ancho contraído (b) cuando el coeficiente de Poisson ν es:

$$\nu = -\frac{\epsilon_y}{\epsilon_x} \quad (7-23)$$

Para materiales comunes, el valor de ν es cercano a 1/3.

7.5.2 Térmicas

Conductividad térmica. El flujo de calor q a lo largo de una varilla es proporcional al gradiente de temperatura dT/dx y a la sección transversal de la varilla A . La constante de proporcionalidad, λ , se conoce como conductividad térmica^{7, 32}.

$$q = -\lambda A \frac{dT}{dx} \quad (7-24)$$

donde λ = conductividad térmica, $W/m \cdot K$

A = área, m^2

dT/dx = gradiente de temperatura.

Tabla 7-9. Propiedades de los alambres de cobre y aluminio

Tamaño (calibre estadounidense de alambres, AWG)	Diámetro (pulg.)	Área (mpulg. cir.) (pulg ²)	Cu por ASTM B3-74 ^a		Al por ASTM B609-81 ^b	
			Peso (#/1000')	Máxima resistivi- dad a la CD a 20 °C (Ω/1000')	Peso (#/1000')	Máxima resistivi- dad a la CD a 20 °C (Ω/1000')
0000	0.4600	211600.	640.5	0.04901	194.7	0.08035
000	0.4096	167800.	507.9	0.06180	154.4	0.1013
00	0.3468	133100.	402.8	0.07793	122.5	0.1278
0	0.3249	105600.	319.4	0.09827	97.15	0.1611
1	0.2893	83690.	253.3	0.1239	77.03	0.2031
2	0.2576	66360.	200.9	0.1563	61.07	0.2562
3	0.2294	52620.	159.3	0.1970	48.43	0.3230
4	0.2043	41740.	126.3	0.2485	38.41	0.4073
5	0.1819	33090.	100.2	0.3133	30.45	0.5136
6	0.1620	26240.	79.46	0.3951	24.15	0.6476
7	0.1443	20820.	63.02	0.4982	19.16	0.8167
8	0.1285	16510.	49.97	0.6282	15.20	1.030
9	0.1144	13090.	39.63	0.7921	12.04	1.299
10	0.1019	10380.	31.43	0.9988	9.556	1.637
11	0.0907	8230.	24.92	1.260	7.571	2.065
12	0.0808	6530.	19.77	1.588	6.008	2.604
13	0.0720	5180.	15.68	2.003	4.771	3.283
14	0.0641	4110.	12.43	2.525	3.871	4.140
15	0.0571	3260.	9.858	3.184	3.001	5.220
16	0.0508	2580.	7.818	4.015	2.375	6.583

Tabla 7-9. Propiedades de los alambres de cobre y aluminio (continuación)

Tamaño (calibre estadounidense de alambres, AWG)	Diámetro (pulg.)	Área (mpulg. cir.)	Cu por ASTM B3-74 ^a		Al por ASTM B609-81 ^b	
			Peso (#/1000')	Máxima resistivi- dad a la CD a 20 °C (Ω/1000')	Peso (#/1000')	Máxima resistivi- dad a la CD a 20 °C (Ω/1000')
17	0.0453	2050.	0.00161	5.063	1.889	8.301
18	0.0403	1620.	0.00128	6.385	1.495	10.47
19	0.0359	1290.	0.00101	8.051	1.186	13.20
20	0.0320	1020.	0.000804	10.15	0.9424	16.64
21	0.0285	812.	0.000638	12.80	0.7475	20.99
22	0.0253	640.	0.000503	16.14	0.5891	26.46
23	0.0226	511.	0.000401	20.36	0.4701	33.37
24	0.0201	404.	0.000317	25.67	0.3718	42.08
25	0.0179	320.	0.000252	32.37	0.2949	53.06
26	0.0159	253.	0.000199	40.81	0.2327	66.91
27	0.0142	202.	0.000158	51.47	0.1856	84.37
28	0.0126	159.	0.000125	64.90	0.1461	106.4
29	0.0113	128.	0.000100	81.83	0.1175	134.2
30	0.0100	100.	0.0000785	103.2	0.09203	169.2
31	0.0089	79.2	0.0000622	130.1	0.07289	213.3
32	0.0080	64.0	0.0000503	164.1	0.05890	270.0
33	0.0071	50.4	0.0000396	206.9	0.04638	339.2
34	0.0063	39.7	0.0000312	260.9	0.03654	427.7
35	0.0056	31.4	0.0000246	329.0	0.02890	539.3
36	0.0050	25.0	0.0000196	414.8	0.02301	680.1

Fuente. Datos de las referencias bibliográficas 14, 24 y 25.

^a En base a las especificaciones ASTM B3-74, B609-81 ($\rho = 2.8264 \mu\Omega\text{-cm}$) y B258-81.

Tabla 7-10. Máxima capacidad de transporte de corriente de alambres de cobre y aluminio (amperes)

MIL-W-5088									
Tamaño (AWG)	Cobre		Aluminio		National Electrical Code	Underwriters Laboratory		American Insurance Associ- ation	500 m pulg cir./A
	Mono- filar	Haz de alambres	Mono- filar	Haz de alambres		60°	80°		
30	—	—	—	—	—	0.2	0.4	—	0.20
28	—	—	—	—	—	0.4	0.6	—	0.32
26	—	—	—	—	—	0.6	1.0	—	0.51
24	—	—	—	—	—	1.0	1.6	—	0.81
22	9	5	—	—	—	1.6	2.5	—	1.28
20	11	7.5	—	—	—	2.5	4.0	3	2.04
18	16	10	—	—	6	4.0	6.0	5	3.24
16	22	13	—	—	10	6.0	10.0	7	5.16
14	32	17	—	—	20	10.0	16.0	15	8.22
12	41	23	—	—	30	16.0	26.0	20	13.05
10	55	33	—	—	35	—	—	25	20.8
8	73	46	58	36	50	—	—	35	33.0
6	101	60	86	51	70	—	—	50	52.6
4	135	80	108	64	90	—	—	70	83.4
2	181	100	149	82	125	—	—	90	132.8
1	211	125	177	105	150	—	—	100	167.5
0	245	150	204	125	200	—	—	125	212.0
00	283	175	237	146	225	—	—	150	266.0
000	328	200	—	—	275	—	—	175	336.0
0000	380	225	—	—	325	—	—	225	424.0

Fuente. Datos de las referencias bibliográficas 14, 24 y 25.

Tabla 7-11. Constantes de elasticidad y propiedades mecánicas de metales y aleaciones policristalinos a temperatura ambiente

Material	Condición o temple	Módulo de Young (GPa)	Punto de fluencia a 0.2% de deformación (MPa)	Resistencia última a la tensión (MPa)	Elongación en 2 pulg (%)	Dureza
Acero al carbono AISI 1010	Laminado en caliente, estirado en frío	200	180 305	325 365	28 20	Brin 95 Brin 105
Acero al carbono AISI 1042	Laminado en caliente, estirado en frío	200	305 515	550 615	16 12	Brin 163 Brin 179
Acero al carbono AISI 1078	Laminado en caliente, estirado en frío	200	380 500	690 650	12 10	Brin 207 Brin 192
Acero inoxidable Tipo 304	Recocido ASTM A666 GrC	195	205 515	515 860	40 10	— —

Tabla 7-11. Constantes de elasticidad y propiedades mecánicas de metales y aleaciones policristalinos a temperatura ambiente (continuación)

Material	Condición o temple	Módulo de Young (GPa)	Punto de fluencia a 0.2% de deformación (MPa)	Resistencia última a la tensión (MPa)	Elongación en 2 pulg (%)	Dureza
Acero al carbono Tipo 450	Acero inoxidable	200	170	415	—	R _B 88
54Fe-29Ni-17Co ASTM F15	trabajado en frío	140	345	515	—	R _B 100
58Fe-42Ni ASTM F30	Recocido	145	275	450	30	R _B 70
48Fe-52Ni ASTM F30	Recocido	165	275	450	35	R _B 70
Oro	Recocido	83	—	131	45	Vick 25
	trabajo en frío		207	221	4	—
Platino	Recocido	172	14	124	40	Vick 40
	trabajo en frío		186	234	2.5	Vick 100
Plata	Recocido	76	55	152	48	Vick 26
	trabajado en frío		303	345	2.5	—
Aleación de aluminio 1100	— O	69	34	90	35	R _B 23 ^a
	— H18		150	165	5	R _B 44 ^a
Aleación de aluminio 2024	— O	72	76	185	20	R _B 47 ^a
	— T3		345	485	18	R _B 120 ^a
Aleación de aluminio 5056	— O	72	150	290	—	R _B 65 ^a
	— H18		405	435	—	R _B 105 ^a
Aleación de aluminio 6061	— O	68	55	125	25	R _B 30 ^a
	— T6		275	310	12	R _B 95 ^a
Aleación de aluminio 7075	— O	71	105	230	17	R _B 60 ^a
	— T6		505	570	11	R _B 150 ^a
Cobre C10100 OFHC	Recocido	115	69	220	45	R _F 40
	— HO2		205	260	35	R _F 70
	— HO4		250	290	14	R _F 84
Cobre C14500 0.5% Be	Recocido	115	69	220	50	R _F 40
	— HO2		275	290	25	R _F 42
	— HO4		305	330	20	R _F 48
Cobre C17200 2.0% Be	Sol. tratada & TD04	125	515	655	20	R _B 92
Cobre C23000 latón rojo	Recocido	115	69	270	48	R _F 56
	— HO2		340	395	12	R _B 65
	— HO4		395	485	5	R _B 77

Fuente. Datos de las referencias bibliográficas 5, 6, 26 y 27.

^a Escala B de Rockwell, 500 kg de carga, bolas de 10 mm.

Ley de Wiedemann-Franz. Se ha determinado que en los metales la razón de la conductividad térmica λ entre el producto de la temperatura T y la conductividad eléctrica σ es una constante, L :

$$L = \frac{\lambda}{\sigma T} \quad (7-25)$$

donde λ = conductividad térmica, W/m · K
 σ = conductividad eléctrica, S/m
 T = temperatura, K

Esta razón, denominada relación de Wiedemann-Franz, es constante para metales huéspedes puros y su valor varía de 20 000 a 30 000 ($\mu\text{V/K}$)².

Calor específico. El calor específico, conocido también como capacidad calorífica, es la cantidad de calor que se requiere para elevar 1 K la temperatura de 1 kg de material a presión constante, y se denota por C_p . La misma cantidad a volumen constante se denota por C_v , y ambas cantidades están relacionadas por:

$$C_p - C_v = \alpha^2 VTK \quad (7-26)$$

donde α = coeficiente de expansión volumétrica
 V = volumen molar
 T = temperatura, K
 K = módulo volumétrico
 Para metales sólidos, $C_p \approx C_v$.

Coefficiente térmico de resistividad eléctrica (CTR). Por encima de la temperatura de Debye, cuando todos los nodos de vibración de la malla están excitados, la componente térmica de la resistividad de los conductores es aproximadamente lineal. Así, la resistividad puede expresarse como:

$$\rho = \rho_0[1 + 2(T - \alpha T) + \dots] \quad (7-27)$$

donde ρ_0 = resistividad a temperatura ambiente
 α = coeficiente térmico de resistividad eléctrica, °C⁻¹

Para los metales puros, el valor de α es aproximadamente 0.004/°C o 4 000 ppm/°C en unidades conocidas. Para las aleaciones, α suele ser menor. Esta ecuación no se cumple para altas temperaturas, con diferentes procesos de dispersión de electrones.

Coefficiente de dilatación lineal. Este coeficiente, α_L , se define como el cambio en longitud, dl , por unidad de longitud, l , por cada grado de aumento en la temperatura, dT :

$$l = l_0[1 + \alpha_L(T - T_0)] \quad (7-28)$$

donde l_0 = longitud, m, a la temperatura T_0
 l = longitud, m, a la temperatura T
 T = temperatura, K

Si se conoce la longitud l_0 de una varilla a la temperatura T_0 , su longitud l a otra temperatura T es:

$$l = l_0[1 + \alpha_L(T - T_0)] \quad (7-29)$$

Coefficiente volumétrico de dilatación. Este coeficiente, α_V , se define como:

$$\alpha_V = \frac{1}{V} \frac{dV}{dT} \quad (7-30)$$

donde V = volumen, m³
 T = temperatura, K
 En general, $\alpha_V = 3 \alpha_L$.

Flujo plástico. El flujo plástico o deslizamiento es la deformación plástica continua de los materiales como función del tiempo bajo un esfuerzo dado. Se acelera bastante a altas temperaturas, cercanas a la temperatura de fusión.

7.5.3 Electroemigración

La **electroemigración** es el transporte de masa bajo la influencia de una corriente eléctrica. Puede verse una explicación exhaustiva en la referencia 30. La electroemigración provoca fallas en las componentes de metales conductores de un circuito integrado. En la tabla

Tabla 7-12. Tiempos medios antes de la falla de conductores de aluminio, oro y cobre como función de las condiciones de aleación ($J = 2 \times 10^6$ A/cm²)

Conductor	Aleación (% en peso)	Tiempo medio antes de falla (h)
Al a 175 °C		
Al		30-45 ^a
	Si 1.8	100-200
	Cu 4	2 500
	Cu 4, Si 1.7	4 000 ^a
	Ni 1	3 000
	Cr	8 300
	Mg 2	1 000
	Cu 4, Mg 2	10 000 ^a
	Cu 4, Ni 2, Mg 1.5 ^b	32 000 ^a
	Au 2	55
	Ag 2	45
Au a 300 °C^c		
	Ta	4 000 ^a
	Ni-Fe	800
	Mo	100 ^a
	W + Ti	25
Cu a 300 °C		
Cu		180
	Al 1	300
	Al 10	6 000
	Be 1.7 ^b	20 000 ^a

Fuente. Datos de las referencias bibliográficas 19, 28 y 33.

^a Extrapolado por medio de valores apropiados de n y ΔH .

^b Depositado de una fuente de aleación por evaporación; la composición indicada es la de la fuente.

^c Las películas de oro en SiO₂ requieren una capa de adhesión, que se hace del material indicado en la tabla. Después del recocido ocurre aleación en cierta medida.

Tabla 7-13. Resultados de la difusión de película delgada medidos con técnicas superficiales y con el método de difusión lateral

Difusor Matriz	Métodos de medición	Temperatura (°C)	Difusividad (cm ² /s)	Energía de activación (eV)
$\frac{\text{Ag}}{\text{Au}}$	Acumulación superficial AES	30 – 260	$10^{-5}(D_b^0)$	0.63
$\frac{\text{Ag}}{\text{Au}}$	Perfilado por clono AES	30	$\sim 2 \times 10^{-18}$	
$\frac{\text{Ag}}{\text{Au}}$	Dispersión de iones He	30	$\sim 10^{-14}$	
$\frac{\text{Ag}}{\text{Au}}$	Acumulación superficial AES	150 – 260	$10^{-6}(D_b^0)$	1.1
$\frac{\text{Cr}}{\text{Au}}$	Acumulación superficial AES	210 – 293	$10^{-3}(D_b^0)$	1.09
$\frac{\text{Cu}}{\text{Au-0.2 Co}}$	Acumulación superficial AES	100	6.4×10^{-15}	
		125	3.6×10^{-14}	
		150	1.3×10^{-13}	
$\frac{\text{Cu}}{\text{Au}}$	Retrodispersión de iones He	25	$\sim 10^{-17}$	
$\frac{\text{Co}}{\text{Au-0.86 Co}}$	Perfilado por clono AES	100	3.3×10^{-18}	
		150	2.4×10^{-16}	
$\frac{\text{Cu}}{\text{Al}}$	Perfilado por clono AES	175	8.6×10^{-13}	
$\frac{\text{Cu}}{\text{Al}}$	Dispersión lateral	150 – 300	$10(D_b^0)$	0.81
$\frac{\text{Au}}{\text{Ag}}$	Acumulación superficial ESCA	141	4.3×10^{-17}	
$\frac{\text{Pt}}{\text{Cr}}$	Acumulación superficial AES	400 – 580	$10^{-2}(D_b^0)$	1.69
$\frac{\text{Ag}}{\text{Cu}}$	Acumulación superficial AES	300	$\sim 10^{-14}$	
$\frac{\text{Pb}}{\text{Cu}}$	Acumulación superficial AES	125 – 230		0.52
$\frac{\text{Ag}}{\text{Ni}}$	Acumulación superficial AES	300	$\sim 10^{-18}$	
$\frac{\text{Au}}{\text{Pt}}$	Acumulación superficial AES	250 – 350		0.96
$\frac{\text{Si}}{\text{W}}$	Acumulación superficial AES	670 – 850		2.6 – 3.2
$\frac{\text{Sn}}{\text{Sn}}$	Barrido cont rizador	142 – 213	1.8×10^{-5}	0.46

Fuente. Datos de la referencia bibliográfica 34.

Tabla 7-14. Datos de difusión con trazador radiactivo de fronteras granulares para sistemas seleccionados

Difusor Matriz ^a	Método de medición	$\delta D^0(\text{cm}^3/\text{s})$	Energía de activación (eV)	Observaciones ^b
¹⁹⁵ Au Au (pelíc. epi)	Bombardeo iónico	1.9×10^{-10}	1.16	Disl. dis.
¹⁹⁵ Au Au (volumen de la p.)	Bombardeo iónico	3.1×10^{-10}	0.88	Fronteras granulares de ángulo grande
¹⁹⁵ Au Au (pelíc. de p.)	Bombardeo iónico	9.0×10^{-10}	1.0	Fronteras granulares de ángulo grande
¹⁹⁵ Au Au-1.2 Ta (volumen de la p.)	Bombardeo iónico	5×10^{-7}	1.26	Fronteras granulares de ángulo grande
¹⁹⁵ Au Au-1.2 Ta (pelíc. de p.)	Bombardeo iónico	1×10^{-9}	1.2	Fronteras granulares de ángulo pequeño
¹⁹⁵ Au Au-1.2 Ta (volumen de la p.)	Bombardeo iónico	1×10^{-9}	1.2	Fronteras granulares de ángulo pequeño
¹⁹⁵ Au Ni-0.5 Co (pelíc. de p.)	Bombardeo iónico	1.4×10^{-10}	1.6	Fronteras granulares de ángulo grande
¹¹⁰ Ag Ag (volumen de la p.)	Corte delgado en torno	1.3×10^{-9}	0.80	Fronteras granulares de ángulo grande
Ni Ni (bicristalino)	Autorradiografía	7×10^{-10}	1.08	45<100>
⁶³ Ni Ni (bicristalino)	Autorradiografía	2.2×10^{-8}	1.77	10°<112> inclinación

Fuente. Datos de las referencias bibliográficas 28 y 34.

^a p. = matriz policristalina; epi = epitaxial.

^b Disl. dis. = dislocaciones disociadas.

7-12 se presenta el tiempo medio antes de falla (MTF) para películas de aluminio, oro y cobre. La electroemigración puede ocurrir en las películas delgadas a través de las fronteras granulares²⁹.

7.5.4 Difusión

La difusión es el transporte de masa en un sólido, líquido o gas. La fuerza impulsora es un gradiente de concentración. La ley de Fick establece que el flujo J de una componente es proporcional a su gradiente de concentración dC/dx .

$$J = -D \frac{dC}{dx} \quad (7-31)$$

donde D = coeficiente de difusión, m^2/s

J = flujo, $\text{s}^{-1} \text{cm}^{-2}$

C = concentración, cm^{-3}

El coeficiente de difusión depende de la temperatura y obedece la ley exponencial:

$$D = D_0 \exp \left(\frac{-Q}{kT} \right) \quad (7-32)$$

donde D = coeficiente de difusión a la temperatura T

D_0 = constante preexponencial o coeficiente de difusión

Q = energía de activación para la difusión, J o eV

k = constante de Boltzmann, J/K o eV/K

T = temperatura, K

En las tablas 7-13 y 7-14 se presentan datos importantes de difusión para algunos pares³⁴⁻³⁶.

7.6 MATERIALES DE ENCAPSULAMIENTO MICROELECTRÓNICO

Encapsulamiento microelectrónico²⁵ es un término amplio que incluye todos los materiales y procesos por medio de los cuales se proporciona protección ambiental y mecánica, control térmico e interconexión eléctrica a un dispositivo semiconductor o ensamblaje de dispositivos semiconductores.

En muchos casos se utilizan alambres de oro, aluminio o ambos²⁸ para la conexión eléctrica con el chip semiconductor. En la tabla 7-15 se presentan algunas propiedades de estos alambres de conexión. Los ma-

Tabla 7-15. Propiedades de alambres desfatigados y de alambres de unión microelectrónicos recocidos

Alambre de unión	Diámetro		Elongación (%)	Resistencia mecánica a la tensión (tracción) (g)	Resistencia eléctrica ($\Omega/0.100''$)	Capacidad de corriente ^a (A)
	(pulg.)	(μm)				
Oro	0.0007	18	2-6 6-10	3.5 3.5	0.225	0.380
	0.001	25	2-27 6-10	9 9	0.110	0.648
	0.00125	32	3-8 7-12	14 12	0.071	0.906
	0.002	51	5-12 12-20	32 30	0.028	1.834
Aluminio, 1% de silicio	0.0007	18	0.75-3 0.75-3.5	6 4	0.273	0.282
	0.001	25	1-3.5 1-4	15 12	0.133	0.481
	0.00125	32	1-4 1-4	22 17	0.086	0.671
	0.002	51	2-7 2-8	62 50	0.034	1.359

^a Admisible por la MIL-M-38510E para alambres de más de 0.040 pulg de longitud.

teriales específicos de encapsulamiento se eligen en base a su capacidad de resolver satisfactoriamente un problema específico. Suele ser necesario hacer diversas consideraciones, y el método elegido de encapsulamiento representa una solución intermedia. Algunos de los factores más importantes y algunas consideraciones relacionadas con ellos que afectan la elección de los materiales de encapsulamiento son:

1. Requerimientos térmicos³⁴:

- Máximas temperaturas permisibles en las uniones.
- Igualación de los coeficientes de dilatación.
- Intervalo de temperatura ambiente de operación.

2. Requerimientos eléctricos:

- Efectos de la frecuencia.
- Protección electrostática o electromagnética.
- Caídas de voltaje en las interconexiones.
- Interferencia o interacciones electromagnéticas (EMI).
- Sensibilidad a la luz.
- Resistencia de aislamiento.

3. Requerimientos de fiabilidad:

- Hermetismo.
- Protección contra el ambiente.

- Resistencia a la radiación.
- Susceptibilidad a las partículas alfa.
- Resistencia a la fatiga.

4. Consideraciones prácticas:

- Costo.
- Equipo disponible.
- Experiencia con que se cuenta.
- Materiales disponibles.
- Programación.
- Seriación del proceso.
- Facilidad de retrabajar.
- Peso.
- Estructura.
- Tamaño.
- Especificaciones impuestas.

Se dan algunos materiales de uso común utilizados en la industria de la microelectrónica, así como sus limitaciones. La mayoría de las aplicaciones están limitadas por los materiales.

En la página 115 se muestran tres tablas en las que se especifican algunos de los materiales resistivos, conductores y para empaque o encapsulamiento respectivamente.

En la tabla 7-16 se pueden consultar algunas propiedades de materiales de encapsulamiento seleccionados.

Tabla 7-16. Propiedades de materiales de encapsulamiento seleccionados

Material	Punto de fusión (°C)		Densidad (g/cm ³)	Conductividad térmica (W/m-K)	Dilatación (10 ⁻⁶ /K)	Resistividad (μΩ-cm)
54Fe-29Ni-17Co ASTM F15	1 450		8.37	17.6	5.5	49
58Fe-42Ni ASTM F30	1 440		8.12	10.9	4.3	72
48Fe-52Ni ASTM F 30	1 435		8.30	13.3	9.8	43
Acero inoxidable tipo 304	1 400		7.9	16.2	17.2	72
Acero al carbono AISI 1010	1 510		7.84	60.2	12.2	16.8
Aluminio	660		2.70	226	23.5	2.87
Cobre (OFHC)	1 083		8.96	391	17.0	1.72
Molibdeno	2 615		10.2	134	5.1	5.7
Níquel	1 455		8.9	88	13.3	6.9
Silicio	1 412		2.34	138	7.6	10 ⁵
Tungsteno	3 400		19.3	159	4.5	5.5
Al ₂ O ₃ , 96%	—		3.7	30.2	6.4	10 ²⁰
Al ₂ O ₃ , 99,5%	2 040		3.9	32.3	6.5	10 ²⁰
BeO, 99,5%	2 550		2.9	215	7.5	10 ²⁰
Soldaduras	Sólido	Líquido				
52In-48Sn	118	118	7.3	34	20	14.7
97In-3Ag	143	143	7.38	73	22	7.5
62Sn-36Pb-2Ag	179	179	8.41	31	25	14.8
63Sn-37Pb	183	183	8.4	50	25	15
50In-50Pb	180	209	8.86	22	27	28.7
96.5Sn-3.5Ag	221	221	7.36	33	30	10.8
Adhesivos						
Epoxy-rellenado con Ag	—	—	—	1.6	—	200
Epoxy-sin relleno	—	—	—	0.35	—	10 ¹³

Fuente. Datos de las referencias bibliográficas 6, 23 y 27.

Aplicación	Material	Desventaja
Coefficiente térmico de resistividad cero	Nicromo	
Resistores alambre de precisión	Nicromo	Envejecimiento por flujo plástico del alambre
Resistores de precisión	Carbono	Incompatibilidad con el procesamiento microelectrónico
Resistencia de película gruesa	Basado en rutenio	Alto costo; tienden a fallar cuando en la atmósfera se encuentran hidrocarburos

Aplicación	Material	Desventajas
Alambres de interconexión	Al	No son fuertes; el Si tiende a precipitarse ³⁷ ; propensos a la electroemigración
	Al-Si	
	Al-Cu-Si	El Cu evita la electroemigración
Alambre de interconexión de 25 µm de diámetro	Au puro	Alto costo; son suaves Son duros y ligables
	Au + 3 ppm de Be	
Películas de interconexión delgadas para altas corrientes y altas temperaturas	Al-Cu-Si	Propensos a la electroemigración y a la degradación ³⁷
	Siliciuros de Mo, siliciuros de W	Sensibles al método de deposición en la fase de vapor
Películas gruesas de interconexión	Películas de Pt-Pd-Ag, Au impurificado con Be, Pt o Cu	Problemas con la formación interfacial de compuestos de silicio ³⁸ ; se liga con las componentes; son costosos

Aplicación	Material	Desventajas
Sustratos	Al ₂ O ₃ , alúmina	La radiactividad en el sustrato provoca inversión en la memoria del dispositivo
Sustratos	BeO	
Unen los dispositivos con los sustratos y disipan el calor	Resinas epóxicas conductoras	
Cubiertas de sellos en el encapsulamiento microelectrónico; las cubiertas son de KOVAR chapeado de oro o de aleación 42	Soldadura de Au-Sn y de Ag-Sn	Defectos estéticos; la soldadura humedece la superficie superior
Cabezales para sujetar dispositivos	Chapeados con níquel, acero laminado en frío; los pasadores son de KOVAR chapeado de níquel	No se pueden chapear de oro; los pasadores chapeados de níquel tienen problemas de unión y soldadura

REFERENCIAS BIBLIOGRÁFICAS

1. American National Standard Metric Practice IEEE Std. 268; ANSI Z210; ASTM E 380-76.
2. R. M. Rose, L. A. Shepard y J. Wolff, "The Structure and Properties of Materials", vol. IV, *Electronic Properties*, Wiley, Nueva York.
3. E. H. Sondheimer, *Adv. Phys.* 1:1.
4. A. N. Gerritsen, "Metallic Conductivity, Experimental Part", vol. 19, p. 137, editado por S. Flügge, en: *Handbuch der Physik*, Springer-Verlag, Berlín.
5. C. J. Smithells, *Metals Reference Book*, 5.^a ed., Butterworth, Londres.
6. B. P. Bardes, ed., *Metals Handbook*, 9.^a ed., American Society for Metals, Metals Park, PA.
7. C. A. Harper, ed., *Handbook of Materials and Processes for Electronics*, McGraw-Hill, Nueva York.
8. E. S. Parker, *Material Data Book*, McGraw-Hill, Nueva York.
9. C. T. Lynch, *Handbook of Materials Science*, CRC, Boca Raton, FL.
10. L. E. Toth, *Transition Metal Carbides and Nitrides*, Academic Press, Nueva York.
11. O. E. Gray, ed., *AIP Handbook*, 3.^a ed., McGraw-Hill, Nueva York.
12. M. Hansen y K. Anderko, *Constitution of Binary Alloys*, McGraw-Hill, Nueva York.
13. R. P. Elliot, *Constitution of Binary Alloys*, 1.^a Sppl., McGraw-Hill, Nueva York.
14. D. G. Fink y D. Christiansen, *Electronics Engineers Handbook*, 2.^a ed., McGraw-Hill, Nueva York.
15. A. S. Bereznoi, *Silicon and Its Binary Systems*, Consultants Bureau, Nueva York.
16. F. A. Trumbore, "Solid Solubilities of Important Elements in Ge and Si," *Bell System Tech. J.* 39:205.
17. D. S. Campbell, *The Use of Thin Films in Physical Investigation*, Academic, Nueva York, p. 299.
18. C. A. Neugebauer, editado por B. Schwartz y N. Schwartz en *Measurement Techniques for Thin Films*, The Electrochemical Society, Nueva York, p. 191.
19. P. B. Ghate y T. C. Blair, *Thin Solid Films* 55:113.
20. F. Mohammadi, *Solid State Tech.* 24(1):65.
21. A. K. Sinha, *J. Vac. Sci. Technol.* 19(3):778.
22. S. P. Murarka, *J. Vac. Sci. Technol.* 17(4).
23. C. A. Harper, ed., *Handbook of Electronic Packaging*, McGraw-Hill, Nueva York.
24. D. G. Fink, *Electronics Engineers Handbook*, McGraw-Hill, Nueva York.
25. D. G. Fink y H. W. Beaty, *Standard Handbook for Electrical Engineers*, McGraw-Hill, Nueva York.
26. L. F. Mondolfo, *Aluminum Alloys, Structure and Properties*, Butterworth, Londres.
27. *Materials Engineering Magazine*, Materials Selector, Penton/IPC.
28. J. M. Poate, K. N. Tu y J. W. Mayer, *Thin Films: Interdiffusion and Reactions*, Wiley, Nueva York.
29. F. J. d'Heurle y P. S. Ho, "Electromigration in Thin Films", en Poate y col.²⁸, p. 243.
30. H. B. Huntington y A. R. Grone, *J. Phys. Chem. Solids* 20: 76.
31. J. R. Black, Electromigration—A Brief Survey and Some Recent Results, *IEEE Trans. Electron Dev.* 16: 338.
32. A. Goldsmith y col., *Handbook of Thermo-physical Properties of Solid Materials*, CRC, Boca Raton, FL.
33. G. T. Meadea, *Electrical Resistance of Metals*, Plenum, Nueva York.
34. D. Gupta, D. R. Campbell y P. S. Ho, "Grain Boundary Diffusion in Thin Films", en Poate y col.²⁸, p. 161.
35. D. Gupta, "Thin Film Couples", en Poate y col.²⁸.
36. Y. Adda y J. Philibert, *La Diffusion dans les Solides*, 1, Vols. 1 y 2, Presse. Univ. France (París).
37. M. Paugh y K. Seshan, "Degradation of Mechanical Properties of Annealed 25 μ m Al-1% Si Microelectronic Interconnect Wire", *Met. Trans.* 14(A), 921.
38. G. Ottaviani y J. W. Meyer, "Mechanisms and Interfacial Layers on Silicide Formation; Thin Film Interactions", en *Reliability and Degradation*, M. J. Howes y D. V. Morgan, eds., Wiley, Nueva York, pp. 105-149.

CAPÍTULO 8

Materiales aislantes y dieléctricos

Sorab K. Ghandhi

J. Keith Nelson

Rensselaer Polytechnic Institute
Troy, Nueva York

8.1 Películas aislantes para dispositivos de estado sólido y circuitos integrados

- 8.1.1 Óxidos nativos
- 8.1.2 Películas depositadas
- 8.1.3 Películas orgánicas

8.2 Materiales dieléctricos para almacenamiento de carga

- 8.2.1 Clasificación de los dieléctricos
- 8.2.2 Aspectos fundamentales

- 8.2.3 Falla de materiales dieléctricos sólidos
- 8.2.4 Aplicaciones de los capacitores
- 8.2.5 Otras aplicaciones

8.3 Materiales dieléctricos como aislantes

- 8.3.1 Contexto y criterios de selección
 - 8.3.2 Aplicaciones de bajo voltaje
 - 8.3.3 Aplicaciones de alto voltaje
 - 8.3.4 Resumen de materiales
 - 8.3.5 Estadísticas de la disrupción eléctrica
-

8.1 PELÍCULAS AISLANTES PARA DISPOSITIVOS DE ESTADO SÓLIDO Y CIRCUITOS INTEGRADOS

Sorab K. Ghandhi

Las películas aislantes se utilizan en la actualidad en muchos puntos de la elaboración de circuitos integrados monolíticos de silicio y arseniuro de galio^{1,2}. Su primer y más importante uso es como máscara protectora contra impurezas durante importantes procesos de manufactura tales como la difusión y la implantación de iones. Por tanto, permiten la introducción silueteada de impurezas al sustrato del semiconductor, de tal forma que es posible fabricar varios dispositivos en un solo chip.

También participan de manera importante en la operación de dispositivos, como el material dieléctrico de los transistores de efecto de campo con compuerta aislada. Asimismo, permiten efectuar conexiones eléctricas entre los dispositivos y sirven como la capa aislante sobre la cual es posible colocar los conductores, y con frecuencia se utilizan varias capas de estas películas en esquemas de interconexión de capas múltiples, que son comunes en la moderna tecnología de los circuitos integrados.

Por último, estas películas tienen una función importante en la protección de todo el microcircuito. Idealmente, su uso puede evitar el costoso encapsulado hermético, dando así por resultado una importante reducción de costos para estos circuitos.

Cada una de las áreas de aplicación mencionadas requiere diferentes propiedades de las películas; además, es posible utilizar muchas técnicas de preparación en cada caso^{3,4}. Debido a consideraciones de espacio, la descripción se limitará a los métodos que se utilizan con mayor frecuencia.

Las películas aislantes nativas se producen por reacción con el conductor en sí⁵. Estas películas, que suelen ser los óxidos de silicio y el arseniuro de galio, están relativamente libres de contaminación; además, pueden formarse *in situ* como parte de otros procesos, como el de difusión. En el caso del silicio, el crecimiento del óxido suele efectuarse sometiendo la oblea o placa de silicio a oxígeno o vapor de agua secos a temperaturas elevadas. Este proceso se efectúa en una cámara de reacción de cuarzo sometida a calentamiento por resistividad en la cual las placas de silicio se sujetan en un carro. El oxígeno o el vapor de agua secos se introducen por la entrada de esta cámara, para dar como resultado la formación de dióxido de silicio amorfo. En las publicaciones especializadas pueden encontrarse los detalles de este tipo de sistema y sus muchas variantes^{1,2}.

Las películas aislantes nativas de arseniuro de galio suelen constar de mezclas de Ga_2O_3 y As_2O_3 , y por lo general se producen por anodización a temperatura ambiente. Con este fin se han utilizado muchas soluciones de anodización.

En la tecnología de los circuitos integrados también se utilizan diversas películas aislantes depositadas. La técnica más común⁶⁻⁸ para generarlas es el depósito

químico en fase de vapor (CVD, del inglés *chemical vapor deposition*), aunque también se utilizan la evaporación al vacío, el bombardeo iónico y técnicas de depósito de iones^{9,10}. Las películas CVD se forman por la reacción química de especies volátiles en sistemas cuyas configuraciones son muy semejantes a las de los sistemas utilizados para la oxidación térmica. Sin embargo, en este caso se utilizan tanto los sistemas de pared caliente calentados por resistencia como los sistemas de pared fría calentados por RF. En muchos casos la formación de la película ocurre a bajas presiones, por lo general del orden de las 10^{-3} atm. Esto incrementa enormemente la velocidad de la difusión gaseosa en el reactor y permite el estrecho apilamiento de obleas a la vez que se mantiene uniforme el grosor de la película.

En otros sistemas CVD se incorpora al proceso de depositación el uso de las especies excitadas electrónicamente, que se crean por medio de un plasma. Los sistemas CVD mejorados con plasma permiten la fabricación de películas aislantes a bajas temperaturas y son de especial utilidad en las etapas finales del procesamiento de circuitos integrados.

Las películas depositadas también pueden formarse mediante centrifugación de un líquido seguida de coqueamiento. De nuevo, la posibilidad de fabricar estas películas a bajas temperaturas las hace atractivas en las fases finales de la manufactura de circuitos integrados.

En el siguiente artículo se analiza la formación de las películas aislantes, así como sus propiedades eléctricas y físicas, y se cita una gran cantidad de referencias.

8.1.1 Óxidos nativos

El óxido de silicio nativo suele producirse haciendo pasar oxígeno o vapor de agua sobre silicio mantenido a una temperatura de 800 a 1 200 °C. El empleo de oxígeno seco da por resultado óxidos densos con una velocidad de crecimiento relativamente lenta^{10,11}. El vapor de agua, que se obtiene haciendo pasar burbujas de oxígeno o nitrógeno por agua desionizada a 95 °C (que corresponde a una presión de vapor de 640 torr o 0.842 atm), da por resultado un óxido un poco menos denso, pero con una velocidad de crecimiento más rápida. Con este objeto suelen utilizarse sistemas pirógenos, en los que se queman hidrógeno y oxígeno, debido a sus ventajas operacionales.

En las figuras 8-1 a 8-4 se muestra la velocidad de crecimiento del óxido como función de la temperatura y de la orientación cristalina.

Como se observa, el crecimiento es mucho más rápido durante la oxidación húmeda². Los sistemas de oxidación a alta presión están actualmente en desarrollo para permitir el crecimiento aún más rápido del óxido o, alternativamente, para que ocurra a temperaturas reducidas^{12,13}.

En la tabla 8-1 se presentan las propiedades del crecimiento de óxidos en condiciones húmedas y secas. Obsérvese que el crecimiento en seco del óxido da por resultado una mejor calidad, con mayores densidad y resistencia dieléctrica, que en el caso del método húmedo.

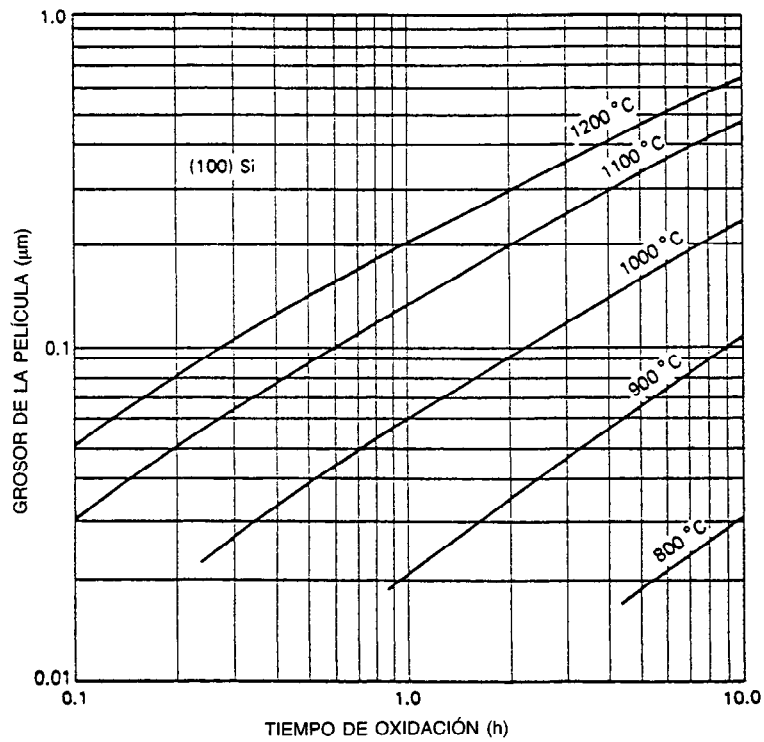


Fig. 8-1. Tasa de crecimiento del óxido para oxígeno seco: (100) silicio.

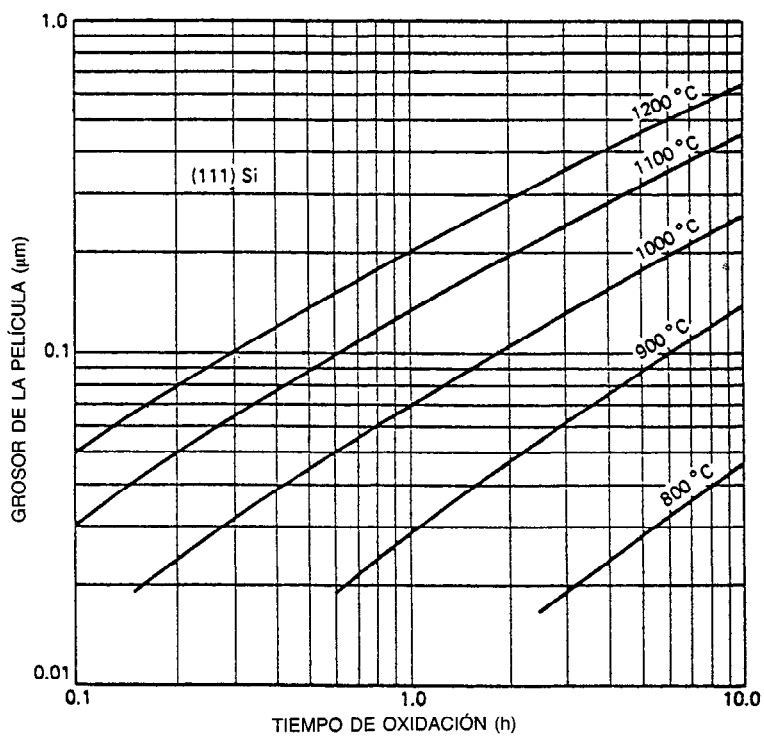


Fig. 8-2. Tasa de crecimiento del óxido para oxígeno seco: (111) silicio.

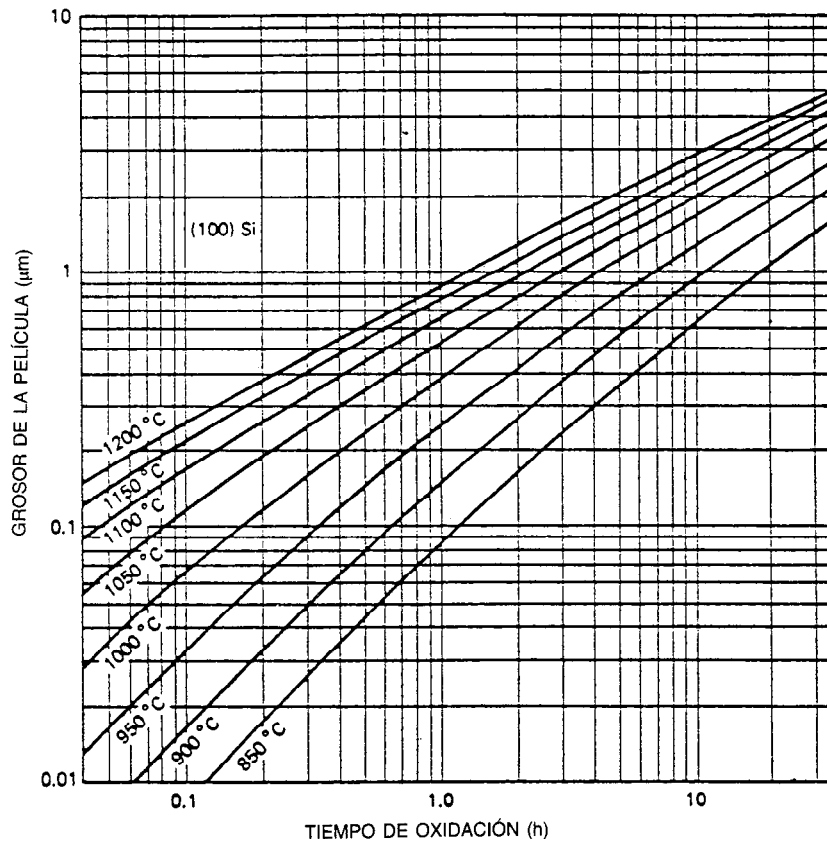


Fig. 8-3. Tasa de crecimiento del óxido para oxígeno húmedo: (100) silicio.

Una propiedad importante de estos óxidos es su capacidad de proteger contra la difusión de impurezas. En la tabla 8-2 se presenta una medida de esta capacidad, así como las difusividades relativas de las impurezas comunes en el silicio. Con excepción del galio, los demás elementos de la tabla 8-2 tienen valores más bajos en unos tres a cuatro órdenes de magnitud para la difusividad de impurezas en el silicio. Por consiguiente, las películas de dióxido de silicio son excelentes protecciones contra la difusión para el boro, fósforo, arsénico y antimonio. En las figuras 8-5 y 8-6 se muestran las propiedades relativas de protección del dióxido de silicio contra las impurezas que se utilizan con mayor frecuencia, el boro y el fósforo. Sin embargo, debe hacerse hincapié en que éstas y otras propiedades del dióxido de silicio dependen en gran medida del método de formación de la película. Por tanto, estas figuras sirven cuando mucho como una aproximación a esta capacidad protectora.

Las propiedades electrónicas de los óxidos formados en seco son superiores a las de los que se forman por oxidación húmeda. Así, se ha demostrado que los óxidos secos tienen menor carga fija de óxido¹⁵, además de menor densidad de carga atrapada en la interficie óxido-silicio que los óxidos húmedos. El empleo de especies halogénicas durante la oxidación ha dado por

resultado mayor reducción de estas cargas y un aumento en la calidad del óxido¹⁶⁻¹⁹. Los óxidos de silicio nativos, desarrollados en condiciones secas, son de esta forma especialmente útiles en la elaboración de óxidos para compuertas de dispositivos a base de semiconductores de metal y óxido (o simplemente dispositivos MOS, del inglés *metal oxide semiconductor*). Los óxidos húmedos, por otro lado, se utilizan en la mayor parte de las aplicaciones generales debido a su crecimiento más rápido.

Los mejores óxidos de arseniuro de galio se desarrollan por medio de técnicas de anodización húmeda, las cuales se efectúan en una celda de anodización a temperatura ambiente. Aquí el arseniuro de galio actúa como ánodo, y como cátodo se utiliza un metal noble como el platino. Se utiliza una amplia gama de electrolitos, y todos dan por resultado una mezcla de Ga_2O_3 y As_2O_3 en proporciones variables.

Los sistemas de electrolitos acuosos para el arseniuro de galio se basan en agua o en mezclas de agua y H_2O_2 (agua oxigenada)^{20,21} con aditivos para el control del pH. Estos aditivos pueden ser ácidos (como H_3PO_4), básicos (como NH_4OH) o neutros [como $(\text{NH}_4)_2\text{HPO}_4$]. A fin de evitar que el óxido se diluya durante su formación, sólo es admisible un pequeño intervalo de pH en cada sistema.

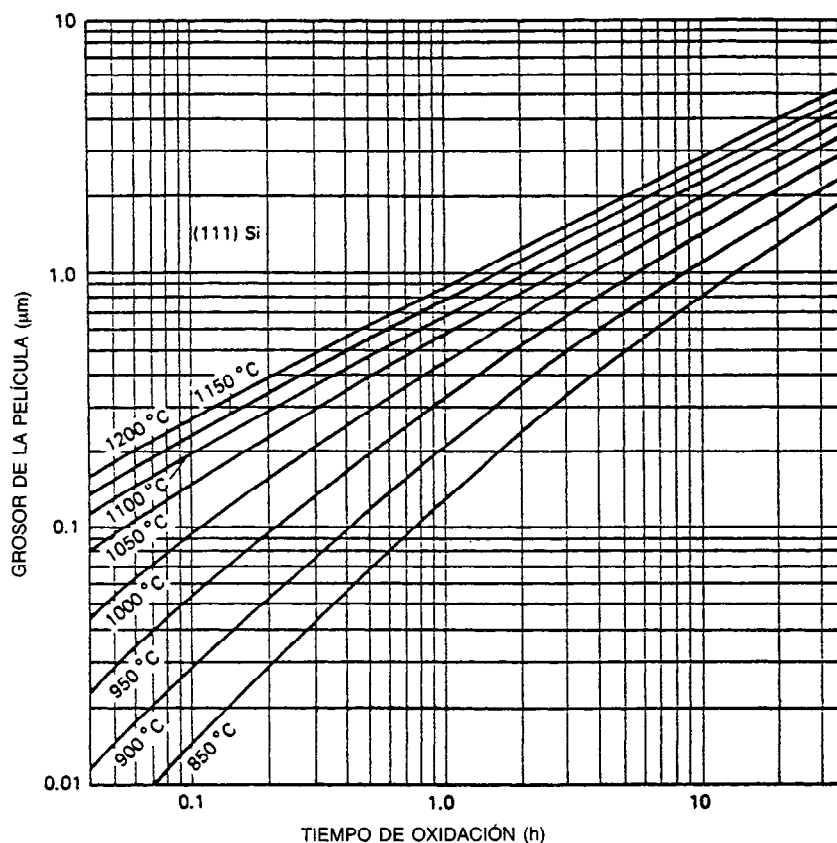


Fig. 8-4. Tasa de crecimiento del óxido para oxígeno húmedo: (111) silicio.

Para este fin también se han utilizado sistemas no acuosos²² y sus características son semejantes a las de los sistemas acuosos.

Recientemente se ha dedicado considerable trabajo a investigar un sistema de anodización de tipo combinado acuoso/no acuoso, empleando una solución de agua y propilén-glicol que contiene ácido tartárico²³. En este caso se utiliza una solución acuosa al 3% de ácido tartárico, con amoníaco para ajustar su pH al intervalo de 5 a 7. A esta solución se agrega propilén-glicol en una relación 1:2 (partes por volumen de solución acuosa/glicol). Esta mezcla da por resultado una tasa de disolución extremadamente baja para el GaAs, típicamente de $4 \text{ a } 6 \times 10^{-2} \text{ Å/s}$, en comparación con las tasas tan elevadas como de 1 a 3 Å/s de los electrolitos acuosos. Los óxidos desarrollados en este sistema se han utilizado con éxito como compuertas en dispositivos MOS²⁴, y también como protecciones de difusión para el arseniuro de galio.

Es posible asimismo desarrollar óxidos de silicio anódicos. Sin embargo, estos óxidos son bastante porosos y no se utilizan actualmente en los procesos de elaboración de dispositivos. Su uso está restringido a aplicaciones diagnósticas en las que deben efectuarse mediciones en capas de silicio eliminadas en forma consecutiva²⁵.

La tabla 8-1 es un breve resumen de las propiedades eléctricas más importantes de las películas de óxidos nativos. El amplio intervalo de parámetros pone de relieve la sensibilidad de estas películas a la técnica de formación que se utilice.

En resumen, las películas aislantes delgadas deben satisfacer varios requerimientos para poder emplearse de manera idónea en microcircuitos. Primordialmente, su resistividad eléctrica debe ser lo suficientemente alta para que la operación del circuito no sea afectada por ellas. Por lo general, la conducción de la mayor parte de los aislantes ocurre por procesos electrónicos, y con comportamiento óhmico, en campos eléctricos por debajo de 10^5 V/cm ; en este caso se considera adecuada una resistividad volumétrica de 10^{12} Ω-cm para la mayor parte de los objetivos prácticos. La conducción en campos intensos se convierte en dominante en el intervalo de 10^6 V/cm y puede presentarse por procesos electrónicos o iónicos. Además, también pueden ocurrir efecto túnel y salto de electrones, con lo que la característica general de conducción asume un comportamiento no óhmico.

Los procesos de conducción en campos intensos son importantes en los óxidos que se utilizan en la región de compuerta de los dispositivos MOS, debido a la extrema delgadez de las capas formadas ($\approx 250 \text{ Å}$).

Tabla 8-1. Propiedades de óxidos nativos de silicio y arseniuro de galio

Método de formación	Densidad (g/cm ³)	Resistividad a 300 K (Ω-cm)	Resistencia dieléctrica (V/cm)
Sílice fundida	2.2	1.5×10^{16}	2.9×10^6
Óxido térmico seco	2.2-2.3	10^{15} - 10^{16}	5.9×10^6
Óxido térmico húmedo	2-2.2	10^{15} - 10^{16}	1.5×10^6
Óxido de silicio anódico	1.8	10^{12} - 10^{15}	1.5×10^6
Óxido anódico (arseniuro de galio)	4.2	10^{14} - 10^{15}	1.5×10^6

A frecuencias elevadas se hacen evidentes los mecanismos de pérdida, que son una función compleja de los fenómenos de captura interfacial y de volumen y que suelen estar directamente relacionados con la técnica de formación del óxido. En este caso los óxidos nativos tienen en general mejor rendimiento que las películas depositadas, debido a su carácter intrínseco más limpio.

Las películas aislantes también deben ser compatibles metalúrgicamente con los metales sustrato y de interconexión. Por tanto, deben adherirse firmemente a ambos y no deben agrietarse ni dañar las capas adyacentes durante los ciclos térmicos. En algunos casos debe utilizarse una capa como almohadilla de algún material intermedio para proporcionar este alivio de esfuerzo. Un caso especial es el empleo de películas de nitruro de silicio como protecciones contra la oxidación. La dilatación térmica notablemente dispar que existe entre el nitruro de silicio y el silicio subyacente provoca la formación de una red de dislocación masiva en el silicio, haciéndolo inadecuado para la operación en dispositivos. A fin de aliviar el esfuerzo creado durante este proceso suele utilizarse una capa de dióxido de silicio de 100 a 200 Å de grosor.

8.1.2 Películas depositadas

Se han investigado varias películas diferentes para su empleo en la tecnología de los circuitos integrados. Sin embargo, las de mayor uso son las de dióxido de silicio, vidrio de fosfosilicato y nitruro de silicio. A continuación se analizará cada una de ellas.

Dióxido de silicio

Este aislante posee muchas propiedades únicas que lo hacen casi ideal para la tecnología de los semiconduc-

tores y los circuitos integrados. Es denso, duro e inerte a muchos agentes químicos que se utilizan durante el procesamiento del dispositivo. Puede usarse como protección en la difusión o implantación de iones contra muchas impurezas comunes, y también como recubrimiento superior para propósitos de recocido después de la implantación. Asimismo, se graba al ácido fácilmente con alta resolución (definición de líneas finas), por lo que puede diseñarse para aplicaciones en microcircuitos. Puede depositarse a temperaturas que varían de 250 a 1 000 °C, de modo que tolera las restricciones térmicas impuestas por procesos previos.

El dióxido de silicio depositado es un excelente aislante, con resistencia específica de alrededor de 10^{14} Ω-cm y rigidez dieléctrica superior a los 10^6 V/cm. Por ello, se utiliza como capa aislante sobre la cual es posible colocar metales de conexión entre los componentes. Su capacidad de adherencia firme con el aluminio (que es el metal de interconexión de mayor uso) lo hace casi ideal para esta aplicación. Por último, sus excelentes propiedades de resistencia a la abrasión lo hacen idóneo como capa protectora de todo el microcircuito.

La técnica más sencilla de depósito que se usa con este material es el método de centrifugación, en el que la oblea de silicio se recubre por rotación sosteniéndola en un mandril al vacío que se hace girar entre 2 500 y 5 000 rpm. A continuación se aplica por goteo una mezcla de dióxido de silicio, a fin de formar una capa delgada (aproximadamente de 5 000 Å) en la oblea de silicio por medio de fuerza centrífuga. Este método puede utilizarse para formar capas razonablemente uniformes, prestando atención especial al control de la viscosidad. Además, las aplicaciones sucesivas pueden permitir que el grosor de estas películas sea hasta de 1 a 2 μm.

Tabla 8-2. Difusividad de impurezas en el dióxido de silicio

Elemento	Difusividad a 1 100 °C (cm ² /s)	Difusividad a 1 200 °C (cm ² /s)
Boro	3×10^{-17} - 2×10^{-14}	2×10^{-16} - 5×10^{-14}
Galio	5.3×10^{-11}	5×10^{-8}
Fósforo	2.9×10^{-16} - 2×10^{-13}	2×10^{-15} - 7.6×10^{-13}
Antimonio	9.9×10^{-17}	1.5×10^{-14}
Arsénico	1.2×10^{-16} - 3.5×10^{-15}	2×10^{-15} - 2.4×10^{-14}

Fuente. Referencia 14, con permiso.

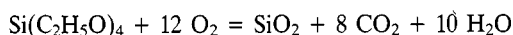
Tabla 8-3. Alcoxisilanos utilizados para el desarrollo de películas de sílice

Material	Fórmula	Peso molecular	Punto de ebullición (°C)	Temperatura de descomposición del vapor (°C)
Tetraetoxisilano	$\text{Si}(\text{OC}_2\text{H}_5)_4$	208	167	728-840
Etiltriethoxisilano	$(\text{C}_2\text{H}_5)_3\text{Si}(\text{OC}_2\text{H}_5)_3$	192	161	650-750
Amiltriethoxisilano	$\text{C}_5\text{H}_{11}\text{Si}(\text{OC}_2\text{H}_5)_3$	234	198	600-740
Viniltriethoxisilano	$\text{CH}_2 = \text{CHSi}(\text{OC}_2\text{H}_5)_3$	190	160	600-700
Feniltriethoxisilano	$\text{C}_6\text{H}_5\text{Si}(\text{OC}_2\text{H}_5)_3$	240	234	610-750
Dimetildietoxisilano	$(\text{CH}_3)_2\text{Si}(\text{OC}_2\text{H}_5)_2$	148	111	760-900

Fuente. Referencia 10, con permiso.

Para este fin existen en el mercado varias fórmulas patentadas. En general, consisten de alcoxisilanos y alquilsiloxanos, con aditivos como butil-carbitol y celulosa para el control de la viscosidad durante la operación de rotación. Una vez aplicadas, y después de secarlas en hornos, estas películas se transforman en dióxido de silicio al hornearlas a unos 250 °C. Cocerlas a temperaturas más elevadas permite su densificación, con una mejora global en su microestructura.

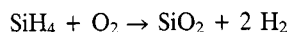
Las películas que se elaboran por centrifugación son útiles como recubrimientos más externos, ya que simplifican el manejo de partes terminadas. Su falta de uniformidad y su porosidad hacen que su uso se restrinja a esta aplicación. Es posible desarrollar películas de mucho mejor calidad por medio de la oxidación pirolítica de varios alcoxisilanos²⁶ en el intervalo de temperaturas de 700 a 800 °C. Uno de estos compuestos es el tetraetilortosilano (TEOS), que se transporta en forma de vapor a la cámara de reacción mediante un mecanismo de burbujeo. La reacción de oxidación se realiza en un sistema CVD de pared fría a 800 °C; dicha reacción es:



En la tabla 8-3 se presentan las propiedades de varios alcoxisilanos que pueden usarse para este fin¹⁰.

En todos los casos, la formación del óxido va acompañada de la formación de una gran cantidad de agua como un subproducto de la reacción, lo cual da por resultado películas con algunas diferencias de calidad. A pesar de ello, las propiedades electrónicas de estas películas son adecuadas para su utilización en varias situaciones en las que debe depositarse un óxido grueso por medio de un proceso a temperatura relativamente baja.

La formación de películas de dióxido de silicio suele efectuarse con mayor frecuencia por medio de la oxidación pirolítica del silano^{27,28}. La reacción del silano ocurre a menores temperaturas que la del alcoxisilano (de 300 a 500 °C), y es la siguiente:



lo que da por resultado películas de sílice de alta calidad. Típicamente se emplea un sistema CVD de pared fría del tipo calentado por resistencia, debido a las bajas temperaturas implicadas. La película que se ob-

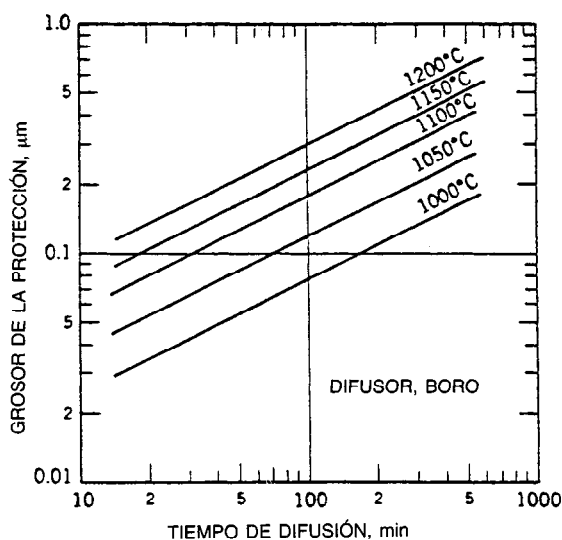


Fig. 8-5. Grosor de la protección para el boro.

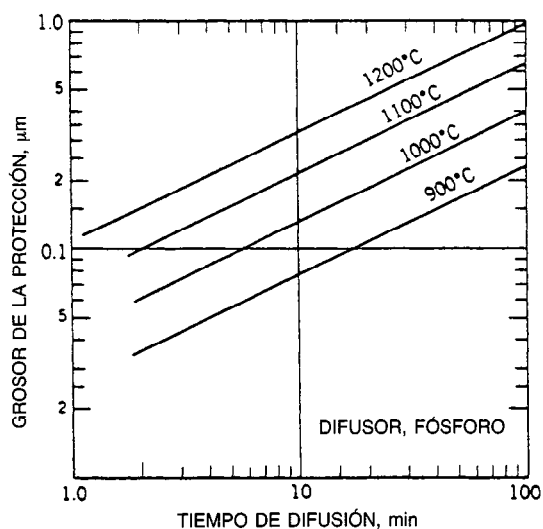


Fig. 8-6. Grosor de la protección para el fósforo.

Tabla 8-4. Propiedades de las películas de dióxido de silicio desarrolladas mediante deposición química en fase de vapor (CVD) y deposición de plasma CVD

Propiedad	CVD	Plasma de CVD
Densidad de poros	1-10 cm ⁻²	≤ 1 cm ⁻²
Densidad de partículas	10-100 cm ⁻²	≤ 0.2 cm ⁻²
Cobertura de la etapa	Deficiente	Conforme
Adhesión al aluminio	Buena	Buena
Resistencia al agrietamiento	Deficiente	Buena
Esfuerzo pelicular intrínseco	De tensión 6-8 × 10 ⁸ dinas/cm ²	De compresión 2 × 10 ⁹ dinas/cm ²
Voltaje de disrupción	8 × 10 ⁶ V/cm	8 × 10 ⁶ V/cm
Resistividad pelicular	> 10 ¹⁷ Ω-cm	> 10 ¹⁷ Ω-cm

tiene posee un esfuerzo integrado de tensión de unas 3×10^9 dinas/cm² para una temperatura de desarrollo de 450 °C.

Los sistemas CVD de baja presión del tipo de pared caliente están siendo utilizados cada vez más para este fin²⁹, y producen películas cada vez más uniformes de oblea a oblea, así como mayor rendimiento. Además, la calidad de la película suele ser superior a la de las películas que se obtienen en un sistema a presión atmosférica, con densidad reducida de poros. Sin embargo, la velocidad de crecimiento es algo menor (de 100 a 150 Å/min, en comparación con los 500 a 600 Å/min de los sistemas a presión atmosférica).

Las películas de dióxido de silicio también pueden desarrollarse a baja presión (aproximadamente de 0.1 a 0.5 torr) y baja temperatura (250 °C) en sistemas mejorados con plasma^{30,31} por medio de reacciones en las que participen mezclas de SiH₄/O₂, SiH₄/CO₂ y SiH₄/N₂O. Con esta técnica se obtiene un esfuerzo integrado de compresión en las películas depositadas y se reduce bastante la tendencia al agrietamiento durante ciclos térmicos subsecuentes. Como resultado, las películas desarrolladas con este procedimiento pueden ser mucho más gruesas que las desarrolladas a presión atmosférica.

En la tabla 8-4 se muestran algunas propiedades selectas de películas depositadas de dióxido de silicio.

Vidrio de fosfosilicato

El vidrio de fosfosilicato (PSG, del inglés *phosphosilicate glass*) se elabora incorporando P₂O₅ en la película de dióxido de silicio (sílice) durante el proceso CVD. Este vidrio es más denso y libre de vacíos que la sílice (dióxido de silicio), y por ello forma mejores capas protectoras³². Además, al incorporar P₂O₅ en las películas de sílice se reduce el esfuerzo tensil integrado de 3×10^9 dinas/cm² para película no impurificada hasta alrededor de 2×10^9 dinas/cm² para películas con el 13% en peso de P₂O₅, y hasta cero para películas con un 20% en peso de P₂O₅. Esto, a su vez, mejora en gran medida su integridad durante los ciclos térmicos.

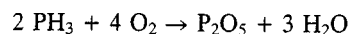
Se obtiene una ventaja más debido a que el coeficiente de dilatación térmica del dióxido de silicio aumenta con rapidez con la incorporación de P₂O₅. Por consiguiente, las películas de PSG pueden adecuarse para un ajuste térmico más idóneo con el semiconductor subyacente. En la figura 8-7 se muestra el coe-

ficiente de dilatación del PSG en función del contenido de P₂O₅, sobre el intervalo de composición de interés práctico³³. En dicha figura también se presentan valores para la sílice no impurificada ($6 \times 10^{-7}/^{\circ}\text{C}$), para el silicio ($2.6 \times 10^{-6}/^{\circ}\text{C}$) y para el arseniuro de galio ($5.9 \times 10^{-6}/^{\circ}\text{C}$).

Se ha encontrado que las capas de PSG son efectivas para inmovilizar iones sodio en la tecnología MOS, y se utilizan ampliamente para esta importante aplicación³⁴. También se emplean para mejorar la estabilidad de dispositivos y microcircuitos bipolares³⁵.

La aplicación de estas películas, seguida de tratamiento térmico, da por resultado la obtención de impurezas metálicas en circuitos integrados lineales y en dispositivos de potencia a base de semiconductor^{36,37}. Por último, se utilizan con frecuencia como recubrimientos en microcircuitos acabados, a fin de proteger la metalización de aluminio contra rayaduras durante el proceso final de adhesión y para proteger de manera permanente contra la inmigración de iones alcalinos. Estas capas de pasivación no deben contener más del 8% en peso de P₂O₅ para impedir reacciones de corrosión con la metalización de aluminio en presencia de humedad³⁸.

Las películas de PSG se desarrollan por medio de la oxidación simultánea de silano y gases fosfamina en el intervalo de 300 a 500 °C, el mismo que para el desarrollo de las películas de sílice. Por ello, su técnica de depósito es una extensión natural del proceso de desarrollo del dióxido de silicio y puede realizarse en el mismo sistema. El gas fosfamina suele suministrarse en una solución al 5 o 10% en argón o nitrógeno para este uso y la reacción de este gas con el oxígeno da por resultado la formación de P₂O₅, que se incorpora como formador de redes en el vidrio resultante. La reacción de oxidación de la fosfamina es:



de modo que se produce una pequeña cantidad de agua como subproducto de este proceso.

Con esta técnica es posible incorporar casi cualquier cantidad de P₂O₅. Sin embargo, las películas se hacen cada vez más higroscópicas, de modo que su contenido de P₂O₅ se limita a alrededor del 2 al 8% en peso para películas que se dejan permanentemente en el producto terminado. Este contenido aumenta hasta el 20% en

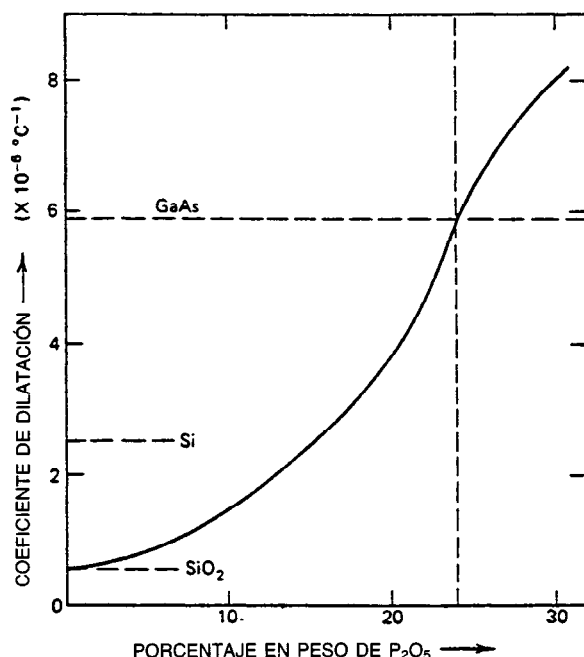


Fig. 8-7. Coeficiente de dilatación térmica de películas de vidrio de fosfosilicato (PSG).

películas que se emplean sólo durante el procesamiento del dispositivo. El porcentaje en peso del P_2O_5 en el vidrio es aproximadamente igual a 1.5 veces la relación molar PH_3/SiH_4 en la fase gaseosa, en este intervalo de composiciones.

Nitruro de silicio

Las películas de este material se utilizan ampliamente tanto en la tecnología de los dispositivos de silicio como en la de los dispositivos de arseniuro de galio. Son más difíciles de desarrollar que las películas de sílice o de PSG, de modo que su empleo está indicado en situaciones en las que imparten mejores propiedades que estos aislantes. Con frecuencia se utilizan junto con películas de sílice para obtener una combinación de características que ninguna de ellas podría proporcionar por sí sola.

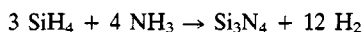
A diferencia del dióxido de silicio, el nitruro de silicio es una excelente barrera contra la inmigración de iones alcalinos, y por ello se utiliza ampliamente en la tecnología MOS como capa de protección. Es superior al PSG debido a que no es higroscópico y a que permite el uso de circuitos no encapsulados en muchas aplicaciones comerciales. También es una protección excelente para la difusión del galio, y por este motivo se utiliza en aplicaciones de dispositivos de potencia. Su capacidad de restringir la difusión del galio permite su uso como material de recubrimiento para el arseniuro de galio durante el proceso de recocido a alta temperatura (900 °C) que debe efectuarse después de la implantación de iones. Es unas 50 a 100 veces más resistente que el silicio a la oxidación térmica, de modo que es posible utilizarlo como protección en varios esque-

mas de circuitos VLSI (de integración a muy grande escala, del inglés *very large scale integration*) a base de silicio, en los que se requiere oxidación selectiva del semiconductor. Se adhiere con facilidad al aluminio metálico por tratamiento térmico, y se emplea como aislante base para la metalización en circuitos integrados de arseniuro de galio. Por último, su resistencia a la radiación es superior a la del dióxido de silicio y por ello se usa en dispositivos endurecidos a la radiación³⁹.

El nitruro de silicio es un aislante de intervalo grande con constante dieléctrica de 5.8 a 6.1, índice de refracción de 1.98 a 2.05 y densidad entre 2.3 y 2.8 g/cm³, y su composición está dada por el Si_3N_4 . Sin embargo, con frecuencia se encuentran desviaciones considerables en su estequiometría durante su depósito por los diversos métodos presentados aquí, y es frecuente encontrar relaciones de los átomos Si/N desde 0.7 hasta 1.1. Aunque es posible desarrollar una película nativa por nitruración directa del silicio, el proceso debe realizarse a temperaturas elevadas (de 1 100 a 1 300 °C) y da por resultado películas extremadamente finas, debido a la lenta difusividad del nitrógeno a través de ellas durante el proceso de desarrollo⁴⁰.

Es posible depositar películas amorfas de nitruro de silicio mediante bombardeo iónico por RF utilizando un objetivo o "blanco" de silicio con una descarga de nitrógeno⁴¹. En este proceso se utilizan temperaturas del sustrato de 200 a 300 °C. Es posible lograr el depósito electrónico directo de un objetivo de nitruro de silicio con un fondo de argón-nitrógeno. Sin embargo, el proceso que se utiliza con mayor frecuencia para el desarrollo de películas es el depósito químico en fase de vapor, en el que participa la reacción de silano ga-

seoso y amoníaco, con nitrógeno gaseoso como diluyente⁴². Esta reacción suele efectuarse a 700 °C, y ocurre de la siguiente forma:



Para este fin es posible utilizar los sistemas ya sea de pared fría o de pared caliente, y el depósito puede lograrse a presión atmosférica o a presión reducida. En todos los casos, la composición y las propiedades de la película son controladas por la relación de amoníaco respecto al silano en la corriente de gas. Son típicas relaciones molares NH_3/SiH_4 de 150 o más para un sistema de pared caliente, con velocidades de crecimiento en el intervalo de 100 a 200 Å/min.

Cada vez más, el depósito de películas se realiza en sistemas de pared caliente a baja presión, que permiten el apilamiento estrecho de obleas, a la vez que se obtiene cobertura uniforme. La operación a baja presión (aprox. 1 torr) se logra reduciendo en gran medida, o eliminando del todo, el uso de gases portadores, de modo que es posible operar estos sistemas con presiones parciales de los reactivos (y tasas de crecimiento) comparables a las que se logran con los sistemas atmosféricos.

Los sistemas de pared fría mejorados con plasma también pueden emplearse para el desarrollo de nitruro de silicio^{43,44}. En este caso se usa un plasma de RF para obtener una o más especies activas de los reactivos, y es posible desarrollar las películas a bajas temperaturas (275 a 300 °C). Esto permite que las películas sean depositadas directamente en los microcircuitos terminados. Las velocidades de crecimiento de estos sistemas son del orden de los 200 Å/min⁴³. Las películas desarrolladas de esta forma tienen grandes cantidades de hidrógeno incorporado, lo que afecta bastante a propiedades tales como la velocidad de grabado o corrosión. Su desarrollo a temperaturas elevadas, o su densificación por tratamiento térmico a 700 °C, da por resultado una notable mejora en la calidad de la película, pero anula la ventaja inicial del desarrollo a baja temperatura para este método.

Las películas depositadas de nitruro de silicio tienen una gran cantidad de esfuerzo tensil integrado (5×10^9 dinas/cm²) cuando se desarrollan a 700 °C, y suelen depositarse en grosores inferiores a 1 000 Å para evitar rotura, desprendimiento o daño del semiconductor subyacente. Por lo general, las películas con relación Si/N de 0.75 presentan esfuerzos mínimos, por lo cual son preferidas.

Las películas desarrolladas por medio de técnicas mejoradas con plasma suelen tener menor esfuerzo (aproximadamente de 2×10^9 dinas/cm²) que las desarrolladas en sistemas de pared caliente; además, este esfuerzo puede ser de tensión o de compresión, dependiendo de las condiciones de desarrollo de la película. Es posible desarrollar películas relativamente gruesas (de 0.5 a 1 µm) con esfuerzo de compresión, y con excelente adhesión con la superficie del semiconductor⁴⁴.

Las propiedades de grabado o corrosión de las películas de nitruro de silicio varían bastante y están

asociadas con su relación Si/N, así como con la forma en que se depositan. El Si_3N_4 puro se corroe a 68 Å/min en HF (49%) aproximadamente, mientras que las películas desarrolladas en CVD o en plasma suelen hacerlo mucho más rápido (250 a 500 Å/min). Con frecuencia se emplea ácido fosfórico en ebullición para corroer de forma selectiva estas películas en presencia de capas de sílice.

8.1.3 Películas orgánicas

Existen dos problemas relacionados con las interconexiones, cuya importancia aumenta con la complejidad de los esquemas de los circuitos integrados. En primer lugar, se hace bastante difícil efectuar dichas interconexiones en forma ordenada y evitar (o minimizar) los cruces; y en segundo lugar, el área ocupada por el metal de interconexión aumenta hasta un punto en que se convierte en el factor dominante en el uso del área del chip. Ambos problemas requieren el empleo de múltiples niveles de interconexión. El lugar común suelen ser tres capas de éstas, aunque en algunos circuitos se requieren más. Con cada nivel, la superficie del circuito se hace cada vez más irregular, y aumenta la altura de la etapa encontrada por el metal de interconexión. Esto da por resultado una pérdida importante de rendimiento durante la fabricación del circuito, debido a la rotura del metal en estas etapas.

Las películas orgánicas se han utilizado con éxito para resolver estos problemas. Dichas películas se colocan como recubrimientos por centrifugación, en forma líquida, y después se alisan las irregularidades mediante tensión superficial. Como resultado, su uso permite muchos niveles de metalización sin un deterioro progresivo de la topología de la superficie⁴⁵⁻⁵⁰.

Estas películas deben satisfacer varios requerimientos para que su aplicación sea exitosa, y su resistividad eléctrica y rigidez dieléctrica a la disrupción deben ser comparables a las del dióxido de silicio. Deben ser capaces de formarse en superficies libres de poros y no deben experimentar alteraciones dimensionales durante el proceso subsecuente de curación, ya que de ser así se dañaría la capa del metal de interconexión sobre la que se encuentran. Deben adherirse firmemente a la superficie en que estén y proporcionar adhesión a la capa siguiente de metalización. Por tanto, deben adherirse a la primera capa aislante de dióxido de silicio en la superficie de la tableta, al metal de aluminio, que es la elección común para las interconexiones, y también al óxido de aluminio que se forma sobre el metal como un óxido nativo delgado. Por último, estas películas deben ser capaces de resistir los tratamientos térmicos que se requieren en los procesos consecutivos de manufactura, como formación de interconexiones entre capas consecutivas, adhesión de la matriz con el sustrato y adhesión del alambre con las terminales del paquete.

Algunos polímeros orgánicos que se emplean con este fin son los cauchos (hules) sintéticos, como el cis-poliisopreno cíclico, resinas de epoxifenol y silicones. Sin embargo, se ha demostrado que las poliimidas son los materiales más estables térmicamente, y el trabajo se ha concentrado en su aplicación a los microcircuitos.

En el mercado existen varias formulaciones de poliimida. La poliimida isoindroquinazolinodiona (PIQ) es una de las más estables, ya que contiene el anillo de isoindroquinazolinodiona, térmicamente estable⁴⁶. Este material presenta estabilidad dimensional cuando es procesado a temperaturas tan elevadas como 450 °C durante períodos tan largos como una hora. En contraste, no es posible utilizar materiales como los silicones en los procesos sometidos a temperaturas superiores a 250 °C durante este período⁴⁷.

En la práctica, las resinas de poliimida se diluyen con solventes extra como metilpirolidinona o dimetilformamida, para que sea posible recubrirlas por centrifugación en capas de 1 a 4 µm de grosor. Después, las películas se curan mediante tratamientos térmicos en el intervalo de temperaturas de 300 a 400 °C. Mediante agentes corrosivos húmedos como la hidracina se perforan orificios para efectuar las interconexiones entre estas películas.

Sin embargo, para las aplicaciones en circuitos de alta densidad se prefiere el método de la corrosión por gases utilizando un plasma de oxígeno, ya que de esta forma se evitan problemas asociados con la acción capilar del líquido corrosivo.

Las películas de poliimida curadas de forma apropiada suelen tener una rigidez dieléctrica de campo de 4×10^6 V/cm, que es comparable a la del dióxido de silicio. Sin embargo, su resistividad eléctrica es casi 20 veces inferior a la del dióxido de silicio, pero esto no representa problema alguno a los niveles de voltaje que se encuentran en los microcircuitos⁴⁸.

Tal vez el inconveniente más importante de estas películas es la adhesión a la primera capa aislante de dióxido de silicio, así como a la metalización. Típicamente esta adhesión se ha mejorado mediante el empleo de varios agentes de acoplamiento, de los cuales el más popular es el hexametildisilazano (HMDS). Sin embargo, recientemente se ha descubierto que son superiores los nuevos materiales de acoplamiento basados en el empleo de quelatos orgánicos de aluminio⁴⁹. El empleo de estos materiales ha permitido la preservación de altas resistencias al desprendimiento (bastante mayores de 200 g/cm), inclusive cuando los circuitos se han mantenido en condiciones de altas temperatura y humedad (120 °C con una presión de 2 atm de vapor de agua).

REFERENCIAS BIBLIOGRÁFICAS

1. R. A. Colclaser, *Microelectronics: Processing and Device Design*, Wiley, Nueva York.
2. S. K. Ghandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*, Wiley, Nueva York.
3. L. I. Maissel y R. Glang, *Handbook of Thin Film Technology*, Mc Graw-Hill, Nueva York.
4. J. L. Vossen y W. Kern, eds., *Thin Film Processes*, Academic, Nueva York.
5. C. J. Frosh y L. Derick, "Surface Protection and Selective Masking During Diffusion in Silicon", *J. Electrochem. Soc.* **104**:547.
6. B. Mattson, "CVD Films for Interlayer Dielectrics", *Solid State Tech.*, p. 60.
7. W. Kern y R. S. Rosler, "Advances in Deposition Processes for Passivation Films", *J. Vac. Sci. Tech.* **14**:1082.
8. W. Kern, G. L. Schnable y A. W. Fisher, "CVD Glass Films for Passivation of Silicon Devices: Preparation, Composition, and Stress Properties", *RCA Rev.* **37**:3.
9. H. V. Schreiber y E. Froschle, "High Quality RF-Sputtered Silicon Dioxide Layers", *J. Electrochem. Soc.* **123**:30.
10. R. M. Burger y R. P. Donovan, *Fundamentals of Silicon Integrated Circuit Device Technology*, Vol. 1, Prentice-Hall, Englewood Cliffs, NJ.
11. B. E. Deal, "The Oxidation of Silicon in Dry Oxygen, Wet Oxygen, and Steam", *J. Electrochem. Soc.* **110**:527.
12. J. R. Ligenza, "Oxidation of Silicon by High Pressure Steam", *J. Electrochem. Soc.* **109**:73.
13. R. J. Zeto, C. G. Thornton, E. Hryckowian y C. D. Bosco, "Low Temperature Thermal Oxidation of Silicon by Dry Oxygen Pressure Above 1 Atm", *J. Electrochem. Soc.* **122**:1409.
14. M. Ghezzi y D. M. Brown, "Diffusivity Summary of B, Ga, P, As, and Sb in SiO₂", *J. Electrochem. Soc.* **120**:146.
15. E. H. Nicollian y J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology*, Wiley, Nueva York.
16. A. Rohatgi, S. R. Butler y F. J. Feigl, "Mobile Sodium Ion Passivation in HCl Oxides", *J. Electrochem. Soc.* **126**:149.
17. B. E. Deal, "Thermal Oxidation Kinetics of Silicon in Pyrogenic H₂O and 5% HCl/H₂O Mixtures", *J. Electrochem. Soc.* **125**:576.
18. B. R. Singh y P. Balk, "Thermal Oxidation of Silicon in O₂-Trichloroethylene", *J. Electrochem. Soc.* **126**:1288.
19. T. Hattori, "Elimination of Stacking Faults in Silicon by Trichloroethylene Oxidation", *J. Electrochem. Soc.* **123**:945.
20. R. A. Logan, B. Schwartz y W. J. Sundburg, "The Anodic Oxidation of GaAs in Aqueous H₂O₂ Solution", *J. Electrochem. Soc.* **120**:1385.
21. B. Schwartz, F. Ermanis y M. H. Brastad, "The Anodization of GaAs and GaP in Aqueous Solution", *J. Electrochem. Soc.* **123**:1089.
22. B. N. Arora y M. G. Bidnukar, "Anodic Oxidation of Gallium Arsenide", *Sol. State Electron.* **19**:657.
23. H. Hasegawa y H. L. Hartnagel, "Anodic Oxidation of GaAs in Mixed Solutions of Glycol and Water", *J. Electrochem. Soc.* **123**:713.
24. H. Tokuda, Y. Adachi y T. Ikoma, "Microwave Capability of 1.5 Micron-Gate GaAs MOSFET", *Electron. Lett.* **13**:761.
25. H. D. Barber, H. B. Lo y J. E. Jones, "Repeated Removal of Thin Layers of Silicon by Anodic Oxidation", *J. Electrochem. Soc.* **123**:1404.
26. C. R. Barnes y C. R. Geesner, "Pyrolytic Deposition of Silicon Dioxide for 600 °C. Thin Film Capacitors", *J. Electrochem. Soc.* **110**:361.

27. N. Goldsmith y W. Kern, "The Deposition of Vitreous Silicon Dioxide from Silane", *RCA Rev.* **28**:153.
28. B. J. Baliga y S. K. Ghandhi, "Growth of Silica and Phosphosilicate Films", *J. Appl. Phys.* **44**:990.
29. R. S. Rosler, "Low Pressure CVD Production Processes for Poly, Nitride, and Oxide", *Solid State Technol.*
30. R. S. Rosler y G. M. Engle, "Plasma Enhanced CVD in a Novel LPCVD-Type System", *Solid State Technol.*
31. P. G. Evert y T. van de Ven, "Plasma Deposition of Silicon Dioxide and Silicon Nitride Films", *Solid State Technol.*
32. M. M. Schlacter, E. S. Schlegel, R. S. Kan, R. A. Lathlaen y G. L. Schnable, "Advantages of Vapor-Plated Phosphosilicate Glass Films in Large Scale Integrated Circuit Arrays", *IEEE Trans. Electron Dev.* **ED-17**:1077.
33. B. J. Baliga y S. K. Ghandhi, "Lateral Diffusion of Zink and Tin in Gallium Arsenide", *IEEE Trans. Electron Dev.* **ED-21**:410.
34. E. Yon, W. H. Ko y A. B. Kuper, "Sodium Distribution in Thermal Oxide by Radiochemical and MOS Analysis", *IEEE Trans. Electron Dev.* **ED-13**:276.
35. M. Yamin, "Observation of Phosphorus Stabilized SiO₂ Films", *IEEE Trans. Electron Dev.* **ED-13**:256.
36. S. P. Murarka, "A Study of the Phosphorus Gettering of Gold in Silicon by the Use of Neutron Activation Analysis", *J. Electrochem. Soc.* **123**:765.
37. S. K. Ghandhi, *Semiconductor Power Devices*, Wiley, Nueva York.
38. R. B. Comizzoli, "Aluminum Corrosion in the Presence of Phosphosilicate Glass and Moisture", *RCA Rev.* **37**:483.
39. J. A. Appels, E. Kooi, M. M. Paffen, J. J. H. Schatorje y W. H. C. G. Verkuylen, "Local Oxidation of Silicon and Its Application in Semiconductor Device Technology", *Philips Res. Repts.* **25**:118.
40. T. Ito, S. Hijiya, T. Nozaki, H. Arakawa, M. Shinoda y Y. Fukukawa, "Very Thin Silicon Nitride Films Grown by Direct Thermal Reaction with Nitrogen", *J. Electrochem. Soc.* **125**:448.
41. G. J. Kominiak, "Silicon Nitride Films by Direct RF Sputter Deposition", *J. Electrochem. Soc.* **122**:1271.
42. R. Ginsburgh, D. L. Heald y R. C. Neville, "Silicon Nitride Chemical Vapor Deposition in a Hot Wall Diffusion System", *J. Electrochem. Soc.* **125**:1557.
43. R. S. Rosler, W. C. Bensing y J. Baldo, "A Production Reactor for Low Temperature Plasma-Enhanced Silicon-Nitride Deposition", *Solid State Technol.*
44. A. K. Sinha, H. J. Levinstein, T. E. Smith, G. Quintana y S. E. Haszko, "Reactive Plasma Deposited Si-N Films for MOS-LSI Passivation", *J. Electrochem. Soc.* **125**:60.
45. K. Kato, S. Harada, A. Saiki, T. Kimura y T. Okubo, "A Novel Planar Multilevel Interconnection Technology Utilizing Polyimide", *IEEE Trans. Part, Hybrids Packag. PHP-9*:176.
46. Y. Homma, H. Nozawa y S. Harada, "Polyimide Lifftoff Technology for High-Density LSI Metallization", *IEEE Trans. Electron Dev.* **ED-28**:522.
47. A. Saiki, S. Harada, T. Okubo, K. Makai y T. Kimura, "A New Transistor with Two-Level Metal Electrodes", *J. Electrochem. Soc.* **124**:1619.
48. L. B. Rothman, "Properties of Thin Polyimide Films", *J. Electrochem. Soc.* **127**:2216.
49. A. Saiki y S. Harada, "New Coupling Method for Polyimide Adhesion to LSI Surface", *J. Electrochem. Soc.* **129**:2278.
50. L. B. Rothman, "Process for Forming Passivated Metal Interconnection System with a Planar Surface", *J. Electrochem. Soc.* **130**:1131.

8.2. MATERIALES DIELECTRICOS PARA ALMACENAMIENTO DE CARGA

J. Keith Nelson

8.2.1 Clasificación de los dieléctricos

La introducción de un material dieléctrico de permitividad relativa ϵ_r entre las placas de un capacitor paralelo plano de área A y separación d incrementa la capacitancia y el almacenamiento de energía en un factor ϵ_r hasta un valor de $\epsilon_0 \epsilon_r A/d$ coulombs, donde ϵ_r es la permitividad del espacio libre (8.85×10^{-12} F/m). El incremento en el almacenamiento de carga es posible debido a que algunas cargas libres se neutralizan por la acción de dipolos electrostáticos en el material que se introduce. En campos aplicados es posible considerar que dichos dipolos, formados por la separación de cargas a escala molecular o atómica, se autoalinean en el campo eléctrico. La naturaleza de los dipolos influye profundamente en el comportamiento del material dieléctrico y lleva a clasificar los materiales en algunos grupos amplios¹:

1. **No polares.** Son materiales cuyas estructuras y enlaces moleculares son tales que no existen grupos químicos dipolares. La única fuente de polarización es electrónica y se induce cuando el centro de carga de los electrones se desplaza del núcleo. Además de la capacidad de polarización electrónica, algunos materiales pueden contener grupos dipolares de átomos cuyo momento dipolar neto es igual a cero, pero que presentan capacidad de polarización iónica *inducida* en un campo aplicado. Algunos ejemplos de materiales que sólo poseen capacidad de polarización electrónica son el hidrógeno, nitrógeno, benceno, polietileno y poliestireno, que suelen tener constantes dieléctricas en el intervalo de 1 a 4 y en los que es posible probar la dominancia de la capacidad de polarización electrónica observando que $\epsilon_r = n^2$, donde n es el índice de refracción. En materiales

como dióxido de carbono, tetracloruro de carbono, dióxido de titanio y en los cristales de haluros alcalinos se observa una capacidad extra de polarización iónica (o atómica) inducida.

2. **Dipolares.** Además de contar con los mecanismos anteriores de polarización *inducida*, las sustancias que contienen grupos *dipolares* permanentes que pueden orientarse en un campo aplicado poseen una fuente importante extra de capacidad de polarización, lo cual hace que los dieléctricos tengan una permitividad relativa mucho mayor. A diferencia de las formas de polarización *inducida*, esta capacidad de polarización se reduce con la temperatura, debido al obstáculo de la agitación térmica. Los líquidos dipolares, como agua y nitrobenzono, pueden tener constantes dieléctricas relativas hasta de 100.

3. **Ferroeléctricos.** Por analogía con el ferromagnetismo, los materiales ferroeléctricos se polarizan de manera espontánea y tienen dominios de alrededor de $1\ \mu\text{m}$ a través de los cuales todos los dipolos permanentes están orientados en la misma dirección.

Este ordenamiento interno por debajo de la temperatura de Curie da por resultado constantes dieléctricas hasta de 20 000. Muchas formulaciones ferroeléctricas se han basado en el titanato de bario (BaTiO_3) y en el fosfato diácido de potasio y materiales asociados.

8.2.2 Aspectos fundamentales

Los materiales dieléctricos reales no son aislantes perfectos, sino que también experimentan pérdidas eléctricas. Por tanto, un dieléctrico de capacitor con pérdidas puede representarse por medio del circuito equivalente simplificado que se muestra en la figura 8-8a, donde C_0 está libre de pérdida y R_0 disipa energía para simular la

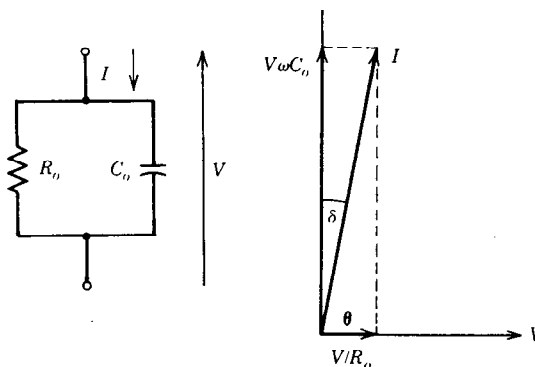


Fig. 8-8. Representación esquemática de un dieléctrico con pérdidas. a) Circuito equivalente. b) Diagrama fasorial.

naturaleza no ideal del capacitor. En condiciones de CD las pérdidas se explican principalmente debido a las conductividades electrónica y iónica. Sin embargo, para campos eléctricos que varían en el tiempo, la energía también se disipará en el proceso de orientación del dipolo (o pérdida por histéresis para los materiales ferroeléctricos). Como elemento de un circuito, la figura 8-8a tendrá el diagrama fasorial de la figura 8-8b, a partir del cual es evidente que R_0 ha introducido un ángulo de fase, descriptivo de las pérdidas. Como θ mide casi 90° , se suelen expresar las pérdidas en términos del ángulo $\delta = 90^\circ - \theta$, donde:

$$\tan \delta = \frac{1}{\omega C_0 R_0}$$

para un campo con frecuencia angular ω . Alternativamente, la pérdida puede incorporarse a la permitividad relativa empleando términos reales e imaginarios:

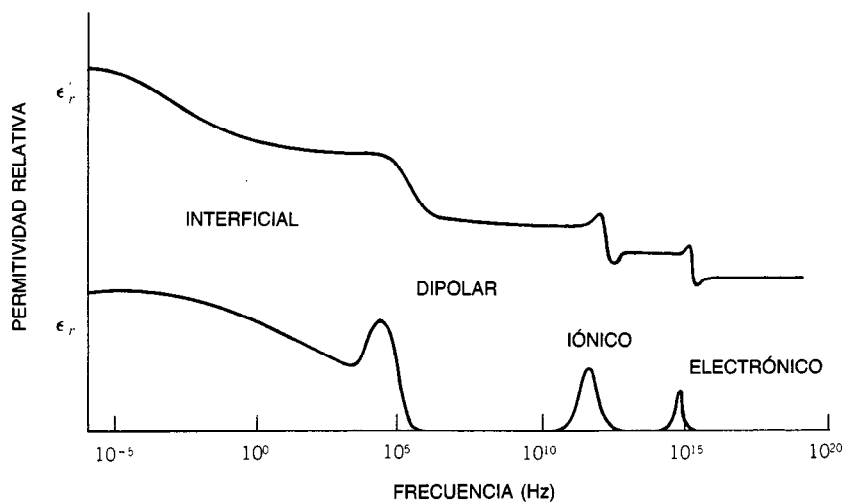


Fig. 8-9. Representación idealizada de la dependencia de la permitividad dieléctrica con respecto a la frecuencia para un intervalo de posibles mecanismos de polarización.

$$\epsilon_r^* = \epsilon_r' - j\epsilon_r''$$

de modo que es posible considerar a $\tan \delta = \epsilon_r''/\epsilon_r'$ la relación de magnitudes de las componentes en fase y en cuadratura (desfasada) de la corriente.

La reorientación del dipolo ocurre con un tiempo de relajación τ , característico de la estructura molecular. Por tanto, la respuesta dieléctrica a un campo eléctrico sinusoidal depende de la frecuencia y, para un sistema sencillo, está dada por las bien conocidas ecuaciones de Debye:

$$\epsilon_r' = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + \omega^2\tau^2}$$

$$\epsilon_r'' = \frac{\omega\tau(\epsilon_s - \epsilon_\infty)}{1 + \omega^2\tau^2}$$

donde ϵ_s es la permitividad estática (CD) y ϵ_∞ es la permitividad a muy alta frecuencia a la que los dipolos están efectivamente "congelados" y no son capaces de contribuir a la constante dieléctrica.

En general, a medida que se incrementa la frecuencia es posible medir varios picos de pérdidas, correspondientes a la operación del mecanismo particular de polarización. Ocurren reducciones concomitantes en la constante dieléctrica hasta que se verifica la última dispersión a frecuencias ópticas debido a la polarización electrónica inducida, como se muestra en la figura 8-9.

8.2.3 Falla de materiales dieléctricos sólidos

Dado que la densidad de energía de un capacitor es proporcional a $\epsilon_r E^2$, evidentemente un incremento en el esfuerzo de trabajo o de fatiga E es una forma atractiva de mejorar el almacenamiento de energía. A pesar de conocidas afirmaciones opuestas, los materiales sólidos no tienen una rigidez dieléctrica bien definida. Muchos materiales poliméricos polares a bajas temperaturas pueden presentar rigidez dieléctrica superior a 10 MV/cm, como se muestra en la figura 8-10, aunque en circunstancias diferentes pueden fallar a esfuerzos 100 veces inferiores. En la raíz de este hecho se encuentra tanto el material como su entorno. La falla de la mayor parte de los materiales en el uso industrial está dominada por imperfecciones físicas e impurezas químicas, y no por el material en sí. Algunos ejemplos de problemas de esta categoría que son bastante difíciles de eliminar por completo los constituyen las microgrietas y fisuras que se forman durante el vaciado o la extrusión de un material, así como los contaminantes iónicos. Aun cuando parámetros ambientales, como temperatura y humedad, influyen en gran medida en la resistencia eléctrica, la forma precisa en que se aplica el esfuerzo dieléctrico también puede ser decisiva para determinar la falla².

El área sometida a esfuerzo, la distribución del campo y la forma de onda del voltaje influirán en el voltaje de resistencia y la disrupción de un dieléctrico altamente esforzado.

Ciertamente el mecanismo de falla cambia a menudo con las circunstancias, y la predominancia de procesos

electrónicos, térmicos, de derivaciones múltiples, de descarga parcial o electroquímicos determina la rigidez dieléctrica de un sólido³.

8.2.4 Aplicaciones de los capacitores

La elección de los materiales y la construcción de capacitores son muy específicas, dependiendo de la aplicación. Aun cuando evidentemente deben preferirse materiales que tengan altas permitividad relativa y resistencia dieléctrica (o aisladora), las aplicaciones que implican elevadas frecuencia o temperatura o las situaciones de baja pérdida muchas veces hacen que sea de suma importancia elegir un dieléctrico con $\tan \delta$ baja.

Aunque se practican otras construcciones, como la de película apilada, por lo general los capacitores pelliculares se manufacturan a partir de un dieléctrico que esté metalizado o que se encuentre entre dos láminas metálicas y enrollado de manera apretada a fin de formar una estructura compacta. Las unidades se encapsulan por moldeo o inmersión en resina con polipropileno o materiales alquídicos. Las unidades de mayor voltaje se impregnan con un fluido para evitar el daño por descargas parciales en los bordes de las láminas o en los vacíos interiores, por lo que siempre se sellan herméticamente en recipientes ("botes") de aluminio. La elección del fluido de impregnación es importante. Después de proibirse el empleo de los policlorobifenilos por razones ambientales, los fabricantes centraron su atención en los aceites minerales de hidrocarburos, aceites sintéticos (como el dodecil-benceno o el fenil-xililetano), ésteres (como el dioctilftalato) o fluidos de silicón⁴. De nuevo, la elección depende de la aplicación y de la naturaleza de la principal película dieléctrica que se utilice. Como parte del esfuerzo por lograr un mejor equilibrio en las propiedades se han creado varias mezclas de fluidos. Por ejemplo, un incremento en el contenido aromático mejora las propiedades para absorber gases, en tanto que algunas formulaciones pueden requerir la adición de estabilizadores químicos.

El papel impregnado se utilizó ampliamente en el pasado como dieléctrico de capacitores, pero dado que su esfuerzo de trabajo está limitado a aproximadamente 20 V/ μ m y su tangente de pérdida típica es 0.003, ha habido un cambio gradual a los materiales poliméricos, que superan al papel y en algunos casos también toleran temperaturas de servicio muy superiores al límite de 100 °C del papel. Para aplicaciones especializadas, como el equipo de radiotransmisión, la mica presenta un intervalo atractivo de propiedades para capacitores de alta frecuencia y baja pérdida. La permitividad relativa de la mica moscovita es de 6 a 7, su tangente de pérdidas varía de 10^{-3} a 10^{-4} y su resistencia dieléctrica supera con mucho los 100 V/ μ m, a lo que se suma su estabilidad química hasta de 500 °C aproximadamente. Los capacitores de mica de precisión también pueden producir estabilidad de un 1% sobre largos períodos, pero su manufactura es relativamente difícil para tolerancias más pequeñas.

Los dieléctricos de película polimérica sintética se han convertido en los materiales de elección para mu-

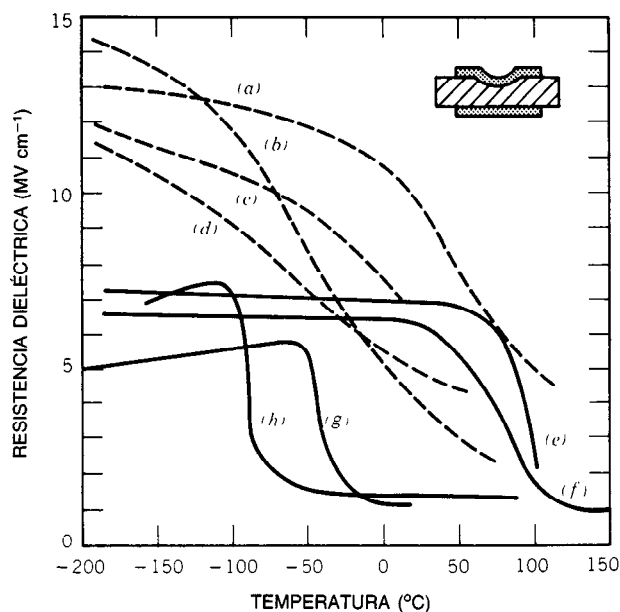


Fig. 8-10. Ejemplos de la dependencia con respecto a la temperatura de la resistencia dieléctrica para CD de un intervalo de polímeros polares (líneas discontinuas) y no polares (líneas continuas). *a)* Polimetilmetacrilato. *b)* Alcohol de polivinilo. *c)* Cloroacetato de polivinilo. *d)* Polietileno clorado al 55%. *e)* Poliestireno atáctico. *f)* Polietileno de baja densidad. *g)* Poliisobutileno. *h)* Polibutadieno. (Reimpreso con permiso de Nelson².)

chas aplicaciones. En la tabla 8-5 se presentan las propiedades relativas aproximadas de películas candidatas, la mayor parte de las cuales existen como películas grado eléctrico libre de poros en el intervalo de grosor de 1 a 10 μm . Con base en la tabla es evidente por qué el polipropileno ha pasado a ser un material de uso común para diversos propósitos, ya que imparte propiedades dieléctricas razonables a bajo costo y es capaz de soportar un esfuerzo de 65 V/ μm a 70 °C para una duración de 20 años.

En general, las películas aptas para una alta temperatura de servicio (como las de poliimida y de polisul-

fona) imponen una gran carga económica. En algunos casos las pérdidas dieléctricas son muy sensibles a la temperatura y la frecuencia. Los diseños de capacitores en los que se utilizan películas tales como las de fluoruro de polivinilideno corren el riesgo de falla por calentamiento en operación con CA sostenida a temperatura elevada, pero son atractivos para trabajos de descarga intermitente.

Cuando las densidades de corriente requeridas son bajas, el empleo de una capa de vapor de aluminio de 300 a 500 Å depositada como electrodo sobre el polímero puede elevar la eficiencia volumétrica hasta

Tabla 8-5. Propiedades representativas de materiales poliméricos para dieléctricos de capacitores

Material	Constante dieléctrica	Factor de disipación (%)	Temperatura de servicio (°C)	Costo relativo ^a
Policarbonato	3.0	0.1-1.0	110	7
Tereftalato de polietileno	3.2	0.2-2.0	85	3.6
Poliimida	3.5	0.2	240	68
Polipropileno	2.2	0.01-0.03	90	1.8
Poliestireno	2.5	0.01-0.1	75	1.3
Politetrafluoroetileno	2.0	0.003-0.025	250	21
Polisulfona	3.1	0.1-0.4	150	23
Fluoruro de polivinilideno ^b	11	0.01-0.12	160	

Fuente. Referencia 4, con permiso.

^a Con respecto al polietileno de baja densidad = 1.

^b No disponible comercialmente en la actualidad como película de grado (calidad) para capacitores.

aproximadamente $1.5 \mu\text{F}/\text{cm}^3$. Por medio del control cuidadoso del grosor de la capa metalizada (de 1 a $5 \Omega/\text{cuad.}$), la unidad del capacitor puede hacerse autorreparable⁵. La falla local provocada por alguna imperfección en la película dieléctrica vaporizará la metalización y así aliviará el esfuerzo eléctrico para permitir la recuperación a expensas de alguna pérdida en la capacitancia.

Además de los polímeros sintéticos, las cerámicas forman la base de la mayor parte de la industria de los capacitores para aplicaciones de señales y de baja potencia. Los capacitores de cerámica suelen estar formados por mezclas sinterizadas de materiales inorgánicos a temperaturas hasta de 1800°C . La estructura policristalina resultante también puede elaborarse como un dispositivo de múltiples capas con electrodos integrados, formados por la impresión con una tinta o pasta de algún metal precioso.

Los capacitores de cerámica basados en materiales de baja pérdida, como el silicato de magnesio (MgSiO_3), con aditivos de unión han encontrado uso en aplicaciones de circuitos resonantes, dado que pueden presentar alta estabilidad y un coeficiente térmico positivo de unas $100 \text{ ppm}/^\circ\text{C}$. Las formulaciones basadas en el dióxido de titanio (TiO_2) o sus derivados pueden ajustarse para adecuar los coeficientes térmicos al intervalo de $+100$ a $-1500 \text{ ppm}/^\circ\text{C}$, pero sus permitividades son menores de 300. Las cerámicas ferroeléctricas pueden presentar permitividades relativas tan elevadas como 15 000, y se utilizan ampliamente en la fabricación de capacitores de alto valor. Sin embargo, los materiales ferroeléctricos tienen bastantes pérdidas inherentes debido a su histéresis ferroeléctrica. El titanato de bario (BaTiO_3), por ejemplo, posee un factor de disipación de 0.02 y una constante dieléctrica aproximadamente de 1 500, que es sensible tanto a la frecuencia como a la temperatura en el intervalo del punto de Curie a los 120°C . Se han desarrollado formulaciones más estables basadas en sustituciones en la estructura del BaTiO_3 , y con cerámicas multifásicas, como el piezoeléctrico circonato-titanato de plomo $\text{PbZrO}_3\text{-PbTiO}_3$ modificado por lantano (PLZT), se han producido materiales con constante dieléctrica que dependen de menos de la intensidad del campo. La falla de los capacitores de cerámica se debe con frecuencia a la formación de grietas y huecos y a la deslaminación del electrodo, por lo que no se recomiendan para trabajos de descarga de pulsos. La resistencia dieléctrica de las cerámicas es limitada en última instancia por la porosidad, aunque la migración iónica y el ingreso de humedad también pueden provocar fallas prematuras.

Para aplicaciones como el suavizado del suministro de energía, en las que se requiere gran capacitancia y las pérdidas tienen importancia secundaria, es posible obtener densidades de alta energía mediante el uso de la tecnología electrolítica. Típicamente es posible formar capas de Al_2O_3 ($\epsilon_r = 8.5$) o de Ta_2O_5 ($\epsilon_r = 25$) tan delgadas como $0.02 \mu\text{m}$ en una lámina grabada de aluminio o tantalio mediante un proceso de anodización. La lámina se enrolla junto con un separador y un electrolito sólido o líquido idóneo [p. ej., MnO_2 formado por pirólisis *in situ* de $\text{Mn}(\text{NO}_3)_2$] para formar

un montaje compacto. La capa activa del dieléctrico es autorreparable, dado que las imperfecciones de la película se reconstruyen de manera continua por la corrosión del metal de apoyo. En consecuencia, tales películas pueden operar a esfuerzos superiores a 10^8 V/m , que se encuentran cercanos al valor de falla. De cualquier forma es difícil fabricar películas coherentes más gruesas de $0.8 \mu\text{m}$, y la tecnología impone la restricción de una polaridad de voltaje fija.

Un concepto que está surgiendo para el almacenamiento de carga, y el cual tiene algunas semejanzas con la tecnología electrolítica, implica la utilización de una doble capa de Helmholtz como capacitor. Tales capas, que separan cargas en la interficie de un electrodo polarizable y un electrolito, miden unos cuantos angstroms de anchura, por lo que son capaces de generar capacitancias muy grandes siempre que el voltaje se limite a aproximadamente 1 V, lo que es determinado por el potencial de descomposición del electrolito. Los primeros dispositivos comerciales⁶ en los que se utiliza este principio mediante un sistema de carbón activado/ácido sulfúrico permiten un considerable ahorro en cuanto a espacio respecto a los dispositivos electrolíticos ordinarios, y presentan algunos atractivos como sustitutos de las baterías. Uno de los problemas primordiales en los dispositivos con capas de Helmholtz, dada su configuración actual, es la alta impedancia interna; una unidad típica de 5 V y 1 F podría presentar una resistencia de fuente equivalente de 5Ω .

8.2.5 Otras aplicaciones

La aplicación de campos eléctricos a materiales dieléctricos a temperaturas elevadas permite el movimiento de cargas. La reducción ulterior de la temperatura en el caso de sustancias de baja pérdida inmoviliza las cargas y produce un material con polarización permanente conocido como electreto. Algunos materiales comúnmente utilizados para los electretos poliméricos son: tereftalato de polietileno, etilén-propileno fluorado, politetrafluoroetileno y polietileno. Las especies cargadas responsables del fenómeno pueden ser transportadores atrapados o dipolos moleculares orientados, pero un electreto es capaz de producir un campo eléctrico de alta duración análogo en cierta medida al de un imán permanente. Esta cualidad ha encontrado aplicaciones en los micrófonos electrostáticos, como el que se muestra en la figura 8-11. Se hace vibrar el electreto polimérico laminar polarizado entre 10^6 y 10^7 V/m a 200°C , provocando que varíe el campo electrostático a través del intervalo de aire o entrehierro h , lo cual produce una pequeña fluctuación en el voltaje a través del dispositivo⁷.

Los electretos también se están aplicando en esquemas avanzados de filtración, en los que el campo eléctrico puede usarse para incrementar electrostáticamente la eficiencia de colección de un filtro. La capacidad de controlar la carga en un material dieléctrico también ha conducido a varias aplicaciones útiles, como el copiado xerográfico y la impresión electrostática⁸.

En el contexto del almacenamiento de carga en dieléctricos, el fluoruro de polivinilideno polimérico

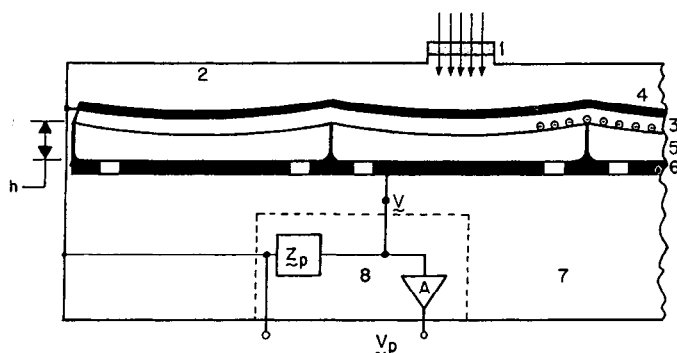


Fig. 8-11. Micrófono multicelular de electreto. 1. Lambra (abertura) de sonido y material de impedancia. 2. Cámara acústica frontal. 3. Película de electreto. 4. Metalización. 5. Película de aire. 6. Electrodo estacionario con orificios. 7. Cámara acústica posterior. 8. Preamplificador. (Adaptado de la referencia 7, según Baumhauer y Brzezinski⁷.)

(PVF₂) es de particular interés. Éste es un polímero semicristalino único con la unidad monomérica CH₂ = CF₂ que tiene por lo menos dos estructuras cristalinas estables⁹, una de las cuales está orientada con un momento dipolar de $7.01 \times 10^{-30} \text{ C} \cdot \text{m}$. Las películas delgadas pueden polarizarse en campos de aproximadamente 10^8 V/m a fin de formar un electreto que presente una respuesta piezoeléctrica y piroeléctrica. Estas propiedades lo hacen interesante como transductor y como dispositivo de visualización por infrarrojo.

REFERENCIAS BIBLIOGRÁFICAS

1. J. C. Anderson, *Dielectrics*, Chapman & Hall, Londres.
2. J. K. Nelson, "Breakdown Strength of Solids", en R. Bartnikas y R. M. Eichhorn, eds., *Engineering Dielectrics Vol. 2A*, ASTM, Filadelfia.
3. W. P. Baker, *Electrical Insulation Measurements*, Newnes, Londres.
4. J. K. Nelson, "Solid Dielectrics for Capacitors", en M. Bever, ed., *Encyclopedia of Materials Science and Engineering*, Pergamon, Oxford.
5. D. G. Shaw, S. W. Cichanowski y A. Yializis, "Changing Capacitor Technology: Failure Mechanisms and Design Innovations", *Trans. IEEE* E1-16:399.
6. K. Sanada y M. Hosokawa, "Electric Double Layer Capacitor, Super Capacitor", *NEC Res. Devel.* 55:21.
7. J. C. Baumhauer, Jr. y A. M. Brzezinski, "The EL2 Electret Transmitter: Analytical Modeling, Optimization and Design", *Bell Sys. Tech. J.* 58:1557.
8. A. D. Moore, *Electrostatics and Its Applications*, Wiley, Nueva York.
9. D. K. Das Gupta y K. Doughty, "On the Nature and Mechanism of Piezo- and Pyro-Electricity in Polyvinylidene Fluoride", editado por Y. Wada, M. M. Perlman y H. Kokado: *Charge Storage, Charge Transport and Electrostatics With Their Applications*, Elsevier, Amsterdam.

8.3 MATERIALES DIELECTRICOS COMO AISLANTES

J. Keith Nelson

8.3.1 Contexto y criterios de selección

La elección de medios gaseosos, líquidos o sólidos para trabajo como aislantes de conductores eléctricos debe hacerse con cuidado, especialmente cuando se diseñan estructuras aislantes compuestas. En base a lo expuesto en el artículo 8.2.3, está claro que la resistencia dieléctrica o aisladora (el campo eléctrico en disrupción) depende en gran medida de las circunstancias en las cuales se mide, y que suele estar dominada por las impurezas y las imperfecciones presentes. Además, los valores citados suelen asumir un campo eléctrico uniforme o bien definido, como el de la especificación IEC 243.

En la práctica, los campos eléctricos rara vez son uniformes, dado que la configuración geométrica de los conductores y las asperezas en las superficies suelen crear divergencias de campo local y volumétrica. También puede ocurrir intensificación del campo eléctrico debido a cargas espaciales, que pueden acumularse tanto en el interior de un dieléctrico como en su superficie. La evaluación de tal distorsión interna del campo es difícil, pero puede efectuarse mediante técnicas complejas, e indica que en circunstancias especiales la distorsión puede ser significativa.

Las propiedades de permitividad relativa (constante dieléctrica) y de tangente de pérdida también son de importancia en la falla de estructuras aislantes. En la configuración de un compuesto aislante esforzado con voltajes alternantes, el campo eléctrico estará distribuido capacitivamente, por lo que será afectado por la permitividad de los materiales utilizados. Por ejemplo, una pequeña burbuja de gas atrapada dentro de un molde de resina epóxica ($\epsilon_r = 3.8$) estará sometida a unas cuatro veces el campo eléctrico del material huésped, lo que puede dar por resultado descargas parciales, como se indica en el artículo 8.3.3. El factor

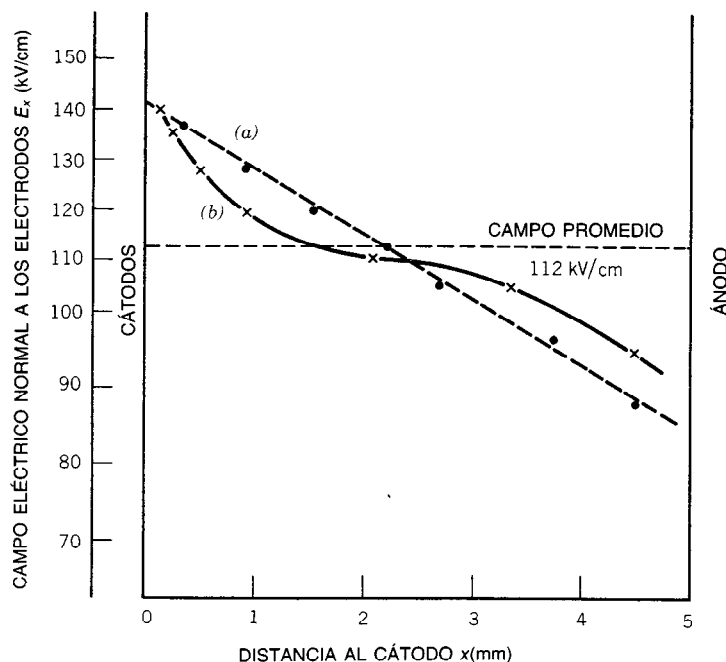


Fig. 8-12. Campo local E_x graficado en función de la posición x en: a) Aroclor 1242 y b) en una interfaz espaciadora fenólica. (Adaptado de Cherney y Cross².)

de disipación es relevante no sólo en las aplicaciones en las que es necesario minimizar las pérdidas, sino también debido a su efecto en los fenómenos de disrupción. Si la velocidad de incremento de energía del campo eléctrico es mayor que aquella a la cual es posible perder calor por el aislamiento, puede resultar falla por calentamiento.

Existen muchos casos en los que la elección de un material eléctrico aislante puede ser determinada por requerimientos no eléctricos. Consideraciones sobre carga mecánica, rigidez o estabilidad dimensional de los sólidos pueden hacer inadecuada la elección de algunos materiales poliméricos. Análogamente, las temperaturas de servicio, la compatibilidad química o ambas cosas suelen ser restricciones para dicha elección. Éste es especialmente el caso al elegir los fluidos de impregnación a utilizar con los sólidos poliméricos (artículo 8.2.4).

Las interfaces entre sólidos y gases o líquidos invariablemente forman un vínculo débil en cualquier sistema de aislamiento. Éste ocurre debido a las mejoras del campo y a las cargas interfaciales y, en el caso de un líquido, puede ser afectado por el movimiento del fluido.

En la figura 8-12 se muestra un ejemplo en el que se compara la distribución del esfuerzo eléctrico en un espaciador fenólico y en el líquido².

Para sistemas de alto voltaje, es posible minimizar las dificultades por medio del diseño geométrico fundamentado, por ejemplo, con trayectorias de flujo plástico creciente y protección de las intersecciones del punto triple (electrodo-sólido-fluido)³. Las superficies

eléctricamente esforzadas que se encuentran expuestas a un ambiente hostil pueden ser una limitación grave. La contaminación por humedad, por agentes atmosféricos o por partículas carbonáceas o metálicas puede perjudicar la integridad del dieléctrico.

8.3.2 Aplicaciones de bajo voltaje

El uso de dióxido de silicio y otros materiales empleados como aislantes en la industria de los semiconductores es el tema del artículo 8.1, pero en general los altos campos eléctricos asociados con películas aislantes cuyo grosor es sustancialmente inferior a $1 \mu\text{m}$ hacen que la operación de inyección de electrones y la ionización produzcan fallas rápidamente, como se puede observar en la referencia 4. La posible falla en tales circunstancias puede modificarse mediante la migración de impurezas iónicas y algunas veces puede suceder por un mecanismo térmico⁵, pero por lo general ocurre en un orden menor del de los microsegundos.

Otros mecanismos de falla son más lentos. Los materiales orgánicos pueden contribuir al envejecimiento por esfuerzo eléctrico a bajos voltajes por medio de mecanismos electroquímicos. Esta forma de falla a largo plazo es de particular importancia para materiales, como las resinas epóxicas, utilizados como moldes de encapsulamiento de paquetes microelectrónicos. Aun cuando la inmigración de iones atmosféricos contaminantes (p. ej., sales, soldaduras, fundentes y humedad) es una causa primordial de falla, el paquete de encapsulamiento en sí puede contribuir al problema.

Los materiales epóxicos pueden contener contaminantes iónicos, halógenos de la degradación térmica y ácidos corrosivos que resultan de la curación incompleta con un anhídrido endurecedor. Una vez que la especie corrosiva llega a la superficie del circuito integrado, la cinética de la corrosión que provoca la falla final dependerá principalmente de la temperatura que prevalezca y del voltaje de polarización, el cual no sólo impulsa la corrosión electrolítica sino que también contribuye a la inmigración de impurezas⁶.

Con frecuencia ocurre deterioro electroquímico en estructuras impregnadas, como capacitores, que contienen iones móviles de impurezas. El movimiento de dichos iones produce reacciones en el electrodo y cambios químicos indeseables en los dieléctricos sólidos o líquidos. Ciertas sustancias orgánicas, como quinonas y compuestos aromáticos azoicos y azóxicos, pueden emplearse como estabilizadores para retardar la falla electroquímica al actuar como aceptores de hidrógeno⁷.

La descarga superficial también puede hacer que el material aislante falle a voltajes comparativamente bajos. La combinación de humedad y contaminantes en una superficie provoca excesiva conductividad superficial. A su vez, las corrientes de conducción provocan calentamiento local y la formación de bandas secas. Por tanto, los esfuerzos excesivos son confinados a áreas locales de las superficies aislantes, con las resultantes degradación y erosión del material orgánico

subyacente⁸. Se han creado pruebas estándares, como la especificación IEC 112, para definir un índice comparativo de descarga para los materiales de pantallas.

8.3.3 Aplicaciones de alto voltaje

Como se analizó en el artículo 8.3.1, las descargas pueden ocurrir localmente sin precipitar de inmediato la falla total del aislamiento. Esto sucede cuando el campo eléctrico de un aislante es altamente divergente debido a gas ocluido, inclusiones extrañas o diseño geométrico. Estas descargas provocan erosión acumulativa en el aislamiento y la generación de subproductos químicos dañinos que provocan falla finalmente. Las estructuras aislantes, como la mica epóxica, utilizadas en los estatores de grandes máquinas giratorias, pueden resistir descargas continuas cuyas magnitudes superan los 1 000 pC y aun así tener una vida útil aceptable, mientras que tales magnitudes en un capacitor podrían provocar fallas en cuestión de horas. Este fenómeno ha llevado al uso difundido de técnicas de detección, localización y medición de descargas en equipos de alto voltaje⁹.

Un caso especial de las descargas parciales es el efecto de arborización "eléctrica" (diferente del "hídrico" o el "electroquímico"¹⁰) en campos altamente no uniformes, como los que se forman en salientes agudas, donde las descargas pueden iniciarse y propagarse en una serie de etapas como las que se muestran en la



Fig. 8-13. Efecto de arborización eléctrica en crecimiento radial desde el escudo semiconductor de un cable de distribución de energía hacia el interior del aislamiento. (Cortesía de Union Carbide Corporation).

figura 8-13. El inicio de tal efecto con frecuencia se atribuye a la formación de microgrietas en la región de campo intenso. Dichas grietas pueden resultar del procesamiento o formarse como consecuencia de la aplicación de esfuerzos. En general, la destrucción de un material aislante por este proceso de canalización es mucho más rápida que la erosión por descarga en cavidades gaseosas. En la referencia 11 pueden consultarse más detalles acerca del proceso.

8.3.4 Resumen de materiales

El lector puede encontrar datos amplios sobre una gran variedad de materiales aislantes en las referencias 12 y 13. Entre los sólidos poliméricos sintéticos con enlaces cruzados se utilizan ampliamente las resinas fenólicas y epóxicas. Ambos grupos suelen emplearse con materiales de relleno o de refuerzo, por lo que sus propiedades varían con el tipo y la forma de carga del material de relleno. Los formaldehídos fenólicos proporcionan resistencia a la tensión en el intervalo de 50 a 70 N/mm² y absorción de humedad entre el 0.1 y el 1%. Pueden tolerarse temperaturas de operación de hasta 120 °C, y la resina bien curada tiene permitividad relativa de alrededor de 4.5 y tangente de pérdida de 0.01 aproximadamente. Sin embargo, estas últimas propiedades son sensibles a la temperatura y a la frecuencia, y el empleo difundido de relleno de celulosa hace que los materiales fenólicos sean susceptibles a la descarga superficial. Es posible obtener una mayor resistencia a esta última por medio de las aminorresinas, como el formaldehído de melamina, pero a un costo extra considerable.

Existen dos tipos de resinas epóxicas: las aromáticas, formadas por la reacción de epiclorhidrina con bisfenol A, y las cicloalifáticas, con estructura anular saturada. Cuando se curan de manera adecuada, las resinas epóxicas presentan muchas características mecánicas y eléctricas atractivas, y son compatibles con una gran variedad de materiales. El aislamiento epóxico suele ser por vaciado o por moldeo, con frecuencia al vacío, y se cura a temperaturas hasta de 180 °C. La resina no rellena tiene permitividad relativa de 3.5 a 5, con tangente de pérdida en el intervalo de 0.005 a 0.02 a 20 °C hasta frecuencias de 1 MHz aproximadamente. Las resinas epóxicas se utilizan ampliamente cuando es posible obtener tableros reforzados con telas de vidrio o basados en celulosa cuyas resistencias a la tensión sean hasta de 400 N/mm² y cuando pueden tolerarse temperaturas sostenidas de hasta 130 °C. Este límite puede elevarse hasta aproximadamente 150 °C para compuestos especiales de mica epóxica resistentes a las descargas, que se utilizan bastante en el aislamiento de maquinaria giratoria. Las resinas cicloalifáticas presentan mayor resistencia a la descarga superficial y se aplican en aisladores externos.

De los polímeros lineales, el polietileno en sus dos formas —de alta y baja densidad— se utiliza ampliamente como material aislante. Se caracteriza por bajas pérdidas ($\tan \delta = 4 \times 10^{-4}$), permitividad de 2.2 aproximadamente y propiedades mecánicas que lo hacen atractivo para aplicaciones en cables, especialmen-

te a altas frecuencias. La temperatura de operación está limitada a alrededor de 80 °C, aunque es posible elevarla en alguna medida por formación de enlaces cruzados. El polietileno es resistente a la descarga superficial, pero es incompatible con algunos disolventes, en especial a altas temperaturas. Aun cuando su absorción de agua es baja, su empleo como aislante de cables de alta tensión es complicado debido a la incidencia del efecto de arborescencia tanto electroquímico como eléctrico¹¹. En la aplicación en cables, los máximos esfuerzos de tensión se limitan a aproximadamente 50 kV/cm, que es alrededor del 1% de la resistencia dieléctrica presentada en una prueba de laboratorio con una muestra pequeña (fig. 8-10 del artículo 8.2). En las aplicaciones que requieren resistencia aún mayor del aislante, como en los circuitos de alta impedancia de los electrómetros, es posible reemplazar el polietileno por politetrafluoroetileno (PTFE) con tangente de pérdida igual a 10^{-5} . El PTFE posee otras propiedades útiles, como alta temperatura de operación (250 °C), resistencia única a los disolventes y no inflamabilidad, pero su precio restringe su empleo a situaciones para las cuales no hay sustituto.

Para el aislamiento de cables de baja tensión se utiliza una gran variedad de materiales. Los copolímeros de etilén-propileno con enlaces cruzados poseen propiedades semejantes a las del hule (caucho) natural, y cuando se agregan al cloruro de polivinilo (PVC) plastificadores y estabilizadores idóneos para minimizar la descomposición a alta temperatura, es un polímero idealmente apropiado para los cables de baja tensión en virtud de su buena adhesión, resistencia al aceite y propiedades mecánicas. Dado que es un material polar (artículo 8.2.1), el PVC experimenta pérdidas eléctricas, con un notable máximo de pérdida a 100 °C, pero este hecho tiene pocas consecuencias en las aplicaciones de bajo voltaje. El PVC puede utilizarse hasta temperaturas aproximadas de 80 °C.

Otra clase importante de materiales aislantes está integrada por los esmaltes para alambre usados en las espiras (devanados) de máquinas giratorias, transformadores y otros dispositivos electromagnéticos. La clase más importante de los materiales empleados para esta aplicación son las poliimidas aromáticas, que pueden usarse hasta 250 °C. Estas poliimidas se aplican como un barniz prepolimérico y se curan a continuación para formar un recubrimiento en el que se combinan flexibilidad y resistencia a disolventes y ácidos. Los materiales de poliimida son costosos, por lo que se han desarrollado formulaciones híbridas (p. ej., la poliimida-iimida), con una gama de propiedades y costos correspondientes.

Aunque los materiales poliméricos dominan la industria de los aislantes, materiales tradicionales como el papel de celulosa siguen siendo los dieléctricos preferidos para los transformadores y cables de muy alto voltaje. El papel es un material de bajo costo con buenas propiedades mecánicas y capacidad de impregnación con aceites minerales. Las propiedades dependen bastante del tipo y la pureza del papel y de la medida en que se elimina la humedad¹². Sin embargo, los papeles bien preparados impregnados con aceite

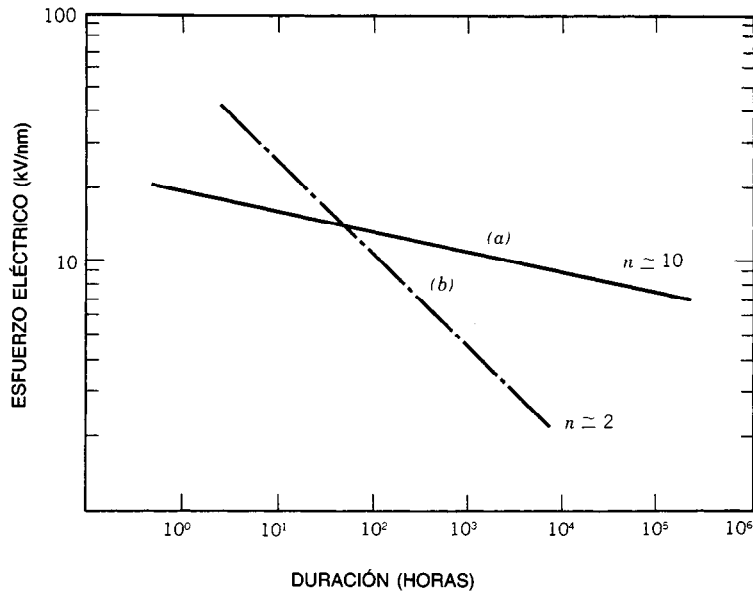


Fig. 8-14. Curvas representativas de la resistencia al voltaje de ruptura para: a) compuesto de epoxi-mica; b) polietileno sometido a descargas parciales.

tienen una tangente de pérdida de aproximadamente 2×10^{-2} a 20 °C. En la referencia 14 se analizan exhaustivamente las propiedades del aceite y otros fluidos de impregnación. Los aceites minerales nafténicos y parafínicos se utilizan ampliamente como enfriadores-aisladores, pero se inflaman con una reacción exotérmica y pueden oxidarse, especialmente en presencia de metales. En las aplicaciones que requieren resistencia al fuego se emplearon bifenilos policlorados fluidos hasta que fue prohibido su uso debido a su inaceptabilidad ecológica. Como una alternativa se utilizan silicones fluidos.

Los sólidos inorgánicos tales como vidrios y cerámicas poseen resistividades de CD que dependen en gran medida de la temperatura, debido a la predominancia de la conducción iónica. Sin embargo, su naturaleza inorgánica los hace resistentes a la degradación térmica y eléctrica. Las pérdidas eléctricas de la mayor parte de los vidrios decrecen con frecuencias de hasta 10^{-6} a 10^{-7} Hz, y después aumentan de manera notable en la región de las microondas. Las mayores aplicaciones del vidrio como aislador se dan en cubiertas de lámparas y tubos al vacío, aunque los materiales de borosilicato de aluminio están siendo empleados como bases para circuitos de películas delgadas¹⁵.

Las cerámicas tienen mayor resistencia al impacto que los vidrios, así como capacidad térmica única. Muchas cerámicas de uno o varios óxidos, como alúmina (Al_2O_3) y berilia (BeO), tienen puntos de fusión mayores de 2 000 °C, y la conductividad de esta última se encuentra en el intervalo de 5×10^4 a 25×10^4 W/K, lo que la hace única como aislante con conductividad térmica parecida a la de un metal. Se utiliza como material de disipación térmica para dispositivos semiconductores de alta potencia. Las porcelanas eléctricas fa-

bricadas con arcilla cocida a temperaturas de 1 000 a 1 400 °C constituyen la base de la mayor parte de los aislantes de alto voltaje para uso en la intemperie, limitados más por la descarga eléctrica sobre la superficie del aislador (contorneamiento externo) que por la falla interna del aislamiento, por lo que siempre cuentan con un vidriado a fin de minimizar la capacidad de la superficie porosa de coleccionar contaminación.

Excepto por el empleo difundido del aire atmosférico como aislante, la aplicación de gases dieléctricos electronegativos, como los clorofluorocarburos y el hexafluoruro de azufre (SF_6), está restringida principalmente a sistemas de potencia de alto voltaje. A una presión de 3 atm, el SF_6 es capaz de resistir esfuerzos eléctricos parecidos a los que resiste un aceite aislante, y tiene la ventaja de ser ligero y no inflamable. Desarrollos recientes en la tecnología del SF_6 están conduciendo al diseño de subestaciones de alto voltaje, completamente encerradas, y se están usando gases a alta presión como opciones a los líquidos como agentes de impregnación para sistemas poliméricos¹⁶. Como en el caso de los dieléctricos líquidos, la resistencia dieléctrica de los gases depende de manera crítica de la presencia de partículas de impurezas que afectan la distribución del campo eléctrico.

8.3.5 Estadísticas de la disrupción eléctrica

La falla del aislamiento no es sólo un fenómeno físico y químico, sino también un evento distribuido estadísticamente. Varias pruebas bajo circunstancias nominalmente idénticas produjeron una distribución de los valores de las disrupciones, que con frecuencia mostró una amplia dispersión, especialmente para sistemas en los que participan líquidos. Además, la probabilidad

de falla depende del tiempo, así como de los esfuerzos. Por lo común (pero no de manera universal) puede descartarse el caso en que un sistema aislante opone mayor resistencia dieléctrica bajo voltajes de impulso de corta duración, debido a que es posible eliminar algunos mecanismos dependientes del tiempo (como la migración de partículas o de carga espacial). Debido a que la falla es controlada por un proceso de "enlace débil"¹⁷, se ha logrado un cierto éxito al describir la falla del dieléctrico mediante una distribución de valor extremo¹⁸, como la del tipo de Weibull, que relaciona la probabilidad de falla con un esfuerzo o con el tiempo.

Para el caso de un esfuerzo E :

$$\text{Pr}(E) = 1 - \exp \left[- \left(\frac{E - \gamma}{\alpha} \right)^\beta \right]$$

donde α , β , y γ son parámetros que especifican la escala, forma y umbral de la distribución, respectivamente. Una de las consecuencias de este tipo de descripción es que la mayor parte de los sistemas aisladores presentan una dependencia aproximadamente logarítmica de la resistencia a la disrupción con respecto al área (o al volumen)¹⁹. Esto resulta del hecho de que, a medida que los sistemas son más grandes, es mayor la probabilidad de que tengan una imperfección o evento iniciador.

La capacidad de predecir la vida útil a largo plazo de una estructura esforzada también es importante. Muchos modelos matemáticos de envejecimiento se han basado en una ley empírica de envejecimiento de la forma:

$$E^n t = \text{constante}$$

donde E es un esfuerzo eléctrico constante y t es el tiempo de falla. El exponente n para el aislamiento orgánico suele variar de 1 a 6, y para sistemas de mica varía de 6 a 13.

En la figura 8-14 se presenta un ejemplo de una de estas curvas de resistencia a la ionización (o resistencia a largo plazo al alto voltaje) para un material de mica epóxica, en comparación con la de muestras de polietileno de 0.25 mm de grosor. Las características revelan el comportamiento típico de los materiales sometidos a descargas parciales, pero los valores precisos dependen bastante de la configuración geométrica individual²⁰. En realidad, la resistencia mencionada del aislamiento depende de sus entornos térmico, mecánico y químico, y en las evaluaciones exactas de la vida útil debe tomarse en cuenta el esfuerzo multifactorial²¹.

BIBLIOGRAFÍA

REFERENCIAS BIBLIOGRÁFICAS

1. D. W. Tong, "Electron Beam Probing of Space Charge in PET Films", *Trans. IEEE* **E1-17**:377.

2. E. A. Cherney y J. D. Cross, "Electric-Field Distortions at Solid-Liquid Dielectric Interfaces", *Trans. IEEE* **E1-9**:37.
3. K. Nakanishi, Y. Shibuya y T. Nitta, "Experimental Study of the Breakdown Characteristics of Large Scale Gas Insulated Systems", editado por L. G. Christophorou en: *Gaseous Dielectrics II*, Pergamon, Elmsford, NY.
4. N. Klein, "A Theory of Localized Electronic Breakdown in Insulating Films", *Adv. Phys.* **21**:605.
5. N. Klein y E. Burstein, "Electrical Pulse Breakdown of Silicon Oxide Films", *J. Appl. Phys.* **40**:2728.
6. L. G. Feinstein, "Failure Mechanisms in Molded Microelectronic Packages", *Semiconductor Internat.*
7. N. Parkman, "Some Properties of Solid-Liquid Composite Dielectric Systems", *Trans. IEEE* **E1-13**:289.
8. P. J. Lambeth, "Effect of Pollution on High-Voltage Outdoor Insulators", *Proc. IEE* **118R**:1107.
9. J. H. Mason, "Discharge Detection and Measurements", *Proc. IEE* **122**:1407 (1965).
10. S. L. Nunes y M. T. Shaw, "Water Treeing in Polyethylene—A Review of Mechanisms", *Trans. IEEE* **E1-15**:437.
11. R. M. Eichhorn, "Treeing in Solid Organic Dielectric Materials", editado por R. Bartnikas y R. M. Eichhorn en: *Engineering Dielectrics Vol 2A*, ASTM, Filadelfia.
12. F. M. Clark, *Insulating Materials for Design and Engineering Practice*, Wiley, Nueva York.
13. R. W. Sillars, *Electrical Insulating Materials and Their Applications*, Peregrinus, Stevenage, Gran Bretaña.
14. A. C. M. Wilson, *Insulating Liquids: Their Uses, Manufacture and Properties*, Peregrinus, Stevenage.
15. C. A. Harper, ed., *Materials for Electronic Packaging*, McGraw-Hill, Nueva York.
16. C. W. Reed, S. F. Philp, M. Kawai y H. M. Schneider, "Partial Discharge Inception and Breakdown Studies on Model Sheet-Wound, Compressed SF₆ Gas-Impregnated Polymer Film-Insulated Windings", editado por L. G. Christophorou en: *Gaseous Dielectrics II*, Pergamon, Elmsford, NY.
17. E. Occhini, "A Statistical Approach to the Discussion of the Dielectric Streight in Electric Cables", *Trans. IEEE* **PAS-90**:2671.
18. G. C. Stone y R. G. van Heeswijk, "Parameter Estimation for the Weibull Distribution", *Trans. IEEE* **E1-12**:253.
19. J. K. Nelson, B. Salvage y W. A. Sharpley, "Electric Strength of Transformer Oil for Large Electrode Areas", *Proc. IEE* **118**:388.
20. J. H. Mason, "Discharges", *Trans. IEEE* **E1-13**:211.
21. L. Simoni, "A General Approach to the Endurance of Electrical Insulation under Temperature and Voltage", *Trans. IEEE* **E1-16**:277.

LIBROS Y REVISTAS

- L. T. Apps, "Metallised Plastics Capacitors for Electronics", *Electronics and Power* **16**:369.
- R. Bartnikas y R. M. Eichhorn, *Engineering Dielectrics Vol. 2A*, ASTM, Filadelfia.
- E. J. Gumbel, *Statistics of Extremes*, Columbia Press, Nueva York.
- P. J. Harrop, *Dielectrics*, Butterworth, Londres.
- E. J. McMahon, "A Tutorial on Treeing", *Trans. IEEE* **E1-13**:277.
- D. A. Nicker, "High Voltage Ceramic Capacitors", *Electrocomp. Sci. Tech.* **1**:113.
- M. M. Perlman, ed., *Electrets, Charge Storage and Transport in Dielectrics*, Electrochemical Society, Princeton, NJ.
- *Proceedings of the 2nd Capacitor and Resistor Technology Symposium*, Components Technology Institute, Huntsville, AL.
- C. W. Reed, ed., *Proceedings of a Symposium on High Energy Density Capacitors and Dielectric Materials*, National Academy of Science, Washington, DC.
- K. Sato y cols., "Characteristics of Film and Oil for All-Polypropylene-Film Power Capacitors", *Trans. IEEE* **PAS-99**:1937.
- Y. Wada, M. M. Perlman y H. Kokado, eds., *Charge Storage, Charge Transport and Electrostatics with Their Applications*, Elsevier, Amsterdam.
- Y. Yoshida y cols., "Evolution of Power Capacitor as a Result of New Material Development", *CIGRE*, 15-01, París.

CAPÍTULO 9

Materiales magnéticos

Alex Goldman

Spang Industries
Butler, Pennsylvania

9.1 Definiciones

9.2 Propiedades eléctricas y magnéticas

- 9.2.1 Propiedades deseables para aplicaciones específicas
- 9.2.2 Intervalos disponibles de propiedades magnéticas

9.3 Propiedades físicas

- 9.3.1 Resistencia a la tensión y a la compresión
- 9.3.2 Dureza
- 9.3.3 Coeficiente de dilatación
- 9.3.4 Temperaturas de transición de fase
- 9.3.5 Resistencia a la corrosión
- 9.3.6 Efectos de la radiación

9.4 Materiales utilizables en la práctica

- 9.4.1 Hierro
 - 9.4.2 Aleaciones de hierro al silicio
 - 9.4.3 Aleaciones de hierro-níquel
 - 9.4.4 Aleaciones de hierro-cobalto
 - 9.4.5 Ferritas blandas
 - 9.4.6 Materiales amorfos
 - 9.4.7 Ferritas duras
 - 9.4.8 Materiales metálicos magnéticos permanentes
 - 9.4.9 Materiales para microondas
 - 9.4.10 Materiales para memorias
 - 9.4.11 Materiales magnetostrictivos
-

9.1 DEFINICIONES

Un material magnético es aquel que puede polarizarse magnéticamente, ya sea a nivel atómico o a nivel masivo o macroscópico. Esta polarización implica el alineamiento (parcial o total) de los momentos magnéticos electrónicos de dicho material por medio de la aplicación de un campo magnético externo. Como en el caso de otros campos (tales como el gravitacional o el eléctrico), la detección del efecto se manifiesta empleando una sonda, que en este caso es el campo magnético de un imán permanente, o bien su equivalente, un campo creado por un alambre conductor de corriente.

El paramagnetismo es el alineamiento en paralelo de spins (giros) atómicos no apareados. El diamagnetismo es causado por la reacción antiparalela opuesta de la cantidad de movimiento orbital electrónica al campo externo. El alineamiento paramagnético causa un ligero aumento de la intensidad de la magnetización, mientras que el efecto diamagnético es negativo y sólo es observable cuando no existen spins no apareados. Son ejemplos de sustancias paramagnéticas el manganeso y el cromo, en tanto que el bismuto y el aluminio son diamagnéticos. En el paramagnetismo se requiere un gran campo externo para el alineamiento de una cantidad relativamente pequeña de spins atómicos. En el ferromagnetismo y en el ferrimagnetismo el efecto es mayor en varios órdenes de magnitud, debido a que existe un efecto cooperativo (conocido como efecto de energía de intercambio) en el que participa una cantidad extremadamente grande de spins en áreas denominadas dominios magnéticos. Todos los spins atómicos en un dominio de los materiales ferromagnéticos son paralelos, y en los materiales ferrimagnéticos (ferritas "duras" y "blandas" o de alta y baja remanencia) esto no necesariamente es cierto, aunque el efecto neto de las interacciones es parecido al de los materiales ferromagnéticos, en el sentido de que los spins descompensados de la combinación iónica se alinean en el dominio. Además, un material ferromagnético (o ferrimagnético) tiene diferentes valores de grados de polarización para un valor fijo del campo aplicado, dependiendo de los antecedentes magnéticos del material. Este efecto se denomina histéresis magnética. El análisis de este capítulo se limita a los materiales ferromagnéticos y ferrimagnéticos, ya que son los únicos materiales de importancia en las aplicaciones eléctricas o electrónicas. Se hace hincapié en los materiales comerciales, aunque muchos otros revisten importancia para fines científicos y especializados. Los materiales ferromagnéticos se clasifican como "duros" o "blandos" o de alta y baja remanencia, respectivamente. Un material ferromagnético "duro" tiene polarización permanente y se utiliza para el campo que genera, y uno "blando" es un material que debe ser energizado durante su uso por medio de un campo externo, especialmente un campo alternante o de CA.

Los materiales ferromagnéticos más importantes contienen hierro, níquel o cobalto, o con frecuencia alguna combinación de ellos. Estos son los elementos en los que la acción cooperativa de la energía de inter-

cambio es más conductiva. Algunos compuestos de manganeso (aleaciones de Heusler) y algunos elementos alcalinotérreos son ferromagnéticos en ocasiones, pero revisten menor importancia.

Además de clasificarse conforme a sus características químicas, los materiales ferromagnéticos también pueden clasificarse por su forma física. Pueden ser masivos (de tamaño relativamente igual en todas las dimensiones), como sucede con los imanes permanentes, pueden ser metales laminados en planchas delgadas con grosor de 0.000125 a 0.030 pulg (de 0.0031 a 0.75 mm), pueden estar hechos en forma de alambre, y por último, pueden estar en forma de polvo u hojuelas. Estas últimas formas pueden compactarse para formar componentes denominados núcleos de polvo o núcleos compactos, o bien pueden emplearse como tintas magnéticas o cintas de grabación sin ser sometidas a compactación. Los materiales de óxido magnético (ferritas) suelen formarse como polvo, compactarse y someterse a la acción del fuego para elaborar cuerpos de componentes prácticamente moldeados. Algunas veces estos materiales se emplean como cristales sencillos (cabezas de grabación), y con frecuencia es posible utilizar en distintas formas el mismo material. El Permalloy 80 puede usarse como material en forma de cinta, así como en forma de hojuelas o polvo. Cada material se analizará en la forma en que se presenta, ya que el montaje de las diversas formas en componentes es tema de otro capítulo.

9.2 PROPIEDADES ELÉCTRICAS Y MAGNÉTICAS

Las unidades de las cantidades magnéticas y de las eléctricas asociadas pueden expresarse en el sistema cgs o en el SI. En este capítulo se usan ambas, y en la tabla 9-1 se presentan conversiones para los dos sistemas.

La característica distintiva de un material magnético es el ciclo de histéresis magnética (ciclo BH; fig. 9-1). En una señal de CA, un ciclo de corriente de onda senoidal envía el material magnético a través de un ciclo de histéresis. Muchas de las propiedades importantes pueden inferirse basándose en esta figura. La abscisa H es la intensidad del campo magnético (que se mide en oersteds o amperes por metro) y es la intensidad del campo de polarización o polarizante. La ordenada B es la densidad de inducción o de flujo, y es la suma de la intensidad de campo y $4\pi M$ [que es la polarización total producida por los momentos magnéticos unitarios (o polos) M a una distancia de 1 cm]; por tanto, $B = H + 4\pi M$. En los materiales blandos para los que $M \gg H$, $B \cong 4\pi M$; en los materiales duros es necesario considerar ambos términos. El flujo magnético total es el producto de la densidad de flujo o de inducción por el área de la sección transversal, $\phi = BA$. La inducción de saturación, B_s , es la máxima polarización que puede obtenerse por unidad de área. Si bien B puede aumentar ligeramente después de este punto, ello se debe por completo al incremento de H . Otra propiedad primordial es el magnetismo remanen-

Tabla 9-1. C cantidades en unidades del sistema centímetro-gramo-segundo (cgs) y el sistema internacional (SI)

Cantidad	Unidades cgs	Unidades SI	Factor de conversión del sistema cgs al SI
ϕ , Flujo	maxwell o línea	weber o $V \cdot s$	10^{-8}
B , Densidad de flujo	gauss o línea/cm ²	tesla o Wb/m^2 o $(V \cdot s)/m^2$	10^{-4}
H , Intensidad de campo	oersted	amp/m	79.6
μ , Permeabilidad relativa	μ (gauss/oersted)	μ_r	1
$(BH)_{\text{máx}}$, Producto de energía	mega-gauss-oersted o gauss-oersted $\times 10^6$	kwatt-s/m ³ o (kJ)/m ³	7.96

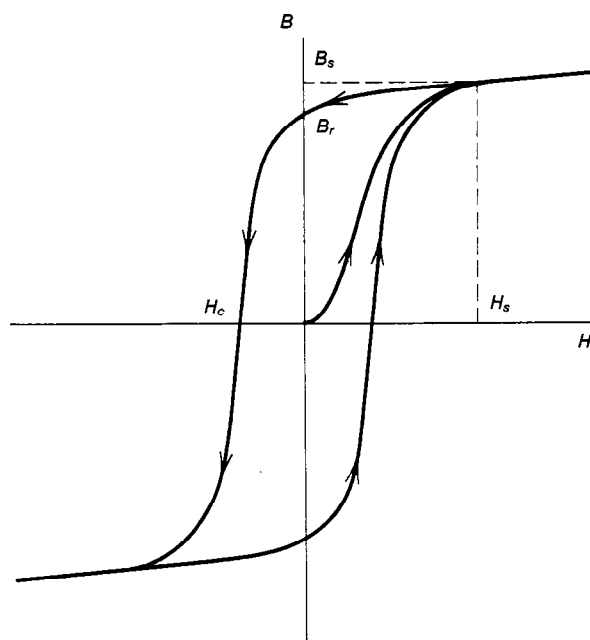
te (remanencia) o inducción residual, B_r , que es la polarización restante una vez que se ha suprimido el campo de polarización. Las unidades de medición de B , B_s y B_r son gauss (G) o teslas (weber/m²). El cociente B_r/B_s se denomina relación de rectangularidad, y puede variar del 10 al 20% hasta casi el 100%. La fuerza coercitiva H_c es el campo inverso necesario para reducir a cero la inducción. Se mide en oersteds o en amperes por metro, así como H_s , la intensidad de campo a la que ocurre B_s .

Una relación de suma importancia es la permeabilidad relativa μ , que es igual a B/H y carece de unidades. Para campos débiles que tienden a cero, esta relación se conoce como μ_i (o algunas veces como μ_0), y en el punto más elevado de la curva de magnetización se denomina $\mu_{\text{máx}}$. Para expresar la permeabilidad a diferentes niveles de conducción se emplean diferentes convenciones. En Estados Unidos, μ_{40} suele significar la permeabilidad a 40 G, y en Europa, μ_5 puede representar la permeabilidad a 5 mA/cm. Por ello, es importante conocer la convención que se utilice.

Una propiedad parcialmente relacionada con el ciclo de histéresis es la pérdida magnética; es decir, la pérdida de energía en el material debido al recorrido completo del ciclo de histéresis. Una parte de esta pérdida, conocida como pérdida por histéresis, puede inferirse a partir del área incluida barrida por el ciclo. Otra parte, denominada pérdida por corrientes parásitas o de Foucault, se debe a los bucles de corriente circulares que se generan en el material debido al campo magnético oscilante. Una tercera componente, denominada pérdida residual (o anómala), representa la parte sin explicación. La suma total de estas pérdidas se conoce como pérdida específica en el núcleo y se expresa en watts por libra o watts por kilogramo.

En el caso de las ferritas, por convención, la pérdida en el núcleo está dada en miliwatts por centímetro cúbico.

A bajos niveles de conducción, como los que se utilizan en las telecomunicaciones, la pérdida suele especificarse como una resistencia de pérdida y es posible establecer una disrupción de pérdida en diferentes for-


Fig. 9-1. Curva inicial de magnetización y ciclo de histéresis.

mas, que de nuevo dependen de la convención. Normalmente se utiliza como base la ecuación de Legg:

$$\frac{R}{\mu_f L} = aB_m + bf + c$$

donde a , b , y c representan los coeficientes de pérdida para histéresis, corrientes parásitas y componentes anómalas, respectivamente, y f es la frecuencia en hertz. Los coeficientes se determinan haciendo variar de manera independiente B_m o f .

Otro criterio de pérdida a bajos niveles (especialmente en las ferritas) es el factor de pérdida (LF, de *loss factor*), que puede expresarse como el cociente de la tangente del ángulo de pérdida, $\tan \delta$, entre la permeabilidad inicial:

$$LF = \frac{\tan \delta}{\mu_i} = \frac{1}{\mu_i Q} = \frac{R_r}{\mu_i X_L} = \frac{R_s}{2\pi\mu_i fL}$$

Algunas veces se especifica el inverso de LF o el producto $\mu_i Q$.

Dos parámetros relacionados con la estabilidad de un material (especialmente ferritas) son el factor de temperatura (TF, de *temperature factor*), y el factor de incapacidad (DF, de *disaccommodation factor*). El TF es el cambio relativo en inductancia por grado por unidad de inductancia:

$$TF = \frac{\Delta L}{L} \times \frac{1}{\mu_i \Delta T}$$

donde L es la inductancia en henrys y ΔT es el cambio de temperatura en grados Celsius. El DF describe el decremento fraccionario en permeabilidad por década de tiempo por unidad de permeabilidad.

$$DF = \frac{\mu_1 - \mu_2}{\mu_1^2 \log \frac{t_2}{t_1}}$$

Una última propiedad eléctrica del material magnético, que no se relaciona con el ciclo de histéresis pero que es de particular importancia para la aplicación de energía eléctrica, es la resistividad, definida como $R = \rho l/A$, donde R = resistencia, l = longitud y A = área transversal. La unidad de ρ es el ohm-cm o el ohm-m. Una propiedad relacionada, la constante dieléctrica, pasa a ser importante a frecuencias de microondas.

9.2.1 Propiedades deseables para aplicaciones específicas

La elección de un material magnético suele hacerse comparando los requerimientos de la aplicación con las características del material, teniendo en cuenta la frecuencia de operación, pérdidas de energía, estabilidad, requerimientos físicos y, por último, pero no por ello menos importante, el costo.

Aplicaciones en el manejo de energía eléctrica. A bajas frecuencias (como 60 Hz) se requieren elevada sa-

turación, bajo costo y baja pérdida en el núcleo. La elevada saturación reduce la cantidad necesaria de vueltas, de tal forma que ocurren bajas pérdidas óhmicas (también llamadas pérdidas por efecto Joule o pérdidas en el cobre). A altas frecuencias, las pérdidas por corrientes parásitas deben mantenerse bajas mediante el empleo de cintas metálicas delgadas o incrementando la resistividad del material.

Aplicaciones en las telecomunicaciones. En este caso se requiere una elevada μ_i a la frecuencia de operación. Es necesario consultar curvas publicadas de μ contra f (fig. 9-2) para evitar exceder la frecuencia crítica de un material (es decir, la frecuencia a la que la permeabilidad desciende abruptamente). Un bajo factor de pérdida, $1/\mu Q$, reduce la amplitud de banda e incrementa la selectividad. Además, la alta estabilidad de μ con respecto a densidad de flujo, tiempo y temperatura evita la desmodulación. Si existe una señal superpuesta de CD, también es de importancia la estabilidad de μ con respecto a la CD. Dado que la frecuencia de operación es alta, un requisito esencial es el empleo de materiales de elevada resistividad (ferritas) o de cintas delgadas.

Aplicaciones en memorias y computadoras. Aquí se requiere una elevada relación de rectangularidad (B_r/B_s). B_s debe ser elevado, y si su velocidad de repetición de pulsos es alta, las pérdidas por corrientes parásitas deben ser bajas (debe usarse material con elevada resistividad o en cintas delgadas).

Aplicaciones en blindajes. Se requiere una permeabilidad inicial elevada. A bajas frecuencias es de utilidad una reducida pérdida por histéresis. La fabricación se facilitará si el material tiene buena conformabilidad.

Materiales magnetostrictivos. Una elevada capacidad de magnetostricción es de primordial importancia. El factor de acoplamiento magnetomecánico debe ser alto, y deben observarse las precauciones mencionadas para impedir altas pérdidas por corrientes parásitas.

Materiales magnéticos permanentes. En este caso, la característica más importante es un elevado producto $(BH)_{\max}$. La fuerza coercitiva y la remanencia deben ser lo más elevadas posible. La estabilidad de estas propiedades contra el tiempo y la temperatura evitará la desmagnetización.

9.2.2 Intervalos disponibles de propiedades magnéticas

Es obvio que las propiedades magnéticas disponibles de los materiales tienen algunos límites prácticos. Es importante saber cuáles son dichos límites y qué otras propiedades están asociadas con ellos.

Inducción por saturación. Ésta es una propiedad intrínseca que depende exclusivamente de la composición química a una temperatura específica cualquiera. Sin embargo, varía con la temperatura de

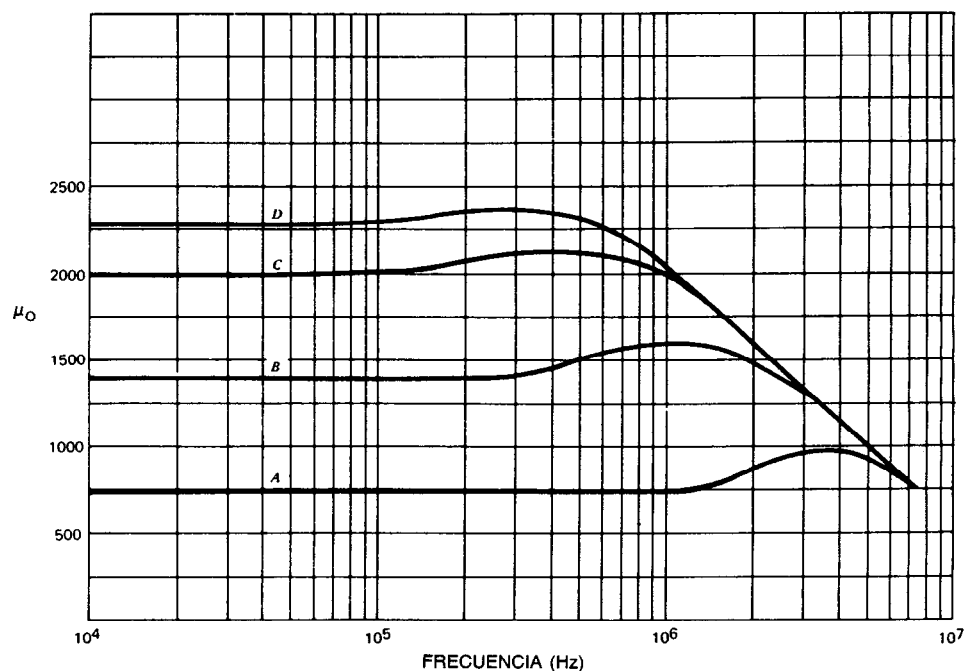


Fig. 9-2. μ_0 contra la frecuencia para cuatro materiales ferríticos. Obsérvese que la frecuencia de decaimiento decrece a medida que aumenta la permeabilidad.

forma parecida para todos los materiales ferromagnéticos. Por tanto, no es sorprendente que la inducción por saturación y el punto de Curie (que es la temperatura a la cual un material pierde sus propiedades ferromagnéticas) suelen ir paralelos.

La inducción por saturación de los metales magnéticos y de las aleaciones suele ser bastante alta, y los valores elevados de saturación (15 000 a 24 000 G o 1.5 a 2.4 T) son importantes en aplicaciones que requieren grandes movilizaciones de flujo para manejo de energía eléctrica (transformadores de potencia, motores). Los elementos ferromagnéticos en orden decreciente de valores de saturación son el hierro (21 500 G o 2.15 T), cobalto (19 000 G o 1.9 T) y níquel (6 000 G o 0.6 T). En general, las mezclas de dos o más de estos elementos producen un valor promedio de los valores de los elementos por separado. Una excepción importante es la aleación 50:50 de hierro-cobalto, cuyo valor de saturación es más alto (24 500 G o 2.45 T) que el de cualquiera de estos dos elementos. Las aleaciones de hierro, níquel o cobalto con adiciones de elementos no magnéticos, como silicio, aluminio, molibdeno y similares, presentan valores de saturación más bajos, pero pueden mejorarse otras propiedades útiles (como resistividad o permeabilidad).

Las saturaciones de estos materiales óxidos son inferiores a las de los metales por dos razones:

1. Los grandes iones oxígeno presentes en la malla cristalina diluyen la contribución de iones magnéticos del metal, ya que los óxidos son no magnéticos.

2. No todos los spins de los iones magnéticos están alineados paralelos en el cristal, como ya se describió.

Como resultado, los máximos valores de inducción por saturación que se encuentran son menores o iguales de 6 000 G o 0.6 T. La magnetita o ferrita ferrosa posee una de las saturaciones más elevadas, pero se utiliza poco debido a su baja resistividad. De las ferritas de uso común, las que presentan las saturaciones más elevadas son las de manganeso-cinc-hierro, cuyas saturaciones son hasta de 5 000 G (0.5 T). Los granates, que son de utilidad para aplicaciones de alta frecuencia (microondas), poseen saturaciones que varían de 500 a 2 000 G (0.05 a 0.2 T).

Permeabilidad. La elevada permeabilidad inicial depende parcialmente de propiedades intrínsecas, como la composición, que afectan los parámetros magnéticos fundamentales conocidos como anisotropía y magnetostricción. El primero es la afinidad de la magnetización por ciertas direcciones cristalinas, y el segundo es la sensibilidad de la magnetización a esfuerzos externos. Además de ser sensible a agentes químicos, la permeabilidad también lo es a características como textura cristalina, esfuerzo externo, inclusiones, dislocaciones y grado de ordenamiento atómico.

La permeabilidad puede medirse a CD o frecuencias mayores. Para blindajes se requiere alta permeabilidad a CD o frecuencias menores. La elevada permeabilidad inicial es de la mayor importancia para aplicaciones electrónicas con muy bajos niveles de conduc-

ción (menores de 100 G). En las aplicaciones de manejo de energía eléctrica ya mencionadas, la característica más importante es la máxima permeabilidad, o permeabilidad a la energía eléctrica. De todos los materiales, las aleaciones de níquel-hierro poseen la permeabilidad inicial más alta, especialmente cuando en la aleación el níquel se encuentra en el intervalo del 80% y se agrega del 4 al 5% de molibdeno. Este hecho se debe a que simultáneamente se logra que la anisotropía y la magnetostricción sean iguales a cero. Son posibles permeabilidades iniciales mayores de 100 000 y permeabilidades máximas hasta de 1 000 000. En el hierro al silicio las permeabilidades iniciales son mucho menores (de 1 800 a 40 G), pero permeabilidades máximas de 50 000 bastan para hacer de éste un material muy importante para aplicaciones de manejo de energía eléctrica. Para hacer hincapié en la necesidad de la pureza química, puede decirse que es posible incrementar la permeabilidad inicial del hierro desde unos millares hasta 100 000 por medio de procesos de refinamiento lentos y costosos.

Originalmente, las ferritas se consideraban sólo para frecuencias muy elevadas, y las permeabilidades eran secundarias. Sin embargo, en años recientes se han colocado en el mercado ferritas de material comercial de alta permeabilidad con μ_i hasta de 15 000. Se han elaborado materiales de laboratorio con μ_i hasta de 40 000. Los materiales para los cuales esto es posible son las ferritas de manganeso-cinc-hierro, que presentan bajas magnetostricción y anisotropía. La magnetostricción se reduce mediante el empleo de hierro en exceso, con lo cual se controla directamente el contenido de Fe^{2+} , y para minimizar la anisotropía se utilizan mecanismos químicos adecuados. Asimismo, la alta pureza es en extremo importante. La disponibilidad de materia prima de alta pureza ha sido, en parte, la causa del logro de las elevadas permeabilidades mencionadas.

Resistividad. Las pérdidas por corrientes parásitas son inversamente proporcionales a la resistividad, y directamente proporcionales al cuadrado de la frecuencia. Para que estas pérdidas sean mínimas debe elevarse la resistividad a medida que se incremente la frecuencia.

Los metales magnéticos puros (níquel, hierro y cobalto) presentan bajas resistividades (aproximadamente $10\mu\Omega\text{-cm}$). Sin embargo, las adiciones de silicio, aluminio y molibdeno incrementan de manera significativa la resistividad (hasta aproximadamente $50\mu\Omega\text{-cm}$). Además, los metales pueden pulverizarse y aislarse, reteniendo muchas características del metal, como la saturación, a la vez que se incrementa de manera importante la resistividad. Las aleaciones recientemente desarrolladas de metales amorfos presentan resistividades aproximadamente tres veces superiores a las de sus contrapartes cristalinas (alrededor de $50\mu\Omega\text{-cm}$).

El principal atractivo de las ferritas es la elevada resistividad, superior a la de los metales en varios órdenes de magnitud. Las resistividades más altas de este grupo se encuentran en los granates y en las ferritas deficientes en hierro, como NiFe_2O_4 , MgFe_2O_4 y

MnFe_2O_4 , cuyas resistividades varían de 10^6 a $10^{10}\Omega\text{-cm}$. La ferrita que presenta menor resistividad es la magnetita, seguida por la ferrita con exceso de hierro (que son principalmente ferritas de manganeso-cinc-hierro), con $\rho \approx 10^2\Omega\text{-cm}$.

9.3 PROPIEDADES FÍSICAS

9.3.1 Resistencias a la tensión y a la compresión

Al elegir un material magnético es necesario considerar sus propiedades mecánicas, debido a los requerimientos de fabricación de la componente (p. ej., estirado, perforación, práctica de incisiones o ranuras, laminado), pero también debido al efecto de las condiciones mecánicas en las propiedades magnéticas. Además, el material puede ser sometido a esfuerzos dinámicos en la aplicación a que se destine (rotores en motores). El permendur 2-vanadio debe poseer alta resistencia a la tensión en generadores de aviones, de modo que se sacrifica algo de las propiedades magnéticas. Las propiedades tensiles de las ferritas son deficientes (como las de la mayoría de las cerámicas), pero su resistencia a la compresión es elevada.

9.3.2 Dureza

Ésta es una característica importante de los metales, ya que indica su labrabilidad (estirado, etc.). Aunque la alta "blandura" magnética suele corresponder a una blandura mecánica total, un material así puede ser demasiado blando para la fabricación (p. ej., con respecto a perforaciones). Otro aspecto importante de la dureza está en la aplicación en que la dureza mecánica se relaciona con la resistencia al desgaste (p. ej., las cabezas de grabación).

Algunas aleaciones como el sendust (hierro-silicio-aluminio) o las aleaciones de hierro al alto aluminio (6.5 o 16% de Al) poseen excelente dureza para esta aplicación. Además y por la misma razón, las ferritas minocristalinas o prensadas en caliente también son excelentes.

9.3.3 Coeficiente de dilatación

Muchas aleaciones magnéticas también son importantes como aleaciones de dilatación o de compensación de la temperatura, lo cual es especialmente cierto para las de níquel-hierro cuyo contenido de níquel varía del 36 al 40%, algunas de las cuales presentan un coeficiente de dilatación cercano a cero.

9.3.4 Temperaturas de transición de fase

Estas temperaturas son importantes en el procesamiento de materiales magnéticos. Por ejemplo, la transición del hierro, que para el hierro puro suele ocurrir a 930°C , se elimina mediante la adición de otros elementos, como silicio. En este tipo de aleación es posible efectuar tratamientos térmicos a alta temperatura que no sería posible aplicar al metal puro.

9.3.5 Resistencia a la corrosión

Muchos materiales metálicos tienen aceptable resistencia a la corrosión, como las aleaciones al alto níquel (Permalloy 80), aunque el permendur vanadio y el mumetal son ligeramente mejores. Las aleaciones al alto hierro, como las de silicio-hierro, se corroen con gran facilidad, por lo que suelen protegerse con algún tipo de recubrimiento aislante, como el fosfato de magnesio o de aluminio. Por supuesto, las ferritas son bastante resistentes debido a su naturaleza de óxidos.

9.3.6 Efectos de la radiación

Casi todos los materiales magnéticos son afectados en alguna medida por los diferentes tipos de radiación. Además, los materiales que contienen cobalto son altamente radiactivos, por lo que en algunas aplicaciones se deben sustituir por materiales más resistentes a la radiación, a expensas de las propiedades magnéticas.

Los metales más afectados por la radiación gamma o la de neutrones son los materiales con altas permeabilidades, como las permalloys. Los más resistentes son los hierros al silicio y los materiales magnéticos permanentes, como los álnicos. Bajo radiación, es posible que los materiales con granos orientados pierdan orientación. Las propiedades sensibles a la estructura, como la permeabilidad μ o la fuerza coercitiva H_c , se ven bastante afectadas, al contrario de las insensibles a la estructura, como B_s . Las ferritas son aceptablemente estables bajo radiación a temperatura ambiente, pero son inadecuadas a altas temperaturas.

9.4 MATERIALES UTILIZABLES EN LA PRÁCTICA

9.4.1 Hierro

De todos los materiales, el menos costoso es el hierro. Con pequeñas cantidades de purificación para reducir el carbono es posible producir un material aceptable con alta saturación y pérdidas moderadas. Puede obtenerse hierro de alta pureza con muy elevada permeabilidad (aproximadamente de 100 000), pero la purificación es demasiado lenta para ser práctica. La inducción por saturación es aproximadamente de 21 000 G (2.15 T), y debido a sus bajas resistencias mecánica y resistividad no es posible utilizarlo en muchos circuitos de CA. El hierro también se expende en núcleos pulverizados comprimidos, y su ventaja principal es su elevada resistividad, así como la capacidad ulterior de operar a altas frecuencias, que llegan al intervalo de los megahertz. Por supuesto, la permeabilidad cae drásticamente a un intervalo que varía de 10 a 100. Además, existen en el mercado núcleos de hojuelas con permeabilidades de alrededor de 100.

9.4.2 Aleaciones de hierro al silicio

El hierro al silicio existe en los grados no orientado y orientado, y la mayor parte de los grados prácticos

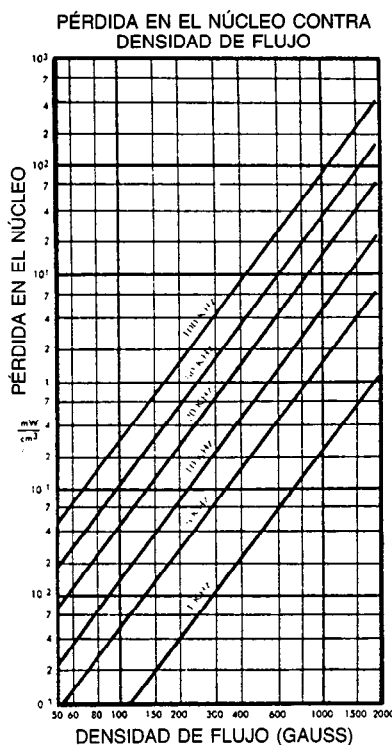


Fig. 9-3. Pérdida en el núcleo de un material ferrítico pulverizado para manejo de energía eléctrica ($\mu_0 = 3000$) como función de la densidad de flujo y de la frecuencia. (Cortesía de Magnetics, División de Spang Industries, Inc., Butler, Pa.)

contienen entre el 0.5 y el 3.5% de silicio. Los niveles superiores hacen demasiado frágil el material para fines de trabajo comercial. La fusión se realiza en hornos básicos de oxígeno (BOF, de *basic oxygen furnaces*) con carga de chatarra. El grosor de las láminas formadas en frío o en caliente a partir de los lingotes varía de 0.001 a 0.017 pulg (0.025 a 0.4 mm).

Si son grados no orientados, se recubren con fosfato y se envían al cliente, y si son orientados, suelen recubrirse con una capa de MgO y someterse a un recocido especial de transformación para desarrollar la textura. A continuación se recubren con fosfato, como los grados no orientados. Con cualquiera de los dos grados, una vez que el cliente ha fabricado el perfil, debe aliviar los esfuerzos del material mediante recocido para desarrollar las mejores propiedades magnéticas.

El método original para especificar la calidad del material consistía en multiplicar por 10 las pérdidas en el núcleo (en watts por libra) del material calibre 29 a 60 Hz y 15 000 G (1.5 T). Por tanto, un material con una pérdida en el núcleo de 2.2 W/lb en dichas circunstancias se denomina M-22. Para obtener la mayor permeabilidad son mejores los bajos niveles de silicio, y las menores pérdidas en el núcleo se obtienen en material orientado con un 3.25% de silicio. Recientemente ha habido adelantos en el procesamiento de material orientado de hierro al silicio. El material para este proceso se denomina Hi-B o de núcleo orientado. Existen

Tabla 9-2. Propiedades de algunos materiales magnéticos blandos

Materiales	Densidad de flujo de saturación (G)	Punto de Curie (°C)	Densidad (gm/cm ³)	Resistividad ($\mu\Omega$ -cm)	Permeabilidad inicial (μ_0)	Permeabilidad máxima	Pérdidas en el núcleo (W/kg)	A la frecuencia	A la inducción
Hierro	21 500	770	7.85	9.6	300	6 000			
Cobalto	19 000	1 121	8.84	9	70	250			
Níquel	6 080	358	8.89	8.7	250	2 500			
Material para manejo de energía eléctrica de baja frecuencia									
Acero al bajo carbono laminado en frío	21 500	770	7.85	9.6	200	5 000	7.3	60 Hz	15 kG
Fe al 3% de Si no orientado	20 000	750	7.65	47	280	8 000	3.2	60 Hz	15 kG
Fe al 3% de Si orientado	20 000	750	7.65	50	1 400	50 000	1.17	60 Hz	15 kG
Fe al 3% de Si de alta B	20 000	750	7.65	50	1 920		1.46	60 Hz	17 kG
Metglas [®] 2605 S2	15 600	415	7.18	130			0.25	60 Hz	14 kG
Metglas 2605 SC	16 100	370	7.32	125			0.3	60 Hz	14 kG
Permendur 2-V	24 500	950	8.3	7	1 000	11 000	25	400 Hz	20 kG
permendur	24 000	980	8.15	26	800	5 000	8.8	50 Hz	18 kG
Supermendur	24 000	940	8.15	26	800	70 000	12.1	400 Hz	15 kG
27% Co-Fe	24 200	970	8.02	19	650	2 800	6	60 Hz	15 kG
Materiales inductores y para blindajes									
50% Ni-Fe	15 500	500	8.2	45	3 500	100 000	6.6	1 kHz	10 kG
Fe al 50% de Ni orientado	16 000	500	8.2	45		200 000	8.8	1 kHz	10 kG
Mumetal	6 500	460	8.25	55	20 000	100 000	5.5	5 kHz	5 kG

Tabla 9-2. Propiedades de algunos materiales magnéticos blandos (continuación)

Materiales	Densidad de flujo de saturación (G)	Densidad de saturación (T)	Punto de Curie (°C)	Densidad (gm/cm ³)	Resistividad (μΩ-cm)	Permeabilidad inicial (μ ₀)	Permeabilidad máxima	Pérdidas en el núcleo (W/kg)	A la frecuencia	A la inducción
Materiales inductores y para blindajes (cont.)										
Metglas ^a 2826MB	8 800	0.88	353	8.02	160		750 000	96	50 kHz	2 kG
4-79 Permalloy	7 500	0.75	460	8.74	55	50 000	300 000	12.2	10 kHz	4 kG
Supermalloy	7 000	0.7	460	8.77	65	6 000	800 000	17.5	10 kHz	4 kG
Materiales inductores a alta frecuencia										
Materiales	Densidad de flujo de saturación (G)	Densidad de saturación (T)	Punto de Curie (°C)	Resistividad (Ω-cm)	Permeabilidad inicial (μ ₀)	Intervalo de frecuencias	Factor de pérdida (1/μ _Q)			
Permalloy pulverizado	8 000	0.8	460	500	14–300	1–300 kHz	300 × 10 ⁻⁶			
Sendust pulverizado	10 000	1.0	500	10	80	1–100 kHz	300 × 10 ⁻⁶			
Hierro carbonilo pulverizado	12 000	1.2	700	100	3–40	0.1–100 MHz	20 × 10 ⁻⁶			
Ferrita de Mn-Zn	3 800	0.38	145	100	2 000	1–200 kHz	2 × 10 ⁻⁶			
Ferrita de Ni-Zn	1 500	0.15	350	10 ⁷	80					
	3 500	0.35	400		1 500	0.1–10 MHz	100 × 10 ⁻⁶			
Material para manejo de energía eléctrica										
Materiales	Densidad de flujo de saturación (G)	Densidad de saturación (T)	Punto de Curie (°C)	Resistividad	Permeabilidad inicial (μ ₀)	Permeabilidad máxima	Pérdidas en el núcleo (W/kg)	A la frecuencia	A nivel de flujo	
Metglas ^a 2605 S3	15 800	1.5	405	125 μΩ-cm	—	20 000	10	20 kHz	4 kG	
Ferrita de Mn-Zn	4 700	0.47	250	100 Ω-cm	2 700	4 500	14	20 kHz	2 kG	
Núcleo de Fe-Ni 50-50 pulverizado	15 000	1.5	500	10 Ω-cm	125	150	88	20 kHz	2 kG (0.2 T)	
Núcleo de hierro pulverizado	19 000	1.9	770	5 Ω-cm	90	250	360	20 kHz	2 kG	

^aMetglas® es una marca registrada de Allied Corporation.

varios fabricantes que producen este tipo de material, que presenta mejor orientación, menores pérdidas en el núcleo y mayores inducciones de operación.

9.4.3 Aleaciones de hierro-níquel

Las aleaciones importantes de hierro-níquel se dividen en dos categorías: aquellas cuyo porcentaje de níquel es aproximadamente el 80% y aquellas cuyo contenido de níquel es aproximadamente el 50%. Las de la primera categoría poseen permeabilidades superiores y saturaciones inferiores, mientras que las de la segunda categoría presentan características inversas. Las aleaciones al 80% de níquel se elaboran a partir de materiales de alta pureza con adiciones de molibdeno o cobre hasta del 4 o 5% (mumetal). Los contenidos de carbono, oxígeno y azufre se mantienen muy bajos, y se utiliza un tratamiento térmico bastante cuidadoso y velocidad de enfriamiento controlada a fin de desarrollar el ordenamiento atómico idóneo. El recocido se efectúa en hidrógeno de alta pureza a temperaturas más bien elevadas.

Debido a que sus permeabilidades son inferiores, las aleaciones al 50% de níquel no requieren tantos cuidados como las del 80%, aunque su saturación es casi dos veces la de éstas. Los granos de algunas aleaciones al 50% también pueden estar orientados a fin de producir relaciones de rectangularidad extremadamente elevadas, en el intervalo del 98 al 99%.

Los materiales al 80% y al 50% de níquel se expanden en el mercado en forma de polvos comprimidos. Un polvo con 81% de material de níquel, 2% de molibdeno y el resto de hierro presenta pérdidas extremadamente bajas, excelente estabilidad a la polarización de CD y muy buena estabilidad a la temperatura. Se utiliza en inductores de baja pérdida y en bobinas de carga. El polvo al 50% de níquel-hierro adquiere importancia creciente en las aplicaciones de manejo de energía eléctrica, principalmente en fuentes de suministro en modo de conmutación. Las pérdidas son aproximadamente dos veces más altas, pero la inducción por saturación también es doble.

9.4.4 Aleaciones de hierro-cobalto

De todos los materiales magnéticos comunes, estas aleaciones tienen la saturación más alta. También son de las más costosas, debido al alto costo del cobalto. La saturación más alta (24 500 G o 2.35 T) ocurre cuando el contenido de cobalto es igual al 35%. Las aleaciones más importantes son las que tienen un 50% de cobalto, seguidas por las de menor contenido (27%) de este metal. Después de formarlo, con frecuencia se aplica recocido al material, con fines de orientación. La mayor parte de las aleaciones de hierro-cobalto se utilizan para aplicaciones en aeronaves o aeroespaciales.

9.4.5 Ferritas blandas

Las ferritas blandas están adquiriendo cada vez mayor importancia, a medida que aumentan las frecuencias

de operación de muchos circuitos eléctricos y electrónicos. A diferencia de los materiales ya mencionados, estas ferritas son cerámicas compuestas por soluciones sólidas de óxido de hierro, Fe_2O_3 , con óxidos de metales bivalentes, como manganeso, níquel, cobalto y magnesio. Su fórmula genérica es $\text{MO} \cdot \text{Fe}_2\text{O}_3$ o MFe_2O_4 , donde M es el ion metálico bivalente. Como ya se dijo, sus elevadas resistividades producen pérdidas muy bajas, especialmente a altas frecuencias.

Las ferritas suelen elaborarse con materias primas de costo más bien bajo, mezclando los óxidos componentes, calcinándolos, moliendo el polvo calcinado, secándolo, granulándolo y vaciándolo en un núcleo. Las propiedades se desarrollan por medio de un proceso de cocido al horno a alta temperatura. Existen otros métodos para formar polvo de ferrita, como la coprecipitación, pero aún no tienen uso comercial. Otras causas del rápido desarrollo de estas ferritas son el costo relativamente bajo de los materiales y su facilidad de formación. Debido a la amplia gama de composiciones, aditivos y condiciones de procesamiento, existen en el mercado muchas ferritas diferentes para una diversidad de aplicaciones. Para aplicaciones electrónicas como las telecomunicaciones, las bajas pérdidas y la permeabilidad relativamente alta permiten su empleo en circuitos LC tales como los filtros y transformadores de alta frecuencia (pulsos, acoplamiento de impedancias). Para aplicaciones de manejo de energía eléctrica a alta frecuencia, sus bajas pérdidas a niveles moderados de flujo y menor costo hacen de ellas materiales de utilidad para transformadores de potencia, transformadores de salida horizontal y bobinas de choque. Es posible elaborar algunas ferritas con relaciones de rectangularidad muy elevadas, lo que las hace aplicables para núcleos de memoria y de conmutación. Los parámetros importantes de las ferritas "blandas" son:

1. **Permeabilidad inicial**, μ_i , que varía aproximadamente de 100 a 15 000.
2. **Factor de pérdida**, $1/\mu Q$, que varía desde un valor bajo de 1×10^{-6} hasta valores mucho más elevados a frecuencias más altas.
3. **Factor de temperatura**, que varía aproximadamente de 1×10^{-6} a 10×10^{-6} .
4. **Factor de incapacidad**, que es bastante bajo, 1×10^{-6} para ferritas con elevada permeabilidad, y llega hasta 10×10^{-6} para ferritas de menor permeabilidad con bajas pérdidas.
5. **Pérdidas en el núcleo**, que suelen fijarse para una frecuencia y nivel de flujo específicos, los cuales hasta hace poco tiempo eran 16 000 Hz y 2 000 G (0.2 T), pero con la tendencia a frecuencias y niveles de flujo superiores, ya se han modificado. De mayor utilidad son las gráficas log-log que se muestran en la figura 9-3, con familias de curvas que presentan todo el intervalo de frecuencias y niveles de excitación.

La gran mejora de las ferritas con el paso de los años se debe a la mayor pureza de la materia prima y al empleo de aditivos como silicio, calcio, titanio y estaño, aunque también han sido de utilidad los hornos con

Tabla 9-3. Propiedades de materiales magnéticos permanentes

Material	Densidad de flujo de saturación, B_s		Fuerza coercitiva, H_c		Máximo producto de energía, $(BH)_{\max}$	
	G	T	Oe	A/m	G-Oe $\times 10^{-6}$	(kJ)/m ³
Acero al carbono	9 000	0.9	51	4 060	0.20	1.6
Acero al tungsteno	10 500	1.05	70	5 570	0.33	2.6
Acero al 37% de Co	10 400	1.04	230	18 300	0.98	7.8
Álnico 5	12 800	1.28	640	50 900	5.5	43.8
Álnico 9	10 500	1.05	1 500	119 000	9.0	71.6
Álnico 8	8 200	0.82	1 650	13 100	5.3	42
Ferrita de Ba orientada	2 300	0.23	3 800	30 200	3.4	27.1
Ferrita de Ba no orientada	2 300	0.23	1 860	14 800	1.05	8.36
Hierro de un solo dominio alargado	7 350	0.735	940	74 800	3.5	27.9
Cunife	5 400	0.54	550	43 800	1.5	11.9
Cunico	3 400	0.34	680	54 100	0.8	6.37
Vicalloy	9 050	0.905	415	33 000	2.3	18.3
Remalloy	8 500	0.85	345	27 500	1.2	9.55
Pt-Co	6 450	0.645	5 400	430 000	9.2	73.2
Tierras raras-cobalto	9 200	0.92	9 000	716 000	21	167
Mn-Al-C	5 200-6 000	0.52-0.6	2 000-6 000	60-48	5-6	40-48
Nd-Fe-B	11 800-12 500	1.18-1.25	9 500-11 500	750 000-915 000	33-36	263-287

mejor sistema de calentamiento, que permiten un mejor control de la temperatura y la atmósfera. Esto es especialmente cierto en el caso de las ferritas blandas más importantes, las ferritas de manganeso-cinc-hierro. El hierro en exceso en la forma de Fe^{2+} debe equilibrarse cuidadosamente con la presión parcial del oxígeno de la atmósfera.

Casi todas las ferritas con ciclo de histéresis rectangular para uso en núcleos de memoria son de magnesio-manganeso. La rectangularidad se logra mediante un ciclo de tratamiento térmico idóneo. Los materiales con rectangularidades elevadas suelen tener alta anisotropía y baja magnetostricción.

Para aplicaciones a frecuencias muy altas, las ferritas de níquel-cinc o de níquel son mejores que las de manganeso-cinc, pese a su permeabilidad más baja.

En la tabla 9-2 se presentan valores típicos de los diversos materiales descritos anteriormente.

9.4.6 Materiales amorfos

En los últimos cinco años ha habido un desarrollo rápido de varios tipos de nuevos materiales denominados metales amorfos o vidrios metálicos. Estos materiales se elaboran mediante el apagamiento o ahogamiento muy rápido de un material fundido, tan pronto como se haya conformado en una cinta delgada. Hasta ahora, las dos categorías de importancia son:

1. $M_{80}Z_{20}$, donde M es una combinación de metales en la que participan níquel, cobalto y hierro, y Z es la combinación no metálica que conforma el

vidrio. Estos materiales pueden diseñarse de modo que tengan baja magnetostricción, y se espera que se emplearán en cabezas de grabación y otras aplicaciones especializadas.

2. $Fe_{80}Z_{20}$, donde Z es el formador de vidrio, que puede ser silicio, boro, carbono o alguna combinación de ellos. Estos materiales pueden procesarse ya sea para aplicaciones de polvo hasta 60 Hz o para aplicaciones de alta frecuencia. Las pérdidas son bastante bajas, mucho menores que las del hierro-silicio. Estos materiales producirán un cambio revolucionario en el empleo de los materiales magnéticos si puede reducirse su precio a un nivel competitivo.

9.4.7 Ferritas duras

Para los materiales magnéticos permanentes, el primer cuadrante del ciclo BH no es el de mayor importancia. Dado que los imanes permanentes tienen una inducción remanente significativa, y que el imán es desmagnetizado aún más por campos externos en las aplicaciones reales, es el segundo cuadrante el de importancia. A diferencia de lo que ocurre en el caso de los materiales blandos, una alta fuerza coercitiva es de utilidad para evitar la desmagnetización; la remanencia también debe ser elevada. Una especificación bastante útil para el material magnético permanente es el producto $(BH)_{\max}$, o sea, el valor máximo del producto de la remanencia y la fuerza coercitiva posible en el segundo cuadrante. A diferencia de las ferritas blandas, las duras poseen elevadas anisotropías

y, como ya se dijo, la B_s de las ferritas es menor que la propia de la mayor parte de los metales magnéticos, pero en las ferritas duras la fuerza coercitiva es superior. La fórmula genérica de las ferritas duras es $MO \cdot 6Fe_2O_3$, donde M es bario, estroncio o plomo. Se forman mediante los mismos procedimientos cerámicos que los usados para las ferritas suaves, y existen dos variedades: las orientadas y las no orientadas. Los materiales no orientados se comprimen de la misma forma que las ferritas blandas. Los materiales orientados se alinean en un campo magnético en una condición húmeda o seca antes de comprimirlos, lo que permite que el eje de fácil magnetización se oriente en forma paralela en alto grado, con lo cual se incrementa la rectangularidad y, por tanto, el producto $(HB)_{máx}$. En la tabla 9-3 se presentan propiedades representativas de las ferritas duras.

9.4.8 Materiales metálicos magnéticos permanentes

Los criterios que se aplican para las ferritas duras también son válidos para los materiales magnéticos metálicos, que son imanes permanentes macizos en forma de barras o herraduras, y los modernos son maleables. Muchos materiales magnéticos permanentes contienen cobalto, esencialmente debido a que éste forma estructuras cristalinas de tipo hexagonal con alta direccionalidad o anisotropía paralela al eje hexagonal. Antes de los recientes aumentos en el precio del cobalto, el alnico V era el material magnético permanente de mayor uso. Se forma por medio de procesos de fundición o sinterización. Antes de que el alnico existiera en el mercado, para los imanes permanentes se empleaban algunos grados de acero al carbono, al cromo-tungsteno o al cobalto. Después se desarrollaron alnicos anisotrópicos de alto grado con productos BH de aproximadamente 9×10^6 G/Oe (72 kJ/m^3). Es posible alinear y comprimir partículas de hierro de un solo dominio alargado, (ESD, de *elongated single-domain*) para formar imanes especiales. Uno de los mejores materiales para imanes permanentes es el platino-cobalto, pero su costo es elevado.

En los últimos 15 años ha aparecido una nueva variedad de imanes metálicos que de alguna manera han revolucionado la industria de la modulación en fase (PM, de *phase modulation*), debido a sus valores superaltos de $(BH)_{máx}$, que llegan hasta 25×10^6 G-Oe (200 kJ/m^3). Estos imanes metálicos se denominan imanes de tierras raras-cobalto y, aunque todavía son algo costosos, su empleo ha proliferado recientemente. En la tabla 9-3 se muestran algunos valores comunes de estos materiales. En la categoría correspondiente a la ductilidad, las aleaciones como vicalloy y remalloy se producen mediante técnicas metalúrgicas como fusión, laminado y otras. Una innovación reciente ha sido el desarrollo de aleaciones de hierro-cobalto-cromo con menores contenidos de cobalto. Últimamente se ha producido comercialmente un nuevo material compuesto de neodimio, hierro y boro con valor $(BH)_{máx}$ de 35×10^6 G-Oe (280 kJ/m^3), y en el laboratorio con un valor correspondiente de 45×10^6 G-Oe (360 kJ/m^3).

9.4.9 Materiales para microondas

Los materiales magnéticos para microondas son sólo óxidos (ferrita y granate), dado que a frecuencias de microondas se requiere alta resistividad. Las ferritas son de níquel o de magnesio-manganeso, y los nuevos materiales de granate, cuya fórmula genérica es $3M_2O_3 \cdot 5Fe_2O_3$, han alcanzado uso a gran escala en componentes de microondas como rotadores, circuladores y defasadores.

Este tema se analiza en otro capítulo.

9.4.10 Materiales para memorias

Las ferritas con ciclo de histéresis rectangular para núcleos de memorias ya han sido analizadas. También se emplean cintas muy delgadas de permalloy arrolladas en pequeñas bobinas para aplicaciones de memorias en las que se requiere alta confiabilidad. Para fines de grabación se han utilizado alambres y cintas magnéticas, pero en la actualidad se han sustituido casi por completo por cinta plástica recubierta de óxido. El óxido es de una variedad semidura que consta de $\gamma\text{-Fe}_2\text{O}_3$, Fe_3O_4 (magnetita) o dióxido de cromo (CrO_2).

Los materiales más recientes desarrollados para memorias son los de memoria de burbuja, que son películas de granate anisotrópico desarrolladas epitaxialmente en sustratos como granates de gadolinio-galio. Los elementos de memoria son los dominios cilíndricos a través del grosor de la película. Estos dominios pueden generarse, almacenarse, moverse y leerse. Las memorias de película delgada de cobalto-níquel, cobalto-fósforo y similares se han manufacturado por electroplastia, reducción directa a partir de la solución, depósito al vacío o bombardeo iónico.

9.4.11 Materiales magnetostrictivos

Las aleaciones procesadas a fin de dotarlas de elevada magnetostricción pueden utilizarse como transductores magnetostrictivos para limpieza o labrado. Es posible emplear níquel puro, hierro-níquel al 50:50, cobalto-níquel al 95:5, o permendur 2V (49Co-49 Fe-2V), que poseen elevada magnetostricción.

LIBROS Y REVISTAS

- R. M. Bozorth, *Ferromagnetism*, Van Nostrand, Nueva York.
- B. D. Cullity, *Introduction to Magnetic Materials*, Addison-Wesley, Reading, MA.
- C. Heck, *Magnetic Materials and Their Applications*, Crane Russack, Nueva York.
- G. T. Rado y H. Suhl, *Magnetism*, Academic, Nueva York.
- J. Smit, *Magnetic Properties of Materials*, McGraw-Hill, Nueva York.
- R. S. Tebble y D. J. Craik, *Magnetic Materials*, Wiley, Nueva York.
- E. P. Wohlfarth, *Ferromagnetic Materials*, vol. 2, North-Holland, Nueva York.

CAPÍTULO 10

Materiales semiconductores

Edward H. Nicollian

University of North Carolina at Charlotte
Charlotte, North Carolina

10.1 Introducción

10.2 Propiedades eléctricas, térmicas y ópticas

- 10.2.1 Bandas de energía
- 10.2.2 Electrones y huecos
- 10.2.3 Semiconductores intrínsecos
- 10.2.4 Semiconductores extrínsecos
- 10.2.5 Concentraciones de portadores libres
- 10.2.6 Tiempo de vida
- 10.2.7 Transporte de portadores libres
- 10.2.8 Ionización por impacto

- 10.2.9 Propiedades térmicas

- 10.2.10 Propiedades ópticas

10.3 Oxidación del silicio

- 10.3.1 Defectos interfaciales y carga de óxido

10.4 Epitaxia, difusión e implantación de iones

- 10.4.1 Crecimiento epitaxial
 - 10.4.2 Difusión
 - 10.4.3 Protección de óxido
 - 10.4.4 Implantación de iones
-

10.1 INTRODUCCIÓN

Los semiconductores elementales y compuestos se utilizan en dispositivos conforme a sus propiedades materiales. El semiconductor básico que se emplea con mayor frecuencia en los circuitos integrados es el silicio (Si). Una razón primordial del predominio del silicio en su aplicación a los circuitos integrados es que sobre éstos puede desarrollarse con facilidad una película de dióxido de silicio delgada, químicamente estable y eléctricamente neutra. Además, su separación energética de 1.12 eV asegura que la densidad de portadores libres o portadores de carga* no depende de la temperatura en el intervalo común (25 a 125 °C) durante la operación de circuitos integrados.

En este capítulo se hace hincapié en el silicio debido a que, con mucho, es el semiconductor con mayor impacto económico. Sin embargo, también se analizan el semiconductor elemental de germanio (Ge) y el semiconductor compuesto de arseniuro de galio (GaAs). En la tabla 10-1 se presenta un resumen de muchas de las propiedades del Si, Ge y GaAs. El germanio incluso se utiliza en aplicaciones comerciales.

Los semiconductores compuestos se usan principalmente en aplicaciones fotónicas y de microondas. Por ejemplo, el GaAs se emplea en dispositivos fotónicos debido a que tiene una separación energética directa (el Si y el Ge son semiconductores con separación energética indirecta) y en dispositivos de microondas debido a la gran movilidad de sus electrones.

En las referencias bibliográficas 1, 2 y 3 se abordan con mayor detalle todos estos temas.

10.2 PROPIEDADES ELÉCTRICAS, TÉRMICAS Y ÓPTICAS

10.2.1 Bandas de energía

El resultado más importante que se obtiene al aplicar a los electrones de un semiconductor los principios de la mecánica cuántica es que a dichos electrones sólo se les permiten ciertos niveles de energía agrupados en bandas con una "separación energética" o "vacío energético" entre ellas. En el silicio, las dos bandas permitidas que participan en el proceso de conducción eléctrica son la "banda de valencia" (en su mayor parte constituida por electrones) y la "banda de conducción" (vacía casi toda) superior. Las concentraciones de portadores libres en las dos bandas dependen de la densidad de los niveles de energía permitidos y de la probabilidad de qué niveles de energía dados estén ocupados por electrones.

Los electrones en la órbita más externa de cada átomo de silicio se denominan electrones de valencia. Cada átomo de silicio tiene cuatro electrones de valencia, que comparte con cuatro átomos vecinos de la malla formando enlaces covalentes.

*Un portador libre o de carga es un electrón o un hueco que tiene libertad de movimiento por todo el cristal semiconductor.

Cada enlace covalente contiene un par de electrones de valencia con spin opuesto. En los enlaces covalentes, los electrones de valencia están "localizados" en átomos de la malla de silicio.

En la figura 10-1 se presenta la dependencia de la energía de los portadores libres (o de carga) con respecto al vector de onda† o cantidad de movimiento. La separación energética E_g es la mínima separación entre las bandas de conducción y de valencia.

Tanto el Si como el Ge son semiconductores de separación energética indirecta debido a que el mínimo de la banda de conducción no ocurre para el mismo vector de onda que el máximo de la banda de valencia. Por consiguiente, las transiciones entre las bandas de valencia y de conducción requieren la absorción o emisión de fonones‡ a fin de conservar la cantidad de movimiento.

Un ejemplo de semiconductor de separación energética directa es el GaAs. En este semiconductor, el mínimo de la banda de conducción ocurre a la misma cantidad de movimiento a la que ocurre el máximo de la banda de valencia. Las transiciones entre las bandas de valencia y de conducción pueden ocurrir por absorción y emisión de fotones, hecho que es importante en muchos dispositivos fotónicos.

La masa de los portadores libres es importante para el transporte eléctrico en los semiconductores y, por ello, en la operación de muchos dispositivos. La masa de un portador libre en un semiconductor es diferente de la masa de un electrón libre en el vacío debido a las fuerzas ejercidas sobre el portador por los núcleos iónicos de la malla. La masa de un portador libre en un semiconductor se denomina "masa efectiva" para distinguirla de la masa de un electrón libre en el vacío. La masa efectiva de un portador libre en un semiconductor depende del recíproco de la curvatura § de las curvas de energía con respecto al vector de onda (fig. 10-1).

En la figura 10-1 se observa que las bandas de valencia del Ge, Si y GaAs constan de varias subbandas. Cerca del máximo de la banda de valencia (mínima energía de huecos) existe una subbanda de huecos grandes (la banda más ancha con la menor curvatura) y otra de huecos pequeños (la banda más estrecha con la mayor curvatura).

Estas dos bandas tienen la misma curvatura en el máximo de la banda de valencia. En la mayor parte de las aplicaciones en circuitos integrados sólo participan huecos en el máximo de la banda de valencia, de modo que no existe diferencia entre los huecos grandes y los pequeños.

En la figura 10-2 se muestran las energías mínimas de las separaciones energéticas y su variación con la temperatura para el GaAs, Si y Ge. La variación de la separación energética $E_g(T)$ con la temperatura satisface la relación empírica aproximada⁵:

† Un vector de onda es la cantidad de movimiento dividida entre $h/2\pi$, donde h es la constante de Planck.

‡ Un fonón es el cuanto de energía de vibración de la malla. § La curvatura es la segunda derivada de la energía con respecto al vector de onda.

Tabla 10-1. Propiedades del germanio, silicio y arseniuro de galio a 300 °K

Propiedades	Ge	Si	GaAs
Átomos/cm ³	4.42×10^{22}	5.0×10^{22}	4.42×10^{22}
Calor específico (J/g-°C)	0.31	0.7	0.35
Coeficiente de dilatación lineal $\Delta L/L_0 \Delta T (^{\circ}\text{C}^{-1})$	5.8×10^{-6}	2.6×10^{-6}	6.86×10^{-6}
Concentración intrínseca de portadores (cm ⁻³)	2.4×10^{13}	1.45×10^{10}	1.79×10^6
Conductividad térmica (W/cm-°C)	0.6	1.5	0.46
Constante dieléctrica	16.0	11.9	13.1
Constante de la malla (Å)	5.64613	5.43095	5.6533
Densidad (g/cm ³)	5.3267	2.328	5.32
Densidad efectiva de estados en la banda de conducción, M_c (cm ⁻³)	1.04×10^{19}	2.8×10^{19}	4.7×10^{17}
Densidad efectiva de estados en la banda de valencia, M_v (cm ⁻³)	6.0×10^{18}	1.04×10^{19}	7.0×10^{18}
Difusividad térmica (cm ² /s)	0.36	0.9	0.44
Electrones	$m_i^* = 0.082$ $m_{ih}^* = 0.044$	$m_i^* = 0.19$ $m_{ih}^* = 0.16$	$m_{ih}^* = 0.082$
Energía de fonones ópticos (eV)	0.037	0.063	0.035
Estructura cristalina	Diamante	Diamante	Esfalerita o blenda de cinc
Huecos	$m_{hh}^* = 0.28$	$m_{hh}^* = 0.49$	$m_{ih}^* = 0.45$
Longitud intrínseca de Debye (μm)	0.68	24	2 250
Masa efectiva, m^*/m_0	$m_i^* = 1.64$	$m_i^* = 0.98$	0.067
Movilidad (deriva) (cm ² /V-s)			
Electrones	3 900	1 500	8 500
Huecos	1 900	450	400
Peso atómico	72.60	28.09	144.63
Presión del vapor (Pa)	1 a 1 330 °C 10^{-6} a 760 °C	1 a 1 650 °C 10^{-6} a 900 °C	100 a 1 050 °C 1 a 900 °C
Punto de fusión (°C)	937	1 415	1 238
Resistividad intrínseca (Ω-m)	47	2.3×10^5	10^8
Rigidez dieléctrica de campo (V/cm)	$\sim 10^5$	$\sim 3 \times 10^5$	$\sim 4 \times 10^5$
Separación energética (eV)	0.66	1.12	1.424
Tiempo(s) de vida de los portadores minoritarios (puros)	10^{-3}	2.5×10^{-3}	$\sim 10^{-8}$
Trayectoria libre media de los fonones λ_0 (Å)	105	76(electrón) 55(hueco)	58

Fuente. Reproducido con permiso de Sze¹.

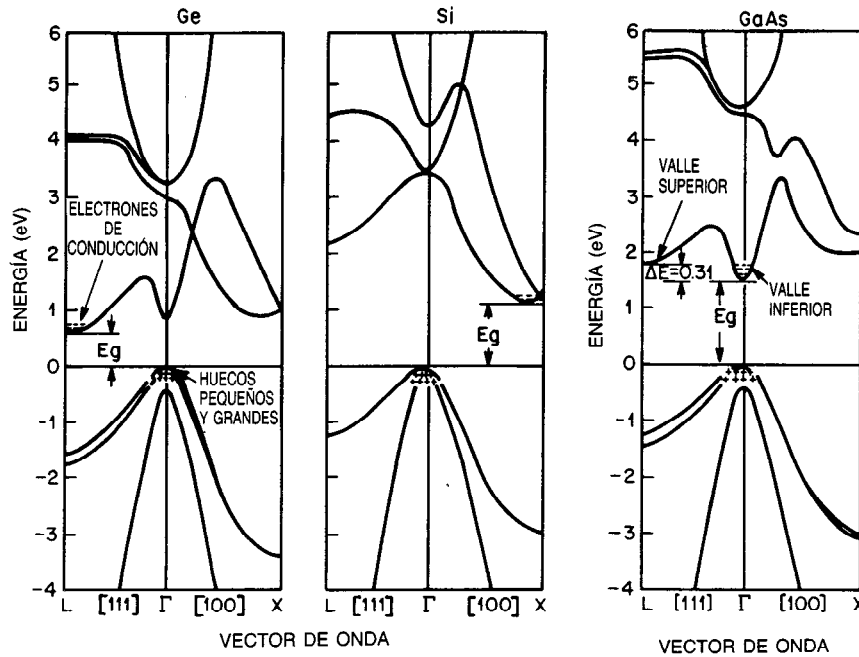


Fig. 10-1. Energía de los portadores libres contra cantidad de movimiento del germanio, silicio y arseniuro de galio. En la tabla 10-1 se muestran los valores de E_g . Los signos (+) indican huecos en las bandas de valencia, y los signos (-) indican electrones en las bandas de conducción. (Tomado de Chelikowsky y Cohen⁴.)

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta} \text{ (eV)} \quad (10-1)$$

donde $E_g(0)$ es la energía de la separación energética a 0 °K, α y β están dadas en la figura 10-2 y T es la temperatura absoluta. El coeficiente de temperatura dE_g/dT , determinado por medición, es negativo para los tres semiconductores, aunque el de algunos es positivo, como en el caso del sulfuro de plomo (PbS), para el cual la energía de la separación energética aumenta de 0.28 eV a 0 °K hasta 0.41 eV a 300 °K. Las separaciones energéticas del Ge y del GaAs medidas a una temperatura próxima a la ambiente aumentan con la presión⁶ [$dE_g/dP = 5 \times 10^{-6}$ eV/(kg/cm²) para el Ge y aproximadamente 12.6×10^{-6} eV/(kg/cm²) para el GaAs]. La energía medida de la separación energética del Si disminuye con la presión, $dE_g/dP = -2.4 \times 10^{-6}$ eV/(kg/cm²).

En las referencias bibliográficas 1 y 7 se analizan detalles de la estructura de la banda de energía de los semiconductores.

10.2.2 Electrones y huecos

Un electrón excitado (p. ej., mediante agitación térmica) dentro de la banda de conducción es una carga libre negativa que se mueve a través de todo el cristal bajo la influencia de gradientes de temperatura o potencial. La deficiencia de carga de valencia producida en la banda de conducción por la excitación de un electrón es una carga positiva, denominada hueco o deficiencia

electrónica, que es libre para desplazarse por todo el cristal bajo la influencia de los gradientes mencionados.

10.2.3 Semiconductores intrínsecos

En un semiconductor puro existen varios pares hueco-electrón a una temperatura dada, y la concentración de electrones n es igual a la concentración de huecos p . Estas concentraciones son intrínsecas al semiconductor e iguales a la concentración intrínseca, n_i (en la tabla 10-1 y en la figura 10-3 se presentan algunos valores relacionados con ello).

Nivel de Fermi. La distribución más probable de energía de un conjunto de portadores libres en equilibrio térmico, sujetos al principio de exclusión de Pauli*, es la distribución de Fermi. La probabilidad f_n de que un nivel de energía E esté ocupado en equilibrio térmico por un electrón está dada por la función de Fermi-Dirac⁸:

$$f_n = \left(1 + \exp \frac{E - E_F}{kT} \right)^{-1} \quad (10-2)$$

donde k es la constante de Boltzmann (8.617×10^{-5} eV/K) y E_F es el nivel de Fermi, la energía a la cual un

* El principio de exclusión de Pauli establece que dos electrones no pueden tener el mismo estado cuántico, y que todos los electrones son exactamente iguales.

estado cuántico –si está presente– tiene el 50% de posibilidades de estar ocupado por un electrón.

La probabilidad f_p de que un nivel E se encuentre vacante o de que esté ocupado por un hueco es $f_p = 1 - f_n$ o bien:

$$f_p = \left(1 + \exp \frac{E_F - E}{kT} \right)^{-1} \quad (10-3)$$

Concentraciones de electrones y huecos. La concentración cero de electrones en el campo aplicado (cm^{-3}) en la banda de conducción es³:

$$n = \int_{E_c}^{\infty} M_c(E) f_n(E - E_F) dE = M_c \exp \left[- \frac{E_c - E_F}{kT} \right] \quad (10-4)$$

donde $M_c(E)$ es la densidad de estados en la banda de conducción (cm^{-3}/J), E_c es el borde de la banda de conducción y $M_c = 5.4 \times 10^{15} T^{3/2} (\text{cm}^{-3})$ para el silicio (en la tabla 10-1 se indican los valores para Ge y

GaAs a 300 °K). Análogamente, la concentración cero de huecos en el campo aplicado (cm^{-3}) en la banda de valencia es³:

$$p = \int_{-\infty}^{E_v} M_v(E) f_p(E_F - E) dE = M_v \exp \left[- \frac{E_F - E_v}{kT} \right] \quad (10-5)$$

donde $M_v(E)$ es la densidad de estados en la banda de valencia (cm^{-3}/J), E_v es el borde o límite de la banda de valencia y $M_v = 2.0 \times 10^{15} T^{3/2} (\text{cm}^{-3})$ para el silicio (en la tabla 10-1 se presentan los valores para Ge y GaAs a 300 °K).

La posición del nivel de Fermi en la separación energética para un semiconductor intrínseco se obtiene por la condición de neutralidad de carga $n = p$ (ref. 3):

$$E_F = E_g/2 + \frac{kT}{2} \ln \frac{M_v}{M_c} = E_i \quad (10-6)$$

donde E_i es el nivel intrínseco de Fermi.

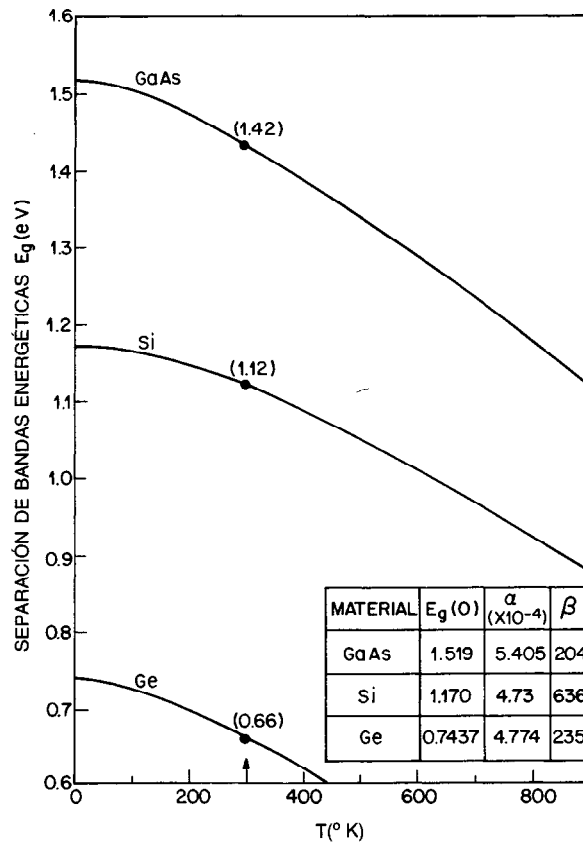


Fig. 10-2. Separaciones energéticas del germanio, silicio y arseniuro de galio en función de la temperatura. (Reproducido de Thurmond⁵, con permiso del editor, The Electrochemical Society, Inc.)

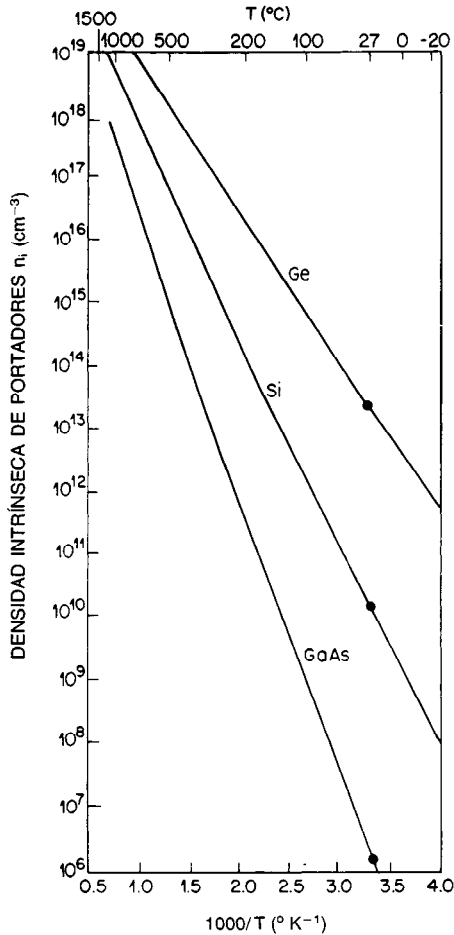


Fig. 10-3. Dependencia de la temperatura respecto de las densidades intrínsecas de los portadores del germanio, silicio y arseniuro de galio. (Reproducido de Thurmond⁵, con permiso del editor, The Electrochemical Society, Inc.)

Dado que $p = n = n_i$ en un semiconductor intrínseco, $pn = n_i^2$. Por tanto³:

$$n_i = (M_c M_v)^{1/2} \exp \left[- \frac{E_g}{2kT} \right] \quad (10-7)$$

En la figura 10-3 se muestra la dependencia de n_i respecto a la temperatura.

10.2.4 Semiconductores extrínsecos

Donadores y aceptores. Cuando se agregan ciertas impurezas (o "impurificadores") elementales al semiconductor, éste se vuelve un semiconductor "extrínseco". Las impurezas se agregan para modificar las densidades de portadores libres del semiconductor intrínseco, controlando la conductividad eléctrica y permitiendo la producción de uniones pn y varias estructuras de transistores unipolares y bipolares.

Los impurificadores efectivos son impurezas elementales que entran por sustitución en la malla cristalina del semiconductor; es decir, toman el lugar de un átomo de silicio en la malla, y tienen niveles de energía muy someros en la separación energética. Debido a que la valencia del silicio es 4, la adición por sustitución de elementos del grupo V de la tabla periódica, como el fósforo, antimonio y arsénico, da por resultado la unión débil de un electrón extra al átomo de la impureza. Este electrón extra ocupa un nivel de energía en la separación energética muy cercano al borde de la banda de conducción (0.05 eV), y a temperaturas normales se excita térmicamente de modo que pasa hacia la banda de conducción, dejando tras de sí un átomo de impureza cargado positivamente, denominado donador. La adición por sustitución al silicio de elementos del grupo III de la tabla periódica, tales como el boro y aluminio, da por resultado la aceptación de un electrón extra de la banda de valencia para completar los cuatro electrones de valencia compartidos con los átomos circundantes de silicio. A temperaturas ordinarias, un electrón se excita térmicamente y sale de la banda de valencia al nivel de energía *somero* (0.05 eV por encima del borde de la banda de valencia) de la impureza, creando un hueco en la banda de valencia y dejando un átomo de impureza cargado negativamente, llamado aceptor. Cuando la concentración del donador N_D (cm^{-3}) es mucho mayor que n_i y la concentración del aceptor N_A (cm^{-3}) es cero, $n = N_D$ y el semiconductor es tipo n . Cuando $N_A \gg n_i$ y $N_D = 0$, $p = N_A$ y el semiconductor es tipo p . Cuando están presentes donadores y aceptores el semiconductor se compensa, pero si $N_D > N_A$, entonces es tipo n , y si $N_A > N_D$, entonces es tipo p .

10.2.5 Concentraciones de portadores libres

La posición del nivel de Fermi en la separación energética de un semiconductor tipo n es³:

$$E_F = E_i + kT \ln \left[\frac{N_D - N_A}{n_i} + \frac{n_i}{N_D - N_A} \right] \quad (10-8)$$

y en la de un semiconductor tipo p es³:

$$E_F = E_i - kT \ln \left[\frac{N_A - N_D}{n_i} - \frac{n_i}{N_A - N_D} \right] \quad (10-9)$$

En la figura 10-4 se muestra $E_F - E_i$ en función de la temperatura para el silicio, con N_D y N_A como parámetros.

En un semiconductor no degenerado*, la concentración de portadores libres es constante, independientemente de la temperatura, siempre que las impurezas permanezcan ionizadas y que su concentración sea mayor que n_i . A temperatura ambiente n_i es pequeño en comparación con la concentración de impurezas. Sin

* El semiconductor es degenerado³ si está tan impurificado que $n \approx M_c$ o $p \approx M_v$, y el nivel de Fermi se encuentra cerca del borde de la banda de conducción o del de la banda de valencia.

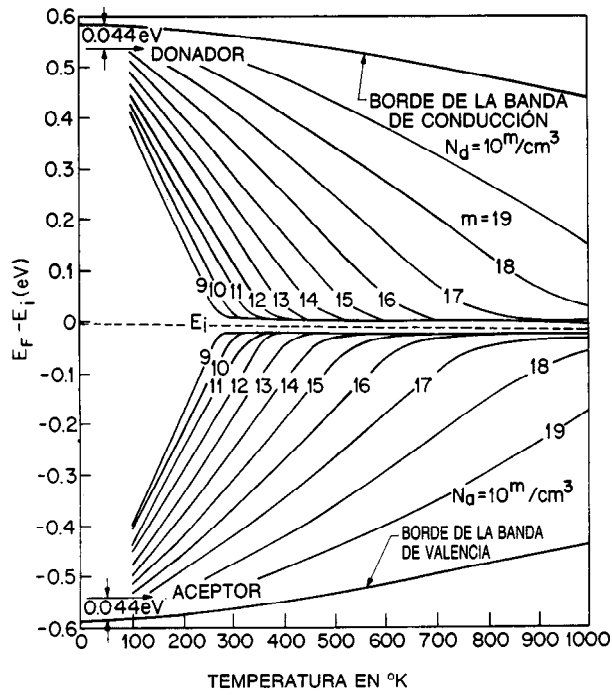


Fig. 10-4. Dependencia del nivel de Fermi (medido en términos de la separación central) respecto a la temperatura, con la concentración de impurezas como parámetro. También se muestra $E_g(T)$ de la ecuación 10-1. (Reproducido con permiso de Beadle y colaboradores⁹. American Telephone and Telegraph Company.)

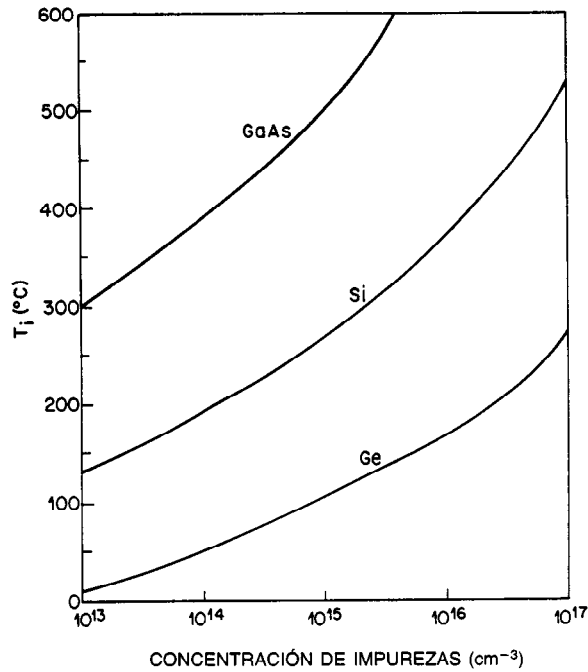


Fig. 10-5. Temperatura intrínseca como función de la concentración de impurezas. (Tomado de Sze¹.)

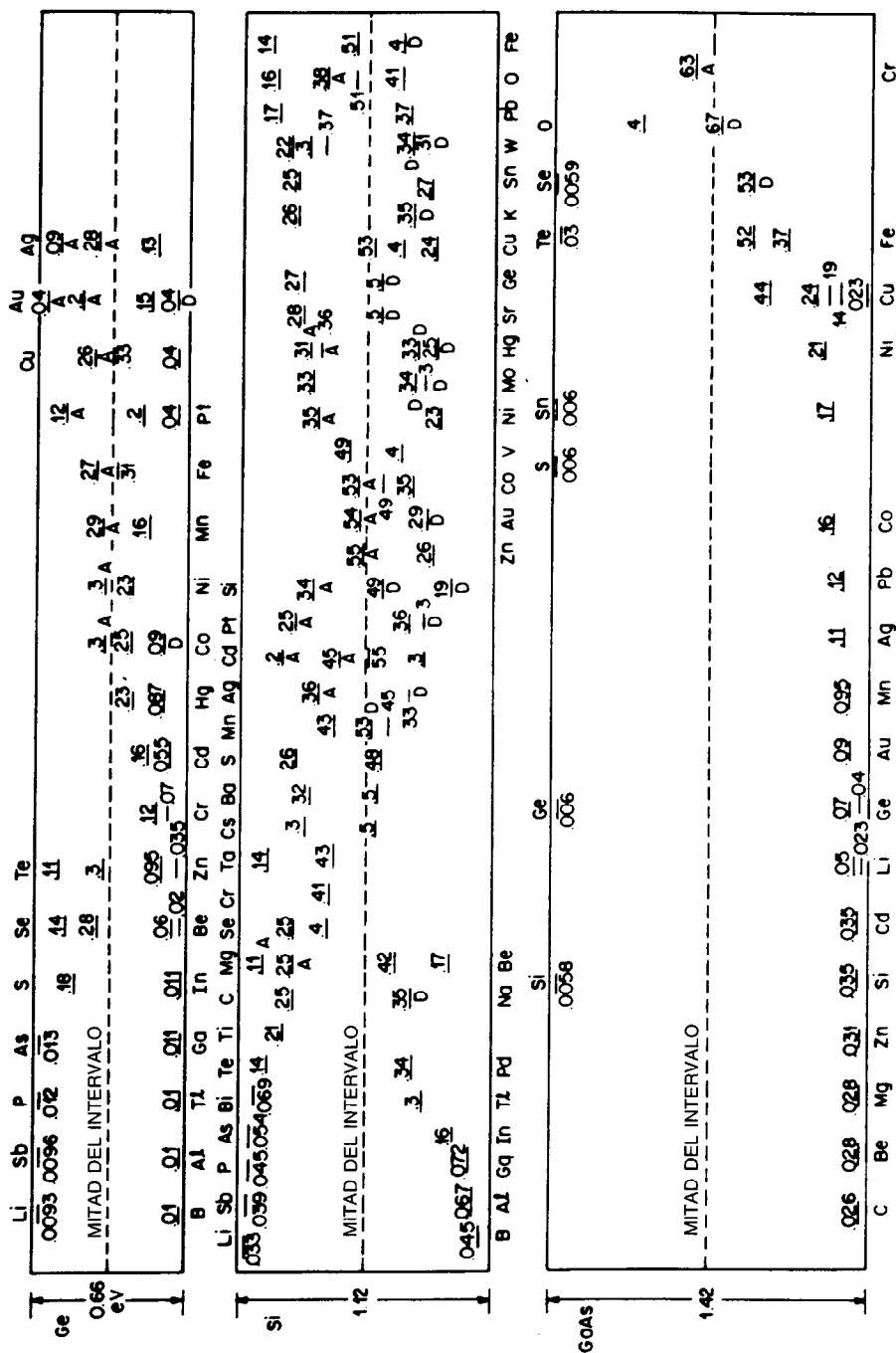


Fig. 10-6. Niveles medidos de energía de ionización para diversas impurezas. Los niveles de energía de la mitad inferior del intervalo de separación energética se miden a partir del borde de la banda de valencia y son del tipo aceptor, a menos que se marquen con D, correspondiente al tipo donador. Los niveles de energía de la mitad superior del intervalo se miden a partir del borde de la banda de conducción y son del tipo donador, a menos que se marquen con A, correspondiente al tipo aceptor. Las separaciones energéticas están a 300 °K. (Tomado de Sze.)

embargo, n_i crece exponencialmente con la temperatura, como se observa en la figura 10-3, por ejemplo, y se duplica cada 11 °C para el silicio. Por consiguiente, existe una temperatura T_i a la cual la concentración de portadores libres es igual a n_i . En la figura 10-5 se representa gráficamente T_i con respecto a la concentración de impurezas calculada para el GaAs, el Si y el Ge. Por encima de T_i la concentración de portadores libres es determinada principalmente por n_i , el semiconductor se vuelve intrínseco y la concentración de portadores libres se eleva exponencialmente con la temperatura, como se observa en la figura 10-3. Debajo de T_i la concentración de portadores libres es determinada por la concentración de impurezas, el semiconductor es extrínseco y la concentración de portadores es relativamente independiente de la temperatura. A medida que la temperatura disminuye aún más, la concentración de portadores permanece constante, hasta que a temperatura relativamente baja las impurezas comienzan a hacerse neutras y la concentración de portadores libres empieza a disminuir con la temperatura decreciente.

El estado de carga de un donador o aceptor depende de la posición del nivel de Fermi con respecto a aquél. Por encima de E_F la carga de los donadores es positiva, pero por debajo es neutra. De manera similar, por debajo de E_F la carga de los aceptores es negativa, pero por encima es neutra. Debido a que los niveles de las impurezas de uso común son muy poco profundos, los donadores y los aceptores permanecen ionizados en cierto intervalo de temperatura, que depende de la concentración de impurezas, como puede observarse en la figura 10-4. Por ejemplo, cuanto más ligera sea la impurificación más baja será la temperatura antes de que el nivel de Fermi se aproxime lo suficiente al nivel de impurificación para que ocurra una desionización significativa. Para concentraciones de impurezas menores de 10^{15} cm^{-3} , los donadores o los aceptores permanecerán ionizados hasta descender a los 45 o 50 °K. La mayor parte de los dispositivos de semiconductores son operados en el intervalo extrínseco de temperatura.

En la figura 10-6 se muestran niveles de energía medidos para varias impurezas en Ge, Si y GaAs. Los elementos elegidos como impurezas tienen niveles de energía someros, de tal modo que el nivel de Fermi no cruza el nivel de impurezas, lo que mantiene constante la concentración de impurezas ionizadas en el amplio intervalo de temperatura y polarización. Sin embargo, no todos los elementos con niveles de energía someros suelen emplearse como impurezas. Otras consideraciones importantes para la elección de impurezas son la solubilidad y difusividad en sólidos del elemento en el semiconductor huésped, como se analiza en el artículo 10.4.

No suelen introducirse de manera deliberada impurezas de nivel profundo, como los metales pesados que se muestran en la figura 10-6. Por el contrario, existen como impurezas residuales a concentraciones varios órdenes de magnitud inferiores a las de las impurezas introducidas en semiconductores "grado dispositivos".

10.2.6 Tiempo de vida

El tiempo de vida de los transportadores libres en un semiconductor es importante para la aplicación en dispositivos, ya que es el tiempo requerido para que la concentración de esos transportadores regrese al equilibrio térmico* después de que ha ocurrido una perturbación, como un cambio repentino del campo eléctrico aplicado. Tanto la concentración de electrones como la de huecos deben cambiar para restablecer el equilibrio térmico. Una forma importante en que se realiza ese restablecimiento es mediante la captura o emisión de portadores libres mediante impurezas de metales pesados de nivel profundo, como el oro y cobre en el silicio.

Para cambios pequeños del equilibrio térmico, el tiempo de vida de los huecos está dado por³:

$$\tau_p = (c_p n_T)^{-1} \quad (\text{s}) \quad (10-10)$$

y el de los electrones por³:

$$\tau_n = (c_n n_T)^{-1} \quad (\text{s}) \quad (10-11)$$

donde c_p y c_n son las probabilidades de captura de huecos y electrones (cm^{-3}/s), respectivamente, y n_T es la densidad volumétrica (cm^{-3}) de los niveles de impurezas profundos.

Los tiempos de vida típicos, en el dispositivo silicio grado varían del intervalo de los milisegundos al de los microsegundos, y pueden modificarse haciendo variar n_T . En particular, el oro se agrega al silicio para acortar el tiempo de vida en algunas aplicaciones. En otras aplicaciones, n_T se reduce mediante la eliminación de impurezas³ para alargar ese tiempo.

En la referencia 3 se describen métodos para medir el tiempo de vida.

10.2.7 Transporte de portadores libres

Movilidad volumétrica. En la aplicación de un campo eléctrico a través de un semiconductor, se superpone una "velocidad de deriva" al movimiento térmico de los portadores libres. Para campos eléctricos bajos, la velocidad de deriva promedio \bar{v}_d de un portador libre es directamente proporcional a la intensidad del campo eléctrico F (V/cm) o bien¹:

$$\bar{v}_d = \mu F \quad (\text{cm/s}) \quad (10-12)$$

donde μ (en $\text{cm}^2/\text{V-s}$), la constante de proporcionalidad definida como la velocidad de deriva promedio por unidad de campo eléctrico, se denomina movilidad o movilidad de conducción. La movilidad de los portadores libres es importante para el rendimiento de transistores bipolares y unipolares.

En un semiconductor policristalino, el movimiento de los portadores libres es entorpecido drásticamente por las fronteras granulares, lo cual reduce la movi-

* Un sistema se encuentra en equilibrio térmico cuando todos sus parámetros macroscópicos son invariantes en el tiempo.

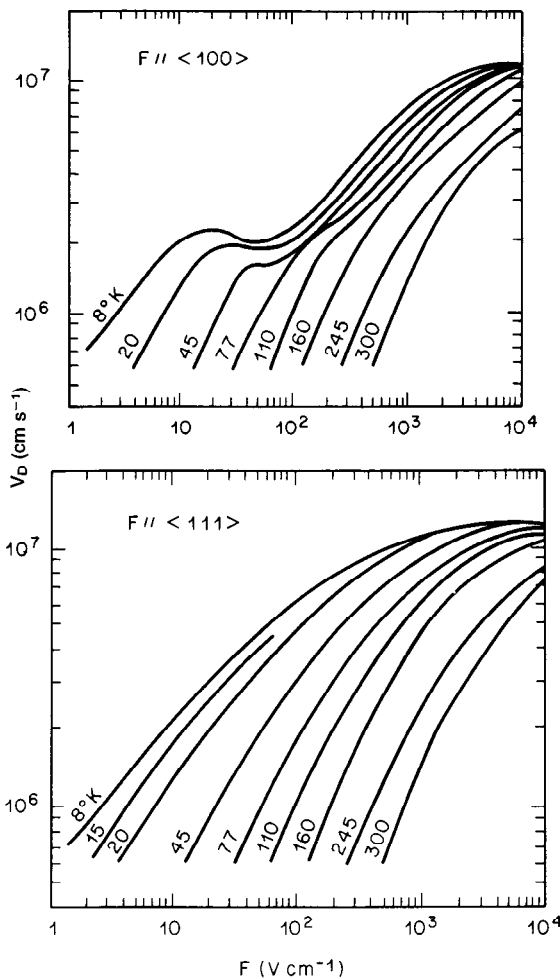


Fig. 10-7. Velocidad de deriva de electrones medida en función del campo eléctrico. (Tomado de Canali y colaboradores¹⁰.)

dad de los portadores libres por debajo del valor característico del semiconductor. Para evitar esta degradación, los transistores se elaboran con un material semiconductor monocristalino.

Los portadores libres ganan energía del campo eléctrico aplicado y a continuación pierden algo de ella debido a la dispersión, lo cual da por resultado una velocidad de deriva finita. Los dos mecanismos primordiales de dispersión en los semiconductores no polares monocristalinos, como Ge y Si, son la dispersión por impurezas y la dispersión en la malla. La dispersión por impurezas ocurre cuando los portadores libres son desviados por iones de impurezas, y su probabilidad depende de la concentración de impurezas ionizadas. La dispersión en la malla ocurre cuando los portadores son desviados por las vibraciones térmicas de los átomos de la malla, y su probabilidad depende de la temperatura. La dispersión por impurezas es el mecanismo dominante para campos débiles, por de-

bajo de la temperatura ambiente, o para altas concentraciones de impurezas. La dispersión en la malla es el mecanismo dominante para campos débiles y a temperaturas igual o mayor que la del ambiente, excepto a altas concentraciones de impurezas. La velocidad de deriva también depende del campo eléctrico y de la orientación cristalográfica.

En la figura 10-7 se muestra la velocidad de deriva medida para un electrón en silicio de alta pureza, en función del campo eléctrico aplicado paralelo a la dirección <100> y paralelo a la dirección <111>, con la temperatura como parámetro. La estructura de las curvas <100> está relacionada con los desplazamientos de la población entre valles en la banda de conducción. A medida que se incrementa el campo eléctrico y \bar{v}_d se vuelve comparable con la velocidad térmica, \bar{v}_d deja de ser directamente proporcional a F , tal como lo predice la ecuación 10-12.

En la figura 10-8 se muestra la dependencia respecto al campo eléctrico de la velocidad medida de deriva de los portadores de carga. La disminución de \bar{v}_d para el GaAs en campos intensos puede explicarse recurriendo a la figura 10-1. A medida que se incrementa el campo eléctrico, la banda de conducción de los electrones se excita desde el valle inferior de alta movilidad (curvatura más pequeña) hasta el valle superior de baja movilidad (curvatura más grande), incrementando la población de electrones del valle superior a expensas de la población del valle inferior. En campos intensos la población de electrones en el valle superior se hace importante, provocando que la velocidad global de deriva disminuya. El efecto de los electrones transferidos provoca una resistencia diferencial negativa, lo que ha sido aprovechado en el GaAs y en el fosforo de indio (InP) para crear amplificadores de potencia de microondas y osciladores locales. En la referencia bibliográfica 1 se hace un análisis más completo del efecto de los electrones transferidos.

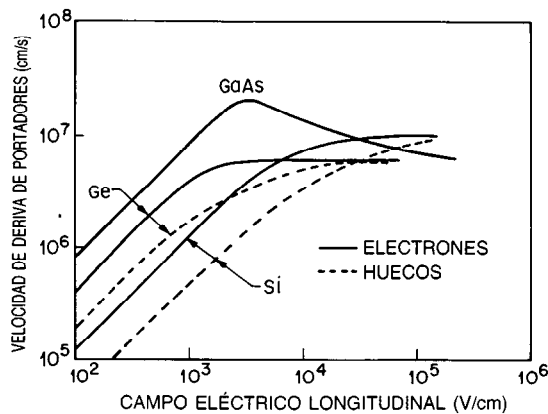


Fig. 10-8. Velocidad de deriva de portadores libres medida para la dirección <111> de arseniuro de galio, germanio y silicio a 300 °K. Para una impurificación intensa, la velocidad de deriva en campos de baja intensidad es menor que la mostrada. La velocidad de deriva en campos de gran intensidad es independiente de la concentración de impurezas. (Tomado de Sze¹.)

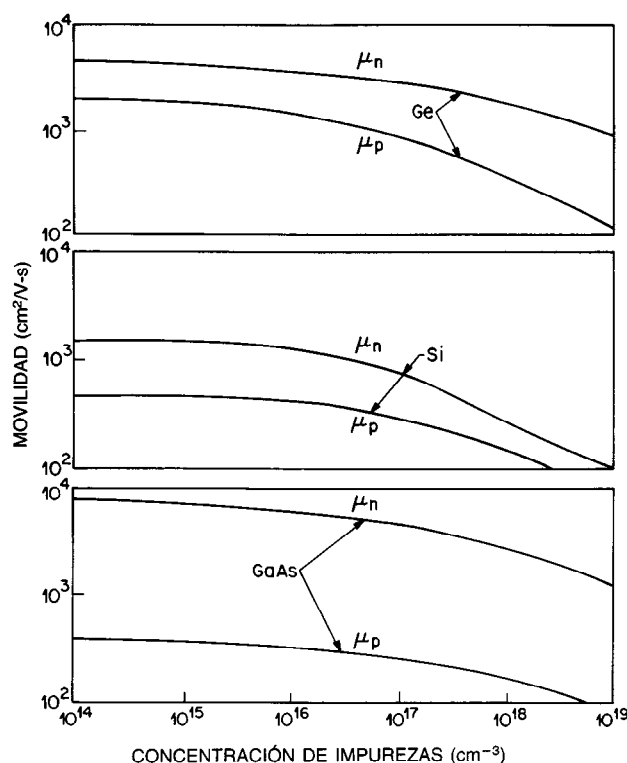


Fig. 10-9. Dependencia de la movilidad de deriva de portadores libres respecto a la concentración de impurezas ionizadas. (Tomado de Sze¹.)

En la figura 10-9 se presentan las movilidades de electrones y huecos en la dirección $\langle 111 \rangle$ para el Ge, Si y GaAs medidas a 300 °K como funciones de la concentración de impurezas ionizadas en campos de baja intensidad. Aquí se aplica la ecuación 10-12. La velocidad de deriva depende muy poco de la concentración en el caso de campos de baja intensidad y a temperatura ambiente, de modo que la movilidad para $\langle 100 \rangle$ es aproximadamente la misma que para $\langle 111 \rangle$. En la figura 10-10 se presenta la dependencia con respecto a la temperatura de las movilidades de electrones y huecos en la dirección $\langle 111 \rangle$ del silicio para campos de baja intensidad en los que se aplica la ecuación 10-12. Para concentraciones menores de impurezas, las pendientes de estas curvas se desvían ligeramente de las pendientes predichas por la dispersión en la malla y por impurezas, debido a otros mecanismos de dispersión¹.

Para campos de baja intensidad, el intervalo entre los choques t_c es independiente del campo eléctrico aplicado, y la relación entre la movilidad y la masa efectiva m^* es²:

$$\mu = \frac{qt_c}{2m^*} \quad (10-13)$$

donde q es la carga electrónica en C. En la tabla 10-1 se presentan las relaciones m^*/m_0 para electrones y para huecos pequeños y grandes, donde m_0 es la masa de un electrón libre en el vacío.

La difusividad D de los portadores libres en un gradiente de concentración del semiconductor está relacionada con su movilidad por medio de la relación de Einstein¹:

$$D = \frac{kT}{q} \mu \quad (\text{cm}^2/\text{s}) \quad (10-14)$$

Movilidad superficial. En un transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET, *metal oxide semiconductor field effect transistor*), los portadores de carga de la capa de inversión también son dispersados por vibraciones de la malla y por impurezas ionizadas. Sin embargo, los portadores libres de la capa de inversión están confinados en un pozo de potencial en la superficie del semiconductor. La movilidad de los portadores de carga en este pozo también es afectada por la anchura de la capa de inversión y por la dispersión debida a las cargas en el óxido próximo a la interficie óxido-semiconductor. Estos efectos extra reducen la movilidad de los portadores libres en el pozo por debajo de la movilidad volumétrica.

La movilidad de los portadores libres de la capa de inversión se denomina movilidad superficial, para distinguirla de la movilidad volumétrica superior.

La anchura de la banda de inversión depende del campo eléctrico transversal F_x aplicado entre la compuerta y la fuente. Esta dependencia hace de la movili-

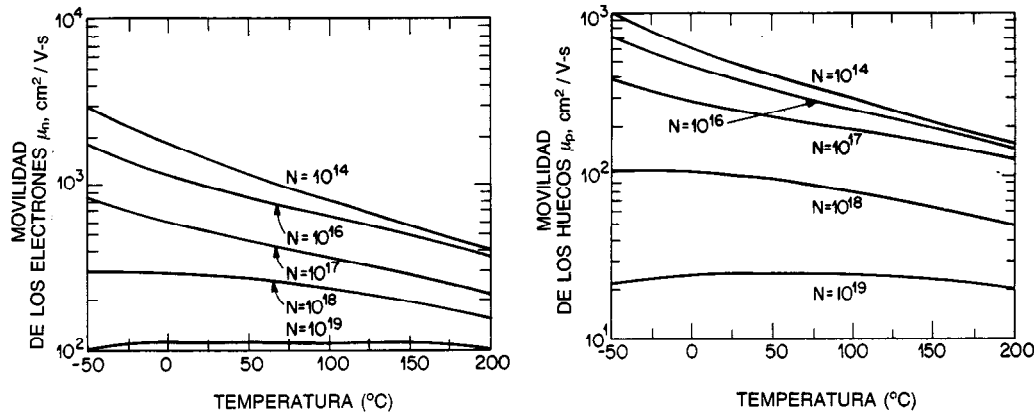


Fig. 10-10. Dependencia respecto a la temperatura de la movilidad de los electrones y huecos en la dirección $\langle 111 \rangle$ del silicio en campos de baja intensidad. (Reproducido con permiso de Beadle y colaboradores⁹. American Telephone and Telegraph Company.)

dad superficial, que depende de la anchura de la capa de inversión, una función de F_x . En la figura 10-11 se muestra la movilidad superficial del electrón medida como función del campo eléctrico transversal efectivo, con la temperatura como parámetro.

La velocidad de deriva de los portadores libres en la capa de inversión también depende del campo eléctrico longitudinal F_y aplicado entre la fuente y el consumo. Para un F_x dado y un F_y bajo, la velocidad de deriva es directamente proporcional a F_y , como se observa en la ecuación 10-12. Sin embargo, a medida que

aumenta F_y la velocidad de deriva tiende a saturarse a campos más bajos que la volumétrica. En la figura 10-12 se presenta la velocidad de deriva medida de los electrones de la capa de inversión en función de F_y , con F_x como parámetro. Se incluye, además, una curva para la velocidad de deriva volumétrica.

Resistividad. La resistividad de un semiconductor es²:

$$R = \rho \frac{L}{A} \quad (\Omega) \quad (10-15)$$

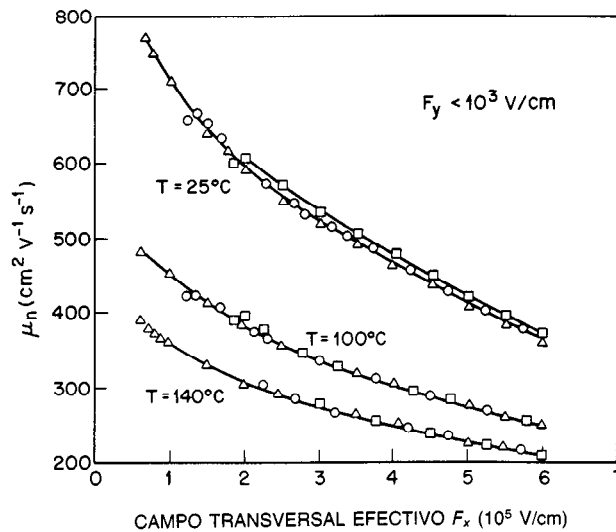


Fig. 10-11. Dependencia de la movilidad superficial de los electrones en la dirección $\langle 100 \rangle$ del silicio tipo p respecto al campo eléctrico transversal efectivo con baja F_y , con la temperatura como parámetro. El campo eléctrico transversal efectivo se define como el campo eléctrico transversal promediado sobre la distribución de electrones en la capa de inversión, y está dado por $(F_x)_{eff} = \epsilon_s^{-1}(Q_B + \frac{1}{2} Q_n)$, donde ϵ_s es la permitividad dieléctrica (C/V-cm), Q_B es la densidad de carga en la capa de agotamiento (C/cm²) y Q_n es la densidad de carga en la capa de inversión (C/cm²). (Tomado de Sabnis y Clemens¹¹. IEEE.)

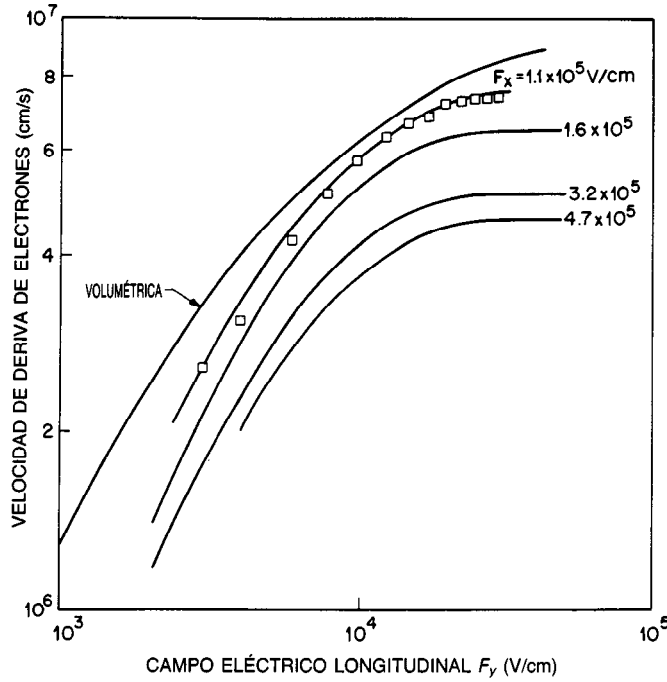


Fig. 10-12. Velocidad de deriva de electrones en la capa de inversión en función de F_y en la dirección $\langle 100 \rangle$ del silicio tipo p a 300 °K, con F_x como parámetro. (Tomado de Sze¹.)

La “resistividad” ρ (en $\Omega\text{-cm}$) de un semiconductor tipo n es²:

$$\rho = \frac{1}{q\mu_n n} \quad (10-16)$$

De forma similar, para un semiconductor tipo p es²:

$$\rho = \frac{1}{\mu_p p} \quad (10-17)$$

donde μ_n es la movilidad de los electrones y μ_p es la movilidad de los huecos.

En general, cuando es necesario tomar en consideración tanto los huecos como los electrones²:

$$\rho = \frac{1}{q(\mu_n n + \mu_p p)} \quad (10-18)$$

Cuando está presente sólo un tipo de impureza, ρ depende sencillamente de la concentración de la impureza ionizada. En la figura 10-13 se muestra la resistividad medida del silicio como función de la concentración de boro ionizado (tipo p) y de la concentración de fósforo ionizado (tipo n).

La concentración de boro $N_B < 10^{20} \text{ cm}^{-3}$ en el silicio como una función de ρ a 300 °K, está dada empíricamente por⁹:

$$N_B = (q\rho)^{-1} [2.13 \times 10^{-3} + 1.73 \times 10^{-2} (1 + 83.02 \rho^{1.105})^{-1}] \quad (10-19)$$

y la concentración de fósforo $N_P < 10^{20} \text{ cm}^{-3}$ en el silicio como una función de ρ a 300 °K, está dada empíricamente por⁹:

$$N_P = (q\rho)^{-1} 10^{u_1} \quad (10-20)$$

donde:

$$u_1 = \frac{-3.1083 - 3.2626x - 1.2196x^2 - 0.13923x^3}{1 + 1.0265x + 0.38755x^2 + 0.41833x^3}$$

y $x = \log(\rho)$.

En la figura 10-14 se muestra la dependencia de la resistividad con respecto a la concentración de la impureza ionizada.

10.2.8 Ionización por impacto

Cuando un campo eléctrico aplicado se incrementa por encima de cierto valor, el impacto de los portadores libres con los átomos de la malla excita los electrones de la banda de valencia y los hace pasar a la banda de conducción, creando pares hueco-electrón. La creación de tales pares se denomina ionización por impacto. La velocidad volumétrica de la formación de pares G mediante ionización por impacto es¹:

$$G = \alpha_n n \bar{v}_n + \alpha_p p \bar{v}_p \quad (\text{cm}^{-3}/\text{s}) \quad (10-21)$$

donde α_n es la velocidad de ionización de los electrones (cm^{-1}) definida como el número de pares electrón-

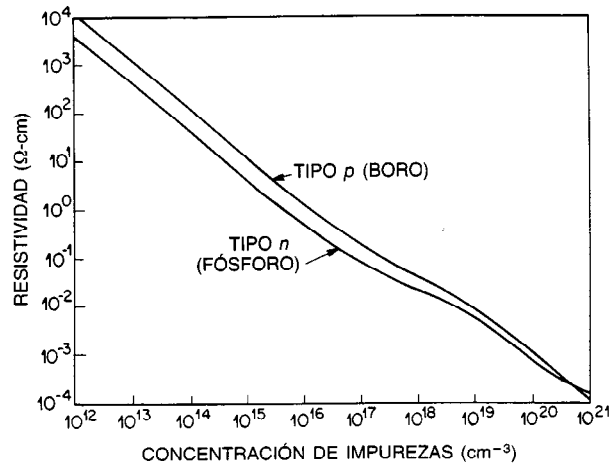


Fig. 10-13. Dependencia de la resistividad del silicio extrínseco a 300 °K con respecto a la concentración de impurezas. (Tomado con permiso de Beadle y colaboradores⁹. American Telephone and Telegraph Company).

hueco generados por un electrón por distancia unitaria recorrida, α_p es la velocidad de ionización definida de forma parecida para los huecos, y \bar{v}_n y \bar{v}_p son las velocidades de deriva de electrones y huecos, respectivamente.

En la figura 10-15 se muestra la dependencia medida con respecto a $1/F$ de α_n y α_p .

Las velocidades de ionización son de utilidad para predecir la multiplicación de los portadores libres y el voltaje de disrupción en avalancha en uniones *pn*. En la referencia 1 se presentan más detalles sobre la ionización por impacto.

10.2.9 Propiedades térmicas

Para un gradiente de temperatura dT/dx a lo largo del eje (dirección x) de una varilla semiconductora, el flujo calorífico resultante Q_x que fluye a través de una unidad de área por unidad de tiempo determina la conductividad térmica del semiconductor como:

$$K = -Q_x / (dT / dx) \quad (10-22)$$

donde las unidades de Q_x son watts por cm² y las de K son W/cm² - °K.

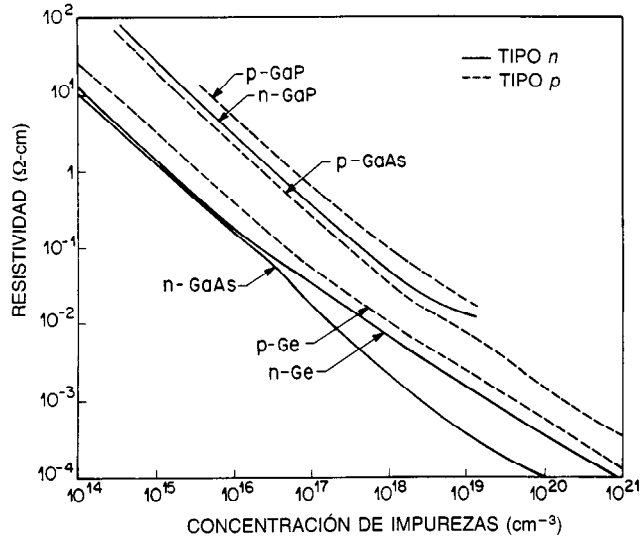


Fig. 10-14. Dependencia de la resistividad del germanio, arseniuro de galio y galio-fósforo extrínsecos a 300 °K respecto a la concentración de impurezas. (Reproducido con permiso de Beadle y colaboradores⁹. American Telephone and Telegraph Company.)

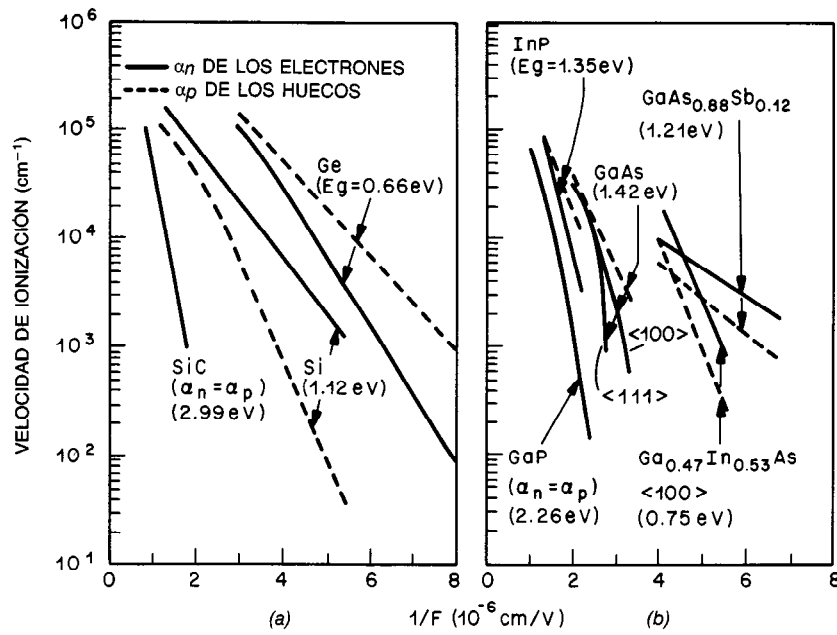


Fig. 10-15. Velocidad de ionización contra campo eléctrico recíproco medido a 300 °K para el germanio, silicio, arseniuro de galio y otros semiconductores compuestos. (Tomado de Sze¹.)

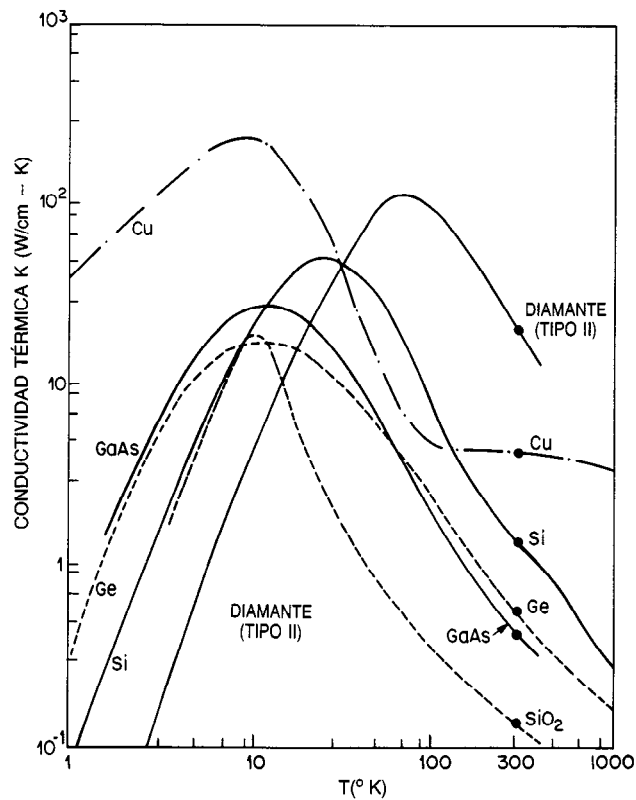


Fig. 10-16. Conductividad térmica medida en función de la temperatura para el germanio puro, silicio, arseniuro de galio, cobre, diamante tipo II y dióxido de silicio. (Tomado de Sze¹.)

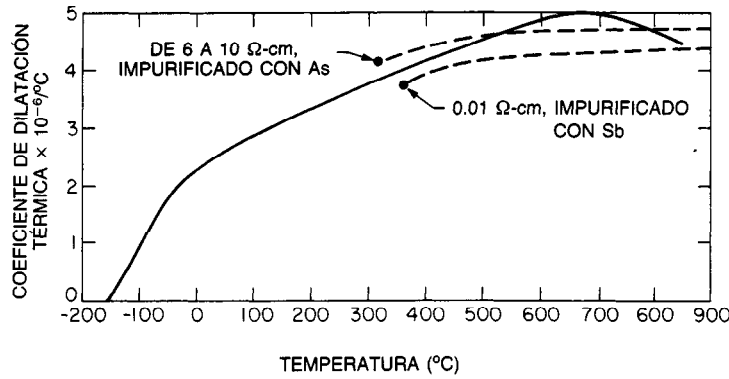


Fig. 10-17. Dependencia respecto a la temperatura del coeficiente de dilatación lineal del silicio. También se incluyen datos para el silicio impurificado con arsénico y para el silicio impurificado con antimonio a temperaturas superiores. (Reproducido con permiso de Beadle y colaboradores⁹. American Telephone and Telegraph Company.)

En la figura 10-16 se muestra la dependencia de la conductividad térmica medida con respecto a la temperatura de la malla. Suele emplearse cobre para disipar el calor de los dispositivos con uniones *pn*; el diamante (tipo II) posee la más alta conductividad térmica a temperatura ambiente conocida hasta hoy, y se utiliza como disipador de calor en láser de unión y en dispositivos de microondas; y el dióxido de silicio es el aislante de mayor uso en los circuitos integrados.

En la figura 10-17 se muestra el coeficiente de dilatación lineal del silicio en función de la temperatura. El cambio de longitud ΔL de una muestra de longitud L_0 a una temperatura T_0 , provocado por un aumento en la temperatura de T_0 a T , está dado por¹²:

$$\Delta L = \alpha L_0 (T - T_0) \quad (10-23)$$

donde α es el coeficiente de dilatación lineal en la figura 10-17.

10.2.10 Propiedades ópticas

Dependiendo de su energía, los fotones pueden interactuar en varias formas con un semiconductor. En todos los casos, el semiconductor absorbe energía de los fotones. Primero, los fotones de la separación energética o con energías superiores excitarán electrones de la banda de valencia y los harán pasar a la banda de conducción por impacto con átomos de la malla. Estas transiciones inducidas por fotones se utilizan para determinar la estructura de la banda de energía⁷ (fig. 10-1). Segundo, los fotones con energías inferiores a la de la separación energética proporcionarán algo de su energía a los portadores libres. Tercero, los fotones pueden excitar electrones de imperfecciones de la malla o niveles energéticos de captura en la separación energética y hacerlos pasar a la banda de conducción, o de la banda de valencia al nivel energético de captura.

En la interacción de los fotones con un semiconductor, se miden el coeficiente de transmisión Tr y el coeficiente de reflexión R . Para incidencia normal, estas dos cantidades están dadas por¹:

$$Tr = \frac{(1 - R^2)\exp(-4\pi x/\lambda)}{(1 - R^2)\exp(-8\pi x/\lambda)} \quad (10-24a)$$

y¹:

$$R = \frac{(1 - n^*)^2 + k^2}{(1 + n^*)^2 + k^2} \quad (10-24b)$$

donde λ es la longitud de onda, n^* es el índice de refracción, k es la constante de absorción y x es el grosor de la muestra. El coeficiente de absorción por unidad de longitud A está dado por¹:

$$A = \frac{4\pi k}{\lambda} \quad (10-25)$$

Tanto n^* como k pueden determinarse a partir de un análisis de los datos de $Tr - \lambda$ o $R - \lambda$ a incidencia normal, u observando R o Tr a diferentes ángulos de incidencia. Entonces es posible relacionarlos con las energías de transición interbanda.

En la figura 10-18 se muestra la dependencia medida de k cerca y por encima del borde fundamental de absorción (transiciones banda-banda) en la energía del fotón.

10.3 OXIDACIÓN DEL SILICIO

El desarrollo de una película delgada de dióxido de silicio (de 0.01 a 1 μm) sobre el silicio es una característica crucial de la tecnología de los circuitos integrados. La estabilidad química y la inactividad eléctrica del dióxido de silicio, así como su preparación relativamente fácil, hacen del silicio el material de mayor uso en la electrónica de los semiconductores. Las capas de dióxido de silicio proporcionan "pasivación" superficial para los transistores dipolares y los diodos de unión, en los que la capa de óxido asegura que las características eléctricas de estos dispositivos estén dominadas por propiedades volumétricas. Una capa de dióxido de silicio también puede servir.

1. Como "protección para la difusión".
2. Para aislar un dispositivo de otro.
3. Para aislar la compuerta respecto del silicio en dispositivos de efecto de campo, que es una de las aplicaciones de mayor importancia.
4. Para aislar múltiples niveles de interconexiones en un circuito integrado.

Los óxidos que satisfacen estos requerimientos pueden elaborarse con diversos métodos, de los cuales los más importantes son la reacción en fase de vapor¹³, la anodización con plasma^{14, 15}, la anodización húmeda¹⁶ y la oxidación térmica tanto en ambientes húmedos como en secos¹⁷. La oxidación térmica produce óxidos que presentan las mejores propiedades eléctricas. Además, los óxidos desarrollados térmicamente cumplen todos los demás requerimientos recién descritos. Por consiguiente, en la actualidad la oxidación térmica es el proceso fundamental de oxidación en la tecnología de los circuitos integrados. Los otros métodos de formación de óxidos amplían las opciones disponibles para el fabricante de dispositivos. En las referencias bibliográficas 2 y 3 se describen detalles de la cinética de la oxidación térmica y de la tecnología del silicio.

10.3.1 Defectos interfaciales y carga de óxido

Defectos interfaciales. Los defectos eléctricamente activos son característicos de la interficie entre un se-

miconductor y un óxido o algún otro aislante. Estos defectos se localizan en la interficie o cerca de ella y sus densidades varían desde $\leq 10^{10} \text{ cm}^{-2}$ hasta $\geq 10^{12} \text{ cm}^{-2}$, dependiendo de las condiciones de formación del óxido o del aislante. Las trampas interfaciales son aquellos defectos que tienen niveles de energía en la separación energética del semiconductor, los cuales modifican las características de ocupación mediante la captura de portadores libres del semiconductor o mediante la emisión de portadores al semiconductor cuando se varía la polarización de la compuerta. La carga fija de óxido son aquellos niveles defectuosos que no modifican la ocupación con variaciones en la polarización de la compuerta. En el sistema Si-SiO₂, la carga fija de óxido suele ser positiva. Ambos tipos de defecto son indeseables a densidades moderadamente altas ($> 10^{11} \text{ cm}^{-2}$), debido a que afectan de manera adversa las características del dispositivo. Por ejemplo, altas trampas interfaciales y densidades de carga fija de óxido degradan la ganancia e incrementan el ruido en los transistores dipolares y en los de efecto de campo a base de semiconductor de metal y óxido (MOSFET), alteran el voltaje umbral en los MOSFET, modifican el voltaje de disrupción en avalancha y degradan algunas características de las uniones *pn*, y reducen la eficiencia de los dispositivos fotónicos. Para optimizar el rendimiento y la estabilidad de los dispositivos, se debe controlar las trampas interfaciales y las densidades de carga fija de óxido durante la fabricación, y mantenerlas sin cambio durante el funcionamiento del dispositivo.

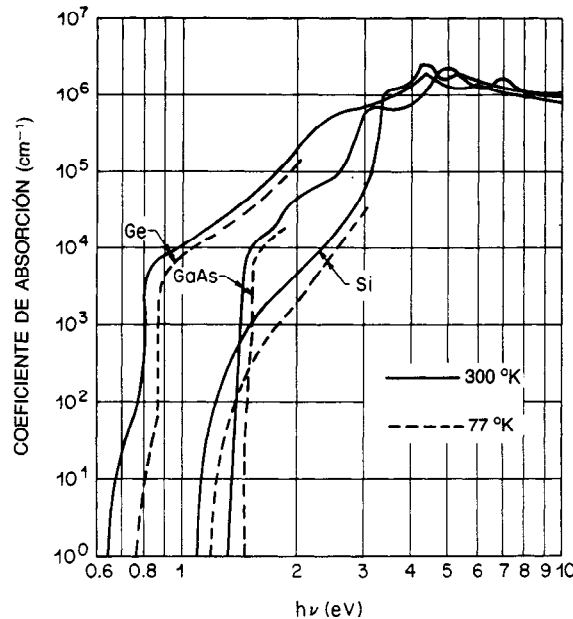


Fig. 10-18. Coeficientes de absorción medidos cerca y sobre el borde fundamental de absorción para germanio puro, arseniuro de galio y silicio a 300 y 77 °K. El desplazamiento de las curvas hacia energías fotónicas superiores a temperaturas más bajas es provocado por la dependencia respecto a la temperatura de la energía de la separación energética [véanse la fig. 10-2 y la ec. 10-1]. (Tomado de Sze¹.)

En la referencia bibliográfica 3 se describe para la interficie del Si-SiO₂ la manera en que estos defectos afectan las características de operación del dispositivo y cómo se miden y controlan sus propiedades durante su fabricación y operación.

Carga de óxido. Existen dos tipos de carga en el óxido: la "carga móvil" y la "carga atrapada de óxido". La carga móvil más común en el SiO₂ es el ion positivo de sodio. El sodio es el metal alcalino contaminante de mayor importancia, debido a su existencia tan difundida en el ambiente. Concentraciones de iones sodio de una parte por millón o menos en el SiO₂ pueden provocar cambios inaceptablemente grandes de las características eléctricas de los dispositivos. Tales iones afectan las características de los dispositivos a través de la gran carga imagen que inducen en el silicio después de la deriva a la interficie Si-SiO₂ bajo un campo eléctrico aplicado*. En la referencia bibliográfica 3 se explica la relación entre el sodio y el SiO₂.

La carga atrapada de óxido se crea por la exposición a radiación ionizante o por inyección de portadores de alta energía. La radiación ionizante produce pares hueco-electrón en el SiO₂. El efecto más importante es la acumulación de carga positiva con el tiempo de exposición. La carga positiva de óxido ocurre cuando los huecos producidos derivan bajo un campo aplicado a la interficie Si-SiO₂, donde quedan atrapados. La carga imagen inducida por los huecos atrapados[‡] en el silicio modifica las características del dispositivo.

Algunos de los electrones que derivan de la fuente al consumo o descarga en un MOSFET ganan suficiente energía de los campos aplicados para vencer la barrera de energía de 3.2 eV entre el silicio y el SiO₂. Una vez en el SiO₂, una pequeña fracción de los electrones queda atrapada por centros relacionados con agua que producen carga negativa de óxido y trampas interfaciales. La inyección de electrones de alta energía se evita operando los MOSFET a voltajes menores de 3.2 V. En la referencia bibliográfica 3 se analizan con más detalle la radiación ionizante, la inyección de portadores de alta energía y la captura.

10.4 EPITAXIA, DIFUSIÓN E IMPLANTACIÓN DE IONES

Los principales procesos de la tecnología de los dispositivos semiconductores son el crecimiento epitaxial de delgadas películas semiconductoras, la implantación de iones, la difusión de impurezas agregadas dentro de los semiconductores y la oxidación.

10.4.1 Crecimiento epitaxial

La **epitaxia** es el proceso por el cual se forma una capa delgada de un semiconductor monocristalino sobre un

sustrato semiconductor monocristalino. El crecimiento epitaxial es importante en la tecnología de los semiconductores por dos razones. En primer lugar, es posible desarrollar películas delgadas con concentración relativamente baja de impurezas sobre sustratos que contienen el mismo tipo de impureza. La concentración de impurezas en el sustrato suele ser mucho mayor que en la película epitaxial a fin de reducir la resistencia en serie asociada con el sustrato, sin alterar de ninguna otra forma las características del dispositivo. En segundo lugar, la concentración de impurezas en la película epitaxial puede controlarse con facilidad e independientemente de la concentración de impurezas en el sustrato, mediante el control de la concentración de impurezas en la fase a partir de la que crece la capa epitaxial. De este modo, es posible utilizar el crecimiento epitaxial para formar una unión *pn* entre la capa epitaxial y el sustrato.

Los grosores de la capa epitaxial suelen variar del intervalo de los micrómetros hasta el de las dimensiones atómicas. Tres métodos de uso común para la formación de capas epitaxiales son:

1. El crecimiento en fase de vapor², utilizado en la tecnología del silicio.
2. El crecimiento en fase líquida¹⁸, utilizado ampliamente en la tecnología de los dispositivos semiconductores compuestos.
3. La epitaxia por haz molecular^{19, 20}, que puede proporcionar un control preciso de composiciones del semiconductor y grosores de capas hasta dimensiones atómicas.

En la referencia bibliográfica se hace un análisis del crecimiento de películas de silicio en fase de vapor.

10.4.2 Difusión

Los dispositivos semiconductores y los circuitos integrados se elaboran mediante la introducción de impurezas de tipo y concentración dados en regiones específicas de un semiconductor. La difusión en estado sólido es una forma práctica de lograr este propósito. Si una impureza cuyo tipo de conductividad es opuesto al tipo de conductividad inicial del semiconductor se difunde dentro de éste, se forma una unión *pn*. Un empleo primordial de la difusión en la moderna tecnología de los circuitos integrados es la impurificación de compuertas de polisilicio.

La difusión suele ser un proceso de dos etapas. La primera, conocida como etapa de predepósito, es aquella en la que se deposita la impureza, usualmente contenida en un gas, sobre la superficie del semiconductor hasta una profundidad que varía de décimas de un micrón a micrones. En la segunda etapa, llamada de transmisión, las impurezas depositadas previamente se difunden a mayor profundidad para formar la distribución de concentración deseada. Para el silicio, la transmisión se efectúa a temperaturas elevadas en un ambiente oxidante a fin de crear una capa de óxido que selle efectivamente las impurezas depositadas con anterioridad.

* Los iones sodio presentan un serio problema de estabilidad porque no pierden su carga en un electrodo.

‡ En comparación con los huecos, pocos electrones quedan atrapados.

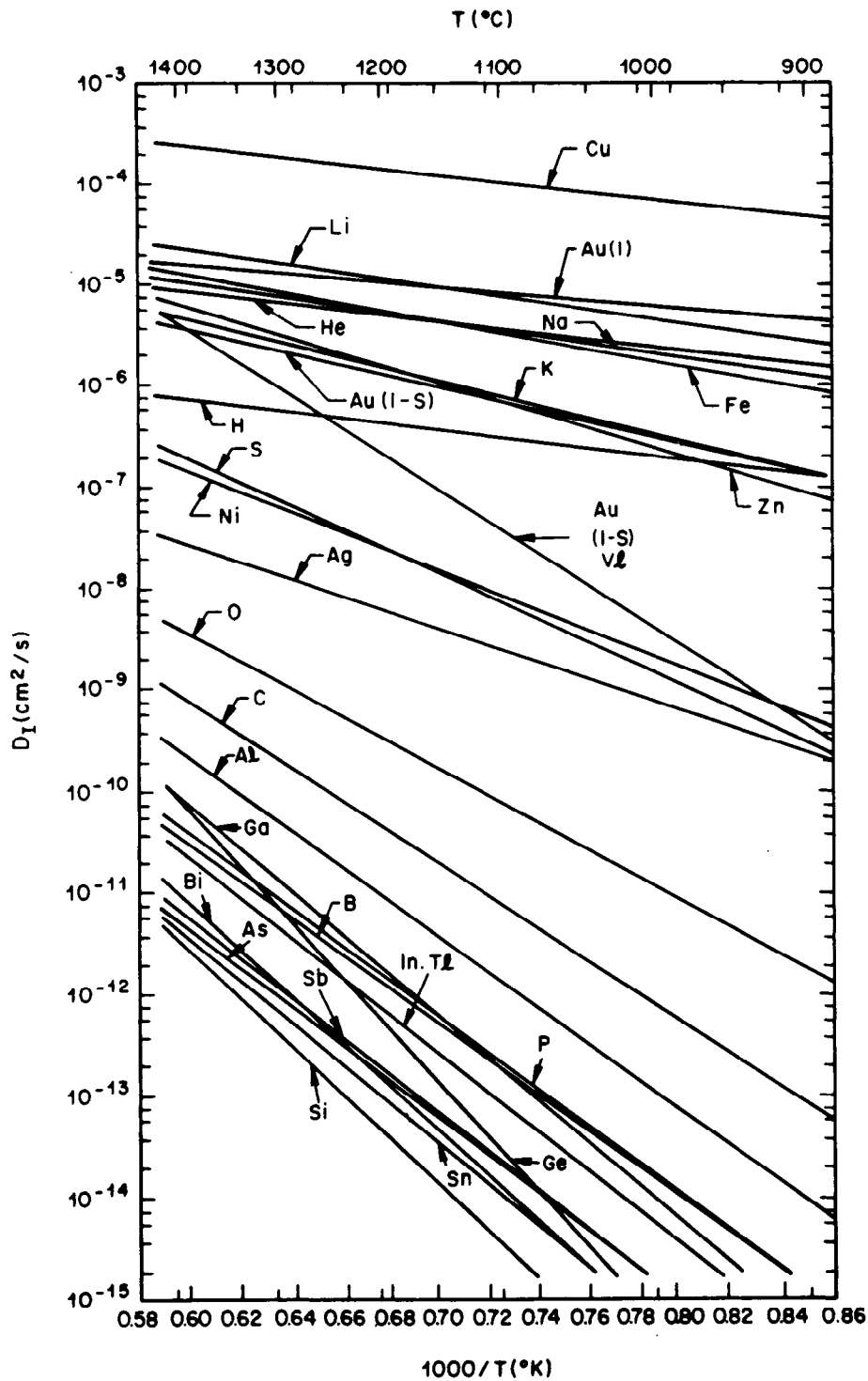


Fig. 10-19. Gráfica de $D_I(T)$ para una baja concentración de impurezas agregadas en el silicio. (Reproducido con permiso de Beadle y colaboradores⁹. American Telephone and Telegraph Company.)

La ecuación unidimensional de difusión que relaciona la velocidad de cambio de la concentración de impurezas con el gradiente de concentración de impurezas es¹:

$$\frac{\partial C(x,t)}{\partial t} = D_I \frac{\partial^2 C(x,t)}{\partial x^2} \quad (10-26)$$

donde $C(x, t)$ es la concentración de impurezas en cm^{-3} y D_I es el coeficiente de difusión de las impurezas. Existen dos condiciones prácticas bajo las cuales es posible resolver la ecuación 10-26: 1) la concentración de impurezas en la superficie del semiconductor, C_s (cm^{-3}), siempre es constante; la solución es:

$$C(x,t) = C_s \operatorname{erfc} \left(\frac{x}{2\sqrt{D_I t}} \right) \quad (10-27)$$

y 2) la concentración inicial de impurezas en la superficie del semiconductor, C_T (cm^{-2}), constituye la cantidad total de impurezas disponibles; la solución es:

$$C(x,t) = \frac{C_T}{\sqrt{\pi D_I t}} \exp \left(-\frac{x^2}{4D_I t} \right) \quad (10-28)$$

Las ecuaciones 10-27 y 10-28 proporcionan buenas aproximaciones a los perfiles de muchas impurezas.

El coeficiente de difusión depende de la temperatura y de la concentración de impurezas. Para bajas concentraciones de impurezas, D_I se vuelve independiente de la concentración. En un intervalo limitado para bajas concentraciones de impurezas¹:

$$D_I(T) = D_\infty \exp(-E_A/kT) \quad (10-29)$$

donde D_∞ es el coeficiente de difusión extrapolado a una temperatura infinita y E_A es la energía de activación en joules. En la figura 10-19 se muestra $D_I(T)$ para el Si, y en la figura 10-20, $D_I(T)$ para el Ge y el GaAs, con diversas impurezas a bajas concentraciones.

La "solubilidad en sólidos" de una impureza es la máxima concentración de dicha impureza que es posible incorporar en un semiconductor a una temperatura dada. En la figura 10-21 se muestra la solubilidad en sólidos de algunas impurezas importantes en el silicio en función de la temperatura. Algunos de los impurificadores más usuales del silicio son el boro, fósforo y arsénico, debido a que presentan solubilidades en sólidos muy elevadas, tal como se observa en la figura

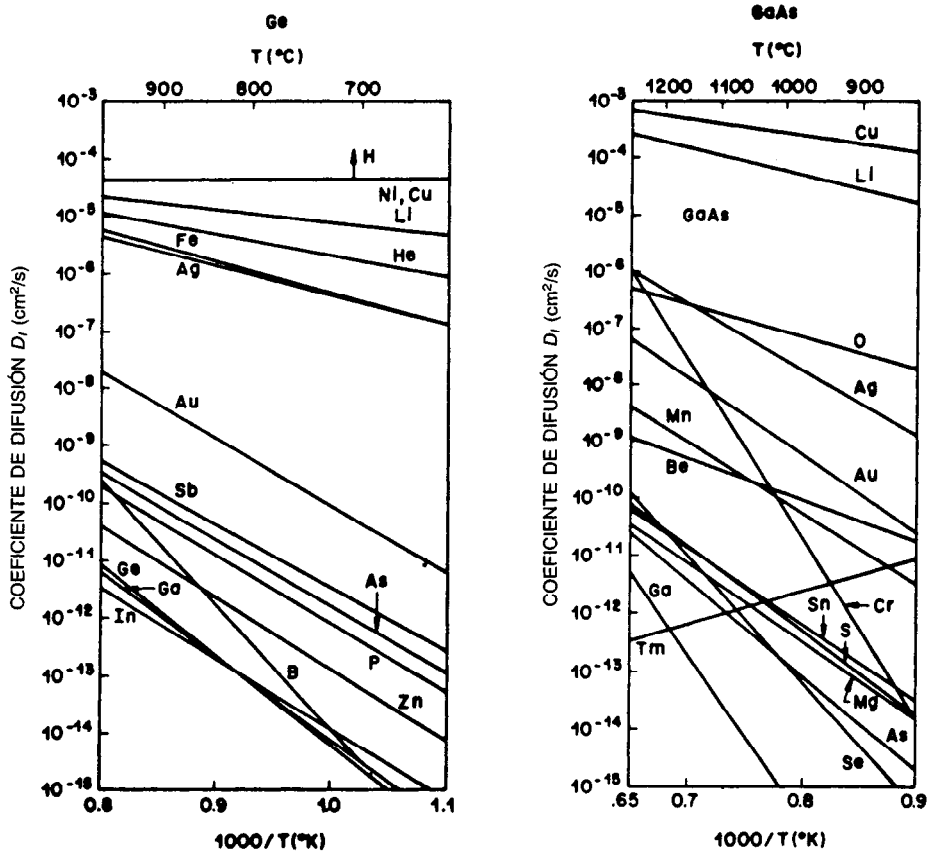


Fig. 10-20. Gráfica de $D_I(T)$ para una baja concentración de impurezas agregadas en el germanio y en el arseniuro de galio. (Tomado de Sze¹.)

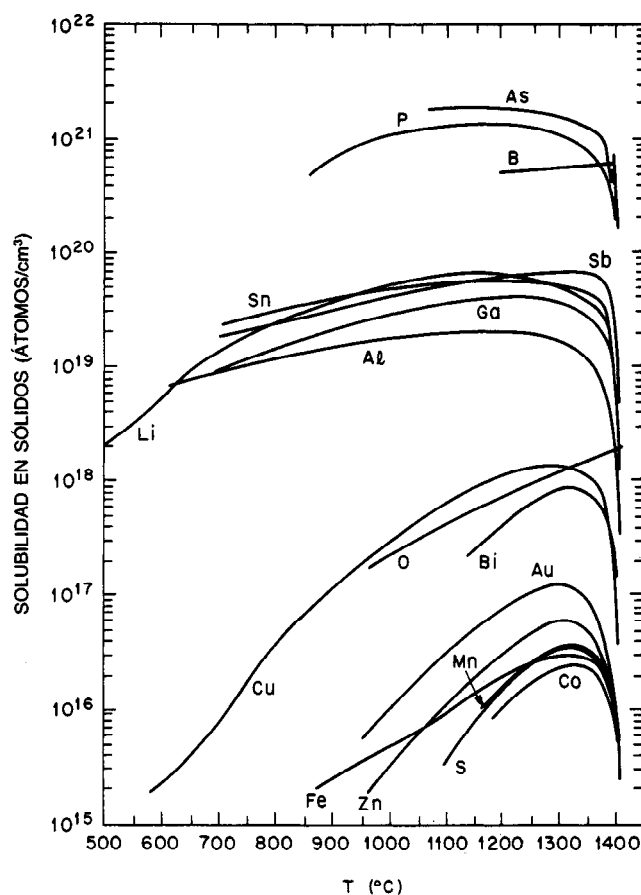


Fig. 10-21. Dependencia respecto a la temperatura de la solubilidad en sólidos de diferentes impurezas agregadas en el silicio. (Tomado de Sze¹.)

10-21, y tienen niveles de energía muy someros, tal como se observa en la figura 10-6. El arsénico es de uso más frecuente que el fósforo como donador en circuitos integrados de muy pequeñas dimensiones, debido a su D_I más bajo*, tal como se observa en la figura 10-19. En las referencias bibliográficas 1 y 2 se hace un análisis más detallado.

10.4.3 Protección de óxido

El SiO_2 es una barrera muy efectiva contra la difusión de muchas impurezas²¹. Al grabar en el óxido con ácido fluorhídrico o con un plasma iónico reactivo agujeros definidos por métodos litográficos, con el óxido restante actuando como protección, es posible incorporar las impurezas controladas dentro de un sustrato de silicio en zonas selectas para formar las regiones activas de unión *pn* en un circuito integrado. Para ser efectiva como protección contra la difusión, la

difusividad de un impurificador en SiO_2 debe ser menor que en el silicio. Los impurificadores de uso común que presentan difusividades más pequeñas en el SiO_2 que en el silicio son el fósforo, boro, arsénico y antimonio.

El grosor mínimo de SiO_2 , x_m , que se requiere para proteger contra la difusión del boro se halla mediante la expresión²²:

$$\frac{x_m^2}{t_m} = 4.9 \times 10^5 \exp(-2.80q/kT_m) \text{ (cm}^2/\text{s)} \quad (10-30)$$

y contra la difusión del fósforo es²³:

$$\frac{x_m^2}{t_m} = 1.7 \times 10^7 \exp(-1.46q/kT_m) \text{ (cm}^2/\text{s)} \quad (10-31)$$

donde t_m es el tiempo en segundos y T_m es la temperatura de difusión en kelvins.

En la figura 10-22 se muestra el grosor x_m para el cual el SiO_2 desarrollado térmicamente en oxígeno seco fracasa como protección contra vapores de P_2O_5 y B_2O_3 en función del tiempo medido, con la temperatu-

* Para mantener reducidas las dimensiones del dispositivo, las impurezas no deben difundirse mucho durante las etapas del procesamiento térmico.

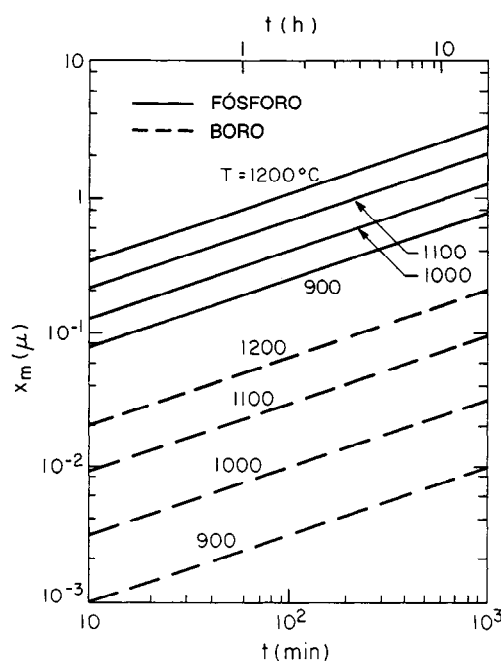


Fig. 10-22. Dependencia respecto al tiempo del grosor mínimo de óxido requerido como protección contra el fósforo y boro, con la temperatura de difusión como parámetro. (Reproducido con permiso de Wolfe²⁴. Pergamon Pres, Ltd.)

ra de difusión como parámetro; ese grosor es descrito por las ecuaciones 10-30 y 10-31. Las difusividades del boro y del fósforo dependen de la concentración. Sin embargo, la figura 10-22 es conservadora porque se basa en las más altas difusividades efectivas del boro y del fósforo que se informan en las publicaciones especializadas.

10.4.4 Implantación de iones

Éste es un método para impurificar un semiconductor bombardeándolo con iones energéticos de la impureza. Esta etapa suele seguirse de tratamiento térmico a temperaturas elevadas, típicamente del orden de los 900 °C, para recocer el daño en la malla producido por los iones energéticos*, activar eléctricamente los átomos contaminantes implantados y difundirlos a mayor profundidad dentro del sustrato[‡].

Las ventajas más importantes sobre el depósito químico son²⁵:

1. Control preciso sobre dosis total, profundidad del perfil y uniformidad del área.

* En necesario recocer el daño en la malla para tener la certeza de que las concentraciones de los portadores libres están determinadas por la concentración de las impurezas, en vez de estarlo por el daño de la red.

‡ En algunas aplicaciones, la implantación de iones es una etapa previa de depósito, seguida de una etapa de oxidación-difusión a alta temperatura en la que se recuece el daño en la malla.

2. Las uniones implantadas pueden autoalinearse con el borde de la protección.

Las energías típicas de los iones varían de 10 a 400 keV, y las dosis, de 10^{11} a 10^{16} iones/cm². En la figura 10-23 se muestra el intervalo proyectado de los iones de boro, fósforo y arsénico en Si y en SiO₂ como función de la energía.

Cuando el haz iónico barre uniformemente la superficie del semiconductor, lo que es una práctica común, la distribución de los átomos de la impureza dentro del semiconductor se aproxima a una curva gaussiana. El perfil impurificado, una vez que se ha completado la fabricación del dispositivo, suele ser distinto del perfil que se observa inmediatamente después de la implantación, debido a los tratamientos térmicos que se hacen con posterioridad.

En la referencia bibliográfica 3 se describen métodos para medir el perfil de impurificación. Pueden emplearse polisilicio, SiO₂ y capas fotorresistentes como protección contra iones impurificantes en la implantación de iones. En la figura 10-23 se muestra la penetración de iones impurificantes de uso común en el SiO₂. Una capa de SiO₂ mucho más gruesa que el intervalo proyectado para una energía y un ion impurificante dados serán una protección efectiva contra el ion impurificante. Por ejemplo, una película de SiO₂ con grosor de 0.1 μm protegerá efectivamente de iones de fósforo a 30 keV. En la referencia bibliográfica 26 pueden encontrarse tablas de estadísticas del intervalo proyectado para una variedad de impurificantes de uso común en semiconductores. En las referencias bibliográficas 25 y 27 se presentan detalles sobre la implantación de iones.

REFERENCIAS BIBLIOGRÁFICAS

1. S. M. Sze, *Physics of Semiconductor Devices*, 2.^a ed., Wiley-Interscience, Nueva York.
2. A. S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, Nueva York.
3. E. H. Nicollian y J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology*, Wiley-Interscience, Nueva York.
4. J. R. Chelikowsky y M. L. Cohen, *Phys. Rev. B* **14**: 556.
5. C. D. Thurmond, *J. Electrochem. Soc.* **122**: 1133.
6. W. Paul y D. M. Warschau, eds., *Solids Under Pressure*, McGraw-Hill, Nueva York.
7. J. C. Phillips, *Bands and Bonds in Semiconductors*, Academic, Nueva York; y F. Herman, *Proc. IRE* **43**: 1703.
8. Max Born, *Atomic Physics*, 5.^a ed., Hafner, Nueva York, Ch. 8, p. 261.
9. W. F. Beadle, R. D. Plummer, y J. C. C. Tsai, eds., *Quick Reference Manual for Semiconductor Engineers*, Bell Laboratories, Murray Hill, NJ.
10. C. Canali, C. Jacoboni, F. Nava, G. Ottaviani y A. Alberigi-Quaranta, *Phys. Rev. B* **12**: 2265.
11. A. G. Sabnis y J. T. Clemens, *IEEE Tech. Dig., Int. Electron Dev. Meet.*, p. 18.

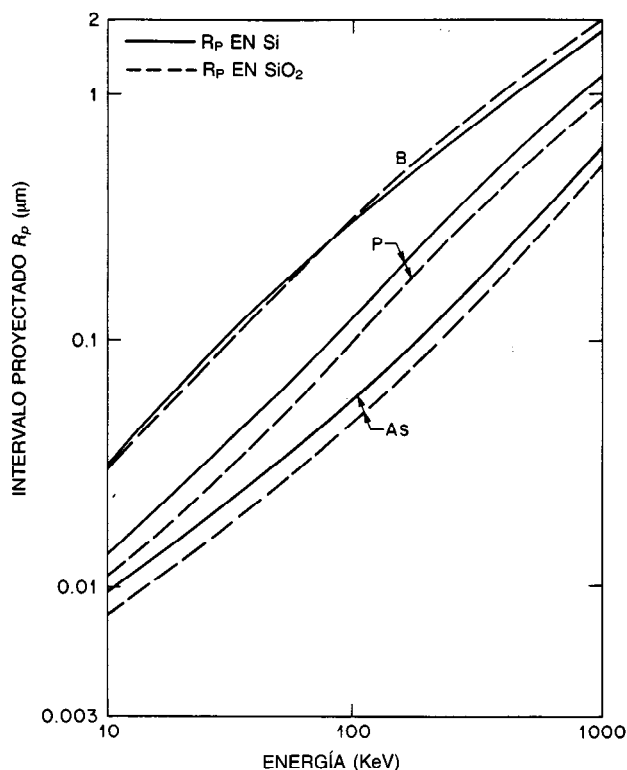


Fig. 10-23. Dependencia del intervalo proyectado de los iones de boro, fósforo y arsénico respecto a la energía del haz incidente de iones. (Tomado de Pickar²⁵.)

12. C. Kittel, *Introduction to Solid State Physics*, 2.^a ed., Wiley, Nueva York, Cap. 6.
13. E. L. Jordan, *J. Electrochem. Soc.* **108**: 478.
14. J. R. Ligenza, *J. Appl. Phys.* **36**: 2703.
15. J. R. Ligenza y M. Kuhn, *Solid-State Technol.* **13**: 33.
16. P. F. Schmidt y W. Michel, *J. Electrochem. Soc.* **104**: 230.
17. M. M. Atalla, en H. Gatos, ed., *Properties of Elemental and Compound Semiconductors*, Vol. 5, Wiley-Interscience, Nueva York, pp. 163-181.
18. H. C. Casey y M. B. Panish, *Heterostructure Lasers*, Academic, Nueva York.
19. A. Y. Cho, *J. Vac. Sci. Technol.* **16**: 275.
20. J. C. Bean, "Growth of Doped Silicon Layers by Molecular Beam Epitaxy", en F. F. Y. Wang, ed., *Impurity Doping Processes in Silicon*, North-Holland, Amsterdam.
21. C. J. Frosch y L. Derick, *J. Electrochem. Soc.* **104**: 547.
22. S. Horiuchi y J. Yamaguchi, *Jap. J. Appl. Phys.* **1**: 314.
23. C. T. Sah, H. Sello y D. A. Tremere, *J. Phys. Chem. Solids* **11**: 288.
24. H. F. Wolfe, *Silicon Semiconductor Data*, Pergamon, Nueva York, p. 601.
25. K. A. Pickar, "Ion Implantation in Silicon-Physics, Processing and Microelectronic Devices", en R. Wolfe, ed., *Applied Solid State Science*, Vol. 5, Academic, Nueva York.
26. J. F. Gibbons, W. S. Johnson y S. W. Mylroie, *Projected Range Statistics*, 2.^a ed., Dowden, Hutchinson, y Ross, Stroudsburg, PA.
27. J. W. Mayer y O. J. Marsh, "Ion Implantation in Semiconductors", en R. Wolfe, ed., *Applied Solid State Science*, Vol 1, Academic, Nueva York.

CAPÍTULO 11

Materiales emisivos, ópticos y fotosensibles

Bernard Smith

Eradcom Delet-BS, US Army
Fort Monmouth, New Jersey

Mansoor A. Saifi

Belcom Research
Murray Hill, New Jersey

11.1 Materiales emisivos

- 11.1.1 Introducción
- 11.1.2 Cátodos de emisión termoiónica
- 11.1.3 Cátodos emisores por campo eléctrico
- 11.1.4 Emisores que funcionan a temperatura ambiente
- 11.1.5 Nuevos cátodos en proceso de desarrollo

11.2 Materiales ópticos y fotosensibles

- 11.2.1 Materiales ópticos e indicatriz
- 11.2.2 Vidrios ópticos
- 11.2.3 Recubrimientos ópticos
- 11.2.4 Materiales electroópticos
- 11.2.5 Materiales elastoópticos
- 11.2.6 Materiales magnetoópticos
- 11.2.7 Materiales ópticos no lineales
- 11.2.8 Materiales fotosensibles

11.1 MATERIALES EMISIVOS

Bernard Smith

11.1.1 Introducción

La emisión de electrones por los sólidos se ha estudiado con intensidad variable desde el siglo pasado. El corazón de cualquier tubo electrónico, ya sea receptor, de microondas o de ondas milimétricas, es la fuente de electrones, que suele denominarse cátodo. Hasta la fecha, los cátodos existentes para empleo en dispositivos electrónicos se clasifican en dos grandes categorías: emisores primarios o secundarios. El funcionamiento idóneo de cualquiera de estos dispositivos depende directamente del funcionamiento adecuado del cátodo. Por tanto, es fácil observar la gran importancia que tiene la elección apropiada del cátodo, para un rendimiento satisfactorio de cualquier tubo electrónico.

Los grandes avances logrados en los emisores electrónicos en los últimos cincuenta años están estrechamente ligados al grado de perfección alcanzado en las técnicas al alto vacío y de fabricación única. El advenimiento de la tecnología de estado sólido ha conducido al desarrollo de equipo moderno de investigación al vacío, como espectroscopios de perforación electrónica, microscopios de barrido electrónico y espectroscopios de masa iónica secundaria. Este equipo se ha empleado para identificar la composición de elementos en la superficie del cátodo, y su uso ha permitido avances significativos para comprender la emisión de electrones de materiales compuestos, como los cátodos de óxidos alcalinotérreos y los de metales puros. Las publicaciones especializadas en el campo de los cátodos son tan numerosas que aquí no se intentará cubrir cada aspecto de los materiales emisivos y su tecnología. Por consiguiente, en esta sección se revisan sólo los emisores más promisorios que existen en el mercado y se analizan brevemente los cátodos en proceso de desarrollo. Se mencionan referencias para quienes estén interesados en un análisis más profundo sobre los aspectos físicos y de funcionamiento de los diferentes emisores.

Para comprender las diferencias entre los diversos materiales emisivos es importante repasar brevemente la ecuación que describe la emisión termoiónica.

La emisión a partir de una superficie a la temperatura T depende del valor de la función trabajo ϕ y está dada por la conocida ecuación de Richardson¹:

$$J = AT^2 (1 - r)e^{-\phi/KT}$$

Dado que en la mayor parte de los casos $r \ll 1$, esta ecuación se reduce a:

$$J = AT^2 e^{-\phi/KT}$$

donde J = densidad de corriente, en A/cm²

r = coeficiente de reflexión

T = temperatura, en °K

ϕ = función trabajo, en eV

K = constante de Boltzmann

Dado que ϕ depende de la temperatura, es posible expresar:

$$\phi = \phi_0 + \alpha T$$

Sustituyendo en la ecuación para J se obtiene:

$$J = (Ae^{-\alpha/K})T^2 e^{-\phi_0/KT}$$

donde $Ae^{-\alpha/K}$ se define como la constante de Richardson A_0 , y J pasa a ser:

$$J = A_0 T^2 e^{-\phi_0/KT}$$

En la tabla 11-1 se enumeran algunas propiedades importantes de ciertos materiales emisivos.

11.1.2 Cátodos de emisión termoiónica

Cátodos recubiertos de óxido

Esta clase de emisores termoiónicos es la que más se ha investigado y empleado de todos los sistemas de cátodos. El cátodo a base de óxido consta de óxidos alcalinotérreos (bario, estroncio, calcio) que se usan para recubrir una base metálica^{2,3}. La base de mayor uso es el níquel o un metal base de alguna aleación de níquel. En estos cátodos la emisión proviene de una monocapa de bario que se forma en la superficie cuando se disocian los óxidos. En las publicaciones especializadas se han informado diversas variantes del cátodo ordinario recubierto de óxido, incluyendo el de níquel prensado³, el Medicus⁴ y el de polvo recubierto⁵. Estos cátodos se consideran variantes del cátodo compensado o, con mayor precisión, variaciones entre el cátodo de óxido y el cátodo compensado. En las referencias citadas se encuentran detalles sobre las técnicas de fabricación de estas variaciones del cátodo re-

Tabla 11-1. Propiedades de materiales emisivos

Material	Constante de Richardson, A_R	Función trabajo, ϕ_R (eV)	Densidad de corriente, J (A/cm ²)	Temperatura de operación (K)
Tungsteno (W)	70	4.50	0.25-0.7	2 500-2 600
Níquel recubierto de óxido	< 10	1.0-1.3	0.1-2	1 000-1 150
Tungsteno impregnado (tipo B)	2.4	1.65	1-5	1 325-1 450
Cátodo M	5.67	1.67	1-5	1 225-1 300
Tungsteno toriado	4	2.65	0.25-3	1 900-2 000

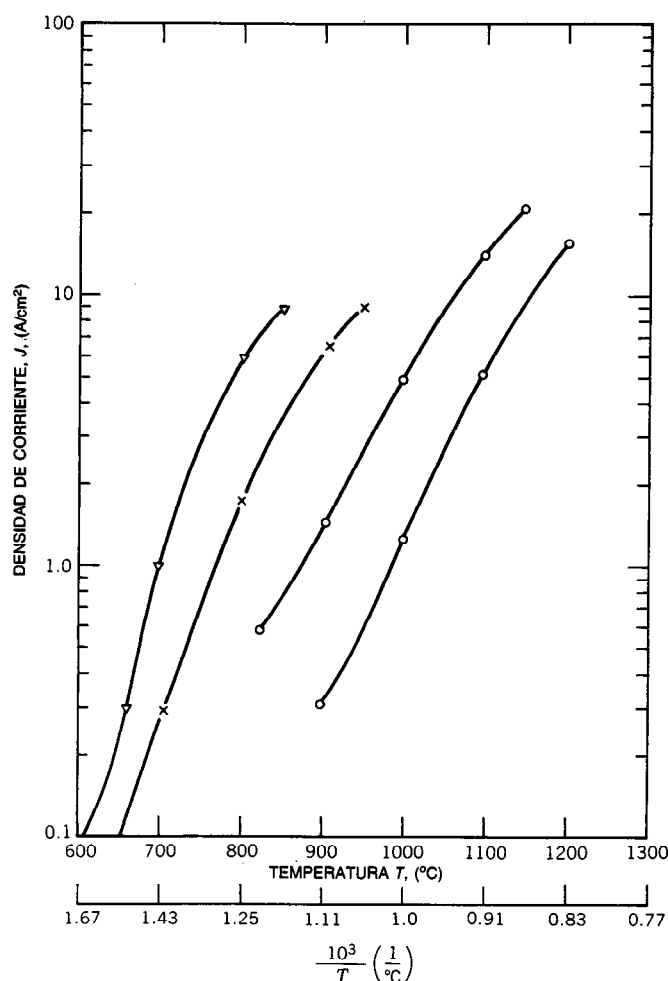


Fig. 11-1. Emisividad en función de la temperatura para cátodos comerciales.

cubierto de óxido. En general, todas las técnicas se ajustan a la teoría de la emisión que se encuentra en Wagner y Herrman², Wright⁶ y Nergaard⁷.

La mayoría de los investigadores están de acuerdo en que la emisión de los cátodos de óxido proviene de la monocapa de bario que se forma en la superficie cuando se descomponen los óxidos. A bajas temperaturas de operación, la reacción química que produce bario libre es lenta. Contaminantes residuales pueden desactivar el cátodo más rápidamente que la liberación de bario, una condición que con frecuencia se denomina envenenamiento o desactivación. A elevadas temperaturas, la duración se acorta debido a la gran evaporación del recubrimiento de óxido del cátodo. Así, para cualquier tubo en el que se utilicen emisores de óxido, el diseño pasa a ser una solución intermedia entre la duración del cátodo y el rendimiento del tubo. A densidades sostenidas de corrientes por encima de 1 A/cm², el recubrimiento de óxido se sobrecalienta y se evapora debido a su resistencia. Este fenómeno ocurre en todas las variaciones del cátodo.

En la figura 11-1 se presentan las características de emisividad de los cátodos recubiertos de óxido. En general, éstos son capaces de producir emisiones de corriente directa de 100 a 400 mA/cm² durante miles de horas de vida útil.

Para pulsaciones muy cortas (aplicaciones de bajo régimen de trabajo) es posible obtener densidades de corriente de 1 A/cm² sin afectar la razonablemente larga duración de los cátodos. Un cátodo bien activado permitirá la obtención de cargas de corriente hasta de 0.5 A/cm² durante miles de horas, y para dichas cargas se han informado valores de emisión máximos⁸ de 50 A/cm².

Emisores de matriz metálica con recubrimiento de metales activos

Los cátodos de matriz metálica o de monocapa de bario constan de una base de metal cubierta por una monocapa electropositiva absorbida. Para los cátodos prácticos, la monocapa electropositiva (que suele ser de bario) se coabsorbe con oxígeno a fin de formar una

capa dipolar que reduzca de manera efectiva la función trabajo⁹. El moderno tubo de microondas o de ondas milimétricas, que requiere densidades de CD superiores a 1 A/cm² durante miles de horas de duración en la mayor parte de los casos, se construye con el cátodo de monocapa de bario como fuente de electrones. La época moderna del cátodo de monocapa de bario se inició con la introducción del cátodo L, que consta de una clavija porosa de tungsteno conectada a un depósito hermético a los vapores que contienen carbonatos de bario y estroncio. Durante el procesamiento se descomponen estos carbonatos para formar bario libre, que se difunde a través de las píldoras de tungsteno poroso para recubrirlas de bario. La función trabajo de estos cátodos es aproximadamente 2.2 eV, valor típico del bario puro sobre tungsteno limpio. Las principales dificultades que se encuentran con estos cátodos son las producidas por una velocidad excesivamente alta de evaporación del bario.

El siguiente mayor avance en importancia de los cátodos compensados o de difusión ocurrió con la introducción de cátodos de aluminato cálcico de bario, que se fabrican impregnando el material activo dentro de una clavija de tungsteno poroso^{10,11}. Este cátodo (tipo B) es capaz de entregar de 1 a 2 A/cm² a temperaturas razonables de operación (de 1 000 a 1 050 °C) y velocidades de sublimación aceptables. Durante muchos años este cátodo fue el caballo de batalla de la industria de los tubos cuando se requerían densidades de corriente directa de 1 a 2 A/cm² para pulsos de régimen de trabajo elevado.

Dadas las demandas en aumento de mayores densidades de corriente, Zalm y Van Stratum¹² desarrollaron lo que suele conocerse como el cátodo M, que consta de un cátodo ordinario impregnado sobre cuya superficie se deposita electrónicamente una delgada capa (<10 000 Å) de osmio-rutenio. El efecto de este recubrimiento metálico delgado encima del cátodo ordinario impregnado es reducir de manera significativa la función trabajo de la superficie. Los primeros cátodos M que existieron en el mercado fueron fabricados por Phillips Metalonic, una compañía que ya no se dedica a la manufactura de cátodos en Estados Unidos. Poco antes de que dicha compañía cerrara sus instalaciones para esa manufactura, hizo un informe¹³ para la National Aeronautics and Space Administration (NASA) en el que se describe el proceso de fabricación del cátodo M. La característica sobresaliente de este cátodo es su capacidad de operación a una temperatura alrededor de 50 a 100 °C más baja que la necesaria por el cátodo ordinario para la misma densidad de emisión.

Aún se realizan amplias pruebas sobre la duración del cátodo M y otros cátodos de matriz metálica por contrato con la NASA¹⁴. Los resultados obtenidos con el cátodo M demuestran que para la operación con una densidad de corriente de 2 A/cm² en un entorno simulado de tubo de ondas progresivas u ondas viajeras (TWT, de *traveling wave-tube*) este cátodo tiene más de 40 000 horas de duración. Es posible lograr una densidad de emisión de 4 A/cm² a una temperatura de funcionamiento aproximada de 1 020 °C durante miles

de horas. La operación del cátodo M a una densidad de emisión mayor de 5 A/cm² no es recomendable debido a que su rendimiento se deteriora con rapidez hasta el nivel del cátodo ordinario impregnado, por efecto de la pérdida del recubrimiento de osmio-rutenio en la parte superior de la matriz de tungsteno.

En la figura 11-1 se muestra la emisividad en función de la temperatura de operación para los cátodos M y B.

Emisores refractarios de alta temperatura

En la búsqueda de un sistema catódico mejorado también se han estudiado las propiedades electrónicas de muchos compuestos refractarios y semirrefractarios. Dichos compuestos se han venido empleando como emisores volumétricos, recubrimientos en sustratos refractarios e impregnantes de cátodos de matriz metálica.

Entre los emisores refractarios se encuentran los elementos de tierras raras y sus compuestos, y en las publicaciones especializadas se exponen emisores refractarios de torio, lantano, cerio, gadolinio, escandio y uranio. Los compuestos y aleaciones de torio se utilizan en dispositivos electrónicos, mientras que el dióxido de torio y sus carburos se usan en aplicaciones especiales como medidores de ionización.

El lantano, en particular su hexaboruro, se ha utilizado como la fuente de emisión en aplicaciones en las que es necesario exponer a la atmósfera el cátodo sin reducir el rendimiento en cuanto a emisividad. Debido a que el torio y el lantano se utilizan con frecuencia como emisores en aplicaciones especiales, sólo estos dos cátodos se analizarán en esta sección. Los cátodos de compuestos de torio y lantano, que tienen funciones trabajo superiores a las de los cátodos de óxido, impregnados o de metal refractario, requieren mayor temperatura de operación y por ello son más susceptibles a fallas del calentador (o calefactor). Los cátodos comunes de torio y lantano funcionan en el intervalo de temperaturas de 1 200 a 1 800 °C. Son resistentes, no se dañan fácilmente con el uso a casi el alto vacío, resisten la descarga disruptiva y son relativamente estables. Son fáciles de activar, tienen resistencia eléctrica muy baja y no forman capas interfaciales de alta resistencia con los metales base. Lafferty¹⁵ ha examinado extensamente los cátodos de hexaboruro en el General Electric Research Laboratory.

Los cátodos de tungsteno toriado se han empleado en tubos de transmisión de radiofrecuencias, son capaces de tolerar altas densidades de corriente con una vida razonablemente larga y se producen mediante la adición de un pequeño porcentaje de óxido de torio a tungsteno pulverizado; a continuación se sinteriza, estampa o estira el material a fin de producir alambre o un filamento que se calienta en una atmósfera de hidrocarburo para formar las capas superficiales de carburo de tungsteno. Los filamentos de tungsteno toriado son más susceptibles al daño por bombardeo de iones de gas residual que los cátodos de hexaboruro.

Se ha desarrollado un método con el cual es posible fabricar por compactación una composición particular de carburo de torio, que es una modificación del cá-

todo de tungsteno toriado, y después incorporarla a un vehículo para prensado en caliente al vacío. El cátodo consta de una copa de tantalio en la que se coloca la mezcla emisora, que se cubre después con una lámina ("hoja") de tantalio. Algunos cátodos cuyo contenido de composición de emisión está integrado por carburo de torio, óxido de torio y carburo de circonio han operado a densidades de corriente continua superiores a 30 A/cm² por periodos mayores de 25 horas. La temperatura común de operación de los cátodos es de 1 740 a 1 780 °C. En algunos artículos de Glascock^{16,17} y en uno de Bondley¹⁸ pueden encontrarse más detalles acerca de los cátodos de tungsteno toriado.

11.1.3 Cátodos emisores por campo eléctrico

Cátodos emisores por campo eléctrico de alto voltaje

El atractivo de esta clase de emisores radica en que se han informado densidades de corriente mayores de 100 A/cm² en puntas de material refractario muy finas de un solo extremo operadas en campos eléctricos mayores de 10⁷ V/cm. A partir de la formulación teórica de los cátodos emisores por campo eléctrico hecha por Fowler y Nordheim¹⁹ se ha logrado mucho en la comprensión y el conocimiento del proceso de emisión por campo eléctrico. Las propiedades físicas deseables para una fuente de emisión son:

1. Alta resistencia tensil, a fin de soportar el esfuerzo electrostático.
2. Alta resistencia al bombardeo iónico y a la contaminación.
3. Elevado punto de fusión.

Los materiales que poseen estas propiedades son:

1. Metales puros.
2. Cátodos heterogéneos que constan de una capa superficial sobre una sustancia sólida.
3. Compuestos o aleaciones metaloides homogéneos.

Estos materiales han encontrado aplicación como emisores por campo termoiónico de alto voltaje²⁰ (TFE, de *thermionic field emitters*) y como fuente de ionización por campo eléctrico (FI, de *field ionization*). Es posible obtener muestras de densidades de corriente muy altas para rayos menores o iguales de 0.1 μm con disposiciones ópticas relativamente sencillas de rayos a partir de fuentes TFE. Es posible lograr densidades de corriente de 1 300 A/cm² o densidades de potencia de 1.6 × 10⁷ W/cm² en un punto de haz de 0.1 μm^{21,22}. Chapman²³ y Spindt²⁴ han demostrado resultados parecidos. Los principales obstáculos para el empleo de estos cátodos en tubos de microondas son los siguientes:

1. Requieren alto voltaje de operación.
2. La operación estable puede lograrse sólo en vacíos menores o iguales a 10⁻⁸ torr.
3. La corriente total es muy baja, del orden de microamperes a miliamperes.

Cátodos emisores por campo eléctrico de bajo voltaje

El cátodo emisor por campo eléctrico de bajo voltaje o de película delgada^{25,26} consta de un conjunto conductor-aislante-conductor con grosor dieléctrico de 1 a 2 μm aproximadamente. El conductor superior o película de la compuerta metálica y la capa dieléctrica están perforados con agujeros que miden aproximadamente 2 μm de diámetro. Los agujeros están escariados en la capa dieléctrica y tienen conos de metal; la emisión por campo eléctrico se obtiene a partir de las puntas de los conos cuando dichas puntas se activan en modo negativo con respecto a la película de la compuerta. Debido a la intensificación por campo eléctrico en las puntas y la estrecha separación entre el perímetro del agujero en la película de la compuerta y las pequeñas puntas, para la emisión por campo se requiere un potencial de sólo 100 a 200 V a través del conjunto emparedado conductor-aislante-conductor.

Tanto los emisores por campo eléctrico de alto voltaje como los de bajo voltaje satisfacen la ecuación de Fowler-Nordheim^{25,27}:

$$J = \frac{AF^2}{t^2(y)\phi} \exp - \left[\frac{BV(y)\phi^{3/2}}{F} \right]$$

donde J = densidad de la corriente de emisión, en A/cm²

F = campo en la punta

ϕ = función trabajo, en eV

A y B son constantes, y $V(y)$ y $t(y)$ son funciones de variación lenta de y , que se encuentran tabuladas en las publicaciones especializadas²⁸. El campo en la punta es:

$$F = BV \text{ (V/cm}^2\text{)}$$

donde V es el voltaje aplicado. Asimismo:

$$B = f(\gamma, R, \theta) \text{ cm}^{-1}$$

donde γ es el radio de la punta, R es la separación entre el ánodo y la punta, y θ es el semiángulo del cono emisor. En general, puede afirmarse que a medida que γ , R y θ se hacen más pequeños, el factor de intensificación por campo eléctrico B se hace más grande.

Hasta la fecha, estos emisores han presentado capacidades de emisión de 10 a 20 A/cm² al operar en vacíos de 10⁻⁸ torr o mejores, y la corriente total típica de los emisores ha sido hasta de 100 mA. El potencial de estos emisores es muy promisorio para aplicaciones en dispositivos militares de rápido calentamiento, en tubos de ondas milimétricas y en aplicaciones comerciales en las que se requieren alta densidad de corriente, tamaño reducido de puntos y corriente total razonable (es decir, de 10 a 100 mA). Para averiguar que estos requerimientos pueden cumplirse, es necesario establecer la posibilidad de operación de estos cátodos en cañones de electrones, así como su funcionamiento en dispositivos de operación práctica. En la actualidad se realizan trabajos intensos para demostrar lo anterior en el Stanford Research Institute²⁵ y en el Georgia Institute of Technology²⁶.

Tabla 11-2. Características de emisión secundaria de materiales de uso común

Material	Rendimiento de la emisión secundaria, δ	Potencial medio
Platino	1.8	Medio-alto
Al ₂ O ₃	3-5	Bajo
BeO	8	Medio
MgO	8-15	Bajo
BaO-SrO	5-12	Medio
W-BaCaAlO ₄	2-3	Medio
W-ThO ₂	2-3	Medio
Au-MgO	3-8	Medio
GaAs	> 3	Bajo

11.1.4 Emisores que funcionan a temperatura ambiente

Cátodos fríos de emisión primaria

El advenimiento de los semiconductores de haces de electrones a fines de la década de 1960 generó bastante interés por el desarrollo de dispositivos catódicos fríos de estado sólido a temperatura ambiente²⁹⁻³².

Debido a que en la mayor parte de los casos estos emisores requerían un recubrimiento de cesio a fin de lograr propiedades de afinidad negativa a los electrones para densidades razonables de corriente (es decir, de 0.1 a 1 A/cm²), la operación en un tubo de vacío normal era difícil, si no imposible. Por consiguiente, debido a la dificultad de mantener una superficie activa de cesio se suspendió el desarrollo de este tipo de emisores para dispositivos de microondas.

Cátodos fríos de emisión secundaria

Bruining³³ y Hachenberg³⁴ han informado en artículos de revisión la emisión secundaria de electrones. Es bien conocido que este fenómeno puede ser deseable (en un emisor) o indeseable (en un multiplicador) en el diseño de dispositivos de electrones. El presente análisis se limita a su potencial como emisor.

La emisión secundaria de electrones ocurre principalmente en la superficie de los materiales. Metales, semiconductores, aislantes y líquidos presentan características de emisión secundaria. El rendimiento de la emisión secundaria δ de un material objetivo ("blanco") se define como la razón del número total de electrones secundarios entre el número de electrones primarios, y depende en gran medida de la energía de los electrones primarios (eV) y del ángulo de incidencia del haz primario. En la tabla 11-2 se enumeran algunas de las características de la emisión secundaria de materiales de uso común. Para los metales, como regla general δ varía de 0.7 a 1.7. Para haluros alcalinos, óxidos alcalinos y óxidos alcalinotérreos, δ puede alcanzar valores de 20 o mayores, pero tiene corta duración. El material ideal para un cátodo de emisión secundaria es aquel que:

1. Presenta un valor de δ mayor de 2.5.
2. Tiene un primer punto de cruce de baja energía (<30 V).
3. Es capaz de soportar bombardeos.
4. Es insensible al ambiente del tubo normal de microondas.
5. Es estable química, mecánica y térmicamente.

En la tabla 11-2 se incluye un metal cerámico (Au-MgO), el cual se ha demostrado que tiene el potencial para satisfacer todos los criterios mencionados.

Para el MgO se han informado 34 valores de δ hasta de 20, pero el uso amplio de este material en tubos de microondas requerirá mejoras a su estabilidad en un ambiente normal de operación de tubos.

La figura 11-2 es la gráfica de una curva representativa de emisión secundaria, donde Ep_1 es el primer punto de cruce; es decir, δ es mayor de 1, Ep_m es el punto en el que δ es máximo y Ep_2 es el punto de la curva en el que δ decrece hasta menos de 1. En todos los casos la energía del primer cruce, Ep_1 , es menor que la del segundo, Ep_2 . En las referencias 33 y 34 pueden encontrarse mayor información al respecto.

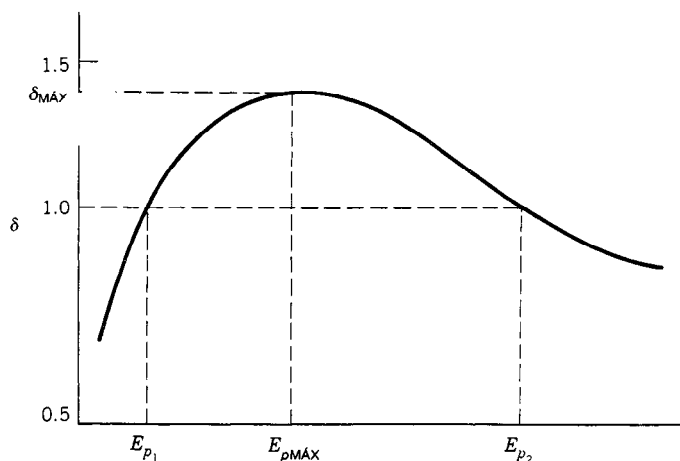


Fig. 11-2. Curva representativa de emisión secundaria.

Tabla 11-3. Cátodos en proceso de desarrollo

Tipo de cátodo	Beneficios	Observaciones	Área de aplicación
Tungsteno-iridio	Alta densidad de corriente	Difícil de fabricar, densidad de corriente superior a 6 A/cm ² , temperatura aproximada de 1 000 °C	ECM ^a , tubos de alta potencia para ondas milimétricas
Densidad de porosidad controlada (CPD, de <i>controlled porosity density</i>)	Densidad de corriente uniforme, potencial para larga duración	Programa de manufactura planificado	Tubos espaciales, tubos para ondas milimétricas
Tungsteno	Densidad de corriente de 2 a 20 A/cm ²	La versión prensada es difícil de fabricar; la versión impregnada está en proceso de desarrollo; se requiere trabajar con los impregnantes a fin de reducir la temperatura de impregnación	ECM, tubos de alta potencia para ondas milimétricas
Cátodo codepositado	Densidad de corriente de 2 a 20 A/cm ² , baja velocidad de sublimación	No disponible comercialmente, se requiere más información	ECM, tubos para ondas milimétricas
Emisores de campo eléctrico de bajo voltaje	No precisa calefactor (catódico), alta densidad de corriente, funcionamiento instantáneo	Corriente total 100 mA, potencial para corrientes mucho mayores; se requieren datos sobre funcionamiento en el ambiente del tubo	Dispositivos espaciales, dispositivos de potencia intermedia, para ondas milimétricas, ECM
Escandato	Baja temperatura de impregnación, fácil de fabricar	Se necesita más trabajo para mejorar su confiabilidad y capacidad de repetición; no se dispone de datos sobre su duración	Tubos de alta potencia para microondas, tubos para ondas milimétricas

^a ECM = contramedidas electrónicas.

11.1.5 Nuevos cátodos en proceso de desarrollo

En Varian Associates se encuentra en proceso de desarrollo una variante del cátodo M recubierto con osmio-rutenio³⁵. Este cátodo consta de una matriz de tungsteno-iridio (W-Ir) impregnada con aluminato cálcico de bario. Con este cátodo se han obtenido densidades de emisión de 4 a 8 A/cm² operando en el intervalo de temperatura de 1 025 a 1 100 °C. La aplicación de este cátodo se ha limitado a tubos del tipo de investigación y desarrollo que requieren densidades de corriente superiores a 4 A/cm².

En los *Philips Laboratories*³⁶ de Eindhoven, Holanda, está en proceso de desarrollo un cátodo basado en una matriz de tungsteno prensada o impregnada con escandato de bario (en vez del aluminio de bario que suele emplearse en la matriz de tungsteno), y se han informado valores de emisión muy elevados, pero aún falta que dichos resultados sean repetidos por los investigadores de Estados Unidos.

También se ha informado sobre cátodos basados en una matriz de tungsteno fabricada con Ba₂Sr(WO₆)₂ (tungstato de estroncio y bario). El trabajo original sobre este cátodo fue realizado por General Electric (Schenectady, EE.UU.) en una matriz prensada, y en la actualidad está siendo efectuado por Smith y Newman³⁷ para fabricar el cátodo de tungstato por impregnación. Se prevé que cuando se haya desarrollado por completo, la versión impregnada será capaz de producir densidades de emisión de 8 A/cm² a 1 025 °C, con una vida superior a las 5 000 h.

El Naval Research Laboratory (NRL)³⁸ de Washington, DC, está desarrollando un cátodo de matriz con densidad de porosidad controlada. Este cátodo consta de una película delgada de BaO sobre una superficie de iridio en la que se emplean técnicas de microprocesamiento para fabricar "papel" (lámina) de iridio de 0.001 pulg con una disposición uniforme de huecos (agujeros). Durante la operación, este "papel" se recubre con una capa muy delgada de Ba-BaO. De esta manera, el cátodo adquiere una superficie muy uniforme y con baja función de trabajo. En Varian se han logrado densidades de emisión de 6 A/cm² de CD en un sistema de colector hundido con un cátodo de Ir-BaO.

Debe señalarse que en la etapa actual ninguno de los nuevos cátodos puede considerarse superior. Queda por realizar bastante trabajo antes de poder establecer su nivel de confiabilidad y la capacidad de repetición de su funcionamiento. En la tabla 11-3 se presentan algunos de los cátodos en proceso de desarrollo y datos sobre su potencial de emisión y posibles aplicaciones.

REFERENCIAS BIBLIOGRÁFICAS

- Herring, Conyers, "Thermionic Emission", *Rev. Mod. Phys.* **21**(2).
- S. Wagner y G. Herrman, *The Oxide Coated Cathode*, Vols. 1 y 2, Chapman and Hall, Londres.
- C. P. Hadley, W. G. Rudy y A. J. Sloechert, "A Study of the Molded Nickel Cathode", *J. Electrochem Soc.* **105**:395-398.
- G. Medicus, Presentación en 1978 Tri-Services Cathode Workshop, Naval Research Laboratory, Washington, D.C.
- D. W. Mauer y C. M. Pleass, "The CPC: A Medium Current Density, High Reliability Cathode", *Bell Syst. Tech. J.* **46**(10).
- D. A. Wright, "A Survey of Present Knowledge of Thermionic Emitters", *Proc. Inst. Elec. Engr.* **100**:1250142.
- L. S. Nergaard, "Studies of the Oxide Cathode", *RCA Rev.* **13**:464-545.
- E. A. Coomes, *J. Appl. Phys.* **17**:647.
- J. Affleck y W. T. Boyd, "Investigation of Various Activator Refractory Substrate Combinations", *Gen. Elec. Tech. Inform. Series*, R66ET 1-9, 12.
- H. J. Lemmens, "A New High Emission Density Thermionic Emitting Cathode", *Phillips Tech. Rev.* **11**:341.
- A. Venema, R. C. Hughes, P. P. Coppola y R. Levi, "Dispenser Cathodes", *Phillips Tech. Rev.* **19**(6):177.
- P. Zalm y A. J. Van Stratum, "Osmium Dispenser Cathode", *Phillips Tech. Rev.* **27**:69.
- A. Gupta, "The Manufacture of M-Type Impregnated Cathode", NASA/Lewis Research Center, NASA Order No. C-2129-D.
- R. Gorske, NASA Contract NAS 3-22335, "Design, Construction and Long Life Endurance Testing of Cathode Assemblies for Use in Microwave High Power Transmitting Tubes".
- J. M. Lafferty, "Boride Cathodes", *J. Appl. Phys.* **22**:299.
- H. Glascock, *Rev. Science Instru.* **43**:698.
- H. Glascock, *Rev. Science Instru.* **47**:90.
- R. J. Bondley, Contract DA28-043-AMC-01719(E), ARPA ORDER n.º 679, "High Current Density, Short Life Cathodes for Linear Beam Tubes".
- R. W. Fowler y L. W. Nordheim, "Electron Emission in Intense Electric Fields", *Proc. Roy. Soc. A* **119**:173.
- L. W. Swanson y N. A. Martin, "Field Emission Cathode Stability Studies: Zirconium/Tungsten Thermal Field Cathode", *J. Appl. Phys.* **46**:2029.
- W. P. Dyke y J. K. Trolan, "Field Emission: Large Current Densities, Space Charge, and the Vacuum Arc", *Phys. Rev.* **89**:799-808.
- L. W. Swansen, J. Orloff y A. E. Bell, "Field Electron and Ion Source Research for High Density Information Storage System", Air Force Contract AFA1-TR-79-1133.
- A. T. Chapman, Arpa Contract NAAH01-71-C-1046 "Melt Grown Oxide Metal Composites".
- C. A. Spindt, *J. Appl. Phys.* **39**:304.
- C. A. Spindt, NASA Contract NAS3-21507, "Development Program on a Cold Cathode Electron Gun".
- Contract DAAK40-77-0096, MICOM, "Manufacturing Method for the Production of Low Voltage Field Emitters".

27. R. H. Fowler y L. W. Nordheim, *Proc. Roy. Soc. (Londres)* **A-119**:73.
28. R. B. Burgess, H. Kroemer y J. M. Houston, *Phys. Rev.* **90**:515.
29. K. R. Faulkner y J. R. Howorth, CVD Research Project RP 28-1, "Semiconductor Cold Cathode Displaying a Sharp High Energy Cut-Off", Contract N/CP16/1711.
30. K. R. Faulkner y cols., "Negative Electron Affinity GaAsP Cold Cathode Silicon Vidicon", 1973 International Electron Device Meeting Technical Digest, Washington, DC.
31. H. Kressel y cols., "An Optoelectronic Cold Cathode Using an AlxGa_{1-x} as Heterojunction Structure", *App. Phys. Lett.* **16**:9.
32. H. Schade y cols., "Optoelectronic Electron Emitter", Contract DAA-V07-71-C-0054.
33. H. Bruining, *Physics and Applications of Secondary Emission*, McGraw-Hill, Nueva York.
34. O. Hachenberg y W. Brauer, "Secondary Electron Emission from Solids", in *Advances in Electronics and Electron Physics*, vol XI, Academic, Nueva York, p. 413.
35. L. K. Falce, Patent n.º 4, 165, 473, "Electron Tube with Dispenser Cathode".
36. A. Van Oostron y L. Augustus, "Activation and Early Life of a Pressed Barium Scandate Cathode", *Appl. Surface Sci.* **2**:2.
37. B. Smith y A. Newman, Patent n.º 4 078 900, "Method of Making a High Current Density Long Life Cathode".
38. L. Falce, Contract 00173-77-C-0186, "Iridium Foil Cathode Development".

11.2 MATERIALES ÓPTICOS Y FOTOSENSIBLES

Mansoor A. Saifi

11.2.1 Materiales ópticos e indicatriz

Un material óptico puede ser gaseoso, líquido o sólido. Los materiales ópticos de mayor interés en ingeniería son sólidos cristalinos y amorfos. Una nueva clase de materiales conocidos como cristales líquidos se está empleando en muchos dispositivos como matriz indicadora o de lectura; estos materiales se analizan con detalle en el artículo 15.3.

Los sólidos cristalinos se caracterizan por una disposición periódica ordenada de los átomos, y los sólidos que no presentan esta periodicidad en su estructura se denominan amorfos.

El mayor grupo de materiales amorfos empleados en aplicaciones ópticas es el de los vidrios y las cerámicas. Una propiedad óptica importante es el índice de refracción.

Cuando la luz pasa del vacío a un medio más denso, como el vidrio, se reduce su velocidad; la relación entre las velocidades de la luz en ambos medios se denomina índice de refracción n del medio más denso, y está dada por:

$$n = \frac{\text{Velocidad en el vacío } (c)}{\text{Velocidad en el medio } (v)} = \sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0}} \quad (11-1)$$

donde ϵ y μ son la permitividad y la permeabilidad del medio y ϵ_0 y μ_0 son la permitividad y la permeabilidad del vacío, respectivamente. La longitud de onda y la dependencia direccional del índice de refracción están relacionadas con propiedades materiales, electrónicas y estructurales como la simetría. En los materiales isotrópicos el índice de refracción no depende de la dirección de propagación y de la polarización de la luz, mientras que en los materiales anisotrópicos dependen de ambas. La anisotropía también puede inducirse por aplicación de campos eléctricos y magnéticos o mediante esfuerzos y deformaciones mecánicos. Los materiales que presentan grandes cambios en sus propiedades ópticas por inducción de campo son de especial interés para varias aplicaciones en dispositivos. Una forma conveniente de describir el índice de refracción en función de la dirección de propagación y de la polarización es por medio del elipsoide de índices o indicatriz¹. La ecuación de la superficie elipsoidal es:

$$B_{11}x_1^2 + B_{22}x_2^2 + B_{33}x_3^2 + B_{23}x_2x_3 + B_{13}x_1x_3 + B_{12}x_1x_2 = 1 \quad (11-2)$$

donde x_i son las coordenadas cartesianas, y los recíprocos de los coeficientes B_{ij} están relacionados con el índice de refracción. El efecto del campo aplicado sobre el índice de refracción se describe entonces por medio de coeficientes adecuados que relacionan los B_{ij} con los campos aplicados. En un artículo posterior se analizarán varios de estos coeficientes. Basándonos en la figura 11-3, es posible obtener el índice de refracción de una onda de propagación a partir del plano perpendicular a la dirección de propagación y que pasa por el origen de la indicatriz. La intersección de este plano con la indicatriz define una elipse cuyos semiejes mayor y menor (OA y OB , respectivamente) representan los índices de refracción de la luz polarizada en esas direcciones. En el caso especial en que la dirección de propagación coincide con el eje mayor de una indicatriz, los índices de refracción están dados por las longitudes de los otros dos ejes. Los índices que corresponden a los tres ejes principales se denominan índices principales de refracción, y en el sistema coordenado de ejes principales la ecuación de la indicatriz es:

$$\frac{x_1^2}{n_1^2} + \frac{x_2^2}{n_2^2} + \frac{x_3^2}{n_3^2} = 1 \quad (11-3)$$

donde los n_i pueden obtenerse a partir de los coeficientes B_{ij} .

11.2.2 Vidrios ópticos

Debido a su transparencia en la región visible del espectro (de 400 a 800 nm de longitud de onda), el vidrio es el material óptico de mayor uso en aplicaciones como lentes, prismas y ventanas. Cuando sobre una

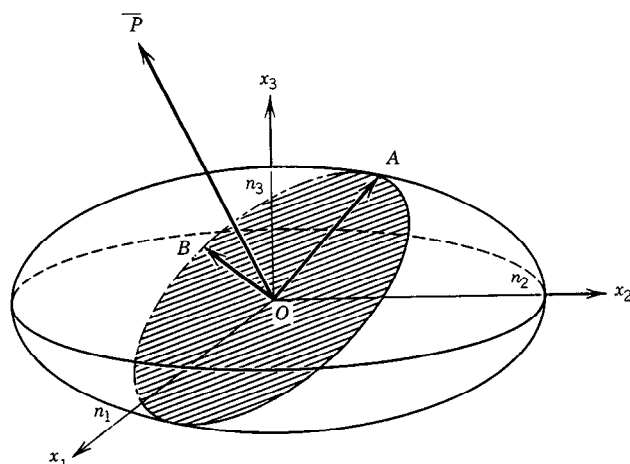


Fig. 11-3. Elipsoide índice o indicatriz. OP = dirección de propagación de la luz. El plano sombreado es perpendicular a OP.

superficie de vidrio incide un haz de luz ocurren los fenómenos de reflexión, refracción y absorción. La proporción de luz reflejada en una superficie sencilla pulida puede determinarse con la ecuación de Fresnel:

$$R = \frac{(n - 1)^2}{(n + 1)^2} \quad (11-4)$$

donde R es el coeficiente de la intensidad de la luz reflejada entre la intensidad de la luz incidente y n es el índice de refracción del vidrio (tabla 11-4). En la región visible del espectro, el índice de refracción de casi todos los vidrios comerciales varía de 1.45 a 2, lo que corresponde a una reflectancia que va del 3.36 al 11.11%. Es posible reducir estos valores de la reflectancia mediante la aplicación de recubrimientos ópticos idóneos (artículo 11.2.8). Un componente primordial de los vidrios ópticos es el dióxido de silicio o sílice (SiO_2), con la adición de óxidos de muchos otros elementos, como bario, calcio, boro, potasio y plomo. Así, el índice de refracción y su dispersión (variación con la longitud de onda) dependen de la composición del vidrio. De los vidrios hechos de un solo componente, el de sílice tiene el menor índice de refracción. La

razón de cambio del índice de refracción con la longitud de onda λ , es decir, $dn/d\lambda$, se conoce como dispersión. Para los vidrios que se utilizan en el espectro visible, las características de refracción suelen especificarse de manera convencional mediante dos números: el índice de refracción para la línea D del sodio ($0.5893 \mu\text{m}$) y el número V de Abbe. El número V , que suele denominarse dispersión relativa recíproca, se define como:

$$V = \frac{n_d - 1}{n_f - n_c} \quad (11-5)$$

donde n_d , n_f y n_c son los índices de refracción para la línea D del sodio, para la línea F del hidrógeno ($0.4861 \mu\text{m}$) y para la línea C del hidrógeno ($0.6563 \mu\text{m}$), respectivamente. La diferencia ($n_f - n_c$) es una buena medida de dispersión. Esta información es de utilidad en el diseño de elementos ópticos tales como las lentes y los instrumentos más complejos que operen en el intervalo visible del espectro. Es interesante anotar que los vidrios que tradicionalmente se encuentran en el intervalo $n_d > 1.6$, $V_d > 50$ y $n_d < 1.6$, $V_d > 55$ se denominan cristales corona o vidrios sin plomo (*crowns*), mientras que a los demás se les denomina cristales de roca o vidrios con plomo (*flints*).

Tabla 11-4. Vidrios ópticos y sus índices de refracción

Tipo de vidrio ^a	N_f	N_d	N_c
FK5-487704	1.492	1.487	1.486
PK1-504669	1.509	1.504	1.501
BK3-498651	1.504	1.498	1.496
SK1-610567	1.607	1.610	1.618
KF6-517522	1.524	1.517	1.514
SF1-717295	1.735	1.717	1.711

^a Los tipos de vidrios son de "Optical Glass", Catálogo núm. 3050/USA, Schott Optical Glass Inc. El código de seis cifras se compone de dos partes. Las tres primeras cifras describen el índice de refracción nominal N_d y las últimas tres, el número de Abbe.

11.2.3 Recubrimientos ópticos

Para modificar las propiedades de reflexión y transmisión de lentes, prismas y ventanas se emplean recubrimientos metálicos o dieléctricos delgados. Dependiendo del uso pretendido, en el mercado existen recubrimientos sencillos o de múltiples capas sobre un sustrato para diversas aplicaciones, con reflectancias desde muy bajas hasta muy elevadas, en un intervalo espectral estrecho o amplio. La terminología de mayor uso para dichas estructuras es AR (antirreflejante), HR (de alta reflexión), BBAR (antirreflejantes de banda ancha) y filtros ópticos (de alta transmisión en

una banda dada). Es evidente que la elección del material de recubrimiento y del sustrato depende de la longitud de onda de operación. En la tabla 11-5 se indican algunos materiales de uso común, como recubrimientos y sus intervalos espectrales útiles. Los recubrimientos se aplican por evaporación térmica, calentamiento por haz de electrones o bombardeo iónico. A continuación se describen brevemente los principios que caracterizan el comportamiento de los recubrimientos sencillos y de varias capas.

En base a la ecuación de Fresnel (11-4), la reflectancia R de un sustrato dieléctrico para una incidencia normal está dada por:

$$R = \frac{(n_s - n_0)^2}{(n_s + n_0)^2} \quad (11-6)$$

donde n_s y n_0 son, respectivamente, los índices de refracción del sustrato y del medio a la longitud de onda λ . La aplicación de una película dieléctrica sencilla con índice n_1 y grosor $d_1 = (\lambda/4n_1)(1 + 2m)$, donde $m = 0$ o cualquier valor entero, modifica⁴ la reflectancia a:

$$R = \frac{(n_1^2 - n_0 n_s)^2}{(n_1^2 + n_0 n_s)^2} \quad (11-7)$$

Para $m = 0$ el grosor d_1 es igual a un cuarto de longitud de onda de la luz en el material de recubrimiento, y los valores enteros de m proporcionan un grosor extra igual a la mitad de la longitud de onda ($\lambda/2n_1$). En esencia, las propiedades de reflexión del sustrato recubierto se modifican por la interferencia de la luz reflejada en las interfaces aire-película y película-sustrato. Existen dos casos de particular interés. Primero, para $n_0 < n_1 < n_s$, en la ecuación 11-7 puede observarse que la reflectancia de un sustrato recubierto siempre es menor que la del sustrato mismo (ec. 11-6). De este modo se forman los recubrimientos antirreflejantes mencionados antes. Es claro a partir de la ecuación 11-7 que para el índice pelicular $n_1 = \sqrt{n_0 n_s}$ el valor teórico de R es 0. Con frecuencia no es posible fabricar películas exactamente con este índice de refracción, por lo que es necesario utilizar recubrimien-

tos de capas múltiples. Los recubrimientos con dos capas proporcionan cuatro parámetros ajustables; dos índices de refracción y dos grosores.

Para un recubrimiento con dos capas, cada una de las cuales es de un cuarto de onda, es posible demostrar que⁴:

$$R = \frac{(n_0 n_2^2 - n_1^2 n_s)^2}{(n_0 n_2^2 + n_1^2 n_s)^2} \quad (11-8)$$

De nuevo, para $n_1/n_2 = \sqrt{n_0/n_s}$, el valor teórico de R es 0. En caso de que no sea posible satisfacer esta condición, se pueden ajustar los grosores del recubrimiento a fin de minimizar la reflexión². Los recubrimientos de varias capas diseñados para baja reflectancia en un intervalo espectral estrecho suelen denominarse recubrimientos V , de los cuales, en la figura 11-4, se muestra una característica común.

Para el segundo caso, cuando $n_0 < n_1 > n_s$, la reflectancia de un sustrato recubierto con una sola película (ec. 11-7) es mayor que la del sustrato mismo. Es posible obtener más del 99% de reflectancia (lo que es de particular utilidad para los espejos de láser) mediante la acumulación de pares consecutivos de recubrimientos de un cuarto de onda con índices de refracción altos (H) y bajos (L). La elevada reflectividad de esta estructura se debe a la interferencia constructiva de reflexión de todas las interfases. Esta estructura puede denotarse por $[a(HL)^m S]$, donde m representa el número de apilamientos y a y S representan el aire y el sustrato, respectivamente. Obsérvese que la película más superficial es la que siempre tiene el índice más alto. Además, si $n_H > n_L < n_s$, a continuación del sustrato se utiliza una capa H extra. Los recubrimientos diseñados para valores intermedios de reflexión y transmisión se denominan separadores de haces.

Los filtros de paso de onda larga presentan alta reflectividad a longitudes de onda corta y transmiten longitudes de onda larga, mientras que los filtros de paso de onda corta hacen lo contrario. En muchos casos se emplean apilamientos de un cuarto de onda a fin de obtener propiedades de alta reflectancia, con películas extra no de un cuarto de onda cerca del aire y del

Tabla 11-5. Materiales para recubrimiento de uso común

Material, índice de refracción nominal	Intervalo espectral				Núm. de ref.
	Ultravioleta	Visible	Infrarrojos (< 5 μ m)	Infrarrojos (> 5 μ m)	
PbF ₂ , 1.75	←————→				2
MgF ₂ , 1.38		←————→			2
Al ₂ O ₃ , 1.62		←————→			2
SiO ₂ , 1.46		←————→			2
ZrO ₂ , 2		←————→			2
SiO, 1.9		←————→			2
Si, 3.5		←————→			2
Ge, 4.05		←————→			2
CdTe, 2.69		←————→			3
ZnSe, 2.44		←————→			3

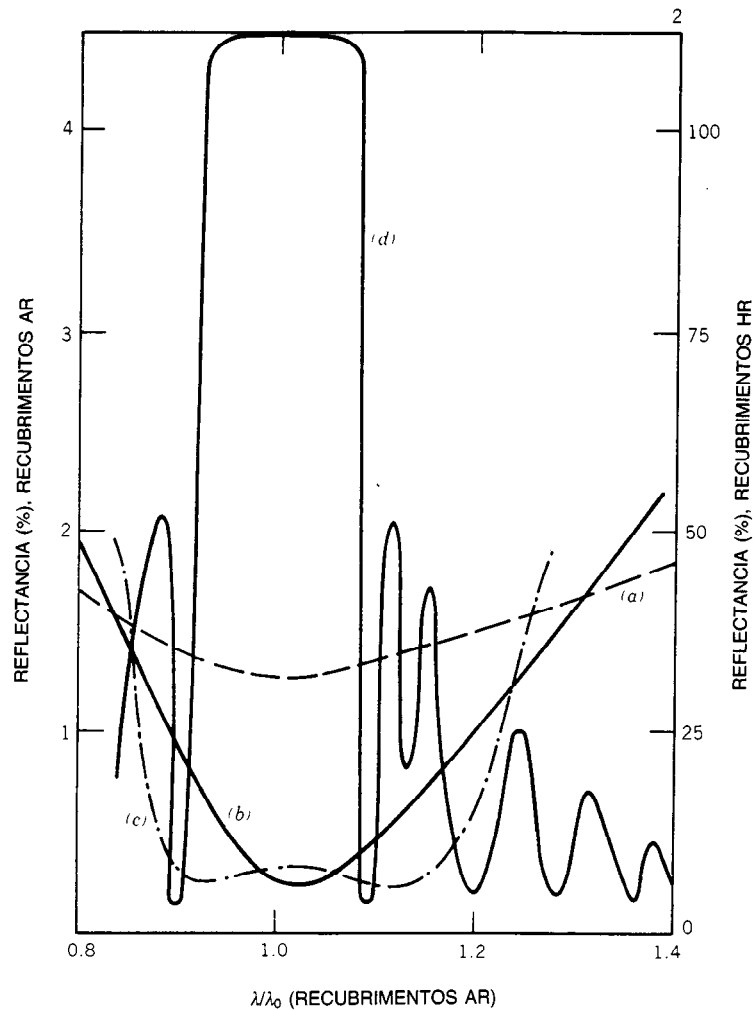


Fig. 11-4. Espectros de reflexión típicos para algunos recubrimientos: *a*) antirreflejantes (AR) de una capa; *b*) recubrimiento AR de dos capas, recubrimiento V; *c*) recubrimiento AR de banda ancha; *d*) recubrimiento de alta reflexión (HR) de varias capas. λ = longitud de onda de operación.

sustrato para lograr elevada transmitancia a las longitudes de onda deseadas. Por ejemplo, las estructuras:

$$S(0.5H)L(HL)^m(0.5H)a, \quad S(0.5L)(HL)^mH(0.5L)a$$

se usan típicamente como filtros de paso de onda larga y onda corta, respectivamente. Aquí $0.5H$ y $0.5L$ son recubrimientos de un octavo de onda con índices alto y bajo. Los recubrimientos extra se diseñan para eliminar los picos de reflexión secundaria que se observan alrededor de la región de alta reflectividad de los apilamientos de un cuarto de onda (fig. 11-4).

Las propiedades de reflectancia y transmitancia dependen en gran medida del ángulo de incidencia, por lo que deben especificarse con claridad. Suele suceder que la polarización de la onda reflejada es bastante diferente de la polarización de la onda transmitida.

11.2.4 Materiales electroópticos

El efecto electroóptico se refiere al cambio en el índice de refracción debido a un campo eléctrico aplicado. Existen dos partes en el origen físico del efecto: una parte electrónica producida por la interacción del campo eléctrico aplicado con los dos electrones, y una parte nuclear por el acoplamiento del desplazamiento nuclear inducido por el campo y los electrones. En muchos casos es posible expresar el índice de refracción como una serie de potencias del campo aplicando:

$$n = n_0 + a_1E + a_2E^2 + a_3E^3 \dots + a_mE^m \quad (11-9)$$

donde n_0 es el índice de refracción en ausencia del campo eléctrico E y a_m son constantes. En materiales isotrópicos como los líquidos y en cristales con centro

Tabla 11-6. Algunas características de los materiales ferroeléctricos

Material	Coefficiente electroóptico	Índice de refracción	Longitud de onda (μm)	Temperatura de Curie, T_c (K)	Núm. de ref.
BaTiO ₃	$g_{11} = 0.12, g_{12} = -0.01$	2.4	0.6328	401	6
KTaO ₃	$g_{11} - g_{12} = 0.16, g_{44} = 0.12$	2.24	0.6328	4	7
KTa _{0.65} Nb _{0.35} O ₃	$g_{11} = 0.136, g_{12} = -0.38, g_{44} = 0.147$	2.29	0.6328	283	7
BaTiO ₃	$r_{51} = r_{42} = 1280$ $r_c^{(b)} = 1.08$	$n_1 = n_2 = 2.44$ $n_3 = 2.37$	0.546	393	6
LiNbO ₃	$r_{13} = 8.6, r_{33} = 30.8$ $r_{51} = r_{42} = 28$ $r_c = 21$	$n_1 = n_2 = 2.29$ $n_3 = 2.2$	0.6328	1470	8
GaP	$r_{41} = 1.06 S$	$n = 3.315$	0.6	—	9
GaAs	$r_{41} = 0.27-1.2 S$	$n = 3.6-3.42$	1-1.8	—	10

^a T = esfuerzo constante.^b $r_c = r_{33} - (n_1/n_2)^2 r_{13}$.^c S = deformación constante.
 r en 10^{-12} m/V, g en m^2/C^2 .

de simetría, el índice de refracción es independiente de la dirección del campo aplicado ($\pm E$), y por ello se eliminan los coeficientes a_m para m entero impar. El cambio en el índice de refracción ($\Delta n = n - n_0$) con el cuadrado del campo eléctrico aplicado y con los términos pares de orden superior se denomina efecto electroóptico cuadrático o efecto Kerr. El cambio en el índice de refracción (Δn) con el campo eléctrico aplicado produce birrefringencia, que en términos del coeficiente K de Kerr está dada por:

$$\Gamma = \frac{(n_p - n_s)l}{\lambda} = IKE^2 \quad (11-10)$$

donde n_p y n_s son índices de refracción paralelo y perpendicular al campo aplicado E a una longitud de onda λ , y l es la longitud de la trayectoria.

Una clase importante de materiales electroópticos son los ferroeléctricos, que poseen elevada constante dieléctrica en la vecindad de la temperatura de Curie y presentan un efecto electroóptico de orden superior cuya magnitud es significativa. Los coeficientes electroópticos y elastoopticos de los materiales cristalinos dependen de la dirección cristalográfica y son cantidades tensoriales⁵. Así, las propiedades ópticas se describen mediante una indicatriz generalizada (ec. 11-2).

El cambio en B_{ij} debido al campo eléctrico aplicado describe el efecto electroóptico. Para el efecto Kerr en los materiales ferroeléctricos es conveniente emplear como variable independiente la polarización eléctrica P en vez del campo eléctrico E . Entonces, el cambio en B_{ij} está dado por:

$$\Delta B_{ij} = g_{ij,k}(P^2)_k \quad (11-11)$$

donde k es la dirección de polarización. En notación tensorial reducida*:

$$\Delta B_1 = g_{1k}(P^2)_k \quad (11-12)$$

* La notación reducida para los subíndices ij es 11 = 1, 22 = 2, 33 = 3, 23 = 32 = 4, 13 = 31 = 5, 12 = 21 = 6.

En su aplicación a dispositivos, estos materiales suelen mantenerse por encima de su temperatura de Curie, en cuyo caso los coeficientes convencionales de Kerr están dados por:

$$K_{100} = \frac{2n^3(g_{11} - g_{12})\epsilon^2}{\lambda} \quad (11-13)$$

y

$$K_{100} = \frac{2n^3g_{44}\epsilon^2}{\lambda} \quad (11-14)$$

donde n es el índice de refracción y ϵ es la constante dieléctrica. Por encima de la temperatura de Curie, T_c , la constante dieléctrica de los materiales ferroeléctricos obedece la Ley de Curie-Weiss $\epsilon = C/(T - T_c)$, donde T es la temperatura de operación y C es la constante de Curie. En la tabla 11-6 se proporcionan los valores de diversas constantes para algunos materiales ferroeléctricos.

Una clase muy importante de materiales electroópticos son los cristales cuyas estructuras moleculares no presentan un centro de simetría. Tales materiales tienen coeficientes diferentes de cero para las potencias impares del campo eléctrico E en la ecuación 11-9. El efecto electroóptico debido al término lineal en el campo eléctrico se denomina efecto Pockell. De nuevo, el efecto electroóptico es descrito por los cambios en los coeficientes B_{ij} debidos al campo eléctrico inducido.

Por convención, los coeficientes r describen el efecto Pockell con el campo eléctrico E , y los coeficientes g describen el efecto Kerr (efecto electroóptico cuadrático) con polarización eléctrica P . En la tabla 11-6 se presentan los valores de los coeficientes de Pockell para diversos materiales.

Los materiales electroópticos se utilizan en dispositivos como moduladores, memorias, placas de retardo variable y conmutadores ópticos. Estas aplicaciones se basan en el fenómeno de la birrefringencia por el campo inducido, que provoca que una onda incidente po-

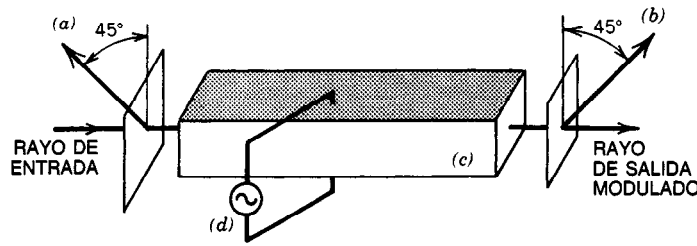


Fig. 11-5. Representación esquemática de un modulador electroóptico: a) y b) polarizador y analizador. Obsérvese que sus ejes son perpendiculares entre sí; c) cristal electroóptico con electrodos evaporados; d) señal de modulación.

larizada plana se vuelve polarizada elípticamente después de pasar a través del material. En los moduladores electroópticos se emplean polarizadores transversales, como los que se muestran en la figura 11-5, para impartir información a una onda portadora por medio de modulación de la intensidad.

Un campo que está creciendo con rapidez es el desarrollo de los dispositivos electroópticos con película delgada, como moduladores y deflectores de guías de ondas^{10,11,12}.

11.2.5 Materiales elastoópticos

El efecto elastoóptico se refiere al cambio en el índice de refracción debido a deformación o esfuerzo mecánicos. El origen físico de este efecto es el corrimiento en la separación nuclear y el acoplamiento entre varios campos cristalinos. Como en el caso de los materiales electroópticos, el efecto se describe por medio de la modificación en la indicatriz inducida por deformación o por esfuerzo. Los cambios en el índice de refracción están dados por un tensor de cuarto orden óptico de deformación u óptico de esfuerzo. Los cambios en el coeficiente B (ec. 11-21) debidos a una o más deformaciones aplicadas están dados en notación reducida por:

$$\Delta B_i = \sum_{j=1}^6 p_{ij} S_j \quad (11-15)$$

Por tanto, las componentes del tensor óptico de deformación B_{ij} están dadas por un arreglo de los elementos de una matriz 6×6 . Dependiendo de la estructura y la simetría cristalinas, muchas de estas componentes son cero o tienen la misma magnitud. En

la tabla 11-7 se presentan los coeficientes ópticos de deformación de algunos materiales. La birrefringencia Γ inducida por una deformación se determina despejando la dirección deseada de propagación de la onda en la ecuación de la indicatriz modificada:

$$\Gamma = n^3(p_{ij}S_j - p_{lk}S_k)\gamma l \quad (11-16)$$

donde n es el índice de refracción, l es la longitud de la trayectoria de propagación y γ es una constante del material casi igual a la unidad.

Los coeficientes ópticos de esfuerzo q_{ij} se definen de manera parecida a como se hace con los coeficientes ópticos de deformación, y ambos están relacionados mediante la ecuación de módulos de elasticidad:

$$q_{ij} = p_{ij}C_{kj} \quad (11-17)$$

Los materiales elastoópticos se emplean en dispositivos ópticos como conmutadores Q de láser, moduladores y exploradores (*scanners*). Todos estos dispositivos se basan en la dispersión de la luz por medio de ondas sonoras^{16,17}. Se lanza una onda acústica en un material elastoóptico transparente mediante un transductor piezoeléctrico (fig. 11-6). El efecto elastoóptico provoca que la onda acústica vaya acompañada por una onda semejante de variación del índice de refracción n . El patrón periódico de los índices de refracción alternados máximos y mínimos actúa como una eficiente rejilla de difracción.

Para la trayectoria de la luz $l < \Lambda^2/\lambda$ (Λ y λ son longitudes de onda acústica y óptica en el medio), la luz se divide en muchos órdenes separados por ángulos aproximados de λ/Λ . Este fenómeno se conoce como la dispersión de Debye-Sears (en un medio líquido) o

Tabla 11-7. Coeficientes elastoópticos de algunos materiales

Material	Longitud de onda (μm)	Coeficiente elastoóptico	M_2^a	Núm. de ref.
Sílice fundida	0.63	$p_{11} = 0.121, p_{12} = 0.27$	1	13
Lucita	0.63	$p_{11} = 0.3, p_{12} = 0.28$	33	14
GaP	0.63	$p_{11} = -0.151, p_{12} = -0.082, p_{44} = 0.074$	29.5	13
GaAs	1.5	$p_{11} = -0.165, p_{12} = -0.14, p_{44} = -0.074$	69	13
Ge	10.6	$p_{11} = 0.27, p_{12} = 0.235, p_{44} = 0.125$	540	15

^a M_2 , cifra de mérito, normalizada en relación con la sílice fundida. Cifra de mérito = $n^6 p^2 / \rho v^3$, donde ρ = densidad, v = velocidad acústica, n = índice de refracción.

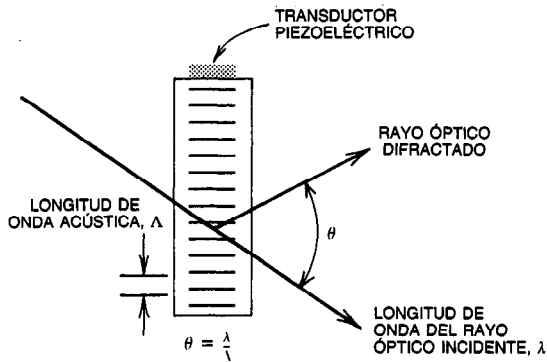


Fig. 11-6. Reflector acústico de Bragg.

de Raman-Nath (en un medio sólido). Para la trayectoria de la luz $l > \Lambda^2/\lambda$ (ondas sonoras de muy alta frecuencia), la luz incidente es fuertemente difractada en un punto (de primer orden). Esto ocurre para un ángulo específico de incidencia de la luz, y debido a su semejanza con la reflexión de Bragg de los rayos X por los planos paralelos de una red cristalina, se denomina reflexión acústica de Bragg. El ángulo de deflexión con respecto al haz de luz incidente (para pequeños ángulos) está dado por $\theta = \lambda/\Lambda$. Por tanto, el ángulo de deflexión puede controlarse con facilidad regulando la frecuencia acústica con un oscilador controlado por voltaje (VCO, de *voltage-controlled oscillator*).

11.2.6 Materiales magnetoópticos

El efecto magnetoóptico de mayor importancia es el efecto Faraday, que se define como la rotación ϕ_f de la dirección de polarización de una luz polarizada linealmente al pasar a través del material bajo el efecto de un campo magnético externo. En los materiales paramagnéticos o diamagnéticos está dado por:

$$\phi_f = VHI \quad (11-18)$$

donde H es la intensidad del campo magnético, l la longitud de la trayectoria paralela al campo magnético y V la constante de Verdet.

En los materiales ferromagnéticos el campo magnético se reemplaza por la densidad de flujo magnético (magnetización), y en este caso el coeficiente de rotación de Faraday (K) se denomina constante de Kundst.

En estos materiales una cantidad de utilidad es la rotación específica de Faraday F , que se obtiene para la magnetización de saturación M_s :

$$F = KM_s \quad (11-19)$$

Wemple¹⁸ ha realizado una revisión de materiales para moduladores magnetoópticos, y Chen¹⁹ ha elaborado una amplia tabla de constantes magnetoópticas. Los materiales magnetoópticos pueden emplearse en moduladores ópticos, en forma parecida a como se utilizan los materiales electroópticos^{20,21}. Sin embargo, en comparación con éstos, adolecen de elevada absorción de frecuencias ópticas.

11.2.7 Materiales ópticos no lineales

Los efectos electroóptico, elastoóptico y magnetoóptico descritos en las secciones previas se refieren a cambios en las propiedades ópticas debidos a la aplicación de campos de baja frecuencia. En esta sección se analizan los materiales ópticos no lineales. La distinción que se hace aquí es que los cambios en las propiedades ópticas son inducidos por los campos electromagnéticos a frecuencias ópticas. Esto da por resultado una multitud de fenómenos nuevos, tales como la generación del segundo armónico, osciladores paramétricos, absorción multifotónica, autoenfoco y dispersión estimulada de Raman, Brillouin y Rayleigh. La observación y aplicación en dispositivos de estos fenómenos se debe en gran medida al advenimiento de los láseres, fuentes de radiación electromagnética monocromática intensa. En muchos casos los materiales cristalinos que se emplean en estos dispositivos son los mismos que se emplean para el efecto electroóptico. Una diferencia importante es que a frecuencias ópticas ya no es posible excitar más el movimiento de la malla, y la polarización inducida a la suma de la frecuencia se debe sólo a las no linealidades del movimiento electrónico. Como en el caso del efecto Kerr, la polarización

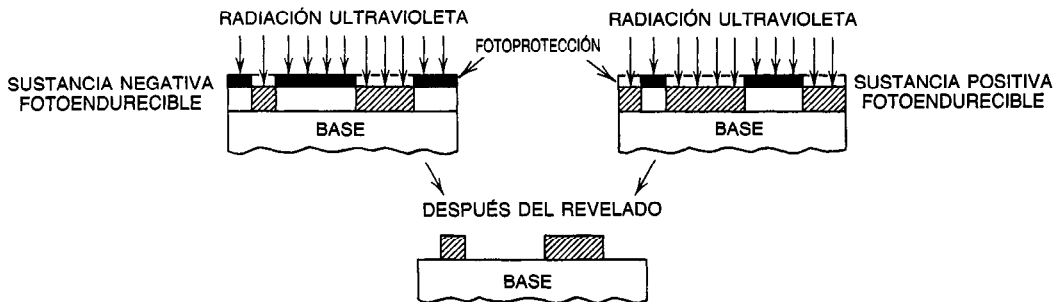
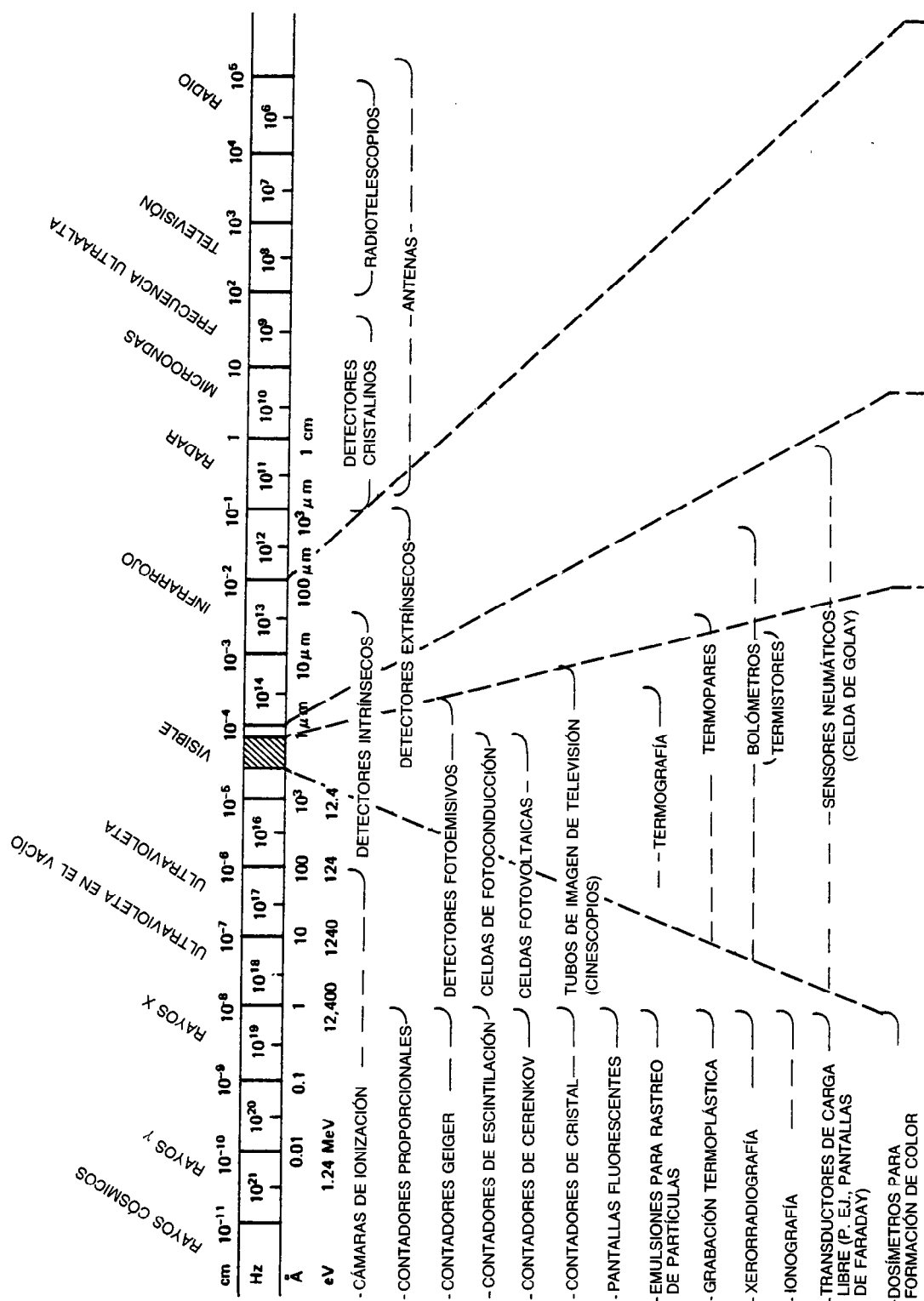


Fig. 11-7. Formación de imágenes con sustancias fotoendurecibles. Las regiones sombreadas son las expuestas a la radiación. Durante el revelado, ambas sustancias fotoendurecibles proporcionan regiones químicamente resistentes para el procesamiento ulterior.



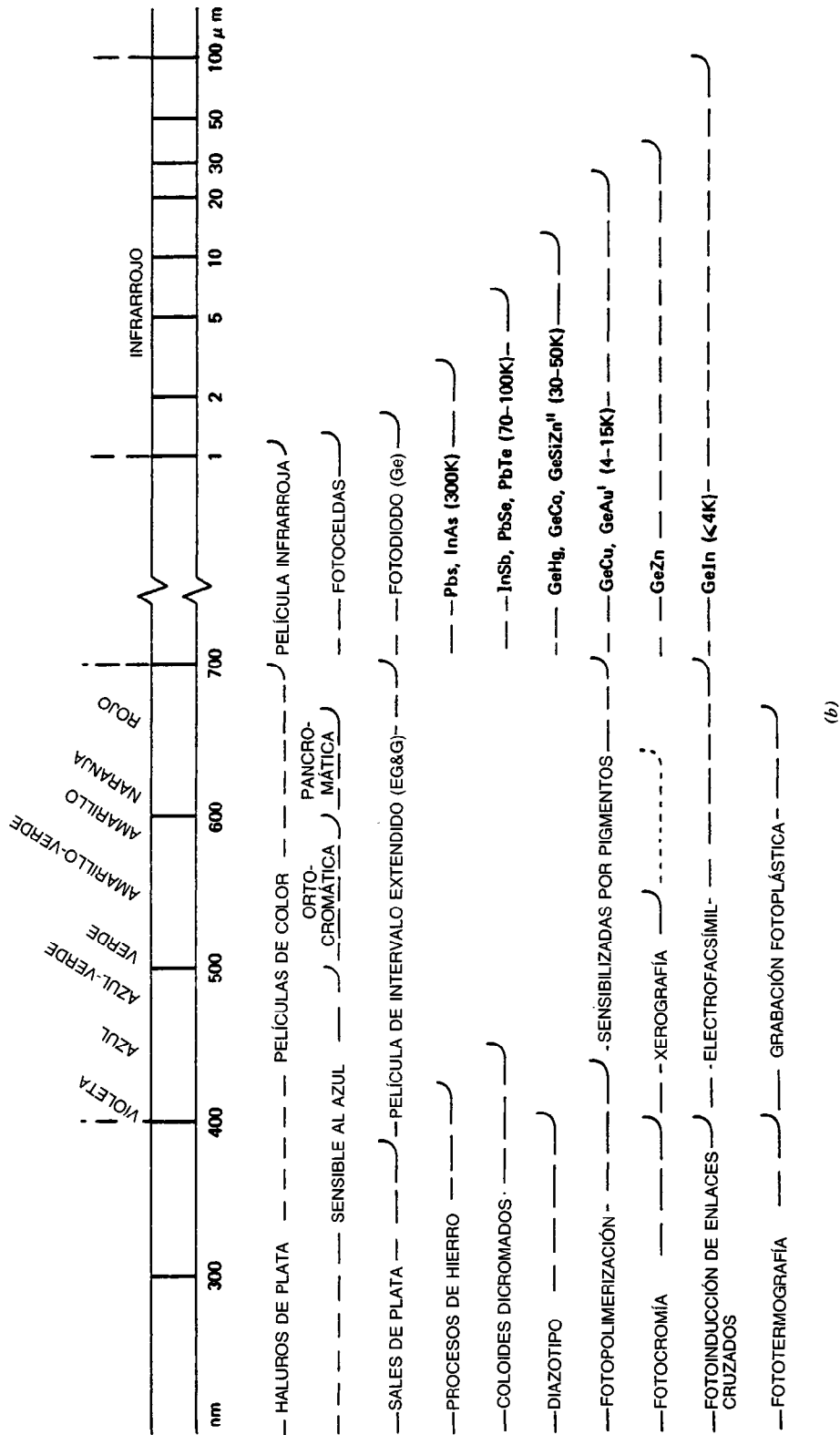


Fig. 11-8. Espectro electromagnético: a) sensibilidades espectrales de los detectores de radiaciones; b) sensibilidades espectrales de los detectores de radiaciones en la región del ultravioleta al infrarrojo.

eléctrica P debida a campos eléctricos intensos a frecuencias ópticas puede representarse en forma de desarrollo de una serie:

$$P = \epsilon_0(\chi_1 \cdot E + \chi_2 \cdot E^2 + \chi_3 \cdot E^3 + \dots) \quad (11-20)$$

donde χ_1 es la susceptibilidad óptica lineal y χ_2, χ_3 son las susceptibilidades ópticas no lineales del material. El término cuadrático en esta ecuación da por resultado fenómenos como el de la generación armónica. El término cúbico de la polarización produce fenómenos como la generación del tercer armónico y varios efectos de dispersión estimulada. En los textos de Yariv²² y Nelson²³ puede encontrarse una revisión de estos fenómenos, y Singh²⁴ presenta una amplia tabla de constantes de materiales y muchas referencias.

11.2.8 Materiales fotosensibles

Los materiales fotosensibles experimentan una reacción química al ser expuestos a la radiación óptica a la que son sensibles. La sensibilidad a la radiación cubre el intervalo que va del infrarrojo al ultravioleta. Los haluros de plata forman la clase más diversificada de materiales fotosensibles, con composiciones para un gran intervalo de sensibilidad espectral y temporal y contraste de imagen. Suelen ser suspensiones coloidales de sales de plata (bromuro, yoduro, cloruro) en una gelatina o un aglomerante sintético. La fotosensi-

bilidad resulta de la descomposición química de los haluros de plata en plata metálica y halógenos libres bajo la exposición a la radiación óptica. La absorción intrínseca en las películas de haluros de plata ocurre a longitudes de onda azul, violeta e inclusive más cortas. Sin embargo, puede lograrse la sensibilidad a la radiación a longitudes de onda mayores agregando colorantes idóneos a la emulsión. Las emulsiones de haluro de plata con grosor de unos micrones hasta varios milímetros se depositan sobre papel, vidrio u otros soportes flexibles, y como protección con frecuencia se emplea un sobrerrecubrimiento de alguna película gelatinosa transparente. El aglomerante de gelatina en la emulsión contiene sensibilizadores para ajustar la velocidad de la película, y también actúa como un medio ideal para la pronta penetración de las soluciones de procesamiento. Otra función del aglomerante de gelatina es actuar como aceptor del halógeno liberado por la exposición a la radiación. Esto es importante, ya que de otro modo el halógeno liberado podría recombinarse con la plata metálica. En la tabla 11-8 se enumeran algunas películas comunes.

Otra clase importante de materiales fotosensibles de amplio uso en la industria electrónica son los fotoendurecibles. Estos materiales forman películas químicamente resistentes y experimentan cambios profundos de solubilidad al ser expuestos a la luz (por lo general a la ultravioleta)²⁵. Los materiales fotoendurecibles pueden dividirse en tres clases:

Tabla 11-8. Películas fotográficas^a

Tipo de película	Sensibilidad espectral ^b (μm)	Resolución (líneas/mm)	Velocidad (ASA)	Contraste	Aplicación
Kodak					
Análisis de espectro núm. 1	250-440	>225	—	Alto	Espectroscopia de emisión
103-0	250-500	56-68	—	Medio	Espectroscopia de baja intensidad
649-F	—	2 000	—	Extra alto	Hologramas
Registro de alta velocidad					
2485	PR	20-50	800	—	Registro por tubo de rayos catódicos (CRT, de <i>cathode ray tube</i>) de alta velocidad
RAR-2495	O	32-100	320	—	Registro por CRT con sensibilidad ampliada al azul
LPF-7	O	—	—	Muy alto	Fotografadores, microfotografía con reducción intermedia
Polaroid					
612	P	15-18	20 000	Alto	Registro por CRT
47, 57, 107	P	22-28	3 000	Medio	Cámaras
552	P	20-25	400	Medio	Fotografías de microscopio electrónico de barrido
1462	—	40-50	200	Alto	Sensibilidad al azul

^a Estas películas son representativas. Para más detalles, consúltense los catálogos de los fabricantes.

^b P = pancromática; PR = pancromática con sensibilidad ampliada al rojo; O = ortocromática.

Tabla 11-9. Sustancias fotoendurecibles típicas

Tipo de sustancia fotoendurecible	Proveedor	Marca comercial
Líquido negativo	Kodak Hunt Dynachem Norland	Microneg Waycoat y SC CMR-5000, OME 83-DC NPR 6, NPR 22, NPR 29
Líquido positivo	Kodak Hunt Shipley	Kodak 809 Waycoat HPR Micropaset 1300
De película seca	Dupont Dynachem Hercules	Riston Laminar Akuamer

1. Endurecibles líquidos positivos.
2. Endurecibles líquidos negativos.
3. Endurecibles de película seca.

Los endurecibles negativos, al ser expuestos a la luz, experimentan la formación de enlaces cruzados que reducen su solubilidad en sus solventes anteriores. Algunas sustancias fotoendurecibles negativas que existen en el mercado están compuestas de derivados de cinamato de polivinilo o de caucho (hule) poliisopreno. Al ser expuestos a la luz, los endurecibles líquidos positivos experimentan descomposición fotoquímica y reacciones de reordenamiento que dan como resultado un aumento en su solubilidad en sistemas de solventes acuosos. Por ejemplo, algunos endurecibles contienen soluciones de disolventes orgánicos de cresol-formaldehído más diácidas de naftoquinona fotodescomponibles. Estas sustancias se hacen solubles en reveladores alcalinos acuosos después de exponerlas a la luz ultravioleta. En la figura 11-7 se muestran las etapas de procesamiento para la formación de imágenes con sustancias fotoendurecibles negativas y positivas.

Las sustancias fotoendurecibles líquidas pueden aplicarse por procesos de inmersión, laminado o centrifugación, en los que la resina se adhiere al sustrato por evaporación del disolvente. Los endurecibles de película seca son sustancias endurecibles negativas compuestas de una capa fotosensible preteñida (de 10 a 125 μm de grosor) colocada entre una película delgada transparente de mylar y un soporte de película de poliolefina, y poseen la ventaja de eliminar varias etapas de procesamiento en la aplicación de la película y proporcionar recubrimientos más uniformes, en comparación con las películas húmedas. En la tabla 11-9 se enumeran algunas sustancias fotoendurecibles que existen en el mercado.

Los materiales que responden a la radiación electromagnética incidente con cambios en sus propiedades eléctricas u ópticas se emplean como detectores de radiaciones, y también pueden clasificarse como materiales fotosensibles.

En la figura 11-8 se presenta un resumen sobre tales detectores, su intervalo espectral útil y sus posibles aplicaciones²⁶.

BIBLIOGRAFÍA

REFERENCIAS BIBLIOGRÁFICAS

1. J. F. Nye, *Physical Properties of Crystals*, Clarendon, Oxford, Inglaterra, pp. 236-240.
2. G. Haas y J. Ritter, *Vac. Sci. Tech.* **4**.
3. D. T. F. Marple, "Refractive Index of ZnSe, ZnTe and CdTe", *J. Appl. Phys.* **33**.
4. D. S. Heavens, "Optical Properties of Thin Solid Films", Butterworth, Londres.
5. Nye, *loc. cit.*, pp. 241-258.
6. A. R. Johnston, "Dispersion of Electro-Optic Effect in BaTiO₃", *J. Appl. Phys.* **42**:3501-3507.
7. J. E. Geusic, y otros, "Light Modulation and Beam Deflection with Potassium Tantalate-Niobate Crystals", *J. Appl. Phys.* **37**:388-398.
8. E. H. Turner, "High Frequency Electro-optic Coefficient of Lithium Niobate", *Appl. Phys. Lett.* **8**:303-304.
9. D. F. Melson y E. H. Turner, "Electro-optic and Piezoelectric Coefficients and Refractive Index of Gallium Phosphide", *J. Appl. Phys.* **39**:3337-3343.
10. L. Ho y C. F. Barber, "Electro-optic Effect of Gallium Arsenide", *Appl. Opt.* **2**:647-648.
11. I. P. Kaminow, "Optical Waveguide Modulator", *IEEE Trans. Microwave Theory Tech.* **MTT-23**.
12. M. Izutu, T. Itoh y T. Sueta, "10 GHz Bandwidth Travelling Wave LiNbO₃ Optical Waveguide Modulator", *IEEE J. Quantum Electr.* **QE-14**(6): 394-395.
13. R. W. Dixon, "Photoelastic Properties of Selected Materials and Their Relevance for Application to Acoustic Light Modulators and Scanners", *J. Appl. Phys.* **38**:5149-5153.
14. T. M. Smith y A. Korpel, "Measurement of Light-Sound Interaction Efficiencies in Solids", *IEEE J. Quant. Electr.* **QE-1**:283-284.
15. R. L. Abrams y D. A. Pinow, "The Acousto-optic Properties of Crystalline Germanium", *J. Appl. Phys.* **41**:2765-2768.
16. C. F. Quate, C. D. W. Wilkinson y D. K. Winslow, "Interaction of Light and Microwave Sound", *Proceed. IEEE* **53**(10):1604-1623.

17. E. I. Gordon, "A Review of Acousto-optical Deflection and Modulation Devices", *Appl. Opt.* **5**(10):1629-1639.
18. S. H. Wemple, "Materials for Magneto-optic Modulators", *J. Electr. Mater.* **3**(1):243-263.
19. D. Chen, "Magneto-optic Materials", *Handbook of Lasers with Selected Data on Optical Technology*, CRC Press, Boca Raton, FL, pp. 460-477.
20. F. S. Chen, "Modulators for Optical Communications", *Proceed. IEEE* **58**(10):1440-1457.
21. P. D. Tien y cols., "Switching and Modulation of Light in Magneto-optic Waveguides of Garnet Films", *Appl. Phys. Lett.* **21**:394-396.
22. A. Yariv, *Quantum Electronics*, Wiley, Nueva York.
23. D. F. Nelson, *Electric, Optic and Acoustic Interactions in Dielectrics*, Wiley, Inglaterra.
24. S. Singh, "Non-Linear Optical Materials", en *Handbook of Lasers with Selected Data on Optical Technology*, Chemical Rubber Co., Cleveland, Ohio.
25. J. Pacansley, "Recent Advances in Photodecomposition Mechanisms of Diazo-Oxides", *Poly. Eng. Sci.* **20**(16).
26. Thomas Woodlief, Jr., ed., *SPSE Handbook of Photographic Science and Engineering*, Wiley, Nueva York.

LIBROS Y REVISTAS

- F. T. Arecchi y E. O. Schulz-DuBois, eds., *Laser Handbook*, vol. I, North-Holland, Amsterdam.
- M. Born y E. Wolf, *Principles of Optics*, 4.^a ed., Pergamon, Nueva York.
- G. W. DeBell y D. H. Harrison, eds., *Optical Coatings*, *Proceed. SPIE*, vol. 50.
- W. S. DeForest, *Photoresist Materials and Processes*, McGraw-Hill, Nueva York.
- W. J. Driscoll y W. Vaughan, eds., *Handbook of Optics*, McGraw-Hill, Nueva York.
- T. Kallard, ed., *Acoustic Surface Wave and Acousto-optic Devices*, Optosonic, Nueva York.
- *Kodak Microelectronics Seminar Proceedings*, Kodak publication No. G-130, Eastman Kodak Co., Rochester, NY.
- M. E. Lines y A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials*, Clarendon, Oxford.
- *Photofabrication Methods with Kodak Resists*, Eastman Kodak Co., Rochester, NY.
- R. F. Potter, ed., *Physical Properties of Optical Materials*, *Proceed. SPIE*, vol. 204.
- Thomas Woodlief, Jr., *SPSE Handbook of Photographic Science and Engineering*, Wiley, Nueva York.

TERCERA PARTE

Componentes

CAPÍTULO 12

Elementos circuitales concentrados pasivos

Henry Domingos
Joseph Scaturro

Clarkson University
Potsdam, Nueva York

12.1 Componentes resistivos

- 12.1.1 Resistores discretos
- 12.1.2 Resistores de circuitos microelectrónicos
- 12.1.3 Termistores
- 12.1.4 Varistores

12.2 Capacitores

- 12.2.1 Introducción
- 12.2.2 Definiciones y propiedades generales de los circuitos
- 12.2.3 Capacitores electrolíticos de aluminio
- 12.2.4 Capacitores electrolíticos de tantalio
- 12.2.5 Capacitores de cerámica
- 12.2.6 Capacitores de papel y plástico
- 12.2.7 Capacitores de mica y vidrio
- 12.2.8 Selección de capacitores
- 12.2.9 Capacitores para circuitos integrados

12.3 Inductores

- 12.3.1 Introducción
- 12.3.2 Inductancia de configuraciones sencillas
- 12.3.3 Bobinas con núcleo de aire
- 12.3.4 Materiales magnéticos
- 12.3.5 Inductores con núcleo de hierro
- 12.3.6 Inductores para aplicaciones en circuitos integrados
- 12.3.7 Evaluación de la aplicación de un inductor

12.4 Transformadores

- 12.4.1 Circuitos acoplados
 - 12.4.2 El transformador ideal
 - 12.4.3 Transformadores sin pérdidas
 - 12.4.4 Transformadores electrónicos
 - 12.4.5 Pérdidas en los transformadores
 - 12.4.6 Efecto corona
 - 12.4.7 Modos y mecanismos de fallo en transformadores e inductores
-

12.1 COMPONENTES RESISTIVOS

Henry Domingos

12.1.1 Resistores discretos¹⁻⁴

Introducción

Los componentes electrónicos de mayor uso son los resistores, cuyo volumen de ventas va de la mano con las fortunas generadas en la industria de los circuitos integrados. El ingeniero de diseño que se enfrenta con la elección de un resistor debe considerar muchos factores: precio, disponibilidad, tolerancia, disipación de potencia, estabilidad, confiabilidad, respuesta de frecuencia, coeficiente térmico, coeficiente de voltaje, tamaño y empaque, por mencionar algunos. Pero además de estos factores es necesario considerar los materiales y la construcción de los diversos tipos.

La mayor parte de los resistores discretos quedan comprendidos en algunas de las siguientes categorías:

Resistores de composición de carbono. Se han utilizado durante casi cien años y aún son populares. Parte de su popularidad debe atribuirse a la inercia de su amplio uso con los años y a la resistencia a los cambios en el diseño. Sin embargo, como resistores de uso general, los de este tipo poseen ciertas ventajas peculiares; a saber, bajo precio, amplio intervalo de resistencia, baja inductancia, excelente capacidad para soportar oscilaciones de carga, buen rendimiento en ciclos de temperatura, buena confiabilidad y fácil disponibilidad. Por otra parte, no existe en el mercado en tolerancias inferiores al 5%, tiene poca estabilidad a largo plazo, es sensible a los niveles ambientales de humedad y tiene valores de ruido elevados (a altos valores de resistencia). Durante los últimos años sus ventas han disminuido gradualmente, a medida que los precios de otros resistores de elevado rendimiento se han vuelto más competitivos. Allen-Bradley Co. produce resistores de carbono, aunque otras compañías venden productos importados.

Resistores de película de carbono. Han sustituido a los de composición de carbono para uso general, debido a que son menos costosos, tienen mayor estabilidad, tolerancia y rendimiento a alta frecuencia, y presentan menor ruido. Se dispone de piezas para aplicaciones electrónicas con potencia nominal de 1/10 a 2 W, con tolerancias hasta del 0.5% y cambios de duración en carga y resistencia a la humedad del 2% o menos. La característica distintiva de los resistores de película de carbono es su coeficiente de resistencia negativo, que suele ser de $-200 \text{ ppm}/^{\circ}\text{C}$ para bajos valores de resistencia hasta $-1\,000 \text{ ppm}/^{\circ}\text{C}$ para tamaños más grandes. Su volumen de ventas en la actualidad es de varios miles de millones de unidades al año, y entre los fabricantes más importantes se encuentran Dale Electronics, R-Ohm, Mepco/Electra y Airco Electronics.

Resistores de película metálica. Pueden constar, en efecto, de una delgada capa de metal sobre un núcleo aislante, o bien pueden ser un óxido o algún otro

compuesto metálico, e incluso un vidrioado metálico. Fabricantes como Corning Glass, Dale Electronics, Mepco/Electra y TRW ofrecen una gran variedad de estilos, empaques y especificaciones para muchas aplicaciones. Estos resistores se fabrican con potencias nominales que llegan a ser tan bajas como 1/20 W, y suelen ser más pequeños que otros resistores equivalentes. Aunque su intervalo de resistencia es un poco inferior que el de los resistores de composición de carbono, su tolerancia es mucho mejor —por lo general de 0,1 a 1%—, pero existen en el mercado algunos productos cuya tolerancia es hasta de 0.025%. Además, son superiores en cuanto a estabilidad, duración en almacenamiento y coeficiente térmico. Por último, su rendimiento a alta frecuencia supera el de otros tipos de resistores.

Debido al amplio intervalo de especificaciones que existe para ellos, los resistores de película metálica se están introduciendo en los mercados tradicionales de los económicos resistores de composición de carbono y de película de carbono, así como en el de los más costosos resistores de alambre arrollado. En la actualidad, su volumen de ventas es el más importante y se observa una tendencia a la baja en sus precios, a medida que su flexibilidad acrecienta la base de su mercado.

Resistores de alambre arrollado se clasifican en tres categorías:

1. De uso general, de bajo costo.
2. De potencia de hilo arrollado.
3. De precisión de hilo arrollado.

Estas dos últimas clases de resistores se emplean siempre que es posible tolerar su elevado costo, gran tamaño y deficiente respuesta de frecuencia, en aras de sus sobresalientes exactitud, estabilidad, factor de ruido, coeficiente térmico y coeficiente de voltaje. Los resistores de alambre arrollado de precisión tienen envoltentes (envases) de gran tamaño envolvente para mantener bajo el aumento de la temperatura interna, lo que a su vez minimiza los cambios en la resistencia. Aun cuando es posible emplear configuraciones de baja inductancia, como los arrollamientos de Ayrton-Perry, la inductancia y la capacitancia distribuida suelen limitar su utilización al intervalo de las audiofrecuencias. Todos los resistores de alambre arrollado están limitados en cuanto a valores de alta resistencia por el pequeño diámetro del alambre y las grandes longitudes que podrían requerirse. Excepto para los tipos de uso general, la tolerancia en el momento de la adquisición suele variar del 0.01 al 1%, aunque se anuncian valores hasta del 0.002%. Algunos de los fabricantes más importantes en Estados Unidos, son TRW, Dale Electronics, RCL y Ohmite.

Fabricación

En los **resistores de composición de carbono** se utiliza grafito o negro de humo calcinado, dispersados en un sistema de resina, como fenol-formaldehído relleno con sílice microcristalina. La resistividad volumétrica del núcleo conductor es controlada por el tamaño y la

densidad de las partículas de carbono. La cubierta aislante está hecha del mismo material sin carbono, lo que da por resultado un montaje robusto con problemas mínimos de desajustes térmicos. En el montaje, la coraza se forma alrededor de un conductor de cobre estañado. Para incrementar la resistencia mecánica del conductor y reducir la resistencia eléctrica de contacto, las puntas de conexión pueden moletearse, estamparse o conformarse en un ahusador invertido o en una cabeza de clavo doble. Se inserta el material del núcleo, se alinea el segundo conductor y se añade la cubierta final. Después se comprime y cura todo el resistor, se impregna al vacío con compuestos aromáticos y, por último, se prueba, clasifica y etiqueta la pieza.

Los **resistores de película de carbono** se fabrican mediante la descomposición pirolítica del carbono sobre la superficie de un sustrato cerámico. La fuente de carbono es un gas orgánico tal como el metano, desintegrado a una temperatura de 1 100 °C. Las características del resistor son sensibles a las condiciones de depósito. El grosor de la película suele variar desde menos de 2 µm hasta aproximadamente 100 µm. Las películas más delgadas tienen resistividades laminares de 10 000 Ω/cuad y presentan el coeficiente térmico más elevado, aproximadamente de -1 000 ppm/°C. Las películas más gruesas tienen resistividades laminares hasta de 10 Ω/cuad y coeficientes térmicos de -100 ppm/°C.

Se hace contacto con cada extremo de la película por medio de cubiertas finales de ajuste, forzado con una capa de cemento de plata, a fin de obtener mejor contacto. En esta etapa se ajustan los resistores al tamaño idóneo mediante un procedimiento conocido como espiralización. Se emplea una muela delgada para cortar una ranura en la película a lo largo de una trayectoria helicoidal, incrementando el valor de la resistencia mediante el cambio de la configuración de la película transportadora de corriente. El corte en espiral puede ocupar desde una vuelta, o más, hasta una docena. La resistencia se mide de forma continua y el proceso se detiene automáticamente cuando se llega al valor deseado. El resistor terminado se recubre con resina de termofraguado, resina epóxica moldeada, recubrimiento cerámico o recubrimiento de conformación de diversos materiales.

Cuando se espiraliza cualquier tipo de resistor de película, el flujo de corriente y la distribución de la potencia dejan de ser uniformes sobre la película. La densidad de corriente es máxima en la punta del corte en espiral, lo cual produce un punto caliente en ese sitio y posibles cambios permanentes en el valor de la resistencia del resistor. Para valores de resistencia elevados, la disipación de potencia es demasiado baja para causar daño térmico, pero el voltaje máximo puede ser limitado por disrupción a través de un corte en espiral. Ambos tipos de daño se han observado durante la operación de pulsos de alta potencia.

En los **resistores de película metálica** se emplean varios materiales y métodos de manufactura. Se utilizan con frecuencia películas evaporadas de aleaciones de níquel-cromo y la resistencia laminar y el coeficiente térmico son determinados por los procesos de depósi-

to, en los que variables como el grosor pelicular, velocidad de depósito, temperatura del sustrato, recocido y tratamientos de oxidación son las más importantes. El material inicial que suele utilizarse es una aleación de níquel al 80%, pero la película depositada es más rica en cromo. Con grosores peliculares de 5 a 100 nm y la inclusión de otros materiales en la película, es posible obtener resistividades laminares que varían de 10 a 10⁴ Ω/cuad. Sin embargo, la mayor parte de las películas presentan resistividades que varían de 100 a 500 Ω/cuad. Los coeficientes térmicos también dependen de las condiciones de deposición, y es fácil obtener valores tan bajos como 25 ppm/°C.

Para el elemento resistivo suelen emplearse mezclas de metales, compuestos metálicos, vidrio y solventes. En este caso, el método usual de aplicación es por inmersión o laminado. Las películas tienden a ser relativamente gruesas, de 5 a 50 µm. Una secuencia común podría ser empezar con un sustrato de alúmina bañado en un vidrioado líquido, compuesto de tantalio, nitrato de tantalio, vidrio y un portador. El sustrato recubierto se somete a calentamiento en horno (vitriificado) a 1 100 °C para fundir las partículas de vidrio y ligar la película con el sustrato. Luego se cortan a tamaño las barras y los extremos se chapean con níquel y se estañan con soldadura fuerte. Los resistores en bruto se espiralizan, y luego se sujetan a los conjuntos de cubiertas de los extremos con los conductores. El ensamblaje se completa con una envoltura, por ejemplo, de alguna sustancia fenólica moldeada.

El alambre para los **resistores de alambre arrollado** se formula especialmente para presentar alta estabilidad y bajo coeficiente térmico, y los tipos más comunes son aleaciones de níquel-cromo (con nombres comerciales como Evanohm, Karma y Moleculoy), con resistividad de 1.33×10^{-4} Ω-cm (800 Ω por mpulg circular-pie), y aleaciones de cobre-níquel (como Cupron, Advance y Neutraloy), con resistividades cercanas a 5×10^{-5} Ω-cm (300 Ω por mpulg circular-pie). Los diámetros del alambre varían de 7 µm a 10 mm. También existen en el mercado barras y cintas con valores de resistencia muy bajos. El alambre se puede obtener recubierto o no con diversos esmaltes, plásticos o tejidos en una o varias capas.

La construcción de arrollados de alambre de potencia y precisión ofrece un contraste interesante. Se pretende que los arrollamientos de potencia toleren temperaturas superiores de operación, por lo cual los materiales y los ensamblajes son diferentes. La construcción empieza con un núcleo cerámico, como esteatita, alúmina o berilia. Las cubiertas de los extremos y los ensamblajes de conductores se sujetan y luego se suelda un alambre de resistencia, descubierto por un extremo. El alambre se devana en una sola capa con bastante separación entre las vueltas. La resistencia se mide de manera continua durante el proceso de devanado, y, cuando se llega al valor deseado, el alambre se corta y suelda a la otra cubierta extrema, y se aplica un recubrimiento de silicón, cerámica o barniz vítreo como aislante. Todos los materiales son capaces de operar a temperatura elevada, y el tipo de devanado facilita la transferencia de calor al ambiente.

Tabla 12-1. Códigos de color de los resistores

Color	1ª banda ^a , primera cifra significativa	2ª banda, segunda cifra significativa	3ª banda, factor de multiplicación	4ª banda ^b , tolerancia (%)	5ª banda ^b , proporción de fallos (%/1 000 h)
Negro	0	0	1	—	—
Café	1	1	10	—	1
Rojo	2	2	10 ²	—	0.1
Naranja	3	3	10 ³	—	0.01
Amarillo	4	4	10 ⁴	—	0.001
Verde	5	5	10 ⁵	—	—
Azul	6	6	10 ⁶	—	—
Violeta	7	7	10 ⁷	—	—
Gris	8	8	10 ⁸	—	—
Blanco	9	9	10 ⁹	—	—
Plateado	—	—	0.01	10	—
Dorado	—	—	0.1	5	—
Ninguno	—	—	—	20	—

^a La primera banda es la más cercana a un extremo del resistor. Una primera banda más ancha que las demás indica resistor de alambre arrollado.

^b Ciertas piezas para uso militar.

Por otra parte, los arrollados de alambre de precisión se diseñan para presentar una resistencia exacta estable. Tal resistencia se logra mejor por medio de grandes longitudes de alambre y oscilaciones de temperatura restringidas. Esto implica un mayor volumen y el empleo de materiales diferentes y menos costosos para el empaque. El alambre recubierto con barniz se corta a tamaño, se arrolla en un carrete de plástico dividido en varios segmentos y se suelda a los conductores. Dado que las variaciones en la resistencia del devanado inicial del alambre pueden ser tan elevadas como del 1% y que el aislamiento impide la medición continua de la resistencia, ésta se ajusta a un valor final mediante abrasión por chorro de arena. Algunas veces se fija el devanado con resinas epóxicas o con caucho (hule) RTV (vulcanizado a temperatura ambiente, de *room temperature vulcanized*), y el montaje se inserta en una envoltura epóxica, que se cierra en un extremo que tiene un agujero para el conductor. Ambos extremos se sellan. Debido a que la operación de devanado induce esfuerzos de flexión y de tensión que modifican la resistencia, las piezas se estabilizan térmicamente por horneado durante 24 a 96 horas a 160 °C.

Características

Algunas de las características de los resistores de interés para los usuarios son: exactitud (o tolerancia), estabilidad, coeficiente térmico, coeficiente de voltaje, efectos de la humedad, disipación de potencia, efectos de las frecuencias y confiabilidad. En artículos posteriores se analizan las tres últimas.

Tolerancia. Los resistores de uso general existen en el mercado con valores ordinarios de tolerancia del 5, 10 y 20%. En los resistores de composición de carbono y en algunos de alambre arrollado se emplea un código de color de tres a cinco bandas. En la tabla 12-1 se muestra el código de color para resistores.

Dado que en una etapa final se clasifican los resistores, la distribución de valores puede ser bimodal, en

vez de normal. Las tolerancias iniciales al adquirir un resistor son del 5 al 20% para los de composición de carbono y típicamente del 0.5 al 10% para los de película de carbono, del 0.1 al 1% para los de película metálica y del 0.01 al 1% para los de alambre arrollado de precisión.

Estabilidad. La estabilidad se refiere al cambio en el valor de la resistencia después de la exposición a un esfuerzo ambiental específico. Entre los esfuerzos pueden considerarse: alta o baja temperatura durante el almacenamiento, larga duración (vida) en almacenaje, aplicación de potencia nominal completa, humedad, calor de soldadura, sobrecarga a corto plazo y exposición a radiación. Los resistores de composición de carbono suelen presentar la estabilidad más deficiente, y esfuerzos relativamente leves durante cierto tiempo pueden provocar el desplazamiento del valor de la resistencia bastante lejos de la tolerancia de adquisición; no es raro encontrar en inventario piezas con tales desplazamientos de resistencia. Los resistores de alambre arrollado son los más estables, seguidos por los de película metálica y los de película de carbono.

Coeficiente térmico. El coeficiente térmico o coeficiente de temperatura se define como $(1/R)(\partial R/\partial T)$, expresado en partes por millón por grado Celsius (ppm / °C). Los resistores de composición de carbono presentan el coeficiente térmico más alto.

Este coeficiente es mayor para resistores de mayor valor, y a temperaturas elevadas puede ser significativo. Puede variar de +1 600 a -800 ppm / °C sobre el intervalo de operación a temperatura ambiente de -55 a 130 o 150 °C, por lo que es posible que un resistor del 5 o 10% se encuentre fuera de la especificación a temperaturas extremas en condiciones de no carga. Para los resistores de película de carbono, el coeficiente térmico suele ser negativo, con variación de -200 ppm / °C para bajos valores de resistencia, hasta más de -1 000 ppm / °C para valores altos. Los resistores

de alambre arrollado poseen el mejor coeficiente térmico, con valores hasta de ± 20 ppm / $^{\circ}\text{C}$, y los de película metálica son casi tan buenos, con valores que varían de ± 20 a ± 200 ppm / $^{\circ}\text{C}$.

Coeficiente de voltaje. Cuando se aplica voltaje a un resistor puede presentarse una ligera disminución en la resistencia (además de los cambios inducidos por la temperatura). El cambio porcentual en resistencia por voltio aplicado se conoce como coeficiente de voltaje. Para los resistores de composición de carbono el cambio es elevado, de casi $-0.05\%/V$, y para los resistores de película de carbono suele ser inferior en un orden de magnitud. Para los restantes tipos de resistores el coeficiente de voltaje es despreciable.

Efectos de la humedad. La humedad puede provocar dos efectos reversibles. Sobre la superficie de un resistor de alto valor puede constituir una vía para la fuga de corrientes, reduciendo de esta forma la resistencia aparente. Si se absorbe a través de la envoltura de un resistor de composición de carbono, el valor de la resistencia se incrementará hasta un 10%. En cualquier caso es posible eliminar la humedad por medio de hornado y su efecto se minimiza a medida que el resistor disipa potencia. Los demás tipos de resistores son susceptibles a las reacciones químicas si se permite que la humedad penetre en el elemento resistivo.

Disipación de potencia. Para aplicaciones electrónicas existen en el mercado resistores discretos cuya disipación de potencia varía de $1/20$ a 2 W , o más. Sin embargo, para altos valores de resistencia, el factor limitante es el voltaje, que varía desde 150 V para los empaques más pequeños hasta aproximadamente 750 V para los tamaños de 2-W . Para bajos valores de resistencia, la aplicación de estos voltajes superará la disipación de potencia nominal. El valor de la resistencia que separa estos dos regímenes se denomina resistencia crítica, y está dado por $R = V^2/P$, donde V y P son el voltaje y la disipación nominales, respectivamente.

La consideración más importante en la disipación de potencia es el incremento de temperatura. Por ejemplo, puede permitirse que un resistor de composición de carbono de 1-W disipe 1 W sólo si está montado de modo que el aire pueda circular a su alrededor y la temperatura ambiente sea menor de 70°C . En estas condiciones, más de la mitad de la potencia se elimina a través de los conductores y la mayor parte del resto se elimina por convección. Si el resistor está montado de modo que se encuentre cerca de otros componentes productores de calor, o no tiene suficiente ventilación, es necesario reducir la disipación de potencia para no alcanzar la máxima temperatura permisible de 150°C . Es obvio que para temperaturas ambiente superiores a 70°C la disipación debe reducirse linealmente a la capacidad nominal para 150°C .

En la operación por pulsos la situación es más compleja⁵. Cuando se aplica un solo pulso, la limitación también suele ser la temperatura máxima (o pico), inclusive cuando el voltaje supere momentáneamente el valor nominal. Inmediatamente después de la apli-

cación de potencia, el incremento de temperatura es limitado por la masa térmica del elemento resistivo. El incremento de temperatura está dado por:

$$T = \frac{1}{\rho c V} \int_0^t P(t) dt$$

donde $P(t)$ = potencia aplicada, W

ρ = densidad, kg/cm^3

V = volumen, cm^3

c = calor específico, $\text{J}/\text{kg } ^{\circ}\text{C}$

Si P es constante, entonces la temperatura es directamente proporcional al tiempo. Con el tiempo, el calor generado se difunde hacia la envoltura, el sustrato y las puntas, y el incremento en la temperatura es proporcional a \sqrt{P} para una potencia aplicada constante. El tiempo en que el incremento de temperatura cambia de esencialmente adiabático a esencialmente por difusión puede estimarse aproximadamente a partir de las dimensiones del elemento resistivo. El cambio ocurre cuando las dimensiones se vuelven comparables con la longitud de difusión de calor del material resistivo, dada por $L = \sqrt{Dt}$, donde D es el coeficiente de difusión (cm^2/s) y t el tiempo en segundos. Para los resistores de composición de carbono con disipación nominal de $1/8\text{ W}$, esta transición ocurre aproximadamente en 1 s , mientras que para los de película metálica puede suceder en un tiempo tan corto como 100 ns .

Dado que el volumen del núcleo resistivo de los resistores de composición de carbono es muy grande, estos resistores pueden absorber grandes cantidades de energía de pulsación antes de que el incremento de temperatura alcance el límite relativamente moderado de 150°C para los materiales plásticos. En la tabla 12-2 se muestran las energías de pulsación para los resistores de composición de carbono que ofrece la Allen-Bradley Co.

La capacidad de absorción de energía de pulsación depende no sólo de la masa térmica, sino también de la temperatura máxima que es posible tolerar antes de que ocurra daño permanente. La temperatura máxima más baja corresponde a los materiales plásticos, y la más alta a las cerámicas. En términos generales, los resistores de composición de carbono presentan la mejor capacidad de manejo de los pulsos (debido a que la masa térmica es bastante grande), seguidos por los de alambre arrollado de precisión, los de película de carbono y los de película metálica.

Tabla 12-2. Capacidades de manejo de pulsos para resistores compuestos de carbono

Disipación nominal (W)	Energía del pulso (J)	Fuente equivalente
1/8	0.45	2 μF @ 670 V
1/4	1.8	10 μF @ 600 V
1/2	6.4	32 μF @ 630 V
1	16	32 μF @ 1 000 V
2	44	32 μF @ 1 650 V

Fuente: Allen-Bradley Co.

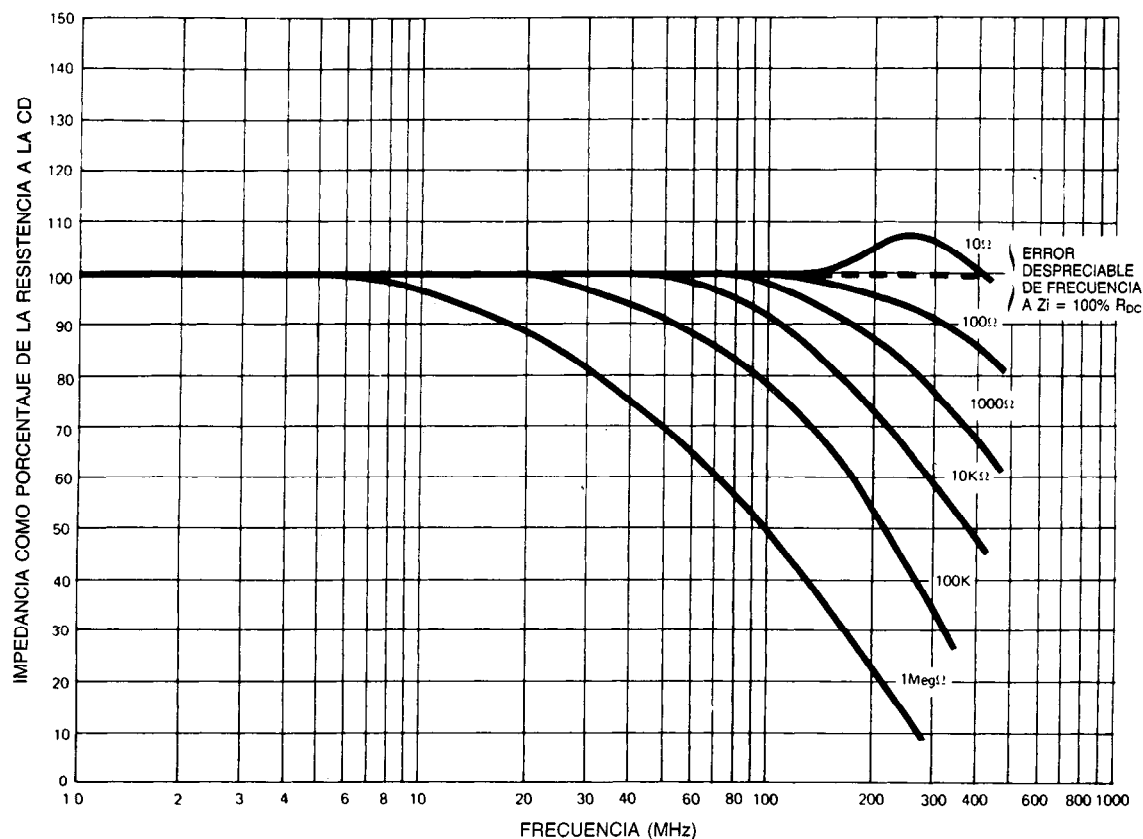


Fig. 12-1. Variación de la magnitud de impedancia total con la frecuencia como porcentaje de la resistencia a la CD para resistores de película. (Cortesía de Mepco/Electra, Inc.)

Cuando la potencia se aplica durante varios minutos se alcanzan las condiciones de estado estable y el aumento de temperatura alcanza un valor constante dado por $T = P/\theta$, donde θ es la resistencia térmica en $W/^{\circ}C$ entre el elemento resistivo y la temperatura ambiente. Es posible obtener θ de manera aproximada a partir de la pendiente de la curva de reducción de la capacidad de absorción nominal. Por ejemplo, para un resistor de composición de carbono de 1-W, $\theta = 1W/(150^{\circ}C - 70^{\circ}C) = 12.5 mW/^{\circ}C$. Con un tren de pulsos o forma de onda periódica, la potencia promedio se emplea a menudo como potencia de estado estable, aunque no es válido a bajos regímenes de trabajo, en los que el alejamiento de temperatura pico (máxima) es mucho mayor que la temperatura promedio.

Efectos de las frecuencias. La resistencia permanece en un valor constante sólo a bajas frecuencias. La respuesta a las frecuencias es afectada por la capacitancia e inductancia distribuidas en la trayectoria resistiva, inductancia de los conductores, capacitancia de la envoltura a tierra, efecto peculiar y pérdidas dieléctricas. Un circuito equivalente exacto no es práctico, por lo que es necesario recurrir a modelos aproximados, cuya

complejidad depende de la exactitud deseada. Por desgracia, los fabricantes rara vez proporcionan parámetros, inclusive para los modelos más sencillos. En vez de ello, el comportamiento a frecuencias típicas está dado por gráficas de impedancia, como la que se observa en la figura 12-1.

Para un tipo dado de resistor, la respuesta a las frecuencias tiende a mejorar a medida que disminuye el valor de la resistencia y se reduce el tamaño de la envoltura. El mejor comportamiento respecto a las frecuencias lo presentan los resistores de película con espiralización mínima. Entre todos los resistores, los de alambre arrollado tienen la peor respuesta a las frecuencias, ya que son inductivos a frecuencias de audio (audiofrecuencias, AF) y capacitivos a frecuencias de radio (radiofrecuencias, RF).

Fiabilidad⁶. La fiabilidad se refiere a la probabilidad de que un resistor se encuentre todavía en el intervalo de cierta especificidad después de un tiempo dado en ciertas condiciones de operación. La fiabilidad no es equivalente a la calidad, que se relaciona con el porcentaje de piezas defectuosas de un envío, aunque es posible que exista alguna correlación. La fiabilidad se expresa en términos de una razón porcentual de falla

por cada mil horas de operación, o en FITS (fallas por cada 10^9 piezas-hora de operación, del inglés *failures per 10⁹ part-hours of operation*). La velocidad de falla para cualquier resistor aumenta con la temperatura ambiente y con la disipación de potencia.

Los resistores pueden fallar catastróficamente (circuitos abiertos o cortocircuitos) o derivar saliéndose fuera de las especificaciones. Los circuitos abiertos pueden deberse a sustratos agrietados, soldaduras abiertas, elementos resistivos rotos o corrosión química. Algunas causas de los cortocircuitos son disrupciones eléctricas, objetos o materiales extraños y emigración de plata. La deriva excesiva se debe a penetración de humedad, corrosión, recocido y estabilización insuficientes, descarga electrostática, otros efectos transitorios y uso indebido.

En el *Military Handbook 217C*⁶ se presentan frecuencias de falla básicas para diferentes tipos de resistores en función de la temperatura ambiente y la disipación de potencia. Los factores multiplicativos se basan en la condición medioambiental de servicio, factor de calidad al adquirir la pieza e intervalo de resistencia. Como medio de comparación, en la tabla 12-3 se muestran las frecuencias de falla para diferentes tipos de resistores con bajo valor de disipación nominal a 25 °C en un ambiente propicio a tierra, calificados a un nivel de velocidad de falla M (1% por 1 000 horas al 60% de confianza).

Redes de resistores y resistores de pastilla (chip)

El empleo de redes de resistores y resistores de pastilla es una forma de reducir los costos a través de un abatimiento del número de componentes y satisfacer la necesidad de menor espacio en los tableros de circuitos impresos, aunque en algunas aplicaciones una ventaja importante es una mejora en el rendimiento. Las redes de resistores pueden ser componentes discretos en un solo paquete, arreglos de redes de película gruesa (aplicados por serigrafía) o redes de película delgada (aplicadas por evaporación al vacío). La mayor parte de los productos se encuentran en una cajetilla de doble línea (DIP, del inglés *dual inline package*) o en una de línea única (SIP, del inglés *single inline package*); también se dispone de componentes planos y blindajes de bote. Los resistores de pastilla tienen terminales de una sola superficie de papel metálico arrollado que permiten unirlos con resinas epóxicas o reflujo de soldadura.

Tabla 12-3. Proporción de fallos de resistores^a

Tipo	Especificación militar	Proporción de fallos (FITS) ^b
Composición fija	MIL-R-39008	0.94
Película fija (aislada)	MIL-R-39017	1.9
Película fija	MIL-R-55182	2.0
De alambre arrollado fijo (exacto)	MIL-R-39005	10

^a Los resistores de bajo valor, con potencia nominal de disipación a 25 °C en un entorno terrestre benigno, se clasifican en el nivel M de proporción de fallas.

^b Fallas por 10^9 partes-hora de operación.

Las redes de resistencia del tipo de película gruesa o de cerámicas metálicas suelen tener tolerancias del 2% y coeficientes térmicos de 100 ppm / °C. Las redes de película delgada tienen tolerancias tan bajas como 0.1% y coeficientes térmicos de resistencia de 25 ppm / °C. Algunas características sobresalientes de ambos tipos de redes son la capacidad de presentar relaciones de resistencia considerablemente más estrechas que las tolerancias absolutas y seguimiento de temperatura hasta ± 5 ppm / °C (± 25 ppm / °C para películas gruesas). Estas ventajas se aprovechan en ciertas aplicaciones, como en las redes celulares para convertidores de digital a analógico. Los resistores de pastilla varían en valores de 5Ω a $5\text{ M}\Omega$, con tolerancias del 1% y coeficientes térmicos de resistencia de 100 ppm / °C.

Las redes pueden adquirirse con resistores aislados por separado o con arreglos especiales de resistores. Algunas aplicaciones importantes son divisores de voltaje, terminaciones de ascenso/descenso de redes celulares, traductores de la familia de lógica digital, redes de opamps, interfaces y atenuadores.

Resumen

En la tabla 12-4 se muestran las características más importantes de cada familia de resistores. Se proporcionan valores típicos, no los disponibles para condiciones especiales o a costo extra. Obsérvese que es posible que los valores extremos no puedan alcanzarse simultáneamente; por ejemplo, en el mercado sólo existen resistores de alambre arrollado de precisión hasta $10\text{ M}\Omega$ en ciertos estilos, no suele disponerse de tolerancia de 0.01% para bajos valores de resistencia, y los altos valores de resistencia tienen baja respuesta a la frecuencia. Si se requieren más detalles es necesario consultar los catálogos de los fabricantes.

12.1.2 Resistores de circuitos microelectrónicos⁷⁻⁹

En este artículo se consideran tres tecnologías de procesamiento: la de película gruesa, en la cual los resistores se depositan por medio de un proceso de serigrafía; la de película delgada, en la cual el depósito se hace mediante evaporación al vacío o bombardeo iónico; y la monolítica de silicio, en la cual los resistores forman parte de la pastilla de silicio.

Resistores de película gruesa

Los resistores de película gruesa se emplean en redes de resistores o en circuitos híbridos debido a su bajo costo, poca inversión en equipo, corto tiempo de reposición, buena efectividad en términos de costo para series cortas, buena flexibilidad y capacidad de combinación con componentes activos o pasivos en el mismo sustrato tales como inductores, grandes capacitores y dispositivos lineales, digitales y de potencia.

Como sustrato se emplea casi de manera universal la alúmina. Antes de depositar la pasta del resistor se coloca una pasta conductora para las terminales que consta de oro, platino-oro, paladio-plata o alguna otra combinación. La presencia del conductor afecta las propiedades del resistor a través de reacciones químicas.

Tabla 12-4. Características típicas de los resistores

Característica	Composición de carbono	Película de carbono	Película metálica	Arrollamiento de alambre de precisión	Pastilla
Resistencia	1 Ω a 100 M Ω	1 Ω a 10 M Ω	1 Ω a 10 M Ω	0.1 Ω a 10 M Ω	5 Ω a 5 M Ω
Disipación de potencia	1/8 a 2 W	1/10 a 2 W	1/20 a 2 W	0.1 W y más	50 a 600 mW
Tolerancia en el momento de la adquisición	5, 10, 20%	0.5 a 10%	0.1 a 1%	0.01 a 1%	1 a 10%
Límite de frecuencia	1 MHz	100 MHz	400 MHz	50 kHz	100 MHz
Coefficiente térmico	-800 a +1 600	-200 a -1 000	± 20 a ± 200	± 20 a ± 200	± 100
Fiabilidad	La mejor	Buena	Buena	Muy deficiente	Buena
Estabilidad	Muy deficiente	Buena	Buena	La mejor	Aceptable
Costo aproximado	1 centavo de dólar	0.05 a 1 centavo de dólar	2 a 20 centavos de dólar	50-100 centavos de dólar	1 a 10 centavos de dólar
Principales ventajas	Bajo costo, alta confiabilidad, elevada capacidad para soportar oscilaciones de carga	El menor costo	Buena exactitud y estabilidad, tamaño reducido	Máxima exactitud y estabilidad	El tamaño más reducido
Principales desventajas	Deficiente exactitud, deficiente estabilidad		Deficiente capacidad para soportar oscilaciones de carga	Muy costoso, deficiente respuesta de frecuencia	Disponibilidad limitada

micas entre los dos materiales, y debido a que el espesor del resistor durante la impresión por serigrafado es afectado a medida que el estarcido (la pantalla de seda) se mueve sobre el conductor. Las pastas del resistor constan de un material conductor, vidrio en polvo y un vehículo orgánico para controlar las propiedades básicas de flujo (reología). Anteriormente, la pasta para resistor de mayor uso fue la del paladio-plata, pero el valor de su resistencia es sensible a las condiciones de horneado debido a que depende de la oxidación parcial del paladio. En la actualidad, es más común el empleo de dióxido de rutenio.

Las tintas para resistores se expanden en familias en las que las resistividades laminares cambian en múltiplos de 10 por lo menos sobre una parte del intervalo que va de 1 Ω /cuad a 100 Ω /cuad para un grosor de 25 μm . El grosor depende de la reología de la pasta, el tamaño de poro del estarcido y la velocidad y presión de la escoba de goma. El valor típico es de 25 μm . El ancho mínimo preferido de la línea impresa es de 1.25 mm, y el buen diseño determina que las terminales del conductor deben superponerse al resistor unos 0.25 mm alrededor de los extremos, y que la relación alto/ancho (cuadrangular) debe estar en el intervalo de 0.2 a 5. Deben evitarse las formas en serpentin y las esquinas agudas. Para evitar el empleo de más de una tinta, algunas veces los resistores se conectan en serie o en paralelo. El ancho de la línea y la relación alto/ancho se diseñan para asegurar una densidad de potencia menor de 7.5 W/cm² aproximadamente.

La tolerancia de un resistor de película gruesa tal como se deposita, rara vez es inferior al 10% y puede llegar al 30%. Sin embargo, es posible lograr tolerancias finales del 1% por medio de desbaste con rayo láser o abrasivo. Los resistores se diseñan para un valor del 70 al 80% del valor deseado, y después se ajustan según las necesidades. En la figura 12-2 se presentan varios ejemplos de recorte de resistores.

Los resistores de película gruesa tienen coeficientes térmicos de 50 a 250 ppm / °C y coeficientes de voltaje de 50 ppm / V. Son útiles por lo menos hasta 1 MHz y presentan excelentes características de envejecimiento y duración en carga: por lo general, menos del 1% de cambio después de miles de horas de operación.

Resistores de película delgada

Los resistores de película delgada se emplean en circuitos híbridos en los que se requiere mayor exactitud y estabilidad, menor tamaño y mejor funcionamiento a alta frecuencia que los proporcionados por los resistores de película gruesa. La tecnología de los resistores de película delgada está bien desarrollada, pero suele ser más costosa que la producción de resistores de película gruesa.

Los resistores, conductores y dieléctricos se depositan por evaporación o bombardeo iónico en un proceso de grabado o corrosión selectiva, o posiblemente con una protección (máscara) para delinear las formas geométricas idóneas. Los anchos y separaciones mínimos de las líneas varían de 25 a 50 μm . Todos los

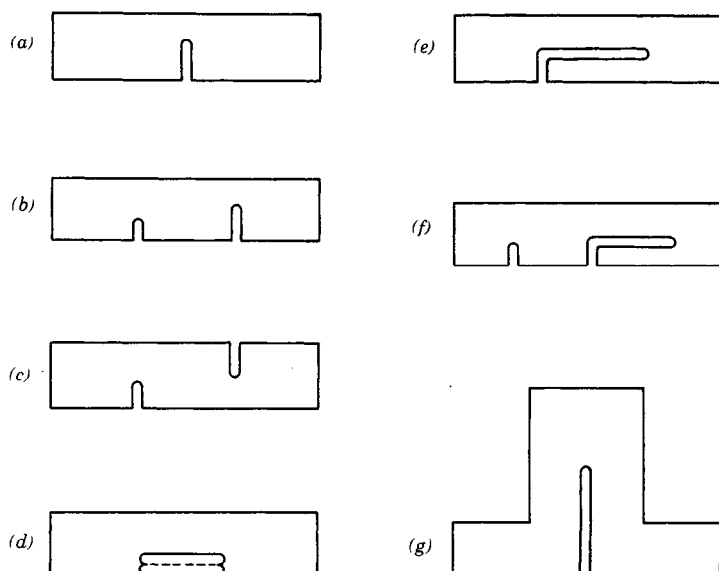


Fig. 12-2. Ejemplos de formas geométricas empleadas en el recorte (ajuste) con rayo láser de resistores de película gruesa: a) corte recto; b) corte doble; c) corte doble invertido; d) corte por barrido; e) corte en L; f) corte en L con vernier, y g) diseño en forma de sombrero de copa.

resistores se colocan en una etapa, lo cual significa que sólo se dispone de un único valor de resistencia laminar, en comparación con los hasta tres valores con que se cuenta en un proceso de película gruesa. Por consiguiente, la configuración geométrica del resistor debe tomar en cuenta amplios intervalos de valores de la resistencia, y los patrones en serpentina son comunes. Para los patrones de resistores en "biblioteca" se emplea una sola capa que es posible recortar sobre un intervalo amplio formando "ojales" y "escaleras" (fig. 12-3).

Las películas resistivas están compuestas de nitruro de tantalio, níquel-cromo o cerámicas metálicas (cermet). Una película típica de cermet es una mezcla de cromo y monóxido de silicio en una capa con grosor aproximado de algunas decenas de nanómetros, resistividad laminar de $500 \Omega/\text{cuad}$ y coeficiente térmico de $100 \text{ ppm}/^\circ\text{C}$.

Las películas de níquel-cromo varían en grosores que van desde menos de 5 nm hasta $1 \mu\text{m}$, con resistividades laminares de 1 a $1000 \Omega/\text{cuad}$. Un valor común es 10 nm , con resistividad laminar de $150 \Omega/\text{cuad}$ y coeficiente térmico de $30 \text{ ppm}/^\circ\text{C}$.

Los resistores de nitruro de tantalio se depositan en una serie de pasos, a fin de producir películas de 20 a 150 nm de grosor sobre un sustrato preparado con capas de titanio, paladio y oro para los contactos eléctricos. El nitruro de tantalio es único en el sentido de que los resistores elaborados con él pueden recortarse por anodización de la película, aunque en la actualidad se emplea más el recorte con láser. Para lograr estabilidad a largo plazo se hornea (p. ej., durante cinco horas a 300°C en aire). El recorte con láser daña el material adyacente a la ranura, por lo cual se requiere una esta-

bilización ulterior. El modelo de biblioteca de la figura 12-3 tiene tal forma para que los cortes del láser en ojales y escaleras no se encuentren en la trayectoria portadora de corriente, y el último recorte micrométrico se realiza de tal modo que el material dañado sea un porcentaje muy pequeño del área de la resistencia. Se han proyectado desviaciones tan bajas como 0.1% durante 20 años a 65°C ¹⁰. Naturalmente, mientras más alta sea la temperatura de operación de la película, mayor será la desviación. Un valor razonable para la disipación de potencia es $3 \text{ W}/\text{cm}^2$ a 25°C de temperatura ambiente.

Resistores monolíticos de silicio

La calidad de los resistores formados como una parte integral de la pastilla de silicio es muy deficiente, con grandes tolerancias, elevados coeficientes térmicos de resistencia, deficiente respuesta a la frecuencia y limitada disipación de potencia.

Por otra parte, estos resistores ocupan una valiosa área, elevan el consumo del suministro de energía, incrementan la disipación de potencia, introducen elementos parásitos y complican la configuración. Sin embargo, constituyen una parte indispensable de los circuitos integrados.

Anteriormente, los resistores se delinearon durante la difusión de la base, empleando en ocasiones la difusión del emisor para bajos valores. Con la proliferación de las tecnologías de los circuitos integrados y con los desacostumbrados requerimientos de la integración a muy alta escala (VLSI, del inglés *very-large-scale-integration*), los diseñadores de pastillas (chips) están aprovechando el amplio intervalo de capacidades producido durante el procesamiento de circuitos nor-

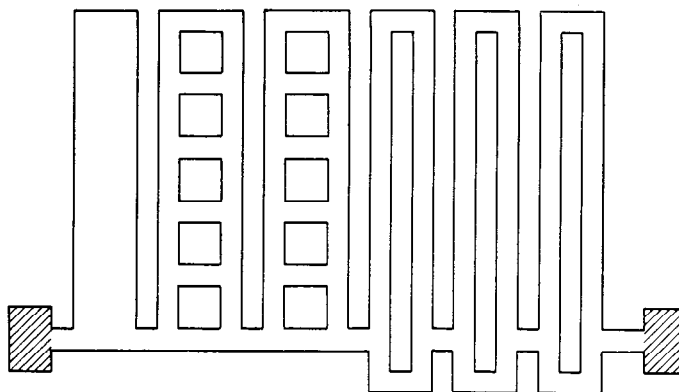


Fig. 12-3. Patrón estándar para obtener un amplio intervalo de valores de resistencia. Es posible agregar grandes incrementos mediante la apertura de "ojales" en la derecha, y valores más bajos mediante la apertura de peldaños en las "escaleras" del centro, con un ajuste continuo de vernier por medio de un corte recto con láser en la columna de la izquierda.

males. Por supuesto que el proceso de fabricación se adapta a los requerimientos de los dispositivos activos, y la optimización de las propiedades del resistor se desplaza a un segundo plano.

Por motivos de espacio, aquí no es posible hacer un análisis completo sobre el diseño de resistores de circuitos integrados.

En términos generales, las tolerancias típicas varían hasta el 20%, los coeficientes térmicos son de varios miles de ppm / °C, la respuesta a las frecuencias está limitada a unas cuantas decenas de megahertz y la disipación de potencia tiene un máximo de unos 10 μ W/ μ m². Sin embargo, las relaciones de resistencia pueden mantenerse por debajo del 1%, y el seguimiento de temperatura es excelente.

Estas importantes ventajas se han incorporado a diseños complejos de convertidores de analógico a digital y de digital a analógico, así como a amplificadores diferenciales, entre otros. Además, puede disponerse

de resistores de película delgada de alta calidad, fabricados por depósito en la superficie de la pastilla de silicio.

En la tabla 12-5 se resumen las propiedades de materiales potencialmente resistivos para su uso en circuitos integrados. En la tabla se observan solamente valores representativos y no se incluyen muchos aspectos del diseño. Para ello, es necesario consultar la lista de referencias bibliográficas que se encuentra al final de esta sección.

12.1.3 Termistores^{11,12}

Los **termistores** son resistores sensibles a la temperatura; es decir, resistores con coeficiente térmico inusualmente grande. La resistencia puede cambiar en dos órdenes de magnitud entre 0 y 100 °C, o hasta siete órdenes de magnitud desde -100 hasta 400 °C. Esto hace de los termistores unos detectores de temperatura

Tabla 12-5. Materiales resistivos para circuitos integrados

Material	Resistividad ρ (Ω -cm) o resistividad laminar R (Ω /cuad)	Comentarios
Sustrato	$\rho = 0.1-100$	Casi nunca se emplea
Capa epitaxial	$\rho = 0.1-50$	Algunas veces se emplea para resistores de alto valor y alto voltaje
Difusión de la base	$R = 100-300$	El más empleado; intervalo práctico de resistencia de 20 Ω a 100 k Ω ; ancho mínimo de 5 μ m
Difusión del emisor	$R = 1-5$	Empleado para resistores de bajo valor
Regiones de implantación de iones	$R = 10^3-10^4$	Empleado para resistores de elevado valor
Capa de polisilicio	$R = 10-200$	
Metalización de aluminio	$R = 0.1-0.05$	Por lo general, se le considera un conductor, aunque la caída de IR (voltaje) debe considerarse en los diseños de I ² L
Nicromo y tantalio	$R = 1-1\ 000$	Resistores de alta calidad, pero requieren etapas extra de procesamiento

extremadamente sensibles, hasta tal punto que las pruebas de calidad o de calibración no pueden efectuarse en el aire a temperatura ambiente, sino en un baño de temperatura constante que se mantiene cuando más en el intervalo de variación de $\pm 0.1^\circ\text{C}$.

Construcción

Los termistores más comunes con coeficiente térmico negativo (NTC, del inglés *negative temperature coefficient*) se elaboran de materiales cerámicos, como los óxidos de manganeso, cromo, níquel, cobalto, hierro, cobre y uranio. Con la adición de pequeñas cantidades de ciertos metales, estos compuestos se convierten en semiconductores (tipo *p* o tipo *n*) que experimentan un decremento en la resistencia a medida que se activan térmicamente portadores extra.

Los termistores con forma de disco se elaboran de la misma forma que los capacitores cerámicos de disco. Las materias primas se pulverizan finamente, y después este polvo se prensa en discos, que se calientan para eliminar el adhesivo orgánico, y luego se sintetizan a elevadas temperaturas. Se aplica pintura de plata, se fijan las puntas de contacto y se proporciona aislamiento. Entre otras formas se incluyen cuentas o perlas (que se elaboran aplicando una gota de suspensión sobre dos alambres conductores), varillas extruidas y arandelas en un amplio intervalo de tamaños. Para un corto tiempo de respuestas se requiere una pequeña masa térmica, y es posible producir cuentas aisladoras tan pequeñas como de $100\ \mu\text{m}$ de diámetro, con tiempos de respuesta de una fracción de segundo en aire tranquilo, o de algunos milisegundos en una inmersión líquida. El material del termistor suele en-

capsularse en un recubrimiento epóxico de conformación, pero también se emplean empaques de vidrio para sondas, y pueden conseguirse con facilidad montajes sensores especiales.

Los termistores con coeficiente térmico positivo (PTC, del inglés *positive temperature coefficient*) se elaboran con titanato de bario, titanato de estroncio o silicio. Las mezclas de titanato producen un incremento en la resistencia con la temperatura, como resultado de un complicado efecto de carga espacial en la superficie de las fronteras granulares. En los dispositivos de silicio, la característica deseada se produce por medio de la disminución en la movilidad. En cada material, el efecto PTC ocurre sólo en un intervalo muy limitado de temperatura; fuera de este intervalo, el coeficiente térmico es pequeño y negativo.

Características

Las características de resistencia en función de la temperatura de una familia típica de termistores NTC se muestran en la figura 12-4. Los valores paramétricos de la resistencia se refieren a la resistencia a una temperatura ambiente de 25°C bajo condiciones de disipación de potencia despreciable en el termistor. Si la disipación de potencia es suficientemente alta para incrementar la temperatura inicial del termistor inclusive en una pequeña cantidad, la resistencia cambia. Entonces, la relación entre el voltaje aplicado y la corriente es altamente no lineal, como se observa en la figura 12-5. Para un termistor dado, estas curvas dependen de la temperatura ambiente.

En las figuras 12-6 y 12-7 se muestran curvas correspondientes para un termistor PTC.

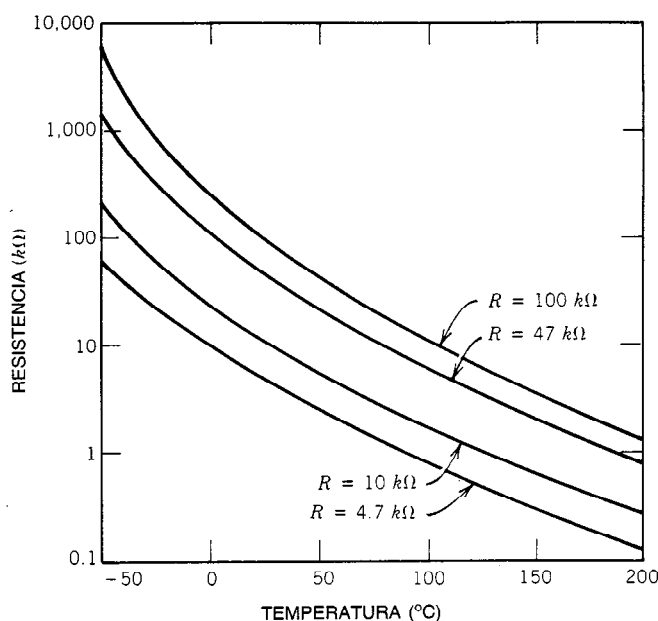


Fig. 12-4. Ejemplo de la variación de resistencia con la temperatura para un termistor con coeficiente térmico negativo (NTC). Los valores de resistencia marcados son los nominales a 25°C .

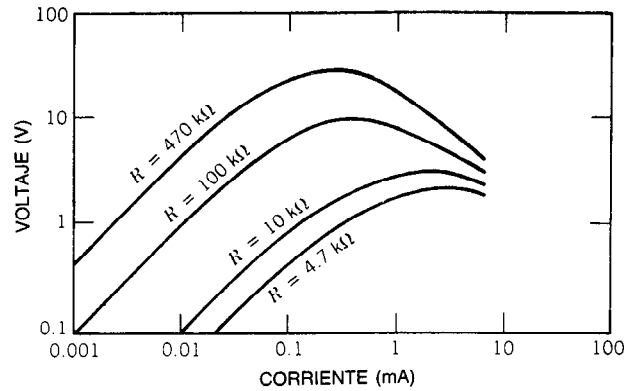


Fig. 12-5. Ejemplo de las características volt-ampere para un termistor con coeficiente térmico negativo (NTC) a una temperatura ambiente de 25 °C. La no linealidad es el resultado del calentamiento interno en alta disipación de potencia.

Algunas veces, las curvas de resistencia que se muestran en la figura 12-4 se aproximan mediante una ecuación de la forma:

$$\frac{R_{T1}}{R_{T2}} = \exp \left\{ \beta \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \right\}$$

donde T está en °K. Entonces, el coeficiente térmico es el siguiente:

$$\alpha = \frac{1}{R} \frac{dR}{dT} = - \frac{\beta}{T^2}$$

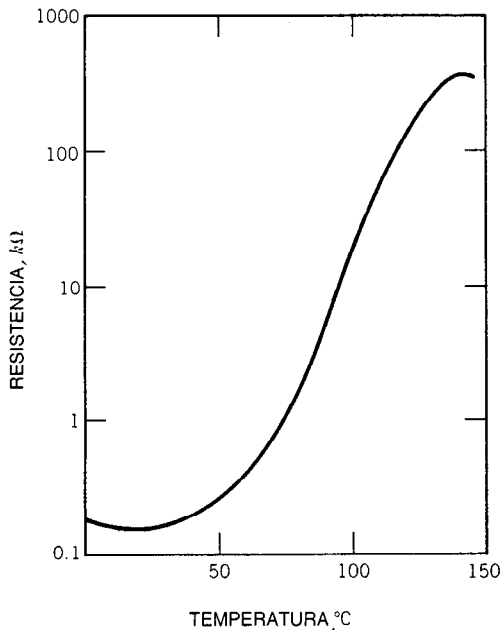


Fig. 12-6. Comportamiento típico de un termistor con coeficiente térmico positivo (PTC).

Los fabricantes especifican la resistencia sin disipación a 25 °C. El intervalo acostumbrado es de 10 Ω a 1 MΩ, con una tolerancia del 10 al 20%. La tolerancia de la resistencia puede traducirse en una tolerancia de temperatura mediante una curva de resistencia como la que se muestra en la figura 12-4. Para indicar la variación de la resistencia con la temperatura, es posible que se proporcione el coeficiente térmico (3 a 6%/°C), o bien el valor β (2 000 a 5 000/K). Con frecuencia se especifica la relación de resistencias a 25 °C y a 125 °C (típicamente, de 5 a 50). La disipación máxima de potencia varía de 10 mW a 2 W, con un factor de disipación térmica (inversa de la resistencia térmica) de 0.1 a 25 mW/°C. Las constantes temporales térmicas varían de 0.5 a 150 s, dependiendo del tamaño y empaque.

Una aplicación importante de los termistores es como patrones de temperatura. La calibración, referida a las normas de la National Bureau of Standards, puede efectuarse a 0.0015 °C. La estabilidad se encuentra dentro de 0.005 °C por año. Además de este empleo, los termistores se utilizan para compensación en circuitos activos, regulación de voltaje y corriente, retrasos de tiempo, detección y control, protección contra sobretensiones y eliminación de chispas.

12.1.4 Varistores¹³⁻¹⁵

Un **varistor** es un resistor sensible al voltaje. La característica de potencia aparente (volt-ampere) es altamente no lineal; a medida que se incrementa el voltaje, la resistencia dinámica decrece y la corriente se eleva con rapidez. El mecanismo se debe a procesos electrónicos y no a efectos térmicos (en contraste con lo que ocurre en los termistores). Los primeros tipos estaban hechos de carburo de silicio, pero ya se han sustituido en gran medida por dispositivos de óxido de cinc. Otros componentes que presentan características altamente no lineales y algunas de las mismas aplicaciones son los supresores de selenio (en la actualidad obsoletos), los diodos de unión *pn* de silicio y los tubos de descarga gaseosa. Aunque este análisis está bastan-

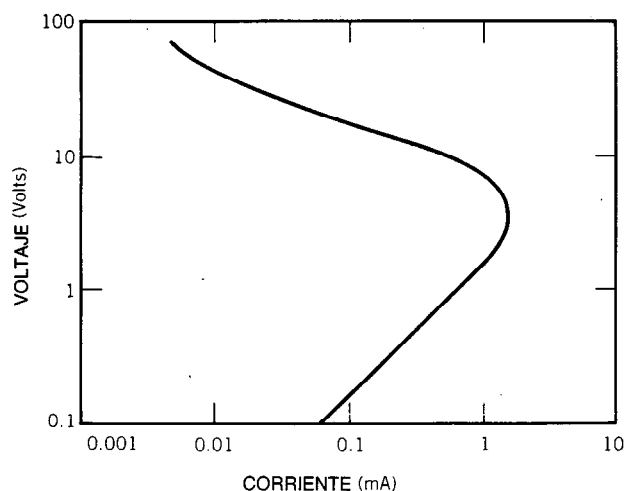


Fig. 12-7. Características de potencia aparente (volt-ampere) de un termistor con coeficiente térmico positivo (PTC). Para grandes valores de voltaje la corriente empieza a disminuir, dado que el incremento en la temperatura interna provoca un gran incremento en la resistencia.

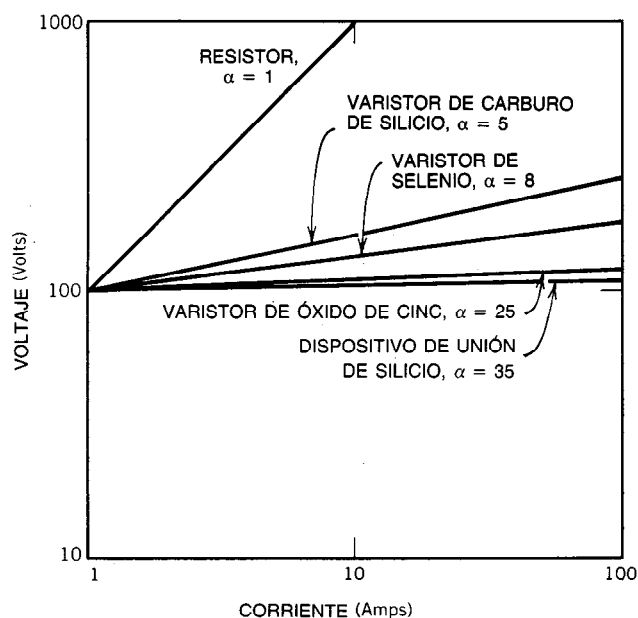


Fig. 12-8. Comparación de diferentes tipos de varistores; α es el exponente del voltaje en la ecuación $I = kV^\alpha$.

te restringido a los varistores de óxido de cinc, se hacen comparaciones con otros dispositivos, como se observa en la figura 12-8. La figura ilustra el rápido incremento de la corriente en estos dispositivos a medida que se eleva el voltaje.

Características

Las características de los varistores de óxido de cinc están controladas por las materias primas, su procesa-

miento y el tamaño y forma del dispositivo terminado. El material inicial es óxido de cinc hasta con un 10% de otros óxidos metálicos.

El proceso de manufactura es exactamente igual que el de los capacitores cerámicos de disco o el de los termistores de disco. El disco completo consta de partículas conductoras de óxido de cinc de 25 μm de diámetro aproximadamente, rodeadas de una capa de material aislante. La conducción del varistor proviene de

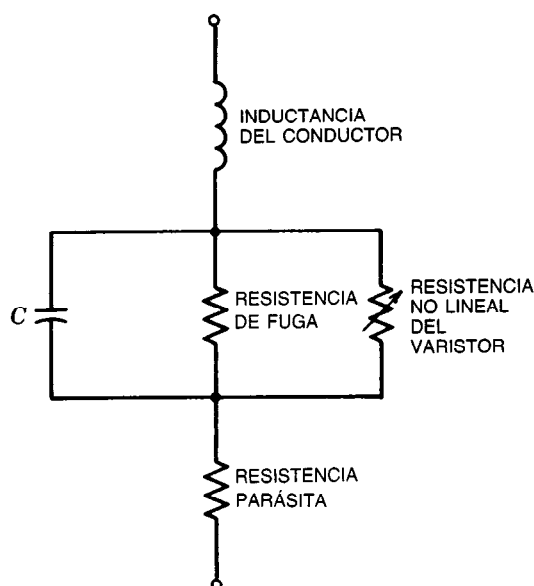


Fig. 12-9. Circuito equivalente para un varistor de ZnO.

la tunelización de electrones a través de la capa de aislante, iniciada por un exceso de portadores en el óxido de cinc. Para esto se requiere un voltaje de 3 a 3.5 V por frontera granular para una densidad de corriente de 1 mA/cm². Controlar el tamaño de grano y el grosor del disco significa que el voltaje al cual empieza a fluir una cantidad apreciable de corriente es fijo. El diámetro del disco determina la corriente total y la capacidad de manejo de corriente del varistor.

La característica de potencia aparente suele aproximarse mediante una relación de la forma:

$$I = kV^\alpha$$

En la figura 12-8 se indican los valores de α para diferentes clases de dispositivos. Para los varistores de óxido de cinc, α varía de 20 a 80. A valores bajos de voltaje, la corriente predicha por la ecuación anterior es acrecentada por fugas, mientras que a niveles elevados de corriente, la caída de tensión a través de la conductividad finita de las partículas de óxido de cinc es demasiado grande para ignorarla. Esto conduce a un circuito equivalente para el varistor, como se muestra en la figura 12-9. La capacitancia es sorprendentemente alta, no debido a que la constante dieléctrica de los materiales constituyentes sea de por sí muy alta, sino porque los granos conductores en una matriz aislante actúan como un dieléctrico artificial e incrementan la constante dieléctrica efectiva hasta aproximadamente 1 000. La capacitancia de 1 MHz varía desde 10 pF para pequeños dispositivos de alto voltaje hasta 0.02 μ F para piezas de elevada corriente y bajo voltaje.

Aplicaciones

Con mucho, la aplicación más común de los varistores es la supresión de oscilaciones momentáneas (efectos

transitorios). En operación normal la resistencia es tan alta que fluye una corriente despreciable. Cuando se presenta una oscilación momentánea potencialmente destructiva, la resistencia disminuye y el dispositivo conduce una corriente mucho mayor, fijando en efecto el voltaje.

Algunas especificaciones importantes son el voltaje al que fluye 1 mA, el voltaje máximo para una corriente de onda específica, la capacidad de absorción de energía, la corriente máxima y la capacitancia. Por ejemplo, un varistor para proteger de oscilaciones momentáneas un aparato en una línea de 120 V debe contar con un voltaje mínimo de 184 V a 1 mA de CD, una CA pico máxima de 273 V a 1 mA, un voltaje de fijación de 390 V a 10 A (de aquí que deba tener una relación de fijación de $390/120 \sqrt{2} = 2.30$ a 10 A), una corriente de pico de 500 A, y una energía de onda (o energía característica) de 10 J.

En general, el voltaje mínimo a 1 mA es determinado por el voltaje nominal de operación, y el voltaje de fijación, la corriente pico y los requerimientos de capacidad para resistir las ondas irruptivas deben estimarse a partir de las características del ambiente transitorio, la impedancia de la fuente y la naturaleza del circuito por proteger.

REFERENCIAS BIBLIOGRÁFICAS

1. Charles A. Harper, ed., *Handbook of Components for Electronics*, McGraw-Hill, Nueva York.
2. Gerald L. Ginsberg, *A User's Guide to Selecting Electronics Components*, Wiley, Nueva York.
3. "Military Standard-Resistors, Selection and Use of", *MIL-STD-199B*.
4. "Choosing the Right Resistor: Three Experts Give Their View", *Electr. Des.* **27**:74-78.
5. Henry Domingos y Donald C. Wunsch, "High Pulse Power Failure of Discrete Resistors", *IEEE Trans. Parts, Hyb. Packag.* **THP-11**:225-229.
6. "Reliability Prediction of Electronic Equipment", *MIL-HDBK-217C*.
7. Douglas A. Hamilton y William G. Howard, *Basic Integrated Circuit Engineering*, McGraw-Hill, Nueva York.
8. Roy A. Colclaser, *Microelectronics: Processing and Device Design*, Wiley, Nueva York.
9. Arthur B. Glaser y Gerald E. Subak-Sharpe, *Integrated Circuit Engineering*, Addison-Wesley, Reading, MA.
10. P. L. Scarff y L. J. Kiszka, en *Proc. 31st Electr. Comp. Conf.* pp. 449-455.
11. *Thermistor Manual*, Fenwal Electronics.
12. *PTC/NTC Thermistors*, Mepco/Electra.
13. *Transient Voltage Suppression Manual*, 2.^a ed., General Electric.
14. Lionel M. Levinson y Herbert R. Philipp, "ZnO Varistors for Transient Protection", *IEEE Trans. Parts, Hyb. Packag.* **THP-13**:338-343.
15. G. D. Mahan, Lionel M. Levinson y H. R. Philipp, "Theory of Conduction in ZnO Varistors", *J. Appl. Phys.* **50**:2799-2812.

12.2 CAPACITORES

Henry Domingos

12.2.1 Introducción

Los capacitores son uno de los componentes electrónicos de mayor uso, y sus ventas se han incrementado de forma continua, hasta el punto de que sólo en Estados Unidos el mercado supera los mil millones de dólares y los seis mil millones de unidades. La tasa de crecimiento (las ventas de unidades se han duplicado en los últimos diez años) se debe al desarrollo explosivo de las industrias de las computadoras y de los circuitos integrados, así como al creciente empleo de la electrónica en nuevas aplicaciones tales como automóviles y aparatos para el hogar. La diversificación de los mercados ha conducido a la creación de una amplia variedad de productos para satisfacer una gran variedad de requerimientos eléctricos, y la industria está mejorando constantemente sus productos a fin de lograr tamaños más pequeños, mayor fiabilidad, mejor rendimiento y costos más bajos. El desarrollo de nuevas líneas de productos ha conducido a realizar cambios en la selección de productos, lo que refleja nuevas demandas. En Estados Unidos, aproximadamente el 50% de los capacitores vendidos tienen muchas capas de cerámica, mientras que en Japón y Europa occidental dominan los capacitores electrolíticos y los capacitores de película de plástico, respectivamente.

Los recientes desarrollos tecnológicos dan testimonio de la vitalidad de la industria de los capacitores, y señalan una evolución continua de líneas de productos en los próximos años.

12.2.2 Definiciones y propiedades generales de los circuitos

La **capacitancia** se define como la relación que existe de la carga al voltaje entre dos conductores. Es decir:

$$C = \frac{q}{v} \quad (12-1)$$

donde C está en farads, q en coulombs y v en volts. Dado que la corriente es la derivada de la carga respecto al tiempo, la ecuación anterior conduce a:

$$i = C \frac{dv}{dt} \quad (12-2)$$

o bien:

$$v = \frac{1}{C} \int i \, dt \quad (12-3)$$

donde la corriente i (amperes) y el voltaje v pueden ser funciones del tiempo. La energía instantánea almacenada en un capacitor está dada por:

$$w = \frac{1}{2} C v^2 \quad (12-4)$$

o bien:

$$w = \frac{1}{2} \frac{q^2}{C} \quad (12-5)$$

donde w está en joules.

En condiciones de CA en estado estable, la corriente y el voltaje en un capacitor ideal están relacionados por:

$$I = 2\pi f C V \quad (12-6)$$

donde I y V son la corriente y el voltaje rcm (o pico), respectivamente, y f es la frecuencia en hertz. La reactancia está dada por:

$$X_c = \frac{1}{2\pi f C} \quad (12-7)$$

y sus unidades son ohms.

En la figura 12-10 se muestran varias configuraciones de conductor, de interés para aplicaciones electrónicas. La capacitancia de los capacitores de dos placas paralelas (parte *a* de la figura) está dada por:

$$C = \frac{\epsilon A}{d} \quad (12-8)$$

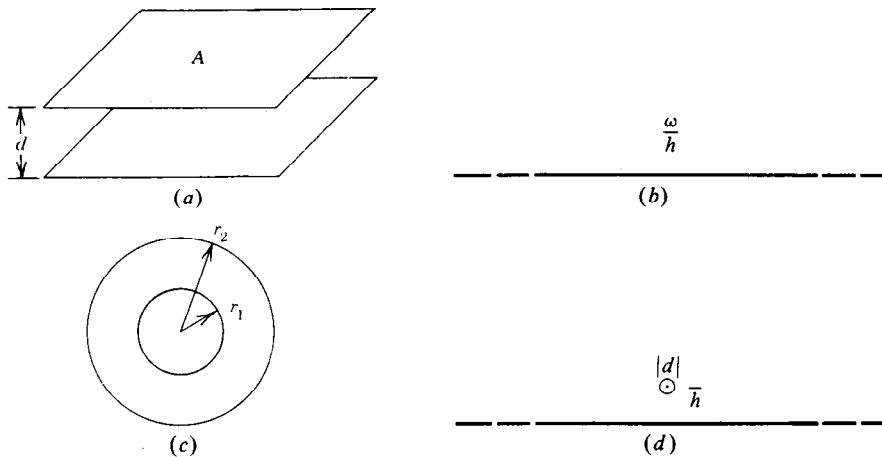
donde $\epsilon = \epsilon_r \epsilon_0$ es la permisividad del dieléctrico, ϵ_r es la permisividad relativa (o constante dieléctrica) y ϵ_0 es la permisividad del espacio libre, cuyo valor es 8.85×10^{-12} F/m. En la ecuación 12-8 se supone que las líneas del campo eléctrico son rectas paralelas con separación uniforme y confinadas a la región entre las placas. Si se consideran sus efectos de borde, el valor de la capacitancia es algo mayor, pero no es posible obtener una expresión exacta para el caso general. Algunas veces se suma una cantidad empírica, 0.44d, a la longitud de cada lado para aproximar el efecto del campo marginal.

La parte *b* de la figura 12-10 es importante para aplicaciones como microcinta, línea de cinta, tableros de circuitos impresos e interconexiones en una pastilla de circuitos integrados. El valor aproximado de la capacitancia por unidad de longitud está dado por:

$$C = \frac{\epsilon w}{h} \left\{ 1 + \frac{2h}{\pi w} \left[1 + \ln \left(\frac{\pi w}{h} \right) \right] \right\} \quad (12-9)$$

para $w/h \gg 1$. El error en esta expresión es menor del 10% para $w = 2h$ y decrece a medida que w/h aumenta. Esta fórmula subestima la capacitancia entre un conductor y un plano de tierra si el grosor del conductor no es despreciable en comparación con w y h . Además, la fórmula sólo se aplica si está presente un material dieléctrico.

Las otras formas geométricas de la figura 12-10 son exactas (despreciando los efectos de extremos) y son útiles para calcular la capacitancia de un cable coaxial, de un conductor próximo a un armazón metálico, o de un par de alambres (en la parte *d* se muestra la mitad de dicha configuración).



Cualquier capacitor real disipa energía además de almacenarla. Esta pérdida de energía, que aparece como calor, es el resultado de varios efectos:

- Conductividad finita de los electrodos, los contactos de éstos y los alambres conductores.
- La resistividad finita del dieléctrico, lo que resulta en una fuga de corriente de CD.
- Pérdidas dieléctricas de CA, que dependen de la frecuencia y del mecanismo de polarización del dieléctrico.
- Efectos corona.
- Absorción dieléctrica (que se manifiesta como una elevación de voltaje a través de un capacitor después de descargado).

El modelo más sencillo para explicar las pérdidas del capacitor consiste en un resistor en serie con un capacitor ideal, tal como se muestra en la figura 12-11. Esta resistencia se denomina resistencia serie equivalente, o ESR (del inglés *equivalent series resistance*). Algunos puentes de medición de capacitores regulan el modelo equivalente de una resistencia y una capacitancia en paralelo. En la figura se presentan las fórmulas de conversión y se hace hincapié en el hecho de que el punto

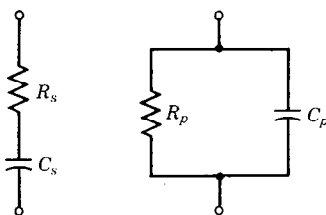


Fig. 12-11. Circuitos equivalentes en serie y en paralelo para capacitores con pérdidas. $R_s = R_p/(1 + R_p^2 \omega^2 C_p^2)$; $C_s = (1 + R_p^2 \omega^2 C_p^2)/R_p^2 \omega^2 C_p^2$; $R_p = R_s + 1/R_s \omega^2 C_s^2$; $C_p = C_s/(1 + R_s^2 \omega^2 C_s^2)$.

C_s no es igual a C_p y que los componentes de cada modelo suelen depender de la frecuencia.

El factor de calidad Q de una red se define como 2π por el producto de la razón de la máxima energía almacenada entre la energía disipada por ciclo a una frecuencia dada. Para un capacitor:

$$Q = \frac{1}{2\pi f C_s R_s} = 2\pi f C_p R_p \quad (12-10)$$

Cuanto mayor es el valor de Q , más ideal es el capacitor.

Con frecuencia, los puentes de medición miden el factor de disipación, DF (del inglés *dissipation factor*), que es el recíproco de Q , a menudo expresado como porcentaje. Algunas veces, las pérdidas se especifican en términos del ángulo de pérdida o factor de potencia. En la figura 12-12 se muestran las relaciones entre estas cantidades.

Un modelo exacto para un capacitor que debe ser válido entre un amplio intervalo de frecuencia es difícil de especificar, ya que cualquier capacitor es una red distribuida. Una representación razonablemente exacta, útil para explicar ciertas características, se observa en la figura 12-13. En esta figura, las pérdidas dieléctricas se representan por R_p y las pérdidas del conductor por R_s . La adición importante es la inductancia en serie L , que surge de la inductancia de los conductores y de la inductancia de la estructura misma del capacitor.

Si se desprecia R_p , el capacitor se comporta como un circuito RLC en serie, cuya impedancia se muestra en la figura 12-14. En términos generales, los grandes capacitores tienen bajas frecuencias de autorresonancia. A una frecuencia mayor que la de autorresonancia, un capacitor se comporta como un inductor, y la impedancia aumenta con la frecuencia. Por esta razón, se coloca un pequeño capacitor, por ejemplo, de $0.01 \mu\text{F}$, en paralelo con un gran capacitor a fin de mantener una baja impedancia combinada a altas frecuencias.

12.2.3 Capacitores electrolíticos de aluminio

Los capacitores electrolíticos de aluminio son populares debido a su bajo costo y gran capacitancia por unidad de volumen. Se emplean en aplicaciones de filtrado, derivación y acoplamiento, en las que se requiere gran capacitancia, pero en las que las tolerancias y la variación de capacitancia con la temperatura no son importantes. Existen en el mercado unidades polarizadas y no polarizadas.

Los capacitores electrolíticos de aluminio son del tipo de hojas metálicas, con un electrolito que puede ser acuoso, en pasta o "seco" (sin agua). El ánodo se forma con lámina de aluminio de alta pureza (99.99%) de 25 a 100 μm de grueso. La lámina puede ser lisa (sin grabar), pero con mayor frecuencia suele grabarse electrolíticamente a fin de incrementar el área superficial, con lo que se eleva la capacitancia. El incremento es en un factor de 8 a 30 veces, dependiendo del grosor del óxido que actúa como dieléctrico.

La lámina se anodiza a fin de producir una capa de óxido de aluminio. El voltaje aplicado durante la etapa final de anodización se conoce como voltaje de formación, y determina el grosor del óxido (aprox. 1.2 nm/V). Un límite superior para el voltaje de formación, denominado voltaje de escintilación (centelleo), se alcanza cuando el campo eléctrico provoca disrupción momentánea local. Este voltaje de escintilación es aproximadamente igual a 600 V para el aluminio, dependiendo de la resistividad del electrolito. El voltaje nominal del capacitor suele ser del 60 al 70% del voltaje de formación.

La conexión del cátodo es una segunda lámina de aluminio. Si el capacitor es del tipo no polarizado, también se anodiza la lámina del cátodo con el mismo voltaje de formación. Para los tipos polarizados, la lámina del cátodo no tiene óxido o la capa que tiene es muy delgada. La lámina del cátodo sirve sólo como contacto eléctrico con el electrolito, que es el cátodo real. El electrolito más común es del tipo de glicol-

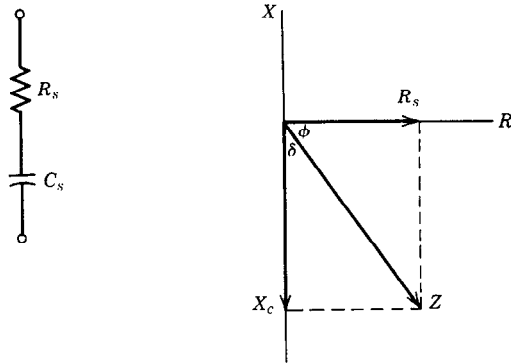


Fig. 12-12. Definiciones para un capacitor con pérdidas. Q es el factor de calidad, DF el factor de disipación, PF el factor de potencia, y δ es el ángulo de pérdida.

$$X_c = -\frac{1}{2\pi f C}; Q = \frac{1}{DF} = \frac{1}{2\pi f C_s R_s}$$

$$DF = \tan \delta = \cot \phi = 2\pi f C_s R_s; PF = \sin \delta = \cos \phi = R_s/Z.$$

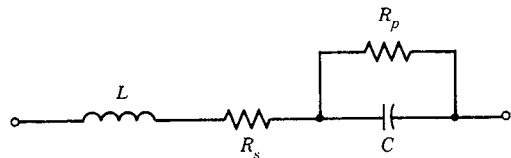


Fig. 12-13. Modelo razonablemente exacto de un capacitor.

borato, elaborado por medio de la disolución de pentaborato de amonio en glicol o de ácido bórico en glicol. Algunas veces se agrega agua a fin de mejorar las características a bajas temperaturas. Algunos otros electrolitos incluyen etilenglicol o dimetilformamida.

Las láminas del ánodo y del cátodo se sueldan a alambres conectores y a continuación se enrollan de

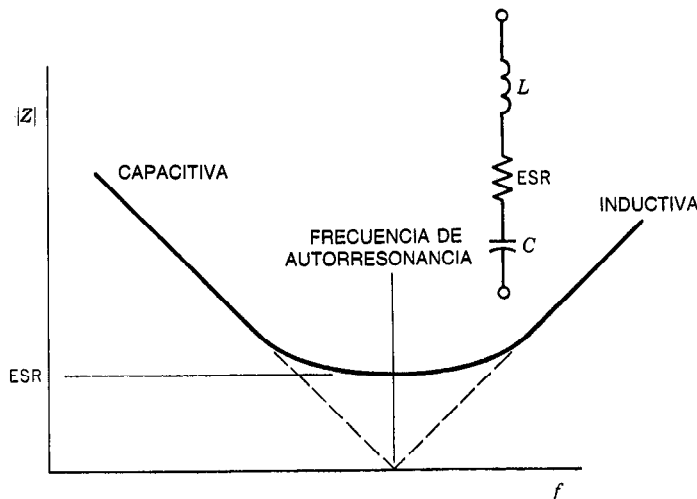


Fig. 12-14. Variación de impedancia de un capacitor en función de la frecuencia.

manera apretada sobre un mandril extraíble con separadores de papel poroso. El papel evita los cortocircuitos entre las láminas, provocados por bordes dentados o superficies ásperas, y absorbe el electrolito a fin de mantener contacto uniforme y estrecho con las superficies de las láminas. En una forma más reciente de construcción, las láminas se disponen en pilas, con una extensión de lengüeta en cada lámina. Esta disposición puede reducir tanto la ESR como la inductancia en un orden de magnitud.

Los capacitores electrolíticos de aluminio se encuentran disponibles en valores que van de $1\ \mu\text{F}$ hasta $1\ \text{F}$, con voltajes nominales de 3 a 475 V. Voltajes superiores implican valores más bajos de capacitancia para un volumen dado, ya que requieren óxidos más gruesos, y con ello menos capacitancia por unidad de área de la lámina. Las tolerancias suelen variar de -10 a $+150\%$. La capacitancia está estrechamente relacionada con la temperatura y puede decrecer en un orden de magnitud desde la temperatura ambiente hasta -55°C . Esta variación se reduce en capacitores de primera calidad y en productos recientes con formulaciones electrolíticas más complicadas.

Los capacitores electrolíticos no están diseñados para aplicaciones a frecuencias elevadas, y la impedancia puede alcanzar un valor mínimo a frecuencias tan bajas como $10\ \text{kHz}$. La ESR es baja a temperatura ambiente, pero a -55°C se eleva de forma aproximada en un orden de magnitud.

La corriente de fuga es considerable en los capacitores electrolíticos de aluminio, por lo que requiere ciertas precauciones. Durante largos periodos de almacenamiento a alta temperatura el óxido puede disolverse parcialmente en el electrolito. Si después se aplica el voltaje nominal, la corriente puede llegar a ser extremadamente alta, conduciendo a la falla catastrófica. El procedimiento adecuado consiste en reformar el óxido limitando inicialmente la corriente con una resistencia en serie o con un voltaje reducido, hasta la restauración del óxido. Es característico de los capacitores electrolíticos de aluminio que la corriente de fuga disminuya durante la operación. En el uso normal, la corriente de fuga aumenta con el voltaje aplicado y con la temperatura. Como guía muy general, la corriente se duplica a medida que el voltaje aplicado se incrementa del 50 al 100% del valor nominal, y se duplica por cada 25°C de aumento en la temperatura.

Los capacitores electrolíticos de aluminio presentan un decremento gradual en capacitancia sobre un largo periodo, debido a la pérdida de electrolito a través de los sellos, aunque con los tipos recientes de empaque se ha reducido de manera significativa este deterioro, y los capacitores presentan en la actualidad un decremento del 10%, o menor, al cabo de 10 000 horas.

Otro problema que debe observarse implica el empleo de ciertos agentes limpiadores en los tableros de circuitos impresos. El cloro de los solventes de hidrocarburos halogenados, como el freón, puede penetrar por los sellos y atacar la estructura interna de aluminio, provocando la falla en poco tiempo. Para la limpieza se recomienda xileno, alcoholes y ciertos tipos de detergentes exentos de cloro.

12.2.4 Capacitores electrolíticos de tantalio

Los capacitores de tantalio son más flexibles y confiables, y presentan mejores características que los electrolíticos de aluminio, pero también su costo es mucho más elevado. Existen tres tipos: de hojas metálicas, húmedo y sólido.

Los capacitores de hojas metálicas (láminas) se elaboran del mismo modo que los electrolíticos de aluminio. La lámina de tantalio, de 10 a $25\ \mu\text{m}$ de espesor, se anodiza para formar pentóxido de tantalio, que funciona como dieléctrico. El grosor es aproximadamente $1.8\ \text{nm/V}$ de voltaje de formación. Para las unidades polarizadas, la lámina del ánodo se anodiza mientras que para los capacitores no polarizados se anodizan ambas láminas. Los alambres conductores de tantalio se sueldan por puntos tanto a la lámina del ánodo como a la del cátodo, las cuales se arrollan después con separadores de papel en un rollo compacto. Este rollo se inserta dentro de una envoltura metálica y, a fin de mejorar el rendimiento, se agrega un electrolito idóneo, como etilenglicol o dimetilformamida con nitrato de amonio, pentaborato de amonio o polifosfatos.

Los capacitores de hojas de tantalio existen en el mercado en tamaños que varían de 0.12 hasta $3\ 500\ \mu\text{F}$, a voltajes hasta de $450\ \text{V}$. Pueden emplearse sobre todo el intervalo militar de temperatura, de -55 a 125°C . La capacitancia aumenta lentamente en este intervalo de temperatura, y suele variar muy poco; aproximadamente $\pm 10\%$ para algunas unidades. La ESR decrece con rapidez a medida que la temperatura se incrementa desde el límite inferior, y la corriente de fuga aumenta drásticamente. A 125°C la corriente de fuga es aproximadamente 40 veces el valor a temperatura ambiente.

La mayor parte de las aplicaciones para este tipo de capacitor se encuentran en los intervalos de voltaje superiores, en los que no es posible aplicar los condensadores de tantalio húmedo, y cuando se requieren calidades superiores a las de los electrolíticos de aluminio, a pesar del mayor costo. Las desventajas, en comparación con otros tipos de capacitores de tantalio, son: gran tamaño, alta ESR, elevadas corrientes de fuga y gran variación en la capacitancia con la temperatura.

El condensador de tantalio húmedo presenta la mayor eficiencia volumétrica de todos los condensadores considerados. La clave de esta característica es la gran área superficial de las escamas porosas de tantalio empleadas como ánodo.

Para formar las escamas, se mezcla polvo de tantalio con un adhesivo y se comprime alrededor de un alambre conductor de tantalio. Se sintetiza el conjunto a muy altas temperaturas al vacío a fin de evaporar el adhesivo y cualesquiera impurezas, y soldar entre sí las partículas de tantalio. Esto produce un ensamblaje anódico fuerte, pero ligero y poroso. El trozo de metal se sumerge en un baño electroquímico y se anodiza para producir una capa uniforme de pentóxido de tantalio. Después, el ánodo se introduce en una envoltura chapeada con plata y platino, que actúa como cátodo.

La envoltura se llena con un electrolito, ya sea cloruro de litio o una solución de ácido sulfúrico en forma de gel, y se sella.

La principal aplicación de los condensadores de tantalio húmedo se encuentra en filtros de fuentes de alimentación. Aunque el máximo voltaje nominal es sólo de 125 V, existen en el mercado empaques de unidades múltiples con voltajes nominales hasta de cientos de volts. Los valores de la capacitancia varían desde una fracción de microfarad hasta 2 000 μF .

Los condensadores de tantalio húmedo presentan dos desventajas principales: una es la posibilidad de fuga del electrolito líquido, altamente corrosivo; la segunda es la necesidad de evitar la aplicación de voltaje inverso incluso por poco tiempo. Cuando fluye corriente inversa se deposita plata sobre el ánodo, lo cual puede provocar un cortocircuito catastrófico. En la actualidad existen en el mercado unidades con envoltura de tantalio con las que se evita este problema.

El condensador de tantalio sólido es parecido a la versión húmeda, en cuanto a sus etapas iniciales de manufactura. Una vez que se ha anodizado la pílora para formar la capa dieléctrica, se aplican varias capas de material catódico. La pílora anodizada se sumerge en una solución de nitrato de manganeso, que se transforma pirolíticamente en dióxido de manganeso. Esta etapa se repite varias veces a fin de producir un buen recubrimiento del dióxido de manganeso, que es un conductor razonablemente bueno y sirve como primera capa del cátodo. La etapa siguiente consiste en la aplicación de varias capas de grafito coloidal, seguida de pintura con plata, a fin de establecer una buena conexión con el cátodo. El conjunto completo se suelda dentro de un envase de metal, o bien se sumerge en resinas epóxicas una vez que se le han unido las puntas. Los condensadores de tantalio sólido se emplean para aplicaciones de filtrado de baja frecuencia, derivación y acoplamiento, en las que son muy importantes una gran capacitancia y pequeño tamaño de la envoltura. Existen en el mercado en valores que van de 0.0047 a 1 000 μF , con voltajes hasta de 125 V. Ningún electrolítico supera su estabilidad y confiabilidad. No hay líquido que se evapore, y el electrolito sólido es estable. La variación de la capacitancia es muy pequeña: $\pm 10\%$ respecto de su valor a temperatura ambiente en todo el intervalo de temperatura desde -55 hasta 125°C . Por desgracia, ni el electrolítico ni el dieléctrico presentan las cualidades de autorreparación asociadas con otros capacitores electrolíticos. En lugares con defectos pequeños, el MnO_2 pueden convertirse en Mn_2O_3 o Mn_3O_4 , que son aislantes, aislando de esta forma cualquier defecto, lo cual produce un decremento de la corriente de fuga con el uso. Sin embargo, pequeños defectos pueden conducir a falla catastrófica por calentamiento local, y los condensadores de tantalio sólido deben protegerse de situaciones en las que pueden ocurrir altas irrupciones de corriente. Esto ha conducido a la práctica recomendada de insertar una resistencia en serie de $3\Omega/\text{V}$ de voltaje nominal. También excluye el empleo de condensadores de tantalio sólido para la mayor parte de las aplicaciones de alimentación. Este mismo mecanismo de falla también hace

susceptibles estos condensadores a los altos picos transitorios de voltaje. Las perturbaciones transitorias, que exceden el voltaje de formación, darán por resultado incrementos permanentes de la corriente de fuga y posible falla catastrófica.

Para proteger los condensadores de fallas tempranas debidas a defectos del óxido y del electrolito se recomienda su envejecimiento conectando durante 100 h a voltaje nominal y temperatura máxima, empleando una fuente de energía de baja impedancia. Además, se recomienda que el voltaje de operación no exceda el 60% del voltaje nominal.

12.2.5 Capacitores de cerámica

Aproximadamente las tres cuartas partes de los capacitores vendidos en los Estados Unidos son de cerámica. Este extenso uso se debe a su bajo costo, reducido tamaño, amplio intervalo de valor de capacitancia y aplicabilidad general en la electrónica. Los capacitores de cerámica son particularmente idóneos para aplicaciones de filtrado, derivación y acoplamiento de circuitos híbridos integrados, en las que es posible tolerar considerables cambios en la capacitancia.

Los capacitores de cerámica se elaboran en forma de disco, como capacitores de capas múltiples o monolíticos, o en forma tubular. El material dieléctrico es principalmente titanato de bario, titanato de calcio o dióxido de titanio con pequeñas cantidades de otros aditivos para obtener las características deseadas. En los capacitores de clase I se emplea titanato de calcio, y éstos se caracterizan por una baja constante dieléctrica (de 6 a 500), buen control de tolerancias, excelente estabilidad, excelentes características de envejecimiento, baja disipación y comportamiento de la capacitancia en función de la temperatura bien controlado. Los capacitores de clase II tienen constantes dieléctricas mucho mayores (de 200 hasta más de 10 000) y, por ello, mucho mayor eficiencia volumétrica. Sin embargo, sus características eléctricas son inferiores.

Los capacitores discoidales se elaboran a partir de polvos molidos para obtener un pequeño tamaño de las partículas. El polvo se comprime en forma de disco delgado que después se somete a cocción a alta temperatura a fin de fundir el material. Los electrodos se imprimen por serigrafía en ambos lados del disco y luego se someten a horneado. Se fijan las puntas y, por último, se encapsula la unidad.

Para fabricar capacitores cerámicos de capas múltiples se vacía una suspensión que consta del polvo dieléctrico, un adhesivo y un disolvente en láminas delgadas sobre una banda de acero inoxidable o una cinta plástica. Después del secado, se imprimen los electrodos en las láminas, que a continuación se apilan y comprimen. Las pilas se cortan en capacitores individuales, se tratan térmicamente y se someten a cocción a alta temperatura. Por último, los electrodos se sujetan en ambos extremos y las partes se encapsulan.

Los capacitores discoidales tienen un intervalo limitado de área del electrodo y de grosor. Por consiguiente, la formulación dieléctrica hace variar sobre todo el intervalo a fin de lograr un amplio intervalo de

valores de capacitancia. Por otra parte, en los capacitores de capas múltiples se emplea un conjunto restringido de dieléctricos. El intervalo de valores de capacitancia se ordena con facilidad por medio de cambios en el área del electrodo y del número de capas en la pila. Pueden emplearse dieléctricos de clase I para valores de capacitancia desde 1 pF hasta aproximadamente 1 nF, con tolerancias tan bajas como $\pm 10\%$ y voltajes nominales hasta el orden de los kilovoltios. Se pretende que la capacitancia sea estable con la temperatura, o bien lograr una característica de temperatura específica para compensar las variaciones de los otros componentes electrónicos. Los coeficientes térmicos varían de P100 (100 ppm / °C positivo) hasta N750 (750 ppm / °C negativo). También se fabrican intervalos extendidos desde N1400 hasta N5600. La especificación del coeficiente térmico cero (NPO) tiene tolerancia de ± 30 ppm / °C. La capacitancia es muy estable en almacenaje o en operación. El factor de disipación para los dieléctricos de clase I es menor del 0.1% en todo el intervalo de temperatura, hasta frecuencias de varias decenas de megahertz. El coeficiente de voltaje es esencialmente cero.

Los dieléctricos de clase II se emplean para aplicaciones generales en las que tiene gran importancia un tamaño pequeño y la estabilidad y la tolerancia son secundarias. Los valores de capacitancia varían de 2 pF a 2 μ F, con capacidades nominales de 50 a 200 V para capacitores de pastilla (chip) de capas múltiples y hasta 1 000 V para otros empaques. Los materiales de clase II son ferroeléctricos que presentan un efecto de envejecimiento; es decir, la capacitancia decrece con el tiempo. La tasa de envejecimiento es máxima para elevados valores de la constante dieléctrica y varía desde 0.8 hasta 8% por década de tiempo medida a partir de una hora después de efectuar un tratamiento a elevada temperatura. La capacitancia puede restaurarse mediante calentamiento del capacitor hasta 150 °C.

El factor de disipación varía del 5 al 10% a -55 °C hasta entre el 1 y 4% a temperatura ambiente, y disminuye lentamente hasta 125 °C. El factor de disipación decrece con un voltaje de CD, pero aumenta con un voltaje de CA. También aumenta con la frecuencia a frecuencias mayores de 1 a 100 kHz.

Los dieléctricos de clase II presentan un pronunciado coeficiente de voltaje. La disminución en capacitancia puede ser hasta del 30% del voltaje nominal a bajas temperaturas. Por otra parte, la aplicación de voltajes de CA incrementa la capacitancia. La variación de la capacitancia con la temperatura es grande y no lineal. La capacitancia suele disminuir tanto a temperaturas muy bajas como a temperaturas muy elevadas; puede reducirse hasta un 80%.

El cambio en la capacitancia con la frecuencia también depende bastante del tipo de dieléctrico. Los materiales de clase II con elevadas constantes dieléctricas presentan un 10% de disminución en capacitancia a una frecuencia de 10 kHz, mientras que para los de bajas constantes dieléctricas la frecuencia correspondiente es superior a 10 MHz. Los materiales de clase I presentan cambios semejantes a frecuencias de cientos de megahertz.

12.2.6 Capacitores de papel y plástico

El papel, el plástico y las combinaciones de ambos se utilizan en una gran variedad de aplicaciones, como filtrado, acoplamiento, derivación, cronometraje y supresión de ruido. Son capaces de funcionar a altas temperaturas, poseen alta resistencia de aislamiento, buena estabilidad y existen en el mercado con tolerancias hasta de un 1%. La propiedad de autorreparación de las películas metálicas es bastante útil en ciertas aplicaciones. La disponibilidad de películas extremadamente delgadas y la gran variedad de materiales proporciona la flexibilidad necesaria para un gran intervalo de aplicaciones.

Estos capacitores se elaboran siguiendo tres técnicas de construcción: hojas arrolladas, película metalizada y película apilada. Un capacitor de hojas arrolladas consta de dos láminas de aluminio, separadas por hojas de dieléctrico y arrolladas formando un cilindro compacto. El contacto con las láminas se establece mediante la inserción de lengüetas durante el devanado o, en los capacitores del tipo de hojas extendidas, permitiendo que las láminas se extiendan más allá del dieléctrico en lados opuestos. Después del devanado se colocan las puntas y se sujetan a las lengüetas, o a los bordes expuestos de las láminas. Si se utilizan hojas de papel, el capacitor suele impregnarse con aceite, resina, cera o algún compuesto sintético. Así se evita la absorción de agua y se reducen el efecto corona y la disrupción a alto voltaje. Los capacitores de película de plástico se impregnan algunas veces, pero con mayor frecuencia son secos. Después se sujetan las puntas y luego el montaje se moldea, encofra, baña o sella en una lata metálica.

Los capacitores de película metalizada se elaboran mediante un depósito al vacío de aluminio sublimado de 0.1 μ m de grosor directamente sobre la película dieléctrica. El patrón es básicamente el de la configuración de hojas extendidas. Después del devanado se establece contacto con cada extremo del rollo por medio de un rocío pulverizado fino de metal fundido, al que por último se sueldan los conductores. Este tipo de construcción reduce el volumen de los capacitores de bajo voltaje y alto valor, y proporciona una propiedad de disrupción de voltaje conocida como autodespejo. Si ocurre un defecto en el dieléctrico, la corriente de descarga a través del defecto genera suficiente calor para vaporizar los delgados electrodos metálicos. Esto aísla el lugar del defecto y permite la restauración para la operación normal. En la construcción con hojas arrolladas los electrodos son mucho más gruesos y una disrupción produce un corto permanente. El autodespejo requiere una energía mínima de 10 a 50 μ J, y los bajos valores de los capacitores de película metalizada empleados en aplicaciones de bajo voltaje y alta impedancia pueden hacer cortocircuito y fallar en vez de despejarse.

El tipo de construcción más reciente es el de película apilada. Se arrollan películas metalizadas sobre un gran cilindro, y luego se cortan en forma de secciones rectangulares. Se efectúan conexiones a fin de alternar los electrodos en los extremos opuestos, lo que da lu-

Tabla 12-6. Propiedades típicas de los dieléctricos para capacitores de película

Dieléctrico	Constante dieléctrica	Factor de disipación a 25 °C (%)	Temperatura máxima (°C)	ΔC^a a -55 °C (%)	ΔC^a a temperatura máxima (%)	Resistencia de aislamiento a 25 °C ($M\Omega \times \mu F$)
Papel	2.2	0.5	125	-10	+4	40 000
Poliéster	3.2	0.4	125	-5	+15	100 000
Policarbonato	2.9	0.1	125	-2	+1	300 000
Poliestireno	2.5	0.05	85	+1	-2	500 000
Polisulfona	3.1	0.1	150	-2	+1	100 000
Polipropileno	2.1	0.05	105	-6	+2	100 000
TFE ^b	2.1	0.02	150	+2	-2	5 000 000

^a ΔC es el cambio a partir del valor a temperatura ambiente; se modifica mediante el empleo de impregnantes.

^b Politetrafluoretileno.

gar a una pila de películas metalizadas conectadas en paralelo. La estructura es semejante a la de un capacitor cerámico de capas múltiples, excepto por el dieléctrico, que es mucho más delgado. Esto compensa la baja constante dieléctrica de los dieléctricos peliculares y hace del capacitor con películas apiladas un contrincente para aplicaciones en las cuales con anterioridad se empleaban los capacitores cerámicos.

En los capacitores de película se emplea una amplia gama de materiales peliculares; en la tabla 12-6 se presentan algunas propiedades de materiales comunes.

Los capacitores de papel y plástico para aplicaciones electrónicas existen en el mercado en tamaños desde 1 nF hasta 10 μF y en capacidades nominales de voltaje desde 50 V hasta varios miles de volts. Las tolerancias son del 10 y 20% para los capacitores de papel, y hasta del 1% para algunos tipos de película. La capacitancia varía con la temperatura de un dieléctrico a otro, tal como se indica en la tabla 12-6.

Los factores de disipación para los capacitores de película son menores del 1% a 25 °C. Las películas cuyos factores de disipación son bajos a temperaturas cercanas a la ambiental suelen tener bajos DF en todo el intervalo de temperatura. Los DF para los dieléctricos de papel y poliéster se incrementan de manera significativa en cualquier extremo y pueden llegar al 2% para ciertos impregnantes.

La resistencia del aislamiento disminuye con la temperatura a partir de los valores que se muestran en la tabla 12-6. La disminución típica es aproximadamente de tres órdenes de magnitud a partir de temperatura ambiente hasta 125 °C.

Los capacitores de papel y plástico pueden emplearse a altas frecuencias, según el tamaño y la longitud de las puntas. La frecuencia de autorresonancia de un capacitor de 1 μF es aproximadamente 1 MHz y aumenta alrededor de un orden de magnitud cuando la capacitancia disminuye en dos órdenes.

12.2.7 Capacitores de mica y vidrio

Los capacitores con dieléctricos de mica y vidrio se aplican cuando se requiere Q alto y excelente estabilidad con respecto a la temperatura, envejecimiento y frecuencia. Se usan en circuitos de acoplamiento, de

rivación y sintonización a alta frecuencia en los que se requieren rendimiento sobresaliente y alta fiabilidad.

Los capacitores de mica se elaboran formando una pantalla de pasta de plata sobre hojas delgadas de mica. Después de someterlas a cocción, las hojas se apilan con láminas, extendidas alternativamente más allá de la mica sobre extremos opuestos de la pila. Las láminas se doblan sobre los extremos, luego se comprimen y se sujetan por medio del montaje de las puntas. Los capacitores se sumergen varias veces en resinas fenólicas o epóxicas, o bien se moldean en sustancias fenólicas o en poliéster.

Los capacitores de vidrio se elaboran a partir de un vidrio de potasa con plomo, que se estira en forma de cinta delgada y después se apila con capas alternadas de hoja de aluminio. Las hojas alternadas se sueldan a las puntas conductoras, se agrega una cubierta de vidrio y todo el montaje se sella al vacío a elevadas temperatura y presión.

Los capacitores de mica existen en el mercado en tamaños desde 1 pF hasta 100 nF con tolerancias hasta del 0.5%. El intervalo de voltajes nominales varía de 50 a 1 000 V, y con construcción especial hasta 35 kV. Los valores de capacitancia para los capacitores de vidrio varían de 0.5 pF a 10 nF con tolerancias del 1 al 10%. Los voltajes nominales son de 300 o 500 V.

Tanto los capacitores de mica como los de vidrio son estables con respecto a la temperatura. Para algunos valores de capacitancia es posible que el coeficiente de temperatura sea cero. El intervalo varía de -20 a ± 200 ppm / °C.

El DF suele estar entre el 0.01 y 0.1% para capacitores de tamaño mediano. Para pequeñas capacitancias puede llegar al 1%; por lo general es más grande para los capacitores de mica que para los de vidrio. El DF aumenta con la temperatura aproximadamente un orden de magnitud sobre todo el intervalo de temperatura, y también aumenta con la frecuencia. La resistencia de aislamiento a temperatura ambiente suele ser de $10^3 M\Omega \times \mu F$ para los capacitores de mica, y aproximadamente de 10^6 para los de vidrio.

Ambos tipos de capacitores pueden operar a alta frecuencia. La frecuencia de autorresonancia es de unos 10 MHz para grandes valores del capacitor y mayor de 100 MHz para valores más pequeños.

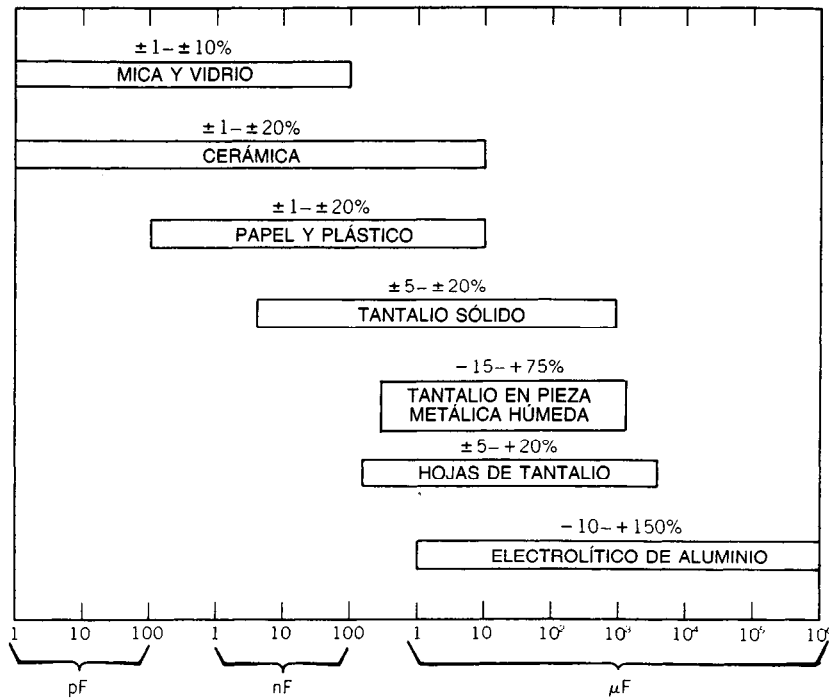


Fig. 12-15. Intervalo de valores de capacitancia y tolerancias para diferentes tipos de capacitores.

12.2.8 Selección de capacitores

El criterio más importante en la elección de capacitores para aplicaciones particulares es el rendimiento, pero también es necesario considerar su disponibilidad y precio. "Disponibilidad bajo pedido" suele significar grandes demoras, cantidades limitadas, inexistencia de una segunda fuente y precios elevados. Las características generales de rendimiento son difíciles de especificar debido a los continuos cambios en las líneas de productos como respuesta a las mejoras tecnológicas y a los cambios en los mercados. También existe una amplia superposición en las especificaciones entre las diversas familias de capacitores. Por estas razones, algunas veces es difícil la elección de capacitores.

En la figura 12-15 se muestra el intervalo de valores de capacitancia y las tolerancias de diferentes clases de capacitores. En los tamaños de intervalo mediano existen varias familias entre las cuales es posible elegir. Sin embargo, factores como corriente de pico, CA ondulatoria y requerimientos de polarización o no polarización pueden limitar la elección.

En la figura 12-16 se presentan los valores nominales de voltaje de diferentes familias. No suele disponerse de grandes valores de capacitancia en los voltajes nominales más elevados, y voltajes elevados pueden implicar grandes dimensiones o empaques diferentes. El voltaje nominal no necesariamente se correlaciona con el voltaje real de disrupción. Valores pequeños de capacitancia requieren áreas de electrodos pequeñas y, por tanto, poco prácticas, a menos que se utilicen múl-

tiples capas de dieléctrico, pero el fabricante puede especificar el mismo voltaje de trabajo que el de otros capacitores en la misma línea de productos.

Por otra parte, debe hacerse hincapié en que la confiabilidad de un capacitor aumenta a medida que se reduce el voltaje.

En la figura 12-17 se muestra el intervalo útil de frecuencias de familias de capacitores. La frecuencia superior está limitada por la frecuencia de autorresonancia (que depende parcialmente de la longitud de la punta), la resistencia equivalente en serie y la disminución en el valor del capacitor.

En la figura 12-18 se muestran variaciones de la capacitancia con la temperatura. Esos valores son sólo representativos, ya que dependen de la magnitud de la capacitancia, voltaje nominal, tipo de electrolito o impregnante, etc.

En la figura 12-19 se presenta la variación del factor de disipación con la temperatura. De nuevo, los valores son representativos para efectos de comparación solamente.

En la tabla 12-7 se resumen las características de los capacitores; en la referencia bibliográfica 3 se ofrecen datos acerca de su fiabilidad, y para información más detallada sobre las características y aplicaciones es necesario consultar las publicaciones de los fabricantes.

12.2.9 Capacitores para circuitos integrados

En los circuitos integrados monolíticos de silicio se emplean tres tipos de capacitores:

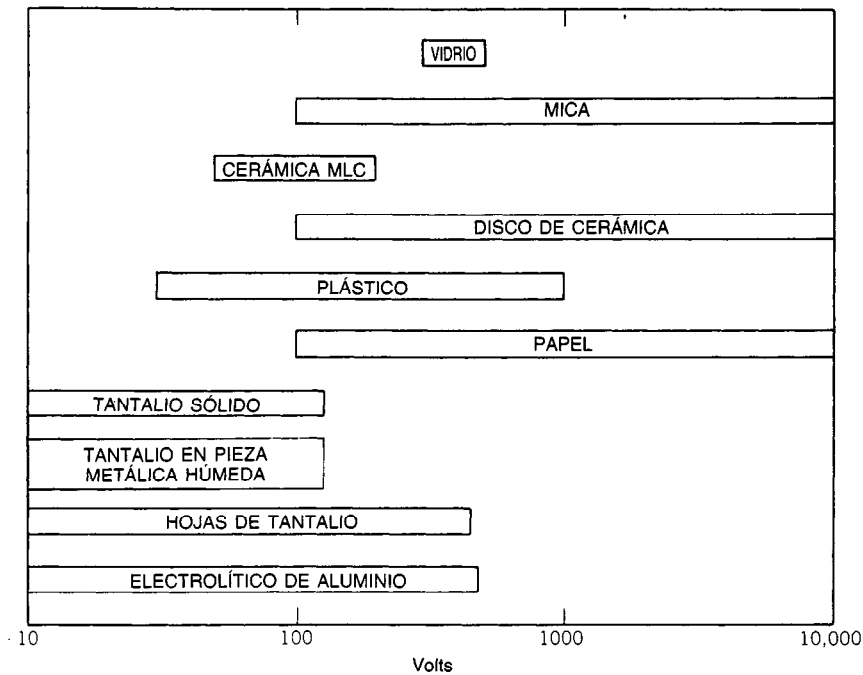


Fig. 12-16. Voltajes nominales para diferentes tipos de capacitores.

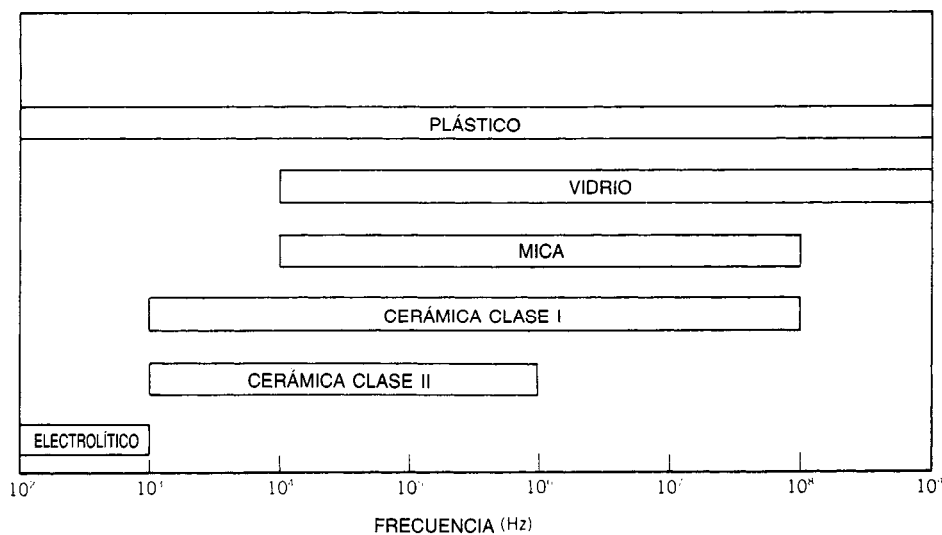


Fig. 12-17. Intervalo de frecuencia útil para familias de capacitores.

1. Los elaborados con una capa altamente difundida, una capa de dióxido de silicio y un electrodo de aluminio.
2. Aquellos en los que se emplea una unión *pn* con polarización inversa.
3. Los que se fundamentan en la capacitancia parásita y en la capacitancia de entrada a la compuerta de los transistores a base de semiconduc-

tores de metal y óxido (MOS, del inglés *metal oxide semiconductor*). Este tipo se utiliza en memorias de semiconductor y en lógica dinámica.

En la figura 12-20 se muestra un corte transversal de un capacitor de óxido. La difusión del emisor tipo *n* forma una placa del capacitor y la metalización de aluminio forma la otra.

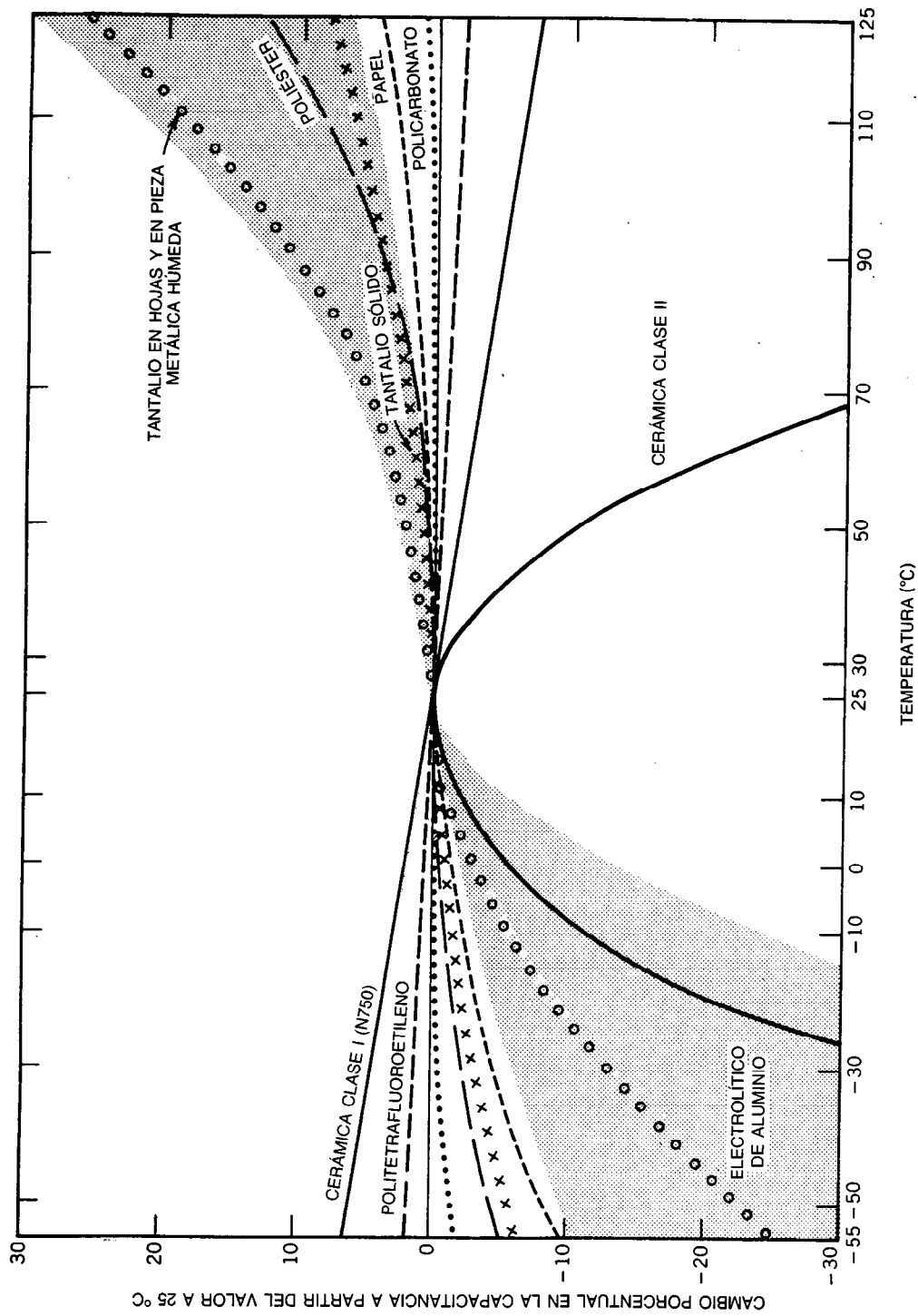


Fig. 12-18. Variación típica de la capacitancia con la temperatura. La parte sombreada muestra la amplia variación que puede ocurrir en los capacitores electrolíticos de aluminio con diferentes valores de capacitancia, voltaje nominal y tipo de construcción. Otros tipos de capacitores presentan variaciones semejantes.

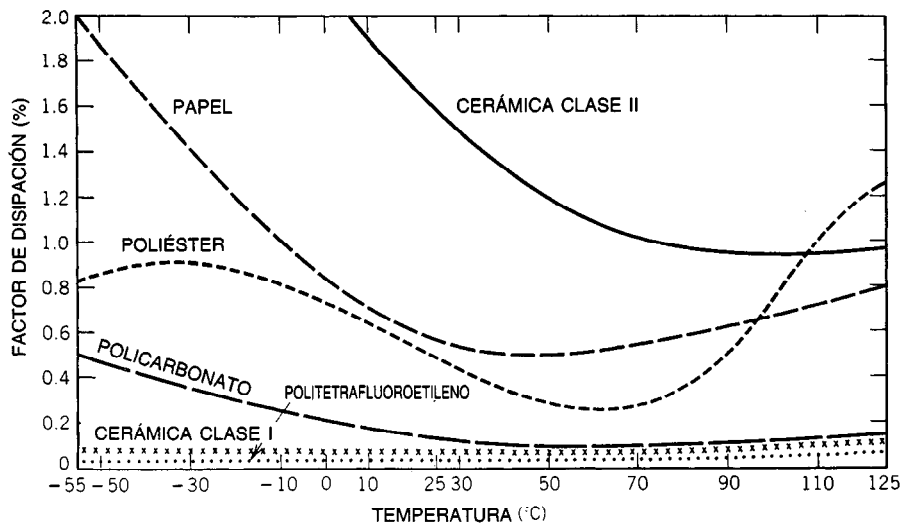


Fig. 12-19. Variación típica del factor de disipación con la temperatura.

A fin de evitar etapas extra de procesamiento, la capa de óxido es la misma que la que cubre las regiones de difusión del emisor. Ocasionalmente se agrega una capa delgada de nitruro de silicio (Si_3N_4) para mejorar la estabilidad y producir una capacitancia algo mayor por unidad de área. La capacitancia está dada por la ecuación 12-8.

La constante dieléctrica para el dióxido de silicio es aproximadamente 3.9, dependiendo de las condiciones de desarrollo; para el nitruro de silicio es aproximadamente 7. Para un grosor típico de óxido emisor de $0.12 \mu\text{m}$, la capacitancia por unidad de área es aproximadamente $3 \times 10^{-4} \text{ pF}/\mu\text{m}^2$.

El capacitor de óxido es aceptable con coeficiente de temperatura de $15 \text{ ppm}/^\circ\text{C}$ aproximadamente, factor de disipación aceptablemente bajo, buenas tolerancias, bajo coeficiente de voltaje y construcción no polarizada (sin embargo, una terminal está conectada al sustrato tipo p mediante una unión pn). La principal desventaja es la gran área requerida para la pastilla, lo que limita los tamaños a aproximadamente 100 pF . La aplicación más importante se dedica a capacitores compensadores de frecuencia en circuitos integrados lineales.

La capacitancia de una unión pn con polarización inversa también está dada por la ecuación 12-8. La

separación d de las placas es el grosor de la capa de agotamiento. Para una unión abrupta pn , d está dada por:

$$d = \left\{ \frac{2\epsilon(\phi/q + V_R)(N_A + N_D)}{qN_A N_D} \right\}^{1/2} \quad (12-11)$$

donde $\epsilon = 11.9$, ϵ_0 es la permitividad del silicio, q la carga del electrón ($1.6 \times 10^{-19} \text{ C}$), V_R el voltaje inverso aplicado y N_A y N_D las concentraciones del aceptor y del donador en cada lado de la unión, respectivamente. ϕ/q es el potencial de la barrera, dado por:

$$\phi/q = \frac{kT}{q} \ln \left(\frac{N_A N_D}{n_i^2} \right) \quad (12-12)$$

k es la constante de Boltzmann, T la temperatura absoluta y n_i la concepción intrínseca de portadores ($1.5 \times 10^{10}/\text{cm}^3$ a temperatura ambiente). ϕ/q es una función de la concentración de impurezas que varía lentamente y tiene un valor aproximado de 0.65 V a temperatura ambiente.

La ecuación 12-11 predice que la capacitancia es proporcional a la raíz cuadrada de la concentración de impurezas. Por tanto, las uniones más contaminadas producen la capacitancia más elevada. La capacitancia

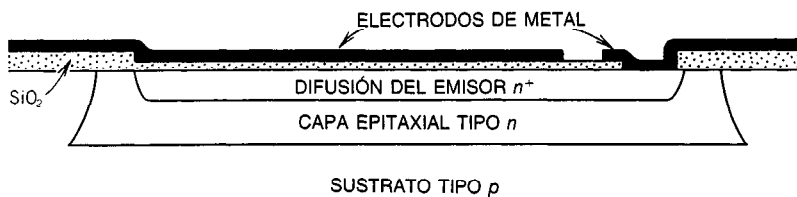


Fig. 12-20. Corte transversal de un capacitor de óxido.

Tabla 12-7. Características de diferentes tipos de capacitores

Dielectrico	Especificación aplicable	Capacitancia			Voltaje nominal de CD (V)
		Intervalo	Tolerancia	Estabilidad después de una prueba de vida de 2 000 h	
Vidrio Fijo	MIL-C-23269(ER)	0.5-10 000 pF	0.25 pF-5%	0.5%-0.5 pF (el que sea mayor)	100, 300 y 500
Variable	MIL-C-14409	0.6-1.8 pF hasta 1-120 pF	Cambio de capacitancia en función de la rotación: <10%	250-1 250
Mica Tipo botón	MIL-C-10950	5-2 400 pF	$\pm 1, \pm 2, \pm 5$ o $\pm 10\%$	$\leq 1\%$ o 0.5 pF (el que sea mayor)	500
De uso general	MIL-C-5	47-27 000 pF	$\pm 1, \pm 2$ o $\pm 5\%$	$\leq 5\%$ o 1 pF (el que sea mayor)	300-2 500
	MIL-C-39001(ER)	1-91 000 pF	0.5 pF- $\pm 5\%$	$\leq 1\%$ o 1 pF (el que sea mayor)	50-500
Electrolítico Aluminio	MIL-C-62	1-1 000 μ F	-10, +50%	$\pm 15\%$	400 y 450
Tantalio (no sólido)	MIL-C-39006(ER)	0.5-1 200 μ F	-15: +30, +50, +75, $\pm 5\%$ a $\pm 20\%$	$\leq 15\%$	6-450
Tantalio (sólido)	MIL-C-39003(ER)	0.0023-330 μ F	$\pm 5, \pm 10$ o $\pm 20\%$	$\leq 2\%$	6-100
Oxido de aluminio	MIL-C-39018	0.68-220 000 μ F	-10: +5, +30, +50, +75%	$\leq 15\%$	5-350
Pastilla de tantalio (sólido)	MIL-C-55365(ER)	0.001-4 700 μ F	$\pm 5, \pm 10$ o $\pm 20\%$	2-50
Papel Impregnado de cera	MIL-C-12889	0.01-0.5 μ F	+0, -30 o +50, -10%	$\leq 10\%$	100: 250 y 500 (AC-CD)

Tabla 12-7. Características de diferentes tipos de capacitores (continuación)

Dieléctrico	Especificación aplicable	Capacitancia			Voltaje nominal de CD (V)
		Intervalo	Tolerancia	Estabilidad después de una prueba de vida de 2 000 h	
Papel-plástico					
Policarbonato	MIL-C-19978(ER)	0.001-1 μ F	± 5 o $\pm 10\%$	$\leq 6\%$	50-600
Metalizado	MIL-C-39022(ER)	0.022-10 μ F	± 5 o $\pm 10\%$	$\leq 10\%$	50-400
Papel y tereftalato de polietileno	MIL-C-19978(ER)	0.001-1 μ F	± 2 , ± 5 o $\pm 10\%$	$\leq 6\%$	200-1 000
Plástico o plástico metalizado	MIL-C-55514(ER)	0.001-10 μ F	± 1 , ± 2 , ± 5 o $\pm 10\%$	$\leq 5\%$	50-600
Tereftalato de polietileno	MIL-C-19978(ER)	0.001-10 μ F	± 2 , ± 5 o $\pm 10\%$	$\geq 6\%$	30-1 000
Cerámica					
Fijo, de uso general	MIL-C-11015	2.2-2 200 000 pF	0.5 pF, ± 10 , $\pm 20\%$	$\leq 10\%$ o $\leq 20\%$	50-1 600
Compensador de temperatura	MIL-C-39014(ER)	10-1 000 000 pF	± 10 , $\pm 20\%$	$\leq 20\%$	50-1 600
Variable	MIL-C-81	1.0-68 000 pF	0.25 pF, ± 1 , ± 2 , ± 5 , $\pm 10\%$	$\pm 3\%$ o 0.5 pF (el que sea mayor)	50-200
		1.5-7 hasta 15-60 pF	200-500
Fijo, pastilla	MIL-C-55681(ER)	10-180 000 pF	± 5 , ± 10 o $\pm 20\%$...	50 y 100
Gas o vacío					
Fijo	MIL-C-23183	6-1 000 pF	$\pm 10\%$...	30 kV
Variable	MIL-C-23183	5-3 000 pF	$\pm 10\%$...	2 y 3 kV

Tabla 12-7. Características de diferentes tipos de capacitores (continuación)

Temperatura		Costo relativo para C, V nominales equivalentes ^a	Tamaño relativo		Factor de disipación (%)		
Temperatura de operación (°C)	Coefficiente térmico (% o ppm / °C)		Varía con	Para C, V nominales equivalentes	60 Hz	1 000 Hz	1 MHz
-55-+125	140, 105, 0 ± 25	Medio	CV ²	Grande	...	<0.001	...
-55-+125 o +150	±50, ±100, ±150 y +50, -0	Medio alto	C	Grande
-55-+85 o +150	±100, -20-+100, y no especificado	Medio alto	CV ²	Grande	...	<0.17	<1.2
-55-+125 o +150	0-+70, -20-+100, ±100, ±200 y no especificado	Medio bajo	CV ²	Grande	...	<0.18	<0.12
-55-+125 o +150	0-+70, -20, -+100, ±200, y no especificado	Medio bajo	CV ²	Grande	...	<0.1	<1 000
-40-+85	La capacitancia cae desde 30-60% a -40 °C	Medio	CV	Muy pequeño	...	15 a 18% a 120 Hz; varía con V	...
-55-+85, desclasificado a +125	La capacitancia cae desde 12-50% a -55°C	Alto	CV	Muy pequeño	...	10-32% a 120 Hz; y varía con C y V	...
-55-+85, desclasificado a +125	La capacitancia cae 10% máx. a -55 °C	Medio	CV	Muy pequeño	...	3-8% a 120 Hz; varía con V	...
-40-+85, desclasificado a +125	...	Medio	CV	Muy pequeño	...	10-35% a 120 Hz; y varía con C y V	...
-55-+125	...	Medio	CV	Muy pequeño	...	0.6-15% a 120 Hz; varía con V	...
-55-+85	La capacitancia cae <30% a -55°C	Medio	CV ²	Pequeño	...	<1	...

Tabla 12-7. Características de diferentes tipos de capacitores (continuación)

Temperatura		Costo relativo para C, V nominales equivalentes ^a	Tamaño relativo		Factor de disipación (%)		
Temperatura de operación (°C)	Coefficiente térmico (% o ppm / °C)		Varía con	Para C, V nominales equivalentes	60 Hz	1 000 Hz	1 MHz
-55—+85 o + 125	La capacitancia cae < 10% a -55 °C	Medio	CV ²	Pequeño	...	<1	...
-55—+125	La capacitancia cambia ±2% a -55 °C	Alto	CV ²	Grande	<1	<0.1	Mayor
-65—+125	±10 %	Alto	CV ²	Medio grande	<0.1	<0.1	...
-55—+85 o + 125	...	Medio	CV ²	Pequeño	...	<2.0	...
-65—+85	-7 a +5%	Alto	CV ²	Pequeño	<0.6	<0.6	...
-55—+ 85 o + 125	La capacitancia cambia <+30, -80% a -55 °C	Muy bajo	CV ² + k	Pequeño	...	<2.5	<2.5
-55—+ 85 o + 125	La capacitancia cambia <+30, -80% a -55 °C	Muy bajo	CV ² + k	Pequeño	...	<2.5	<2.5
-55—+ 125	0. ± 30, 0 ± 60	Muy bajo	CV ² + k	Pequeño	...	0.15	<0.10
-55—+85	La capacitancia cambia <-4.5, + 2% a -55 °C	Media baja	CV ² + k	Grande	0.2
-55—+ 125	0 ± 30 o 0 ± 15	Bajo	CV	Pequeño	...	<2.5	<2.5
-55—+ 85	...	Alto	...	Grande	...	<0.001	...
-55—+ 85	...	Alto	...	Grande	...	<0.001	...

Fuente: Reproducido de MIL-STD-198D
^a"C" = capacitancia y "V" = voltaje.

también es inversamente proporcional a la raíz cuadrada del voltaje inverso para $V_R > \phi/q$, una característica que se utiliza en la sintonización electrónica controlada por voltaje. Los valores típicos de la capacitancia por unidad de área varían de 1×10^{-3} a 1×10^{-4} pF/ μm^2 aproximadamente.

Por desgracia, estos capacitores tienen altos factores de disipación, y además sólo es posible operarlos con polarización inversa. Asimismo, la capacitancia está limitada a pequeños valores para evitar grandes áreas de unión. También debe observarse que las uniones en los circuitos integrados son uniones difundidas, cuyo comportamiento no es necesariamente aproximado por la ecuación 12-11. Para obtener expresiones más exactas es necesario consultar textos estándares de física de semiconductores.

Los capacitores para circuitos integrados de película delgada se elaboran con monóxido de silicio u óxido de tantalio. El monóxido de silicio ($\epsilon_r = 6$) se evapora sobre una capa de aluminio que es el electrodo inferior. Una segunda capa de aluminio en la parte superior del dieléctrico funciona como el otro electrodo. Los capacitores de tantalio se elaboran por depósito electrónico de una capa de tantalio puro sobre un sustrato. Esta capa puede emplearse también para resistores de película delgada. Después, la capa de tantalio se anodiza parcialmente para formar Ta_2O_5 ($\epsilon_r = 22$). Una segunda capa de metal se evapora para formar el otro electrodo y el patrón de interconexión. El capacitor así formado se polariza, y la capa inferior debe tener polaridad positiva con respecto a la de la parte superior.

Los capacitores de película delgada tienen capacitancia por unidad de área de 10^{-3} a 10^{-4} pF/ μm^2 , coeficiente de temperatura de 100 a 200 ppm / °C, y factor de disipación de 0.001 a 0.01. Aunque la tolerancia inicial es del 10% aproximadamente, los valores pueden reducirse a menos del 1%.

También existen en el mercado diversas composiciones dieléctricas para la impresión serigráfica de capacitores de película gruesa. Sin embargo, el grosor del dieléctrico es demasiado grande para altos valores de la capacitancia, a menos que se utilicen materiales con altas constantes dieléctricas. Suelen utilizarse más los capacitores cerámicos discretos de capas múltiples o de pastilla de tantalio.

REFERENCIAS BIBLIOGRÁFICAS

1. "Selection and use of Capacitors", MIL-STD-198D.
2. *Parts Application and Reliability Information Manual for Navy Electronic Equipment*, NAVSEA 0967-LP-597-1011.
3. "Reliability Prediction of Electronic Equipment", MIL-HDBK-217C.
4. Gerald L. Ginsburg, *A User's Guide to Selecting Electronic Components*, Wiley, Nueva York.
5. Charles A. Harper, ed., *Handbook of Components for Electronics*, McGraw-Hill, Nueva York.
6. *Component Technology and Standardization Manuals*, General Electric Co.

12.3 INDUCTORES

Joseph Scaturro

12.3.1 Introducción

De forma general, la **inductancia** puede definirse como la propiedad de un elemento de un circuito mediante la cual es posible almacenar energía en un campo magnético. Sin embargo, sólo tiene importancia en un circuito eléctrico cuando la corriente cambia con respecto al tiempo. Cuando la corriente aumenta o disminuye, el efecto que se opone a este cambio se denomina inductancia (L) o autoinductancia (L_s). Por consiguiente, la inductancia es provocada por un campo magnético cambiante, producido por una corriente cambiante. La inductancia en henrys (H) por vuelta de la bobina, puede expresarse como:

$$\frac{L_s}{N} = \frac{d\phi}{di}$$

donde ϕ = flujo magnético, webers

i = corriente, amperes

N = número de vueltas

En general, los componentes inductivos son únicos en comparación con los resistores y los capacitores, que existen en el mercado como productos estándares, ya que aquellos suelen diseñarse para una aplicación específica. Sin embargo, recientemente ha aparecido en el mercado una amplia variedad de inductores como productos estándares debido a la tendencia a la miniaturización. Los inductores con bajos valores de inducción suelen estar devanados en formas no inductivas (núcleos de aire o fenólicos); los de valores medios, en núcleos de hierro pulverizado; y los de elevados valores, en núcleos de ferrita.

12.3.2 Inductancia de configuraciones sencillas

Un **inductor** ideal es una bobina devanada de alambre sin resistencia (o sin pérdidas). Cuando en el alambre existe movimiento de cargas, la energía se almacena en el campo magnético alrededor de la bobina. Una corriente directa (CD) de estado estable o corriente constante a través del alambre de la bobina produce un voltaje cero a través de la bobina. Una corriente que varíe con respecto al tiempo produce una fuerza electromotriz (fem) o voltaje inducidos. La fem es resultado de la interacción del voltaje, corriente y campo magnético. La expresión para el voltaje es la siguiente:

$$v = L \frac{di}{dt}$$

donde v = voltaje, voltios

i = corriente, amperes

L = inductancia, henrys

t = tiempo, segundos

Obsérvese que un cambio repentino en la corriente en un tiempo cero es imposible en un circuito que contiene inductancia, ya que se requeriría un voltaje in-

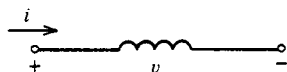


Fig. 12-21. Direcciones de referencia para la corriente y el voltaje de un inductor.

finito. En la figura 12-21 se muestra la representación esquemática de la inductancia. La ecuación dimensional para la inductancia es la siguiente:

$$1H = 1 \frac{V \cdot s}{A}$$

La inductancia también puede expresarse como la pendiente de la curva cuando se representa gráficamente el voltaje en función del cambio en el movimiento de la carga (o aceleración de la carga).

La mayor parte de los inductores constan de una bobina de alambre con un núcleo apropiadamente elegido para la aplicación específica. Sin embargo, todos los circuitos presentan alguna inductancia (inductancia parásita), debido a que el movimiento de la carga a través de cualquier conductor, inclusive un alambre recto, da por resultado un campo magnético a su alrededor. En general, la mayor parte de las configuraciones de bobinas no se prestan a cálculos matemáticos sencillos. Como apoyo se emplean fórmulas empíricas, tablas y nomogramas para determinar la inductancia (o autoinductancia). A continuación se presentan algunas ecuaciones empíricas para configuraciones sencillas:

1. Bobinas de una sola capa o solenoides, donde la longitud es mayor que la mitad del diámetro y el núcleo es de algún material no ferromagnético:

$$L = \frac{\mu_0 N^2 A}{l + 0.45d} \text{ (henrys)}$$

donde N es el número de vueltas, A el área transversal en metros cuadrados, l la longitud en metros, d el diámetro de la bobina en metros, y μ_0 la constante $4\pi \times 10^{-7}$, conocida como permeabilidad del espacio libre.

2. La inductancia a baja frecuencia de la bobina toroidal de N vueltas, con un devanado de sección transversal circular de diámetro d en pulg y radio medio de revolución del núcleo $D/2$ pulg, es

$$L_s = 0.01595N^2[D - (D^2 - d^2)^{1/2}] 10^{-6} \text{ (henrys)}$$

12.3.3 Bobinas con núcleo de aire

Las bobinas con núcleo de aire suelen ser de una sola capa o de múltiples capas devanadas en formas cilíndricas. Algunas veces se emplean toroides, espirales planas y devanados de una o múltiples capas sobre formas en cuadro o hexagonales. La inductancia de muchas bobinas con núcleo de aire puede calcularse mediante fórmulas empíricas en las que participen las dimensiones de la bobina y el número de vueltas (como se vio en el artículo 12.3.2). Estas relaciones pueden

encontrarse en el US Department of Commerce Circular del National Bureau of Standards C74. Sin embargo, la inductancia de una bobina con núcleo de aire a elevadas frecuencias no puede calcularse con exactitud, debido al efecto pelicular y a la capacitancia distribuida de la bobina.

Las bobinas con núcleo de aire de una sola capa (o solenoides) se emplean para circuitos resonantes a frecuencias de transmisión de radio (de 535 kHz a 1 605 kHz) o mayores, y para reactores de radiofrecuencia a frecuencias muy altas (de 30 MHz a 300 MHz). Para grandes valores de inductancia en un pequeño espacio se emplean las bobinas de capas múltiples.

Las pérdidas del núcleo de aire pueden expresarse de varias formas: como la resistencia serie equivalente de la bobina, como el factor de potencia de la bobina, o en términos del factor de calidad, Q .

La resistencia serie equivalente suele ser mayor que la resistencia de CD, excepto a frecuencias extremadamente bajas. A frecuencias superiores la distribución de la corriente sobre la sección transversal del conductor es modificada de tal modo que ocurre una concentración creciente de densidad de corriente en las capas superficiales de los conductores; esto se debe a que la reactancia asociada con las posibles trayectorias de corriente es más baja cerca de la superficie del conductor. Este fenómeno se denomina efecto superficial o pelicular, y es igual a $\sqrt{\sigma\mu f}$, donde σ es la conductividad, f la frecuencia de la corriente y μ la permeabilidad del medio. Esta ecuación determina la profundidad a la que la densidad de la corriente ha decrecido a $1/e$ (37%) de su valor en la superficie del conductor. La proximidad de los conductores en la bobina también puede afectar la distribución de la corriente. Las pérdidas dieléctricas ocurren esencialmente en la forma de la bobina. Pueden existir pérdidas por corrientes parásitas en objetos metálicos circundantes. Los efectos de estas pérdidas a medida que aumenta la frecuencia producen sólo un pequeño cambio en el factor de calidad Q , que se define como la relación de la reactancia inductiva entre la resistencia efectiva ($2\pi fL/R$) con variaciones en la frecuencia.

El factor de potencia de un inductor es la relación de la resistencia entre la impedancia de la bobina:

$$pf = \frac{R}{\sqrt{R^2 + (\omega L)^2}}$$

Dividiendo entre R :

$$pf = \frac{1}{\sqrt{\frac{(\omega L)^2}{R^2} + 1}}$$

Pero $Q = \omega L/R$, y por consiguiente:

$$pf = \frac{1}{\sqrt{Q^2 + 1}}$$

Si $Q > 5$, es posible suponer que el factor de potencia es igual al recíproco de Q .

12.3.4 Materiales magnéticos

En términos de las características magnéticas, la gran mayoría de los materiales son diamagnéticos o paramagnéticos. La permeabilidad relativa de los materiales diamagnéticos, como la plata y el cobre, es ligeramente menor que la unidad. Los materiales paramagnéticos, como el aluminio y el platino, tienen permeabilidad relativa ligeramente superior a la unidad. La permeabilidad relativa (μ_r) es el cociente de la permeabilidad absoluta (μ) entre la permeabilidad del espacio libre (μ_0). Las unidades de permeabilidad son henrys por metro, y en el SI, $\mu_0 = 4\pi \times 10^{-7}$. (Véase también el capítulo 9.)

Los materiales más importantes en la ingeniería eléctrica son los conocidos como ferromagnéticos, que principalmente están compuestos de hierro o sus aleaciones con níquel, cobalto o aluminio. La permeabilidad relativa de los materiales ferromagnéticos es muchas veces mayor que la del espacio libre. Las diferencias en las características magnéticas de los materiales pueden atribuirse a la forma en la cual un campo magnético afecta los spins o giros de los electrones y protones y los movimientos orbitales de los electrones de cada átomo del material. En los materiales ferromagnéticos, los spins se encuentran alineados entre sí de forma paralela. El material se divide en dominios magnéticos, cada uno de los cuales tiene una magnetización neta incluso sin un campo externo. Sin embargo, no existe una magnetización neta global, ya que las magnetizaciones de los diferentes dominios se cancelan mutuamente. La aplicación de un pequeño campo magnético provoca el desarrollo de dominios favorables, lo que da por resultado una alta magnetización. Estos materiales se transforman en paramagnéticos por encima de una temperatura crítica conocida como temperatura de Curie.

La permeabilidad de los materiales ferromagnéticos no es constante. La **permeabilidad** es la pendiente de la curva normal de magnetización, que es una gráfica de la densidad de flujo magnético (B) en función de la intensidad magnética (H). Esta curva no es lineal, ya que en los materiales ferromagnéticos existen condiciones de saturación. Las pérdidas más sobresalientes que ocurren en los materiales ferromagnéticos son las producidas por histéresis y por corrientes parásitas. Muchas aplicaciones implican el empleo de CA. El flujo magnético variante que resulta induce fuerzas electromotrices en el material, lo que a su vez crea corrientes parásitas. El calentamiento debido a estas corrientes es una pérdida de energía, denominada pérdida por corrientes parásitas, que puede reducirse mediante el empleo de hojas laminadas para los núcleos, a fin de confinar las corrientes dentro de las resistividades eléctricas relativamente altas. A medida que el material se magnetiza durante cada medio ciclo, la cantidad de energía almacenada en el campo magnético excede la que es liberada en la desmagnetización. La diferencia entre las dos densidades de energía puede determinarse integrando las áreas de un ciclo de histéresis que representan la energía liberada. La diferencia representa la energía que no regresa a la fuente, sino que se

disipa como calor a medida que los dominios se realinean en respuesta a la intensidad cambiante del campo magnético. Esta disipación de energía se conoce como pérdida por histéresis. La suma de todas las pérdidas en un núcleo se denomina pérdida total en el núcleo.

Los materiales magnéticos pueden clasificarse en materiales "blandos" (no retentivos o de baja remanencia) y materiales "duros" (retentivos o de gran remanencia). Aquellos son materiales de baja pérdida empleados en núcleos de motores, generadores y transformadores, dispositivos electromagnéticos y núcleos de memoria. El hierro no aleado se utiliza en aplicaciones de baja frecuencia, mientras que las ferritas, el permalloy, el hierro al silicio y los granates en las aplicaciones de alta frecuencia. Los hierros purificados se emplean en aplicaciones de CD tales como relevadores y otros núcleos electromagnéticos.

Algunos fabricantes de materiales magnéticos son: Allen-Bradley Co., 1201 S. Second St., Milwaukee, WI 53204; Ceramic Magnetics, Inc., 87 Fairfield Rd., Fairfield, NJ 07006; Fair-Rite Products Corp., Wallkill, NJ 12589; Ferronics Inc., 69 North Lincoln Rd., East Rochester, NY 14445; Ferroxcube Div., Ampere Electronic Corp., 5083 Kings Hwy., Saugerties, NY 12477; Indiana General Div., EM&M Corp., Crows Mill Rd., Kensby, NJ 08832; Krystinel Corp., 126 Pennsylvania Ave., Paterson, NJ 07509; Magnetics Div., Spang Industries, PO Box 391, Butler, PA 16001; y Stackpole Carbon Co., Electronic Components Div., Stackpole St., St Mary's, PA 15857.

12.3.5 Inductores con núcleo de hierro

Los núcleos de hierro tienen mayores permeabilidades que los de aire, por lo cual concentran flujo magnético. Este efecto incrementa la inductancia del devanado, lo que reduce el costo, debido a que se reduce el número de vueltas y con ello el tiempo requerido para el devanado a un valor dado de inductancia. También resulta un aumento en el valor de Q , ya que la resistencia es evidentemente menor. Por consiguiente, para aplicaciones de CA en las que se requieren elevados valores de inductancia, los devanados suelen colocarse sobre núcleos laminados de hierro y las vueltas se hacen tan cerca del núcleo como sea posible.

El cálculo de la inductancia para una bobina con núcleo de hierro, incluyendo un entrehierro sujeto a CA, correspondiente a una máxima densidad de flujo (B_M) en el hierro, es:

$$L = \frac{1.256N^2}{\frac{l_i}{\mu\mu_i} + \frac{l_a}{a_a}} \times 10^{-8} \text{ (henrys)}$$

donde N = número de vueltas

l_i = longitud media del circuito magnético en el hierro, cm

l_a = longitud media del circuito magnético en el entrehierro, cm

μ = permeabilidad (H/cm) del material magnético, evaluada a una densidad máxima de flujo alterno idóneo $B_{m\acute{a}x}$

a_i = área transversal (cm^2) del hierro (no se incluye el área del aislamiento entre las laminaciones; en ocasiones se emplea el factor de apilamiento de 0.9, multiplicado por el área verdadera)

a_a = área transversal efectiva (cm^2) del entrehierro (para entrehierros pequeños, la longitud de éstos se suma a cada dimensión que afecte el área)

Para dimensiones en pulgadas:

$$L = \frac{3.19N^2}{\frac{l_i}{\mu a_i} + \frac{l_a}{a_a}} \times 10^{-8} \text{ (henrys)}$$

Para un entrehierro de 0.1 pulg y un rectángulo de 2×3 pulg, el área efectiva = 2.1×3.1 o bien 6.51 pulg^2 . Este cálculo corrige el efecto del flujo marginal.

Los inductores toroidales producen valores de inductancia mucho mayores que los inductores de conductor axial o del tipo de solenoide. Dado que el núcleo es redondo, la fuga de flujo (flujo en el aire alrededor del núcleo) tiende a minimizarse y los requerimientos de blindaje se reducen. De hecho, el inductor toroidal se considera autoprotegido. Los valores de Q de los toroides suelen ser altos, debido a las bajas pérdidas en el núcleo. Sin embargo, son considerablemente más costosos que los tipos de conductores axiales. Para un toroide de una sola capa de sección rectangular sin entrehierro:

$$L = 0.0020N^2b/\mu_d \ln(r_2/r_1)$$

donde L = inductancia, μH

b = ancho del núcleo, cm

μ_d = permeabilidad incremental promedio

r_1 = radio interior, cm

r_2 = radio exterior, cm

N = número total de vueltas

Algunos fabricantes de inductores son: Dale Electronics Inc., Yankton Div., East Highway 50, Yankton SD 57078; Bourns Inc., Trimpot Products Div., Magnetic Products, 28151 Hwy. 74, Romoland, CA 92380; Delevan Div., American Precision Industries, Inc., 270 Quaker Rd., East Aurora, NY 14052; TRW, VTC Transformers, 150 Varick St., New York, NY 10013; Litton TRIADUTRAD, 305 No. Briant St., Huntington, IN 46750; Thordarson, Electronic Center, Mt. Carmel, IL 62863; Mecrotran Co. Inc., PO Box 236, Valley Stream, NY 11582.

12.3.6 Inductores para aplicaciones en circuitos integrados

La exactitud de fabricación de los componentes integrados es del orden del 10%. Sin embargo, es posible mantener relaciones entre diversos componentes hasta el 3% aproximadamente. Por ejemplo, si deben elaborarse dos resistores y la relación de resistencia debe ser 4:1, esta relación puede obtenerse con un 3% de exactitud, aun cuando puede haber errores del 10% en el valor de las resistencias mismas.

La fabricación de inductores integrados no ha sido satisfactoria. Una de las limitaciones de la tecnología de los circuitos integrados es la carencia de inductores integrados, por consiguiente, siempre que es posible se evitan. En muchos casos es posible eliminar la necesidad de elementos inductivos mediante el empleo de una técnica conocida como síntesis RC. Si se requieren inductores con Q mayor de 5 μH , se utilizan inductores discretos y se conectan de manera externa con la pastilla de silicio. El tamaño físico de estos inductores suele ser mucho mayor que el de la pastilla.

La Delevan Division of American Precision Industries, Inc., fabrica microinductores. Los microinductores "i" (marca registrada de Delevan) presentan inductancias de 0.010 μH hasta 1 000 μH en diferentes empaques de bajo perfil (0.075 pulg) para su uso en empaques híbridos de película gruesa. Estas piezas tienen conectores de cinta o zonas terminales para soldadura por reflujo. Bourns Inc. Trimpot Products Division, Magnetic Products, también diseña inductores micro-miniatura en empaques con dimensiones desde 0.125 \times 0.125 \times 0.125 hasta 0.310 \times 0.410 \times 0.465 pulg, con contactos de circuitos integrados o conectores de cinta.

12.3.7 Evaluación de la aplicación de un inductor

Para evaluar la aplicación de un inductor se requiere la siguiente información:

1. El valor de la inductancia y la tolerancia requerida.
2. Las dimensiones y tolerancias mecánicas.
3. El Q deseado.
4. La corriente nominal del inductor.
5. La frecuencia a la cual operará el inductor.

Con respecto al punto 3, a medida que Q sea mayor, el rendimiento del inductor será más alto. Dado que en la mayor parte de los casos es aconsejable un elevado valor de Q , este valor suele expresarse como un mínimo. Con respecto al punto 5, cada inductor posee una frecuencia por encima de la cual dejará de funcionar de la manera deseada. Se comporta como un capacitor por encima de la frecuencia de autorresonancia (SRF, del inglés *self-resonant frequency*). Por consiguiente, la SRF suele especificarse como un mínimo. En la producción de un inductor la SRF se mantiene al máximo nivel alcanzable, con lo cual se amplía el intervalo de frecuencias utilizables del inductor.

LIBROS Y REVISTAS

- R. B. Angus, *Electrical Engineering Fundamentals*, 2.^a ed., Wesley, Reading, MA.
- D. G. Fink y cols., eds., *Electronics Engineers Handbook* 2.^a ed., McGraw-Hill, Nueva York.
- F. Langford-Smith, ed., *Radiotron Designers Handbook*, 4.^a ed., Wireless Press, Harrison, NJ.
- Reproducido y distribuido por RCA Victor Division, Radio Corporation of America.
- F. E. Terman, *Radio Engineers Handbook*, 3.^a ed., McGraw-Hill, Nueva York.

12.4 TRANSFORMADORES

y

Joseph Scaturo

12.4.1 Circuitos acoplados

El término "circuitos acoplados" suele referirse casi siempre al acoplamiento magnético; es decir, al efecto de interacción magnética entre dos partes de un circuito o entre dos circuitos ajenos y diferentes. En 1830 Faraday y Henry observaron que, si se colocaba una bobina de alambre en la vecindad de otra bobina conductora de una corriente variable, entonces en la primera se inducía un voltaje. Este fenómeno puede expresarse de la siguiente forma:

$$v = N \frac{d\phi}{dt}$$

Esta ecuación se conoce como **ley de Faraday** y establece que, si una bobina de alambre con N vueltas se coloca en un flujo magnético variante en el tiempo ($d\phi/dt$), se induce un voltaje (v) en la bobina. Por ejemplo, un circuito acoplado sencillo puede constar de dos bobinas. Supóngase que el voltaje aplicado a una de las bobinas es sinusoidal, $v = V_m \sin \omega t$. Esta fuente de voltaje produce una corriente con variación sinusoidal que, a su vez, crea un flujo magnético variante en el tiempo. Una parte del flujo de esta bobina pasa a través de la otra bobina o enlaza con ella. Dado que el flujo varía con el tiempo, sobre esa bobina se induce un voltaje. Este voltaje inducido está en función del número de vueltas de la bobina y de la cantidad de flujo variante. Debe tenerse presente que también se induce un voltaje en la bobina a la que se aplica el voltaje.

El voltaje autoinducido se opone al voltaje aplicado o de excitación (ley de Lenz).

La inductancia de la bobina debida a la corriente variable en la otra bobina se denomina inductancia mutua. De nuevo, supóngase un circuito sencillo acoplado que consta de dos bobinas, tal como se muestra en la figura 12-22. Conforme a la figura, si:

$$v = V_m \sin \omega t$$

entonces:

$$L_1 = N_1 \frac{d\phi_1}{di_1}$$

$$L^2 = N_2 \frac{d\phi_2}{di_2}$$

donde L_1 y L_2 = autoinductancia de cada bobina, H

ϕ = flujo, Wb

N = número de vueltas

i = corriente, A

Además:

$$M_{21} = N_2 \frac{d\phi_{21}}{di_1}$$

donde M_{21} es la inductancia mutua de la bobina 2, debida a la corriente de la bobina 1 y:

$$M_{12} = N_1 \frac{d\phi_{12}}{di_2}$$

donde M_{12} es la inductancia mutua de la bobina 1, debida a la corriente de la bobina 2.

Cuando la permeabilidad del medio es constante, entonces:

$$M_{21} = M_{12} = M$$

Las bobinas pueden estar estrechamente acopladas o un poco sueltas. Por ejemplo, dos bobinas con núcleo de aire, colocadas entre sí a unos cuantos centímetros de separación, se encuentran acopladas de manera suelta, debido a que sólo una pequeña parte del flujo de la bobina 1 enlazará con la bobina 2. Si están devanadas una sobre la otra o si se devanan juntas sobre un material ferromagnético, están estrechamente acopladas.

En los transformadores de potencia comerciales, las bobinas están acopladas tan estrechamente que más del 98 % del flujo generado en una bobina enlaza con la otra. El coeficiente de acoplamiento K , que establece la proximidad magnética de dos bobinas, se define como:

$$K = \frac{M}{\sqrt{L_1 L_2}}$$

donde M es la inductancia mutua de las bobinas en henrys y L_1 y L_2 son autoinductancias. Un coeficiente de acoplamiento igual a la unidad indica acoplamiento máximo; como ya se ha dicho, para transformadores

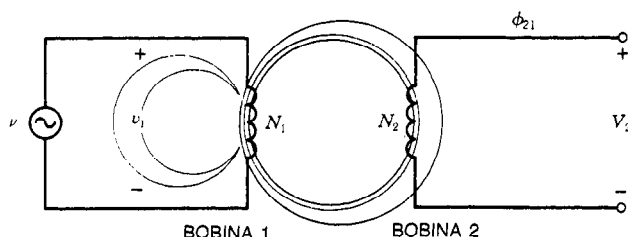


Fig. 12-22. Bobinas acopladas.

de potencia se tiene $K \geq 0.98$. Para bobinas acopladas un poco sueltas, como suele ser el caso de los radios, K puede ser menor o igual que 0.1. El grado deseado de acoplamiento depende de las necesidades particulares en un circuito dado.

12.4.2 El transformador ideal

Un **transformador** es un dispositivo que transfiere energía de un circuito a otro mediante inducción electromagnética. La inductancia mutua, previamente mencionada, describe el principio básico implicado; es decir, los circuitos mantienen su acoplamiento de tal modo que cualquier cambio de corriente en la primera bobina, o bobina primaria, provoca un cambio de flujo que induce un voltaje en la segunda bobina, o bobina secundaria. Cuando se conecta una carga a la segunda bobina, este voltaje del secundario provoca una corriente de carga, o corriente secundaria, que a su vez crea un contraflujo que provoca el incremento de la corriente de la primera bobina en un intento de proporcionar más flujo. Esta acción, denominada acción de transformador, provoca el paso de energía del primario al secundario a través del medio del campo magnético cambiante. Es posible utilizar un núcleo ferromagnético para obtener un acoplamiento más estrecho.

El análisis del transformador es una tarea difícil, a menos que se hagan suposiciones o idealizaciones iniciales. El modelo del circuito real o circuito equivalente debe incluir la resistencia de las bobinas, la fuga de flujo, las pérdidas en el núcleo y el hecho de que el núcleo tiene permeabilidad finita. Por consiguiente, las suposiciones para el transformador ideal son:

1. La resistencia de los devanados es despreciable.
2. La energía perdida o la pérdida en el núcleo son despreciables.
3. Todo el flujo enlaza cada vuelta en cada bobina, o la fuga de flujo es cero.
4. El material del núcleo tiene permeabilidad infinita, o la curva de magnetización (B contra H) del material del núcleo es el eje B (un cortocircuito magnético).
5. Los efectos de la capacitancia resultan despreciables.

El circuito que represente el transformador ideal puede dibujarse como se muestra en la figura 12-23. El voltaje secundario inducido es producido por el mismo flujo que el voltaje primario inducido y está en fase con el voltaje primario, o 180° fuera de fase con él, dependiendo de la manera en que las bobinas se encuentren devanadas sobre el núcleo.

Para ser específicos en este sentido, se colocan puntos cerca de las terminales del transformador en los diagramas. Los puntos indican que los aumentos en el voltaje inducido en los dos devanados están en fase cuando ambos se definen como incrementos desde la terminal sin marcar hasta la terminal señalada con el punto en cada bobina.

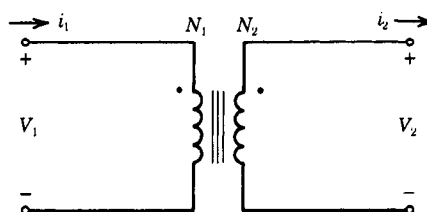


Fig. 12-23. Transformador ideal.

Para analizar la acción del transformador ideal se aplica la ley de conservación de la energía y se extiende para incluir la reactancia, así como la resistencia:

$$v_1 i_1 = v_2 i_2$$

$$v_1 = N_1 \frac{d\phi}{dt} \quad \text{y} \quad v_2 = N_2 \frac{d\phi}{dt}$$

$$\frac{v_1}{v_2} = \frac{N_1(d\phi/dt)}{N_2(d\phi/dt)}$$

o bien:

$$\frac{v_1}{v_2} = \frac{N_1}{N_2} = a$$

donde a se denomina razón de transformación. Por consiguiente, para el transformador ideal el voltaje a través de los devanados es proporcional al número de vueltas del devanado.

Dado que:

$$v_1 i_1 = v_2 i_2$$

entonces:

$$\frac{v_1}{v_2} = \frac{i_2}{i_1} = \frac{N_1}{N_2}$$

o bien:

$$\frac{i_1}{i_2} = \frac{N_2}{N_1} = \frac{1}{a}$$

Es decir, la corriente a través de los devanados es inversamente proporcional al número de vueltas.

Si se conecta una impedancia Z_2 a través de las terminales del secundario, entonces:

$$Z_2 = \frac{v_2}{i_2}$$

pero:

$$Z_1 = \frac{v_1}{i_1}$$

$$\frac{Z_1}{Z_2} = \frac{v_1/i_1}{v_2/i_2} = \frac{v_1}{v_2} \frac{i_2}{i_1} = \frac{N_1}{N_2} \frac{N_1}{N_2}$$

o bien:

$$\frac{Z_1}{Z_2} = \left(\frac{N_1}{N_2}\right)^2 \text{ o bien } a^2$$

Es decir, la impedancia de carga aparente o equivalente vista desde (o reflejada por) los devanados del transformador es proporcional al cuadrado de la razón de transformación (o razón de vueltas).

12.4.3 Transformadores sin pérdidas

Las pérdidas de potencia de los transformadores reales consisten en pérdidas óhmicas o en el devanado, debidas a la resistencia de cada devanado y en pérdidas en el núcleo cuando el material del núcleo es ferromagnético, que producen un calentamiento del material ferromagnético. Por consiguiente, es posible simularlas mediante un valor idóneo de resistencia en el modelo del circuito (o circuito equivalente). Los transformadores reales también difieren de los ideales en que el flujo producido por un devanado no se acopla con todas las vueltas del segundo devanado. Es decir, existe fuga de flujo principalmente en el aire, que puede simularse mediante un valor apropiado de inductancia. Además, los núcleos tienen permeabilidad finita, por lo que se requiere una corriente para producir una fuerza electromotriz a fin de establecer un flujo. La relación entre la corriente que produce el flujo y el voltaje inducido del primero es de 90° , donde la corriente queda detrás del voltaje. Por tanto, el elemento que simula este efecto es una inductancia en derivación con el devanado primario (L_m).

Si se desprecian las pérdidas de potencia, el modelo circuital para el transformador sin pérdidas, que incluye el transformador ideal, puede ser como se muestra en la figura 12-24.

12.4.4 Transformadores electrónicos

Transformadores de potencia

Los transformadores electrónicos de potencia suelen operar a una sola frecuencia. Por lo general, las frecuencias son 50, 60 o 400 Hz. En Europa, 50 Hz es lo común; 400 Hz es la frecuencia de alimentación de mayor empleo en aeronáutica.

Sin embargo, los futuros transformadores aeronáuticos operarán a mayores frecuencias a fin de reducir peso y tamaño.

Las consideraciones para el diseño de transformadores de potencia están controladas por:

1. **Eficiencia.** Ésta es la relación entre la potencia de salida y la potencia de entrada. El factor de calidad de los transformadores de potencia está en función de las pérdidas del núcleo y de las pérdidas óhmicas (por efecto Joule o en el cobre):

$$\% \text{ Ef} = \frac{P_o}{P_i} (100)$$

Las eficiencias típicas varían del 70 al 98%. Los grandes transformadores suelen tener eficiencias superiores.

2. **Factor de potencia.** El factor de potencia reviste particular interés en grandes transformadores, en los que se consume una gran cantidad de potencia. El **factor de potencia** es sencillamente el coseno del ángulo de fase o la razón de la potencia verdadera o potencia real disipada (watts), entre la potencia aparente (volt-amperes):

$$\text{Factor de potencia (PF)} = \cos \theta =$$

$$= \frac{\text{potencia disipada} \left(\frac{\text{W}}{\text{VA}} \right)}{\text{potencia aparente}}$$

3. **Incremento de temperatura.** Ésta es una especificación importante, ya que estipula la temperatura de operación del dispositivo. Las pérdidas de potencia provocan aumento de temperatura.
4. **Regulación del voltaje.** Esta regulación se define como el cambio en magnitud del voltaje secundario a medida que la corriente cambia desde carga cero hasta carga total, mientras se mantiene fijo el voltaje primario. Por consiguiente, cuando la carga es cambiante, se espera que el voltaje a través de la carga permanezca dentro de ciertos límites, y la regulación se convierte en un factor importante en el diseño.
5. **Desfase.** Éste es un factor importante en el diseño de transformadores de referencia. El desfase permisible para condiciones específicas se establece en muchas especificaciones. Es una función de la resistencia de CD del devanado primario, la inductancia de fuga, y la impedancia del generador de activación. Por lo general, es deseable un desfase mínimo.

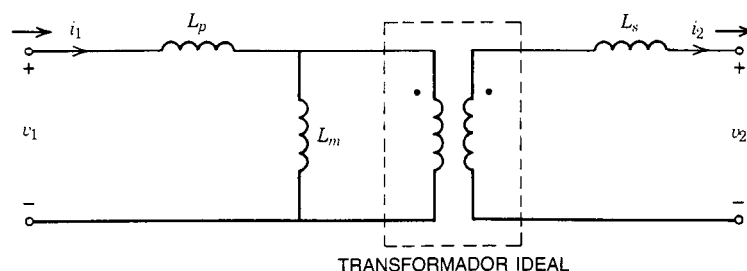


Fig. 12-24. Modelo de transformador sin pérdidas.

6. **Corriente de irrupción.** Ésta es una función de la conmutación o el valor del voltaje. Pueden producirse corrientes pico varias veces mayores que la corriente normal de excitación. Esta corriente suele tener naturaleza transitoria y dura sólo unos cuantos ciclos de la frecuencia de potencia. Después se estabiliza en la corriente normal de excitación.
7. **Inductancia de fuga.** Como ya se ha dicho, la inductancia de fuga no representa una pérdida de potencia. Consta de líneas de fuerza magnética que no cortan o acoplan ninguna vuelta del devanado, por lo que no producen un voltaje utilizable. La inductancia de fuga afecta la regulación de voltaje. Mientras mayor sea dicha inductancia, más deficiente será la regulación.
8. **Pérdidas en el núcleo y efecto corona.** Estos factores también deben considerarse en el diseño de transformadores (art. 12.4.5).

Transformadores de audio

Estos transformadores se emplean para el acoplamiento de voltajes, corrientes, e impedancias sobre un intervalo de frecuencia nominal de 20 a 20 000 Hz. También pueden utilizarse para otras funciones auxiliares, tales como proporcionar una trayectoria para la corriente directa a través del primario, al mismo tiempo que la mantienen fuera del circuito secundario.

La primera consideración para el diseño, tamaño y costo de los transformadores de audio es el intervalo de la respuesta de frecuencia. En transformadores de audio de banda ancha, la cobertura de frecuencias suele separarse en tres intervalos independientes para fines de análisis. En el intervalo de alta frecuencia, la inductancia de fuga y la capacitancia distribuida (C_p , derivación primaria y capacitancia distribuida; C_s , derivación secundaria y capacitancia distribuida; y C_{ps} , capacitancia de primario a secundario) son las más importantes. En el intervalo de baja frecuencia es importante la inductancia del circuito abierto (L_m). En el intervalo de mediana frecuencia, aproximadamente 1 000 Hz, el efecto del transformador en la frecuencia de respuesta puede despreciarse. El circuito equivalente del transformador de banda amplia con una carga resistiva se observa en la figura 12-25.

Los transformadores de audio miniaturizados poseen excelente respuesta a alta frecuencia, debido a que, a medida que disminuye el tamaño del transformador, la inductancia de fuga y la capacitancia distribuida de los devanados también disminuyen. Sin embargo, el pequeño tamaño de estos transformadores aumenta el costo y la degradación de la respuesta a baja frecuencia (que depende de la inductancia del circuito primario abierto, L_m).

Transformadores de pulsos

El transformador de pulsos (o impulsos) es un componente magnético diseñado para generar o transferir un pulso de energía eléctrica con características eléctricas específicas. Básicamente, el pulso es de alta o baja potencia. Los transformadores de pulsos de baja potencia suelen emplearse en circuitos disparadores (ac-

tivadores o de mando) y de acoplamiento. En un circuito disparador, el transformador de pulsos se emplea ya sea para iniciar otra función o para formar otro pulso en un circuito del tipo de regeneración. En los circuitos de acoplamiento, los transformadores de pulsos se emplean para la optimización del circuito por acoplamiento de impedancias, inversión de fase o aislamiento. Los transformadores de pulsos de alta potencia (pico mayor de 300 W) suelen emplearse en moduladores para sistemas de radar. Su función es el acoplamiento de impedancias entre la red formadora de pulsos y el magnetrón. El principal interés es la transformación del pulso con un mínimo de distorsión.

Las especificaciones para un transformador de pulsos deben contemplar: la amplitud, usualmente expresada en términos de volts dentro de una carga específica; el ancho del pulso; el excedente por arriba, expresado en voltaje o como un porcentaje del pulso al 100%; y la caída, expresada en volts o como un porcentaje del pulso al 100%.

Transformadores de radiofrecuencia (RF) de banda ancha

Los transformadores RF son medios poco costosos, sencillos y compactos para la transformación de impedancia a frecuencias superiores. Los devanados bifilares (una forma de devanado no inductivo) y los núcleos de hierro o ferrita pulverizados proporcionan un acoplamiento óptimo. El empleo de núcleos con alta permeabilidad a frecuencias más bajas reduce el número de vueltas necesarias y la capacitancia distribuida. A frecuencias superiores aumenta la reactancia, aun cuando la permeabilidad puede disminuir.

Transformadores de doble sintonía

El transformador de doble sintonía es una de las configuraciones de circuitos de mayor uso para sistemas de frecuencias intermedias en el intervalo de frecuencias de 250 kHz a 50 MHz. Consta de un primario y un secundario sintonizados a la misma frecuencia y acoplados inductivamente a un grado deseado para la curva de selectividad. Las bobinas pueden estar subacopladas, críticamente acopladas o sobreacopladas.

12.4.5 Pérdidas en los transformadores

Las pérdidas en los transformadores consisten en pérdidas en el devanado u óhmicas, y pérdidas en el núcleo. Las pérdidas en el devanado se deben a la resistencia de los devanados primario y secundario. Dichas pérdidas pueden expresarse como el producto del cuadrado de la corriente primaria y la resistencia equivalente del devanado primario, y como el producto del cuadrado de la corriente secundaria y la resistencia equivalente del devanado secundario, $I_1^2 r_1$ e $I_2^2 r_2$.

Como ya se ha dicho en el artículo 12.3.4, existen dos tipos de pérdidas en el núcleo; por histéresis y por corrientes parásitas. El término histéresis se refiere a la pérdida por calor en el núcleo generado por el movimiento de las moléculas de hierro dentro del material del núcleo. La CA que fluye en el transformador provoca que las moléculas se reorienten a sí mismas du-

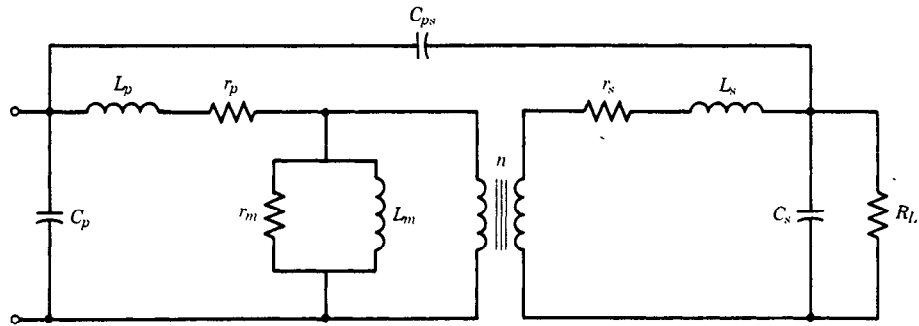


Fig. 12-25. Modelo de transformador de banda ancha.

rante un ciclo de CA. La resistencia que oponen las moléculas a esta reorientación provoca pérdidas en forma de calor. Las corrientes parásitas son las producidas en los materiales de los núcleos que se establecen en ángulos rectos con respecto al campo normal de magnetización. Estas corrientes no cortan las vueltas de la bobina y, por tanto, no contribuyen a los campos normales. La energía utilizada para establecer estas corrientes se pierde en forma de calor. Las laminaciones de materiales ferromagnéticos reducen las corrientes parásitas debido a la resistencia aumentada del núcleo. Un incremento en la frecuencia de potencia requiere que las laminaciones sean más delgadas para mantener las pérdidas en un mínimo. En transformadores de muy alta frecuencia, las laminaciones se convierten en cintas delgadas. Las pérdidas en el núcleo se simulan como una resistencia en derivación a través de la inductancia del transformador.

12.4.6 Efecto corona

El efecto corona es un fenómeno de alto voltaje en el que el aire que rodea una terminal o alambre de alta tensión se ioniza. El aire ionizado parece luminiscente y produce sonidos audibles de agrietamiento. También produce ruido RF (de radiofrecuencia) y es un agente destructivo para los materiales aislantes eléctricos. Largos periodos de exposición al efecto corona pueden conducir a disrupciones.

12.4.7 Modos y mecanismos de fallo en transformadores e inductores

Las principales causas de falla en transformadores e inductores pueden atribuirse al grado de aislamiento, la temperatura ambiente y el esfuerzo eléctrico. A continuación se analizan algunos de estos mecanismos:

1. **Voltaje primario excesivo.** Si el voltaje primario es suficientemente alto puede producirse una perforación inmediata del aislamiento. Un sobrevoltaje moderado (aproximadamente un 20% del voltaje nominal) conducirá a la rotura prematura del aislamiento.
2. **Fluctuación de la frecuencia de entrada.** Frecuencias menores del valor nominal provocarán baja reactancia y el flujo de corrientes más altas

que las nominales. Frecuencias mayores del valor nominal provocarán pérdidas más altas en el núcleo. Cualquiera de los dos casos provocará un sobrecalentamiento, o el incremento de temperatura del dispositivo será mayor que la máxima temperatura de diseño. Esto conducirá finalmente a rotura del aislamiento.

3. **Corriente secundaria excesiva.** Corrientes secundarias mayores que los valores nominales producen un sobrecalentamiento del transformador, lo que debilita la resistencia dieléctrica del aislamiento. Esto puede provocar que los devanados se abran o cortocircuiten. Asimismo, si los recipientes son de compuestos de impregnación o de relleno, el sobrecalentamiento puede provocar rotura o distorsión.
4. **Efecto corona.** El efecto corona ocurre en puntos de alto esfuerzo potencial. Provoca envejecimiento acelerado del aislamiento por la liberación de ozono y por incremento de la temperatura. Por tanto, en el aislamiento se crean puntos débiles, lo que a la larga producirá su rotura.

Las bobinas moldeadas y encapsuladas tienden a fallar debido a vueltas cerradas, terminaciones abiertas y degradación del factor de calidad Q . Aunque el moldeo o el encapsulamiento mejoran la resistencia de la bobina a la humedad, los impactos y la vibración, el flujo indebido del compuesto de moldeo puede aplastar el devanado, lo cual puede provocar circuitos abiertos o fallos en la rigidez dieléctrica. En los casos en que existe una degradación de Q , también habrá una mayor incidencia de fallos en bobinas con núcleos de hierro pulverizado. Las piezas de metal de hierro pulverizado se degradarán en caso de sobrecalentamiento, lo que ocurre a 105 °C.

LIBROS Y REVISTAS

- V. del Toro, *Principles of Electrical Engineering*, Prentice Hall, Englewood Cliffs, NJ.
- D. G. Fink y cols., eds., *Electronics Engineering Handbook*, 2.^a ed., McGraw-Hill, Nueva York.
- A. E. Fitzgerald y D. E. Higginbotham, *Basic Engineering*, 2.^a ed., McGraw-Hill, Nueva York.
- "Reliability Stress and Failure Rate Data for Electronic Equipment", *MIL HDBK-217A*.

CAPÍTULO 13

Elementos circuitales concentrados activos

Mark B. Barron

General Electric Microelectronics Center
Research Triangle Park, North Carolina

Henry Domingos

Clarkson University
Potsdam, Nueva York

George H. Ebel

Singer Company
Wayne, New Jersey

13.1 Diodos, rectificadores, varistores

- 13.1.1 Diodos de unión *pn*
- 13.1.2 Rectificadores de potencia

13.2 Transistores bipolares

- 13.2.1 Acción de transistor
- 13.2.2 Parámetros de los transistores
- 13.2.3 Tipos de transistores de potencia
- 13.2.4 Transistores Darlington
- 13.2.5 Consideraciones para las aplicaciones de los transistores

13.3 Transistores de efecto de campo

- 13.3.1 Transistores de efecto de campo con compuerta aislada
- 13.3.2 Transistor de efecto de campo de unión (JFET)
- 13.3.3 Variaciones de los MOSFET
- 13.3.4 Consideraciones de aplicación

- 13.4.2 Fabricación de circuitos integrados

- 13.4.3 Circuitos integrados bipolares

- 13.4.4 Circuitos integrados MOS

13.5 Circuitos híbridos: de película gruesa y de película delgada

- 13.5.1 Descripción general
- 13.5.2 Definición de términos
- 13.5.3 Procesos de manufactura
- 13.5.4 Especificaciones de corriente
- 13.5.5 Consideraciones térmicas
- 13.5.6 Consideraciones sobre frecuencias
- 13.5.7 Aplicaciones
- 13.5.8 Referencias generales

13.6 Otros dispositivos electrónicos

- 13.6.1 Tiristores
- 13.6.2 Transistores uniunión
- 13.6.3 Otros dispositivos semiconductores
- 13.6.4 Tubos de vacío
- 13.6.5 Tiratrones

VOLUMEN 2

13.4 Circuitos integrados

- 13.4.1 Introducción

13.1 DIODOS, RECTIFICADORES, VARISTORES

Mark B. Barron

13.1.1 Diodos de unión pn

La columna vertebral de la mayor parte de los dispositivos semiconductores, que amplifican, conmutan o emiten radiación, es la unión pn . Esta unión, que se forma colocando un semiconductor tipo p adyacente a un semiconductor tipo n , tiene la propiedad de impedir el flujo de corriente en una dirección, al tiempo que permite su paso en la otra dirección. Aunque es posible utilizar materiales diferentes para los semiconductores tipo p y tipo n , formando de este modo una heterounión, la mayor parte de las uniones pn se forman del mismo material, por ejemplo, silicio o germanio.

El funcionamiento físico de una unión pn puede visualizarse si se recuerda que el material tipo n tiene portadores de carga que en su mayor parte son electrones (con carga negativa) y que el material tipo p tiene portadores de carga que en su mayor parte son huecos (con carga positiva). Por consiguiente, si en la unión se aplica una polarización, de modo que el lado p sea

positivo y el n sea negativo, tal como se muestra en la figura 13-1a, los electrones serán atraídos por el material n hacia el lado positivo de la fuente, y los huecos serán atraídos por la terminal negativa. Así, la corriente fluirá de un lado a otro de la unión (fig. 13-1c).

Si se invierte la polarización (fig. 13-1b) y el material n adquiere polarización positiva con respecto al material p , los electrones serán atraídos hacia afuera del material p por el lado positivo de la fuente. De manera similar, los huecos serán atraídos hacia afuera del material p por el lado negativo de la fuente. Dado que en el material p sólo existen algunos electrones, la corriente atribuida a estos portadores, que fluye de un lado a otro de la unión, será pequeña.

La corriente ideal a través de una unión pn está dada por la siguiente ecuación¹:

$$I = I_0(e^{qV/kT} - 1) \quad (13-1)$$

donde V es el voltaje a través de la unión, T la temperatura en $^{\circ}\text{K}$ y q/k una constante cuyo valor es $11.6 \times 10^4 \text{ K/V}$. En dispositivos reales habrá una recombinación de los huecos y los electrones a medida que fluyen a través de la unión, y la corriente se reducirá respecto al caso ideal.

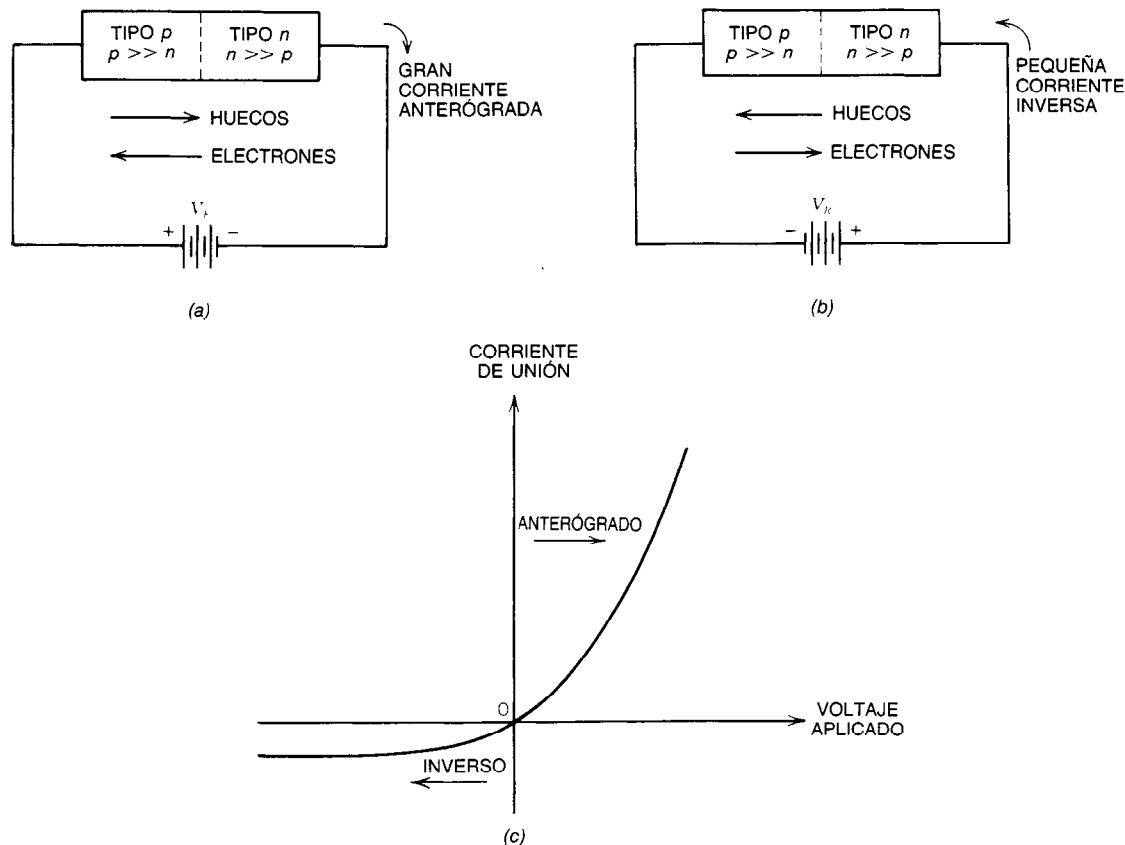


Fig. 13-1. a) Unión pn con polarización directa. b) Unión pn con polarización inversa. c) Característica corriente-voltaje de una unión pn .

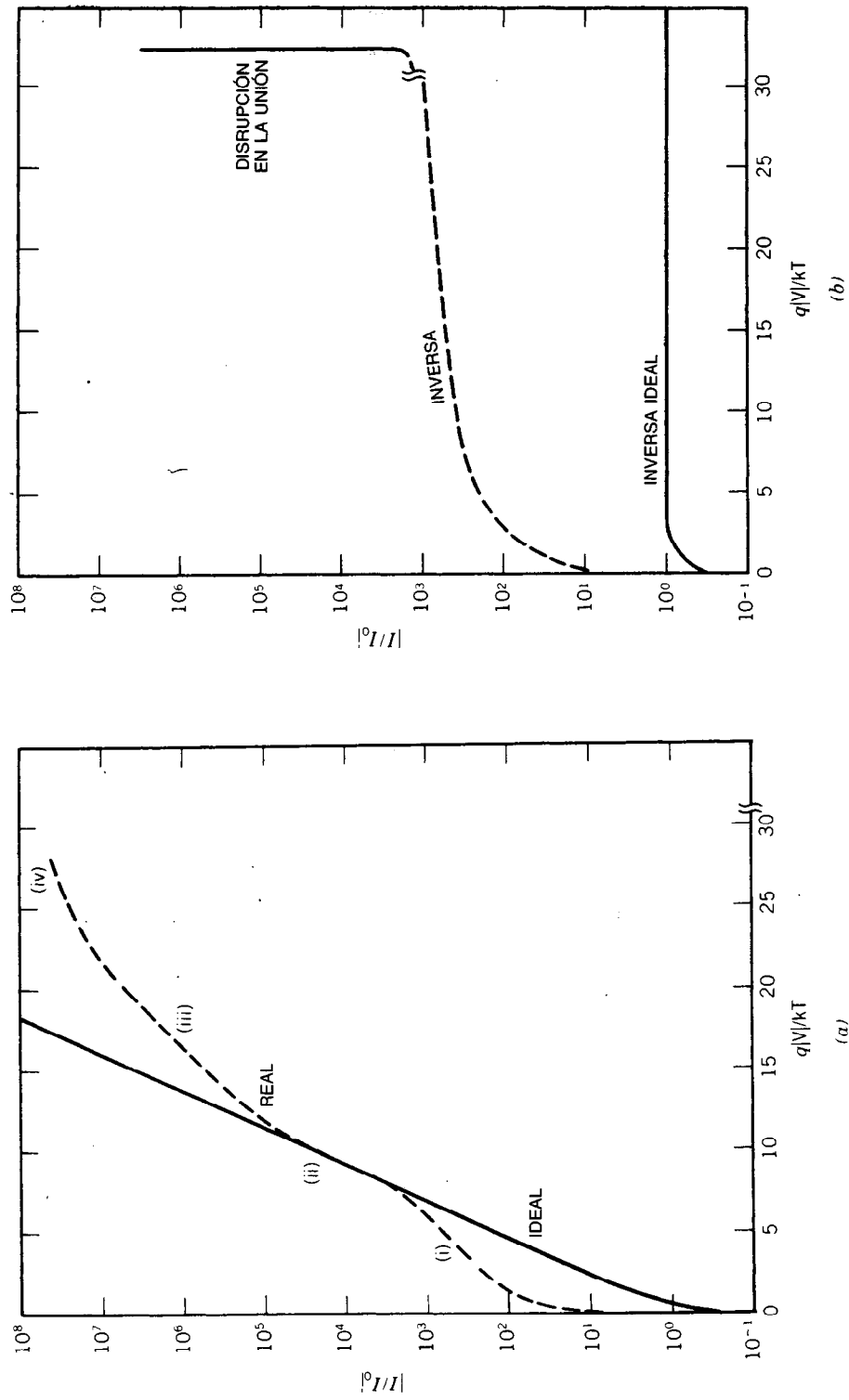


Fig. 13-2. a) Características corriente-voltaje con polarización directa de un diodo de silicio: (i) corriente de generación-recombinación, (ii) corriente de difusión, (iii) alta inyección, (iv) resistencia en serie. b) Características corriente-voltaje con polarización inversa de un diodo de silicio.

La ecuación anterior puede modificarse a fin de incluir este efecto mediante un factor de idealidad n :

$$I = I_0(e^{qV/nkT} - 1) \quad (13-2)$$

donde n varía entre 1 y 2, dependiendo del material, de la temperatura y del nivel de corriente. Para el germanio, n se acerca a 1, mientras que para el silicio a bajas corrientes es casi igual a 2. A niveles superiores de corriente, el valor de n para el silicio es aproximadamente 1, y para niveles muy elevados de corriente existe una modificación significativa en cuanto a la concentración de portadores fuera de las regiones de la unión, y de un lado a otro de dichas regiones se presenta una caída de potencial. Para elevados niveles de inyección, la corriente varía conforme a $I \propto V^2$:

$$I = I_0 e^{qV/2kT} \quad (13-3)$$

donde I_0 es una constante algo mayor que I_0 de la ecuación 13-2. En la figura 13-2a se muestran las características de corriente-voltaje con polarización directa de un diodo típico de silicio.

Cuando a través del diodo se establece una polarización inversa, V es negativo y la corriente a través de la unión ideal es $I = -I_0$. Los huecos y electrones son atraídos hacia afuera de la unión para formar una región agotada de portadores móviles. La región de agotamiento se forma de un lado a otro de la unión, de modo que la carga asociada con los donadores ionizados de un lado de la unión sea igual a la carga asociada con los aceptores ionizados del otro lado. Por tanto:

$$N_D X_n = N_A X_p \quad (13-4)$$

donde X_n es la penetración de la capa de agotamiento en el material n y X_p es la penetración de la capa de agotamiento en el material p . Como se describió en el capítulo 10, N_D y N_A son las concentraciones de las impurezas agregadas en los materiales n y p , respectivamente. Dentro de la región de agotamiento se forman pares hueco-electrón mediante la excitación térmica del cristal semiconductor, y se crea corriente a medida que los portadores son atraídos a las terminales. Esta corriente se expresa como sigue³:

$$I_{\text{gen}} = \frac{1}{2} q \frac{n_i}{\tau_0} W A \quad (13-5)$$

donde n_i es la concentración intrínseca de portadores, τ_0 es el tiempo de vida efectiva de los portadores antes de la recombinación, W es el ancho de la región de agotamiento ($W = X_n + X_p$) y A es el área transversal de la unión. Para una unión abrupta como la que se formaría en una aleación, el semiconductor cambia abruptamente desde el material p de concentración N_A hasta el material n de concentración N_D , y el ancho de la región de agotamiento se expresa como:

$$W = \sqrt{\frac{2\epsilon(V_0 - V)}{q} \left(\frac{N_A + N_D}{N_A N_D} \right)} \quad (13-6)$$

donde V_0 es el voltaje integrado de la unión pn , expresado como:

$$V_0 = \frac{kT}{q} \ln \frac{N_A N_D}{n_i^2} \quad (13-7)$$

Para grandes voltajes inversos en una unión abrupta:

$$I_{\text{gen}} = \frac{n_i q A}{2\tau_0} \cdot \sqrt{\frac{2\epsilon V}{q} \cdot \left(\frac{N_A + N_D}{N_A N_D} \right)} \quad (13-8)$$

En la figura 13-2b se muestran las características de corriente-voltaje con polarización inversa de un diodo típico de silicio.

La capacitancia en una unión pn tiene dos componentes; una asociada con el dipolo en la región de agotamiento alrededor de la unión (capacitancia de la unión) y otra asociada con el almacenamiento de carga. La capacitancia de la unión domina bajo condiciones de polarización inversa, y la del almacenamiento de carga domina cuando la polarización de la unión es directa. La capacitancia de la unión es idéntica a la de un capacitor de placas paralelas, y se expresa como:

$$C_j = \frac{\epsilon A}{W} \quad (13-9)$$

donde W es el ancho de la región de agotamiento. La capacitancia de almacenamiento (o difusión) de carga, atribuida al almacenamiento de electrones en el material p y de huecos en el material n , se expresa como:

$$C_D = \frac{q I \tau_0}{kT} \quad (13-10)$$

La aproximación de unión abrupta es aceptable para las uniones de aleaciones o uniones difusas poco profundas, pero no lo es para uniones pn elaboradas mediante la difusión profunda de una impureza de un tipo de material en otro tipo de material. Para tales dispositivos, una unión linealmente graduada es una buena aproximación en intervalos limitados de voltaje. En una unión linealmente graduada se supone que la densidad de impureza es:

$$N_D - N_A = ax \quad (13-11)$$

Es posible calcular el ancho de la región de agotamiento como⁴:

$$W = \left(\frac{12\epsilon(V_0 - V)}{qa} \right)^{1/3} \quad (13-12)$$

donde V_0 es el potencial integrado que se encuentra al resolver la ecuación:

$$V_0 = \frac{kT}{q} \ln \left[\frac{a}{2n_i} \left(\frac{12\epsilon V_0}{qa} \right)^{1/3} \right]^2 \quad (13-13)$$

Con el ancho de la región de agotamiento dado por la ecuación 13-13, todavía es posible aplicar la corriente

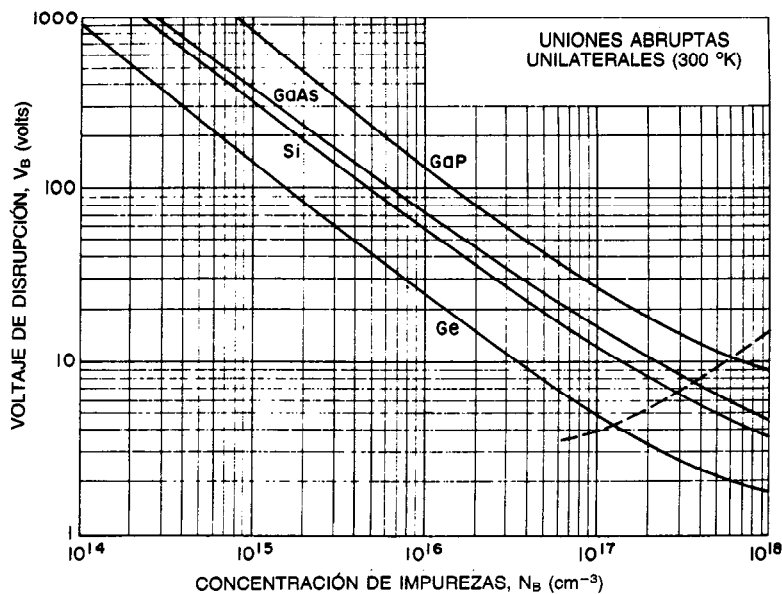


Fig. 13-3. Voltaje de disrupción en avalancha en función de la concentración de impurezas de uniones abruptas unilaterales.

te de generación I_{gen} y la capacitancia C_j , que vienen proporcionados por la ecuación 13-9.

Voltaje de disrupción

Cuando a través de una unión pn se aplica una polarización inversa, la mayor parte del voltaje se encuentra a través de la región de agotamiento, cuyo ancho varía sólo en $V^{1/2}$ o $V^{1/3}$, dependiendo de si la unión es abrupta o graduada. Por consiguiente, el campo eléctrico $E = dv/dx \approx V/W$ aumenta con el voltaje. En un campo eléctrico suficientemente elevado, los portadores generados en la región de agotamiento adquirirán suficiente energía cinética del campo, de modo que cuando choquen contra la malla de silicio liberarán otros pares hueco-electrón. Esta multiplicación de portadores se denomina proceso en avalancha y produce una corriente:

$$I_r = MI_{r0} \quad (13-14)$$

donde I_{r0} es la corriente sin avalancha y:

$$M = \frac{1}{1 - (V/V_{Br})^n} \quad (13-15)$$

V_{Br} es el voltaje de disrupción de la unión y n varía entre 3 y 6, dependiendo del semiconductor. A partir de la ecuación 13-15 es posible observar que, cuando $V = V_{Br}$, el factor de multiplicación $M \rightarrow \infty$. La corriente está limitada sólo por la resistencia del semiconductor fuera de la región de la unión. Para uniones abruptas, V_{Br} está dado por:

$$V_{Br} = \frac{\epsilon(N_A + N_D)E_{crit}^2}{2qN_A N_D} \quad (13-16)$$

donde E_{crit} es el campo en el cual ocurre la disrupción. Para una unión graduada, el voltaje de disrupción es⁵:

$$V_{Br} = \sqrt{\frac{32eE_{crit}^3}{9qa}} \quad (13-17)$$

En la figura 13-3 se muestran los voltajes de disrupción de uniones escalonadas p^+n unilaterales en silicio, germanio y arseniuro de galio como función de la concentración de impurezas en el material tipo n ligeramente impurificado. En la figura 13-4 se muestran los voltajes disruptivos en los mismos materiales para uniones linealmente graduadas.

Para materiales altamente impurificados, con concentraciones de impurezas mayores de $5 \times 10^{17} \text{ cm}^{-3}$, puede ocurrir otro mecanismo disruptivo, en lugar del de avalancha; éste es el efecto Zener. En la disrupción de Zener, el campo es tan intenso que los electrones se separan de los átomos cristalinos sin la necesidad de choques de portadores, tal como se requiere en la disrupción en avalancha. Las condiciones para que ocurra la disrupción de Zener sólo se dan en uniones abruptas con una estrecha región de agotamiento, por ejemplo, en semiconductores altamente impurificados con bajos voltajes de disrupción. Asimismo, la disrupción de Zener ocurre más fácilmente en semiconductores con estrecha separación energética, como el germanio, que en materiales con mayor separación energética, como el fosforo de galio. De forma experimental se ha determinado que la disrupción de Zener ocurre en campos con valor aproximado de 10^6 V/cm^6 .

Disrupción en diodos de unión pn reales. En las uniones pn prácticas fabricadas con tecnología moderna,

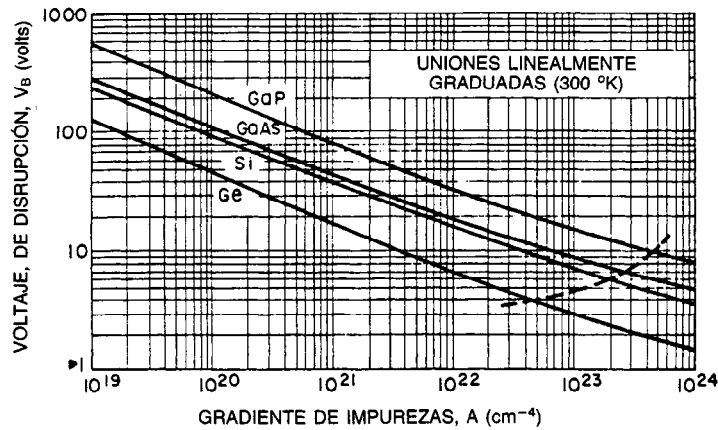


Fig. 13-4. Voltaje de ruptura en avalanche en función de la concentración de impurezas de uniones p graduadas linealmente.

los voltajes de ruptura están dominados por efectos superficiales, por la curvatura de las uniones o por ambas cosas. La mayor parte de los diodos pn se fabrican mediante procesos de difusión (descritos en el cap. 10). La localización del diodo es determinada ya sea por

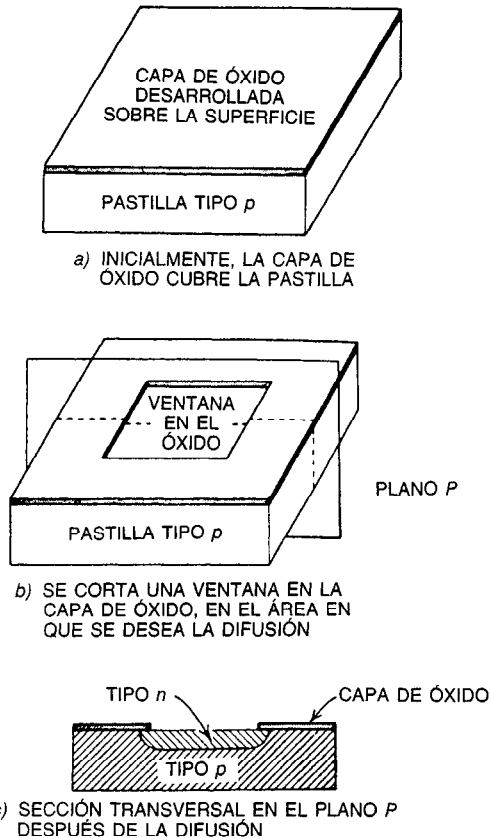


Fig. 13-5. Fabricación del diodo de unión pn plana.

ventanas de óxido a través de las cuales se difunden las impurezas (fig. 13-5) o, en el caso de los rectificadores de potencia, por el corte o grabado de islas fuera de una región difundida protegida (fig. 13-6). En el primer caso, el voltaje disruptivo es reducido por la curvatura en los bordes de la unión difundida. En el segundo caso, las imperfecciones de la superficie reducen de manera efectiva la E_{crit} en el borde de la unión, y $V_{Br sup} < V_{Br vol}$.

La fig. 13-7 muestra el voltaje de ruptura de las uniones de silicio difundido en función de la profundidad de difusión, con la concentración de impurezas en el sustrato y de la superficie de difusión como parámetros.

Características transitorias. Cuando el voltaje a través de una unión pn se conmuta repentinamente de positivo a negativo, la corriente no adquiere de manera inmediata el valor de polarización inversa de estado estable. La causa de este hecho es que existe un exceso de portadores minoritarios en cada lado de la unión en condiciones de polarización directa. Cuando la polarización del voltaje cambia de directa a inversa, el exceso de huecos en el lado n de la unión es atraído hacia la terminal negativa en el material p , y de forma parecida, el exceso de electrones en el lado p de la unión fluye hacia la terminal positiva en el material n . Es común que la impedancia del circuito accionador limite la corriente inversa a algún valor pico, I_R , durante el

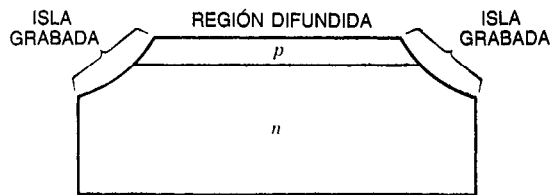


Fig. 13-6. Diodo de meseta resultante del proceso de grabado.

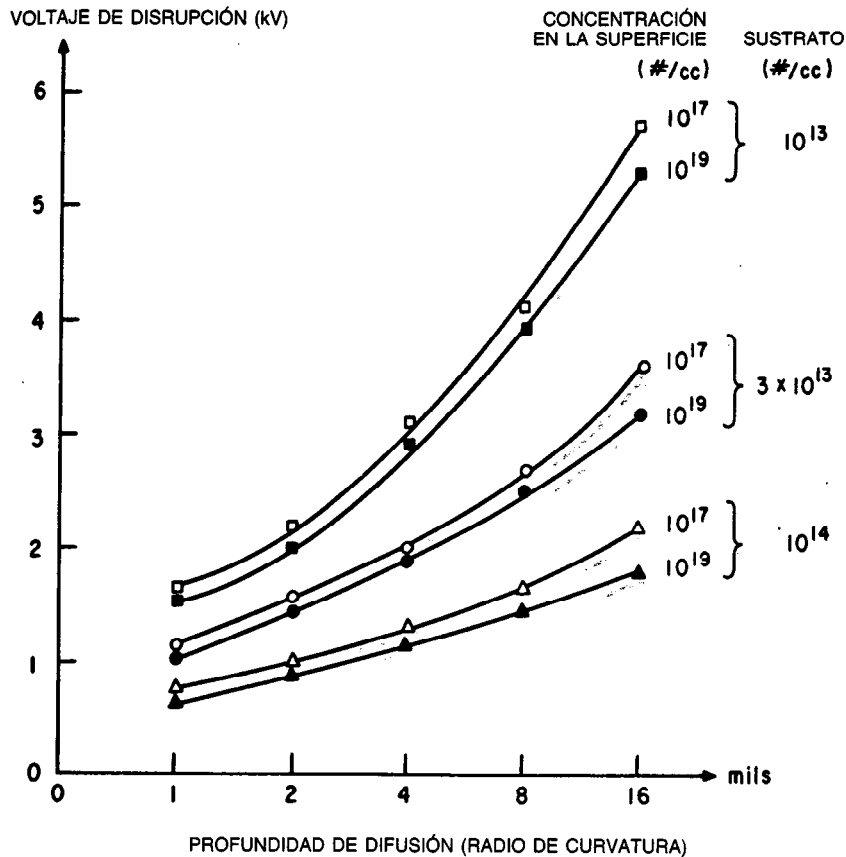


Fig. 13-7. Voltaje de ruptura en avalancha de uniones de silicio difundido en función de la profundidad de difusión, con la concentración de impurezas y la concentración superficial de la difusión como parámetros.

cambio. Como se observa en la figura 13-8, la corriente permanece en el valor I_R un tiempo de almacenamiento t_s hasta que el exceso de portadores minoritarios alcanza su valor de equilibrio. A partir de ese momento la corriente cae al valor de estado estable asociado con el voltaje inverso, a medida que los portadores minoritarios cercanos a la unión se agotan por debajo de sus valores de equilibrio.

Un análisis del problema⁷ revela que el tiempo de almacenamiento está determinado por la siguiente expresión:

$$t_s = \tau_0 \left[\operatorname{erf}^{-1} \left(\frac{I_F}{I_F + I_R} \right) \right]^2 \quad (13-18)$$

donde τ_0 = tiempo de vida efectivo

I_F = corriente en sentido directo

I_R = corriente inversa

Para uniones p^+n , τ_0 podría ser el tiempo de vida de los huecos en el material n antes de que se recombinen. De manera parecida, el tiempo de descenso, t_f , puede obtenerse despejando dicho valor en la siguiente ecuación:

$$\operatorname{erf} \sqrt{\frac{t_f}{\tau_0}} + \frac{e^{-t_f/\tau_0}}{\sqrt{\frac{\pi t_f}{\tau_0}}} = 1 + 0.1 \left(\frac{I_R}{I_F} \right) \quad (13-19)$$

En ocasiones es aconsejable que el cambio sea rápido, y los fabricantes reducen el tiempo de vida τ_0 impurificando el semiconductor con oro o platino, o bien irradiándolo con electrones de alta energía.

Sin embargo, existe una posición intermedia entre conmutación rápida y baja corriente de fuga inversa ya que, como se indica en la ecuación 13-5, la corriente inversa aumenta al disminuir el tiempo de vida τ_0 .

Efectos térmicos. A medida que aumenta la temperatura en una unión pn polarizada a un voltaje directo fijo, la corriente de conducción puede aumentar o disminuir, dependiendo del nivel de la corriente de operación. A niveles bajos y medios de esta corriente, la corriente de conducción aumentará con la temperatura. A inyección de nivel alto, la corriente de conducción disminuye, como se muestra en la figura 13-9. En condiciones de polarización inversa, la corriente

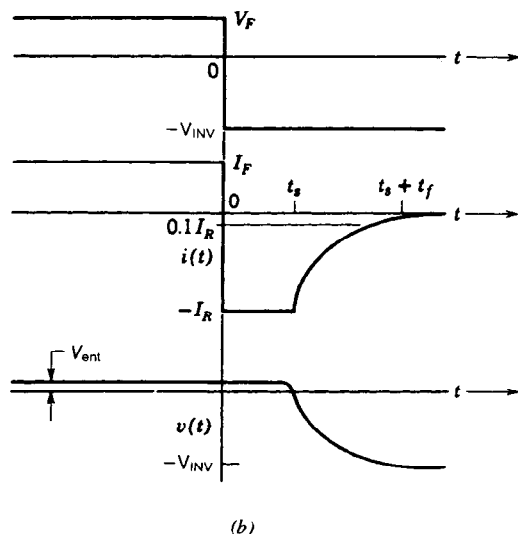
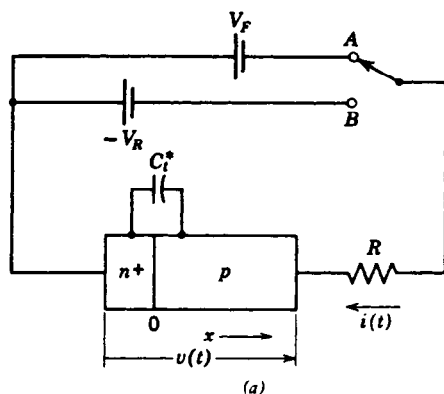


Fig. 13-8. Voltaje y corriente transitorios en un diodo de unión.

inversa aumenta exponencialmente con la temperatura (fig. 13-10). El voltaje de disrupción en avalancha aumenta ligeramente con la temperatura (fig. 13-11). Por otra parte, debido a que la separación energética se estrecha a temperaturas superiores, el voltaje de disrupción del diodo Zener disminuye con la temperatura. En dispositivos de mayor voltaje, la corriente puede hacerse tan grande a altas temperaturas que un voltaje elevado sostenido generará tantos watts que el dispositivo mismo se calentará y la corriente aumentará hasta que el dispositivo se autodestruya. Al aplicar diodos a altas temperaturas es necesario considerar la posibilidad de fuga térmica y diseñar los circuitos con dispositivos sensores de temperatura apropiados.

Tipos de diodos

Diodos varactor. Mientras que en algunas aplicaciones la capacitancia de la unión pn es una amenaza, es una característica útil para aplicaciones en las que se re-

quiere capacitancia controlada por voltaje. Para tales aplicaciones es aconsejable que la capacitancia varíe más rápidamente que el $V^{-1/2}$ que se obtiene con uniones abruptas. Es posible lograr una variación más rápida con el voltaje mediante el empleo de una unión hiperabrupta, como la que se muestra en la figura 13-12. Con este dispositivo, construido con técnicas epitaxiales controladas, es posible producir varactores cuya capacitancia varía como $(V_0 - V)^{-2}$.

Diodos *pin*. Un diodo *pin* se construye con una capa de alta resistividad (silicio intrínseco) colocada entre el material *p* y el material *n* (fig. 13-13a). Tales diodos se caracterizan por una capacitancia relativamente constante de la capa de agotamiento, dada por:

$$C = \frac{\epsilon A}{W_i} \quad (13-20)$$

donde W_i es el ancho de la región intrínseca. El voltaje disruptivo está dado por:

$$V_B = E_{\text{crit}} W_i \quad (13-21)$$

donde E_{crit} es el campo eléctrico crítico al cual ocurre la disrupción. Para una estructura *pin* de silicio con

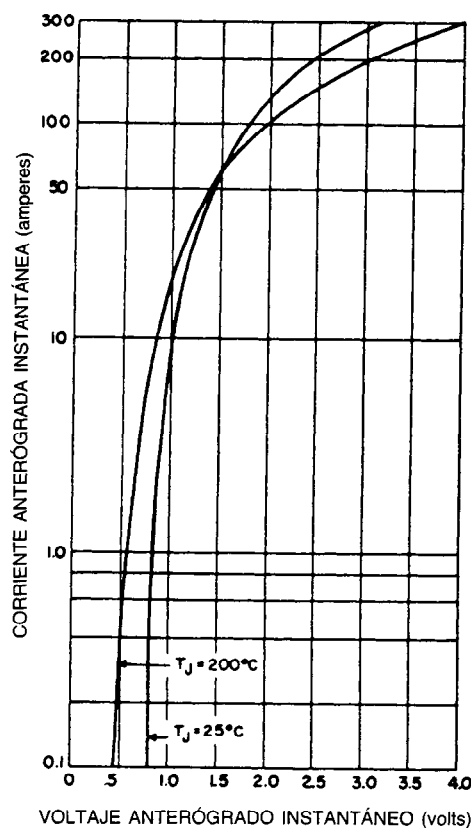


Fig. 13-9. Dependencia respecto a la temperatura de las características de un diodo con polarización directa.

$W_i = 10\mu\text{m}$, el voltaje disruptivo sería $V_B = 2 \times 10^5 \times 10 \times 10^{-4} = 200\text{ V}$. Dada la alta resistencia de la región intrínseca, la mayor parte de la caída de voltaje ocurre a través de esta región. Para niveles bajos y moderados de corriente en sentido directo, la resistencia varía inversamente con la corriente (fig. 13-13b), por lo cual el diodo *pin* suele emplearse como un atenuador variable en circuitos de microondas.

Fotodiodos. En el análisis de las uniones *pn* se observó que los pares hueco-electrón se generan térmicamente, y que con polarización inversa contribuyen a la corriente de polarización inversa. A polarización cero, la corriente de generación se encuentra en equilibrio con la corriente de recombinación, y la corriente neta es igual a cero. Es posible proyectar luz sobre el semiconductor y generar más pares hueco-electrón que los generados en equilibrio térmico; el único requisito es que la energía de la luz sea superior a la energía de la separación energética del semiconductor: $h\nu > E_g$. Para el silicio, dicha separación es aproximadamente de 1.11 eV, lo que significa que la longitud de onda de la luz debe ser menor de $\lambda_T = hc/E_g = 1.12\mu\text{m}$. Por consiguiente, el silicio es sensible a la luz en la zona del infrarrojo cercano y en la región visible. Las características *VI* de un fotodiodo se presentan en la figura 13-14. Cuando el fotodiodo opera en el cuarto cuadrante, suministra potencia al circuito externo. En los cuadrantes primero y tercero, se entrega potencia del circuito externo al fotodiodo. La corriente de cortocircuito, I_{SC} , es determinada por el flujo de luz, en tanto que el voltaje de circuito abierto, V_{OC} , es limitado por la separación energética del semiconductor. Para el silicio, $V_{OC} < 1.1\text{ V}$.

Dado que los fotodiodos pueden generar potencia en el cuarto cuadrante, se emplean para generar electricidad, y en aplicaciones como éstas se les denomina celdas solares. Las celdas solares de silicio convierten sólo alrededor del 20% de la luz incidente con longitud de onda menor de $1.1\mu\text{m}$, y nada de la luz con mayor longitud de onda. En el caso de la luz solar, parte de la energía tiene mayores longitudes de onda, donde $h\nu < E_g$. Por consiguiente, la eficiencia compuesta de las celdas solares de silicio para la conversión de luz solar en energía eléctrica es de sólo 10%. Ya que aproximadamente 1 kW/m^2 de radiación incide sobre la superficie terrestre en un día soleado, con celdas solares es posible generar 100 W/m^2 .

Diodos Zener. Estos diodos constituyen una clase de dispositivos de unión *pn* con un voltaje de disrupción especificado; se pretende que operen a ese voltaje como fijadores del nivel de tensión. En realidad, el mecanismo de disrupción es más a menudo en avalanche que el de Zener, pero a dichos dispositivos se aplica el nombre genérico de diodos Zener aun cuando esta designación no precisa la causa física de la disrupción.

Los diodos Zener se aplican como reguladores de voltaje o como referencias de voltaje. El voltaje de disrupción es un parámetro de especificación, como lo es la resistencia dinámica del dispositivo. Estos diodos son fabricados para ciertos intervalos de puntos

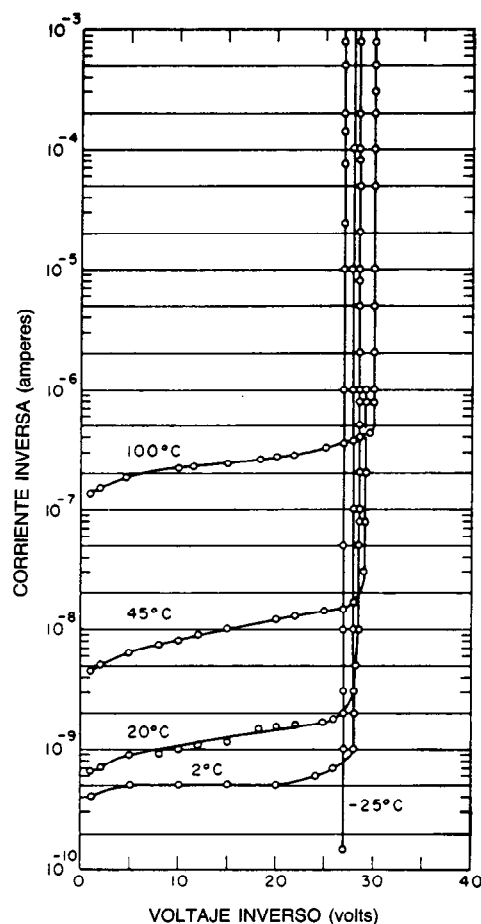


Fig. 13-10. Dependencia respecto a la temperatura de las características de un diodo con polarización inversa.

de voltaje de descarga disruptiva, y después se seleccionan por medio de equipo automático de prueba para intervalos más pequeños (por lo general $\pm 5\%$).

Los diodos Zener también pueden emplearse para la supresión de voltajes transitorios. En esta aplicación el diodo no se opera en la región de disrupción, sino hasta que ocurre una oscilación de voltaje. El voltaje a través del diodo en condiciones normales de operación suele ser menor que el 90% del valor de disrupción. Cuando ocurre una oscilación momentánea, el diodo fija el nivel de voltaje y actúa como elemento de derivación a fin de absorber la energía de la corriente de irrupción. A la capacidad normal para la máxima corriente transitoria no repetitiva se la denomina I_{RSM} ; suele ser un pulso con cierta duración.

La capacidad normal de controlar el voltaje de descarga disruptiva para los diodos Zener suele variar de 4 a 100 V. Con tales diodos puede disiparse una capacidad normal de potencia de supresión de hasta 20 kW para pulsos cortos, y en ciertas especificaciones se da información acerca de la potencia inversa pico repetitiva que es posible disipar en diversas condiciones.

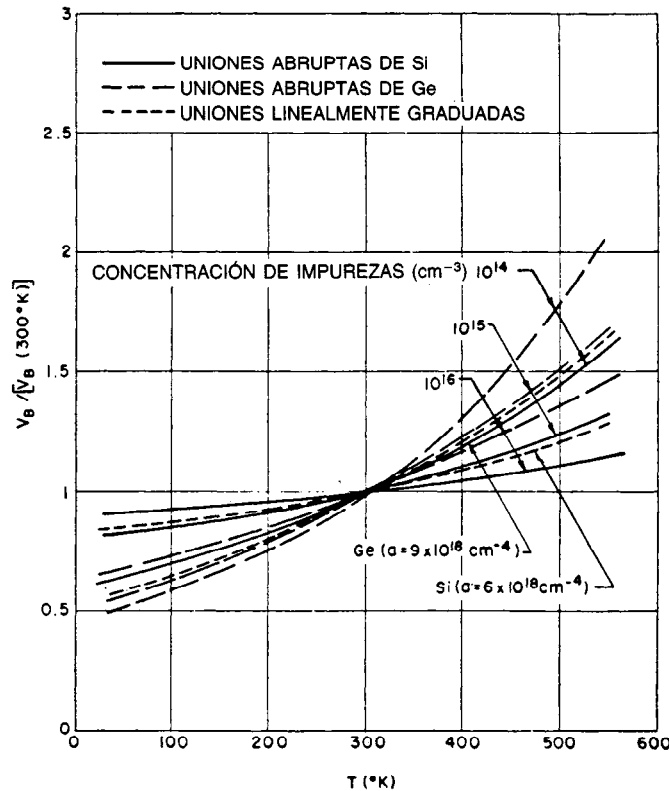


Fig. 13-11. Dependencia respecto a la temperatura del voltaje de disrupción en avalancha en el germanio y el silicio.

Varistores de óxido metálico. Para aplicaciones en las que se presentan voltajes superiores, como en los circuitos que funcionan a 120 V de CA, o con posibles oscilaciones momentáneas de potencia que exceden la capacidad de los diodos Zener, una opción práctica puede ser el varistor de óxido metálico (MOV, de *Metal Oxide Varistor*). Los MOV se fabrican empleando polvos de óxido de cinc o de óxido de plomo con diversos elementos impurificadores. Desde el punto de vista

eléctrico, los varistores de óxido metálico son uniones *pn* espalda con espalda (fig. 13-15), y tienen valores nominales a voltaje aplicado eficaz máximo, voltaje devastado pico recurrente, máximo voltaje DC aplicado, energía (joules), disipación media de potencia y corriente pico para una anchura designada de pulso. Los varistores de óxido metálico pueden manejar corrientes pico de varios miles de amperes, miles de volts y energía de cientos de joules. Como los diodos Zener, los MOV tienen una resistencia dinámica especificada que depende de la corriente.

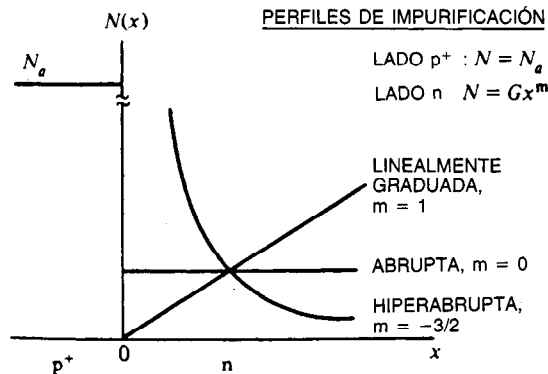


Fig. 13-12. Perfiles de impurificación de una unión.

Diodos de túnel. Un diodo de túnel (también llamado de efecto túnel o diodo Esaki) es una unión *pn* con regiones tipo *p* y tipo *n* extremadamente impurificadas, y por lo general se elaboran de germanio, aunque también existen de arseniuro de galio. Una explicación detallada de los diodos de túnel requiere el empleo de la mecánica cuántica, lo que está más allá de los objetivos de este manual. Para quien esté interesado en una descripción física exhaustiva del proceso existen varias referencias útiles⁸.

El diodo de túnel es capaz de conducir considerable corriente en polaridad inversa, debido a que los electrones y huecos pueden moverse a través de la unión sin experimentar los efectos normales del campo eléctrico. Para bajos voltajes, positivos o negativos, la ca-

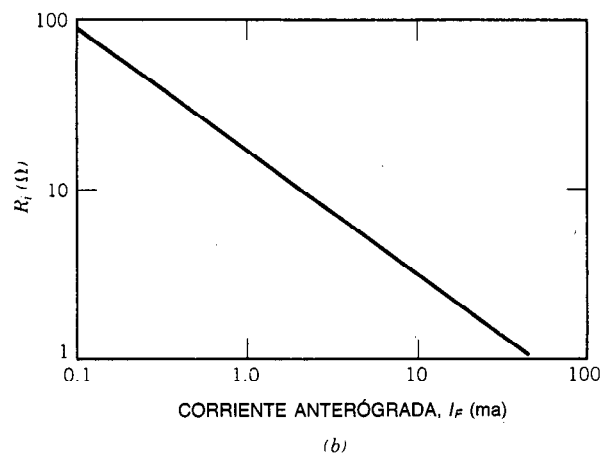
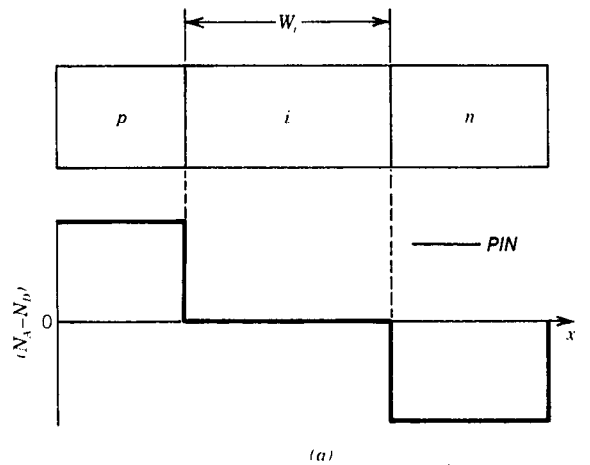


Fig. 13-13. a) Diodo *pin*. b) Resistencia en función de la corriente para un diodo *pin* con polarización directa.

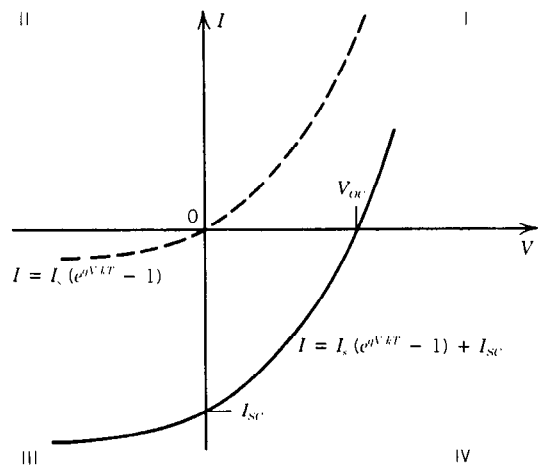


Fig. 13-14. Características corriente-voltaje de un fotodiodo.

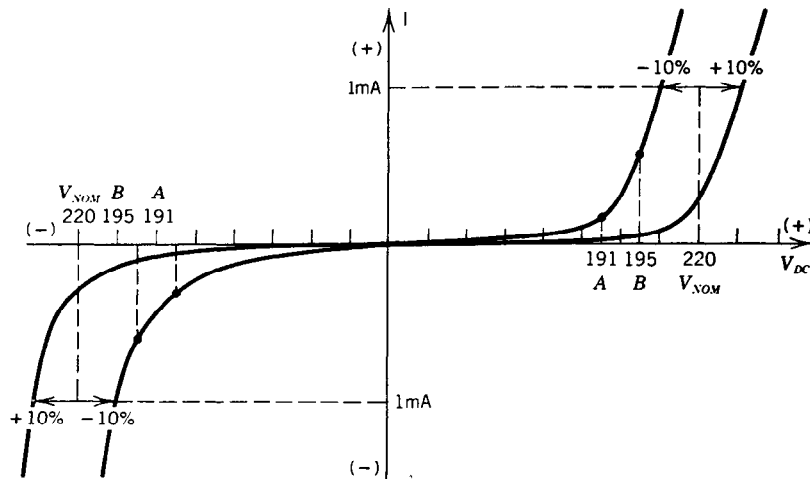


Fig. 13-15. Características corriente-voltaje de un varistor de óxido metálico (MOV) de cinc. A = máximo voltaje CD de estado estable aplicado permisible; B = máximo voltaje pico recurrente de estado estable aplicado permisible; V_{NOM} = voltaje nominal del varistor a 1 mA de CD.

racterística V - I de un diodo de túnel es casi lineal (fig. 13-16). Para un voltaje directo de V_p , el dispositivo empieza a perder sus propiedades de filtración cuántica (efecto túnel) y la corriente disminuye en realidad con el aumento de voltaje. Esta región de resistencia negativa continúa hasta que la corriente normal del diodo se hace lo suficientemente grande para provocar que la resistencia sea positiva de nuevo. Es esta región controlada de resistencia negativa la que hace útil al diodo de túnel.

El valor de la corriente pico de efecto túnel, I_p , y el de la corriente de valle, I_v , determinan el valor de la resistencia negativa. Por consiguiente, las cantidades I_p/I_v y V_p/V_v son cifras de mérito para un diodo de túnel. Los diodos de túnel son de utilidad para aplicaciones en osciladores de alta frecuencia controlados por voltaje y para aplicaciones de conmutación rápida.

Estos diodos son uno de los dispositivos de conmutación más rápidos que se conocen, con tiempos de transición hasta de 25 ps.

13.1.2 Rectificadores de potencia

Los **rectificadores** son diodos capaces de manejar niveles de corriente superiores a 1 A. Mientras que los primeros rectificadores se fabricaban con óxido de cobre o selenio, casi todos los rectificadores modernos son semiconductores, y en la actualidad el rectificador de potencia predominante es la unión pn de silicio. El rectificador difiere de los diodos de baja potencia principalmente en el tamaño y en los métodos de fabricación. Los rectificadores generan cantidades sustanciales de calor que es necesario eliminar del semiconductor y su encapsulamiento o empaque. De otra

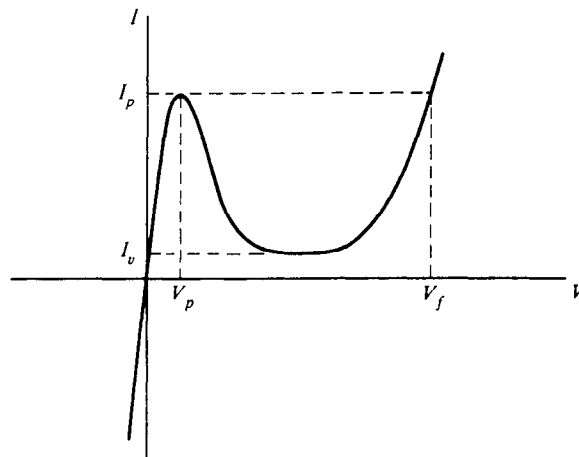


Fig. 13-16. Característica corriente-voltaje de un diodo de túnel.

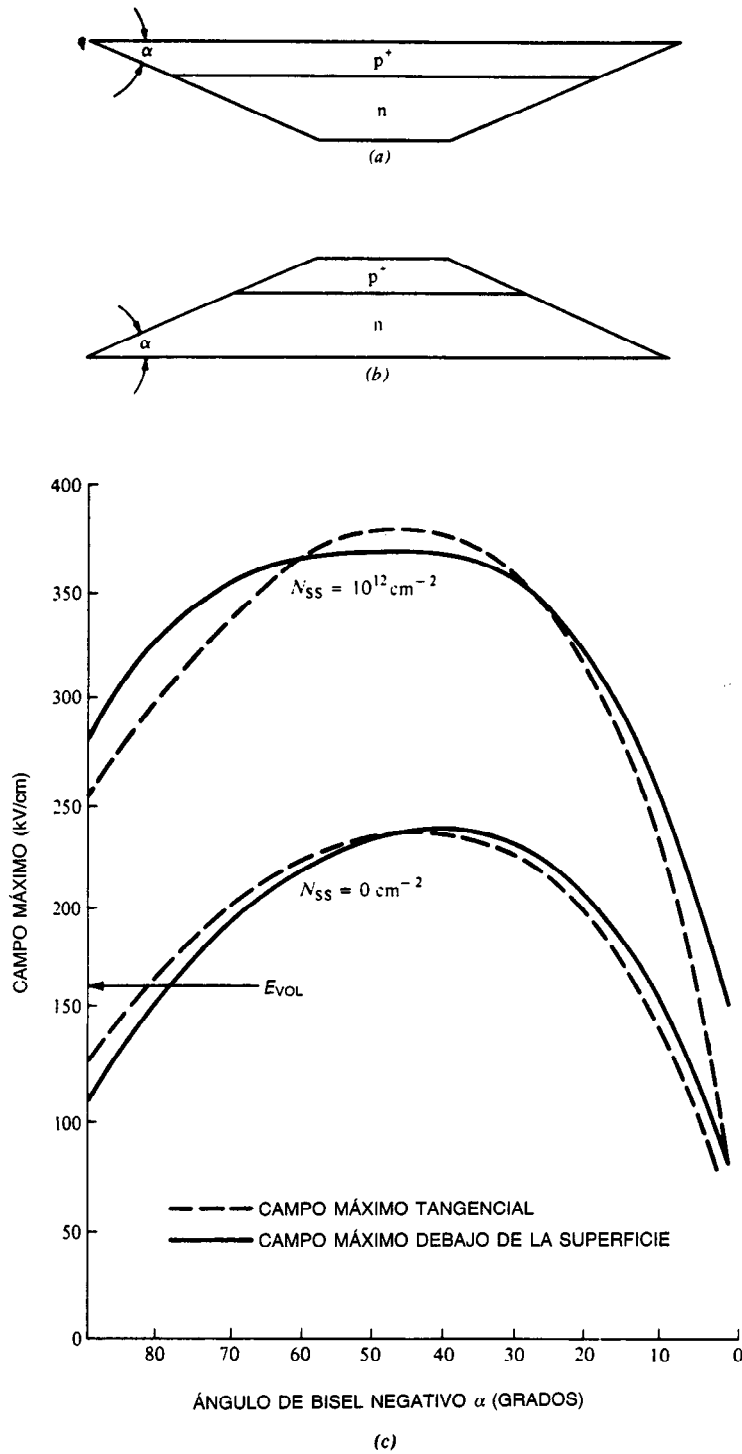


Fig. 13-17. a) Unión positiva biselada (la región altamente impurificada tiene mayor área que la región de baja impurificación); b) unión negativa biselada (la región altamente impurificada tiene menor área que la región de baja impurificación); c) máximo campo eléctrico en función del ángulo del bisel para dos valores de carga superficial.

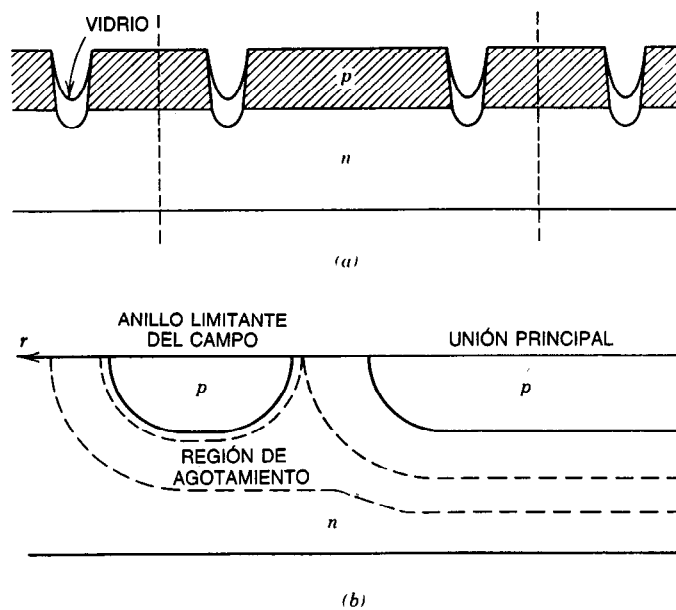


Fig. 13-18. a) Diodo pasivado con vidrio, con muescas en "foso". b) Unión pn con anillo limitante de campo.

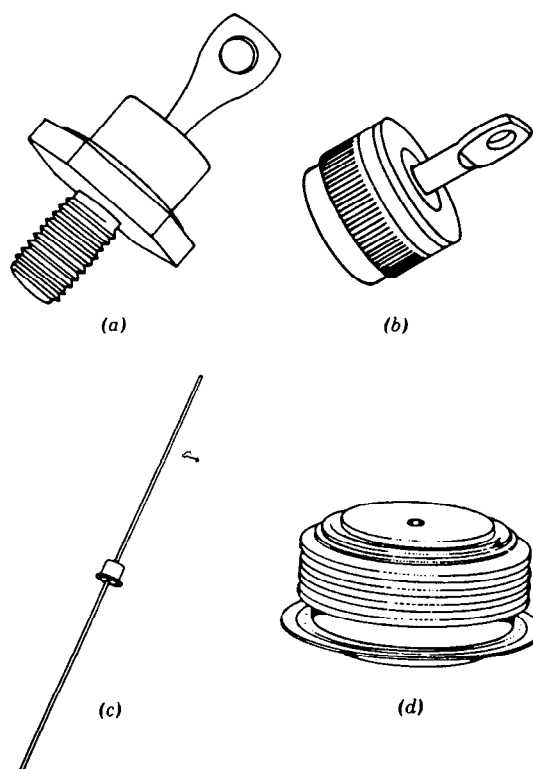


Fig. 13-19. Tipos de empaques de rectificadores de potencia: a) tipo perno; b) de base plana (o de ajuste a presión); c) de conductor axial; d) de contacto a presión.

forma, las temperaturas excesivas pueden provocar autodestrucción. La transferencia de calor del semiconductor es crítica en todos los dispositivos de potencia, por lo que se emplean encapsulamientos de cobre y se tiene bastante cuidado a fin de conectar uniformemente el semiconductor a éste. Durante muchos años se han construido rectificadores de alta potencia (> 35 A) mediante el empleo de placas metálicas refractarias para aliviar esfuerzos entre el diodo de silicio y el encapsulamiento de cobre. El coeficiente de dilatación lineal de estas placas de alivio de esfuerzos se aproxima al del silicio, de modo que cuando el dispositivo se somete a ciclos térmicos se minimiza el esfuerzo sobre el silicio.

Los rectificadores de alta corriente también presentan un bisel en el borde de la unión a fin de reducir el campo eléctrico en la superficie del borde del semiconductor, con lo cual se minimiza la posibilidad de ruptura local en la superficie (fig. 13-17).

Los dispositivos de baja corriente suelen tener una ranura pasivada con un vidrio (fig. 13-18a), o bien presentan anillos de campo⁹ que distribuyen el campo por la superficie (fig. 13-18b).

Los empaques de los rectificadores de potencia son ya sea del tipo perno (fig. 13-19a), de base plana (fig. 13-19b) o, en el caso de dispositivos de menos de 3 A, de conductor axial (fig. 13-19c). Para corrientes mayores de 200 A, con frecuencia se utilizan empaques de contacto a presión (fig. 13-19d).

Una de las principales preocupaciones al usar cualquier empaque es asegurar que la resistencia térmica entre empaque y sumidero de calor se minimice. La mayor parte de los fabricantes especifican la máxima resistencia térmica entre la unión y el recipiente. De este modo, al conocer la potencia generada en el semiconductor, es posible hacer una estimación del límite superior de temperatura del empaque. Para un rectificador con capacidad normal de 5 A, caída directa de 1 V y resistencia térmica de 7 °C/W entre unión y recipiente, el dispositivo debería experimentar un incremento en la temperatura del empaque de $5 \times 1 \times 7 = 35$ °C. Si la capacidad nominal del dispositivo está dada a una temperatura máxima de operación en la unión de 200 °C, entonces el empaque debe mantenerse a menos de 165 °C cuando a través del sumidero de calor se disipan 5 W de potencia. Con una temperatura ambiente de T_A , la resistencia térmica del empaque al ambiente debe ser menor de $R_{\theta CA} = (165 - T_A)/5$ W. Para una temperatura ambiente máxima de 50 °C, $R_{\theta CA} \leq 25$ °C/W.

REFERENCIAS BIBLIOGRÁFICAS

1. E. S. Yang, *Fundamentals of Semiconductor Physics*, Cap. 4, McGraw-Hill, Nueva York.
2. S. K. Ghandi, *Semiconductor Power Devices*, Cap. 2 y 3, Wiley, Nueva York.
3. S. M. Sze, *Physics of Semiconductor Devices*, Cap. 3, Wiley, Nueva York.
4. B. G. Streetman, *Solid State Electronics Devices*, Sec. 5.5.3, Prentice-Hall, Englewood Cliffs, NJ.
5. D. A. Fraser, *The Physics of Semiconductor Devices*, Oxford University Press, Nueva York.
6. P. E. Gray y C. L. Searle, *Electronic Principles—Physics, Models, and Circuits*, Sec. 6.5, Wiley, Nueva York.
7. B. G. Streetman⁴, Sec. 5.5.2.
8. S. M. Sze³, Cap. 4.
9. Y. C. Kao y E. D. Wolley, "High-Voltage Planar p-n Junctions", *Proc. IEEE* **55**(8):1409-1414.
10. *Transient Voltage Suppression Manual*, General Electric Company, Syracuse, NY.
11. J. F. Gibbons, *Semiconductor Electronics*, caps. 5-8, McGraw-Hill, Nueva York.
12. A. Blicher, *Thyristor Physics*, caps. 14 y 15, Springer-Verlag, Nueva York.
13. M. S. Adler y V. A. Temple, *Semiconductor Avalanche Breakdown Design Manual*, General Electric Company, Schenectady, NY.
14. A. S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, Nueva York.
15. *Semiconductor Data Handbook*, General Electric Company, Syracuse, NY.

13.2 TRANSISTORES BIPOLARES

Mark B. Barron

13.2.1 Acción de transistor

Amplificación de corriente y flujo de portadores

En los artículos dedicados a los diodos de unión pn se observó que cuando un diodo está polarizado en sentido directo conduce corriente conforme a la expresión:

$$I = I_0 e^{qV/nkT} \quad (13-22)$$

donde I_0 = constante

V = voltaje directo aplicado

n = factor de "idealidad", cuyo valor suele aproximarse a 1

Cuando la unión directamente polarizada se coloca cerca (en una vecindad de micrones) de una unión con polarización inversa, es posible obtener un dispositivo con tres terminales, denominado transistor. En la figura 13-20a se ilustra un semiconductor tipo p colocado entre dos semiconductores tipo n. Si la polarización de la región 3 es negativa con respecto a la región 2, y la de ésta es positiva con respecto a la región 1, los electrones inyectados desde la región 1 experimentarán un potencial más positivo en la terminal 3 que en la terminal 2. En consecuencia, los electrones emitidos de la región 1 a la región 2 serán barridos a la región 3. Cambiando el voltaje de polarización a través de la unión entre las regiones 1 y 2 es posible hacer que cambien la corriente en la región 3. Por razones históricas, a la región 2 se la denomina base, a la región 1, emisor, y a la región 3, colector. Para que esta disposición produzca un dispositivo efectivo de amplificación, la corriente en la terminal de la base debe ser pequeña con respecto a la corriente total del emisor. El número de huecos inyectados de la base al emisor es

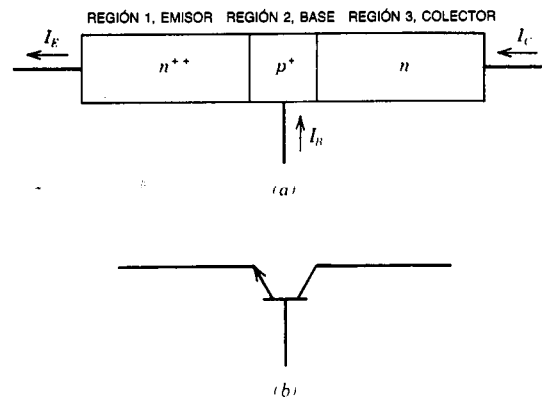


Fig. 13-20. a) Transistor *npn*; b) símbolo del transistor *npn*.

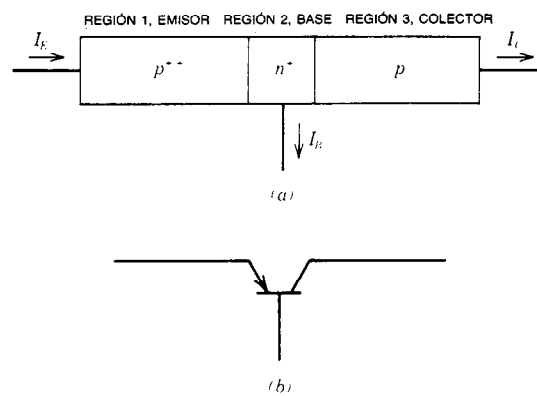


Fig. 13-21. a) Transistor *pnp*; b) símbolo del transistor *pnp*.

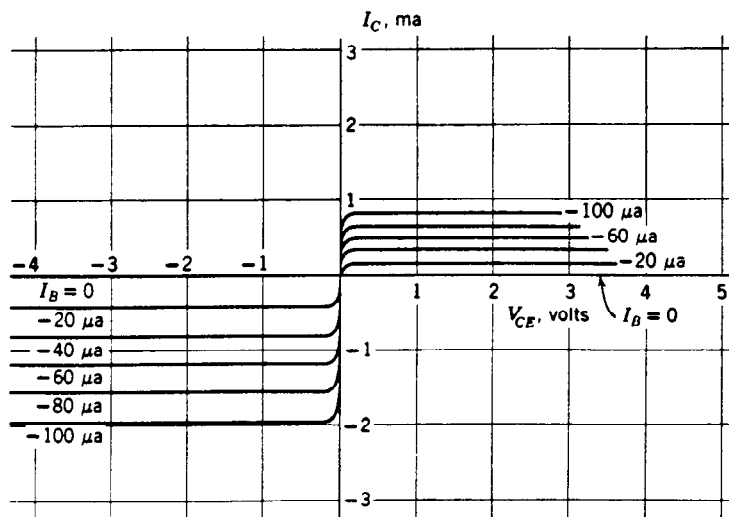


Fig. 13-22. Características corriente-voltaje de un transistor *pnp* ideal.

proporcional al inverso de la concentración de impurezas en el emisor, N_E , y el número de electrones inyectados del emisor a la base es proporcional al inverso de la concentración de impurezas en la región de la base, N_B . Por consiguiente, para la máxima razón de inyección de electrones del emisor sobre huecos de la base, el emisor debe estar fuertemente impurificado y la base debe estarlo sólo ligeramente. La razón de número de electrones inyectados desde el emisor sobre corriente total a través de la unión se denomina eficiencia de inyección, γ :

$$\gamma = \frac{I_n}{I_p + I_n} \quad (13-23)$$

npn y pnp. El análisis anterior se dedicó a un emparejado de tres capas: n, p, n . En el modo normal de operación, el colector tipo n está polarizado positivamente con respecto a la base y al emisor, y cuando se aplica una polarización positiva a la base con respecto al emisor, los electrones fluyen hacia la terminal del emisor o, de manera equivalente, fluye corriente positiva hacia afuera del emisor. Por tanto, la representación simbólica de un dispositivo npn denota corriente hacia afuera del emisor (fig. 13-20b). Para construir un transistor también puede emplearse una estructura alternativa con un emisor y un colector tipo p y una base tipo n . Esta estructura pnp opera en el modo normal con el colector polarizado negativamente, de modo que al aplicar un voltaje de base ligeramente negativo los electrones fluirán fuera del emisor, lo que representa una corriente positiva hacia la terminal. En la figura 13-21 se representa una estructura pnp .

Características del dispositivo ideal

En semiconductores con largo tiempo de vida de los portadores y base estrecha, los electrones se difunden inalterados a través de la región de la base. En semiconductores con corto tiempo de vida de los portadores, base amplia o ambas cosas, los electrones se recombinan con los huecos en la región de la base, y se requiere corriente extra en la terminal de la base a fin de proporcionar estos huecos. Por consiguiente, los mejores transistores de elevada ganancia tienen base estrecha y largo tiempo de vida de los portadores. El comportamiento de los transistores de este tipo está descrito por las siguientes ecuaciones:

$$I_E = I_{ES} (e^{qV_{EB}/kT} - 1) - \alpha_R I_{CS} (e^{qV_{CB}/kT} - 1) \quad (13-24)$$

$$I_C = -\alpha_F I_{ES} (e^{qV_{EB}/kT} - 1) + I_{CS} (e^{qV_{CB}/kT} - 1) \quad (13-25)$$

$$I_B = -(I_E + I_C) \quad (13-26)$$

donde α_R , α_F , I_{ES} , I_{CS} son constantes que dependen de las características del material semiconductor y de las propiedades dimensionales del dispositivo. En la figura 13-22 se muestran las características VI asociadas con las ecuaciones 13-24 y 13-25.

Regiones de operación

Un transistor posee cuatro regiones básicas de operación, ya que cada una de las dos uniones puede tener polarización directa o inversa.

Región de corte. Cuando la polarización de ambas uniones es negativa, sólo fluye corriente de fuga, por tanto:

$$I_E = -I_{ES} + \alpha_R I_{CS} \quad (13-27)$$

$$I_C = \alpha_F I_{ES} - I_{CS} \quad (13-28)$$

Región activa normal. En este caso la unión emisor-base tiene polarización directa y la unión base-colector tiene polarización inversa. Si la unión base-colector está polarizada con un voltaje suficientemente grande, las corrientes están dadas por:

$$I_E = -I_{ES} (e^{qV_{EB}/kT} - 1) + \alpha_R I_{CS} \quad (13-29)$$

$$I_C = -\alpha_F I_E - (1 - \alpha_F \alpha_R) I_{CS} \quad (13-30)$$

Región inversa. Aquí, la unión base-colector tiene polarización directa y la unión base-emisor tiene polarización inversa. En estas condiciones:

$$I_E = -I_{ES} - \alpha_R I_{CS} (e^{qV_{CB}/kT} - 1) \quad (13-31)$$

$$I_C = -\frac{I_E}{\alpha_R} + \alpha_F I_{ES} \left(1 - \frac{1}{\alpha_R \alpha_F}\right) \quad (13-32)$$

Región de saturación. Ambas uniones tienen polarización directa. Hasta donde concierne a las corrientes en las terminales, es posible concebir la corriente en la región de saturación como una superposición de funcionamiento en la región normal activa y en la región inversa. En esta última el flujo de corriente está caracterizado por las ecuaciones 13-24 a 13-26.

13.2.2 Parámetros de los transistores

En la mayor parte de las aplicaciones de los transistores, el emisor es el electrodo de referencia (tierra) y la base es el electrodo de control (fig. 13-23). En esta configuración de "emisor común", la corriente del colector es una variable dependiente, determinada por la corriente de la base. La razón de corriente del colector sobre corriente de la base es la "ganancia" de corrientes del dispositivo, que suele representarse con la letra griega β . Es posible demostrar, basándose en las ecuaciones 13-29 y 13-30, que la ganancia de CA está relacionada con la constante α_F por la siguiente ecuación:

$$\beta = \frac{\alpha_F}{1 - \alpha_F} \quad (13-33)$$

En la mayor parte de los transistores modernos, α_F se aproxima bastante a 1, de modo que la ganancia de corriente puede ser muy grande. Los valores comunes son mayores de 100.

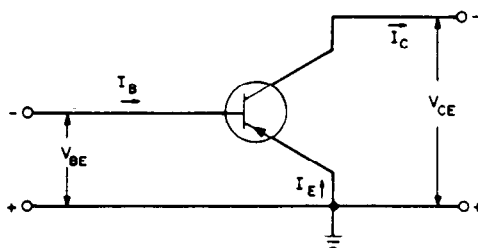


Fig. 13-23. Transistor *pn*p en configuración de emisor común.

Parámetros *h*

Con frecuencia, los transistores están caracterizados por parámetros *h* que representan las ganancias directa e inversa, así como la impedancia y la admitancia de entrada y salida. El primer subíndice denota el parámetro, con letra mayúscula para los parámetros de CD y con letra minúscula para los parámetros de CA. El segundo subíndice denota emisor común, base común o colector común. En el caso del emisor común $i_c/i_b = h_{fe}$ es la ganancia de corriente CA en sentido directo, que es igual a β .

$$h_{re} = v_{be}/v_{ce}$$

es la ganancia de corriente inversa, $v_{be}/i_b = h_{ie}$ es la impedancia de entrada y $h_{oe} = i_c/v_{ce}$ es la admitancia de entrada.

Ganancia contra corriente

Las ecuaciones ideales 13-24 a 13-26 indican que h_{fe} es una constante igual a $\alpha_F/(1 - \alpha_F)$. Sin embargo, en los transistores reales la ganancia de corriente disminuye para corrientes muy bajas y altas (fig. 13-24). La razón de este hecho es que, a bajas corrientes, las imperfecciones superficiales del semiconductor provocan que los portadores inyectados del emisor se recombinen en la superficie con los portadores de la base, por lo que no hay oportunidad de que sean captados por el colector. Para altas corrientes, los huecos y electrones se recombinan con mayor frecuencia en la base debido a choques entre portadores, lo que reduce el tiempo de vida. También la eficiencia de inyección disminuye debido a los efectos de modulación de la conductividad.

Además, a altas corrientes existe una caída de voltaje resistivo entre el conductor de la base y la parte central de la región del emisor (fig. 13-25). Por consiguiente, en medio del emisor la unión emisor-base tiene menor polarización directa que el borde. Este efecto reduce esencialmente el área efectiva portadora de corriente del transistor.

Corrientes de fuga

Las ecuaciones para el transistor ideal indican que para $V_{EB} = 0$, la corriente del colector es igual a $-I_{CS}$ para altos voltajes inversos en el colector. Para $I_B = 0$, I_C es:

$$I_C|_{I_B=0} = \frac{-I_{CS} + \alpha_F \alpha_R I_{CS}}{1 - \alpha_F} \quad (13-34)$$

En los transistores reales existe un exceso de corriente de fuga asociado con las imperfecciones cristalinas en la región superficial de las uniones. Tales corrientes suelen indicarse en la hoja de especificaciones del vendedor como I_{CBO} , corriente colector-base con el emisor abierto, o como I_{CEO} , corriente colector-emisor con la base abierta. De manera aproximada, la relación entre I_{CEO} e I_{CBO} es:

$$I_{CEO} = h_{FE} I_{CBO} \quad (13-35)$$

De este modo, la corriente de fuga de la unión colector-base es aumentada significativamente por la ganancia del transistor. En especial a elevadas temperaturas, la disipación significativa de potencia puede asociarse con la corriente de fuga.

Ganancias contra frecuencia

Los cambios de voltaje base-emisor en un transistor requieren un tiempo finito para propagarse hacia la unión base-colector, ya que los portadores de la región de la base tienen que recombinarse antes de alcanzar un estado estable. Así, a elevadas frecuencias la ganancia de corriente se degrada. También se degrada por efecto de la capacitancia parásita base-emisor, que deriva corriente de la base directamente al emisor.

En la figura 13-26 se muestran h_{fe} y h_{fb} en función de la frecuencia en un transistor de uso general. Obsérvese que mientras que la ganancia del emisor común a baja frecuencia h_{fe} es mucho mayor que la ganancia de la base común h_{fb} , el punto de atenuación a 3 dB $f_{h_{fb}}$ para la base y configuración comunes se encuentra a una frecuencia mucho más elevada que el punto a 3 dB para un emisor común $f_{h_{fe}}$. Por consiguiente, para aplicaciones de alta frecuencia, la base común ofrece algunas ventajas. Es interesante observar que la frecuencia a la que h_{fe} es igual a 1 (f_T) se aproxima bastante a $f_{h_{fb}}$:

$$\frac{1}{2\pi f_T} = \frac{kT}{qI_C} C_{IE} + t_B + r_C C_{IC} + \frac{x_C}{V_{lim}} \quad (13-36)$$

donde C_{IE} = capacitancia promedio del emisor

C_{IC} = capacitancia promedio de transición del colector

t_B = tiempo de tránsito de la base

r_C = resistencia del cuerpo del colector

x_C = ancho de la capa de agotamiento colector-base

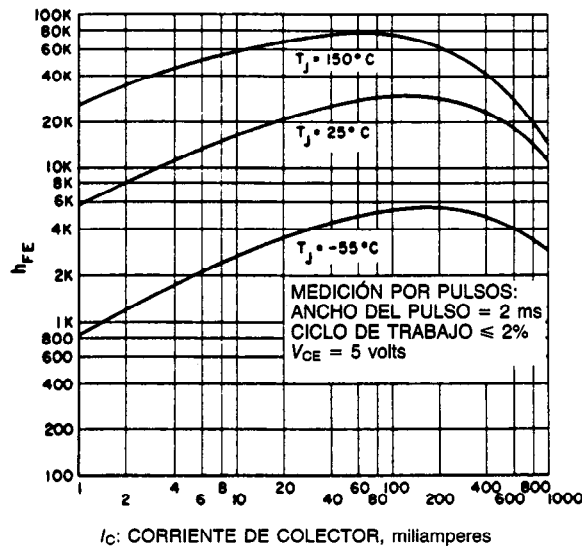


Fig. 13-24. Ganancia de corriente del transistor contra corriente de colector.

V_{lim} = velocidad de deriva de emisión de electrones a través de la capa de agotamiento del colector

Respuesta transitoria

Como ya se mencionó, cuando la corriente de la base cambia repentinamente, se requiere un tiempo antes de que la corriente del colector alcance un estado estable. En aplicaciones digitales o de conmutación de potencia, el transistor a menudo está en el modo de saturación cuando se enciende a fin de minimizar las pérdidas por conducción en los dispositivos. En el modo de saturación, el exceso de carga se crea en la base debido a que la unión base-colector está polarizada directamente. Cuando, de forma repentina, la corriente de la base se reduce a cero, la corriente continúa fluyendo al mismo valor hasta que el exceso de carga se elimina de la base. Este retraso se caracteriza por un tiempo de almacenamiento, t_s . Una vez que se ha eliminado el exceso de carga de la base, la corriente del colector disminuye, y el tiempo necesario para que se haga cero es el tiempo de descenso, t_f . Cuando se vuelve a aplicar la corriente a la base, la corriente del colector no aumenta de manera instantánea, sino que requiere un intervalo de tiempo denominado tiempo de ascenso t_r .

Con referencia a la figura 13-27, si el transistor se coloca en un circuito con una resistencia de carga R_L , el tiempo de conmutación para el transistor está dado por las siguientes ecuaciones:

$$t_r = h_{fe} \left(\frac{1}{6f_r} + 1.7 R_L C_{IC} \right) \ln \left(\frac{h_{fe} I_{Benc}}{h_{fe} I_{Benc} - 0.91 I_C} \right) \quad (13-37)$$

$$t_f = h_{fe} \left(\frac{1}{6f_r} + 1.7 R_L C_{IC} \right) \ln \left(\frac{I_C - h_{fe} I_{Bap}}{0.1 I_C - h_{fe} I_{Bap}} \right) \quad (13-38)$$

$$t_s = \tau_s \ln \left(\frac{I_{Benc} - I_{Bap}}{I_C/h_{fe} - I_{Bap}} \right) \quad (13-39)$$

donde C_{IC} = capacitancia de transición del colector
 I_{Bap} = corriente forzada hacia afuera de la base durante el apagado
 I_{Benc} = corriente forzada hacia adentro de la base durante el encendido
 τ_s = tiempo de almacenamiento característico
 Para un transistor dado, es posible mejorar la rapidez de conmutación, impidiendo que el dispositivo se

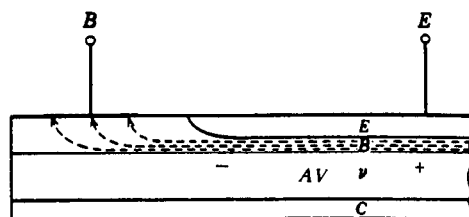


Fig. 13-25. Caída de voltaje resistivo entre el centro del emisor y el borde del emisor, debida a la corriente lateral de la base.

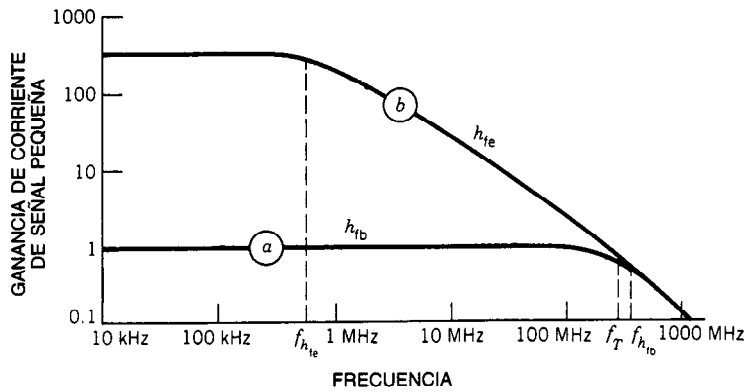


Fig. 13-26. Ganancia en función de la frecuencia para configuraciones con base común (h_{fb}) y emisor común (h_{fe}).

sature, lo que significa $t_s \approx 0$. Esto puede lograrse derivando la unión colector-base con un diodo de germanio o Schottky que conduzca la corriente a un voltaje más bajo que cuando se emplea silicio.

En los casos en que la unión base-colector se polariza directamente a medida que el transistor entra en saturación, el voltaje es fijado por los diodos de derivación y el exceso de carga se desvía a la base del transistor.

La desventaja de un diodo de fijación es que la caída de voltaje del transistor será mayor que la caída de uno en verdadera saturación.

La velocidad de conmutación del transistor también mejora durante el proceso de manufactura mediante la degradación del tiempo de vida de los portadores, la difusión de metales pesados como oro o platino dentro del semiconductor, o la irradiación con electrones de alta energía.

Sin embargo, el precio que hay que pagar por hacer esto es que las corrientes de fuga se incrementan y la ganancia de corrientes disminuye.

Características de disrupción

El mecanismo de descarga disruptiva en los transistores es igual que en los diodos. La diferencia más importante es que el efecto de ganancia en un transistor de base abierta significa que la corriente de fuga que proviene de la unión colector-base se amplifica a medida que pasa a través de la unión emisor-base. Por ello, el efecto de transistor incrementa el efecto de multiplicación en avalancha, esencialmente a base de reducir el voltaje al que $M \rightarrow \infty$.

En la figura 13-28 se muestran las características VI de un transistor en distintas condiciones de polarización de la base. Obsérvese que el voltaje disruptivo con la base abierta, BV_{CEO} , es sustancialmente menor que el voltaje disruptivo con el emisor abierto, BV_{CBO} . De hecho:

$$BV_{CEO} = \frac{BV_{CBO}}{(h_{fe})^{1/n}} \quad (13-40)$$

donde $n \approx 4$ para transistores nnp , y $n \approx 6$ para los pnp .

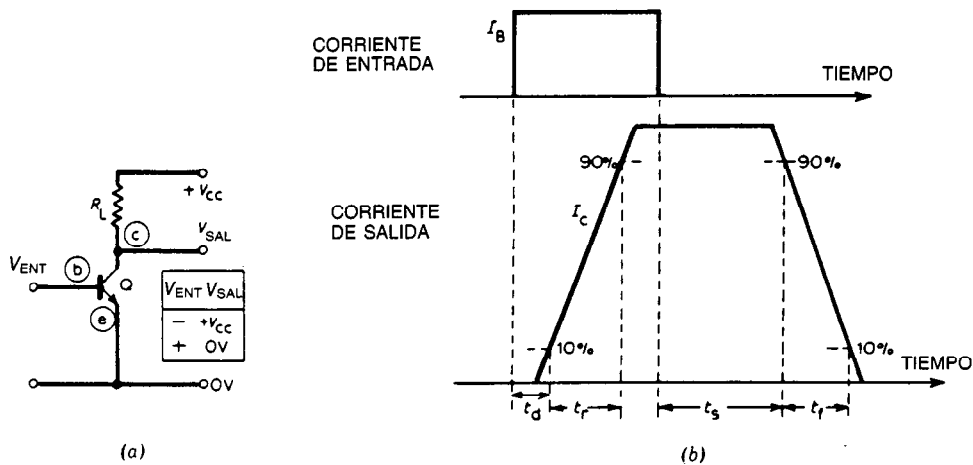


Fig. 13-27. Conmutación de un transistor: a) disposición del circuito para conmutación; b) corriente transitoria durante la conmutación.

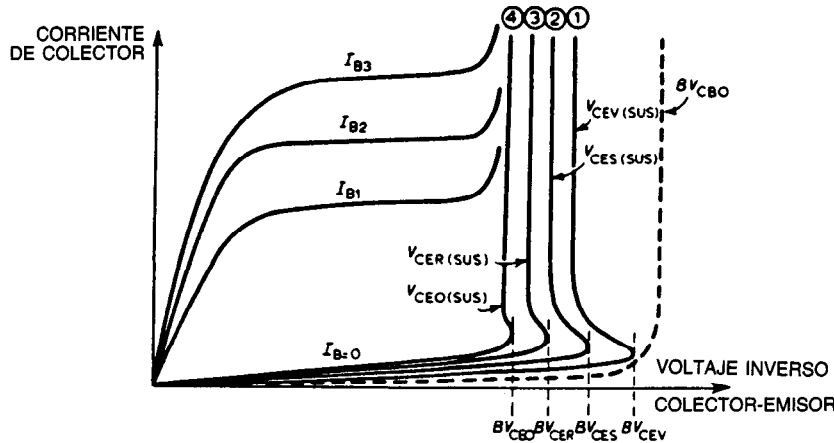


Fig. 13-28. Disrupción en avalancha de un transistor en diferentes condiciones de polarización de la base. 1) V_{CEV} : emisor con la base en polarización inversa; 2) V_{CES} : cortocircuito base-emisor; 3) V_{CER} : resistencia externa en paralelo con la unión base-emisor; 4) V_{CEO} : base en circuito abierto.

El voltaje disruptivo con la base en cortocircuito, BV_{CES} , es aproximadamente igual a BV_{CBO} a bajas corrientes del colector, pero para corrientes más elevadas, la resistencia interna en serie con la base permite que la unión base-emisor adquiera polarización directa, con lo que la curva de BV_{CES} se aproxima a la de BV_{CEO} . Dado que la mayor parte de las aplicaciones tienen una impedancia de activación finita en la base, debe tenerse cuidado cuando se opere un transistor con el voltaje del colector en la vecindad de BV_{CEO} . Para que sea segura, la operación debe diseñarse de modo que ocurra bastante por debajo de BV_{CEO} , con algún margen de seguridad por debajo de los valores probados.

Efectos de la temperatura

Como en un diodo de unión pn , el aumento de la temperatura puede provocar que las corrientes de fuga se incrementen. Dado que la corriente de fuga de la unión base-colector con polarización inversa I_{CBO} es amplificada por la ganancia de corriente del transistor (ec. 13-35), es posible generar una cantidad significativa de potencia mediante la corriente de fuga I_{CEO} que desciende a través de la caída de potencial de V_{CE} .

El aumento de la temperatura también eleva la ganancia de corriente, hasta el punto en que el número de portadores generados térmicamente en la región de la base empieza a ser igual a la concentración de impureza en la base. Entonces, el transistor empieza a perder sus características de amplificación. Dado que la degradación del transistor ocurre a cierta temperatura de unión $T_{j\max}$, la potencia que puede disiparse en el transistor está dada por:

$$P_{dis} = \frac{T_{j\max} - T_{amb}}{R_{\theta jc}} \quad (13-41)$$

donde T_{amb} es la temperatura ambiente de la envoltura del transistor y $R_{\theta jc}$ es la resistencia térmica entre la

unión del transistor y la envoltura (en los artículos 13.1 y 13.2 puede consultarse un análisis de la resistencia térmica).

Características del transistor real

Las características del transistor ideal que se muestran en la figura 13-22 cambian significativamente durante la operación del transistor real (fig. 13-29).

Los fenómenos físicos responsables de los cambios son:

1. La caída del voltaje resistivo a través del colector cuando fluye la corriente.
2. El estrechamiento de la región efectiva de la base a medida que la región de agotamiento del colector penetra en la región de la base.
3. La caída de voltaje resistivo lateralmente a través de la base, que provoca que la corriente inyectada del emisor se acumule hacia el borde del emisor.

En los transistores de mayor voltaje, la región del colector está ligeramente impurificada. En la región de operación de 0 a A' el transistor está altamente saturado y la unión base-colector, bajo fuerte polarización directa, inyecta muchos portadores en la región del colector, con lo que se reduce la resistencia del colector. En la región de A' a A , la unión base-colector se hace menos fuertemente polarizada en sentido directo y hay menos modulación de la conductividad en el colector.

De esta forma, la resistencia del cuerpo del colector aumenta y se dice que el transistor se encuentra en casi-saturación. En el punto A de la curva, el transistor deja el estado de saturación y la unión base-colector se polariza inversamente. Puede observarse que la casi-saturación en un transistor reduce significativamente la ganancia de corriente a bajos voltajes de colector-emisor.

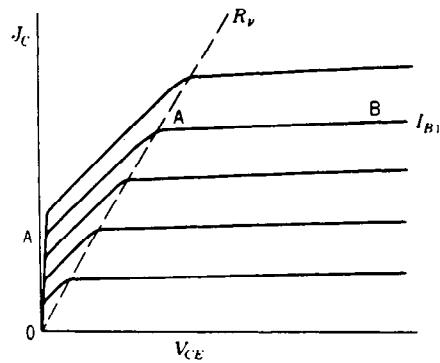


Fig. 13-29. Características de un transistor real.

Disrupción secundaria

Otro fenómeno que limita el rendimiento del transistor es la **disrupción secundaria**, que es la usualmente destructiva disrupción de la capacidad del transistor de soportar el voltaje, y que se encuentra debajo del valor normalmente asociado con la disrupción en avalancha. Por lo general se observan dos tipos de disrupción secundaria: la disrupción secundaria con polarización di-

recta, en la cual la corriente de base es positiva y enciende el transistor que se encontraba bloqueando altos voltajes; y la disrupción secundaria con polarización inversa, en la que el transistor se encuentra en una condición de "encendido" y la unión base-emisor se polariza inversamente para apagar el transistor. La disrupción secundaria suele estar asociada con regiones localizadas de exceso de calor debido a una resis-

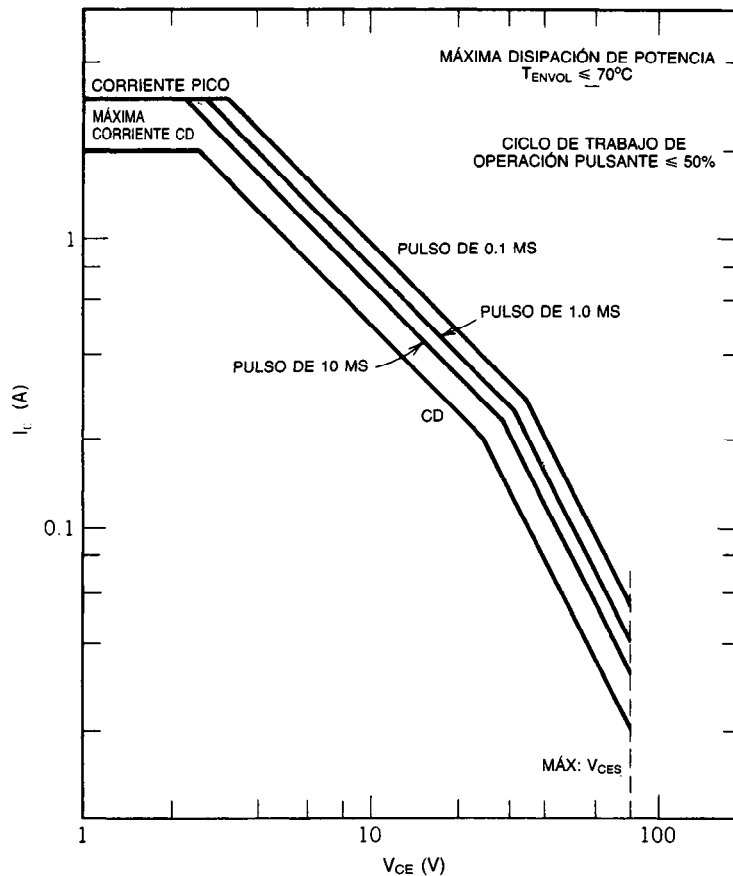


Fig. 13-30. Área segura de operación (SOA) de un transistor.

tencia térmica anormalmente alta (p. ej., sobre huecos de la soldadura) o debido a la acumulación localizada de corriente. La acumulación de corriente puede ocurrir durante el apagado cuando los portadores en la base se remueven primero cerca del borde del emisor, forzando que la corriente se estreche bajo el centro del emisor. También puede ocurrir durante el encendido, debido a que la periferia del emisor se enciende antes que el resto de dicho dispositivo y es posible disipar grandes cantidades de potencia en el borde del emisor. El efecto de la disrupción secundaria es reducir el área segura de operación (SOA, de *safe operating area*) dentro de la cual el transistor puede operarse confiablemente. En la figura 13-30 se proporciona un ejemplo de región segura de operación para un transistor.

13.2.3 Tipos de transistores de potencia

La construcción de los transistores varía con la edad del tipo de dispositivo y con las aplicaciones que se pretende dar al dispositivo.

De difusión única

Probablemente la construcción más simple es la del transistor de difusión única, que se fabrica mediante la difusión simultánea de los impurificadores del emisor y del colector dentro de una rebanada delgada de material base. Este proceso está limitado a los dispositivos de bajo voltaje, ya que la región de agotamiento de base-colector se extiende dentro de la región de la base. La gruesa región de la base que se requiere para fabricar este transistor da por resultado una ganancia de corriente relativamente baja en esta estructura. La gruesa región de la base combinada con la ausencia de un campo de deriva (o campo interno) en la región de la base uniformemente impurificada también hace que esta estructura sea algo más lenta.

De triple difusión

Para dispositivos de mayor voltaje, como material inicial se emplea material tipo n (v) de alta resistividad. Con él se forma la región del colector masivo, que soporta el voltaje de la unión base-colector. Después el colector de baja resistencia n^+ se difunde dentro del material, con la superficie superior protegida. Después de la difusión se realiza una difusión de base p^+ en la superficie superior y una difusión de emisor n^+ . La región de agotamiento asociada con la unión base-colector con polarización inversa se encuentra principalmente en la región v . De este modo la región de la base puede estrecharse mucho más que en el transistor de difusión única. La base más delgada combinada con el campo de deriva asociado con la base difundida hace que el dispositivo de triple difusión sea más rápido que el de difusión única. Sin embargo, las dislocaciones cristalinas asociadas con la difusión de la base, así como el almacenamiento de carga en la región v durante el apagado, tienden a hacer al transistor de triple difusión menos fuerte eléctricamente que el dispositivo de difusión única. Es decir, la disrupción secundaria puede ser más problemática en el transistor de triple difusión que en el de difusión única.

Desde el punto de vista de la manufactura, ambas estructuras son poco atractivas, en virtud de que el material semiconductor inicial debe ser delgado a fin de mantener bajas las caídas de voltajes resistivos de la base y del colector.

Epitaxial de doble difusión

Para facilitar la manufactura de los transistores, con frecuencia se emplea silicio grueso de baja resistividad como material de inicio, y sobre la parte superior de este sustrato se desarrolla epitaxialmente una capa de mayor resistividad. A continuación, las regiones de la base y el emisor se difunden dentro de la capa epitaxial. Los parámetros eléctricos son semejantes a los de los transistores de triple difusión, excepto que el máximo voltaje nominal es más bajo en el dispositivo epitaxial. La razón de esta capacidad reducida es que en las capas epitaxiales resulta difícil lograr resistividades tan altas como las del silicio macizo.

Base epitaxial

Un proceso más antiguo que aún se emplea para los dispositivos de bajo voltaje es iniciar con un sustrato n^+ grueso, como en el dispositivo de doble difusión. Después, sobre el sustrato se desarrolla epitaxialmente una capa n de alta resistividad, seguida de una capa p más densamente impurificada, que se emplea como la base. En dispositivos más antiguos, la capa p suele desarrollarse directamente sobre el sustrato n^+ sin la capa n extra. En cualquier caso, el emisor se difunde dentro de la capa epitaxial p .

La ventaja de la base epitaxial, como en el transistor de difusión única, es que la resistividad de la región de la base es superior a la de la base difundida. Por consiguiente, existe una resistencia extra entre los contactos de la base y las regiones del emisor, que ayuda a estabilizar los segmentos internos del transistor. Esto contribuye a asegurar la operación uniforme durante el encendido y el apagado, lo que a su vez reduce la posibilidad de disrupción secundaria.

13.2.4 Transistores Darlington

Para transistores de alta corriente, que por lo general tienen una ganancia de corriente (a corriente nominal) de sólo 10, es común colocar en cascada dos o más transistores (fig. 13-31). En un transistor Darlington la corriente de la primera etapa de transistor acciona la base de la segunda etapa. La ganancia combinada de corriente es la siguiente:

$$\frac{I_C}{I_{B1}} = \frac{I_{C1} + I_{C2}}{I_{B1}} = h_{FE1} + h_{FE2} + h_{FE1} + h_{FE2} \quad (13-42)$$

Los transistores Darlington monolíticos suelen construirse de modo que la unión base-emisor de los transistores sea derivada por resistores. Algunas veces también existe un diodo antiparalelo de derivación alrededor de la unión base-emisor de la primera etapa a fin de eliminar la carga de la segunda base durante el apagado.

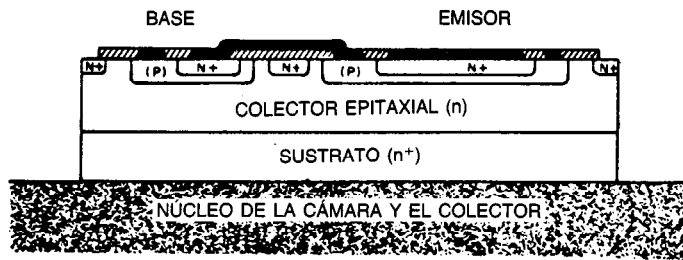


Fig. 13-31. Construcción de un transistor Darlington monolítico.

Si bien la ganancia de un transistor Darlington es considerablemente superior a la de un transistor de una sola etapa, debe reconocerse que para un nivel dado de corriente el transistor Darlington disipará más potencia que un dispositivo de una sola etapa. La razón de este hecho es que el V_{CEsat} de un transistor Darlington será cuando menos de 0.5 V, debido a que el emisor de la primera etapa debe tener por lo menos una caída de diodo por encima del emisor de la segunda etapa a fin de encender la segunda etapa. Por consiguiente, el colector de la primera etapa, que es común con la segunda etapa, debe estar a un voltaje cuando menos igual al de la unión base-emisor de la segunda etapa, que es aproximadamente de 0.5 V o más. Un transistor de una sola etapa, por otra parte, puede tener un V_{CEsat} cercano a cero para corrientes bajas.

13.2.5 Consideraciones para las aplicaciones de los transistores

Operación en paralelo

Si se requieren corrientes superiores a las que puede manejar un solo transistor, es frecuente conectar varios transistores en paralelo. Debe procurarse que los transistores se acoplen cuidadosamente, ya que de otra forma el dispositivo con el menor V_{CEsat} puede "acaparar" toda la corriente. Dado que los transistores poseen un coeficiente térmico positivo, el dispositivo sobrecargado puede sobrecalentarse y conducir aún más corriente en un proceso regenerativo, lo que a la larga puede conducir a su destrucción.

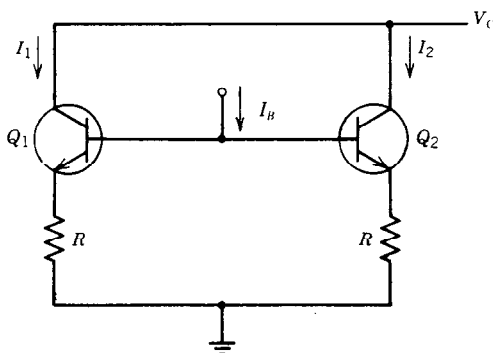


Fig. 13-32. Estabilización del emisor para equalizar la corriente en transistores en paralelo.

Una forma de equilibrar la carga a través de transistores en paralelo, además de probar y hacer equivaler las características V_I , consiste en colocar un resistor de bajo valor en el alambre externo del emisor de cada transistor (fig. 13-32). Con esta configuración existe una retroalimentación negativa. A medida que la corriente aumenta en cualquiera de los alambres externos, el voltaje se incrementa a través del resistor, reduciendo el voltaje de la unión base-emisor, lo que a su vez reduce la corriente a través del transistor "culpable".

Operación en serie

En algunas aplicaciones puede ser necesario que un transistor bloquee más voltaje del que le permite su voltaje nominal. En estos casos es posible conectar más de un transistor en serie si se toman ciertas precauciones; es decir, se se ecualizan el voltaje compartido por los transistores individuales en estado estable y en condiciones transitorias de operación. Debido a las

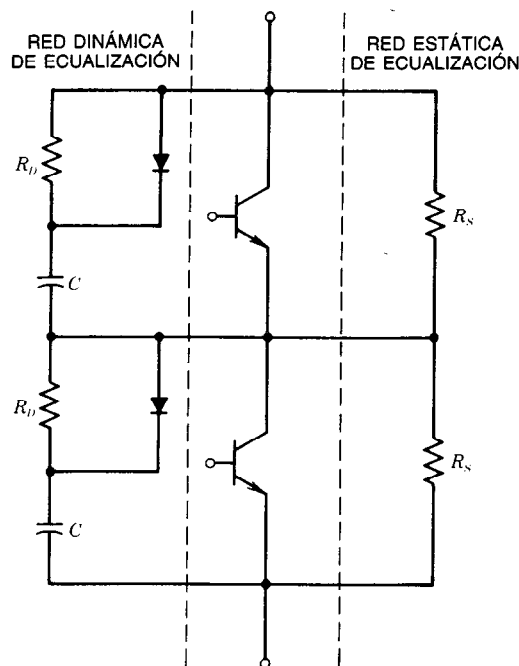


Fig. 13-33. Red de equalización para la conexión de transistores en serie.

diferencias en capacitancias de las uniones, tiempos de retardo y corrientes de fuga para los transistores individuales, se requieren redes externas de equalización y consideraciones especiales para los circuitos de accionamiento de la base. En la figura 13-33 se puede observar una de estas redes de equalización. Los resistores R , equalizan las corrientes de fuga estática. Los capacitores C equalizan la tensión durante el encendido y apagado, mientras que los resistores limitan la disipación de potencia en la red RC .

Redes de protección

La red RC a través de los transistores, que se presentó en el ejemplo de operación en serie del artículo anterior, se emplea algunas veces con transistores simples en aplicaciones en las que las oscilaciones momentáneas de voltaje pueden potencialmente llevar al transistor fuera de su área segura de operación. Por ejemplo, cargas inductivas como las de motores pueden crear oscilaciones momentáneas del voltaje si la corriente se interrumpe repentinamente. Los valores de R y C a menudo se determinan empíricamente para lograr la menor pérdida de potencia y la mayor rapidez de conmutación consistentes con las posibles oscilaciones en el circuito.

LIBROS Y REVISTAS

- F. C. Fitchen, *Transistor Circuit Analysis and Design*, Van Nostrand, Nueva York.
- S. K. Ghandi, *Semiconductor Power Devices*, cap. 4, Wiley, Nueva York.
- J. F. Gibbons, *Semiconductor Electronics*, cap. 9, McGraw-Hill, Nueva York.
- A. B. Glaser y G. E. Subak-Sharpe, *Integrated Circuit Engineering—Design, Fabrication, and Applications*, cap. 2, Addison-Wesley, Reading, MA.
- P. E. Gray y C. L. Searle, *Electronic Principles—Physics, Models, and Circuits*, cap. 7, Wiley, Nueva York.
- J. L. Moll, *Physics of Semiconductors*, McGraw-Hill, Nueva York.
- A. B. Phillips, *Transistor Engineering*, McGraw-Hill, Nueva York.
- *Power Transistor Users Guide*, General Electric Company, Syracuse, Nueva York.
- *SCR Manual*, cap. 6, General Electric Company, Syracuse, NY.
- C. L. Searle, A. R. Boothroyd, E. J. Angelo, P. E. Gray y D. D. Pederson, *SEEC Notes 2—Elementary Circuit Properties of Transistors*, Wiley, Nueva York.
- *Semiconductor Data Book*, General Electric Company, Syracuse, NY.
- B. G. Streetman, *Solid State Electronic Devices*, cap. 7 Prentice-Hall, Englewood Cliffs, NJ.
- S. M. Sze, *Physics of Semiconductor Devices*, cap. 6, Wiley, Nueva York.
- T. D. Towers y S. Libes, *Semiconductor Circuit Elements*, caps. 5-9, Hayden, Rochelle Park, NJ.
- E. S. Yang, *Fundamentals of Semiconductor Devices*, cap. 9, McGraw-Hill, Nueva York.

13.3 TRANSISTORES DE EFECTO DE CAMPO

Mark B. Barron

Existen dos tipos básicos de transistores de efecto de campo: con compuerta aislada y con compuerta de unión. Estas últimas pueden dividirse en uniones pn y en uniones Schottky.

13.3.1 Transistores de efecto de campo con compuerta aislada

Los transistores que prevalecen en las aplicaciones actuales son los pertenecientes a la variedad de compuerta aislada, esencialmente debido a que son los más fáciles de fabricar y aplicar en circuitos integrados. En un dispositivo con compuerta aislada se coloca una capa metálica en la parte superior de un aislante que se encuentra sobre la superficie de un semiconductor. Un voltaje aplicado entre el metal y el semiconductor establece un campo eléctrico a través del aislante y dentro del semiconductor. El campo eléctrico en la superficie del semiconductor atrae y repele huecos y electrones; el portador atraído depende de la polaridad del campo. Por ejemplo, si se aplica un voltaje positivo al metal, serán atraídos electrones a la superficie del semiconductor y los huecos serán repelidos. Si el semiconductor es tipo p , un pequeño voltaje positivo repelerá los huecos de la superficie del semiconductor, y el campo eléctrico será determinado por los átomos impurificadores aceptores ionizados. A medida que aumenta el voltaje en la superficie se agotan completamente los huecos, y los electrones en movimiento empiezan a acumularse en la superficie. Cuando la región superficial se puebla densamente de electrones, el semiconductor se "invierte". Si se han difundido o implantado regiones tipo n dentro del semiconductor tipo p en el borde del metal (fig. 13-34a) y se aplica un voltaje entre ellos, fluirá corriente entre estas regiones n si se ha invertido la superficie del semiconductor (fig. 13-34b). Cuando el voltaje positivo en el metal se reduce por debajo de un nivel umbral V_t , el semiconductor tipo p ya no se invierte y la corriente deja de fluir entre las regiones n . El valor del voltaje umbral está determinado por el espesor del aislante, la concentración de impurificador en el sustrato del semiconductor y la cantidad de carga atrapada que pueda estar presente en el aislante.

La mayor parte de los transistores con compuerta aislada se fabrican empleando dióxido de silicio como aislante. Así, la estructura cuyo uso prevalece en la actualidad es el transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET, de *metal oxide semiconductor field effect transistor*). En la figura 13-34 se muestra un MOSFET de "canal n ". El dispositivo consta de dos regiones tipo n formadas en un sustrato tipo p . Una de estas regiones se denomina *fuerza* (debido a que es la fuente de portadores cuando se "enciende" el transistor), y la otra se denomina *drenaje* (debido a que actúa como drenaje para hacer que fluyan los portadores). El metal situado sobre el óxido

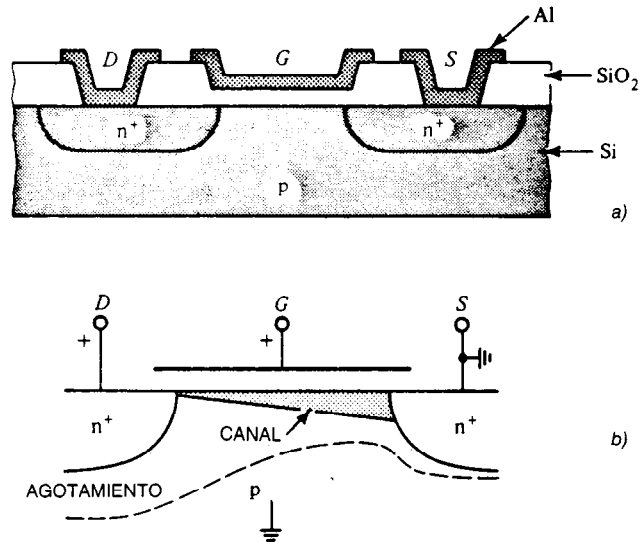


Fig. 13-34. a) Construcción de un transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET). S = fuente; D = drenaje; G = compuerta. b) Región de inversión (canal) en un MOSFET bajo polarización de voltaje.

se denomina *compuerta*, porque se emplea para controlar el paso de las señales de encendido y apagado del dispositivo. Un material común para la compuerta es el aluminio, aunque algunas veces se emplean otros, como el silicio policristalino, y metales refractarios (o siliciuros metálicos), como tungsteno o molibdeno.

Un MOSFET de canal *p* es idéntico al dispositivo de canal *n*, excepto que se emplea el material tipo *n* como sustrato inicial, y para formar la fuente y el drenaje se emplean regiones *p*. En un transistor de canal *p* se requiere un voltaje negativo para invertir la superficie del sustrato tipo *n*, y con ello permitir que la corriente fluya entre la fuente y un drenaje con polarización inversa.

La carga en la región de la superficie invertida de un dispositivo MOSFET puede deducirse con facilidad comparando el dispositivo con un capacitor sencillo. La carga en el capacitor es $Q = CV$. En un dispositivo MOSFET, la capacitancia de la compuerta por unidad de área cuando la superficie se encuentra fuertemente invertida es la siguiente:

$$C_g = \frac{\epsilon_{ox}}{x_o} \quad (13-43)$$

donde x_o es el grosor del óxido y ϵ_{ox} es la permitividad del óxido.

La carga móvil por unidad de área en la superficie del MOSFET es por consiguiente:

$$q = \epsilon_{ox} (V_g - V_t) \quad (13-44)$$

donde V_g es el voltaje del sustrato de la compuerta y V_t es el voltaje umbral que es necesario aplicar a fin de encender el dispositivo. Para bajos voltajes del drenaje

a la fuente ($V_D \ll V_g - V_t$), es posible demostrar que la corriente a través del canal del MOSFET es:

$$I_D = \frac{\mu_s \epsilon_{ox}}{x_o} \frac{W}{L} (V_g - V_t) V_D, \quad V_D \ll (V_g - V_t) \quad (13-45)$$

donde μ_s = movilidad superficial de los portadores en el canal

W = ancho de la región de la compuerta

L = longitud del canal entre la fuente y el drenaje

La movilidad superficial μ_s suele ser aproximadamente la mitad del valor en el material volumétrico del semiconductor, debido a la dispersión superficial de los portadores.

Cuando el voltaje del drenaje es mayor de algunas decenas de volts, la caída de tensión a lo largo del canal afecta el campo eléctrico a través del óxido. La corriente a través del transistor está dada por las siguientes ecuaciones:

$$|V_g - V_t| > V_D : I_D = \frac{\mu_s \epsilon_{ox} W}{x_o L} [(V_g - V_t) V_D - \frac{1}{2} V_D^2] \quad (13-46)$$

Región de no saturación

$$|V_g - V_t| < V_D : I_D = \frac{\mu_s \epsilon_{ox} W}{2 x_o L} (V_g - V_t)^2 \quad (13-47)$$

Región de saturación

En la figura 13-35 se muestra la corriente de drenaje en función del voltaje de drenaje para algunos voltajes de compuerta, y en la figura 13-36a se muestra la gráfica de la corriente de saturación en función del voltaje de compuerta. Puede observarse que la corriente varía

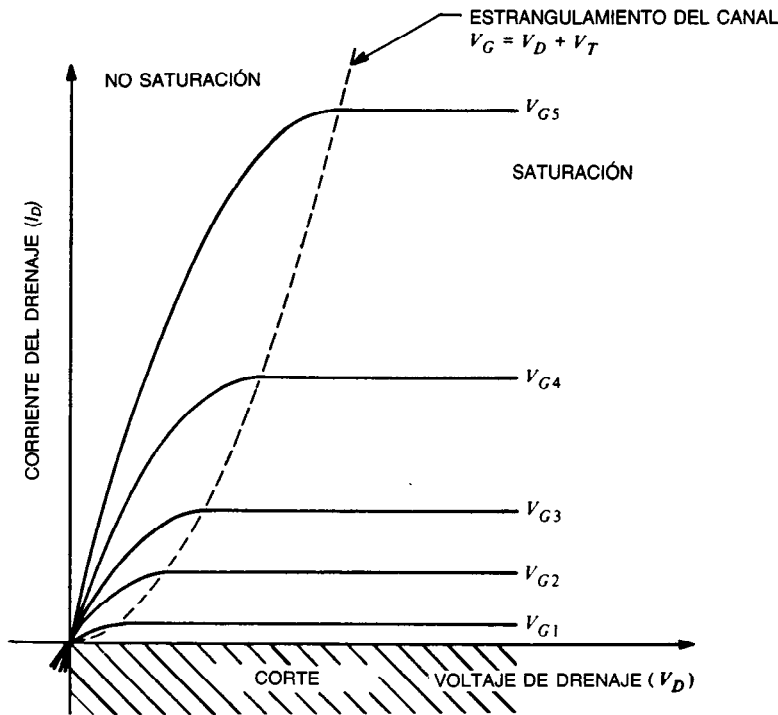


Fig. 13-35. Características corriente-voltaje de un transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET).

con el cuadrado del voltaje de compuerta a altas corrientes, cuando la saturación por movilidad causa un incremento más lento en la corriente con el voltaje. Asimismo, tal como se observa en la figura 13-36b, a voltajes de compuerta cercanos a V_t , la corriente varía con mucha mayor rapidez que V_g^2 . De hecho, es posible demostrar que para bajos valores, la corriente está dada por:

$$I_D = I_o \exp(qV_g/nkT) \cdot [1 - \exp(-qV_D/kT)] \quad (13-48)$$

donde I_o y n son constantes que dependen de las características de la superficie del semiconductor, así como de las propiedades físicas y geométricas del transistor.

A partir de las ecuaciones 13-45 y 13-46 puede observarse que la transconductancia o ganancia del MOSFET:

$$g_m = \left| \frac{\partial I_D}{\partial V_g} \right|_{V_D > V_t - V_i} = \left| \frac{\mu_s \epsilon_{ox}}{x_o L} \right| W (V_g - V_t) \quad (13-49)$$

se incrementa haciendo pequeño el grosor del óxido, x_o , grande el ancho del canal, W , y pequeña la longitud del canal, L . Existen límites prácticos en cuanto al valor máximo de W y el mínimo de x_o y L . Por debajo de 50 Å de grosor, los electrones pueden penetrar por efecto túnel a través del dióxido de silicio, de modo que x_o se limita a valores mayores que 50 Å para el

dióxido de silicio. La longitud del canal, L , es limitada por la penetración de la región de agotamiento del drenaje hacia la región de la fuente. Al primer orden, el voltaje de penetración (o de perforación) está dado por:

$$V_{PT} = \frac{L^2 q N_{ch}}{2 \epsilon_s} \quad (13-50)$$

donde N_{ch} es la concentración de impurificador en la región de la superficie del semiconductor y ϵ_s es la permitividad del material de silicio. Para una impurificación del canal de 10^{16} , el voltaje de penetración debería ser aproximadamente 7 V para un canal con longitud de 1 μm .

El voltaje umbral, V_t , es determinado por:

$$V_t = \phi_{ms} + \frac{Q_{fs}}{C_{ox}} + 2\psi_B + \frac{x_o}{\epsilon_{ox}} \cdot \sqrt{2q\epsilon_s N_{ch}(2\psi_B + V_{ss})} \quad (13-51)$$

donde ϕ_{ms} = diferencia de función trabajo entre el material de la compuerta y el semiconductor

Q_{fs} = carga parásita en la interficie óxido-semiconductor

ψ_B = diferencia de potencial entre el nivel de Fermi asociado con N_{ch} y el nivel intrínseco de Fermi

V_{ss} = polarización de la fuente al sustrato

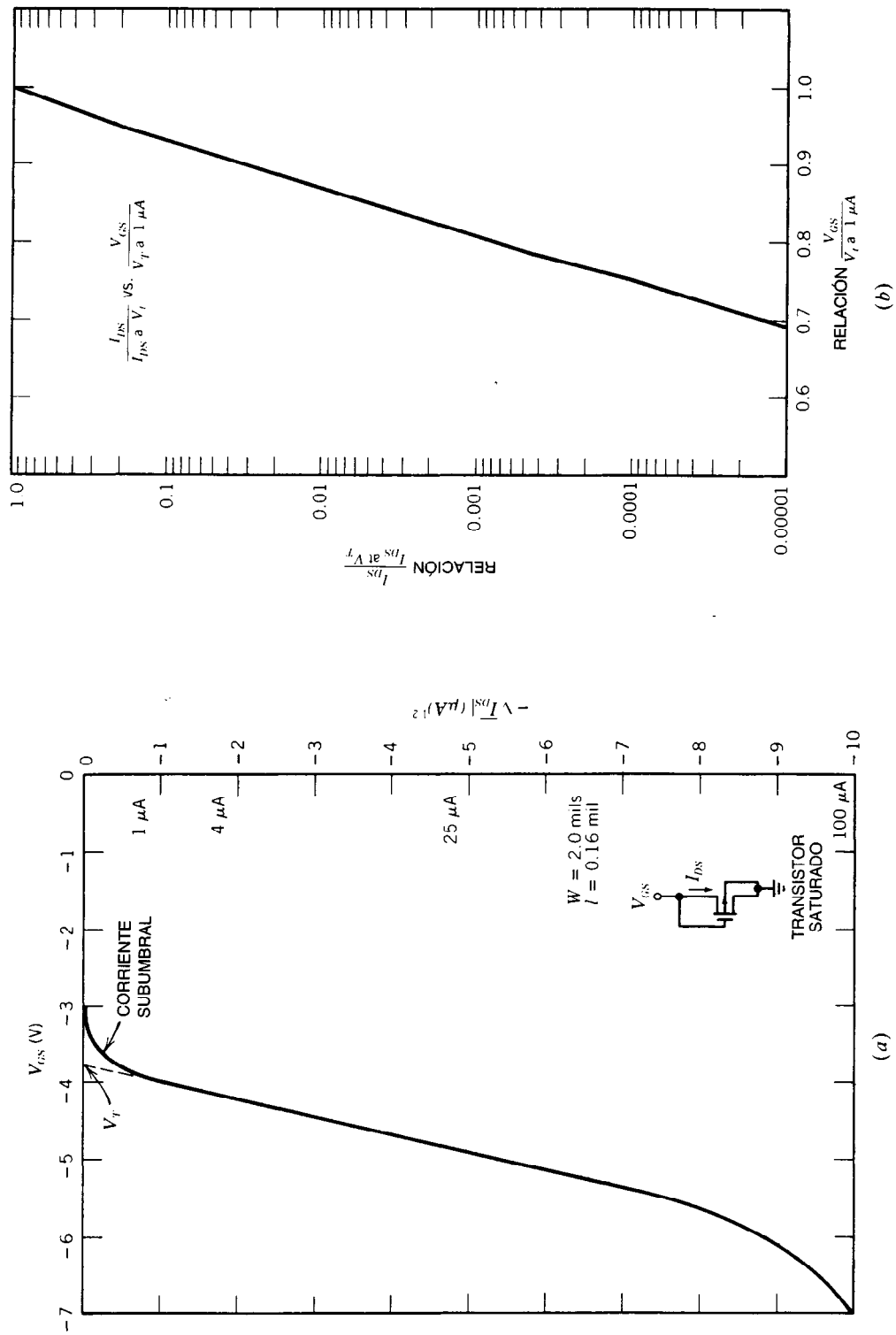


Fig. 13-36. a) $\sqrt{\text{Corriente de drenaje}}$ en función del voltaje de compuerta para un transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET) en saturación. b) Corriente subumbral en función del voltaje de compuerta.

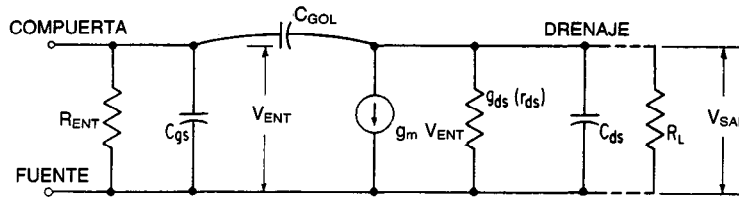


Fig. 13-37. Circuito equivalente de señal pequeña de un transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET).

Para los dispositivos MOSFET de buena calidad, Q_{fs} no es mayor que $q \times 10^{11}/\text{cm}^3$, y ψ_B puede aproximarse con mucho a la separación energética del semiconductor, lo que para el silicio significa $\psi_B \approx 0.5$ V.

En la figura 13-37 se muestra el circuito equivalente del MOSFET. La frecuencia máxima de operación del dispositivo es la frecuencia a la cual la corriente a través de C_{ent} es igual a $g_m(V_G - V_t)$:

$$f_{\max} = \frac{g_m}{2\pi C_{ent}} \approx \frac{\mu_s}{2\pi L^2} V_D, \quad |V_D| < |V_G - V_t| \quad (13-52)$$

$$f_{\max} = \frac{\mu_s(V_G - V_t)}{2\pi L^2}, \quad |V_D| > |V_G - V_t| \quad (13-53)$$

La conductancia de salida del drenaje, g_D , se asocia con la modulación de la longitud efectiva del canal, L , mediante la región de agotamiento del drenaje de la siguiente forma:

$$\frac{1}{r_D} = g_D = \frac{\partial I_D}{\partial V_D} \approx \frac{\mu_s C_{ox}}{\pi L^2} \frac{(V_G - V_t)^2}{(V_D + V_t - V_G)}, \quad |V_D| > 2|V_G - V_t| \quad (13-54)$$

La ganancia de un dispositivo MOSFET, cuya resistencia de realimentación de la fuente es negativa, está dada por:

$$A'v = \frac{-g_m r_D / (G_L + g_D)}{r_D + R_L + (g_m r_D + 1)R_s} \approx \frac{-g_m R_L}{1 + g_m R_s} \quad (13-55)$$

si se supone que r_D y g_m son grandes.

Dado que un MOSFET es un dispositivo monopolar (es decir, sólo un tipo de portador está asociado con el flujo de corriente), la velocidad de conmutación sólo está limitada por la corriente que el dispositivo puede transportar y por la carga capacitiva por accionar.

Para un MOSFET en saturación, el cálculo del tiempo de tránsito de los portadores de la fuente al drenaje está dado por:

$$t_T = \frac{4}{3} \left(\frac{L^2}{\mu_s} \right) \frac{1}{(V_G - V_t)} \quad (13-56)$$

Para un canal con longitud de $3 \mu\text{m}$ y $V_G - V_t = 1$ V, el tiempo de tránsito es aproximadamente 0.2 ns. Siempre que las constantes de tiempo del circuito RC excedan este valor, es posible ignorar el tiempo de

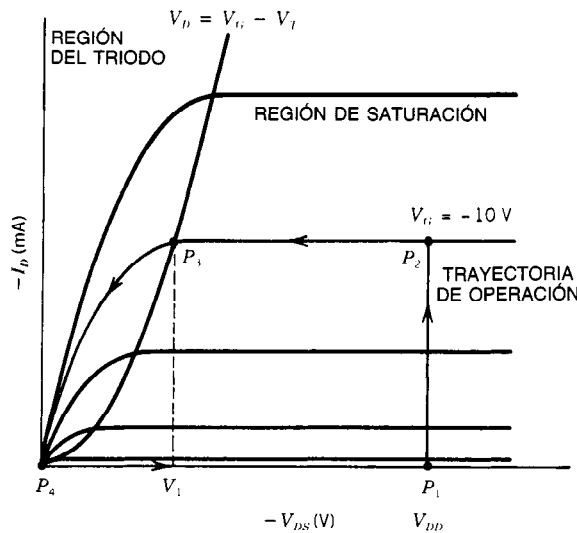


Fig. 13-38. Trayectoria transitoria ($P_1 \rightarrow P_4$) de la operación de un transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET).

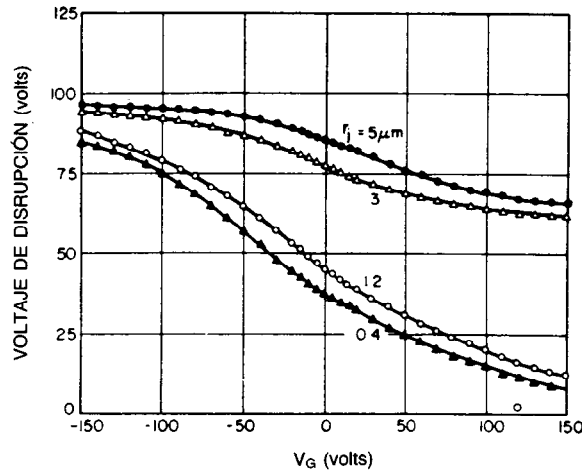


Fig. 13-39. Voltaje de disrupción entre drenaje-sustrato en función del voltaje de compuerta para cuatro diferentes profundidades de difusión del drenaje.

tránsito al evaluar el comportamiento transitorio de un MOSFET en el circuito. Para un MOSFET que acciona una carga capacitiva C_L , el tiempo de conmutación se determina sumando el tiempo requerido para que el voltaje a través de C_L disminuya de V_{DD} a $V_G - V_t$ y el tiempo que le lleve disminuir de $V_G - V_t$ a aproximadamente cero.

En la figura 13-38 el tiempo requerido para que el voltaje de la carga decaiga de V_{DD} a $V_G - V_t$ es:

$$t_1 = \frac{2C_L |V_{DD} - V_g + V_t|}{g_m (V_g - V_t)} \quad (13-57)$$

El tiempo t_2 necesario para que el voltaje de carga decaiga de $(V_G - V_t)$ a cero se encuentra despejando t en la siguiente ecuación:

$$\frac{v(t)}{(V_g - V_t)} = \frac{2e^{-t/\tau}}{1 + e^{-t/\tau}} \quad (13-58)$$

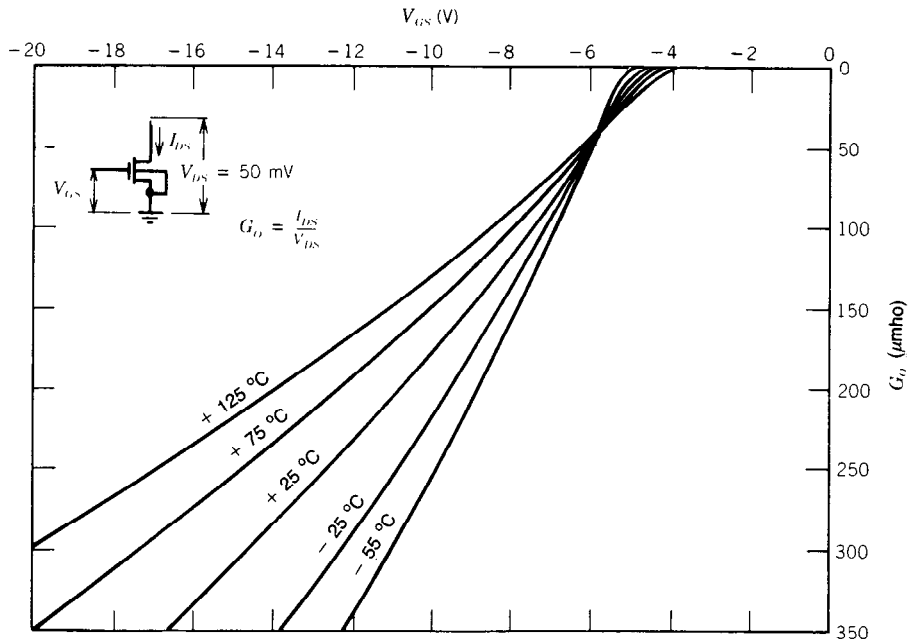


Fig. 13-40. Variación de la conductancia del canal con la temperatura en un transistor de efecto de campo a base de semiconductor de metal y óxido (MOSFET).

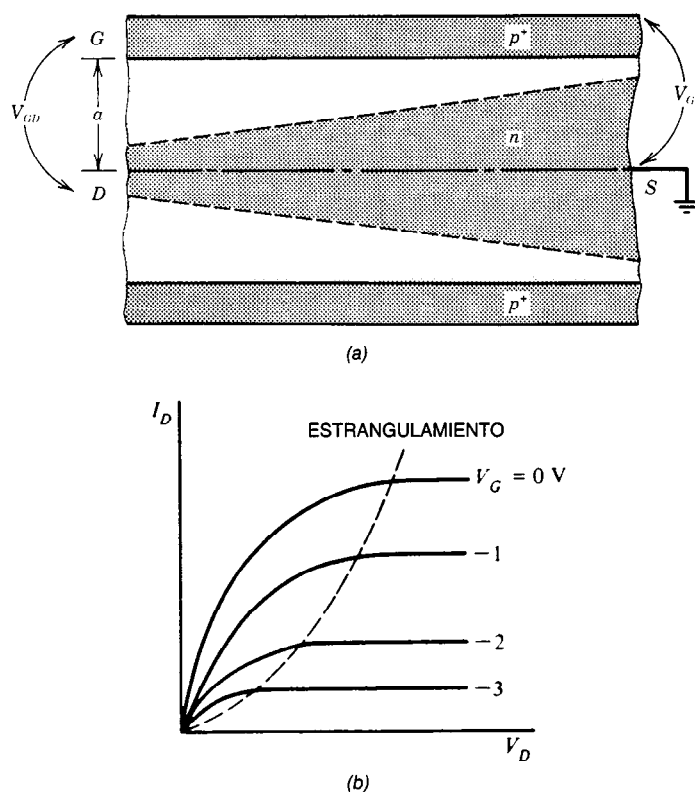


Fig. 13-41. a) Construcción de un transistor de efecto de campo de unión (JFET).
b) Características corriente-voltaje en un JFET de canal n.

donde:

$$\tau = C_L / g_m, \quad g_m = \frac{\mu_s C_{ox} W}{L} (V_G - V_t)$$

Los voltajes disruptivos en las uniones de los dispositivos MOSFET suelen ser menores que en las uniones *pn* no cubiertas con una estructura MOS, ya que el metal de la compuerta provoca un congestionamiento del campo alrededor de la unión de drenaje. En la figura 13-39 se muestra el efecto del voltaje de compuerta en la disrupción de la unión de cuatro profundidades.

Por lo general, los MOSFET se conciben para que presenten dependencia de la corriente respecto a la temperatura, con coeficiente negativo, de modo que la corriente disminuya cuando la temperatura aumente. Esta dependencia con respecto a la temperatura ayuda a asegurar una igual distribución de corriente en los dispositivos MOSFET que se operan en paralelo. Sin embargo, en los MOSFET existen efectos opuestos que pueden producir un coeficiente térmico positivo o negativo, dependiendo de las condiciones de operación. Dado que la movilidad disminuye como $T^{-3/2}$, existe una tendencia a que la corriente disminuya al aumentar la temperatura. Por otra parte, el voltaje umbral V_t varía como una función de $1/T$, de modo que

existe presión para que la corriente aumente, ya que $(V_G - V_t)^2$ aumentará con la temperatura. En la figura 13-40 se muestra que para valores de V_G cercanos a V_t , la variación del voltaje umbral es dominante y la conductancia aumenta con la temperatura. Para grandes valores de V_G , la variación en la movilidad es dominante y la conductancia disminuye con el aumento de la temperatura.

El ruido eléctrico en un MOSFET limita las señales mínimas que es posible amplificar. Tres de los tipos de ruido son el ruido térmico asociado con variaciones en la carga de los portadores en el canal, el ruido por generación-recombinación asociado con la carga en la región de agotamiento debajo del canal, y el ruido $1/f$ que se debe a la ocupación aleatoria de centros de captura en la interficie del óxido y del semiconductor. El ruido térmico es de importancia a altas frecuencias, y el ruido $1/f$ es dominante a bajas frecuencias.

13.3.2 Transistor de efecto de campo de unión (JFET)

Un transistor de efecto de campo de unión (JFET, de *junction field effect transistor*) es semejante a un MOSFET, excepto en que la compuerta es sustituida por una unión *pn* (fig. 13-41a) y el dispositivo opera a través del agotamiento de un canal ya existente.

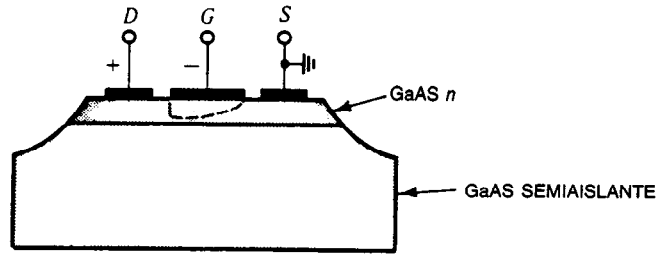


Fig. 13-42. Transistor de efecto de campo a base de semiconductor de metal (MESFET) de arseniuro de galio.

El rendimiento máximo se logra mediante el empleo de una compuerta $p+$ y de un canal tipo n , debido a que los electrones presentan mayor movilidad que los huecos. Con referencia a la figura 13-41, la corriente a través de un JFET está dada por:

$$I_D = \frac{2aWV_p}{\rho_c L} \left[\frac{V_D}{V_p} + \frac{2}{3} \left(-\frac{V_G}{V_p} \right)^{3/2} - \frac{2}{3} \left(\frac{V_D - V_G}{V_p} \right)^{3/2} \right] \quad (13-59)$$

donde ρ_c = resistividad del canal
 a = la mitad de la separación entre las regiones de compuerta
 V_p = voltaje de estrangulamiento

El voltaje de estrangulamiento puede calcularse como el voltaje al que la región de agotamiento de la compuerta es igual a a . Este voltaje es:

$$V_p = \frac{-a^2 q N_D}{2\epsilon_s} \quad (13-60)$$

Cuando el JFET entra en saturación, la corriente está dada por:

$$I_{Dsat} = \frac{2aWV_p}{\rho_c L} \left[\frac{V_G}{V_p} + \frac{2}{3} \left(-\frac{V_G}{V_p} \right)^{3/2} + \frac{1}{3} \right] \quad (13-61)$$

y el voltaje de drenaje en saturación es:

$$V_{Dsat} = V_G + V_p \quad (13-62)$$

La transconductancia g_m de un JFET en la región de saturación es:

$$g_{m sat} = \frac{2aW}{\rho_c L} \left[1 - \left(-\frac{V_G}{V_p} \right)^{1/2} \right] \quad (13-63)$$

Los JFET fueron elegidos alguna vez para aplicaciones en las que se requería una alta impedancia de entrada, como en los amplificadores del extremo delantero para radios y electrómetros. Sin embargo, debido a que son difíciles de fabricar y debido a las características mejoradas de los MOSFET, actualmente los JFET se emplean menos que los MOSFET.

Los MOSFET no son prácticos en algunos materiales semiconductores, como el arseniuro de galio, debido a que la interficie entre el GaAs y un aislante superficial es tal que un gran número de trampas su-

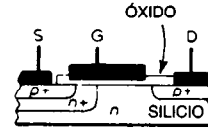


Fig. 13-43. Construcción de un transistor de efecto de campo a base de semiconductor de metal y óxido de doble difusión (DMOSFET).

perficiales inhibe la acción compuerta. Además, las compuertas de unión no se difunden con tanta facilidad dentro del GaAs debido a la inestabilidad del material semiconductor a altas temperaturas. Sin embargo, dado que el GaAs tiene mayor movilidad que el silicio, su empleo es aconsejable en aplicaciones en las que se requieren rápidas velocidades de conmutación. Por consiguiente, se utilizan estructuras del tipo JFET, excepto que un diodo Schottky sustituye a la unión pn como compuerta. En la figura 13-42 se muestra tal transistor de efecto de campo a base de semiconductor de metal (MESFET, de *metal semiconductor field effect transistor*).

Con frecuencia, los MESFET se fabrican sobre una delgada capa epitaxial depositada sobre un sustrato semiaislante. También es posible utilizar la implantación de iones para formar esta delgada región, y algunas veces se emplea además para formar los contactos fuente-drenaje. Ya que sólo se emplea una compuerta unilateral en tales estructuras, la corriente está dada por la siguiente ecuación:

$$I_D = \frac{x_c W V_p}{\rho_c L} \left[\frac{V_G}{V_p} + \frac{2}{3} \left(-\frac{V_G}{V_p} \right)^{3/2} + \frac{1}{3} \right] \quad (13-64)$$

donde x_c es el grosor del canal, ρ_c es la resistividad del canal, y el voltaje de estrangulamiento V_p es:

$$V_p = \frac{q x_c^2 N_D}{2\epsilon_s} \quad (13-65)$$

13.3.3 Variaciones de los MOSFET

MOSFET con compuerta doble

Para aplicaciones de alta frecuencia en las que se requiere baja distorsión, algunas veces se emplean estructuras de transistores de efecto de campo con compuerta doble. Estos MOSFET tetrodos operan de

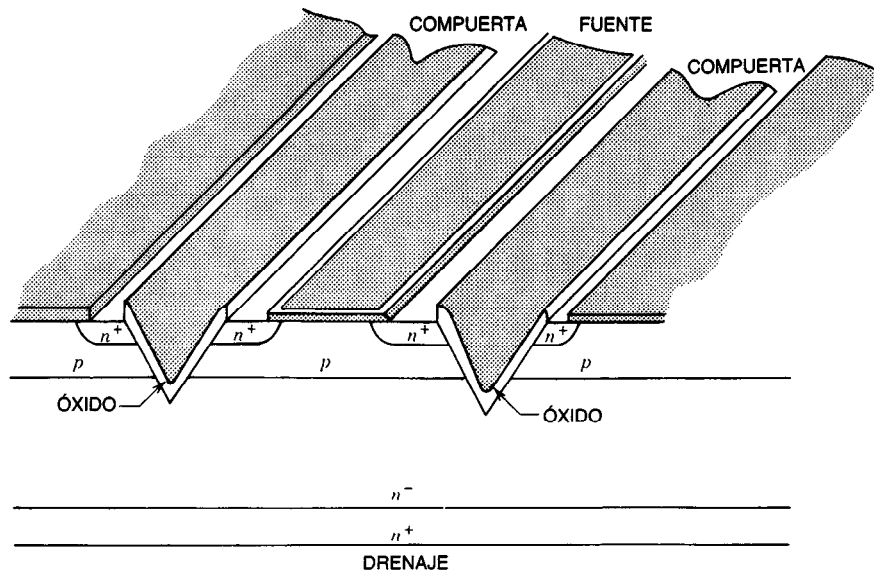


Fig. 13-44. Construcción de un semiconductor de metal y óxido con ranura en V (VMOS).

forma bastante parecida a como lo hacen los MOSFET ordinarios, excepto porque la compuerta primaria está desacoplada de la región de drenaje por la presencia de la compuerta secundaria. El beneficio de la compuerta secundaria es que se reduce sustancialmente la capacitancia del drenaje de compuerta. También es posible emplear tales dispositivos en aplicaciones como mezcladores y de control de ganancias automático.

MOSFET de doble difusión

Otro método para obtener MOSFET de alta velocidad es fabricar el dispositivo en un sustrato de alta resistividad. La región activa del canal se forma por la difusión de una región tipo n seguida de una difusión p^+ , que forma la fuente; la longitud del canal es determinada entonces por la diferencia en las profundidades de la unión, tal como se indica en la figura 13-43. Esta estructura se denomina DMOSFET. La región n de alta resistividad entre el drenaje p^+ y la fuente se agota cuando al drenaje se aplican unos cuantos volts. Debido a que es posible mantener altos valores de transconductancia incluso con distancias apreciablemente grandes entre el drenaje y la región n^+ , se pueden obtener mayores voltajes de ruptura de los que se obtendrían en el caso de que la región p^+ fuese inmediatamente adyacente a la región n^+ del canal.

MOSFET de potencia

La estructura DMOS recién descrita puede aplicarse a la construcción de dispositivos MOSFET de potencia. Sin embargo, para controlar cantidades significativas de potencia con un MOSFET, es necesario contar con grandes anchos del dispositivo (W en la ec. 13-47). Para una máxima utilización del área de silicio, los MOSFET de potencia suelen fabricarse con el contacto del drenaje en la parte inferior del silicio y los contactos de la fuente y la compuerta en la superficie su-

perior. Los primeros dispositivos MOS de potencia que se introdujeron en el mercado se denominaban VMOS debido a la ranura en forma de V que se efectuaba en el silicio (fig. 13-44). Estas estructuras son útiles para aplicaciones de bajo voltaje (100 V), ya que la resistencia del drenaje puede minimizarse empleando la "V" grabada a fin de reducir la distancia que los portadores deben recorrer a través de la región no modulada n de alta resistividad. Sin embargo, la estructura en V provoca acumulación del campo eléctrico en la punta de la región grabada, y esta concentración de campo reduce el voltaje disruptivo drenaje-fuente. Para dispositivos de mayor voltaje, una estructura preferida es la del DMOS plano (fig. 13-45a), o bien la del VMOS truncado (fig. 13-45b). La solución intermedia dominante en los MOSFET de potencia es entre la resistencia de encendido y el voltaje disruptivo. Resulta que la resistencia de encendido ideal varía como el voltaje disruptivo ideal elevado a la potencia 2.5; es decir:

$$R_{\text{ideal}} \propto V_{\text{NB}}^{2.5} \quad (13-66)$$

La resistencia de dispersión desde la región del canal hasta el contacto del drenaje incrementa la resistencia real. En la figura 13-46 se muestran algunas mediciones de dispositivos reales comparadas con la resistencia ideal como función del voltaje disruptivo.

13.3.4 Consideraciones de aplicación

Para una área dada de semiconductor, un MOSFET tiene mayor voltaje de encendido, a un nivel específico de corriente, que un transistor bipolar. En la figura 13-47 se muestra la eficiencia de conmutación de un MOSFET y de un transistor Darlington, ambos con capacidad nominal de 450 V, y cada uno de los cuales

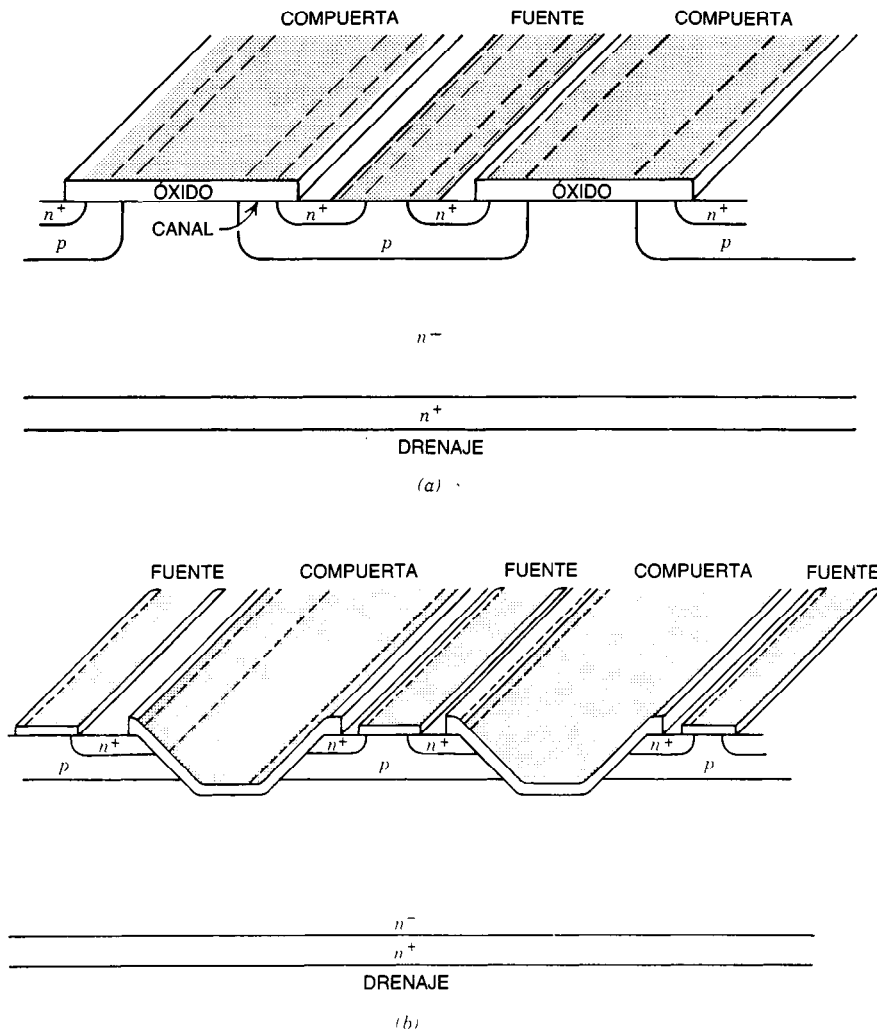


Fig. 13-45. a) Transistor de potencia (DMOSFET) de efecto de campo a base de semiconductor de metal y óxido de doble difusión. b) MOSFET de potencia con ranura en V truncado.

se ha conmutado empleando una corriente pico de conmutación igual a la décima parte de la corriente de salida. A bajas frecuencias, el transistor Darlington presenta un tercio de las pérdidas del transistor MOS de potencia. Sin embargo, a frecuencias mayores de 13 kHz las menores pérdidas de conmutación del MOSFET compensan sobradamente las mayores pérdidas por conducción, y el MOSFET es el dispositivo más eficiente. Para los MOSFET de bajo voltaje (< 100 V), las pérdidas por conducción se aproximan mucho más a las de los dispositivos bipolares.

Es necesario tomar algunas precauciones al aplicar MOSFET en circuitos de conmutación de alto voltaje, debido al transistor bipolar parásito que existe en paralelo con el MOSFET (fig. 13-48). Dado que la tensión de drenaje aumenta con rapidez, la capacitancia de la unión pn puede permitir el flujo de carga en la región tipo p , que es la base del transistor parásito npn .

Si la región es suficientemente ancha, la corriente capacitiva que fluye horizontalmente bajo la región n^+ de la fuente generará suficiente caída de potencial para polarizar en sentido directo la unión $n + p$, y es posible que bastante corriente fluya a través del transistor parásito. Este problema se ha minimizado con los diseños modernos, pero existen limitantes acerca de la máxima razón de cambio de la tensión de drenaje permitida (dv/dt), incluso para el mejor de los diseños. El valor de $(dv/dt)_{\max}$ se presenta en las hojas de aplicación.

Una ventaja de los MOSFET de potencia es que la resistencia de entrada a la compuerta es muy alta, de modo que se requiere muy poca potencia para encender el dispositivo y mantenerlo. Sin embargo, la capacitancia de entrada para un dispositivo de 10 A y 400 V puede ser hasta de 4 000 pF. Si el voltaje de compuerta se conmuta de 0 a 10 V en 10 ns, la corriente pico que fluirá a través del capacitor de entrada es:

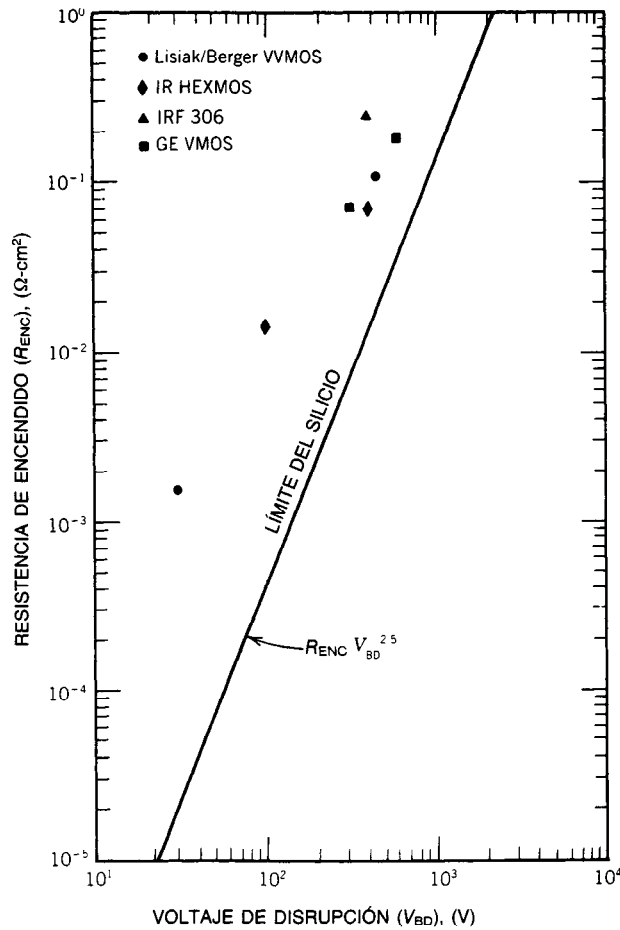


Fig. 13-46. Resistencia de encendido en función del voltaje en algunos transistores de potencia de efecto de campo a base de semiconductor de metal y óxido (MOSFET).

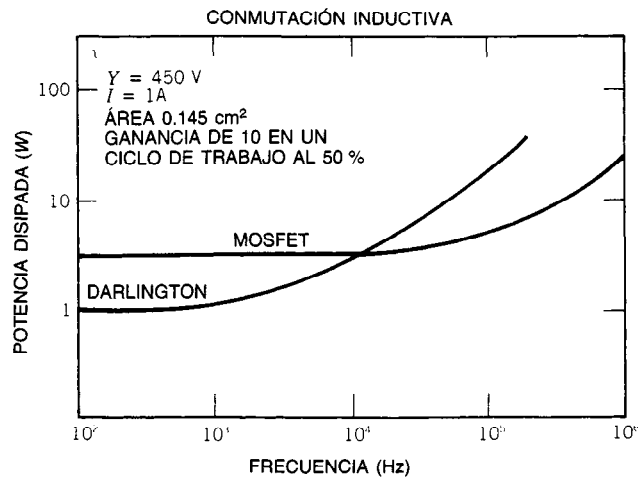
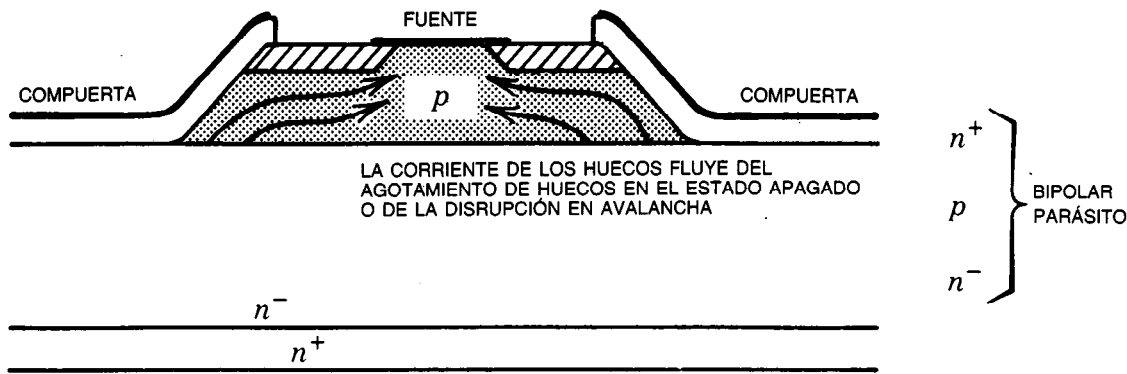


Fig. 13-47. Comparación del transistor de potencia de efecto de campo a base de semiconductor de metal y óxido (MOSFET) (IR330) con un transistor Darlington bipolar de corriente nominal equivalente (ZJ499D).



CIRCUITO EQUIVALENTE INCLUYENDO EL BIPOLAR

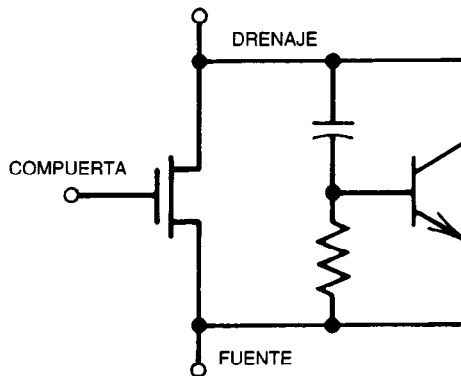


Fig. 13-48. Transistor de potencia de efecto de campo a base de semiconductor de metal y óxido (MOSFET), que presenta un transistor bipolar parásito.

$$I = C \frac{dv}{dt} = \frac{4\,000 \times 10^{-12} \times 10}{10^{-8}} = 4\text{ A}$$

Por consiguiente, se requiere un accionador de compuerta bastante rígido para aprovechar la velocidad potencial de conmutación del MOSFET.

LIBROS Y REVISTAS

- M. S. Adler y S. R. Westbrook, "Power Semiconductor Switching Devices—A Comparison Based on Inductive Switching", *IEEE Trans. Electr. Dev.* **ED-29**(6):947-952.
- B. J. Baliga, "Switching Lots of Watts at High Speeds", *IEEE Spectr.* **18**(12):42-48.
- M. B. Barron, "Low Level Currents in Insulated Gate Field Effect Transistors", *Solid State Electr.* **15**:293-302.
- R. S. C. Cobbold, *Theory and Applications of Field Effect Transistors*, Wiley, Nueva York.
- H. W. Collins y B. Pelly, "HEXFET—A New Power Technology", *Electr. Des.*

- R. H. Crawford, *MOSFET in Circuit Design*, McGraw-Hill, Nueva York.
- A. B. Glaser y G. E. Subak-Sharpe, *Integrated Circuit Engineering—Design, Fabrication, and Applications*, cap. 3, Addison-Wesley, Reading, MA.
- A. B. Grebene, *Analog Integrated Circuit Design*, Van Nostrand Reinhold, Nueva York.
- A. S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, Nueva York.
- W. M. Penney y L. Lau, *MOS Integrated Circuits*, cap. 2, Van Nostrand Reinhold, Nueva York.
- B. G. Streetman, *Solid State Electronic Devices*, cap. 8, Prentice-Hall, Englewood Cliffs, NJ.
- S. M. Sze, *Physics of Semiconductor Devices*, caps. 7-10, Wiley, Nueva York.
- T. D. Towers y S. Libes, *Semiconductor Circuit Elements*, cap. 10, Hayden, Rochelle Park, NJ.
- J. T. Wallmark y H. Johnson, *Field Effect Transistors*, Prentice-Hall, Englewood Cliffs, NJ.
- E. S. Yang, *Fundamentals of Semiconductor Devices*, caps. 7 y 8, McGraw-Hill, Nueva York.