

12
29.



**UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO**

FACULTAD DE CIENCIAS

EL FISICO Y LA COMPUTACION

T E S I S
Que para obtener el Título de:
F I S I C O
p r e s e n t a
SILVIA ESPINOSA WOLF



México, D. F.

1994

**TESIS CON
FALLA DE ORIGEN**



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CIUDAD UNIVERSITARIA



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

FACULTAD DE CIENCIAS
División de Estudios
Profesionales
Exp. Núm. 55

M. EN C. VIRGINIA ABRIN BATULE
Jefe de la División de Estudios Profesionales
Universidad Nacional Autónoma de México.
P r e s e n t e .

Por medio de la presente, nos permitimos informar a Usted, que habiendo
revisado el trabajo de tesis que realizó la pasante Silvia
Espinosa Wolf
con número de cuenta 8153281 8 con el título: El físico y
la computación

Consideramos que reúne los méritos necesarios para que pueda conti-
nuar el trámite de su Examen Profesional para obtener el título de
Física

GRADO NOMBRE Y APELLIDOS COMPLETOS

FIRMA

Dr. HECTOR GERARDO RIVEROS ROTGE

Director de Tesis

Dr. JUAN ANTONIO COGORDAN RAMIREZ

Dr. ERNESTO JOSE MARTA DE LA SALET REIMONT MORDINO

Dra. MARIA DE LOS ANGELES ORTIZ FLORES

Suplente

Fig. JESUS ARMANDO LARA VELAZQUEZ

Suplente

Ciudad Universitaria, D.F., a de

de 199

Para:

Nelly, quien me dio la vida

Enrique, con quien comparto la vida

Aldo, a quien le di la vida

AGRADECIMIENTOS

Para mi director de tesis, el Dr. Héctor Riveros Rotgé:

“... to whom you may impart griefs, joys, fears, hopes, suspicions, counsels, and whatsoever lieth upon your heart to oppress it, in a kind of civil shrift or confession”.

Francis Bacon

Deseo agradecer también los valiosos comentarios y sugerencias de mis sinodales: Dra. Ma. de los Angeles Ortíz Flores, M. en C. Jesús Armando Lara Velázquez (con su inigualable paciencia), Dr. Ernesto José Marfa de la Salet Belmont Moreno y Dr. Juan Antonio Cogordan Ramírez.

Contenido

| | |
|--------------------------|----|
| PRÓLOGO | ix |
|--------------------------|----|

PRIMERA PARTE: HARDWARE

| | |
|----------------------------------------------------------|----|
| INTRODUCCIÓN | 3 |
| Historia de las computadoras | 3 |
| Computadoras PC compatibles | 6 |
| Componentes internas de una computadora | 7 |
| La unidad de control | 7 |
| La unidad lógica aritmética | 8 |
| El reloj del sistema | 8 |
| La memoria principal | 9 |
| Tipos de memoria | 9 |
| Características comerciales de una computadora | 10 |
| Sistema operativo | 11 |

| | |
|--------------------------------------------------|--------|
| TECLADO | 15 |
| Teclado alfanumérico | 15 |
| Teclas de funciones programables | 17 |
| Teclas de control del cursor | 18 |
| Teclado numérico | 18 |
| Funciones especiales | 18 |
| Indicadores | 19 |
| Operación del teclado | 19 |
| Comunicación con el sistema operativo | 20 |
| Traducción de códigos de barrido | 20 |
| Programación de teclas de funciones programables | 21 |
| ANSI.SYS | 21 |
| MEMORIA | 25 |
| Unidades | 26 |
| Direcciones de memoria | 26 |
| Mapas de memoria | 26 |
| Tipos de memoria | 27 |
| Memoria ROM | 27 |
| Memoria RAM | 28 |
| Registros internos | 28 |
| Memoria cache | 28 |
| Memoria principal | 29 |
| Memoria shadow | 32 |
| Memoria virtual | 33 |
| Memoria expandida | 33 |
| Memoria extendida | 34 |

| | |
|-----------------------------------|-----------|
| BUSES | 37 |
| Bus ISA | 41 |
| Bus Micro-canal | 42 |
| Bus EISA | 43 |
| Bus SCSI | 43 |
| Bus PCI | 43 |
| Consideraciones generales | 44 |
| MONITORES | 45 |
| Sistema de video | 46 |
| Modo texto | 46 |
| Modo gráfico | 47 |
| Resolución | 47 |
| Tipos de monitores | 48 |
| Tarjetas aceleradoras de video | 51 |
| Consideraciones generales | 54 |
| UNIDADES DE ALMACENAMIENTO | 55 |
| Unidades de disco flexible | 55 |
| Tipos de disco flexible | 57 |
| Discos duros | 58 |
| Tipos de discos duros | 59 |
| Discos FM y MFM | 61 |
| Discos RLL | 62 |
| Discos SCSI | 62 |
| Discos IDE | 62 |

| | |
|---------------------------------------------|-----------|
| SCSI vs. IDE | 64 |
| Discos ópticos | 65 |
| Discos flópticos | 69 |
| Organización de un disco | 69 |
| Estructura de un disco | 71 |
| Estructura física | 71 |
| Estructura lógica | 71 |
| Organización interna de un disco | 72 |
| Boot | 73 |
| Directorio | 74 |
| FAT | 76 |
| MODELOS DE COMPUTADORAS PC | 79 |
| 8088 | 80 |
| 8086 | 81 |
| 80286 | 81 |
| 80386SX | 82 |
| 80386 | 83 |
| 80486 | 83 |
| Consideraciones generales | 85 |

SEGUNDA PARTE: LAS COMPUTADORAS EN LA FÍSICA DE SUPERFICIES

| | |
|---------------------------------------------------------------------------------|------------|
| INTRODUCCIÓN | 91 |
| Emulación de multicanal | 92 |
| DETECTORES | 95 |
| Detector de barrera superficial | 95 |
| Semiconductores | 95 |
| Modelo de enlaces de valencia | 95 |
| Modelo de bandas de energía | 100 |
| Junturas | 103 |
| Polarización inversa | 105 |
| Generación de pares | 106 |
| Detectores de radiación gama | 107 |
| TEORÍA: RETRODISPERSIÓN DE RUTHERFORD | 109 |
| Técnica de Retrodispersión de Rutherford para análisis de superficies | 110 |
| Factor cinemático | 111 |
| Sección transversal | 114 |
| Geometría | 114 |
| Frenamientos nuclear y electrónico | 115 |
| Generación del espectro | 122 |
| Análisis de espectros | 125 |
| Aproximación superficial | 125 |
| Aproximación para películas gruesas | 126 |

| | |
|-----------------------------------------------------------------------|------------|
| Altura de pico | 127 |
| LAS COMPUTADORAS EN LA RETRODISPERSIÓN DE RUTHERFORD | 129 |
| Sistemas de adquisición de datos por computadora | 129 |
| Programas para análisis de espectros | 132 |
| Un ejemplo | 135 |
| Discusión | 136 |
| TECNOLOGÍA DE FABRICACIÓN DE CIRCUITOS INTEGRADOS | 139 |
| CONCLUSIÓN | 145 |
| REFERENCIAS | 149 |

APÉNDICES

| | |
|--------------------------------------------|------------|
| I. Aritmética hexadecimal | 153 |
| II. Caracteres | 157 |
| III. Boot | 161 |
| BIBLIOGRAFÍA | 165 |

PRÓLOGO

En la época actual el uso de la computadora indudablemente se ha convertido en una herramienta cotidiana e indispensable. La incidencia de éstas se encuentra diseminada en todos los ámbitos de nuestra sociedad. El factor primordial ha sido el desarrollo tan acelerado de las computadoras tipo PC. Este avance tecnológico tan acelerado ha permitido, en pocos años, diseñar, desarrollar, fabricar y volver obsoletas cuatro generaciones de procesadores, desde el primitivo 8088 hasta los muy recientes 486, gracias al modernísimo PENTIUM. Al paralelo que los microprocesadores, la tecnología asociada a las computadoras ha creado dispositivos o periféricos más eficaces y de mejor desempeño. Un ejemplo palpable son las impresoras, que han escalado desde las impresoras de matriz de punto hasta las Laser. La misma celeridad en el avance tecnológico ha causado una disminución en los costos de producción y en consecuencia, un abaratamiento en el precio de las computadoras. Este factor económico es uno de los responsables de la diseminación masiva de las computadoras.

Todo este desarrollo ha sido posible, además, con el soporte dado por los fabricantes de programas, que han trabajado al parejo, para ofrecer paquetes de programas cada vez más sofisticados, y que abarcan todas las necesidades potenciales del usuario medio: hojas de cálculo, procesadores de palabras, programas de diseño gráfico o de maquinaria y planos, juegos, etc.

En particular, en el área científica resulta imprescindible esta herramienta, tanto para la adquisición como para el procesamiento de datos; sobre todo, debido a la filosofía con que fue diseñada la PC, llamada tecnología abierta, es posible conectar a una de estas computadoras gran cantidad de dispositivos que se utilizan en la física experimental, haciendo menos tediosos algunos procesos de control de experimentos.

El propósito, en el caso más general, ha sido crear una obra en la cual esté compilada toda la información que un usuario requiera acerca de qué es, de qué consta, y en algunos casos cómo opera su computadora o las componentes que tenga, temas que resultan imprescindibles en la formación de un físico hoy en día.

En esta tesis se incluyen desde los temas básicos —como es el manejo del teclado— hasta la manera en que interactúan las diferentes partes de la computadora, y la forma en que se comunican, de modo que se pueda tener una comprensión global de cómo trabaja el sistema. Lo anterior es de vital importancia para el científico, ya que el tener estos conocimientos le permite explotar al máximo su computadora.

La tesis está organizada en dos partes: la primera, se refiere a todo aquello que tiene que ver con los dispositivos físicos, también conocidos como *hardware*. En la segunda se presenta un caso particular: la aplicación de las computadoras personales tipo PC para la adquisición y análisis de datos en la física de superficies, en particular al usar la técnica conocida como Retrodispersión de Rutherford.

Espero que este texto cubra con su propósito: dar al físico los conocimientos básicos para que pueda explotar al máximo la que hoy en día constituye una herramienta básica de trabajo: la computadora personal.

Crafty men contemn studies, simple men admire them, and wise men use them; for they teach not their own use; but that is a wisdom without them and above them won by observation. Read not to contradict and confute; nor to believe and take for granted; nor to find talk and discorse; but to weight and consider. Some books are to be tasted, others to be swallowed, and some few to be chewed and digested; that is, some books are to be read only in parts; others to be read, but curiously; and some few to be read wholly, and with diligence and attention.

Francis Bacon

PRIMERA PARTE:

HARDWARE

INTRODUCCIÓN

Una computadora es una máquina que permite hacer cierto tipo de tareas rutinarias en forma eficiente. Sus primeras aplicaciones fueron el ejecutar cálculos numéricos, desde sumas y restas hasta la solución de ecuaciones diferenciales.

Más adelante se desarrollaron programas que permiten a la computadora hacer otro tipo de tareas, como el procesamiento de textos, graficación de datos, estadísticas, dibujo, y muchas otras cosas más.

HISTORIA DE LAS COMPUTADORAS

Desde la prehistoria, el hombre ha tenido necesidad de contar. Inicialmente no tenía el concepto de número, sino que comparaba un conjunto de objetos que deseara contar (como podía ser el número de animales que constituían su ganado) con un conjunto de piedras. Poco a poco, conforme fue avanzando la civilización, se fue introduciendo el concepto de número: primero, ya no era necesario cargar con las piedras para poder darse cuenta de si faltaba o no un animal de su ganado, sino que el hombre asignaba un número diferente a cada parte de su cuerpo, como podían ser los dedos de una mano, luego el antebrazo, brazo, los dedos de la otra mano, etc. No fue sino hasta que se hubo desarrollado la civilización como tal (historia —inicio de la escritura) en que se empieza a "formalizar" la aritmética. Es en este momento cuando surgen los números como una abstracción; ya no es necesario comparar dos conjuntos de igual magnitud. Basta con decir "cuatro" (o el equivalente en el lenguaje que tuviesen) para que se supiera la cantidad a la que se referían. Desde las primeras culturas que aparecieron sobre la tierra ya se encuentran representaciones y

sistemas numéricos. Los caldeos, babilonios y egipcios usaron cotidianamente los números. Los mayas llegaron a tener la noción del cero. Los árabes introdujeron el sistema decimal y fueron los iniciadores del álgebra. Junto con este desarrollo, se fueron introduciendo máquinas que simplificaban los cálculos. La primera máquina que se usó para facilitar los cálculos numéricos, como multiplicaciones, fue la mano. Otra, la más importante de ellas y que sobrevive hasta nuestros días es el ábaco, que viene del griego "abax", que quiere decir tableta.

El ábaco se usa aún en la actualidad, sobre todo en los países orientales. Permite hacer operaciones de suma, resta, multiplicación y división¹.

En el año de 1612, el escocés John Napier crea una tabla de logaritmos, para que se puedan hacer más rápidos algunos cálculos.

En 1632 se inventa la regla de cálculo, usada hasta hace unos años. Sus principios son las propiedades de los logaritmos; de esta manera, una multiplicación se convierte en una suma y la división en resta.

En 1642, Blaise Pascal inventa la primer máquina capaz de llevar a cabo operaciones aritméticas elementales (sumas y restas). El sistema mecánico consistía en un conjunto de engranes (o ruedas dentadas) que representaban las unidades, decenas, centenas, etc. Sobre cada rueda estaban los dígitos del cero al nueve: un dígito para cada posición que podía tomar la rueda. La rotación completa de una rueda producía el avance en una posición de la siguiente a la izquierda. A esto se le denomina acarreo, es decir, cuando queremos sumar dos números (por ejemplo, 45 y 37) y la suma de los dos primeros dígitos es mayor que diez, decimos "cinco más siete igual a dos y 'llevamos uno'". Este uno es el acarreo.

Hasta que Pascal inventó su máquina, nada más el hombre hacía acarreo.

En 1671, Leibnitz proyectó una máquina con la que intentó realizar multiplicaciones y divisiones mediante sumas y restas. Debido a la complejidad de las piezas y la falta de herramientas para construirlas en la época, no se pudo llevar a cabo la idea de Leibnitz a nivel comercial.

1 En 1945 se llevó a cabo una competencia entre un japonés y un americano; se trataba de realizar cinco operaciones aritméticas, el japonés con un ábaco —y siete años de entrenamiento en su uso— y el americano con una calculadora eléctrica —y cuatro años de entrenamiento. La competencia consistía en realizar una suma (números con tres a seis dígitos), resta (seis a ocho dígitos), multiplicación (cinco a 12 dígitos), división (cinco a 12 dígitos) y un evento compuesto por treinta sumas, tres restas, tres multiplicaciones y tres divisiones de números con seis a 12 dígitos. El japonés derrotó al americano en cuatro de los cinco eventos: el americano sólo ganó en la multiplicación.

En 1801, Joseph Marie Jacquard crea la primera máquina de tarjetas perforadas, que le permitían hacer tejidos complicados. Las perforaciones de la tarjeta proporcionaban las instrucciones para la selección de hilo y ejecución de los diseños. Actualmente se sigue usando el mismo principio en las máquinas tejedoras. Estas fueron las primeras máquinas programables que existieron. Es el método de programación más primitivo, pero que se sigue utilizando incluso hoy en día, en el sector textil.

Charles Babbage, en 1812, diseña la primera máquina analítica. Su idea era incluir una unidad de almacenamiento de datos con tarjetas perforadas, una unidad aritmética que llevara a cabo las operaciones aritméticas, y una unidad de control, que dirigiese las operaciones. Resultó de una complejidad mecánica tan grande que no se pudo construir. En 1993 se construyó esta máquina, y se encuentra en el British Museum, en Inglaterra.

Hollerith crea la primera máquina calculadora de tarjetas perforadas, para poder procesar la información del censo de 1880 en Estados Unidos. La información se codificaba como perforaciones: una respuesta afirmativa o negativa, como presencia o ausencia de perforación; una cantidad numérica como perforaciones en ciertas posiciones —la posición de la perforación determina el valor numérico. En 1896, Herman Hollerith fundó la compañía International Business Machines (IBM).

En 1936, Konrad Zuse construye la primera calculadora electromecánica (la Z1) e introduce la representación binaria de los números.

En 1944, Howard y Aiken, en Harvard, diseñan la Mark 1, que usa una cinta perforada que contiene las instrucciones de la máquina. Puede sumar dos números de 23 dígitos en 1/3 de segundo y los multiplica en seis segundos.

Entre los años de 1942 y 1946 se desarrolla la primera computadora electrónica, sin partes mecánicas, para los cálculos. El resultado sale en tarjetas perforadas. Se trata de la ENIAC (Electronic Numerical Integrator and Calculator).

Un poco más adelante, en 1952, John von Neumann construye la EDVAC (Electronic Discrete Variable Automatic Calculator), que ya es capaz de almacenar datos y programa; también puede tomar decisiones lógicas y actuar en consecuencia. Esta máquina constituye la primera generación de computadoras. La electrónica es con bulbos (tubos al vacío).

En 1951 se construye la UNIVAC (Universal Automatic Computer), que es la primera computadora comercial. Sus características son: uso de cinta mag-

nética para suministro de datos y almacenamiento de resultados; usa datos alfanuméricos (letras y números); y usa un programa especial capaz de traducir el lenguaje de programación a un código particular, de modo que pueda ser entendido por la computadora, y así ejecutar las instrucciones.

La segunda generación de computadoras cambia la tecnología en uso: aparece el transistor. Con esta nueva generación se economiza en espacio, pero fundamentalmente las computadoras tienen las mismas características funcionales que las de la primera generación.

En 1964 aparece la tercera generación de computadoras: usa circuitos integrados (chips). También se introducen adelantos como: la computadora se autogobierna y maneja pasando de una tarea a otra; la máquina se puede autodiagnosticar e indica cuál es el problema. También se presenta una evolución en los lenguajes de programación y aumenta la velocidad de los procesos.

La cuarta generación surge a raíz de que se van integrando más circuitos en un chip, por lo que se requiere de menos espacio físico para la computadora. Finalmente, se llega a desarrollar el microprocesador (principios de los 70's).

Con el advenimiento del microprocesador fue posible la construcción de microcomputadoras, siendo una de las primeras la Radio Shack, que contaba con 64 kb de memoria y el almacenamiento de información se hacía en cintas magnéticas (cassettes).

Otro uso reciente de los microprocesadores ha sido a nivel industrial, controlando procesos. A estos dispositivos se les conoce como "controladores", aunque en realidad son computadoras con un programa pregrabado y que, además, pueden sensar variables físicas —como temperatura, presión, etc.— y tienen control sobre actuadores, como válvulas, relevadores, motores, etc., de modo que los datos de salida sean los establecidos en el programa.

Computadoras PC compatibles. En 1981, IBM lanza al mercado una computadora personal, conocida como IBM PC (*Personal Computer*), que es muy rudimentaria, ya que tiene 128 kb de memoria RAM² y una sola unidad de diskette de baja densidad (su capacidad de almacenamiento es de 180 kb). Poco después aparece una computadora con disco duro integrado, aunque de pequeña capacidad, denominada IBM PC XT. A raíz del éxito que tienen estas máquinas, un gran número de fabricantes diseñan sus propias computadoras siguiendo los lineamientos de IBM, y cuidando la compatibilidad, es decir, que un

2 Es la memoria disponible para la ejecución de programas y almacenamiento temporal de datos.

programa elaborado para las IBM también se puede ejecutar en éstas. A partir de ese momento, IBM marca el camino a seguir, en tanto que los demás fabricantes se limitan a reproducir o copiar los modelos desarrollados por esta firma. En 1984 nace la IBM PC AT, que es una computadora más poderosa y rápida que la XT. En la primera mitad de 1987 surge la serie IBM PS/2 (*Personal System*); inicialmente aparecen los modelos 25 y 30, que son computadoras tipo AT, pero con una nueva tecnología para el manejo de los datos y señales de control; continúa con los sistemas 386, que ya incluyen una unidad de control más avanzada y veloz. En 1993 se da a conocer la primera computadora personal con un procesador que incorpora parte de la filosofía de los sistemas grandes (*mainframes*), y que involucra el "multi-tasking", que consiste en tener dos o más tareas ejecutándose simultáneamente.

COMPONENTES INTERNAS DE UNA COMPUTADORA

Para poder hacer uso de todas sus capacidades, una computadora requiere de: dispositivos o elementos físicos que conforman una computadora, como son: el teclado, pantalla y CPU (Unidad Central de Procesos); a este conjunto de dispositivos físicos que conforman una computadora se le denomina *hardware*; y de programas para poder operar dichos dispositivos³. Generalmente, los programas comerciales reciben el nombre de *software*. Dentro de los programas se incluyen dos grupos fundamentales, que son el Sistema Operativo, que es el conjunto de programas básicos para interactuar con la computadora; y los programas de aplicación, como pueden ser: procesadores de texto, hojas electrónicas, programas de graficación y otros.

En la figura 1 se muestra un diagrama esquemático de una computadora típica y sus componentes, que son el procesador, la memoria y dispositivos de entrada/salida.

El CPU consta de cuatro elementos fundamentalmente: la unidad de control, la unidad lógica aritmética, el reloj del sistema y la memoria.

La unidad de control Es una de las partes más importantes de una computadora. Ésta activa a las demás unidades en el orden correcto para llevar a cabo cada instrucción. Para esto, tiene que codificar las instrucciones a una secuencia de comandos de transferencia y/o manipulación. Traduce al lenguaje de máqui-

³ Es necesario contar con los programas básicos que sean capaces de interpretar el conjunto de instrucciones que se le dan a la máquina, ya que de otra forma no es posible ni siquiera introducir comandos o instrucciones desde el teclado.

na (instrucciones definidas en términos de unos y ceros) las operaciones que se quieren llevar a cabo.

El operar en lenguaje de máquina da más rapidez, pero se pierde la claridad en su uso, por lo complicado que resulta el trabajar con secuencias de unos y ceros; hay más riesgos de cometer errores.

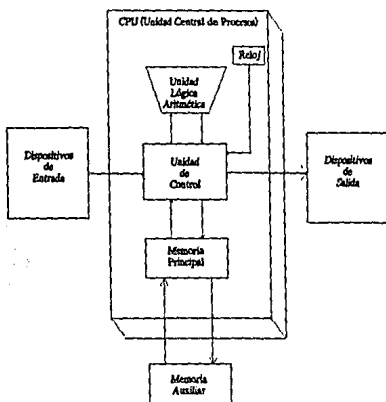


Fig. 1. Estructura básica de una computadora

La unidad lógica aritmética (ALU) Combina dos cantidades binarias (es decir, números en base dos, o sea, que consisten en unos y ceros únicamente) para producir un resultado binario. Los tipos de operaciones que puede realizar son aritméticas (suma, resta, multiplicación y división) y lógicas (AND, OR, XOR, NOR).

El reloj del sistema Es un elemento pocas veces mencionado, pero fundamental para el funcionamiento de cualquier computadora, ya que es el que sincroniza todas las operaciones. El reloj se encarga de emitir pulsos de la siguiente manera: durante un cierto lapso de tiempo, produce un voltaje positivo y en otro lapso un voltaje nulo; la electrónica del sistema lleva a cabo una función cualquiera durante las transiciones entre voltaje positivo y nulo.

En cada pulso del reloj se realiza una de las funciones que se soliciten al microprocesador. La velocidad con que trabaje el reloj repercute en la velocidad de ejecución de los cálculos o programas.

La memoria principal Es simplemente un arreglo de grupos de ocho dígitos binarios⁴ ordenados secuencialmente y que son accesables aleatoriamente, cada uno de los cuales se identifica por su dirección. Cada dirección está dada por un valor numérico entre cero y 65,535. Cabe hacer notar que esta definición de memoria principal es aplicable cuando se trabaja con el sistema operativo DOS, y que, al estar usando otros tipos de sistemas operativos —y usando procesadores 80386SX o superiores— no existe este límite.

En cada dirección de memoria se guarda un dato. Los datos pueden ser de varios tipos, entre los que se encuentran porciones de instrucciones o valores numéricos que se requieran.

Tipos de memoria La memoria principal de una computadora puede ser de dos tipos: memoria de lectura/escritura, en la cual se pueden almacenar datos y direcciones para ser leídas en cualquier momento; y memoria de sólo lectura, con la que se puede leer el contenido de cualquier dirección, pero los datos se almacenan una sola vez, con un proceso llamado "quemado".

Memoria RAM. Así se denomina generalmente a la memoria de lectura/escritura. Las siglas significan Random Access Memory (memoria de acceso aleatorio). En esta porción de la memoria principal se "cargan" los programas que se van a ejecutar, se almacenan los datos que requiera el programa y también, en el caso de estar trabajando con el sistema operativo DOS, se guarda ahí una parte del archivo COMMAND.COM, para permitir el control de los dispositivos⁵ de entrada/salida que se requieran. Es una memoria dinámica que siempre está siendo usada con diferentes programas y datos. Por ejemplo, todo lo que se introduce desde el teclado se almacena en una porción de esta memoria y de ahí la lee el procesador para poder ejecutar lo que se requiera.

Memoria ROM. Las siglas significan Read Only Memory, es decir, memoria de sólo lectura. Es una memoria permanente, es decir, no se puede alterar su

4 A un conjunto de ocho dígitos binarios —es decir, que pueden tomar los valores 1 y 0— se les denomina byte.

5 Un dispositivo es cualquier tipo de equipo adicional conectado a la computadora, como puede ser una impresora, el teclado, monitor, etc.

contenido. En ésta se almacena el programa BIOS, que es el primer programa que lee la computadora al encenderla.

Memoria externa. Permite almacenar y recuperar la información mediante unidades de memoria auxiliar, como discos duros, diskettes o unidades de cinta magnética. Estas últimas se usan más como unidades de respaldo de información.

CARACTERÍSTICAS COMERCIALES DE UNA COMPUTADORA

Cuando uno adquiere una computadora, el vendedor da una serie de especificaciones, es decir, qué contiene la computadora. En éstas se incluye:

Microprocesador (o CPU), que es la parte fundamental de una computadora y que es el circuito capaz de ejecutar todos los procesos que se requieran, además de controlar a todos los periféricos y puertos.

Memoria RAM, que es el área total disponible para almacenar la información que se esté usando en un momento dado (es la serie de circuitos donde se "cargan" o guardan los programas para poder ser ejecutados), así como la capacidad máxima de memoria que se puede llegar a instalar.

Slots de expansión, que es la capacidad que se tiene para aumentar tarjetas a la computadora. Casi todos los dispositivos nuevos que se agreguen a la computadora requieren de una tarjeta, la cual se conecta a los *slots* o ranuras de la parte posterior de la computadora, para poder ser reconocida por la máquina.

Fuente de alimentación, que limita la cantidad de dispositivos que se pueden conectar ya que, si la demanda de corriente es mayor a la que puede dar la fuente, la computadora no funciona.

Tipo de teclado, que depende del modelo de computadora que se adquiera, aunque actualmente se maneja el teclado de 101 teclas; puede estar en inglés o español (para los sistemas que se comercializan en México). Los teclados de diferentes modelos de computadoras no son intercambiables.

Puertos serie y paralelo. Usualmente las computadoras tienen dos puertos "serie", para conectar el ratón en uno de ellos y el modem para comunicación en red o por vía telefónica en el otro, y uno "paralelo", para conectar la impresora.

Reloj con batería, para que se tenga un registro de la fecha y hora en que fueron creados los archivos de datos del usuario (es una manera de reconocer cuál es la versión más reciente del archivo). Si no se cuenta con reloj, hay que darle la fecha y hora a la computadora en el momento de encenderla.

Unidades de disco flexible (*floppy disk drive*), que es imprescindible en cualquier computadora, ya que es el dispositivo que nos permite copiar programas al disco duro, cargar el sistema operativo, almacenar información y en general, para poder usar la computadora. Las presentaciones que se manejan actualmente son:

- » Unidades de disco de 5 1/4" de doble lado-doble densidad
- » Unidades de disco de 5 1/4" de alta densidad
- » Unidades de disco de 3 1/2" de doble lado-doble densidad
- » Unidades de disco de 3 1/2" de alta densidad

Unidades de disco duro, cuya forma de operación es muy similar a las unidades de disco flexible, pero que permiten almacenar mayor cantidad de información. Usualmente los discos duros tienen una capacidad de 30 millones de bytes (abreviado Mb) en adelante (hasta cientos de millones de bytes). En la actualidad existen dos tipos fundamentales de discos duros, que son los IDE y los SCSI, cada uno de ellos con diferentes características en cuanto a su funcionamiento. Los primeros de ellos son los más comunes y menos costosos.

Debido a que la mayor parte de los programas comerciales que existen actualmente son muy extensos, para poderlos usar, requieren de un disco duro, donde se "instalan" antes de poder ser ejecutados.

Monitor, que puede ser en "blanco y negro" (usualmente en verde o ámbar), o a color. Existe una gran variedad de monitores; los más populares en la actualidad son los tipo TTL, que son monocromáticos; y los VGA, que pueden ser monocromáticos o a color.

SISTEMA OPERATIVO

Veamos ahora con más detalle el sistema operativo.

Un sistema operativo es un conjunto de programas diseñado para permitir y facilitar la comunicación con la computadora.

La comunicación entre la computadora y el usuario se lleva a cabo mediante los periféricos: como el teclado, que nos permite darle órdenes; la pantalla, que nos despliega la información; los controladores de disco, que nos ayudan a almacenar la información de manera permanente; y los puertos, que nos permiten imprimir, usar el ratón o modems.

El sistema operativo no es independiente del tipo de computadora que se use, ya que, dependiendo del tipo de máquina y sus periféricos, se necesitan diferentes programas para su control y adecuado manejo; también varía el CPU, por lo que, tanto el conjunto de instrucciones que cada uno reconoce, así como su codificación, son diferentes para cada tipo de computadora. Por ejemplo, una computadora VAX 11/780, que admite una serie de terminales, unidades de cinta magnética, etc., requiere que se asignen prioridades a los diferentes procesos; claves y memoria disponible para cada usuario, y de la protección del sistema operativo, de modo que sólo un *system manager* pueda tener acceso a él para modificarlo y darle mantenimiento, entre otras cosas.

Las computadoras personales PC (compatibles con el modelo XT de IBM) pueden funcionar con tres sistemas operativos diferentes, que son DOS, OS2 y UNIX⁶. El más difundido, tanto por antigüedad como por su facilidad de manejo, es el DOS (Disk Operating System). Este sistema operativo está orientado fundamentalmente al manejo de archivos. Desde esta perspectiva, DOS no es tanto un sistema operativo como un sistema de programas y funciones diseñados para ser usados por los programas de aplicación y los programas desarrollados por el usuario.

En las PC donde, en principio, existe un solo usuario, éste debe responsabilizarse de todos los programas y archivos de datos que se encuentren instalados; de dar mantenimiento, borrando los archivos que ya no necesite (o los guardará como respaldos en diskettes) o de agregar los programas que requiera. Por tanto, el sistema operativo no es tan sofisticado como en los equipos grandes (con varios usuarios y terminales). Existe, claro, la posibilidad de conectar las

6 Proximamente se liberará un nuevo sistema operativo, llamado Windows NT, desarrollado por MicroSoft, y con el cual se prevé que haya compatibilidad entre las computadoras tipo PC y otros tipos de computadora.

computadoras personales en red, de modo que una de ellas actúe como *server* o servidor y varias estaciones de trabajo⁷, que puedan ejecutar parte de los procesos que, si se vuelven complejos, los manden al *server*. En este caso el sistema operativo MS-DOS ya no es suficiente para dar servicio y controlar a todos los periféricos, por lo que se requiere de *software* especializado para este fin, que supla las carencias del sistema operativo (a partir de la versión 3.0 de MS-DOS se agregan algunos comandos para poder manejar archivos compartidos, dentro de una red, aunque sigue requiriéndose de programas especiales —y el *software* correspondiente— para poder entrar a la red).

Existen otros sistemas operativos de aparición reciente para las computadoras PC, como OS/2 o UNIX, que permiten, entre otras cosas, el control de redes y el *multitasking*, es decir, procesar simultáneamente varios programas. El *software* desarrollado para que funcione con esos sistemas operativos es poco y no para el público en general, por lo que ha tenido poca difusión entre los usuarios particulares de PC.

En los próximos capítulos veremos las componentes principales de la computadora y la manera en que funcionan e interactúan con el resto de la máquina.

7 En las estaciones de trabajo se llevan a cabo la mayor parte de los procesos que requiere el usuario. El servidor se usa para almacenar programas y bases de datos compartidas.

TECLADO

El teclado es el dispositivo de entrada que se usa fundamentalmente para comunicarnos con la computadora. Desde él podemos dar las instrucciones y comandos, editar textos en procesadores de texto, o crear bases de datos.


En términos generales todos los teclados son similares, y se ajustan al estándar de IBM. Algunas variantes consisten en añadir teclas (como las teclas de función programable F11 y F12 para los teclados tipo AT) o son versiones reducidas del teclado estándar (como la omisión de las teclas de control del cursor; en este caso se usa el teclado numérico).

Siempre que se escribe alguna línea de comandos en DOS y se presiona la tecla Enter, la computadora almacena la secuencia de caracteres en un *buffer*, además de ejecutar los comandos indicados.

El teclado está dividido en seis partes, según se muestra en la figura 2, que se presenta en la siguiente página:

TECLADO ALFANUMÉRICO

Esta parte es muy similar a una máquina de escribir. Se usa para teclear la mayor parte de los datos. Los caracteres especiales y signos de puntuación se despliegan en pantalla cuando se oprime la tecla Shift junto con la tecla del signo correspondiente. También cuenta con una serie de teclas especiales que se describen a continuación.

Enter (o Intro)  Equivale a la tecla de retorno de carro de una máquina de escribir. Cuando se oprime, el cursor se desplaza al inicio de la siguiente

línea. Dentro de DOS se usa para indicar a la computadora que ya está la instrucción o comando completo y que se inicie su ejecución.

Backspace (o retroceso) ← Esta tecla desplaza un lugar la posición del cursor hacia la izquierda y borra el caracter inmediatamente anterior.

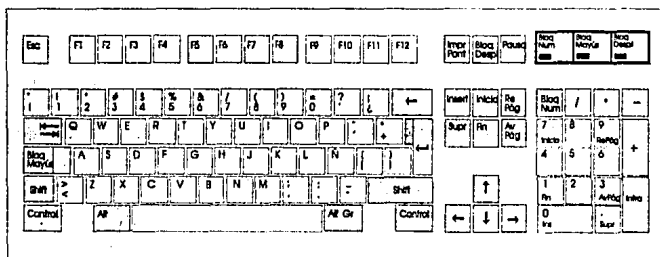


Fig. 2. Teclado de una computadora PC AT

Tab (tabulador) ⇔ Esta tecla desplaza el cursor un número dado de espacios a la derecha, donde la cantidad exacta de espacios depende del procesador de textos o programa que se esté usando. Algunos procesadores de texto toman el tabulador como una serie de espacios y otros asignan un código especial (en éstos no es equivalente un tabulador al número correspondiente de espacios).

Shift Las dos teclas de **Shift** del teclado tienen la misma función. Se usan para escribir minúsculas y mayúsculas y desplegar los símbolos superiores de las teclas que tienen dos caracteres diferentes.

Caps Lock (o Bloq Mayús) Se usa para escribir todo un conjunto de caracteres en mayúsculas. Al oprimirlo, se enciende el indicador correspondiente en la región de indicadores. Al contrario de las máquinas de escribir, los

caracteres especiales y algunos de los signos de puntuación no se afectan al estar habilitado el modo **Caps Lock**. Para desactivarlo, basta volver a oprimir la tecla **Caps Lock**, ya que actúa como un interruptor de encendido/apagado.

Ctrl o Control Esta tecla se usa siempre en conjunto con otra(s) para llevar a cabo un comando o función específica. Para sus usos dentro de cada programa hay que referirse al manual correspondiente, ya que no hay un estándar para estos comandos de control. En sistema operativo, al oprimir simultáneamente la tecla **Ctrl** y la tecla **C**, se puede suspender la ejecución de un programa y con las teclas **Ctrl** y **S** se detiene temporalmente el proceso que se esté ejecutando.

Alt Esta tecla es la "tecla alterna". Es decir, si se conoce el código ASCII¹ de un carácter, se puede teclear desde el teclado alfanumérico o bien oprimir la tecla **Alt** y, sin soltarla, teclear en el *pad* numérico el código ASCII decimal correspondiente. También se usa cuando no hay un carácter disponible desde el teclado alfanumérico. Sin embargo, hay que consultar el manual del procesador de texto que se esté usando para ver las equivalencias del conjunto de caracteres ASCII extendido, ya que cada procesador asigna a un mismo número de código un carácter ASCII extendido (superior a 127) de manera diferente. Este método permite meter cualquier carácter ASCII entre 1 y 255. El único carácter ASCII que no se puede teclear directamente es el 0, ya que está reservado para indicar caracteres no ASCII, como las teclas de control del cursor y las teclas de funciones programables. También se usa, en algunos programas de aplicación, en conjunto con las teclas alfabéticas, para dar comandos o instrucciones especiales.

TECLAS DE FUNCIONES PROGRAMABLES

Están marcadas con las etiquetas F1 a F10 en las computadoras tipo XT y F1 a F12 en las tipo AT, y tienen distintas funciones dependiendo del programa que se esté usando e incluso, en los lenguajes de programación, puede uno mismo definir su función. Esencialmente se usan en sustitución de una serie de instrucciones, para ahorrar teclazos y tiempo. Dentro de DOS, las funciones asignadas son las siguientes:

1 Ver apéndice II.

| <i>Tecla</i> | <i>Función</i> |
|--------------|----------------------------------------------------------------------------------------------|
| F1 | Copia un caracter del <i>buffer</i> a la línea de comandos |
| F2 | Copia los caracteres del <i>buffer</i> a la línea de comandos hasta el caracter especificado |
| F3 | Copia todos los caracteres del <i>buffer</i> a la línea de comandos |
| F4 | Copia los caracteres restantes en el <i>buffer</i> a la línea de comandos |

TECLAS DE CONTROL DEL CURSOR

Estas teclas, también llamadas "teclas de navegación" están etiquetadas con flechas hacia arriba, abajo, derecha e izquierda. Se usan para desplazar el cursor por la pantalla, siempre que el programa que estamos usando lo permita.

TECLADO NUMÉRICO (o *pad* numérico)

Está ubicado en la parte inferior derecha del teclado y tiene dos funciones, dependiendo de si está o no habilitado el interruptor **Num Lock** (o **Bloq Num**). Si el indicador de **Num Lock** está apagado, se puede usar este teclado para desplazarse, un espacio a la vez, en cualquiera de las cuatro direcciones indicados en la parte inferior de las teclas 2, 4, 6 y 8. Si está encendido, se usa en forma similar a las teclas numéricas del teclado alfanumérico. En este teclado también hay una tecla **Enter**, con las mismas funciones que en el teclado alfanumérico. Otras teclas con funciones especiales son el **O (Ins)**, que se usa para insertar caracteres a partir de la posición del cursor; en algunos procesadores de texto esta tecla sirve para cambiar entre modo "inserción" y modo "sobrescritura". La tecla **Del** borra el carácter sobre el cual se encuentra colocado el cursor. También existen un conjunto de teclas grises con los signos aritméticos de suma (+), resta (-), multiplicación (*) y división (/). Además, las teclas 1, 3, 7 y 9 tienen las leyendas **End** (o **Fin**), que sirven para colocar el cursor en el extremo derecho de la línea en la que se encuentra el cursor; **Pg Dn** (o **AvPag**), para ir una página hacia el fin del documento; **Home** (o **Inicio**), para ir hacia el extremo izquierdo de la línea donde está el cursor y **Pg Up** (o **RePag**) para ir una página hacia el inicio del documento. La definición de "una página" varía dependiendo del programa que se use.

FUNCIONES ESPECIALES

Las funciones especiales son las siguientes:

Print Screen Si se usa junto con la tecla Shift, todo lo que está desplegado en pantalla se manda a la impresora. En caso de estar desplegando gráficas en pantalla, se requiere de ejecutar previamente el comando **graphics de** sistema operativo para que se puedan imprimir las gráficas; si no se da este comando, se imprimirá únicamente el texto que se encuentre en pantalla. Si se usa junto con la tecla **Ctrl**, cada línea de datos que se teclee se manda a la impresora.

Esc Su función depende del programa que se use. Por ejemplo, en algunos procesadores de texto sirve para saltar de la zona de texto al menú de opciones y viceversa.

Scroll Lock En algunos programas de aplicación, esta tecla se usa para determinar los movimientos del texto en la pantalla una vez que el cursor ha alcanzado el extremo superior o inferior de la pantalla. En otros casos no tiene ninguna aplicación.

Pause Esta tecla sirve para detener temporalmente un proceso. Para continuar, basta con oprimir cualquier tecla. Usada junto con la tecla **Ctrl**, interrumpe definitivamente el proceso que se esté ejecutando

INDICADORES

El teclado tiene tres o cuatro luces indicadoras en la parte superior derecha, dependiendo del tipo de teclado que se tenga. Estas son el indicador de encendido, que se omite en algunas marcas de computadora (como la Acer); **Caps Lock**, para indicar cuando el teclado está en mayúsculas (encendido) o minúsculas (apagado); **Num Lock**, que indica si el *pad* numérico está en modo numérico (si está apagado, se puede usar como teclas de control del cursor); y **Scroll Lock** (o **Bloq Despl**) para indicar que está activado dicho modo (en la mayoría de los programas de aplicación no importa si se activa o no).

OPERACIÓN DEL TECLADO

El teclado contiene en su interior un controlador, cuya principal función es el supervisar las teclas y reportar al sistema cuando se ha oprimido o soltado una tecla. Si una tecla permanece oprimida por más de medio segundo, el controlador manda una acción de repetición a intervalos fijos (a esto se conoce como *Auto-repeat*). Este controlador contiene también un *buffer* que puede almacenar hasta 20 "teclazos". Si la computadora está ocupada llevando a cabo algún proceso, esta serie de teclas oprimidas se registran en el *buffer* para que, cuando otra vez esté libre la computadora, las despliegue en pantalla y, si alguna de las teclas oprimidas fue un *Enter*, ejecuta los comandos indicados.

Cada que se oprime o libera una tecla, se genera un número de 1 byte, llamado código de barrido (*scan code*) que identifica unívocamente a cada tecla. Cuando se oprime una tecla, se genera un byte con código de barrido entre 1 y 83, y cuando ésta se libera, se genera un byte con código 128 mayor que el código que se generó al oprimirla. Por ejemplo, al oprimir Z se genera un código de 44; al soltarla, se genera un código de 172 (128 + 44). En algunos manuales de DOS se presenta un diagrama del teclado y los códigos de barrido asignados a cada tecla.

Cuando escribimos algo, el teclado no decodifica el significado de los teclazos; simplemente los reporta a la computadora.

COMUNICACIÓN CON EL SISTEMA OPERATIVO (BIOS²)

Cada vez que se oprime o libera una tecla, la acción se reporta al BIOS, donde se llama a una subrutina para averiguar qué acción se llevó a cabo. El código de barrido se regresa al BIOS y se traduce a un código de 2 bytes. El primer byte (menos significativo) contiene el valor ASCII de la tecla y el segundo el código de barrido. Las teclas especiales tienen un cero como byte menos significativo (es decir, no tienen un valor ASCII asignado). Luego BIOS manda los códigos traducidos a un *buffer* ubicado a partir de la localidad de memoria RAM 0000:041E. Ahí se almacenan los códigos hasta que son solicitados por algún programa, como DOS, BASIC o los procesadores de texto, que están en espera de recibir una entrada por teclado.

2 Basic Input Output System. Para su descripción, consultar el apéndice III.

TRADUCCIÓN DE LOS CÓDIGOS DE BARRIDO

La tarea de traducir los códigos de barrido resulta complicada debido a que existen varias opciones que cambian el significado de una tecla, como las teclas **Shift**, **Alt** y **Ctrl** (se denominan genéricamente *shift state*). Cuando se oprime una de estas tres teclas y no se libera, BIOS reconoce que todas las acciones subsecuentes estarán influenciadas por ese *shift state*.

Otras teclas que afectan el teclado son **Caps Lock** y **Num Lock**. Cuando se activa **Caps Lock**, se invierte el significado de la tecla **Shift** para las teclas de las letras del alfabeto y únicamente para éstas.

Cuando se recibe el código de barrido de una tecla cualquiera, BIOS chequea primero si existe un *shift state* y luego procede a traducirlo al código de 2 bytes adecuado.

Conforme se están traduciendo los códigos de barrido, BIOS chequea continuamente si se está oprimiendo alguna de ciertas combinaciones de teclas que demandan acción inmediata (como las teclas **Alt-Ctrl-Del**, que reinician el sistema o **Ctrl-Pause**, que suspende la ejecución de un proceso). Estas teclas hacen que BIOS tome acción inmediatamente, en lugar de mandar la secuencia al *buffer*.

PROGRAMACIÓN DE TECLAS DE FUNCIONES PROGRAMABLES

El sistema operativo nos permite programar las teclas numeradas de F1 a F10 (en teclados tipo XT) y hasta F12 (en teclados tipo AT). Para poder hacerlo, es necesario primero instalar un dispositivo (*driver*) llamado ANSI.SYS en el archivo CONFIG.SYS. Para ésto, basta con editar el archivo antes mencionado y agregarle la línea:

```
DEVICE=C:\DOS\ANSI.SYS
```

Si se cuenta con una línea similar, pero que instala el dispositivo HIMEM.SYS, al inicio del archivo, se debe respetar el orden, es decir, la instalación de ANSI.SYS no puede ser la primera línea de CONFIG.SYS.

Con esta modificación es necesario apagar y volver a encender la computadora o presionar simultáneamente las teclas **Alt**, **Ctrl** y **Del**, para que surta efecto.

Hablemos un poco sobre este dispositivo.

ANSI.SYS

El *driver* ANSI.SYS —American National Standards Institute— modifica la operación estándar de la computadora (teclado y pantalla). Monitorea tanto la salida a pantalla como la entrada desde teclado que pasan a través de los servicios estándar de DOS: busca códigos especiales que identifican los comandos del dispositivo. Toma nota de cualquier carácter o conjunto de caracteres que reconozca como un comando y lo procesa, para luego quitarlo, de modo que no aparezca en pantalla. Por ejemplo, en cuanto al manejo de la pantalla, DOS no nos permite desplazarnos por toda la pantalla (por ejemplo, si teleamos una instrucción y oprimimos < ENTER >, no podemos regresar a esa línea con las teclas de navegación para modificarla y que se vuelva a ejecutar la instrucción. El *driver* ANSI nos permite desplazar el cursor, limpiar la pantalla, cambiar los atributos del cursor, cambiar de modo texto a gráfico y viceversa e incluso guardar la posición actual del cursor de manera que éste se pueda desplazar para desplegar información y posteriormente se regresa a su posición original. En cuanto al teclado, tiene dos funciones primordiales: cuando se usa uno de sus comandos —siguiendo la sintaxis adecuada de las llamadas "secuencias de escape"—, se puede monitorear las entradas por teclado y reemplazar un carácter con otro o con una cadena de caracteres; también permite la redefinición de las teclas de funciones programables. La desventaja de este *driver* es su lentitud para desplegar información en pantalla y el espacio que ocupa en memoria RAM. Si se llega a presentar un problema de espacio en RAM, puede ser necesario evaluar qué tanto se está usando este dispositivo, o si vale o no la pena el removerlo.

En resumen, tiene cinco tipos de funciones:

- » Control del cursor
- » Borrado de pantalla
- » Despliegue de video
- » Atributos de video
- » Redefinición de teclas de funciones programables

Todos los comandos de ANSI tienen el mismo formato: primero, un carácter de "escape", que se abrevia ESC —corresponde al carácter ASCII 27 decimal o 1B hexadecimal— y cuya única función es indicar a la computadora que los caracteres que se mandan a continuación tienen un significado especial, es decir, implican que se va a llevar a cabo una acción que requiere de la interpretación por parte de ANSI, y no se deben desplegar como cualquier comando de DOS. Cabe notar que este carácter de escape no equivale a pulsar la tecla **Esc**.

Después del carácter de escape, va un carácter "[" (abrir corchetes) y, a continuación, los parámetros. Si se incluye más de un parámetro, éstos se deben separar por el símbolo "punto y coma" (;). Finalmente, viene un código de función, que debe estar en mayúscula o minúscula, según se muestra en la siguiente tabla. No son intercambiables las mayúsculas con las minúsculas.

Para poder usar los comandos de ANSI, es necesario antes encontrar una manera de enviar el comando a la pantalla. No se puede teclear el carácter de escape desde el prompt de DOS; si presiona la tecla **Esc**, DOS lo interpreta como un mensaje para borrar la línea del prompt y dejar la máquina lista para recibir un comando. Como alternativa, se puede crear un archivo texto con un procesador de texto ASCII que admita códigos de control embebidos. Algunos procesadores permiten introducir el código de control pulsando simultáneamente las teclas **Ctrl** y [, que puede aparecer como una "flecha hacia la izquierda" o un acento circunflejo con un corchete que abre (^ [). Después, es necesario usar la función **type** para desplegar el archivo y activar las funciones ANSI.

Otra manera es a través de archivos por lotes, usando la función **echo** para activar las funciones ANSI.

También se puede hacer uso del comando **prompt**, que interpreta los caracteres \$E como código de escape y lo envía a ANSI.

A continuación se presenta una tabla donde se muestran las funciones ANSI para la redefinición de teclas programables.

| Redefinición de teclas programables | | | | |
|--------------------------------------------|-------------|------------------|-----------------|----------------|
| <i>Tecla</i> | <i>Sola</i> | <i>Con Shift</i> | <i>Con Ctrl</i> | <i>Con Alt</i> |
| F1 | 0;59 | 0;84 | 0;94 | 0;104 |
| F2 | 0;60 | 0;85 | 0;95 | 0;105 |
| F3 | 0;61 | 0;86 | 0;96 | 0;106 |
| F4 | 0;62 | 0;87 | 0;97 | 0;107 |
| F5 | 0;63 | 0;88 | 0;98 | 0;108 |
| F6 | 0;64 | 0;89 | 0;99 | 0;109 |
| F7 | 0;65 | 0;90 | 0;100 | 0;110 |
| F8 | 0;66 | 0;91 | 0;101 | 0;111 |
| F9 | 0;67 | 0;92 | 0;102 | 0;112 |
| F10 | 0;68 | 0;93 | 0;103 | 0;113 |

Por ejemplo,

```
ESC[0;63;"cd \trabajo";13p
```

asigna a la tecla F5 el cambio del directorio activo al llamado TRABAJO. Por tanto, cada que se pulse la tecla F5, automáticamente la computadora cambiará al subdirectorio TRABAJO, en la unidad de disco activa. El caracter 13 es equivalente a oprimir la tecla < ENTER> —el valor ASCII de < ENTER> es 13. La letra p se debe incluir siempre al final de un comando de reasignación de teclas.

En el manual de DOS se presentan los códigos necesarios para poder reasignar casi todo el teclado.

MEMORIA

La memoria de la computadora es el área de trabajo a la cual se copian los programas antes de ejecutarse. Tomando una analogía, equivale a la mesa de trabajo de un carpintero. Esta analogía ayuda a entender de qué manera depende la cantidad de memoria (o el tamaño de la mesa) del tipo de tareas que se pueden llevar a cabo. La cantidad de memoria fija un límite práctico a los tipos de trabajos que puede realizar la máquina.

La memoria es el área donde el procesador encuentra los programas y datos para poder llevar a cabo una instrucción dada; es solamente un espacio temporal, donde la computadora almacena datos temporales conforme lleva a cabo su tarea. En este sentido, no equivale a las unidades de disco, ya que éstas son de carácter más permanente. Por el contrario, la memoria de una computadora provee sólo de un lugar para poder llevar a cabo los cálculos que se solicitan.

La memoria está formada por pequeñas celdas, cada una de las cuales puede guardar un uno o un cero (que corresponde, en cuanto a señales eléctricas, a que hay o no voltaje). Estas celdas reciben el nombre de **bits**, que es un acrónimo de **binary digit** o dígito binario. A un conjunto de ocho bits se le denomina **byte**, que es la unidad básica de medida para la memoria o las unidades de almacenamiento, como son los diskettes, discos duros o unidades de respaldo.

Sin importar qué tipo de información estamos almacenando en la memoria, toda ésta se codifica como una secuencia (o patrón) de bits; éstos son interpretados de diferentes maneras dependiendo del contexto —pueden corresponder a instrucciones, datos, letras, etc.

UNIDADES

Igual que para las unidades de distancia, donde la unidad básica es el metro, existen múltiplos. Por ejemplo, el kilómetro equivale a mil metros; en el caso de los gramos, un kilogramo son mil gramos. De manera análoga, aunque no directamente equivalente, un kilobyte —abreviado kb— correspondería a mil bytes. Sin embargo, debido a que no se está trabajando directamente en base diez (sistema decimal) sino en binario (base dos), un kilobyte tiene un valor de 2^{10} bytes ($2^{10} = 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 1,024$). Un megabyte —abreviado Mb— corresponde aproximadamente a un millón de bytes ($2^{20} = 1,048,576$); un gigabyte (Gb) corresponde a 2^{30} bytes ($=1,073,741,800$). En la actualidad, los gigabytes son la máxima medida para la capacidad de memoria en las computadoras tipo PC.

DIRECCIONES DE MEMORIA

Cada byte ocupa un lugar especial, conocido como localidad de memoria. A este lugar se le denomina dirección. Cada dirección localiza unívocamente a una localidad o byte específico.

De esta manera se puede guardar un programa a partir de una localidad dada.

Mapas de memoria. Para poder organizar la memoria y poder presentar la documentación de una computadora, incluyendo su sistema operativo, de una manera comprensible, se establecieron ciertas convenciones. Primero, para describir qué áreas de memoria está ocupando una aplicación dada (por ejemplo, el área destinada a almacenar la información que se teclee, el área donde se va a almacenar un programa o los datos que éste requiere, como puede ser la rutina del sistema operativo para decodificar o traducir las instrucciones que le damos a la computadora) se usa un diagrama, en el cual la localidad más baja se encuentra en el extremo superior. De ahí, conforme aumenta el número de la dirección de memoria, baja la posición en el diagrama. A continuación se presenta un diagrama típico de un mapa de memoria.

| | |
|---------------|--------------|
| 0-9FFFF | Convencional |
| A0000-FFFF | Superior |
| 100000-10FFEE | HMA |
| 10FFEF - | Extendida |

Este mapa puede tener el detalle que se desee. Por ejemplo, si deseamos hacer énfasis en el área ocupada para el despliegue en pantalla (video), se puede

bosquejar el resto y detallar únicamente esta zona. Por ejemplo, tomemos el mapa anterior y, en la región de la memoria superior, podemos ampliar la información, para tener el mapa que se muestra a continuación:

| | |
|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 0-9FFFF | Convencional |
| A0000-AFFFF | Para VGA: Buffer para el adaptador de video. Región usada en todos los modos gráficos. Zona libre en los modos texto. Para CGA y monocromtáico: Vacante. |
| B0000-B0FFF | Fin del buffer para el adaptador monocromtáico. |
| B8000- B8FFF | Buffer de video CGA |
| BFFFF | Buffer de video VGA |
| C0000-DFFFF | Región para diversos adaptadores RAM y ROM (como tarjetas de fax, redes, etc.) |
| E0000-FFFFF | Memoria shadow |
| 100000-10FFEE | HMA (memoria superior) |
| 10FFEF- | Extendida o expandida |

TIPOS DE MEMORIA

En cuanto a la memoria, existen dos tipos fundamentales, que son:

Memoria ROM. La memoria ROM es una memoria de sólo lectura (**Read Only Memory**). Este tipo de memoria se encuentra físicamente en un circuito integrado o chip dentro de la computadora. Viene programado de fábrica e incluye el programa BIOS, que sirve para arrancar la computadora. Su tamaño es de 64 kb. Esta memoria no se puede modificar desde sistema operativo o ningún otro programa.

Esta memoria se usa para almacenar programas que no van a ser modificados posteriormente, como los cartuchos de juegos (tipo **Atari**, **Nintendo** u otros videojuegos). Existen diferencias entre los ROM de las computadoras IBM y las otras marcas, ya que una computadora IBM, al momento de arrancar, verifica que exista un disco desde el cual cargar el sistema y, de no encontrarlo, automáticamente carga ROM BASIC, que es el lenguaje de programación BASIC incluido de fábrica en el ROM de la máquina. En cambio las otras marcas, de no encontrar el disco con sistema operativo, despliegan un mensaje de error, solicitando al usuario que introduzca el diskette con sistema operativo y pulse cualquier tecla.

Memoria RAM. Estas siglas significan random acces memory, es decir, memoria de acceso aleatorio. En realidad, tanto la memoria ROM como la RAM se pueden acceder de manera aleatoria.

En una computadora tipo PC existe una subclasificación para la memoria RAM: registros internos o archivos de registros, memoria cache, memoria principal, memoria expandida y memoria extendida. A continuación describiremos brevemente cada una de ellas.

Registros internos. Antes de que una instrucción pueda invocar a alguna de las unidades que conforman el CPU y antes de poder manejar los datos, el procesador deberá obtener tanto la instrucción como los datos de la memoria, y almacenarlos en los registros internos del procesador. La función de los registros es el permitir al procesador tener "a la mano" la información que va a requerir en un futuro inmediato. En estos registros se almacenan también los datos de salida —es decir, los resultados de las operaciones que se hayan llevado a cabo— antes de ser transferidos a la memoria RAM o a alguna unidad de almacenamiento permanente.

Dependiendo del tipo de procesador que tenga la computadora varía la capacidad de almacenamiento interno.

Memoria cache. Usualmente, aunque un procesador trabaje internamente a muy altas velocidades (actualmente hasta 66 MHz), los canales de comunicación con el resto del sistema no es tan eficiente: el tiempo para que se obtenga un dato de la memoria y se transmita al procesador, para poder ser procesado, es mucho mayor debido a que los avances en la velocidad de la memoria principal no han mantenido el paso con los avances de la velocidad de los procesadores.

Para resolver este problema, los diseñadores de sistemas se han orientado a la técnica de memoria cache, para mantener al procesador funcionando lo más congruentemente posible.

La idea fundamental en la que se basa la concepción de la memoria cache es el llamado principio de localidad, que establece que un programa tiende a acceder a las localidades de memoria a las cuales ha hecho referencia más recientemente (llamada dimensión temporal); además, tiende a acceder a las localidades de memoria cercanas a las más recientemente accedidas (dimensión espacial). Por tanto, un sistema cache mueve los datos recientemente accedidos y los cercanos a éstos a un medio de almacenamiento más rápido que la memoria principal, para que el procesador los pueda acceder más versátilmente.

El principio de localidad se basa en las técnicas actuales de programación (programación estructurada), ya que, tradicionalmente, un programa sigue una secuencia lineal, a menos que se encuentre con un bucle (*loop*) —conjunto de instrucciones que se ejecutan repetidamente, en tanto se cumpla con una condición dada. Entonces, la información que se requerirá en un momento dado será ésta misma (en el caso de un bucle) o bien la contigua a la que se está usando (en caso de seguirse una secuencia lineal). La computadora no puede saber *a priori* qué datos usará a futuro, ya que, para ésto, tendría que revisar antes el programa. Es por lo anterior que su única manera de anticiparse a las demandas del procesador sea usando los lineamientos generales que se siguen para programar.

Los sistemas cache se clasifican dependiendo de cómo actualizan la memoria principal. Un cache de escritura (*write-throug cache*) escribe a la memoria principal cada vez que se escribe una localidad de la memoria cache. Por el contrario, un cache de retro-escritura (*write-back cache*) escribe a la memoria principal sólo cuando se elimina una localidad en particular de la memoria cache.

Generalmente la capacidad de una memoria cache es de hasta 256 kb.

Memoria principal. También se le conoce como memoria “real”. Es la memoria RAM “convencional”, donde se almacenan los programas y datos y es la parte de la memoria a la cual el sistema operativo DOS tiene acceso.

Esta memoria recibe su nombre debido al diseño estructural del microprocesador 8088, que solamente permite direccionar hasta 1 Mb; originalmente se suponía que ésta era el área total que requeriría una computadora para los programas y datos.

La memoria principal se divide en dos partes: la memoria baja, que tiene una capacidad máxima de 640 kb, y la memoria alta, entre los 640 kb y 1 Mb. No toda la memoria alta es utilizable por las aplicaciones comerciales; dentro de los 640 kb más bajos se almacenan los programas a ejecutar y los datos que éstos requieren. En el resto se puede almacenar parte del sistema operativo y

1 Cabe destacar la diferencia entre la memoria cache del procesador y la memoria cache del disco duro, ya que la segunda no puede ser usada directamente por el procesador: simplemente da al sistema un área de memoria RAM en la cual se almacena parte del contenido de un archivo, de modo que se haga más rápido el acceso al disco duro. En cambio, la memoria cache del procesador es una memoria RAM rápida donde se guardan datos o instrucciones que probablemente sean los siguientes que solicite el procesador.

algunas de las variables declaradas en el archivo de configuración de sistema (CONFIG.SYS). También pueden quedar ahí programas residentes².

Aunque el 8088 puede direccionar todo el megabyte de memoria, la parte alta estaba destinada a la memoria ROM que contienen parte del sistema operativo.

El uso más frecuente para la memoria superior es el adaptador de video. También hay espacios reservados para otros adaptadores del BIOS. Si alguno de ellos no es instalado, el espacio permanece sin usarse.

Cualquier programa que corra en la memoria baja se dice que está en "modo real".

Veamos con más detalle la manera en que la computadora puede acceder 1 Mb de memoria en modo real. El microprocesador 8088 —y el 8086, que es equivalente— tiene 16 líneas para direccionar una localidad de memoria cualquiera, lo que equivale a 2 bytes o, en decimal, $2^{16} = 65,536 = 64$ kb. Con 16 líneas se podría llamar únicamente a 65,536 localidades diferentes; el límite de memoria RAM sería entonces de tan solo 64 kb. Para no tener que incluir más líneas, Intel³ diseñó la técnica de "direcciones segmentadas". Estas se construyen combinando dos palabras de 16 bits de manera que les permite direccionar 1,048,578 (= 1 Mb) localidades diferentes. Para entender el cómo se hace esto, hay que ver dos cosas: la aritmética involucrada para combinar las dos palabras de una dirección segmentada y la manera en la cual trata el procesador estas direcciones.

La aritmética involucra lo que podemos llamar una "suma recorrida", que permite crear un número binario de 20 dígitos a partir de dos números de 16 bits. Supóngase que tenemos los valores hexadecimales ABCD y 1234. Cada uno de ellos representa un valor binario de 16 bits⁴. Si tomamos una de estas palabras, por ejemplo ABCD, y le agregamos un cero a la derecha, tenemos el número ABCD0. En efecto, esto recorre el número un lugar hexadecimal (o 4 lugares binarios), lo que equivale a haber multiplicado el número por 16^1 . Ahora

-
- 2 A partir de la versión 5 del DOS se declara directamente en el CONFIG.SYS qué programas residentes (incluyendo parte de DOS) se desea cargar en la memoria superior. En el archivo por lotes AUTOEXEC.BAT se puede elegir si se desean cargar en la memoria superior parte de las variables de DOS.
 - 3 Intel es el diseñador y fabricante principal de todos los microprocesadores que usan las computadoras tipo PC, y es quien marca el ritmo de evolución de los procesadores. Actualmente existen más fabricantes de microprocesadores, pero sus nombres no se corresponden uno a uno con las características de los de Intel.
 - 4 Consúltase el apéndice "Aritmética hexadecimal", que trata sobre números binarios y hexadecimales, así como su codificación.

el número tiene 5 dígitos hexadecimales (o 20 dígitos binarios), lo que permite ya tener acceso a 1 Mb de memoria. Sin embargo, no puede servir como una dirección de memoria completa, ya que tiene un cero como dígito menos significativo, por lo que no podría acceder localidades cuyo último dígito fuese diferente de cero; en realidad, sigue pudiendo acceder hasta 64 kb de memoria.

Para completar el esquema de las direcciones segmentadas, tomamos el otro número de 16 bits —en este ejemplo 1234— y lo sumamos al número recorrido de la siguiente manera:

$$\begin{array}{r} \text{ABCD0} \\ + 1234 \\ \hline \text{ACF04} \end{array}$$

Al combinar estos dos números de 16 bits, terminamos con un número de 20 bits que ya puede direccionar 1 Mb de memoria, ya que puede tomar valores entre 0 y 1,048,577. Esta es la manera en la cual, con 16 líneas de direccionamiento, se puede tener acceso a 1 Mb de memoria en modo real.

Los dos valores involucrados reciben el nombre de “segmento” y “offset” (o desplazamiento relativo). En nuestro ejemplo el número ABCD es el segmento y 1234 el offset. El segmento especifica una localidad de memoria que es múltiplo de 16. A estas localidades de memoria se les denomina “fronteras de párrafos”.

El offset de una dirección segmentada especifica la localización exacta de un byte dado dentro de cada párrafo. Debido a que el offset puede tomar 65,356 valores diferentes, un párrafo tiene un tamaño de 64 kb.

La manera en que la computadora nos presenta el valor de una dirección segmentada es el segmento, luego dos puntos (:) y a continuación el offset. En nuestro ejemplo, desplegará el valor ABCD:1234.

Si dividimos la memoria en bloques de 64 kb de manera que cada uno de ellos tenga un dígito hexadecimal más significativo diferente, la memoria queda dividida en bloques: el primero tiene como dígito más significativo el 0 y cualquier dirección dentro de este bloque se puede escribir en forma segmentada como 0xxx:xxxx; el segundo bloque tendrá como dígito más significativo el 1 y las direcciones que incluye son de la forma 1xxx:xxxx, etc. El uso que se da a cada bloque es el siguiente:

5 Análogamente, en base decimal, si agregamos un cero a la derecha de un número cualquiera, equivale a multiplicarlo por 10.

| Bloque | Uso | Bloque | Uso |
|--------|----------------------|--------|----------------------|
| 0 | Aplicaciones y datos | 8 | Aplicaciones y datos |
| 1 | Aplicaciones y datos | 9 | Aplicaciones y datos |
| 2 | Aplicaciones y datos | A | Video |
| 3 | Aplicaciones y datos | B | Video |
| 4 | Aplicaciones y datos | C | Extensión de ROM |
| 5 | Aplicaciones y datos | D | Extensión de ROM |
| 6 | Aplicaciones y datos | E | ROM-BIOS |
| 7 | Aplicaciones y datos | F | ROM-BIOS y BASIC |

De acuerdo con la tabla anterior, se puede ver que únicamente se pueden usar los primeros 10 bloques para aplicaciones y datos; es por ésto que, aunque se cuente con 1 Mb de memoria RAM, la computadora reconoce hasta 640 kb para el usuario⁶. El resto se usa para algunas partes del sistema operativo y el despliegue en pantalla de la información. En modo real existe un límite de 640 kb como espacio total disponible para el usuario, lo que incluye el programa que se use y los datos que dicho programa requiera. Existen programas administradores de memoria que permiten, hasta cierto punto y dependiendo del tipo de tarjeta de video (y monitor) con que se cuente, usar las porciones libres de memoria, siempre y cuando éstas sean múltiplos de 16 kb. Por ejemplo, de acuerdo con el mapa de memoria de video que se presenta a principios de este capítulo, se puede ver que existen varias zonas reservadas para video tipo EGA o VGA, que no se usan si se está usando un monitor CGA o monocromático. Estos programas administradores de memoria pueden hacer uso de las áreas libres en la parte alta de la memoria.

Memoria shadow. A partir del procesador 286, al diseño de los microprocesadores se añaden circuitos que pueden administrar la memoria superior —es decir, la memoria que está entre los 640 kb y 1 Mb— y pueden usar alternadamente memoria RAM y el BIOS en ROM conforme se requiera, de modo que se tiene acceso a los 384 kb adicionales de la memoria principal⁷.

La técnica que se usa para ésto se llama "shadowing". Por ésto, la memoria RAM que se usa, se conoce como memoria "shadow". Para poder ser implementada, el contenido del BIOS y/o los adaptadores de video se tiene que copiar a la memoria RAM que reside en las mismas direcciones que en el ROM (por compatibilidad interna del sistema). La copia se conoce como la sombra

6 De hecho, el área para usuarios es menor, ya que una parte de ésta (aunque muy pequeña) la requiere el sistema operativo. Ahí se almacenan algunos parámetros de configuración de sistema, existe un área para guardar lo que se tecléa, etc.

7 Esto se debe a que, según se muestra en el segundo mapa de memoria de este capítulo, la memoria alta se usa para el despliegue en pantalla (video) y para el ROM-BIOS, y no para datos y aplicaciones.

(*shadow*). Cuando el sistema hace referencia a alguna de estas localidades no usa ya el ROM, sino su sombra en RAM. La memoria que queda disponible puede ser usada para almacenar información que no pertenece directamente al sistema operativo, sino a programas de aplicación.

Memoria virtual. Esta no es realmente una memoria RAM, sino que se usa una parte del disco duro como si fuese memoria principal. La manera en que se usa es la siguiente: cuando se va a correr el programa, el sistema operativo genera un espacio de memoria virtual, que es un modelo de la cantidad de memoria y de direcciones de memoria que se tienen a disposición. Después, una porción de la memoria física se asigna para la manipulación de la memoria virtual. Usando un elemento que es parte integral del procesador, el programa de sistema operativo que permite el uso de memoria virtual le indica al procesador que suponga que la memoria real que se asignó a éste programa está en otra dirección —la dirección virtual que usará el programa. De esta manera, se está usando un área de memoria diferente a la real. Cuando el programa intenta usar alguna localidad de memoria que no está redireccionada, el programa controlador de memoria virtual entra en acción y elige una porción de la memoria virtual que está en la memoria real en ese momento y guarda su contenido en disco. Esa parte de memoria real liberada se puede reciclar conforme se requiera.

Su desventaja es el tiempo de acceso (casi 200,000 veces mayor que a la memoria principal).

Adicionalmente, se puede contar con más memoria, para lo cual hay que añadir más chips a la tarjeta principal o "motherboard" —en caso de que tenga la capacidad de aceptarla— o bien con tarjetas de expansión de memoria. Toda la memoria superior a 1 Mb se conoce como memoria extendida, aunque se puede configurar por software para que funcione como memoria expandida. Veamos con más detalle estos tipos de memoria.

Memoria expandida. Físicamente, las memorias expandida y extendida son lo mismo, pero varían en cuanto a la forma de configurarla. La memoria expandida se configura declarando en el archivo CONFIG.SYS el driver LIM EMS (o alguno equivalente, como puede ser ABOVE o EMM386), que es el estándar de Lotus, Intel y Microsoft (de ahí su nombre). EMS significa Expanded Memory Specification.

La manera en que funciona la memoria expandida es la siguiente: usted adquiere una tarjeta de expansión de memoria que contiene chips adicionales

de memoria RAM. La tarjeta tiene una cierta cantidad de bancos de memoria de 64 kb cada uno. Estos bancos no tienen asignada una dirección permanente; se accesan uno a la vez, usando un "page frame" o página. Esto se puede hacer al instalarse, junto con la tarjeta, un programa especial llamado administrador de memoria extendida (Expanded memory manager, o EMM). Este programa requiere de 64 kb en la memoria principal para sí mismo.

Cuando un programa desea usar la memoria expandida, llaman al EMM para su administración. Este programa envía una señal a la tarjeta de memoria para habilitar el banco que se va a usar y ayuda al programa a accesar el banco, vía el "page frame". Cuando el programa desea hacer referencia a un banco de memoria en la tarjeta, toma como dirección de memoria el área de los 64 kb que tiene reservados el EMM. Por tanto, cuando se usa la memoria expandida se está "engañando" al sistema operativo, que ve a esta memoria como si se tratase siempre de los 64 kb reservado, cuando en realidad está usando un banco dado de la tarjeta.

Para que pueda trabajar el EMM, los programas de aplicación deben incluir la capacidad de usar esta memoria expandida. No todos los programas cuentan con esta opción (principalmente los programas antiguos).

Generalmente, todas las computadoras a las que se les hace una expansión de memoria (en realidad una extensión) se tienen que configurar por *software* para ser reconocida y usada por programas de las marcas LIM. Este *software* se provee junto con el sistema operativo MS-DOS a partir de la versión 4.01. También existen tarjetas explícitamente de memoria expandida (que se pueden configurar como extendida si así se desea). En este caso, la tarjeta contiene su propio BIOS que contiene el administrador de memoria expandida. Por tanto, la diferencia entre configurar por *software* o por tarjeta puede ser decisiva únicamente cuando se usan aplicaciones que requieren un espacio en modo real superior a aquel que queda libre después de haberse instalado el administrador de memoria expandida. En este caso, conviene contar con la tarjeta de memoria expandida. En cualquier otro caso, no importa cual de las dos técnicas se use.

Memoria extendida. Como ya se mencionó antes, el microprocesador 8088, usado en las PC XT, tiene una longitud para direccionamiento de memoria hasta 1 Mb, de los cuales 640 kb se destinan para aplicaciones y los 384 kb adicionales están reservados para el sistema operativo (ahí se guarda la porción residente de COMMAND.COM o bien se puede usar como "disco RAM", que es un área de memoria que se usa para emular un disco real). La memoria expandida se desarrolla para romper esta barrera de los 640 kb, usando una técnica denomi-

nada "block switching", que consiste en partir en bloques la memoria, para poder direccionarlos.

Es importante hacer notar que no se puede usar la memoria extendida en computadoras tipo XT, ya que éstas requieren que los chips de memoria estén ubicados física y lógicamente en la misma dirección.

A partir del procesador 286, que puede direccionar hasta 15 Mb, se tiene la opción de correr los programas en modo real o bien en "modo protegido", es decir, haciendo uso de la memoria expandida (o extendida), mediante un programa administrador de memoria (*DOS extender*). Estos programas actúan como interruptores entre modo real y modo protegido. Accesan la memoria extendida mientras se encuentran en modo protegido y regresan a la memoria principal para dar los códigos necesarios a la aplicación que se esté usando. Cabe hacer notar que si una aplicación no puede usar el modo protegido, no hay manera de usar la memoria extendida.

Otra manera de acceder la memoria extendida mientras se está en modo real es a través del uso de *drivers*, que son adiciones al sistema operativo que establecen el protocolo de comunicación entre la aplicación y la memoria expandida.

Cuando se introdujo el procesador 80286, con 32 bits para el buffer de direcciones, se pudo romper la barrera de los 640 kb, con lo cual ya no es necesario un emulador (o controlador de memoria expandida) para poder usar la memoria adicional. A este tipo de memoria se le llama memoria extendida, que no tiene un tamaño fijo de bloque, sino que pueden definirse segmentos de tamaño variable, para ajustarse a las necesidades del programa que se ejecute.

Igual que con la memoria expandida, para que un programa pueda hacer uso de ella, necesita incluir en su programación la habilidad para accederla.

Debido a que la especificación de memoria extendida es más reciente que la de expandida, antes de la popularización de **Windows**, eran menos los programas comerciales que podían usarla (por motivos históricos). Sin embargo, desde la aparición de este nuevo ambiente, cada vez son más los programas que usan memoria expandida.

En los diskettes de sistema operativo se encuentra un programa administrador de memoria extendida: el **HIMEM.SYS**.

En la actualidad, casi todas las computadoras tipo PC incluyen ya sea diskettes de **Windows** o bien este programa está ya en la memoria ROM; **Windows** es un nuevo ambiente⁸, que suple algunas de las deficiencias de DOS y que resulta muy amigable para el usuario si éste cuenta con un ratón. Existen varios programas que solamente funcionan (corren) dentro de **Windows**. Estos programas generalmente tienen la capacidad de explotar la memoria extendida de la misma manera que lo hace **Windows**, es decir, reconocen y aprovechan toda la memoria superior a 1 Mb para almacenar datos. En este sentido, si se desea adquirir una nueva máquina, vale la pena detenerse un poco a pensar si se va a trabajar directamente en sistema operativo DOS o si se va a trabajar dentro del ambiente **Windows** y, en función de esto, elegir tanto el modelo de computadora (idealmente un sistema 386SX o superior), la cantidad de memoria RAM, los periféricos (si se va a tener ratón o no), así como los programas que se instalarán en el sistema.

Por compatibilidad con DOS, **Windows** —en su versión 3.1— ya es capaz de reconocer algunas de las aplicaciones DOS y añadir los iconos para poder tener acceso a ellos con solo pulsar el botón izquierdo del ratón dentro del ambiente **Windows**.

8 Ambiente es el entorno que requiere un programa para poder trabajar. En este sentido, los sistemas operativos funcionan como ambientes para los programas de aplicaciones. Un caso muy especial es **Windows**, que requiere a DOS como ambiente para poder ser ejecutado, pero no es nada más un programa, sino que también se utiliza, a su vez, como ambiente para otras aplicaciones.

BUSES

Un elemento importante en toda computadora es la manera en que se interconectan las diferentes partes. Por ejemplo, al insertar una nueva tarjeta en un slot de expansión, la computadora debe ser capaz de reconocerla y poderla utilizar. Además, internamente, el microprocesador requiere comunicarse con toda la electrónica asociada, como son los controladores de disco, monitor, o memoria.

Toda la comunicación interna se lleva a cabo a través de unas líneas llamadas *buses*.

Hay dos formas diferentes de considerar los *buses*. La primera, desde un punto de vista local, interno a la computadora y sin considerar su comunicación con el exterior, donde es posible separar cada conjunto de líneas de acuerdo con su función específica:

- » Existe un *bus* de datos —es decir, por esas líneas sólo transitan datos— y está conectado con todos los módulos que conforman la computadora, como son: el microprocesador, la tarjeta de video, disco duro, unidades de diskette, memoria y los slots de expansión
- » *Bus* de direcciones —a través del cual se especifica la localidad de memoria a la cual se desea acceder
- » *Bus* de protocolo, que permite establecer la comunicación con un módulo específico de la computadora, como puede ser una tarjeta de fax, emulador de multicanal, etc.

A continuación se presenta un diagrama esquemático de los *buses* de datos y direcciones. Los números que se encuentran arriba de una diagonal indican la cantidad de líneas en cada *bus*.

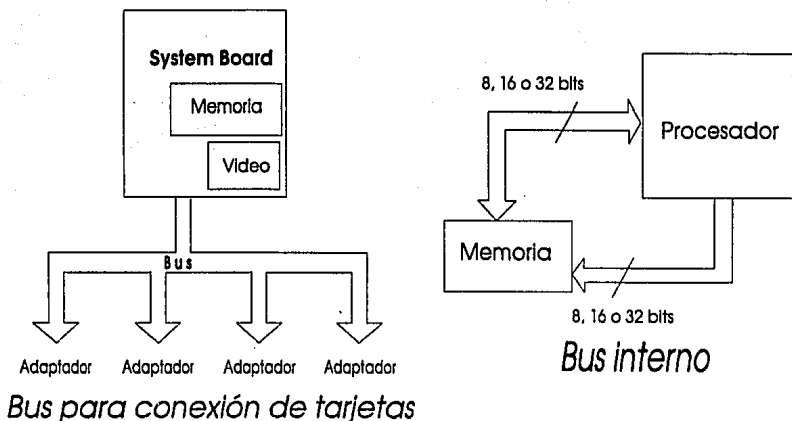


Fig. 5. Diagrama esquemático de los buses

Desde otra perspectiva, podemos considerar un *bus* como un conjunto de líneas de datos, direcciones y protocolo, que satisfacen un estándar específico. Así, encontramos que, para especificar un *bus*, entendido en el sentido más amplio, debemos definir el número de líneas para el *bus* de datos, direcciones y protocolo, así como el tipo de señales de protocolo que incluye.

¿Qué hay en un *bus*? Cada *bus* está conformado por diferentes líneas, cada una de ellas (o conjunto de ellas) con diferentes funciones. Dentro de las señales que se manejan en el *bus* están:

- Señales analógicas, originadas en los transductores o aparatos de medición que se estén usando en laboratorio, líneas telefónicas, etc.
- Señales digitales, que se pueden subdividir en varios grupos:
 - Datos
 - Direcciones, que pueden ser de memoria o para identificar a uno de los módulos de la computadora
 - Señales de control, utilizadas para determinar cual de los módulos es el que tiene el control de los recursos en un momento

dado y para definir la manera en que se cede el control de uno a otro módulo

- Señales de reloj, protocolo y sincronía, ya que, cada vez que dos módulos intercambian datos, debe existir un protocolo para decidir qué señales se pueden transferir en las líneas de datos, direcciones o control
- Interrupciones: En muchos casos, es importante tener la posibilidad de interrumpir al procesador cuando uno de los módulos requiere de atención inmediata. Para ésto, se debe contar con líneas que permitan detener el proceso que se está llevando a cabo y poder dedicar el tiempo al dispositivo que solicitó la atención inmediata. Además, es conveniente contar con una línea que autorice el acceso a uno de los módulos o dispositivos periféricos sin requerir de la atención del procesador. En general, se cuenta con una línea llamada DMA (*Direct Memory Access*), que permite utilizar porciones de la memoria sin que se suspenda el proceso que se esté llevando a cabo
- Adicionalmente, el *bus* cuenta con líneas de alimentación y tierra, para mantener los mismos niveles de voltaje en todo el sistema.

Existen diferentes tipos de especificaciones para los *buses*. Cada uno de ellos está determinado por el número de líneas de datos, de direcciones y de protocolo. Adicionalmente, los protocolos de comunicación pueden variar de uno a otro *bus*. También se debe especificar la capacidad de voltaje y corriente, el tipo de lógica utilizada¹, la velocidad de operación, tipo de conectores usados, dimensiones físicas de las tarjetas que acepta, etc.

Todas las computadoras, sean personales o *mainframes*, requieren de estos conjuntos de líneas para interconectar a todos los grupos de componentes, y se les conoce genéricamente con el nombre de *bus* local. Lo más relevante respecto al bus local, es que permite la interconexión de los módulos de alto rendimiento de la computadora, de modo que se pueda aprovechar al máximo la velocidad con que puede trabajar la electrónica interna de la computadora.

Sin embargo, con el advenimiento de la PC, surge un estándar inicial, llamado ISA (*Industry Standard Architecture*), al cual se ajustan todas las primeras computadoras PC. Con el tiempo, al desarrollarse procesadores más poderosos dentro de la familia 80*** (es decir, los procesadores utilizados en las PC), cada

1 La lógica puede ser positiva, cuando 1=positivo y 0=negativo o negativa, cuando 1=negativo y 0=positivo.

uno con la posibilidad de admitir el direccionamiento de más memoria, se han tenido que ampliar los *buses* de direcciones y los de datos, debido a que el procesador puede usar cada vez más datos en cada ciclo de reloj. Así, con el Pentium, es posible manejar dos datos simultáneamente. Como la velocidad con que se puede tener acceso a un dato es menor a la velocidad con que éste se procesa, resulta necesario duplicar la capacidad del *bus* de datos, para transmitir hasta cuatro datos de un byte simultáneamente, con lo que se hacen más rápidos los cálculos y disminuyen los estados de espera.

También se han hecho mejoras en el rendimiento de los discos duros. El *bus* más veloz que existe actualmente es el llamado VL-*bus*, que es un estándar de interfaz de periféricos establecido en 1993 por la VESA (*Video Electronic Standard Association*). Este estándar surge como respuesta a la diversidad de *buses* que existían en el mercado para el manejo de video; antes de éste, no había un estándar para las tarjetas de video y hay que recordar que uno de los puntos que han dado mayor versatilidad a los sistemas PC es justamente el que existen estándares definidos en la industria, de modo que un fabricante cualquiera de tarjetas para PC —con la función que sea— debe ser capaz de adaptarse a las tecnologías de *bus* existentes, para que el usuario la pueda usar sin requerir una computadora especialmente diseñada para dicha función.

Adicionalmente, IBM diseñó un *bus*, llamado micro-canal, con especificaciones diferentes a las que se manejaban previamente. Este *bus* es de uso exclusivo de IBM. Como respuesta, un grupo de competidores, modificó su estándar ISA, y lo registró como EISA (*Extendid ISA*).

En la actualidad, los dos *buses* más populares son el micro-canal y el EISA, en cuanto a computadoras de alto desempeño se refiere, y el ISA para las computadoras personales de uso doméstico o estaciones de trabajo.

Lo anterior es importante por lo siguiente: En principio, cuando uno adquiere una computadora, puede no tener previstas expansiones futuras. Si, en algún momento, decide adquirir una tarjeta adicional, es necesario que ésta se conforme a los estándares de la computadora con que se cuenta, ya que de otra manera, no funcionará al ser conectada. Por tanto, si se cuenta con una computadora con *bus* micro-canal, será necesario instalar tarjetas con el mismo *bus*. Lo mismo pasa con los *buses* ISA y EISA.

Uno de los grandes avances en el diseño conceptual en las computadoras personales tipo PC, es el de este tipo de sistemas que se conforman a estándares en los *buses*, llamados sistemas con tecnología abierta. De este modo, lo más relevante en el diseño de estas máquinas, es la posibilidad de que terceros

fabricantes creen sus propias tarjetas, adaptables a la computadora personal y que baste con apearse a las especificaciones estándar de *bus* para que pueda operar cualquier computadora que use el mismo tipo de *bus*. Incluso, con los nuevos *buses* micro-canal y EISA, la computadora es capaz de reconocer, al momento de encenderla, que se ha agregado una tarjeta y, ejecutando un programa llamado *setup*, que se incluye en un disco adicional —diskette de configuración— para el caso de la tecnología micro-canal o se encuentra en ROM, en el caso del *bus* EISA, se configura automáticamente, sin necesidad de modificar interruptores —*jumpers*— en la tarjeta principal.

De hecho, haciendo uso de estas tecnologías abiertas, es posible el ir adquiriendo tarjetas con diferentes funciones según las necesidades del usuario en un momento dado. Basta con insertarlas en los slots de expansión y configurar la computadora —generalmente, el proveedor de la tarjeta incluye el *software* asociado para poder configurar el sistema y que funcione adecuadamente.

Cabe hacer notar que, además de los *buses* de computadora, existen otros, como los *buses* en chasis para instrumentación —como CAMAC, VXI y FastBus, diseñados para la interconexión de módulos en un chasis común, donde el chasis provee el voltaje de alimentación para todos los módulos— o los *buses* para módulos “*stand alone*” —como IEEE-488, RS-232 y SCSI, utilizado cuando los módulos son relativamente independientes; más aún, algunos de estos *buses*, como el estándar IEEE-488 se diseñó como *bus* dedicado para aplicaciones de instrumentación, mientras que otros no, como el RS-232, que es un *bus* estándar para comunicación de datos en “serie” (es decir, enviando bit por bit; contrasta con la comunicación en “paralelo”, donde se envía un conjunto de bits simultáneamente).

A continuación se presentan, en términos generales, las especificaciones de los *buses* más populares: ISA, EISA, micro-canal, SCSI y PCI, que es de reciente aparición.

BUS ISA

Las primeras PC contaban con un *bus* de 8 bits para datos y 16 bits de direcciones. Al aparecer la PC-AT, que usa un procesador 80286 de 16 bits, en 1984, se extiende el *bus* para cubrir las necesidades de datos del procesador, y se crea el estándar ISA, con un bus de 98 líneas, de las cuales 16 son de datos y 32 de direcciones. Inicialmente, la velocidad de transmisión de datos en este *bus* era de 6 MHz, aunque posteriormente se actualizó a 8 MHz. Esto quiere

decir que el rendimiento máximo de este *bus* es la transmisión de ocho megabytes por segundo (Mbps).

Aún hoy en día, sigue siendo el *bus* más común en las computadoras PC, debido en parte a su bajo costo y a la gran cantidad de tarjetas que existen en el mercado con esta tecnología.

Tiene 16 líneas para interrupciones, pero doce son dedicadas a funciones del sistema, como teclado, unidades de disco duro y flexible, coprocesador matemático, y puertos serie y paralelo. Las cuatro restantes pueden ser compartidas por una tarjeta de red, unidad de respaldo en cinta o CD-ROM, dependiendo de los requerimientos del sistema.

BUS MICRO-CANAL

En 1987, IBM presenta los *buses* micro-canal, que son incompatibles con la tecnología ISA. Existen dos versiones de este *bus*: 16 y 32 bits de datos. Permite llegar hasta una velocidad de transmisión de información de 33 Mbps. Este *bus* es considerablemente más costoso que el ISA, por lo que no ha tenido tanta aceptación para el público en general, sino que más bien se ha destinado a servidores de red, donde resulta vital la velocidad de transmisión de información, ya que, de otra manera, el poder acceder al servidor puede resultar una operación muy lenta. Otro factor que resulta importante para que no haya proliferado este *bus* es la disponibilidad de tarjetas —e incluso de computadores— que son capaces de utilizarlo: únicamente IBM tiene la patente para poder usarlo.

Las ventajas que ofrece son: la configuración automática de tarjetas nuevas que se instalan en los slots de expansión; menor interferencia eléctrica, lo que conlleva una mayor integridad de los datos; tiene adaptadores especiales, llamados "*bus masters*", que tienen su propio procesador, de modo que pueden efectuar parte del trabajo independiente del procesador principal; de existir una falla, se puede desactivar esa parte, y continuar trabajando con el resto del sistema.

El bus micro-canal de 16 bits cuenta con 116 líneas, de las cuales 77 son de señales, 12 de alimentación, 17 de tierra, 1 tierra de audio y las restantes están reservadas.

De las primeras 77, 24 son de direcciones, 16 de datos y las demás de control. Una de ellas se usa para saber si, en las líneas de direcciones, se está haciendo referencia a una localidad de memoria o a un dispositivo. Otro conjunto de las

líneas de control se usa para el manejo de interrupciones y otra lleva la señal de reloj (14.318 MHz).

En el bus micro-canal de 32 bits, la distribución de líneas es análoga, salvo que en total son 186 líneas, de las cuales 32 son de datos y otras 32 de direcciones, por lo que puede acceder a 4 Gb de memoria.

Bus EISA

Como respuesta de la industria de las microcomputadoras ante la aparición del micro-canal, surge el *bus* EISA, en un acuerdo entre 17 de los principales competidores de IBM. Los slots EISA mantienen el diseño básico de los ISA, pero se hacen más profundas, para acomodar una segunda fila de contactos. Las tarjetas EISA pueden llegar a ambas filas de contactos.

La velocidad de transmisión de datos en un *bus* EISA es de 33 Mbps máximo, con lo que se tiene un desempeño similar al micro-canal y su bus de direcciones es de 32 bits, por lo que se puede tener acceso hasta 4 Gb de memoria.

Bus SCSI

El *bus* SCSI (*Small Computer Systems Interface*) es uno de los más populares para el manejo de unidades de disco duro, además de proveer la interface para el manejo de unidades lectoras de disco óptico (CD-ROM). Este *bus* es el estándar para conectar dispositivos inteligentes en computadoras pequeñas; cuenta con su propio BIOS, para proveer la interface con el CPU y la memoria RAM. Por tanto, todas las operaciones que requieran de toma de decisiones en los dispositivos periféricos, los lleva a cabo el *bus*, sin requerir tiempo de micro-procesador.

Utiliza un *bus* de datos de 9 bits (ocho de datos y uno de paridad, para verificar la integridad de los mismos), un bit de protocolo usando la línea de señal de entrada/salida, la de solicitud (*request*) y de reconocimiento (*acknowledge*).

Bus PCI

En la actualidad, se está trabajando aún para mejorar los *buses*. De este modo, está por aparecer en el mercado un nuevo estándar: el PCI (*Peripheral*

Component Interconnection). Esta tecnología permitirá una velocidad de transmisión de datos de hasta 264 Mbps, usando un *bus* de datos de 64 bits, lo que permitirá presentar gráficos de alta resolución en tiempo real sin deficiencias en la imagen y con muy buena calidad de color.

Surge como respuesta a la necesidad de mejorar la calidad de los sistemas "multimedia", que consisten en un sistema integrado de audio (sonido) y video (imágenes) controladas por computadora. Los datos pueden provenir de video cassetteras, discos compactos (tanto de música como CD-ROM), disco duro, etc. e incluso, hoy en día, sistemas de video-conferencia por computadora.

El *bus* PCI de 32 bits que ya desarrolló Intel y que está aún a prueba, puede mover datos a una razón máxima de 132 Mbps, con un rendimiento promedio de 80 Mbps al desplazar grandes bloques de información.

Para no perder compatibilidad entre el *bus* PCI de 32 y 64 bits, el diseño permite que una tarjeta de 64 bits trabaje en un slot de 32 bits y viceversa.

Este *bus* está diseñado para trabajar junto con procesadores de alto rendimiento, como el Pentium y requiere de una interface ISA, EISA o micro-canal. Los datos se pueden transmitir sin necesidad de ocupar tiempo de microprocesador, por lo que éste se libera de algunos procesos y puede continuar con los cálculos que esté llevando a cabo sin que haya pérdidas de ciclos de reloj.

CONSIDERACIONES GENERALES

En resumen, como se puede apreciar de las especificaciones para los distintos *buses*, dependiendo de la computadora que se tenga y las necesidades futuras, será conveniente elegir uno u otro tipo de *bus*. Por lo general, los modelos "antiguos" de computadoras, incluyen un *bus* ISA, mientras que, para las más actuales, es posible hacer la elección entre micro-canal, ISA o EISA. Una limitante posible es el costo y la disponibilidad de tarjetas para cada uno de esos *buses*. De este modo, el *bus* ISA, por ser el más antiguo, es el más usado, aún en la actualidad, sobre todo en sistemas personales. En cambio, cuando se va a demandar una gran velocidad por parte de la máquina, es conveniente considerar el *bus* EISA. Generalmente, las computadoras que incluyen este bus también pueden tener slots de expansión para tarjetas ISA, lo que les da mayor versatilidad, aunque el costo aumenta.

MONITORES

Hasta hace aproximadamente tres lustros, el dispositivo de salida más común para las computadoras era la impresora. La manera en que funcionaban era la siguiente: se tenía un dispositivo que tenía un teclado y en la parte posterior una impresora. A estos dispositivos se les llama teletipo. Cualquier comando que se diese a la computadora era a través del teclado. Estos comandos se enviaban a la máquina y también aparecían impresos en papel. Lo mismo ocurría con las respuestas que daba la computadora. Unos años después se empezaron a usar monitores para las computadoras grandes (o *mainframes*).

Las primeras computadoras personales —como la **Radio Shack** y la **Apple**— se podían ya conectar a una televisión, evitando así el uso de la impresora. Con la aparición de la PC, también sale al mercado el monitor para computadoras personales. Su funcionamiento es casi igual a un televisor, salvo por las frecuencias de barrido, que para la televisión es fija y para los monitores varía dependiendo de su resolución. Otra diferencia entre las televisiones y los monitores son los blindajes que se incluyen en la pantalla de los segundos, no tanto con el fin de proteger al usuario contra la radiación, sino para tener una mayor nitidez en la imagen, ya que, si un haz se desenfoca (es decir, en lugar de golpear directamente sobre los puntos de fósforo del color que le corresponden, pega entre dos puntos adyacentes), la imagen se hace ligeramente borrosa y se pierde la calidad del color, que hoy en día resulta muy importante para los diseñadores por computadora.

Conforme ha ido evolucionando la calidad de los monitores, cada vez se pueden desplegar imágenes más claras y con mayor resolución. Para explotar esto, existen nuevos programas que permiten la interacción con el usuario mediante pantallas que presentan las opciones disponibles en forma de gráficas y la manera de dar un comando es guiar el ratón hasta el comando deseado y pulsar

uno de sus botones. De esta manera, se evita el tener que memorizar el conjunto de interruptores que se necesitan para ejecutar un proceso. Sin embargo, no se tienen todas las posibilidades que permite DOS directamente. A este tipo de interface entre la computadora y el usuario se le llama GUI (*Graphic User Interface*) y es usado por **Windows** y por **OS/2**, dentro del mundo de las PC IBM compatibles, y por todas las computadoras personales de la marca **MacIntosh**.

SISTEMA DE VIDEO

El monitor, por sí solo, no es capaz de reconocer la información que envía la computadora para ser desplegada, sino que se requiere de una interface que reciba el código binario de la máquina y lo traduzca a señales que pueda entender el monitor. Esta interface se llama "tarjeta de video". Cada tipo de monitor requiere de una tarjeta de video diferente, excepto en el caso de los monitores "multi-scanning" o "multi-sync, que están diseñados para trabajar con cualquier tipo de tarjeta de video. Al conjunto de tarjeta de video y monitor se le denomina sistema de video.

La tarjeta de video tiene una memoria RAM que se usa para almacenar toda la información que envía la computadora y es independiente de la memoria de la computadora en sí. Dependiendo de la resolución del monitor que se tenga, es necesario ocupar más o menos memoria.

Para entender ésto, es necesario analizar la manera en que se maneja la memoria de video y los modos en que puede operar el monitor. La computadora es capaz de manejar dos "modos" de video, que son:

MODO TEXTO

En este modo, se hace uso de los caracteres ASCII, y se puede desplegar únicamente caracteres alfanuméricos (es decir, letras, números, signos de puntuación) y "caracteres semigráficos", que se denominan así porque no son realmente gráficos, pero se pueden usar para hacer dibujos simples. Entre ellos están las líneas verticales, horizontales, sombreados (donde no se puede superponer además un carácter alfanumérico), etc. Si se está trabajando en este modo, como es el caso cuando se encuentra uno en DOS, se aprovecha un chip incluido en la tarjeta de video, que se llama generador de caracteres ASCII, y que incluye la forma de cada carácter y la manera en que ha de desplegarse éste en la pantalla. La manera en que se construyen los caracteres de modo texto es la siguiente: La pantalla se divide en cuadros de un número dado de puntos,

llamada matriz, y cuya dimensión depende del tipo de monitor que se use. Esta es la unidad mínima para despliegue de caracteres de texto, y cada caracter ocupa una de estas matrices; no se pueden superponer dos o más caracteres. Usualmente, se puede desplegar hasta 80 caracteres por renglón y 25 renglones por pantalla —se puede modificar estos valores si se instala el dispositivo ANSI, con lo que se modifica el tamaño por omisión de la matriz. Sin embargo, a menos que sea indispensable, conviene no cambiar estos valores por omisión, ya que se pierde nitidez en el despliegue de la pantalla —y ocupa una porción de memoria que puede ser requerida por programas extensos.

MODOS GRÁFICO

En modo gráfico la técnica para desplegar en pantalla es diferente:

La computadora forma la imagen a desplegar en una memoria especial, separada de la memoria RAM de la computadora, y que se ubica en la tarjeta de vídeo, aunque la computadora la trata como si fuese parte de la memoria principal del sistema. El CPU se encarga de enviar todos los datos correspondientes a cada punto de la pantalla, codificados en varios bits, dependiendo del número de colores que se permita desplegar de modo simultáneo. La tarjeta de vídeo es la responsable de decodificar estos bits —para controlar los haces de electrones— e introducir los pulsos de sincronía horizontal (para indicar un fin de línea) y vertical (que indica el fin de pantalla). Esta señal es enviada al monitor.

La pantalla de modo gráfico está formada por puntos, llamados píxeles (que es un acrónimo de *picture elements*). Cada pixel, a su vez, en los monitores a color, está constituido por tres puntos de fósforo de diferente color: azul, verde y rojo. La manera en que se forma un color es haciendo incidir un haz de electrones (de manera similar a una televisión) sobre cada uno de estos puntos de fósforo. La intensidad de cada uno de ellos determinará el color de cada pixel que se presente en la pantalla.

RESOLUCIÓN

La resolución de una pantalla se define por el número de líneas de barrido y el número de píxeles en cada línea.

Mientras mayor sea el número de píxeles por línea y haya más líneas de barrido, mayor es la resolución de un monitor y mayor es el espacio de memoria

que se requiere para almacenar todos los datos de una pantalla. Si no se maneja color o diferentes tonos de gris, como en algunos monitores monocromáticos, basta un bit para codificar los datos sobre cada punto de la pantalla, ya que cada bit indicará si el punto está encendido o apagado. Por el contrario, si se usa un monitor a color, la pantalla está formada por conjuntos de tres puntos, uno verde, otro rojo y el último azul. Entonces, cada dato para formar el punto en pantalla debe contener cuatro bits, tres de los cuales indican si el punto de un color dado está encendido o apagado y el cuarto bit se usa para la intensidad o brillantez del punto. De esta manera podemos tener acceso a un total de 16 colores diferentes. Si se introduce la opción de manejar dos intensidades para cada uno de los colores del pixel, podemos tener hasta 64 colores diferentes. Mientras más intensidades diferentes tengamos para cada uno de los tres puntos que forman el pixel, mayor será el conjunto de colores, pero también crecerá la cantidad de memoria requerida para almacenar la información referente a cada punto de la pantalla.

Es por esto que algunas tarjetas de video actuales cuentan con 512 kb de memoria RAM y, en algunos casos, ésta se puede ampliar al doble.

Otro factor que se debe tomar en cuenta es que, mientras más datos se almacenen por pantalla, mayor será el tiempo que se requiera para poder formar la imagen completa. Por esto, no es recomendable adaptar un monitor de muy alta resolución a una computadora lenta, como las XT, ya que se va a ocupar demasiado tiempo en formar toda la imagen, lo que hace que el sistema sea más lento.

TIPOS DE MONITORES

Los monitores se clasifican dependiendo de sus características, como son resolución (número de píxeles por pulgada), si son a color o monocromáticos, frecuencia de barrido —de hecho, existen dos frecuencias, íntimamente relacionadas, para caracterizar un monitor: frecuencia de barrido horizontal (que es una medida de cuantas líneas se trazan por segundo, y se miden en kilo Hertz, que se abrevia kHz, o miles de veces por segundo) y la frecuencia de barrido vertical (que determina cuantas pantallas completas se despliegan por segundo, y se miden en Hertz o pantallas por segundo)— y memoria (en la actualidad) que requieren para poder desplegar el máximo número posible de colores en cada caso.

Cabe hacer notar que es diferente la tecnología de los monitores de baja y alta resolución por dos razones fundamentales:

Primero, en los monitores de baja resolución los píxeles tienen un área mayor, por lo que es más difícil que se desenfocan el haz de electrones.

Segundo, el tipo de fósforo con que se construyen tienen diferente persistencia (que quiere decir que, una vez que el haz de electrones incidió sobre un píxel dado, el fósforo brilla, y ese brillo se mantiene por un cierto tiempo). Para monitores de fósforo con mayor persistencia se requiere refrescar la imagen menos frecuentemente.

Monocromáticos Existen tres tipos de monitores monocromáticos. El primero y más antiguo es el MDA (monochrome display adapter), que solamente puede desplegar texto. Después surgieron los monitores TTL (*transistor-transistor logic*), que se conectan a una tarjeta de video llamada Hércules y que permite desplegar tanto texto como gráficas. Sin embargo, casi todos los programas de aplicación que manejan gráficas no son compatibles con este tipo de monitor, por lo que se requiere de "emuladores" —programas que traducen el código de los gráficos para poder ser desplegados en estos monitores— para poder ver en pantalla los gráficos de dichos programas.

Las características generales de los monitores TTL son:

Resolución de 720 x 348 píxeles, la matriz de texto es de 9 x 14 píxeles y requieren de una memoria RAM de 64 kb para almacenar toda la información de lo que se está desplegando en pantalla. La velocidad de barrido es de 18.4 kHz.

El tercer tipo de monitor monocromático es el RGB (*red, green, blue*) monocromático, que funcionan con la tarjeta de video CGA -2 (*color graphics adapter*) y cuyas especificaciones son las siguientes:

Resolución de 640 x 200 píxeles con dos colores y de 320 x 200 píxeles con 4 colores, matriz de texto de 8 x 8 y requieren de 64 kb de memoria RAM. La velocidad de barrido es de 15.75 kHz.

CGA La tarjeta de video CGA puede operar tanto en modo texto como gráfico, produciendo dibujos y caracteres en diferentes formatos y colores. Está diseñada para trabajar con diferentes tipos de monitores, desde televisiones hasta monitores de alta resolución. Para poder trabajar con T.V. se requiere adicionalmente de un dispositivo, llamado modulador de frecuencia. Estos monitores tienen las mismas características que los RGB, pero despliegan las imágenes a color, a escoger en un conjunto de 16 colores, cuatro simultáneos. Prácti-

camente todos los programas comerciales son compatibles con este tipo de monitor. En su época, hace aproximadamente un lustro, fueron los monitores a color más populares y, precisamente para mantener la compatibilidad de los nuevos programas con las máquinas antiguas, se ha mantenido a este tipo de monitor como una opción de despliegue.

EGA Estos monitores combinan la versatilidad de los monitores CGA con mayor calidad para el despliegue de texto, cuentan con más colores y mayor resolución gráfica. Fueron populares en algún tiempo, pero poco después de su aparición surgió el VGA, que se hizo mucho más popular.

Sus características son:

Resolución variable hasta 640 x 350 píxeles con 16 colores, matriz de texto de 8 x 14, requieren de 256 kb de RAM y la velocidad de barrido es de 21.8 kHz.

VGA Es el monitor de mayor resolución que existe, junto con el VGA mejorado (Super VGA, VGA plus y VGA extendido) y con manejo de más colores (hasta 256). Es compatible con los monitores EGA y CGA, es decir, el *software* desarrollado para estos dos tipos de monitor también funciona en el VGA. El monitor VGA es el estándar actual —de hecho casi cualquier computadora de las que hay actualmente en el mercado, incluyen este tipo de monitor.

El estándar a futuro son los monitores SuperVGA, aunque recién está entrando al mercado nacional. La diferencia entre los monitores VGA y SuperVGA es la resolución, ya que los últimos pueden desplegar hasta 1,024 por 768 píxeles y la cantidades de colores disponibles, que en VGA y SuperVGA actuales alcanza los 16.78 millones de colores diferentes.

Las especificaciones de estos monitores son:

Resolución de 640 x 480 píxeles, matriz de 8 x 14 para caracteres de texto, requiere de 256 kb de memoria RAM y la velocidad de barrido es de 31.5 kHz.

Monitores multiscaning o Multi-sync Estos monitores tienen la capacidad de poder ajustar la velocidad de video en función de la tarjeta con que se cuente, y son compatibles con todas las tarjetas de video. Por tanto, dependiendo de la tarjeta de video con la que se estén usando, varía su resolución y demás especificaciones. Cabe hacer la aclaración de que los monitores multi-sync pueden ser de dos tipos: el primero se adapta a toda la gama de frecuencias

en un intervalo dado (por ejemplo, entre 50 y 90 Hz), mientras que el segundo está diseñado para operar a frecuencias específicas, por ejemplo 56, 60, 70 y 72 Hz. Este último tipo de monitores resultan menos versátiles, sobre todo si se considera la evolución que se está presentando en los monitores, ya que, mientras más resolución se tenga, se requieren frecuencias de barrido mayores.

TARJETAS ACELERADORAS DE VIDEO

El sistema de video comprende tres partes: adaptador de video, *software* de manejador de video y monitor. El adaptador puede ser una tarjeta añadida de 8 o 16 bits, o bien puede estar incluida en la tarjeta principal, que es lo más usual ya que, de otra manera, requiere utilizar uno de los slots de expansión, restando capacidad de expansión para otros dispositivos.

El adaptador de video determina el tipo de monitor que se puede conectar a la computadora. Dentro de sus especificaciones se incluye la resolución que puede manejar, la capacidad de colores y la manera en que la imagen pasa del sistema al adaptador de video, que puede ser a través del *bus* de expansión o mediante una conexión directa de alta velocidad con el procesador.

La resolución y el número de colores disponibles están limitados por la memoria de video, que se encuentra en el adaptador de video. Conforme más colores se requieran, mayor debe ser la memoria; lo mismo ocurre con la resolución. De este modo, para poder desplegar un punto o pixel en la pantalla utilizando 16 colores, son necesarios 4 bits, para 256 colores se requiere de 8 bits por pixel, etc. Cada que se duplica el número de colores, se requiere de un bit más para su codificación. Además, mientras mayor sea la resolución, hay más píxeles, por lo que se requiere de mayor capacidad de memoria. Por ejemplo, con un pixel, basta con 4 bits en cuatro colores, 1 byte en 256 colores, etc. Un monitor VGA puede desplegar 640 x 480 píxeles, por lo que, para poder presentar una pantalla completa en cuatro colores se requiere de 144,000 bytes, en 256 colores se requieren 288,000 bytes, etc.

De considerar estos valores, realmente no es "demasiada" la cantidad de memoria que se requiere pero, ¿qué pasa si deseamos usar más colores o mejorar la resolución, como es el caso si la labor principal de la computadora es en el área de las artes gráficas? Por ejemplo, si manejamos 16.78 millones de colores, requerimos de 24 bits para su codificación, lo que equivale a 3 bytes por cada punto en pantalla. Para un monitor VGA, con este número de colores, requeriremos de 432,000 bytes para poder trazar una pantalla; en SuperVGA, cuya

resolución es de 800 x 600 píxeles, requeriremos de 1,440,000 bytes para poder desplegar una sola pantalla!

Debido a lo anterior, es muy importante el decidir, antes de elegir el sistema de video que se vaya a adquirir, el tipo de programas de aplicaciones que se utilizarán. Para el usuario medio, que requiere de procesadores de texto, hojas electrónicas, contabilidades, análisis financiero, etc., bastará con monitores VGA o SuperVGA; incluso, este tipo de monitor bastará si se requiere de algunos programas de diseño gráfico básicos, como **PaintBrush**, **Artes y Letras** o **Corel Draw**, o bien programas de autoedición como **Ventura Publisher** o **PageMaker**. En cambio, si se requiere de la edición por computadora de fotografías, es recomendable tener la mayor cantidad de colores y la mejor resolución posible.

Mientras más bytes se requieran para poder formar una imagen, es mayor la cantidad de información que hay que trasladar de la memoria RAM de la computadora a la de video, para que se despliegue a razón de 60 a 72 veces por segundo.

En el sistema de video tradicional, el microprocesador debe encargarse de transferir la información al sistema de video. La velocidad de esta operación depende de la potencia del microprocesador, la velocidad con que se puede transmitir la información a través del *bus* y la memoria de video disponible. Existen varias maneras de optimizar el rendimiento del sistema de video:

1. Un procesador más rápido, ya que, de este modo, es posible procesar la información de video con mayor velocidad.
2. Manejadores de *software* actualizados. Aún cuando se cuente con determinada tarjeta de video, se puede actualizar el *software* asociado (siempre verificando previamente la compatibilidad entre *hardware* y *software*). Uno de los ejemplos más populares es el ambiente de trabajo **Windows**, que demanda de una mayor cantidad de memoria de video debido a que opera únicamente en modo gráfico. Recordemos que, en modo texto, el sistema de video cuenta con un generador de caracteres, que convierte cada byte en el carácter correspondiente y que ocupa 9 x 16 píxeles. Cualquier programa que trabaja bajo **Windows** puede ser instalado sin necesidad de especificar (más que en la instalación inicial de **Windows**) el tipo de monitor o impresora que se esté utilizando. De este modo, si se actualiza el manejador de video, éste tendrá efecto sobre todas las aplicaciones **Windows** que se utilicen. Es por esto que resulta vital que el manejador de video sea lo más eficiente posible. Típicamente, al dibujar imágenes a la memoria del sistema, no se hace uso

de la tarjeta de video pero sí del manejador de video, por lo que el desempeño se basa únicamente en este último.

3. Modificar la arquitectura de memoria del adaptador de video, instalando una tarjeta aceleradora de video que incluya memoria RAM de video —llamada VRAM—, y que contiene adicionalmente otro puerto de conexión con el *bus*, de modo que se duplica la velocidad de transmisión de información. En este marco, se han desarrollado diversas tarjetas, llamadas “aceleradores de video”, que se instalan en uno de los slots de expansión y cuya exclusiva función es permitir un manejo más dinámico y eficiente del sistema de video. A continuación analizaremos con más detalle los aceleradores de video.

Los aceleradores de video logran mayor velocidad debido al *bus* doble y a que el procesador no requiere enviar toda la información pixel por pixel, sino que el sistema de video está preparado para recibir instrucciones codificadas, que describen la manera en que deben generar la imagen en pantalla. Por ejemplo, para el trazo de una línea, en vez de cambiar los atributos de todos los pixeles que la formarán, se indicarán los puntos inicial, final y color de la línea. La tarjeta aceleradora de video está programada para decodificar estos datos y poder trazar la línea en pantalla.

Existen dos tipos de tarjetas aceleradoras de video: las aceleradoras y los coprocesadores. Las primeras consisten en un procesador paralelo al principal, con funciones fijas, con la habilidad de manejar funciones gráficas específicas relacionadas con programas o aplicaciones determinadas. En cambio, el coprocesador permite mayor flexibilidad de programación, es decir, las funciones que puede realizar no están pre-programadas o fijas, sino que se pueden modificar, lo que les da mayor versatilidad, aunque su costo es significativamente mayor a las aceleradoras.

Las partes de una tarjeta aceleradora son las siguientes:

1. Memoria de video. Todas las tarjetas aceleradoras de video requieren su propia memoria RAM, independiente de la principal.
2. Acelerador de video. Este es un circuito integrado dedicado a la decodificación de la información que envía el sistema para ser desplegada y, en algunos casos, también se encarga de la conversión de la señal digital a analógica, para que salga directamente al monitor; en otros casos, existe otra parte de la tarjeta que se encarga de esta tarea.

3. BIOS de video. El BIOS es la parte donde residen los programas que dan soporte a las funciones implementadas para poder codificar y decodificar la información a ser desplegada.
4. Adaptador para video. Es la parte de la tarjeta que tiene salida por la parte posterior del slot de expansión que ocupa, y es justamente ahí donde se debe conectar el monitor.
5. Interface para el *bus*. El tipo de conector que se tenga, así como las especificaciones del fabricante sobre el *bus* para el cual fue diseñada la tarjeta, determinan si se puede o no conectar la tarjeta en los slots de expansión de una computadora determinada. Los *buses* más populares para video son: ISA, EISA, Micro-canal y VL-BUS.

CONSIDERACIONES GENERALES

Conforme se han ido popularizando las computadoras en todas las áreas del quehacer humano, se han desarrollado nuevos programas orientados a diferentes actividades. Una de las que mayor relevancia ha tomado es la del diseño —tanto gráfico como técnico— por computadora. De hecho, es posible ya el controlar máquinas herramienta con la computadora, de modo que se puede diseñar cualquier parte mecánica con la computadora y luego conectarla a la máquina herramienta para que se construya físicamente.

Por todo esto, cada vez es más necesaria la calidad más alta posible —técnicamente hablando— para el despliegue. Lo anterior incluye no solamente la resolución, sino el tener monitores más grandes. Inicialmente, los monitores tenían una única presentación: 12". Sin embargo, conforme las necesidades por una mejor calidad de imagen van aumentando, se han popularizado monitores cada vez más grandes —actualmente entre 14" y 16". De hecho, pese a que ocupan más espacio sobre el escritorio, son más comunes los monitores de 14" que los de 12". Además, existen monitores especiales para la formación de textos por computadora, que permiten el despliegue de una página completa (es decir, cambia la proporción entre la longitud horizontal y vertical del monitor).

Otro de los factores que se evalúan actualmente de un monitor es la distorsión de la imagen en los bordes e incluso los reportes de las revistas especializadas ya han definido diferentes parámetros de evaluación. En términos generales, para el usuario común, que no requiere de una gran precisión en sus diseños o despliegue, no requieren de monitores más caros y sofisticados que lo que puede presentar un VGA o SuperVGA.

UNIDADES DE ALMACENAMIENTO

Una unidad de disco es un dispositivo que permite el leer o guardar información. Cada unidad de disco requiere de un controlador para poder conectarse con la computadora. Usualmente un controlador puede dar servicio hasta a dos unidades de disco. Cada unidad de disco tiene un nombre asignado, que es una letra de la A a la Z.

UNIDADES DE DISCO FLEXIBLE

En el caso de las unidades de disco flexible, reciben el nombre de A o B, y se requiere de diskettes, que se insertan en las ranuras de la parte frontal de la computadora como se muestra en la figura 4. Los diskettes son discos plásticos magnetizados, con una cubierta de plástico. La forma de guardar información en los diskettes es la siguiente: el diskette está formado por un disco de mylar recubierto con un material magnetizable —usualmente óxido de hierro. (El sistema es muy parecido al que se usa para grabar música en cintas magnéticas —cassettes). Esta cubierta es muy delgada. La superficie es tratada, por la computadora, como un arreglo de puntos, cada uno de los cuales se fija al equivalente magnético de "1" o "0". Debido a que la ubicación exacta de estos puntos no está determinada precisamente, es necesario incluir ciertas marcas que ayudan a la cabeza lectora/grabadora a encontrar las posiciones adecuadas. Esta es parte de la razón por la cual se tiene que formatear (dar formato) a los discos antes de poder usarlos.

Un diskette de 5 1/4" de doble densidad puede almacenar cientos de páginas de texto. Estos son los diskettes más comunes, y se usan en todas las computadoras PC (modelos XT, AT, 386SX y 386). Existen también los diskettes de "alta densidad", que pueden guardar casi cuatro veces más información que los de doble densidad, pero se requiere de modelos de computadora más avanzados que la XT (o del controlador y la unidad de disco) para poder ser usados (AT o 386). Todos los programas y archivos de datos que requieren los programas que se usen se pueden guardar en diskettes. Cuando se desea modificar una serie de datos, como por ejemplo una carta que se escribió en un procesador de textos, la computadora lee el archivo del diskette y lo "carga" o guarda en memoria, para poderlo procesar.

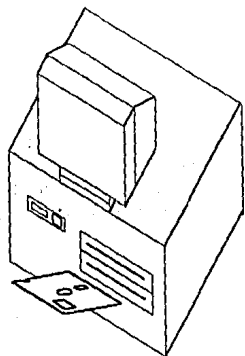


Fig. 4. Modo de insertar los diskettes

Los cuidados que se deben tener al manejar los diskettes son:

- Guardarlos en sus fundas de papel.
- Pegarles etiquetas en la parte superior delantera y escribir en ellas qué tipo de información contiene el diskette.
- Escribir en las etiquetas con plumón con punta de fieltro; no usar plumas o lápices.
- No doblarlos.
- No tocar la superficie magnética expuesta (ranura central inferior).

- No exponerlos a campos magnéticos (como imanes, televisiones, teléfonos).
- No exponerlos al sol ni humedad.
- No colocar sobre ellos objetos pesados ni filosos.
- De preferencia, guardarlos en sus cajas, para que no se contaminen de polvo.

TIPOS DE DISCOS FLEXIBLES

Los diskettes vienen en dos presentaciones: 5 1/4" y 3 1/2". No es fácil distinguir entre los diskettes de doble y alta densidad a simple vista, por lo que hay que verificar las especificaciones en la caja original o bien guiarse por los detalles siguientes: para los diskettes de 5 1/4", observe si tiene un anillo en color claro en la parte central; si lo tiene, se trata de un diskette de doble densidad y, en caso contrario, es un diskette de alta densidad. Para los diskettes de 3 1/2", vistos por la parte posterior, si tienen un "cuadrado" oscuro en la parte izquierda, entonces son de alta densidad. La razón para ésto es la siguiente: en las primeras unidades de disco de alta densidad no había un controlador que pudiese determinar por medios electrónicos si el disco era de alta o baja densidad. Para evitar esta limitación, se diseñó un sensor mecánico que permitía a la unidad de disco determinar si se estaba trabajando con diskettes de 720 kb o 1.44 Mb y, de esta manera, poder fijar los parámetros adecuados. En la actualidad los drives de 3 1/2" ya no requieren de este cuadrado, por lo que se puede llegar a usar discos de baja densidad como de alta. Sin embargo, existe el riesgo de perder información. En la figura 5 se presenta un dibujo de cada uno de ellos.

Los diskettes de 5 1/4" tienen en el extremo superior derecho una ranura. Si se cubre, se protege el diskette contra escritura, es decir, sólo se puede leer información del diskette, pero no se puede escribir nada en él. Esto es útil para los programas de aplicación, pero no para los diskettes que contienen datos o textos del usuario, ya que no se pueden actualizar o modificar. Esta protección sirve para prevenir el borrado accidental de algún archivo importante. También se pueden proteger archivos cambiando sus atributos, usando el comando `attrib`, de DOS.

De manera análoga, los diskettes de 3 1/2" tienen un "tab" de plástico deslizable, de manera que se puede abrir o cerrar una ventana cuadrada. Si ésta está cerrada, el diskette acepta que se escriba información en él; si está abierta, se encuentra protegido contra escritura.

Estos discos están dentro de una cubierta rígida, pero se consideran flexibles debido a que el medio en que se almacena la información es plástico recubierto de una sustancia magnetizable, igual que los diskettes de 5 1/2".

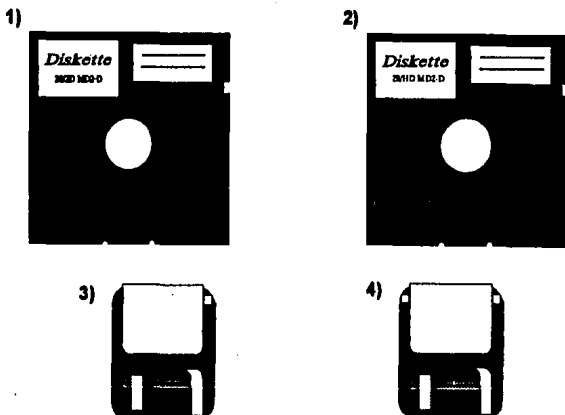


Fig. 5. Tipos de diskettes:

- 1) 5 1/4" doble densidad. 2) 5 1/4" alta densidad. 3) 3 1/2" doble densidad.
4) 3 1/2" alta densidad.

DISCOS DUROS

Algunas computadoras (casi todas actualmente), además de tener una unidad de diskettes, tienen integrados "discos duros", que tienen capacidades de almacenamiento de información considerablemente mayores que los diskettes. Consisten en una serie de discos metálicos (*platters*) cubiertos con superficies magnéticas (otra vez, el mecanismo es similar al de los cassettes). La velocidad de acceso a un disco duro es mucho mayor que a un diskette, por lo que la interacción entre el usuario y la computadora se vuelve más eficiente, sobre todo cuando los programas que se estén usando requieran estar leyendo constantemente información del disco. Para poder lograr que operen adecuadamente, se requiere de un ambiente "limpio", es decir, libre de polvo y otros contaminantes, por lo que están sellados.

Actualmente el disco duro de menor capacidad que existen el mercado es de 30 Mb, el mínimo recomendable si se manejan muchos programas es de 40 Mb y la capacidad máxima de un disco duro es de 2048 Mb. En principio, puede uno adquirir el disco duro que más le convenga (en función del tipo de uso que se le vaya a dar, ya que, dependiendo del tipo de disco duro, varían los tiempos de acceso y los costos), para cargar ahí todos los programas de aplicación y/o sistemas operativos que desee.

Si se desea tener dos sistemas operativos diferentes en el mismo disco, éste se tiene que "partir" en dos unidades lógicas, esto es, un mismo disco duro se formatea de modo que el sistema lo reconozca como dos discos duros diferentes (por ejemplo C y D); en este caso, si se va a usar la computadora como servidor (*server*) de una red, y también como estación de trabajo, será necesario crear al menos dos particiones lógicas: una para DOS y la otra para OS/2. Otro motivo por el cual se requiere este tipo de particiones lógicas es la combinación capacidad del disco duro/versión del DOS que se esté usando ya que, para versiones anteriores a la 4.01, el sistema operativo no puede acceder más allá de 32 Mb.

A las unidades de disco duro se les asigna como nombre una letra que puede ser C, D, y hasta Z.

TIPOS DE DISCOS DUROS

En la actualidad, existen varios tipos fundamentales de disco duro, entre los que destacan los MFM, RLL y los SCSI.

Los últimos se denominan también "discos inteligentes".

Para poder entender las diferencias entre los tipos de discos duros que existen, es necesario aclarar antes un punto. El disco duro, en sí, consiste en un conjunto de platos recubiertos de material magnetizable, el eje sobre el cual giran y las cabezas de lectura-escritura. Estas cabezas se mueven simultáneamente, es decir, en un momento dado todas se encontrarán apuntando a exactamente el mismo punto de los diferentes platos. Adicionalmente, se debe incluir un "controlador de disco duro". El controlador de disco duro, que se debe incluir siempre y cuando se cuente con una unidad de disco duro (es opcional), es el dispositivo que controla el funcionamiento de las unidades para almacenamiento de información no removible, montados dentro de la computadora y con mucha mayor capacidad que los diskettes. Es la electrónica encargada de colocar las cabezas en la posición adecuada para que se lleve a

cabo la lectura o escritura, codifica la información antes de ser escrita al disco, insertando información necesaria para que sepa exactamente dónde está escribiendo, traduce los códigos a datos para el procesador cuando se lee la información de disco y lleva a cabo una verificación de errores, llamada "verificación cíclica de redundancias" (CRC—*cyclic redundancy check*)¹. Para poder llevarse a cabo la codificación de información y verificación de errores, la tarjeta controladora cuenta con una pequeña área de memoria RAM (conocida como *buffer*), que no puede ser usada por la computadora, sino únicamente para las funciones antes mencionadas.

Debido a que se tiene que almacenar gran cantidad de información en regiones muy pequeñas del disco, el controlador de disco se perdería si no se agregara información adicional, para poderse orientar. Por ejemplo, si en un sector hubiese gran cantidad de ceros, no existiría ninguna variación del campo magnético. Si de pronto, al estarse leyendo la información aparece un uno lógico, no sabría la posición exacta de éste, es decir, no sabría a cual de los 512 bits del sector asignar el valor "1".

Para poder entender la manera en que se codifica y decodifica la información en un disco duro, es conveniente hacer un paréntesis para tratar la manera en que se graba ésta. La cabeza del disco es un electroimán, que genera un campo magnético al pasar una corriente por él. Este campo ocasiona que el material que recubre los platos se oriente en un sentido dado, de la misma manera que una brújula se orienta con el norte magnético de la Tierra. La corriente puede fluir en dos direcciones en el electroimán, por lo que los campos magnéticos pueden tener una de dos orientaciones diferentes. Cuando se invierte el sentido de la corriente en el electroimán, se graba un patrón similar en el medio del disco. Al leer los datos, el proceso es inverso al de escribir; el campo magnético en la superficie del disco genera una corriente —que recibe el nombre de corriente inducida— en la cabeza lectora. La presencia o ausencia de una corriente inducida determina el valor del punto, que puede ser "1" o "0". Debido a que las cabezas pueden sentir únicamente variaciones de campo, para que haya mayor integridad en la información almacenada, se codifican los "1" consecutivos como un campo magnético, pero alternando la polaridad (es decir,

¹ El CRC consiste en que a un bloque de datos que se van a grabar en disco se les aplica una fórmula matemática que da como resultado un número de muchos dígitos. Este resultado se escribe en el disco a continuación de los datos del sector. Al ser leídos los datos, se vuelve a aplicar la fórmula a los datos y se compara el resultado con lo que se había grabado previamente. Si no coinciden, el controlador sabe que ha ocurrido un error. Esta técnica es lo suficientemente sofisticada como para detectar errores múltiples. Dependiendo de la severidad del error, el controlador puede o no ser capaz de recuperar la información. La probabilidad de que ocurra un error recuperable de este tipo es muy baja ($1/10^{10}$) y la probabilidad de que ocurra un error no recuperable (cuando la información de ese sector se ha perdido definitivamente) es de $1/10^{11}$.

si ocurren tres "1" consecutivos, el primero tendrá polaridad norte-sur, el segundo sur-norte y el tercero norte-sur. De este modo, es más fácil para las cabezas el poder distinguir el valor de los datos y la cantidad de bits con un valor igual. El problema más severo se presenta con los bits que tienen un valor "0". Los datos se deben codificar, de modo que no se presenten "muchos" ceros consecutivos —la cantidad admisible varía entre dos y siete, dependiendo del tipo de disco duro y formato que se esté usando para su codificación y los algoritmos de decodificación.

El proceso de conversión de bits a campos magnéticos en el disco involucra tres tipos de procesos diferentes: primero, a nivel de sistema operativo, el paso de información a la electrónica asociada con el disco. Segundo, al recibir el controlador de disco los datos, los codifica e introduce información adicional usando técnicas de corrección de errores (llamados ECC o *error-corrected codes*). Los ECC introducen bits de "redundancia", que permiten al controlador reconstruir la información en caso de que se presente un error. Tercero, esta serie de bits se convierte a códigos RLL (aunque existen otras técnicas, según se discute un poco más adelante), que son los que añaden suficientes bits de redundancia como para garantizar que no haya más que cierta cantidad de bits con valor "0" consecutivos. El proceso de escribir la información a disco genera un nuevo campo magnético cada que ocurre un "1" y, si no ocurren lo suficientemente seguidos, el reloj que intenta llevar un registro de la posición en el disco se puede "perder". Por supuesto, el introducir bits adicionales para lograr la recuperación eficiente de información del disco lleva a un límite en la cantidad de información que se puede guardar, entendiendo por información los datos que está generando el usuario. En general, todos estos códigos y datos adicionales, aumentan en un 50% el número de bits almacenados.

Existen varios tipos de disco duro. Cada uno de ellos utiliza una técnica diferente para resolver el problema de posible pérdida de información y sincronía al leer el disco.

Discos FM y MFM La primera forma en que se resolvió esto fue usando uno de cada dos bits como pulso de reloj —es decir, cada sector incluía 1,024 bits, de los cuales 512 eran datos y los otros señales para el controlador de disco duro. A esta técnica se le llamó Frecuencia Modulada (FM). Se presentaba un problema, que era la baja capacidad de almacenamiento de estos discos.

Para resolverlo, se diseñó otra manera de codificar los datos, que depende de la información que precede al pulso de reloj —o señal para el controlador— y se denominó Frecuencia Modulada Modificada (MFM —*modified frequency modulation*); casi duplica la capacidad de almacenamiento de los discos FM.

Esta técnica fue la más usada hasta hace uno o dos años; casi todos los fabricantes de discos duros establecían dentro de las especificaciones que se trataba de discos MFM.

Discos RLL Otra técnica de codificación que se hizo popular en estas épocas, pero considerablemente más costosa (entre un 50% y 100% más cara que la MFM) es la RLL (*run length limited*); traduce los datos a una serie de códigos especiales, de modo que no se admita más de un cierto número de ceros consecutivos. Por ejemplo, si se usa RLL 2,7 los códigos se eligen de manera que las secuencias de ceros en los códigos siempre varía entre dos y siete.

Discos scsi En los discos duros tipo SCSI (*Small Computer Systems Interface*, que se pronuncia "escosi" y que utilizan el *bus* del mismo nombre) todas las funciones de la tarjeta controladora están incorporadas directamente en la unidad de disco, incluyendo el buffer de memoria temporal donde se almacena el contenido del sector que se esté leyendo o escribiendo. Más aún, incluyen una memoria RAM "cache" de disco, de modo que no se lee únicamente un sector del archivo, sino varios a la vez, para que el proceso de suministro de datos al procesador sea más eficiente², ya que la información se tiene lista para ser enviada a la memoria RAM convencional cuando la requiere el procesador. Estos discos requieren de un puerto SCSI, al cual se pueden conectar hasta siete periféricos, entre discos duros, unidades de disco flexible de alta densidad, discos ópticos, *scanners*³ e impresoras, lo que permite el tener conectados más periféricos a una computadora que no tiene ranuras de expansión libres.

Discos IDE (*Integrated Drive Electronics*) Este tipo de disco duro es uno de los más populares en la actualidad. Se originó a partir del estándar ST-506, que manejaba IBM en sus primeras computadoras tipo AT. Un controlador de disco IDE puede dar soporte únicamente a dos unidades de disco, que deben ser magnéticos, es decir, no admite una unidad de CD-ROM. Otra de sus limitaciones es la capacidad máxima de disco duro que puede admitir (528 Mb), debido a la manera en que se configura. Veamos ésto con detalle. Los controladores IDE identifican que se ha instalado un disco y determinando su capacidad tomando en cuenta la cantidad de cabezas, cilindros y sectores por track que contiene. Debido a que el sistema operativo DOS especifica que hay 512 bytes por sector,

- 2 Esto se debe a que, de otra manera, al momento en que el procesador solicita información del disco duro, se tiene que iniciar la búsqueda, para lo cual antes hay que colocar las cabezas en el lugar adecuado, luego iniciar la lectura de un sector, verificar que no haya errores de lectura, decodificar, y enviarla a la memoria RAM convencional para que pueda ser usada por el procesador.
- 3 Un *scanner* es un dispositivo que permite la "digitalización de imágenes", traduciendo patrones de claro y oscuro impresos en una hoja a información que puede entender la computadora.

la capacidad del drive está dada por el producto de cabezas por cilindros por sectores por 512. El número máximo que puede tomar cada uno de estos parámetros es: 65,536 cilindros, 16 cabezas y 255 sectores. Adicionalmente, debido a que los dispositivos IDE utilizan un servicio del BIOS llamado INT 13h (el 13 es un valor hexadecimal, no decimal, y está indicado por la "h"), se requiere consistencia en los datos que este último puede admitir y los de IDE. Los valores que puede manejar el BIOS son: 1,024 cilindros, 255 cabezas y 63 sectores por track. Por tanto, para que haya compatibilidad en las cantidades que se manejan, queda como capacidad máxima los valores más pequeños para cada uno de los elementos que determinan la capacidad de disco: 1,024 cilindros, 16 cabezas y 63 sectores, lo que resulta en 528,482,304 bytes, equivalentes a 528 Mb como capacidad máxima reconocible por los discos tipo IDE.

Actualmente, existe una propuesta dual de los fabricantes originales de este tipo de discos para poder admitir capacidades mayores, dependiendo del ambiente o sistema operativo que se esté usando. La primera, que opera bajo Unix, consiste en saltarse el BIOS al acceder los discos duros. Utiliza una tabla de parámetros de dispositivos (DPT, que significa *Drive Parameter Table*) para pasar la información sobre la capacidad del disco del drive al sistema operativo.

La segunda técnica, si se opera bajo DOS o **Windows**, consiste en crear una tabla de parámetros de dispositivos mejorada (EDPT, o *Enhanced Drive Parameter Table*) durante el POST (*Power On Self Test*⁴). Durante el proceso de arranque de la computadora, la tabla es generada y se colocan en ella todos los valores que describen las características del disco, junto con datos adicionales (*checksum*⁵) que se usan para verificar la integridad de los mismos. Esta tabla se envía a BIOS —que finalmente es la parte del sistema capaz de reconocer los dispositivos periféricos conectados a la computadora— no en forma de tabla, sino como la dirección de un bloque de memoria, donde está residente la tabla de valores para los parámetros del disco. De este modo, la capacidad máxima para un disco IDE puede llegar hasta 8.4 Gb.

Otro factor importante al considerar el tipo de disco duro que se va a adquirir es la velocidad con que se puede almacenar y recuperar información de él. Por supuesto, las limitantes que se tengan no son únicamente debidas al disco, sino al tipo de *bis* con que se cuente. Los discos IDE transfieren datos de dos formas:

-
- 4 Consúltense el apéndice III, donde se trata en más detalle el proceso de inicialización del sistema al encender la computadora.
 - 5 El *checksum* consiste en sumar todos los bytes de datos y tomar uno o los dos bytes menos significativos del resultado; este valor es el *checksum*. Si se modifica un bit, inmediatamente será detectado al no coincidir la suma total con el valor de *checksum* almacenado.

- PIO (*Processor I/O*), que depende del procesador para el control de la transferencia de datos; es el sistema tradicional usado en los discos IDE, ya que, cuando apareció IDE era la forma más rápida de transmisión de información.
- DMA (*Direct Memory Access*), que permite al disco ser el administrador de la transmisión, una vez que el procesador pre-fija ciertos parámetros. Su velocidad depende de la del *bus* utilizado. Actualmente, existen dos tipos de transmisión usando el DMA: tipo B, con velocidad máxima de 4 Mbps (Mega bytes por segundo) y tipo F, que se puede usar en *buses* PCI y que extienden este límite hasta 8.33 Mbps. También, con el nuevo *bus* PCI, se hace posible una nueva forma de transmisión de datos usando el DMA, llamado *Scatter-Gather*. Con ésto, es posible transferir cuatro bloques de 4 kb cada uno, formando un bloque de 16 kb, con una sola petición de entrada/salida, en lugar de las cuatro que tomaría por métodos convencionales.

SCSI vs. IDE Debido a que SCSI es un *bus*, no sólo un controlador, permite mayor versatilidad que IDE. Por ejemplo, SCSI es muy flexible en el tipo de dispositivos conectados a él: soporta discos duros, scanners, CD-ROM, e impresoras, entre otros; además, admite hasta siete dispositivos en un solo adaptador SCSI y todos los dispositivos conectados a este *bus* son inteligentes; cada uno de ellos tiene su propio ROM que contiene los parámetros básicos de operación. Otra de sus ventajas es que la longitud de cable para conexión de dispositivos externos es de hasta 6 metros, mientras que IDE admite hasta 18 pulgadas, lo que restringe su uso para dispositivos internos únicamente.

Otra diferencia importante radica en las velocidades de transmisión: IDE admite hasta 8.33 Mbps, en tanto que SCSI acepta hasta 20 Mbps.

Sin embargo, los discos IDE son los más populares hoy en día, en parte debido a su costo. Por último, con las mejoras que se están haciendo a IDE, ya no va a haber limitaciones en cuanto a la capacidad de disco se refiere y al tipo de dispositivos que se puedan conectar.

Cada modelo de computadora puede tener cierto tipo de disco duro o flexible. Por ejemplo, las XT no pueden usar unidades de diskette de alta densidad, tanto por la velocidad de la electrónica como por el tamaño del *bus* (estas máquinas aceptan solamente tarjetas controladoras de disco de 8 bits y las unidades de diskette de alta densidad tienen *bus* de 16 bits).

DISCOS ÓPTICOS

Los discos ópticos, también llamados CD (*compact disk*), surgen para cubrir una importante necesidad: la posibilidad de trasladar grandes volúmenes de información —generalmente bases de datos, que pueden contener catálogos de bibliotecas, colecciones de imágenes, o enciclopedias, entre otros— de manera compacta, a poder ser accesibles a un gran público.

La capacidad de almacenamiento de un CD de baja densidad, es de alrededor de 550 Mb, que equivaldría a 190,000 páginas de texto o más de 5,000 imágenes a color en alta resolución y para uno de alta densidad, es de 1 Gb (el doble de la capacidad del disco de baja densidad).

Inicialmente, los CD eran de sólo lectura, es decir, se requería contar con un lector de discos ópticos (parecido a un *Compact Disk Player*) y los discos ya pre-grabados —que se han usado para comercializar fundamentalmente bibliotecas de imágenes o bases de datos. El usuario puede llamar una imagen dada o consultar la base de datos que contiene el disco. Sin embargo, no se puede modificar la información del disco. A este tipo de discos se les llama CD-ROM, ya que es información no volátil y no modificable (equivale a la memoria de sólo lectura o ROM, y de ahí su nombre).

Los CD-ROM se basan en la misma tecnología usada para los CD de música. Los datos son codificados y almacenados en una superficie metálica altamente reflectiva que está protegida por una cubierta transparente de policarbonato. Los datos —que pueden ser texto, sonido, video o una combinación de éstas— se almacenan en una espiral continua, que inicia en el centro del disco y continúa hacia el borde externo. En este sentido, equivale al zurco de un disco de acetato, sólo que girando en sentido inverso. Cada vuelta sucesiva de la espiral está separada 1.6 micras de las adyacentes.

Un laser —generalmente de galio-arsénico— ubicado en la cabeza de lectura pasa por una lente colimadora (que hace que los rayos de luz sean paralelos), luego por una lente convergente para enfocar el haz en un solo punto sobre el disco. Incidentalmente, el sistema que hace girar al disco es de velocidad variable, de manera que en tiempos iguales se lea la misma cantidad de información independientemente de si se está en el extremo interior o exterior del disco. Por ésto, el disco rota más rápidamente cuando se está leyendo el interior que el exterior.

La información se guarda en forma de "huecos" y "llanuras". Los huecos dispersan la luz de manera que sólo una pequeña porción se refleja a la cabeza de lectura. Por el contrario, las llanuras reflejan una cantidad de luz mucho

mayor. Un fotodetector (detector de luz) sensa la intensidad de la luz reflejada y la convierte a corriente, asignando una baja corriente a la proporcionalmente baja reflectividad de los hoyos y asignando un valor lógico de "1". La alta reflectividad de las llanuras resulta en un valor lógico de "0", debido al aumento de corriente.

Estos valores lógicos ya los puede entender la computadora. La velocidad de lectura de un CD es de 150 kb por segundo aproximadamente.

Además de los CD-ROM, están apareciendo en el mercado los CD que permiten al usuario almacenar y modificar la información. Estos discos requieren de tecnologías mucho más sofisticadas, sobre todo en cuanto al tipo de materiales usados y la potencia del láser utilizado.

Hagámos un breve paréntesis para analizar con más detalle la manera en que están fabricados los CD de lectura/escritura.

Un disco compacto (CD) está formado por una película de un material especial —en general son aleaciones amorfas de tierras raras y metales de transición (hierro o cobalto)— embebidos en un polímero transparente. En la parte inferior se deposita una capa metálica, de modo que, al incidir el láser, éste sea reflejado lo más eficientemente posible.

El material amorfo es la región donde se almacena la información, y se elige debido a que este tipo de material no presenta variaciones no deseadas en la reflexión de la luz.

Existen dos tipos de discos ópticos sobre los que se puede grabar información: los magneto-ópticos (MO), y los que usan cambios de fase (PC).

En esencia, son iguales, pero cada uno aprovecha una característica diferente del medio para poder almacenar la información. En los PC, el laser debe ser capaz de elevar la temperatura del punto sobre el cual se desea grabar un bit (y que recibe el nombre de dominio) más allá del punto de fusión del material, que se encuentra en un arreglo amorfo; al enfriarse, solidifica formando una estructura cristalina, que tiene diferentes propiedades ópticas que las regiones amorfas. Por tanto, al leer la información, resulta posible detectar los bits cuyo valor es "1".

En cambio, en los discos magneto-ópticos, el laser, al momento de escribir información, calienta el dominio hasta la temperatura de Curie, y el enfriamiento se lleva a cabo en un campo magnético, por lo que las moléculas del material

se imantan, cambiando así sus propiedades ópticas. De hecho, bajo este esquema, cuando el laser incide sobre la superficie y se refleja, al leer la información, la luz reflejada tiene una polarización elíptica. El ángulo de rotación del eje principal de la elipse varía dependiendo de la orientación magnética del medio —entre 0.2 y 0.4 grados— por lo que resulta posible decodificar la información al medir estos ángulos. De este modo, si se usa esta técnica para almacenar información en un disco óptico, es necesario escribir todos los bits: aquellos que tienen un valor "1" estarán magnetizados en un sentido y los que tienen un valor "0" estarán magnetizados inversamente.

En la actualidad, los discos MO son los más usados, y sobre los que se ha avocado la mayor parte de las investigaciones.

Los avances en cuanto al aumento de capacidad de almacenamiento de estas unidades se ha llevado a cabo lentamente. No es sino hasta hace alrededor de uno o dos años que se logró alcanzar los 650 Mb. Actualmente, la capacidad de almacenamiento de una unidad magneto-óptica es de 1.3 Gb, pero se preve que se duplique antes de 1995, otra vez en 1996 y otra más antes de 1998. De este modo, a principios del año 200 se espera que la capacidad de almacenamiento de estas unidades llegue a los 10.4 Gb.

La clave para poder pasar de los 650 Mb a los 1.3 Gb de almacenamiento fue el desarrollo de un método, llamado ZBR (*zone bit recording*), que permite almacenar diferentes cantidades de información en regiones internas y externas del disco, a diferencia de los discos magnéticos, que almacenan 512 bytes por sector, independientemente de que sea un sector en la porción más externa del disco o en la más interna. De este modo, también se debe tener un control sobre la velocidad de rotación del disco, de manera análoga a los CD-ROM, que se presentaron en el apartado anterior.

En los discos magneto-ópticos actuales, la información se almacena como campo magnético presente, para un bit cuyo valor es "1", y ausencia de campo, para uno que vale "0". La nueva tecnología que se está desarrollando codificará los datos de manera diferente: tomará en cuenta las transiciones entre existencia y ausencia de campo como "1" (en caso de existir la transición, sin importar que se trate de campo a no campo o viceversa). Los ceros se determinan por ausencia de transición entre presencia y ausencia de campo, es decir, si existe un campo por una cierta distancia recorrida por la cabeza, se asigna un valor de cero al bit. Lo mismo ocurre si no hay campo en la misma distancia. Varios ceros consecutivos estarán almacenados como la presencia o ausencia de campo por una distancia equivalente a lo que se requiere para almacenar ese número

de bits. Mientras menores puedan ser estas regiones donde se pueda llevar a cabo una transición, mayor será la capacidad del disco.

Una de las limitaciones en cuanto al tamaño de los dominios es su dependencia con la longitud de onda del laser utilizado y el inverso de la "apertura numérica" (NA) —equivalente al inverso del "número f" de las cámaras fotográficas— de la lente utilizada. Actualmente, se usa un laser con longitud de onda de 780 nanómetros (un nanómetro equivale a 10^{-9} metros, es decir 0.000000001 metros) y lentes con NA alrededor de 0.55. El diámetro de un dominio está dado por:

$$d = 1.22 \lambda / NA$$

Si se desea aumentar la capacidad de almacenamiento de un disco MO, es necesario reducir la longitud de onda del laser utilizado; se están llevando a cabo investigaciones para poder construir diodos laser con la longitud de onda deseada, ya que el usar otro tipo de laser implicaría tener que transmitir la luz a través de fibras ópticas desde la fuente hasta la cabeza lectora-grabadora..

Para poder usar un CD-ROM es necesario, aparte de instalarlo adecuadamente en la computadora siguiendo las instrucciones del manual, el instalar el *software* necesario para que el sistema lo reconozca. Usualmente, los programas que se requieren están incluidos en el paquete. Una vez instalado todo, se le asigna una unidad de disco y se puede acceder a toda la información de igual manera que a un diskette o disco duro.

En la actualidad ya existen unidades de disco óptico que permiten leer y grabar información. Los más comunes son los llamados WORM (*write once read many*), que permiten al usuario grabar una vez la información, pero no se puede borrar. Los discos CD que permiten grabar y borrar datos del disco compacto son aún costosos (alrededor de 8,000 dólares), pero tienen la ventaja de poder almacenar entre 550 Mb y 2 Gb de datos en cada disco!

Otro tipo de unidades ópticas que ya están en el mercado son las llamadas "unidades ópticas de respaldo", que usan discos encerrados en cartuchos. Tienen dos presentaciones, análogas a los diskettes, solamente que tanto los discos ópticos de 5 1/4" como los de 3 1/2" tienen una cubierta (cartucho) rígida. Estos discos se usan básicamente como unidades de respaldo, y no pueden leer los CD-ROM que no tengan el cartucho. Esto se debe a que la electrónica es más delicada y requiere de un medio completamente rígido para su adecuado funcionamiento. El costo de cada cartucho es elevado, ya que, a final de cuentas, reemplaza un disco duro de hasta 600 Mb.

Están apareciendo ya en el mercado los discos ópticos de lectura escritura para que el usuario pueda generar su propio "master" (igual que en los discos de acetato) de CD, que puede contener bases de datos o enciclopedias que no se actualizan muy frecuentemente, de modo que se puedan reproducir en grandes volúmenes y comercializarse. El costo de un disco "grabable" es de alrededor de 80 USD, aunque se espera que, en unos dos o tres años, su precio baje hasta los 20 USD.

DISCOS FLÓPTICOS

El nombre es una contracción de "floppy" (o disco flexible) y "óptico". Es una tecnología joven, que usa diskettes parecidos a los de 3 1/2". Usa técnicas ópticas para colocar la cabeza de lectura/escritura. El medio en que se almacena la información contiene pistas que crean áreas de contraste (claros y oscuros). Un haz de luz reflejado de estas pistas puede encontrar la posición adecuada. En caso de no encontrar estos patrones de claros y oscuros, el sistema supone que se trata de un diskette convencional de 3 1/2". Si existen los patrones, acciona una cabeza de alta densidad, capaz de leer la información almacenada flópticamente.

La óptica se usa únicamente para colocar la cabeza de alta densidad, no para almacenar o leer la información; ésto se hace por medios magnéticos, al igual que en los discos flexibles o duros.

Su capacidad de almacenamiento es de alrededor de 20 a 25 Mb.

ORGANIZACIÓN DE UN DISCO

Un disco, sea duro o flexible (diskette), se puede considerar como un archivero, donde se guarda una serie de folders, cada uno de ellos etiquetado de manera que se sepa qué información contienen.

En el caso de los diskettes, el archivero cuenta, usualmente, con un solo cajón, llamado directorio raíz (no se acostumbra usar subdirectorios en los diskettes debido a su baja capacidad de almacenamiento), y los folders que contiene se denominan archivos. Un disco duro, debido a que puede almacenar muchísimos más archivos que los diskettes, se debe organizar en directorios, que se pueden ver como los diferentes cajones de un archivero. Cada uno de estos directorios puede a su vez contener separadores, o subdirectorios. Es recomendable crear un directorio para cada programa o tipo de documento con que

se vaya a trabajar, ya que, de este modo, cuando se requiera dar mantenimiento al disco duro, se pueda localizar exactamente qué programa o grupo de archivos de trabajo se va a borrar y que al hacerlo se borren todos los archivos secundarios que éste use.

Es importante recalcar que no conviene mezclar en un mismo directorio el programa y los archivos de trabajo, tanto por los problemas que se pueden tener al dar mantenimiento al programa —como el actualizar la versión que se esté manejando— como por eficiencia, ya que un directorio con gran cantidad de archivos es más lento que si se cuenta con pocos archivos.

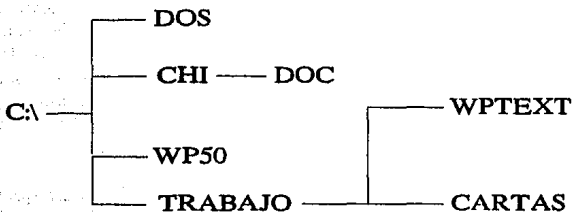


Fig. 6. Organización de un disco

Se conoce como directorio raíz a la unidad de disco (que equivale a referirse a un archivero en particular), y como subdirectorios a cada uno de los cajones que contiene. Hay una excepción a la analogía que estamos usando, ya que en el directorio raíz se pueden guardar archivos, mientras que en el archivero se tiene que usar siempre algún cajón (no se guardan documentos encima del archivero!).

Una buena manera de organizar el disco duro es, antes que nada, cargar en el directorio raíz únicamente los programas que son vitales para el arranque del sistema y generar los subdirectorios que se vayan necesitando, tanto para el almacenamiento de programas como para los archivos de trabajo. Si se cuenta con dos procesadores de texto, es más conveniente generar un subdirectorio para cada procesador y, dentro de éste un sub-subdirectorio para los archivos de trabajo. De esta manera, no será necesario estar recordando en qué procesador de texto se escribió cierto documento para poder editarlo.

ESTRUCTURA DE UN DISCO

Aquí se presenta una visión global de cómo están estructurados internamente los discos, sean duros o flexibles, y la forma en que éstos almacenan la información.

Estructura física Los *drives* y sistema operativo de la computadora establecen la capacidad del disco usado, pero la estructura es esencialmente la misma, independientemente del formato. Los datos se guardan en la superficie del disco en una serie de círculos concéntricos, llamados *tracks*. El número de *tracks* varía dependiendo del tipo de disco; por ejemplo, los discos de 3 1/2" de 1.44 Mb tienen 80 *tracks*, los discos de baja densidad (tanto de 3 1/2" como 5 1/4") usan 40 *tracks* y los discos duros típicamente tienen entre 300 y mil *tracks*. Cada *track* está dividido en segmentos, llamados sectores. El tipo de disco y su formato determinan cuántos sectores hay en cada *track*. Los discos de baja densidad tienen 17 sectores/*track*, los de alta densidad 18 y los más recientes, de 2.88 Mb, tienen 36 sectores/*track*. La cantidad de datos que se pueden almacenar en cada lado de un disco depende del número de *tracks* (densidad), el número de sectores y el tamaño de los sectores. La densidad varía considerablemente en función del drive. El estándar de doble lado doble densidad puede almacenar 40 *tracks* de datos. En ocasiones —sobre todo cuando tratamos con discos duros— resulta más cómodo el referirse a todos los *tracks* que están a la misma distancia del centro del disco. A éstos se les denomina cilindros.

La ubicación, tamaño y número de sectores por *track* se controlan por *software*. Esto se fija al momento de formatear el disco. Usualmente se usan sectores de 512 bytes en el ambiente DOS.

La configuración para XT es 2 lados, 40 *tracks* y 9 sectores por *track*, con lo que se pueden almacenar 360 kb. Los diskettes formateados en AT o sistemas 386SX y 386 se formatean con 80 *tracks* (se conocen como *quad-density*), 9 sectores por *track* y 2 lados. En la figura 7 se muestra la estructura física de un diskette.

Estructura lógica Independientemente del tipo de disco que se usa, todos los discos que trabajan bajo sistema operativo DOS están formateados de la misma forma lógica: los lados, *tracks* y sectores se identifican numéricamente usando la misma notación, y algunos sectores siempre se reservan para programas especiales e índices que requiere DOS para manejar el disco.

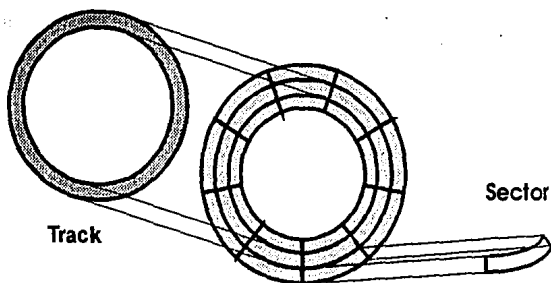


Fig. 7. Estructura física de un disco

Para un diskette de 360 kb con 40 *tracks*, éstos están numerados del 0 al 39, empezando por la parte exterior del disco. Los lados están numerados como 0 y 1. Los sectores se numeran del 1 al 9. Nótese que los *tracks* y lados empiezan en 0, en tanto que los sectores en 1.

El BIOS (*Basic Input/Output System*) localiza los sectores en un disco usando coordenadas tridimensionales (*track*, lado y sector). En cambio, DOS la ubica por sector, y numera los sectores secuencialmente de afuera hacia adentro. La secuencia inicia con el primer sector del disco: sector 1 del lado 0 y *track* 0, siguiendo con los siguientes sectores del mismo lado y *track*, luego cambia de lado y finalmente de *track*.

ORGANIZACIÓN INTERNA DE UN DISCO

El proceso de dar formato divide los sectores del disco en cuatro secciones, para diferentes usos: Boot, FAT, directorio y datos. El tamaño de cada sección varía dependiendo del formato, pero la estructura y el orden de las secciones permanece igual.

El primer sector (sector 1, *track* 0, lado 0) contiene información de sistema operativo y no se tiene acceso directo, más que mediante DEBUG.COM o utilerías

especiales, como **Norton Integrator**, **PC-Tools**, etc. Esta es la zona de **Boot**, que es un programa que inicia el proceso de cargar sistema operativo desde diskette. Todos los diskettes tienen el área de boot aunque no contengan sistema operativo.

El *File Allocation Table* (**FAT**) inicia en el sector 2 del *track* 0 y lado 0. Contiene la lista oficial del formato de disco y el mapa de la ubicación de los sectores usados por los archivos. DOS usa el FAT para llevar un registro del uso de espacio en el disco. Cada entrada en la tabla contiene un código específico para indicar qué espacio se está usando, cuánto hay disponible y cuál no es usable (sectores dañados). Debido a que ahí se encuentra la información vital para poder manejar los archivos, hay 2 copias idénticas del FAT, y están ubicadas consecutivamente en el disco, ya que, en caso de que una se dañe, se puede usar la otra y no se pierde la información. El tamaño del FAT varía en función del formato del disco y las particiones lógicas que se tengan.

A continuación están los sectores que forman el directorio, que se usa como índice, identificando cada archivo en el disco con una entrada que contiene el nombre y extensión del archivo. Una parte de la entrada es un número que apunta al primer grupo de sectores que tienen la información del archivo (o su contenido). El tamaño del directorio también varía en función del formato; en diskettes de doble lado-doble densidad ocupa 7 sectores.

Luego sigue la zona de datos que es donde se almacena físicamente el contenido de los archivos y ocupa el resto del disco. Aquí la información se almacena en *clusters*, formados por 2 sectores en diskettes de 360 kb, y cuyo tamaño ha ido variando en función de la capacidad de los discos duros: conforme se hacen mayores, se requiere que los clusters sean más grandes —en la actualidad hasta de 32 kb.

Veamos estas regiones en detalle:

Boot. El *boot* consiste en un programa en lenguaje máquina que inicia el proceso de cargar DOS en RAM. Primero, verifica que el disco tenga sistema operativo. Si no se encuentra el sistema, despliega un mensaje de error, que también se almacena en el *boot*, y que se puede modificar, siempre y cuando se respete el número de caracteres del mensaje. De hecho, el sistema operativo no está ubicado físicamente en este sector, sino que ahí reside un conjunto muy pequeño de instrucciones que indican a la computadora el lugar donde residen los programas de inicialización del sistema.

Directorio. El *directorío* contiene la mayor parte de la información de los archivos almacenados en disco, incluyendo el nombre, tamaño, entrada inicial en FAT, fecha y hora en la que fue creado y atributos especiales. Lo único que no contiene es la ubicación exacta de los clusters individuales que forman el archivo; ésto lo contiene el FAT.

Hay una entrada para cada archivo en el disco, incluyendo entradas para los subdirectorios y la etiqueta del disco. Cada entrada tiene 32 bytes, por lo que un sector puede almacenar 16 entradas (=512/32). Un disco para XT puede contener hasta 112 entradas, ya que el espacio disponible para el directorio es de siete sectores. Los subdirectorios se tratan como archivos, por lo que no hay límite para el número de archivos que puede contener un subdirectorio. Cada entrada se divide en 8 campos:

| <i>Campo</i> | <i>Offset</i> | <i>Descripción</i> | <i>Tamaño (bytes)</i> | <i>Formato</i> |
|--------------|---------------|-----------------------|-----------------------|------------------------|
| 1 | 0 | Nombre del archivo | 8 | ASCII |
| 2 | 8 | Extensión del archivo | 3 | ASCII |
| 3 | 11 | Atributos | 1 | codificado por bits |
| 4 | 12 | Reservado | 10 | ceros |
| 5 | 22 | Hora | 2 | codificado por palabra |
| 6 | 24 | Fecha | 2 | codificado por palabra |
| 7 | 26 | Cluster inicial | 2 | palabra |
| 8 | 28 | Tamaño del archivo | 4 | entero |

Cuando un archivo se borra, sólo se afectan dos cosas en el disco: la letra inicial del nombre —se fija a E5 hexadecimal y aparece como una σ si se despliega el sector del directorio donde se encontraba el nombre del archivo usando **debug**, **Norton Utilities** o un programa equivalente— y la cadena del FAT se borra. Todo lo demás permanece igual, incluyendo la información en la zona de datos. La información se puede recuperar si no se ha reescrito en los clusters que ocupaba el archivo (por ésto hay que tener cuidado, ya que, una vez borrado un archivo, DOS hará que los primeros clusters que se ocupen con un archivo nuevo o al actualizar uno ya existente, serán precisamente éstos).

El campo de atributos se codifica de la siguiente manera:

| Bit | | | | | | | | Valor | | |
|-----|---|---|---|---|---|---|---|-------|-----|----------------|
| 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Dec | Hex | Significado |
| | | | | | | | 1 | 1 | 1 | Sólo lectura |
| | | | | | | 1 | | 2 | 2 | Archivo oculto |
| | | | | | 1 | | | 4 | 4 | Sistema |
| | | | | 1 | | | | 8 | 8 | Etiqueta |
| | | | 1 | | | | | 16 | 10 | Subdirectorío |
| | | | | | | | | 32 | 20 | Archivo |
| | | | | | | 1 | | 64 | 40 | No se usa |
| 1 | | | | | | | | 128 | 80 | No se usa |

La hora está codificada de acuerdo con la fórmula:

$$\text{Hora} = \text{Horas} * 2048 + \text{minutos} * 32 + \frac{\text{segundos}}{2}$$

donde la hora está en formato cero a 23 hrs. y el valor resultante se convierte a hexadecimal.

La fecha, que es un valor hexadecimal, está codificada de acuerdo con la fórmula:

$$\text{Fecha} = (\text{año} - 1980) * 512 + \text{mes} * 32 + \text{día}$$

El tamaño del archivo es un entero sin signo (positivo) de cuatro bytes, que le permite al archivo tener una longitud mayor que la capacidad del disco! Para DOS, éste es el tamaño real del archivo, aunque no es necesariamente cierto, ya que, por ejemplo, con algunos procesadores de texto, el indicador de fin de archivo no coincide con el de DOS, por lo que el tamaño puede reportarse como hasta 128 bytes mayor de lo que en realidad es.

La zona de datos se va asignando de un cluster a la vez. Usualmente, un archivo se va ubicando en bloques contiguos; sin embargo, hay casos en que el archivo se parte en bloques no contiguos, especialmente cuando se añade información a un archivo existente o cuando un archivo nuevo se almacena en el espacio que dejó disponible un archivo borrado. Para hacer contiguos los bloques que ocupa un sector, hay que copiarlos a un diskette nuevo, recién

formateado, o utilizar programas especiales, como PCTools o defrag, que es un comando de DOS 6.

FAT. El FAT lleva un registro que muestra cómo se está usando el espacio en disco. Hay dos formatos: 12 y 16 bits; el primero para diskettes o discos de hasta 16 Mb y el segundo para discos duros. Sin embargo, cuando la capacidad del disco excede de 32 Mb, se tienen que buscar alternativas, ya que el sistema operativo, con los formatos standard, no es capaz de reconocer más allá de 32 Mb. Esto se debe a que, por un lado, el tamaño standard de un sector es de 512 bytes y, por el otro, que los sectores se numeran secuencialmente. Inicialmente, DOS almacenaba estos números en un formato de enteros positivos de 16 bits, por lo que se puede numerar hasta 65,536 (64 k) sectores diferentes; y como el tamaño de un sector es de 512 bytes, tenemos que DOS sólo podía reconocer hasta 32 Mb (= 64 k * 512). A partir de la versión 4 del sistema operativo cambia el standard para la capacidad de un cluster, de modo que se pueda tener acceso a discos duros de mayor capacidad. A continuación se presenta una tabla donde se muestra el tamaño de los clusters para diferentes capacidades de discos.

| <i>Capacidad</i> | <i>Tamaño de cluster</i> | <i>Número de bits en FAT</i> |
|---------------------|--------------------------|------------------------------|
| <i>Diskette</i> | | |
| 3 1/2": 2.88 Mb | 1 k = 1024 bytes | 12 bits |
| 3 1/2": 1.44 Mb | 0.5 k = 512 bytes | 12 bits |
| 3 1/2": 720 kb | 1 k = 1024 bytes | 12 bits |
| 5 1/4": 1.2 Mb | 0.5 k = 512 bytes | 12 bits |
| 5 1/4": 360 kb | 1 k = 1024 bytes | 12 bits |
| <i>Discos duros</i> | | |
| 16 Mb o menos | 4 k | 12 bits |
| 17 Mb a 128 Mb | 2 k | 16 bits |
| 129 Mb a 256 Mb | 4 k | 16 bits |
| 257 a 512 Mb | 8 k | 16 bits |
| 513 Mb a 1024 Mb | 16 k | 16 bits |
| 1025 Mb a 2048 Mb | 32 k | 16 bits |

Como ya se mencionó antes, cada entrada en el área de directorio incluye un campo que indica el número del cluster en el cual inicia el archivo. El FAT tiene un espacio reservado para cada cluster. Si vemos el valor del FAT para el primer cluster de un archivo, éste contendrá el número para el siguiente cluster del archivo. De esta manera, los valores en FAT están encadenados para dar a

DOS una forma de saber dónde se encuentra la continuación de un archivo. Veámoslo con más detalle.

El FAT está organizado como una tabla de hasta 4,086 números (si se están usando 12 bits) que varían (en valor) de cero a 4,095 (cero a FFF hexadecimal), con una entrada para cada cluster en la zona de datos. El número en cada entrada indica el estatus y uso del cluster que corresponde a la entrada en FAT. Si la entrada tiene un valor de cero, indica que el cluster está disponible para su uso. En cambio, si la entrada es 4,087 (FF7 hex) indica que el cluster no puede ser usado (está dañado). La información del FAT se puede ver como una cadena de datos, donde el valor de la entrada correspondiente a un cluster de un archivo apunta al siguiente cluster donde hay información de un archivo. Por ejemplo, si la entrada 29 tiene un valor de 32 indica que el siguiente cluster con información del archivo es el cluster 32. Si el valor de la entrada es FFF hex, indica que ése es el último cluster que contiene al archivo.

Hasta aquí, aparentemente no hay problemas para ubicar el siguiente cluster de un archivo. Sin embargo, la información está codificada de modo que se hace más eficiente el manejo con el costo de pérdida de claridad. Cada número se codifica con tres dígitos hexadecimales. Cada entrada del FAT se organiza en pares, donde cada par ocupa tres bytes. Éstos se decodifican de acuerdo con el siguiente patrón: si un par de entradas del FAT es hex 123 y 456, entonces los tres bytes que los contienen sería 23 61 45. Análogamente, si tres bytes en el FAT son AB CD EF, entonces los valores son DAB y EFC.

La primera entrada del FAT es un byte que contiene la información sobre el formato del disco, de acuerdo con la siguiente tabla:

| <i>Formato</i> | <i>Valor del Byte</i> |
|----------------|-----------------------|
| 360 kb | FD |
| 720 kb | F9 |
| 1.2 Mb | F9 |
| Disco duro | F8 |

Cabe hacer notar que, debido a la manera en que DOS almacena la información en disco, el espacio mínimo que puede ocupar un archivo es de 1 cluster, por lo que existe espacio desperdiciado, que en promedio asciende a medio cluster por archivo. Por ejemplo, un archivo que contenga un solo byte (un carácter), ocupará todo un cluster.

MODELOS DE COMPUTADORAS TIPO PC

Se conoce genéricamente como "PC compatibles" a todas las computadoras personales que son compatibles con las microcomputadoras IBM. Dentro de este género se encuentran fundamentalmente cuatro tipos de sistemas diferentes entre sí y se explican más adelante. Cada uno de ellos recibe su nombre por el microprocesador sobre el cual está basado; todos estos microprocesadores forman parte de la familia 80***, que tienen características similares. Por ejemplo, un sistema 286 es aquel que contiene como CPU un microprocesador 80286.

Las especificaciones generales de una computadora o sistema que determinan su velocidad y capacidades son:

1. La frecuencia máxima del reloj del sistema. Este punto es importante debido a que existen computadoras, como las XT, cuyo reloj trabaja más de tres veces más lento que el de algunos sistemas 386. La medida para la frecuencia del reloj son los megaHertz, es decir, millones de ciclos por segundo. Cada pulso de reloj se supone que representa el tiempo que requiere el procesador para ejecutar una función, pero éste no es siempre el caso. Los procesadores 8086, 8088, 80286 y 80386DX, en ocasiones, requieren varios pulsos de reloj para ejecutar instrucciones complejas. Por otro lado, los 80386SX y 80486 siempre completan una función en un ciclo de reloj. El que los 386DX no lo hagan y los SX sí, se debe a que el DX fue el primero en aparecer en el mercado, mientras que el SX es de aparición más reciente que el 486, e incorpora parte de esa nueva tecnología en su diseño.
2. La cantidad de memoria que puede direccionarse. Cada modelo de microprocesador tiene un cierto número de líneas para establecer la dirección (direccionar) de una localidad de memoria. A estas líneas se les llama bus de

direcciones. Por ejemplo, un microprocesador que cuente con 20 líneas de direcciones puede hacer referencia hasta a 2^{20} localidades de memoria diferentes. Mientras mayor sea el bus de direcciones, se tiene acceso a más localidades de memoria. La ventaja de poder manipular más memoria es que se puede tener programas muy complejos y extensos o bien poder ejecutar simultáneamente dos o más programas.

3. El ancho del bus de datos. Para poder tener acceso a un cierto dato, el CPU requiere primero direccionar la memoria —apuntar al punto en el cual se encuentra la información. Una vez hecho esto, se puede leer el dato. Para que no se mezclen direcciones y datos, existe un bus de datos, análogo al de direcciones. Entonces, al leerse un dato, éste es enviado a través de las líneas de datos hacia el microprocesador. Mientras más grande sea el bus de datos, más bits de información se pueden transferir por vez entre el CPU y la memoria, por lo que la computadora será más eficiente que una con bus de datos menor. Los buses de datos pueden ser de 8, 16, 32 o 64 bits.
4. Tamaño del registro. Cada microprocesador puede leer y almacenar internamente un cierto número de datos, antes de trabajar con ellos. Al número de bits que puede almacenar se le conoce como tamaño del registro. Mientras mayor sea su capacidad para almacenar información, menos referencias tendrá que hacer a la memoria para buscar datos faltantes para ejecutar el proceso que se le demanda.
5. Memoria caché. Todos los procesadores que trabajan a velocidades superiores a 25 MHz son más rápidos que la memoria, por lo que es necesario almacenar la información en la memoria caché, de modo que el sistema sea más eficiente. Debido a que la memoria caché interna está físicamente dentro del procesador, son más eficientes que la memoria caché externa, ya que estas últimas dependen del diseño global del sistema. Los sistemas 486 tienen integrada una memoria caché de 8 kb, mientras que el Pentium tiene dos memorias caché internas de 8 kb cada una. La memoria caché externa es opcional. Además, dependiendo del diseño de la tarjeta principal, es o no posible el agregar memoria cache externa a la computadora, así como la máxima cantidad de memoria adicional que admite.

8088

El primer grupo lo conforman todas las computadoras PC de IBM y compatibles cuyo CPU es el microprocesador 8088. Estas son las máquinas más antiguas y se conocen como "PC".

La velocidad de operación de estos procesadores y su electrónica asociada (sin incluir los buses o líneas, tanto de direcciones como de datos) es de 8 MHz, es decir de ocho millones de ciclos por segundo. En cada ciclo se ejecuta una instrucción, se lee un dato o se almacena un byte de información. El bus de datos es de 8 bits y el tamaño de registro de 16 bits. Cuenta con cinco slots de expansión, es decir, se pueden conectar hasta cinco tarjetas adicionales, que pueden ser controladores de disco, expansiones de memoria, modems, fax, digitalizadores o cualquier otro tipo de tarjeta cuyo bus de datos sea de 8 bits.

Estos microprocesadores tienen 20 líneas en el bus de direcciones, por lo que pueden direccionar hasta 1 Mb, es decir, si la computadora cuenta con 2 Mb de memoria RAM, el microprocesador puede tener acceso únicamente al primer millón de bytes.

Debido a la velocidad de operación de estas máquinas, no es conveniente conectarlo a un monitor de alta resolución tipo VGA o SuperVGA, ya que el trazado de la imagen en pantalla resulta muy lento.

8086

Este grupo de computadoras tienen un microprocesador muy parecido al 8088 en cuanto al conjunto de instrucciones y tamaño de registro, pero el bus de datos es de 16 bits, y su velocidad máxima de operación es de 10 MHz. Pueden direccionar, al igual que el 8088, hasta 1 Mb de memoria RAM. Cuentan con ocho slots de expansión.

A este modelo de computadoras se les conoce con el nombre de "XT". Dentro de este conjunto se encuentran los primeros modelos de las PC y compatibles (es decir, aquellas no fabricadas por IBM) y los sistemas PS/2 modelos 25 y 30 de IBM.

80286

El conjunto de computadoras que contienen un microprocesador 80286 se denominan genéricamente PC AT compatibles.

La velocidad máxima de operación de estas computadoras es de 20 MHz, aunque los sistemas típicos trabajan a 12 o 16 MHz. Puede direccionar hasta 16 Mb de memoria RAM, ya que cuenta con 24 líneas para el bus de direcciones.

El bus de datos es de 16 bits, que es el máximo número de bits con que puede trabajar el sistema operativo DOS.

El conjunto de instrucciones del procesador 80286 es más poderoso que los dos anteriores, lo que lo hace aún más veloz.

Las computadoras que contienen este procesador son todas las AT, sean o no de IBM, así como los sistemas PS/2 modelos 30/286, 50, 50z y 60.

Los siguientes procesadores en aparecer vienen en dos presentaciones diferentes. Los denominados SX tienen un bus de datos de 16 bits y los DX de 32 bits. Esta es la diferencia fundamental entre los SX y DX para procesadores 386 —además de la velocidad de operación, como se menciona más adelante. En cambio, para los procesadores 486 hay una diferencia más importante —aparte del tamaño del bus de datos y velocidad de operación máxima— y es que el DX incluye un coprocesador matemático, que acelera los cálculos numéricos, mientras que el SX no lo incluye.

80386SX

El microprocesador 386SX es más poderoso que el 286 y menos que el 386 (también llamado 386DX). Surge para acelerar cierto tipo de procesos, donde no se requiere de todo el potencial de los sistemas 386 pero donde ya no es suficiente un sistema 286. Cronológicamente es posterior al microprocesador 80386.

El 386SX puede direccionar hasta 4 Gb de memoria RAM, ya que cuenta con 32 líneas para el bus de direcciones. Tiene 16 bits para el bus de datos y el tamaño de registro es de 32 bits. La velocidad máxima de operación de estos microprocesadores es de 25 MHz.

En la actualidad existe un gran número de marcas de computadoras que manejan este microprocesador, entre las que se encuentran **Swan**, **Gama** y **Compaq** y las IBM modelo 55SX. Estas últimas tienen una tecnología diferente, ya que, en lugar del bus de datos "normal", llamado ISA, utilizan la tecnología "microcanal", exclusiva de IBM. El microcanal no es compatible con los buses ISA, por lo que las tarjetas que se desce agregar a una computadora con una tecnología dada deben ser compatibles con el resto de la máquina; por ejemplo, no se puede añadir una tarjeta ISA para red a una computadora con microcanal o viceversa. Actualmente, en EUA, ya se discontinuó la fabricación de "motherboards" o tarjetas principales que usan este procesador.

80386

Hasta hace un par de años, era el microprocesador más poderoso y rápido de los que se encontraban disponibles en el mercado nacional. Su velocidad máxima de operación es de 33 MHz, aunque los sistemas típicos operan a 20 y 25 MHz. Los buses de datos y direcciones son de 32 bits cada uno, por lo que puede direccionar hasta 4 Gb de memoria, y su tamaño de registro es de 32 bits.

De toda la familia de microprocesadores de la serie 80***, éste es el primero capaz de llevar a cabo multiprocesos (**multitasking**), dividiendo toda la memoria disponible en bloques de 640 kb, en cada uno de los cuales puede ejecutarse un programa diferente.

Dentro de este grupo se encuentran los sistemas PS/2 modelos 70 y 80, que usan una tecnología nueva para los buses, llamada "microcanal"; y otras marcas en el mercado, pero que no cuentan con esta tecnología, ya que está patentada por IBM.

Actualmente, todos los programas nuevos se están desarrollando para los sistemas 386 o superiores, por lo que, en poco tiempo, todos los modelos anteriores serán obsoletos. Claro que, al paso que avanzan actualmente los desarrollos tecnológicos, en unos cuantos años (unos cinco a lo más), también estos sistemas serán obsoletos. De hecho, ya se está llevando a cabo toda una campaña publicitaria, por parte de los fabricantes de componentes electrónicas, para los nuevos microprocesadores que operarán a velocidades de 100 MHz, aunque éstos se encuentran aún en fase experimental.

A partir de este modelo de microprocesador, se pueden tener dos sistemas operativos simultáneos (**dual boot**) en la misma computadora, lo que permite que actúen como servidores de una red y como estaciones de trabajo simultáneamente.

80486

Son los procesadores más modernos a precios accesibles actualmente. Los buses de datos y direcciones son de 32 bits cada uno, por lo que puede direccionar hasta 4 Gb de memoria, y su tamaño de registro es de 32 bits.

Se usan básicamente como servidores para redes, que son computadoras tipo PC —aunque también se pueden usar máquinas tipo *mainframe* para esta función— encargadas de proveer de servicios especiales, como puede ser el

almacenar grandes volúmenes de información de modo que cada una de las estaciones de trabajo (computadoras personales) conectadas a la red pueda tener acceso a ella, o como "servidor de comunicaciones", cuya función es el permitir la comunicación entre la red local y otras redes o servicios externos; o bien como estaciones de trabajo. Su velocidad máxima de operación es de 66 MHz, aunque los más comunes operan a 33 MHz y también los hay que funcionan a 25 y 50 MHz. Existen dos versiones de este procesador, igual que del 386: 486SX y 486DX. Las dos diferencias fundamentales entre ellos son el número de líneas de bus (16 y 32 respectivamente) y que el último incluye el coprocesador numérico. De hecho, cuando se construyen los circuitos integrados y se evalúan, si alguna parte del circuito coprocesador no funciona adecuadamente, se cortan las líneas de conexión de éste y se etiqueta como 486SX. Si pasa adecuadamente por el control de calidad, sale a la venta como 486DX.

Es importante notar que los procesadores 486DX que operan a 66 MHz, para ser eficientes, requieren de una memoria caché de procesador "alta" (128 kb es bueno, y 256 kb es casi ideal). Esto se debe a que, aunque el procesador trabaje a velocidades tan vertiginosas, el resto de la electrónica no lo hace, sino que opera a 33 MHz (de hecho, se usa el mismo tipo de tarjeta principal). Las razones son: costo —resulta mucho más complejo el diseñar tarjetas principales mientras mayor sea la velocidad de funcionamiento que se requiera; y tiempo —en la actualidad, apenas y sale al mercado un nuevo procesador y, en cuestión de semanas, aparecen las primeras computadoras capaces de usarlos y los precios de modelos anteriores baja sustancialmente. Los fabricantes de computadoras no pueden "darse el lujo" de mejorar el resto de la electrónica que compone una máquina, además de los costos que eso representa en cuanto a diseño.

Para hacer más rápido el funcionamiento de las computadoras que contienen estos procesadores, se incluyen dos memorias caché: una en el procesador (generalmente de 8 kb, expandible a 256 kb) y otra para el disco duro. Cabe hacer notar que también los periféricos que se tienen en una de estas máquinas trabajan a mayores velocidades, como son las memorias, discos duros, unidades de respaldo, etc.

Hace poco tiempo se empezó a hablar de "tecnología escalable", que consiste en reemplazar únicamente el procesador anterior por uno más reciente, si se desea actualizar el sistema con que se cuenta. En términos generales no resulta así, ya que si se desea actualizar de un procesador SX a un DX, "faltan" 16 líneas del bus de datos. En cambio, cuando se cuenta con un sistema 486DX a 33 MHz, y si la tarjeta principal cuenta con el espacio (llamado socket) para conectar el procesador a 66 MHz (overdrive), se puede hacer el cambio. Un

ejemplo de lo anterior son las computadoras Compaq: sus modelos Prolinea no admiten el procesador a 66 MHz, ya que no cuentan con el socket para el overdrive, mientras que el modelo DeskPro sí lo tiene. Es conveniente verificar todo este tipo de detalles antes de elegir no sólo la marca, sino el modelo de computadora que se va a adquirir.

CONSIDERACIONES GENERALES

En términos generales, cada nuevo microprocesador es más poderoso que los anteriores, y todos son compatibles "hacia atrás", es decir, si un programa se desarrolla para que funcione con un microprocesador 8088, también se puede usar en cualquiera de los demás. Lo contrario no es cierto, ya que un programa desarrollado para un sistema 286 no funcionará adecuadamente en los modelos anteriores, ya que hace uso del conjunto de instrucciones de este procesador y algunas de éstas no las reconocen el 8088 ni el 8086.

Todas las especificaciones anteriores de los sistemas son fundamentales para determinar la velocidad de respuesta de la computadora, aunque el usuario medio no alcanza a percatarse de las diferencias en tiempos debido a que usan generalmente programas en que el usuario introduce constantemente información a la computadora via el teclado, como procesadores de texto, hojas electrónicas y bases de datos. Sin embargo, cuando se desarrolla un programa en un lenguaje de alto nivel y se compila, el tiempo de procesamiento —es decir, el tiempo que ocupa la computadora para generar el programa ejecutable a partir del programa fuente (el escrito por el programador)— puede resultar enorme en un sistema lento.

Otro punto importante es que, independientemente del sistema que se tenga, hay factores que determinan su funcionamiento global. Una computadora AT puede ser más veloz para ciertas tareas que un sistema 386; esto depende de la velocidad con que se manejan los buses y los estados de espera del sistema¹.

En la mayoría de las computadoras, los buses de expansión trabajan a velocidad diferente que la tarjeta principal (que es la tarjeta donde se encuentra

1 Los estados de espera son ciclos de reloj "desperdiciados". Surgen debido a que, en algunos casos, el microprocesador es mucho más veloz que el resto de la electrónica, por lo que en procesos que involucran sólo al procesador, éste realiza el proceso sin interrupción, pero cuando se requiere llamar a componentes más lentas, se requiere dejar pasar ciclos de reloj con el procesador inactivo. A estos ciclos de reloj se les denomina estados de espera. Se eliminan al usar memoria cache, ya que ésta permite que el procesador no esté ocioso, en espera de un dato o instrucción.

el microprocesador). Uno de los motivos es que el bus no está diseñado para trabajar a altas velocidades o bien los periféricos no requieren velocidades tan altas —como cuando se está trabajando con modems— o bien las tarjetas ya existentes no pueden estarse actualizando cada que se desarrolla la tecnología necesaria para hacer un CPU más veloz —todas las tarjetas adicionales con que cuenta la computadora se tendrían que cambiar al actualizar la computadora, lo que resultaría en un costo excesivo.

Los estados de espera se insertan cuando la memoria es más lenta que el microprocesador. Si se requiere obtener un dato de la memoria y ésta no lo puede tener a la velocidad que demanda el CPU, se inserta un pulso de reloj en el cual el CPU permanece ocioso, en lo que le llega la información para poder llevar a cabo la instrucción solicitada. Mientras menos estados de espera hayan, más rápido es el sistema, ya que es menor el tiempo en que está inactivo el microprocesador.

Otro factor que determina la velocidad de operación del sistema en conjunto es el número de bits que se pueden “desplazar” simultáneamente entre la memoria y el CPU. Si el bus de datos es pequeño (8 bits) se puede transferir sólo parte de la información por vez, en tanto que si es de 4 bytes (32 bits), pueden pasar hasta dos datos por vez. Si existen adicionalmente tiempos de espera, en un bus de 8 bits el tiempo muerto antes de tener todo el dato y poder ejecutarse la instrucción es mucho mayor que si se tiene un bus de 32 bits, ya que se hacen menos referencias a la memoria.

Cabe hacer la aclaración de que están apareciendo en el mercado nuevos microprocesadores, como el llamado Pentium —siguiente modelo después del 486. La velocidad de operación de éste es de 66 MHz, pero se espera que próximamente aparezcan procesadores a 100 MHz. Las cuatro modificaciones más importantes entre los procesadores anteriores y el Pentium son las siguientes:

- » El bus externo² de datos se duplica, es decir, ahora es de 64 bits, mientras que en los sistemas anteriores era de 32 bits —para los modelos DX— y de 16 bits para los SX. El bus interno se mantiene en 32 bits, igual que en el 486. Lo que se logra es el duplicar la velocidad de transmisión de datos entre el procesador y la memoria RAM.

2 El bus externo es aquel que permite la transmisión de información de y hacia la memoria (bus de datos) y el bus interno es el número de bits que maneja internamente el procesador.

- » La segunda innovación del Pentium es su "pipeline" doble. Un pipeline es, básicamente, una línea que acelera el procesado de información a través de cinco etapas: *prefetch* (toma de datos), decodificación de la instrucción, generación de dirección de memoria, ejecución y escritura de del resultado. El 486 ya incluye un pipeline, lo que le permite procesar una instrucción en cada ciclo de reloj. De manera análoga, el Pentium puede procesar dos instrucciones por cada ciclo de reloj, lo que lo hace dos veces más eficiente que el sistema 486DX a 66 MHz. Adicionalmente a los dos pipelines, se implementó una "unidad predeterminadora de salto" (*branch predictor unit*), que garantiza que no va a haber ningún problema con la ejecución de bucles (*loops*). Esta unidad determina inmediatamente la dirección que va a tomar el programa en ejecución.
- » Tercero, se requiere de una memoria caché interna, que le alimente datos constantemente, y a la cual escribir los resultados, sin tener que perder ciclos de reloj en activar las direcciones de memoria necesarias, enviar la información, etc. El caché interno de un 486 es de 8 kb, mientras que el Pentium contiene dos memorias caché de 8 kb cada una, una para datos y la otra para instrucciones. Cada caché se puede comunicar con ambos pipelines a la vez, en un solo ciclo de reloj.
- » Cuarto, se rediseñó la FPU o unidad de punto flotante, que es el elemento del procesador dedicado a cálculos con números no enteros, que se usa, por ejemplo, para graficar posiciones de puntos individuales en diseños gráficos, generar regresiones en Excel, etc. Con este nuevo diseño, la FPU es entre cinco y diez veces más rápida que la del 486.

Dentro de los próximos dos años se espera que el Pentium sea el procesador por excelencia para los servidores de red.

Más adelante, se prevé que una máquina basada en Pentium permita aceptar un segundo procesador en el socket de overdrive, sin deshabilitar el primero, por lo que se tendrán dos procesadores paralelos, es decir, trabajando simultáneamente.

A continuación se presenta una tabla que resume la evolución de la familia de procesadores 80***, utilizados en las computadoras tipo PC, a lo largo del tiempo:

| <i>Procesador</i> | <i>Fecha de aparición</i> | <i>Precio inicial (USD)</i> | <i>Precio actual (USD)</i> | <i>MIPs*</i> | <i>MIPs**</i> |
|-------------------|---------------------------|-----------------------------|----------------------------|--------------|---------------|
| 8086 | Junio 78 | 360 | N.D. | 0.33 | 0.75 |
| 80286 | Febrero 82 | 360 | 8 | 1.20 | 2.66 |
| 80386 | Octubre 85 | 299 | 91 | 5.00 | 11.40 |
| 80486 | Agosto 89 | 950 | 317 | 20.00 | 54.00 |
| Pentium | Marzo 93 | 900 | | 112.00 | |

Por todo lo anterior, si es importante aclarar que, si no se actualiza la computadora con que se cuenta frecuentemente, se corre el riesgo de que los programas nuevos no puedan ser usados. La desventaja de esto es que cada vez son más poderosos los programas y pueden ahorrar gran cantidad de tiempo al usuario. Finalmente, la computadora se hizo para simplificar las cosas, no para complicarle a uno más la vida, y vale más la pena cambiar cada dos o tres años de máquina que el tiempo que se invierte desarrollando uno mismo los programas que requiere, ya que los que existen en el mercado no sirven para las computadoras viejas, o adecuando los programas existentes para que puedan aprovechar las capacidades de los procesadores más recientes.

* Millones de instrucciones por segundo, cuando el procesador aparece en el mercado.

** Millones de instrucciones por segundo actualmente.

N.D. No disponible, descontinuado.

SEGUNDA PARTE: LAS COMPUTADORAS

EN LA FÍSICA DE SUPERFICIES

INTRODUCCIÓN

Conforme ha ido evolucionando la tecnología, se han desarrollado computadoras cada vez más poderosas y veloces, de dimensiones menores, y con reducciones significativas en los costos, sobre todo en las máquinas que están un paso atrás de la tecnología de punta. Hace unas décadas, solamente se tenía acceso a estas máquinas en los centros de investigación en ciencia básica y de aplicaciones militares. Sin embargo, conforme se reducen los costos, se difunde más su uso, hasta que, hoy en día, es una herramienta cotidiana en oficinas, industria e incluso en las escuelas.

La primera área en la que entró la computadora como herramienta básica es en la física, e incluso ésta ha permitido el desarrollo de gran cantidad de algoritmos y métodos numéricos para la solución de ecuaciones diferenciales, integrales, diagonalización de matrices, etc., que han sido aprovechadas en otros campos del conocimiento, sobre todo la estadística aplicada.

La manera de trabajar en física teórica con computadoras es la siguiente: se cuenta con modelos matemáticos complejos, desarrollados por el investigador, y se requiere conocer los resultados de aplicar el modelo, para poder analizarlos y ver si el modelo reproduce adecuadamente el fenómeno que se está estudiando. Más aún, haciendo uso del modelo, se pueden predecir resultados no observados previamente, para después diseñar un experimento y poder evaluar si las predicciones fueron correctas o no. En este caso, la computadora ha resultado ser invaluable, ya que, una vez que se ha desarrollado el programa correspondiente, se puede ejecutar modificando parámetros, de modo que se tiene un buen control sobre las variables de las cuales depende el modelo.

En la física experimental, además de la necesidad de desarrollar modelos teóricos que puedan explicar los resultados observados en un experimento, se requiere de computadoras capaces de controlar las variables independientes que se consideran

importantes en el experimento en tiempos relativamente cortos comparados con los intervalos de medición de las variables dependientes, de modo que el sistema se mantenga estable. En este caso, la computadora tiene dos funciones simultáneas: la primera, sensar las variables en tiempo real, es decir, se debe contar con sensores adecuados, una interface para comunicarse con la computadora y el *software* adecuado para poder ir almacenando la información proveniente de cada sensor, además de que la máquina debe estar lista siempre para recibir los datos provenientes de los sensores. Segundo, tener acceso a los actuadores necesarios para poder mantener las variables en los valores requeridos. Veamos un ejemplo:

EMULACIÓN DE MULTICANAL

En el área de física de superficies, parte de los análisis consisten en introducir una muestra dentro de un acelerador de partículas, ser bombardeada por un haz y detectar los productos del bombardeo. Dependiendo de la energía de las partículas incidentes (y del tipo de partículas, que pueden ser electrones o iones positivos), se producen diferentes fenómenos físicos. En el caso de radiar la muestra con iones positivos, los fenómenos físicos que se pueden llevar a cabo son: dispersión de Rutherford (abreviado RBS, que significa *Rutherford Back Scattering*), cuando el ion incidente golpea a un átomo de la muestra, rebota y sale del blanco; ERDA (*Energy Recoil Dispersive Analysis*) [1], cuando el ion incidente golpea a un átomo de la muestra, tomando su lugar en la estructura y liberando al ion del blanco, que sale y es detectado; reacción nuclear; channeling, que es un fenómeno físico que ocurre en muestras cristalinas y consiste en que el ion incidente tenga una trayectoria tal que es canalizado por la estructura cristalina; surge por la interacción elástica con los potenciales nucleares, que lo repelen del arreglo; o PIXE, que consiste en la emisión de rayos X debido a que el ion crea un hoyo en alguna capa interna del átomo de la muestra y, cuando un electrón cubre esa vacancia, emite los rayos X, entre otros.

En todos estos casos, las partículas o radiación que sale del blanco es detectada por un dispositivo físico, llamado detector, que traduce la información que le llega a impulsos eléctricos. Cada detector es capaz de procesar diferentes tipos de fenómenos. Por ejemplo, en los casos de RBS o ERDA, el detector más común es el de "barrera superficial", que describiremos en el siguiente capítulo.

Una vez que el detector recibe una partícula y la convierte en una señal eléctrica, ésta es amplificada y después pasa a un multicanal, que es un dispositivo electrónico capaz de discriminar los niveles de voltaje que recibe e ir contando estas señales en diferentes canales, cada uno de los cuales tiene un

intervalo de energía asociado; la información se despliega en una pantalla y se cuenta con un cursor y una perilla, para poder desplazarse por la gráfica que se formó en la pantalla. además de una zona de despliegue en la pantalla donde se muestra la altura de la gráfica en el punto en que se encuentra colocado el cursor. Para poder analizar los datos que se despliegan en el multicanal, es necesario primero irlos anotando en papel, para después poder trabajar con ellos. Actualmente, con el advenimiento de la computadora de arquitectura abierta, se han desarrollado una serie de tarjetas capaces de emular multicanales, con lo cual, una vez conectada la tarjeta a la computadora y al amplificador, es la computadora misma la que se encarga de ir almacenando los datos y desplegándolos en pantalla. De esta manera, el investigador cuenta con un archivo que contiene los resultados de su experimento, para poder analizarlos, graficarlos, etc., lo que puede hacer con métodos tradicionales o usando programas computacionales —algunos comerciales y otros desarrollados también por investigadores, de uso específico— que usan aproximaciones numéricas para el análisis de resultados.

Esto es, en términos generales, un ejemplo del uso de las computadoras en la física, usada como aparato de análisis. Ahora veremos un caso particular: la retrodispersión de Rutherford (RBS), donde abundaremos sobre algunos de los puntos tratados en este capítulo.

Al final de esta sección se presenta un capítulo que ilustra la tecnología planar para fabricación de circuitos semiconductores. Ahí se bosqueja la manera de fabricar circuitos integrados, tomando uno de los ejemplos más sencillos: un transistor de efecto de campo (MOSFET). También se presentan algunas fotografías de circuitos.

DETECTORES

Como se mencionó en el capítulo anterior, los detectores son dispositivos capaces de registrar los fenómenos físicos y producir una señal eléctrica como respuesta. Estas señales son proporcionales a la energía de la partícula incidente. Trataremos únicamente dos tipos de detector, que son: el de "barrera superficial", para iones positivos; y el contador de centelleo, para detectar radiación gama.

DETECTOR DE BARRERA SUPERFICIAL

El detector de barrera superficial es una juntura $n-p$ en polarización inversa.

Para entender como opera, hagamos un paréntesis para tratar sobre semiconductores.

Semiconductores Existen dos modelos para describir a los semiconductores: el modelo clásico, llamado modelo de enlaces de valencia; y el modelo cuántico, llamado modelo de bandas de energía.

Modelo de enlaces de valencia De acuerdo con el primer modelo, de enlaces de valencia, tenemos que, en un semiconductor, los electrones de valencia no están libres para moverse a través del volumen del material. En cambio, tienen enlaces covalentes, que mantienen a todos los átomos en un arreglo cristalino periódico. La forma general del arreglo se presenta en la figura 8, de acuerdo con el modelo bidimensional para enlaces covalentes. Cada átomo está rodeado por cuatro vecinos equidistantes. Los

electrones de valencia (cuatro por cada átomo) se comparten con los cuatro vecinos. De este modo, cada enlace entre dos átomos contiene dos electrones.

Existen otros tipos de enlaces, además del covalente, que son:

- Enlace iónico, que ocurre cuando un átomo con valencia entre uno y tres cede sus electrones de capa abierta a otro u otros átomos, de modo que en este último se completa la última capa y en el primero se pierden los electrones de capa abierta. La molécula se forma por la atracción eléctrica entre los iones.
- Enlace metálico, donde cada átomo cede al electrón o electrones de la última capa abierta a la región entre los átomos vecinos. Estos electrones quedan libres para moverse por el material, siendo los portadores de la corriente eléctrica cuando se aplica un campo eléctrico.

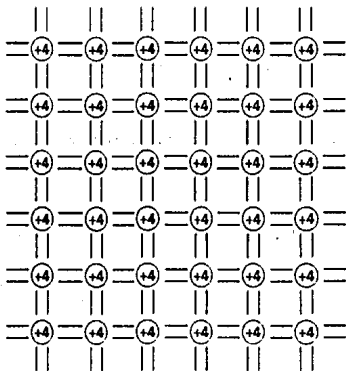


Fig. 8. Modelo de un semiconductor

Cuando todos los electrones de valencia están en enlaces covalentes, no es posible la conducción eléctrica, ya que no hay portadores de carga libres. En cambio, en un metal existen grandes concentraciones de electrones "libres"; son libres en el sentido de que no están ligados a los átomos del metal, por lo que se pueden desplazar por todo el volumen del mismo. En muchos metales, se puede considerar a estos electrones libres como un "mar de electrones"¹ que permea el volumen del metal. Usualmente, un átomo contribuye con sólo uno o dos electrones (electrones de valencia) a este

1 Cabe notar que la definición formal del mar de electrones incluye a todos los electrones.

mar. Los demás están restringidos a permanecer asociados con un núcleo particular. El metal es eléctricamente neutro porque la carga negativa de los electrones libres está exactamente balanceado por la carga positiva neta asociada con los núcleos.

La distribución real de electrones de valencia en un semiconductor difieren de la figura anterior en que, a toda temperatura superior al cero absoluto, unos cuantos enlaces covalentes están incompletos. Los electrones faltantes no están confinados a la región de enlace, sino que están libres, como se muestra en la figura 9:

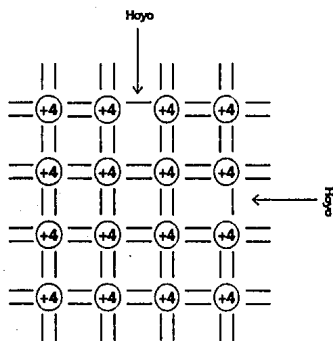


Fig. 9. Enlaces covalentes rotos

Los enlaces rotos a temperaturas normales (alrededor de 300°K) se deben a la vibración térmica de los electrones de valencia. Unos cuantos de ellos adquieren la energía suficiente para romper el enlace. La fracción de electrones de valencia libres es muy pequeña. Por ejemplo, para el germanio a temperatura ambiente, existen alrededor de 10^{13} enlaces rotos por centímetro cúbico; como hay alrededor de 10^{23} átomos por centímetro cúbico, sólo un átomo de cada diez mil millones carece del enlace. Estos enlaces faltantes, con sus electrones libres asociados, son los que hacen posible la conducción eléctrica en un semiconductor. En el caso de un aislante, existen únicamente un enlace roto por cada 10^{15} átomos y en un conductor hay más de 10^{23} electrones libres por centímetro cúbico.

Como consecuencia de estos enlaces covalentes rotos, existen dos grupos de portadores de carga que pueden permitir la conducción de corrientes eléctricas en el semiconductor. El primer grupo son los electrones libres, llamados electrones de conducción, que portan una carga $-e$. Los otros portadores de carga están asociados con los electrones de valencia que permanecen atados en enlaces covalentes. Un enlace roto se asocia con una región de carga positiva; en la vecindad del enlace roto, existe un exceso de carga positiva, con un valor de $+e$. A esta región se le denomina hoyo, ya que proviene de una vacancia en la estructura de enlaces. El movimiento de hoyos se lleva a cabo cuando un electrón de valencia en un enlace cercano llena la vacancia, ocasionando que el hoyo se desplace en sentido contrario. Se presenta un dibujo esquemático en la figura 10.

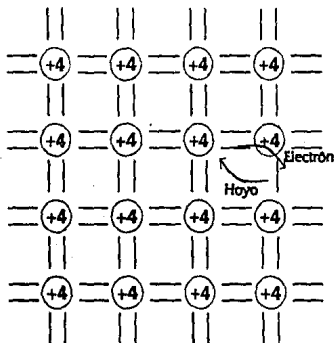


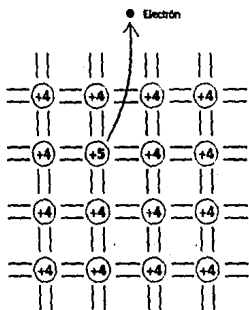
Fig. 10. Desplazamiento de un hoyo

El modelo de un semiconductor que hemos presentado requiere que exista el mismo número de hoyos y electrones de conducción. Cada uno de los portadores de carga es el resultado de un enlace covalente roto; en consecuencia, los hoyos y electrones libres se producen en pares. Si un conductor contiene una cantidad despreciable de impurezas, la concentración de electrones y hoyos son iguales. En este caso, el semiconductor se denomina intrínseco, ya que la posibilidad de conducción eléctrica es intrínseca al material.

Sin embargo, no siempre existe la misma cantidad de electrones de conducción y hoyos, debido a pequeñas impurezas, que se añaden al semiconductor al ser fabricado.

Estas impurezas o dopantes son de dos tipos:

1. *Donantes*, que tienen cinco electrones de valencia en lugar de cuatro. Estos átomos se acomodan en la red cristalina, pero como tienen cinco en vez de cuatro electrones, uno queda fuera de los enlaces, como electrón de conducción. La impureza queda en forma de ion y representa una carga positiva iónica inmóvil. En resumen, una impureza donante contribuye con un electrón de conducción pero no con hoyos, ya que la carga positiva es inmóvil. Ejemplos de los donantes más comunes son: fósforo, arsénico y antimonio. A los cristales dopados con estas impurezas se les denomina "cristal tipo *n*".



Un átomo donador tiene cinco electrones de valencia; contribuye con un electrón de conducción

Fig. 11. Cristal tipo *n*

2. *Aceptores*, que son elementos con tres electrones de valencia y que al entrar a la red cristalina y formar los enlaces covalentes, hay uno faltante, por lo que hay un hoyo o vacancia asociado con ese átomo. De manera análoga a la situación anterior, se produce un hoyo móvil y una carga negativa inmóvil, ya que no contribuye con ningún electrón para completar el enlace covalente de la estructura. Los aceptores más usados en la industria electrónica son: boro, indio y aluminio. A los cristales con impurezasceptoras se les llama "cristal tipo *p*" (figura 12).

En caso de que el semiconductor contenga una cantidad considerable de impurezas donantes y no aceptoras, tiene más electrones que hoyos y se conoce como cristal tipo *n* (negativo); si hay más hoyos que electrones, se

dice que el cristal es tipo *p* (positivo). Los materiales tipo *p* y tipo *n* son llamados conductores extrínsecos, ya que sus propiedades eléctricas se deben más a las impurezas que a las propiedades inherentes al semiconductor. A los portadores con mayor concentración (sean hoyos o electrones) se les llama portadores mayoritarios y los de menor abundancia son denominados portadores minoritarios.

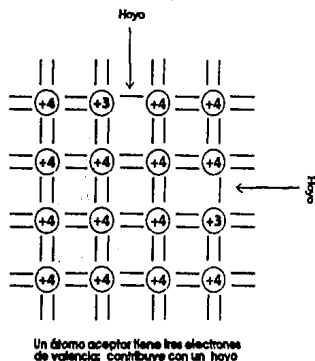


Fig. 12. Cristal tipo *p*

Modelo de bandas de energía. Las propiedades de un sólido dependen de la naturaleza de los átomos que lo forman y de la manera en que están agrupados. En un átomo, hay un núcleo con carga positiva, rodeado de electrones localizados en órbitas discretas. Sólo una cierta cantidad de electrones pueden ocupar un nivel dado. Si colocamos otro átomo cercano al anterior, la estructura de niveles se modifica, por la presencia del campo eléctrico debido al otro átomo. De esta manera, se obtienen los orbitales moleculares, a la distancia adecuada.

Conforme se van agregando más átomos, los niveles están tan juntos que se convierten en bandas, por lo que se puede ilustrar según se muestra en la figura 13, que se presenta a continuación. Por supuesto, la idea de ir "agregando átomos" es solamente una manera conceptual de ilustrar la manera en que se modifican las órbitas electrónicas permitidas.

Una manera de entender, desde el punto de vista conceptual, la forma en que se distribuyen los electrones en el material es mediante el modelo

unidimensional de bandas de energía, que se muestra en la figura 13. En la parte derecha de este diagrama se presenta la gráfica de distribuciones de densidad de la nube electrónica como función de la energía de manera esquemática.

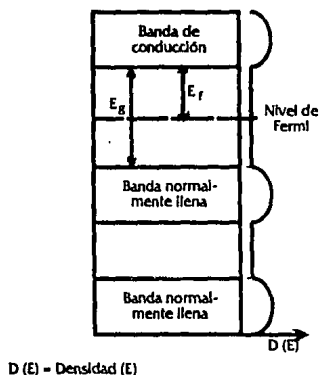


Fig. 13. Modelo de bandas de energía

Del lado derecho del diagrama se gráfica la densidad de electrones. Como se observa, sólo hay electrones presentes en las bandas de valencia. Si el material se encuentra a una temperatura de 0°K , las bandas de valencia estarán pobladas y la de conducción no. Al aumentar la temperatura, por vibraciones térmicas, algunos electrones adquirirán la energía suficiente para saltar de la banda de valencia más alta a la de conducción. Se define el nivel de Fermi como el nivel de energía más alto que está lleno a una temperatura de cero absoluto. En la figura 14 se muestra el modelo de bandas para metales, semiconductores y aislantes.

Como se observa, para un metal (conductor) las bandas de valencia y conducción están unidas, por lo que se requiere una energía casi nula para que puedan saltar los electrones de la banda de valencia a la de conducción; a temperaturas superiores al cero absoluto, siempre habrá electrones en la banda de conducción. En el otro extremo tenemos los aislantes, que requieren de una energía relativamente alta para que un electrón pueda saltar de la banda de valencia a la de conducción. Por tanto, a temperatura ambiente, la banda de conducción no estará poblada. El caso

intermedio son los semiconductores (de ahí su nombre), que requieren de una energía mayor que los conductores pero menor que los aislantes para que los electrones puedan saltar de la banda de valencia a la de conducción (del orden de unos cuantos eV).

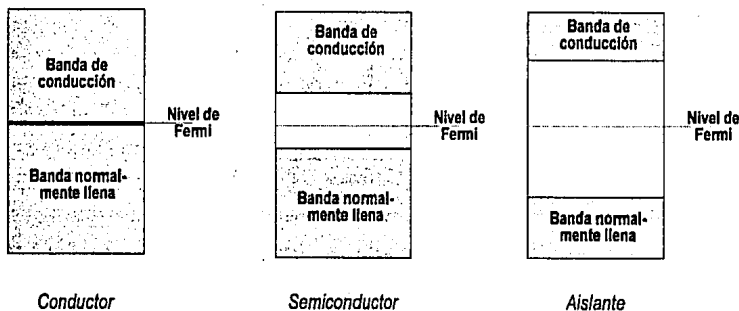


Fig. 14. Modelo de bandas para conductores, semiconductores y aislantes

Si el semiconductor es un cristal puro, el número de hoyos es igual al número de electrones libres ya que, por cada electrón que sube a la banda de conducción, se crea un hoyo en la banda de valencia (por conservación de carga); a éstos se les llama acarreadores intrínsecos.

Sin embargo, casi todos los semiconductores que se usan en dispositivos electrónicos tienen una cierta cantidad de impurezas capaces de donar electrones a la banda de conducción (llamados cristal tipo *n*) o aceptar electrones de la banda de valencia, creando hoyos en ésta (llamados cristal tipo *p*). Al igual que en el modelo de enlaces de valencia, a estas impurezas se les llama portadores (o acarreadores) extrínsecos.

Las impurezas se concentran en un nivel de energía único, cercano a la banda de conducción, pero más abajo (donantes) o muy cerca de la banda de valencia (aceptadores). Igual que para los acarreadores intrínsecos, el nivel de Fermi para los acarreadores de las impurezas está entre la banda de conducción y el nivel de la impureza.

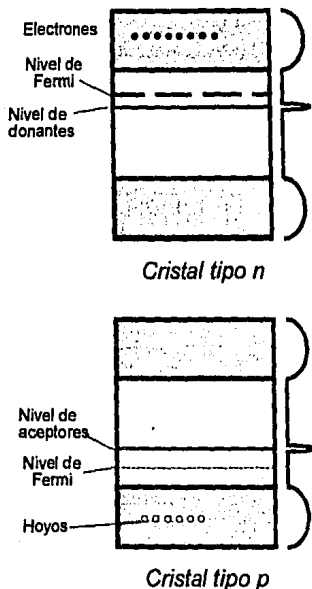


Fig. 15. Modelo unidimensional de bandas para cristales tipo n y p

Junturas Veamos ahora lo que ocurre si se unen dos cristales, uno tipo n y el otro tipo p . A esta unión se le denomina *juntura*, y es la base de todos los dispositivos electrónicos que existen en la actualidad, incluyendo las componentes internas de una computadora, como puede ser la memoria o el microprocesador.

De acuerdo con el modelo de enlaces de valencia, tendremos que algunos de los electrones libres del cristal n , en la vecindad de la unión, tenderán a migrar hacia el cristal tipo p , para completar los enlaces covalentes faltantes en este último cristal, por lo que se formarán dos regiones no neutras eléctricamente, una positiva en el cristal n y una negativa en el p , de acuerdo con la siguiente figura.

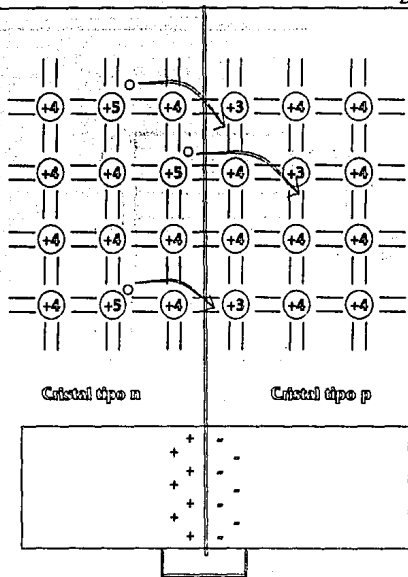


Fig. 16. Juntura n-p

A la región que no es eléctricamente neutra localmente se le denomina "zona de depleción", y se origina únicamente por el hecho de unir los dos cristales; el ancho de la zona de depleción depende de la cantidad de impurezas en el cristal. En equilibrio, el campo eléctrico se opone a la difusión de los portadores. De hecho, al alcanzarse el equilibrio van a existir cuatro corrientes: dos debidas a la energía térmica de los electrones del cristal n al p , hoyos del p al n , compensados por rupturas — también por causas térmicas — en los enlaces débiles en la zona de depleción.

De acuerdo con el modelo de bandas de energía, tenemos que, cuando se unen dos materiales con diferente estructura de bandas, el nivel de Fermi debe estar en la misma posición (energía), si no hay un campo eléctrico externo. Solamente los electrones con $E_e > dW_e$ podrán cruzar la unión del material n al p , y hoyos con $E_n > dW_n$ podrán pasar de la región p a la n . Los hoyos presentes en la región n y los electrones en la región p reciben el nombre de portadores minoritarios. De hecho, va a existir una difusión de portadores minoritarios a través de la unión pero, como no hay un campo eléctrico externo presente, éstos se mantendrán en

la vecindad de la unión entre los cristales n y p ; a esta unión se le llama *juntura*, donde el campo separa cualquier formación de pares, ya sean de origen térmico o por ionización.

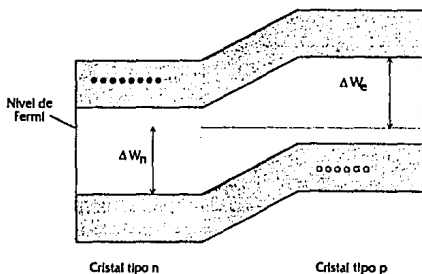


Fig. 17. Juntura n - p

Polarización inversa Si ahora se aplica una diferencia de potencial de modo que se oponga al movimiento de los portadores minoritarios (polarización inversa), de acuerdo con el modelo de bandas de energía, el nivel de Fermi se desplazará proporcionalmente con la polarización, lo que hace muy poco probable el movimiento de portadores minoritarios.

Usualmente, los detectores de barrera superficial se forman con un sustrato tipo n y, sobre éste, se deposita una capa o película delgada tipo p . Posteriormente, se cubre con una capa de oro, para proteger la juntura y contacto eléctrico. Esta última película debe ser muy delgada, para no frenar significativamente las partículas antes de llegar a la zona sensible del detector (zona de depleción). La polarización del detector siempre es inversa.

Los electrones son los portadores mayoritarios a la región n y los hoyos en la p .

No hay movimiento de electrones a la región p ni de hoyos a la n , debido a las diferencias de energía: la banda de conducción en la región p está a un potencial más alto (negativo) y los hoyos no se mueven a la región n , ya

que la banda de valencia está a un potencial más alto (positivo). En consecuencia, hay repulsión de portadores mayoritarios en la juntura (zona de depleción). El ancho de la zona de depleción depende del voltaje aplicado.

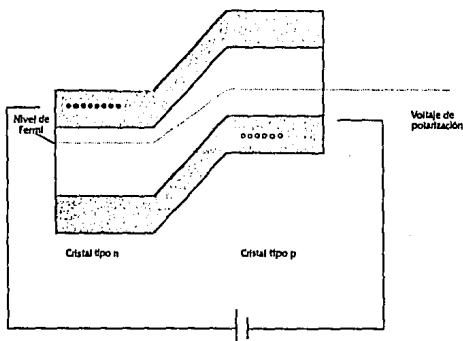


Fig. 18. Juntura n-p con polarización inversa

De acuerdo con el modelo de enlaces de valencia, al aplicar un voltaje en polarización inversa, el campo eléctrico permitirá que más electrones libres migren de la región p a la n para completar más enlaces covalentes en la estructura del cristal, por lo que la zona de depleción se hará mayor. En cambio, si la polarización es directa, el campo eléctrico ocasionará que parte de los electrones que se desplazaron cuando no había campo de la zona n a la p regresen a la región n , por lo que la zona de depleción se hará más pequeña.

Generación de pares Como se mencionó previamente, un detector de barrera superficial es una juntura $n-p$ en polarización inversa, de modo que la zona de depleción sea "grande". Al incidir una partícula cargada, se generan pares electrón-hoyo. Si éstos están en la zona de depleción, el campo eléctrico es tal que acelera la carga negativa hacia la zona n , donde tendrá gran movilidad, y el hoyo se acelerará hacia la región p . Entonces, se logra una alta eficiencia en la recolección. Este es el origen del pulso eléctrico que después pasa al amplificador, para mejorar sensibilidad, y de

ahí al multicanal, La altura del pulso es proporcional a la energía de la partícula incidente ya que, a mayor energía, más pares se generan.

El mecanismo de generación de pares hoyo-electrón es el siguiente:

Cuando incide un ion en el detector, el tipo de interacción que éste tiene con la juntura es análoga a la que ocurre en la muestra: existe frenamiento nuclear y electrónico, colisiones con los átomos del detector, etc. Con el paso del ion, se rompen algunos de los enlaces covalentes de la estructura del semiconductor. Debido al campo eléctrico presente, al ocurrir una ruptura de enlace, se genera un par electrón-hoyo que es acelerado hacia los polos del detector. A su vez, este electrón interactúa con los demás átomos del detector, generando más pares. A este fenómeno se le conoce como "efecto avalancha" y la cantidad de pares generado depende de la energía de la partícula incidente. Por ejemplo, un protón de 1 MeV creará alrededor de 300,000 pares hoyo-electrón.

Para mejorar la profundidad de la zona de depleción, se puede aumentar el voltaje de polarización, pero no indefinidamente; el voltaje de ruptura (break-down voltage) es el límite, y depende del material del detector. Si el campo eléctrico en la zona de depleción es "grande" (del orden de 5×10^6 V/cm, aunque depende de la densidad de impurezas en el semiconductor), la fuerza que ejerce sobre los electrones de valencia es suficiente como para romper el enlace, creando pares hoyo-electrón sin que exista un ion incidente que detone el proceso. Si el campo es demasiado grande, puede romperse la estructura y dañar permanentemente al detector. A este fenómeno de generación de pares por el campo eléctrico se le conoce como "efecto zenner".

Para dar una idea de las dimensiones con que se opera, tenemos que la profundidad de la zona sensible del detector (es decir, la región donde se pueden generar pares) es del orden de 10^{-2} a 1 mm.

DETECTORES DE RADIACIÓN GAMA

En caso de observar radiación, provenientes de una reacción nuclear, generalmente se usan contadores de centelleo, asociados con fotomultiplicadores. La idea básica del funcionamiento de un contador de centelleo es que, al incidir radiación en un gas o un cristal, los electrones se excitarán e incluso se podrá llegar a ionizar el átomo. Luego, los electrones tienden a regresar al estado base, y emiten fotones con una energía igual a la que

absorbieron al excitarse. En algunos materiales, los fotones emitidos caen en el rango del espectro visible. Estos centelleos luminosos son sensados en un fotomultiplicador, cuyo principio de operación es el efecto fotoeléctrico: la mayor parte de los metales, al ser radiados con luz, emiten electrones. Los electrones son acelerados en varias etapas, hasta llegar a un colector (cátodo), del cual sale la señal hacia un amplificador y de ahí a un multicanal.

TEORÍA: RETRODISPERSIÓN DE RUTHERFORD

En esta parte, describiremos la manera en que se han aprovechado estas máquinas en un área específica de la física: el análisis de superficies mediante la técnica denominada retrodispersión de Rutherford (RBS).

Esta técnica se ha usado desde hace ya varias décadas para:

- Medir la penetración de iones pesados en sólidos
- Determinar perfiles de concentración de impurezas superficiales
- Medir grosores de películas delgadas
- Análisis de interdifusión

La técnica RBS lleva el nombre de Lord Rutherford que, junto con Marsden y Geiger, fue el primero en usarla: encontró que el modelo del átomo era diferente de lo que se creía en ese entonces (modelo de Thompson). De los resultados de su experimento, concluyó que el átomo estaba compuesto de un núcleo pesado rodeado de electrones, con masa mucho menor que la de los núcleos. El primero en reportar el uso de esta técnica fue S. Rubin, en 1950 [2]. A partir de entonces, se ha hecho cada vez más popular —y otras que involucran la interacción de iones con materia, como reacciones nucleares o PIXE (Proton Induced X-Ray Emission). Su difusión como técnica de análisis se debe, básicamente, a:

- La gran variedad de información que se puede obtener, como composición y estructura atómica.

- Es un análisis no destructivo, ya que la muestra se bombardea a energías bajas (entre 750 keV y 2 MeV) y no hay deterioro del material ni alteraciones en su composición.
- El creciente interés por analizar las propiedades superficiales.
- La creciente disponibilidad de aceleradores diseñados específicamente para este tipo de análisis.

RBS es la técnica más sencilla de aplicar entre las que involucran interacción de iones con materia, debido a que los fenómenos físicos involucrados no incluyen términos de mecánica cuántica o relativista; dentro de esta familia de técnicas que aprovechan la interacción de iones con materia, además de los fenómenos comunes a todos (frenamientos nucleares y electrónicos), conceptualmente, el fenómeno más sencillo a considerar es el choque elástico entre dos cuerpos. Se ha convertido en herramienta rutinaria para la caracterización de semiconductores y otras películas delgadas, tan importantes en el desarrollo de la tecnología moderna.

Veamos en términos generales en qué consiste el método.

TÉCNICA DE RBS PARA ANÁLISIS DE SUPERFICIES

Se hace incidir un haz de iones sobre un blanco. Se presentan varios tipos de interacciones entre los iones incidentes y los átomos del blanco:

- Choques elásticos entre el ion y los núcleos atómicos del blanco.
- Choques inelásticos entre el ion y los electrones de la muestra.
- Excitación nuclear y reacciones nucleares, que son muy poco probables si el ion incidente tiene una energía de unos pocos MeV — generalmente se usan aceleradores de 600 keV a 6 MeV.
- Bremsstrahlung, que es proporcional a $Z_1^2 Z_2^2 e^4 / m_1^2$ — donde Z_1 se refiere a la carga del ion incidente y Z_2 a la del blanco, e es la carga unitaria y m indica la masa del ion incidente — por lo que es despreciable salvo para partículas cargadas muy ligeras, i.e. electrones.
- Radiación Cerenkov. Para que se presente, se requieren proyectiles muy energéticos (arriba de 100 MeV).

Los dos primeros son los mecanismos más importantes de pérdida de energía al interactuar el ion con la muestra, y contribuyen aditivamente al

poder de frenamiento. Algunos de los iones chocan elásticamente contra los átomos superficiales y los demás penetran en la muestra. Conforme se van adentrando en ella, van perdiendo energía (pérdida de energía por frenamiento nuclear y electrónico). En el trayecto, algunos iones van chocando con los átomos del blanco, y son deflectados. Muy pocos de ellos son retrodispersados. Estos últimos, aparte de perder energía en el choque elástico, durante el trayecto de salida siguen perdiendo energía, hasta que salen de la muestra y llegan al detector. Veamos con detalle los fenómenos físicos.

Factor cinemático Primero, al chocar el ion elásticamente con un átomo del blanco, que supondremos en reposo, para fines del análisis, hay conservación de momento y de energía. Dependiendo de la masa del átomo de la muestra con el cual choque el ion, la transmisión de momento (y energía) será mayor o menor. Se puede definir un "factor cinemático" (K) como el cociente de la energía final del ion entre la energía inicial. Este cociente depende de las dos masas y el ángulo de salida respecto a la trayectoria de entrada —llamado ángulo de dispersión—, de acuerdo con la figura 21, que se presenta en la sección sobre la geometría del experimento.

El factor cinemático está dado por la siguiente ecuación, donde

$$E_1 = K E_0$$

$$K = \left\{ \frac{m_1 \cos(\theta) + [m_2^2 - m_1^2 \sin^2(\theta)]^{\frac{1}{2}}}{m_1 + m_2} \right\}^2$$

con m_1 la masa del ion incidente y m_2 es la masa del átomo del blanco con el cual choca y θ el ángulo entre las trayectorias de entrada y salida del haz. A continuación se presenta un programa en BASIC para la obtención de factores cinemáticos para retrodispersión de Rutherford. Este programa obtiene los datos de cargas y masas atómicas de una base de datos, llamada BASE1, cuyo contenido es muy similar a los datos que se presentan en el programa de Rickards que se lista a continuación. Incluye el nombre del elemento, carga, masa y otros términos adicionales que se usan para el cálculo de frenamiento electrónico.

0 ' Factor cinemático para retrodispersión de Rutherford

10 ' obtiene el factor cinemático para los diferentes


```

elementos del
20 ' blanco y los datos los obtiene de la base de
   datos BASE1
30 PI=4*ATN(1) : CLS
40 OPEN "r", #1, "base1", 13
50 FIELD #1, 8 AS MASS$, 2 AS CHARGE$, 2 AS ELEM$, 1
   AS ISOTOPE$
60 INPUT "Z del proyectil";I
70 CODE%=I
80 GET #1, CODE%
90 QP%=CVI(CHARGE$)
100 MASAP#=CVD(MASS$)
110 ELEMENTP$=ELEM$
120 INPUT "Cuántos elementos hay en el blanco ";K
130 DIM MASAT#(K)
140 FOR J=1 TO K
150 PRINT "Z de los posibles elementos del blanco ";J;
   : INPUT I
160 CODE%=I
170 GET #1, CODE%
180 QT%(J)=CVI(CHARGE$)
190 MASAT#(J)=CVD(MASS$)
200 ELEMENTT$(J)=ELEM$
210 NEXT
220 ' obtiene el factor cinematico
230 INPUT "angulo=" ;ANGLE
240 ANGLE=(ANGLE/180)*PI
250 FOR J=1 TO K
260 KF(J)={(MASAP# * COS(ANGLE) + SQR(MASAT#(J) *
   MASAT#(J) - (MASAP# * SIN(ANGLE))^2)) / (MASAT#(J)
   + MASAP#)^2}
270 PRINT "el factor cinematico del elemento  $\gamma$ ;
   ELEMENTT$(J);" es ";KF(J)
280 NEXT

```

La manera en que varía K , dependiendo de la masa de los átomos del blanco y del ángulo de dispersión se presentan en las siguientes gráficas, donde se consideró como proyectil un ion de hidrógeno ($m = 1$ amu).

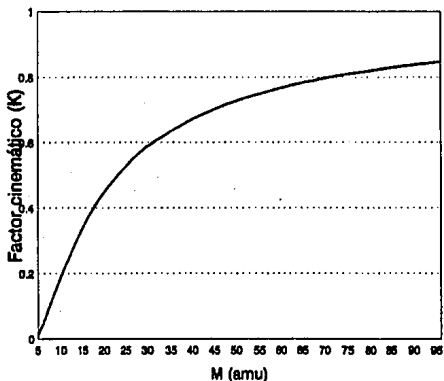


Fig. 19. Factor cinemático para el hidrógeno en función de la masa de los átomos del blanco, para un ángulo de 170° .

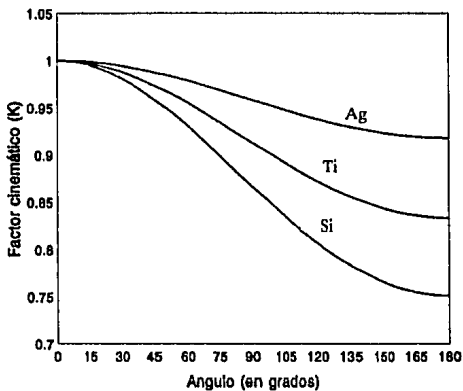


Fig. 20. Factor cinemático para hidrógeno en plata, titanio y silicio en función del ángulo de dispersión.

Sección transversal La cantidad de átomos por unidad de área del blanco determina la probabilidad de que exista una colisión entre los iones incidentes y los átomos del blanco. Esta probabilidad está dada por la "sección transversal", que es el área efectiva presentada por cada núcleo a los iones incidentes. La sección recta diferencial determinará el número de partículas que serán dispersadas en un ángulo dado y se define como:

$$A = \sigma_{\text{RBS}} N t \Omega Q$$

donde A es el número total de partículas detectadas, Nt el número de átomos por unidad de área en el blanco ($N =$ densidad volumétrica, $t =$ espesor), $\Omega =$ ángulo sólido que subtende el detector y Q el número total de iones incidentes.

Rutherford calculó la sección transversal —en coordenadas de centro de masa—, tomando en cuenta únicamente la fuerza coulombiana entre el ion y el átomo, considerándolos como cargas puntuales [3]. Tres años después, Darwin [4] hace la transformación a coordenadas de laboratorio. La expresión que obtiene es la siguiente (en cm^2):

$$\sigma_{\text{RBS}} = \left(\frac{Z_1 Z_2 e^2}{4 E_0} \right) \left(\frac{4}{\sin^2(\theta)} \right) \frac{(1 - u_1^2)^{\frac{1}{2}} + \cos(\theta)}{(1 - u_1^2)^{\frac{1}{2}}}$$

donde

$$u_1 = \frac{m_1 \sin(\theta)}{m_2}$$

Se han efectuado correcciones, tomando en cuenta el efecto de apantallamiento debido a la nube electrónica circundante al núcleo, quedando la siguiente expresión:

$$\sigma_{\text{RBS corregida}} = \sigma_{\text{RBS}} \left(1 - 0.049 Z_1 Z_2^{\frac{3}{4}} / E_1 \right)$$

donde E_1 está expresada en keV. De acuerdo con esta ecuación, es evidente la dependencia de la sección con la energía del ion.

Geometría Al tratar con retrodispersión, el ángulo formado por el haz incidente y las partículas dispersadas es mayor que 90° . Generalmente, se

coloca el detector a ángulos cercanos a 180° (típicamente entre 165° y 175°). La geometría del experimento se presenta en la siguiente figura.

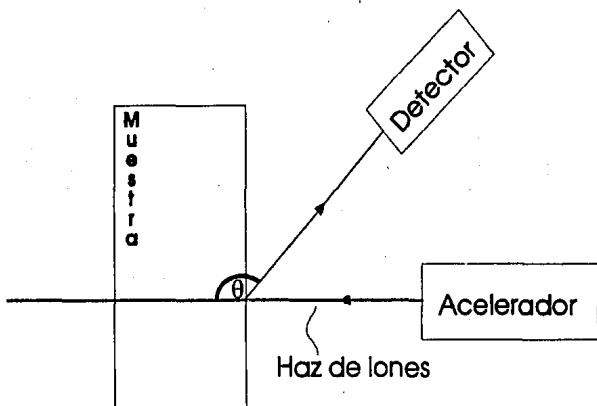


Fig. 21. Geometría del experimento

Frenamientos nuclear y electrónico Existen dos mecanismos básicos de pérdida de energía conforme el ion atraviesa la muestra: frenamiento nuclear —interacción elástica con el núcleo apantallado, que conlleva una transferencia de energía al núcleo atómico— y frenamiento electrónico —excitación e ionización inelástica.

Debido a que, de acuerdo con un criterio meramente estadístico, las colisiones elásticas son menos probables que las inelásticas, se pueden considerar por separado. Sin embargo, lo anterior puede no ser necesariamente cierto, ya que, al ocurrir una colisión elástica, conlleva excitaciones electrónicas, que afectarán el apantallamiento electrónico y la transferencia de energía elástica. Además, cuando la velocidad relativa es comparable con la velocidad orbital de los electrones, existe un término adicional debido al intercambio de carga entre el ion incidente y el átomo del blanco, que puede representar una pérdida adicional de energía del 10% a 20%[5].

Los choques elásticos son más importantes a bajas energías y los inelásticos a altas energías, según se muestra en las siguientes gráficas, obtenidas de acuerdo con Ziegler, Biersack y Littmark [6] (frenamiento nuclear) y Montenegro, Cruz y Vargas [7] (frenamiento electrónico). Esto se debe a que, a altas energías —que equivale a altas velocidades— el ion incidente está completamente ionizado, por lo que interactúa con la nube electrónica de los átomos del blanco —esta interacción es a más larga distancia que los choques elásticos. Conforme se va reduciendo la velocidad del ion, va atrapando electrones, por lo que el frenamiento electrónico no es tan importante; de hecho, para obtener el valor del frenamiento electrónico en un momento dado, en función de la carga del ion, se considera un potencial apantallado, no la carga completa de la nube electrónica. Por otro lado, conforme el ion va capturando electrones, presenta un área mayor para la colisión elástica con los átomos de la muestra.

En las dos referencias se presentan aproximaciones numéricas, ya que no existe una expresión analítica para el cálculo de estas pérdidas de energía. J. Rickards desarrolló un programa en BASIC que permite obtener el frenamiento total (nuclear y electrónico) que se presenta a continuación.

```

0 ' dedx
50 REM programa dedx, calcula dedx nuclear segun zbl
  + electronico segun mcv en compuestos
60 REM resultados en kev/mg/cm2
100 INPUT "z y m del proyectil";Z1,M1
110 INPUT "cuantos elementos tiene el blanco";NE
115 FOR IE=1 TO NE
116 PRINT "z del elemento numero ";IE;" ?"
117 INPUT Z(IE)
118 INPUT "su numero por molecula";C(IE)
119 NEXT IE
124 REM selecciona constantes para mcv
125 FOR ID=1 TO NE
126 RESTORE
127 FOR IG=1 TO 92
128 READ A1,A2,A3,A4,A5
129 IF A1=Z(ID) THEN M(ID)=A2 : ZS(ID)=A3 : KA(ID)=A4
  : PI(ID)=A5 : IG=92
130 NEXT IG
131 NEXT ID
140 INPUT "dar energias inicial, final, e intervalo
  en kev";E1,E2,E3

```

```
145 PRINT:PRINT " E(keV) dE/dx en kev/mg/cm2":PRINT
146 INPUT "nombre de la base";BASE$
147 OPEN BASE$ FOR OUTPUT AS #1
150 FOR EK=E1 TO E2 STEP E3
200 FOR IB=1 TO NE
205 M2=M(IB):Z2=Z(IB):ZA=ZS(IB):K=KA(IB):IP=PI(IB)
210 GOSUB 9000
215 GOSUB 8000
220 TE(IB)=SE*602.3/M(IB):TN(IB)=SN*602.3/M(IB)
225 DT(IB)=TE(IB)+TN(IB)
230 NEXT IB
349 DN=0:NM=0
350 FOR IC=1 TO NE
355 NM=NM+M(IC)*C(IC)*DT(IC)
360 DN=DN+M(IC)*C(IC)
365 NEXT IC
370 DD=NM/DN
380 PRINT EK,DD
385 PRINT#1,EK,DD
390 NEXT EK
400 END

8000 REM frenamiento nuclear
8005 EP=32.53*M2*EK/(Z1*Z2*(M1+M2)*((Z1^.23)+
(Z2^.23)))
8010 IF EP30 THEN GOTO 8030
8020 SM = LOG(1+1.1383*EP) *.5/ (EP + .01321 *
(EP ^ .21226) + .19593*SQR(EP)) : GOTO 8040
8030 SM=LOG(EP)*.5/EP
8040 SN=8.462*Z1*Z2*M1/((M1+M2)*((Z1^.23)+(Z2^.23)))
8045 SN=SN*SM
8050 RETURN

9000 REM frenamiento electronico
9010 AM=9.5616*Z2
9011 BM=54.38/IP
9012 CM=AM*BM
9013 DM=(K/CM)+(BM/2)+(ZA/(9*Z2))
9014 EN=3*BM*BM*DM
9015 ED=((Z2^1.4)*BM*DM*DM)+(BM*BM)+(DM*DM)
9016 EM=EN/ED
9017 FM=CM*EM*EM*EM/DM
9018 GM=2/(3*CM)
9020 U=.1997*SQR(EK/M1):U2=U*U
```

```
9025 S1=K*U*EXP(-U)
9026 S2=AM*LOG(1+BM*U2)/U2
9027 S3=CM/(1+DM*U2)
9028 S4=FM*U2*U2/((1+EM*U2)^3)
9029 S5=((GM*U2)^1.5)/(6*BM*BM*((1+GM*U2)^4))
9030 SP=S1+S2-S3+S4-S5
9040 AL=Z1^(-.66666667)
9050 GN=1-EXP(-AL*U)-(AL*U*EXP(-2*AL*U))/6
9051 GD=1-EXP(-U)-(U*EXP(-2*U))/6
9052 G2=GN*GN/(GD*GD)
9060 SE=G2*Z1*Z1*SP
9065 IF Z1=1 THEN SE=SP
9070 RETURN
9101 DATA 1,1,1,6.305,17
9102 DATA 2,4,2,6.14,42
9103 DATA 3,6.94,3,7.05,47
9104 DATA 4,9.01,4,11.232,63
9105 DATA 5,10.81,3,12.361,75
9106 DATA 6,12.01,4,13.145,79
9107 DATA 7,14.01,5,14.759,86
9108 DATA 8,16,6,13.25,99
9109 DATA 9,19,7,10.417,118.8
9110 DATA 10,20.18,8,9.748,135
9111 DATA 11,22.99,9,12.7,141
9112 DATA 12,24.32,10,18.946,149
9113 DATA 13,26.98,11,20.754,162
9114 DATA 14,28.09,10,20.734,159
9115 DATA 15,30.98,5,16.148,168.9
9116 DATA 16,32.07,6,17.222,179.2
9117 DATA 17,35.46,7,25.216,170.3
9118 DATA 18,39.94,8,28.633,180
9119 DATA 19,39.1,9,25.736,189.4
9120 DATA 20,40.08,10,27.584,195
9121 DATA 21,44.96,11,25.985,215
9122 DATA 22,47.9,12,24.292,228
9123 DATA 23,50.95,13,22.383,237
9124 DATA 24,52.01,14,19.9,257
9125 DATA 25,54.94,15,17.332,275
9126 DATA 26,55.85,16,17.582,284
9127 DATA 27,58.94,17,15.688,304
9128 DATA 28,58.69,18,17.752,314
9129 DATA 29,63.54,19,18.466,330
```

9130 DATA 30,65.38,18,21.034,323
9131 DATA 31,69.72,19,25.186,335.4
9132 DATA 32,72.6,20,27.749,323
9133 DATA 33,74.91,15,26.595,354.7
9134 DATA 34,78.96,16,29.348,343.4
9135 DATA 35,79.92,17,28.034,339.3
9136 DATA 36,83.8,18,32.031,347
9137 DATA 37,85.48,19,28.449,349.7
9138 DATA 38,87.63,10,31.671,353.3
9139 DATA 39,88.92,11,32.011,365
9140 DATA 40,91.22,12,33.645,382
9141 DATA 41,92.91,13,34.484,391.3
9142 DATA 42,95.95,14,32.101,393
9143 DATA 43,99.15,33.969,416.2
9144 DATA 44,101.1,16,30.517,428.6
9145 DATA 45,103.91,17,29.598,436.4
9146 DATA 46,106.7,18,26.17,456
9147 DATA 47,107.88,19,28.094,470
9148 DATA 48,112.41,20,29.048,466
9149 DATA 49,114.76,21,31.127,479
9150 DATA 50,118.7,20,32.126,511.8
9151 DATA 51,121.76,21,37.472,491.9
9152 DATA 52,127.61,22,34.869,491.3
9153 DATA 53,126.91,21,38.596,452.4
9154 DATA 54,131.3,18,41.124,459
9155 DATA 55,132.91,19,36.408,484.8
9156 DATA 56,137.36,20,39.465,485.5
9157 DATA 57,138.92,21,40.175,493.8
9158 DATA 58,140.13,22,37.417,512.7
9159 DATA 59,140.92,23,36.428,520.2
9160 DATA 60,144.27,24,35.463,540
9161 DATA 61,145,25,34.524,537
9162 DATA 62,150.43,16,33.615,545.9
9163 DATA 63,152,17,32.73,547.5
9164 DATA 64,156.9,18,33.67,567
9165 DATA 65,158.93,19,31.037,577.2
9166 DATA 66,162.46,20,27.564,578
9167 DATA 67,164.94,21,26.075,612.2
9168 DATA 68,167.2,22,25.336,583.3
9169 DATA 69,168.94,23,24.612,629.2
9170 DATA 70,173.04,24,23.917,637
9171 DATA 71,174.99,25,24.447,655.1

9172 DATA 72,178.6,26,25.121,662.9
 9173 DATA 73,180.95,27,23.672,682
 9174 DATA 74,183.92,28,22.853,695
 9175 DATA 75,186.31,29,25.98,713.6
 9176 DATA 76,190.2,30,25.331,726.6
 9177 DATA 77,192.2,31,24.706,743.7
 9178 DATA 78,195.23,32,22.363,760
 9179 DATA 79,197.33,24.262,742
 9180 DATA 80,200.61,34,21.524,768.4
 9181 DATA 81,204.39,33,23.597,764.8
 9182 DATA 82,207.21,20,26.575,761
 9183 DATA 83,209,21,29.758,762.9
 9184 DATA 84,210,22,30.767,765.1
 9185 DATA 85,210,23,30.997,761.7
 9186 DATA 86,222,24,30.882,733.1
 9187 DATA 87,223,19,34.719,762.3
 9188 DATA 88,226.05,20,37.502,760.1
 9189 DATA 89,227,21,38.216,767.9
 9190 DATA 90,232.05,22,38.521,776.4
 9191 DATA 91,231,23,37.007,807
 9192 DATA 92,238.07,24,36.423,808

De acuerdo con el programa anterior, se obtienen las siguientes gráficas para los frenamientos nuclear y electrónico.

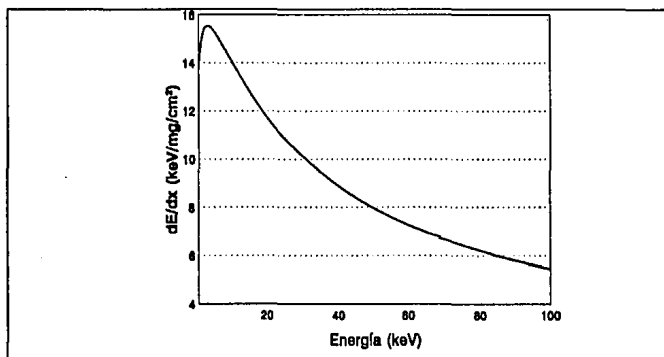


Fig. 22. Frenamiento nuclear de He^{++} en Ag en función de la energía del ion. En la gráfica se presenta un acercamiento en la región de baja energía, para apreciar su comportamiento, que es muy similar cualitativamente al electrónico.

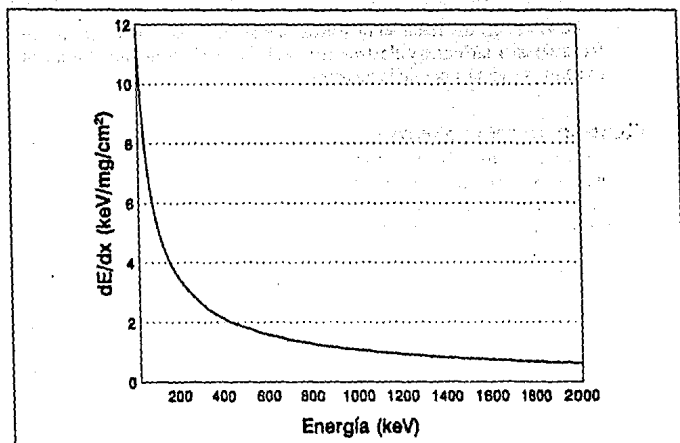


Fig. 23. Frenamiento nuclear de He^{++} en Ag en función de la energía del ion

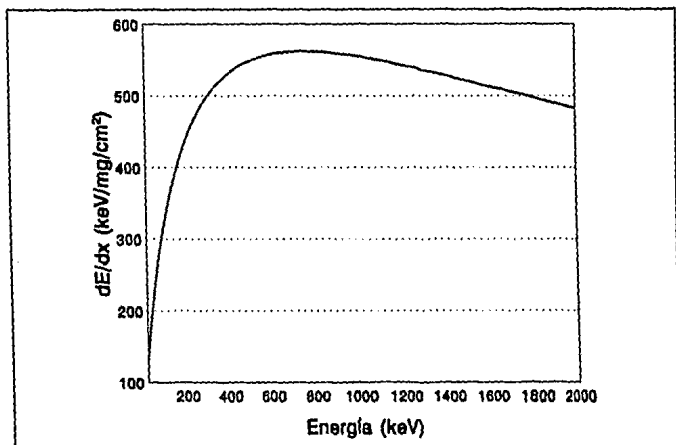


Fig. 24. Frenamiento electrónico de He^{++} en Ag en función de la energía del ion

En resumen, los fenómenos físicos que se están llevando a cabo son frenamientos nucleares y electrónicos y colisiones elásticas entre los iones incidentes y los átomos de la muestra.

Generación del espectro Como ya se mencionó antes, el detector tiene como función detectar las partículas del haz de iones que fueron dispersadas hacia él; más aún, también debe ser capaz de traducir la energía de la partícula en una señal eléctrica proporcional a dicha energía. Usualmente son detectores de estado sólido, que generan un número de pares hoyo-electrón proporcional a la energía de la partícula incidente. Estos pares son atraídos hacia los polos del detector, generando así una carga que dependerá de la energía del ion dispersado. Esta señal tradicionalmente pasaba a un amplificador y de ahí a un multicanal, que es un dispositivo electrónico capaz de discriminar los niveles de voltaje que recibe e ir contando estas señales en diferentes canales, cada uno de los cuales tiene un intervalo de energía asociado.

Conforme van llegando partículas con diferentes energías al detector, en el multicanal se van contando los eventos en los canales correspondientes. De esta manera, se va formando el espectro experimental, que no es más que una gráfica de número de cuentas vs. energía.

Las energías de las partículas deflectadas que se observan en el detector forman un continuo, con una energía máxima que depende de la energía inicial del ion y de la masa de los átomos que conforman la muestra. De hecho, de acuerdo con las ecuaciones de factor cinemático para retrodispersión y la gráfica 20, que se presentó en el apartado sobre factor cinemático, se observa que mientras mayor sea la masa del átomo del blanco, menor será la energía que pierde el ion en el choque, por lo que la energía de salida será mayor, de acuerdo con las gráficas 19 y 20, que se presentaron al tratar el factor cinemático.

Ahora veamos cómo se genera el llamado "espectro experimental". Un espectro experimental se puede describir como la gráfica de la cantidad de iones detectados en cada intervalo de energía.

Las deducciones teóricas detalladas para la generación de espectros de RBS se presentan en las referencias [8, 9, 10, 11, 12, 13, 14, 15, 16, 17], junto con algunos resultados para casos particulares.

Un espectro de RBS se forma de la siguiente manera: de las partículas del haz de iones que inciden sobre la muestra, algunas chocan con los

átomos superficiales y son dispersados; se pierde una parte de la energía debido a la colisión. La relación entre las energías antes y después del choque está dada por el factor cinemático. Los iones que chocan con átomos de la capa superficial del blanco y son retrodispersados hacia el detector contribuirán con los valores más altos de energía al espectro experimental. Los iones van penetrando a diferentes profundidades antes de que se presente el choque, y sólo una pequeña parte de éstos serán dispersados hacia el detector. Mientras se adentran en la muestra, van perdiendo energía (colisiones inelásticas o frenamientos electrónico y nuclear), de modo que, al momento en que se lleva a cabo el choque, su energía será menor conforme más adentro ocurra. Luego, al ser dispersados, atraviesan una porción del blanco, por lo que siguen perdiendo energía, hasta que salen y llegan al detector. Conforme mayor sea la profundidad a la que choca un ion con un átomo de la muestra, llegará con menor energía al detector, por lo que generará menos pares hoyo-electrón en éste y, en consecuencia, el voltaje de salida será menor; el multicanal lo registrará en un canal menor.

Cabe hacer notar que, si la muestra está formada por átomos de un solo elemento, el espectro experimental será similar al que se presenta en la gráfica 25. Sin embargo, conforme más elementos tiene la muestra, dependiendo de las masas, los espectros pueden ser muy complejos en su análisis. Se presentan dos ejemplos de diferentes espectros.

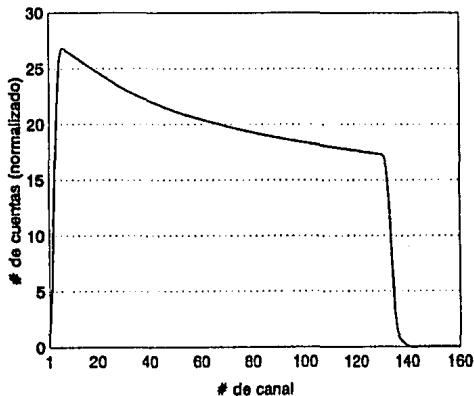


Fig. 25. Espectro simulado en RUMP de silicio puro bombardeado con un haz de He^+ de 2 MeV

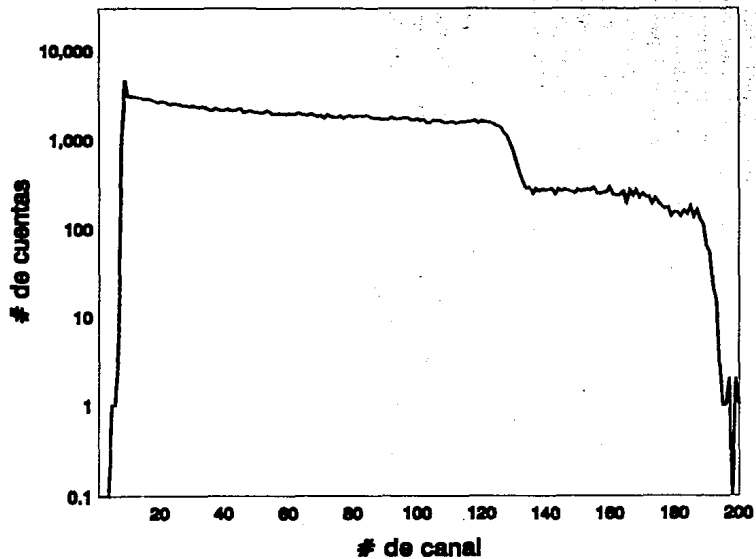


Fig. 26. Espectro de RBS de Al 7075 atacado con F por 5 hrs. a 300° y 1 atmósfera; bombardeado con partículas α de 2 MeV y detectadas a un ángulo de 165° .

Como se puede apreciar en los diferentes espectros, algunos de ellos permiten un análisis relativamente sencillo de la composición de la muestra, ya que los picos correspondientes a cada uno de los elementos que lo componen están separados. En cambio, en otros, se tiene una serie de picos sobre un fondo de niveles tales que los opacan.

A continuación presentamos las maneras más comunes de analizar espectros.

ANÁLISIS DE ESPECTROS

Conforme los iones van atravesando el blanco, pierden energía a razón de dE/dx , por lo que un ion que ha penetrado a una profundidad t tendrá una energía dada por

$$E = E_0 - \frac{1}{\cos(\theta_1)} \int_0^t \frac{dE}{dx} dx = E_0 - \frac{N}{\cos(\theta_1)} \epsilon_1(E)$$

donde θ_1 es el ángulo entre la trayectoria de entrada y la normal a la superficie del blanco, N es la densidad atómica de la muestra y $\epsilon_1(E)$ es la sección de frenamiento que presentan los átomos de la muestra (incluye frenamiento nuclear y electrónico). Si trabajamos a incidencia normal, tenemos que $\cos(\theta_1) = 1$ y

$$E = E_0 - \int_0^t \frac{dE}{dx} dx$$

Aproximación superficial Si consideramos t "pequeño", la integral se puede aproximar como:

$$\int_0^t \frac{dE}{dx} \approx \frac{dE}{dx} \Big|_{in} \cdot t$$

Al ocurrir un choque elástico a la profundidad t , el ion rebotará y saldrá de la colisión con una energía $KE(t)$ y en el trayecto de salida seguirá perdiendo energía y saldrá de la muestra con una energía dada por

$$E_1(t) = KE(t) - \frac{t}{|\cos(\theta)|} \frac{dE}{dx} \Big|_{out}$$

donde θ es el ángulo de dispersión.

Por tanto, el ancho en energía ΔE de la señal de una película con grosor Δt es

$$\Delta E = \Delta t \left(K \left. \frac{dE}{dx} \right|_{in} + \frac{1}{|\cos(\theta)|} \left. \frac{dE}{dx} \right|_{out} \right) = \Delta t [S]$$

Al término [S] se le denomina "factor de pérdida de energía de retrodispersión".

La suposición de tomar valores constantes para la pérdida de energía lleva a considerar una relación lineal entre ΔE y la profundidad t a la cual ocurre la dispersión. Esta aproximación es válida cuando se está trabajando con películas delgadas, y se usa la "aproximación superficial", en la cual se evalúa $(dE/dx)_{in}$ en E_0 y $(dE/dx)_{out}$ se evalúa en KE_0 . En esta aproximación, el ancho en energía debido a una película de grosor t es

$$\Delta E = \Delta t \left[K \left. \frac{dE}{dx} \right|_{E_0} + \frac{1}{|\cos(\theta)|} \left. \frac{dE}{dx} \right|_{KE_0} \right] \quad (1)$$

donde los subíndices denotan la aproximación superficial. Esta ecuación puede usarse para determinar la profundidad a la que se encuentra determinado elemento de la muestra, o para analizar difusión de una película en un sustrato.

Aproximación para películas gruesas Cuando el grosor de la trayectoria recorrida por el ion se torna apreciable, ya no es posible usar la aproximación superficial (para películas delgadas), debido a que ya no es válido considerar dE/dx como constante. Entonces, se considera una energía media E entre los extremos del trayecto. Para la partícula incidente, se tiene que

$$E_{in} = \frac{1}{2} [E(t) + E_0]$$

Y para la de salida se considera

$$E_{out} = \frac{1}{2} [E_1 + KE(t)]$$

Para materiales cuya estructura es diferente a interfaces definidas, como puede ser el caso de películas delgadas interdifundidas o aleaciones, no se puede usar ninguna de las dos aproximaciones anteriores, lo que implica que se debe evaluar la pérdida de energía exactamente para cada punto (o intervalo, en caso de estar utilizando aproximaciones numéricas).

A continuación se presentan dos espectros. En el primero, se puede usar la aproximación superficial y en el segundo no, debido a los grosores de las muestras en cada caso. Sustituyendo estos valores en la ecuación (1), se puede determinar la profundidad a la que se encuentra un elemento dado en la muestra.

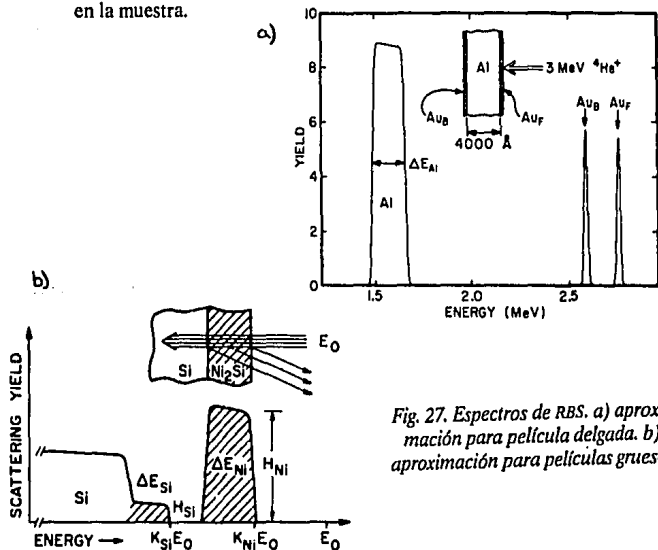


Fig. 27. Espectros de RBS. a) aproximación para película delgada. b) aproximación para películas gruesas

Altura del pico La altura que alcanzará cada uno de los picos — correspondientes a cada elemento de la muestra — estará dado por la siguiente ecuación para películas delgadas:

$$Y = \sigma(\theta) \Omega Q N t$$

donde Q es el número de partículas incidentes y Nt es el número de átomos en el blanco (en átomos/cm²) y Ω es el ángulo que subtende el detector.

Para películas gruesas, los proyectiles son dispersados a cualquier profundidad, dando como resultado un espectro continuo. El número de cuentas que se tendrá en un canal dado está dado por la ecuación:

$$Y(t) = \left[\frac{Z_1 Z_2 e^2}{4 E(t)} \right] \Omega Q N t$$

donde $E(t)$ es la energía de la partícula a la profundidad t y N es la densidad atómica.

A partir de estas ecuaciones, se puede determinar la densidad atómica de cada uno de los elementos que conforman el blanco, que depende de la relación entre las alturas de los picos. De este modo, se tiene que, en una primera aproximación —donde se ignora la diferencia entre las secciones de frenamiento entre los dos tipos de átomos—, el cociente entre las alturas de dos picos da como resultado la composición de la capa, de acuerdo con la siguiente ecuación

$$\frac{N_1}{N_2} \approx \frac{H_1 \sigma_2}{H_2 \sigma_1} \approx \frac{H_1 Z_2^2}{H_2 Z_1^2}$$

donde los subíndices se refieren al primer y segundo elemento del blanco, respectivamente.

Una mejor aproximación se tiene cuando se considera que la relación que existe entre $Y(t)$ y ΔE es lineal, por lo que la ecuación queda como:

$$\frac{N_1}{N_2} \approx \frac{H_1 \Delta E_1 \sigma_2}{H_2 \Delta E_2 \sigma_1}$$

En el próximo capítulo presentaremos, de manera general, la manera de analizar espectros con las bases teóricas aquí presentadas, y haciendo uso de la computadora como herramienta para procesar la información.

LAS COMPUTADORAS EN LA RETRODISPERSIÓN DE RUTHERFORD

Hasta este momento, hemos visto los aspectos teóricos y consideraciones experimentales de la técnica RBS para análisis de superficies. Ahora veremos el impacto que ha tenido la computadora personal, tanto en la parte experimental (toma de datos) como el apoyo que brinda para el análisis de los espectros y la determinación de la composición superficial de una muestra usando RBS.

SISTEMAS DE ADQUISICIÓN DE DATOS POR COMPUTADORA

Inicialmente, en el IFUNAM se contaba con un acelerador de partículas — iones positivos a 750 keV — y su electrónica asociada. La salida era a través de un multicanal, lo que implicaba el tener que irse desplazando de canal en canal, e ir anotando las lecturas. Luego se incluye una computadora HP, cuya salida era por teletipo o cinta perforada. De esta manera, se abrevia el proceso de ir leyendo los datos de cada canal y anotando en papel, para poder hacerse el análisis de los resultados del experimento. Esta computadora tenía 32 kb de memoria RAM y, en lugar de contar con un microprocesador en forma de circuito integrado — que no existía en aquellas épocas — éste se construía a partir de lo que ahora conforman las componentes internas del procesador: memoria ROM, donde se almacena el “microcódigo” y la lógica asociada para controlar el flujo de información. La memoria RAM es externa, es decir, está ubicada físicamente en otras tarjetas (en módulos de 8 kb cada una). El programa se almacena en RAM y, si no hay fallas de alimentación, se mantiene almacenado aunque se apague la computadora. Sin embargo, si llegase a faltar la corriente eléc-

trica, el programa se pierde. El medio de almacenamiento permanente con que se cuenta son las cintas perforadas.

Cuando aparecen en el mercado las PC, con la filosofía de tecnología abierta, y se popularizan, empiezan a surgir en el mercado nuevas tarjetas (sobre todo con bus ISA), dedicadas a funciones tales como la emulación de diferentes aparatos de medición, que incluyen desde voltímetros y osciloscopios hasta tarjetas más sofisticadas, para la toma de datos de experimentos en "tiempo real", es decir, conforme ocurre un evento, inmediatamente es detectado, codificado y almacenado en la computadora.

Una de estas tarjetas es la MCA, fabricada por Ortec, que es un emulador de multicanal —existen otras tarjetas multicanal en el mercado, como la fabricada por Camberra, pero nos limitamos a la ORTEC, que es con la que se realizó parte de este trabajo. Cada tarjeta incluye su(s) programa(s) de control; en el caso de la MCA Ortec, el software se llama MCB. Así, para abreviar la captura de datos en computadora, se elige una computadora tipo PC-XT, debido a dos factores: el primero, que no se requiere una máquina muy veloz para la adquisición de datos y basta con una máquina a 10 MHz; y el segundo fue el momento del cambio del equipo, ya que solamente había en el mercado computadoras con procesador 8086 y 80286, y estas últimas eran muy costosas.

La función básica del MCA es el traducir las señales provenientes del amplificador, que son voltajes analógicos, y traducirlos a un código binario (ADC —Analog to Digital Converter), además de proveer el protocolo necesario para que la computadora pueda entender los datos y desplegar el espectro en pantalla. El programa es el encargado de:

- Traducir el valor binario de la tarjeta a una cuenta en el canal correspondiente e ir sumando las cuentas por canal para formar el espectro.
- Desplegar en pantalla (monitor) el espectro.
- Dar el control al usuario para iniciar / detener el conteo.
- Permitir la graficación con diferentes escalas (lineal y semi-log).
- Almacenar los espectros.
- Permitir al usuario seleccionar un canal dado y mostrar el número de cuentas acumulado en él.
- Abrir archivos de espectros almacenados previamente para su despliegue en pantalla.

El programa puede ejecutarse sin necesidad de la tarjeta, en caso de que se desee desplegar un espectro previamente almacenado, quedando inactivas las opciones para captura de datos.

Como ya se mencionó, en el IFUNAM, el grupo de Colisiones cuenta con una tarjeta MBA Ortec y su programa asociado (MCB), que codifica de una manera particular los datos, y no es compatible con las hojas electrónicas comerciales (como Lotus o Excel), de modo que no es posible almacenar un espectro experimental e importarlo a una hoja electrónica para su análisis. Por tanto, se desarrolló un programa que permitiese traducir el código del MCB a ASCII, que es un formato universal, exportable a procesadores de texto, hojas electrónicas, etc.

Para poder hacer el programa, fue necesario usar Norton Utilities, para explorar los archivos de los espectros experimentales y encontrar cual era la codificación de datos que se usa. El listado se presenta a continuación.

```

0 ' LEESPEC
10 ' Programa para leer los valores de un espectro
   generado con
20 ' el MCB y cuya extensión es .CHN. Escribe los
   valores a un archivo
30 ' en ASCII llamado DATOS.PRN.
40 OPEN "datos.prn" FOR OUTPUT AS #2
50 OFFSET%=8 : NUMCANAL=0
60 INPUT "Nombre del archivo del espectro (con
   extensión) ", FILENM$
70 OPEN "x", #1, FILENM$,4
80 FIELD #1,2 AS VAL1$,2 AS VAL2$
90 CODE%=OFFSET%+NUMCANAL%
100 GET #1, CODE%
110 LONGITUD=CVI(VAL2$)
120 PRINT "La longitud en canales es de ";LONGITUD
130 ' Actualizamos el offset para que el primer canal
   que lea sea el cero
140 OFFSET%=OFFSET%+1
150 FOR NUMCANAL%=0 TO LONGITUD-1
160 CODE%=OFFSET%+NUMCANAL%
170 GET #1, CODE%
180 VALOR1#=CVI(VAL1$) : VALOR2# = CVI(VAL2$) : VALOR#
   = VALOR2# * 65536! + VALOR1#
190 PPRINT #2, NUMCANAL%, VALOR#

```

```
200 NEXT  
210 END
```

El programa se desarrolló en BASIC, debido a que, si se cuenta con el sistema operativo DOS, siempre se incluye este lenguaje de programación, sin importar la versión de DOS que se tenga.

Se requiere este programa porque el MCB no tiene la opción de imprimir los espectros más que haciendo uso de las teclas Shift y PrtScrn, con lo que se envía a impresora toda la información que se está desplegando en pantalla, y el espectro aparece en un tamaño fijo de casi media página, junto con todas las funciones del MCB, por lo que no resulta práctico para trabajar directamente sobre la gráfica — además de que su presentación no es la más adecuada para conferencias, publicaciones, etc.

De este modo, contando con una PC y la tarjeta emuladora de multicanal se pueden simplificar considerablemente la captura de datos y su posterior presentación. De hecho, la PC representa un avance importante con respecto a la computadora HP que se tenía antes, ya que, por primera vez con ese sistema, la captura de datos se lleva a cabo de manera automática. Más aún, debido a que se presenta el espectro en pantalla — además de almacenarse los datos en medios electrónicos — el investigador puede ir tomando decisiones de momento, como el aumentar el tiempo de radiación para eliminar ruido de fondo.

PROGRAMAS PARA ANÁLISIS DE ESPECTROS

Conforme se fue popularizando la técnica de RBS, se fueron desarrollando programas de computadora que permiten generar espectros teóricos. En un principio, estos programas se diseñaron para computadoras grandes (*mainframes*), debido a que eran esas las máquinas que tenía disponible un investigador, pero cada vez es mayor su adaptación a computadoras personales. Actualmente, aún existen objeciones al uso de las PC, debido a la resolución que se puede obtener con los procesadores de 16 bits, que son los más populares. Sin embargo, con la proliferación de los nuevos procesadores de 32 bits, se pueden lograr resoluciones mejores. Si consideramos que la computadora puede almacenar información en diferentes formatos, dependiendo de las necesidades del usuario, cuando se llevan a cabo operaciones con números no enteros, se usa el llamado "formato de punto flotante". Usándolo, la computadora descompone el número en tres partes: signo, mantisa y exponente.

El formato de un número cualquiera escrito en punto flotante es el siguiente:

$$\text{signo} * (1.a_0a_1\dots a_{23}) \times 2^{\wedge} \text{exponente}$$

En una PC se usa un bit de signo, 24 de mantisa y ocho de exponente. Si sumamos el número de bits, encontraremos que se requeriría de 33 bits para la codificación; sin embargo, basta con 32, ya que la computadora codifica eligiendo siempre un 1 como bit más significativo en la mantisa, por lo que se puede omitir y usar realmente 23 bits de memoria, aunque la mantisa tenga 24.

De este modo, tenemos que el número más grande que se puede almacenar en este formato es de aproximadamente 3.4×10^{38} ($= 2 \times 2^{128}$). El más pequeño, si no se trabaja con signos, es de aproximadamente 1×10^{-38} —ya que el bit de signo se aprovecha en la mantisa. Si se incluye signo, tenemos que el mayor es 2^{128} y el menor es -3.4×10^{-38} .

La PC también permite llevar a cabo cálculos con números de “doble precisión”, es decir, que requieren el doble de bits para su codificación. Si usamos este formato, tenemos que la asignación de bits es la siguiente:

1 bit de signo, 11 bits de exponente y 53 bits de mantisa (igual que en el caso anterior, se utilizan físicamente 52 y el más significativo es “asumido”, ya que siempre tiene un valor de 1 y se puede omitir).

En la siguiente tabla se presenta un cuadro indicando los valores máximos y mínimos (con y sin signo) que se pueden almacenar en formato de punto flotante.

| | <i>Exponente (bits)</i> | <i>Mantisa (bits)</i> | <i>Signo (bits)</i> | <i>Número mayor</i> | <i>Número menor</i> |
|--------------------|-----------------------------|---------------------------|-------------------------|-------------------------|-------------------------|
| Precisión sencilla | 8 | 25 | 0 | 3.4×10^{38} | 1×10^{-38} |
| | 8 | 24 | 1 | 1.7×10^{38} | -3.4×10^{-38} |
| Precisión doble | 11 | 53 | 0 | 1.8×10^{308} | 1×10^{-309} |
| | 11 | 52 | 1 | 9×10^{307} | -1.8×10^{308} |

De este modo, resulta que la resolución al usar estos procesadores es mucho mayor de lo que se puede obtener en cualquier experimento, e incluso es más lo que se pierde por aproximaciones numéricas (en la

solución numérica de ecuaciones que no tienen una expresión cerrada) y en los valores estimados para poderes de frenamiento nuclear y electrónico o cualquier otra variable física que intervenga en el experimento o modelo que se esté desarrollando.

Es importante notar que, pese a la gran precisión que se puede obtener, se pierden dígitos poco significativos cuando las operaciones que se llevan a cabo involucran números con exponentes muy diferentes. Esto se debe a que, al sumar dos cifras con diferentes exponentes, la computadora ajusta mantisas y exponentes de modo que estos últimos sean iguales; suma los dígitos de la mantisa y, en caso de que exista un acarreo, suma uno al exponente y en caso contrario, permanece igual. En términos generales, se estima la precisión es de aproximadamente 17 dígitos en la mantisa al trabajar con doble precisión.

Regresando al tema, todos estos programas de análisis de espectros de RBS parten de la idea de que el investigador sabe "más o menos" la composición atómica en cada capa de la muestra que desea analizar. Introduce estos datos al programa, que se encarga de generar el espectro correspondiente a esa composición atómica. Si no son iguales (dentro de los límites establecidos por las aproximaciones numéricas, resolución del detector y electrónica asociada usados, ruido en las señales, fenómenos físicos no considerados, etc.), se puede modificar la composición, y volver a generar el espectro, hasta que coincidan. En general, de acuerdo con Rickards *et al*, si la diferencia máxima entre los dos espectros es menor o igual al 7% [18], se considera que la composición atómica supuesta en el programa, es igual a la real. Lo anterior se debe a las aproximaciones que se hacen para los cálculos numéricos. Cabe hacer notar que este error se refiere a la desviación máxima entre los valores obtenidos por los diferentes métodos [7, 19, 20, 21] respecto al promedio de dichos métodos.

Los métodos son todos interactivos, por lo que, si la decisión inicial en cuanto a la composición de la muestra no fue adecuada, puede tomar un tiempo relativamente grande el ajuste del espectro experimental.

A partir de los años 70's, empezaron a aparecer artículos presentando diferentes algoritmos para la generación de espectros teóricos. Inicia en 1975, con la publicación del artículo de Chu y Ziegler llamado "An analytic solution to elastic backscattering" [11], donde se presenta una aproximación analítica para la predicción de espectros de RBS de $^4\text{He}^+$ en Al, obteniendo diferencias máximas del 2% entre espectro teórico y experimental y evitando las integraciones numéricas que tradicionalmente se requerían para el cálculo.

Poco después aparece un artículo de Ziegler, et al. [22] donde se discute el uso potencial de la computadora para el análisis de espectros de RBS por usuarios no expertos. Presentan los resultados de un programa desarrollado por ellos (en Vax, aunque luego se adapta a las PC) y que incluye el efecto debido a isótopos, resolución del detector, pérdida de energía del proyectil en la superficie del detector, y la interdifusión de dos o más capas.

Durante los años 80's, se presentaron cada vez más de estos programas, que aún se utilizan. Entre los más usados actualmente se encuentran:

- RUMP (R. L. Doolittle, 1985) [23, 24]
- ION-SCAT (J. W. Buttler, 1986) [25]

El programa de Buttler —llamado ION-SCAT— fue diseñado especialmente para aplicaciones metalúrgicas, y permite realizar análisis de blancos con hasta ocho elementos en 55 capas con grosores definidos por el usuario.

RUMP permite calcular espectros tanto de RBS como de ERDA, ya que está basado en el método interactivo presentado por Chu, *et al.* [10]. Este programa presenta algunos problemas con los cálculos de pérdida de energía, ya que ajustan la sección de frenamiento por un polinomio de quinto grado y no consideran el apantallamiento.

Climent-Font *et al.* [26] presentan una crítica a las aproximaciones que involucran todos estos programas computacionales, llevando a errores de alrededor del 7% entre los espectros generados con la computadora y los observados experimentalmente.

UN EJEMPLO

A continuación se presenta un ejemplo de un espectro experimental obtenido con el MCA y cuyo ajuste llevé a cabo utilizando el programa RUMP.

De acuerdo con un análisis de la muestra por espectrometría de masas, el material contiene 89.89% de Al, 5.9% de Zn, 2.3% de Mg, 1.5% de Cu, 0.15% de Fe, 0.15% de Cr, 0.09% de Si y 0.02% de Mn. El espectro que se obtiene se presenta a continuación.

De acuerdo con el mejor ajuste usando el programa RUMP, se obtiene la composición atómica que se presenta en la discusión, un poco más adelante.

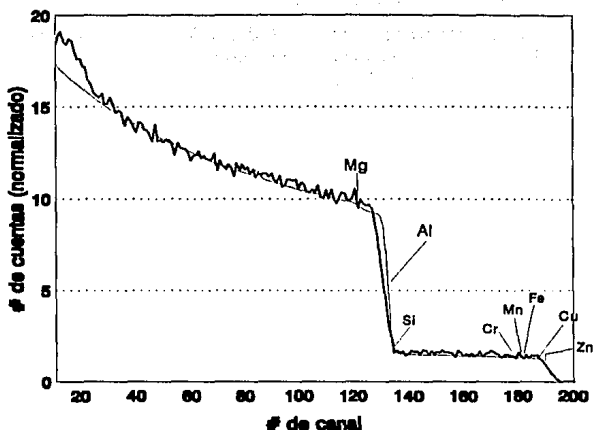


Fig. 28. Espectro experimental (línea irregular) y espectro ajustado con RUMP (línea regular). Se muestra cada uno de los elementos que contiene la muestra y el canal donde aparecen

Un punto importante que hay que notar es el siguiente: el aluminio solo reacciona con el oxígeno del ambiente, lo que provoca un oxidamiento superficial de la muestra. En caso de haber oxidación superficial del aluminio, ésta no se alcanza a detectar debido a su baja concentración relativa y al límite de detección en este espectro.

De acuerdo con las experiencias que se adquirieron al hacer estos ajustes, hay que hacer notar lo siguiente:

Primero, el programa RUMP es muy poco "amigable", en el sentido de que no es claro para el usuario novel (aparte de que el manual no tiene un índice, lo que hace muy complicada su consulta).

Segundo, las rutinas PERT, que se supone se aplican una vez que se ha logrado reproducir "adecuadamente" el espectro experimental, admiten únicamente la modificación automática de ocho variables; sin embargo, si se intenta hacer un ajuste relativamente complejo, el programa despliega un mensaje de error y termina. Si no se ha tenido la precaución de almacenar la información previamente, es posible que se pierda el trabajo realizado.

Tercero, existen dos maneras de almacenar los datos: en formato RBS, que es el usado por RUMP, y en ASCII; además de que hay que tener mucho cuidado de almacenar los datos del espectro simulado en lugar del experimental, si se desea hacer cálculos adicionales con los resultados del ajuste en una hoja electrónica, es conveniente que, la primera vez que se les vaya a dar formato para poder trabajarlos —ya que RUMP los almacena a cuatro columnas; incrementando primero la columna, hasta cuatro y para el quinto canal baja al siguiente renglón, etc.— se diseñe un “macro”, es decir, un pequeño programa que permita el organizar automáticamente los datos, ya que, de otra manera, resulta muy tedioso el proceso.

DISCUSIÓN

El “mejor ajuste” es ambiguo, ya que, si se varían las concentraciones relativas de los elementos en porcentajes que dependen del elemento, según se presenta en el siguiente cuadro, el espectro simulado sigue pareciéndose “bastante” al experimental.

| Elemento | Composición promedio | Error* | Elemento | Composición promedio | Error* |
|----------|----------------------|--------|----------|----------------------|---------|
| Al | 92.31% | ± 0.12 | Mg | 3.69% | ± 0.6 |
| Cr | 0.16% | < 0.5% | Mn | 0.11% | < 0.5% |
| Cu | 0.62% | ± 0.15 | Si | 0.77% | < 1.5% |
| Fe | 0.09% | < 0.5% | Zn | 2.31% | ± 0.15% |

* El error mostrado corresponde a una estimación gráfica del máximo error posible; la desigualdad muestra la estimación del límite de detección.

Hay que hacer notar que el programa es mucho más sensible a las concentraciones relativas de ciertos elementos —sobre todo los más pesados— por lo que, en un espectro como el del ejemplo anterior, la región más compleja de ajustar es la de altas energías. Por tanto, se puede concluir que aún existen problemas por resolver en este campo, y cada vez son más las aproximaciones numéricas que se desarrollan para poder simular los espectros experimentales, incluyendo fenómenos físicos no considerados anteriormente, como *straggling* o apantallamiento y cada vez más se van refinando los métodos de modo que los errores se minimicen. Conforme se vayan entendiendo a mayor profundidad los procesos que intervienen y simulándolos en computadora para abreviar el tedioso trabajo de análisis, se podrá llegar a determinar con mayor exactitud el perfil de concentraciones de diferentes muestras, lo que redundará, finalmente, en un mejor control de

calidad en el proceso de caracterización de diferentes materiales compuestos.

La computadora ha resultado ser una herramienta invaluable para poder entender —y así simular— los diferentes fenómenos físicos involucrados en la espectrometría por RBS, además de simplificar el análisis de los espectros experimentales. Más aún, este tipo de programas de simulación de espectros resultan interesantes para el estudiante, ya que puede ver gráficamente lo que ocurre al hacer cualquier modificación a los distintos parámetros involucrados, por lo que puede aclarar los conceptos.

Adicionalmente, con el fin de mejorar la resolución en estos programas, cotidianamente se actualizan los valores usados para el frenamiento de iones en materia, ya que es la falta de determinación exacta de ellos lo que introduce errores entre los espectros experimentales y teóricos, además de las polémicas que se han desatado al elegir una u otra aproximación para las pérdidas de energía.

Existe un programa, llamado TRIM, desarrollado por Ziegler, que presenta una simulación de cómo se van adentrando los iones en la materia —con la composición atómica que elija el usuario—, hacia donde son dispersados y sus subsecuentes colisiones, todo de manera gráfica. Para poder correr esta simulación, el mismo autor se encarga de ir tomando medidas más refinadas para los frenamientos y modificar los algoritmos, de modo que el programa es cada vez más eficiente. Sin embargo, no genera datos numéricos que puedan ser utilizados en otros programas, por lo que no resulta de gran utilidad, más que con fines ilustrativos. Si se le añadiese la posibilidad de obtener los valores numéricos junto con la demostración gráfica, podría usarse también como subrutina en programas que se hayan desarrollado por otros investigadores, en caso de que los disponibles no lo satisfagan. Esto último es frecuente, debido a que cada quien tiene sus preferencias personales en cuanto a los métodos numéricos y algoritmos.

Generalmente los autores de cada programa son los responsables de su distribución —gratuita— y de su documentación, en forma de artículos. Las bases teóricas y aproximaciones para la simplificación de los cálculos en computadora también se reportan en artículos.

Son importantes las aproximaciones debido a que hacen más veloz el cálculo en computadoras y abrevian el código de programación.

TECNOLOGÍA DE FABRICACIÓN DE CIRCUITOS INTEGRADOS

Usualmente, se usa la llamada "tecnología MOS" (*Metal Oxide Semiconductor*) para crear transistores especiales (como el FET, por ejemplo) y todos los circuitos de gran escala de integración, como son los microprocesadores.

Para crear el circuito, se requiere primero crecer un monocristal de silicio. Después, este cristal se corta en rebanadas circulares, llamadas *wafers*. Sobre estos *wafers* se van a manufacturar los circuitos integrados (*chip*), por un proceso de máscaras fotolitográficas. De hecho, en un solo *wafer* se fabrican varios circuitos iguales, que se evalúan para ver si su funcionamiento es el adecuado y, en caso afirmativo, se cortan, se añaden las terminales eléctricas (*pin*s) y se encapsulan en plástico. Los que no sirvan se desechan.

La manera de producir un circuito semiconductor con el proceso planar (usando las máscaras fotolitográficas) es la siguiente:

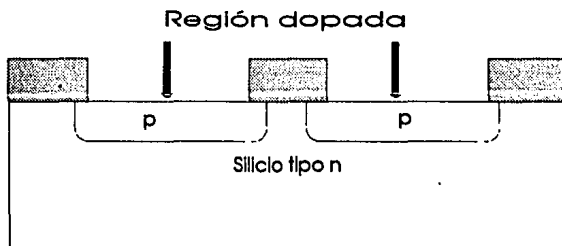
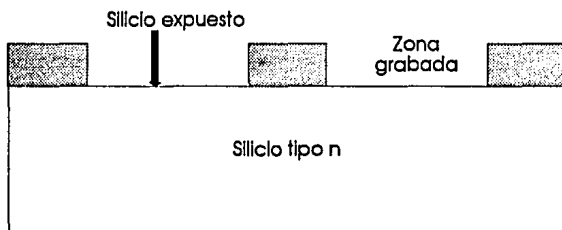
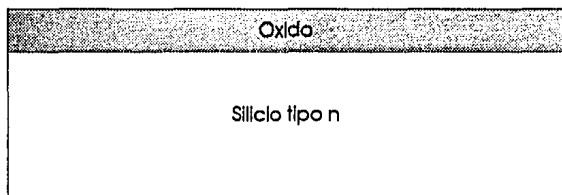
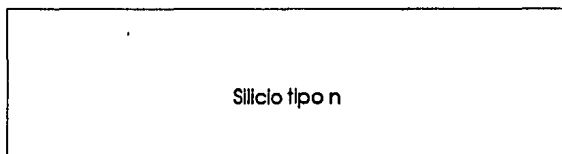
Vamos a tomar como ejemplo la fabricación de un transistor.

Primero, sobre el *wafer* de silicio — de hecho, ya es un cristal tipo *p* o tipo *n* y se distribuyen comercialmente —, se deposita una capa de entre 5,000 y 6,000 Å de óxido de silicio. Se deposita una emulsión fotosensible encima del óxido de silicio y se superpone una máscara para dejar libres las superficies donde se desea contaminar el sustrato; se revela, de modo similar a las fotografías, y se remueve el óxido de las áreas negras (siguiendo

con la analogía fotográfica), con un proceso llamado "grabado" (*etching*). El grabado puede ser de varios tipos: por corrosión química, *sputtering* (erosión iónica) o plasma. Una vez que se deja expuesta la superficie del monocristal, se dopa, ya sea por difusión térmica o por implantación iónica. Después se vuelve a crecer la superficie completa de óxido de silicio, ahora con un grosor de entre 10,000 y 15,000 Å. Se repite el proceso de grabado y dopado. Finalmente, se repite el proceso con una capa de óxido de silicio de alrededor de 1,000 Å, pero la corrosión no es completa; permanece una pequeña capa de óxido de silicio que será el "puente" (*gate*) del transistor. Luego, se protege esta área y se vuelve a grabar, para exponer las regiones de "fuente" y "drenado" (que son las otras dos terminales del transistor). Se metaliza (típicamente con aluminio, aunque se presentan algunos problemas en circuitos de gran escala de integración, debido a su baja temperatura de fusión —aprox. 525°C; en la actualidad también se usa titanio, que tiene menor movilidad, por lo que no se difunde en el recocido) y se le da un tratamiento térmico, para poder hacer el enlace entre el metal y el semiconductor; a este tratamiento se le llama recocido. Estas terminales metálicas son las que permiten la conexión entre el transistor y el resto del circuito integrado o con las terminales externas.

Con ésto, se tiene el wafer con varios transistores, que son probados, cortados —separados del wafer, término que en inglés se llama *chip* y de ahí el nombre genérico con que se conoce actualmente a los circuitos integrados—, se conectan las terminales, se encapsulan, se vuelven a evaluar y, si funcionan, están listos para ser usados.

A continuación se presenta un diagrama de los pasos a seguir para la fabricación de un transistor, y dos ejemplos de diferentes técnicas de dopado, que logran diferentes definiciones en las regiones contaminadas e incluso permiten disminuir capacitancias parásitas que pueden llegar a ser importantes a altas frecuencias. Finalmente, se presentan algunas fotografías de algunos circuitos integrados antes de ser encapsulados.



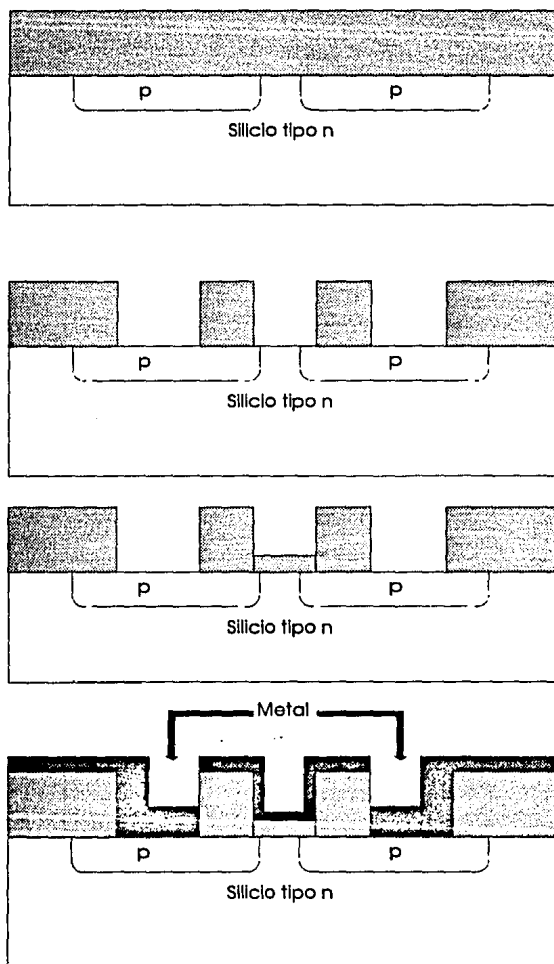


Figura 29. Proceso de elaboración de un transistor de efecto de campo

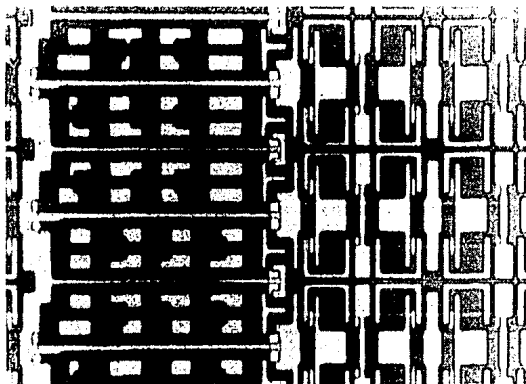


Figura 30. Fotografía de un circuito integrado en escala 50X

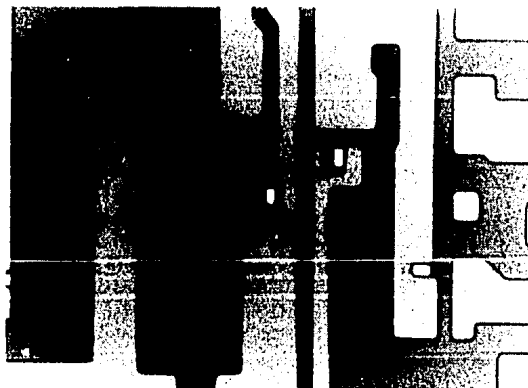


Figura 31. Fotografía de un circuito integrado en escala 100X

La implantación iónica es un proceso con el cual se pueden depositar cantidades controladas de impurezas en la región superficial de un semiconductor para la fabricación de chips. Por supuesto, se puede usar esta misma técnica con otros fines, pero comercialmente, este es el uso más popular. Hasta hace una década, se consideraba a esta técnica como poco eficiente en la fabricación de circuitos integrados. Sin embargo, las necesidades de circuitos que pudiesen operar a mayores velocidades ha llevado al desarrollo y perfeccionamiento de esta técnica para controlar con mayor detalle el dopado. Las ventajas principales de la implantación iónica sobre la difusión térmica son:

1. La concentración de impurezas como función de la profundidad se puede controlar con la energía del ion, por lo que es posible incluso llegar a crear una capa "enterrada" de dopantes.
2. Se puede controlar mejor la cantidad y pureza del material implantado.
3. El proceso ocurre a relativamente bajas temperaturas, por lo que se minimiza la difusión de impurezas no deseadas.
4. Los iones entran a la muestra en forma de haz directo, por lo que se puede lograr buena definición lateral.

Cabe hacer notar que la tecnología de fabricación de circuitos integrados requiere de condiciones ambientales especiales, ya que la atmósfera puede contaminar los chips en el proceso y hacerlos inservibles. Dentro de los cuidados que se debe tener están el contar con "cuartos limpios" y herramientas especiales para su manejo y traslado.

Existen tres tecnologías principales, que son PMOS, cuando los portadores mayoritarios son hoyos y el sustrato es tipo n . NMOS, donde el sustrato es tipo p y los portadores mayoritarios son electrones, lo que hace que sean más rápidos en su respuesta. CMOS, que combina sustratos tipo p y n , tienen un menor consumo de potencia y excelente inmunidad al ruido. La CMOS es la más utilizada en la fabricación de componentes para PC, aunque actualmente se utiliza más el silicio amorfo que el cristalino, que tiene las mismas propiedades eléctricas, pero mejores propiedades mecánicas y es más barato.

CONCLUSIÓN

Con las primeras computadoras personales que aparecieron en el mercado, debido a sus limitaciones en cuanto a capacidad de memoria (en general 64 kb antes de encender, de los cuales quedaban aproximadamente 24 kb libres al cargar BASIC, que era el único lenguaje de programación que manejaban) y capacidad de almacenamiento (el almacenamiento era en cassettes, con acceso secuencial), no tuvieron gran relevancia en la investigación científica —aunque ya empezaba a difundirse su uso para juegos y algunos programas administrativos, generalmente publicados en revistas sobre el tema. Sin embargo, conforme fueron evolucionando, y se desarrollaron procesadores de texto, empezaron a entrar a la comunidad científica, para ahorrar problemas y errores en la captura de los artículos. También empezó su uso como controles sencillos en algunos procedimientos de laboratorio. Más adelante, con la comercialización de tarjetas para PC con diversas aplicaciones a aparatos de medición, tuvieron más difusión, hasta llegar a hoy en día, en que un investigador, desde su terminal, se puede comunicar a través del "correo electrónico" con cualquier parte del mundo, solicitar información y acceder a bases de datos sobre casi cualquier campo del quehacer humano.

Más aún, las primeras PC eran lentas (4.7 MHz), por lo que un proceso complejo tomaba días (o hasta semanas) de procesamiento ininterrumpido. Ahora se cuenta con computadoras con velocidades de 66 MHz —y promesa de llegar en un futuro cercano a 100 MHz—, que equivale a ser aproximadamente 14 veces más rápidas; sin embargo, en realidad son aún más veloces, ya que se ha mejorado considerablemente el manejo de los códigos de programación internos y toda la lógica asociada.

Cada vez resulta más necesario el estar familiarizado con las computadoras, ya que están ocupando lugares importantes en nuestra vida cotidiana, no sólo en áreas especializadas.

Más aún, conforme se ha ido diversificando el uso de las computadoras personales, cada día crece la necesidad de ellas y no sólo eso, sino que cada vez son más los microprocesadores que se ofrecen en el mercado, creando gran confusión entre los usuarios. Por ejemplo, existen procesadores con nombres que aparentemente los hacen equivalentes al 486 de Intel, y en realidad su funcionamiento está más cercano al 386, otros trabajan a velocidades no estándar (40 MHz), etc. Debido a que el líder en los desarrollos de procesadores para computadoras personales tipo PC ha sido Intel —que a fines de enero de 1994 anunció la aparición del nuevo procesador, llamado P6 para fines de 1995 o principios de 1996— es conveniente usarlo como referencia.

Otra aplicación de las PC que cada vez está más difundida, es el uso de "simuladores", utilizados en la capacitación de personal que requiere habilidades particulares, como puede ser la operación de una planta hidroeléctrica o nuclear, pilotos de aviones, etc. En estos casos, es muy importante la velocidad a la cual opera el sistema, ya que, de presentarse una "emergencia", debido a condiciones críticas ocasionadas por fallas simuladas o por operación inadecuada de parámetros por parte del usuario, lo que se debe tomar en cuenta es la velocidad de respuesta de la persona que se está capacitando, evitando rezagos ocasionados por los tiempos que tome a la computadora el procesar la información.

La computadora y la investigación científica han ido apoyándose mutuamente, ya que, conforme se pueden ir entendiendo a mayor profundidad los fenómenos físicos, se desarrolla tecnología cada vez más moderna, que permite mayor escala de integración y velocidad en las componentes que forman la computadora, además de desarrollar nuevos productos, como unidades de almacenamiento, comunicaciones, etc. Y, ¿cuánto tiempo se tardaría un investigador en llevar a cabo sus cálculos sin las computadoras? Actualmente, conforme evoluciona la tecnología, se van refinando los aparatos de medición y análisis, lo que permite tener cada vez medidas más refinadas. Un ejemplo claro de esto son los datos experimentales para frenamientos de iones en materia. Mientras menos error se tenga en estas medidas, se podrá ir considerando otro tipo de fenómenos y determinar su relevancia para el análisis de películas delgadas.

Es por todo lo anterior que se escribió el presente texto, para dar un panorama general de lo que han sido y son, hoy en día, las computadoras personales, ya que el mundo del futuro —y del no tan futuro— pertenece a la computadora personal y a todas aquellas personas que sean capaces de utilizarla de manera eficiente (no se olvide que, a nivel mundial, el criterio primordial dentro de cualquier empresa para evaluar a los empleados —o a sí misma— es la llamada y nunca bien definida “productividad”, que versa sobre el manejo racional y eficiente de los recursos.

De hecho, existen más de dos millones de computadoras personales tipo PC en uso actualmente, que trabajan con el sistema operativo DOS, mientras que la cultura computacional no ha proliferado en la misma escala, debido en parte al tipo de textos que existen en el mercado, que son o muy básicos —es decir, un usuario con un poco de curiosidad puede llegar mucho más allá que lo que permiten este tipo de textos— o bien entran en la categoría de “avanzados”, que son casi incomprensibles, aún para el usuario experto. Con esta tesis espero subsanar las carencias antes mencionadas de textos para el público en general, y para el físico en particular, interesados en conocer a fondo (o someramente) su computadora, y poder explotarla al máximo.

No cabe duda de que, cada vez más, es indispensable tener una cultura computacional y sobre todo en las áreas científicas, ya que los estudiantes de hoy en día tendrán que hacer uso de ellas en su vida profesional —independientemente de su campo de trabajo.

REFERENCIAS

- [1] J. RICKARDS, R. TREJO-LUNA, M. E. ORTIZ, E. ANDRADE, E. CHAVEZ, E. P. ZIRONI, H. del CASTILLO y M. SANCHEZ. *Mat. Sci. and Technol.* **9** (1993) 536.
- [2] S. RUBIN. *Phys. Rev.* **78**, (1950) 83.
- [3] E. RUTHERFORD. *Phil. Mag.* **21** (1911) 669.
- [4] C. G. DARWIN, *Phil. Mag.* **28** (1914) 499.
- [5] G. DEARNALEY. *IAEA-SMR 15/44* (1975) 167.
- [6] J.F. ZIEGLER, J. P. BIRSACK y U. LITTMARK. **The Stopping and Range of Ions in Solids** (Pergamon Press, New York, 1985).
- [7] E. C. MONTENEGRO, S. CRUZ y C. VARGAS-ABURTO. *Phys. Lett.* **92A** (1982) 195.
- [8] J. E. E. BAGLIN y J. S. WILLIAMS, "High Energy Ion Scattering Spectrometry", en J. R. BIRD y J. S. WILLIAMS, eds. **Ion Beams for Materials Analysis**. Academic Press, New York, 1989.
- [9] P. BØRGESEN, R. BEHRISCH y M. M. U. SCHERZER. *App. Phys* **A27**, (1982) 183.
- [10] W. K. CHU, J. W. MAYER y M-A. NICOLET. **Backscattering Spectrometry**. Academic Press, New York, 1978.
- [11] W. K. CHU y J. F. ZIEGLER. *App. Phys.* **42**, (1971) 2031.

- [12] W. LI y Z. AL-TAMIMI. Nucl. Instr. and Meth. **B15**, (1986) 241.
- [13] K. MICHAELIAN y E. ANDRADE. Nucl. Instr. and Meth. **B74**, (1993) 443.
- [14] D. POWERS y W. WHALING, Phys. Rev. **126**, (1962) 61.
- [15] P. A. SANDERS y J. F. ZIEGLER. Nucl. Instr. and Meth. **218**, (1983) 67.
- [16] L. C. FELDMAN y J. W. MAYER. **Fundamentals of Surface and Thin Film Analysis**. North Holland, New York, 1986.
- [17] J. F. ZIEGLER y J. E. E. BAGLIN, App. Phys. **42**, (1971) 2031.
- [18] J. RICKARDS, A. OLIVER, J. MIRANDA y E. P. ZIRONI. App. Surf. Sci. **45** (1990) 155.
- [19] J. F. JANNI. Atomic Data Nucl. Data Tables **27** (1982) 147.
- [20] L. C. NORTHCLIFF y R. F. SCHILLING. Nucl. Data Tables **A7** (1970) 233.
- [21] J. F. ZIEGLER. **Handbook of Stopping Cross Sections for Energetic Ions in all Elements** (Pergamon Press, New York, 1980).
- [22] J. F. ZIEGLER, R. F. LEVER y J. K. HORVONEN. "Computer analysis of Nuclear Backscattering", en **Ion Beam Surface Layer Analysis**, O. Meyer, G. Linker y F. Käppeler eds, Plenum Press, New York (1976) Vol. I, p.
- [23] L. R. DOOLITTLE. Nucl. Instr. and Meth. **B9**, (1985) 344.
- [24] L. R. DOOLITTLE. Nucl. Instr. and Meth. **B15**, (1986) 227.
- [25] J.W. BUTTLER. Nucl. Instr. and Meth. **B15**, (1986) 232.
- [26] A. CLIMENT-FONT, U. WÄTJEN y H. BAX. Nucl. Instr. and Meth. **B71** (1992) 81.

APÉNDICES

APÉNDICE I

ARITMÉTICA HEXADECIMAL

La computadora es una máquina "digital", es decir, trabaja con números binarios (o dígitos binarios = *bits*, que viene del inglés binary digit; un *byte* es un conjunto de 8 bits); los únicos valores admitidos en sistema binario, también llamado base 2, son 0 y 1. Esto se debe a que, desde que se desarrollaron los "circuitos lógicos" (que trabajan en binario), y conforme se pudieron ir integrando más y más de estos circuitos dentro de un "circuito integrado" o chip, las computadoras fueron reduciendo su tamaño, hasta que aparecieron las computadoras personales. Las primeras máquinas digitales ocupaban cuartos enteros mientras que las computadoras personales ocupan ya sólo una porción de un escritorio. A raíz de este desarrollo tecnológico —empezando por la invención de los chips— y usando el "álgebra booleana" (un campo especial de las matemáticas), que comprende el sistema binario y algunas operaciones "lógicas", las computadoras se desarrollan de manera acelerada.

La técnica de "codificación" de información se basa en este sistema binario, ya que los circuitos que la conforman sólo pueden usar este sistema.

La memoria de la computadora, que es la parte de donde se leen todas las instrucciones que tiene que ejecutar, está formada por una serie de "interruptores". Cada interruptor corresponde a un bit.

La codificación de una instrucción consiste en asignarle un número, traducirlo a binario y en función de éste, encender o apagar una serie de

interruptores. Si un interruptor está encendido, se dice que tiene un valor de 1 lógico, y si está apagado vale 0. Una instrucción o dato puede ocupar hasta 16 dígitos.

Los números hexadecimales se construyen en una base de 16 (es decir, hay 16 dígitos diferentes: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E y F), igual que los números decimales se construyen en una base de diez dígitos (del cero al nueve). Usaremos indistintamente las palabras hexadecimal y hex (para abreviar). En notación hex se usan los números del cero al nueve para representar valores del cero al nueve, y los restantes seis valores, del 10 al 15, se representan con los símbolos de la A a la F (siempre se escriben con mayúsculas).

En el sistema decimal, la posición de un dígito en un número dado le da un cierto "peso", como por ejemplo en el símbolo 123 el 1 tiene un peso de centenas ($1 \cdot 10 \cdot 10$), el dos de decenas ($2 \cdot 10$) y el tres de unidades ($3 \cdot 1$), de manera que $1 \cdot 10 \cdot 10 + 2 \cdot 10 + 1 = 123$.

Análogamente, en sistema hex, el símbolo 123 significa $1 \cdot 16 \cdot 16 + 2 \cdot 16 + 3 = 291$ decimal

Para distinguir entre números decimales y hex, usualmente se escribe el número y a continuación una H.

El uso de los números hexadecimales en computación surge de manera casi natural, ya que una computadora solamente es capaz de trabajar con números binarios, pero el representar un valor en binario resulta muy extenso y se presta a muchos más errores que el usar una notación compacta. Cada número hex representa cuatro dígitos binarios. Véase la siguiente tabla:

| Hex | Bits | | | | Hex | Bits | | | | Hex | Bits | | | | Hex | Bits | | | |
|-----|------|---|---|---|-----|------|---|---|---|-----|------|---|---|---|-----|------|---|---|---|
| | 4 | 3 | 2 | 1 | | 4 | 3 | 2 | 1 | | 4 | 3 | 2 | 1 | | 4 | 3 | 2 | 1 |
| 0 | 0 | 0 | 0 | 0 | 4 | 0 | 1 | 0 | 0 | 8 | 1 | 0 | 0 | 0 | 12 | 1 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 1 | 5 | 0 | 1 | 0 | 1 | 9 | 1 | 0 | 0 | 1 | 13 | 1 | 1 | 0 | 1 |
| 2 | 0 | 0 | 1 | 0 | 6 | 0 | 1 | 1 | 0 | 10 | 1 | 0 | 1 | 0 | 14 | 1 | 1 | 1 | 0 |
| 3 | 0 | 0 | 1 | 1 | 7 | 0 | 1 | 1 | 1 | 11 | 1 | 0 | 1 | 1 | 15 | 1 | 1 | 1 | 1 |

Para convertir un número de hex a decimal hay que asignar a cada dígito su valor y multiplicarlo por 16 tantas veces como dígitos haya a su derecha y luego sumarlos todos. Por ejemplo el número AF2D hex vale

$$(10 * 16 * 16 * 16) + (15 * 16 * 16) + (2 * 16) + (13) = \\ 40,960 + 3,840 + 32 + 13 = \\ 44,845$$

ya que A = 10 y hay tres dígitos a su derecha, F = 15 y hay 2 dígitos a su derecha, 2 = 2 y hay un dígito a su derecha, y D = 13

Para convertir un número de decimal a hex el procedimiento es un poco más complicado de describir:

1. Se divide el número entre 16. Luego, se toma la parte entera (quitando los dígitos después del punto decimal) y se multiplica por 16 y se resta esta cantidad del número original. El residuo de la resta es el dígito menos significativo (del extremo derecho)

Veamos un ejemplo:

$$44,845 / 16 = 2802.8$$

tomamos la parte entera, es decir 2802, y lo multiplicamos por 16:

$$2802 * 16 = 44,832$$

lo restamos del número original:

$$44,845 - 44,832 = 13$$

es el dígito de la derecha, que en hex equivale a D.

2. Se repite el mismo procedimiento, tomando ahora la parte entera de la división inicial del paso 1 como el número original:

$$2802 / 16 = 175.1$$

tomamos la parte entera, es decir 175, y lo multiplicamos por 16:

$$175 * 16 = 2800$$

lo restamos del número original:

$$2802 - 2800 = 2$$

es el segundo dígito de derecha a izquierda. Llevamos ya el número 2D.

Volvemos a tomar ahora 175 como número original y repetimos el proceso:

$$175 / 16 = 10.9$$

$$10 * 16 = 160$$

$$175 - 160 = 15$$

y 15 equivale a F en hex. Tenemos ya el número F2D.

Otra vez repetimos el proceso:

$$10 / 16 = 0.6$$

$$0 * 16 = 0$$

$$10 - 0 = 10$$

que equivale a A hex. Como el número original ya llegó a cero, se termina el proceso y el número hex resultante es AF2D (como era de esperarse).

APÉNDICE II

CARACTERES

Se entiende por caracter todo aquel símbolo que es posible desplegar en pantalla en modo texto. Estos símbolos pueden ser caracteres alfanuméricos, que comprenden todas las letras, números y signos de puntuación, como caracteres semigráficos —aquéllos con forma de líneas o sombreados—, caracteres de control —generalmente se despliegan como caritas, corazoncitos, etc. Los procesadores de texto y hojas electrónicas, entre otros programas, permiten el despliegue de caracteres con atributos, como pueden ser negritas, cursivas, subrayados, diferentes tamaños o tipos, letras griegas, símbolos de marca registrada o *copyright*, símbolos musicales o de gografía, entre otros. La manera en que la computadora forma todos estos caracteres es mediante el uso de los caracteres de control, que indican a la computadora que, al aparecer uno de ellos, el despliegue de los caracteres que siguen no sea el estándar, sino que su apariencia en pantalla debe ser modificada. Sin embargo, todos los caracteres se adaptan a un estándar, llamado ASCII (*American Standard Code for Information Interchange*), que no es más que una manera de codificar los símbolos disponibles a través del teclado.

Con la popularización de la computadora en todo el mundo, se ha hecho necesario añadir cada vez más caracteres, por lo que se está trabajando en el desarrollo de nuevos estándares, codificados en dos bytes en lugar de uno —como es el caso de ASCII— para poder integrar caracteres árabes, chino o hebreos. Sin embargo, ASCII sigue siendo el estándar mundial.

CARACTERES ASCII

Existen 256 caracteres diferentes en DOS, cada uno representado por un número entre cero y 255 o entre cero y FF hex. Cada carácter ocupa un byte. Los primeros 128 caracteres corresponden a los llamados "caracteres ASCII estándar", que requiere únicamente de siete bits para su codificación. Los 128 caracteres restantes forman el conjunto de "caracteres ASCII extendidos", y se forman al añadir el octavo bit, de modo que se requiere de un byte completo para codificar cada uno de estos caracteres. En el grupo de caracteres ASCII extendido se encuentran las letras requeridas por lenguajes diferentes al inglés, como son el alemán, francés o español. A continuación se presenta una tabla de los caracteres ASCII.

| Decimal Value | • | 0 | 16 | 32 | 48 | 64 | 80 | 96 | 112 | 128 | 144 | 160 | 176 | 192 | 208 | 224 | 240 |
|---------------|-----------|---|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| • | Hex Value | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
| 0 | 0 | | ▶ | | 0 | @ | P | ' | p | Ç | É | á | | ¸ | ¸ | ¸ | ≡ |
| 1 | 1 | ☺ | ◀ | ! | 1 | A | Q | α | q | ü | æ | í | î | ¸ | ¸ | ¸ | ± |
| 2 | 2 | ☹ | ! | " | 2 | B | R | b | r | ó | Æ | ó | ¸ | ¸ | ¸ | ¸ | ≥ |
| 3 | 3 | ♥ | !! | # | 3 | C | S | c | s | á | ó | ú | ¸ | ¸ | ¸ | ¸ | ≤ |
| 4 | 4 | ♦ | ! | \$ | 4 | D | T | d | t | ä | ö | ü | ¸ | ¸ | ¸ | ¸ | ¸ |
| 5 | 5 | ♣ | ! | % | 5 | E | U | e | u | á | ó | Ë | ¸ | ¸ | ¸ | ¸ | ¸ |
| 6 | 6 | ♠ | = | & | 6 | F | V | f | v | á | ú | ¸ | ¸ | ¸ | ¸ | ¸ | + |
| 7 | 7 | • | ! | ' | 7 | G | W | g | w | ç | ù | ¸ | ¸ | ¸ | ¸ | ¸ | ¸ |
| 8 | 8 | ◼ | ! | (| 8 | H | X | h | x | é | y | ¸ | ¸ | ¸ | ¸ | ¸ | ° |
| 9 | 9 | ◊ | ! |) | 9 | I | Y | i | y | ë | Ö | ¸ | ¸ | ¸ | ¸ | ¸ | • |
| 10 | A | ◼ | - | . | : | J | Z | j | z | è | Û | ¸ | ¸ | ¸ | ¸ | ¸ | • |
| 11 | B | ♂ | - | + | : | K | [| k |] | í | ¸ | ¸ | ¸ | ¸ | ¸ | ¸ | ✓ |
| 12 | C | ♀ | ¸ | . | < | L | \ | l | ! | í | ¸ | ¸ | ¸ | ¸ | ¸ | ¸ | ¸ |
| 13 | D | ♫ | ↔ | - | = | M | | m | | í | v | ¸ | ¸ | ¸ | ¸ | ¸ | ? |
| 14 | E | ♫ | ▲ | . | > | N | ~ | n | ~ | À | ¸ | ¸ | ¸ | ¸ | ¸ | ¸ | ◼ |
| 15 | F | ♫ | ▼ | / | ? | O | - | o | ◊ | À | ¸ | ¸ | ¸ | ¸ | ¸ | ¸ | ◼ |

Los primeros 32 caracteres ASCII tienen dos funciones: primero, se usan para control de impresora y de comunicación; y segundo, forman un conjunto especial de caracteres, que incluyen caras, corazones y rombos, entre otros. En general, estos caracteres se programan en función de las necesidades del programa que se esté usando y lo que se despliega en pantalla puede variar dependiendo del lenguaje o programa que se use.

Formato de los archivos de textos

Muchos programas — como los procesadores de texto, hojas electrónicas o sistemas de diseño por computadora — trabajan con archivos de datos y texto. Como resultado, la mayor parte de ellos han adoptado una convención para que sea más fácil usar un archivo de datos con varios programas¹. Los formatos se definen por caracteres embebidos en el texto y que llevan a cabo ciertas funciones (como cambiar de renglón una vez que se excede cierto márgen).

Los archivos en “formato ASCII” contienen únicamente el conjunto estándar de caracteres ASCII y, en este contexto, los primeros 32 caracteres ASCII (del 0 al 31) no tienen funciones especiales; en cambio, cuando se está trabajando con procesadores de texto que permiten dar atributos al texto, los primeros 32 caracteres ASCII sí tienen una función específica, definida por el programa que se esté usando: algunos de ellos se usan para dar formato al texto y los otros para controlar la impresión del texto. No se despliegan en pantalla.

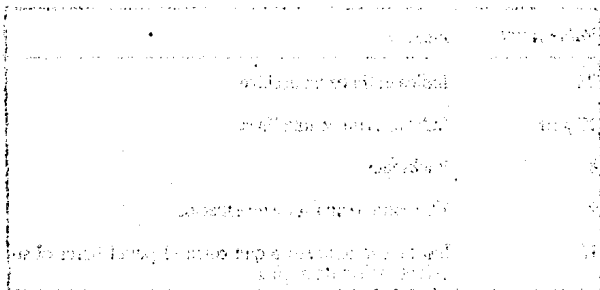
Algunos ejemplos son los siguientes:

| <i>Valor ASCII</i> | <i>Función</i> |
|--------------------|------------------------------------------------------------------------------|
| 26 | Indica el fin de un archivo |
| 13 y 10 | Indican el fin de una línea |
| 8 | Backspace |
| 9 | Tabulador o una serie de espacios |
| 12 | Indica a la impresora que corra el papel hasta el siguiente inicio de página |

Los procesadores de texto usan los caracteres de control de forma diferente e incluso usan el conjunto ASCII extendido, algunos de los caracteres son para hacer “dibujos” (solamente líneas y cuadros, ya sea con línea sencilla y doble) y los demás son caracteres de control.

¹ A raíz de esto, ha surgido el término “hipertexto”, cuya idea fundamental es el poder trasladar de manera transparente para el usuario, los archivos entre diferentes tipos de programas, como puede ser el tener un cuadro en Excel y trasladarlo a WinWord sin tener que hacer más que marcar el área deseada en Excel, dar un comando a la computadora, entrar a WinWord, ubicar el cursor en la posición donde se desea introducir el cuadro y dar otro comando.

Es importante recalcar que aca programa codifica de manera distinta los atributos de texto, por lo que, si se desea traducir, por ejemplo, un archivo de **Word** a **WordPerfect**, se encontrará que se han perdido algunos de los atributos. Por tanto, es conveniente usar las utilerías de conversión que se distribuyen junto con los programas. Algunos paquetes, como Excel, permiten la importación directa de archivos creados con Lotus u otras hojas electrónicas, o bien almacenar los archivos de Excel en otros formatos. Todo ésto ha tenido mayor difusión con el surgimiento de Windows y toda la paquetería que trabaja bajo este ambient — y también es una de las causas por las que este ambiente ha sido tan popular: permite el uso de archivos antiguos con los nuevos programas.



El texto que sigue es una descripción de un producto o servicio, pero es demasiado tenue para transcribirlo con precisión. Parece describir características técnicas o de uso.

Este texto también es muy tenue y parece ser una continuación de la descripción o un comentario adicional. No se puede leer con claridad.

APÉNDICE III

BOOTING

Es el nombre que recibe el proceso que se lleva a cabo cuando se enciende una computadora.

Inicialmente, se ejecuta un programa llamado BIOS (*Basic Input/Output System*), que se encarga de todo lo que se refiere a entrada y salida de información hacia y desde los periféricos y prepara a la computadora para su operación. Para ésto, primero se hace una autoprueba (POST = *Power-On Self-Test*), para verificar que la memoria esté funcionando adecuadamente e inicializar los valores para los vectores de interrupción; luego, tiene que verificar y activar los dispositivos y periféricos existentes (*polling*), de manera que después se pueda trabajar con ellos.

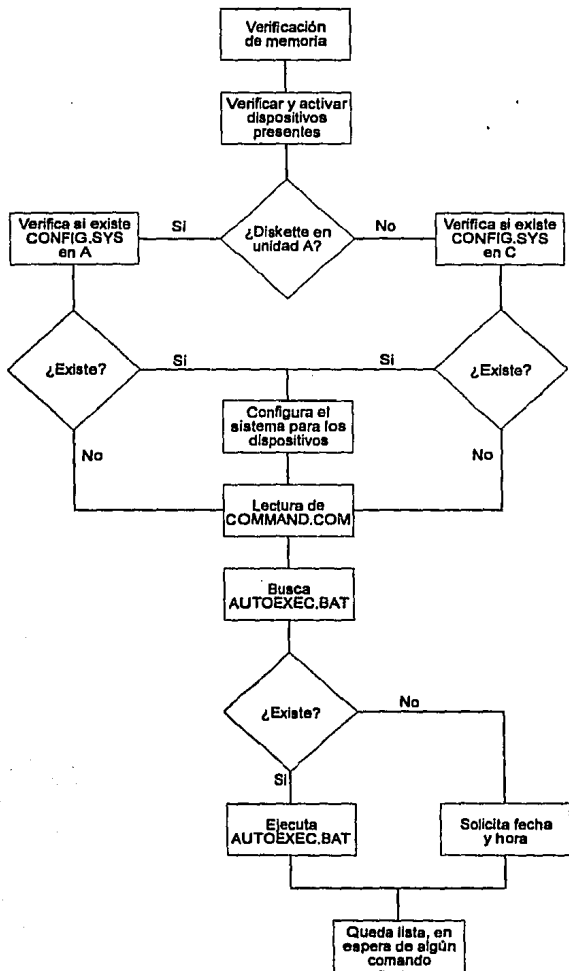
Hagamos un paréntesis aquí para destacar la importancia del BIOS. Toda la filosofía del diseño de las computadoras personales se centra alrededor de las rutinas de servicio del BIOS, que son programas que proveen de todas las funciones de control y operaciones que se consideran necesarias (por los diseñadores, por supuesto). BIOS trabaja directamente con el *hardware* y los dispositivos periféricos, llevando a cabo algunas de las tareas más fundamentales, como el leer o escribir datos en pantalla o disco. La filosofía básica es: dejar que BIOS lo haga; que el programador no tenga que lidiar con el control directo del *hardware*. Lo único que no provee BIOS es la generación de sonido. Por ejemplo, todos los procesadores de texto recurren a la rutina de BIOS para imprimir un documento (con el comando interno **print**), en lugar de desarrollar su propio programa que cumpla con la misma función. Es importante destacar que, al desarrollarse un nuevo

programa, se recomienda usar todos los servicios y funciones tanto de BIOS como de DOS que sean necesarios, ya que el desarrollar rutinas equivalentes puede resultar tanto en una pérdida de tiempo para el programador como en una posible pérdida de compatibilidad, es decir, corre el riesgo de que el programa no funcione más que en una marca específica de computadora PC. Los diseñadores de los sistemas operativos (que actualmente son IBM, MicroSoft y Digital), están obligados a mantener las mismas llamadas a las rutinas básicas de BIOS —rutinas, funciones e interrupciones documentadas. Aunque se cuenta con funciones no documentadas que puede usar el programador, éstas pueden modificarse o desaparecer en versiones posteriores, por lo que no se recomienda el uso de ellas.

Después, carga el sistema operativo, para lo cual checa si existe un diskette en la unidad A. En caso afirmativo, desde ahí lee los archivos IO.SYS, MSDOS.SYS (que son archivos a los que no se tiene acceso directo, por protección) y COMMAND.COM (que es el programa que permite la comunicación entre la máquina y el usuario, como control de video, teclado, unidades de disco flexibles y duros, puertos serie y paralelo, y los comandos internos, además de verificar que todos los comandos se ejecuten en forma adecuada y, en caso de no ser así, desplegar los mensajes de error pertinentes). De no haber un diskette en A, transfiere el control al disco duro (unidad C) y reinicia la búsqueda de estos archivos. Luego carga el archivo CONFIG.SYS, donde se encuentra la información referente a los dispositivos que va a usar el sistema y configuraciones adicionales, tanto de memoria expandida como buffers, etcétera.

Finalmente, busca el archivo AUTOEXEC.BAT. Si no existe, la computadora pide al usuario la fecha y hora. En caso de estar presente, ejecuta los comandos de este archivo.

A continuación se presenta un diagrama de flujo que muestra este proceso en detalle, desde que se enciende la máquina hasta que está lista para recibir alguna instrucción por el teclado.



BIBLIOGRAFÍA

ACER 710 User's Guide, Acer, 1985, s.p.

Este manual contiene los fundamentos para poder conectar y configurar adecuadamente un sistema tipo XT. Describe en forma detallada el funcionamiento del teclado.

Ayala San Martín, Gerardo (coord.). *Computación I. Introducción a la computación*, 2a ed., UNAM-Porrúa, México, 1989, 90 pp.

Libro de texto para el curso de "Introducción a la computación" que imparte la Dirección General de Servicios de Cómputo Académico de la UNAM. Presenta de manera detallada la estructura y funcionamiento de la computadora digital y *software* asociado. También incluye una introducción al procesador de textos **WordStar**.

Houpls, Constantine H. y Gary B. Lamont. *Digital Control Systems. Theory, Hardware, Software*, McGraw-Hill Book Company, New York, 1985, 667 pp.

Libro de texto sobre teoría de control que enfatiza el uso de la computadora digital como controlador. La primera parte presenta una introducción a la arquitectura de computadoras.

DeVoney, Chris with Norman Hale, *DOS Tips, Tricks, and Traps*, Que Corporation, USA, 1989, 521 pp.

Presenta una serie de conceptos avanzados sobre la memoria de una PC y sus direccionamientos; da algunas técnicas para el manejo de discos duros

y diskettes; desarrollo de archivos por lotes; configuración de sistemas; manejo de los drivers de video y *troubleshooting* tanto del hardware como de problemas lógicos.

Holtzman, Jeff. "Computer bits", en *Popular Electronics*, vol. 8, núm 8, agosto de 1991, pp. 70-71.

En este artículo se discuten someramente las diferencias entre los diferentes modelos de computadoras y la actualización de equipos viejos.

Jamsa, Kris. *DOS. Guía para usuarios expertos*, McGraw-Hill, México, 1988, 864 pp.

Este es un libro dedicado a comparar DOS y OS/2. También incluye técnicas para programación de archivos por lotes y un panorama general sobre OS/2.

Lima Tobon, Bertha. *Introducción a la Computadora*, CIESS-División de Informática, México, 1988, 74 pp.

Notas de un curso de sistema operativo MS-DOS, que se puede usar también como tutor: "Tiene como objetivo orientar a los futuros usuarios de micro-computadoras que comienzan su aprendizaje en el conocimiento, historia, funcionamiento y operación de la computadora personal basada en el Sistema Operativo MS-DOS". Incluye un apartado sobre programación estructurada y otros sobre *Basic* y *Lotus 1-2-3*.

McCawley, Mike, Jeanette Gibbson and Alex Schmidt. *386SX User's manual*, SWAN Technologies, USA, 1989-1990.

Este manual presenta de manera muy detallada el tipo de discos duros y de tarjetas de video existentes y sus características. También contiene información general sobre el proceso de encendido de una computadora. Va presentando un glosario para cada tema a tratar, por lo que el contenido de cada capítulo se hace más claro para el lector que no domina el lenguaje técnico.

Microsoft *MS-DOS User's Guide and User's Reference*. Phoenix Computer Products Inc., USA, 1987-1988, 64 y 430 pp.

Manual que describe el modo de operación de los comandos del sistema operativo. Incluye una sección sobre el archivo AUTOEXEC.BAT, otra sobre cómo configurar el sistema y otros tópicos avanzados.

Microsoft MS-DOS Shell User's Guide. Phoenix Computer Products Inc., USA, 118 pp.

El manual de Shell es un tutor para introducir al usuario a su manejo. No presenta un orden lógico adecuado, por lo que, para el principiante, es recomendable seguir el texto paso por paso.

Microsoft MS-DOS User's Guide and User's Reference, Microsoft Press, USA, 1987, 398 pp.

En este manual se presentan todos los comandos de DOS y una introducción a la programación de archivos por lotes.

Morgan, Christopher L. y Mitchell Waite, *Introducción al microprocesador 8086/8088 (16 bit)*, McGraw-Hill, México, 1984, 339 pp.

Un texto que presenta en forma detallada la estructura de una microcomputadora PC tipo XT. Incluye un curso de ensamblador para el procesador 8086.

Murray, William H. y Chris H. Pappas. *80386/80286 Programación en lenguaje ensamblador*, Osborne-McGraw-Hill, México, 1987, 546 pp.

El autor plantea tres objetivos: "(1) introducirle en el mundo potente de la programación en lenguaje ensamblador, utilizando los microprocesadores 80386/80286 y coprocesadores 80387/80287; (2) enseñarle cómo escribir sencillos y avanzados programas en lenguaje ensamblador, y (3) servirle como libro de referencia de instrucciones de programación y programas ejemplos".

Estos son los microprocesadores que usan las computadoras PC AT y sistemas PS2.

Norton, Peter. *The Peter Norton Programmer's Guide to the IBM PC*, Microsoft Press, USA, 1985, 426 pp.

Este libro es útil tanto para el principiante como para el programador avanzado. Aquí se presentan todos los dispositivos que se pueden manejar a través de DOS y las interrupciones y servicios de BIOS para el desarrollo de programas de aplicación.

Norton, Peter. *The hard disk companion.* Brady Books, Prentice Hall, New York, 1992, 325 pp.

Libro que presenta con todo detalle el funcionamiento de un disco duro, tecnologías, instalación, etc.

Norton, Peter. *Inside the IBM PC and the PS/2.* Brady Books, Prentice Hall, New York, 1991, 398 pp.

En este texto se presenta con detalle todo lo que se refiere al hardware de la computadora PC: teclado, monitores, unidades de almacenamiento y la manera en que la computadora maneja los archivos: números binarios y hexadecimales y caracteres alfanuméricos.

Norton, Peter. *Norton Peter. The problem solver.* Brady Books, Prentice Hall, new York, 1991, 268 pp.

Fundamentalmente, es un texto con tutoriales pequeños sobre los temas que se debe conocer si se desea operar una computadora personal de manera eficiente. Básicamente, incluye temas relacionados a sistema operativo (hasta la versión 5.0) y **Windows**, que recientemente ha venido a "sustituir" a DOS para el usuario inexperto.

Sheldon, Kenneth M. "ASCII goes global", en **BYTE**, vol. 16, núm. 7, julio de 1991, pp. 108-116.

Artículo que plantea las necesidades de un nuevo código que supla al código ASCII para agregar todos los símbolos que requiere el mundo actual.

Varios Autores. *Actualización en Informática*, ILCE-SEP/Telesecundaria, México, s.p.i.

En este libro se presenta en forma esquemática el funcionamiento y las aplicaciones de una computadora así como las técnicas para desarrollar programas educativos de diferentes áreas.

Sheldon, Thomas. *PC-DOS y MS-DOS*, Byte Books/MGraw-Hill, México, 1985, 386 pp.

Texto introductorio, en forma de tutor, para introducir al usuario novel al uso básico de DOS. Incluye hasta la versión 3.0.

Udel, Jon. "The all-in-one DOS", en *BYTE*, vol. 16, núm. 7, julio de 1991, pp. 36-40.

Discute las ventajas y desventajas del sistema operativo DOS versión 5.0.

Veale, David S. "The DOS version that's right for you", en *DOS Resource Guide*, sin fecha, pp. 8-11.

En este artículo se presenta una historia muy general sobre el sistema operativo DOS, desde la versión 1.0 hasta la 4.01.

Wakerly, John F. *Microcomputer Architecture and Programming*, John Wiley & Sons, USA, 1981, 692 pp.

Es una introducción a los conceptos básicos subyacentes en el diseño de una computadora. Incluye la arquitectura de algunos de los microprocesadores de propósito general más populares.

Worman, John. "Microsoft's MS-DOS 5.0", en *Computer Craft*, vol. 1, núm 5, agosto de 1991, pp. 26-32.

Discute las ventajas y desventajas del sistema operativo DOS versión 5.0.