

4 ENTORNO DE LOS AUTÓMATAS PROGRAMABLES

En esta cuarta parte se estudia el entorno de los autómatas programable en sus distintos aspectos.

El capítulo 7 está dedicado a los sensores industriales que suministran información del entorno a los sistemas electrónicos de control. En él se estudian los parámetros comunes a los diferentes tipos de sensores y se hace especial énfasis en los sensores detectores de objetos y de medida de distancias, que se utilizan en los procesos de fabricación (*Manufacturing processes*).

El capítulo 8 se dedica a los interfaces de entrada y salida, tanto de conexión de los sistemas electrónicos de control con los procesos como con los usuarios. Especial énfasis se hace en los equipos de interfaz usuario-máquina conocidos actualmente por las siglas inglesas HMI (*Human Machine Interface*) y los sistemas de supervisión y adquisición de datos conocidos por las siglas SCADA (*Supervisory Control And Data Acquisition*).

En el capítulo 9 se describen primero los diferentes conceptos asociados con la automatización de los procesos productivos tanto continuos (*Processes*) como de fabricación (*Manufacturing processes*) y el papel que juega el computador en ellos. A partir de dichos conceptos se describe la pirámide de la fabricación integrada por computador (pirámide CIM) y se establecen sus necesidades de comunicación que han dado lugar al área de la tecnología conocida como "Comunicaciones Industriales". Se analizan finalmente las "Comunicaciones Industriales" y, para facilitar su asimilación por parte del lector, en los apéndices 1, 2, 3 y 4, se describen respectivamente las Comunicaciones Digitales, el bus de campo AS-i, la familia de redes de campo PROFIBUS y la red industrial universal Profinet.

CAPÍTULO 7

Sensores Industriales

7.1 Introducción

Tal como se indica en los capítulos 1 y 4, para que un sistema electrónico pueda controlar un proceso o un producto industrial es necesario que reciba información de la evolución de determinadas variables físicas del mismo, que en su mayoría no son eléctricas (como por ejemplo la temperatura, la presión, el nivel de un líquido o de un sólido, la fuerza, la radiación luminosa y la posición, velocidad, aceleración o desplazamiento de un objeto, etc.). Por ello, el acoplamiento entre el sistema electrónico y el proceso productivo se debe realizar a través de dispositivos que convierten las variables no eléctricas en eléctricas. Dichos dispositivos reciben diversos nombres, como captador (*Receiver*), detector (*Detector*), transductor (*Transducer*), transmisor (*Transmitter*), sonda y sensor (*Sensor*), aunque es este último el más utilizado por los fabricantes de sistemas electrónicos de control en general y por los fabricantes de autómatas programables en particular.

No existe una única definición de sensor aceptada de manera universal. Se considera que, en general, un sensor es todo dispositivo que, situado en un cierto medio, genera una señal (función de alguna característica de dicho medio) de una determinada forma física (presión, nivel, temperatura, etc.), convertible en otra señal de una forma física diferente. El elemento que realiza dicha conversión se suele denominar transductor (*Transducer*). Cada vez es más usual denominar sensor o elemento sensor al conjunto formado por el dispositivo sensor, anteriormente descrito, y el transductor acoplado a él. Por otro lado, la existencia de sistemas que memorizan, amplifican, y en general procesan señales eléctricas, hace que la mayoría de los transductores conviertan variables no eléctricas en eléctricas. Por todo lo expuesto, cada vez se utiliza más la palabra sensor (*Sensor*) para definir al dispositivo o elemento que convierte una variable física no eléctrica en otra eléctrica, que en alguno de sus parámetros (nivel de tensión, nivel de corriente, frecuencia, etc.), contiene información correspondiente a la primera. Por otra parte, el tipo de señal eléctrica portadora de información y sus parámetros varía de un dispositivo sensor a otro y por ello es necesario acoplar la salida de éste a un circuito que, de acuerdo con las características de aquélla, realice al menos una de las siguientes operaciones:

Amplificación de la señal (*Signal amplification*).

Filtrado (*Filtering*).

Corrección (*Correction*).

Conversión (*Conversion*) en otra diferente.

Dicho circuito recibe el nombre de circuito acondicionador o de acondicionamiento (*Signal conditioner*) y su utilización conjunta con el elemento sensor da lugar a un sistema como el representado en la figura 7.1 que puede ser denominado “Sistema Sensor”, aunque normalmente se le denomina simplemente sensor. Dicho sistema genera una señal normalizada, ya sea por el fabricante o siguiendo pautas establecidas por organismos de normalización, como por ejemplo la IEC (acrónimo de *International Electrotechnical Commission*), el IEEE (acrónimo de *Institute of Electrical and Electronic Engineers*), etc.

De todo lo expuesto se concluye que el análisis de los sensores es una tarea compleja y la selección del más adecuado para una aplicación determinada obliga a tener en cuenta múltiples factores (tipo de magnitud a medir y sus características, principio de funcionamiento del sensor, etc.), que se estudian en sucesivos apartados.

A los sistemas sensores adecuadamente contruidos para trabajar en las condiciones existentes en un entorno industrial (temperatura elevada, presencia de polvo, humedad relativa alta, etc.) se les denomina en este libro “Sensores Industriales” (*Industrial Sensors*). Los sensores industriales son, en la gran mayoría de los casos, sistemas sensores como el representado en la figura 7.1 o sensores inteligentes (*Smart sensors*) como los descritos en el apartado 7.2.7 (Figura 7.15), adecuadamente encapsulados. Los sensores inteligentes forman parte de los dispositivos electrónicos inteligentes conocidos como IED (acrónimo de *Intelligent Electronic Devices*).

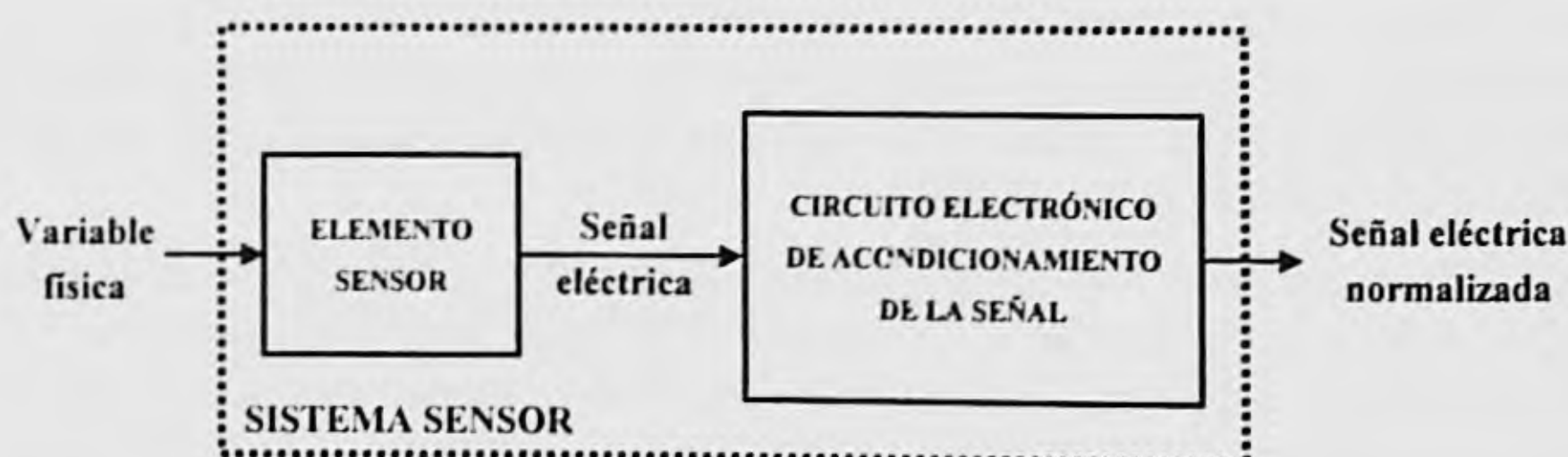


Figura 7.1 Componentes básicos de un sensor.

En este capítulo se trata de proporcionar al lector la formación adecuada para aplicar los sensores industriales comercializados por diferentes fabricantes. Por ello, en general, no se entra en los detalles del circuito de acondicionamiento. Al lector interesado se le remite a la bibliografía [FRAD 96] [PALL 98] [PALL 01] [PERE 03] [WILS 05].

7.2 Características de los sensores industriales

7.2.1 Introducción

Tal como se indica en el apartado anterior, son muy numerosas las variables o magnitudes físicas susceptibles de ser transformadas en señales eléctricas. Pero, además, las señales eléctricas pueden contener la información en un número elevado de parámetros diferentes (la amplitud

de la tensión, la frecuencia, etc.) y por ello los sensores se pueden clasificar de acuerdo con un conjunto de características diferentes y no excluyentes que se indican en la tabla 7.1 y se analizan en sucesivos apartados.

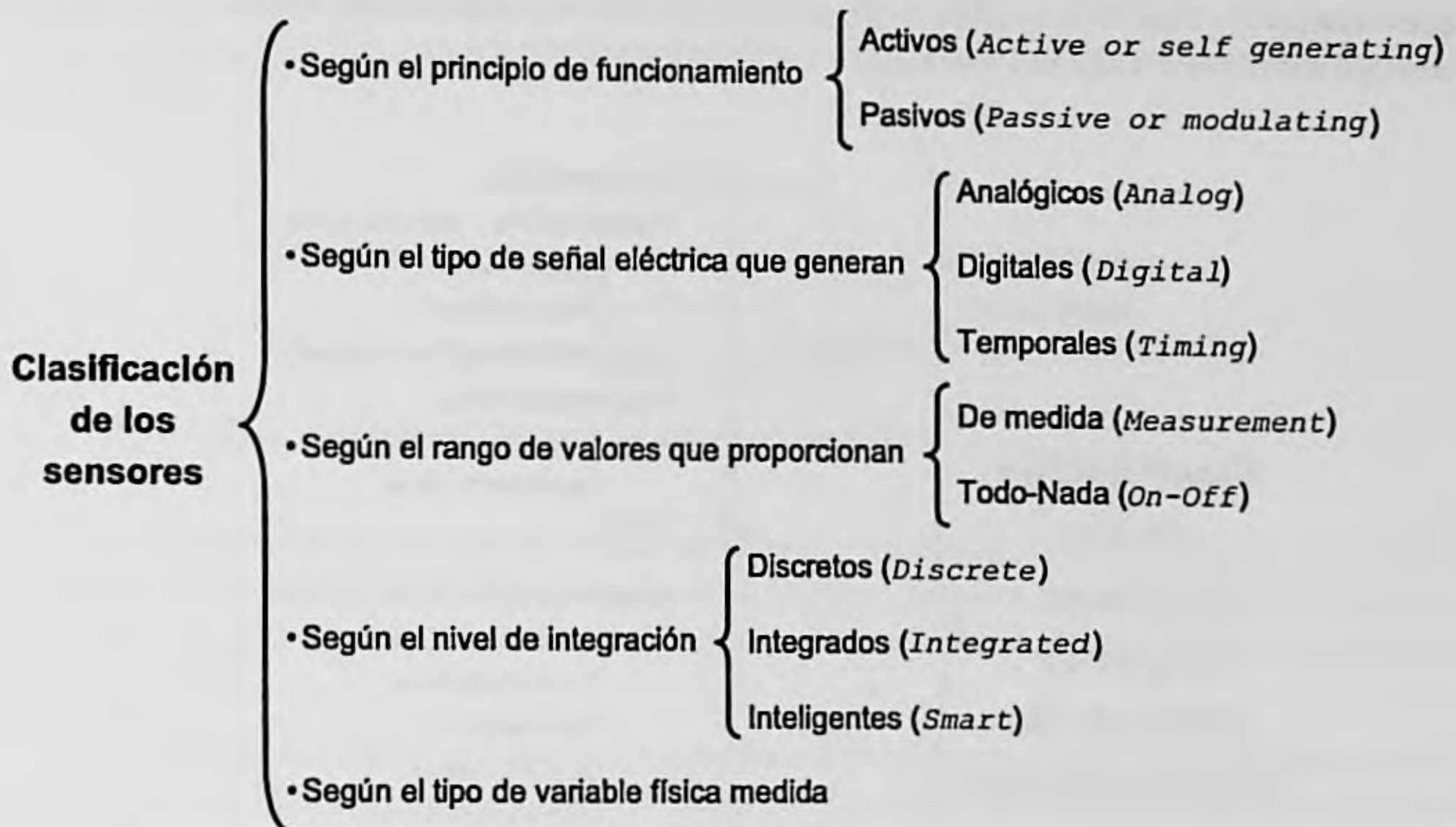


Tabla 7.1 Clasificación de los sensores.

7.2.2 Clasificación de los sensores industriales según el principio de funcionamiento del elemento sensor

La existencia de numerosas variables o magnitudes físicas susceptibles de ser convertidas en señales eléctricas da lugar a que también sean muy variadas las tecnologías o principios de conversión (inductivo, piezoeléctrico, extensiométrico, capacitivo, etc.) que se emplean actualmente.

La naturaleza del principio de funcionamiento del elemento sensor empleado determina la clasificación de los sensores en:

- **Sensores activos**

Se consideran *activos o generadores* (*Active or self generating*) los sensores en los que la magnitud física a medir proporciona la energía necesaria para la generación de la señal eléctrica de salida. Son un ejemplo de sensores activos los basados en los efectos piezoeléctrico y termoelectrónico.

- **Sensores pasivos**

Son *pasivos o moduladores* (*Passive or modulating*) los sensores en los que la magnitud física a medir se limita a modificar alguno de sus parámetros eléctricos característicos como por ejemplo la resistencia, la capacidad, etc. Los sensores de este tipo se caracterizan por necesitar una tensión de alimentación externa. Son ejemplo de sensores

pasivos los basados en las resistencias cuyo valor depende de la temperatura (termorresistivos) o de la luz (fotorresistivos).

En la tabla 7.2 se indican los principales sensores pasivos y activos. Es conveniente resaltar que la mayoría de las variables físicas se pueden convertir mediante varios sensores que utilizan principios de funcionamiento distintos.

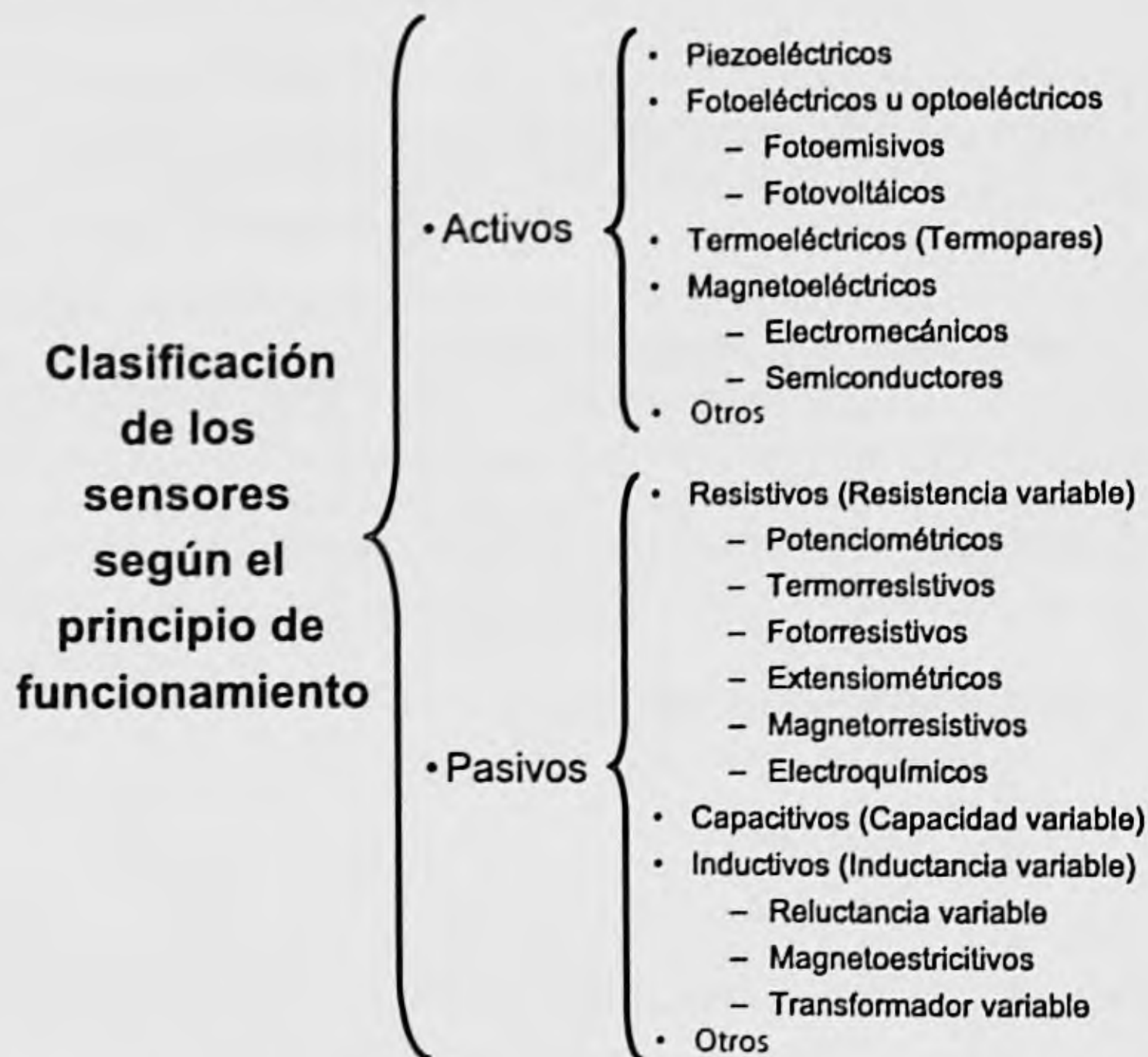


Tabla 7.2 Clasificación de los sensores según el principio de funcionamiento.

7.2.3 Clasificación de los sensores según el tipo de señal eléctrica que generan

Tal como se indica en el apartado 7.1, los sistemas sensores proporcionan a su salida señales eléctricas. Una señal eléctrica puede estar constituida por una tensión o voltaje entre dos puntos, la corriente a través de un componente electrónico o la energía representada por la tensión en bornes de un componente, multiplicada por la corriente que pasa a través de él. Pero tanto una tensión como una corriente eléctrica pueden variar de distinta forma a lo largo del tiempo y dar lugar a diferentes maneras, denominadas formatos o dominios de datos (*Data domains*), de representar información mediante algún parámetro de una señal eléctrica. Por ello, según el formato de la señal o señales eléctricas que genera, un sistema sensor puede ser analógico, digital o temporal, tal como se indica en la tabla 7.1. A continuación se estudia cada uno de ellos.

7.2.3.1 Sensores analógicos

Los *sensores analógicos* (*Analog sensors*) generan señales eléctricas denominadas analógicas (*Analog signals*) que pueden tomar cualquier valor dentro de unos determinados márgenes y que llevan la información en su amplitud. Su diagrama de bloques es el representado en la figura 7.2. Las señales eléctricas analógicas se pueden clasificar a su vez en variables o continuas.



Figura 7.2 Diagrama de bloques de un sensor analógico.

Las señales analógicas variables son aquellas que equivalen a la suma de un conjunto de senoides de frecuencia mínima mayor que cero. Un caso típico es la señal senoidal de frecuencia constante que representa la información mediante su amplitud (Figura 7.3a) y las señales de audiofrecuencia.

Las señales analógicas continuas son aquellas que se pueden descomponer en una suma de senoides cuya frecuencia mínima es cero. Se trata de señales que pueden tener un cierto nivel fijo durante un tiempo indefinido (Figura 7.3b), y que representan también la información mediante su amplitud.

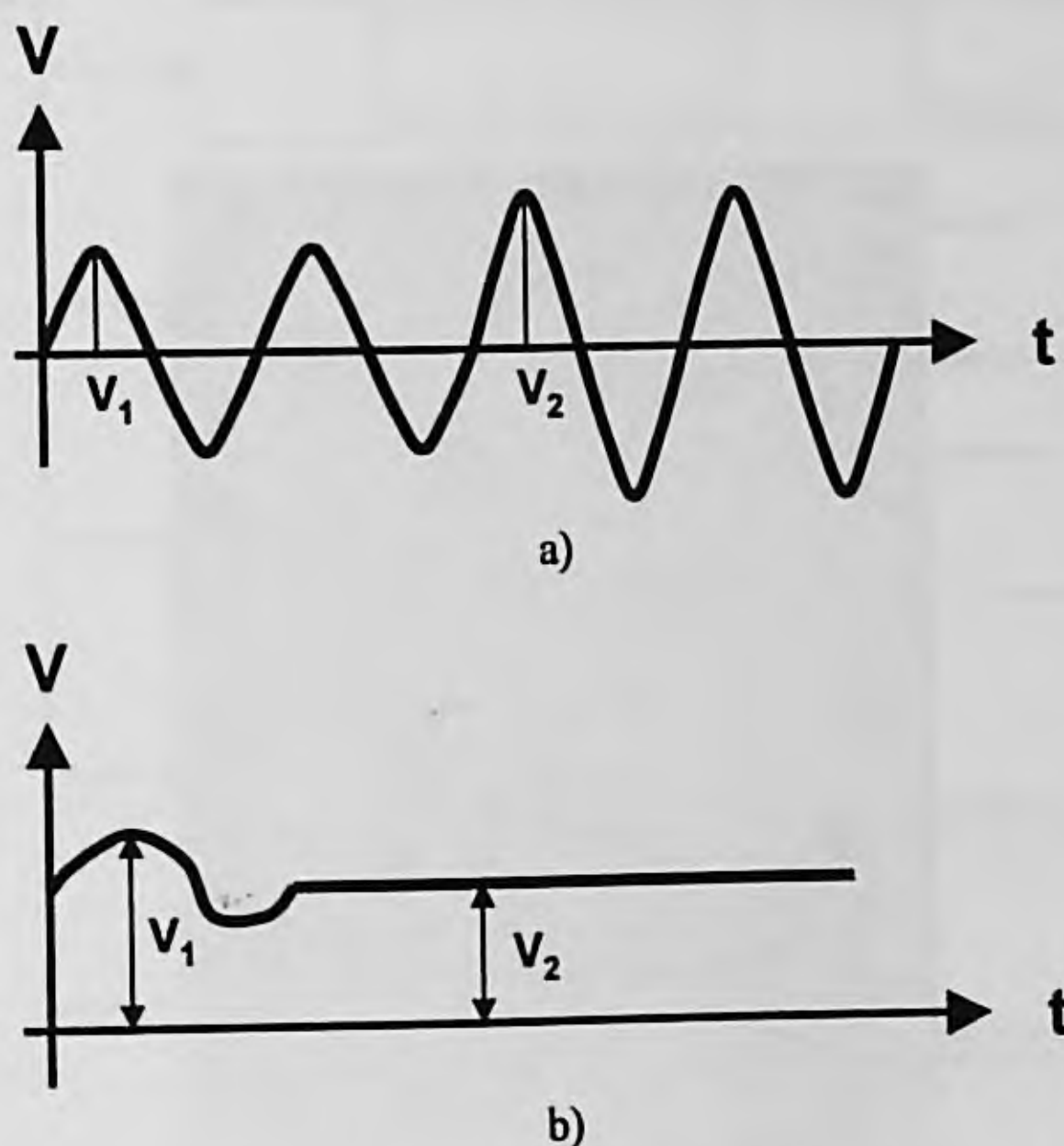


Figura 7.3 Señales analógicas: a) Senoidal de frecuencia constante; b) Continua.

Las señales analógicas pueden ser también unipolares o bipolares. Las unipolares sólo pueden ser positivas o negativas con respecto a un terminal de referencia, mientras que las bipolares pueden ser tanto positivas como negativas (Figura 7.4).

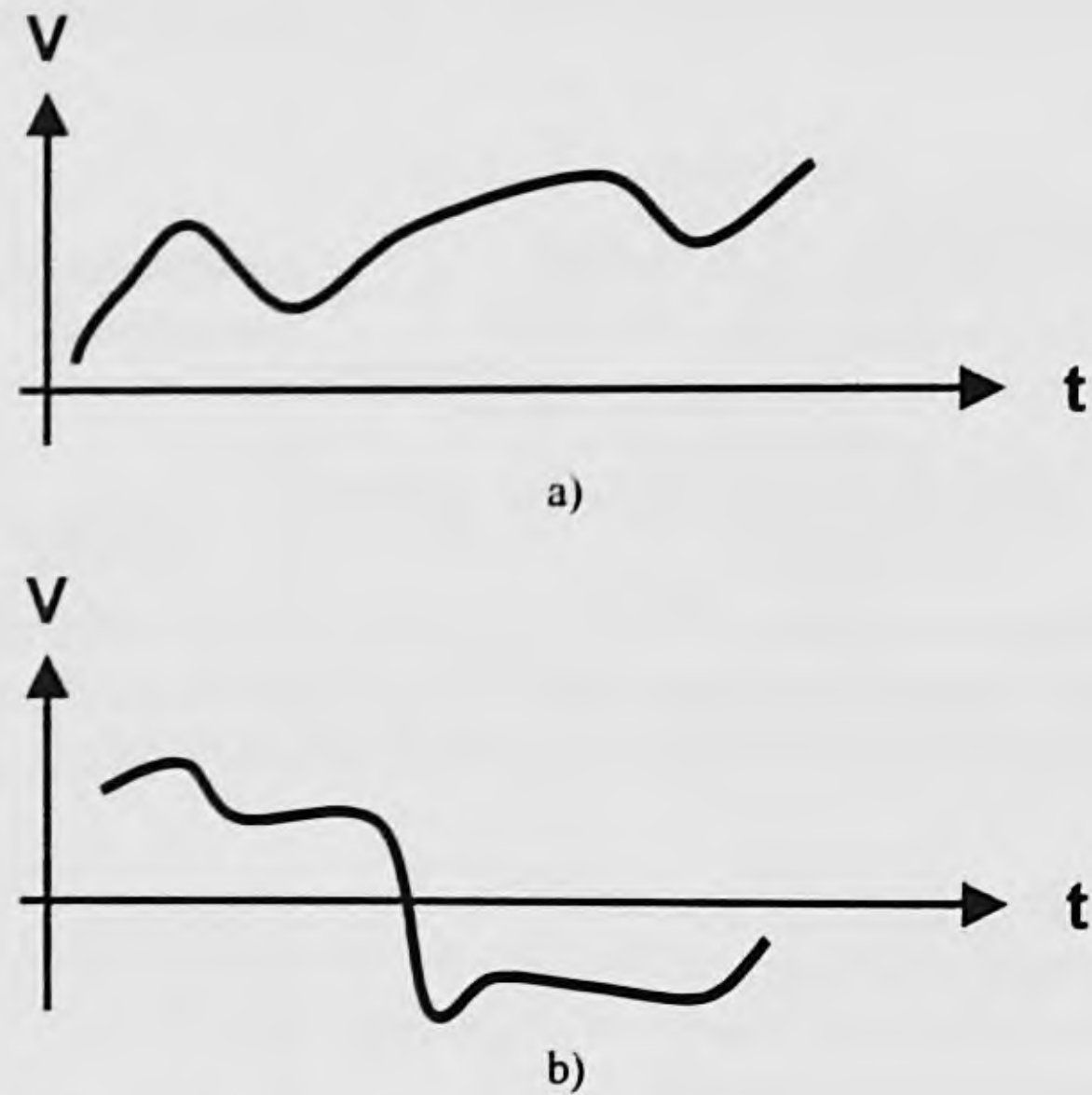


Figura 7.4 Señales analógicas: a) Unipolar; b) Bipolar.

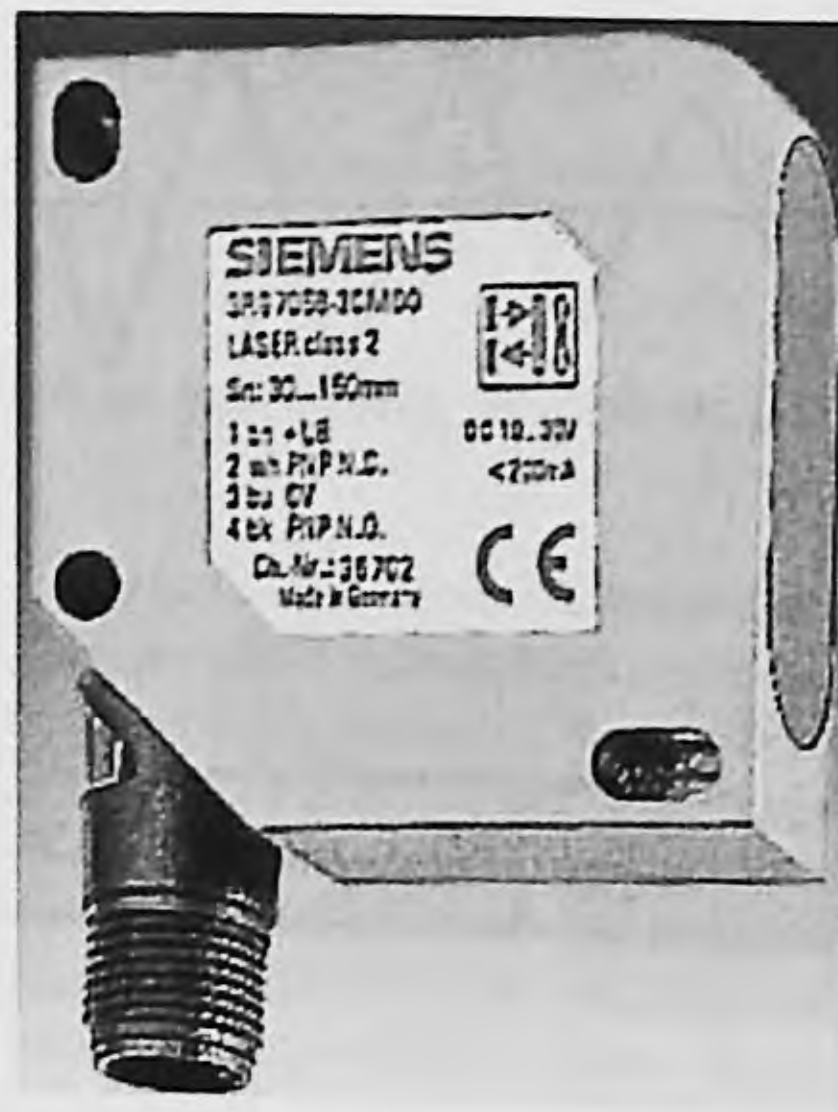


Figura 7.5 Sensor optoelectrónico láser de proximidad 3RG70-1CM00 de Siemens.

El mundo físico es en general analógico y por ello la mayoría de los elementos sensores proporcionan señales analógicas. Sin embargo, las señales eléctricas generadas por los elementos de los sensores analógicos adolecen de problemas relacionados con la presencia de ruido, interferencias y distorsión, sobre todo si se han de transmitir a distancias relativamente grandes, y por ello es imprescindible el circuito electrónico de acondicionamiento de la señal indicado en la figura 7.1 La figura 7.5 muestra un sensor que genera una variable analógica en el rango de 0 a 10 V.

7.2.3.2 Sensores digitales

Los *sensores digitales* (*Digital sensors*) generan señales eléctricas que sólo toman un número finito de niveles o estados entre un máximo y un mínimo, y por ello reciben el nombre de digitales. Las más utilizadas son las binarias, que sólo pueden tener dos niveles de tensión, que se asignan a los números binarios 0 y 1. Una variable binaria recibe el nombre de "bit". El criterio de asignación de los estados es totalmente arbitrario. Si se asigna el valor 1 a la tensión más alta y el valor 0 a la tensión más baja, la lógica utilizada se llama "positiva" y si, por el contrario, se asigna el valor 0 a la tensión más alta y el valor 1 a la tensión más baja, la lógica recibe el nombre de "negativa". Para representar una información se necesita un cierto número n de variables binarias cuyo valor depende de la precisión que se desee. Las n variables binarias se pueden representar de dos formas diferentes:

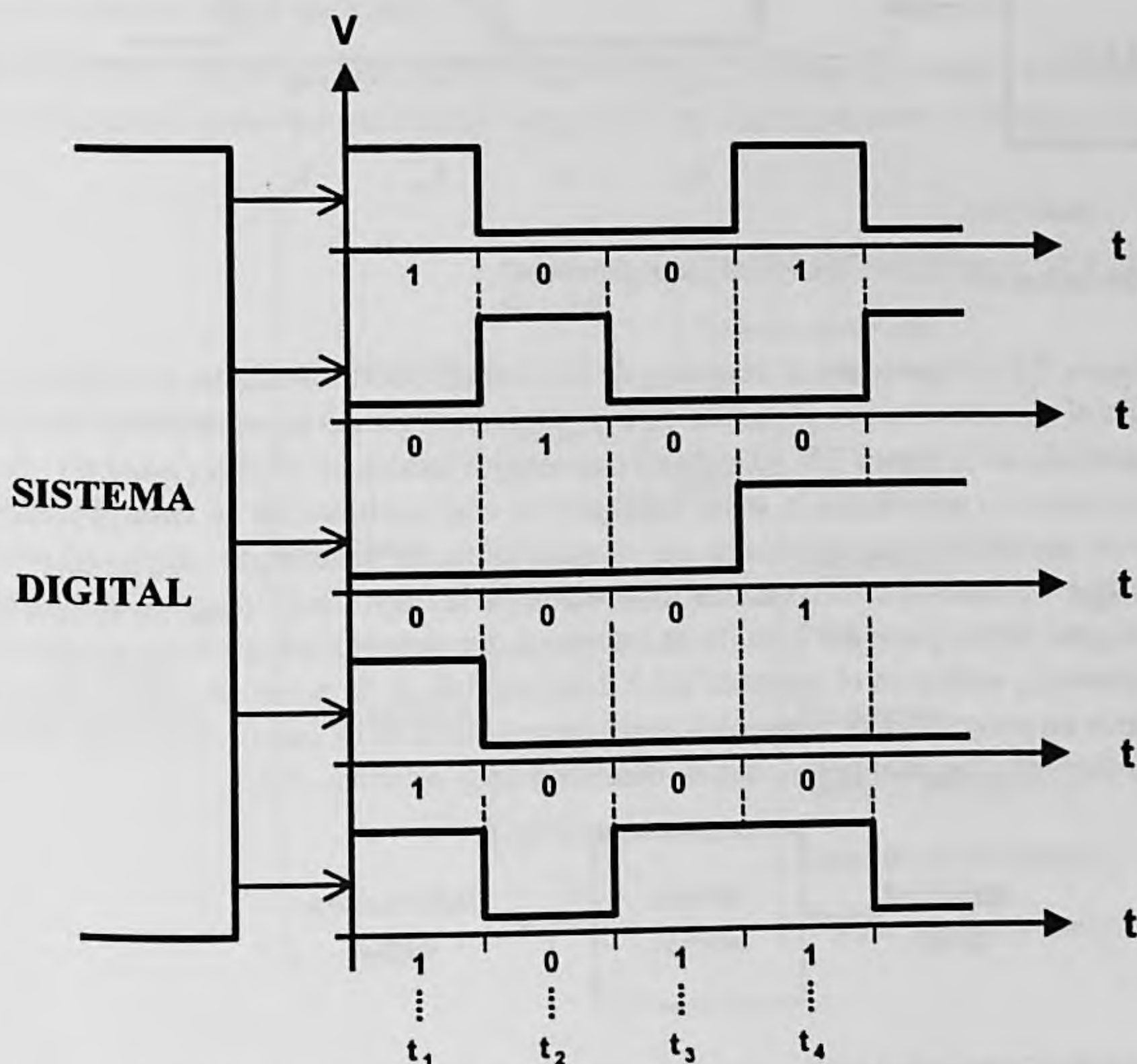


Figura 7.6 Señal digital binaria en formato paralelo.

- **Mediante señales binarias independientes**

Se tiene así, por ejemplo, el número 10011 en un único instante t_1 (Figura 7.6). En instantes sucesivos se pueden tener números diferentes (por ejemplo, 01010 en t_2 , en la figura 7.6). Este formato recibe el nombre de paralelo (*Parallel output*).

- **Mediante una secuencia de niveles cero y uno de una señal digital**

En la figura 7.7 se indica una señal digital binaria que representa el número 10011 en el sistema de numeración binario. Este formato recibe el nombre de *serie* (*Serial output*).

El formato serie se utiliza para transmitir a distancia una información digital, mientras que el paralelo es el utilizado por los procesadores digitales [MAND 07] [MAND 08] y el que proporcionan algunos tipos de sensores, como por ejemplo, los codificadores absolutos angulares de posición, descritos en el apartado 7.7.2.4.2.1 En general, el formato serie se genera a partir del paralelo, mediante un procesador de comunicaciones, que constituye un circuito de interfaz de un procesador digital [MAND 07]. Se tiene así un sensor con capacidad de comunicación (véase el apartado 7.2.7).



Figura 7.7 Señal digital binaria en formato serie.

En la figura 7.8 se representa el diagrama de bloques de un sensor digital de salida en paralelo. Cuando el elemento sensor proporciona una señal analógica el esquema de bloques típico es el representado en la figura 7.9. El circuito convertidor analógico-digital (*Analog to Digital converter*) transforma la señal analógica en una combinación de variables digitales. El circuito de acondicionamiento adapta las características de la señal de salida del elemento sensor a las que necesita el convertidor en su entrada. En la mayoría de los casos el convertidor analógico-digital forma parte del circuito de interfaz de un sensor analógico con un procesador digital, tal como se indica en el apartado 8.2.2.5 del capítulo 8. Si al sensor digital de la figura 7.9 se le añade un procesador de comunicaciones, se convierte en un sensor con salida serie. Son sensores de este tipo los inteligentes que se describen en el apartado 7.2.7.

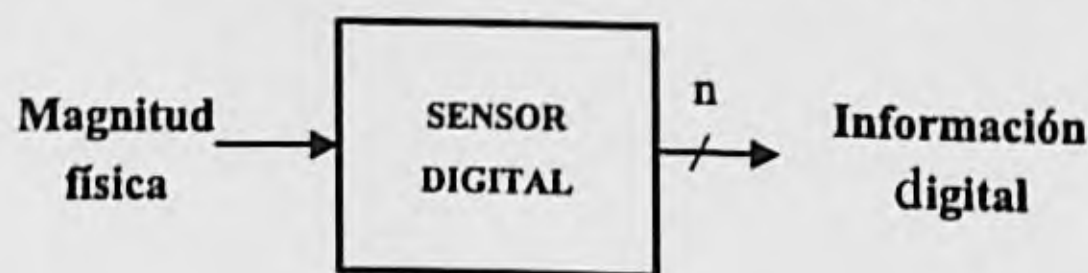


Figura 7.8 Diagrama de bloques de un sensor digital de salida en paralelo.

El formato digital presenta la característica de que la precisión de la medida sólo depende del número de variables binarias. Su utilización se ha generalizado gracias al progreso de la Microelectrónica [MAND 07] [MAND 08], que ha permitido realizar procesadores electrónicos digitales de coste reducido y capacidad adaptada a las especificaciones de cada aplicación concreta.

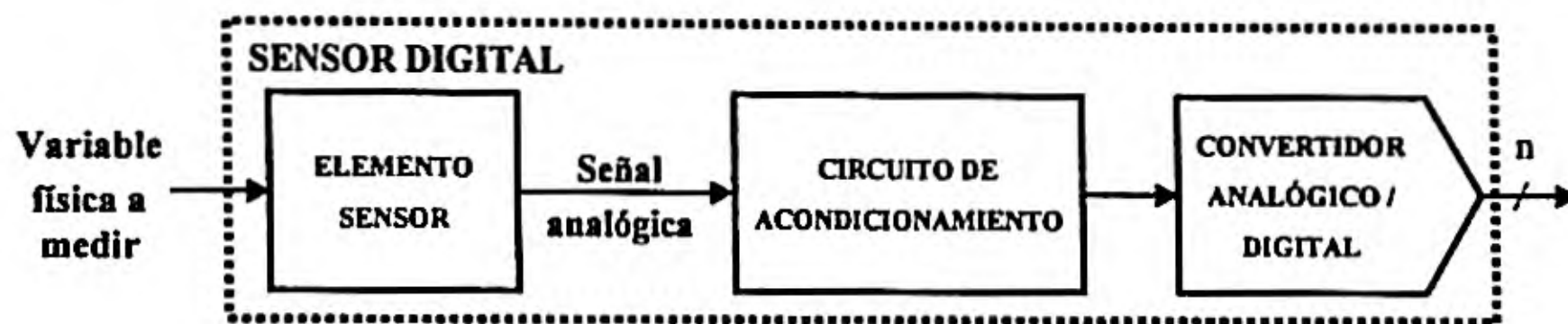


Figura 7.9 Esquema de bloques típico de un sensor digital cuyo elemento sensor proporciona una señal analógica.

7.2.3.3 Sensores temporales

Los sensores temporales proporcionan a su salida señales eléctricas en las que la información está asociada al parámetro tiempo (*Timing signals*). Las señales eléctricas temporales pueden ser senoidales o cuadradas. Según la forma de la señal y el tipo de parámetro, se clasifican tal y como se indica en la tabla 7.3.

Las señales temporales senoidales suelen recibir el nombre de señales moduladas porque se obtienen modificando un parámetro temporal de una señal senoidal generada por un circuito

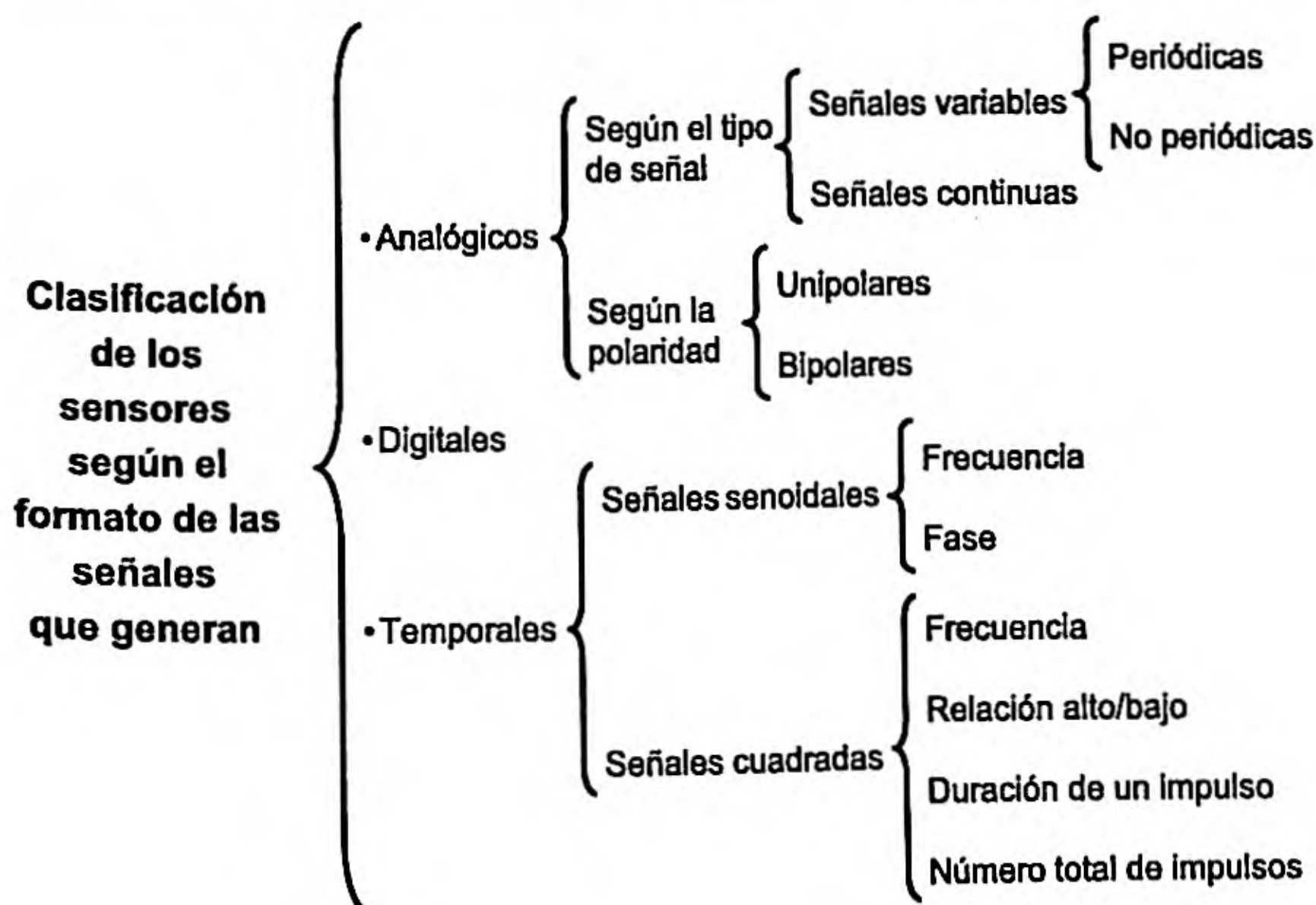


Tabla 7.3 Clasificación de los sensores según el tipo de señales eléctricas que generan.

oscilador o una onda cuadrada generada por un generador de impulsos, mediante un circuito electrónico denominado modulador.

Las señales temporales cuadradas tienen una amplitud fija y un parámetro temporal variable que puede ser (Tabla 7.3):

- La frecuencia o su inverso, el periodo. La información está representada por el valor de cualquiera de ellos. Un ejemplo de sensor temporal es el sensor de ultrasonidos BERO 3RG62 32 (Figura 7.10). Tiene una salida que es una onda cuadrada cuya frecuencia varía entre 250 Hz y 1500 Hz cuando la distancia de un objeto situado frente a él lo hace entre 50 y 300 mm.
- La relación uno/cero o alto/bajo. La señal utilizada posee un periodo de duración fija y la información está contenida en el valor del tiempo durante el cual la información se encuentra en nivel uno con relación a aquél durante el cual está en nivel cero (Figura 7.11). Se suele decir que la señal está modulada en anchura de impulsos y se la conoce como PWM (acrónimo de *Pulse Width Modulation*).
- La duración de un impulso (*Pulse duration*) que se genera en el instante en que se desea conocer el valor de una variable (Figura 7.12).

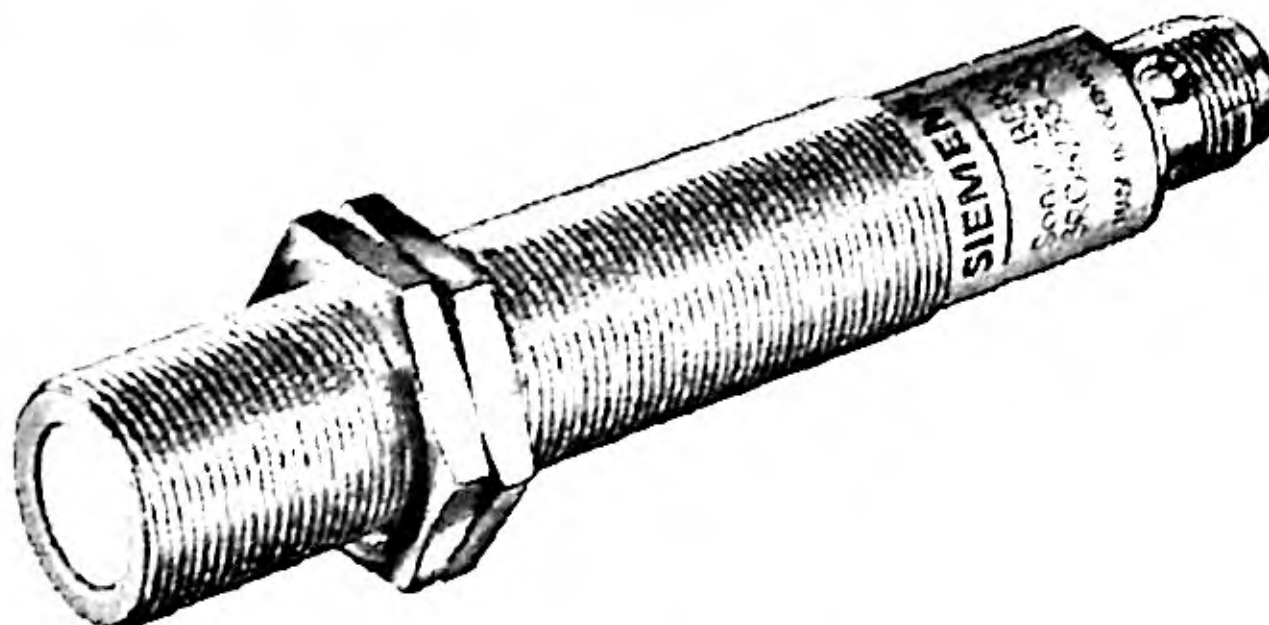


Figura 7.10 Sensor de ultrasonidos BERO 3RG62 32 (cortesía de Siemens).

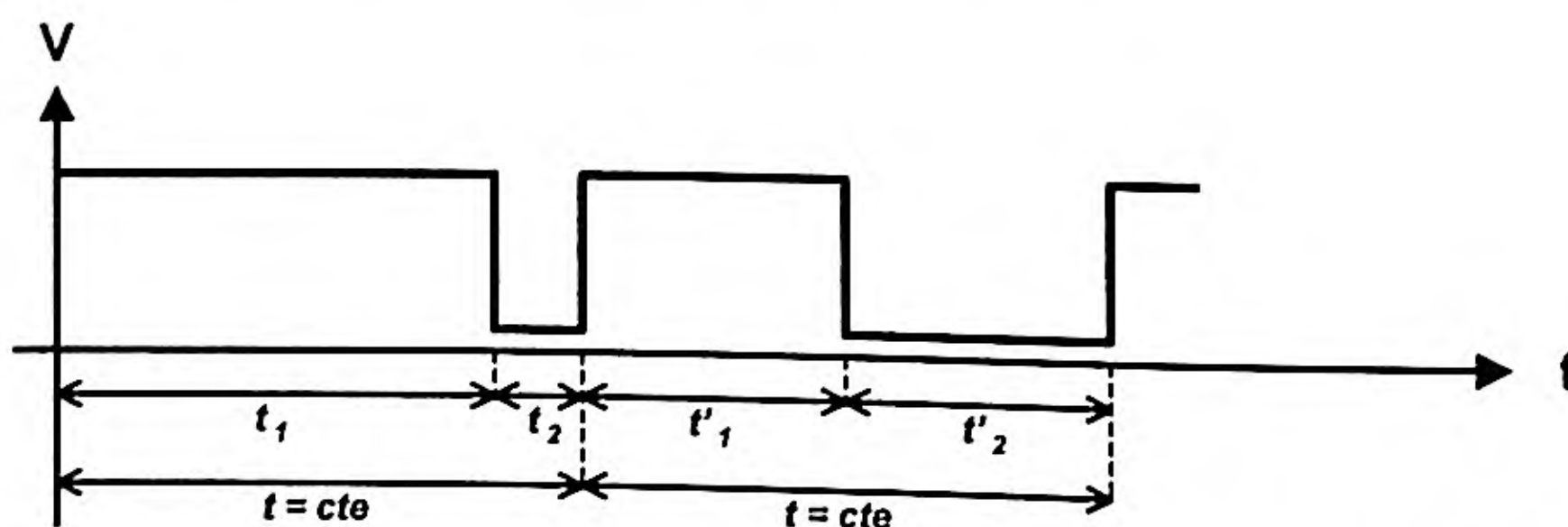


Figura 7.11 Señal temporal que contiene la información en la relación entre la duración de los niveles uno y cero (alto/bajo).

- El número total de impulsos que se generan en su salida a partir de un determinado instante. Un ejemplo típico es el codificador incremental de posición (*Incremental position encoder*) que se describe en el apartado 7.7.2.4.2.2.

Una información codificada mediante alguno de estos parámetros se ve mucho menos afectada por el ruido eléctrico y por la atenuación inherente a cualquier tipo de transmisión, debido a su menor dependencia del valor absoluto de la amplitud de la señal. Algunos ejemplos de esta forma de codificar la información son, entre otros, la radiodifusión en frecuencia modulada conocida como FM (acrónimo de *Frequency Modulation*), la modulación de anchura de impulsos conocida como PWM (acrónimo de *Pulse Width Modulation*), utilizada en sistemas de radio control, y los teclados multifrecuencia empleados en telefonía.

Pocos elementos sensores dan a su salida la información en formato temporal, pero éste se obtiene fácilmente a través de un circuito convertidor de formato. Por ejemplo, la señal analógica proporcionada por un elemento sensor se puede convertir en una señal temporal que lleva la información en la frecuencia mediante un oscilador controlado en tensión conocido como VCO (acrónimo de *Voltage Controlled Oscillator*), tal como se indica en la figura 7.13.

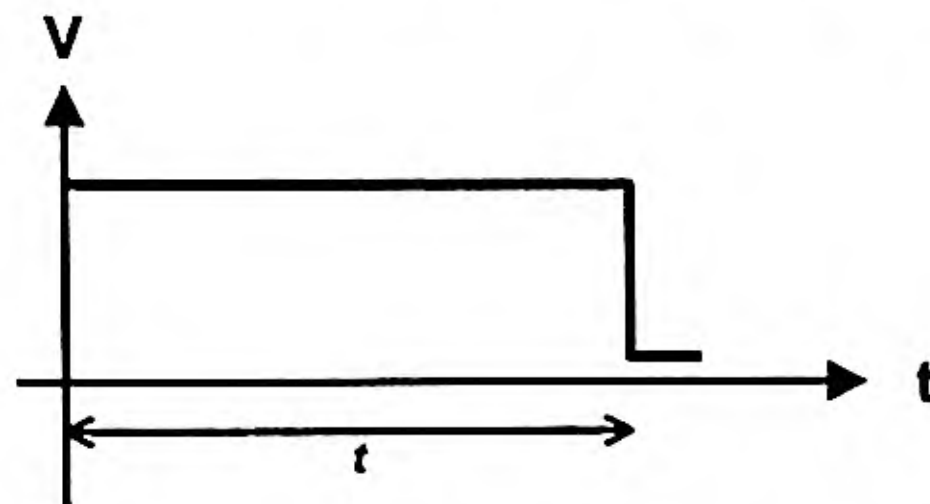


Figura 7.12 Señal temporal que contiene la información en la duración de un impulso.

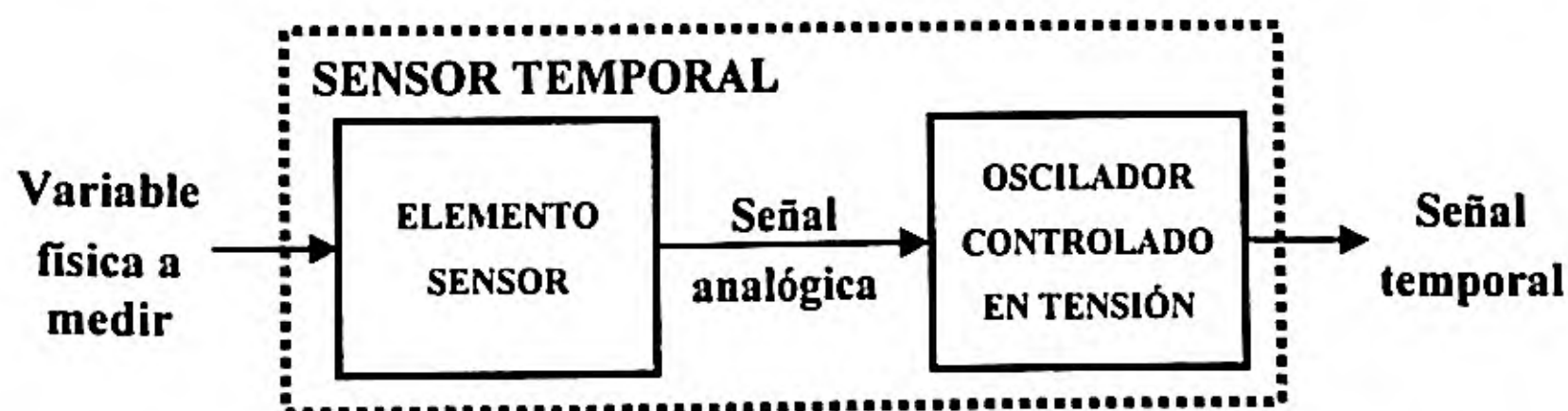


Figura 7.13 Sensor temporal realizado con un oscilador controlado en tensión.

7.2.4 Clasificación de los sensores según el rango de valores

Según el rango de los valores de la señal de salida que proporcionan, los sensores pueden ser de medida (*Measurement sensors*) o todo-nada (*On-off sensors*).

Sensores de medida

Los sensores de medida se caracterizan por proporcionar a la salida todos los valores posibles correspondientes a cada valor de la variable de entrada dentro de un determinado rango.

Pueden ser analógicos, digitales o temporales. Ejemplos de sensores de medida son un sensor analógico resistivo de temperatura y un sensor temporal incremental de posición (*Incremental position encoder*), descrito en el apartado 7.7.2.4.2.

Sensores todo-nada

Los sensores todo-nada se caracterizan por detectar solamente si la magnitud de la variable de entrada está por encima o por debajo de un determinado valor. Proporcionan a la salida una señal eléctrica que sólo puede tomar dos valores y por ello a veces se les denomina, con un exceso de simplificación, sensores digitales.

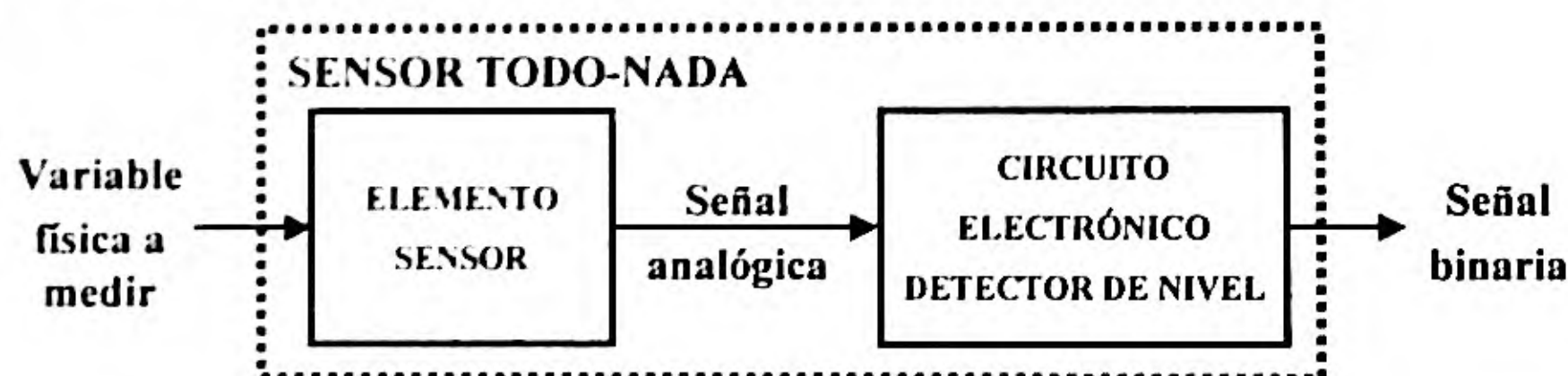


Figura 7.14 Esquema de bloques de un sensor todo-nada cuyo elemento sensor proporciona una señal analógica.

En la figura 7.14 se representa el diagrama de bloques típico de un sensor todo-nada que indica si el valor de una magnitud física está por encima o por debajo de un determinado nivel. La salida del elemento sensor es una señal analógica cuyo nivel es proporcional al valor de la magnitud física a medir y el circuito acondicionador es fundamentalmente un circuito electrónico detector de nivel cuya salida toma un valor u otro, según que el nivel de la señal aplicada a su entrada esté por encima o por debajo de un cierto nivel. Son ejemplos de sensores de este tipo los presostatos, los termostatos, los interruptores de caudal, etc.

Son también sensores todo-nada los que indican si se ha producido o no una determinada circunstancia, como por ejemplo la presencia o ausencia de un objeto en las proximidades del sensor (*Proximity sensors*), que se describen en el apartado 7.7.2.1.

7.2.5 Clasificación de los sensores industriales según el modo de operación

De acuerdo con el modo de operación, los sensores de medida pueden ser de deflexión o de comparación.

En los sensores de deflexión (*Deflection sensors*), la variable medida produce otra similar en el sensor pero opuesta y relacionada directamente con ella. También se denominan sensores de medida directa (*Direct measurement*).

En los sensores de comparación (*Null-type sensors*), la señal a medir se resta de la de un patrón. Las medidas por comparación (*Comparison measurement method*) son más precisas porque se pueden calibrar con un patrón de calidad. El detector de desequilibrio debe medir alrededor del cero y en ocasiones ha de ser muy sensible. Tienen, en general, menor respuesta dinámica que los de deflexión.

7.2.6 Clasificación de los sistemas sensores según la función de transferencia

La función de transferencia es la relación matemática que existe entre la magnitud de entrada del sensor y la señal eléctrica de salida. De acuerdo con ella los sensores pueden ser lineales o no lineales.

Un sensor es lineal cuando la función de transferencia es una ecuación diferencial lineal de coeficientes constantes. Según el orden de la ecuación diferencial los sensores lineales pueden ser de orden cero, uno o dos [PALL 01].

La función de transferencia de los sensores lineales de orden cero es una línea recta que pasa por el origen, es decir, su ecuación es:

$$y = k \cdot x$$

y se caracterizan por tener un elemento que disipa energía.

Los sensores lineales de orden uno poseen un elemento capaz de almacenar energía, como por ejemplo un sensor de temperatura que posee una masa M y un coeficiente de transmisión del calor CT , que genera un retardo a la detección de los cambios de la temperatura del entorno del sensor. Su función de transferencia es:

$$a_1 \cdot \frac{dy}{dt} + a_0 \cdot y = x$$

y se caracterizan por introducir un retardo (constante de tiempo) en la respuesta a una variación en escalón de la magnitud de entrada.

Los sensores lineales de orden dos poseen un elemento que consume energía y dos que la almacenan como por ejemplo un sensor de temperatura que además de poseer una masa M está protegido por una cubierta que añade una resistencia térmica a las características del sensor. Su función de transferencia es de la forma:

$$a_2 \cdot \frac{d^2 y}{dt^2} + a_1 \cdot \frac{dy}{dt} + a_0 \cdot y = x$$

y en su respuesta a un escalón pueden generar una sobreoscilación.

Un sensor no es lineal cuando la función de transferencia no es una ecuación diferencial lineal de coeficientes constantes. Un ejemplo típico de sensores no lineales son los termistores o resistencias de coeficiente de temperatura negativo, conocidas por el acrónimo NTC (*Negative Temperature Coefficient*) que tienen una función de transferencia exponencial. Este tipo de sensor exige una linealización que es realizada por el procesador analógico o digital incluido en el circuito de acondicionamiento.

7.2.7 Clasificación de los sistemas sensores según el nivel de integración

Según el nivel de integración los sistemas sensores pueden ser discretos, integrados e inteligentes.

- **Sensores discretos**

Reciben la denominación de sensores discretos (*Discrete sensors*) los sistemas sensores en los que el circuito de acondicionamiento se realiza mediante componentes electrónicos separados e interconectados entre sí.

- **Sensores integrados**

Los sensores integrados (*Integrated sensors*) se caracterizan porque el elemento sensor y el circuito acondicionador, o al menos este último, están contruidos en un único circuito integrado monolítico o híbrido. Son ejemplos típicos muchos sensores que están basados en las características de los semiconductores y miden temperatura, humedad, presión, etc. [SZE 94].

- **Sensores inteligentes**

No existe un consenso generalizado en relación con la definición de sensor inteligente (*Intelligent or Smart Sensor*). En numerosas ocasiones, la salida del circuito acondicionador debe ser modificada para llevar a cabo una o más de las siguientes tareas:

- Corregir no linealidades.
- Verificar el correcto funcionamiento del elemento sensor y el circuito acondicionador asociado.
- Transmitir la información a distancia.

Para ello es de gran utilidad añadir al sistema sensor de la figura 7.1 un procesador electrónico que realice un algoritmo de corrección, y/o diagnóstico, y/o verificación del funcionamiento, y/o comunicación con otros procesadores. Cuando dicho conjunto se realiza en un solo circuito integrado monolítico o híbrido se tiene un sensor con capacidad de proceso de información al que se suele denominar “sensor inteligente” cuyo diagrama de bloques se representa en la figura 7.15.

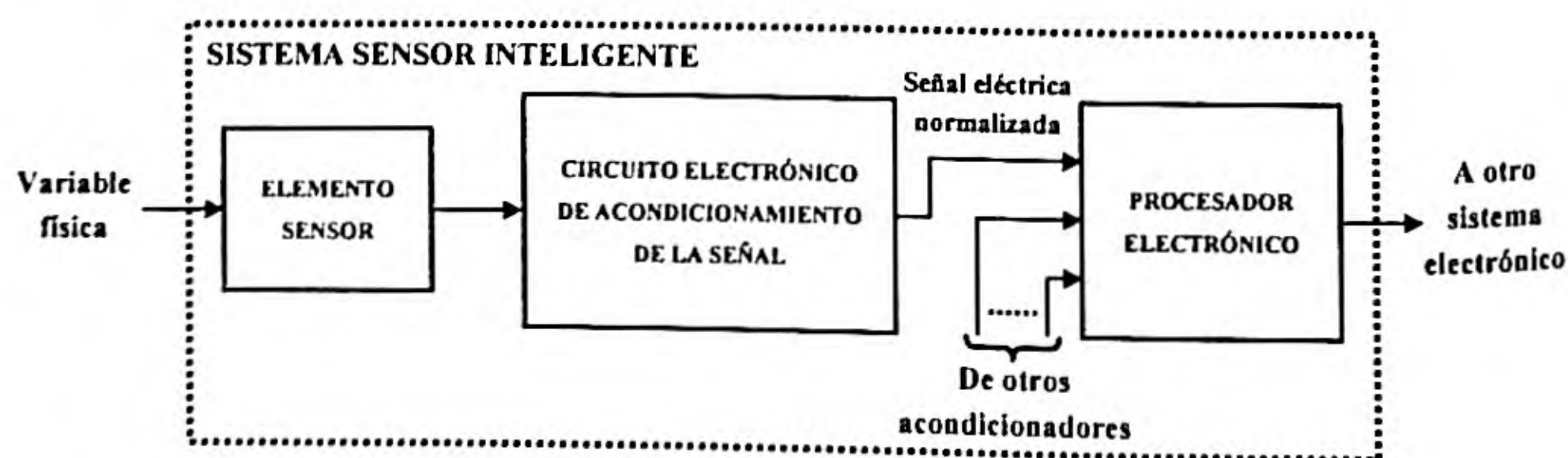


Figura 7.15 Esquema de bloques de un sensor inteligente.

Por ello, aunque no existe un consenso general en la definición de sensor inteligente, se admite que un sensor inteligente tiene capacidad para realizar la mayoría de las funciones siguientes:

- Cálculos numéricos.
- Comunicación en red (no una simple conexión punto a punto).
- Autocalibración y autodiagnóstico.
- Múltiples medidas con identificación del sensor.

Existen sensores industriales de los tres tipos. Inicialmente los sensores industriales eran discretos y tenían sus componentes encapsulados en un bloque único, pero el progreso de la Microelectrónica ha hecho que tanto el elemento sensor como el circuito acondicionador se realicen en un único circuito integrado monolítico o híbrido.

Por otra parte, los sensores discretos e integrados deben conectarse a un procesador digital, como por ejemplo un autómatas programable, a través de conexiones independientes, tal como se indica en la figura 7.16. Ello hace que se complique excesivamente el cableado cuando el número de sensores y/o su distancia al autómatas programable son elevados. Por ello, los fabricantes de equipos de automatización comercializan sensores industriales inteligentes, que poseen un procesador electrónico digital con capacidad de comunicación con otros procesadores a través de una línea única de comunicaciones, que suele recibir el nombre genérico de bus de campo (*Field Bus*) (Figura 7.17). Existen numerosos buses de campo diferentes creados por distintos fabricantes, algunos de los cuales han sido normalizados [AFNO 90] [GARC 97] [MARS 97] [SIEM 96] [WOOD 95]. Los buses de campo forman parte de la estrategia de Automatización Integrada Total (*Totally Integrated Automation*), que ha sido posible gracias al progreso de las comunicaciones digitales industriales descritas en el apartado 9.3 del capítulo 9.

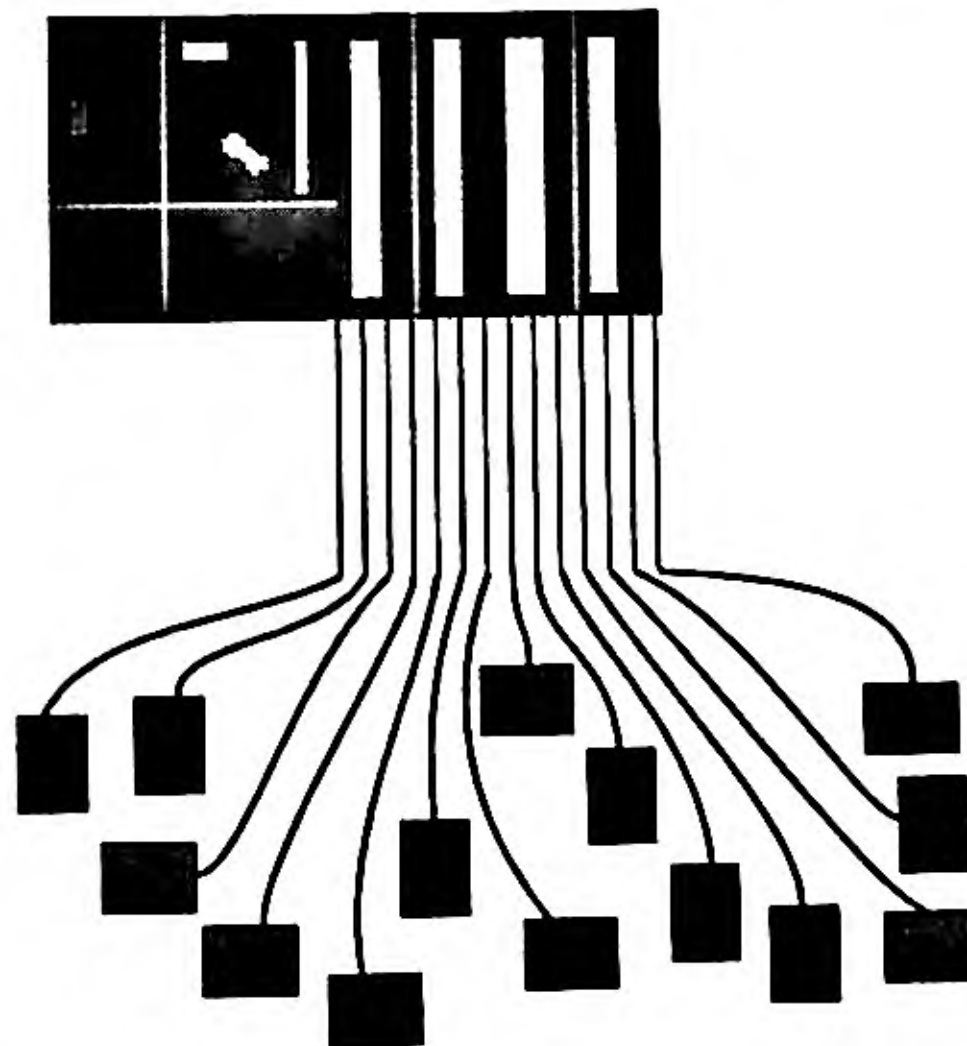


Figura 7.16 Conexión de sensores discretos y/o integrados a un autómatas programable (cortesía de Siemens).

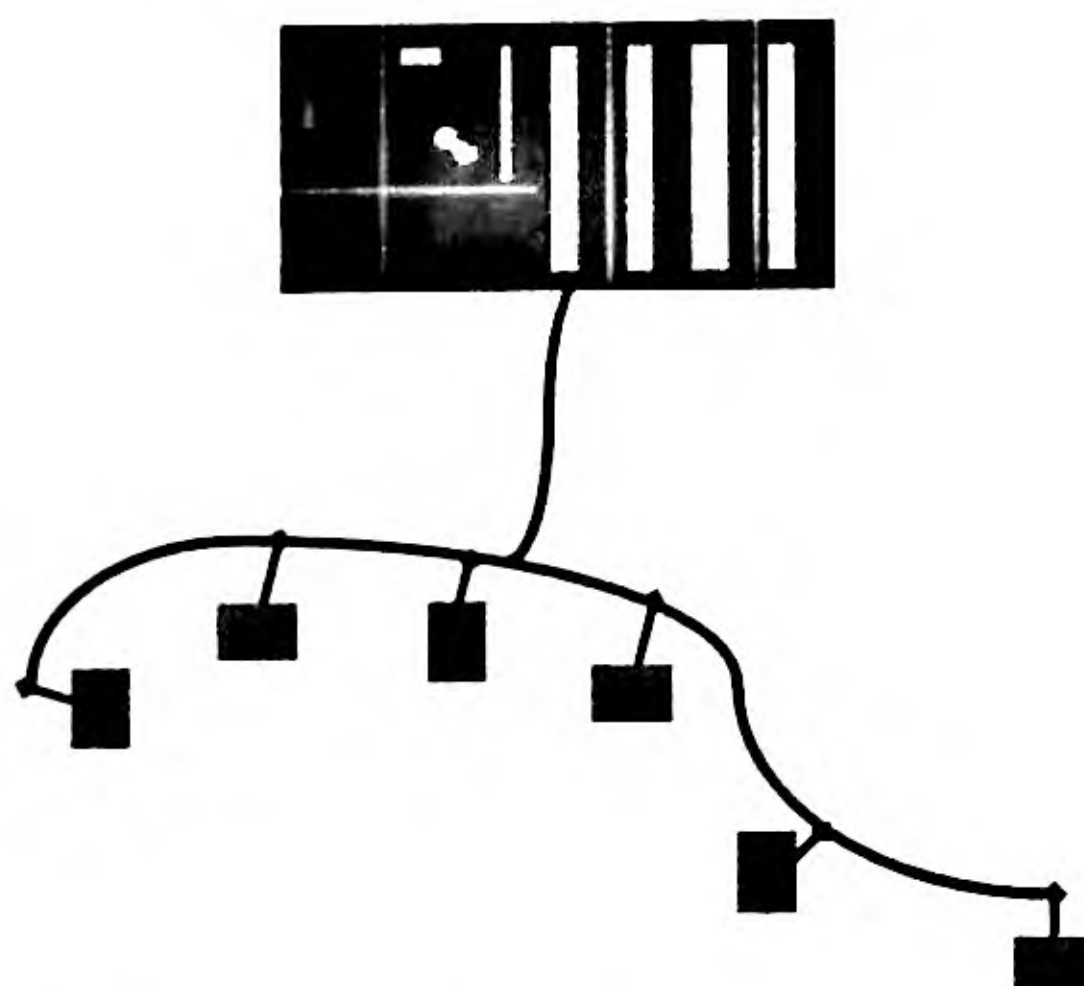


Figura 7.17 Conexión de un conjunto de sensores inteligentes a un autómata programable mediante un bus de campo (cortesía de Siemens).

Los sensores industriales con capacidad de conexión a un bus de campo suelen especificarse mediante la función que realizan (sensor de proximidad, célula fotoeléctrica, etc.) seguida en su caso de un nombre de marca propio (por ejemplo BERO de Siemens) y el apelativo del bus al que se pueden conectar, como por ejemplo el bus AS-i (acrónimo de *Actuator-Sensor-Interface*), que algunos fabricantes denominan “AS-Interface”, (descrito en el apéndice 2) o el bus HART. Son ejemplos de sensores con capacidad de conexión a un bus de campo el sensor optoelectrónico de proximidad BERO 3RG673 01-1RM00 que tiene incorporado un interfaz con el bus de campo AS-i (Figura 7.18) y el sensor de caudal MG711/A de la familia SITRANS FM conectable al bus Hart (Figura 7.19).



Figura 7.18 Sensor optoelectrónico de proximidad BERO 3RG673 01-1RM00 conectable al bus de campo AS-Interface (cortesía de Siemens).

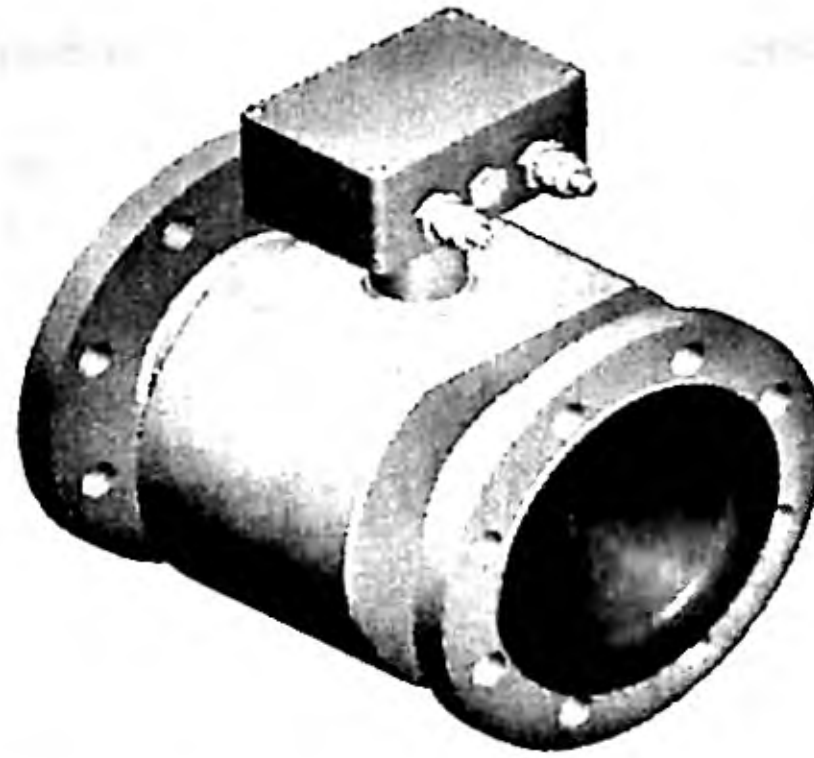


Figura 7.19 Sensor de caudal MG711/A conectable al bus de campo Hart (cortesía de Siemens).

7.2.8 Clasificación de los sensores según la variable física medida

Otra clasificación útil de los sensores es la realizada en función del tipo de variable física que convierten en eléctrica. En la tabla 7.4 se representa dicha clasificación, en la que se incluyen las variables físicas más importantes que es necesario medir en los procesos y productos industriales. Al lector interesado en ampliar esta clasificación se le remite a la bibliografía [FRAD 96] [PALL 98] [WILS 05].

Clasificación de los sensores según el tipo de variable física medida	• Presión
	• Temperatura
	• Humedad
	• Fuerza
	• Aceleración
	• Velocidad
	• Caudal
	• Presencia y/o posición de objetos
	• Nivel de sólidos o líquidos
	• Desplazamiento de objetos
	• Químicos

Tabla 7.4 Clasificación de los sensores según el tipo de variable física medida.

Mediante la combinación de la tabla 7.2 y la tabla 7.4 se obtiene la tabla 7.5 que constituye una guía general para seleccionar el sensor utilizable en diferentes aplicaciones.

		Variable física medida										
		POSICIÓN	DESPLAZAMIENTO	VELOCIDAD	ACELERACIÓN	TAMAÑO	NIVEL	PRESIÓN	FUERZA	PROXIMIDAD	TEMPERATURA	RADIACIÓN LUMINOSA
Principio de funcionamiento	MICRORRUPTORES	X										
	FINALES CARRERA	X				X						
	EXTENSIOMÉTRICOS	X	X	X	X			X	X			
	TERMORRESISTIVOS										X	
	MAGNETORRESISTIVOS	X	X	X								
	CAPACITIVOS	X	X		X		X	X	X	X		
	INDUCTIVOS	X	X	X	X			X	X	X		
	OPTOELECTRÓNICOS	X	X	X						X		
	PIEZOELECTRÍCICOS		X	X	X			X	X			
	FOTOVOLTÁICOS											X
	ULTRASÓNICOS	X					X					

Tabla 7.5 Tabla que indica las variables físicas que se pueden medir con sensores basados en distintos principios de funcionamiento.

7.3 Características de entrada de los sensores industriales

Como se indica en el apartado anterior, la tabla 7.5 proporciona una visión general de los sensores industriales. Pero al elegir el sensor industrial más adecuado para medir una determinada variable o magnitud física, es necesario conocer las características de ésta. Por ello, a continuación se analizan las más importantes.

7.3.1 Campo o rango de medida

Generalmente, un sensor se diseña para responder solamente a una magnitud de entrada determinada (temperatura, presión, etc.) y por ello se especifica por la magnitud que mide y por el rango de las medidas que efectúa.

El campo o rango de medida (*Measurement range or span*) de un sensor se define como el conjunto de valores de la magnitud a medir que están comprendidos dentro de los límites superior e inferior de la capacidad de medida del sensor. Se indica mediante la especificación de los valores extremos. El rango puede ser unidireccional (por ejemplo, "0 a 5 cm"), bidireccional simétrico (por ejemplo, " $\pm 45^{\circ}\text{C}$ "), bidireccional asimétrico (por ejemplo, "-20 a $+80^{\circ}\text{C}$ ") o desplazado (por ejemplo, "50 a 100 Kg/cm²").

El valor máximo que se puede medir con el sensor suele recibir el nombre de salida a fondo de escala [*Full scale output (FSO)*].

El máximo valor de la magnitud a medir que se puede aplicar al sensor sin ocasionarle un cambio de sus características que rebase una tolerancia determinada se denomina sobrerango (*Overrange*) o, a veces, sobrecarga.

7.3.2 Forma de variación de la magnitud de entrada

Otra característica a tener en cuenta en la elección de un sensor es la forma de variación en el tiempo de la magnitud que se desea medir. La importancia de esta característica reside en que pueden producirse grandes errores si la velocidad de los cambios de la variable a medir se sale del margen aceptado por el sensor.

Los sensores se comportan de forma similar a los ojos y el oído humano, que sólo son sensibles a las luces y a los sonidos de un determinado rango de frecuencias. En función del principio de medida en que se basan, hay sensores que sólo detectan magnitudes de variación lenta y en cambio hay otros que responden a variaciones rápidas.

De acuerdo con el modo de variación en función del tiempo, las variables (datos) que se aplican a un sensor se pueden clasificar (Figura 7.20) en estáticos (a), dinámicos (b), transitorios (c) y aleatorios (d), que se analizan seguidamente.

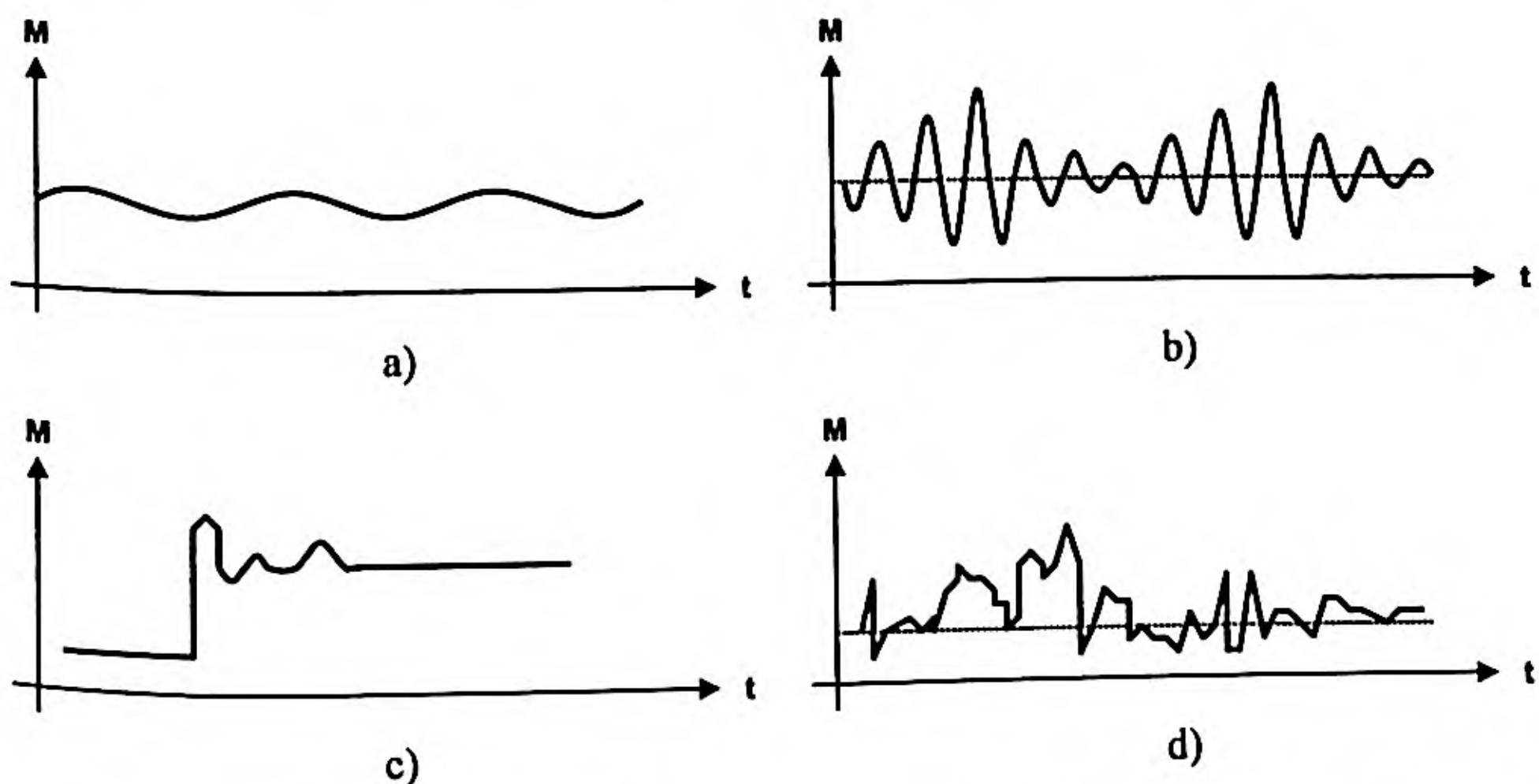


Figura 7.20 Formas de variación a lo largo del tiempo de la magnitud física que debe medir un sensor.

Datos estáticos

Se caracterizan por no presentar variaciones rápidas o discontinuidades y, en general, van asociados a magnitudes cuya evolución en el tiempo implica un contenido en armónicos de baja frecuencia (no superior a una decena de ciclos por segundo). En la figura 7.20a se representa una forma de onda que corresponde a este tipo de datos. Cuando en un sistema se miden varias magnitudes de este tipo no es necesario tratar individualmente cada una de ellas y se pueden utilizar técnicas de muestreo cuya frecuencia está determinada, en la práctica, por la precisión necesaria al tratar de reconstruir las señales eléctricas representativas de cada una de ellas. Estas circunstancias permiten la utilización de un único canal de tratamiento de las señales procedentes de varios sensores cuya salida se mide de forma periódica, lo cual simplifica, y como consecuencia abarata, el sistema de adquisición de datos, descritos en el apartado 8.2.2.5 del capítulo 8.

Datos dinámicos

Son de naturaleza casi siempre periódica y se generan durante el funcionamiento estable y continuo de los sistemas como manifestación de su propio funcionamiento periódico (por ejemplo, órganos giratorios de máquinas, elementos mecánicos en movimiento alternativo, etc.). La medida de este tipo de datos se debe realizar en algunas aplicaciones importantes como, por ejemplo, en el estudio de vibraciones, variaciones de presión en la cámara de combustión de un motor, impactos de prensas, etc. Un ejemplo de forma de onda que corresponde a datos de esta naturaleza es la representada en la figura 7.20b.

Datos transitorios

Se producen en algunos sistemas físicos como consecuencia de alguna perturbación que aparece en un determinado punto de los mismos. Están caracterizados por presentar cambios bruscos que contienen la información principal del comportamiento del sistema. La forma de onda mostrada en la figura 7.20c constituye un ejemplo representativo de una magnitud de este tipo.

Datos aleatorios

Se caracterizan por presentar variaciones no previsibles en sus parámetros fundamentales (amplitud y frecuencia), por lo que su análisis precisa de criterios estadísticos y del estudio de funciones de probabilidad. En la figura 7.20d se representa una forma de onda asociada a datos de esta naturaleza.

Los fabricantes de sensores suelen proporcionar información tanto del rango de medida como de la respuesta a las variaciones de la variable de entrada.

7.4 Características eléctricas

7.4.1 Características eléctricas de salida

Las características eléctricas de salida están ligadas al tipo de formato utilizado y tienen una gran importancia porque de ellas depende la compatibilidad entre un sensor industrial y el sistema acoplado a él. A continuación se analizan para cada uno de los formatos.

7.4.1.1 Sensores de salida analógica

El parámetro de la señal de salida de un sensor analógico (*Analog sensor*) al que está asociada la información puede ser la tensión o la corriente. En el primer caso la *impedancia de salida* (Z_o) del sensor, que se define como la impedancia a través de sus terminales de salida, debe ser mucho menor que la impedancia de entrada (Z_i) de la carga conectada a él, que puede ser un aparato de medida o un procesador electrónico (Véase el apartado 8.2.2.1.2 del capítulo 8). Se logra así que resulte despreciable la caída de tensión en los cables que unen el sensor con la carga pero surge el inconveniente de que se tiene una baja inmunidad al ruido y además se puede producir la destrucción de la carga debido a la presencia de señales espúreas de elevada tensión en los cables, en especial si su longitud es elevada.

Los márgenes de tensión de salida más comunes de los sensores analógicos son, de 0 a 10 V, de 1 a 5V, de -5 a +5 V, y de -10 a + 10 V. En ocasiones, el circuito de acondicionamiento forma parte del circuito de interfaz entre el sensor y el procesador digital. En este caso, el sensor (por ejemplo un termopar) está formado solamente por el elemento sensor propiamente dicho que proporciona una señal analógica de un nivel de tensión del orden de milivoltios.

En general, la salida en tensión es adecuada para transmitir la información cuando la distancia que separa al sensor del equipo electrónico conectado a él es reducida. Los inconvenientes de la transmisión en tensión hacen, más adecuado técnicamente que la salida del sensor constituya una fuente de corriente que transmita a los cables una corriente proporcional a la magnitud a medir. Para ello la impedancia de salida del sensor ha de ser mucho más elevada que la de entrada de la carga (Véase el apartado 8.2.2.1.2 del capítulo 8). La ventaja de este método consiste en que, por ser la corriente constante para cada valor de dicha magnitud, la transmisión no está influenciada por la variación de la impedancia de los cables porque la caída de tensión en ellos no tiene ningún efecto, ya que la corriente en todos los puntos de un circuito serie es la misma. Por otra parte, el hecho de que la impedancia de entrada de la carga sea mucho menor que la de salida del sensor hace que las señales parásitas (ruido) que aparecen en ella, debido a la presencia de interferencias (ruido) sobre los cables, sean de muy pequeña amplitud. Los sensores analógicos con salida por corriente están contruidos de tal manera que admiten un cortocircuito permanente en su salida.

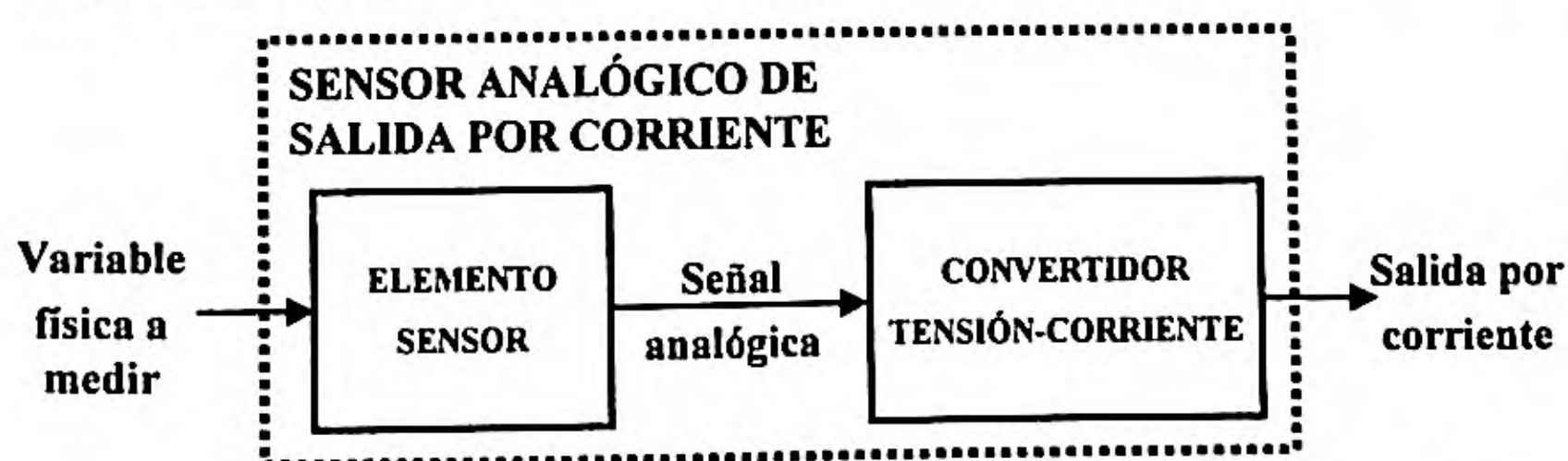


Figura 7.21 Esquema de bloques típico de un sensor analógico de salida por corriente.

El interés de la transmisión por corriente propició la realización de convertidores tensión-corriente y su desarrollo en circuito integrado, así como su incorporación en el propio sensor, formando parte del circuito electrónico acondicionador (Figura 7.21). Se han normalizado los niveles de corriente -20 a +20mA, 4 a 20mA y 0 a 20mA. En la figura 7.22 se representa el sensor ultrasónico de proximidad de tipo eco (descrito en el apartado 7.7.2.3.7.4) 3RG61 13-3BF01 de salida analógica de corriente de 4 a 20mA.

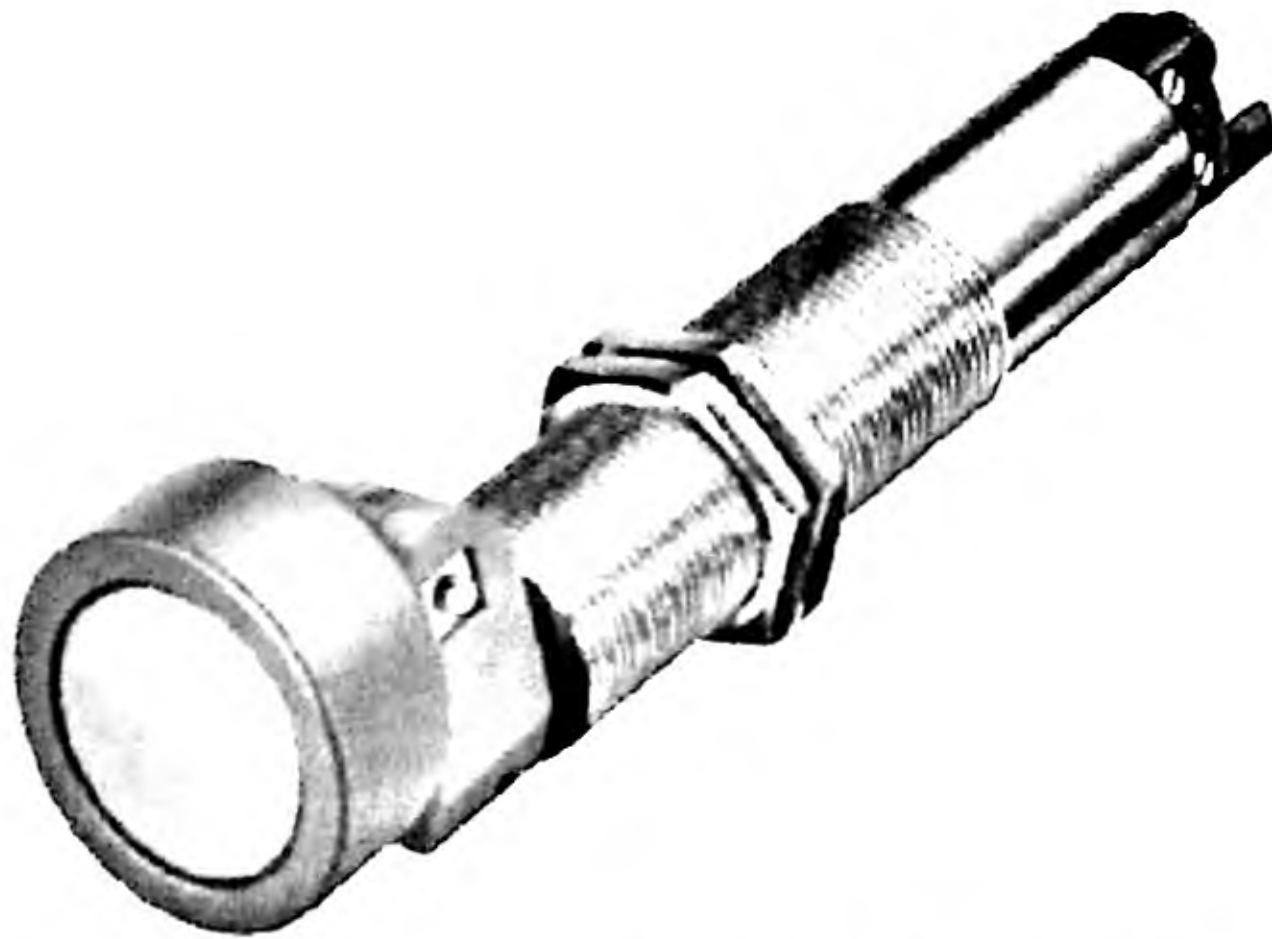


Figura 7.22 Sensor ultrasónico de proximidad del tipo 3RG61 13-3BF01, que posee una salida analógica por corriente de 4 a 20ma (cortesía de Siemens).

7.4.1.2 Sensores de salida digital

El parámetro más importante de los sensores cuya señal de salida posee el formato digital (*Digital sensors*) es la *corriente de carga máxima* que se define como la máxima corriente que puede circular, en uno u otro sentido, a través del terminal de salida. Dicha corriente depende, en general, de la tensión de alimentación y suele recibir la denominación de *cargabilidad de salida*. Los niveles de tensión correspondientes a los estados diferenciados que puede adoptar la salida constituyen otra característica de gran importancia cuando la señal de salida del sensor se aplica a la entrada de otros circuitos electrónicos digitales (TTL, CMOS, etc.), en los que los niveles del 0 y del 1 sólo pueden tomar valores comprendidos dentro de un determinado margen de tensión. Para asegurar la compatibilidad entre los niveles de tensión y corriente de la salida del sensor y la entrada del circuito electrónico conectado a él es necesario conocer también la configuración de su etapa de salida que puede presentar diversas variantes, de las que las más utilizadas son:

- **Salida con transistor NPN y resistencia de carga**

La configuración de cada terminal de salida es la representada en la figura 7.23a. Se caracteriza por proporcionar niveles de tensión y de corriente compatibles con las dos familias lógicas de uso más extendido que son la TTL y la CMOS.

- **Salida con transistor NPN y colector abierto**

Corresponde al esquema de la figura 7.23b que se diferencia de la anterior en que no tiene incorporada la resistencia R, debido a lo cual recibe la denominación de colector abierto (*Open collector*). Mediante la colocación de una resistencia R externa de valor adecuado, la salida se puede conectar a circuitos lógicos de diferentes tecnologías (TTL, CMOS, etc.).

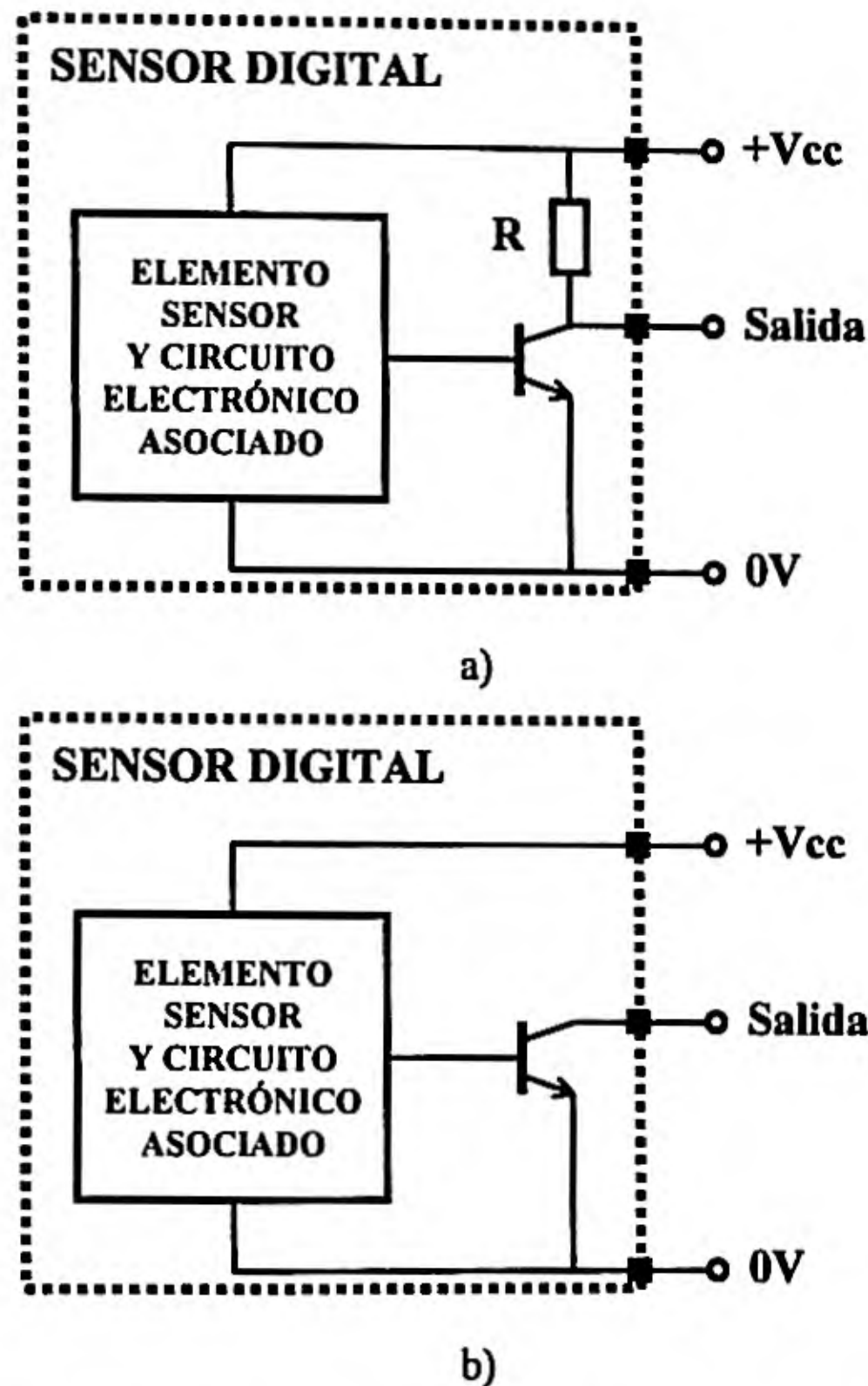


Figura 7.23 Etapas de salida de sensores digitales: a) Con transistor NPN y resistencia de carga; b) Con transistor NPN en montaje de colector abierto.

7.4.1.3 Sensores de salida todo-nada

Los sensores de salida todo-nada (*On-off sensors*) proporcionan solamente, tal como se indica en el apartado 7.2.4, dos niveles de tensión a su salida y por ello la configuración de esta última guarda cierta similitud con la de los sensores digitales. Los parámetros eléctricos más importantes de este tipo de sensores son:

- **La corriente de carga máxima (*Maximum load current*)**
Se define como la máxima corriente que puede circular, en uno u otro sentido, a través del terminal de salida.
- **La tensión de alimentación (*Supply voltage*)**
Puede ser continua o alterna. Los alimentados en alterna llevan incorporado un rectificador y un circuito estabilizador que genera la tensión de alimentación del sensor y del circuito asociado con él. Los alimentados en continua pueden ser polarizados o no polarizados. A los polarizados se les debe aplicar la tensión de alimentación en un sentido determinado y, en el caso de los no polarizados, dicho sentido es indiferente.

- **La configuración de salida**

Además de conectarse a un procesador digital, los sensores todo-nada deben poder hacerlo a otros tipos de cargas (por ejemplo un relé, una electroválvula, etc.). Por ello existen sensores todo-nada con salida de tipo relé o con salida electrónica que se estudian a continuación.

Por otra parte, la forma de conectar los sensores todo-nada a un autómata programable se estudia en los apartados 8.2.2.2 y 8.2.2.3 del capítulo 8.

Sensores todo-nada de salida de tipo relé

Las salidas de estos sensores son contactos libres de potencial que se representan gráficamente en la figura 7.24. Pueden poseer un solo contacto no conmutado que recibe la denominación de SPST (*Single Pole Single Through*) normalmente abierto o normalmente cerrado. El contacto SPST normalmente abierto denominado NO (*Normally Open*) está abierto cuando no pasa corriente por la bobina del relé y está cerrado en el caso contrario. El contacto SPST normalmente cerrado denominado NC (*Normally Close*) está cerrado cuando no pasa corriente por la bobina del relé y abierto en el caso contrario. También pueden poseer un contacto que conmuta entre dos terminales (doble vía) y recibe la denominación de SPDT (*Single Pole Double Through*). Tanto los contactos no conmutados como los conmutados pueden ser dobles o incluso triples. Los dobles no conmutados se denominan DPST (*Double Pole Single Through*) y los conmutados DPDT (*Double Pole Double Through*). La figura 7.24 resulta autoexplicativa.

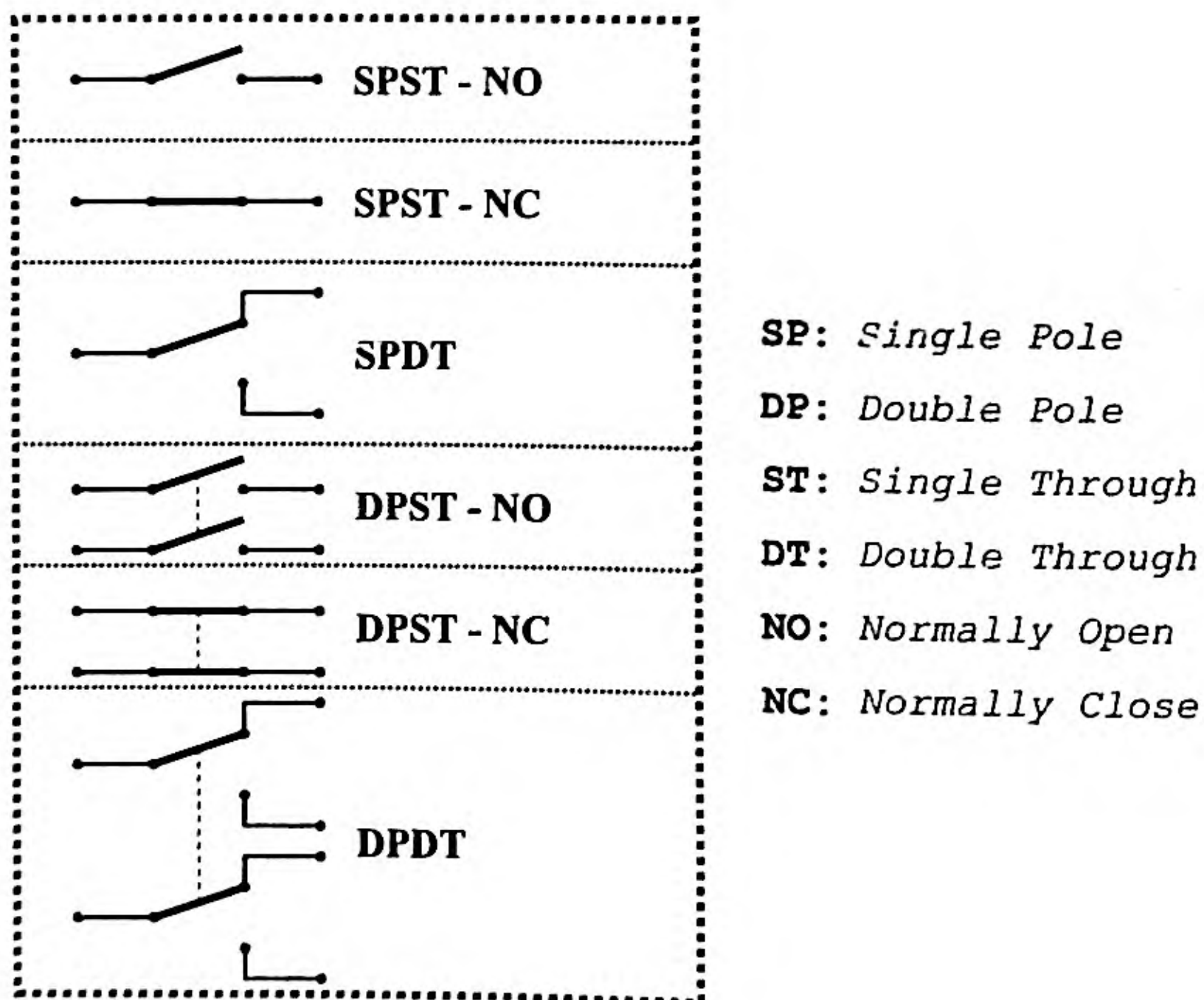


Figura 7.24 Diferentes tipos de contactos de un relé.

El circuito típico de un sensor de este tipo se indica en la figura 7.25. El circuito del sensor activa o desactiva al relé y en consecuencia éste conmuta sus contactos. Los sensores cuya salida es de este tipo se caracterizan porque el tiempo de activación o desactivación es mayor (del orden de 30 ms) que el de los que tienen una salida del tipo transistor. Si la salida de este sensor se conecta a un sistema electrónico es necesario eliminar los rebotes que se producen al conmutar el contacto. Dicha eliminación se puede realizar mediante el circuito adecuado [MAND 95]. Un sensor de salida del tipo relé es el sensor optoelectrónico de proximidad BERO 3RG72 11-6MC00 de Siemens (Figura 7.26).

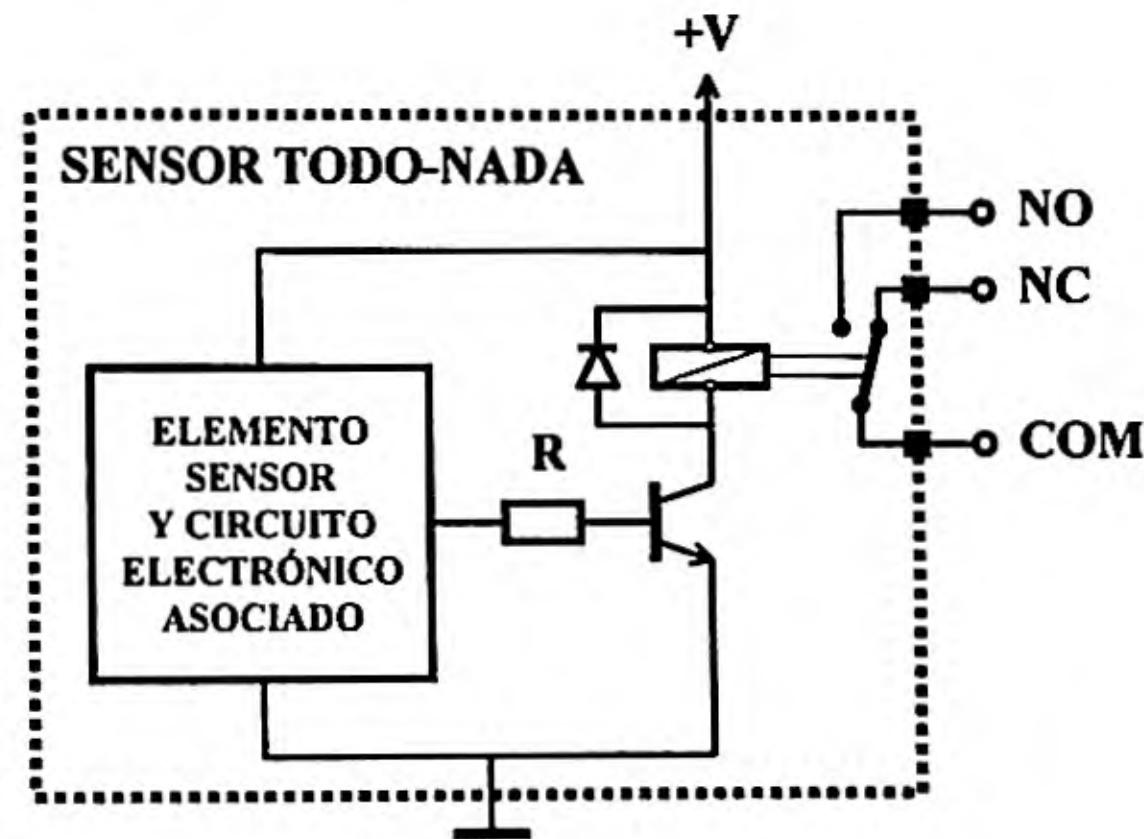


Figura 7.25 Sensores todo-nada con salida mediante relé.

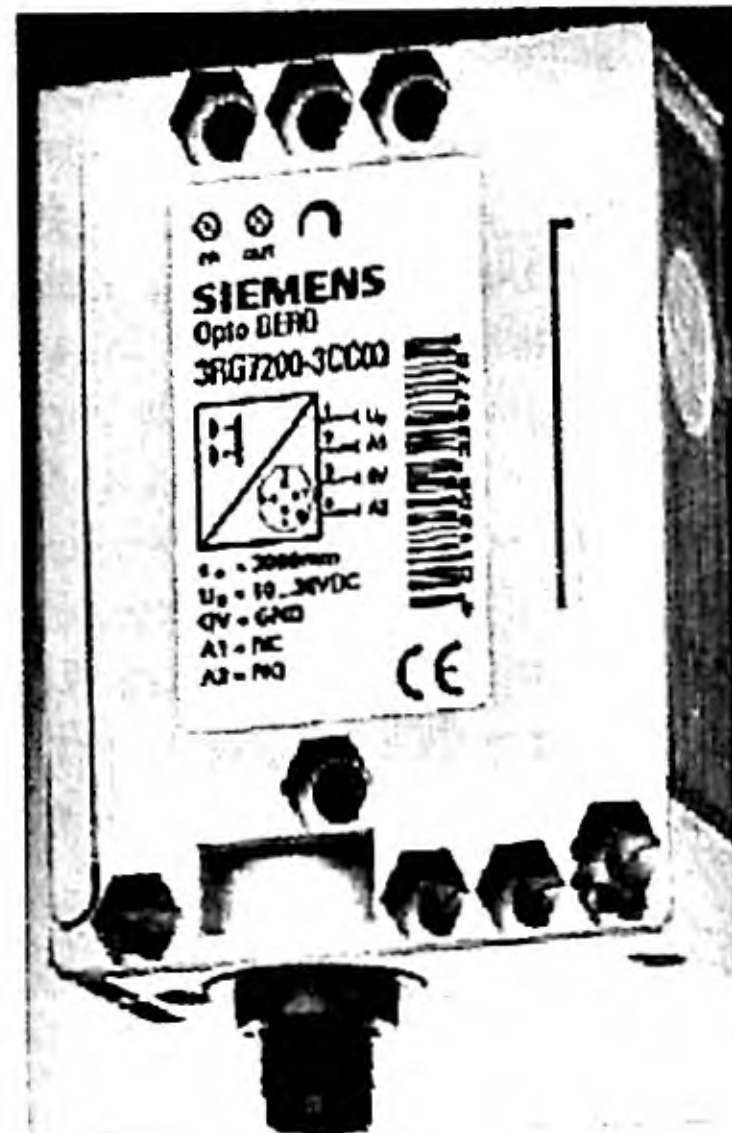


Figura 7.26 Sensor optoelectrónico de proximidad BERO 3RG72 11-6MC00 (cortesía de Siemens).

Otra característica importante de este tipo de sensores es que permiten la conexión de cargas tanto en alterna como en continua. Si la carga es inductiva es necesario proteger el contacto mediante un circuito RC o un diodo de libre circulación, tal como se indica en el apartado 8.2.2.4.2 del capítulo 8. Si la carga es una lámpara de filamento se debe tener en cuenta la máxima corriente que la atraviesa cuando el filamento está frío en el instante en que se aplica la tensión.

Sensores todo-nada de salida electrónica

Los sensores todo-nada de salida electrónica utilizan, como su nombre indica, dispositivos electrónicos en su salida en lugar de un relé. Existen diversas variantes que se distinguen por:

- **El tipo de dispositivo electrónico utilizado**
Puede ser un transistor NPN o PNP en los sensores alimentados en continua y un tiristor o un triac en los alimentados en alterna.
- **La asignación de los niveles de la salida a los valores de la variable que actúa sobre el sensor**
Se suelen utilizar las denominaciones de normalmente abierto (NO) y normalmente cerrado (NC) por similitud con los sensores con salida de tipo relé. La citada asignación depende del tipo de sensor y en el apartado 7.7.2.3.2 se describe para los sensores de proximidad sin contacto.
- **El número de terminales de salida del sensor**
Pueden ser dos, tres o cuatro. Los sensores correspondientes reciben los nombres de sensores de dos hilos, de tres hilos y de cuatro hilos, respectivamente.

Sensores todo-nada de dos hilos

Se caracterizan por tener dos terminales a través de uno de los cuales no sólo se alimenta el sensor sino que además se conecta la carga (Figura 7.27). En este tipo de sensores la corriente que circula por la carga I_L es en todo momento igual a la corriente I_S que circula por el sensor. Cuando la carga no debe activarse (porque el nivel de la magnitud de entrada del sensor es el adecuado para ello) la corriente de alimentación del sensor sigue pasando por ella. Si en este último caso la carga conectada al sensor es un relé la corriente de alimentación del sensor es insuficiente para activarlo. Pero, si por el contrario, la carga es la entrada de un autómata programable, puede ocurrir que la corriente de alimentación del sensor sea suficiente para que el autómata considere que la salida del sensor está activada cuando en realidad no es así. Los fabricantes de autómatas programables suelen indicar, para sus diferentes modelos de tarjetas o placas de entrada, cual es el valor máximo de la citada corriente. Por ejemplo Siemens establece que la corriente del sensor debe ser menor de 1,5mA para que no llegue a activar la entrada de un autómata programable cuando la salida del sensor está desactivada.

Los sensores de dos hilos se pueden diseñar para funcionar igual que un contacto normalmente abierto "NO" (*Normally Open*) o que un contacto normalmente cerrado "NC" (*Normally Close*). En el primer caso la corriente no circula por la carga cuando la variable física que mide el sensor tiene el nivel correspondiente a "nada" (es decir, está desactivada) y por el contrario circula por la carga cuando tiene el nivel correspondiente a "todo" (es decir, está activada). En el segundo se produce la situación contraria. Por ejemplo, en un sensor inductivo de proximidad NO no pasa corriente por la carga cuando no detecta la presencia de un objeto y lo hace en el caso contrario.

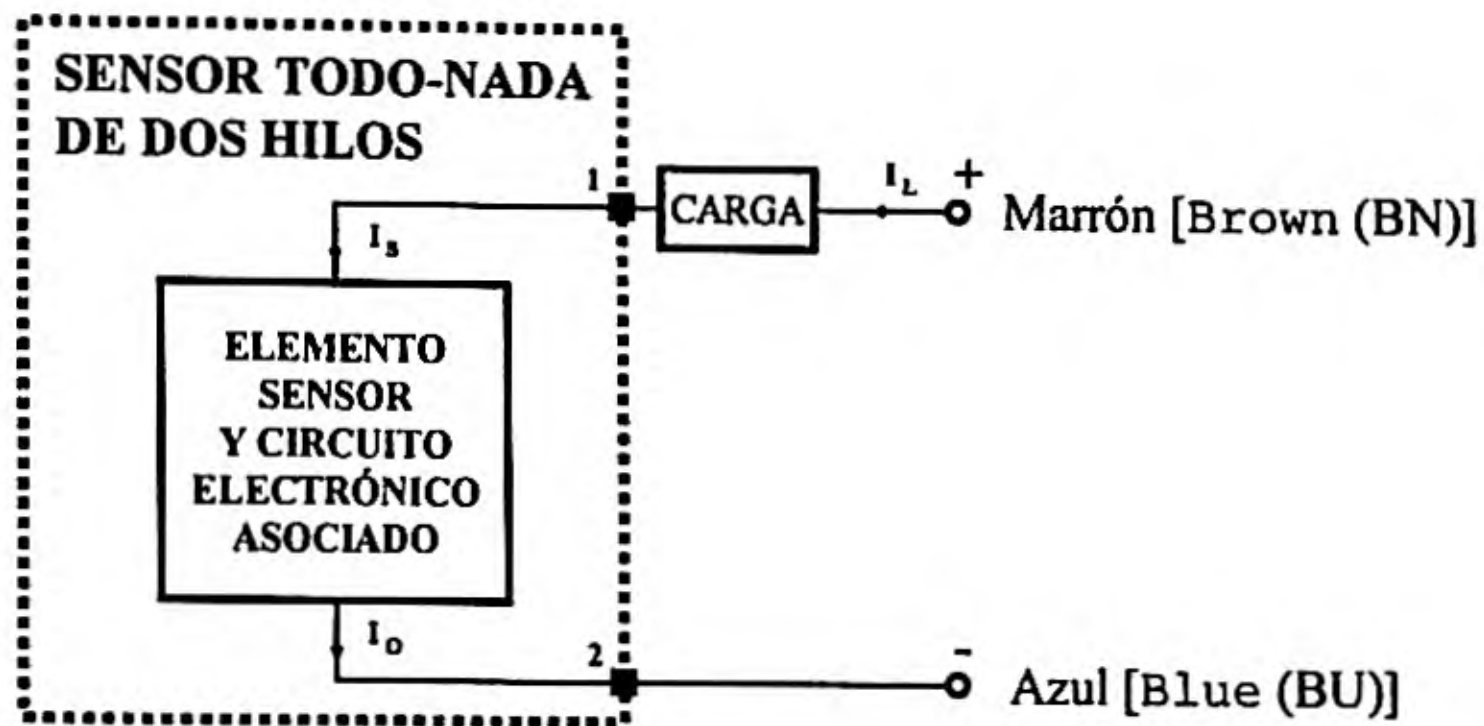


Figura 7.27 Sensor todo-nada de dos hilos.

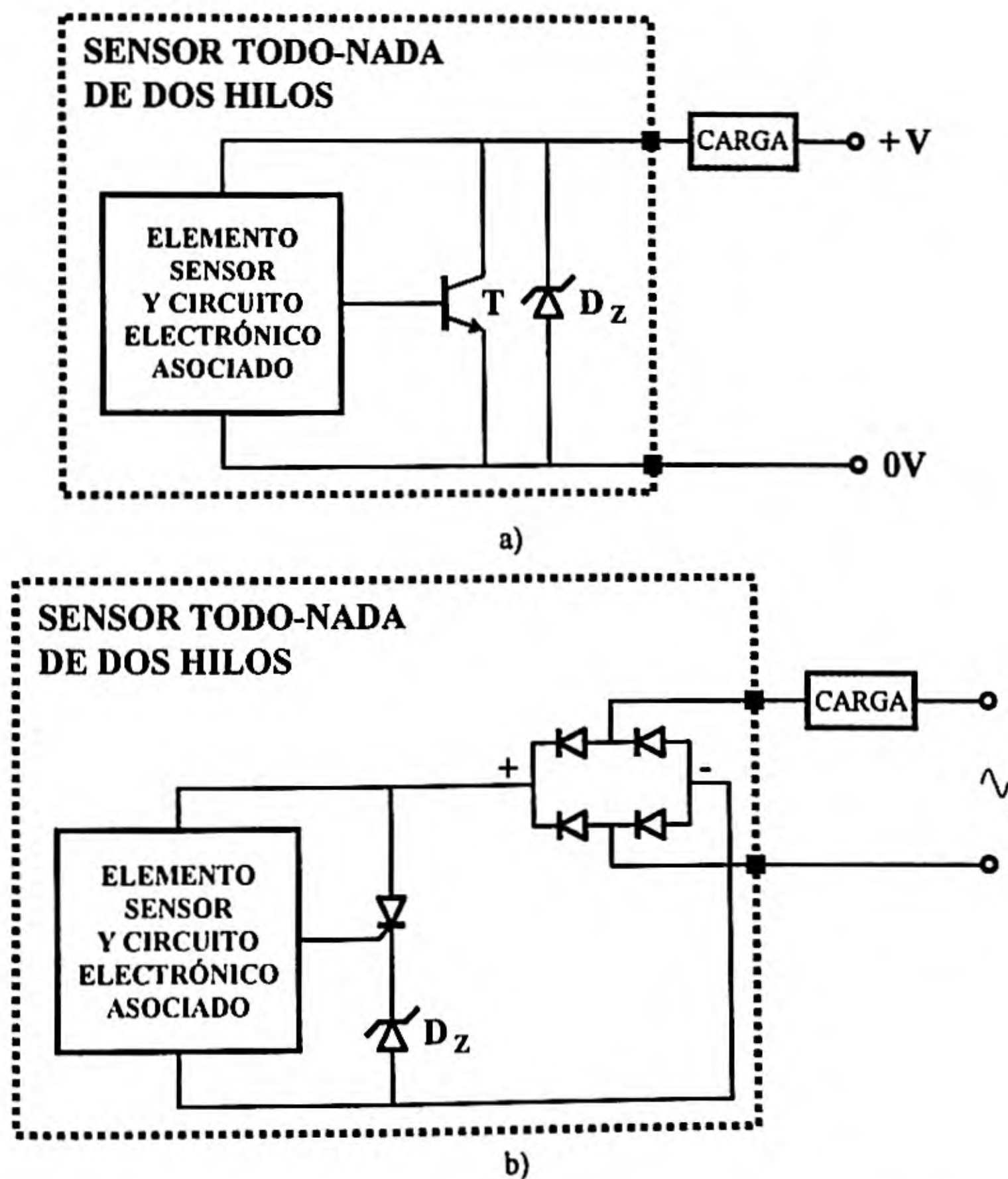


Figura 7.28 Sensores todo-nada de dos hilos: a) Alimentado en continua; b) Alimentado en alterna.

Los sensores de dos hilos pueden ser realizados para ser alimentados en continua o en alterna. En la figura 7.28a se representa el circuito típico de salida de un sensor de dos hilos alimentado en continua en el que el transistor T no alcanza la saturación cuando se activa. De esta forma la tensión entre su colector y su emisor mantiene alimentado el sensor. En la figura 7.28b se representa el circuito de salida de un sensor de dos hilos que se alimenta en alterna a través de un puente rectificador. El tiristor constituye el elemento conmutador accionado por el sensor, mientras que el diodo zener D_z mantiene la tensión de alimentación del sensor en el nivel adecuado para su funcionamiento cuando conduce el tiristor. Un ejemplo de sensor de dos hilos con alimentación en continua (CC) y salida normalmente abierta (cuando el sensor se activa deja pasar más corriente) es el sensor inductivo de proximidad 3RG40 11-7JB00 (Figura 7.29). Un ejemplo de sensor de dos hilos con alimentación en corriente alterna (CA) o continua (CC) y salida normalmente cerrada es el sensor inductivo de proximidad 3RG40 12-0KA00 (Figura 7.30). Los sensores inductivos de proximidad se estudian en el apartado 7.7.2.3.5.

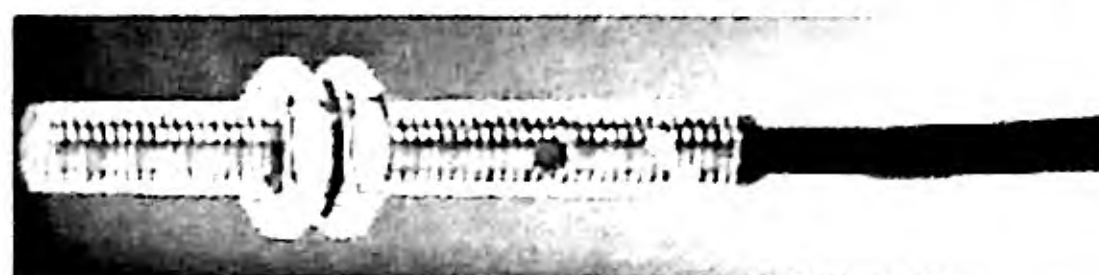


Figura 7.29 Sensor inductivo de proximidad BERO 3RG40 11-7JB00 con alimentación en continua y salida normalmente abierta (cortesía de Siemens).

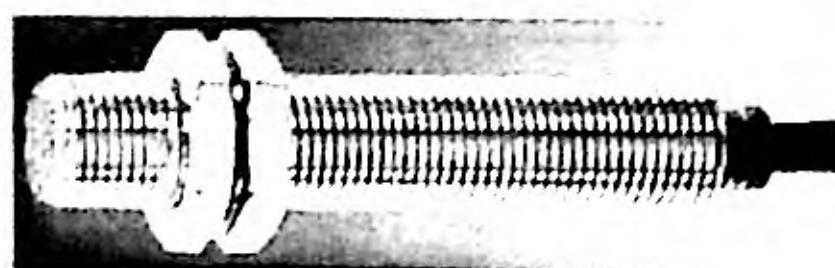


Figura 7.30 Sensor inductivo de proximidad BERO 3RG40 12-0KA00 con alimentación en alterna o en continua y salida normalmente cerrada (cortesía de Siemens).

Sensores todo-nada de tres hilos

Se caracterizan por tener tres terminales de salida (Figura 7.31), a través de dos de los cuales se proporciona la tensión de alimentación al sensor y la carga se conecta entre el tercero y uno de los otros dos. La corriente I_L que circula a través de la carga es prácticamente nula o tiene un valor apreciable según que el sensor esté o no activado respectivamente. Según el tipo de dispositivo situado a la salida del sensor, tal como se indica más adelante, la carga se conecta entre la salida y el positivo de la fuente de alimentación (Figura 7.31a) o entre ella y el negativo de la fuente de alimentación (Figura 7.31b). El primer caso se puede realizar con un transistor bipolar NPN, tal como se indica en las figuras 7.32a y 7.32b, y el segundo con un transistor bipolar PNP (Figuras 7.33a y 7.33b). La figura 7.32a y la 7.33a se diferencian de la 7.32b y la 7.33b, respectivamente, en que poseen una resistencia R , que garantiza una corriente mínima a través del transistor de salida. Los cuatro circuitos de salida de las figuras 7.32 y 7.33 tienen sendos diodos de protección D y D_z . El diodo D protege al sensor contra la inversión de la polaridad de la tensión de alimentación y el diodo zener D_z impide que la tensión de salida se eleve por encima del valor de la tensión de zener del mismo.

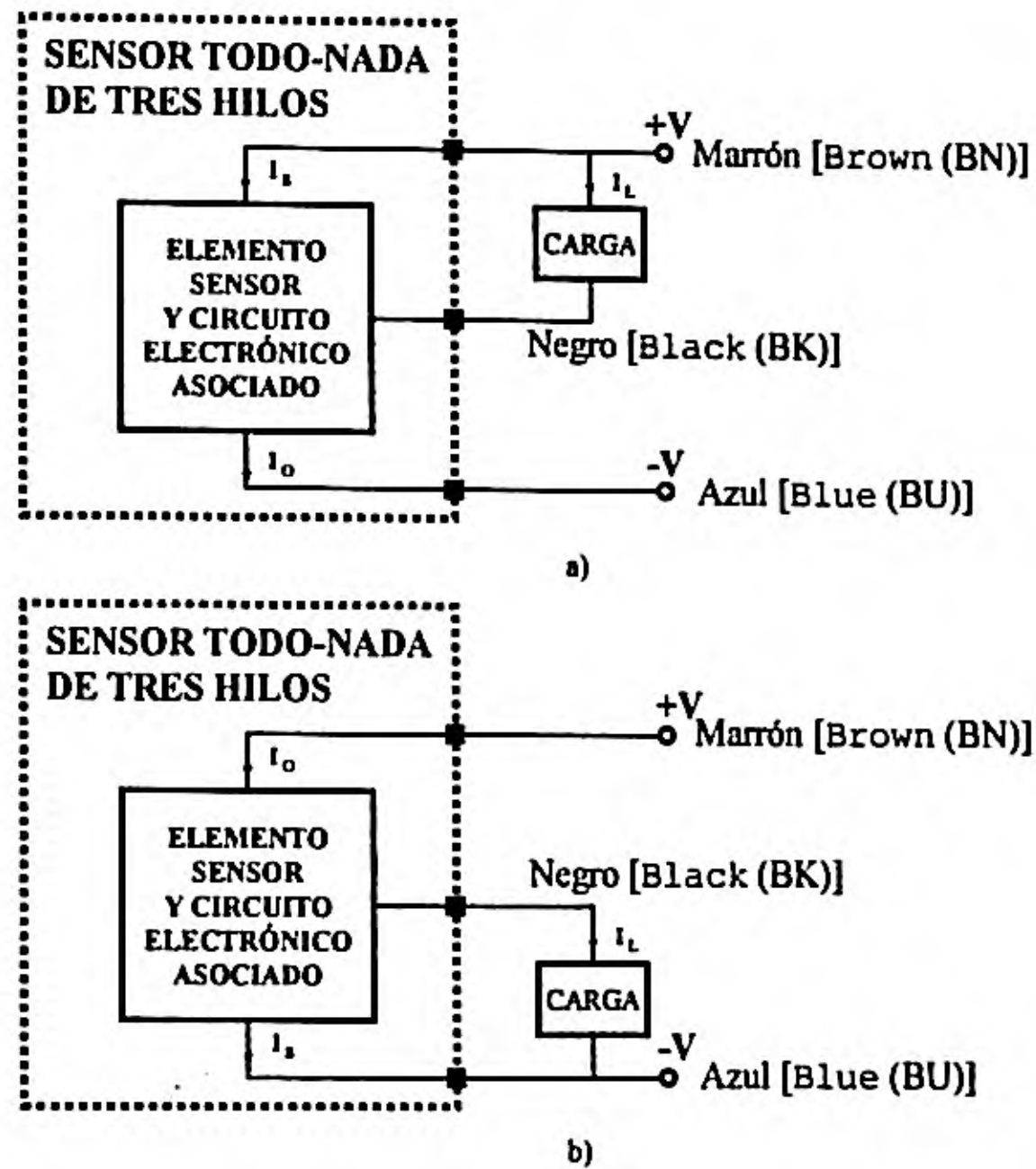


Figura 7.31 Sensores todo-nada de tres hilos: a) Carga conectada al positivo de la fuente de alimentación; b) Carga conectada al negativo de la fuente de alimentación.

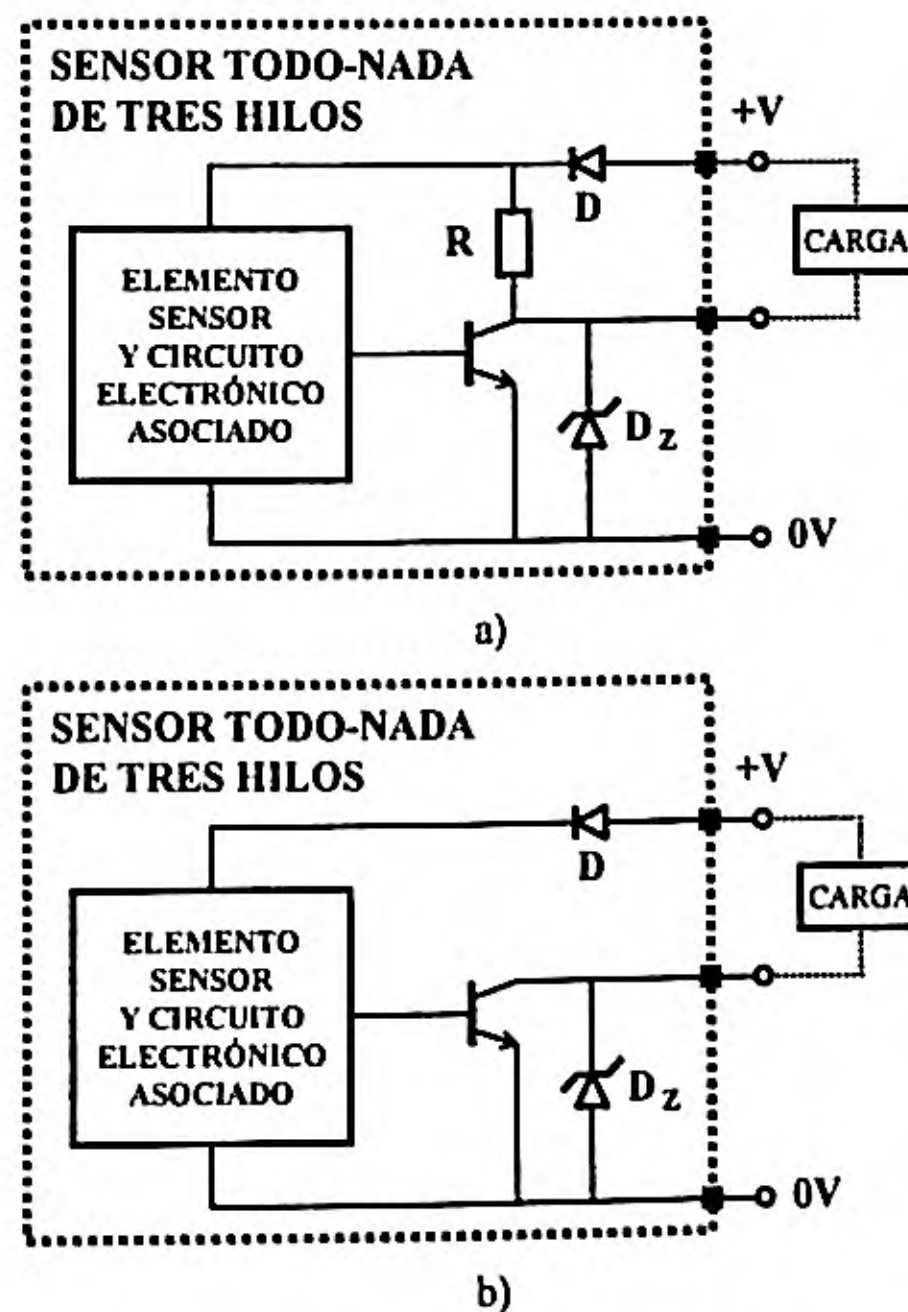


Figura 7.32 Sensores todo-nada de tres hilos cuya salida está realizada con transistores NPN: a) Con resistencia de carga; b) Sin resistencia de carga (colector abierto).

Tanto los sensores implementados con transistores NPN como PNP se pueden diseñar para funcionar igual que un contacto normalmente abierto "NO" (*Normally Open*) o que un contacto normalmente cerrado "NC" (*Normally Close*).

En el primer caso, el transistor está cortado (la corriente no circula por la carga), cuando el sensor no está activado. Por ejemplo, si es un sensor inductivo de proximidad, no detecta el objeto. Por el contrario, el transistor está saturado (la corriente circula por la carga) cuando el sensor detecta el objeto. En el segundo caso se produce la situación contraria. Por ejemplo, en un sensor inductivo de proximidad NO no pasa corriente por la carga cuando no detecta la presencia de un objeto y lo hace en el caso contrario.

La figura 7.34 muestra un sensor de proximidad inductivo de tres hilos BERO 3RG46 10-76G00 cuya salida utiliza un transistor NPN, y la figura 7.35 muestra un sensor de proximidad inductivo de tres hilos BERO 3RG46 00-7AB00 cuya salida utiliza un transistor PNP.

Los sensores de tres hilos se caracterizan porque su corriente I_s de alimentación no circula por la carga (Figura 7.31). Por ello se pueden conectar directamente a la entrada de un sistema electrónico digital de control, como por ejemplo a un autómata programable, tal como se indica en el capítulo 8 dedicado a los circuitos de interfaz de los autómatas programables. En ambos casos la corriente que suministra la fuente de alimentación es la suma de la corriente I_s que consume el sensor y la I_L que pasa por la carga.

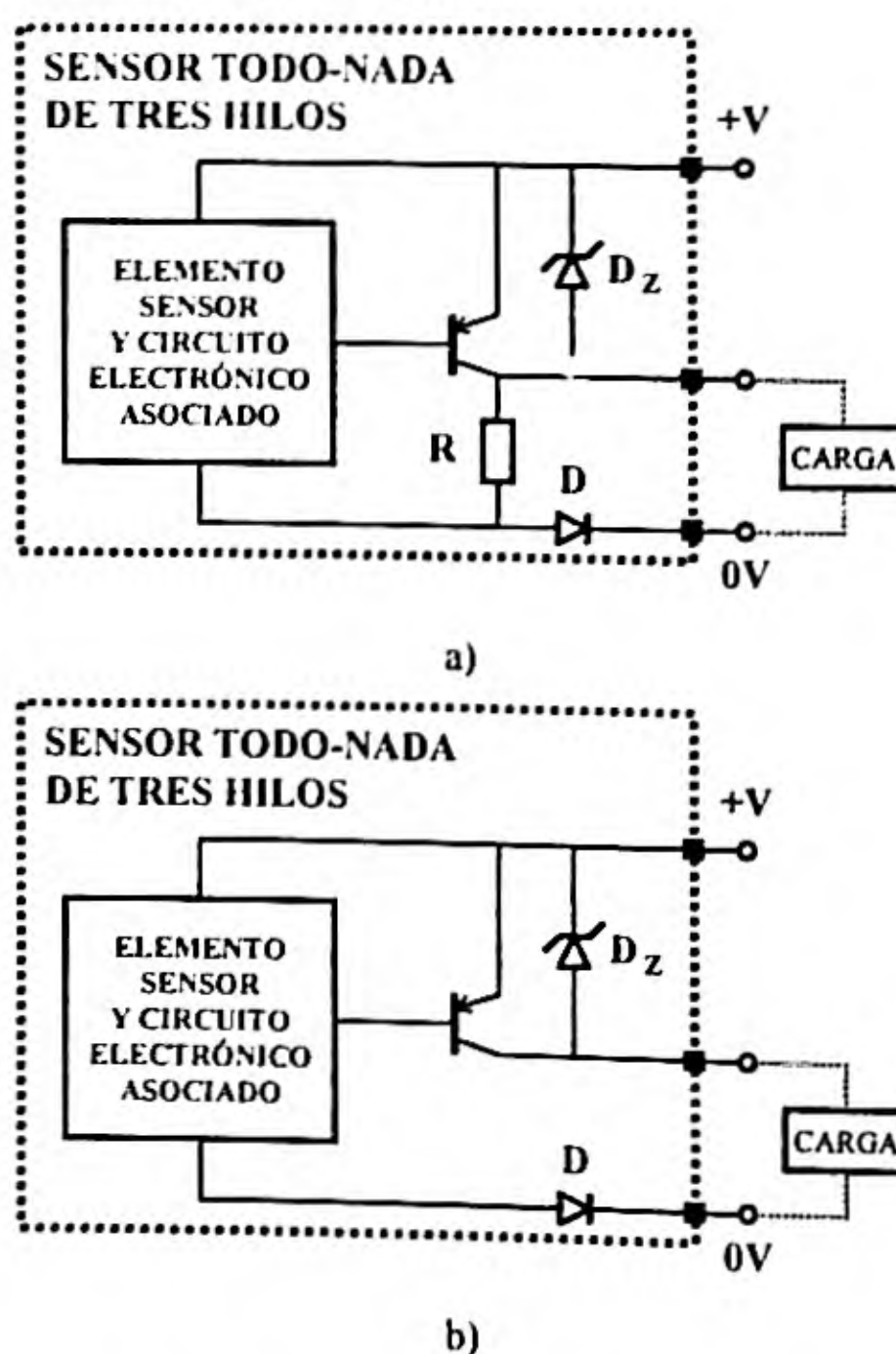


Figura 7.33 Sensores todo-nada de tres hilos cuya salida esta realizada con transistores PNP: a) Con resistencia de carga; b) Sin resistencia de carga (colector abierto).

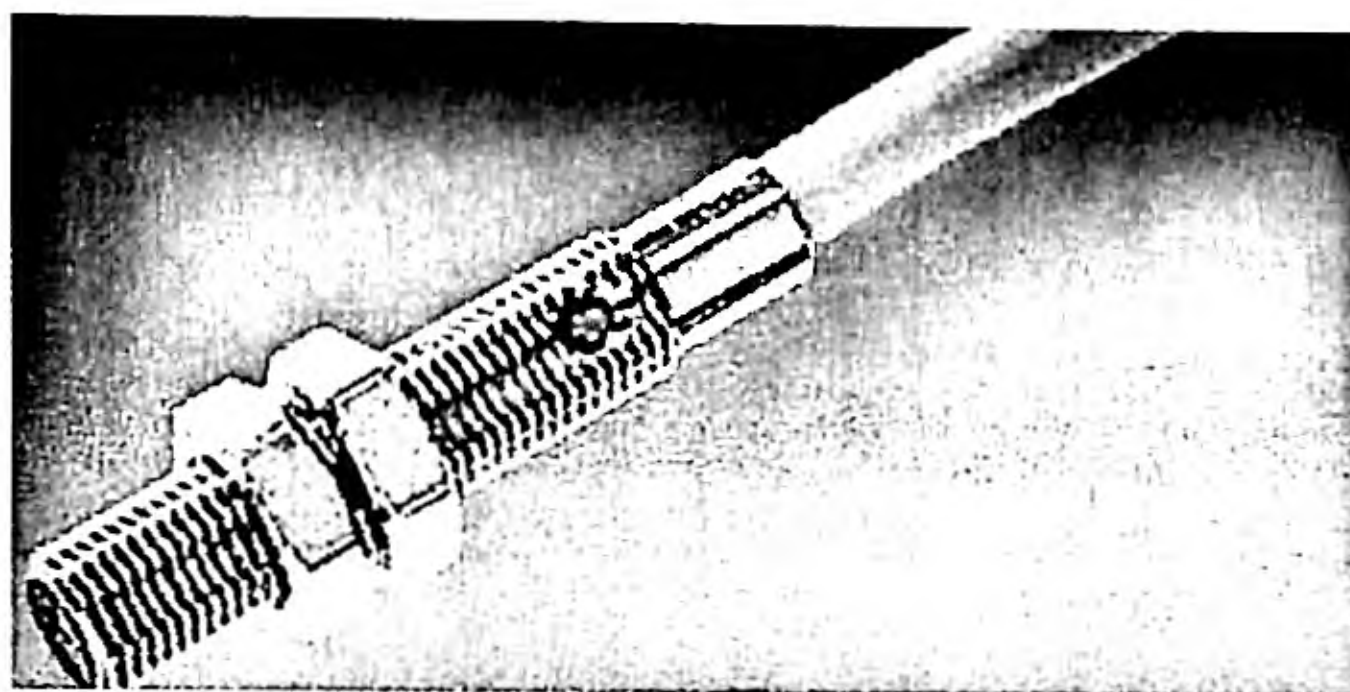


Figura 7.34 Sensor de proximidad inductivo de tres hilos 3RG46 10-7GB00 realizado con un transistor NPN (cortesía de Siemens).

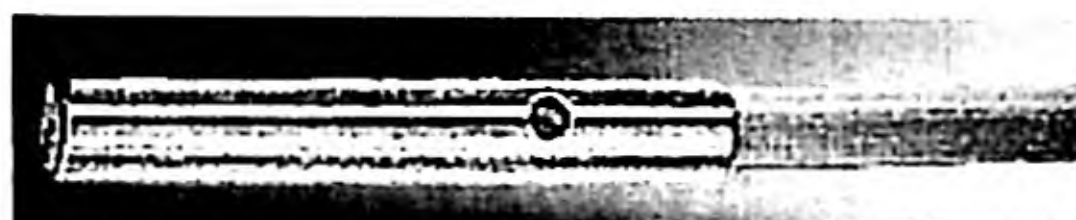


Figura 7.35 Sensor de proximidad inductivo de tres hilos 3RG46 00-7AB00 realizado con un transistor PNP (cortesía de Siemens).

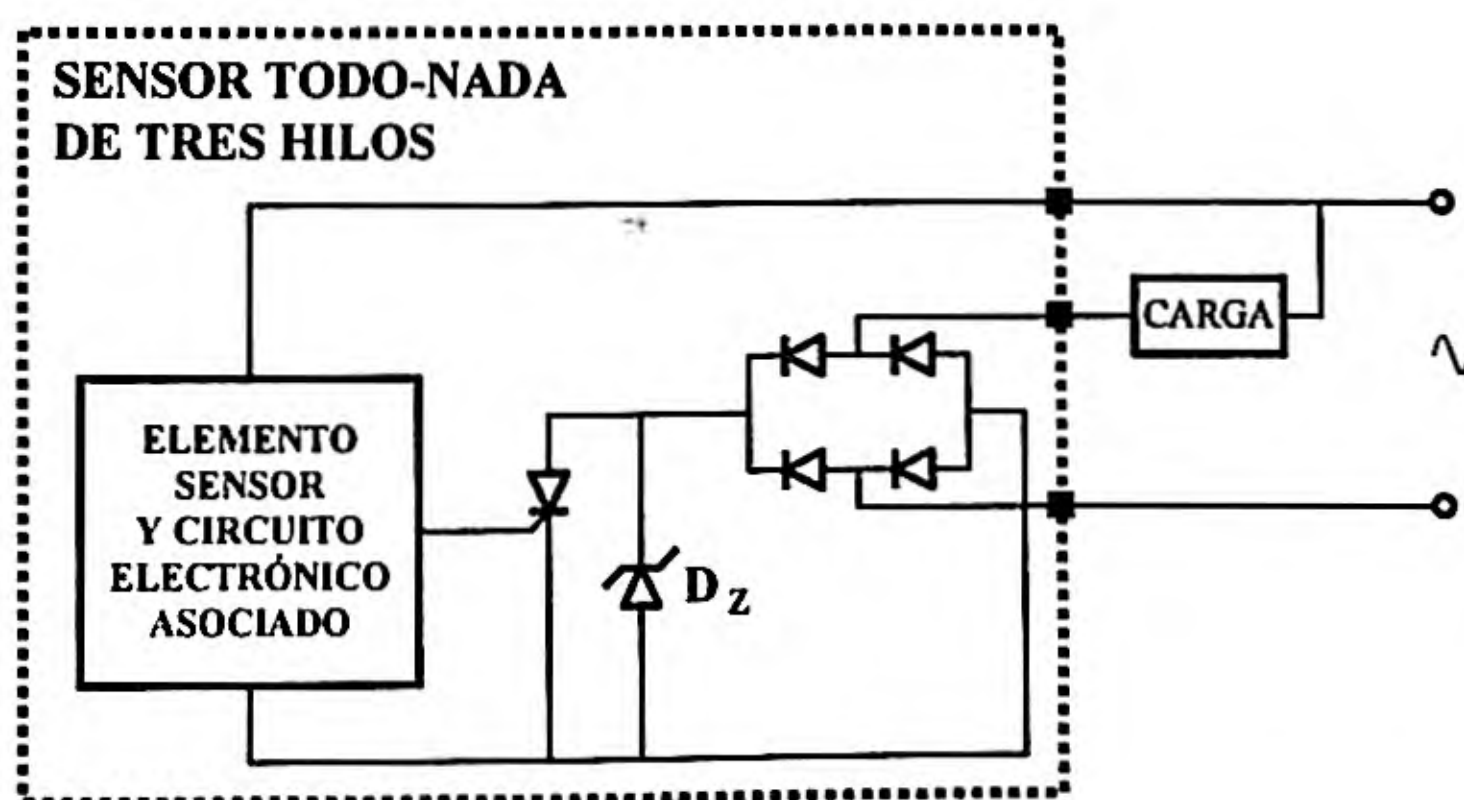


Figura 7.36 Sensor todo-nada con salida de tres hilos y alimentación en alterna.

Los sensores de tres hilos también pueden ser alimentados en alterna. Un circuito típico es el indicado en la figura 7.36. En este caso el circuito electrónico asociado al sensor incluye un rectificador junto con el circuito de filtraje y estabilización. Además la carga se alimenta a través de un rectificador. Un ejemplo de sensor de tres hilos alimentado en alterna es el sensor capacitivo de proximidad 3RG16 14-0LB00 de la familia BERO (Figura 7.37). Los sensores capacitivos de proximidad se estudian en el apartado 7.7.2.3.6.



Figura 7.37 Sensor capacitivo de proximidad BERO 3RG16 14-0LB00 de tres hilos alimentado en alterna (cortesía de Siemens).

Sensores todo-nada de cuatro hilos

Se caracterizan por disponer de dos terminales de salida. Su configuración es similar a la de los de tres hilos, con la particularidad de que en este caso tienen dos salidas, que pueden estar realizadas ambas con transistores bipolares del tipo PNP (Figura 7.38a) o del tipo NPN (Figura 7.38b), una de las cuales es normalmente cerrada (NC) y la otra normalmente abierta (NO). Un ejemplo de sensor de este tipo es el sensor capacitivo de proximidad 3RG16 30-6AC00 de Siemens (Figura 7.39), que posee dos salidas de tipo PNP. Otra opción en este tipo de sensores es que dispongan de una salida PNP y otra NPN (Figura 7.40). Un ejemplo de este tipo lo constituye el sensor de proximidad optoelectrónico 3RG76 00-3RH00 (Figura 7.41).

Otras características de interés, relacionadas con la salida de los sensores todo-nada, son las siguientes:

- **Corriente residual**

La corriente residual (*Off-state leakage current*) se define como la corriente que circula a través de la carga cuando el sensor está desactivado (el transistor de salida no conduce).

- **Tensión residual**

La tensión residual (*Voltage drop*) es el valor máximo de la tensión que puede haber en la salida del sensor cuando se encuentra activado (el transistor de salida conduce).

- **Corriente de carga mínima**

En algunos sensores es necesario, para que funcionen correctamente, que circule a través de ellos una corriente de carga mínima. Si no se alcanza dicho valor el fabricante no garantiza su correcto funcionamiento. Un ejemplo de ello es una salida en alterna realizada con un tiristor que necesita una corriente mínima de cebado.

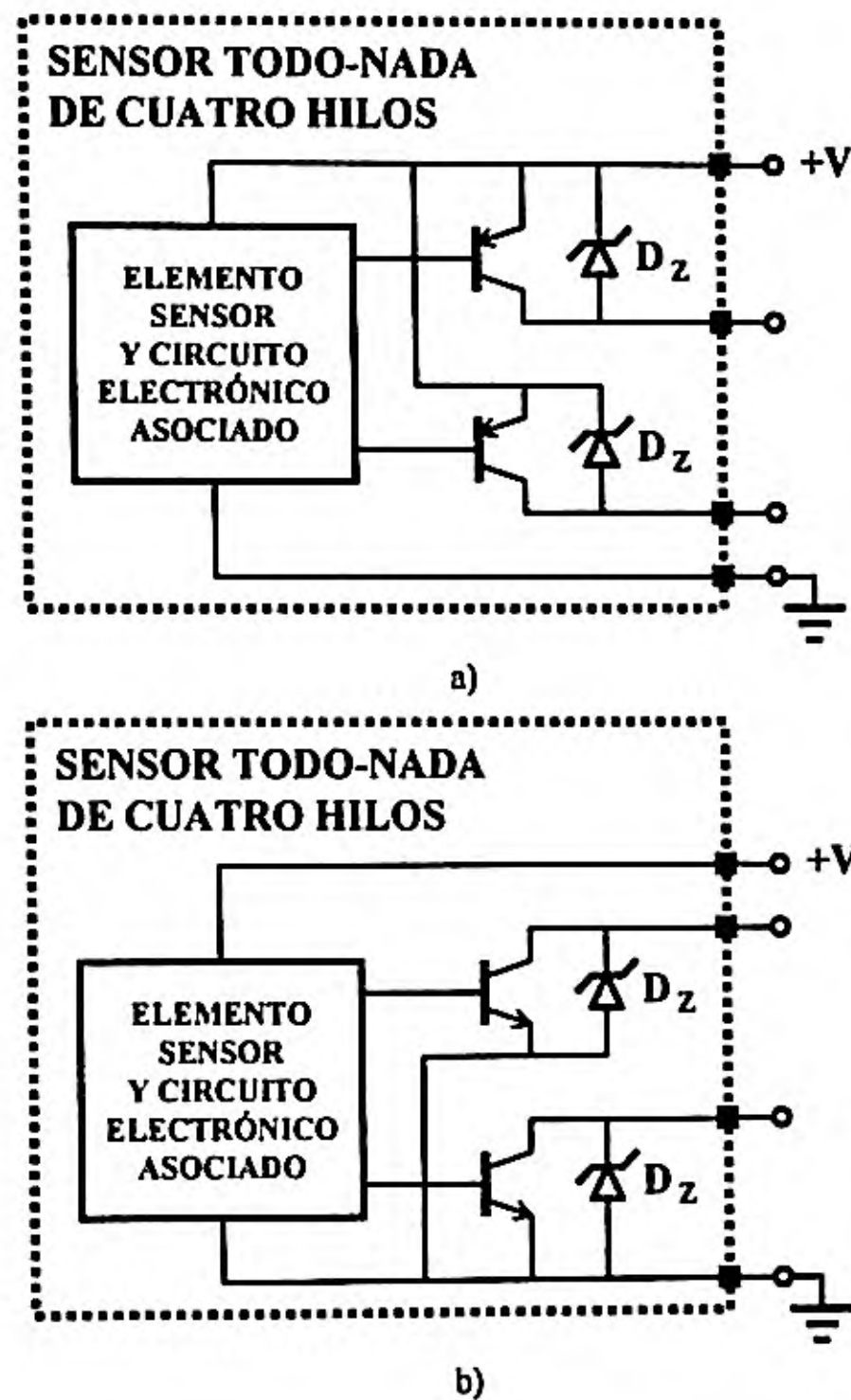


Figura 7.38 Sensores todo-nada de cuatro hilos: a) Realizado con transistores PNP; b) Realizado con transistores NPN.



Figura 7.39 Sensor capacitivo de proximidad BERO 3RG16 30-6LC00 con dos salidas de tipo PNP (cortesía de Siemens).

- **Consumo del sensor**

Los fabricantes suelen especificar el consumo del sensor (*Power or current consumption*) mediante la corriente de alimentación [asociada a un rango reducido de tensión de alimentación (por ejemplo entre 20 y 30 voltios)] o mediante la potencia en vatios o voltamperios, según el sensor se alimente en continua o en alterna respectivamente.

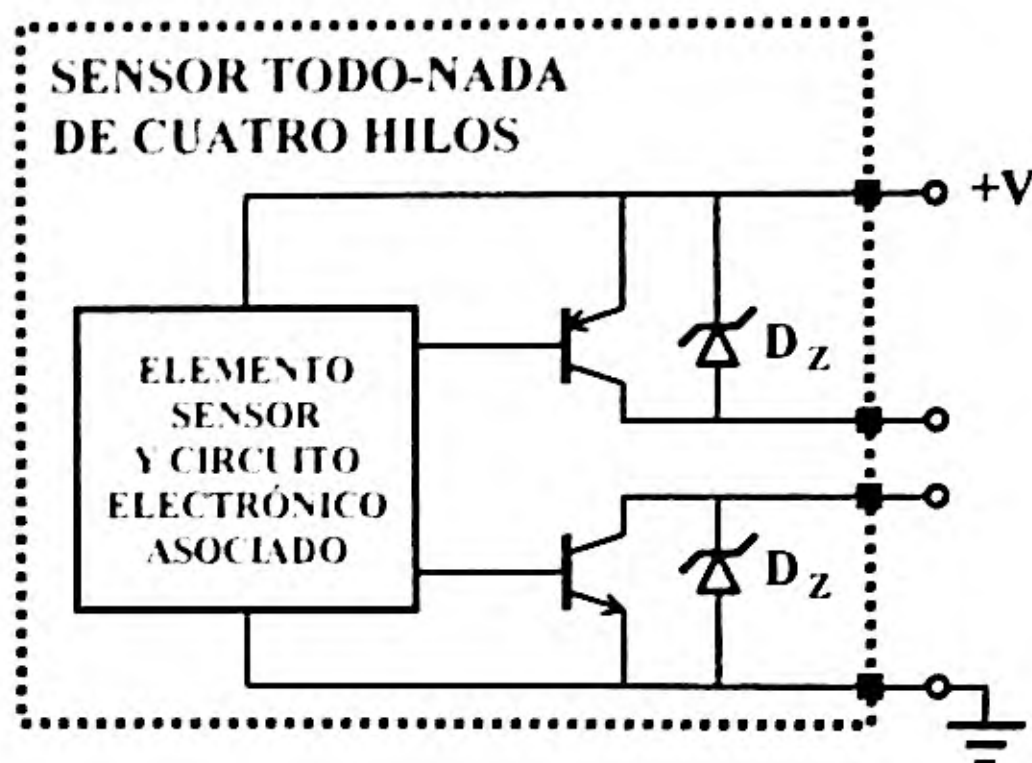


Figura 7.40 Sensor todo-nada de cuatro hilos que posee una salida realizada con un transistor PNP y otra con un transistor NPN.

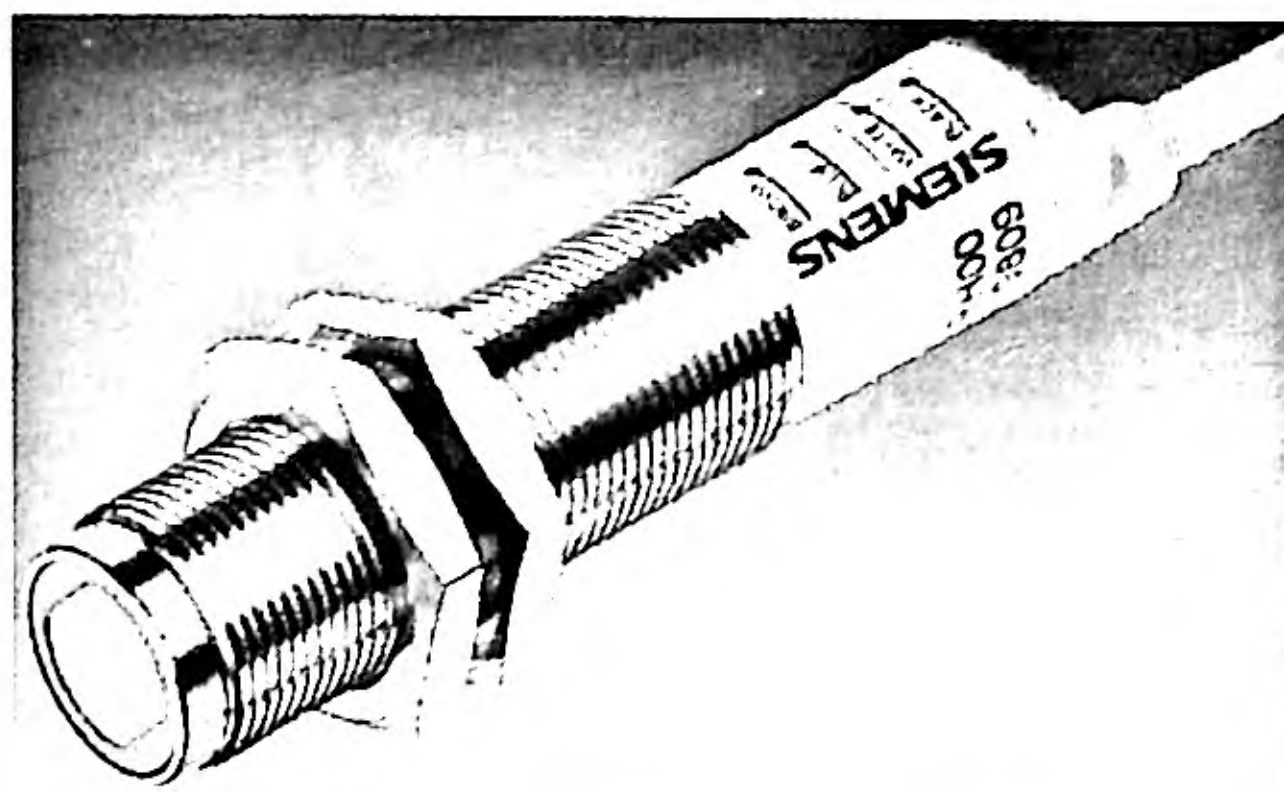


Figura 7.41 Sensor optoelectrónico de proximidad BERO 3RG76 00-3RH00 con dos salidas, una del tipo PNP y otra del tipo NPN (cortesía de Siemens).

7.4.1.4 Sensores de salida temporal

Al igual que en los sensores de salida de formato digital, el parámetro más importante es la cargabilidad de salida y la compatibilidad entre los niveles de tensión y corriente de la salida del sensor y de la entrada del sistema electrónico conectado a él.

7.4.2 Características de alimentación

La gran mayoría de los sensores industriales, a excepción de los autogeneradores o activos, necesitan para su funcionamiento una fuente de alimentación que les proporcione la tensión o la corriente eléctrica adecuadas. La característica que indica el margen de tensiones dentro del cual el sensor funciona adecuadamente, de acuerdo con las especificaciones del fabricante, se denomina *tensión de funcionamiento* (*Operating voltage*) o *tensión de alimentación* (*Supply voltage*). La tensión de alimentación puede ser continua o alterna, aunque generalmente es continua. Cuando la alimentación se realiza mediante una tensión continua es necesario conocer también el máximo rizado u ondulación admisible en ella.

Se define la *ondulación residual* (*Ripple*) como la máxima tensión alterna, pico a pico, superpuesta a la tensión continua de alimentación, admisible para que el sensor industrial funcione correctamente. Se indica en forma porcentual y en muchos casos es admisible una ondulación residual del 10%.

Los fabricantes suelen especificar también el *consumo de corriente en vacío*, que, como su nombre indica, representa el máximo valor de corriente que el sensor puede demandar de la fuente de alimentación cuando no se conecta una carga a la salida.

Algunos sensores industriales que se alimentan en continua, como por ejemplo los todonada, incorporan una protección contra inversión de polaridad (Figuras 7.32 y 7.33). Dicha protección tiene como objetivo evitar que una conexión incorrecta de la tensión de alimentación produzca la destrucción del sensor.

La impedancia presentada al sensor por la fuente de alimentación se denomina *impedancia de la fuente* (Z_s), y en ella se considera incluida generalmente la propia de los cables de conexión. La impedancia presentada a la fuente por el sensor se denomina *impedancia de entrada* (Z_e).

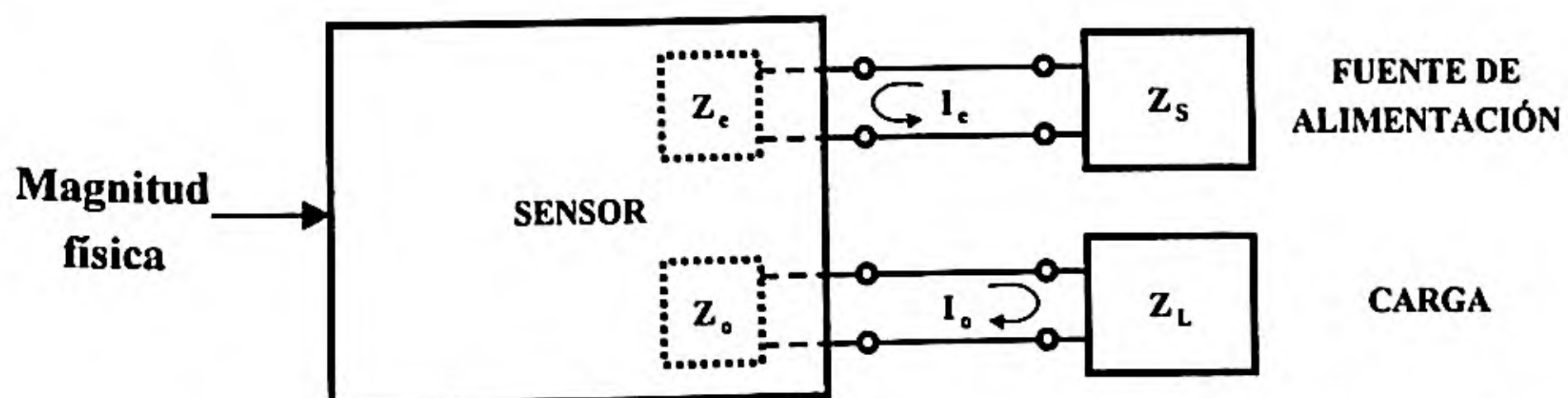


Figura 7.42 Esquema genérico de conexión entre un sensor analógico pasivo, la fuente de alimentación y la carga.

En la figura 7.42 se representa la conexión de un sensor analógico pasivo a su fuente de alimentación y a la carga. Las líneas de retorno de la alimentación y de la salida pueden estar aisladas o conectadas entre sí, lo cual se indica mediante una conexión con línea discontinua. También es corriente que las líneas de retorno estén aisladas eléctricamente del encapsulado del sensor. Ambas líneas pueden estar conectadas a una masa cercana a la fuente y a la carga, o pueden quedar sin conectar (masa flotante).

Los cables de salida (y, a veces, también los de alimentación) pueden estar blindados para evitar que en ellos se induzcan interferencias de tipo electromagnético o electrostático, lo cual es probable si tienen una longitud considerable o cuando los niveles de tensión de salida son relativamente bajos.

7.4.3 Características de aislamiento

Cuando dos o más partes de un sensor industrial están aisladas eléctricamente es importante, a veces, conocer el grado de aislamiento entre ellas. Los parámetros que proporcionan dicha información suelen ser la resistencia de aislamiento y la tensión de ruptura o rigidez dieléctrica que, aunque expresan conceptos distintos, están íntimamente relacionadas.

La *resistencia de aislamiento* es la resistencia entre las partes, medida mediante la aplicación de una tensión continua de un determinado valor. La *tensión de ruptura* o *rigidez dieléctrica* se suele definir como la máxima tensión que se puede aplicar entre las partes aisladas sin que se produzca arco eléctrico o sin que la corriente que circule entre ambas supere un valor determinado.

7.5 Características mecánicas

7.5.1 Conceptos generales

Como su nombre indica, las características mecánicas hacen referencia a los aspectos de tipo mecánico relacionados con el sensor industrial y sus condiciones de manejo e instalación.

Generalmente los fabricantes especifican siempre:

- La configuración constructiva y sus dimensiones externas.
- Las instrucciones de montaje.
- El tipo, tamaño y localización de las conexiones eléctricas y mecánicas.
- La forma de realizar los ajustes externos (en caso de ser necesarios).
- El material de la carcasa.
- El grado de protección de la carcasa ante agentes externos.

Salvo raras excepciones, los sensores industriales incluyen también un conjunto de informaciones inscritas en la carcasa o en una placa unida a ella. La información incluida hace referencia generalmente a la nomenclatura del propio sensor (marca, modelo, número de serie, etc.), nombre y dirección del fabricante, y otras características, como por ejemplo el rango, la tensión de alimentación, la identificación de las conexiones eléctricas, el grado de protección ambiental, etc. Este último es un parámetro característico común a todos los sensores industriales y por su importancia se analiza en el apartado siguiente.

La especificación de las características mecánicas de un sensor industrial es de gran importancia, porque facilita su manejo e instalación, previene utilizaciones inadecuadas y permite la interacción correcta entre la magnitud a medir, el sensor y el sistema de medida.

7.5.2 Grado de protección ambiental de los sensores industriales

El grado de protección ambiental (*Ingress Protection*), también denominado grado de sellado, de los sensores industriales es una característica mecánica de especial importancia en todas aquellas aplicaciones, que constituyen los casos más frecuentes, en las que el sensor forma parte de un sistema hermético o está sometido a determinadas condiciones ambientales de humedad, polvo, etc.

La existencia de diferentes situaciones prácticas y el distinto coste de la construcción mecánica de los sensores para adaptarse a ellas ha hecho que diversas organizaciones de normalización hayan desarrollado y establecido normas aplicables a los sensores cuando deben operar en condiciones como las descritas en el párrafo anterior. Entre ellas destaca la Comisión Electrotécnica Internacional (IEC), perteneciente a ISO (Organización Internacional de Normalización), que tiene la responsabilidad principal para establecer normas electrotécnicas, debido a lo cual muchas de las normas desarrolladas por ella tienen su equivalente en otras organizaciones, como por ejemplo DIN y CENELEC.

Una de las normas desarrolladas por la IEC, aplicable a los sensores industriales, es la IEC 144, que especifica el grado de protección a la entrada de agentes externos sólidos o líquidos. Es conveniente resaltar que, aunque esta norma tiene cierta relación con las técnicas de protección en zonas con peligro de explosión, no es directamente aplicable a las mismas.

El grado de protección ambiental o grado de sellado se indica mediante las siglas IP (acrónimo de *Ingress Protection*) seguidas de dos cifras decimales. La primera cifra indica el grado de protección frente al contacto y entrada de cuerpos sólidos externos y la segunda cifra el grado de protección frente a la entrada de líquidos. Cuanto más alto es el número de la primera y segunda cifras, mayor es el grado de protección de la carcasa. Por ejemplo, un grado de protección IP55 engloba a todos los grados inferiores, tales como IP22, IP23, IP34 e IP54, por citar algunos de los más utilizados.

En la tabla 7.6 se indica el tipo de protección que corresponde a los distintos valores de cada una de las cifras.

7.6 Características de funcionamiento

7.6.1 Introducción

Las características de funcionamiento de un sensor industrial hacen referencia a su respuesta en unas condiciones determinadas. De acuerdo con la naturaleza de dichas condiciones, las características de funcionamiento se pueden clasificar en estáticas (*Static characteristics*), dinámicas (*Dynamic characteristics*), ambientales (*Environmental characteristics*) y de fiabilidad (*Reliability characteristics*).

En sucesivos apartados se analiza cada una de ellas, pero antes es conveniente definir y aclarar un concepto común a todas como es el de *error*. El error es la diferencia entre el valor obtenido y el valor real de la magnitud objeto de medida. Toda medida lleva asociada inevitablemente un error y el conocimiento del mismo es de la mayor importancia para la interpretación y evaluación de los resultados de la misma.

1ª	Grado de protección	2ª	Grado de protección
0	El equipo no está protegido contra la entrada de cuerpos sólidos externos	0	Sin protección.
1	Protección contra la entrada de cuerpos sólidos externos grandes (mayores de 50 mm de diámetro)	1	Protección contra la condensación de gotas de agua.
2	Protección contra la entrada de cuerpos sólidos externos de tamaño medio (mayores de 12 mm de diámetro).	2	Protección contra gotas de líquido; la caída de gotas de líquido no tiene efectos perjudiciales si la carcasa tiene una inclinación de hasta 15° desde la vertical.
3	Protección contra la entrada de cuerpos sólidos externos mayores de 2,5 mm de diámetro.	3	Protección contra lluvia o agua en forma de lluvia, para un ángulo menor o igual a 60° con respecto a la vertical.
4	Protección contra la entrada de cuerpos sólidos externos pequeños (mayores de 1 mm de diámetro).	4	Protección contra salpicaduras de líquido en cualquier dirección.
5	Protección contra depósitos perjudiciales de polvo. La entrada de polvo no se evita totalmente, pero éste no puede entrar en cantidades suficientes como para interferir en el adecuado funcionamiento del equipo.	5	Protección contra chorros de agua. El agua no produce efectos perjudiciales cuando la proyecta un inyector en cualquier dirección bajo condiciones especificadas.
6	Protección contra la entrada de polvo. Protección total frente al contacto con partes móviles situadas dentro de la carcasa.	6	Protección contra condiciones del tipo de las de cubierta de barco (equipos herméticos de cubierta). El agua procedente de un fuerte oleaje no entra en la carcasa bajo condiciones especificadas.
7		7	Protección contra la inmersión en agua bajo condiciones especificadas de presión y tiempo.
8		8	Protección contra la inmersión indefinida en agua bajo condiciones especificadas de presión.

Tabla 7.6 Significado de las cifras del grado de protección de acuerdo con la norma IEC 144.

Por otra parte, aunque la definición de error parece no dejar lugar a dudas, el término error es muy amplio y hasta, en cierto modo, un poco ambiguo. Ello es debido a que el valor final del error o *error total* no es consecuencia únicamente de una sola fuente de error, sino que, aparte de las inexactitudes propias de la respuesta del sensor (linealidad, histéresis, etc.), existen otras que son captadas en el entorno físico en el que tiene lugar el proceso (por ejemplo, ruido eléctrico, interferencias, etc.) o que se deben a una utilización inadecuada del sensor (por ejemplo, montaje incorrecto, alimentación inestable, etc.). En definitiva, todas las componentes de la señal de salida no relacionadas con la medida de la magnitud de entrada se consideran errores. Cada tipo de error se analiza al estudiar en sucesivos apartados las diferentes características de funcionamiento.

7.6.2 Características estáticas

Las características estáticas describen el comportamiento del sensor en unas condiciones ambientales determinadas, con cambios muy lentos de la magnitud a medir y en ausencia de con-

diciones externas duras, tales como golpes, vibraciones, aceleraciones (a no ser que ésta sea la magnitud a medir), etc. A continuación se indican las características estáticas más importantes.

7.6.2.1 Exactitud

La exactitud (*Accuracy*) es la cualidad de un sensor, que establece su capacidad para proporcionar a su salida una señal que coincida con la que corresponde exactamente al valor verdadero (*true, exact or ideal*) de la variable que se mide, cuando se realiza una secuencia de medidas del mismo valor de la variable de entrada en las mismas condiciones ambientales (temperatura, humedad, etc.). Se determina mediante calibración estática que se describe en el apartado 7.6.2.3. Cualquier diferencia entre el valor de la salida correspondiente al valor verdadero o exacto de la variable que se mide y el proporcionado por el sensor constituye un error. La diferencia entre el resultado de la medida y el valor verdadero se denomina error absoluto y, en ocasiones, se expresa mediante un porcentaje del máximo valor que se puede medir con el sensor, o con respecto a la diferencia entre los valores máximo y mínimo que establecen el rango de medida (*Span*):

$$\text{Error Absoluto} = \text{Resultado de la medida} - \text{Valor verdadero}$$

Es usual en la práctica, indicar el error relativo que es el cociente entre el error absoluto y el valor verdadero:

$$\text{Error relativo} = \frac{\text{Error absoluto}}{\text{Valor verdadero}}$$

La exactitud es una de las características estáticas básicas a tener en cuenta al seleccionar un sensor industrial porque define su *categoría* o *clase*. Todos los sensores que pertenecen a la misma clase tienen el mismo valor del error relativo cuando se aplica el mismo valor de la variable de entrada dentro del rango de medida y depende en general del principio de transducción utilizado y del tipo de magnitud a medir. Por ejemplo, un sensor de medida de distancias de clase 0,4 cuyo fondo de escala es 10 cm proporciona la medida, en unas determinadas condiciones, con un error inferior 0,4 mm.

7.6.2.2 Precisión, repetibilidad y reproducibilidad

La precisión (*Precision*) es la característica de un sensor que indica su capacidad para proporcionar el mismo valor de la salida cuando se realiza una secuencia de medidas del mismo valor de la variable de entrada, en las mismas condiciones ambientales (temperatura, humedad, etc.) independientemente de la exactitud de la medida. Por lo tanto una precisión elevada implica una diferencia mínima entre varias medidas y es una condición necesaria pero no suficiente para tener una elevada exactitud. La precisión se puede expresar de distintas maneras, entre las que cabe citar, el % del valor a fondo escala, la máxima desviación entre salidas o el valor cuadrático medio de dicha desviación.

La precisión está ligada con la repetibilidad (*Repeatability*) y la reproducibilidad (*Reproducibility*). La repetibilidad es el valor máximo que, con una determinada probabilidad, tiene la diferencia entre varias medidas llevadas a cabo en las condiciones indicadas anteriormente, en un intervalo corto de tiempo. La figura 7.43 representa una curva de calibración en la que se determina el error de repetibilidad.

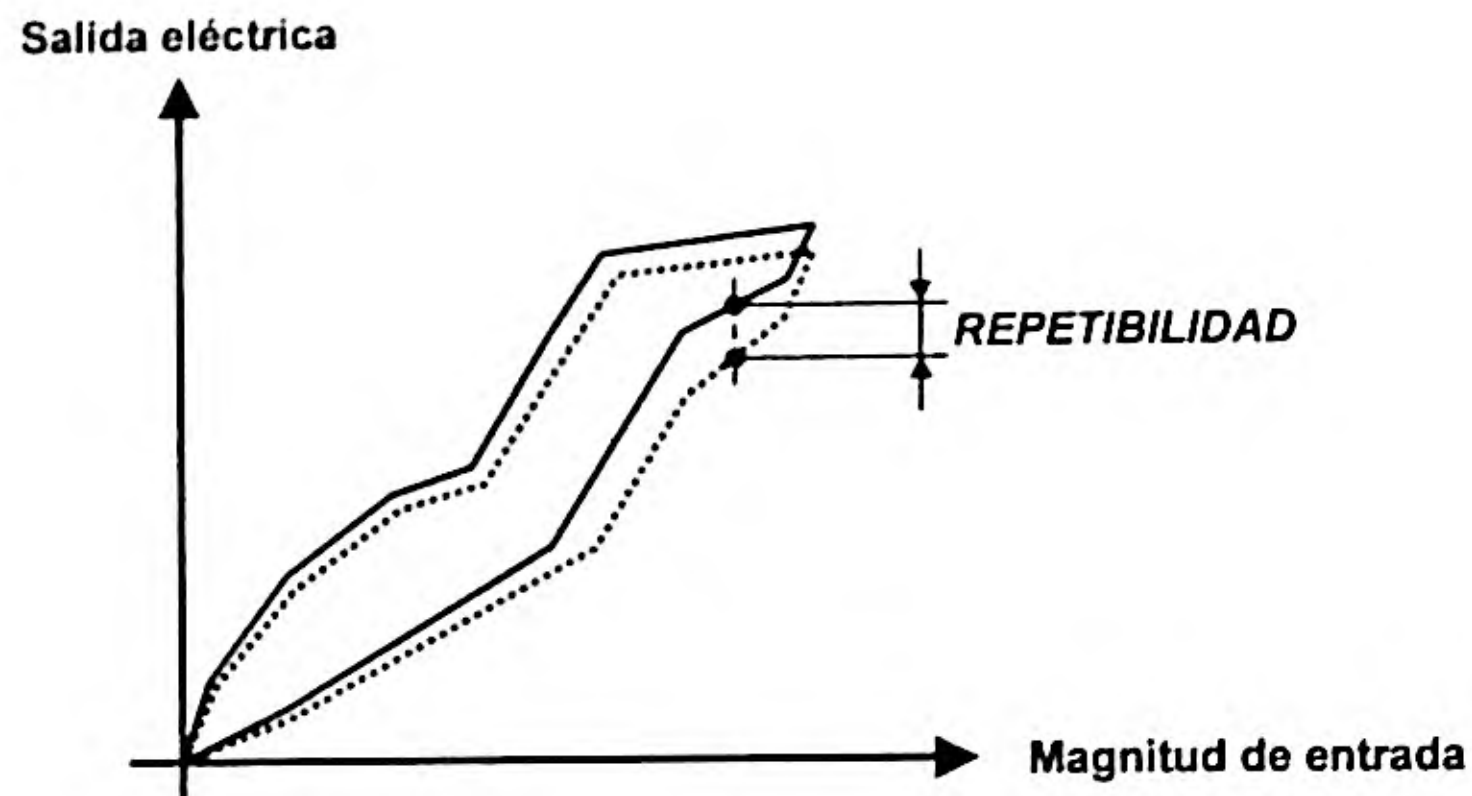


Figura 7.43 Repetibilidad de la respuesta de un sensor.

Dicho valor máximo se denomina reproducibilidad cuando las medidas se realizan a lo largo de un intervalo largo de tiempo.

7.6.2.3 Calibración

La calibración es el procedimiento mediante el cual se ajusta la salida de un sensor sobre su rango completo de medida, de tal forma que coincida al máximo posible con una serie de valores conocidos derivados de los patrones de medida correspondientes a la magnitud a medir. La representación gráfica de los valores obtenidos se denomina *curva de calibración* (*Calibration curve*). En la figura 7.44 se muestra una curva de calibración típica de un sensor lineal, en la que los ajustes para obtener un error nulo se realizan en los puntos de cero y fondo de escala. Generalmente la mayor parte de los errores de calibración dependen del método empleado para realizar el contraste entre el sensor y los patrones. A partir de la curva de calibración se pueden determinar algunos de los errores individuales que influyen en el error total.

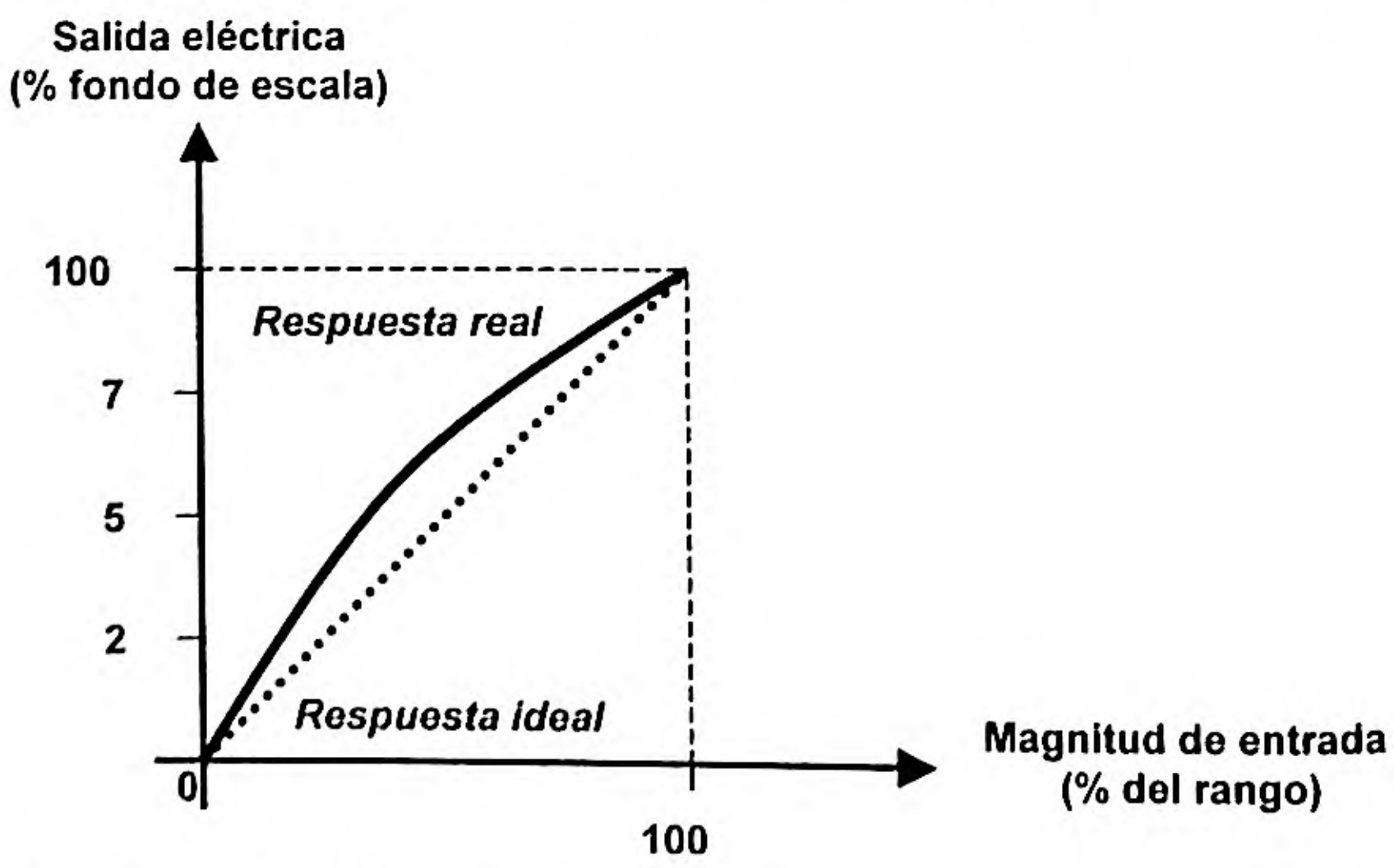


Figura 7.44 Curva de calibración de un sensor.

7.6.2.4 Histéresis

Se define la histéresis (*Hysteresis*) como la máxima diferencia entre las señales de salida, correspondientes a un mismo valor de la magnitud a medir (perteneciente al rango), que se obtienen cuando dicho valor se obtiene mediante el aumento y la disminución de dicha magnitud. Se expresa en % del valor a fondo de escala. Son causas normales de este fenómeno las holguras entre partes mecánicas, rozamientos, etc., y constituye una fuente de error cuyos efectos están limitados únicamente por la calidad del sensor y por la perfección de su diseño.

En la figura 7.45 se muestra la curva de calibración de un sensor en la que se determina la histéresis. La histéresis es una característica imprescindible de los sensores todo-nada para evitar las oscilaciones que se pueden producir a la salida si el nivel de la variable de entrada está justamente en el valor en el que la variable de salida debe pasar de un estado al otro.

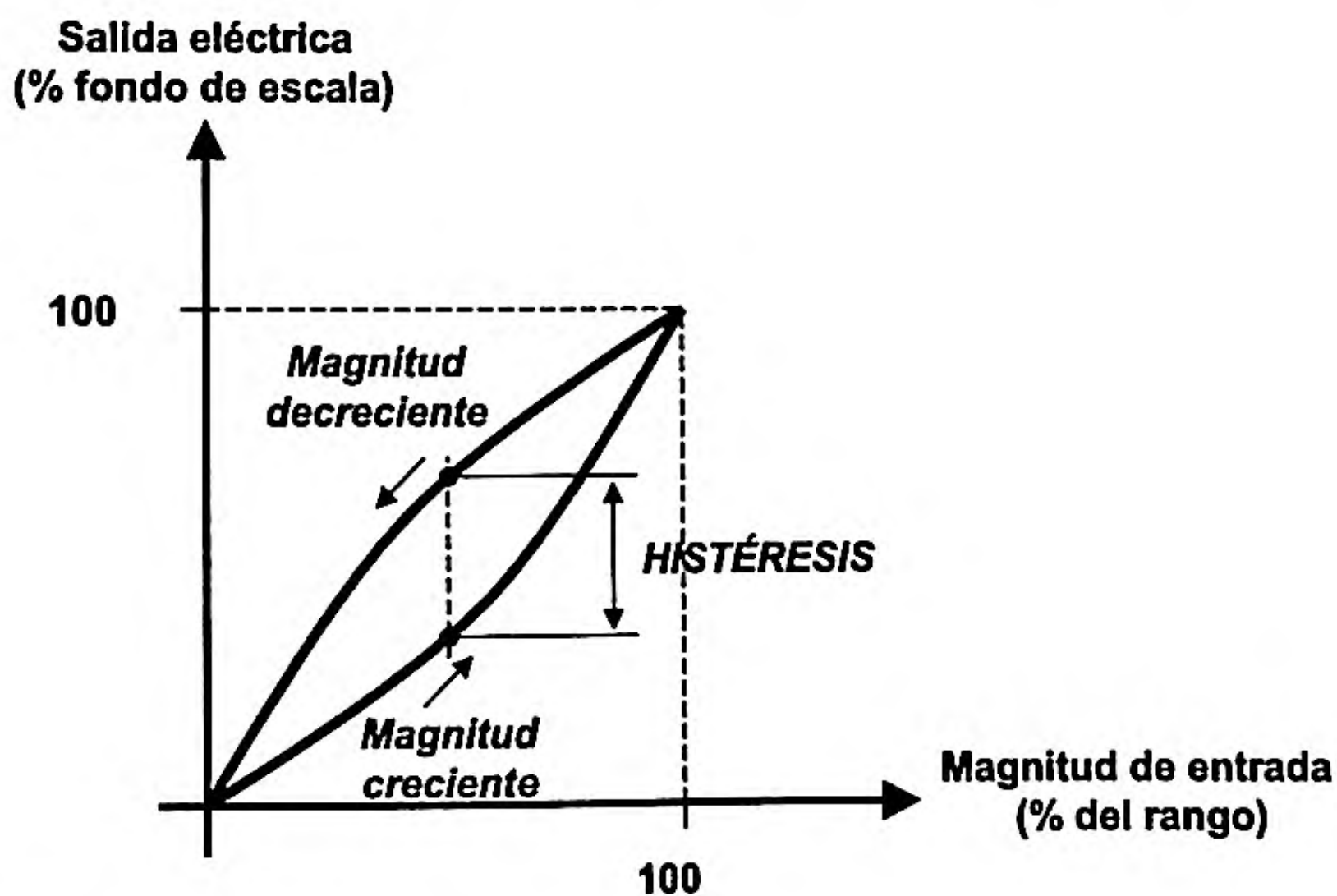


Figura 7.45 Curva de calibración de un sensor que posee histéresis.

7.6.2.5 Linealidad

La no-linealidad (concepto opuesto a la linealidad) es la máxima desviación de la curva de calibración (en su tramo ascendente o descendente) respecto de la línea recta que une los puntos extremos correspondientes al 0% y al 100% del valor a fondo de escala. Se expresa en % del valor a fondo de escala, siempre que el *cero* de la curva de calibración se haga coincidir con el origen de coordenadas (Figura 7.46).

Aunque el concepto de no-linealidad, que se acaba de definir, es el más utilizado en la práctica, es posible también expresar la no-linealidad con relación a una línea de referencia diferente de la que une los puntos extremos. En el caso de que la linealidad ideal no corresponda a una línea recta, sino a una curva (sensor no lineal), este error se denomina *de conformidad*.

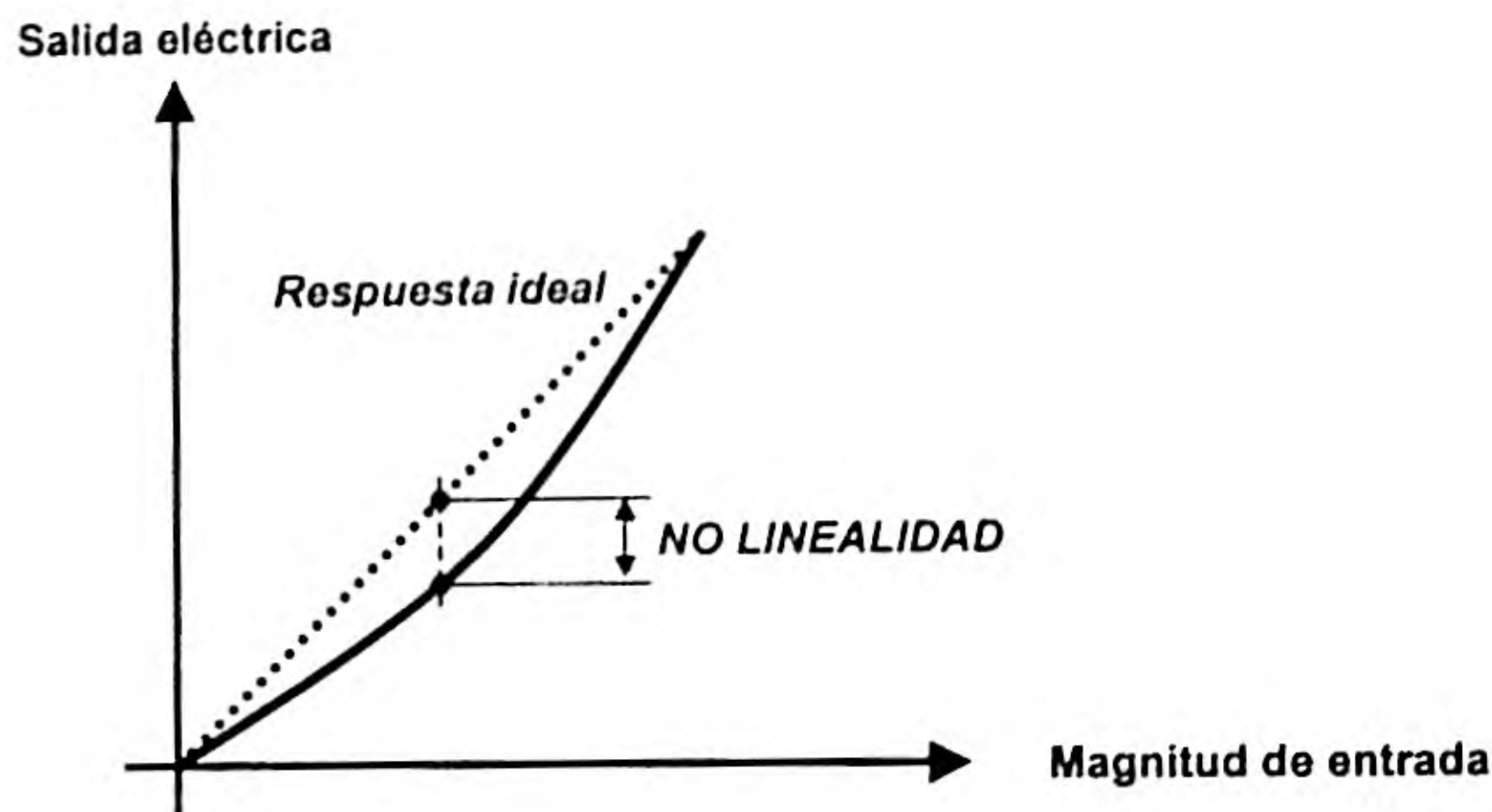


Figura 7.46 Determinación de la no linealidad de la respuesta de un sensor.

7.6.2.6 Mínimo valor medible o umbral

El *mínimo valor medible o umbral* (*Threshold*) es el cambio más pequeño de la magnitud de entrada que produce un cambio medible en la salida. Se define en términos de la magnitud a medir y puede tener diferentes valores en las distintas zonas del rango. Depende del principio de funcionamiento del elemento sensor.

7.6.2.7 Resolución

La salida de algunos sensores no varía de forma continua cuando lo hace la magnitud de entrada, sino que cambia en forma de pequeños incrementos o escalones (por ejemplo, en un potenciómetro bobinado o un codificador de posición de tipo absoluto). Se define la *resolución* (*Resolution*) de un sensor como el cambio o escalón más pequeño de la salida cuando la magnitud a medir varía continuamente dentro del rango y se expresa en % del valor a fondo de escala. La resolución de un sensor de salida digital se expresa mediante el número de bits de la misma. La resolución de un sensor de salida temporal, que proporciona un número de impulsos proporcional al valor de la variación de la magnitud física de entrada, se mide mediante el número de impulsos por unidad de la citada variación. Por ejemplo, en un sensor de desplazamiento de tipo incremental la resolución se expresa mediante el número de impulsos que es capaz de proporcionar por unidad de longitud, ángulo o revolución.

Cuando el valor del cambio de la salida no es medible se acostumbra a decir que el sensor tiene *resolución infinita*, aunque lo más apropiado es *resolución infinitesimal*. En este caso los límites los impone el aparato de medida o el equipo electrónico encargado de procesar la información procedente del sensor. Por ejemplo, si se utiliza un visualizador digital de 3 dígitos (0 a 999), es decir, 1000 puntos, la resolución es 1/1000.

El *mínimo valor medible* y la *resolución* son características diferentes que describen el comportamiento del sensor ante variaciones pequeñas de la señal que se mide.

7.6.2.8 Sensibilidad

La sensibilidad (*Sensitivity*) representa la variación de la señal de salida del sensor cuando se produce un cambio del valor de la magnitud a medir. Establece la pendiente de la curva de calibración en cada punto de la misma. Puede verse afectada por diferentes factores, por ejemplo, la tensión y la frecuencia de la alimentación, la temperatura, la frecuencia de las variaciones de la magnitud a medir, etc. La sensibilidad se expresa de diferente forma para los sensores activos que para los pasivos.

En los sensores *activos*, que no necesitan alimentación, se expresa mediante la relación *magnitud eléctrica (salida)/magnitud física (entrada)*, por ejemplo, mV/g, $\mu\text{V}/^{\circ}\text{C}$, etc. En los sensores *pasivos* se expresa mediante la relación entre la *tensión de salida a fondo de escala y la tensión de alimentación*. Por ejemplo, un sensor extensométrico o célula de carga (*Strain gage*), con un rango de 0 a 50 Kg y una sensibilidad de 2 mV/V, proporciona una salida de 20 mV para 50 Kg si se alimenta a 10 V.

7.6.3 Características dinámicas

7.6.3.1 Introducción

Especifican la respuesta del sensor al variar la magnitud a medir. Los sensores, como dispositivos reales que son, presentan una cierta incapacidad para seguir las variaciones de la magnitud de entrada que superan determinada pendiente de cambio, lo que da lugar a la aparición de errores cuando dichas variaciones son rápidas. Según esto, es necesario conocer las características dinámicas del sensor en aquellos casos en los que trabaja con una magnitud de entrada que varía de forma rápida o en los que se pueden producir incrementos bruscos de la misma.

Con objeto de definir cuantitativamente las características de respuesta dinámica de un sensor industrial, se analiza la forma de onda de la señal de salida cuando la magnitud de entrada varía de acuerdo con unas determinadas formas de onda (variaciones senoidales o cambios en forma de escalón).

Los parámetros que definen la respuesta dinámica de un sensor industrial se analizan a continuación.

7.6.3.2 Respuesta en frecuencia

La respuesta en frecuencia (*Frequency response*), representa la variación de la relación entre las amplitudes de las señales de salida y de entrada, en función de la frecuencia, dentro de un rango definido de frecuencias y variaciones senoidales de la magnitud a medir. La figura 7.47 muestra dos posibles curvas de respuesta en frecuencia.

7.6.3.3 Tiempo de respuesta

El tiempo de respuesta (*Response time*) se define como el tiempo transcurrido desde que se aplica un cambio en escalón de la magnitud a medir hasta que la salida alcanza un porcentaje determinado de su valor final, que suele ser el 10% (Figura 7.48). Se denomina " t_r ", y en los sensores todo-nada se suele especificar como el valor medio del tiempo que tarda la salida en pasar de todo a nada y viceversa.

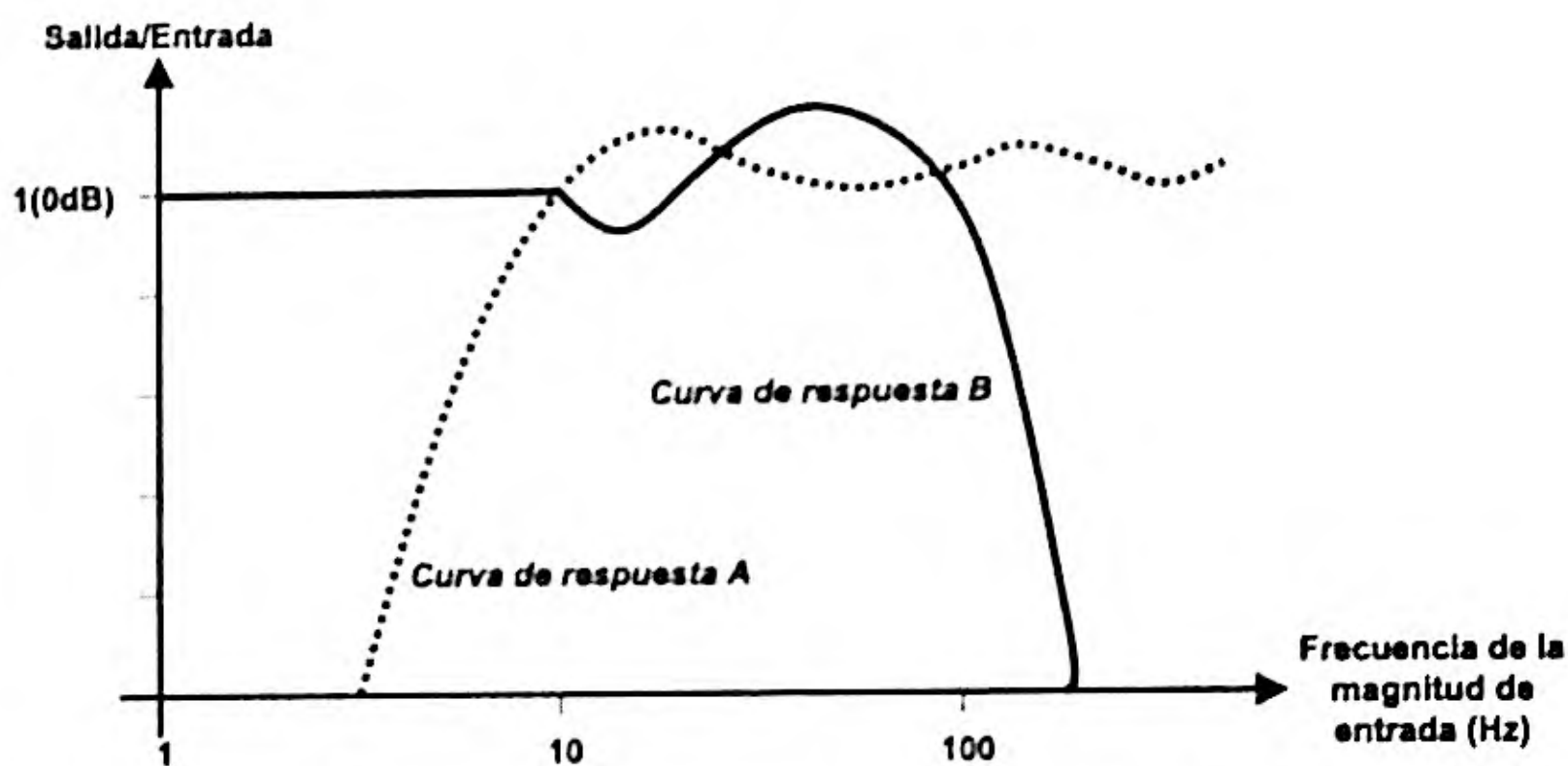


Figura 7.47 Curvas de respuesta en frecuencia de dos sensores industriales diferentes.

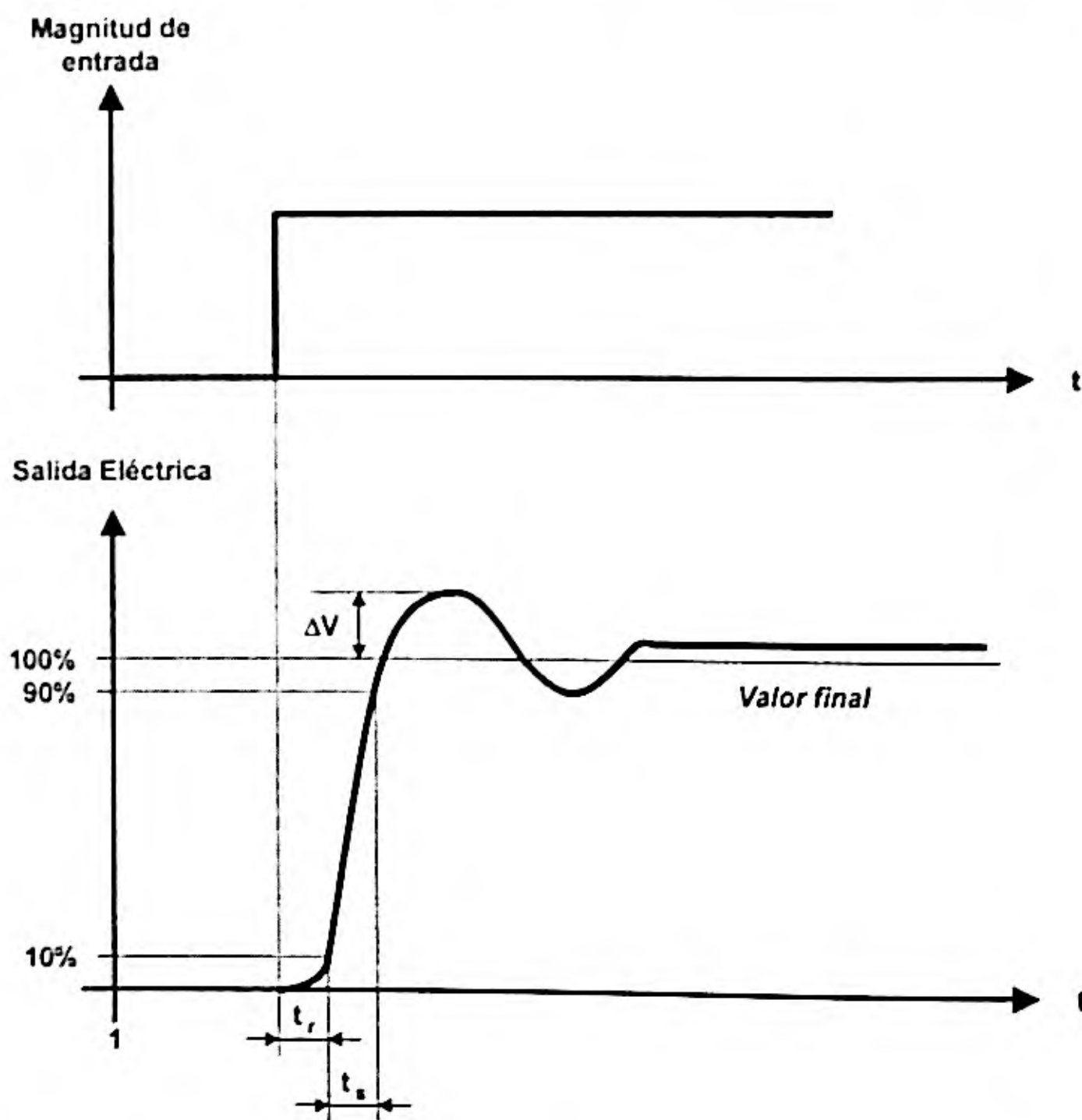


Figura 7.48 Respuesta transitoria de un sensor.

7.6.3.4 Tiempo de subida

El tiempo de subida (*Rise time*) determina el intervalo comprendido entre el instante en el que la señal de salida alcanza el 10% de su valor final y aquél en el que alcanza el 90%, como resultado de un cambio en escalón de la magnitud de entrada y se denomina “*t_r*” (Figura 7.48). Si el escalón es de variación negativa (descendente) este tiempo es de bajada.

7.6.3.5 Constante de tiempo

La constante de tiempo (*Time constant*), simbolizada por τ , representa un caso particular del tiempo de respuesta y se define como el tiempo necesario para que la salida alcance el 63% (concretamente, el 63,2%) de su valor final cuando se le aplica una señal en escalón.

7.6.3.6 Amortiguamiento o sobreoscilación (ΔV)

Es la diferencia entre el valor de pico de la señal de salida y su amplitud final, y suele expresarse en % de esta amplitud. Si, en respuesta a un cambio en escalón de la magnitud de entrada, la señal de salida de un sensor oscila alrededor del valor final antes de permanecer en el mismo, se dice que se trata de un *sensor subamortiguado* (*Underdamped*). Por el contrario, si la salida alcanza el valor final sin superarlo en ningún momento, se dice que es *sobreamortiguado* (*Overdamped*). El amortiguamiento es un factor determinante del límite superior de la respuesta en frecuencia.

El conjunto de características que se acaba de analizar establece la respuesta de un sensor a un cambio en escalón de la magnitud de entrada y constituye lo que se denomina *respuesta transitoria*. Se representan en la figura 7.48.

7.6.4 Características ambientales

Las características estáticas y dinámicas estudiadas anteriormente se especifican en unas condiciones ambientales (*Environmental conditions*) determinadas y en ausencia de factores externos que afecten al funcionamiento del sensor.

Puede suceder, sin embargo, que en determinadas ocasiones el sensor industrial opere bajo condiciones externas (temperatura, humedad, golpes, vibraciones, etc.) distintas de aquellas para las que fue calibrado. Por ello, es necesario conocer las desviaciones de las características estáticas que se producen por efecto de las variaciones de las condiciones ambientales. A continuación se analizan las más importantes.

7.6.4.1 Efectos térmicos

Los *efectos térmicos* constituyen una característica ambiental que se debe especificar en todos los tipos de sensores industriales. Existen diferentes formas de hacerlo. Se pueden especificar mediante el rango de *temperatura* (*Ambient operating temperature*), que representa el intervalo de temperatura ambiental en el que el sensor opera de acuerdo con las especificaciones establecidas por el fabricante. Otro valor importante es el de la *temperatura máxima* a la que el sensor puede estar expuesto sin sufrir daño.

Una forma usual de especificar los efectos térmicos viene dada por el *error de temperatura* (*Thermal error*), que define la variación máxima de la salida cuando la temperatura varía desde la temperatura ambiente hasta una temperatura extrema especificada.

En algunos sensores los efectos de la temperatura se especifican solamente como *deriva térmica del cero* y *deriva térmica de la sensibilidad*, que ocasionan un desplazamiento paralelo y un cambio en la pendiente de la curva de calibración, respectivamente.

El concepto de *deriva térmica del cero* se entiende fácilmente. Teóricamente, un sensor sin carga y en condiciones normales de temperatura, al que se aplica una magnitud de entrada

nula, proporciona salida *cero* o el valor que se le haya hecho corresponder (por ejemplo, 4mA en los sensores analógicos con salida de corriente normalizada de 4 a 20mA). Si se mantiene sin carga, la salida no debe variar en todo el rango de temperaturas de funcionamiento. Por ello, se define la *deriva térmica del cero* (*Zero thermal drift*) como la máxima variación de la salida obtenida al variar la temperatura dentro del rango de temperaturas de funcionamiento, con magnitud de entrada nula y sin carga (Figura 7.49). La *deriva térmica de la sensibilidad* (*Sensitivity thermal drift*) representa la máxima variación de la pendiente de la curva de calibración observada a la salida cuando varía la temperatura de operación, dentro del rango de temperaturas de funcionamiento.

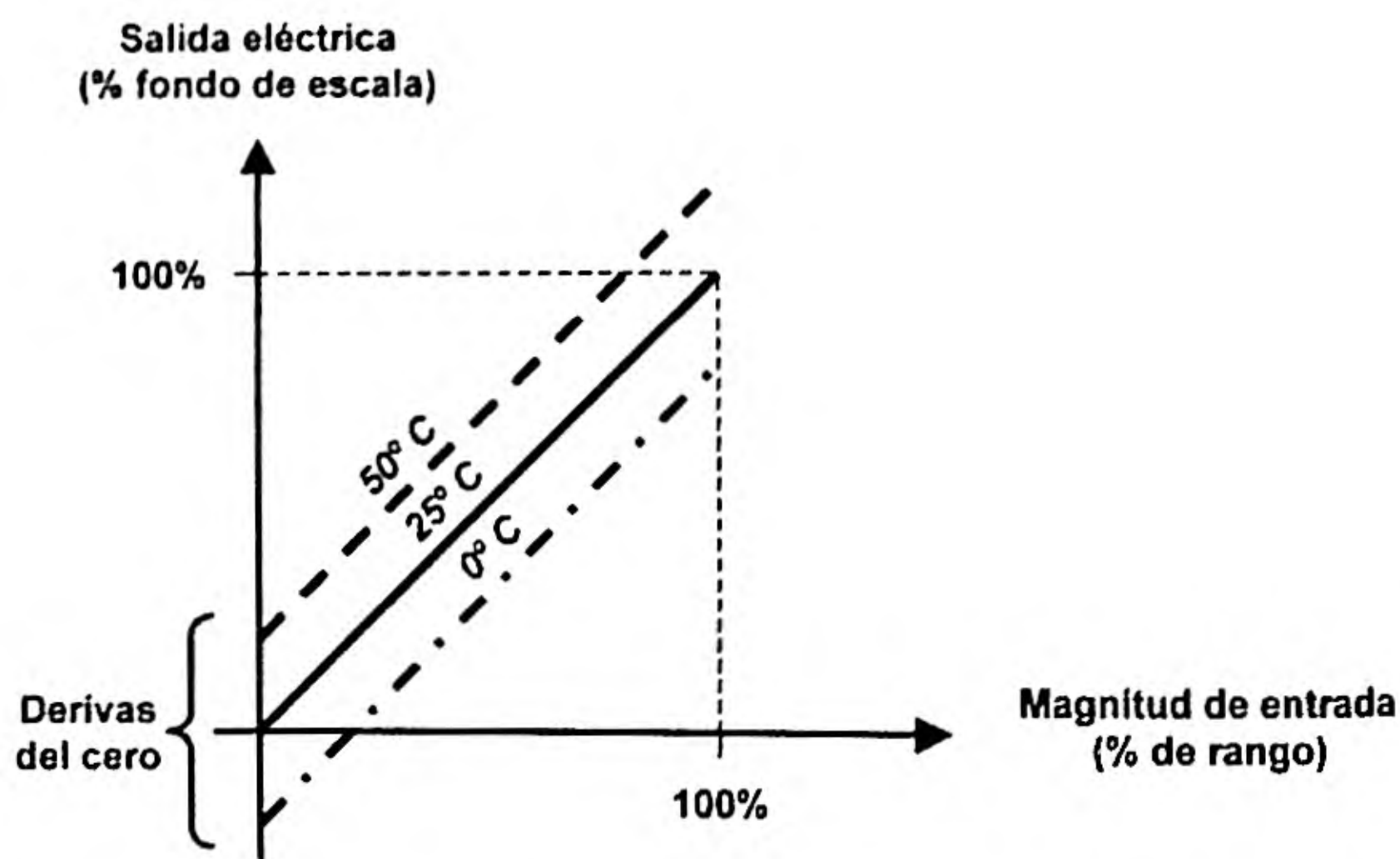


Figura 7.49 Determinación de las derivas térmicas del cero y del fondo de escala de un sensor.

7.6.4.2 Efectos de la aceleración y las vibraciones

Cuando un sensor industrial se utiliza en aplicaciones en las que está sometido a aceleraciones frecuentes, se debe considerar la posibilidad de que existan errores. Los efectos de la aceleración se especifican mediante el *error de aceleración*, que se define como la diferencia máxima (para un valor de la magnitud a medir, dentro del rango) entre los valores de la salida obtenidos con la aplicación de una aceleración constante a lo largo de un eje determinado, y en ausencia de ella.

Las vibraciones afectan a los sensores industriales de forma similar a las aceleraciones pero con efectos mayores, porque dentro de un rango de frecuencias de vibración determinado muestran determinados factores de amplificación (resonancias) debido a los propios componentes del sensor. Se especifican mediante el *error de vibración*, que define la variación máxima de la salida para unos determinados niveles de amplitud y frecuencia de vibración, bajo unos determinados ejes y en unas condiciones ambientales concretas.

7.6.4.3 Efectos de la presión ambiental

Los efectos de la presión ambiental se observan en algunos sensores cuando están calibrados a una presión barométrica normal y se utilizan, por ejemplo, en altitudes elevadas, lugares de presión cercana al vacío, etc., que pueden ocasionar deformaciones o degradaciones importantes. El parámetro que especifica dichos efectos es el *error de presión ambiental*, que se define como la máxima variación de la señal de salida cuando la presión ambiental cambia dentro de unos valores específicos.

7.6.4.4 Efectos de las perturbaciones eléctricas

Una fuente importante de error son los efectos originados por perturbaciones eléctricas no deseadas que son captadas por el sensor, procedentes del entorno físico en el que se utiliza. Se les denomina *errores de interferencia* (*Interference errors*) y se manifiestan en forma de señales parásitas superpuestas sobre la verdadera señal de salida, tal como se observa en la figura 7.50. En esta categoría se incluyen los errores producidos por el ruido eléctrico, los transitorios de conmutación, las interferencias de radio, las perturbaciones en la tensión de alimentación, la captación de zumbido de red (50 Hz), etc. Sus efectos se pueden minimizar mediante técnicas apropiadas de conexión, apantallamiento y puesta a tierra.

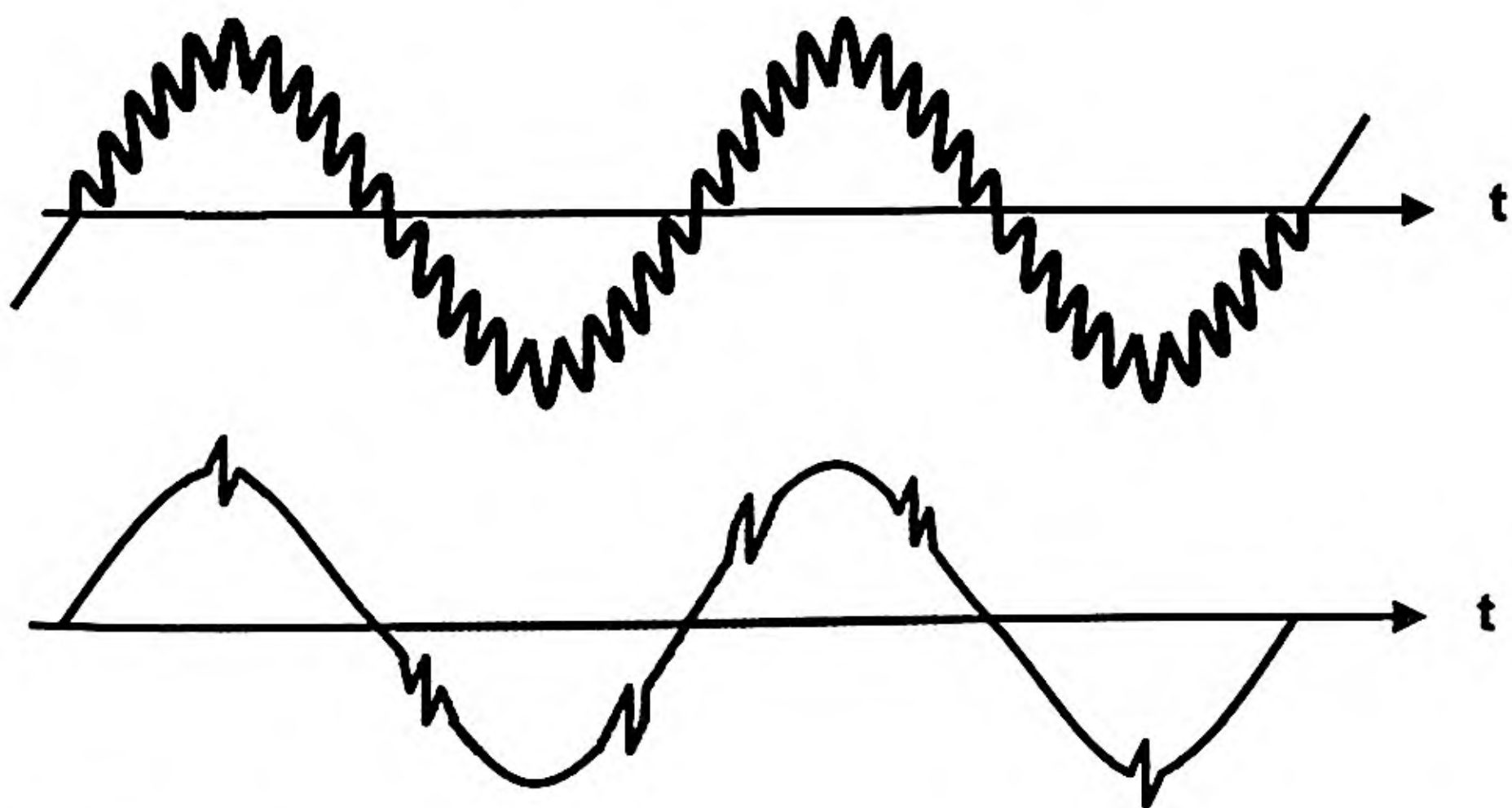


Figura 7.50 Ejemplo de señales eléctricas con errores producidos por interferencias electromagnéticas.

7.6.4.5 Otros efectos

Existen otros factores ambientales que influyen en el comportamiento de un sensor industrial, que se deben tener en cuenta si se manifiestan en el entorno en el que está instalado. Entre otros factores cabe citar la humedad, la corrosión, la atmósfera salina, etc. El fabricante debe especificarlos en las hojas de características correspondientes.

Por último, y aunque no guardan relación directa con el ambiente, pero sí con el funcionamiento bajo condiciones distintas a las de calibración, se deben considerar los efectos sobre el sensor de un montaje inadecuado o *efectos de montaje*. El *error de montaje* (o instalación) (*Mounting error*) es el error resultante de la deformación mecánica del sensor que se produce en la fase de instalación del mismo y en la realización de las conexiones eléctricas o del acoplamiento con la magnitud a medir. La precisión de la medida está condicionada, además de por la precisión propia del sensor, por el grado de exactitud con el que se cumplan las instrucciones de montaje establecidas por el fabricante. En ocasiones, los errores de montaje no se suelen incluir en las especificaciones, lo que hace necesario verificar su ausencia mediante las adecuadas comprobaciones. En general los fabricantes de sensores proporcionan reglas prácticas de instalación o montaje (*Mounting precautions*) que ayudan al técnico a utilizarlos correctamente. En cada uno de los apartados dedicados a los diferentes sensores detectores de objetos, se indican las mismas.

7.6.5 Características de fiabilidad

Aunque algunas características ambientales pueden considerarse, en cierta medida, definitivas de la fiabilidad de un sensor, las características de fiabilidad (*Reliability*) propiamente dichas son las que hacen referencia a la vida útil de un sensor y a los errores que, como consecuencia del envejecimiento del mismo, pueden aparecer con el transcurso del tiempo.

El envejecimiento de un sensor es algo con lo que se debe contar, a más o menos largo plazo, según el tipo de sensor de que se trate (ningún sensor dura eternamente). Produce una pérdida progresiva de las características indicadas en las especificaciones del fabricante, que hace que se deban realizar recalibraciones periódicas.

La *vida operativa o vida útil* de un sensor se puede definir como el mínimo tiempo durante el cual el sensor funciona (continuamente o sobre un número determinado de ciclos) sin sufrir modificaciones de sus características, dentro de un margen de tolerancia especificado.

En algunas ocasiones es también importante conocer la *vida de almacenamiento* de un sensor, que se define como el tiempo durante el cual un sensor puede estar almacenado en unas condiciones determinadas sin sufrir modificaciones de sus características, dentro de un margen de tolerancia especificado y en condiciones de trabajo definidas.

Otras tres características, suficientemente dependientes del tiempo como para ser consideradas como características de fiabilidad, son la *estabilidad temporal de la salida*, el *desplazamiento del cero* y el *desplazamiento de la sensibilidad*. Las dos últimas no se deben confundir con sus homónimas estudiadas al describir los efectos térmicos, aunque su contribución al error final se manifiesta de la misma forma.

La *estabilidad temporal de la salida* representa la variación del valor de la señal de salida durante un tiempo determinado, cuando la magnitud que se mide, al igual que las condiciones ambientales, se mantienen constantes.

El *desplazamiento del cero* se define como la variación del valor de la señal de salida cuando la magnitud de entrada es cero, durante un periodo de tiempo especificado y a temperatura ambiente. Se caracteriza, como se indica anteriormente por provocar un desplazamiento paralelo de la curva de calibración respecto de la curva original.

El *desplazamiento de sensibilidad* se especifica como el máximo cambio de la sensibilidad observado durante un periodo de tiempo determinado a temperatura ambiente. Se manifiesta por un cambio de la pendiente de la curva de calibración.

Además, es necesario tener en cuenta que los efectos no deseados, que inciden sobre el sensor industrial cuando está instalado, dependen, generalmente, del tipo de aplicación. Por ello en las características de fiabilidad se incluyen todas las características consideradas importantes que determinan los efectos de un fallo después de que se produce el mismo (cortocircuitos internos, ruptura de elementos internos o del encapsulado, etc.).

7.7 Sensores industriales de aplicación general en procesos de fabricación

7.7.1 Introducción

En los procesos de fabricación (*Manufacturing processes*) se utilizan sensores para medir diferentes tipos de magnitudes y, entre ellos, son muy utilizados los sensores detectores de objetos y los sensores de medida de distancia o de posición de objetos. Por ello a continuación se estudian estos tipos de sensores.

7.7.2 Sensores detectores de objetos

7.7.2.1 Introducción

Reciben la denominación de sensores detectores de objetos los sensores todo-nada cuya señal de salida depende de la presencia o no de un objeto en sus cercanías [SALO 98].

La detección de la presencia de un objeto depende de diversos factores, entre los que cabe citar la distancia a la que debe estar el objeto para ser detectado, su velocidad y el tipo de material del mismo. La variación de estos factores en amplios márgenes hace que existan diferentes principios físicos de detección que dan lugar a diversos tipos de sensores. Esto hace que los sensores de proximidad constituyan una tecnología compleja para cuyo estudio es conveniente realizar un mapa conceptual que relacione adecuadamente los diferentes conceptos relativos a los mismos.

A partir de su experiencia, los autores han obtenido el mapa conceptual representado en la tabla 7.7 en el que los sensores se clasifican en función de la distancia entre el sensor y el objeto y del tipo de material de que está constituido el objeto.

Según la distancia a la que detectan el objeto, los sensores se clasifican en sensores con contacto y sin contacto. En los sensores con contacto el objeto toca físicamente a aquél y cierra o abre uno o más circuitos eléctricos. Por el contrario, los sensores sin contacto detectan, como su nombre indica, la presencia de un objeto, o miden la distancia a la que se encuentra el objeto, sin necesidad de que exista un contacto físico entre ambos. Los sensores con contacto se pueden clasificar según su robustez en finales de carrera y microinterruptores.

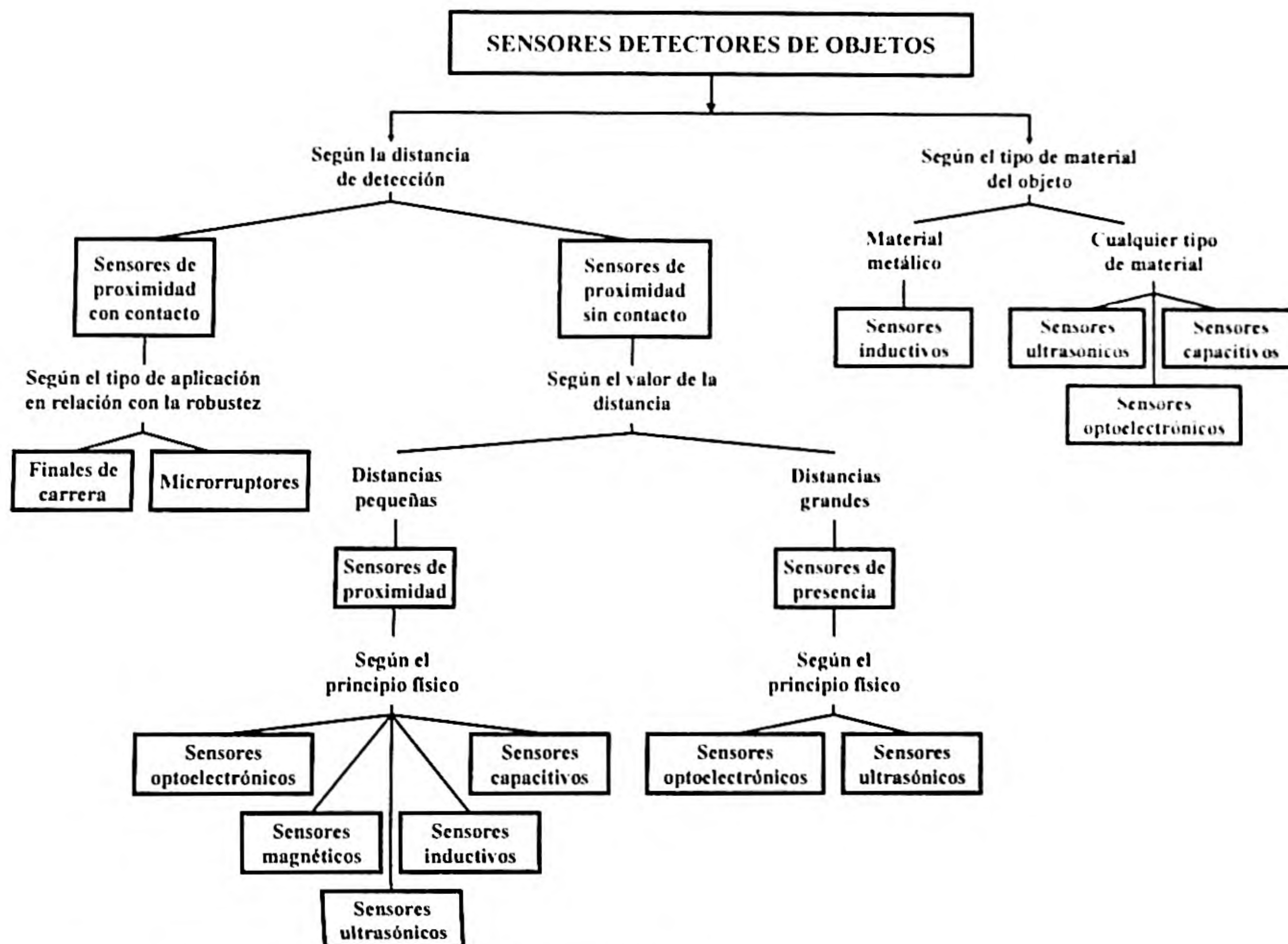


Tabla 7.7 Mapa conceptual de los sensores detectores de objetos.

Los sensores sin contacto se pueden clasificar, según el valor de la distancia a la que detectan el objeto, en sensores de proximidad (*Proximity sensors*), conocidos también como detectores de proximidad (*Proximity detectors*) [MORR 93], que detectan objetos a una distancia inferior aproximadamente a 1 cm, y sensores de presencia (*Presence sensors*) que detectan objetos a una distancia superior a 1 cm. Para referirse a los sensores de proximidad se utilizan también otras denominaciones, como por ejemplo interruptores de proximidad [SIEM 08], y sensores de proximidad digitales (*Digital Proximity switches*). En apartados sucesivos se utilizan indistintamente las denominaciones “Sensores detectores de objetos”, “Sensores de presencia” y “Sensores de proximidad”.

De acuerdo con el principio físico en el que basan su funcionamiento los sensores de proximidad sin contacto pueden ser optoelectrónicos, magnéticos, inductivos, capacitivos y ultrasónicos. Por su parte los sensores de presencia, de acuerdo también con el principio físico en el que basan su funcionamiento, pueden ser optoelectrónicos o ultrasónicos (Tabla 7.7).

Según el tipo de material del que está constituido el objeto los sensores detectores de objetos se clasifican en sensores inductivos que detectan objetos metálicos y sensores optoelectrónicos, ultrasónicos y capacitivos que detectan objetos de diferentes tipos de materiales.

Algunos fabricantes y autores [SIEM 08] denominan también sensores de proximidad a los que proporcionan una señal analógica cuyo nivel es función de la distancia entre el objeto a detectar y el sensor. Un ejemplo de ello es la fotocélula PXO500 L50HF que mide la distancia al objeto y proporciona a la salida una corriente variable entre 4 y 20 mA. Tampoco es raro utilizar otras denominaciones, como por ejemplo sensores de proximidad analógicos, para referirse a los sensores de proximidad de salida analógica continua cuya tensión o corriente de salida es proporcional a la distancia a la que se encuentra el objeto.

Existen además sensores de proximidad especialmente diseñados para automatizar procesos en los que es preciso minimizar el nivel de riesgo de accidentes. Estos sensores, denominados sensores de seguridad, deben cumplir determinadas normas y pueden ser, a su vez, con o sin contacto. Este tipo de sensores, así como sus características y aplicaciones, se describen en el apartado 10.2.3 del capítulo 10.

De los diferentes tipos de detectores de objetos indicados en el mapa conceptual de la tabla 7.7, se hace especial énfasis, en sucesivos apartados, en las características técnicas y las aplicaciones de los sensores con contacto del tipo final de carrera, y los sensores sin contacto optoelectrónicos, magnéticos, inductivos, capacitivos y ultrasónicos, por ser sensores electrónicos muy utilizados en la automatización de los sistemas de fabricación controlados mediante autómatas programables.

7.7.2.2 Sensores de proximidad con contacto

7.7.2.2.1 Introducción

Tal como se indica en el apartado 7.7.2.1, los sensores de proximidad con contacto se caracterizan porque la detección se produce cuando existe contacto físico entre el objeto a detectar y el sensor.

En múltiples instalaciones industriales es necesaria la utilización de sensores detectores de objetos sencillos, robustos, fiables y de coste reducido. Estas características son típicas de los

sensores de proximidad con contacto o de actuación mecánica, que son sensores mecánicos o de recepción de señal por contacto que presentan algunas ventajas con respecto a los sensores sin contacto:

- Pueden detectar cualquier objeto independientemente del material con el que estén implementados.
- No les afectan las interferencias procedentes del medio exterior, como por ejemplo los ruidos eléctricos, las fuentes de luz, las radiaciones electromagnéticas, etc.
- Su salida está constituida por uno o más contactos libres de potencial. Esto hace que se puedan utilizar, incluso simultáneamente, para proporcionar una variable binaria a un sistema electrónico digital y para controlar directamente diversos tipos de cargas eléctricas (un pequeño motor, un contactor, etc.) cuyos parámetros eléctricos (tensión de operación, corriente máxima, etc.) no superen los valores máximos que pueden soportar los contactos.
- Su funcionamiento es exclusivamente mecánico y por lo tanto no necesitan alimentación eléctrica para el funcionamiento del propio sensor.

Por otra parte, presentan el inconveniente de que en su salida se producen rebotes que es necesario eliminar cuando se conectan a un sistema electrónico y además no se pueden utilizar cuando es necesario detectar un objeto a una cierta distancia o si la velocidad de desplazamiento del mismo es elevada.

Existen dos grandes tipos de sensores con contacto que reciben la denominación de finales de carrera (*Limit switches or position switches*) [SIEM 00] y microinterruptores (*Microswitches*). La denominación de final de carrera se asocia, especialmente en la industria, a un sensor con contacto que se utiliza para detectar el final del desplazamiento de un elemento móvil en una máquina o instalación industrial. Por ello, los finales de carrera suelen ser robustos, tener capacidad para soportar golpes y trabajar en ambientes agresivos, y poseer contactos con una elevada capacidad de interrupción de corriente (varios amperios).

Por su parte, reciben el nombre de microinterruptores los sensores con contacto de reducidas dimensiones y elevada precisión, que se utilizan en aplicaciones en las que existe un espacio pequeño y además no es necesario conmutar corrientes elevadas. Este tipo de sensores suelen, en la mayoría de las aplicaciones, suministrar información a un sistema electrónico, debido a lo cual, la calidad eléctrica del contacto es una característica muy importante.

En apartados sucesivos se analizan la forma de operación y las características de los finales de carrera y de los microinterruptores. Se hace énfasis especial en las de los finales de carrera porque, como se indica en el apartado 7.7.2.1, son muy utilizados en la automatización de los sistemas de fabricación controlados mediante autómatas programables.

7.7.2.2 Finales de carrera

7.7.2.2.1 Introducción

Reciben la denominación de finales de carrera y también la de interruptores de posición (*Limit switches*) (Figura 7.51), los sensores con contacto que detectan que un objeto en movimiento alcanza una determinada posición. Se utilizan en muchas aplicaciones como elementos de seguridad y por ello tienen que ser robustos y de gran fiabilidad. En numerosas ocasiones

controlan directamente cargas eléctricas en alterna o continua y por ello han de poseer cierta capacidad de interrupción de corriente. Suelen instalarse en ambientes industriales agresivos (polvo, humedad, gases corrosivos, etc.) o a la intemperie y se pueden dividir en dos grandes grupos:

- Finales de carrera diseñados para conectarlos a un sistema electrónico que ejecuta, a través de sus salidas, todas las acciones necesarias. En general poseen un único contacto con pequeña capacidad de ruptura y por ello se les suele denominar finales de carrera de control.
- Finales de carrera diseñados para controlar directamente cargas eléctricas. Suelen estar dotados de varios contactos y se les suele denominar finales de carrera de mando.

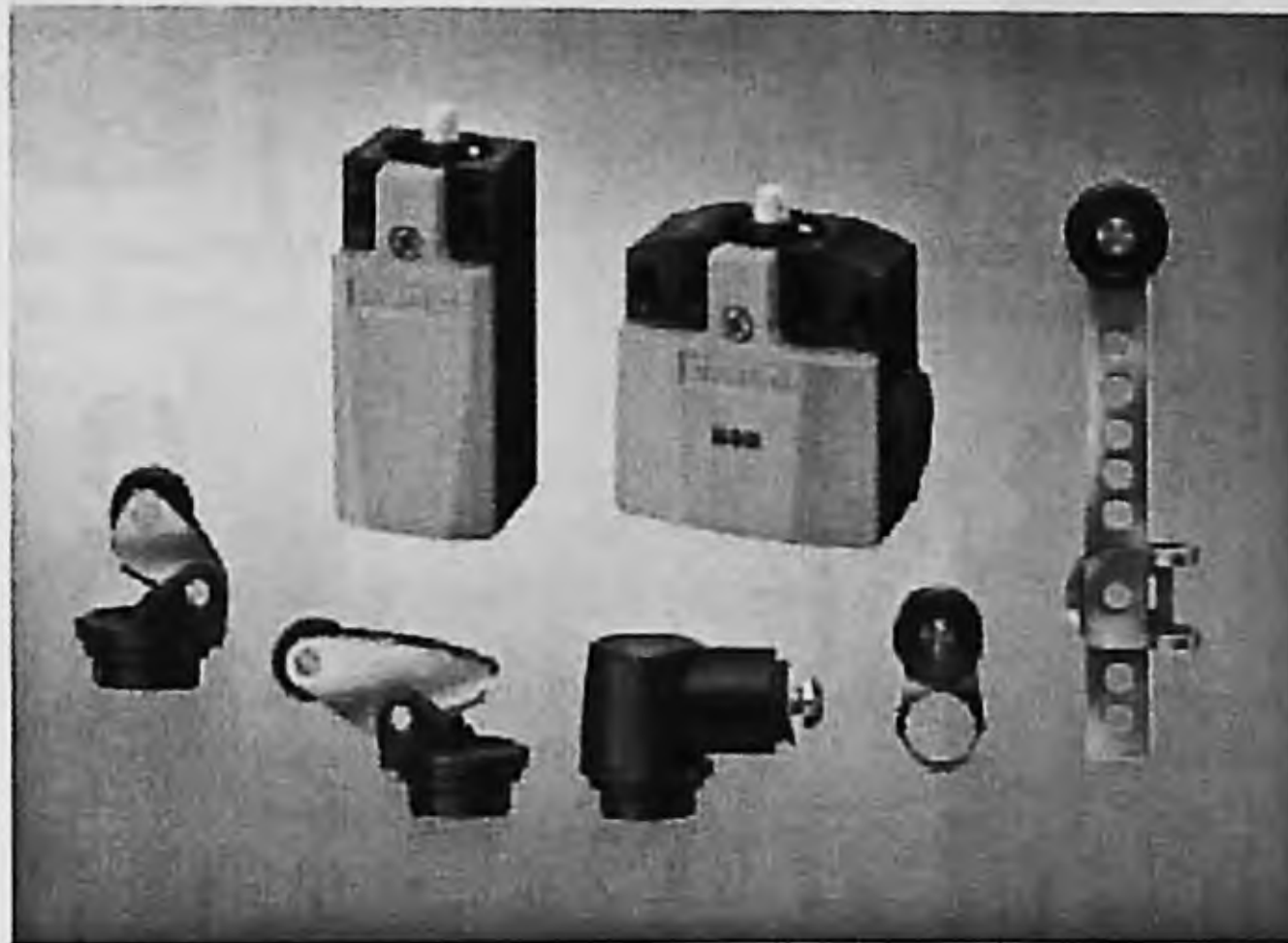


Figura 7.51 Fotografía de diversos sensores de final de carrera (cortesía de Siemens).

Los principales conceptos característicos de un final de carrera son:

- Los elementos que lo componen.
- El modo de operación.

7.7.2.2.2 Elementos de un final de carrera

Un final de carrera está formado por un conjunto de elementos que se describen a continuación.

► Actuador

Es el elemento sobre el que incide el objeto que se desea detectar. Según la forma y trayectoria del objeto a detectar pueden tener distintas formas constructivas, entre las que cabe citar las siguientes (Figura 7.51):

- **Actuadores de movimiento rectilíneo o lineales**

Son actuadores en los que el objeto a detectar incide en la misma dirección en la que se desplaza el actuador. Existen diversas variantes, de las que las más importantes son:

- *Actuadores lineales de émbolo* (Figura 7.52)
- *Actuadores lineales de rodillo*

A su vez presentan diversas variantes según que el rodillo esté colocado en la dirección del final de carrera (Figura 7.53) o en posición perpendicular a él (Figura 7.54a y b).



Figura 7.52 *Actuador de émbolo*
(cortesía de Siemens)



Figura 7.53 *Actuador de rodillo*
(cortesía de Siemens)



a)



b)

Figura 7.54 *Actuador de palanca con rodillo.* (cortesía de Siemens).

- **Actuadores de movimiento circular**

Son actuadores en los que el objeto a detectar incide lateralmente de tal forma que el actuador se mueve en la dirección del objeto con un movimiento circular. Existen también diversas variantes de las que las más importantes son:

- *Actuador circular de palanca con rodillo fijo* (Figura 7.55), o ajustable (Figura 7.56)
- *Actuador circular de tipo varilla ajustable* (Figura 7.57)
- *Actuador circular con dos rodillos fijos* (Figura 7.58)

- **Actuadores de movimiento espacial**

Son actuadores en los que el objeto a detectar puede incidir sobre ellos en cualquier dirección. Un actuador de este tipo, como por ejemplo el de varilla elástica (Figura 7.59), combina las características de los lineales y los circulares.

► **Cabeza**

Constituye el elemento sobre el que se ancla el actuador. Su función es transmitir el movimiento del actuador al mecanismo del final de carrera.



Figura 7.55 *Actuador circular de palanca de rodillo fijo (cortesía de Siemens).*



Figura 7.56 *Actuador circular de palanca de rodillo ajustable (cortesía de Siemens).*



Figura 7.57 *Actuador circular de tipo varilla ajustable (cortesía de Siemens).*



Figura 7.58 *Actuador circular con dos rodillos fijos (cortesía de Siemens).*

► Mecanismo

Es el elemento mecánico accionado por el actuador, que provoca el cierre o apertura de los contactos del final de carrera.

► Caja

Es la cápsula que contiene el mecanismo, los contactos y los terminales. Suele estar realizada con materiales resistentes a la oxidación, como por ejemplo el plástico y algunos tipos de metales como el aluminio.

► Salida

Es el orificio a través del cual se introducen los cables de conexión. Se realiza en la propia caja del final de carrera y suele estar roscado para introducir un prensaestopas que asegure el sellado del conjunto. Los finales de carrera de alto grado de protección ambiental suelen ser suministrados por el fabricante con el cable y el sistema de sellado incorporados.

► Taladros de sujeción

Son los orificios que permiten el anclaje mecánico del final de carrera. A veces son alargados en lugar de redondos para facilitar un mejor ajuste final. Los redondos están normalmente roscados para tornillos de métrica cuatro (M4) o métrica cinco (M5).



Figura 7. 59 Actuador de varilla elástica (cortesía de Siemens).

7.7.2.2.3 Modo de operación

El desplazamiento final del contacto móvil de un final de carrera se produce como consecuencia de la incidencia de un objeto sobre el actuador. La relación entre dicha fuerza y el desplazamiento se puede representar gráficamente tal como se indica en la figura 7.60, para un final de carrera típico. En ella se observan diversos puntos significativos que corresponden a los siguientes parámetros característicos:

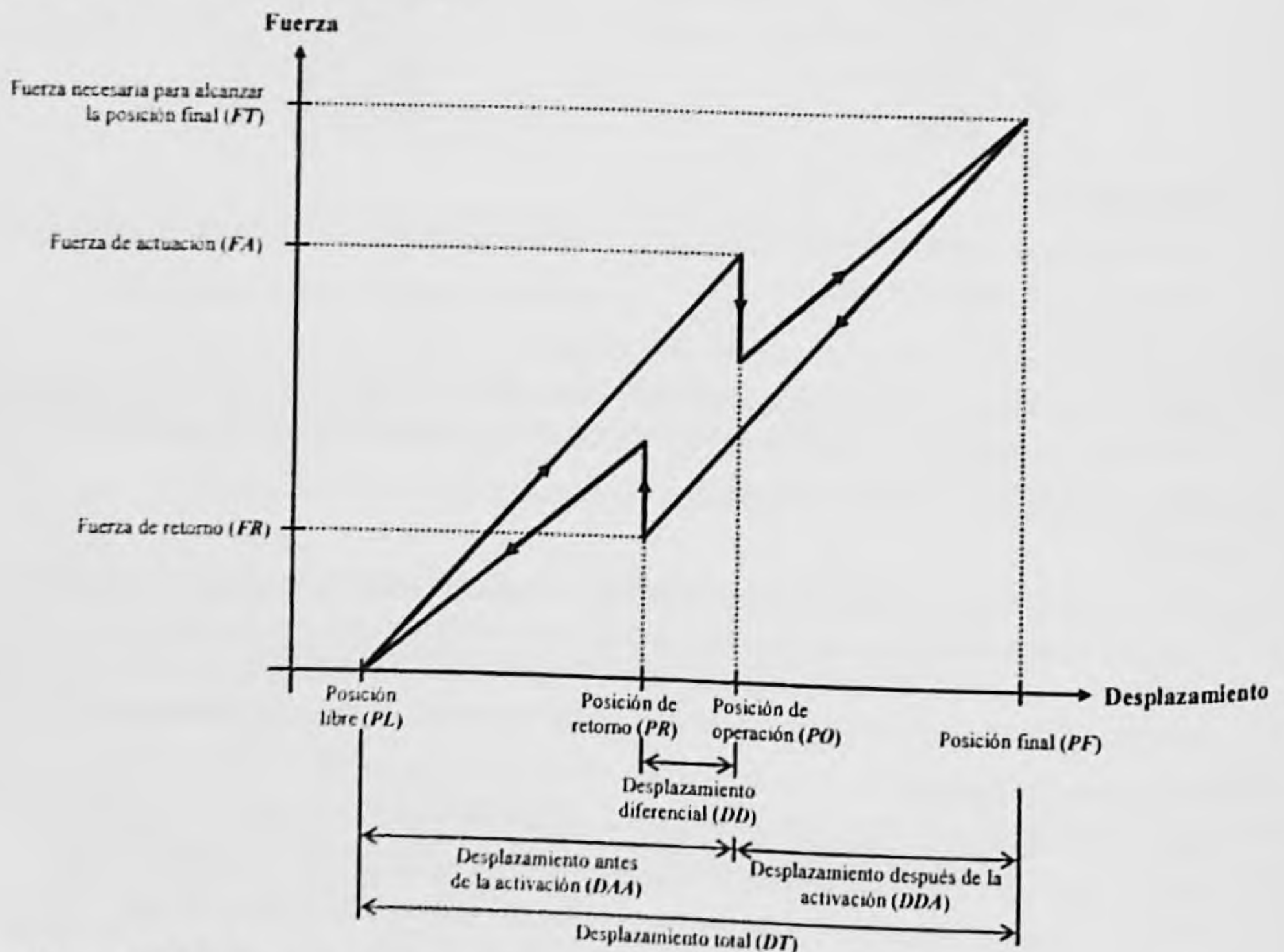


Figura 7.60 Relación fuerza-desplazamiento de un final de carrera.

- **Fuerza total FT (*Total force*)**
Es la fuerza necesaria para que el actuador realice el recorrido total.
- **Fuerza de actuación FA (*Operation force*)**
Es la fuerza necesaria, cuando crece a partir de un valor nulo, para que se produzca la actuación del contacto.
- **Fuerza de desactivación o retorno FR (*Return force*)**
Es la fuerza necesaria, cuando decrece después de haberse activado el contacto, para que se produzca la desactivación del mismo.
- **Posición libre PL (*Free position*)**
Es la posición en la que se encuentra el actuador cuando no se ejerce ninguna fuerza sobre él. Este punto se encuentra a una cierta distancia del origen de coordenadas debido a que inicialmente el mecanismo está sometido a una cierta presión para evitar holguras y mantener el contacto eléctrico normalmente cerrado o abierto.
- **Posición de desactivación o retorno PR (*Return position*)**
Es la posición en la que se encuentra el actuador cuando se produce la desactivación del final de carrera al decrecer la fuerza.
- **Posición de operación o activación PO (*Operating position*)**
Es la posición en la que se encuentra el actuador cuando se produce la activación del final de carrera al crecer la fuerza.
- **Posición final PF (*Total travel position*)**
Es la posición en la que se encuentra el actuador al final del recorrido.
- **Histéresis o desplazamiento diferencial DD (*Differential movement*)**
Representa la diferencia entre la posición de activación u operación (PO) y la de desactivación o retorno (PR) y equivale por lo tanto a la histéresis del final de carrera, debido a que el contacto se activa y desactiva en puntos distintos.
- **Desplazamiento o carrera antes de la activación DAA**
Es el espacio que recorre el actuador desde la posición libre hasta que se produce la activación del final de carrera cuando alcanza la posición de operación (PO). También se le denomina precarrera.
- **Desplazamiento o carrera después de la activación DDA**
Es el espacio que recorre el actuador desde la posición de activación hasta la posición final (PF). También se denomina sobrecarrera.
- **Desplazamiento o carrera total DT**
Es el espacio total que recorre el actuador desde la posición libre (PL) hasta la posición final (PF).
- **Trabajo antes de la activación TAA**
Representa el trabajo que es necesario realizar para que el actuador llegue al punto de operación (PO).

- **Trabajo después de la activación TDA**

Representa el trabajo que es necesario realizar para que el actuador llegue desde el punto de operación (*PO*) hasta la posición final (*PF*).

- **Trabajo total TT**

Representa el trabajo total que es necesario realizar para que el actuador llegue desde la posición libre a la posición final.

7.7.2.2.2.4 Contactos eléctricos

Los finales de carrera, al igual que los sensores electrónicos de salida de tipo relé, descritos en el apartado 7.4.1.3, pueden poseer un solo contacto o polo SP (*Single Pole*) no conmutado o de simple vía ST (*Single Through*) normalmente abiertos (NA) [*Normally open* (NO)] o normalmente cerrados (NC) [*Normally close* (NC)]. El contacto es normalmente abierto cuando está en abierto en ausencia de objeto y es normalmente cerrado en caso contrario.

También pueden poseer un contacto que conmuta entre dos terminales o de doble vía [*Double through* (DT)].

El número de contactos [independientemente de que sean de simple vía (ST) o de doble vía (DT)] del final de carrera establece el número de circuitos distintos con los que puede operar el final de carrera. Cuando el sensor tiene un solo contacto o polo se denomina unipolar o de simple polo (SP), tal como se indica anteriormente. Si tiene dos contactos se denomina bipolar o de doble polo (DP) y si tiene más se sustituye la primera letra por el número de polos en decimal. En la figura 7.24 se representa el circuito eléctrico, y la denominación correspondiente, de diferentes tipos de contactos de un relé, que es también válida para los finales de carrera. Las dos primeras siglas indican el número de polos y las dos últimas el número de vías.

En función del tipo de carga conectada al contacto, puede ser necesario protegerlo mediante alguno de los circuitos indicados en el apartado 8.2.2.4.2 del capítulo 8.

7.7.2.2.2.5 Características técnicas

Las principales características técnicas de los finales de carrera son:

- Las características eléctricas.
- Las características mecánicas.
- Las características de fiabilidad.
- Las características ambientales.

A continuación se estudia cada una de ellas.

Características eléctricas

Definen el comportamiento del final de carrera desde el punto de vista de la señal eléctrica que se aplica al contacto y las más importantes son:

- **Tensión máxima de conmutación**

Suele recibir el nombre de tensión máxima. Representa la máxima tensión que puede interrumpir el contacto eléctrico del interruptor. Este parámetro se especifica en los manuales

técnicos de los fabricantes en voltios de corriente continua (V_{CC}) y en voltios de corriente alterna (V_{CA}). Su valor está relacionado con el de la corriente máxima que puede interrumpir el contacto eléctrico.

- **Corriente de pico máxima en régimen estático**

Suele recibir el nombre de corriente de pico máxima. Es el máximo valor de la corriente que puede pasar por el contacto del final de carrera. Esta característica es fundamental para elegir el tipo de interruptor cuando la carga no es constante, como sucede con las lámparas de incandescencia (muy utilizadas en el pasado) y los motores eléctricos que consumen una corriente varias veces superior a la nominal, en el instante de la conexión.

- **Corriente máxima de conmutación**

Suele recibir el nombre de corriente máxima. Representa la máxima corriente que el contacto eléctrico del interruptor es capaz de interrumpir. En los manuales técnicos de los fabricantes se especifica su valor para diversos valores de la tensión y del tipo de carga (resistiva, inductiva, etc.). Para un mismo valor de la tensión, se puede interrumpir una corriente mucho menor en continua que alterna. Esto se debe a la energía de autoinducción, acumulada por la red de la que forma parte el contacto, que se opone a la anulación de la corriente. Este problema es de menor importancia en alterna porque la corriente se anula 100 veces por segundo en una red de 50 Hz. La corriente que pasa a través del contacto depende de la clase de carga controlada por él. Por ello, este parámetro es uno de los más importantes a tener en cuenta en la selección de un interruptor.

- **Resistencia del contacto**

Especifica la resistencia que opone el interruptor al paso de la corriente eléctrica. Se mide a través de la caída de tensión en el contacto cuando pasa por él una corriente determinada. Su valor se eleva a lo largo de la vida útil y constituye un indicativo del fin de la misma. Un valor típico de la resistencia de contacto es 50 mΩ.

- **Frecuencia de operación**

Representa el máximo número de veces que un final de carrera puede conmutar por unidad de tiempo sin que se deteriore su contacto. Su rebasamiento reduce la vida útil del interruptor, porque se produce una elevación de la resistencia del contacto debido al calor generado, que puede incluso producir la soldadura del mismo. Un valor típico de la frecuencia de operación es de 60 conmutaciones por minuto.

- **Resistencia de aislamiento**

Este parámetro indica el valor mínimo de la resistencia que debe existir entre los terminales de un interruptor y las partes no conductoras del mismo. Su medida se suele realizar con tensiones continuas dado que para tensiones alternas la resistencia es mayor debido al efecto pelicular. Un valor típico de este parámetro es de 100 MΩ para una tensión continua de 500 V.

- **Rigidez dieléctrica**

Representa la máxima tensión aplicable entre distintos elementos del interruptor sin que se produzca la rotura del material que actúa como dieléctrico. Su comprobación se suele realizar mediante tres tipos de pruebas:

- La rigidez entre terminales metálicos se comprueba aplicando una tensión de 1000 V_{CA} a 50/60 Hz durante un minuto (600 V_{CA} en el caso de interruptores que se conectan a un sistema electrónico).
- La rigidez entre las partes metálicas aisladas y cualquiera de los terminales se comprueba aplicando una tensión de 1500 V_{CA} a 50/60 Hz durante un minuto.
- La comprobación de la rigidez entre las partes aislantes y los terminales se realiza aplicando una tensión de 1500 V_{CA} a 50/60 Hz durante un minuto.

Características mecánicas

Definen el comportamiento mecánico del final de carrera. Entre las características más importantes de este tipo destacan las siguientes:

- **Resistencia a las vibraciones**

Indica la máxima vibración mecánica a la que puede ser sometido el final de carrera. Se especifica mediante la frecuencia en Hz de la vibración y su amplitud pico a pico en milímetros. Por ejemplo, la especificación de 10 a 55 Hz y 1,5 mm indica que el final de carrera correspondiente puede soportar una vibración de frecuencia comprendida entre 10 y 55 Hz y de valor pico a pico 1,5 mm. Cuanto mayor es el actuador, menor es la resistencia del interruptor a las vibraciones. Este parámetro se suele especificar en los actuadores del tipo émbolo, que son los de mayor solidez.

- **Resistencia a los golpes**

Este parámetro característico fija el impacto al que se puede someter el final de carrera sin que se produzca su destrucción. Normalmente se expresa en unidades de aceleración (m/s^2) y representa la fuerza máxima con la que el final de carrera puede impactar sobre una superficie inmóvil sin que deje de funcionar correctamente. También resulta frecuente expresar este parámetro en unidades de G (aceleración de la gravedad, $9.81 m/s^2$). Este parámetro hace referencia a otros dos que reciben el nombre de durabilidad mecánica y durabilidad de mal funcionamiento. La durabilidad mecánica representa la resistencia máxima a los golpes por parte del final de carrera sin que llegue a romperse físicamente. Un valor típico de durabilidad mecánica de los finales de carrera es $980 m/s^2$ equivalente a 100 G. La durabilidad de malfuncionamiento representa la resistencia máxima del final de carrera a los golpes sin que deje de funcionar correctamente. Un valor típico de durabilidad de mal funcionamiento es de 20 G.

- **Peso**

Se suele especificar en gramos y constituye un parámetro de gran importancia cuando el final de carrera se instala en un sistema mecánico estructuralmente ligero o sometido a algún tipo de vibraciones.

- **Dimensiones**

Las dimensiones son importantes para realizar el diseño del sistema del que forma parte el final de carrera y constituyen un factor clave en las aplicaciones en las que el espacio disponible es reducido.

- **Velocidad de actuación**

Indica los límites de velocidad del actuador para que opere correctamente. Un rango típico es de 0,1 mm/s a 1 m/s.

- **Frecuencia de actuación**

Define el número máximo de veces por unidad de tiempo que puede conmutar el contacto del final de carrera sin deteriorarse por causas mecánicas. La frecuencia de actuación mecánica es mayor que la frecuencia de actuación eléctrica y por tanto esta última es la que define la frecuencia máxima de actuación del contacto del final de carrera. Si la frecuencia de actuación mecánica supera los valores especificados por el fabricante, se producen rebotes y vibraciones en los contactos, que hacen que el interruptor no opere correctamente en esa aplicación. El funcionamiento del final de carrera en esas condiciones produce un desgaste prematuro que trae consigo una disminución de su vida útil. Un valor típico de este parámetro es de 400 operaciones por minuto.

- **Fuerza de operación**

Indica la fuerza que se debe aplicar sobre el accionador del final de carrera para activar su contacto eléctrico. Es superior a la que se necesita para mantener el contacto activado una vez accionado el final de carrera. Normalmente los fabricantes especifican el valor mínimo de la fuerza de operación que garantiza la actuación del final de carrera. Su valor varía entre algunos gramos y algunos cientos de gramos.

- **Fuerza total**

Es la fuerza necesaria para llevar el actuador desde la posición libre (PL) hasta la posición límite o final de la carrera.

- **Carrera total**

Es la distancia longitudinal o angular entre la posición libre (PL) y la posición límite o final de la carrera (PF). Es la suma del desplazamiento antes de la activación (precarrera) y el desplazamiento después de la activación (sobrecarrera).

Características de fiabilidad

Cuantifican el período de vida útil que cabe esperar del final de carrera en condiciones de funcionamiento especificadas. Las más importantes son:

- **Vida útil eléctrica**

Este parámetro establece el número mínimo de maniobras (aperturas y cierres) que es capaz de realizar el final de carrera sin que se deteriore su contacto eléctrico. Su valor depende principalmente del de la corriente eléctrica que interrumpe el final de carrera y disminuye al incrementarse el valor de la misma. Igualmente para un mismo valor de la corriente, la vida útil es mayor si la corriente es alterna que si es continua. Los fabricantes suelen especificar este parámetro de forma gráfica en función de la corriente alterna para distintos valores de la tensión, la frecuencia de operación y el $\cos \varphi$, en determinadas condiciones ambientales.

- **Vida útil mecánica**

Indica el número mínimo de maniobras mecánicas que garantiza el fabricante siempre y cuando el valor de todos los demás parámetros característicos no rebase el límite especi-

ficado. Su valor típico es de 10^6 operaciones, pero se reduce, principalmente, al rebasar la posición final o límite de carrera del final de carrera.

Características ambientales

Definen las condiciones ambientales de funcionamiento del final de carrera. Las más importantes son:

- **Temperatura ambiente**

Establece el rango de temperaturas en el que puede trabajar el final de carrera. Frecuentemente los fabricantes ofrecen modelos especiales que pueden trabajar correctamente en un margen de temperaturas muy superior al de los modelos normalizados. Un rango típico es de -15°C a $+65^{\circ}\text{C}$, siempre que no haya escarcha.

- **Humedad ambiente**

Establece el máximo valor de la humedad relativa con la que puede trabajar el final de carrera. Un valor típico es el 70%.

- **Grado de protección ambiental o sellado**

Establece la resistencia que el final de carrera presenta a la entrada de agentes externos sólidos o líquidos y se especifica de acuerdo con la norma IEC144 descrita en el apartado 7.5.2.

7.7.2.2.2.6 Criterios de selección de los finales de carrera

La selección del final de carrera más adecuado para una determinada aplicación es función de los valores de los diferentes parámetros eléctricos, mecánicos y medioambientales que la misma impone. Por ello, los criterios a utilizar corresponden a los citados parámetros y tienen como finalidad garantizar una adecuada vida útil del final de carrera.

Criterios eléctricos

Los parámetros eléctricos que más influyen en la selección de un final de carrera son la corriente de pico máxima y la corriente máxima de conmutación que se estudian en el apartado 7.7.2.2.2.5. Por ello el tipo de carga controlada por el final de carrera constituye uno de los factores más importantes a tener en cuenta en la selección del mismo.

Cuando el final de carrera controla directamente una carga en continua o en alterna, la capacidad de interrupción se convierte en un factor clave. Como norma general se debe seleccionar un final de carrera con una capacidad mínima de conducción y conmutación de corriente 1,5 veces superior a la de la carga conectada a él y con capacidad de interrupción de las tensiones que se generen en la conmutación de aquella.

Por lo tanto, para seleccionar el final de carrera es necesario conocer el tipo de carga y sus características. Las cargas se pueden clasificar en cuatro grandes tipos:

- **Resistivas de valor constante**

Si la carga es resistiva, la corriente en el instante de cerrarse el final de carrera es idéntica a la del régimen permanente y no se produce ningún tipo de sobretensión en el contacto, tanto si la alimentación es continua como si es alterna.

- **Resistivas de tipo filamento**

Si la carga es resistiva pero está constituida por una lámpara de filamento, se produce un pico de corriente (*Rush current*) del orden de 10 veces superior al nominal, debido a la baja resistencia del filamento en frío en el instante en el que se cierra el contacto del final de carrera. Este pico se produce tanto si la lámpara se alimenta en continua como en alterna.

- **Inductivas estáticas**

Si la carga es inductiva pura o estática (bobina de un relé, devanado de un transformador, etc.), se genera una fuerza electromotriz que tiende a mantener la corriente en el instante en el que se abre el contacto del final de carrera. Dicha fuerza electromotriz da lugar a la generación de un pico de tensión entre el contacto, tanto en continua como en alterna, que puede provocar la aparición de un arco. Para eliminar dicho arco se deben incorporar circuitos de protección de los contactos como los indicados en el apartado 8.2.2.4.2 del capítulo 8.

- **Inductivas tipo motor**

Si la carga es un motor el comportamiento del circuito depende del tipo de motor utilizado. El carácter inductivo de cualquier motor produce siempre una sobretensión al abrirse el contacto del final de carrera tanto si la alimentación es alterna como si es continua. Las diferencias aparecen al cerrarse el final de carrera. Si la tensión es alterna y el motor es monofásico con arranque bifásico (por ejemplo por medio de un condensador) para aumentar el par de arranque, se produce un pico de corriente que se amortigua lentamente. Si por el contrario la tensión es continua, el comportamiento de la corriente depende de si el motor es de tipo derivación o de tipo serie. En este último tipo se produce un pico de corriente al cerrarse el contacto, del orden de seis veces la corriente nominal, similar a la carga de tipo filamento. Por otra parte, en alterna el arco debido a las sobretensiones de apertura con cargas inductivas se extingue de forma natural al pasar la corriente por cero, al contrario de lo que sucede en continua en que esa situación no se produce. En ambos casos se deben incorporar circuitos de protección de los contactos como los indicados en el apartado 8.2.2.4.2 del capítulo 8.

Para ayudar al usuario a elegir el final de carrera más adecuado, teniendo en cuenta todo lo que se acaba de exponer, el fabricante especifica los valores máximos de la corriente nominal que puede controlar un final de carrera para cada tipo de carga y con diferentes valores de la tensión de alimentación.

Criterios mecánicos

El parámetro mecánico más importante a tener en cuenta en la selección de un final de carrera, a fin de asegurar su correcto funcionamiento, es la carrera o desplazamiento total del actuador. En general se deben tener en cuenta las siguientes consideraciones en relación con ella:

- La carrera o desplazamiento total DT se debe elegir de forma que el actuador recorra todo el camino y llegue hasta la posición libre (PL) cuando se desactiva el final de carrera (Figura 7.60).
- En las aplicaciones en las que el actuador no alcanza la posición libre (PL) al desactivarse, se debe utilizar un final de carrera en el que la operación del actuador esté invertida.

- La carrera o desplazamiento después de la activación DDA, a partir del punto PO en el que se activa el final de carrera, en la figura 7.60, debe estar comprendida entre el 70% y el 100% del valor especificado por el fabricante. De esta forma se consigue que el final de carrera continúe activado aunque se produzca una pequeña reducción de la fuerza que el objeto ejerce sobre el actuador y que la presión entre los contactos sea la adecuada para garantizar un correcto funcionamiento. Si el valor de DDA es, por el contrario, inferior al 70% del valor especificado por el fabricante, se disminuye la presión entre los contactos, lo cual contribuye a su carbonización (especialmente cuando la corriente a través de carga es elevada) y a la consiguiente reducción de su vida útil.

Otro parámetro mecánico importante es la velocidad de operación, dada su influencia en la forma en que se produce la conmutación de los contactos del final de carrera. Con relación a la velocidad de operación se deben tener en cuenta los siguientes criterios:

- Una velocidad de operación excesivamente baja produce un alargamiento del intervalo de tiempo durante el cual la presión entre el contacto móvil y el fijo es reducida, lo que provoca vibraciones. Esto da lugar a la generación de arcos que, especialmente con cargas en continua, producen un deterioro de los contactos que reduce la vida útil del final de carrera, y además aumenta el desgaste de los contactos porque durante el tiempo de vibración se produce el deslizamiento del contacto móvil sobre el fijo.
- Una velocidad de operación excesivamente alta aumenta los rebotes en el instante en que el contacto móvil toca al fijo. Por otra parte provoca un golpe fuerte en el momento en que el objeto toca al actuador que trae consigo también una disminución de la vida útil del final de carrera.

Las dimensiones del final de carrera resultan además importantes en las aplicaciones en las que se dispone de un espacio reducido para su ubicación.

Criterios ambientales

El fabricante proporciona entre las especificaciones de los finales de carrera las relativas a las condiciones ambientales que aseguran su correcto funcionamiento, en especial en lo referente a temperatura y humedad relativa.

En ambientes hostiles, como son las atmósferas corrosivas (existencia de H_2S , SO_2 , Cl_2 , NH_3 o NO_2), las atmósferas con polvo en suspensión, las atmósferas explosivas, la existencia de chorros de agua o la presencia de aceite, es necesario utilizar finales de carrera estancos que posean el grado de protección IP adecuado.

También es posible que en ciertas aplicaciones se exijan determinadas homologaciones, otorgadas por laboratorios o centros de homologación.

7.7.2.2.3 Microrruptores

Los microrruptores o microinterruptores son dispositivos detectores de objetos de reducidas dimensiones y elevada precisión, especialmente idóneos para ser utilizados en máquinas que tienen un elevado nivel de automatización (fotocopiadoras, fax, etc.) y en las que hay que detectar la posición de un determinado mecanismo. En la figura 7.61 se representa la forma física típica de un microrruptor.

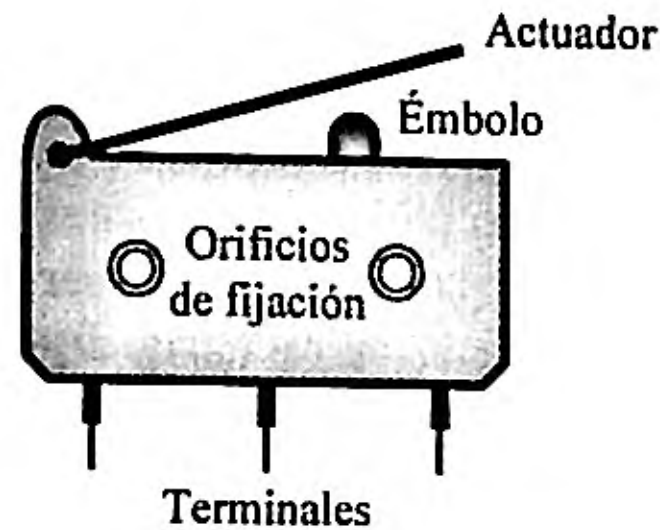


Figura 7.61 Forma física típica de un microinterruptor.

En cuanto a las características técnicas, modo de operación, etc., lo indicado para los finales de carrera es válido para los microinterruptores, con la salvedad de que son mucho menos robustos que aquéllos y de que, en general, no soportan ambientes tan agresivos.

En cuanto a los contactos eléctricos, deben ser de gran calidad para garantizar que su resistencia es muy pequeña cuando se cierran, aunque no es necesario que conmuten corrientes elevadas, porque normalmente se utilizan para dar información de un movimiento a un sistema electrónico. Por ello, los contactos eléctricos de los microinterruptores se suelen realizar en forma de botón o de barras cruzadas. A su vez, cada uno de ellos presenta dos variantes según que el elemento móvil posea un único contacto (botón o barra cruzada única) o dos (botón o barra cruzada bifurcada). Los contactos de remache o botón se utilizan básicamente en los microinterruptores diseñados para controlar corrientes relativamente elevadas durante intervalos de tiempo largos. Su fabricación se realiza normalmente con aleaciones de plata, lo que proporciona una elevada conductividad y un valor reducido de la resistencia específica del contacto. En ellos se suele formar una película sobre los contactos por oxidación, sulfuración o carbonización, que se puede eliminar fácilmente haciendo pasar a su través impulsos de corriente de 1 a 5A con una tensión del orden de 100 V_{CA}. Los contactos de barras cruzadas se suelen utilizar con cargas a través de las que circula una corriente reducida durante un pequeño intervalo de tiempo (cargas impulsionales). Esta situación es típica cuando se conecta un microinterruptor a un sistema electrónico de control (dispositivo lógico programable, microprocesador, etc.). En estas aplicaciones es necesario eliminar los rebotes que se producen en los contactos por medio del circuito de filtrado adecuado [MAND 95]. Esta forma constructiva de los contactos presenta la ventaja de que reduce la influencia del polvo o de cualquier otra sustancia que pueda situarse sobre ellos. En este tipo de contactos también se pueden desarrollar películas de oxidación y sulfurización, debido en gran medida al paso de una corriente reducida a través de ellos. Para evitar este inconveniente y alargar su vida útil, los contactos se recubren de una película de oro o de alguna de sus aleaciones, que poseen una elevada resistencia a la oxidación y a la sulfurización.

Sus aplicaciones son similares a las de los finales de carrera y se utilizan para detectar que un objeto alcanza una determinada posición. La diferencia con los finales de carrera, en lo que se refiere a las aplicaciones, es el tipo de objeto y el ambiente de trabajo. A continuación se indican sus aplicaciones más habituales:

- **Detección de apertura y cierre de puertas**

Se utilizan en diferentes máquinas industriales, como por ejemplo máquinas electrónicas de oficina, electrodomésticos, automóviles, etc.

- **Detección de presencia de objetos**

Se utilizan en diferentes máquinas electrónicas, como por ejemplo, reproductores de vídeos, máquinas expendedoras, máquinas recreativas, etc. Las principales características que deben poseer los microinterruptores utilizados en este tipo de aplicaciones son reducidas dimensiones, pequeña fuerza de operación para poder ser activados por piezas que forman parte de mecanismos complejos y fácil adaptación de diversos tipos de actuadores e incluso posibilidad de diseño del actuador a medida.

- **Detección de piezas en movimiento**

Se utilizan en diferentes tipos de máquinas tales como máquinas de fabricación automática, máquinas electrónicas de oficina, vehículos, etc. para detectar la puesta en marcha, la parada y el inicio o finalización de operaciones concretas. Normalmente los microinterruptores utilizados en este grupo de aplicaciones están sometidos a condiciones de trabajo mecánicas y medioambientales más severas (existencia de vibraciones, elevada frecuencia de operación, presencia de polvo, etc.). Por todo ello, las principales características que deben poseer los microinterruptores utilizados en este tipo de aplicaciones son elevada precisión de las posiciones de activación y desactivación, elevada vida útil mecánica y fácil mantenimiento.

- **Detección de órdenes manuales**

En este grupo de aplicaciones los microinterruptores sustituyen con ventaja a los conmutadores que retornan a la posición de reposo (denominados monoestables) porque son más fiables, seguros y más agradables al tacto.

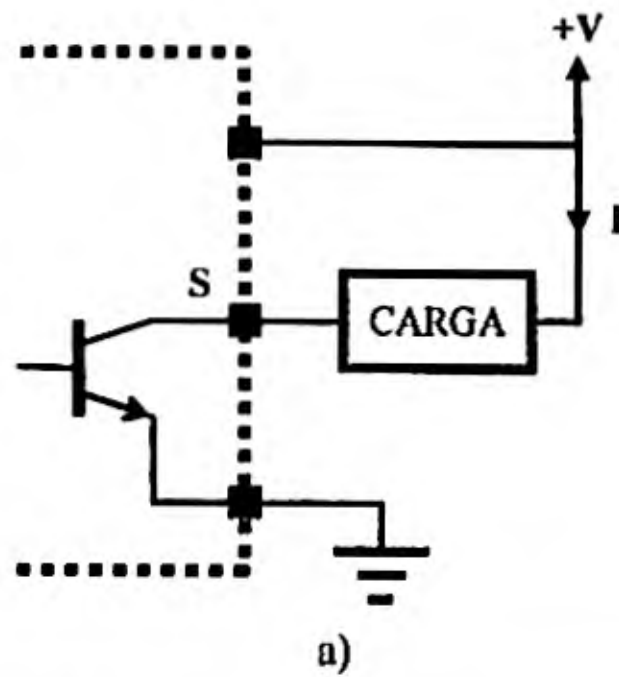
7.7.2.3 Sensores de proximidad sin contacto

7.7.2.3.1 Conceptos generales

Los sensores de proximidad (*Proximity sensors*) sin contacto son sensores todo-nada que detectan el objeto sin necesidad de que exista un contacto físico entre ambos. En el mapa conceptual de la tabla 7.7 se indican, de acuerdo con la distancia y el tipo de material del objeto, los tipos más utilizados, que se analizan con detalle en sucesivos apartados a continuación.

Los sensores de proximidad sin contacto tienen un conjunto de características comunes a todos ellos, independientemente del principio físico en el que se basa su funcionamiento. Entre ellas figuran las características eléctricas, que, además, son comunes a todos los sensores todo-nada y se describen en el apartado 7.4.1.3.

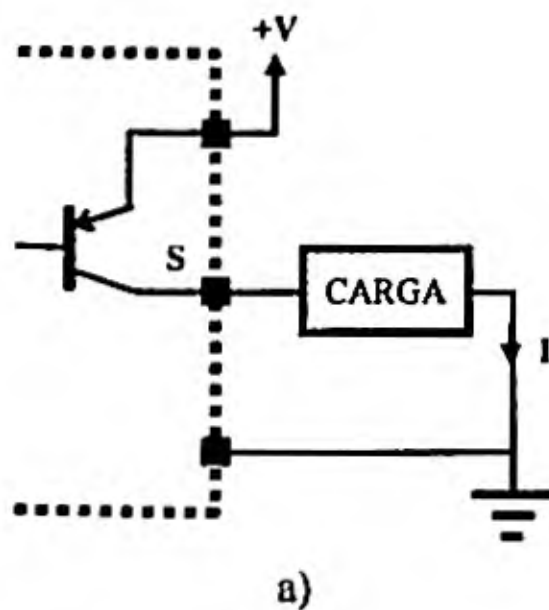
Los sensores de proximidad pueden tener salida todo-nada normalmente abierta NO o normalmente cerrada NC y estar implementados con relés o con transistores NPN o PNP. Si la salida está implementada con un relé es normalmente cerrada cuando el contacto está cerrado en ausencia de objeto y normalmente abierta en caso contrario, al igual que los sensores de proximidad con contacto. Si el sensor es de tres hilos y está implementado con un transistor NPN, su salida es normalmente abierta si está al nivel de cero voltios en presencia de objeto y a + V voltios en ausencia de objeto y es normalmente cerrada en caso contrario, tal como se indica en la figura 7.62. Si el sensor es de tres hilos y está implementado con un transistor PNP, su salida es normalmente abierta si está al nivel de cero voltios en ausencia de objeto y a + V voltios en presencia de objeto y es normalmente cerrada en caso contrario, tal como se indica en la figura 7.63.



Salida NO		Salida NC	
Presencia de objeto	Ausencia de objeto	Presencia de objeto	Ausencia de objeto
$S = 0V$	$S = +V$	$S = +V$	$S = 0V$
$I > 0$	$I = 0$	$I = 0$	$I > 0$

b)

Figura 7.62 Funcionamiento de la salida con transistor NPN de un sensor de proximidad sin contacto de tres hilos: a) Esquema del circuito; b) Descripción del comportamiento.



Salida NO		Salida NC	
Presencia de objeto	Ausencia de objeto	Presencia de objeto	Ausencia de objeto
$S = +V$	$S = 0V$	$S = 0V$	$S = +V$
$I > 0$	$I = 0$	$I = 0$	$I > 0$

b)

Figura 7.63 Funcionamiento de la salida con transistor PNP de un sensor de proximidad sin contacto de tres hilos: a) Esquema del circuito; b) Descripción del comportamiento.

7.7.2.3.2 Símbolos normalizados

Los sensores de proximidad sin contacto se comercializan con terminales o con cables y por ello se han desarrollado símbolos normalizados, incluidos en la norma UNE EN 60947-5-2 [AENO 05], en los que, además del símbolo propiamente dicho, se establece el tipo de sensor, la numeración de los terminales y el color de los cables conectados a los distintos terminales. En la tabla 7.8 se indican los colores que se utilizan para identificar los cables conectados a los distintos terminales y su correspondiente acrónimo en inglés.

En la figura 7.64 se representa el símbolo normalizado básico, que consta de dos partes. En la parte inferior izquierda se coloca un rombo que indica que se trata de un sensor de proximidad y sobre él se indica, mediante una letra X, el tipo de sensor de que se trata de acuerdo con la tabla 7.9.

Son numerosos los fabricantes que no cumplen la norma en su totalidad. Por ejemplo, es habitual indicar los inductivos mediante el símbolo de una bobina en lugar de la letra I y los capacitivos mediante el símbolo de un condensador en lugar de la letra C. Debido a su complejidad, los sensores optoelectrónicos de proximidad poseen además algunos tipos de señales especiales, como por ejemplo de selección del funcionamiento en luz o en oscuridad, de inhibición, etc. Por ello, su símbolo se estudia en el apartado 7.7.2.3.3.1.

COLOR	ABREVIATURA
Negro (Black)	BK
Marrón (Brown)	BN
Rojo (Red)	RD
Amarillo (Yellow)	YE
Verde (Green)	GN
Azul (Blue)	BU
Gris (Grey)	GY
Blanco (White)	WH
Dorado (Gold)	GD
Verde/Amarillo (Green/Yellow)	GNYE

Tabla 7.8 Códigos de colores de los cables de los sensores detectores de objetos sin contacto.

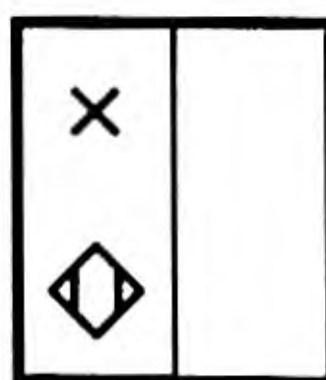


Figura 7.64 Símbolo lógico normalizado básico de los sensores todo/nada sin contacto.

Letra	Tipo de sensor
T	Emisor y receptor. Barrera fotoeléctrica (Sensor optoelectrónico de barrera de luz)
R	Fotoeléctrico réflex (Sensor optoelectrónico con reflector)
D	Emisor-receptor y objeto. Sensor fotoeléctrico de detección directa. (Sensor optoelectrónico de reflexión directa o de reflexión en objeto)
I	Sensor inductivo
C	Sensor capacitivo
U	Sensor ultrasónico

Tabla 7.9 Letra asignada a los diferentes tipos de sensores detectores de objetos en la norma UNE EN 60947-5-2.

En la parte derecha del símbolo se representa, con letras (color del cable) o con números (número del terminal) y gráficamente, el tipo de salida normalmente abierta (NO) o normalmente cerrada (NC). Igualmente se indica mediante letras y números si el sensor se alimenta en alterna o en continua y en este caso, si es o no polarizado. Si el sensor se alimenta en alterna, la fase se representa mediante la letra L y el neutro mediante la letra N. Si se alimenta en continua el positivo se denomina L+ y el negativo L-. Si la alimentación se puede realizar indistintamente en alterna o en continua, la fase de alterna se indica mediante L1.

A continuación se describen los símbolos normalizados de los distintos tipos de sensores de proximidad de dos, tres y cuatro hilos. En todos ellos se coloca un rectángulo, que representa la carga externa, en el terminal o terminales a los que se conecta la misma (Figuras 7.65 a 7.73).

Símbolos de los sensores de proximidad de dos hilos

La carga externa se puede conectar a cualquiera de los dos terminales y por ello se representa tanto con trazo continuo como discontinuo (Figuras 7.65 y 7.66). Según el sensor sea o no polarizado se representan mediante los siguientes símbolos:

- **Sensores de proximidad no polarizados de dos hilos**

Se dice que no es polarizado el sensor de proximidad que se puede alimentar tanto con corriente continua (CC) como con alterna (CA). En este tipo de sensores las denominaciones L1, N, L+ y L- se colocan, todas ellas, en ambos terminales (Figura 7.65). Si la salida del sensor es normalmente abierta (NO), a los terminales se les asignan los números 3 y 4. Si la salida del sensor es normalmente cerrada (NC), a los terminales se les asignan los números 1 y 2. El color de los cables puede ser cualquiera excepto el verde, el amarillo y el verde/amarillo, pero la norma aconseja que los dos sean del mismo color. La figura 7.65a muestra el símbolo normalizado de un sensor de proximidad inductivo de dos hilos no polarizado de salida equivalente a un contacto normalmente abierto, en el que los terminales se identifican con números. En la figura 7.65b se indica el mismo tipo de sensor con salida equivalente a un contacto normalmente cerrado en el que ambos hilos de conexión son negros (BK). Para indicar que la carga se puede conectar en cualquiera de los dos terminales, se representa mediante dos rectángulos, uno en trazo continuo y el otro discontinuo.

- **Sensores de proximidad polarizados de dos hilos**

Se alimentan en CC. El terminal al que se conecta el positivo debe ser el cable marrón (BN) y el terminal al que se conecta el negativo, el cable azul (BU). Al primero se le asigna siempre el número 1 y al segundo el 2 ó el 4 según que la salida sea equivalente a un contacto normalmente cerrado o abierto, respectivamente. En la figura 7.66 se muestra el símbolo de un sensor de proximidad capacitivo polarizado cuya salida es equivalente a un contacto normalmente abierto, indicado mediante números (Figura 7.66a) y mediante letras (Figura 7.66b). Al igual que en los no polarizados, la carga se representa mediante dos rectángulos, uno en trazo continuo y el otro discontinuo, para indicar que puede estar conectada en cualquiera de los dos terminales.

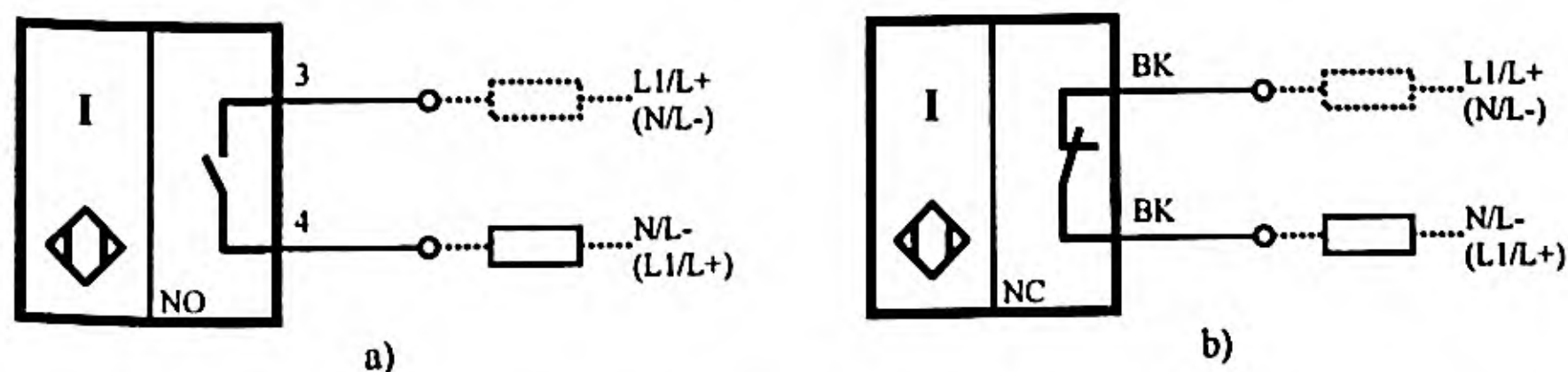


Figura 7.65 Símbolo de un sensor inductivo no polarizado de dos hilos con terminales de salida: a) Salida normalmente abierta (NO) indicada con los números de los terminales; b) Salida normalmente cerrada (NC) indicada con las letras del color del cable.

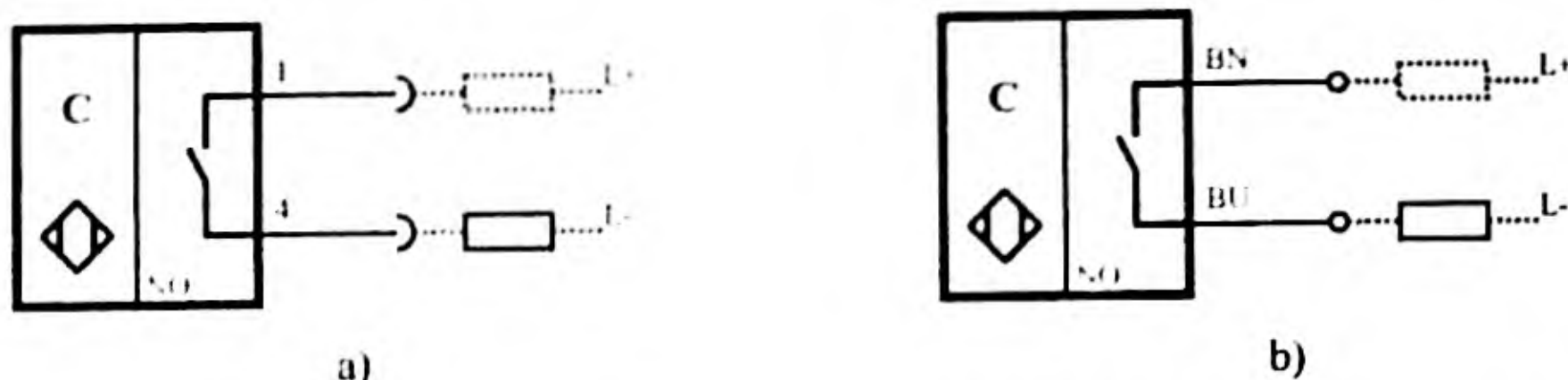


Figura 7.66 Símbolo de un sensor capacitivo polarizado de dos hilos con salida equivalente a un contacto normalmente abierto; a) Indicada con números; b) Indicada con letras.

Símbolos de los sensores de proximidad de tres hilos

Pueden ser polarizados o no y la carga externa se conecta entre dos de los tres terminales, tal como se indica en los símbolos correspondientes:

- Sensores de proximidad no polarizados de tres hilos**

Tal como se indica en la figura 7.67a y b, la alimentación se aplica siempre entre el terminal 1 [cable marrón (BN)] y el 3 [cable azul (BU)]. La salida es siempre el [Negro (BK)] y además si es NO se asigna al terminal el número 4 y si es NC el número 2.

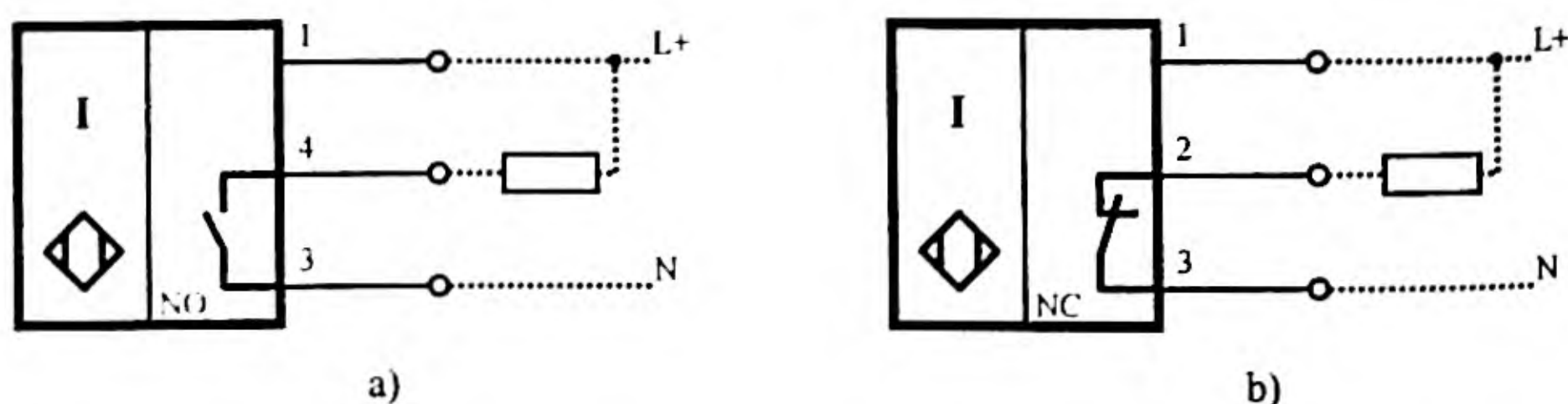


Figura 7.67 Símbolo de un sensor inductivo no polarizado de tres hilos indicados con los números de sus terminales; a) Con salida equivalente a un contacto normalmente abierto; b) Con salida equivalente a un contacto normalmente cerrado.

- Sensores de proximidad polarizados de tres hilos**

El positivo de la alimentación se aplica al terminal 1 [cable marrón (BN)] y el negativo al terminal 3 [cable azul (BU)]. La salida siempre se asigna al terminal 4 [el cable negro (BK)] (Figuras 7.68a y 7.69a).



Figura 7.68 Símbolo de un sensor capacitivo polarizado de tres hilos implementado con un transistor PNP y equivalente a un contacto NO; a) Indicado con las letras de los cables; b) Indicado con los números de los terminales.

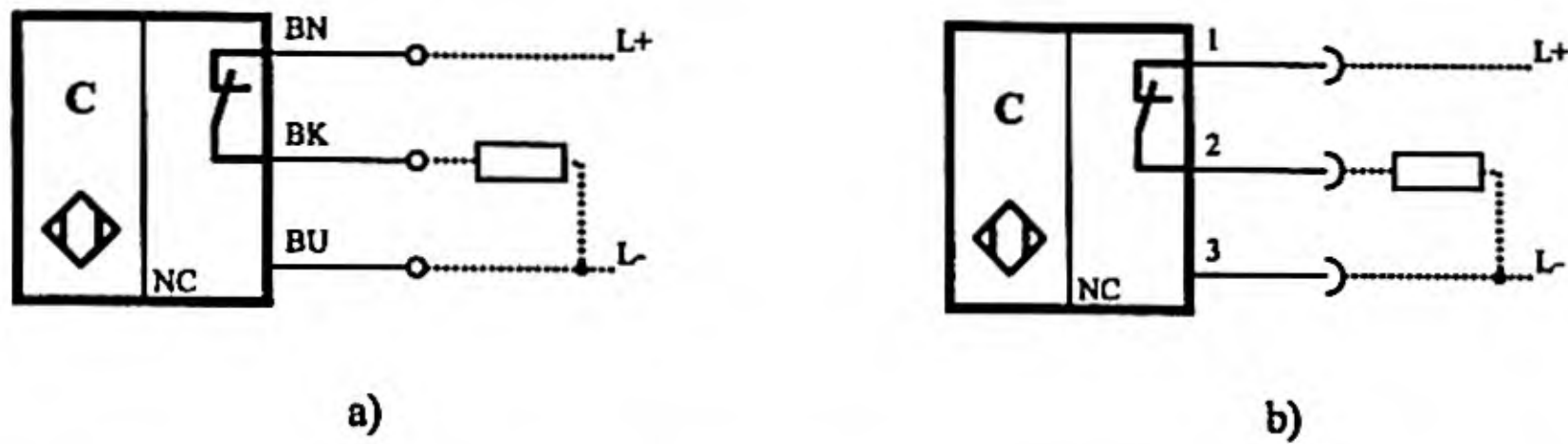


Figura 7.69 Símbolo de un sensor capacitivo polarizado de tres hilos implementado con un transistor PNP y equivalente a un contacto NC: a) Indicado con letras; b) Indicado con números.

Si la salida está implementada con un transistor PNP la carga se conecta entre la salida y el negativo. En este caso, si la salida es NO el número del terminal es el 4, tal como se indica en las figuras 7.68b y si la salida es NC el número es el 2 (Figuras 7.69b).

La salida se puede implementar también con un transistor NPN. En este caso, la carga se conecta entre la salida y el positivo, tal como se indica en las figuras 7.70a y b (contactos normalmente abiertos) y 7.71a y b (contactos normalmente cerrados).

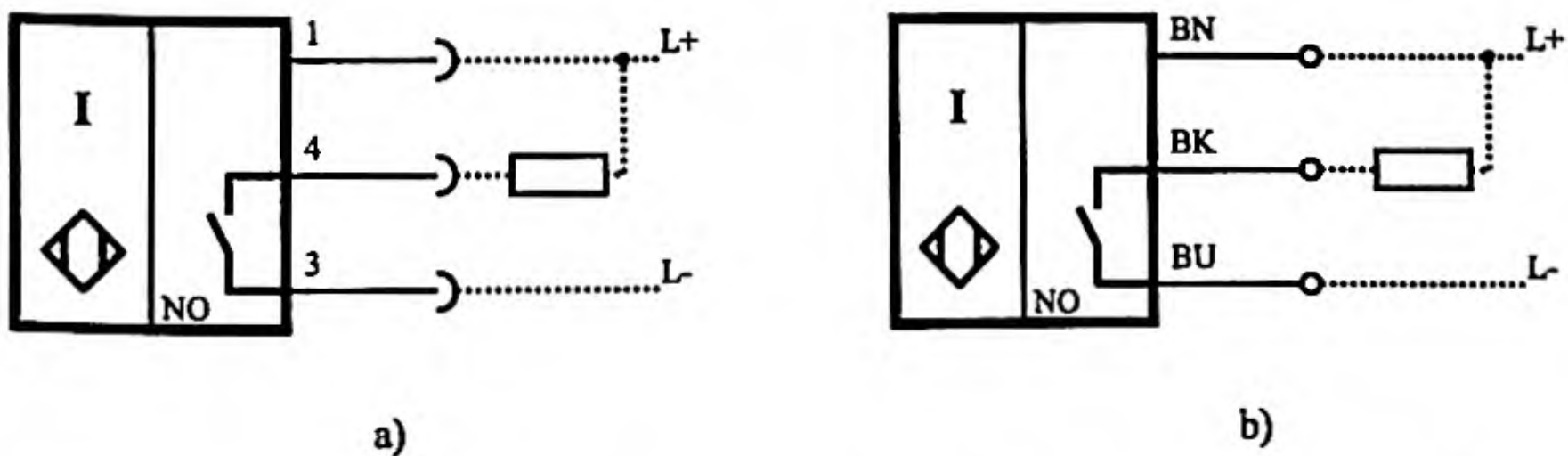


Figura 7.70 Símbolo de un sensor inductivo polarizado de tres hilos implementado con un transistor NPN equivalente a un contacto NO: a) Indicado con números; b) Indicado con letras.

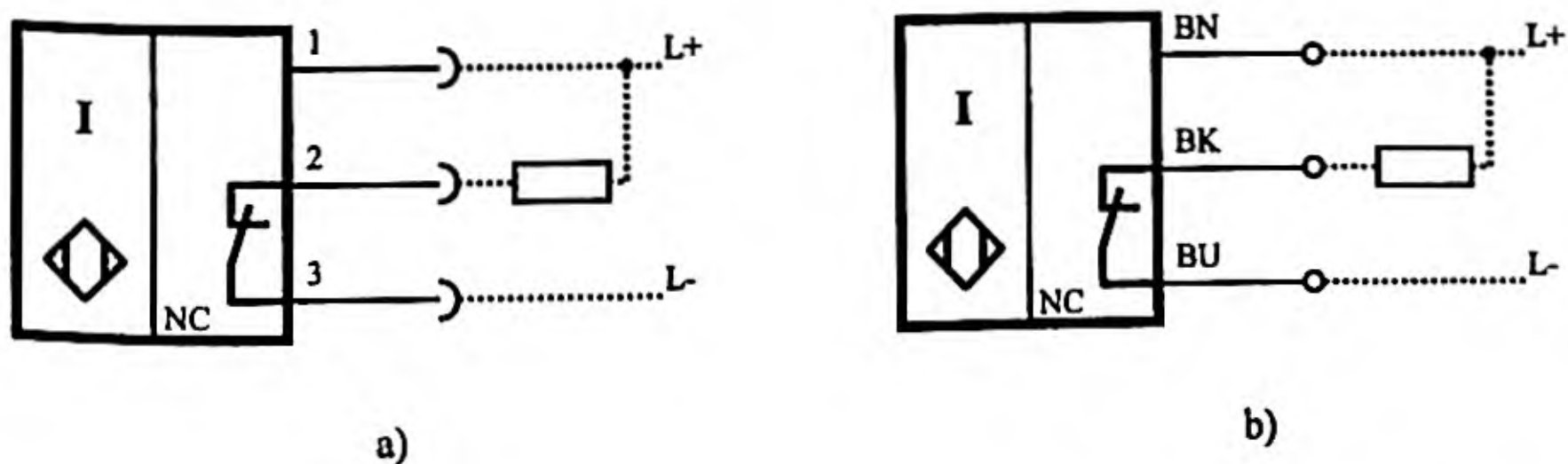


Figura 7.71 Símbolo de un sensor inductivo polarizado de tres hilos implementado con un transistor NPN y equivalente a un contacto NC: a) Indicado con números; b) Indicado con letras.

Símbolos de los sensores de proximidad de cuatro hilos

En la figura 7.72 se muestra el símbolo normalizado de los sensores de cuatro hilos. El positivo de la alimentación se aplica al terminal "1" [cable marrón (BN)] y el negativo al 3 [cable azul (BU)]. La salida normalmente abierta (NO) es siempre el terminal 4 [Negro (BK)] y la normalmente cerrada NC es siempre el terminal 2 [cable blanco (WH)]. Además las salidas pueden ser de tipo PNP o NPN y la conexión de la carga es igual que en el apartado anterior.

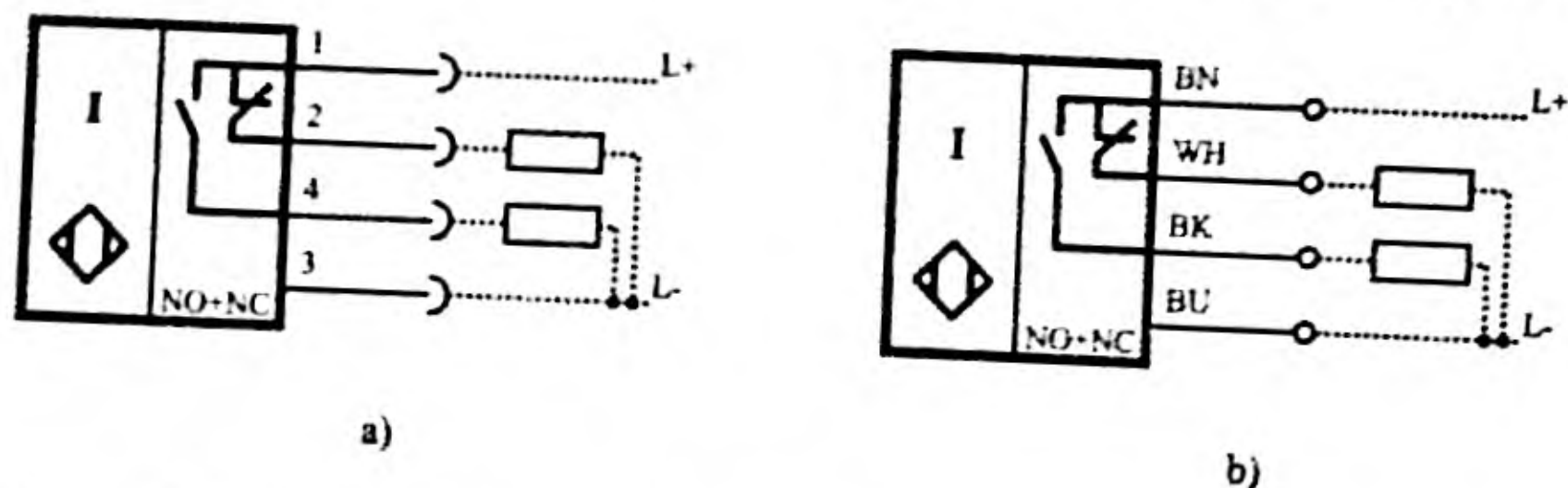


Figura 7.72 Símbolo de un sensor inductivo polarizado de cuatro hilos implementado con transistores PNP: a) Indicado con números. b) Indicado con letras.

7.7.2.3.3 Sensores optoelectrónicos de proximidad

7.7.2.3.3.1 Introducción

Los sensores optoelectrónicos de proximidad detectan la presencia de un objeto mediante fenómenos relacionados con la luz. Suelen recibir diferentes nombres, entre los que cabe citar los de fotocélulas (*Photoelectric Switches or photocells*), detectores optoelectrónicos (*Optoelectronic detectors*), sensores de proximidad ópticos (*Optical proximity sensors*) y detectores de proximidad fotoeléctricos (*Proximity photoelectric detectors*). Se componen, en general, de un emisor y un receptor. El emisor genera un rayo de luz dentro del espectro visible, infrarrojo cercano o láser. El receptor recibe o no el rayo emitido por el emisor, o lo recibe con algún cambio específico de sus características, según que en su trayectoria encuentre o no el objeto a detectar. Existen numerosas formas de realizar los sensores optoelectrónicos de proximidad, lo que da lugar a un gran número de variantes.

En sucesivos apartados se describen los tipos más habituales de sensores optoelectrónicos que funcionan como sensores de presencia y son conocidos habitualmente como fotocélulas.

Sensores optoelectrónicos	Barrera de luz		
	Réflex	Reflexión sobre espejo	
		Reflexión sobre objeto	Reflexión difusa
			Reflexión definida

Tabla 7.10 Clasificación de los sensores optoelectrónicos detectores de objetos.

De acuerdo con la forma en que se disponen el emisor y el receptor y con el método de detección utilizado, los sensores optoelectrónicos se pueden clasificar tal como se indica en la tabla 7.10.

Según la forma en la que se disponen el emisor y el receptor las fotocélulas son:

- **Fotocélulas de barrera de luz (*Through-beam photocells*)**

En ellas el emisor y el receptor están físicamente separados y el objeto interrumpe el haz.

- **Fotocélulas de reflexión (*Reflex photocells*)**

En ellas el emisor y el receptor están colocados en la misma carcasa, y el rayo de luz sale del emisor, se refleja y vuelve al receptor. La luz se puede reflejar en un espejo (reflexión sobre espejo), o en el propio objeto (reflexión sobre objeto o reflexión directa) (*Retroreflective photocell*). En este último caso, y según la constitución física de la fotocélula, la reflexión sobre el objeto puede ser difusa o definida.

El elevado número de variables utilizadas en las fotocélulas hace que algunos fabricantes no sigan de forma estricta la norma UNE EN 60947-5-2. En la figura 7.73 se representa el símbolo de una fotocélula de reflexión de cuatro hilos, utilizado por Siemens.

Los fabricantes suelen utilizar los siguientes colores de los terminales especiales:

- Gris (GR) para las entradas de desinhibición o habilitación.
- Rojo (RD) para la entrada de control del funcionamiento en luz o en oscuridad.
- Naranja (OR) para las salidas analógicas.

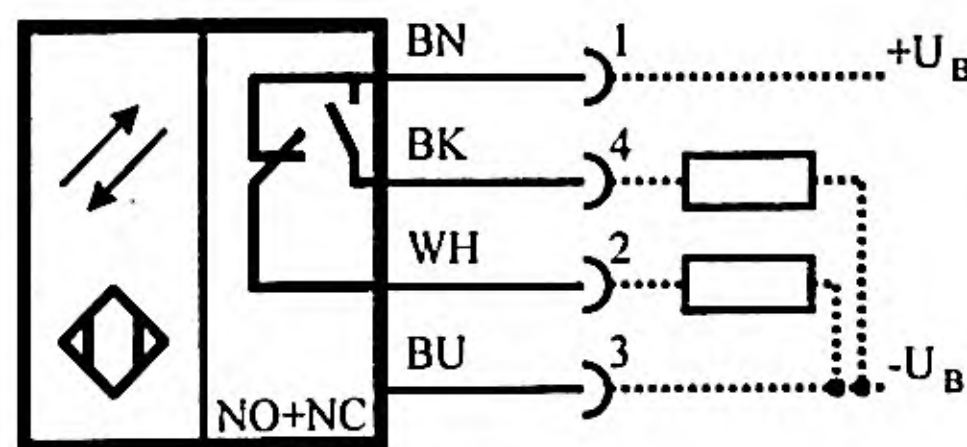


Figura 7.73 Ejemplo de símbolo de una fotocélula de cuatro hilos (cortesía de Siemens).

Las fotocélulas pueden tener también salida normalmente cerrada o abierta y estar implementadas con transistor NPN o PNP en la salida, al igual que los sensores inductivos y capacitivos, pero en ellas esta característica se combina con la activación cuando el receptor recibe luz (activación con luz) o cuando no la recibe (activación con oscuridad). Las tablas 7.11 y 7.12 indican el valor de la salida para las distintas combinaciones de los diferentes parámetros (tipo de transistor de salida, tipo de salida, tipo de activación de la salida y presencia o ausencia de objeto) de los sensores de barrera de luz y reflexión sobre espejo (Tabla 7.11) y de reflexión sobre objeto (Tabla 7.12). El circuito de las salidas NPN y PNP se representa en la figura 7.62a y 7.63a respectivamente.

Tipo de transistor de salida	Tipo de salida	Activación de la salida	Presencia de objeto	Ausencia de objeto
PNP	NO	Con luz	$S = 0V$ $I = 0$	$S = +V$ $I > 0$
		En oscuridad	$S = +V$ $I > 0$	$S = 0V$ $I = 0$
	NC	Con luz	$S = +V$ $I > 0$	$S = 0V$ $I = 0$
		En oscuridad	$S = 0V$ $I = 0$	$S = +V$ $I > 0$
NPN	NO	Con luz	$S = +V$ $I = 0$	$S = 0V$ $I > 0$
		En oscuridad	$S = 0V$ $I > 0$	$S = +V$ $I = 0$
	NC	Con luz	$S = 0V$ $I > 0$	$S = +V$ $I = 0$
		En oscuridad	$S = +V$ $I = 0$	$S = 0V$ $I > 0$

Tabla 7.11 Descripción del comportamiento de los sensores de barrera de luz y de reflexión sobre espejo que tienen salida de tres hilos.

Tipo de transistor de salida	Tipo de salida	Activación de la salida	Presencia de objeto	Ausencia de objeto
PNP	NO	Con luz	$S = +V$ $I > 0$	$S = 0V$ $I = 0$
		En oscuridad	$S = 0V$ $I = 0$	$S = +V$ $I > 0$
	NC	Con luz	$S = 0V$ $I = 0$	$S = +V$ $I > 0$
		En oscuridad	$S = +V$ $I > 0$	$S = 0V$ $I = 0$
NPN	NO	Con luz	$S = 0V$ $I > 0$	$S = +V$ $I = 0$
		En oscuridad	$S = +V$ $I = 0$	$S = 0V$ $I > 0$
	NC	Con luz	$S = +V$ $I = 0$	$S = 0V$ $I > 0$
		En oscuridad	$S = 0V$ $I > 0$	$S = +V$ $I = 0$

Tabla 7.12 Descripción del comportamiento de los sensores de reflexión sobre objeto que tienen salida de tres hilos.

7.7.2.3.3.2 Características constructivas de las fotocélulas

Las fotocélulas constan de un elemento emisor de luz y otro receptor de luz, que pueden estar montados sobre la misma carcasa o sobre carcasas diferentes. Además de los dispositivos y circuitos electrónicos, las fotocélulas poseen, en general, lentes, filtros de luz, etc. La figura 7.74 muestra, utilizando como ejemplo una fotocélula de barrera de luz, los elementos que componen una fotocélula, que se describen a continuación.

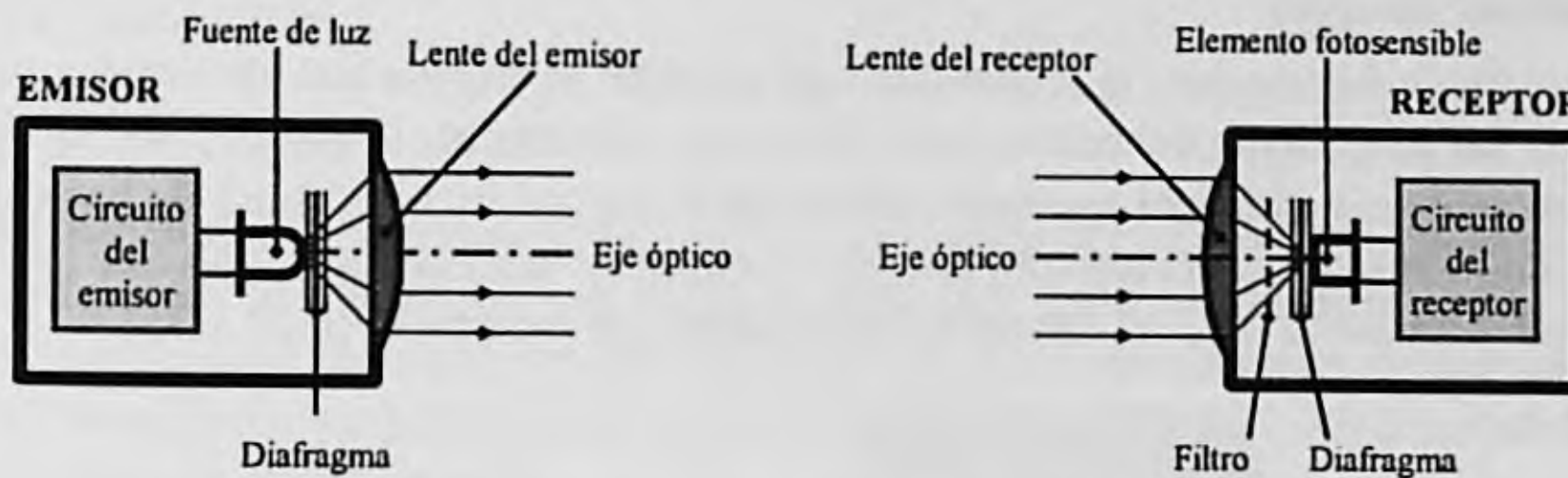


Figura 7.74 Estructura general de una fotocélula de barrera de luz.

Elementos del emisor

- **Circuito del emisor**

Muchas fotocélulas de las existentes en el mercado no trabajan con luz continua sino que lo hacen con luz modulada. En los sistemas con luz modulada el circuito del emisor envía a la fuente luminosa impulsos de corriente para que salgan de ella impulsos de luz. Se puede, de esta forma, aplicar una mayor potencia instantánea a la fuente de luz y se logra un mayor alcance. Además, el receptor se puede hacer más sensible porque las señales alternas son más fáciles de amplificar que las continuas. Por otra parte, la utilización de luz modulada facilita la colocación de fotocélulas muy próximas sin que interfieran entre sí.

- **Fuente de luz**

Las fotocélulas pueden trabajar con luz visible o con luz infrarroja. Muchas lo hacen con luz del infrarrojo cercano [normalmente de un diodo de infrarrojos (*IR-LED*) de arseniuro de galio (*GaAs*)], porque en el rango de longitudes de onda correspondiente es en el que se obtiene mayor potencia y rendimiento. Los sensores de marcas (véase el apartado 7.7.2.3.3.6), las fotocélulas de luz polarizada (véase el apartado 7.7.2.3.3.5) y algunas fotocélulas de tipo barrera (véase el apartado 7.7.2.3.3.4) suelen utilizar luz visible procedente de un diodo luminiscente conocido por el acrónimo *LED* (*Light Emitting Diode*), normalmente de color rojo o verde. La utilización de luz visible favorece el alineado de la fotocélula. En las fotocélulas ya no se utilizan las lámparas de incandescencia como fuentes luminosas debido a su corta vida, su bajo rendimiento, y la dificultad que supone para el receptor distinguir su luz de la del ambiente.

- **Lente del emisor**

Debido a que la fuente de luz emite una radiación que se propaga en todas las direcciones, se coloca una lente y la fuente luminosa se sitúa en el foco de la misma. Se logra así, que los rayos luminosos que la atraviesan salgan del dispositivo emisor paralelos al eje óptico de la lente que constituye el eje óptico del emisor.

- **Diafragma del emisor**

Con el fin de incrementar el carácter puntual de la fuente luminosa y conseguir así una mayor focalización del haz, se suele colocar un diafragma entre la fuente de luz y la lente del emisor.

Elementos del receptor

- **Lente del receptor**

Con el fin de incrementar la sensibilidad del receptor se emplea una lente que concentra el haz de luz procedente del emisor en el elemento fotosensible (Figura 7.74). El eje óptico de esta lente constituye el eje óptico del receptor que, en las fotocélulas de barrera de luz, debe coincidir con el del emisor para que los rayos de luz procedentes de este último incidan correctamente sobre el elemento fotosensible del receptor (Figura 7.75).

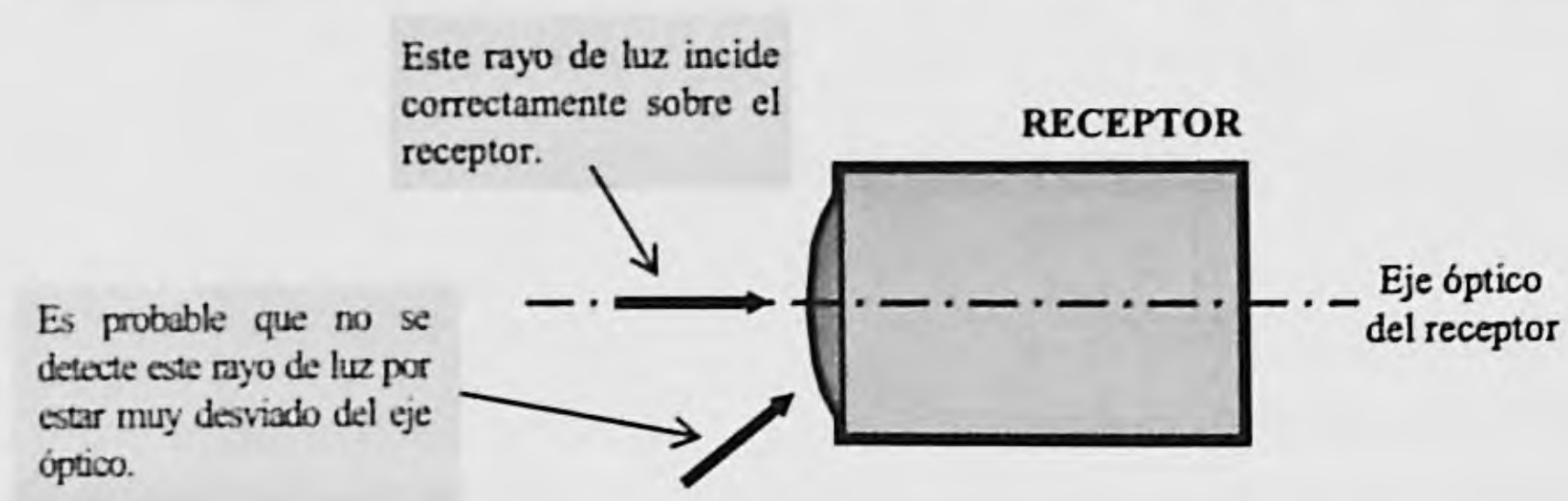


Figura 7.75 Efecto de la desviación del ángulo de incidencia de la luz con relación al eje óptico del receptor.

- **Elemento fotosensible**

Cuando la luz incide sobre él se modifica el valor de alguno de sus parámetros característicos. Los dispositivos electrónicos fotosensibles más utilizados son el fotodiodo (*Photodiode*), el fototransistor (*Phot transistor*), el fotodarlington (*Photodarlington*), el fototiristor controlado de silicio conocido como LASCR (acrónimo de *Light Activated Silicon Controlled Rectifier*). En el pasado se utilizó también la fotorresistencia conocida como LDR (acrónimo de *Light Dependent Resistor*).

- **Circuito del receptor**

El circuito del receptor amplifica la señal eléctrica generada por el elemento fotosensible y establece la señal de salida a partir de ella. Las fotocélulas que tienen capacidad para evitar la interferencia mutua trabajan con luz modulada (pulsante) y en ellas el circuito receptor activa la salida, si la frecuencia de los impulsos que recibe coincide con la frecuencia de los impulsos que envía el emisor.

- **Diafragma del receptor**

El diafragma limita el ángulo de recepción para evitar, en lo posible, la luz que no procede del emisor. Esta medida no evita el problema que supone la existencia de una fuente de luz ajena al sensor y situada cerca del emisor. Para evitarlo es necesario utilizar filtros ópticos.

- **Filtro óptico**

Evita el paso de las componentes del espectro luminoso procedentes del entorno, susceptibles de causar interferencias, y sólo deja pasar la luz del emisor.

7.7.2.3.3 Características técnicas de las fotocélulas

Las principales características técnicas comunes a los distintos tipos de fotocélulas son las indicadas a continuación.

Histéresis

Una fotocélula tiene histéresis (*Hysteresis*) cuando la posición del objeto en la que se activa no coincide con la posición en la que se desactiva. La distancia entre ambas posiciones se denomina también distancia diferencial. Se puede presentar histéresis cuando el objeto se mueve sobre el eje óptico (movimiento axial) de la fotocélula (Figura 7.76) o cuando lo hace en dirección transversal al eje óptico.

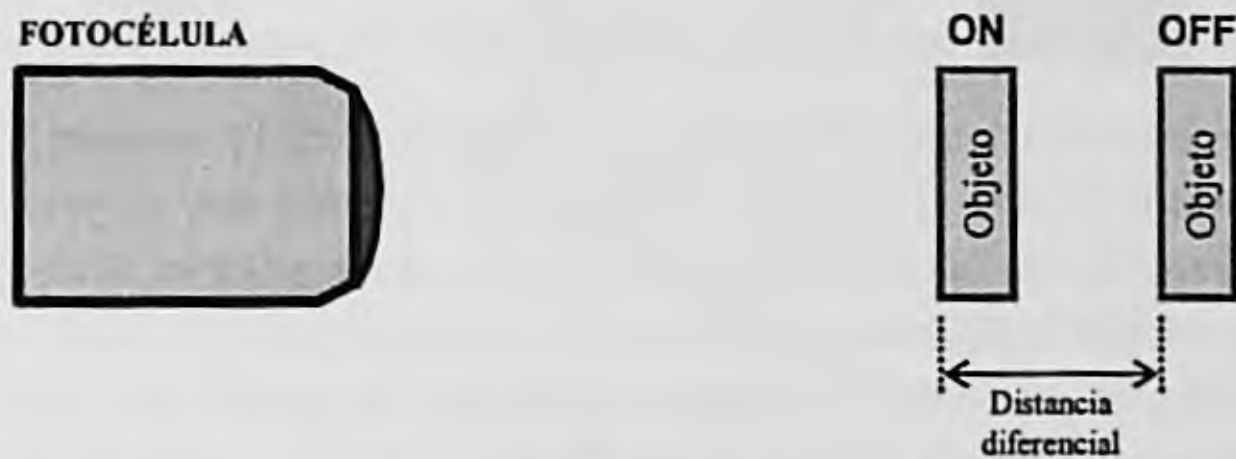


Figura 7.76 Histéresis de una fotocélula cuando el objeto se mueve en dirección axial.

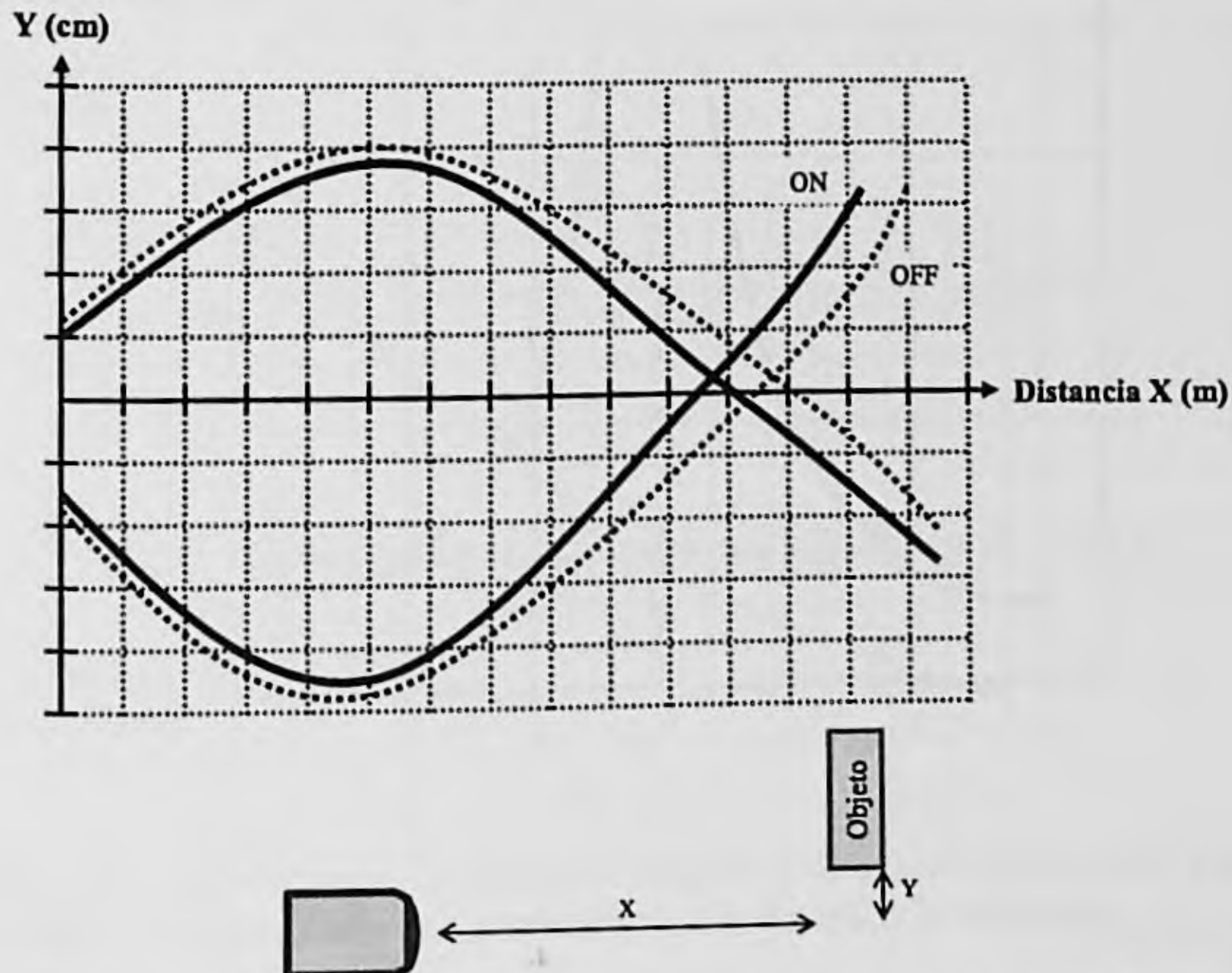


Figura 7.77 Histéresis de una fotocélula cuando el objeto se mueve transversalmente.

En el primero de los casos se suele especificar la histéresis en valor absoluto o como un porcentaje de la distancia máxima de detección. En el segundo de los casos se suele adjuntar una gráfica como la que se muestra en la figura 7.77, en la que, mediante trazo continuo se representa el punto en el que se activa el sensor a medida que se acerca el objeto, y la de trazo discontinuo, el punto en el que se desactiva el sensor a medida que se aleja el objeto.

Desalineación angular

La desalineación angular o ángulo direccional es una característica típica de las fotocélulas de tipo barrera o de reflexión sobre espejo (Figura 7.78a). Indica el valor del ángulo α que puede estar girado el eje óptico del emisor, del receptor o del espejo con respecto al alineamiento perfecto. Tal como se indica en la figura 7.78b, la distancia de detección disminuye con la desalineación angular.

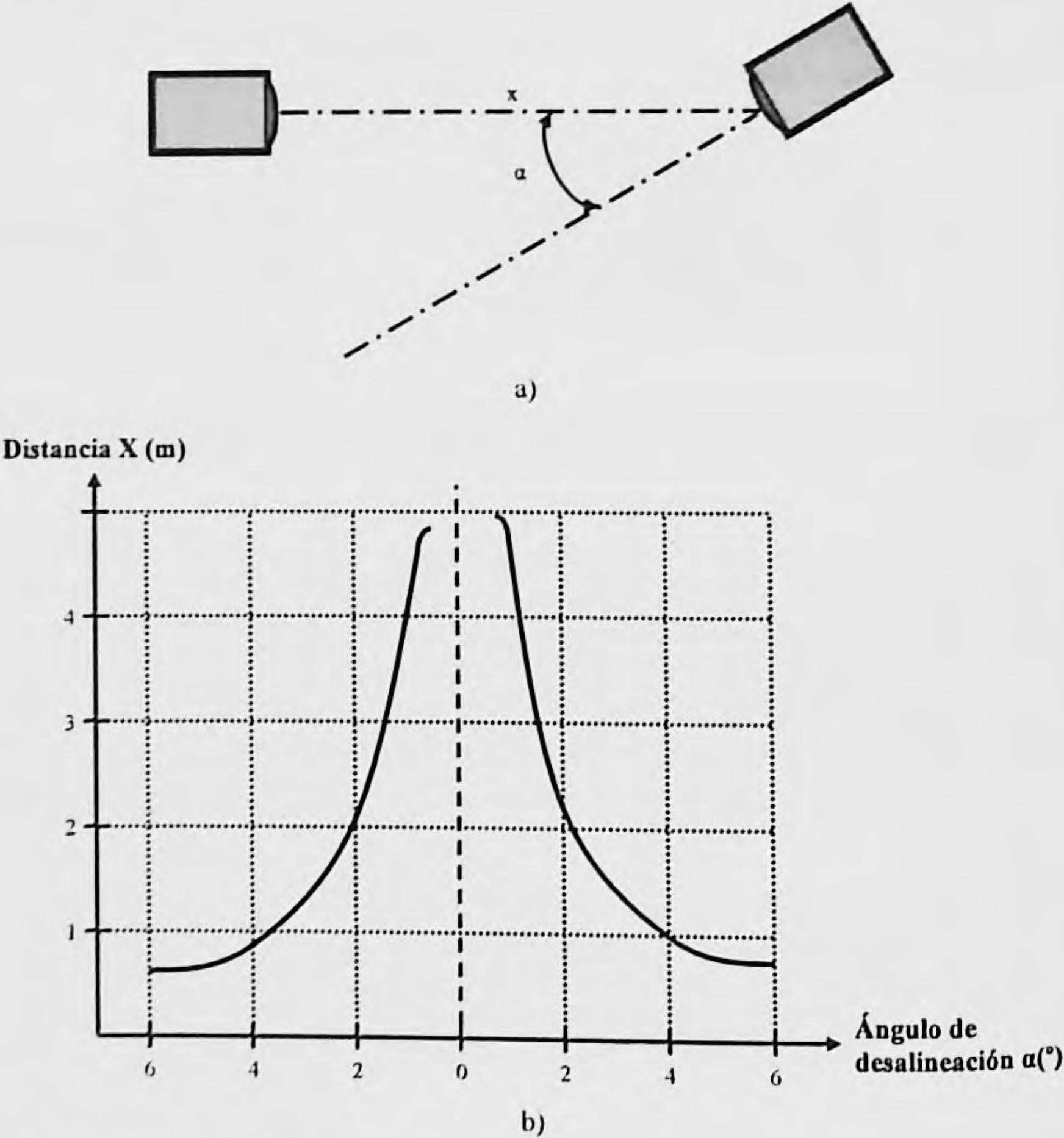


Figura 7.78 *Ángulo direccional: a) Desalineación angular entre el emisor y el receptor; b) Relación entre el máximo ángulo de desalineación angular admisible y la distancia.*

Exceso de ganancia

El circuito del receptor de una fotocélula activa o desactiva la salida según que la señal que recibe el elemento fotosensible del receptor esté por encima o por debajo de un determinado nivel umbral. Cuando las condiciones de trabajo de la fotocélula son ideales (lentes y ambiente limpios, ejes perfectamente alineados, objeto de detección estándar, etc.), dicha señal supera el nivel umbral con un determinado margen, que es el que permite que la fotocélula siga funcionando correctamente cuando las condiciones no son ideales. El exceso de ganancia, ganancia excedente o razón de la ganancia (*Excess gain*) [AENO 05] [BART 03] se define como el nivel de la señal presente en el receptor en condiciones ideales. El umbral da una idea de la capacidad que tiene la fotocélula para superar pérdidas de señal y por ello, por ejemplo, cuanto más sucio se prevé que sea un determinado ambiente, más exceso de ganancia se le debe exigir a la fotocélula. Como orientación se pueden indicar los valores que debe tener el exceso de ganancia en función de la suciedad del ambiente:

- Ambiente ligeramente polvoriento (>5).
- Ambiente polvoriento, entorno contaminado, ligera niebla (>10).
- Ambiente extremadamente contaminado, niebla y vapores (>50).

El exceso de ganancia varía con la distancia de detección en una misma fotocélula, y por ello se suele expresar en una gráfica semilogarítmica como la que se muestra en la figura 7.79. Dicha figura representa el valor de la relación de la ganancia de una fotocélula que sólo funciona en condiciones ideales (relación igual a 1) cuando la distancia al objeto es aproximadamente de 5 metros. El valor del exceso de ganancia es máximo cuando la distancia entre la fotocélula y el objeto es de 1 metro. De acuerdo con los valores del exceso de ganancia anteriormente indicados, esta fotocélula sólo detecta el objeto en un ambiente ligeramente polvoriento cuando está colocada a una distancia próxima a 1 metro.

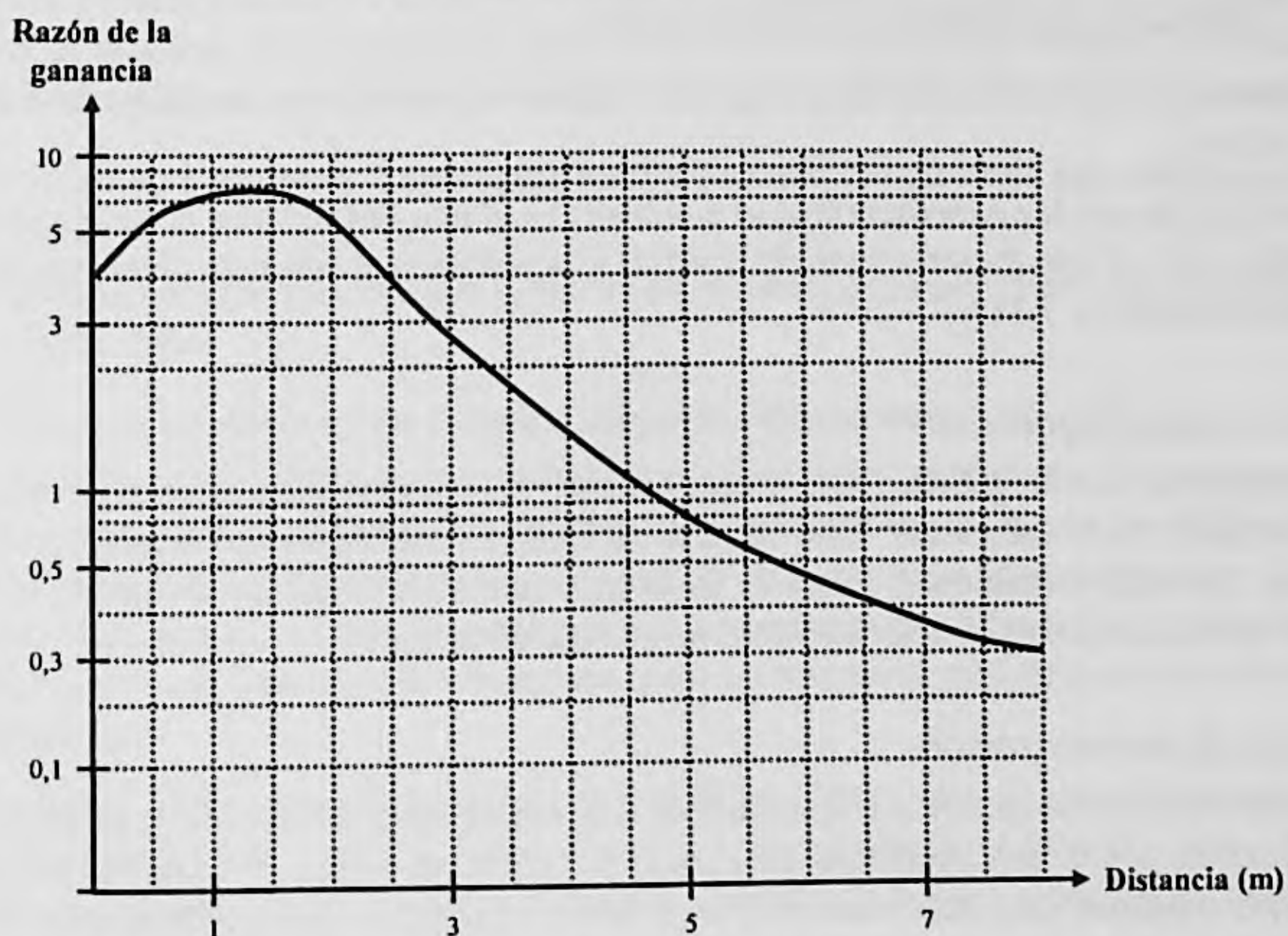


Figura 7.79 Relación entre el exceso de ganancia y la distancia máxima de detección.

Iluminación ambiental

Indica la iluminación ambiental máxima que puede recibir el receptor sin que se produzca un mal funcionamiento del sensor. Normalmente se proporcionan los valores para luz solar y para el tipo de lámpara más desfavorable, que suele ser la de incandescencia. Valores típicos de luz ambiental admisible son 10.000 luxes para iluminación solar y 3.000 luxes para incandescencia.

Activación de la salida

Las fotocélulas pueden tener dos modos de operación según su salida se active cuando el receptor recibe luz o cuando no la recibe (en oscuridad) tal como se indica gráficamente en la figura 7.80. La selección del modo de operación se puede realizar de diferentes maneras entre las que cabe citar las siguientes:

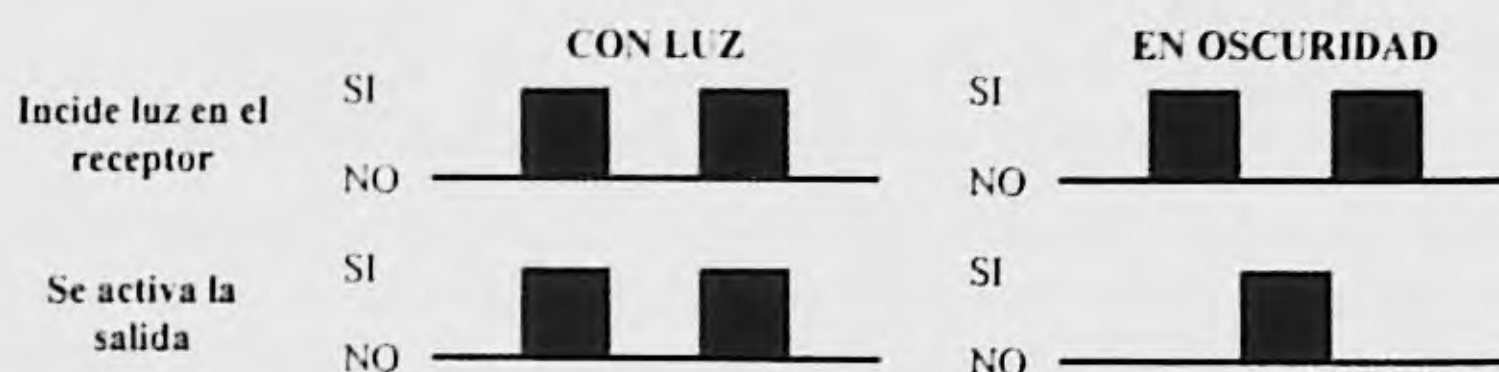


Figura 7.80 Modos de operación de una fotocélula según la salida se active con luz o en oscuridad.

- Mediante un conmutador.
- Cambiando la polaridad de la alimentación.
- Mediante un cuarto cable que se puede conectar al positivo o al negativo de la alimentación.

La activación con luz o en oscuridad se combina con el tipo de transistor de salida, el tipo de salida NO o NC, el tipo de activación de la salida y la presencia o ausencia de objeto tal como se indica en las tablas 7.11 y 7.12.

Entrada de autodiagnóstico

La entrada de autodiagnóstico o de prueba (cuando existe) sirve para comprobar el correcto funcionamiento de la fotocélula. Esta entrada verifica el funcionamiento cuando el receptor recibe luz. En estas condiciones, el haz de luz se interrumpe al activar la entrada de prueba y el estado de la salida cambia si la fotocélula funciona correctamente.

Funciones de temporización

Algunas fotocélulas incorporan funciones de temporización de la variable de salida. El valor de la temporización es independiente del tiempo de respuesta, se puede ajustar externamente, por ejemplo mediante un potenciómetro, y en general, varía entre algunas décimas de segundo y unos pocos segundos.

Indicadores luminosos

Algunas fotocélulas disponen de indicadores luminosos. Normalmente son diodos luminiscentes (*LED*), que indican el estado de la fotocélula al operador humano. Existen varios tipos de indicadores:

- **Indicador de detección**

Se ilumina cuando el receptor detecta luz.

- **Indicador de operación**

Se ilumina cuando se activa o se desactiva la salida según la fotocélula se active cuando recibe luz o en oscuridad. Se diferencia del de detección cuando la fotocélula tiene funciones de temporización.

- **Indicador de estabilidad**

Se ilumina cuando el nivel de luz que recibe el receptor está por encima de un umbral que se considera nivel de luz estable y también cuando está por debajo de un nivel de oscuridad estable.

7.7.2.3.3.4 Sensores optoelectrónicos de barrera de luz

Introducción

Este tipo de sensores está constituido por un emisor y un receptor dispuestos físicamente de tal forma que, en ausencia de un objeto, los rayos de luz generados por el emisor alcanzan el receptor y constituyen una barrera de luz tal como se indica en la figura 7.81a. Reciben por ello el nombre de sensores de barrera de luz (*Through-beam optical sensors*) o fotocélulas de barrera de luz (*Through-beam photocells*) y detectan el objeto cuando se sitúa entre el emisor y el receptor, de tal forma que interrumpe el haz de luz (Figura 7.81b). Los sensores de barrera de luz poseen una elevada sensibilidad y se implementan de dos formas principales:

- **Mediante un emisor y un receptor unidos por una carcasa que los soporta mecánicamente.**

Constituyen un dispositivo con forma de herradura (*Grooved head*) que se representa en la figura 7.82.

- **Mediante un emisor y un receptor separados físicamente.**

La instalación es más difícil en este caso por tener que colocar dos elementos separados cuyos ejes ópticos deben estar alineados, pero presenta la ventaja de que la luz sólo tiene que atravesar el espacio de trabajo una vez, lo que permite el funcionamiento a distancias elevadas. Son apropiadas para condiciones ambientales poco favorables, como suciedad, humedad, o utilización a la intemperie, pero no son aptas para detectar materiales transparentes.

La figura 7.83 muestra la fotografía de uno de los elementos que constituyen un sensor de barrera de luz PXO200M18. Tanto el emisor como el receptor tienen la misma apariencia externa. En sucesivos apartados se analiza el campo de trabajo y las precauciones de instalación de este tipo de fotocélulas.

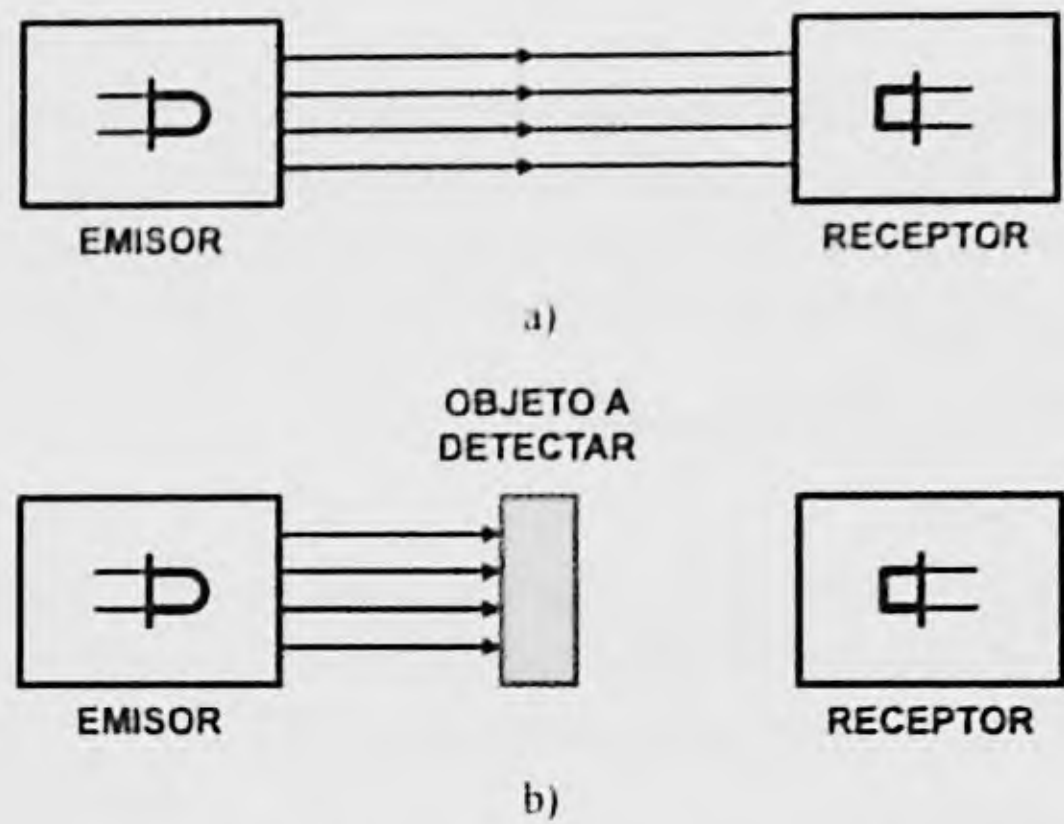


Figura 7.81 Funcionamiento de un sensor optoelectrónico de barrera.

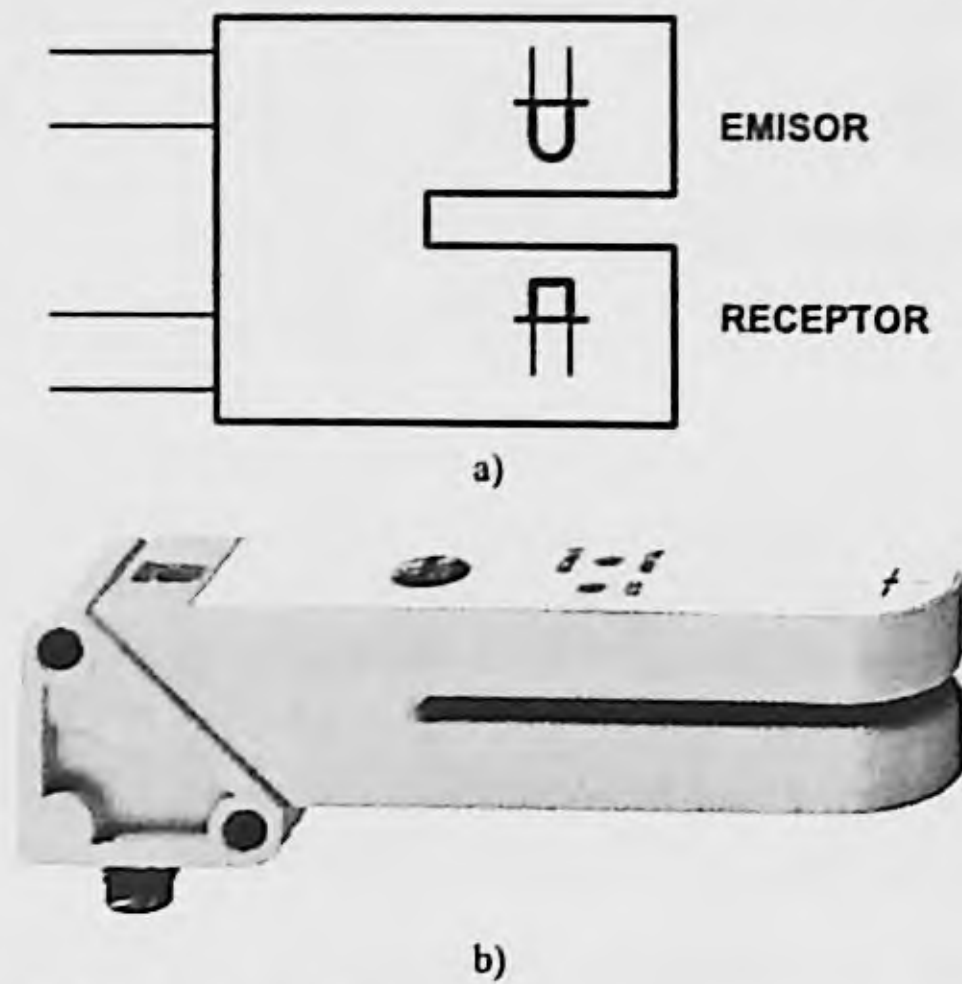


Figura 7.82 Barrera de luz en herradura: a) Diagrama de bloques; b) Fotografía (cortesía de Siemens).



Figura 7.83 Fococélula de tipo barrera de luz modelo PXO200 M18 (cortesía de Siemens).

- **Campo de trabajo**

El campo de trabajo o rango de operación es la parte del espacio en la que se puede detectar un objeto. Es el cilindro que une al emisor con el receptor, y su diámetro coincide prácticamente con el de las lentes utilizadas, tal como se indica en la figura 7.84. En función del exceso de ganancia que tenga la fotocélula el objeto se puede detectar antes de que ocupe toda la anchura del campo de trabajo. Si el exceso de ganancia es muy bajo, próximo a la unidad, la detección del objeto tiene lugar cuando ocupa una pequeña porción del campo de trabajo. Por el contrario cuando el exceso de ganancia es muy elevado la detección solo se produce cuando el objeto ocupa la mayor parte de la anchura del campo de trabajo. El haz del emisor es ligeramente divergente lo que hace que lo sea también el campo de emisión. De la misma forma el receptor posee también un campo de recepción. El efecto conjunto de ambos campos hace que el sensor funcione correctamente aunque los ejes ópticos del emisor y del receptor no estén perfectamente alineados.



Figura 7.84 Campo de trabajo de la fotocélula de barrera de luz.

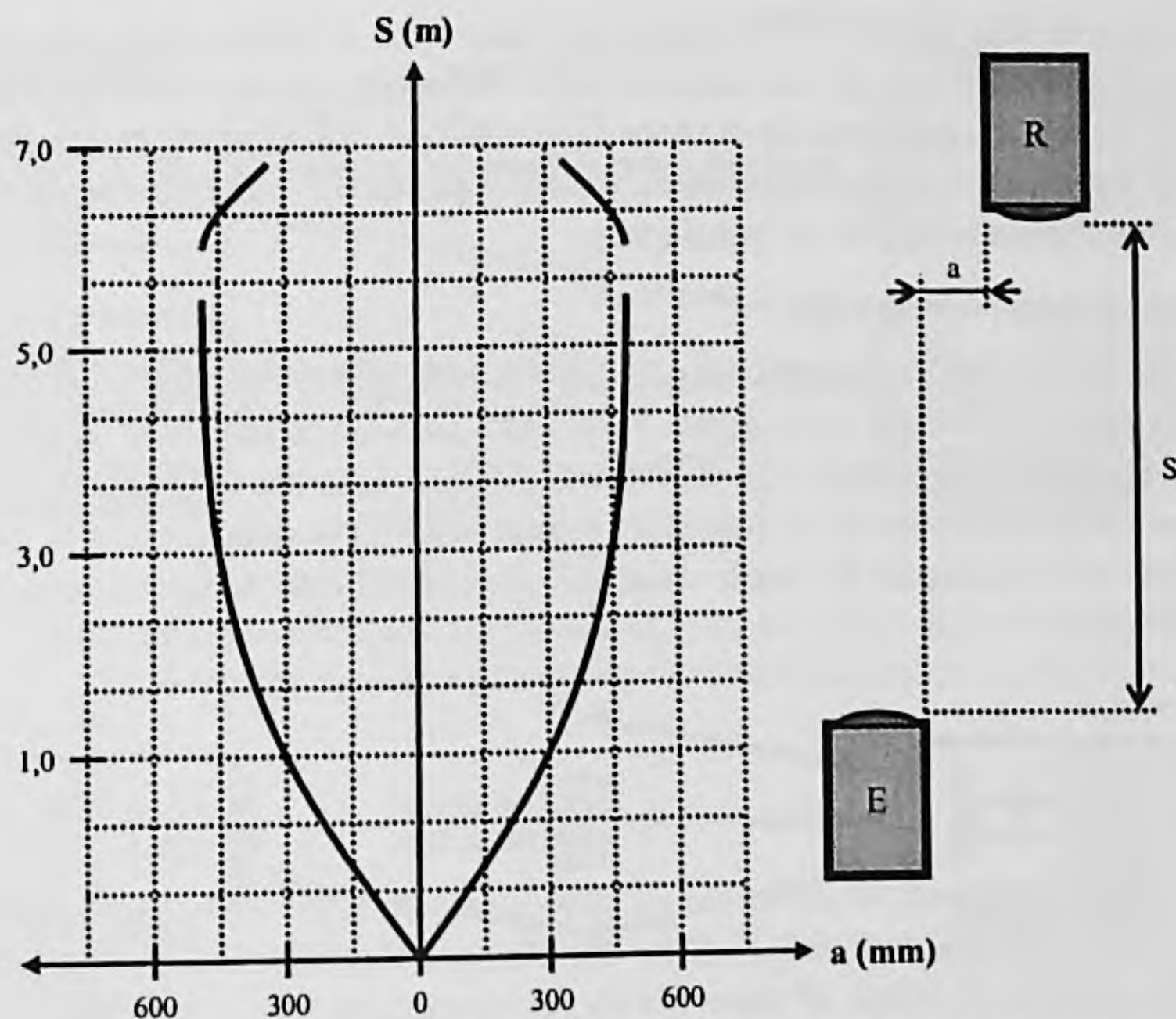


Figura 7.85 Campo de trabajo de la fotocélula PXO200M 18 de la familia Bero en configuración barrera de luz (cortesía de Siemens).

Los fabricantes suelen proporcionar información gráfica del campo de trabajo. La figura 7.85 muestra gráficamente el diagrama del campo de trabajo de la fotocélula PXO200M18 en configuración de barrera de luz. Dicha figura indica la relación que existe entre la distancia de detección S expresada en metros y la desalineación a expresada en milímetros y establece de esta forma la región del espacio en la que se puede detectar un objeto.

Para facilitar la detección de objetos que solo cubren una pequeña parte del campo de trabajo se utilizan diafragmas externos (que no se deben confundir con los diafragmas que incorporan internamente tanto el emisor como el receptor) (Figura 7.86). Dichos diafragmas pueden consistir en una ranura o un orificio cilíndrico y reducen la anchura del campo de trabajo aunque también disminuyen la distancia de detección.

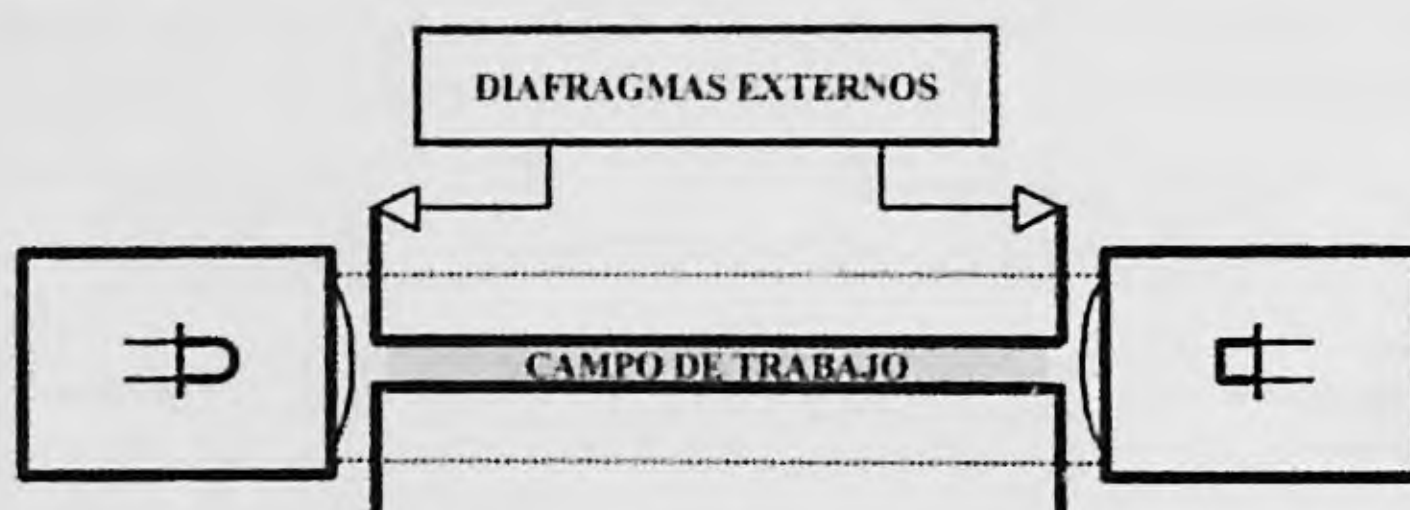


Figura 7.86 Reducción del campo de trabajo de una fotocélula de barrera de luz de emisor y receptor separados físicamente.

Normas de instalación

Para que una fotocélula de barrera de luz, que tiene el emisor y el receptor instalados en puntos distintos con sus correspondientes cableados funcione correctamente, hay que tener en cuenta un conjunto de normas de instalación. Además se debe tener especial cuidado en tener una buena alineación de los ejes ópticos. En particular se deben adoptar las siguientes normas o reglas de instalación:

- **Superficies reflectantes**

Si el haz de luz transcurre muy cercano a una superficie reflectante, puede suceder, tal como se muestra en la figura 7.87, que parte de él se refleje sobre ella debido al comportamiento ligeramente divergente del haz. Esta luz reflejada puede provocar un mal funcionamiento de la fotocélula y, para evitarlo, se debe separar la fotocélula tanto como sea posible de la citada superficie, o recubrir ésta de un material que impida la reflexión.

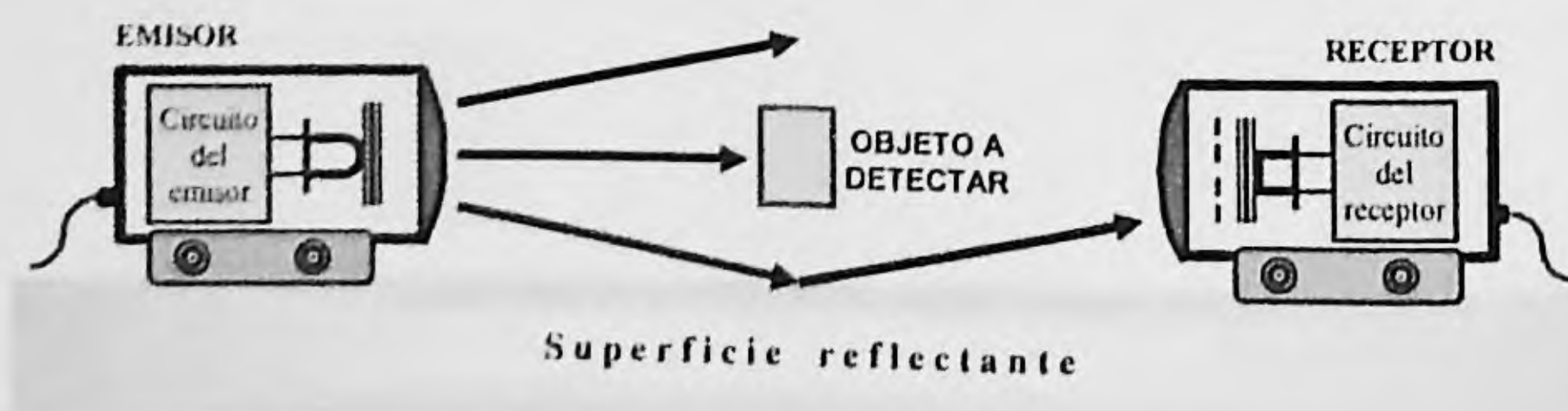


Figura 7.87 Malfuncionamiento de una fotocélula de barrera de luz debido a una superficie reflectante.

• Interferencias mutuas

Cuando dos o más fotocélulas se colocan adosadas o muy cercanas entre sí, puede ocurrir que un receptor reciba luz del emisor que no está asociado con él, tal como se indica en la figura 7.88. Esto puede provocar un mal funcionamiento de las fotocélulas, y para evitarlo se pueden separar las fotocélulas tanto como sea necesario, o intercambiar el emisor y el receptor de una de ellas (Figura 7.89). Otra solución es utilizar fotocélulas con luz modulada, tal como se indica anteriormente al describir el circuito del emisor.

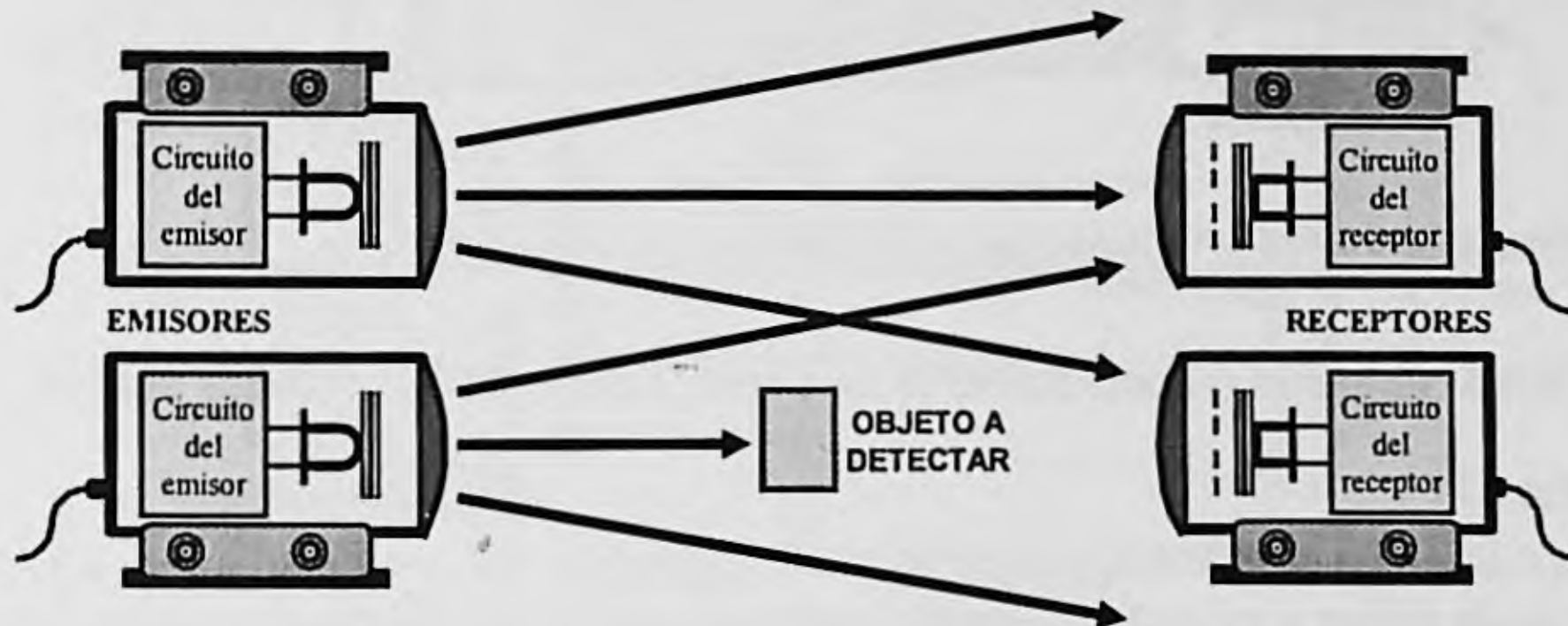


Figura 7.88 Malfuncionamiento de una fotocélula de barrera de luz debido a las interferencias mutuas.

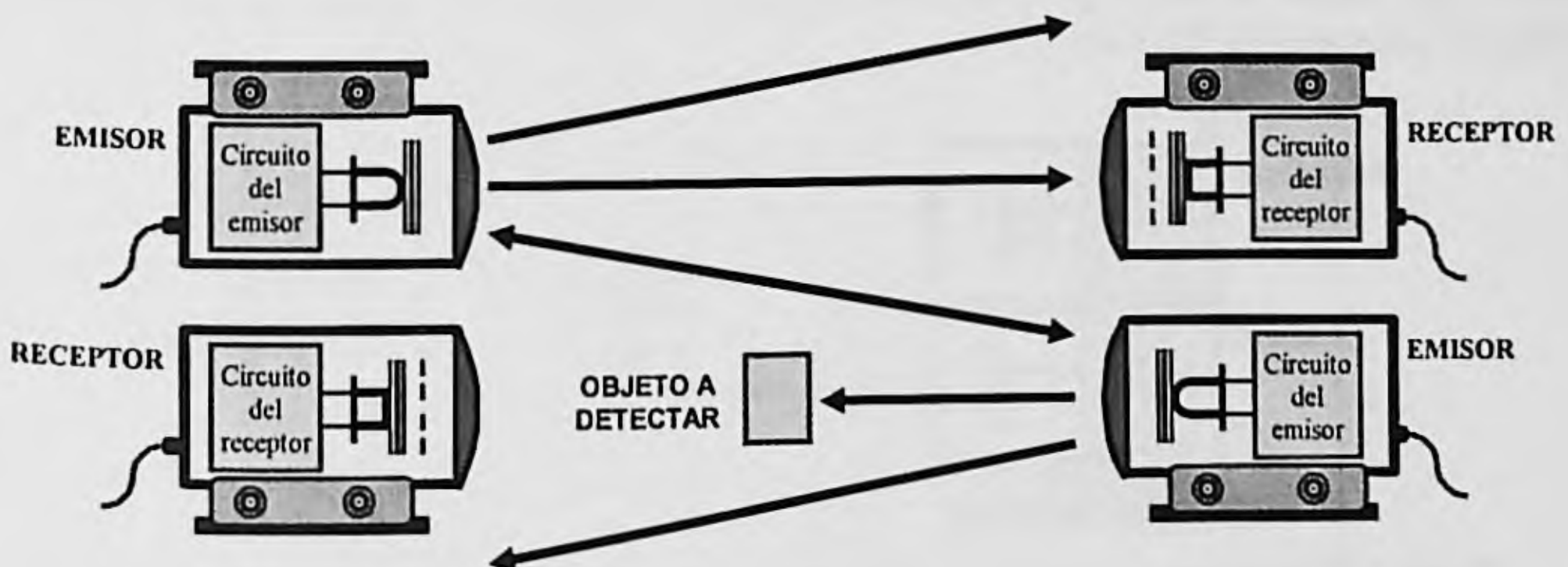


Figura 7.89 Eliminación del malfuncionamiento de una fotocélula de barrera de luz debido a las interferencias mutuas, mediante el intercambio del emisor y el receptor.

Ambiente sucio

Cuando el grado de suciedad de los ambientes en los que se tienen que colocar el emisor y el receptor son diferentes, es conveniente poner el receptor en el ambiente más sucio porque la suciedad tiene efectos más negativos sobre la lente del emisor (Figura 7.90a) que sobre la del receptor (Figura 7.90b), debido al carácter levemente divergente del haz de luz.

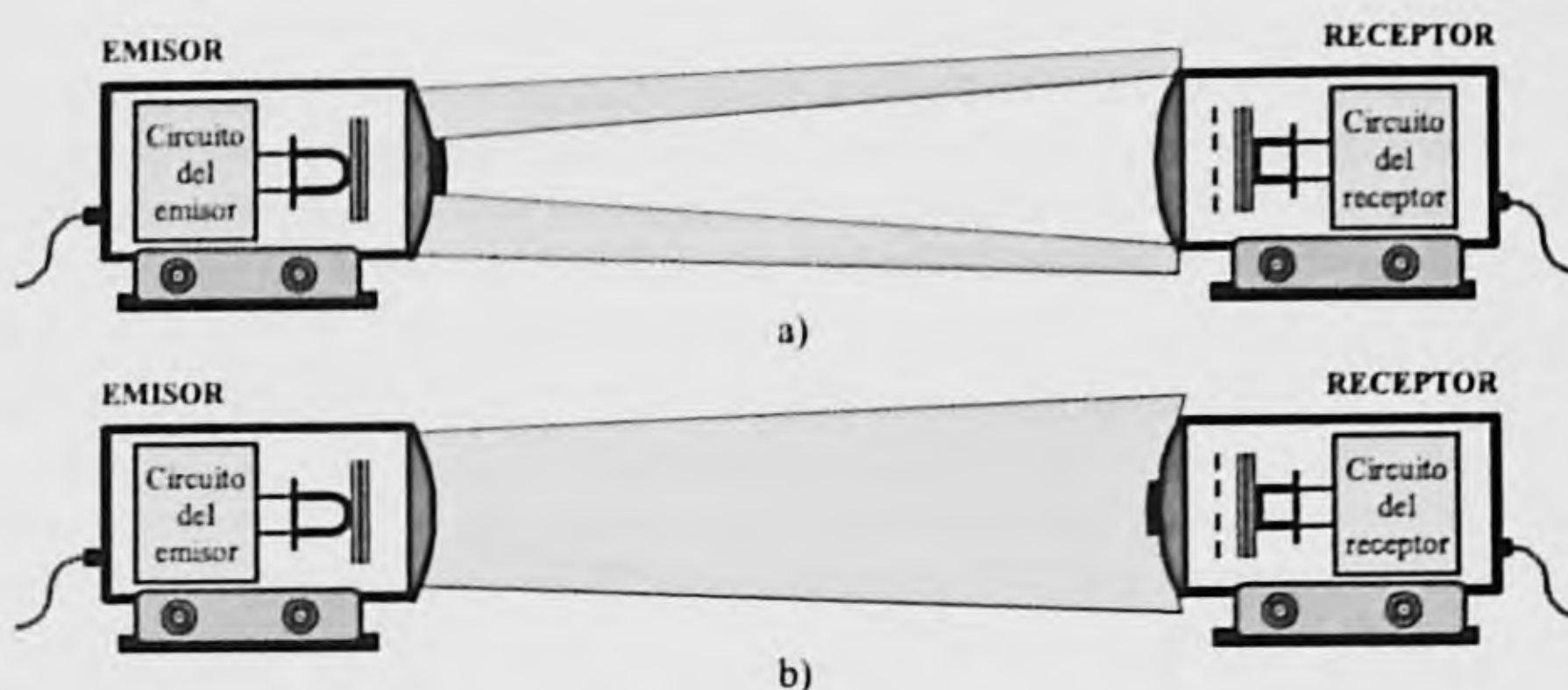


Figura 7.90 Efecto de la suciedad: a) En el emisor; b) En el receptor.

7.7.2.3.3.5 Sensores optoelectrónicos de proximidad de reflexión sobre espejo

Introducción

Las fotocélulas de reflexión sobre espejo (*Retroreflective optical sensors*) se componen de un emisor y un receptor montados en la misma carcasa, y un espejo que se coloca en frente de ellos. El espejo es fácil de instalar y, además, carece de cables. Si no existe un objeto que interrumpa su camino, el haz de luz sale del emisor, se refleja en el espejo y vuelve al receptor (Figura 7.91). Además de detectar objetos opacos, se pueden utilizar también para detectar objetos que tienen un cierto grado de transparencia. La figura 7.92 muestra la fotocélula de reflexión sobre espejo PXO400 K31.

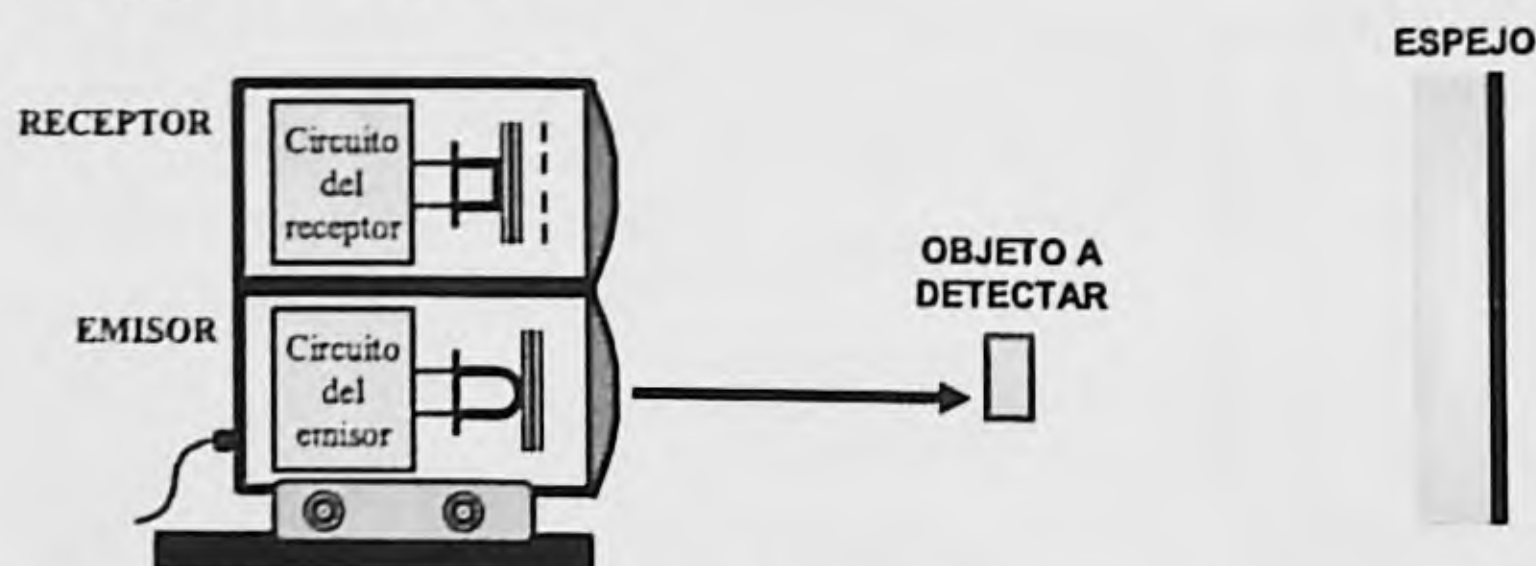


Figura 7.91 Funcionamiento de una fotocélula de reflexión sobre espejo.



Figura 7.92 Fotocélula de reflexión sobre espejo PXO400 K31 (cortesía de Siemens).

Campo de trabajo

El campo de trabajo es la zona en la que se puede colocar el espejo reflector. Tal como se indica en la figura 7.93, coincide con la zona espacial de solapamiento de los campos de emisión y recepción, del emisor y receptor respectivamente. En las proximidades de la fotocélula, existe una región, denominada zona muerta, en la que el fabricante de la fotocélula no garantiza el funcionamiento correcto de la fotocélula cuando se coloca el espejo en ella. El tamaño de la zona muerta depende del tipo de espejo. El espejo reflector debe colocarse centrado con el eje óptico de la fotocélula y perpendicular al mismo.

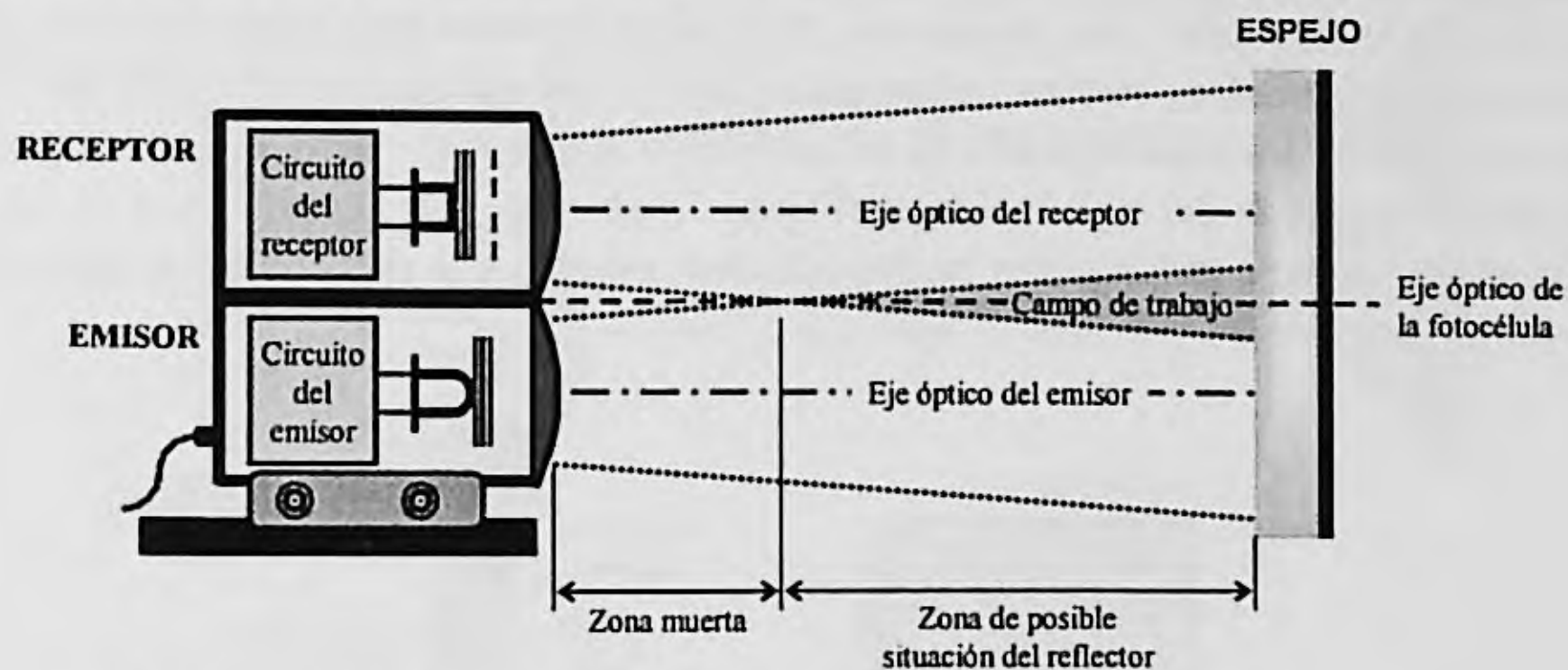


Figura 7.93 Campo de trabajo de una fotocélula de reflexión sobre espejo.

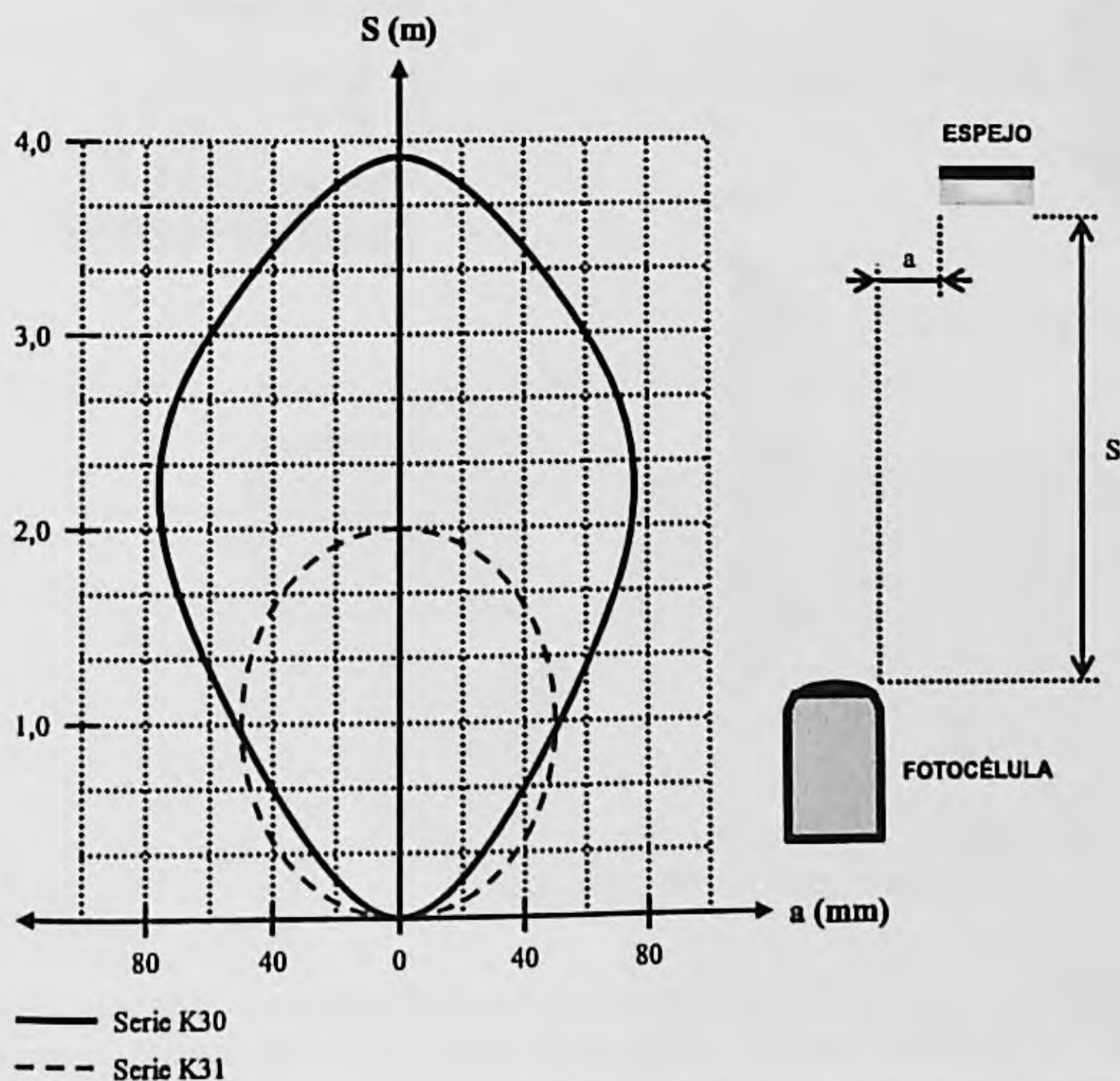


Figura 7.94 Campo de trabajo de las fotocélulas de reflexión sobre espejo PXO400 K31 y K30 (cortesía de Siemens).

La figura 7.94 muestra el campo de trabajo de las fotocélulas PXO400 K31 y K30 de reflexión sobre espejo. En esta gráfica se puede observar que no existe una zona muerta, gracias a la forma constructiva del sistema óptico de la fotocélula.

Espejos retrorreflectores

El espejo es muy importante para lograr el correcto funcionamiento de la fotocélula. Si es liso (espejo reflector), el ángulo de incidencia es igual al de reflexión (Figura 7.95), lo que hace que una pequeña variación de la inclinación del mismo, con respecto a la perpendicular al eje óptico, impida la detección. Para evitar este efecto se utiliza un espejo retrorreflector formado por un conjunto de triedros de caras reflectantes, que forman ángulos de 90° entre ellas (Figura 7.96a), igual que en los catadióptricos de un automóvil. Este espejo envía el haz reflejado en la misma dirección que la del rayo incidente (Figura 7.96b) y de esta forma el haz de luz incide sobre el sensor aunque exista cierta inclinación con respecto a la perpendicular con el eje óptico.

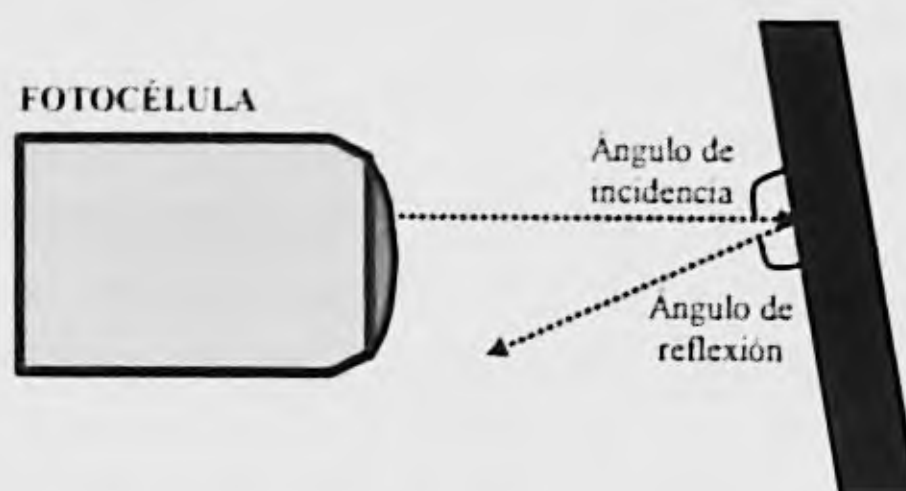


Figura 7.95 Efecto de un espejo retrorreflector liso.

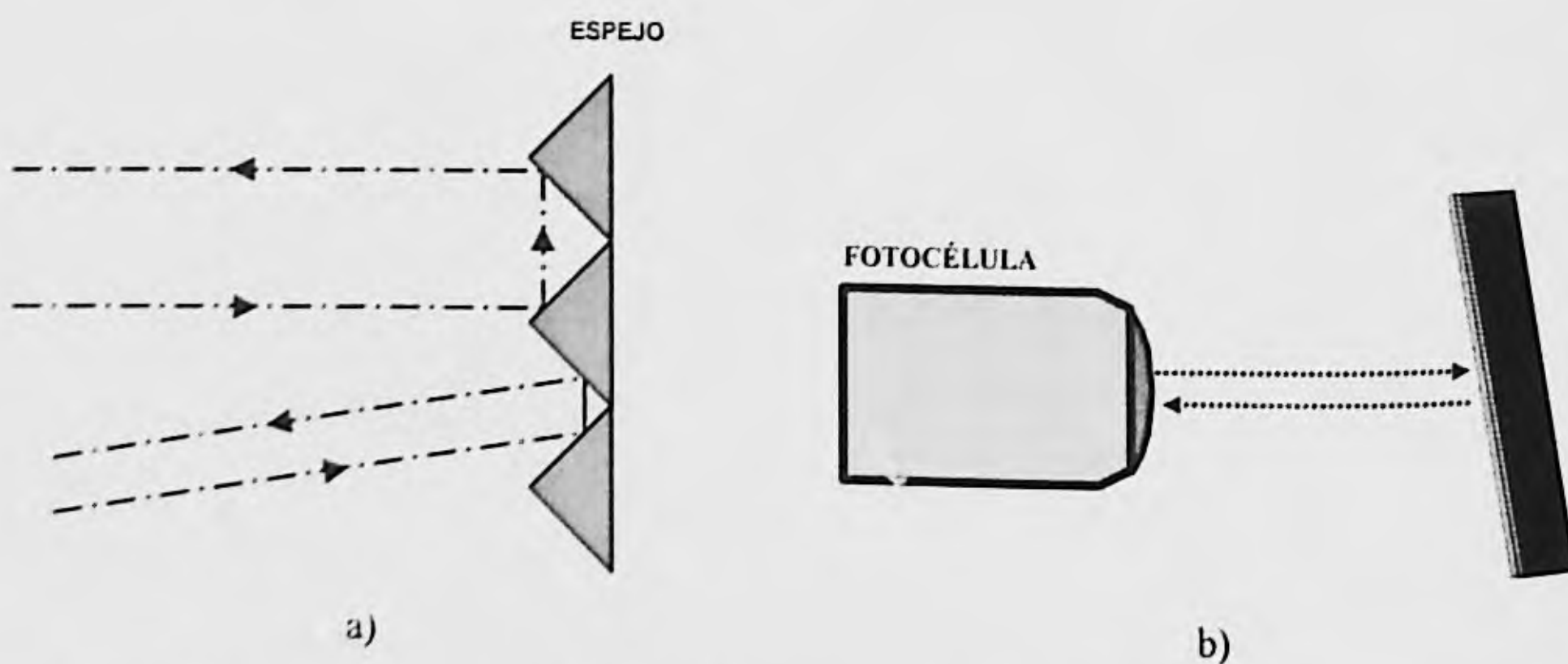


Figura 7.96 Espejo retrorreflector formado por un conjunto de triedros de caras reflectantes, que forman ángulos de 90° entre ellas.

Los fabricantes de fotocélulas suministran al usuario distintos tipos de espejos con diferentes tamaños de los catadióptricos. En la figura 7.97a se puede observar que, cuando el espejo está próximo al sensor, es necesario que los catadióptricos tengan un tamaño mínimo porque en caso contrario (Figura 7.97b) no se garantiza que los rayos reflejados lleguen al receptor. De

igual forma, a medida que aumenta la superficie del espejo, independientemente del tamaño del catadióptrico, aumenta la distancia de detección. La figura 7.98 muestra algunos tipos de espejos para fotocélulas, que se diferencian tanto en su forma física como en su tamaño.

En este tipo de fotocélulas se suele utilizar luz polarizada para detectar objetos brillantes. Tal como se indica a continuación, este tipo de espejos presenta la propiedad de abatir el plano de polarización.

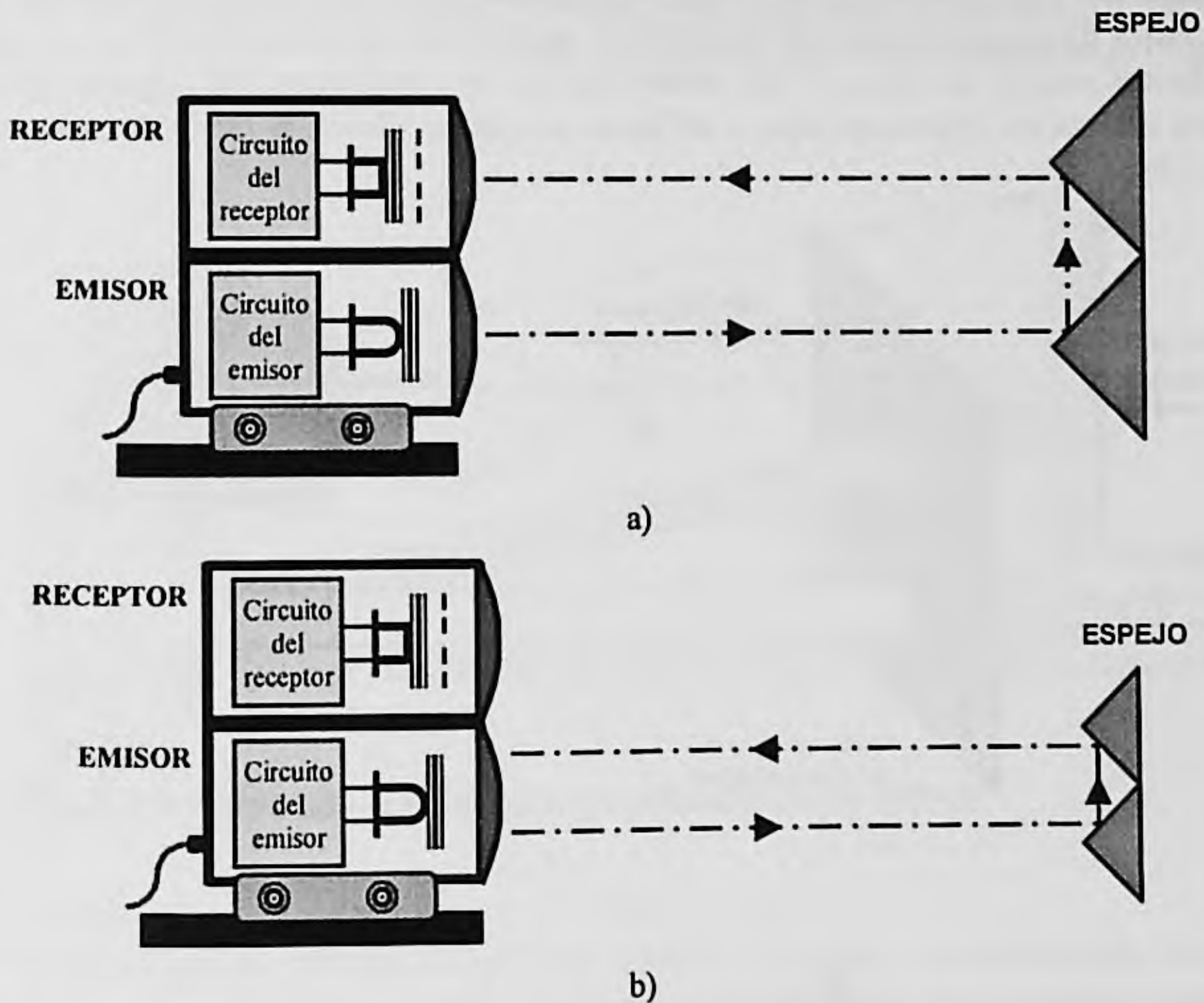


Figura 7.97 Distintos tamaños de los triédros de un espejo y su efecto.

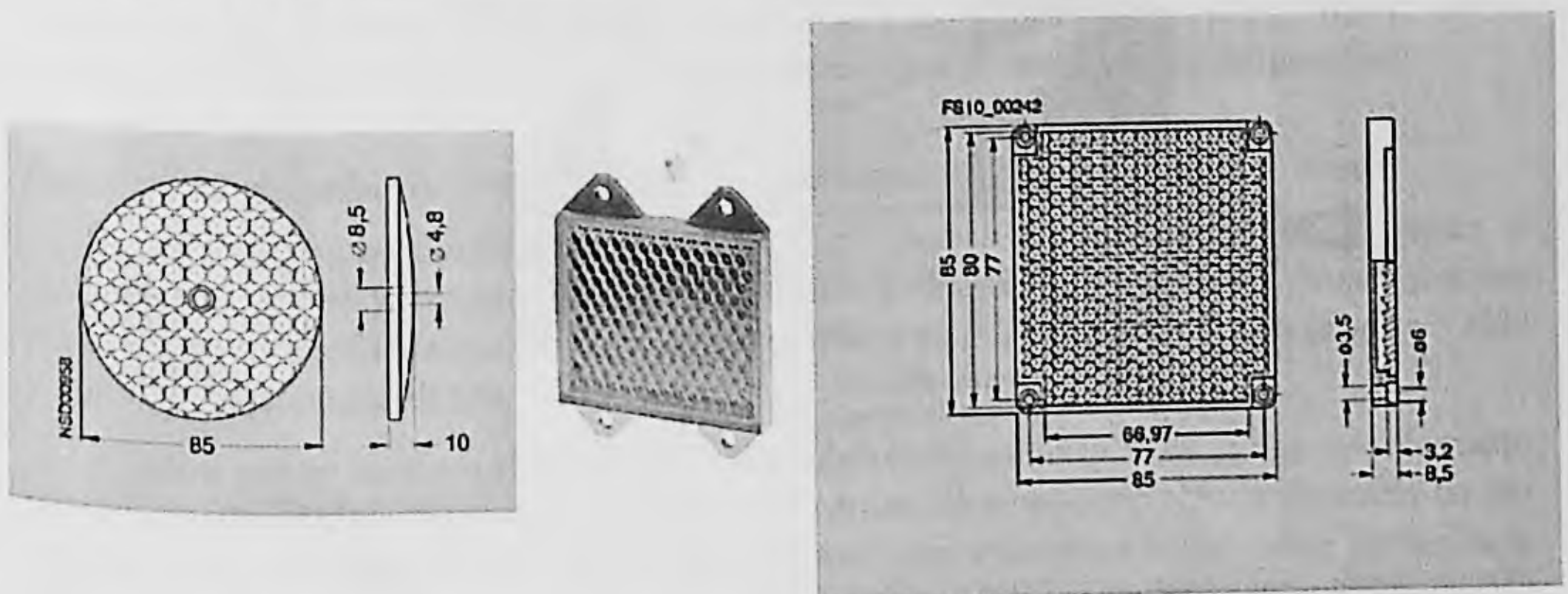
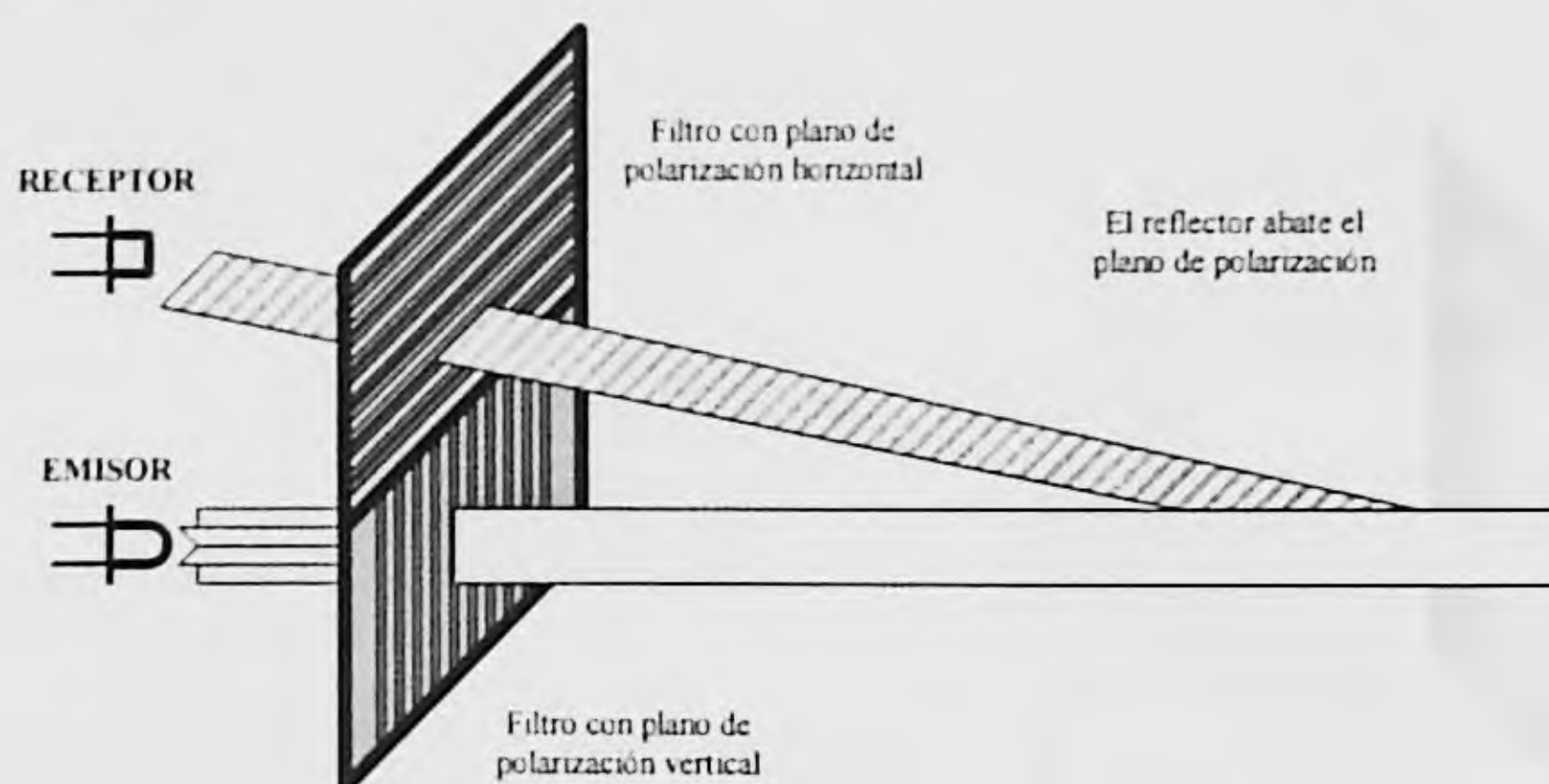


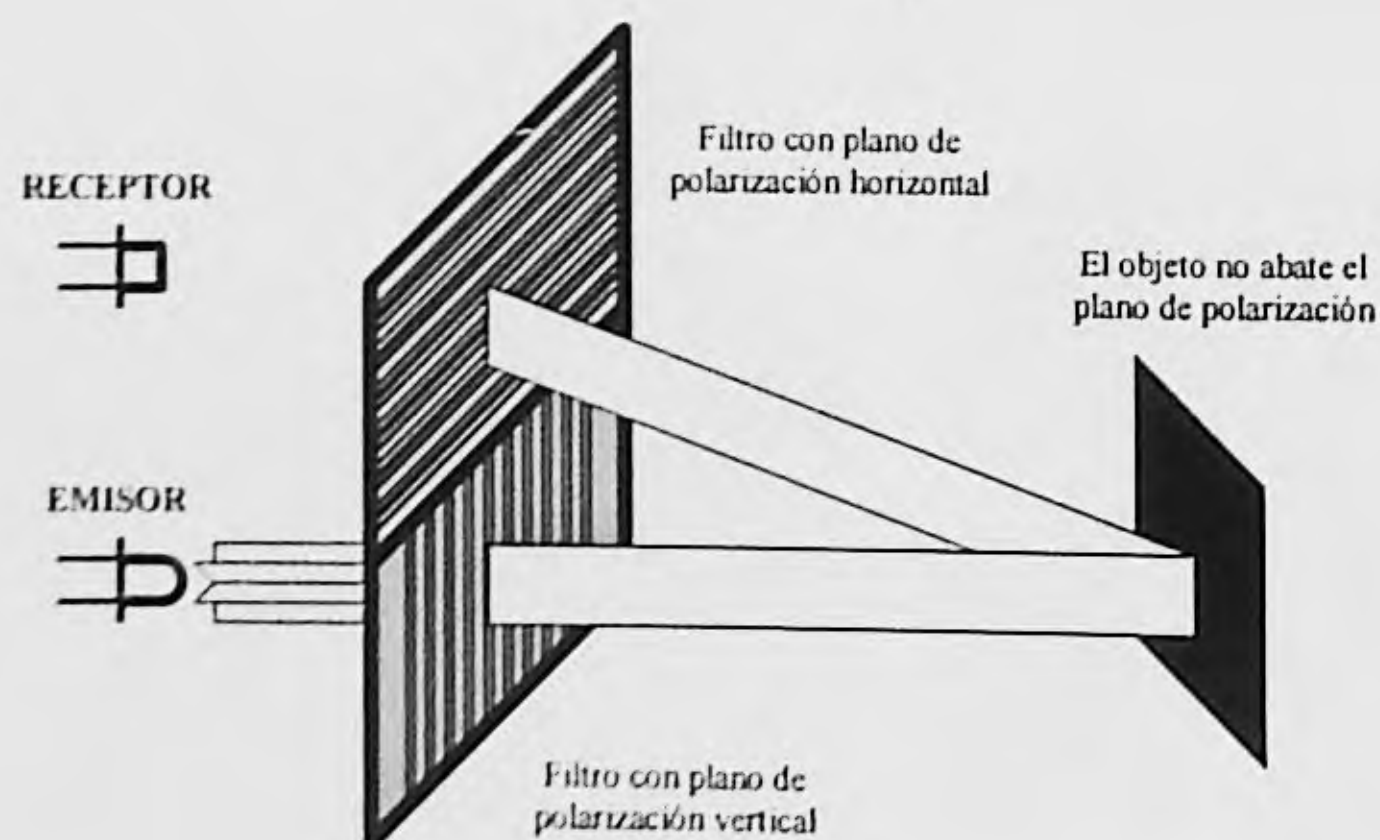
Figura 7.98 Distintos tipos de espejos (cortesía de Siemens).

Funcionamiento con luz polarizada

Los objetos muy brillantes, como por ejemplo las piezas de metal pulidas, pueden provocar errores de detección cuando se utiliza una fotocélula de reflexión sobre espejo porque, en ocasiones, no distinguen si la luz la refleja el espejo o el objeto. Este problema lo resuelven las fotocélulas que trabajan con luz polarizada. En ellas se hace pasar el haz a través de filtros polarizadores, que sólo dejan pasar la luz cuyas ondas oscilan en un plano determinado (una luz sin polarizar está compuesta por ondas electromagnéticas que oscilan en todas las direcciones). Estos filtros se colocan con los planos de polarización girados 90° entre sí. El espejo retrorreflector, además de reflejar el haz, abate su plano de polarización 90° (Figura 7.99a). La fotocélula detecta los objetos pulidos o brillantes porque no abaten el plano de polarización (Figura 7.99b).



a)



b)

Figura 7.99 Funcionamiento de la fotocélula de luz polarizada: a) El espejo retrorreflector abate el plano de polarización del haz; b) El objeto no abate el plano de polarización.

Normas de instalación

En este tipo de sensores se deben tener en cuenta también varias normas de instalación similares a las indicadas para los sensores de tipo barrera de luz en el apartado 7.7.2.3.3.4. En particular, es necesario tener en cuenta que, tal como se indica en el apartado anterior, las fotocélulas de reflexión sobre espejo que no actúan con luz polarizada pueden provocar errores de detección de los objetos brillantes. Se puede evitar esto, tal como se muestra en la figura 7.100, haciendo que la superficie del objeto a detectar tenga una cierta inclinación con respecto a la perpendicular al eje óptico de la fotocélula. Se logra así que la luz reflejada en el objeto salga desviada y no llegue al receptor.

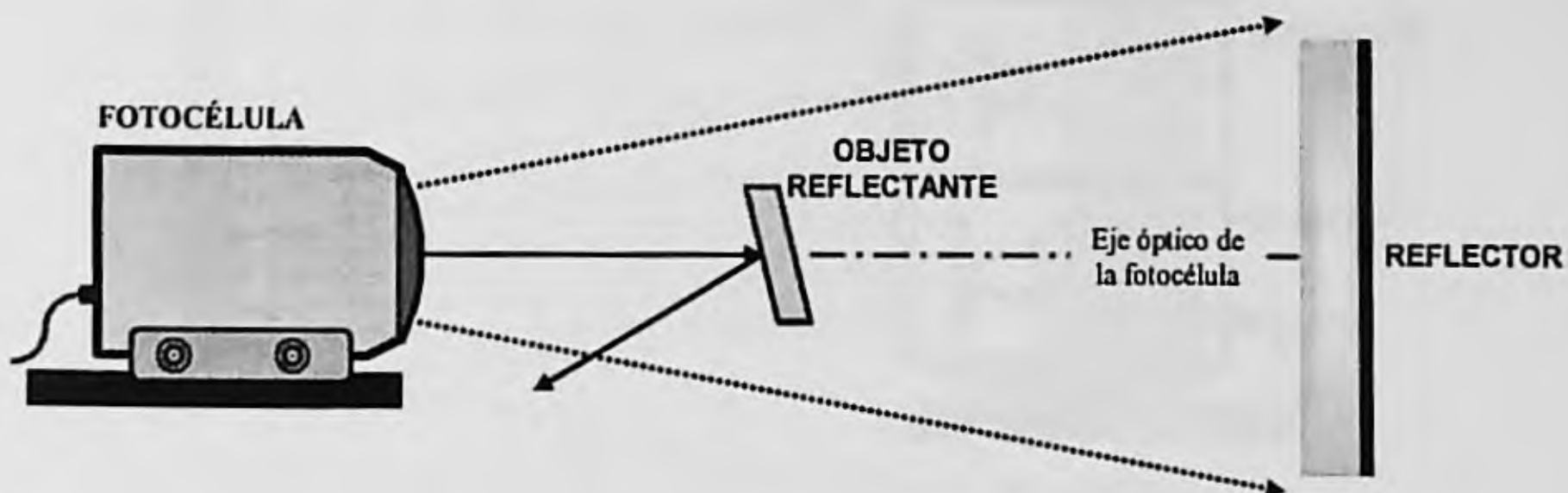


Figura 7.100 Efecto de la inclinación del objeto con respecto al eje óptico de la fotocélula.

7.7.2.3.3.6 Sensores optoelectrónicos de reflexión sobre objeto

Introducción

Las fotocélulas de reflexión sobre objeto (también conocidas como fotocélulas de reflexión directa) se componen únicamente de un emisor y un receptor montados en una misma carcasa. Se diferencian de las fotocélulas de reflexión sobre espejo en que el haz de luz se refleja directamente sobre el objeto que se desea detectar, para lo cual éste debe tener unas características mínimas de reflectividad. Según su constitución física las fotocélulas de reflexión sobre objeto pueden ser de reflexión difusa o de reflexión definida, que se describen a continuación.

Fotocélulas de reflexión difusa. Modo de operación y campo de trabajo

En las fotocélulas de reflexión difusa (*Diffuse Reflective Sensors*) sobre objeto, el emisor emite un haz de luz que se refleja en el objeto y vuelve hacia el receptor. Se caracterizan por el hecho de que la luz que recibe el objeto se refleja en múltiples direcciones (Figura 7.101) y por ello se denomina difusa.

Al igual que en las fotocélulas de reflexión sobre espejo, el campo de trabajo está formado por la intersección del campo de emisión y del campo de recepción, pero la distancia de detección es mucho más pequeña. Existe también una zona muerta en la que no se garantiza la detección del objeto, tal como se indica en la figura 7.102. Dicha zona debe ser tomada en cuenta cuando se pretende detectar objetos que están situados muy próximos a la fotocélula y por ello,

con el fin de eliminarla, se fabrican fotocélulas que poseen una configuración especial en la que los ejes ópticos del emisor y del receptor coinciden (Figura 7.103). Al igual que en las fotocélulas de barrera de luz, el fabricante suministra una gráfica, que representa el rango de operación o campo de trabajo y especifica la zona de detección. Además, debido a que la reflectividad del objeto influye en gran medida en la distancia de detección, los fabricantes especifican, en general, la distancia de detección para un objeto (*target*) normalizado, que suele ser de papel blanco mate. La figura 7.104 muestra el diagrama del rango de operación de la fotocélula de reflexión difusa sobre objeto PXO600. La figura 7.105 muestra la fotocélula de reflexión difusa sobre objeto PXO560.

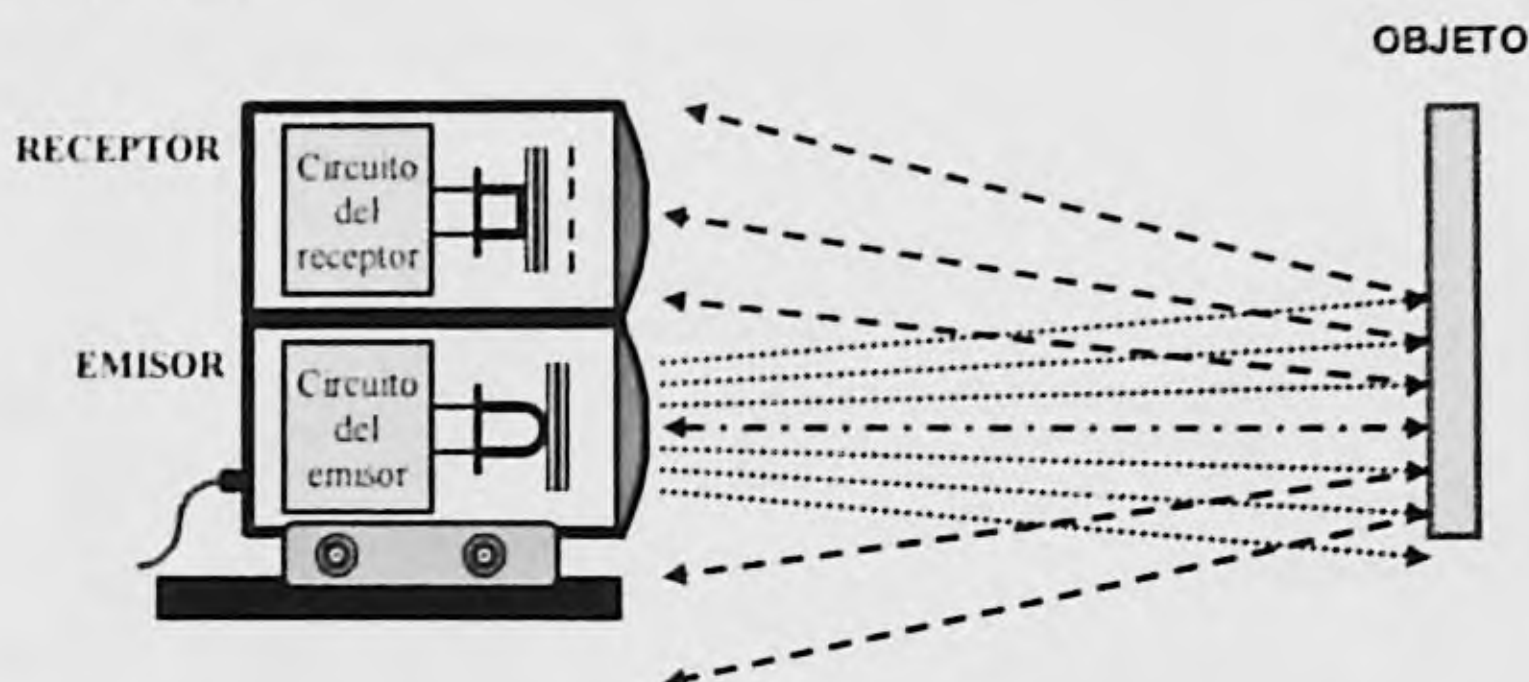


Figura 7.101 Funcionamiento de una fotocélula de reflexión difusa sobre objeto.

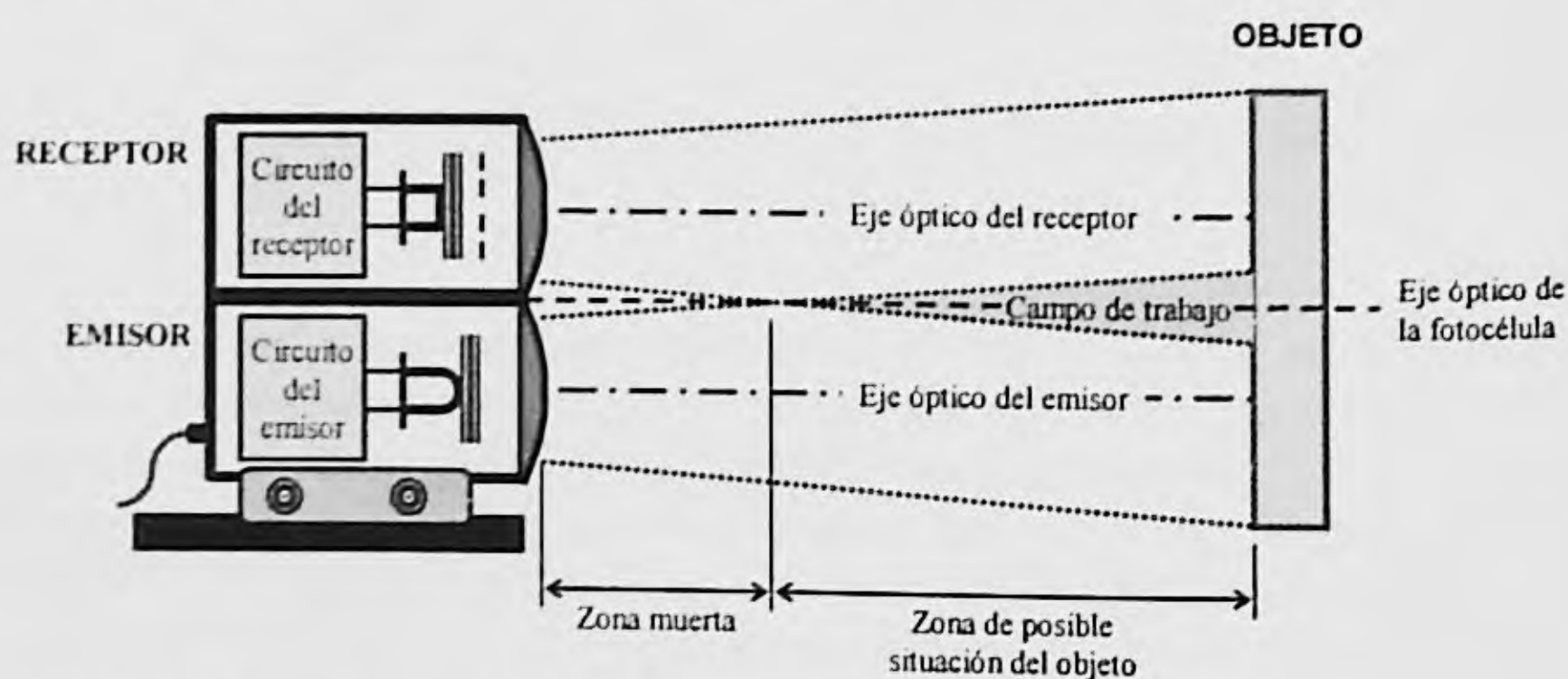


Figura 7.102 Campo de trabajo con zona muerta.

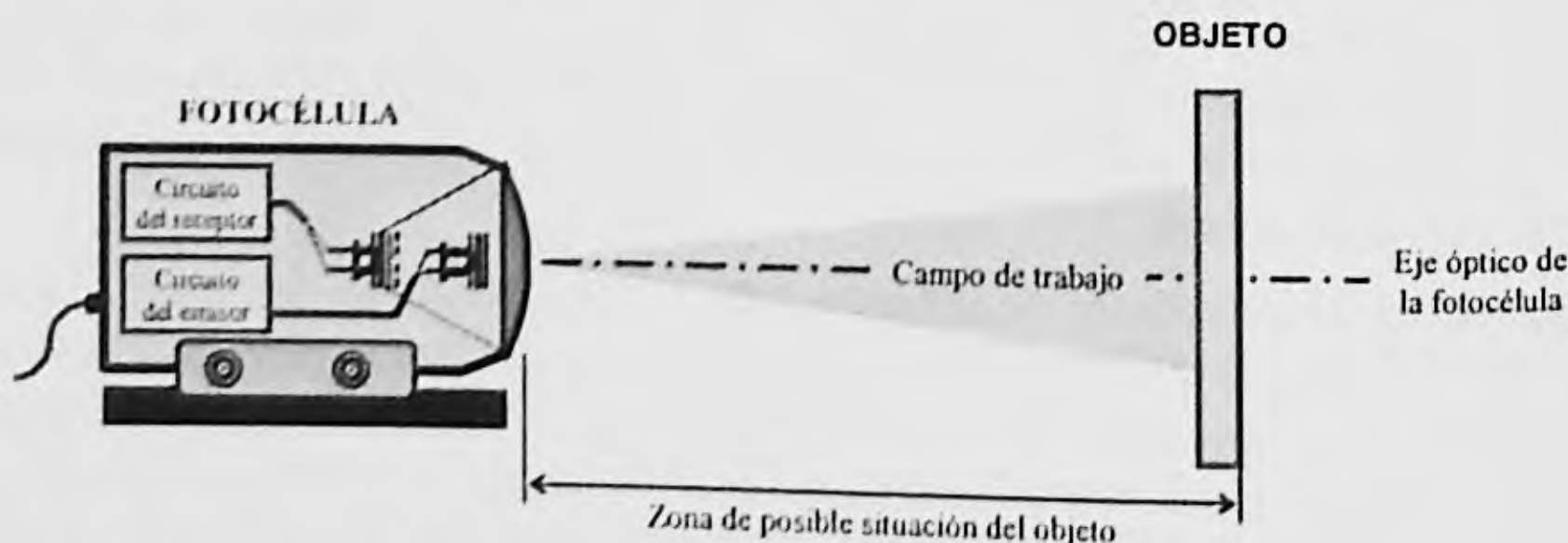


Figura 7.103 Campo de trabajo de una fotocélula que carece de zona muerta.

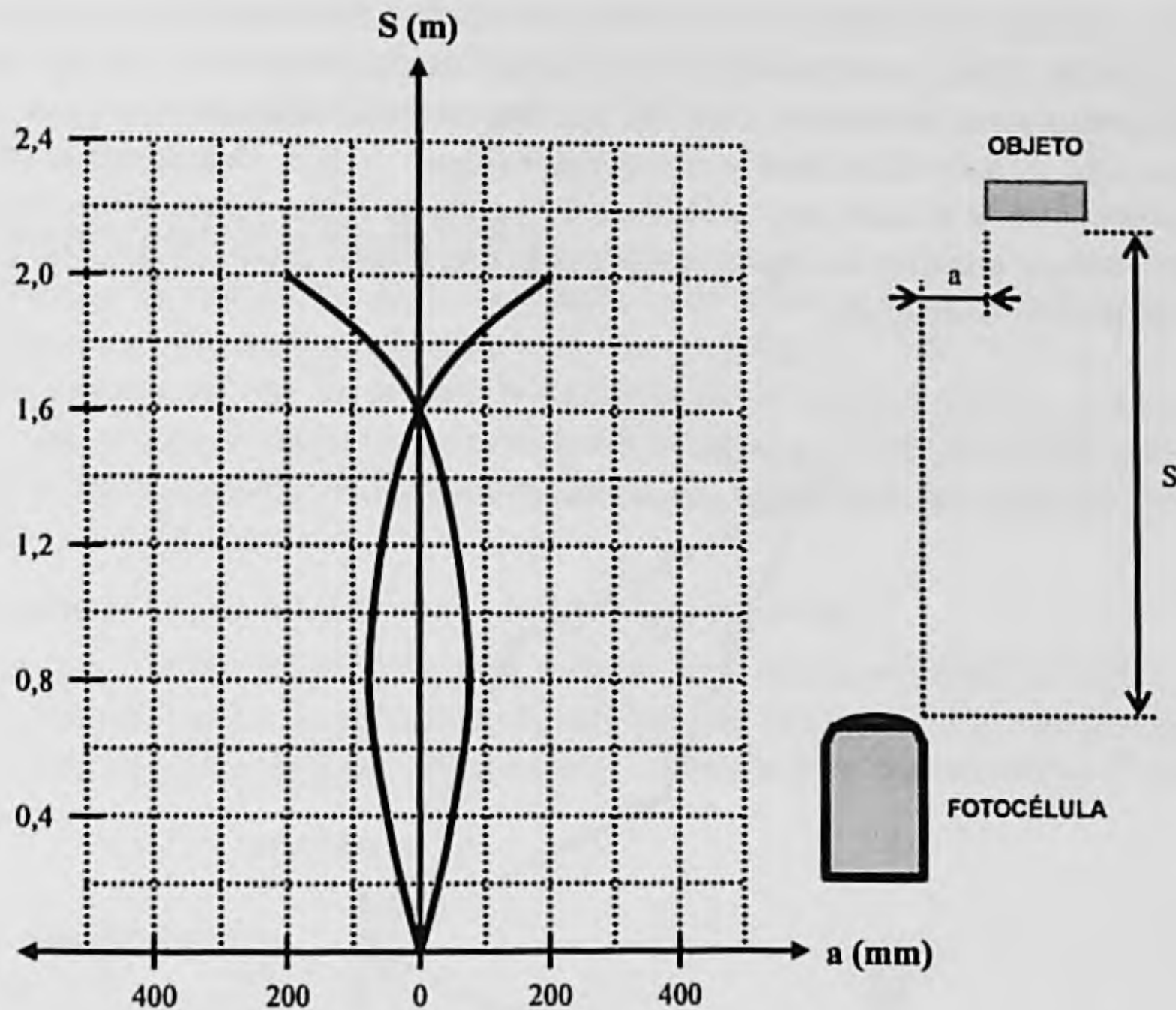


Figura 7.104 Campo de trabajo de la fotocélula de reflexión difusa sobre objeto PXO600 (cortesía de Siemens).



Figura 7.105 Fotocélula de reflexión difusa sobre objeto PXO560 (cortesía de Siemens).

Fotocélulas de reflexión definida. Modo de operación y campo de trabajo

Las fotocélulas de reflexión definida (*Definite reflective sensors*) se caracterizan también porque el haz de luz se refleja en el objeto y vuelve hacia el receptor, al igual que en las de reflexión difusa, pero se diferencian de estas últimas en que su sistema óptico es diferente. En las fotocélulas de reflexión definida la fuente de luz (E) está situada en el eje óptico de la lente del emisor pero a mayor distancia que el foco (F) de la misma. Lo mismo ocurre con el receptor

(R), tal como se indica en la figura 7.106 en la cual el emisor está situado a una distancia de la lente igual al doble de la distancia focal de la misma. Esto hace que los rayos enviados por el emisor solamente lleguen al receptor si el objeto a detectar está justamente en el punto en el que se cruzan los ejes ópticos del emisor y del receptor (figura 7.107). Debido a ello el sensor no detecta el objeto si está situado más adelante o más atrás de dicho punto. Estas fotocélulas se utilizan para detectar posición de objetos con elevada precisión, además de detectar la presencia de objetos pequeños, marcas, etc.

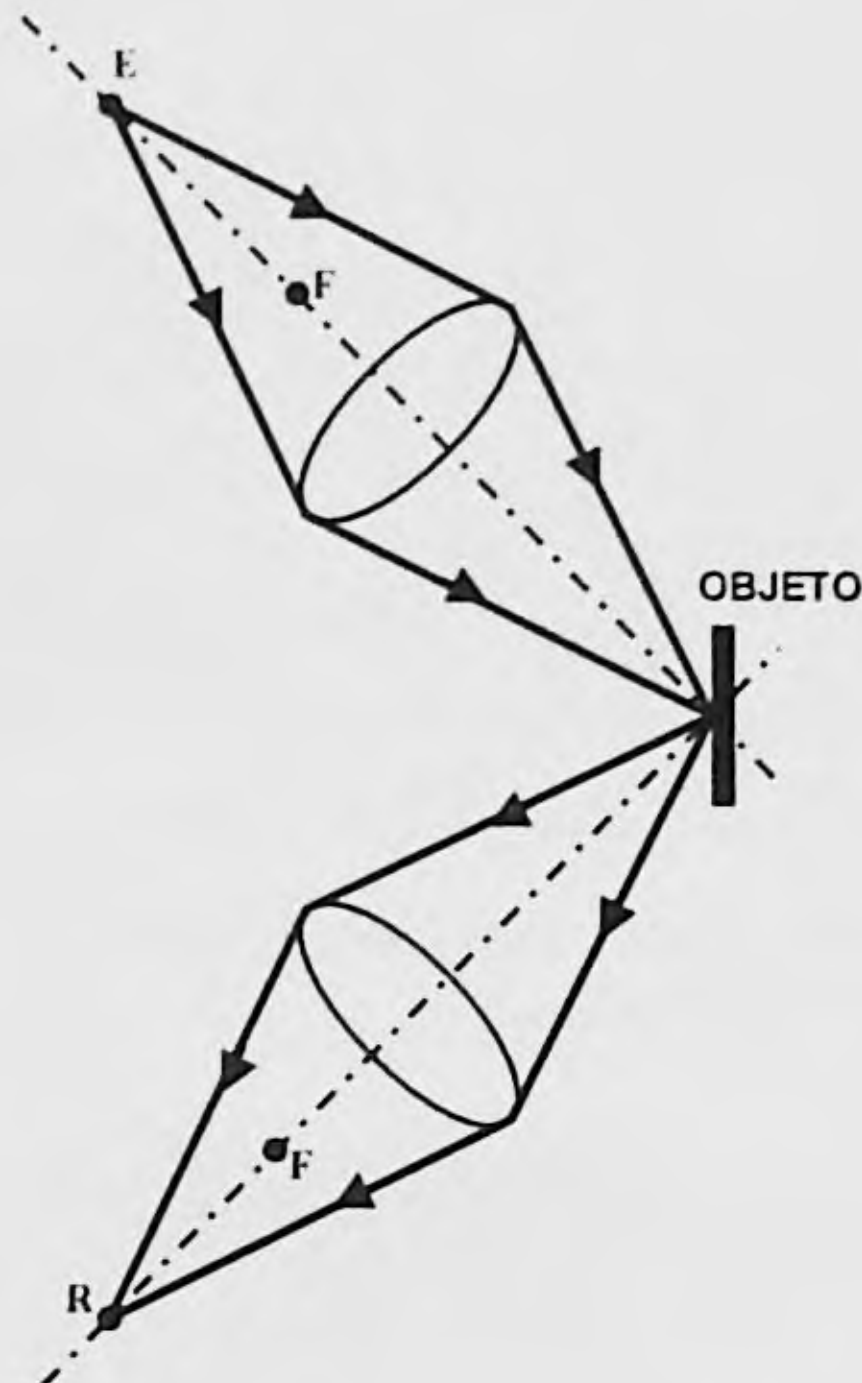


Figura 7.106 Principio de funcionamiento de una fotocélula de reflexión definida.

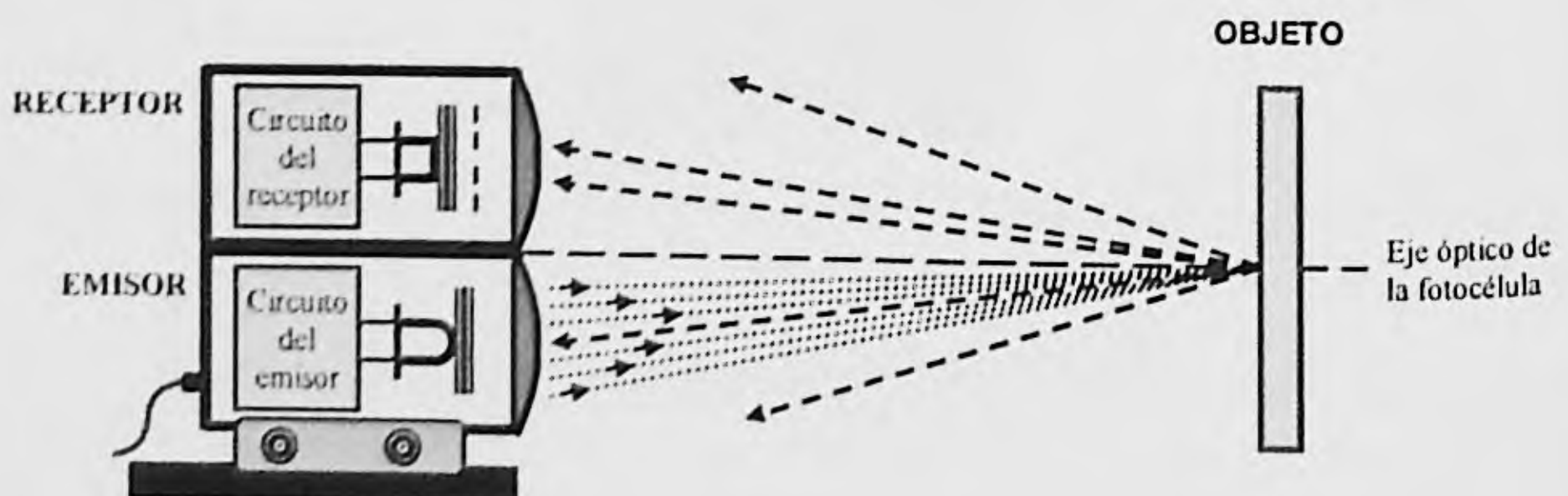


Figura 7.107 Funcionamiento de una fotocélula de reflexión definida.

Para aumentar la precisión del posicionamiento, el campo de trabajo de los sensores de reflexión definida es pequeño (Figura 7.108). En algunos modelos se puede variar dicho campo, lo que se traduce en una distancia de detección ajustable. Dicha variación se puede conseguir de dos formas distintas:

- **Mediante la variación del eje óptico**

En este caso la inclinación de los ejes ópticos del emisor y del receptor se varía mediante un dispositivo mecánico, tal como se indica en la figura 7.109, en la que se representa esquemáticamente una fotocélula de distancia de detección variable, en dos posiciones diferentes. Como se puede observar en la citada figura, el campo de trabajo (área sombreada) se estira al aumentar la distancia de detección, lo que hace que aumente la longitud del mismo (zona de detección).

- **Mediante la variación del campo de emisión-recepción**

En este caso, la fotocélula tiene dos ranuras y otras tantas pantallas deslizantes cuya posición se puede variar simultáneamente mediante un dispositivo mecánico (Figura 7.110). Se modifican así los ángulos de emisión y recepción y en consecuencia el campo de detección.

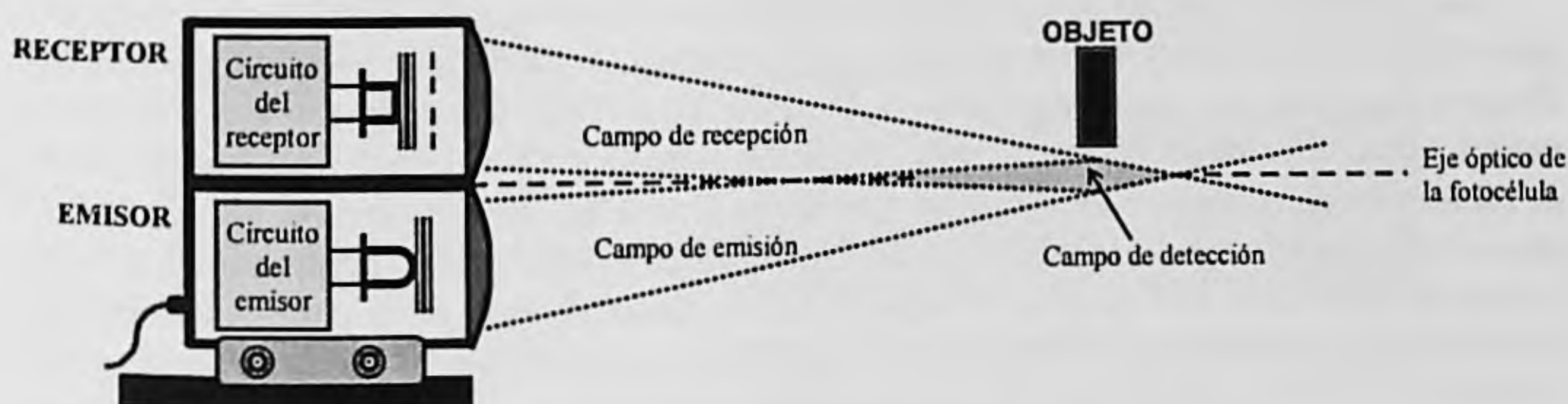


Figura 7.108 Campo de trabajo de una fotocélula de reflexión definida.

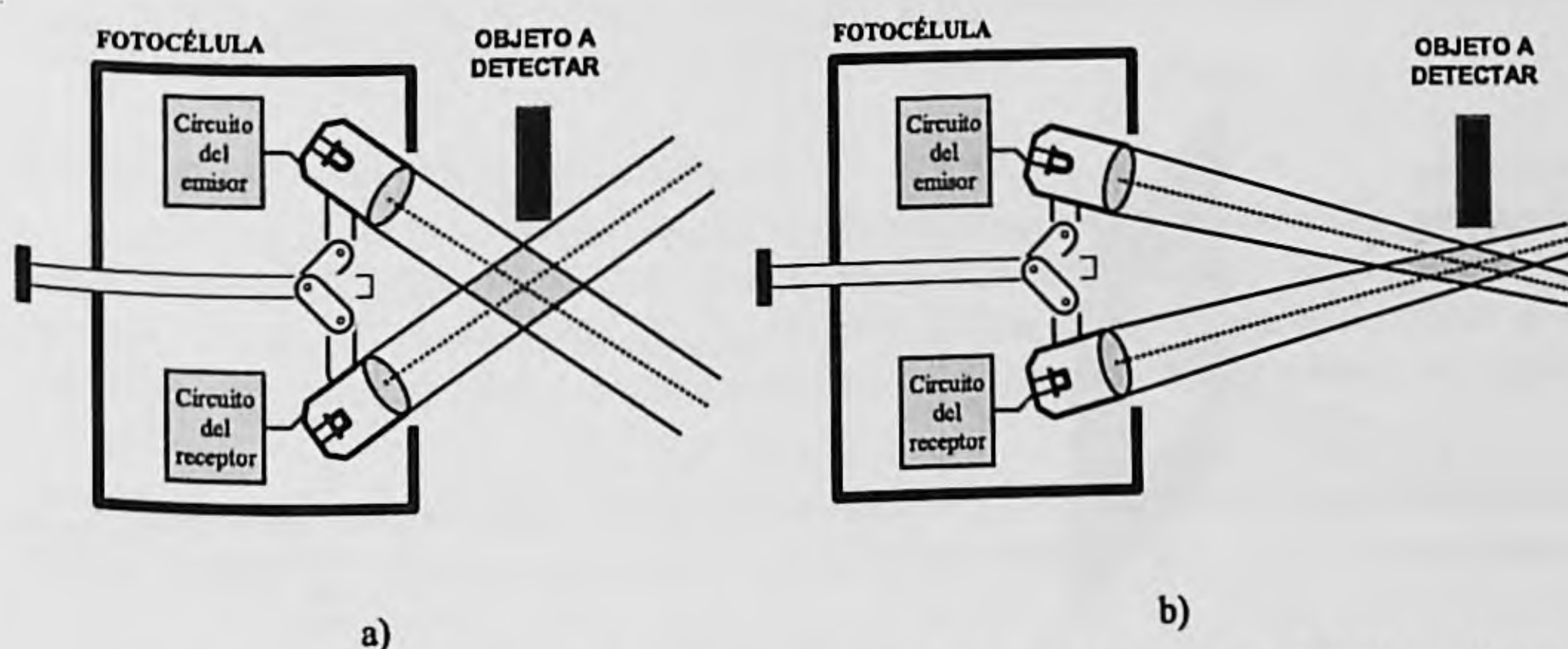


Figura 7.109 Modificación del campo de trabajo de una célula de reflexión definida, mediante variación del eje óptico.

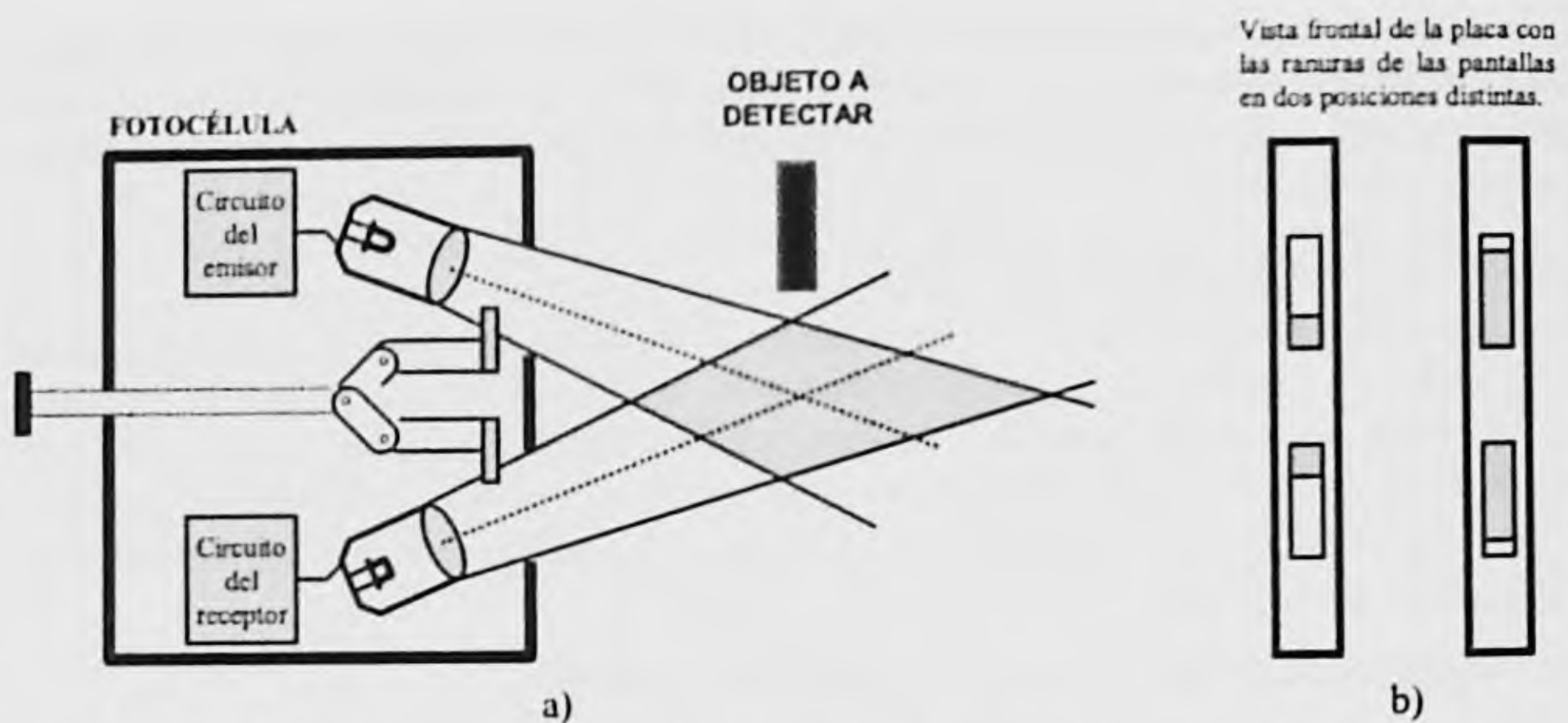


Figura 7.110 Modificación del campo de trabajo de una célula de reflexión definida, mediante variación del campo de emisión-recepción.

Sensores de marcas

Los sensores de marcas son fotocélulas de reflexión definida que detectan una marca que tiene diferente color que el objeto sobre el que está situada (Figura 7.111a). Para ello se utilizan distintas fuentes de luz, implementadas con diodos luminiscentes (normalmente de color rojo o verde), según el color de las marcas y del fondo. Se debe elegir la fuente de luz más adecuada, de tal forma que el contraste entre la luz que refleja la marca y la que refleja el fondo sobre el que se sitúa la misma sea máximo. De la gráfica de la figura 7.111b se deduce que con luz verde se detecta fácilmente una marca roja sobre un fondo blanco, porque presentan gran diferencia de reflectividad (90% para el fondo frente a 10% para la marca). Por el contrario con luz roja la diferencia es muy pequeña (90% del fondo frente al 85% para la marca) y la detección no es posible.

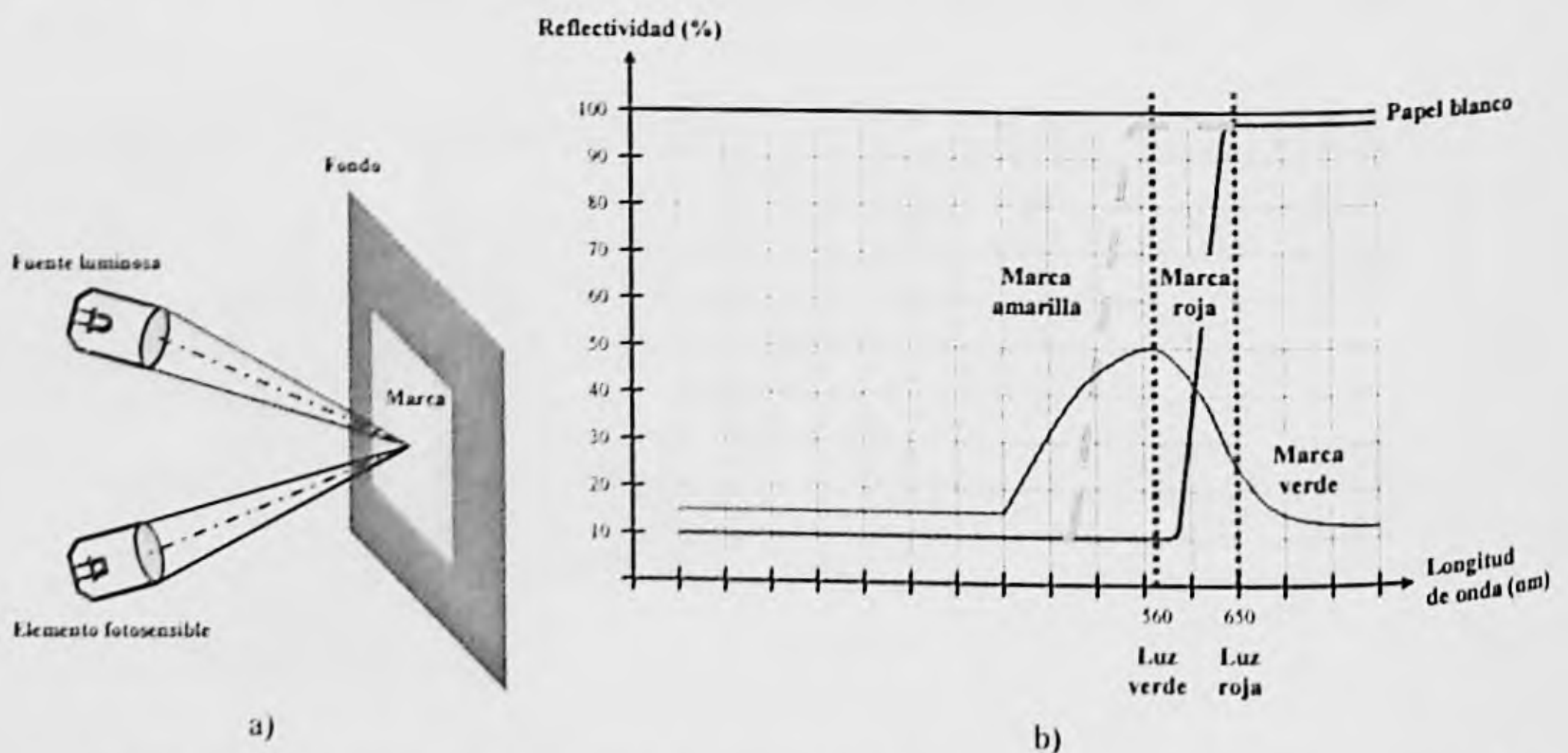


Figura 7.111 Fotocélula de reflexión definida que detecta marcas de colores: a) Representación del funcionamiento; b) Representación gráfica de la influencia del color de las marcas.

También se observa que la marca verde sobre fondo blanco es más fácilmente detectable con la luz roja y que la amarilla no se puede detectar con luz roja ni con luz verde. Por ser fotocélulas de reflexión definida, las marcas pueden ser de tamaño muy reducido.

7.7.2.3.3.7 Fotocélulas de fibra óptica

Introducción

La fibra óptica (*Optical fiber*) es una guía de ondas luminosas que se propagan a través de ella mediante reflexiones sucesivas. Está constituida por dos cilindros coaxiales de material transparente a la radiación y de diferente índice de refracción, tal como se indica en la figura 7.112. El cilindro interior, denominado núcleo (*Core*), es el medio de propagación del rayo luminoso con una atenuación y distorsión mínimas. El segundo cilindro, denominado revestimiento o corteza (*Cladding*) refleja el rayo de luz debido a que su índice de refracción es distinto que el del núcleo. El tercer cilindro exterior, denominado funda o recubrimiento (*Jacket*) es una protección primaria a base de barniz de acrilato o resina de silicona, que protege a la fibra contra la abrasión, rotura y agresiones químicas.

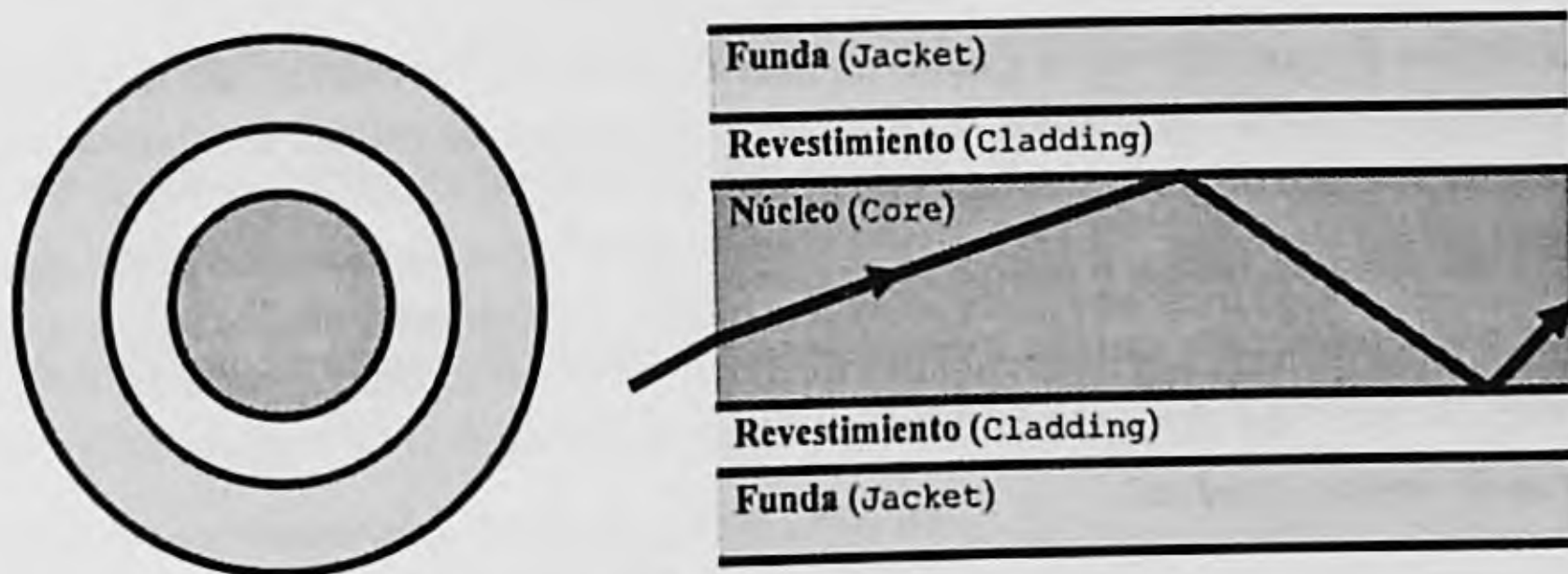


Figura 7.112 Modo de operación de una fibra óptica.

La combinación de las fotocélulas con una fibra óptica facilita la utilización de las fotocélulas en sistemas en los que el espacio es reducido.

Modos de operación de las fibras ópticas

El funcionamiento de las fibras ópticas está basado en el cambio de dirección que experimenta un rayo de luz cuando atraviesa la superficie que divide dos medios de diferente índice de refracción. Dicho cambio en la dirección de propagación del rayo luminoso depende de la naturaleza de los materiales que forman los medios (índice de refracción) y además, del ángulo de incidencia.

Se define el índice de refracción (*Refraction index*) n de un medio como la relación entre la velocidad C de la luz en el vacío y v en el medio considerado:

$$n = C / v$$

y su valor es siempre mayor que la unidad.

En las fibras ópticas, el índice de refracción n_1 del núcleo es mayor que el índice de refracción n_2 del revestimiento, lo que hace que se propaguen a través del núcleo de la fibra los rayos que inciden en el límite entre el núcleo y el revestimiento con un ángulo α inferior al ángulo β_c , denominado ángulo crítico, tal como se indica en la figura 7.113. Por el contrario, los rayos que inciden en dicho límite con un ángulo superior al ángulo crítico, pasan por refracción al revestimiento y se anulan después de algunas reflexiones.

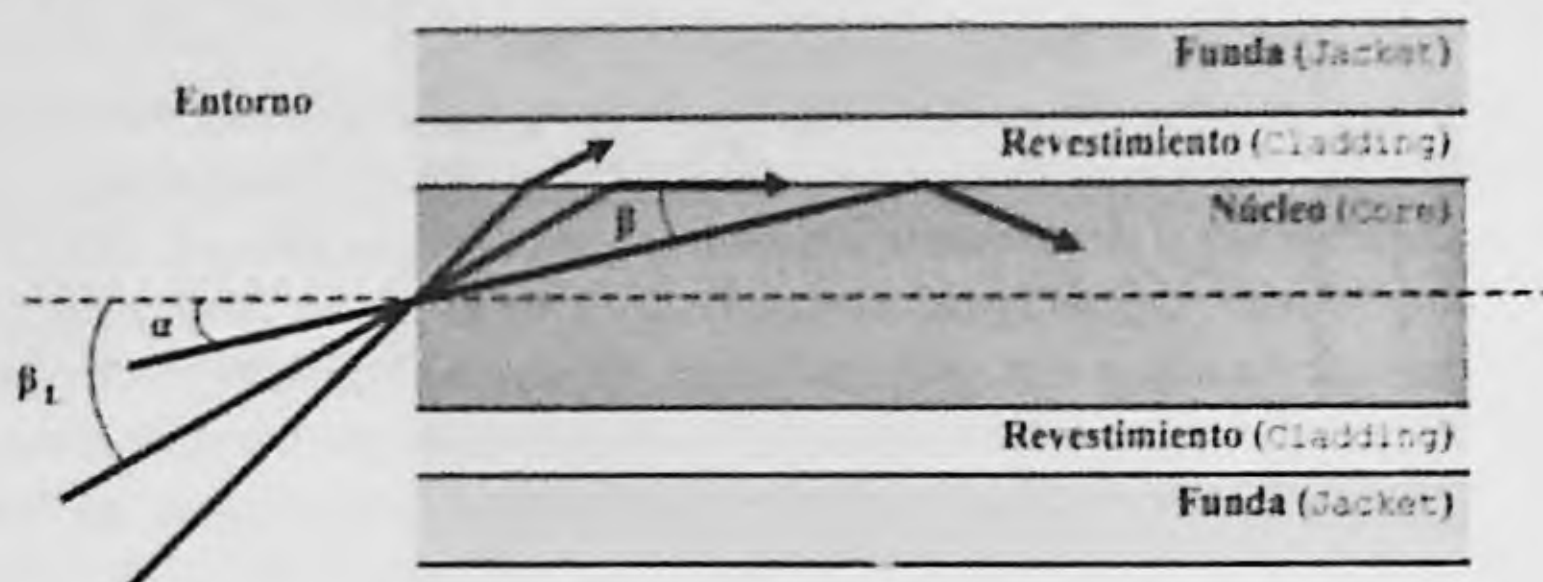


Figura 7.113 Transmisión de la luz a través de una fibra óptica.

La condición de que $n_1 > n_2$ se puede realizar con un cambio brusco del índice o con un cambio continuo. Esto da lugar a dos tipos de perfil del índice de refracción del núcleo, figura 7.114, que a su vez dan origen a dos tipos de fibras:

- **Fibras de salto de índice o índice en escalón**

El índice de refracción cambia bruscamente del núcleo al revestimiento. En este caso la trayectoria seguida por el rayo de luz es lineal.

- **Fibras de índice gradual**

El índice del núcleo es variable y depende de la distancia al eje de la fibra, y es constante en el revestimiento. Esto da lugar a que la trayectoria del rayo sea curva.

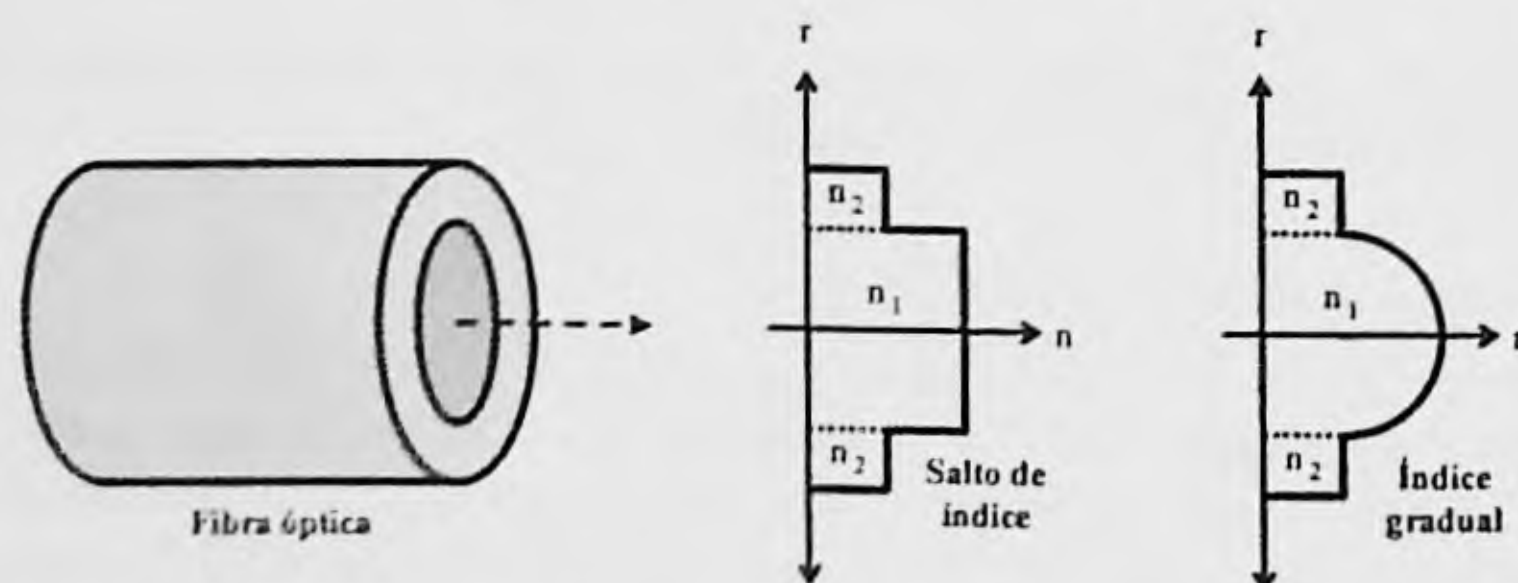


Figura 7.114 Perfiles de fibras ópticas.

La luz se propaga dentro de las fibras ópticas en monomodo o en multimodo. En el caso de propagación monomodo la luz recorre una única trayectoria en el interior del núcleo, lo que proporciona un gran ancho de banda. Para minimizar el número de reflexiones en la superficie externa del núcleo, éste debe ser lo más estrecho posible. Las dificultades de fabricación y empalme de las fibras monomodo llevaron al desarrollo de las fibras multimodo cuyo núcleo

tiene un diámetro mayor y el número de trayectorias que pueden recorrer los rayos de luz en el interior es también mayor. Esto da lugar a la dispersión de las componentes de los impulsos de luz, debido a que cada una de ellas recorre una trayectoria distinta y alcanza su destino en un instante diferente, tras haber sufrido un número de reflexiones internas diferente, lo que disminuye la velocidad de transmisión y aumenta la distorsión de la señal.

En el mercado existen tres tipos de fibras ópticas:

- **Fibras multimodo de salto de índice**

Presentan una distorsión elevada, que limita el ancho de banda entre 10 y 50 MHz. Por ello se utilizan en conexiones relativamente cortas hasta (1Km) en enlaces de datos de baja velocidad.

- **Fibras monomodo de salto de índice**

La dispersión es baja y se pueden lograr anchos de banda de varios GHz/Km.

- **Fibras multimodo de índice gradual**

La luz se refracta desde las partes más externas del núcleo, con lo cual se evitan las reflexiones que aparecen en las fibras de salto de índice. La velocidad es mucho mayor que en las fibras multimodo de salto de índice porque se reduce la dispersión, y el ancho de banda puede superar el GHz (1 Km).

Sensores optoelectrónicos de fibra óptica

Los sensores optoelectrónicos de fibra óptica (*Fiberoptic sensors*) están formados por una fuente de luz (diodo luminiscente, infrarrojo, láser, etc.), una fibra óptica, un modulador que altera la luz de acuerdo con el valor de una determinada magnitud física y un fotodetector. De acuerdo con la función de la fibra óptica este tipo de sensores puede ser:

- **Sensores extrínsecos o de modulación externa**

Son sensores de fibra óptica en los que la luz transportada por la fibra óptica se modifica en un elemento externo bajo la acción de la magnitud a medir. Se utilizan, entre otras aplicaciones, para detectar objetos.

- **Sensores intrínsecos o de modulación interna**

Son sensores de fibra óptica en los que la magnitud a medir modifica las características ópticas de la fibra, directamente o a través de un recubrimiento. Dichas modificaciones producen variaciones de algún parámetro de la radiación luminosa.

- **Sensores evanescentes**

Son sensores de fibra óptica en los que una parte de la misma está desprovista de la capa reflectora externa. Esto permite detectar ciertos elementos químicos (como por ejemplo el hidrógeno) o la densidad del medio en el que está situado el sensor, a través de su influencia en el campo evanescente, que está constituido por la cantidad de luz que sale de la fibra, que depende del índice de refracción del medio al que pasa. Se utilizan especialmente en óptica integrada.

Implementación de las fotocélulas de fibra óptica

Las fotocélulas de fibra óptica son sensores extrínsecos, implementados, en general, con fibras multimodo de salto de índice. Su carcasa se instala en la zona más adecuada y la luz se

transmite al punto de detección mediante una fibra óptica. En lugares de difícil acceso, las fibras ópticas se pueden utilizar con fotocélulas de barrera de luz, de reflexión sobre espejo y de reflexión sobre objeto. En la figura 7.115 se representa una fotocélula de fibra óptica (Figura 7.115a) y su aplicación como barrera de luz (Figura 7.115b) y como sensor de objetos (Figura 7.115c) mediante reflexión directa.

Las figuras 7.116 y 7.117, muestran fibras ópticas para barrera de luz y de reflexión. La figura 7.118 muestra la fotocélula para fibras ópticas PXO840 LV70.

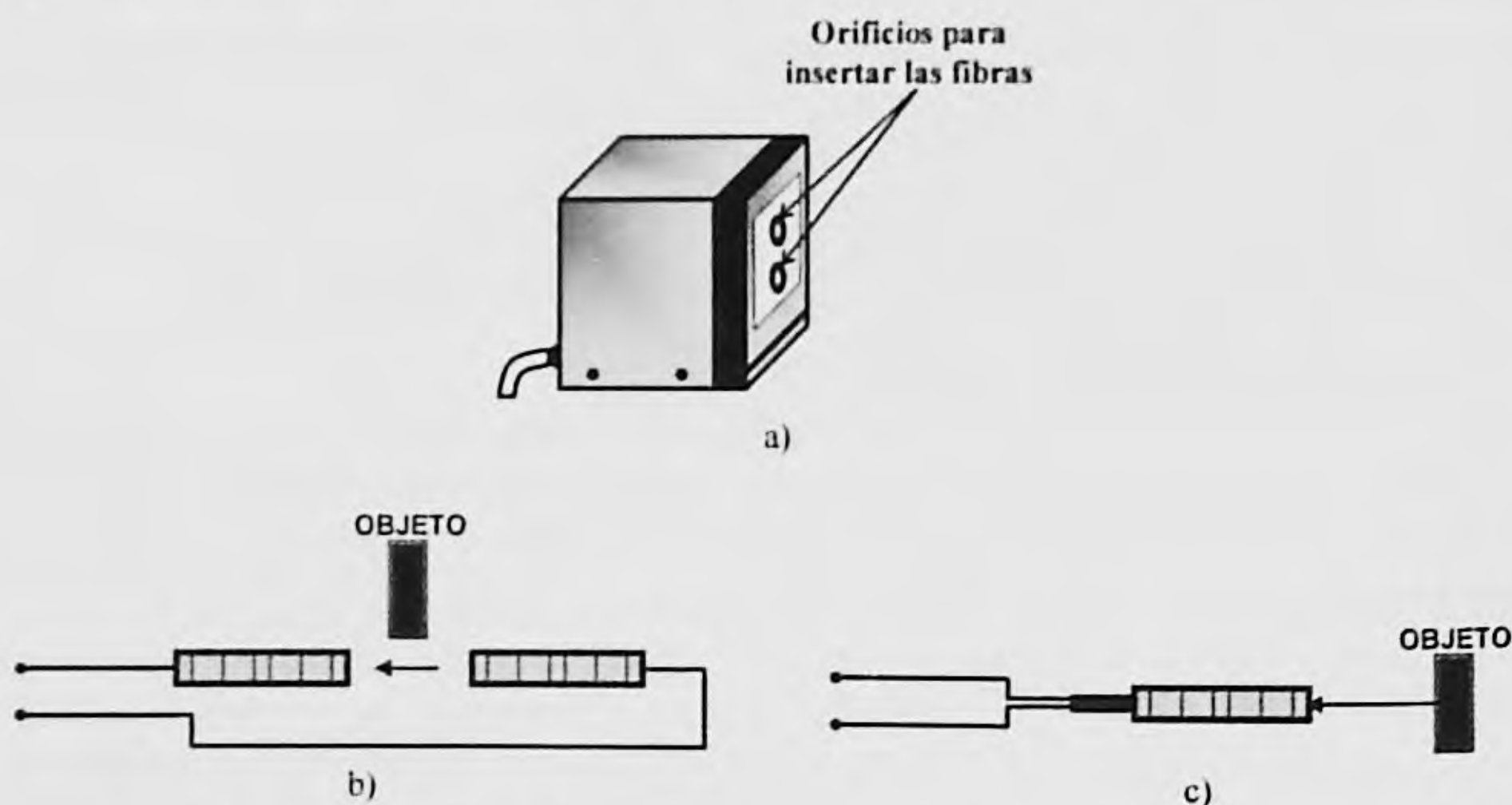


Figura 7.115 Fotocélula de fibra óptica.

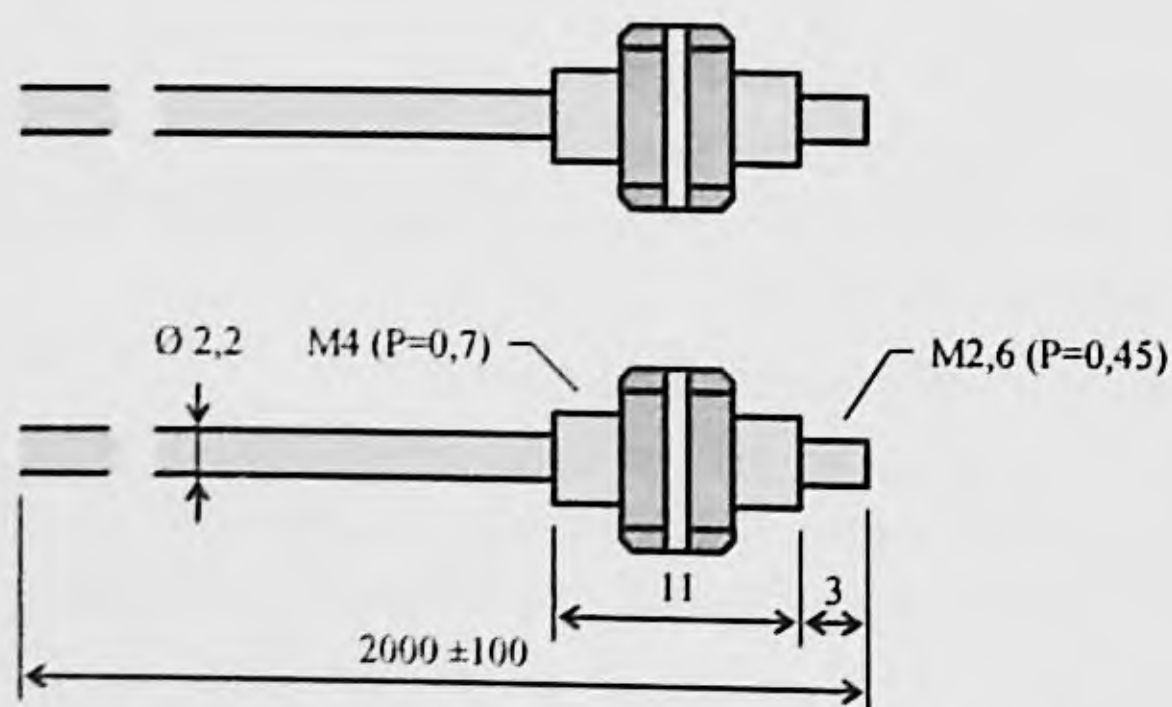


Figura 7.116 Fibra óptica para barrera de luz (cortesía de Siemens).

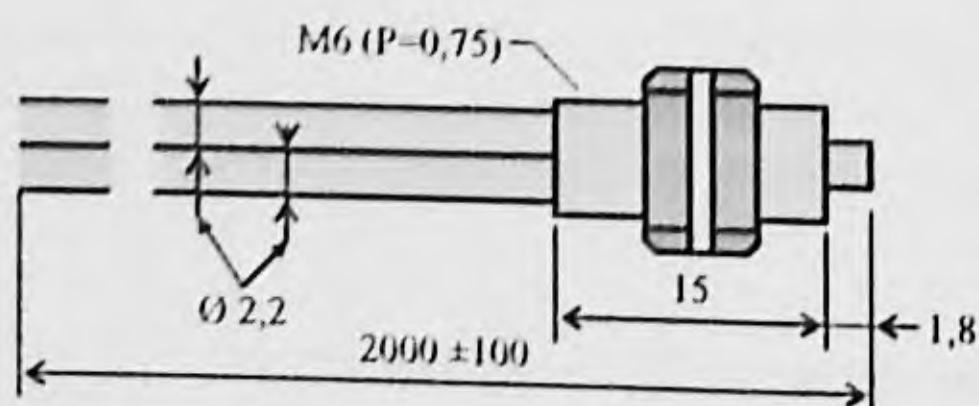


Figura 7.117 Fibra óptica para fotocélula reflex (cortesía de Siemens).

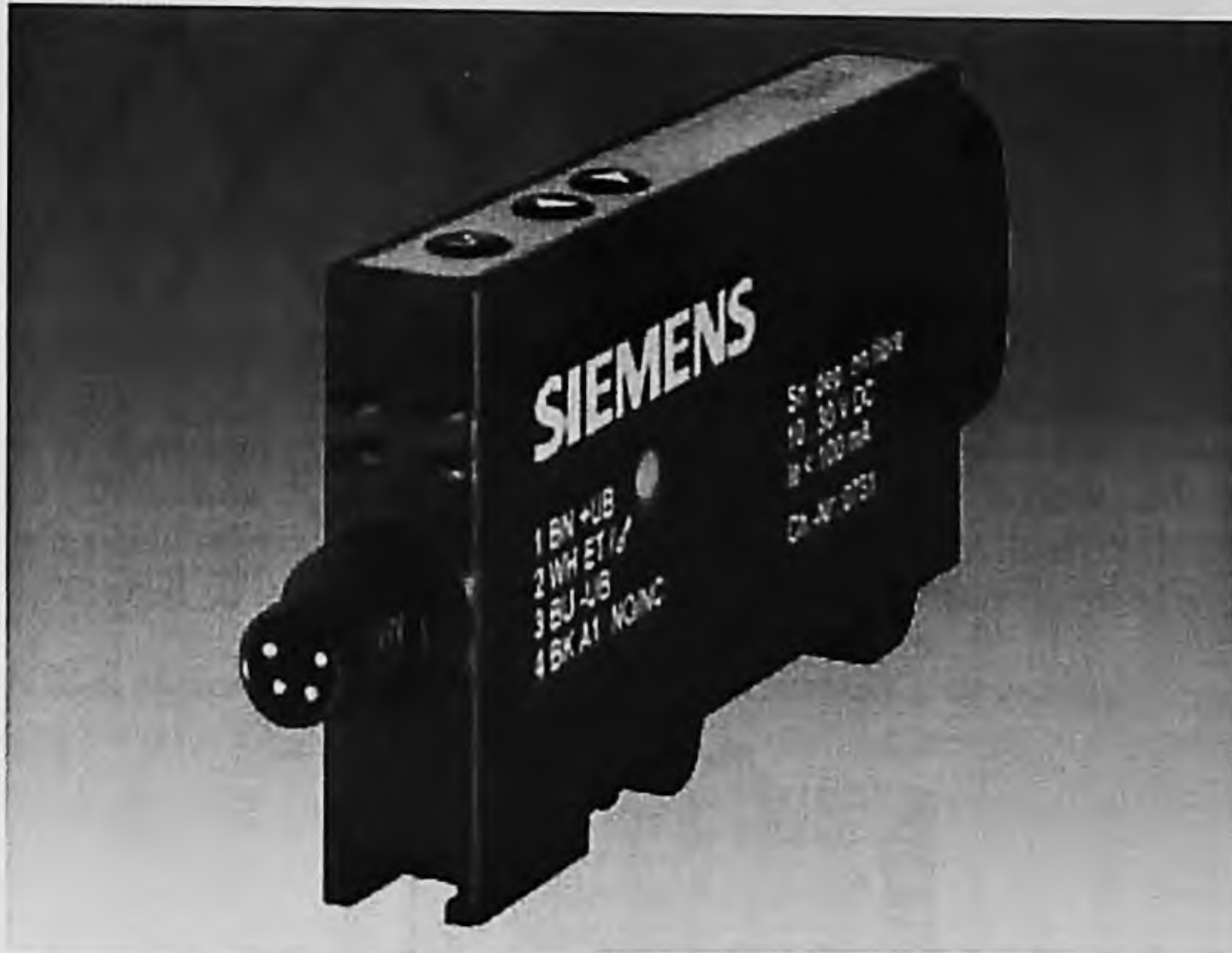


Figura 7.118 Focélula para fibras ópticas PXO840 LV70 (cortesía de Siemens).

7.7.2.3.4 Sensores magnéticos de proximidad

Los sensores de proximidad magnéticos basan su funcionamiento en la influencia del campo magnético sobre algunos metales y semiconductores. Un ejemplo es el representado en la figura 7.119, que consta de una ampolla de vidrio en la que se colocan dos láminas metálicas. Dichas láminas constituyen un contacto que está abierto cuando no se le aplica un campo magnético y se cierra al aproximarle un imán adosado al objeto a detectar. Otro ejemplo son los sensores de proximidad basados en el efecto Hall [PALL 98]. Los sensores magnéticos se utilizan como sensores de seguridad, tal como indica en el apartado 10.2.3.2 del capítulo 10.



Figura 7.119 Funcionamiento de un sensor de proximidad magnético: a) Sin aplicar un campo magnético; b) Al aplicar un campo magnético.

7.7.2.3.5 Sensores inductivos de proximidad

7.7.2.3.5.1 Conceptos generales

Los sensores inductivos de proximidad (*Inductive proximity detectors*) basan su funcionamiento en la interacción entre el objeto a detectar y el campo electromagnético generado por el propio sensor.

El campo magnético creado por un imán o por una bobina, a través de la cual pasa una corriente, es una región del espacio en la que existe un conjunto de líneas de fuerza que se cierran sobre dos polos sin fuentes ni sumideros, tal como se representa en la figura 7.120 para el caso de un imán. La cantidad de líneas magnéticas se denomina flujo magnético (Φ) y se mide en Webbers [Wb]. El flujo magnético por unidad de superficie se denomina inducción magnética (B) y se mide en Webbers por m^2 [Wb/m^2]. 1 Wb/m^2 recibe el nombre de Tesla [T].

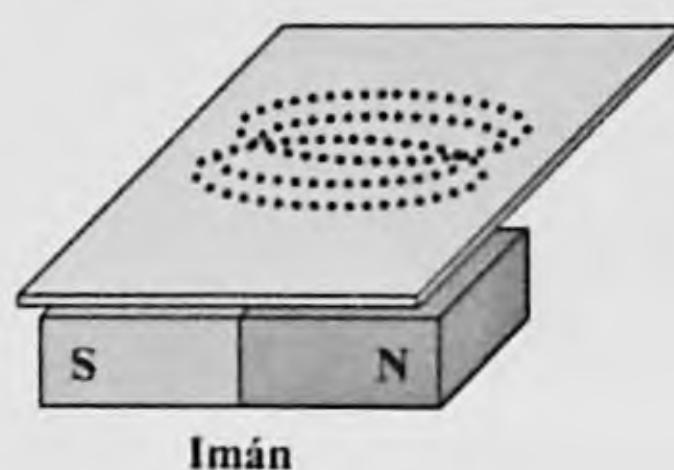


Figura 7.120 Campo magnético creado por un imán.

La resistencia que opone cada material al paso de las líneas de fuerza se denomina reluctancia magnética (R), se mide en Amperios por Webber [A/Wb] y es análoga a la resistencia que oponen los conductores al paso de una corriente cuando se les aplica una tensión eléctrica. La reluctancia es inversamente proporcional a la permeabilidad magnética del material (μ), que a su vez es análoga a la conductividad de los conductores eléctricos y se mide en [$Wb/(A \cdot m)$]. La permeabilidad magnética de un material (μ), es el producto de la permeabilidad del vacío (μ_0), que es una constante, por la permeabilidad relativa (μ_r) de ese medio con relación al vacío. La permeabilidad magnética es una característica de cada material y cuanto mayor es su valor mejor conductor del campo magnético es el material. De acuerdo con el valor de la permeabilidad los materiales son:

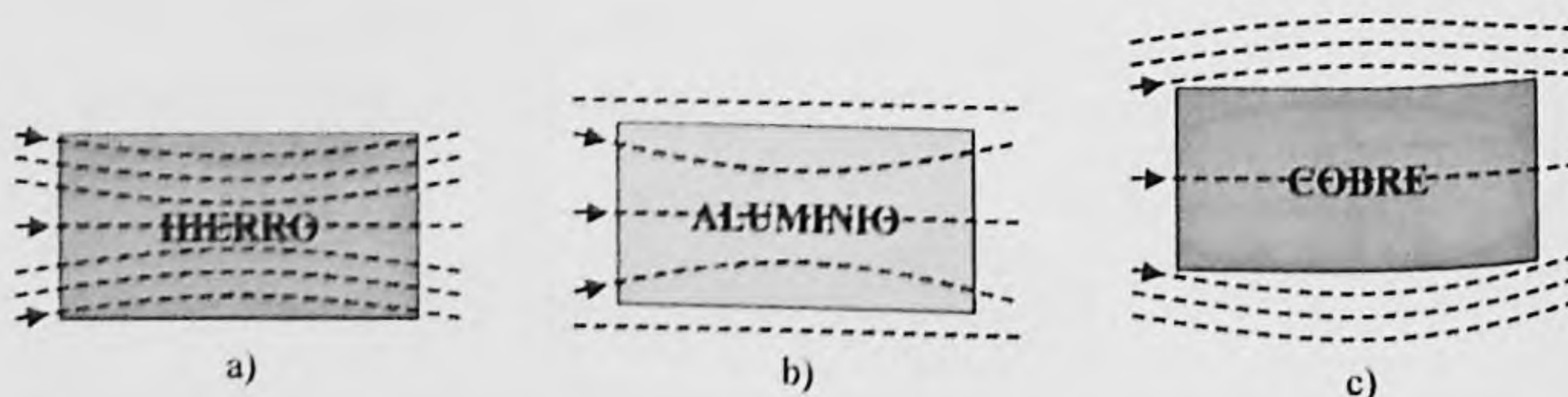


Figura 7.121 Comportamiento de diferentes tipos de materiales ante un campo magnético: a) Material ferromagnético; b) Material paramagnético; c) Material diamagnético.

- Ferromagnéticos, como por ejemplo el hierro, cuya permeabilidad magnética relativa es muy superior a la unidad ($\mu_r \gg 1$) (Figura 7.121a).
- Paramagnéticos, como por ejemplo el aluminio, cuya permeabilidad magnética relativa es superior a la unidad ($\mu_r > 1$), (Figura 7.121b).
- Diamagnéticos, como por ejemplo el cobre, cuya permeabilidad magnética relativa es inferior a la unidad ($\mu_r < 1$), (Figura 7.121c).

Cuando se hace pasar una corriente I a través de una bobina se genera un flujo magnético Φ que depende de la inductancia L de la bobina, que es función de las características físicas de la misma y de la permeabilidad del núcleo de acuerdo con la expresión:

$$L = N \frac{\Phi}{I} = \mu \frac{N^2}{l} A$$

en la que N es el número de espiras, l la longitud de la bobina y A su sección.

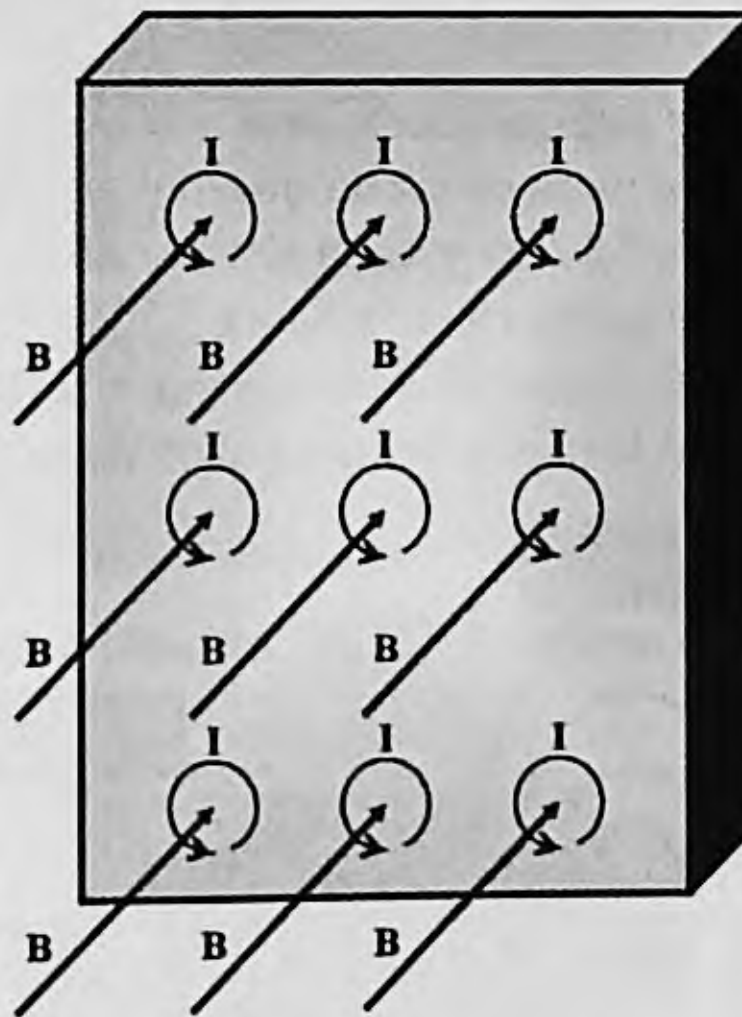


Figura 7.122 Corrientes de Foucault producidas por un campo magnético.



Figura 7.123 Diagrama de bloques de un sensor de proximidad inductivo.

Si a la bobina se le acerca un material ferromagnético de permeabilidad μ_r , las líneas de fuerza lo cortan y hacen que se induzcan en él corrientes I de Eddy o Foucault (Figura 7.122). Dichas corrientes crean, a su vez, un campo magnético que se opone al generado por la bobina y hacen que varíe la inductancia L de la misma. En este efecto basan su funcionamiento los sensores inductivos de proximidad cuyo diagrama de bloques se representa en la figura 7.123, en la que se puede observar que están formados por:

- Una bobina, que constituye el elemento sensor, y un oscilador LC del que forma parte.
- Un rectificador con un filtro paso-bajo que proporciona una tensión continua a partir de la señal alterna de la salida del oscilador.
- Un comparador que proporciona dos tensiones distintas a la salida, una en presencia de objeto y otra en ausencia del mismo. Dicho comparador posee histéresis para garantizar que la salida del sensor no oscila cuando el objeto está situado justamente en la zona en la que comienza a ser detectado por el sensor.
- Una etapa de salida que puede ser cualquiera de las descritas en el apartado 7.4.1.3.

El circuito LC del oscilador está calculado para que oscile en ausencia de material ferromagnético. Por ello, cuando una pieza de dicho material se acerca a la bobina se modifica el campo electromagnético (Figura 7.124a) y varía el valor de L . Como consecuencia de ello el circuito deja de oscilar, tal como se indica en la figura 7.124b. Si el material del objeto es paramagnético, la detección se produce a menor distancia que en el caso de los ferromagnéticos y si es diamagnético disminuye todavía más, e incluso puede llegar a no detectarse.

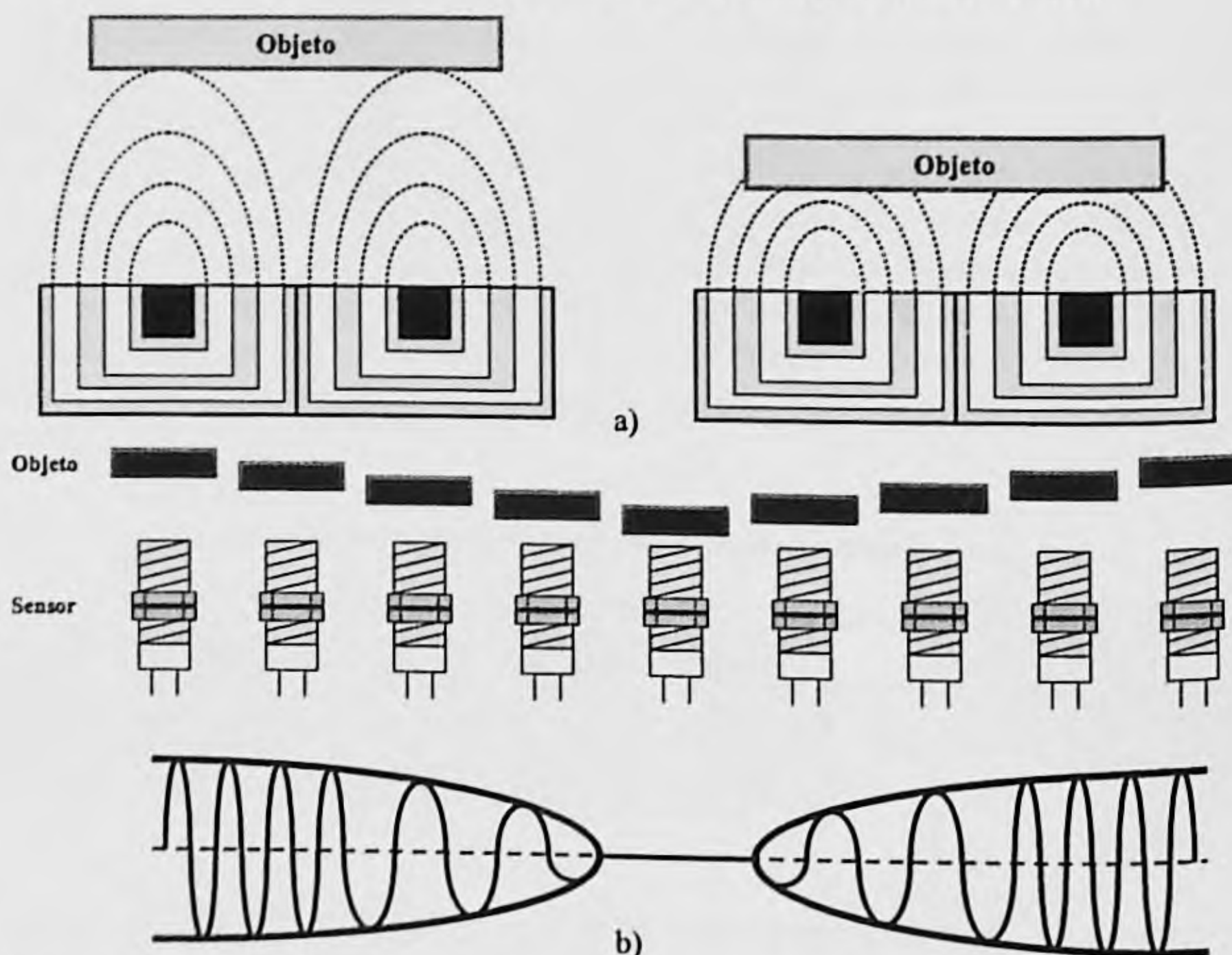


Figura 7.124 Comportamiento de un sensor de proximidad inductivo: a) Modificación del campo magnético al acercarse un objeto; b) Respuesta del circuito oscilador.

7.7.2.3.5.2 Tipos de sensores inductivos de proximidad

Los sensores inductivos de proximidad se pueden clasificar de acuerdo con el formato externo y con la forma constructiva.

Según el formato externo los sensores inductivos de proximidad pueden ser cilíndricos (Figura 7.125) o cúbicos (Figura 7.126). En ambos casos la forma de operación es la misma.

Según la forma constructiva los sensores inductivos de proximidad pueden ser enrasables y no enrasables que se describen a continuación.



Figura 7.125 Sensor inductivo cilíndrico 3RG4014 (cortesía de Siemens).



Figura 7.126 Sensor inductivo de sección cuadrada (cúbico) 3RG4648 (cortesía de Siemens).

- **Sensores enrasables**

Los sensores enrasables o apantallados (*Shielded*) son sensores inductivos de proximidad en los que el cuerpo metálico que rodea a los distintos bloques que constituyen el sensor, se prolonga hasta la bobina sensora. Esto evita la dispersión del flujo electromagnético y lo concentra en la parte frontal tal como se indica en la figura 7.127. De esta forma se anula prácticamente el efecto de los metales circundantes y las interferencias mutuas cuando se instalan varios sensores inductivos de proximidad contiguos. La figura 7.128 muestra un sensor de proximidad inductivo enrasable de la familia 3RG4014.

- **Sensores no enrasables**

Los sensores no enrasables son sensores cuyo recubrimiento metálico externo no llega hasta el borde de la cabeza sensora, es decir no están apantallados (*Unshielded*), lo que

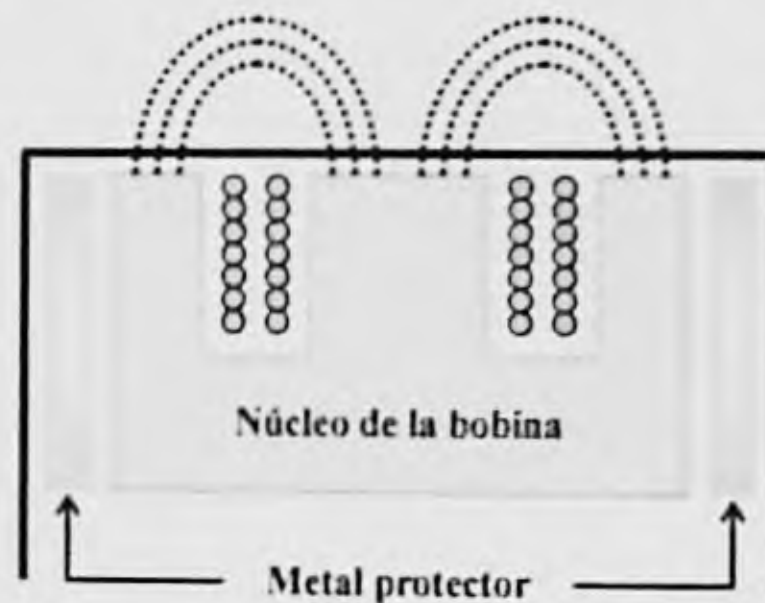


Figura 7.127 Flujo magnético generado por un sensor de proximidad inductivo enrasable.



Figura 7.128 Sensor de proximidad inductivo enrasable 3RG4014 (cortesía de Siemens).

hace que el flujo se disperse por los laterales de la misma tal como se indica en la figura 7.129. Debido a ello el sensor no se puede enrasar con el soporte metálico que lo sustenta porque lo detecta y estaría siempre activado. La figura 7.130 muestra un sensor de este tipo de la familia 3RG4024.

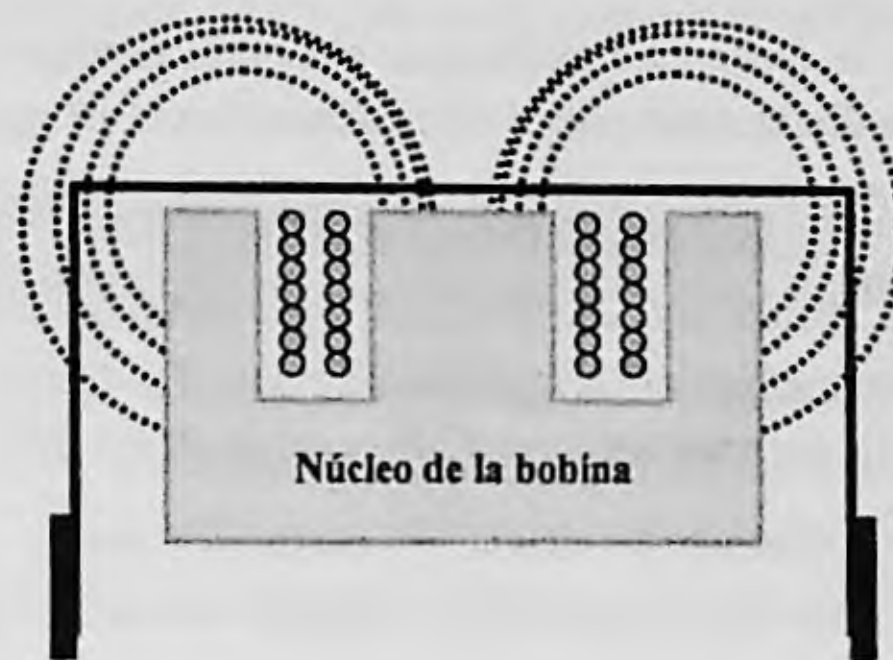


Figura 7.129 Flujo magnético generado por un sensor de proximidad inductivo no enrasable.



Figura 7.130 Sensor de proximidad inductivo no enrasable 3RG4024 (cortesía de Siemens).

Se debe tener en cuenta que, en un sensor enrasable el flujo magnético es bastante restringido, lo que provoca una disminución de la distancia de detección. Para detectar el objeto a la misma distancia con un sensor enrasable que con uno no enrasable, la superficie de detección del primero debe ser aproximadamente el doble que la del segundo.

7.7.2.3.5.3 Campo de trabajo

El campo de trabajo o rango de operación (*Sensing range*) es la zona, situada en las proximidades de la cara activa del sensor, en la que el campo electromagnético creado por la bobina es suficientemente intenso para que el fabricante garantice que el sensor detecta el objeto normalizado o patrón, que es un cuadrado de acero ST37 de un milímetro de espesor y un lado de longitud igual al diámetro de la cara sensible del sensor (la mayoría de los que se utilizan en la industria son cilíndricos) o a tres veces la distancia de detección nominal del sensor (el mayor de ambos).

La figura 7.131a muestra el campo de trabajo típico de un sensor inductivo. En ella se puede observar que:

- Es un cono cuya base es aproximadamente igual a la superficie de la cara activa del sensor.
- El campo de trabajo se reduce a medida que el objeto se aleja de la cara activa del sensor hasta alcanzar la distancia de detección nominal que es igual a la altura del cono.
- La superficie de la cara activa del sensor que debe cubrir el objeto aumenta al hacerlo la distancia entre ambos.

Los fabricantes suelen representar el campo de trabajo en dos dimensiones, tal como se indica en la figura 7.131b cuyo eje de ordenadas coincide con la altura del cono de la figura 7.131a.

Se define por lo tanto la distancia de detección nominal (*Sensing distance*) como la distancia teórica máxima a la que el sensor de proximidad detecta el objeto (*Target*) normalizado o patrón.

Aunque es posible que el sensor detecte al objeto patrón a una distancia superior a la máxima, el fabricante no lo garantiza.

La distancia de detección varía con el tamaño del objeto. Debido a ello, los fabricantes suelen suministrar factores de corrección para calcular la distancia de detección de objetos más pequeños que el patrón. La tabla 7.13 muestra el factor de corrección que se debe utilizar para diferentes reducciones del tamaño del objeto patrón, tanto en el caso de los sensores enrasables como de los no enrasables. La distancia de detección se calcula como el producto de dicho factor por la distancia nominal de detección, suministrada por el fabricante para el objeto patrón.

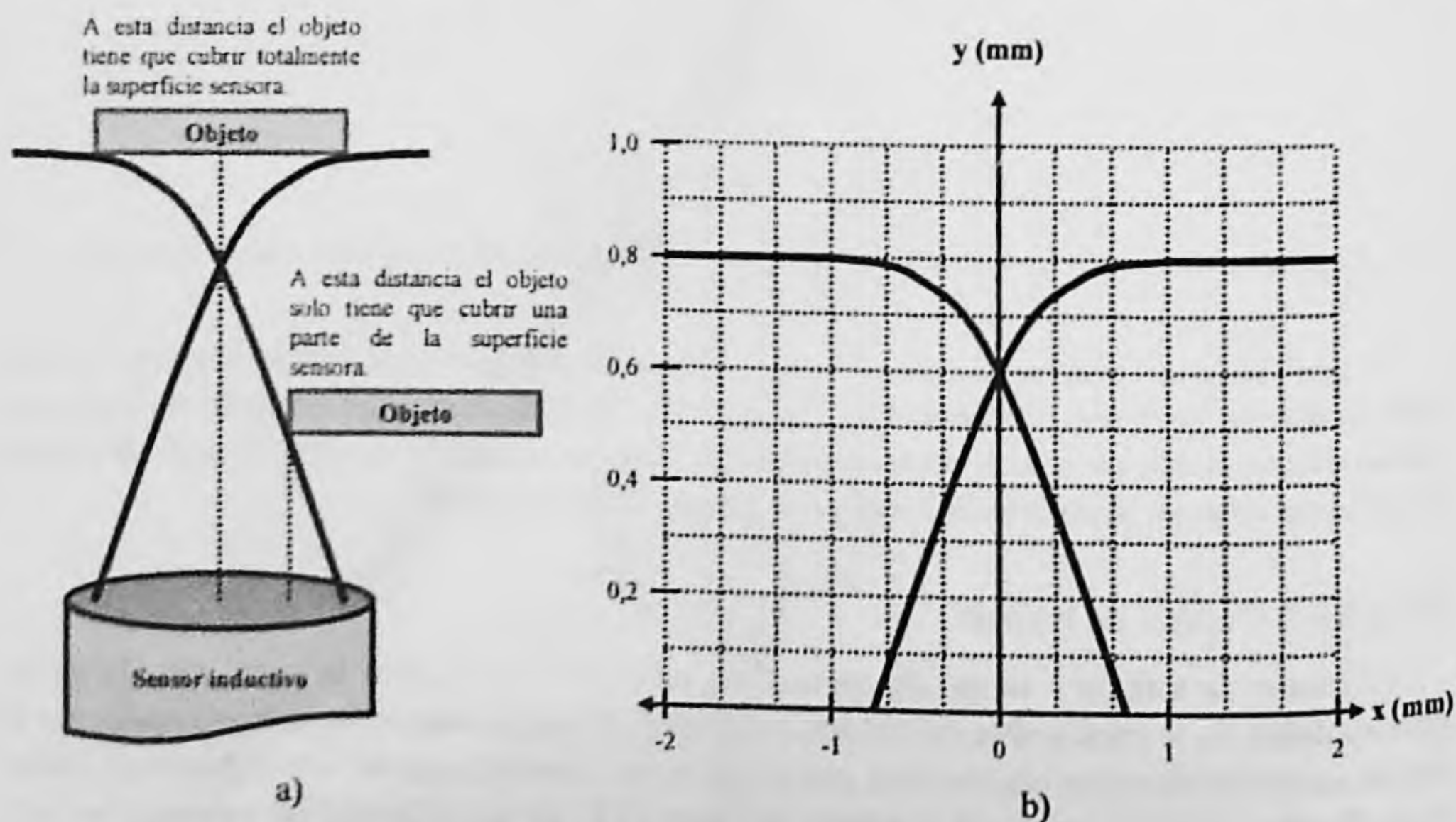


Figura 7.131 Campo de trabajo de un sensor de proximidad inductivo (cortesía de Siemens).

Frecuentemente los fabricantes suministran estos datos gráficamente para distintos tipos de materiales, de lo que es un ejemplo la figura 7.132.

Los fabricantes suelen proporcionar también información sobre el factor de corrección que se debe aplicar al valor de la distancia nominal de detección en función del espesor del objeto para diferentes tipos de material, de lo que es un ejemplo la gráfica de la figura 7.133. En ella se puede observar que:

Reducción del objeto	Factor de corrección	
	Sensor enrasable	Sensor no enrasable
25%	0,56	0,50
50%	0,83	0,73
75%	0,92	0,90
100%	1,00	1,00

Tabla 7.13 Factores de corrección de los sensores de proximidad inductivos para diferentes reducciones del tamaño en relación con el objeto patrón.

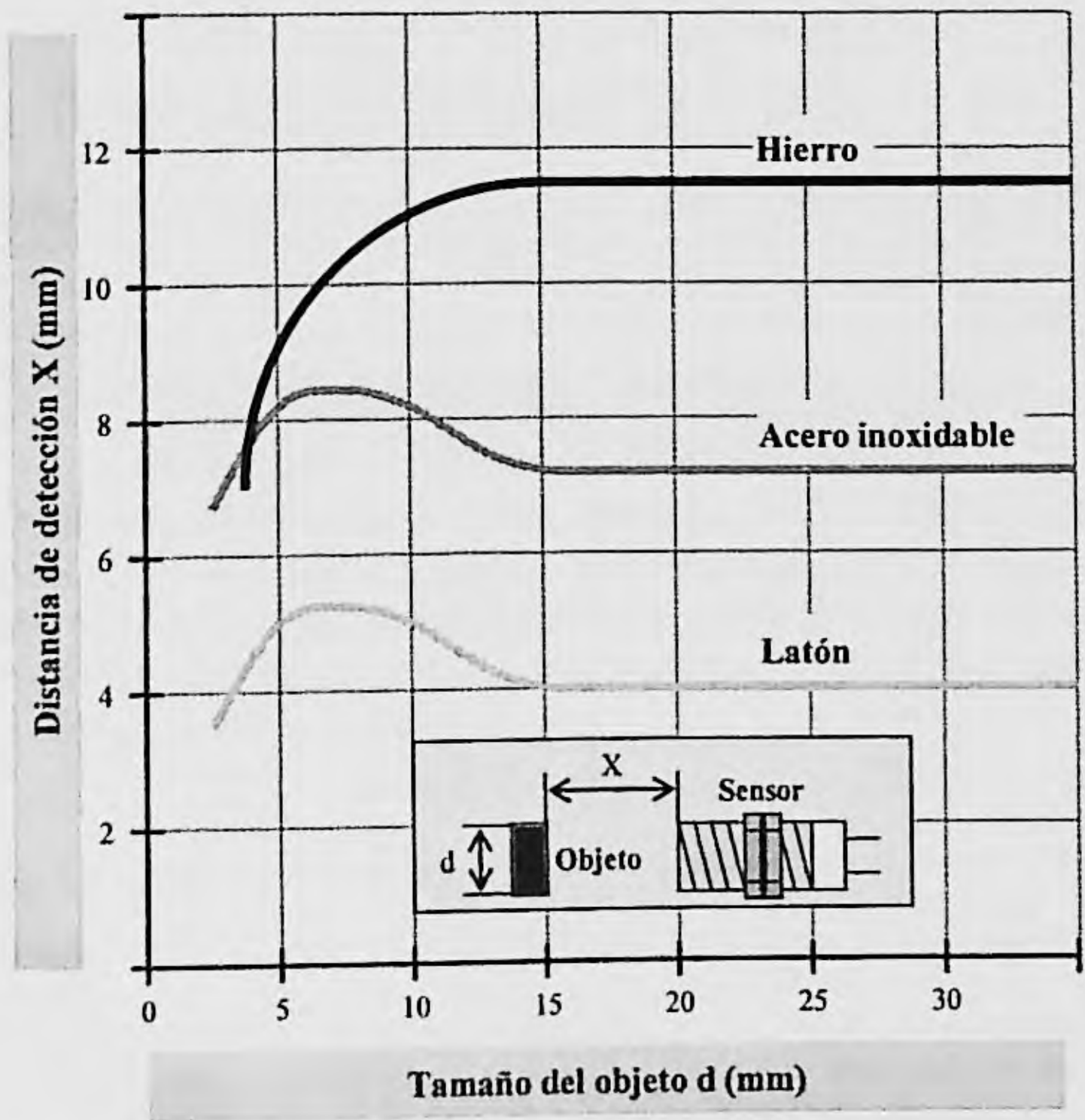


Figura 7.132 Relación entre el tamaño del objeto y la distancia nominal de detección para distintos materiales.

- La distancia de detección aumenta al disminuir el espesor del objeto por debajo de un determinado valor del mismo, debido al efecto pelicular (*skin effect*) [SIEM 00].
- A partir de dicho valor del espesor del objeto la distancia de detección permanece invariable.

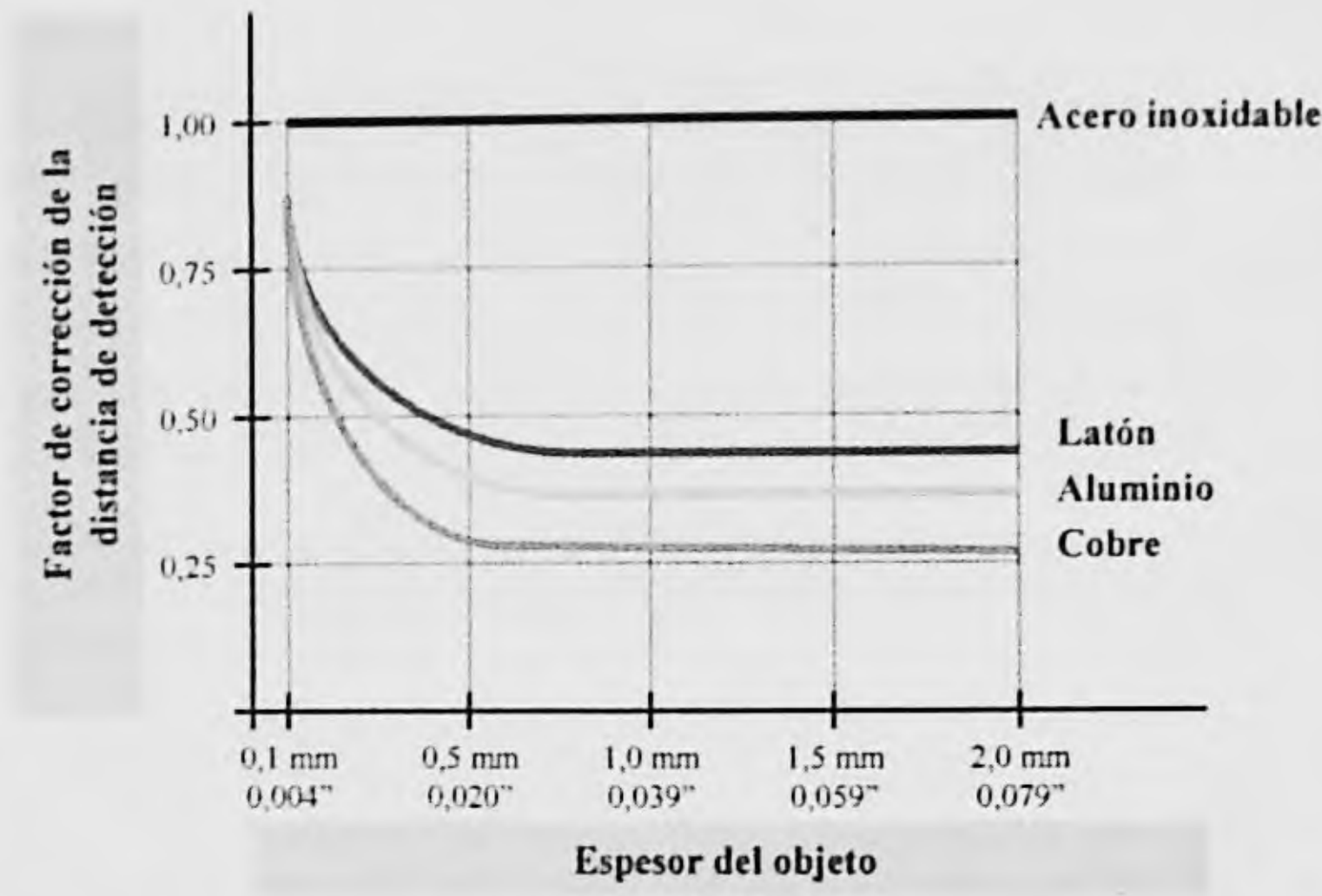


Figura 7.133 Relación entre el factor de corrección que se debe aplicar a la distancia nominal de detección de un sensor de proximidad inductivo y el espesor del objeto para diferentes tipos de materiales.

Material	Factor de corrección	
	Enrasable	No enrasable
Acero ST 37	1,00	1,00
Lámina de aluminio	0,90	1,00
acero inoxidable	0,70	0,80
Bronce	0,40	0,50
Aluminio	0,35	0,45
Cobre	0,30	0,40

Tabla 7.14 Factores de corrección de los sensores de proximidad inductivos según el tipo de material del objeto.

Los fabricantes suelen incluir en sus hojas de características, tablas con el factor de corrección que se debe utilizar para calcular la distancia de detección nominal para diferentes tipos de materiales. Un ejemplo se muestra en la tabla 7.14, a partir de la cual se puede obtener la distancia de detección mediante la ecuación:

$$Sn_{(Material)} = Sn_{(Acero...ST...37)} * Factor_{(Material)}$$

7.7.2.3.5.4 Características técnicas

Además de las características técnicas generales de los sensores todo-nada, indicadas en el apartado 7.3, y las específicas de los sensores de proximidad sin contacto, indicadas en el apartado 7.7.3, existen las siguientes características específicas de los sensores inductivos de proximidad:

- **Distancia de detección**

Tal como se indica en el apartado anterior, el fabricante suele especificar la distancia S_n de conmutación o detección nominal (*Sensing distance*), que es la distancia teórica máxima a la que el sensor de proximidad detecta el objeto normalizado o patrón. La distancia de detección varía con las tolerancias de fabricación, la temperatura ambiente, la tensión de alimentación (márgenes de variación que el fabricante suele indicar en % de la distancia de detección nominal), etc. Por ello existen otras distancias de conmutación [BART 03] que se relacionan a continuación:

- **La distancia S_r de detección real (*Effective operating distance*)**

Es la distancia de conmutación de un sensor de proximidad inductivo a la que su salida cambia de nivel al aproximarle el objeto cuando se le aplica la tensión de alimentación nominal, a la temperatura ambiente de $23^\circ\text{C} \pm 0,5^\circ\text{C}$ y con el objeto normalizado. Está comprendida en el rango de $\pm 10\%$ de la distancia nominal:

$$0.9 S_n \leq S_r \leq 1.1 S_n$$

- **La distancia S_u de detección útil (*Useful switching distance*)**

Es la distancia de conmutación medida con el objeto patrón, en condiciones reales de funcionamiento. Esta distancia debe mantenerse en el rango de $\pm 10\%$ de S_r :

$$0.9 S_r \leq S_u \leq 1.1 S_r \Rightarrow 0.81 S_n \leq S_u \leq 1.21 S_n$$

- **La distancia S_a de detección de trabajo (*Guaranteed operating distance*)**

Es el campo de funcionamiento seguro del sensor, dentro del cual se garantiza la detección del objeto patrón, independientemente de las variaciones de la tensión de alimentación, de la temperatura o de las tolerancias de fabricación. Está comprendida entre 0 y el 81% de la distancia nominal:

$$0 \leq S_a \leq 0.81 S_n$$

- **Velocidad del objeto**

La velocidad v a la que el objeto a detectar pasa a través de la zona de trabajo del sensor no debe superar un valor máximo para que el oscilador y el resto del circuito electrónico puedan actuar adecuadamente tal como se indica gráficamente en la figura 7.134. El valor máximo de v viene dado por la expresión algebraica:

$$\frac{1}{f} < \frac{l_1 + l_2}{v} \Rightarrow v < \frac{(l_1 + l_2)}{f}$$

en la que l_1 es la longitud del objeto, l_2 es la distancia que debe recorrer el objeto (según el punto en el que entre en la zona de trabajo, es decir, a la distancia de trabajo) y f representa la máxima frecuencia de operación del circuito electrónico del sensor.

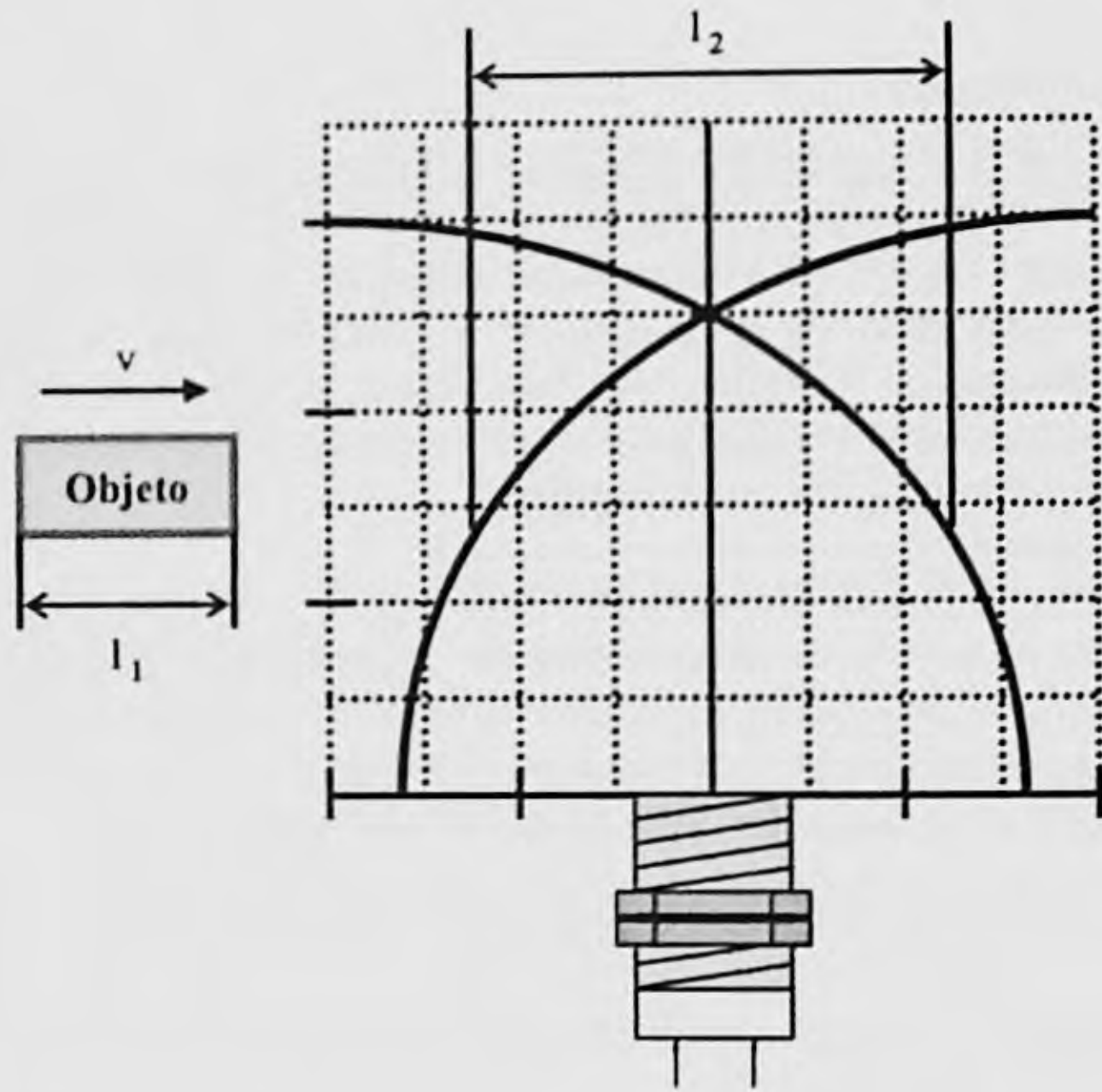


Figura 7.134 Factores que influyen en la máxima velocidad a la que se puede mover un objeto metálico para ser detectado por un sensor inductivo de proximidad.

• **Distancia mínima entre dos objetos consecutivos**

La distancia entre dos objetos consecutivos, debe tener un valor mínimo para que al sensor le de tiempo a conmutar entre ambos. El valor mínimo de dicha distancia está limitado por:

- El campo electromagnético creado por la bobina, que suele exceder el campo de trabajo dado por el fabricante, tal como se puede observar en la figura 7.135. Debido a ello, si la distancia entre los objetos es muy pequeña, es posible que un objeto no haya salido de la zona de detección cuando el siguiente ya esté entrando en ella, tal como se indica en la figura 7.136. En ese caso la salida del sensor no cambia de nivel y los dos objetos se detectan como si fuera uno solo.
- El retardo a la desactivación, que es el tiempo que transcurre desde que el objeto sale de la zona activa hasta que la señal de salida cambia de nivel (Figura 7.137).

Para evitar esta situación, los fabricantes recomiendan que la distancia entre dos objetos consecutivos, sea como mínimo igual a 1.5 veces el diámetro de la cara activa del sensor, si es cilíndrico y a 1.5 veces el lado si es rectangular, tal como se indica en la figura 7.138.

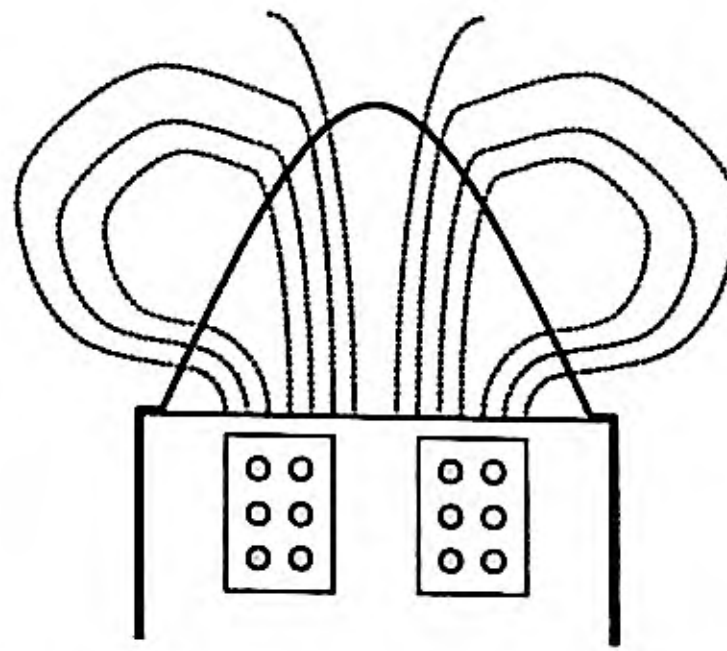


Figura 7.135 Campo electromagnético real creado por la bobina de un sensor inductivo de proximidad.

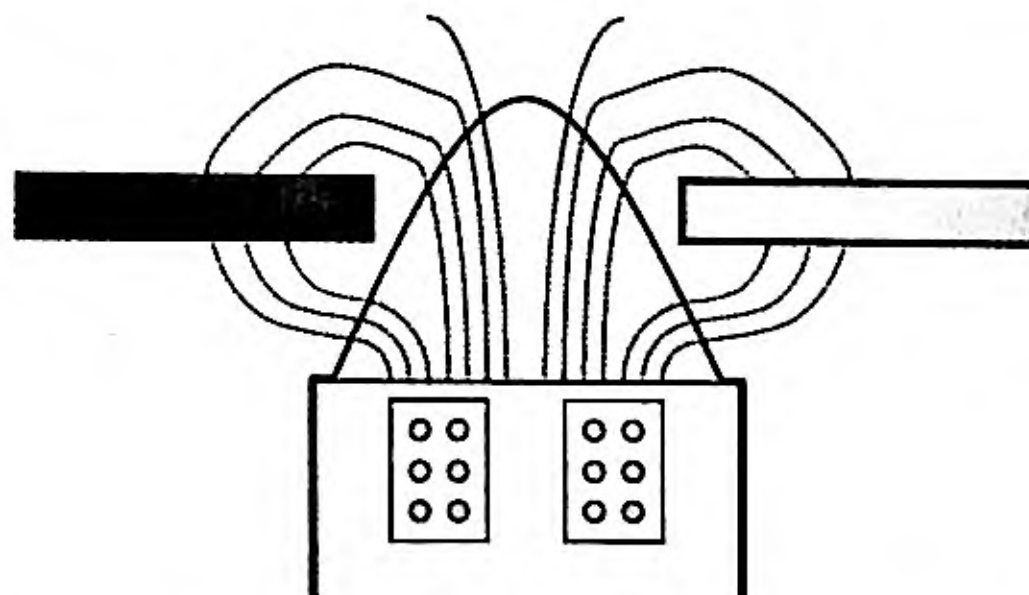


Figura 7. 136 Situación en la que un objeto no ha salido de la zona de detección inestable cuando el siguiente ya está entrando en ella.

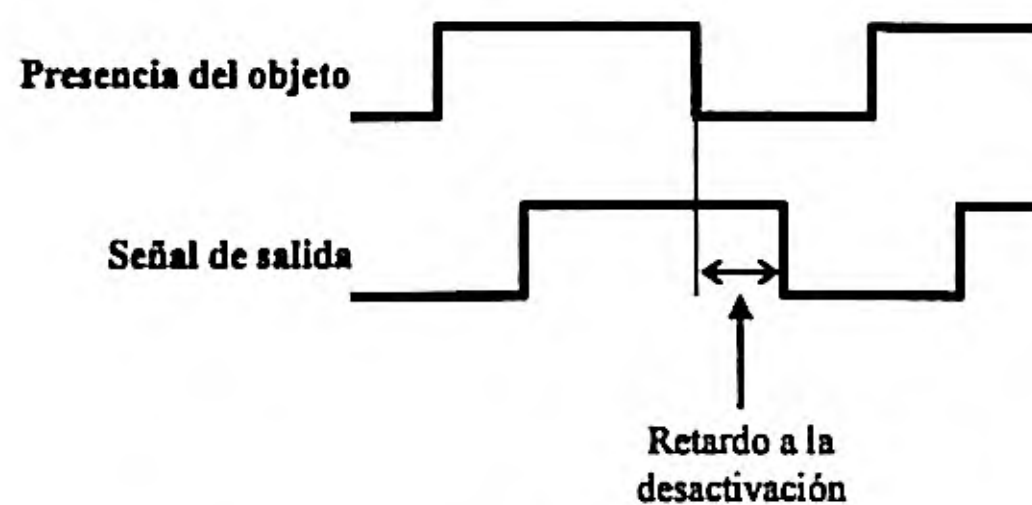


Figura 7.137 Retardo a la desactivación de la salida de un sensor inductivo de proximidad.

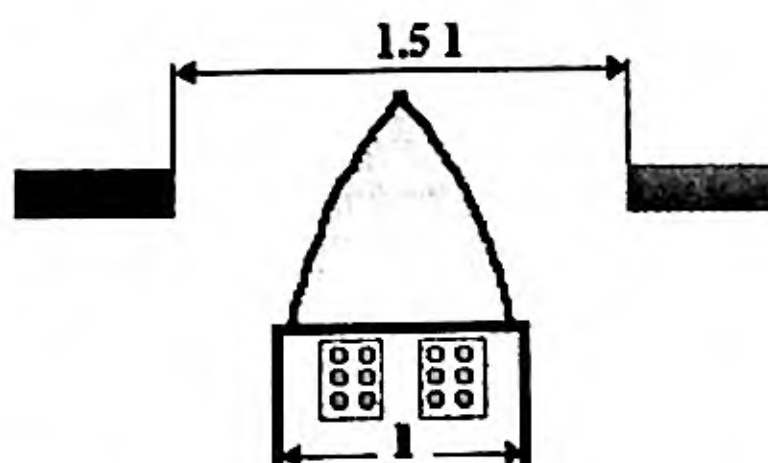


Figura 7.138 Distancia mínima que debe haber entre dos objetos consecutivos que deben ser detectados por un sensor inductivo de proximidad.

• **Histéresis**

La histéresis o distancia diferencial es la diferencia entre la posición de actuación, cuando se acerca el objeto (Figura 7.139a), y la posición de desactivación o reposición, cuando se aleja (Figura 7.139b), tal como se indica en la figura 7.139. Se expresa, frecuentemente, en % de la distancia nominal. A veces los fabricantes la indican gráficamente tal como se muestra en la figura 7.140. De acuerdo con esta definición, en la figura 7.141 se muestran las distancias de detección analizadas anteriormente. En dicha figura, H representa el valor de la histéresis.

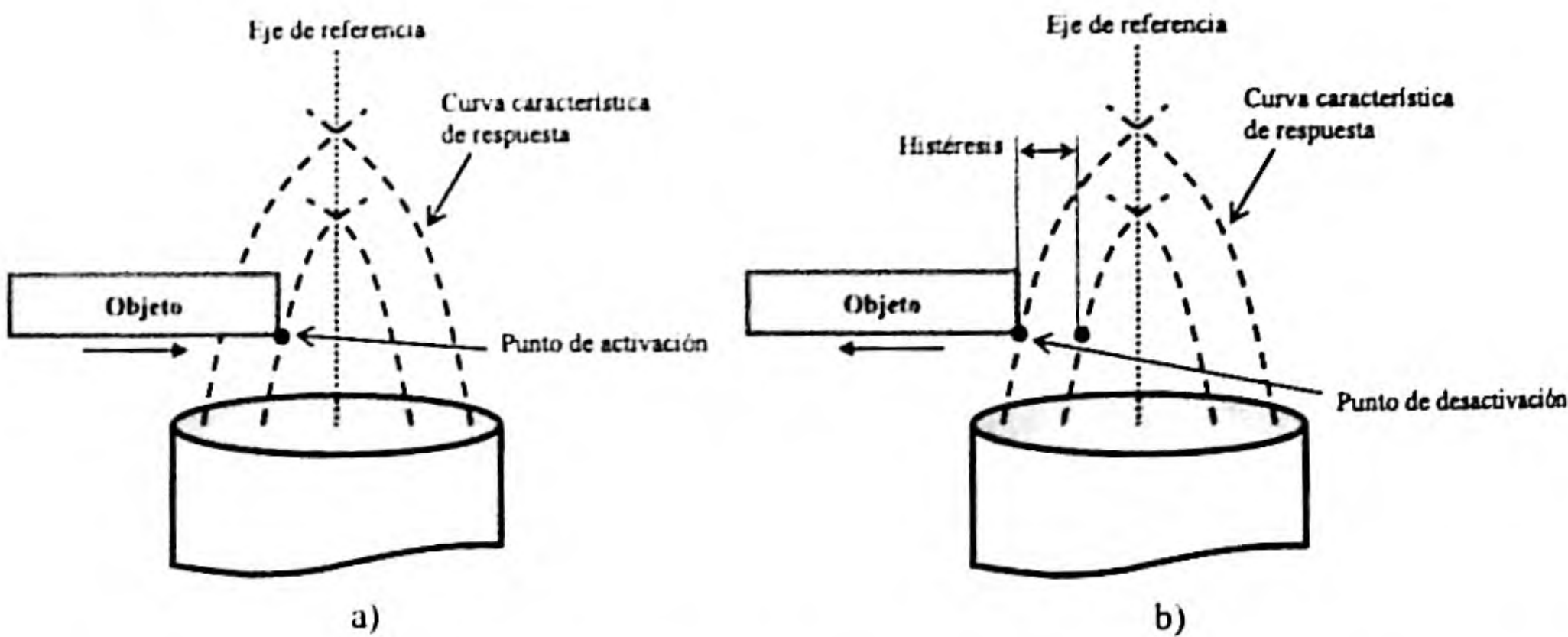


Figura 7.139 *Histéresis en la activación y desactivación de la salida de un sensor inductivo de proximidad.*

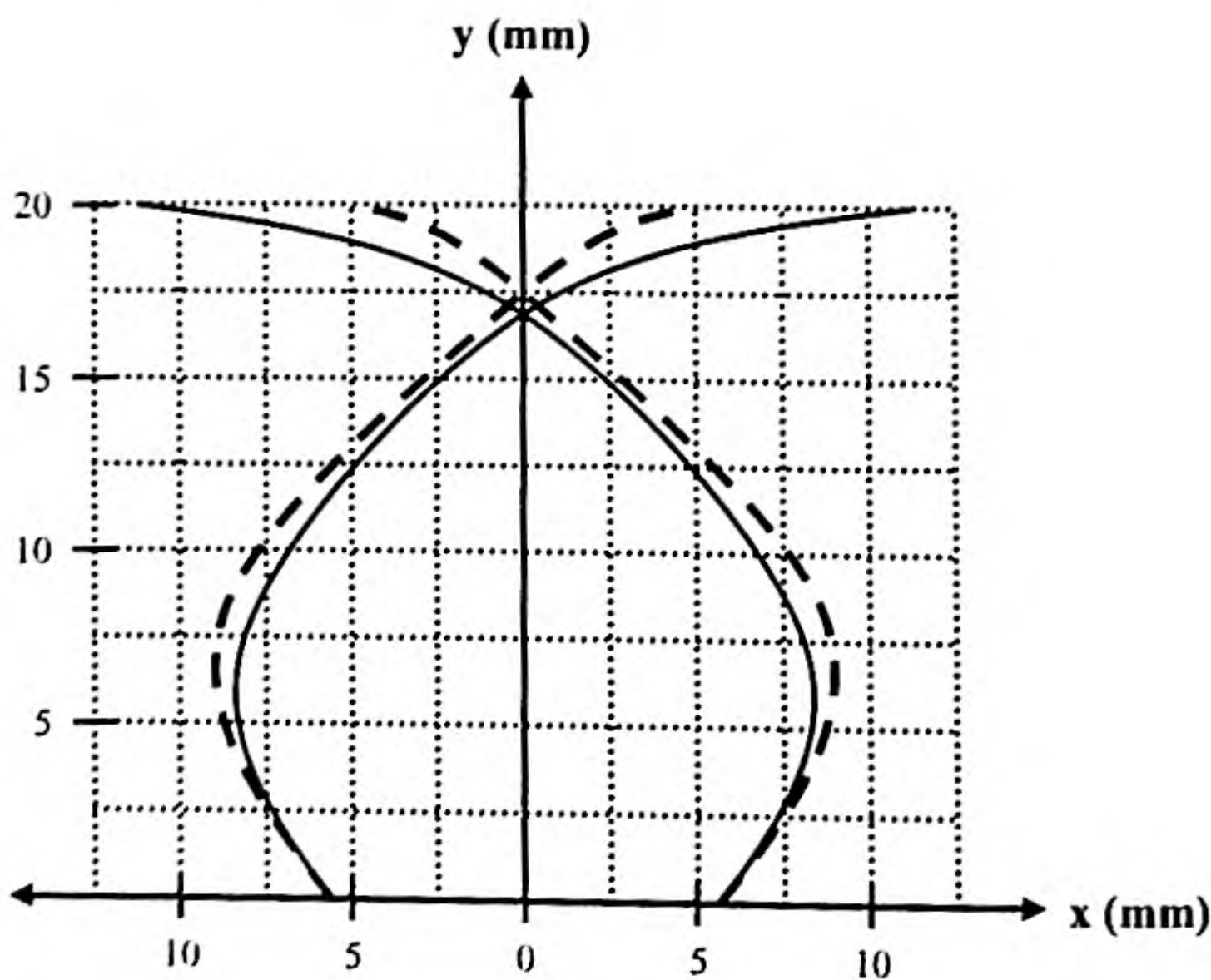


Figura 7.140 *Representación gráfica de la histéresis de un sensor inductivo de proximidad.*

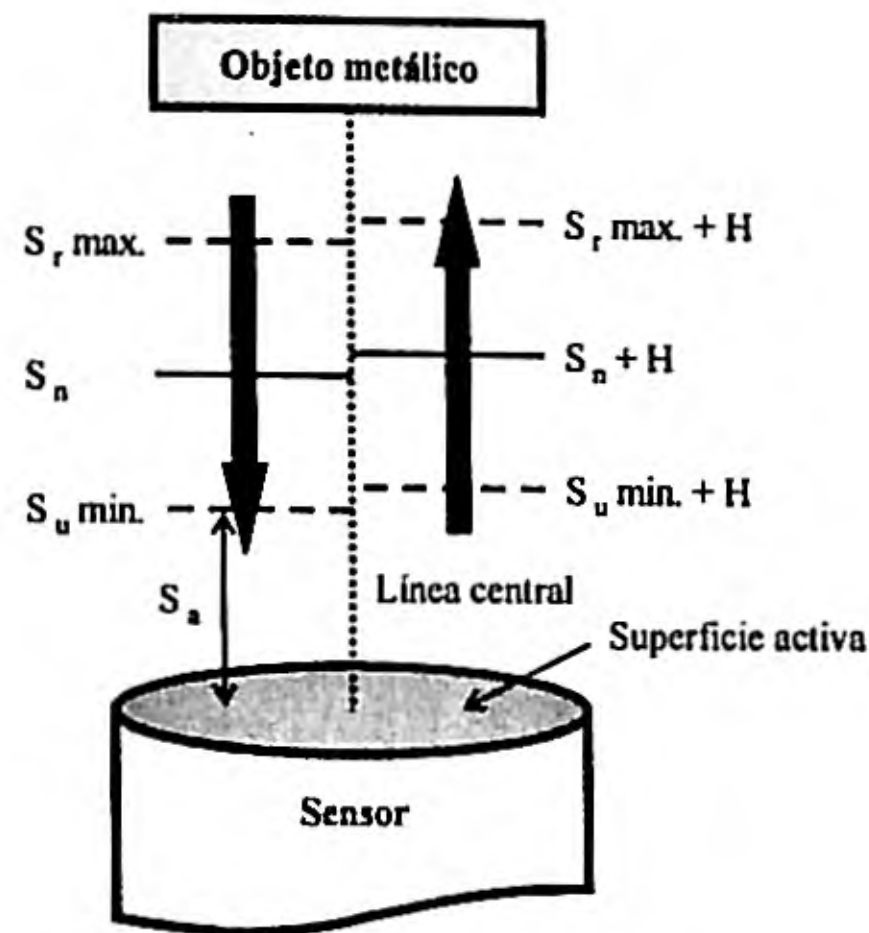


Figura 7.141 Distancias de detección de un sensor inductivo de proximidad.

- **Frecuencia de conmutación**

Para medir la frecuencia de conmutación se utiliza un sistema formado por un disco de material aislante, con salientes de forma cuadrada de lado x , sobre cada uno de los cuales se coloca una pletina del objeto estándar, de las mismas dimensiones que el saliente. A su vez la distancia entre salientes debe ser el doble que el lado del saliente. El sensor se pone a la mitad de la distancia de detección nominal S_n . La figura 7.142 muestra el sistema utilizado por los fabricantes para determinar la frecuencia de conmutación del sensor. Esta frecuencia se calcula como el inverso de la suma del tiempo que el sensor está activado T_{ON} y el que está desactivado T_{OFF} .

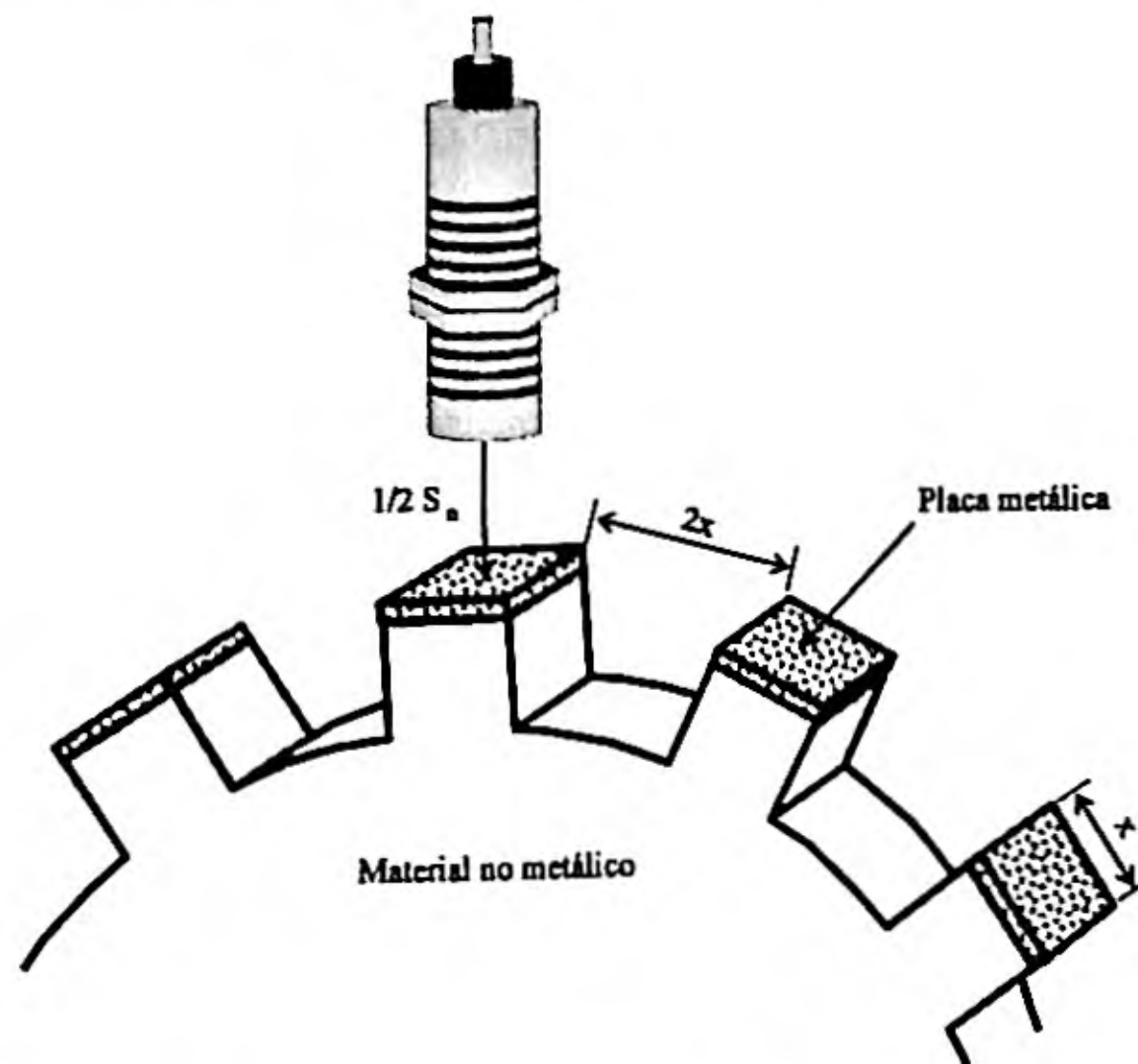


Figura 7.142 Sistema utilizado por los fabricantes para determinar la frecuencia de conmutación de un sensor inductivo de proximidad.

7.7.2.3.5.5 Normas de instalación

En el montaje de los sensores inductivos de proximidad hay que tener en cuenta un conjunto de características relacionadas con el entorno de trabajo [AENO 105], entre las que cabe citar:

- Las interferencias mutuas.
- Los metales circundantes.
- Las interferencias electromagnéticas.

A continuación se analiza cada una de ellas.

• Interferencias mutuas

Las interferencias mutuas se producen cuando dos o más sensores inductivos de proximidad se montan demasiado cerca, ya sea enfrentados o situados lateralmente. El efecto hace variar la distancia de detección o genera rebotes. Para evitarlo se deben montar los sensores con una distancia de separación mínima entre ellos, tal como se indica en la figura 7.143a para sensores enfrentados y en la figura 7.143b para sensores próximos entre sí lateralmente. En las citadas figuras se establecen recomendaciones, tanto para sensores enrasables como no enrasables. Dichas distancias las proporcionan los fabricantes en sus catálogos. Si las limitaciones de espacio impiden separar los sensores la distancia necesaria, se pueden alternar detectores con diferentes frecuencias de oscilación.

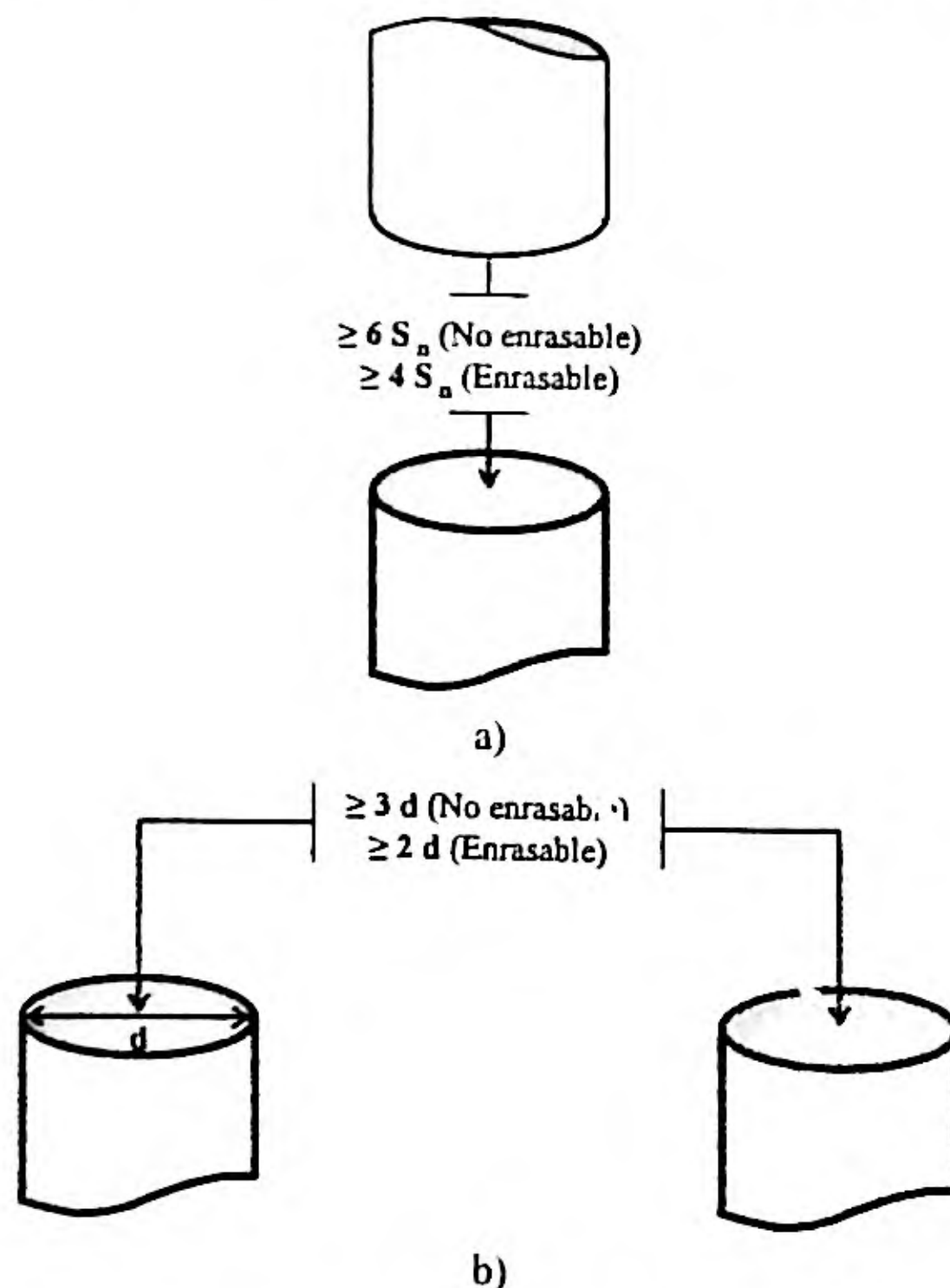


Figura 7.143 Distancia de separación mínima entre dos sensores de proximidad inductivos: a) Cuando los sensores están enfrentados; b) Cuando los sensores están colocados lateralmente.

- **Metales circundantes**

Para no alterar las características del sensor, las superficies metálicas próximas deben estar situados a una distancia mínima establecida por el fabricante. Para sensores cilíndricos dicha distancia suele ser del orden de magnitud que se indica en la figura 7.144 para sensores no enrasables. Para sensores enrasables se recomienda una separación, entre el sensor y el material metálico del fondo, de tres veces la distancia nominal.

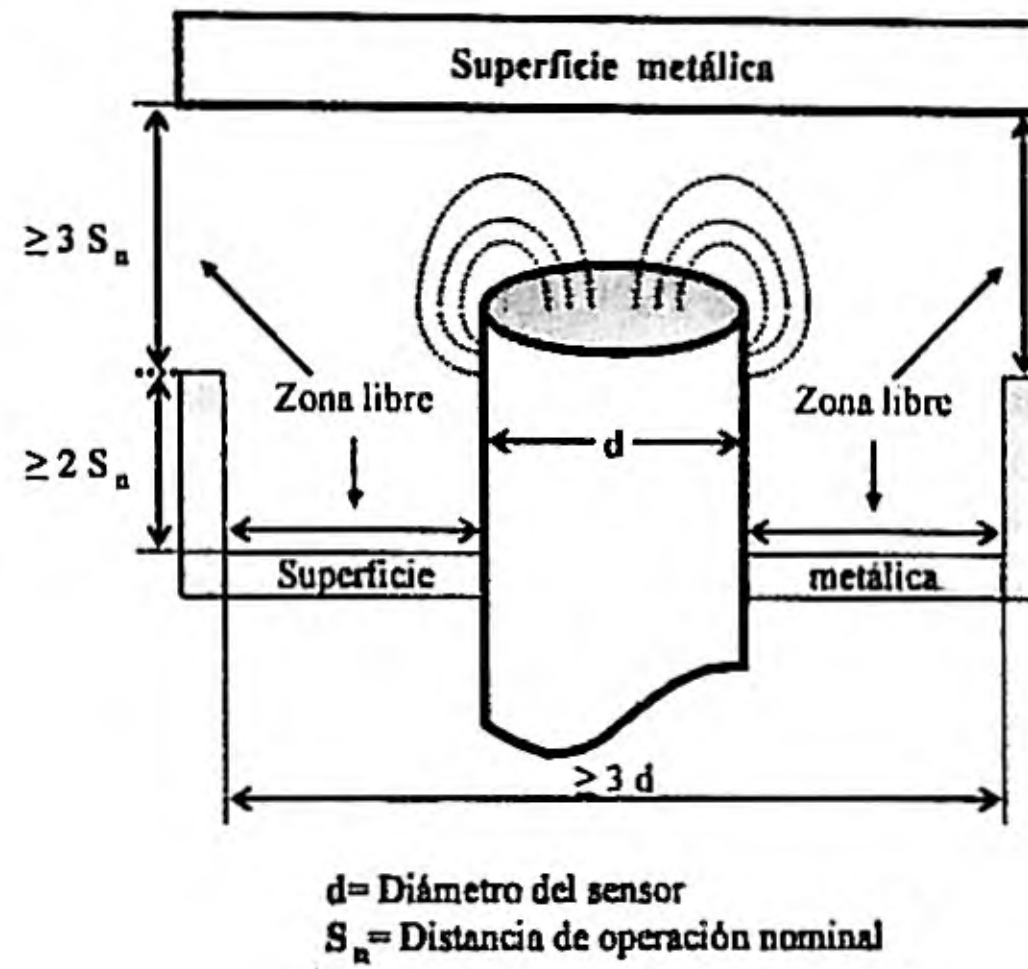


Figura 7.144 Distancias mínimas a las que debe estar una superficie metálica circundante de un sensor inductivo de proximidad no enrasable.

- **Interferencias electromagnéticas**

Los sensores inductivos de proximidad son sensibles a los campos electromagnéticos existentes en el entorno en el que se instalan. Debido a ello, para reducir el riesgo de un mal funcionamiento, se deben tomar las siguientes precauciones:

- Conectar el sensor mediante un cable apantallado.
- Mantener separados los cables del sensor de otros de potencia.
- Reducir al máximo la emisión de interferencias en el entorno del sensor.

7.7.2.3.6 Sensores capacitivos de proximidad

7.7.2.3.6.1 Conceptos generales

Los sensores capacitivos de proximidad (*Capacitive proximity sensors*) basan su funcionamiento en la interacción que se produce entre el campo electrostático que ellos mismos generan y el objeto a detectar. Su principio de funcionamiento es análogo al de los inductivos y se diferencian de ellos en que el elemento sensor es un condensador.

La capacidad C de un condensador de placas paralelas viene dada por la expresión:

$$C = \epsilon \frac{A}{d} = \epsilon_0 \cdot \epsilon_r \frac{A}{d}$$

en la que A representa el área de las placas del condensador, d la separación entre las mismas, ϵ_0 la permitividad del vacío ($8'854 \cdot 10^{-12}$ [F/m]) y ϵ_r la permitividad relativa del material utilizado como dieléctrico, que indica el número de veces que la permitividad de dicho material es mayor que la permitividad del vacío.

La permitividad representa la capacidad que tiene un material para polarizarse ante la presencia de un campo eléctrico y también es conocida como constante dieléctrica.

Si las dos caras del condensador están en el mismo plano, en lugar de estar enfrentadas, se forma un condensador abierto. Si además dichas caras, A_1 y A_2 , son concéntricas, el condensador tiene la forma representada en la figura 7.145. En este condensador las líneas del campo eléctrico se cierran entre ambas caras. Si se supone que A_1 es la cara positiva y A_2 la negativa, el campo eléctrico tiene la forma que se indica en la figura 7.146.

Tal como se representa en la figura 7.147, cuando en el campo eléctrico creado por el condensador abierto entra un objeto, la capacidad del condensador se incrementa por el efecto que el mismo provoca en el campo eléctrico. Esta variación de la capacidad del condensador permite que se pueda detectar el objeto mediante un circuito oscilador LC.



Figura 7. 145 Condensador de placas concéntricas.

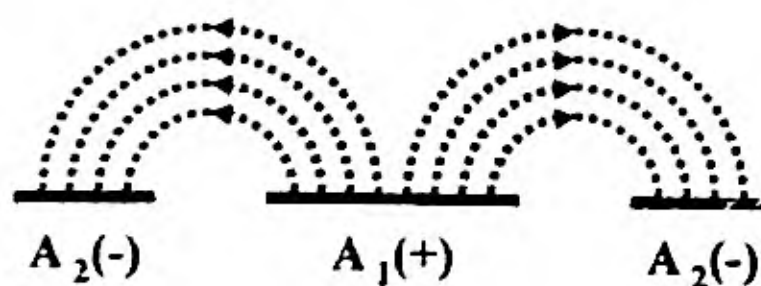


Figura 7.146 Campo eléctrico entre las caras de un condensador de placas concéntricas.

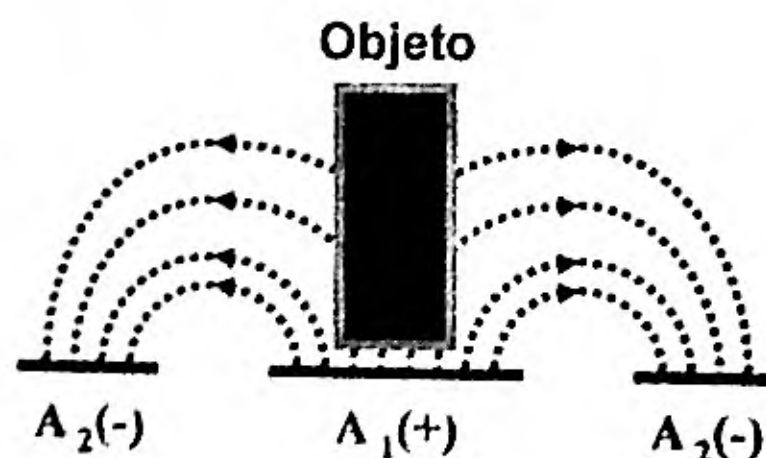


Figura 7. 147 Efecto producido por un objeto que penetra en el campo eléctrico creado por el condensador abierto.

La figura 7.148 muestra el diagrama de bloques de un sensor capacitivo de proximidad, en el que se indican las capacidades más significativas que intervienen en el proceso de medida, que son las siguientes [BAXT 97]:

- C_A : Capacidad entre las placas del condensador abierto.
- C_B : Capacidad entre la placa A_1 del condensador abierto y el objeto.
- C_C : Capacidad entre la placa A_2 del condensador abierto y el objeto.
- C_D : Capacidad entre el objeto y la masa del circuito.

El diagrama de bloques de un sensor de proximidad capacitivo se representa en la figura 7.149, en la que se puede observar que está constituido por:

- El condensador, equivalente a los analizados anteriormente, que constituye el elemento sensor que forma parte de un oscilador LC.
- Un rectificador con filtro, un comparador con histéresis y una etapa de salida, idénticas a las de los sensores inductivos de proximidad.

El oscilador LC no oscila en ausencia de objeto y pasa a hacerlo cuando el objeto se aproxima al sensor (Figura 7.150a), tal como se indica en la figura 7.150b, porque aumenta la capacidad existente entre las placas A_1 y A_2 .

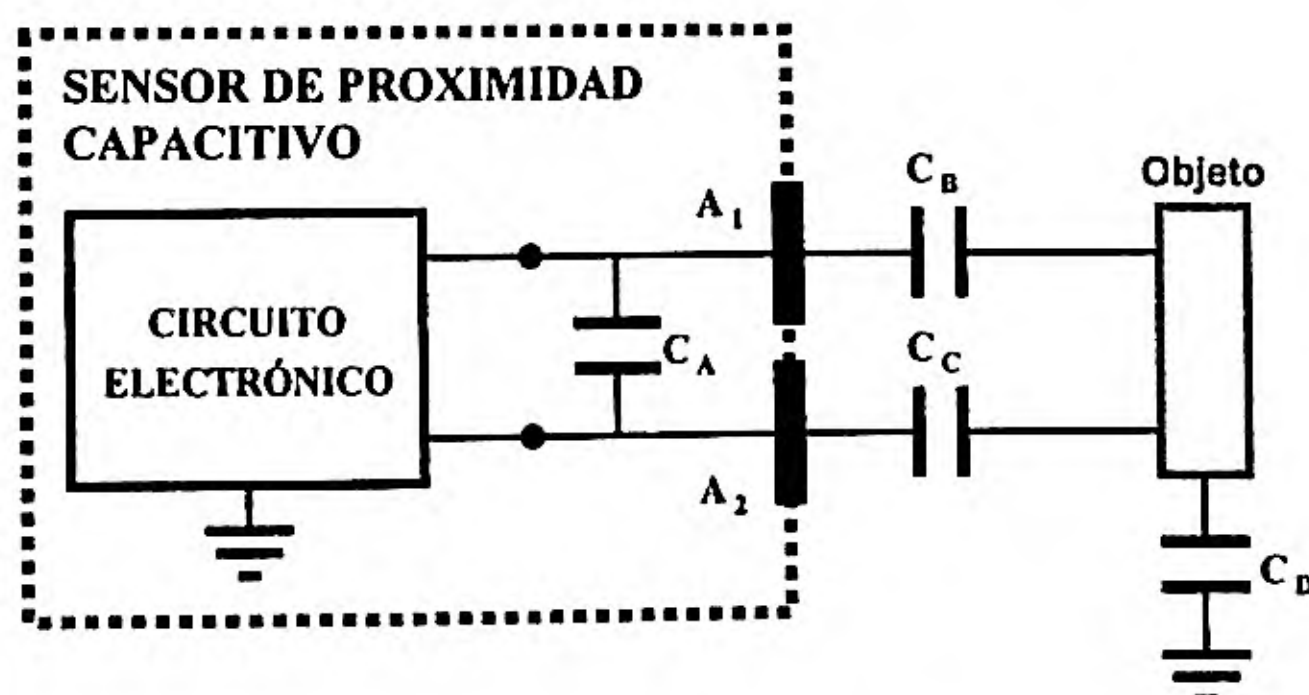


Figura 7.148 Diagrama de bloques de un sensor capacitivo de proximidad en el que se indican las diferentes capacidades que intervienen en el proceso de detección de un objeto.

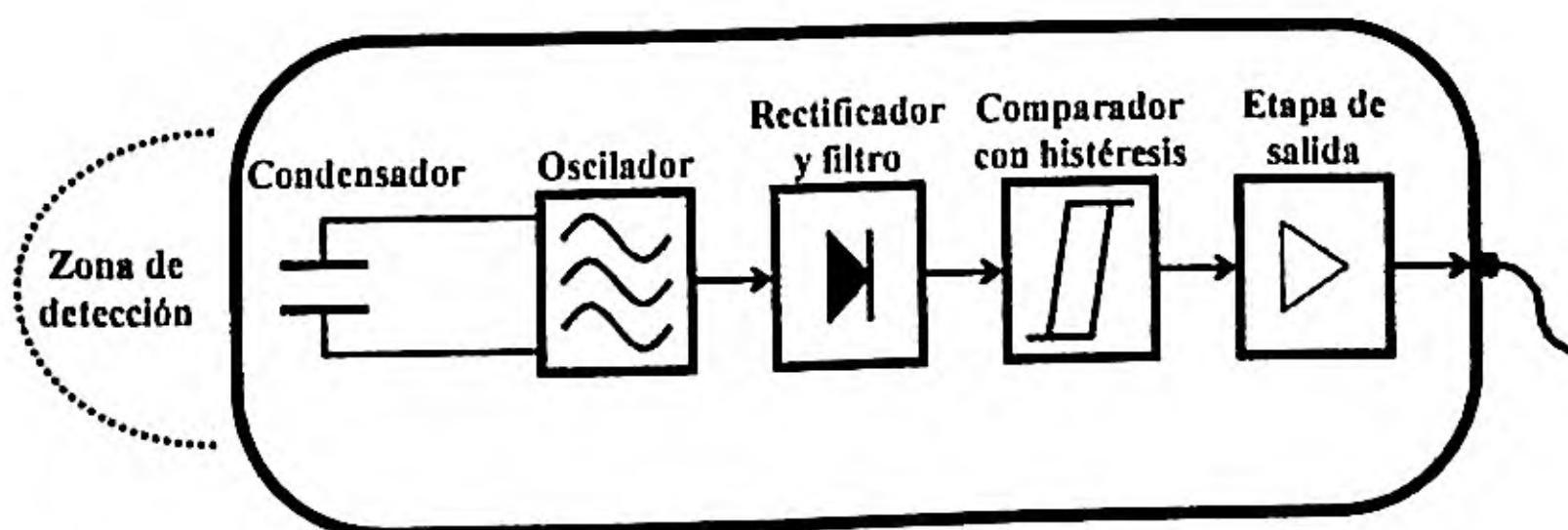


Figura 7. 149 Diagrama de bloques completo de un sensor capacitivo de proximidad.

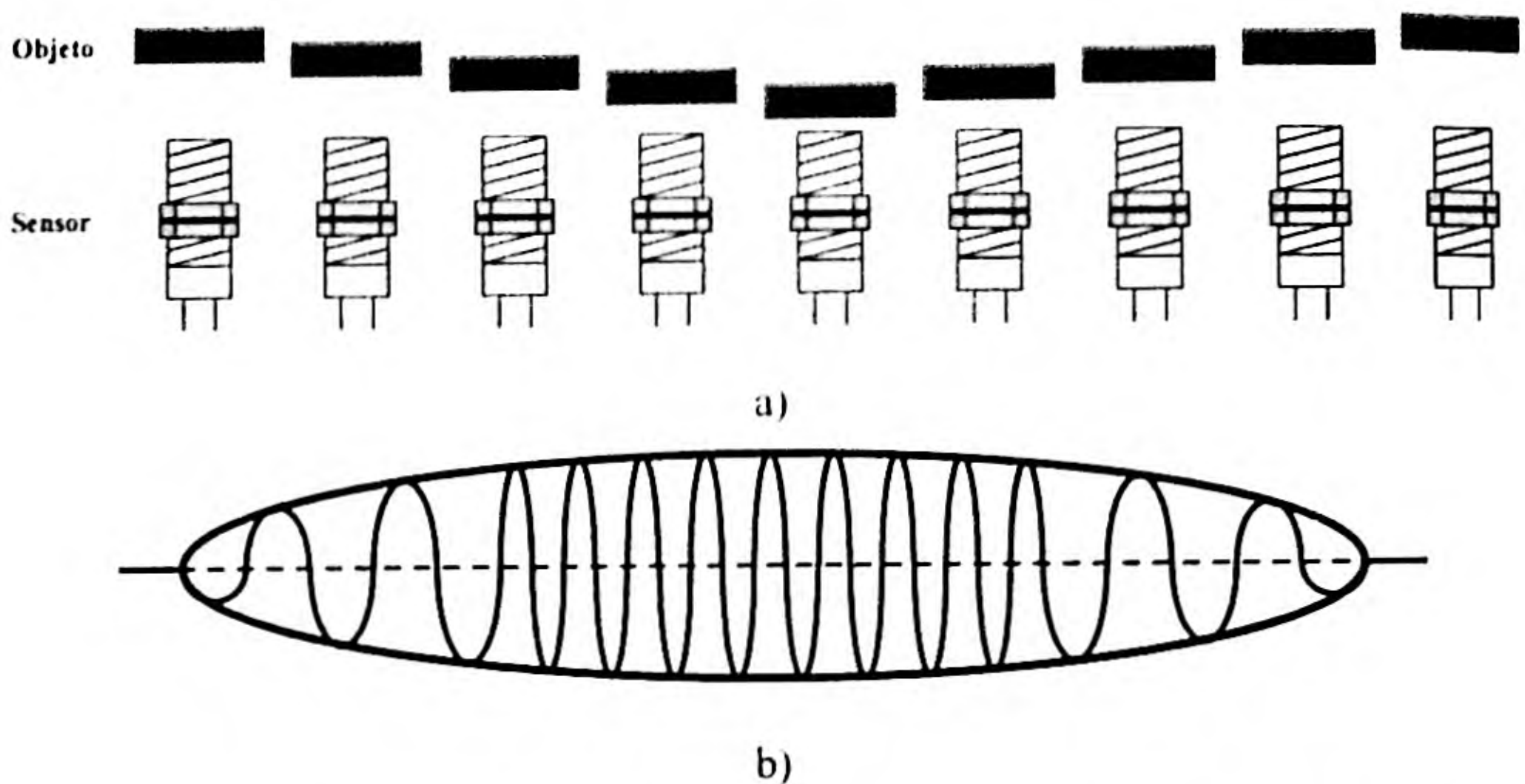


Figura 7.150 *Funcionamiento de un sensor capacitivo de proximidad: a) Evolución de la distancia entre el objeto y el sensor; b) Señal generada por el oscilador que forma parte del circuito de acondicionamiento del sensor.*

7.7.2.3.6.2 Tipos de sensores capacitivos de proximidad

Al igual que los sensores inductivos de proximidad, los capacitivos se pueden clasificar de acuerdo con el formato externo y con la forma constructiva.

La forma física de estos sensores suele ser cilíndrica, con el cuerpo roscado con la métrica correspondiente (Figura 7.151a). Sin embargo también existen otros formatos externos como el cúbico (cara activa cuadrada) o el prismático (cara activa rectangular). De igual forma la carcasa puede ser de plástico o metálica. La figura 7.151b muestra diversos sensores capacitivos.

De acuerdo con la forma constructiva los sensores capacitivos de proximidad, al igual que los sensores inductivos, pueden ser enrasables o no enrasables, que se describen seguidamente.

- **Sensores enrasables**

Los sensores enrasables, también denominados blindados como en el caso de los inductivos, se fabrican con una malla metálica alrededor de la cara activa del sensor, tal como se indica en la figura 7.152. Esto concentra el campo electrostático del condensador abierto en la parte frontal del sensor, lo que permite montarlo a ras del material que lo rodea sin que el sensor lo detecte. Debido a la alta concentración del campo electrostático, este tipo de forma constructiva es adecuada también para detectar objetos implementados con un material de baja permitividad. Por la misma razón los sensores enrasables son sensibles a la suciedad y a la humedad depositada en la cara activa del sensor, lo que puede dar lugar a activaciones erróneas de la señal de salida del sensor.

- **Sensores no enrasables**

En este tipo de sensores, también denominados no blindados como en el caso de los inductivos, el campo electrostático no está tan concentrado como en los enrasables. Por ello son adecuados para la detección de materiales de elevada permitividad y también para discriminar entre materiales de alta y baja permitividad. Este tipo de sensores incluyen frecuentemente un condensador auxiliar que compensa el efecto de las condiciones ambientales que modifican la capacidad del sensor, con el objetivo de que funcione correctamente incluso en presencia de agua pulverizada, polvo, suciedad, etc. Además el campo electrostático de compensación es de baja intensidad y en consecuencia solo existe en las proximidades de la cara activa del sensor. Esto hace que mediante dicho campo se detecte la suciedad, humedad, etc., que se deposita sobre la cara activa del sensor, pero que no se detecte el objeto. La figura 7.153 muestra un sensor capacitivo no enrasable con el condensador de compensación incorporado.

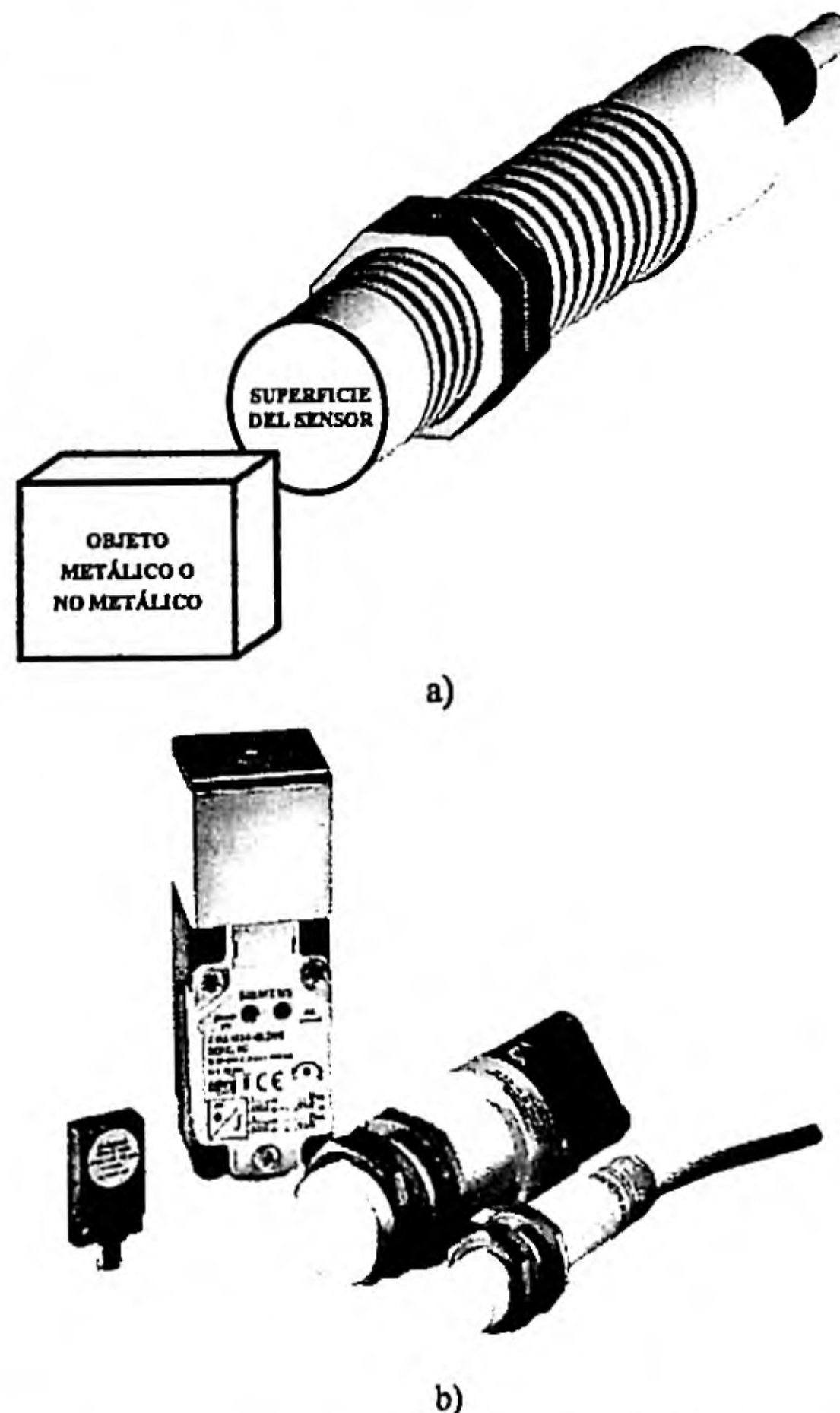


Figura 7.151 Sensor capacitivo de proximidad: a) Forma física y situación del objeto; b) Fotografía de diversos sensores capacitivos de proximidad (cortesía de Siemens).

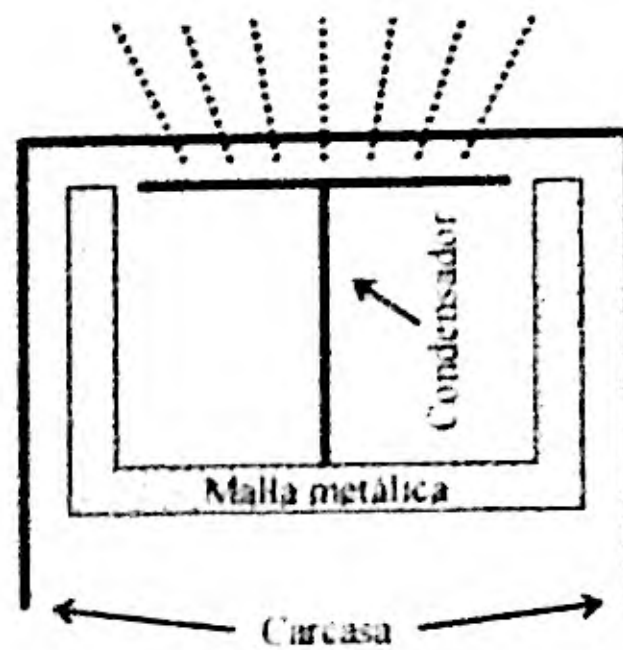


Figura 7.152 Sensor capacitivo de proximidad enrasable.

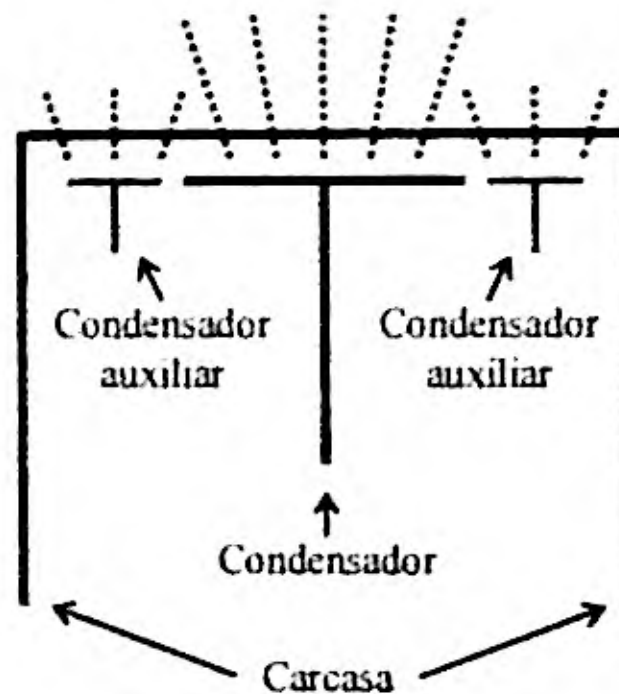


Figura 7.153 Sensor capacitivo de proximidad no enrasable.

7.7.2.3.6.3 Campo de trabajo

Al igual que en los sensores inductivos de proximidad, el campo de trabajo de los sensores capacitivos es un cono, situado en las proximidades de la cara activa del sensor, en cuyo interior el fabricante garantiza que el sensor detecta un objeto normalizado o patrón. Dicho objeto es un cuadrado de acero ST37 que está conectado a masa y tiene un milímetro de espesor y un lado de longitud igual al diámetro de la cara sensible del sensor (la mayoría de los que se utilizan en la industria son cilíndricos) o a tres veces la distancia de detección nominal del sensor (el mayor de los dos).

La figura 7.154 muestra el campo de trabajo típico de un sensor capacitivo. Como se puede observar en ella, el campo de trabajo abarca una región espacial delimitada por las dos curvas, y el punto de corte de ambas coincide con la distancia de detección nominal del sensor. Pero, en las aplicaciones reales los objetos metálicos no se conectan a masa y en múltiples ocasiones los objetos son aislantes. Por ello, es necesario tener en cuenta que en las capacidades indicadas en la figura 7.148 se producen distintos efectos según el tipo de material con que esté realizado el objeto y si está o no conectado a la masa del circuito. En la práctica se pueden tener tres tipos de situaciones:

- **Objeto metálico flotante**

Se trata de un objeto metálico que no está conectado a masa. En este caso el circuito que forman los condensadores de la figura 7.148 es equivalente al representado en la figura 7.155. El objeto constituye la placa común de los condensadores C_B y C_C . Esto hace que, debido a que disminuye la distancia entre las placas de los condensadores C_B y C_C , el valor del condensador equivalente aumente a medida que el objeto se acerca al sensor y que disminuya si se aleja.

- **Objeto metálico conectado a masa**

En este caso el circuito que forman los condensadores es el indicado en la figura 7.156. El condensador C_D es un cortocircuito y el condensador C_A disminuye a medida que el objeto

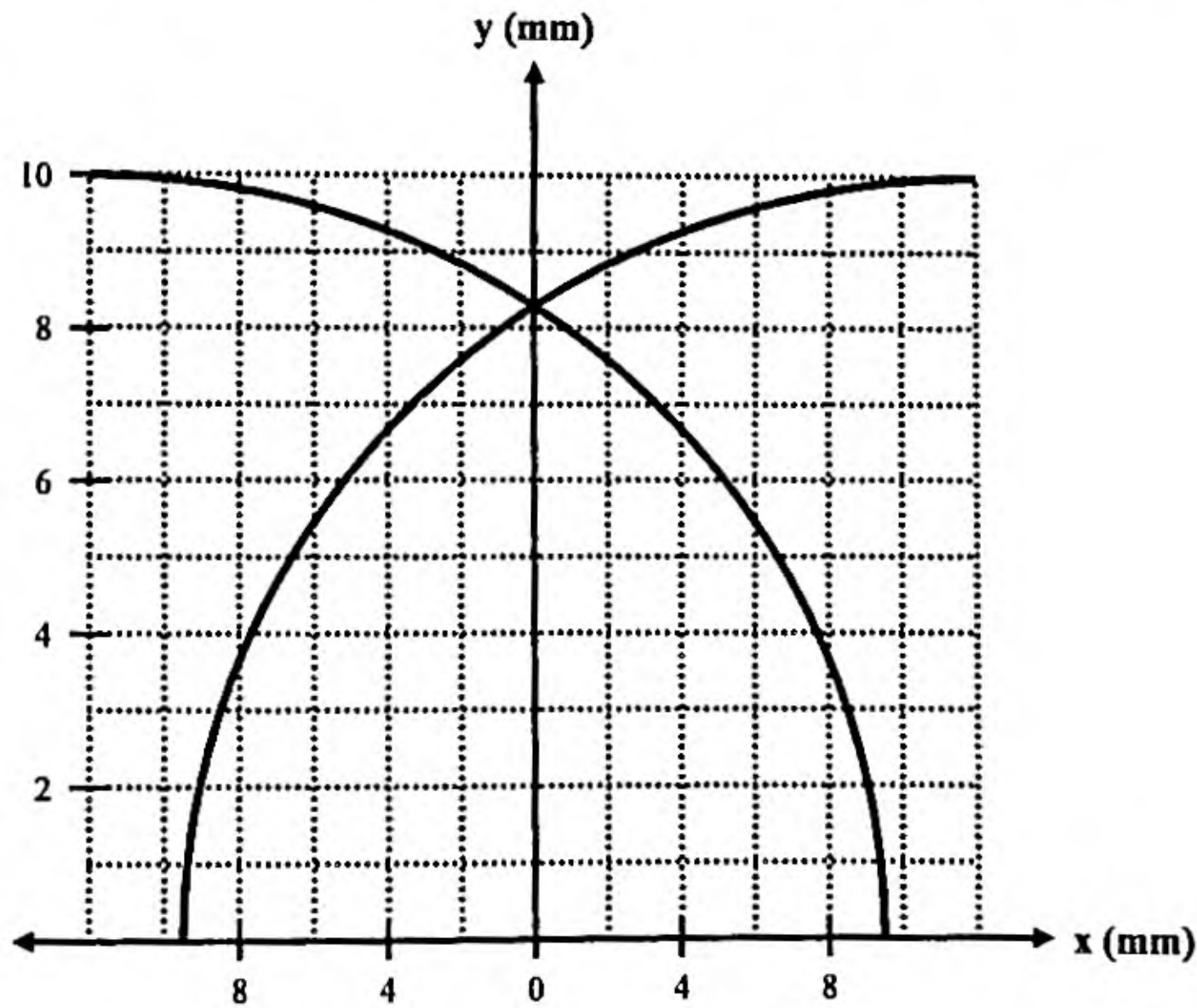


Figura 7. 154 Campo de trabajo del sensor capacitivo de proximidad PXC200 M30 (cortesía de Siemens).

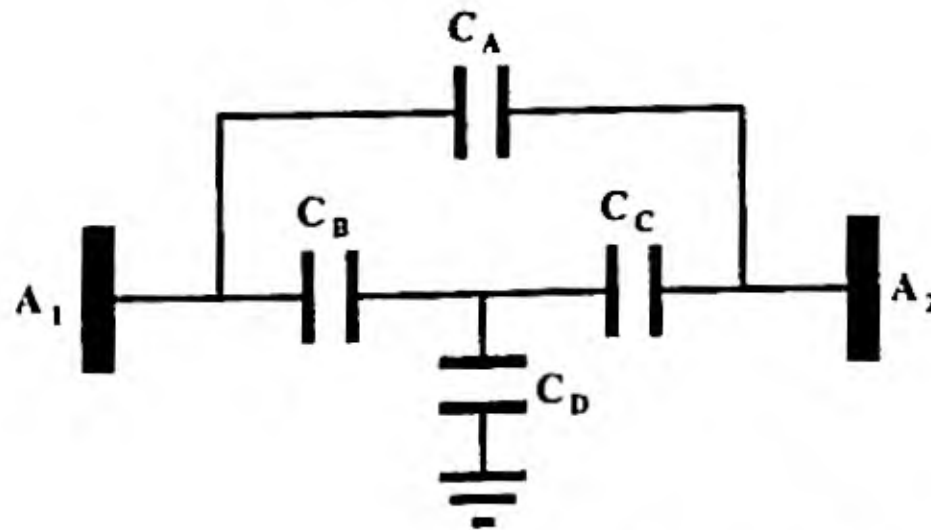


Figura 7. 155 Condensadores que intervienen en el circuito de un sensor capacitivo de proximidad cuando el objeto es metálico y está aislado.

se acerca y su efecto se suma al de los condensadores C_B y C_C que aumentan su capacidad a medida que el objeto se acerca. En este caso la variación de capacidad total es máxima y por eso este tipo de objeto se detecta a mayor distancia.

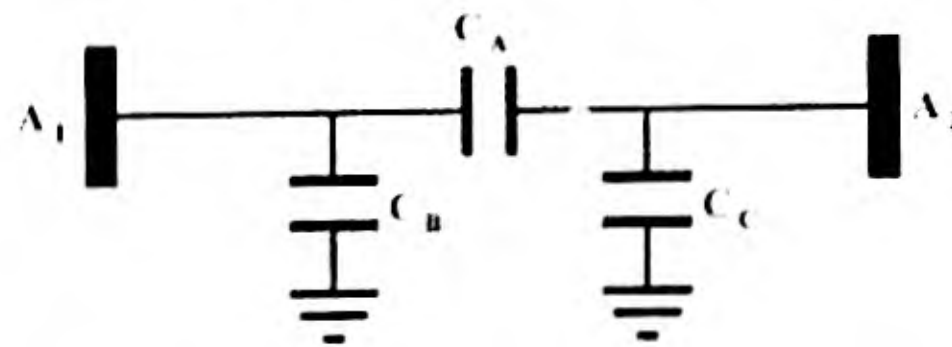


Figura 7.156 Condensadores que intervienen en el circuito de un sensor capacitivo de proximidad cuando el objeto es metálico y está conectado a masa.

- **Objeto aislante (no conductor)**

En este caso el condensador C_D es un circuito abierto y el circuito que forman los condensadores es el de la figura 7.157. El efecto más importante es el del condensador C_A cuya capacidad aumenta al acercarse el objeto porque se eleva la permitividad del dieléctrico debido a que el aire pasa a ser sustituido por el objeto.



Figura 7.157 Condensadores que intervienen en el circuito de un sensor capacitivo de proximidad cuando el objeto es aislante.

Como resumen de todo ello se puede indicar:

- A medida que aumenta el tamaño del objeto, se eleva la capacidad resultante. La distancia de detección no aumenta a partir de un determinado tamaño mínimo del objeto.
- A medida que disminuye la distancia a la que se encuentra el objeto, aumenta la capacidad resultante y se eleva la posibilidad de detección.
- El parámetro más importante que hace variar la capacidad del condensador cuando el objeto a detectar es metálico y está aislado (flotante) o conectado a masa, es la distancia entre el objeto y el sensor.
- Si el objeto es aislante el parámetro que más influye es la permitividad del mismo, debido a que a medida que se acerca al sensor aumenta la cantidad de aire sustituida por el objeto, lo que hace que el dieléctrico del condensador pase paulatinamente de ser aire a ser el objeto, que tiene una permitividad mayor que la del aire. Debido a ello, es necesario realizar un ajuste cuidadoso para detectar objetos realizados con materiales de baja permitividad.
- Se pueden detectar materiales de permitividad elevada a través de las paredes de un depósito realizado con un material de permitividad baja. Por ejemplo un sensor capacitivo se puede utilizar para detectar alcohol ($\epsilon_r = 25,8$) a través de la pared de un recipiente de vidrio ($\epsilon_r = 3,7$), pero no detecta la harina ($\epsilon_r = 1,5$) a través del mismo recipiente.

La tabla 7.15 muestra los valores de la permitividad relativa ϵ_r de distintos tipos de materiales. En la bibliografía [CRC 08] se indican los valores de la permitividad de otros materiales. Dado un objeto de unas dimensiones determinadas, cuanto mayor es el valor de la permitividad mayor es la capacidad del condensador y en definitiva puede ser detectado a mayor distancia. En la figura 7.158 se muestra la relación entre la permitividad relativa y la distancia de detección, en % de la distancia real de detección que se consigue con el objeto patrón.

Material	ϵ_r	Material	ϵ_r
Alcohol	25	Leche en Polvo	3,5-4
Aceite	2-3,5	Madera Húmeda	10-30
Aceite de Soja	2,9-3,5	Madera Seca	2-7
Aceite para transformadores	2,2	Mármol	8-8,5
Aceite de trementina	2,2	Mica	5,7-6,7
Acetona	19,5	Nitrobenceno	36
Agua	80	Naylon	4-5
Aire o vacío	1	Papel	1,6-2,6
Alcohol	25,8	Papel Oleosaturado	4
Amoníaco	15-25	Parafina	1,9-2,5
Anilina	6,9	Perspex	3,2-3,5
Arena	3-5	Petróleo	2-2,2
Azúcar	3	Plexiglás	3,2
Azufre	3,4	Poliacetal	3,6-3,7
Baquelita	3,6	Poliamida	5
Barniz de Laca	2,5-4,7	Poliestireno	3
Barniz de Silicona	2,8-3,3	Polietileno	2,3
Benceno	2,3	Polipropileno	2-2,3
Cartón Prensado	2-5	Polivinilo	2,9
Caucho Blando	2,5	Porcelana	4,4-7
Caucho Duro	4	Resina Acrílica	2,7-4,5
Celuloide	3	Resina Epóxida	2,5-6
Ceniza	1,5-1,7	Resina de Estireno	2,3-3,4
Cereales	3-5	Resina Fenólica	4-12
Cloro líquido	2	Resina de Poliester	2,8-8,1
Cristal de Cuarzo	3,7	Resina de Urea	5-8
Dioxido de Carbono	1	Resina Melamínica	4,7-10,2
Ebonita	2,7-2,9	Sal	6
Etanol	24	Silicona	2,8
Etilenglicol	38,7	Soluciones Acuosas	50-80
Freón R22 y 502 (Líquido)	6,1	Teflón	2
Gasolina	2,2	Tetracloruro de Carbono	2,2
Glicerina	47	Tolueno	2,3
Goma	2,5-35	Vaselina	2,2-2,9
Harina	1,5-1,7	Vidrio	3,7-10

Tabla 7.15 Permitividad relativa de distintos tipos de materiales.

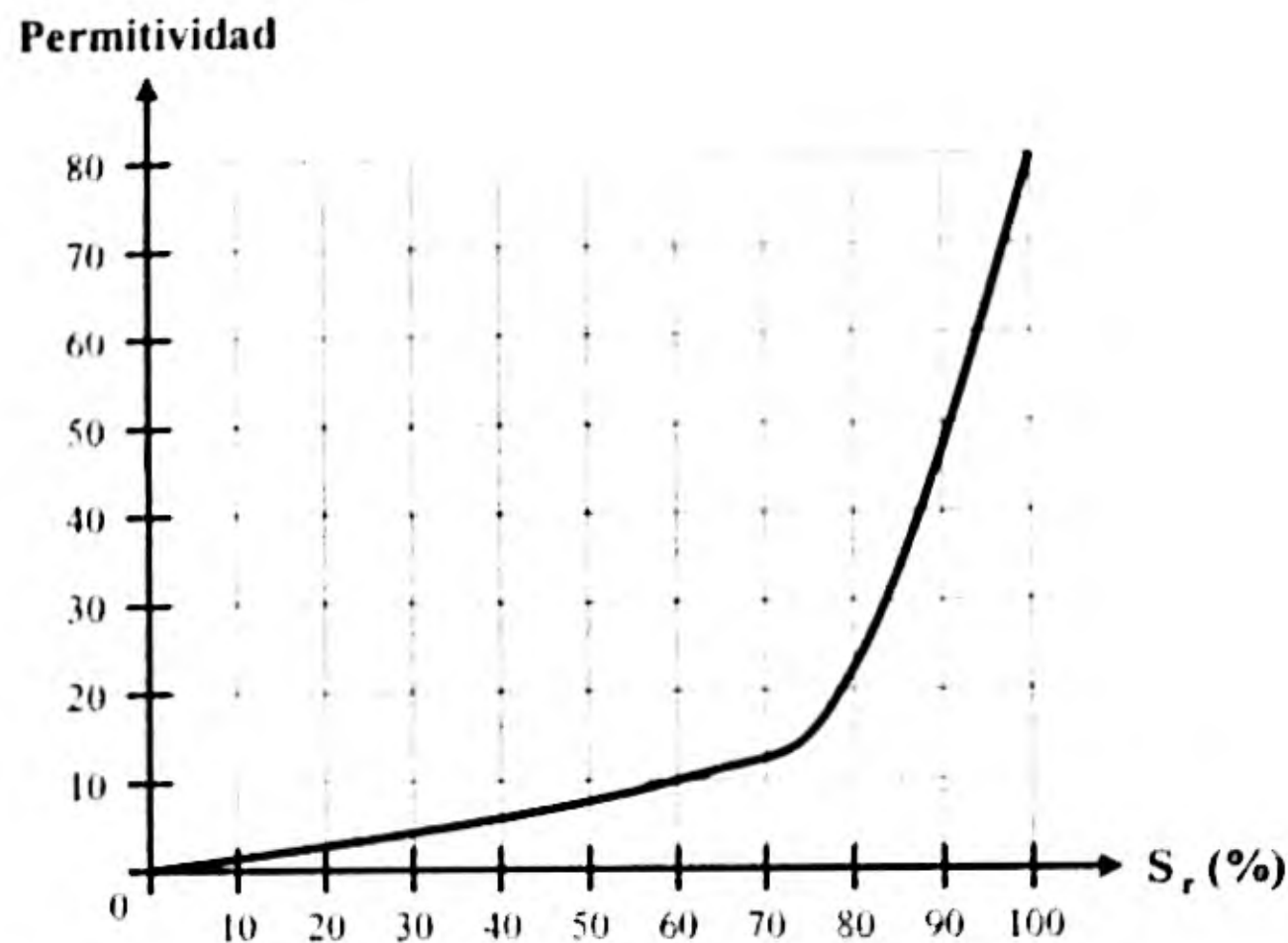


Figura 7.158 Relación entre la permitividad relativa y la distancia de detección real en % de la distancia real de detección que se consigue con el objeto patrón.

Material	Factor
Aceite	0,1
Acero ST 37	1
Acero Inox.	0,6 - 1
Agua	1
Aluminio	0.35 - 0.5
Cobre	0.25 - 0.45
Latón	0.35 - 0.5
Madera	0,2 - 0,7
PVC	0,6
Vidrio	0,5

Tabla 7.16 Factores de corrección de los sensores capacitivos de proximidad según el tipo de material.

7.7.2.3.6.4 Características técnicas

Las características técnicas de los sensores capacitivos de proximidad son similares a las de los sensores inductivos de proximidad y lo dicho para aquellos es válido, en general, para éstos. Por ello a continuación solamente se indican algunos detalles específicos de los sensores capacitivos.

Distancia de detección

En los sensores capacitivos de proximidad se consideran las mismas distancias de detección que en el caso de los inductivos. Además, al igual que en ellos, se puede determinar de forma aproximada la distancia de detección para objetos no metálicos y metálicos conectados a masa, distintos del normalizado, mediante la aplicación de un factor de corrección, de acuerdo con la ecuación:

$$Sn_{(Material)} = Sn_{(AceroST37)} * Factor_{(Material)}$$

En la tabla 7.16 se indica el factor de corrección para diferentes materiales conductores y aislantes.

7.7.2.3.6.5 Normas de instalación

Al instalar los sensores capacitivos de proximidad hay que tener también las siguientes precauciones:

- **Interferencias mutuas**

Las interferencias mutuas se producen cuando dos o más sensores capacitivos de proximidad se sitúan demasiado próximos, enfrentados o lateralmente.

Cuando están enfrentados, la distancia mínima es igual a ocho veces el diámetro d de la cara activa si es circular, o el lado si es cuadrada, tanto si el sensor es enrasable como si no lo es (Figura 7.159a).

Cuando los sensores se sitúan lateralmente, la distancia mínima a la que se deben colocar (Figura 7.159b) es del orden de:

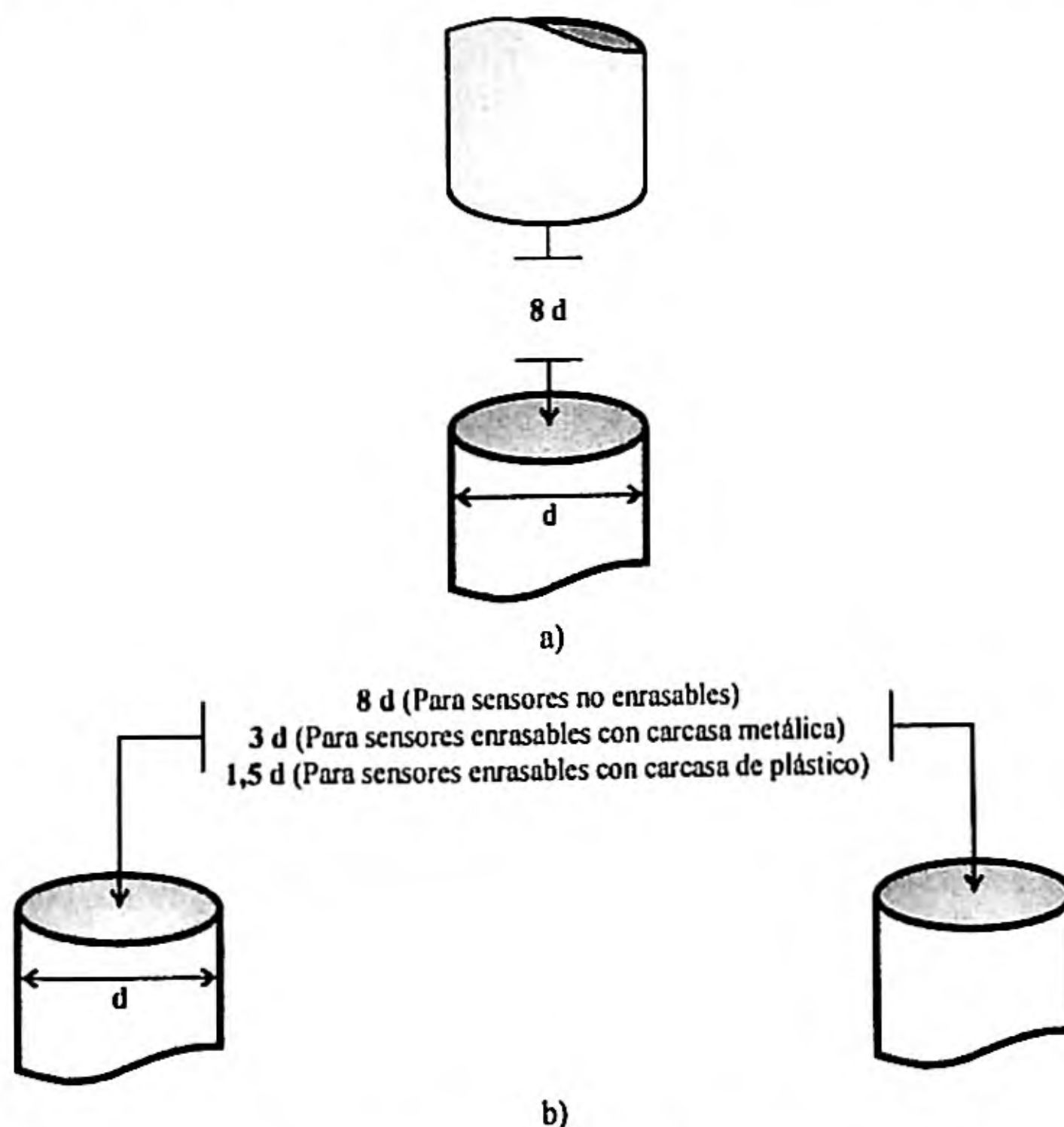


Figura 7.159 Distancia de separación mínima entre dos sensores capacitivos de proximidad: a) Cuando los sensores están enfrentados; b) Cuando los sensores están colocados lateralmente.

- Ocho veces el diámetro d de la cara activa si el sensor no es enrasable.
 - Tres veces el diámetro d de la cara activa si el sensor es enrasable y su carcasa es de plástico.
 - Una vez y media el diámetro d de la cara activa si el sensor es enrasable y su carcasa es metálica.
- **Objetos circundantes**
 Para que un sensor capacitivo de proximidad no detecte los objetos próximos al mismo, éstos deben estar situados a una distancia mínima igual a tres veces la distancia de detección nominal S_0 , tanto si el sensor es enrasable como si no lo es, tal como se indica en la figura 7.160.

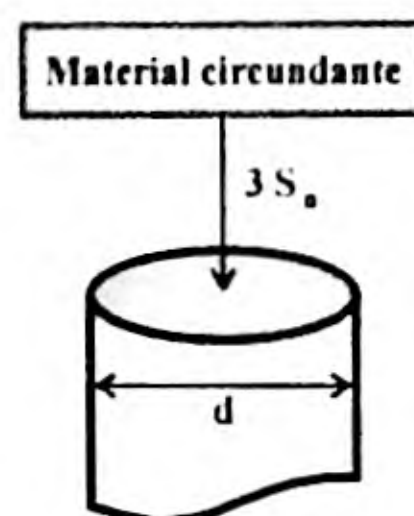


Figura 7.160 Distancia mínima a la que debe situarse un objeto que está cerca de un sensor capacitivo de proximidad.

7.7.2.3.7 Sensores ultrasónicos de proximidad

7.7.2.3.7.1 Introducción

Los sensores ultrasónicos de proximidad (*Ultrasonic proximity sensors*) basan su funcionamiento en las características de las señales sonoras de frecuencia superior al rango audible por el ser humano ($>20\text{Khz}$), que reciben por ello el nombre de ultrasonidos. Constan de un elemento que emite la señal de ultrasonidos o ultrasónica (emisor de ultrasonidos) y otro que la capta (receptor de ultrasonidos).

Los sensores ultrasónicos se pueden utilizar, además de para detectar la presencia de objetos, para medir la distancia que los separa del sensor. La figura 7.161 muestra diversos sensores ultrasónicos de proximidad (*Sonar proximity switches*) de la familia Bero de Siemens. Los sensores de este tipo disponen, en general en la actualidad, de un circuito electrónico basado en un microcontrolador, para, entre otras funciones, medir el tiempo que tarda el elemento receptor en recibir los impulsos enviados por el emisor. La figura 7.162 muestra un sensor ultrasónico de proximidad de la familia Bero, conectado a un computador personal portátil y configurable mediante el mismo.

A continuación se analizan las características generales de los ultrasonidos y la forma de generarlos y detectarlos así como las características técnicas y modo de operación de los sensores detectores de objetos que los utilizan. Posteriormente se describen los diferentes tipos de sensores de ultrasonidos y sus características.

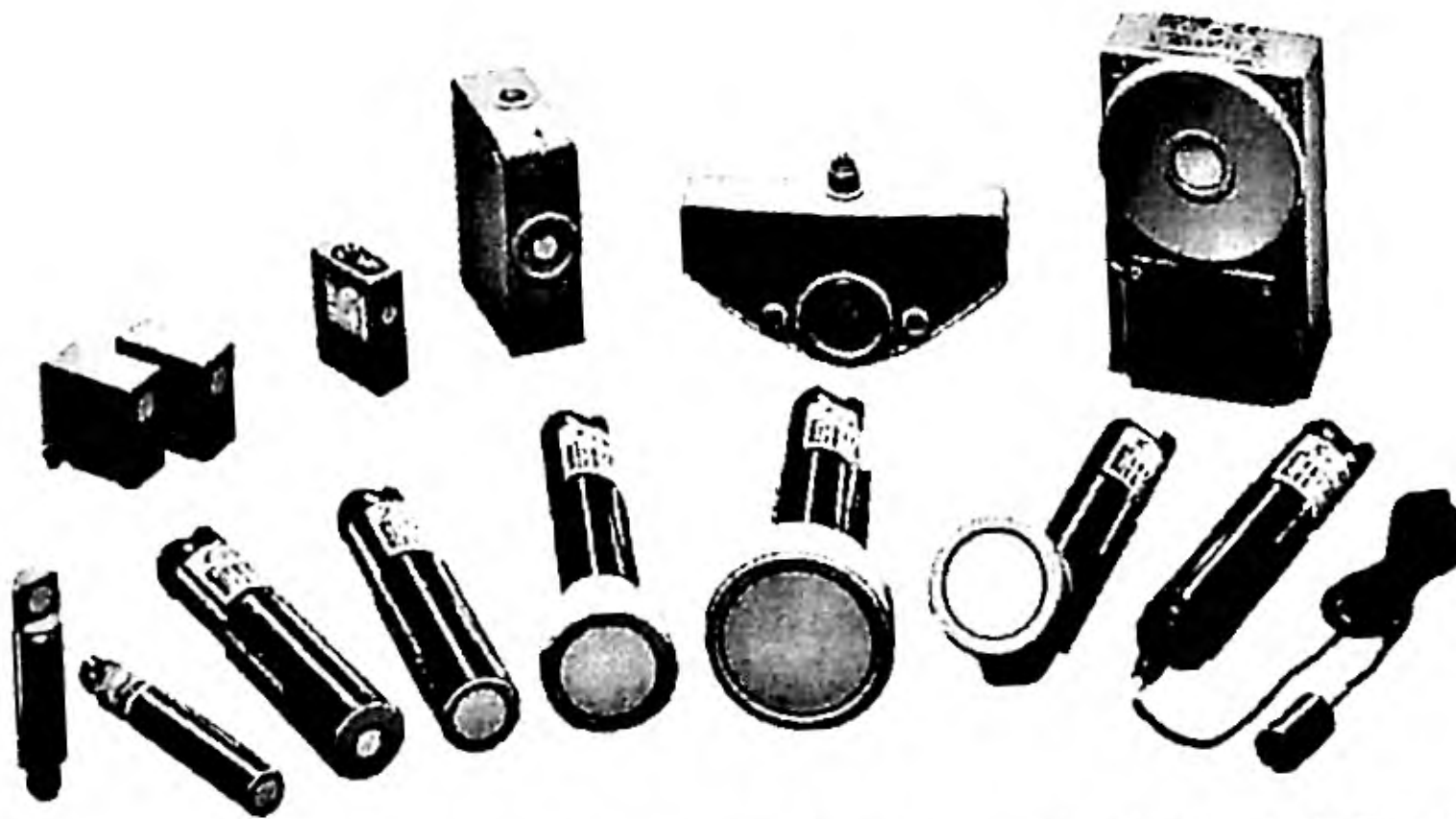


Figura 7.161 Diversos sensores de proximidad de ultrasonidos (cortesía de Siemens).



Figura 7.162 Sensor de ultrasonidos configurable mediante un computador personal.

7.7.2.3.7.2 Ultrasonidos

Los ultrasonidos (*Ultrasonic waves*) utilizados en aplicaciones industriales son, en general, señales acústicas de frecuencia comprendida entre 20 KHz y 250 KHz, aunque la mayor parte de los sensores de ultrasonidos que miden distancias, trabajan en el rango de frecuencias comprendidas entre 40 y 60 KHz. La figura 7.163 muestra el espectro de las señales acústicas.

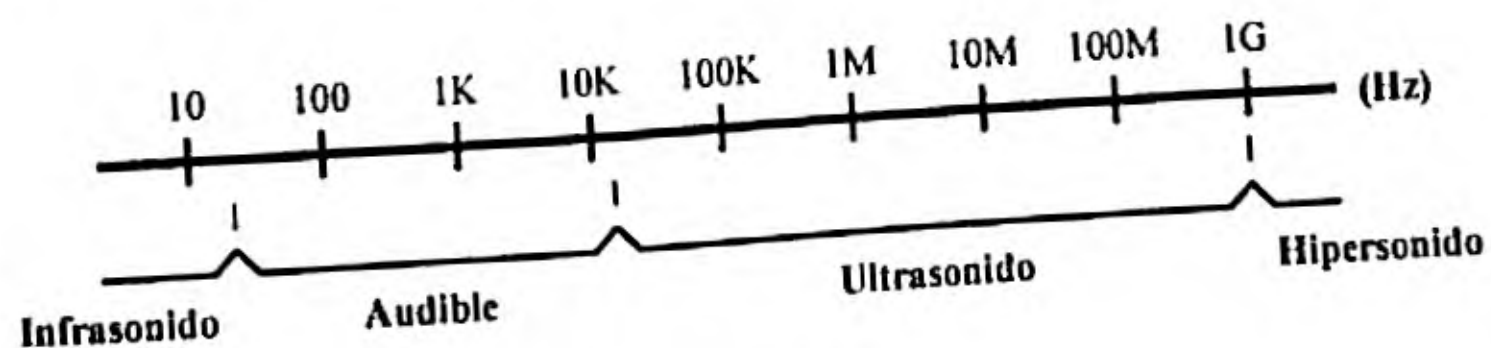


Figura 7.163 Espectro de las señales acústicas.

Aunque los ultrasonidos se pueden propagar mediante distintos tipos de ondas (longitudinales, transversales, de superficie, etc.), es la onda longitudinal o de tipo L la más habitual en los sensores de proximidad, porque se desplaza fácilmente en líquidos, sólidos y gases. En este tipo de ondas, el movimiento de las partículas en un medio es paralelo a la línea de propagación de la onda. Además poseen una elevada velocidad de desplazamiento en la mayoría de los medios, con longitudes de onda pequeñas en relación con la superficie activa del elemento sensor. Esto permite que la energía se enfoque en un haz agudo y de pequeña divergencia.

Los sensores ultrasónicos se ven afectados por el medio que utilizan para su transmisión, que, en el caso de los sensores de proximidad, habitualmente es el aire. Otro aspecto a tener en cuenta es el tipo y forma del objeto. Por todo ello, y para utilizar correctamente los sensores ultrasónicos de proximidad, es necesario tener en cuenta las siguientes características o efectos:

- La variación de la velocidad del sonido en función de la temperatura y de la composición del medio transmisor. Esto afecta a la exactitud y a la resolución del sensor.
- La variación de la longitud de onda del ultrasonido en función de la velocidad y de la frecuencia del mismo. Esto afecta a la resolución, a la exactitud, al tamaño mínimo del objeto y a las distancias mínima y máxima a las que puede ser detectado.
- La variación de la atenuación del sonido en función de la frecuencia y de la humedad del aire. Esto afecta a la distancia máxima de detección.
- La variación de la amplitud del ruido de fondo en función de la frecuencia. Esto puede afectar a la distancia máxima de detección y al tamaño mínimo del objeto.
- La variación del ángulo del haz ultrasónico. Esto afecta a la máxima distancia de detección y reduce la posibilidad de detectar objetos no deseados.
- La variación de la amplitud del haz reflejado en un objeto en función de la distancia, de la geometría, del tipo de superficie y del material y tamaño del objeto. Esto afecta a la distancia máxima de detección.

A continuación se analizan todos estos factores.

Generación de los ultrasonidos

Para generar ultrasonidos se utilizan habitualmente materiales piezoeléctricos (Figura 7.164a), en los que se generan tensiones eléctricas al aplicarles una presión mecánica (Figura 7.164b). El efecto piezoeléctrico (*Piezoelectric effect*) se debe a que la presión provoca una deformación de la red cristalina que da lugar, a su vez, a un desplazamiento de las cargas eléctricas moleculares, lo que hace que aparezcan diferencias de potencial entre las caras del material (Figura 7.164b). El signo de esta diferencia de potencial se invierte cuando se invierte el sentido de la presión, que puede ser de tracción o de compresión. El efecto piezoeléctrico es reversible, es decir, si se aplica una diferencia de potencial entre las caras del material sus cargas moleculares se desplazan, lo que provoca una deformación del mismo.

El cuarzo y la turmalina son materiales piezoeléctricos naturales, pero, debido al valor reducido de sus parámetros característicos, fueron sustituidos por otros materiales piezoeléctricos sintéticos implementados con titanatos y circonatos de plomo. Estos materiales son más estables que los naturales, aunque presentan el inconveniente de su dependencia de la temperatura y la pérdida de sus propiedades piezoeléctricas a medida que su temperatura se acerca al punto de Curie.

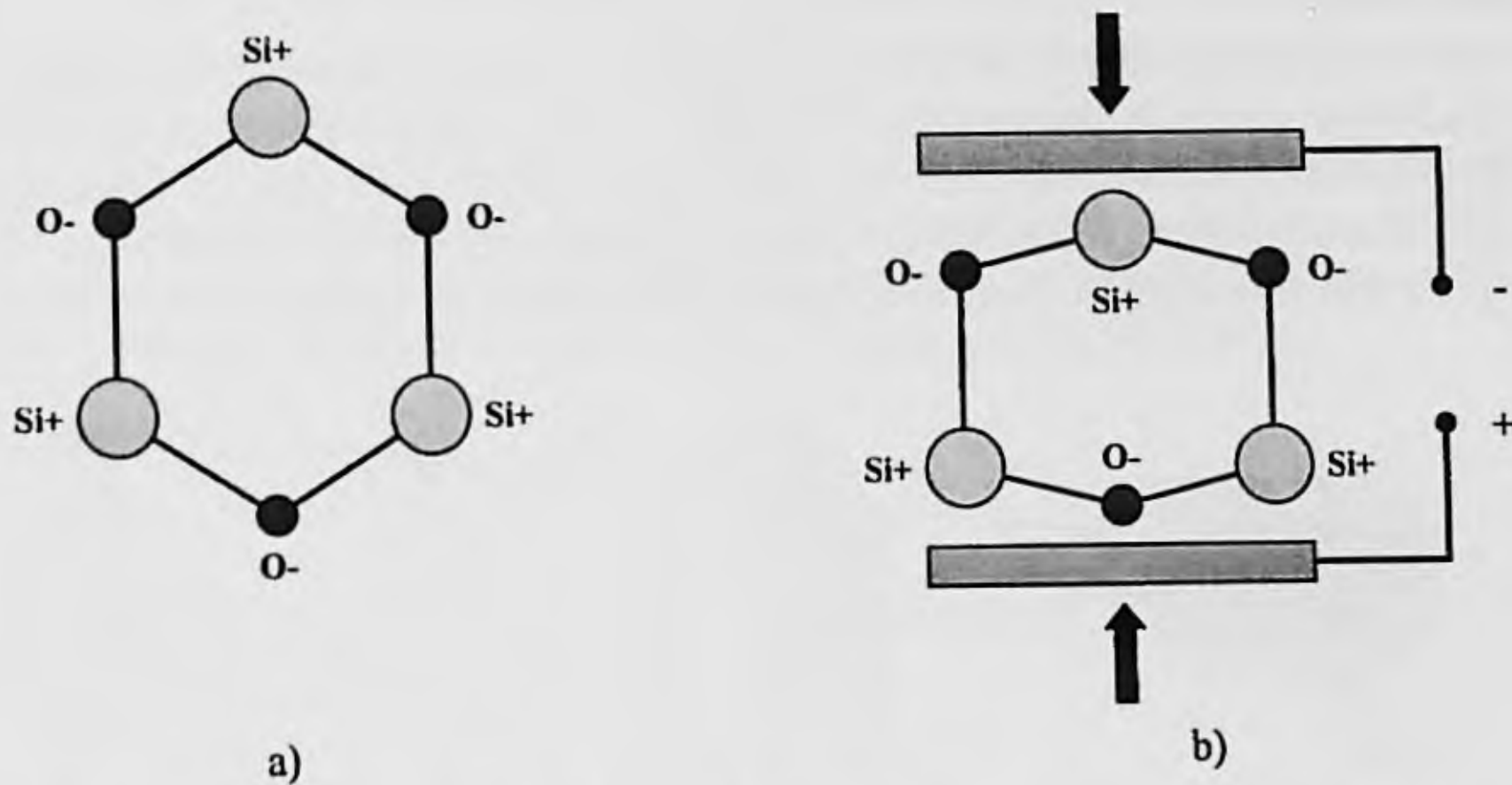


Figura 7.164 Comportamiento del material piezoeléctrico: a) Cuando no se le aplica una fuerza; b) Cuando se le aplica una fuerza.

Utilizando el efecto piezoeléctrico se pueden realizar emisores y detectores de ultrasonidos. Mediante dos electrodos se aplica una señal eléctrica de la frecuencia adecuada a un material piezoeléctrico que se deforma y produce señales de presión de aire de la misma frecuencia que la señal eléctrica aplicada. De igual forma, el mismo material se puede utilizar como receptor, dado que las señales de presión del aire provocan una deformación del mismo y hacen que aparezca una señal eléctrica entre los electrodos. La figura 7.165 muestra la estructura de un sensor piezoeléctrico que puede funcionar como emisor o como receptor de ultrasonidos. Además del material piezoeléctrico y de los electrodos, el emisor-receptor incluye un material elástico que le proporciona estabilidad. El sensor posee también un cono de radiación, que transmite al aire la señal de ultrasonidos generada por el material piezoeléctrico y transmite a éste la señal de presión del aire procedente del objeto.

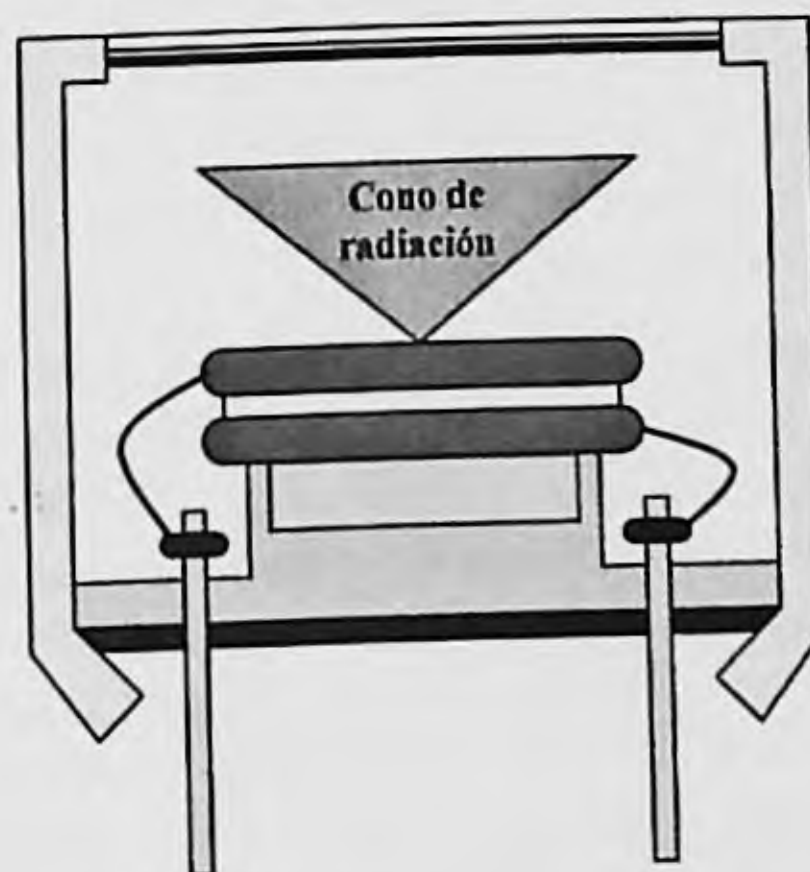


Figura 7.165 Estructura de un elemento sensor piezoeléctrico que puede funcionar como emisor o como receptor de ultrasonidos.

Cono de radiación

El cono de radiación es la zona en la que el sensor emite los ultrasonidos (Figura 7.166). Dicha zona tiene la forma característica de un cono que se expande a medida que aumenta la distancia al emisor. La forma del cono depende de diversos factores como por ejemplo la frecuencia de funcionamiento y el tamaño, la forma constructiva y las características acústicas de la superficie elástica que vibra. Las características del cono de recepción son recíprocas.

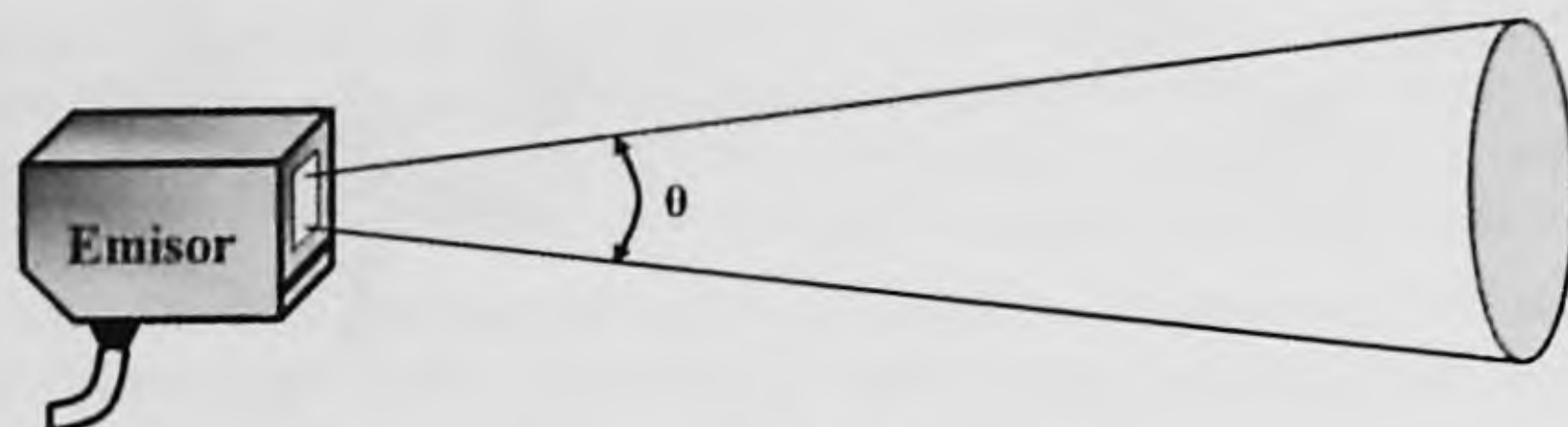


Figura 7.166 Cono de radiación de un emisor de ultrasonidos.

Los emisores de ultrasonidos se pueden diseñar para emitirlos en todas las direcciones (omnidireccionales) o en conos de radiación muy estrechos. La forma física más habitual de los sensores de ultrasonidos está constituida por una superficie de emisión circular y elástica cuyos puntos vibran en fase.

Cuando la superficie de emisión es circular, que es lo más habitual, el ángulo de emisión es función del cociente entre el diámetro D de la superficie de emisión y la longitud de onda λ de la señal de sonido. A su vez, la longitud de onda λ depende de la velocidad del sonido v y de la frecuencia f del mismo, de acuerdo con la ecuación:

$$\theta = \arcsen \left(\frac{\lambda}{D} \right) = \arcsen \left(\frac{v/f}{D} \right)$$

que establece que, cuanto más grande es el diámetro de la superficie de emisión con respecto a la longitud de onda del sonido, más estrecho es el cono de radiación. Por ejemplo, si el diámetro es igual a dos veces la longitud de onda, el ángulo del haz es aproximadamente igual a 30° , pero si se aumenta el diámetro o la frecuencia, de tal modo que el cociente se convierta en $1/10$, el ángulo del haz es aproximadamente igual a 6° .

La ecuación anterior proporciona el valor del ángulo total del haz de emisión. En general se necesitan haces estrechos a fin de evitar reflexiones indeseadas, para lo cual la superficie de emisión debe ser grande con respecto a la longitud de onda del ultrasonido emitido. En este caso, además del cono de radiación fundamental se generan un conjunto de lóbulos secundarios en forma de anillos. En la figura 7.167 se muestra una representación en tres dimensiones (3D) de un caso de este tipo. El ángulo del haz se define como el ángulo total en el que el nivel de presión del ultrasonido se reduce en 3 dB a ambos lados del eje central del haz. Normalmente los fabricantes representan el ángulo del haz de ultrasonidos en una gráfica de dos dimensiones (2D) que muestra la sensibilidad y la zona de emisión del sensor en función del ángulo θ , y recibe el nombre de diagrama de directividad. La figura 7.168 muestra un diagrama de directividad que se obtiene mediante un corte longitudinal del haz en tres dimensiones. Dicha figura muestra

el haz ultrasónico generado por un sensor circular cuyo diámetro es igual a dos veces la longitud de onda. Se puede observar que los puntos en los que la presión del ultrasonido se reduce en 3dB están situados en $+15^\circ$ y -15° respecto del eje, lo que supone un ángulo total de 30° . El ángulo entre los lóbulos secundarios es de aprox. 70° y el ángulo entre los picos de los lóbulos es de 55° y -55° (305°). Por ello, en las aplicaciones industriales de los sensores de ultrasonidos, se debe tener en cuenta que el sensor puede detectar objetos con ángulos mayores que el correspondiente al haz principal, que es el que suelen proporcionar los fabricantes.

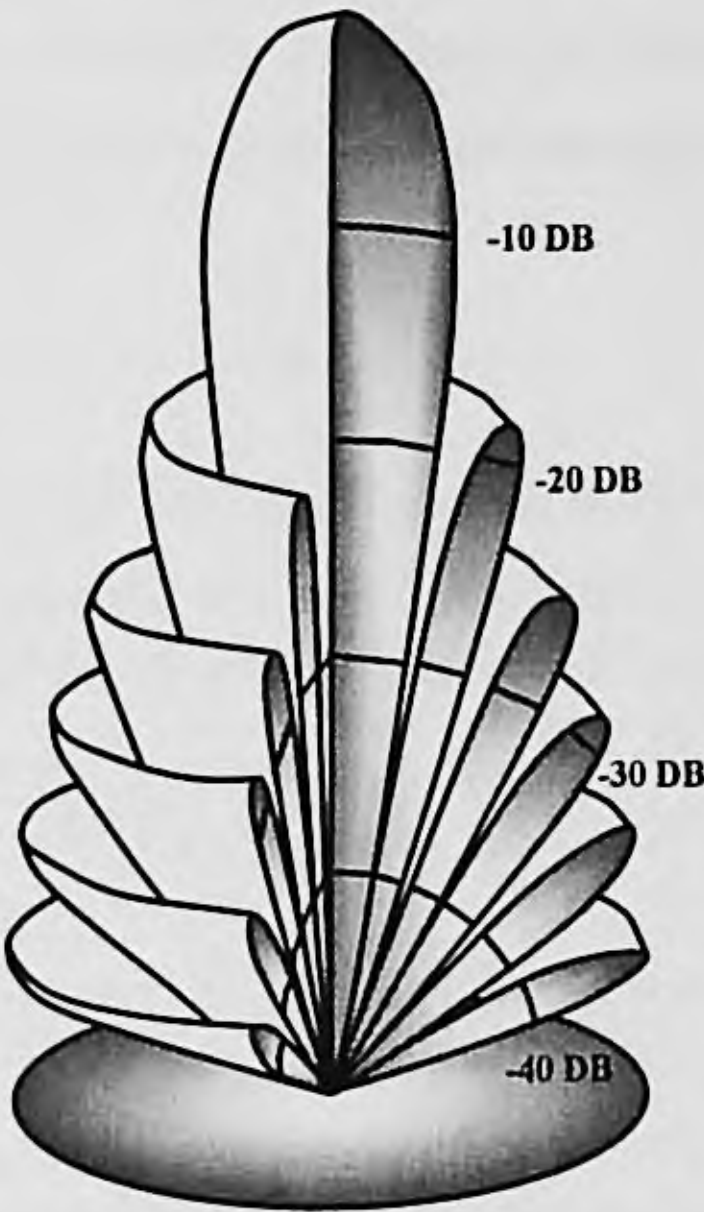


Figura 7.167 *Haz de ultrasonidos representado en tres dimensiones (3D).*

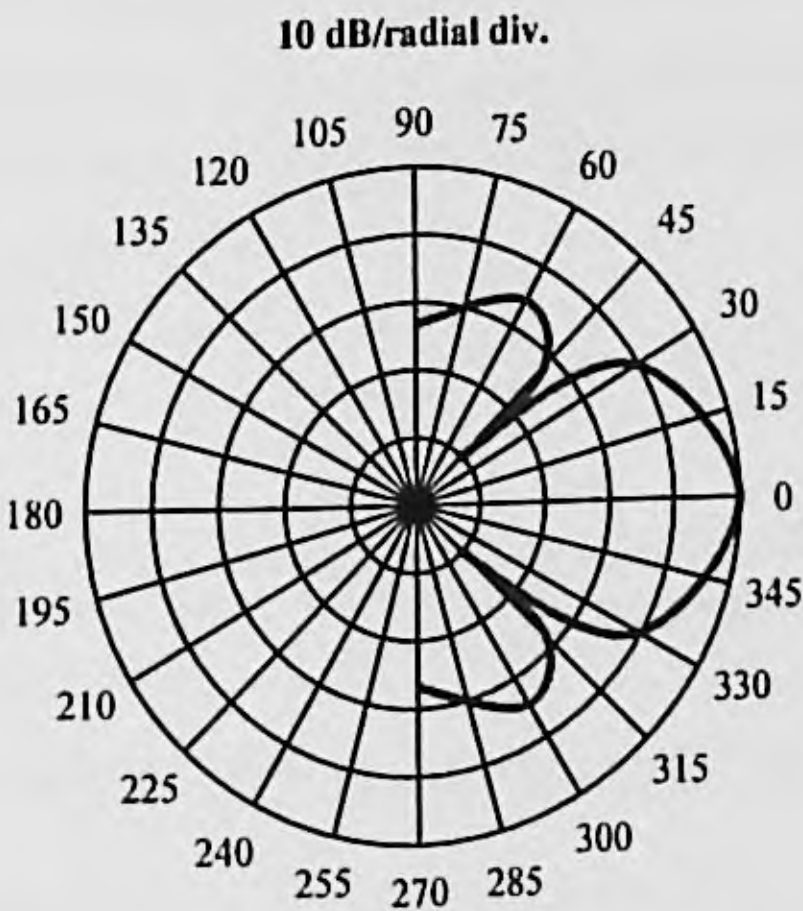


Figura 7.168 *Diagrama de directividad de un haz ultrasónico.*

Velocidad del ultrasonido

Las señales ultrasónicas se transmiten a través del medio a una velocidad que depende de las características del mismo. Para los sólidos la velocidad del sonido v_s viene dada por la ecuación:

$$v_s = \sqrt{\frac{E}{\rho}}$$

en la que E representa el módulo de Young y ρ la densidad del sólido.

Para los líquidos la velocidad del sonido v_L cumple la ecuación:

$$v_L = \sqrt{\frac{B}{\rho}}$$

en la que B representa el módulo de compresibilidad del líquido y ρ su densidad.

La mayor parte de los sensores de ultrasonidos utilizan el aire como medio de transmisión de la señal ultrasónica. La velocidad v_G del sonido en los gases responde a la expresión:

$$v_G = \left(\frac{KP}{\rho} \right)^{\frac{1}{2}} = \sqrt{KRT}$$

en la que K es el exponente adiabático del gas, P es la presión del gas, ρ es la densidad del gas, T es la temperatura, y R una constante que depende del tipo de gas.

El exponente adiabático K depende del coeficiente del calor específico a presión constante C_p y del calor específico a volumen constante C_v . El valor de K para los gases diatómicos como el oxígeno, el nitrógeno y el aire es 1,4. Esta ecuación demuestra que la velocidad de propagación de las ondas ultrasónicas en un gas, depende mayoritariamente de la temperatura y no de la presión del gas. La tabla 1.17 muestra la velocidad del sonido en distintos gases a 0°C.

Como la gran mayoría de los sensores de ultrasonidos trabajan en el aire, se puede calcular la velocidad v [m/seg] del ultrasonido en el aire seco, en función de la temperatura T [°C], mediante la fórmula:

$$v = v_0 \left(1 + \frac{T}{273,2} \right)^{\frac{1}{2}}$$

en la que v_0 representa la velocidad del sonido en el aire a 0° C, cuyo valor es de 331,6m/seg, y a partir de la cual se obtiene la ecuación:

$$v[m/seg] \approx v_0 + 0,58 \left[\frac{m}{seg \cdot ^\circ C} \right] \cdot T[^\circ C]$$

que establece que la velocidad del sonido en el aire varía con la temperatura del orden del 0,2% por °C.

Gas	Velocidad [m/s]
Aire	331
Amoníaco	414
Argón	301
Bióxido de Carbono	257
Clorina	205
Dioxido de Carbono	258
Disulfuro Del Carbón	184
Etileno	313
Gas Que ilumina	490
Helio	969
Hidrógeno	1269
Metano	431
Monóxido De Carbono	337
Neón	434
Nitrógeno	334
Óxido Nítrico	324
Óxido Nitroso	261
Oxígeno	317
Vapor (100°C)	403

Tabla 7.17 Velocidad del sonido para diferentes tipos de gases.

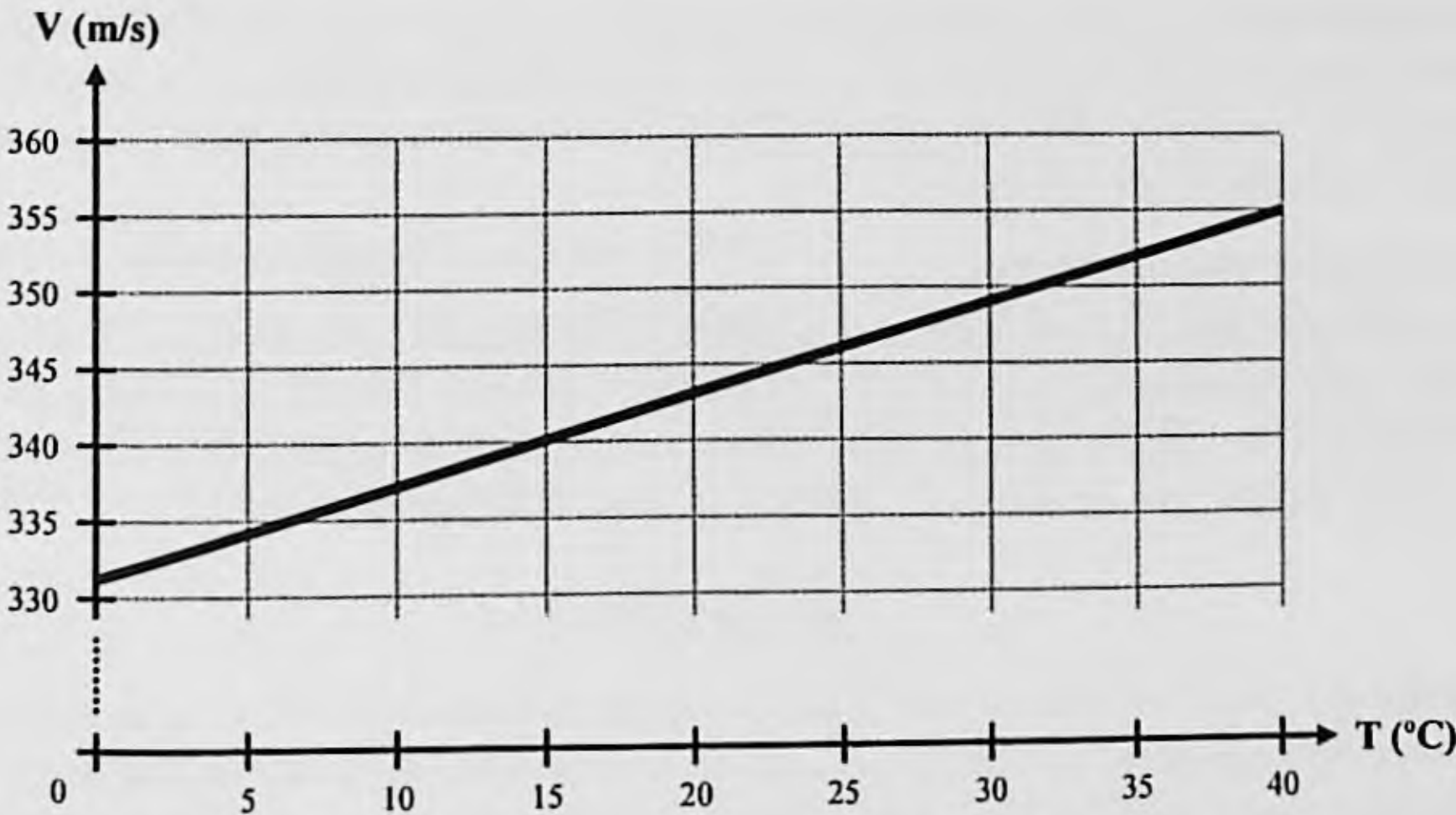


Figura 7.169 Variación de la velocidad del sonido en el aire con la temperatura.

La figura 7.169 muestra la variación de la velocidad del ultrasonido en el aire en función de la temperatura. Debido a ello son frecuentes los sensores ultrasónicos de tipo eco, que se describen en el apartado 7.7.2.3.7.4, que miden la temperatura para compensar su efecto al calcular el valor de la distancia.

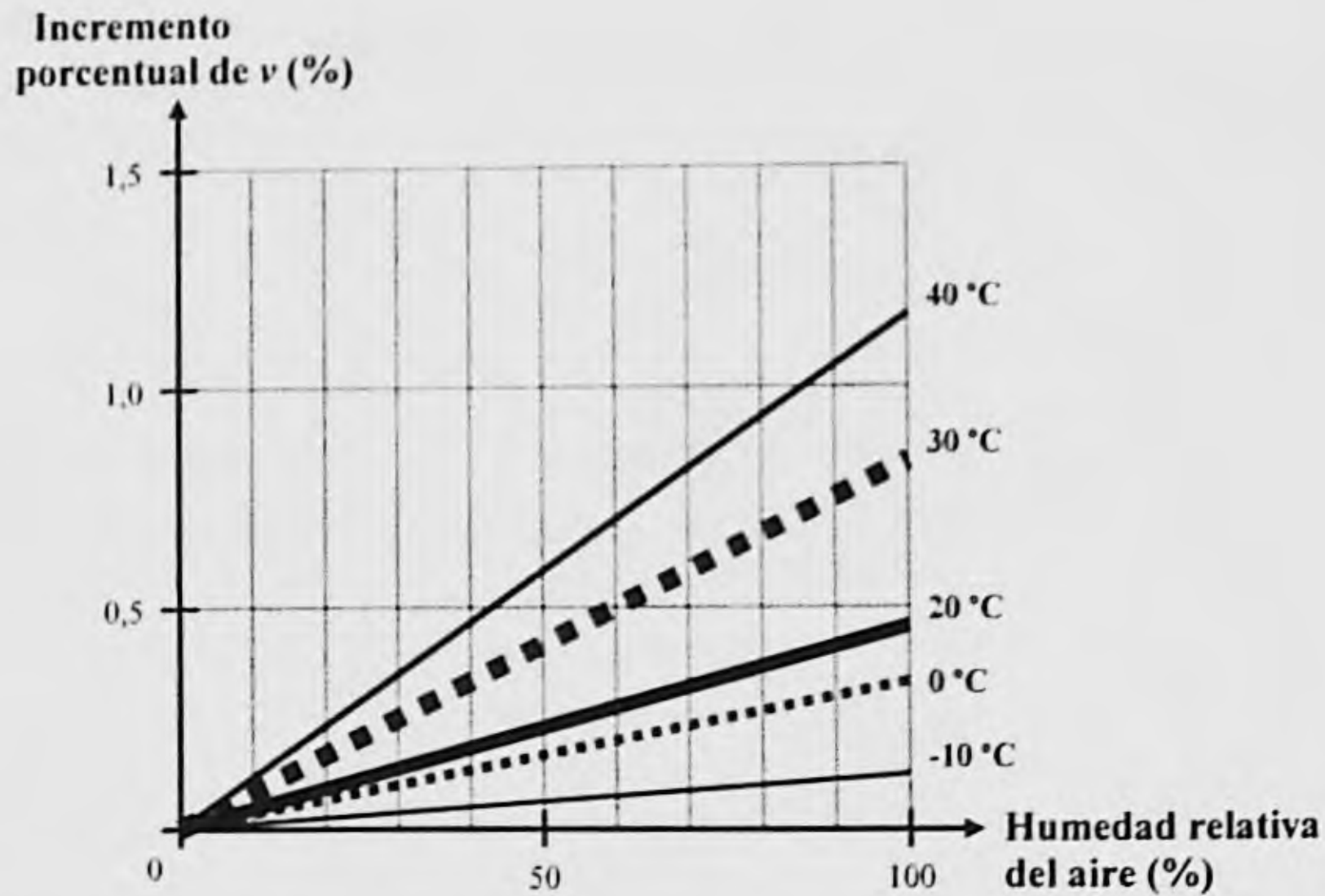


Figura 7.170 Variación de la velocidad del sonido en el aire en función de la humedad relativa y la temperatura.

La velocidad del sonido en el aire depende también de la humedad relativa del mismo tal como se indica en la figura 7.170.

Atenuación del sonido en el aire

La presión sonora P_s de una señal ultrasónica se mide en micropascales (μPa) y se reduce cuando se propaga a través de un medio. En particular en el aire, la presión sonora se reduce con la distancia por ser absorbida por el mismo y debido también a la expansión del haz ultrasónico, tal como se describe a continuación:

• Atenuación por absorción

Los ultrasonidos se atenúan en el aire debido a las pérdidas por fricción. El valor de la atenuación aumenta con la frecuencia de la señal y con la humedad, lo que debe ser tenido en cuenta al utilizar los sensores ultrasónicos. La amplitud de la señal de presión sonora P_s decrece exponencialmente con la distancia d de acuerdo con la ecuación:

$$P_s = P_0 \cdot e^{-\alpha d}$$

en la que P_0 representa la presión sonora a la salida del emisor ($d = 0$) y “ α ” es el coeficiente de atenuación que se mide en m^{-1} . Este coeficiente crece linealmente con la frecuencia y también depende de la humedad, tal como se indica en la figura 7.171, que muestra la variación del coeficiente de atenuación con la frecuencia para distintos valores de la humedad relativa del aire.

El valor máximo del coeficiente de atenuación en el aire a temperatura ambiente, para cualquier humedad y frecuencias inferiores a 50 KHz, viene dado aproximadamente por la ecuación:

$$\alpha(f) = 0,0328 \cdot f$$

en la que el coeficiente de atenuación “ α ” se mide en dB/Km y la frecuencia f en KHz.

El valor máximo del coeficiente de atenuación para señales de ultrasonidos comprendidas entre 50 KHz y 300 KHz viene dado aproximadamente por la ecuación:

$$\alpha(f) = 0,7217 \cdot f$$

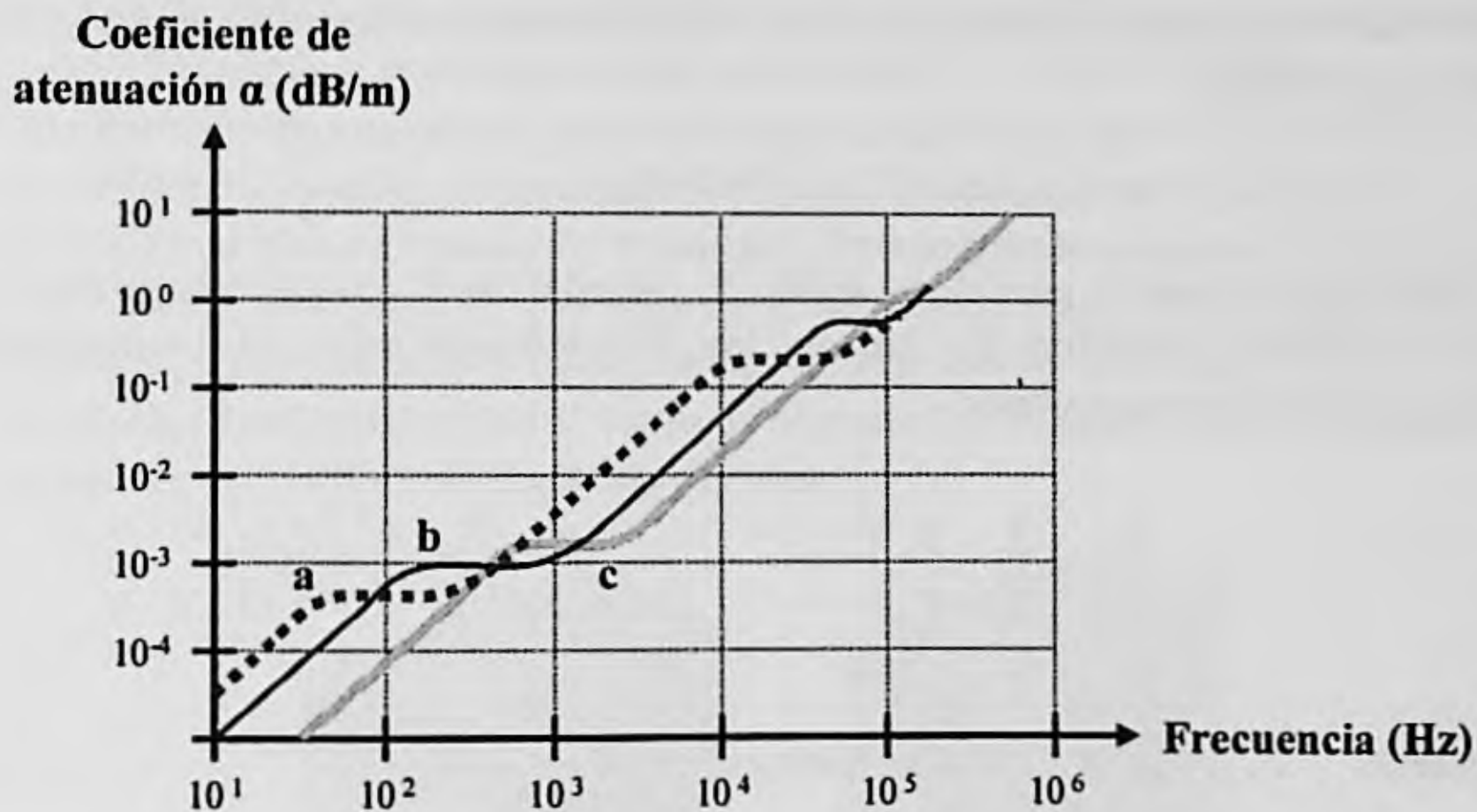


Figura 7.171 Valor del coeficiente de atenuación de los ultrasonidos en función de la frecuencia: a) Con humedad relativa del aire del 10%; b) Con humedad relativa del aire del 40%; c) Con humedad relativa del aire del 80%.

- **Atenuación por expansión del haz sónico**

La presión del sonido se atenúa también por expansión del haz ultrasónico. Esto es debido a que, a medida que aumenta la distancia al emisor, el haz ultrasónico se abre en forma de cono tal como se indica anteriormente. La presión sonora P_s , también se puede expresar en decibelios mediante la ecuación:

$$P_s[dB] = P_i \cdot 20 \cdot \log\left(\frac{d}{d_i}\right)$$

en la que P_i representa la presión sonora en decibelios a una distancia determinada d_i del sensor, que suele ser del orden de 30 cm y P_s representa la presión sonora a la distancia d del sensor.

P_i se obtiene a partir de la presión sonora p en micropascales (μPa), medida a la distancia d_i del sensor, de acuerdo con la expresión:

$$P_i[dB] = 20 \cdot \log(p)$$

Si se tienen en cuenta ambos efectos, la presión del ultrasonido en el aire responde a la ecuación:

$$P_s[dB] = P_i \cdot 20 \cdot \log\left(\frac{d}{d_i}\right) \cdot \alpha \cdot d$$

en la que α es el coeficiente de atenuación y P_i , que representa la presión del sonido a una distancia d del emisor, se mide en decibelios (dB).

Reflexión de los ultrasonidos

Se define la impedancia acústica Z de un medio como el producto de la densidad ρ por la velocidad del sonido c :

$$Z = \rho \cdot c$$

Si una señal ultrasónica pasa de un medio de impedancia Z_1 a otro de impedancia Z_2 , además de ser absorbida, se refleja. En este caso los coeficientes de reflexión R y transmisión T se obtienen mediante las expresiones:

$$R = \frac{I_r}{I_i} = \left(\frac{Z_1 - Z_2}{Z_1 + Z_2} \right)^2 \quad T = \frac{I_t}{I_i} = \frac{4Z_1Z_2}{(Z_1 + Z_2)^2}$$

en las cuales I_i : Intensidad del ultrasonido incidente

I_r : Intensidad del ultrasonido reflejado

I_t : Intensidad del ultrasonido transmitido

Cuanto mayor es la diferencia de impedancia entre los dos medios mayor es la intensidad de la onda reflejada.

Ruido de fondo

El nivel del ruido de fondo disminuye a medida que aumenta la frecuencia, debido a que en el entorno se produce menos ruido a altas frecuencias y el que se produce lo atenúa mucho el aire.

Los sensores ultrasónicos emiten una secuencia de impulsos en forma de un tren que se repite periódicamente y según el modo de funcionamiento pueden ser de dos tipos que se describen seguidamente.

7.7.2.3.7.3 Sensores ultrasónicos de barrera

Los sensores ultrasónicos de proximidad del tipo barrera (*Through-beam ultrasonic sensor*) están constituidos por un emisor y un receptor enfrentados y separados una determinada distancia que depende del tipo de sensor. El receptor deja de recibir la señal ultrasónica emitida por el emisor, cuando se sitúa entre ambos el objeto a detectar. Son, por lo tanto, sensores detectores de objetos cuya salida es del tipo todo-nada.

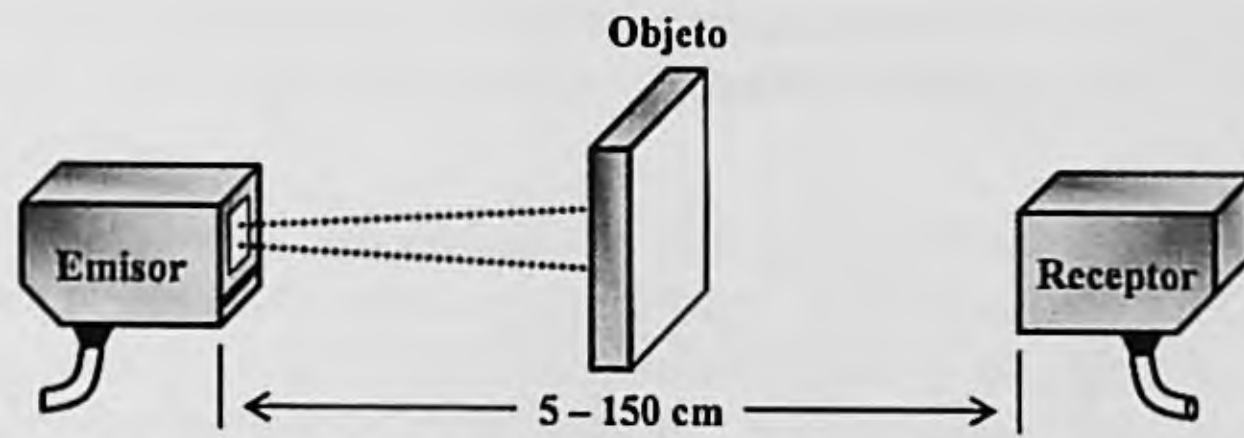


Figura 1.172 Modo de operación de un sensor ultrasónico de barrera.

La figura 7.172 muestra el modo de funcionamiento de un sensor de este tipo, y la figura 7.173 el sensor de barrera ultrasónica 3RG6243 de la familia Bero, en el que la distancia de detección entre el emisor y el receptor puede variar entre 5 y 150 cm. La figura 7.174 muestra el rango de operación de este sensor, que representa la zona en la que se puede situar el receptor respecto del emisor cuando ambos están alineados. Debido al cono de radiación, dicha zona depende de la distancia entre el emisor y el receptor. Para lograr un campo de recepción óptimo, el eje del receptor se debe inclinar con respecto al del emisor a medida que aquél se desplaza transversalmente, porque la señal ultrasónica incide así perpendicularmente sobre él. La figura 7.175 muestra la mejora obtenida en el rango de operación del sensor 3RG6243 cuando se optimiza la posición del receptor con respecto al emisor.

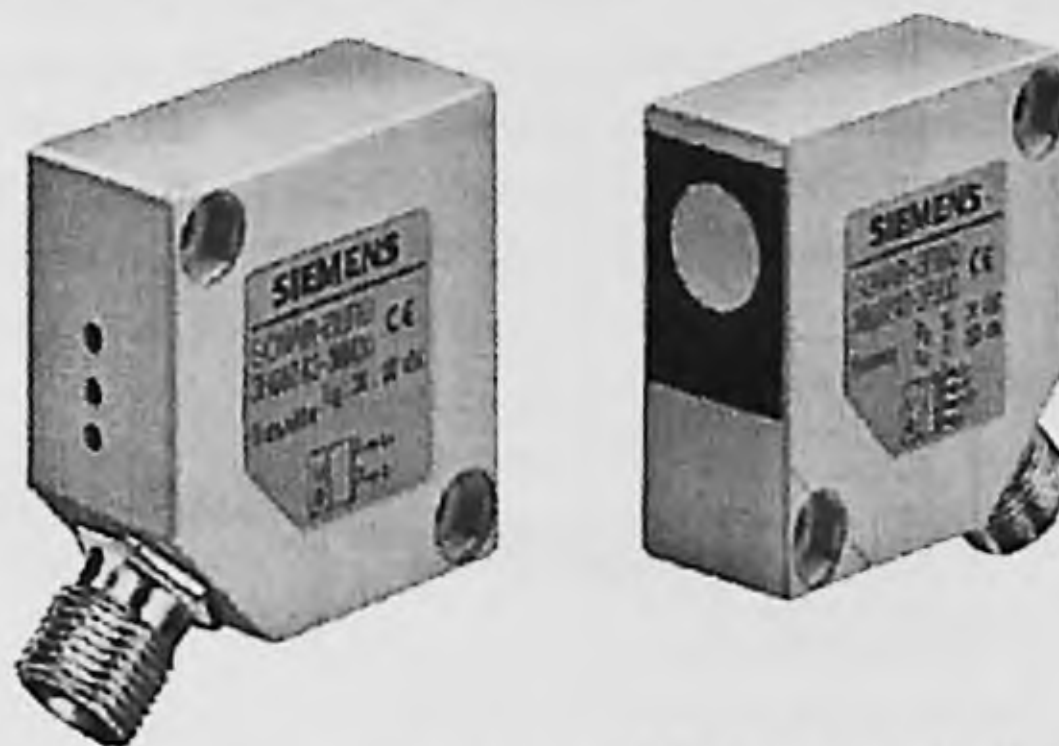


Figura 7.173 Sensor de barrera ultrasónica tipo 3RG6243 (cortesía de Siemens).

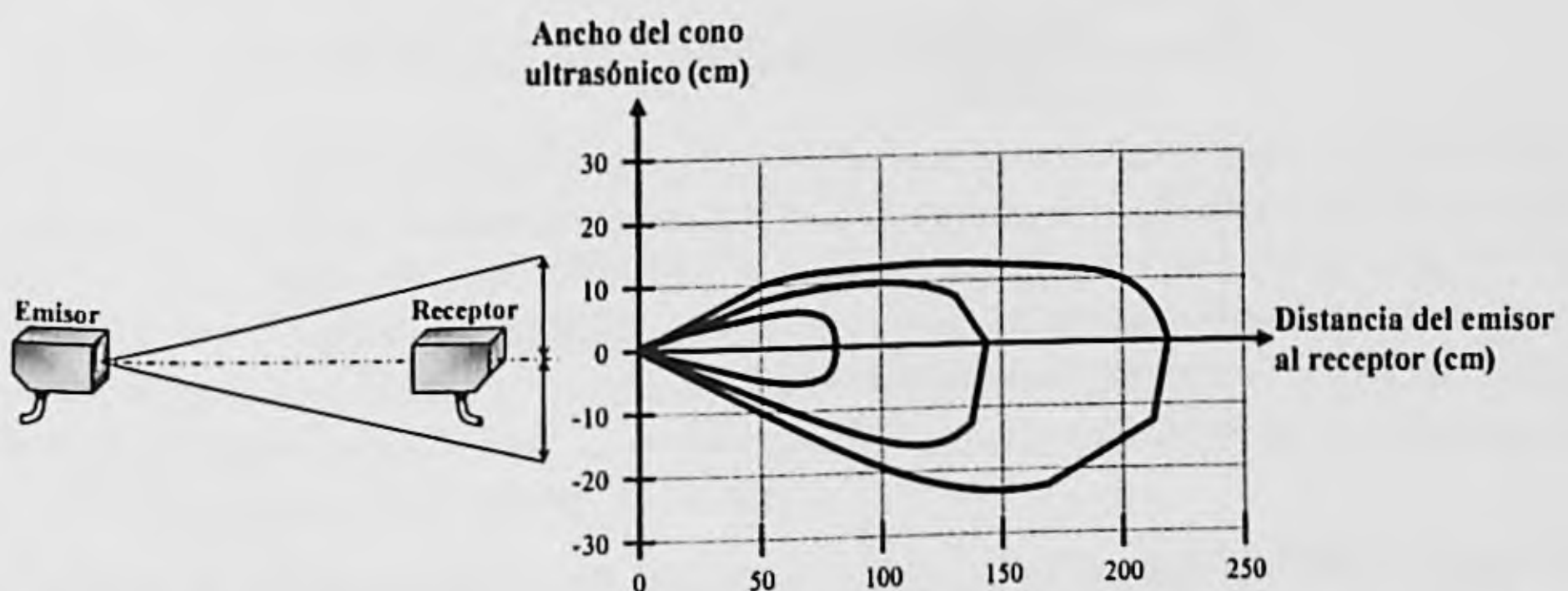


Figura 7.174 Rango de operación del sensor 3RG6243 con el emisor y el receptor alineados (cortesía de Siemens).

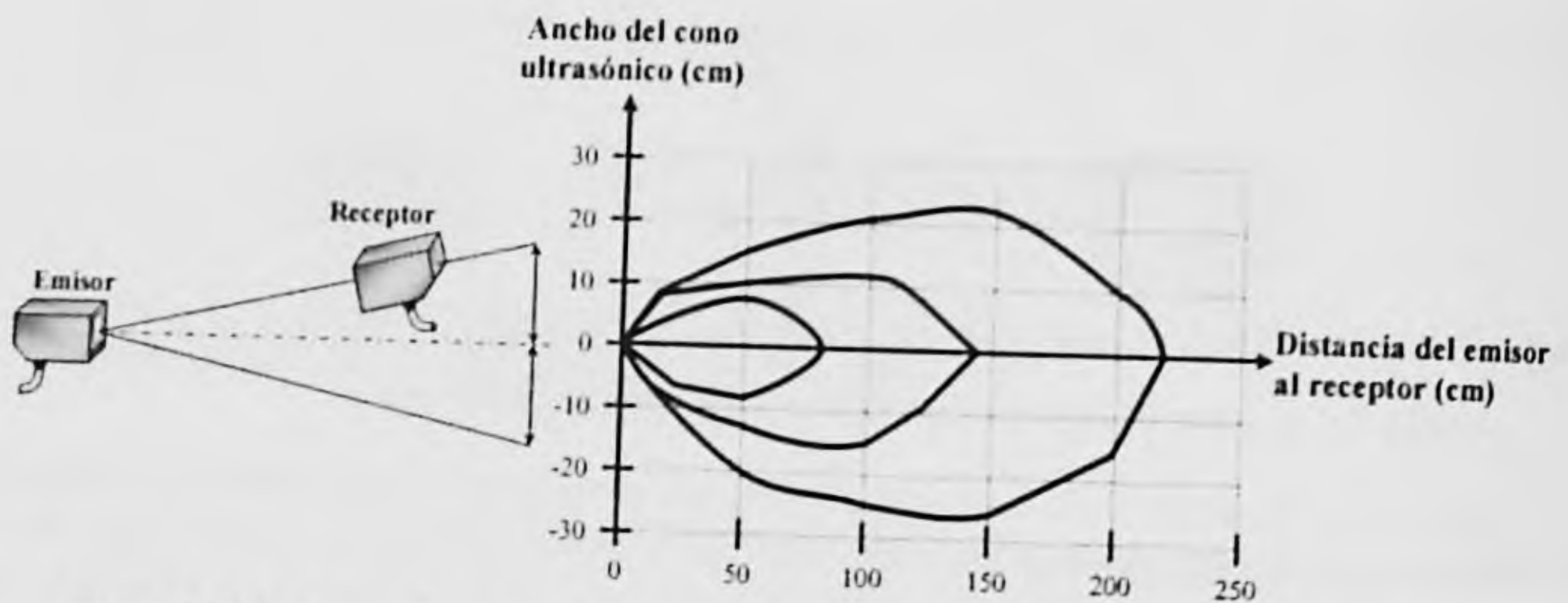


Figura 7.175 Rango de operación del sensor 3RG6243 con el emisor y el receptor desalineados (cortesía de Siemens).

7.7.2.3.7.4 Sensores ultrasónicos detectores de eco

Los sensores ultrasónicos de tipo eco (*Ultrasonic distance measurement sensors*), no solo detectan la presencia de un objeto, sino que, además, miden la distancia a la que se encuentra, para lo cual el emisor y el receptor están ubicados en la misma carcasa. La señal ultrasónica que emite el emisor se refleja en el objeto a detectar y el eco se recibe en el receptor tal como se muestra en la figura 7.176a. El receptor está asociado a un circuito electrónico que mide el tiempo que tarda en recibir el eco (Figura 7.176b). La máxima frecuencia de repetición de los trenes de impulsos depende del tipo de sensor y suele variar entre 10 Hz y 150 Hz.

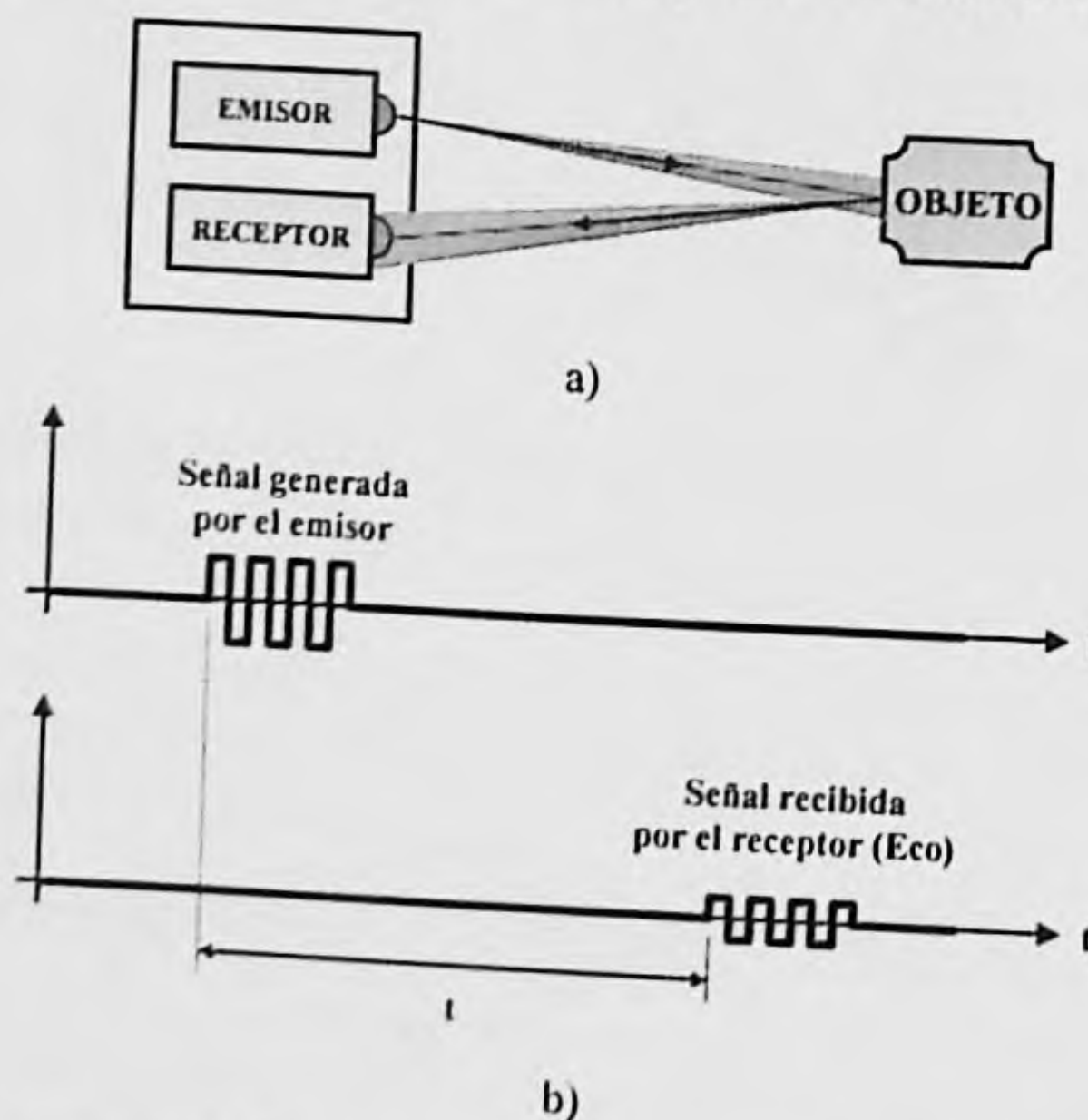


Figura 7.176 Descripción del funcionamiento de un sensor de ultrasonidos de tipo eco. a) Emisión y recepción del haz ultrasónico; b) Señal generada por el emisor y señal recibida por el receptor.

Conocida la velocidad del sonido en el aire (del orden de 340 m/seg) se determina la posición del objeto. En la figura 7.177 se muestra el diagrama de bloques del sensor que consta de los siguientes elementos:

- Un procesador digital que, en este caso es un microcontrolador [MAND 07], que genera trenes de impulsos que se aplican a un transmisor y recibe la señal de eco procedente de un receptor. Además recibe señales de configuración que suelen proceder de un computador y genera una señal todo-nada que indica que el haz ultrasónico se ha reflejado en un objeto y una señal analógica que indica la distancia a la que se encuentra el mismo.
- Un circuito receptor formado por un cristal piezoeléctrico, un amplificador y un comparador cuya salida se conecta al procesador digital.
- Un circuito transmisor que consta de un amplificador, que recibe los trenes de impulsos generados por el procesador digital, y un cristal piezoeléctrico (que puede ser el mismo utilizado en el receptor), que los convierte en una señal ultrasónica.

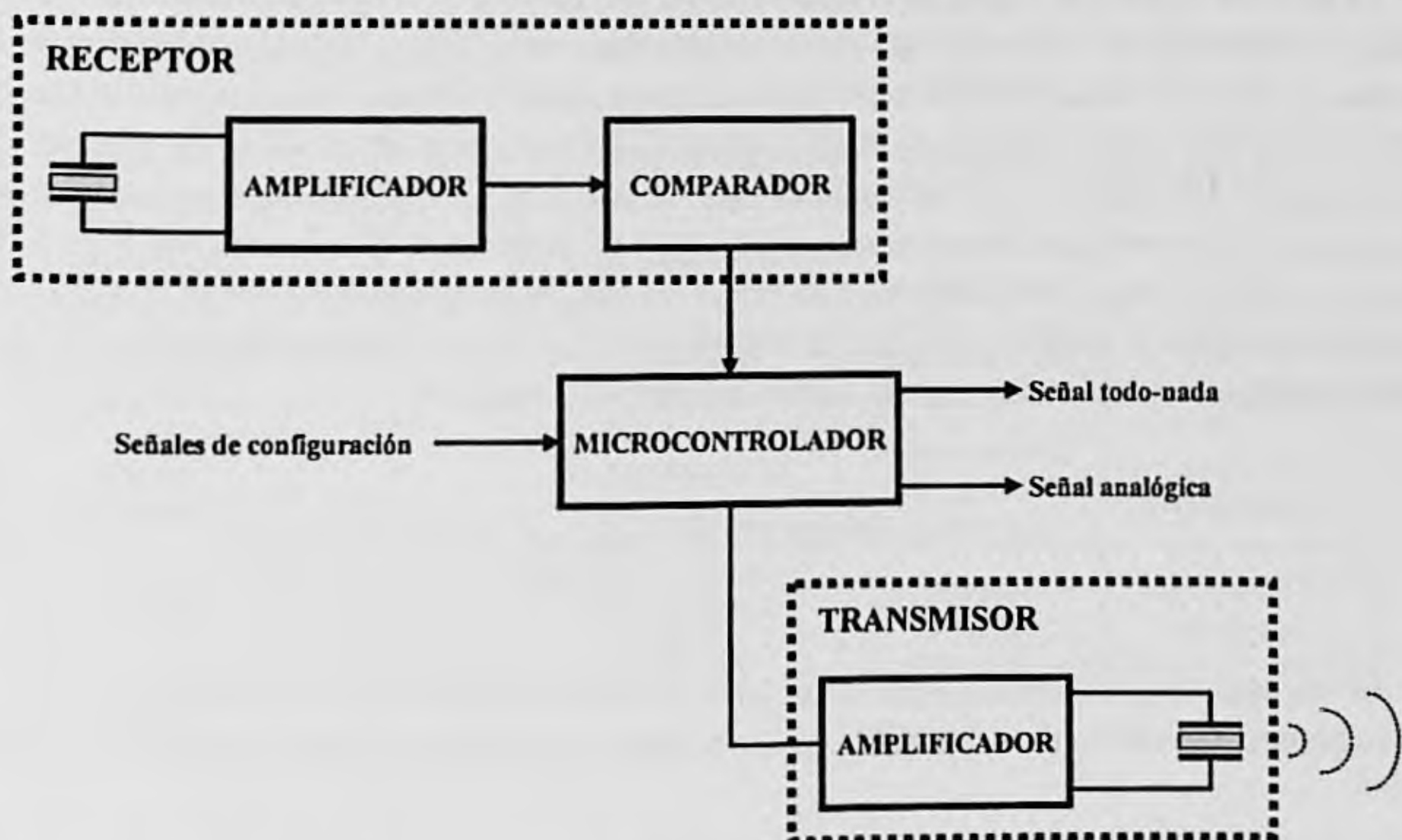


Figura 7.177 Diagrama de bloques de un sensor ultrasónico.

En la figura 7.178 se indica el campo de trabajo (*Sensing range*) típico de un sensor de ultrasonidos, que está comprendido entre la distancia máxima y la mínima especificadas por el fabricante. Los fabricantes establecen una zona muerta o ciega (*Blind zone*) en la que no se garantiza la detección del objeto porque el eco se mezcla con el haz transmitido. Por estar además implementado con un microcontrolador, el usuario puede configurar el sensor, tal como se indica anteriormente (Figura 7.162), y, entre otros parámetros, establecer la zona de detección de acuerdo con las exigencias de la aplicación.

Mediante sensores ultrasónicos de tipo eco se pueden detectar líquidos, sólidos o materiales polvorientos de distintas formas y colores, y solamente es necesario que tengan unas mínimas características de reflexión de los ultrasonidos.

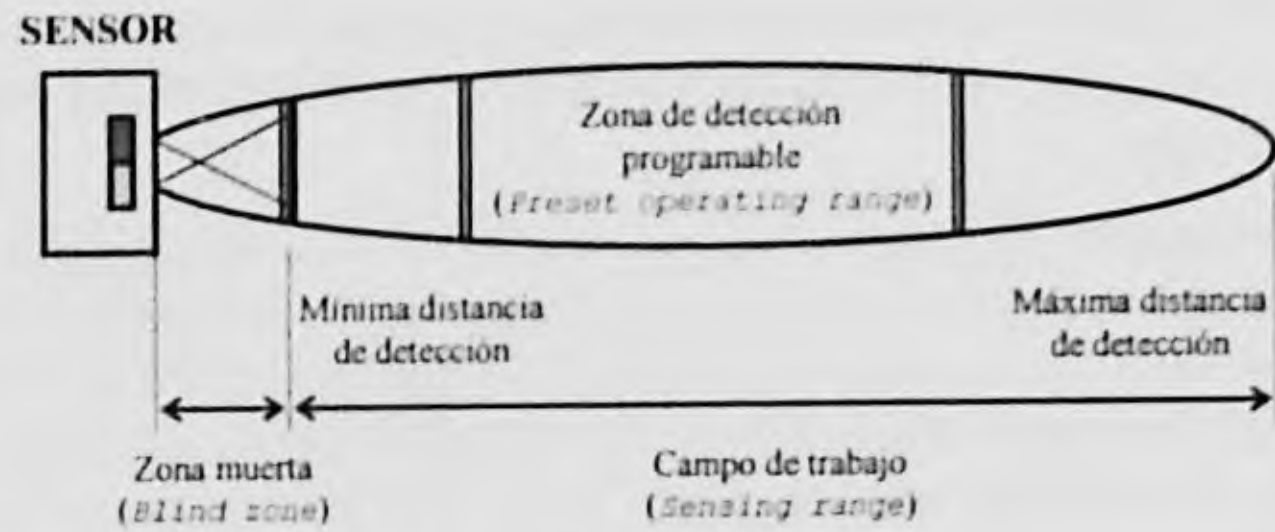


Figura 7.178 Campo de trabajo de un sensor ultrasónico.

En la medida de distancias mediante sensores ultrasónicos del tipo eco influyen un conjunto de parámetros que se indican seguidamente.

Influencia del objeto

El tipo de objeto en el que se refleja la señal ultrasónica y el tipo de superficie del mismo influye en el rango de operación del sensor ultrasónico de tipo eco. Tal como se indica anteriormente, el objeto a detectar debe tener unas mínimas características de reflexión del sonido. La reflexión de una señal ultrasónica en una gran superficie plana, se produce tal como se indica en la figura 7.179. El haz total reflejado es equivalente a un emisor virtual situado al doble de la distancia de la superficie. Para que esto se cumpla es importante que la superficie de reflexión sea mayor que el haz, para asegurar una reflexión total y perpendicular al haz del ultrasonido. Cuando se refleja la totalidad del haz, la presión del sonido en el receptor, teniendo en cuenta el recorrido de ida y vuelta del sonido, viene dada por la ecuación:

$$P_s[dB] = P_i \cdot 40 \cdot \log\left(\frac{d}{d_i}\right) \cdot \alpha \cdot d$$

En las aplicaciones industriales los objetos poseen distintas formas y tamaños, y en general no se comportan como una superficie plana de gran tamaño que refleja todo el haz. El tamaño

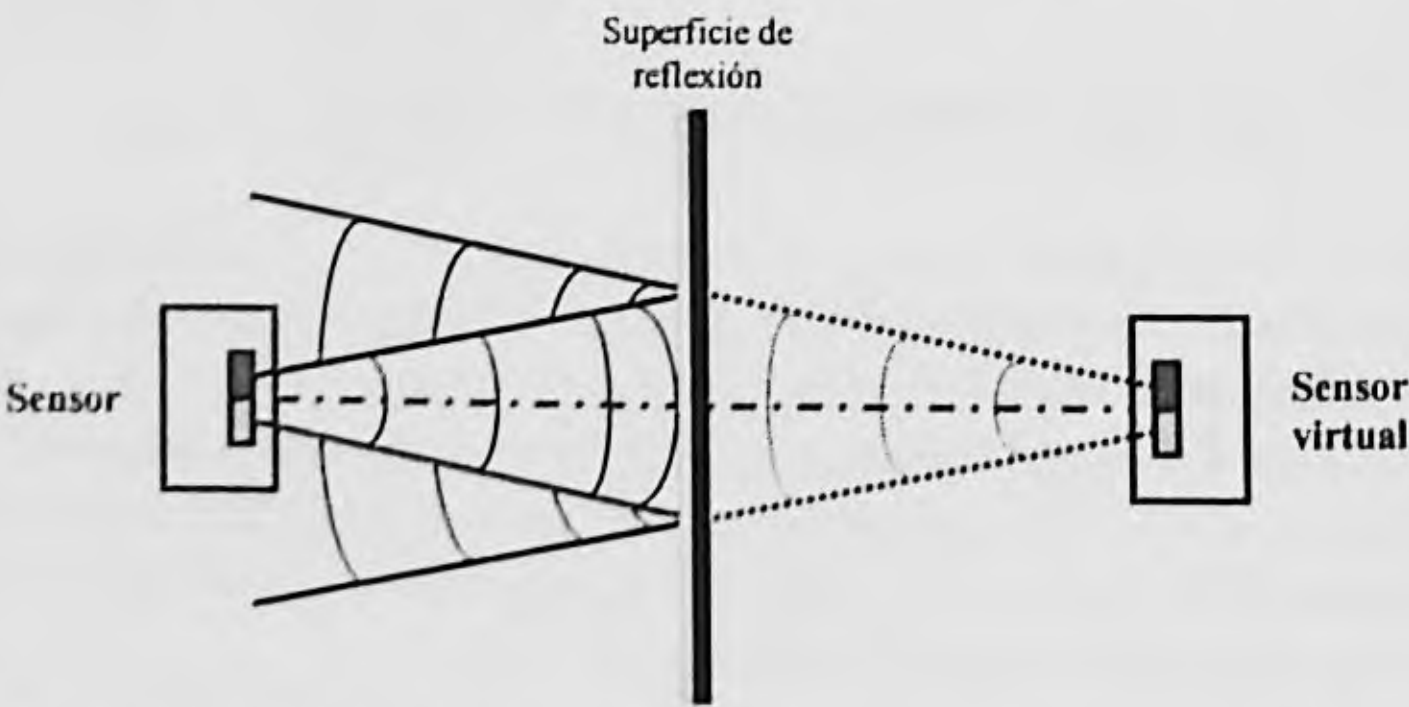


Figura 7.179 Reflexión de una señal ultrasónica sobre una gran superficie plana.

y la forma del objeto influyen en la reflexión del haz ultrasónico y por tanto la distancia de detección depende de ellos. En este caso se debe tener en cuenta el índice de reflexión del objeto (*Target Strength*) IR , que depende de la impedancia acústica del objeto y representa la capacidad del mismo para reflejar el ultrasonido. IR se define mediante la ecuación:

$$IR = 10 \log \frac{\text{Intensidad del sonido reflejado}}{\text{Intensidad del sonido emitido}}$$

La tabla 7.18 muestra los valores teóricos de IR para objetos de diferentes formas físicas. En dicha tabla el parámetro K responde a la expresión:

$$K = \frac{2 \cdot \pi}{\lambda}$$

en la que λ es la longitud de onda de la señal ultrasónica.

Forma física	Índice de Reflexión (IR)	Definiciones	Dirección incidente del sonido	Condiciones
Esfera	$\frac{R^2}{4}$	R = Radio de la esfera	Cualquiera	$KR > 1$ $d > R$
Cilindro infinitamente largo	$\frac{R \cdot d}{2}$	R = Radio del cilindro	Perpendicular al eje del cilindro	$KR > 1$ $d > R$
Cilindro de longitud finita	$\frac{R \cdot L^2}{2 \cdot \lambda}$	L = Longitud R = Radio	Perpendicular al eje del cilindro	$KR > 1$ $d > L^2/\lambda$
Objeto convexo	$\frac{S}{16 \cdot \pi}$	S = Superficie total del objeto	Valor medio de todas las direcciones	Todas las dimensiones $> \lambda$
Elipsoide	$\left(\frac{b \cdot c}{2 \cdot a}\right)^2$	a, b, c => Semiejes del elipsoide ($a \geq b \geq c$)	Perpendicular al eje mayor	$Ka, Kb, Kc > 1$ $d > a, b, c$

Tabla 7.18 Índice de reflexión de objetos de diversas formas físicas.

En la realidad los objetos presentan formas compuestas, debido a lo cual los datos de dicha tabla sólo se deben considerar orientativos.

Teniendo en cuenta todos los factores que se acaban de indicar, se obtiene la ecuación:

$$P_s(d)[dB] = P_i(d_i) \cdot 40 \cdot \log\left(\frac{d}{d_i}\right) \cdot \alpha \cdot d + IR$$

que establece el valor de la presión sonora en el receptor de un sensor de tipo eco.

La rugosidad del objeto es otra característica a tener en cuenta. Las señales de ultrasonidos son de corta longitud de onda. Por ejemplo, a una señal de ultrasonidos de 200KHz le corresponde una longitud de onda de 1,7 mm para una velocidad del sonido de 340m/seg. Dada su corta longitud de onda, las ondas ultrasónicas se comportan de forma similar a las ondas de luz,

y, por ello, también se les aplican las leyes de la óptica geométrica que establecen que el ángulo de incidencia es igual al de reflexión. De igual forma, la estructura de la superficie del objeto a detectar, así como la dirección en la que se refleja el haz ultrasónico, son características importantes a tener en cuenta en la utilización de estos sensores. El objeto cuya superficie es lisa y regular debe ser perpendicular al cono ultrasónico y su inclinación máxima respecto a él debe estar comprendida entre $\pm 3^\circ$ y $\pm 5^\circ$, tal como se muestra en la figura 7.180. Si por el contrario, la superficie del objeto tiene una rugosidad de valor comprendido entre $1/4$ y $1/6$ de la longitud de onda del sonido, las señales se reflejan de forma difusa. Por ello, cuando la superficie del objeto a detectar es muy rugosa, como por ejemplo arena, áridos, etc., el ángulo de inclinación máximo de la misma con respecto al eje del cono ultrasónico puede llegar a ser de 45° , tal como se indica en la figura 7.181.

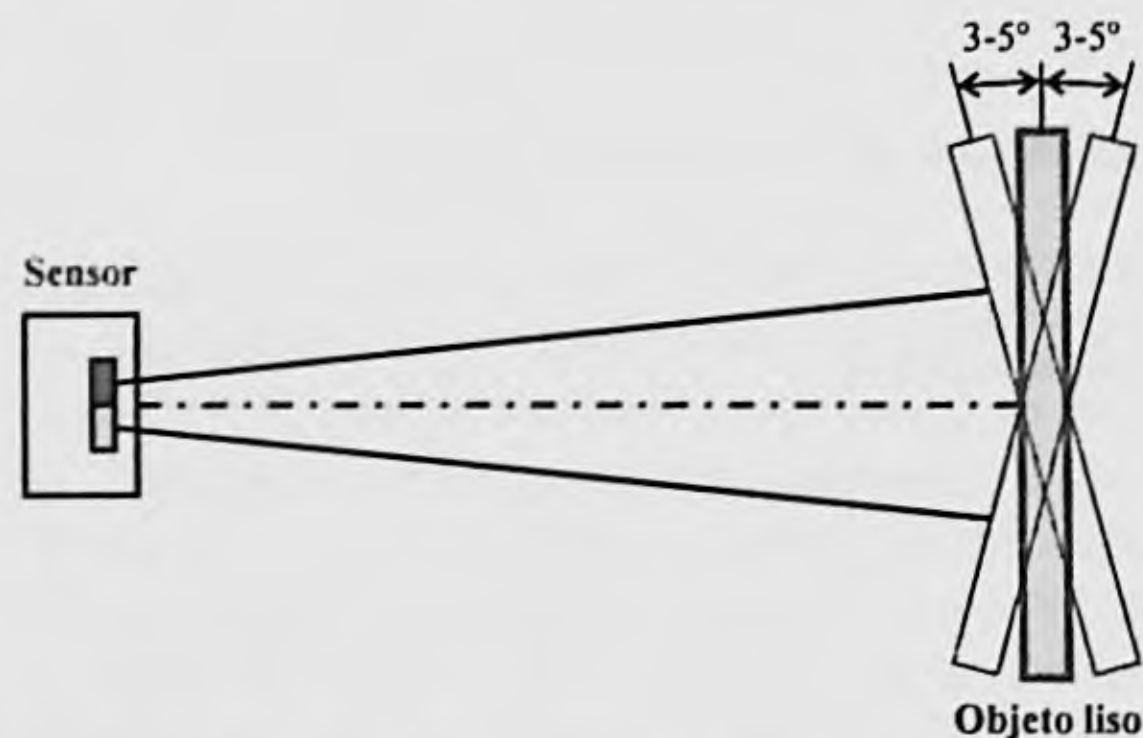


Figura 7.180 Efecto de la inclinación de un objeto liso sobre una señal ultrasónica.



Figura 7.181 Efecto de la inclinación de un objeto muy rugoso sobre una señal ultrasónica.

Los objetos porosos, cuya densidad es en general reducida, como por ejemplo las telas, las espumas, etc., reflejan mal los ultrasonidos, lo que hace que solo se puedan detectar a distancias pequeñas. En la figura 7.182, se muestra la influencia del tamaño de tres objetos de distinta densidad en la distancia de detección.

De todo lo indicado anteriormente se deduce la dificultad que tienen los fabricantes para proporcionar el campo de trabajo o rango de operación de sus sensores ultrasónicos porque en ello influyen, además de las condiciones ambientales, la forma, rugosidad e inclinación del objeto con respecto al cono ultrasónico. Los fabricantes deben tener en cuenta, además, las dispersiones de fabricación y por todo ello optan, en general, por proporcionar, para unas determinadas condiciones ambientales, gráficas del campo de trabajo en las que se representa el cono

ultrasónico mínimo, medio y máximo para distintas distancias entre el objeto y el sensor. Como, además, la forma del objeto también tiene una influencia importante, los fabricantes suelen suministrar una gráfica diferente para objetos de distintas formas como por ejemplo cilíndrica y plana. Por ejemplo, los objetos estándar de 2cm × 2cm se suelen utilizar para distancias nominales de detección que no superan los 130cm, mientras que para distancias mayores se suelen utilizar objetos de 10cm × 10cm.

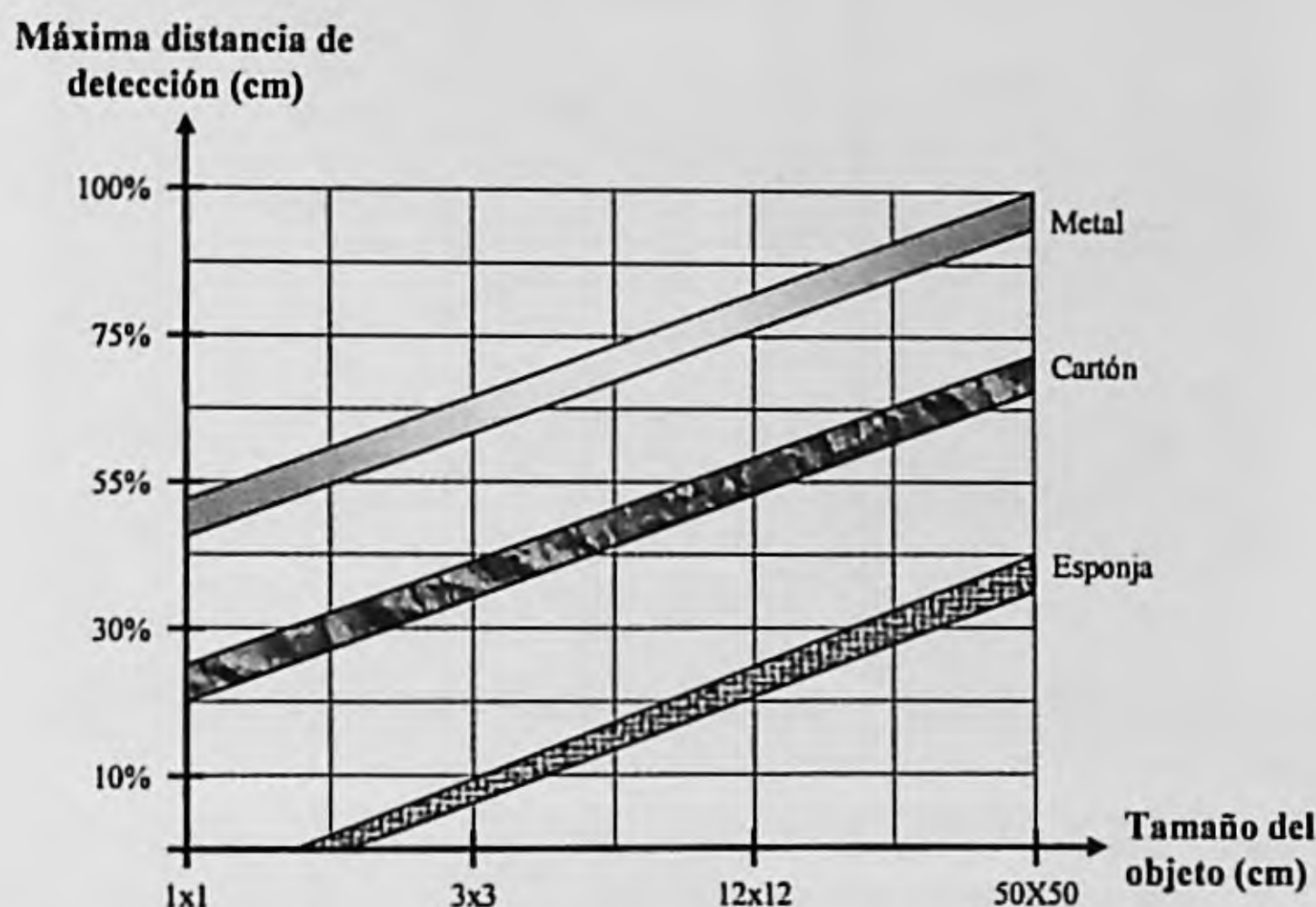


Figura 7.182 Influencia del tamaño del objeto y del tipo de material del mismo en la distancia de detección.

Como ejemplo, las figuras 7.183 a 7.186 muestran el campo de trabajo de los sensores ultrasónicos de tipo eco de la familia Bero de Siemens para un objeto plano de 2cm × 2cm, y las 7.187 y 7.188 para un objeto cilíndrico de 8cm de diámetro. Las figuras 7.183 y 7.184 indican el campo de trabajo de un sensor, cuya distancia nominal máxima es de 37cm, en función de la inclinación con respecto al eje central del cono de emisión. En la figura 7.183 el campo de trabajo es máximo porque el objeto está colocado perpendicularmente al haz que recibe y en la 7.184 el campo de trabajo disminuye debido a que la reflexión no es óptima. Las figuras 7.185 y 7.186 hacen lo mismo para una distancia de detección nominal máxima de 120cm. Las figuras 7.187 y 7.188 se diferencian también en la distancia de detección del objeto cilíndrico.

Influencia de la temperatura

En el apartado 7.7.2.3.7.2 se indica la influencia que las variaciones de la temperatura tienen en la velocidad del sonido y por tanto en el tiempo que el receptor tarda en recibir el eco. Si la temperatura aumenta también lo hace la velocidad del sonido y, en consecuencia, disminuye el tiempo que tarda el ultrasonido en recorrer una determinada distancia, que es el parámetro utilizado para medirla. En efecto, la velocidad del sonido en el aire es de 331,6m/s a 0° C y de 332,206m/s a 1 °C. Por ello si no se tiene en cuenta este cambio se comete un error de 0,0018 m por °C y por cada metro que separa el sensor del objeto a detectar, o lo que es lo mismo un error e del 0,18% por cada grado centígrado. Por lo tanto el error en función de la temperatura y la distancia se obtiene mediante la ecuación:

$$e = 0,0018 \cdot \Delta T [^{\circ}C] \cdot d [m]$$

Debido a ello diversos sensores ultrasónicos de proximidad de tipo eco miden el valor de la temperatura y mediante la fórmula anterior obtienen la distancia real. Son un ejemplo de este tipo de sensores ultrasónicos los M30 K2 de la familia PXS300 (Figura 7.189).

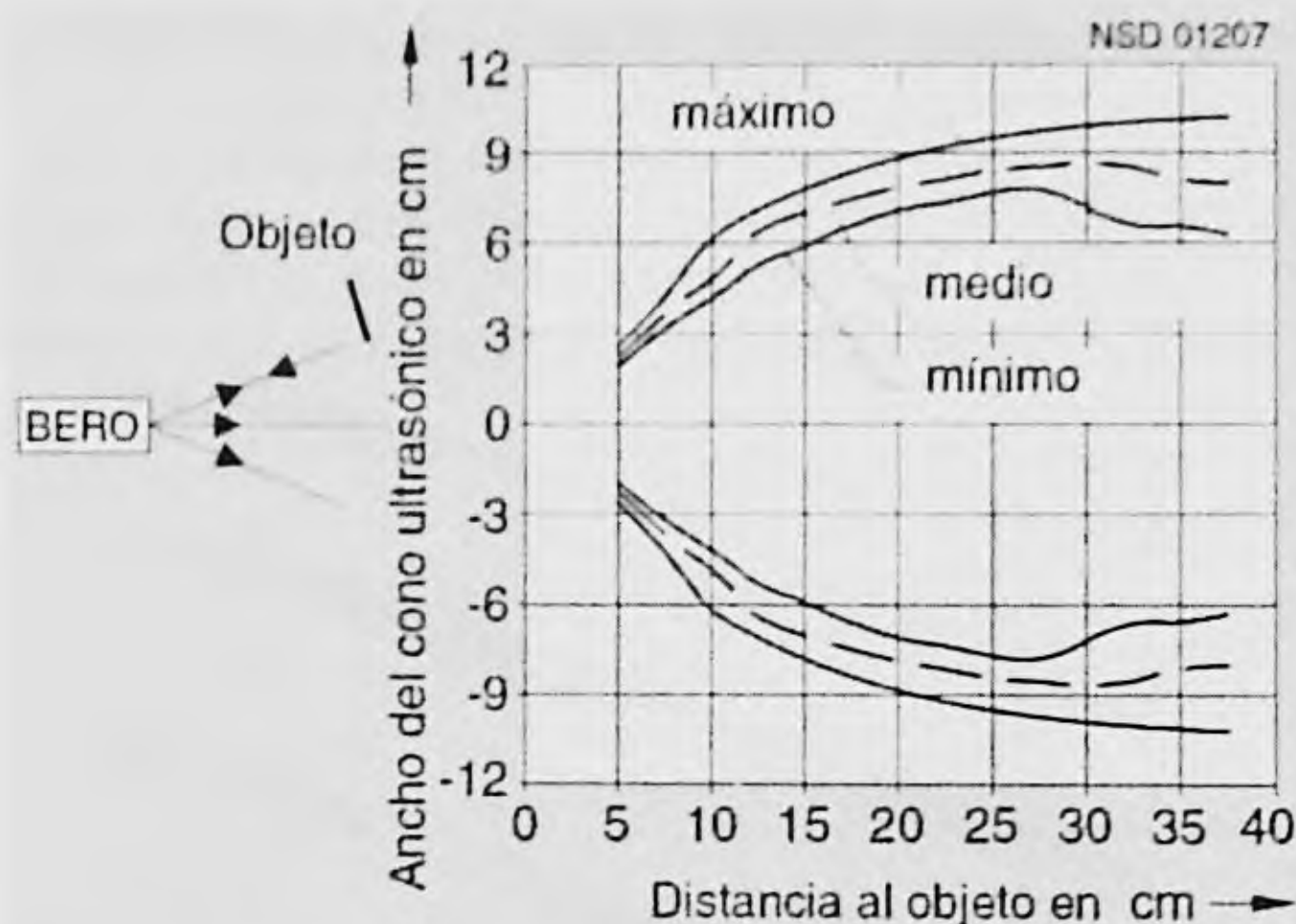


Figura 7.183 Campo de trabajo de un sensor ultrasónico de la familia Bero cuya distancia nominal de detección está comprendida entre 5 y 37 cm, utilizado en una aplicación en la que el objeto es plano y esta colocado perpendicularmente a la línea que lo une con el sensor (cortesía de Siemens).

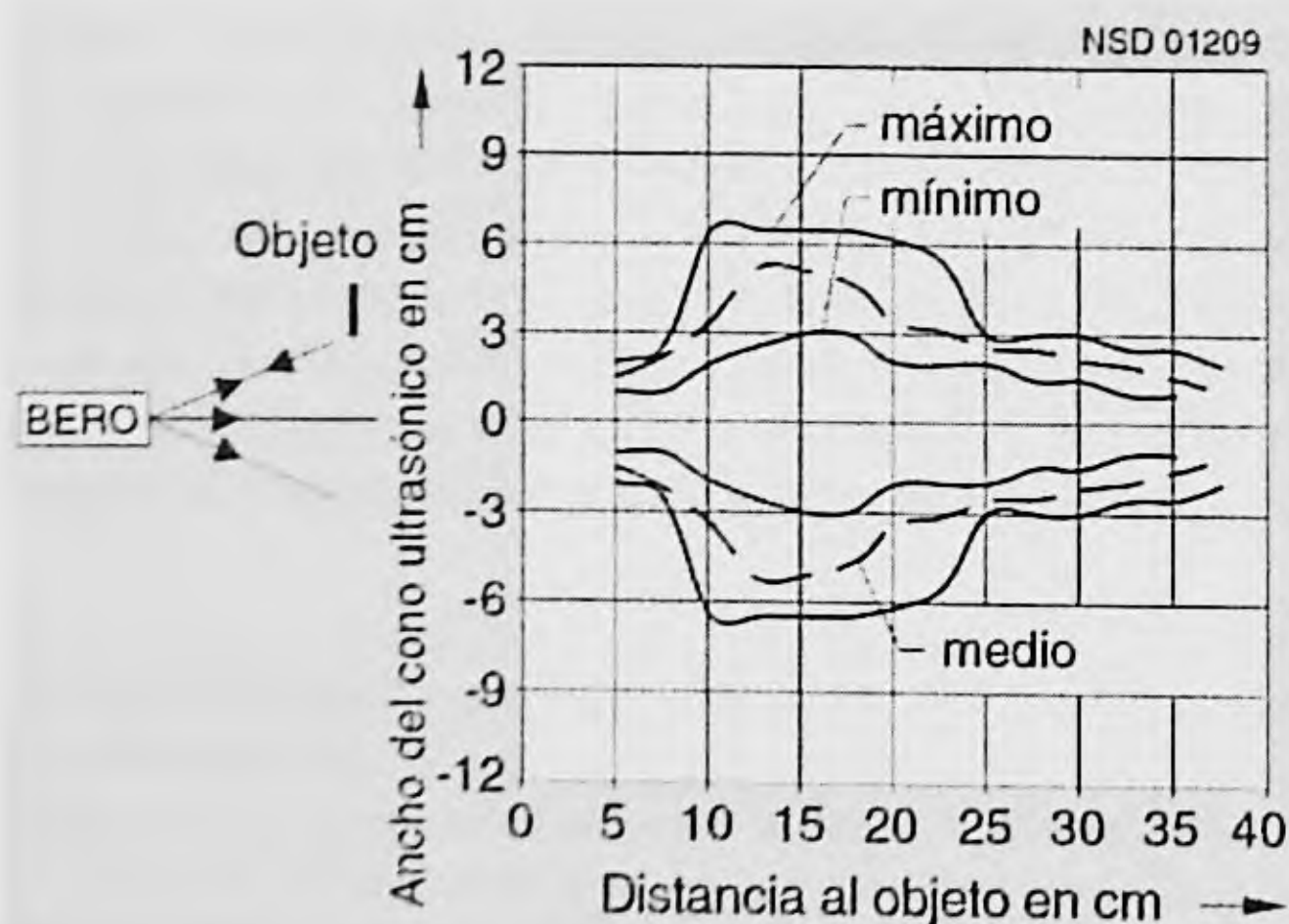


Figura 7.184 Campo de trabajo de un sensor ultrasónico de la familia Bero cuya distancia nominal de detección está comprendida entre 5 y 37 cm, utilizado en una aplicación en la que el objeto es plano y esta colocado perpendicularmente al eje del cono de emisión del sensor (cortesía de Siemens).

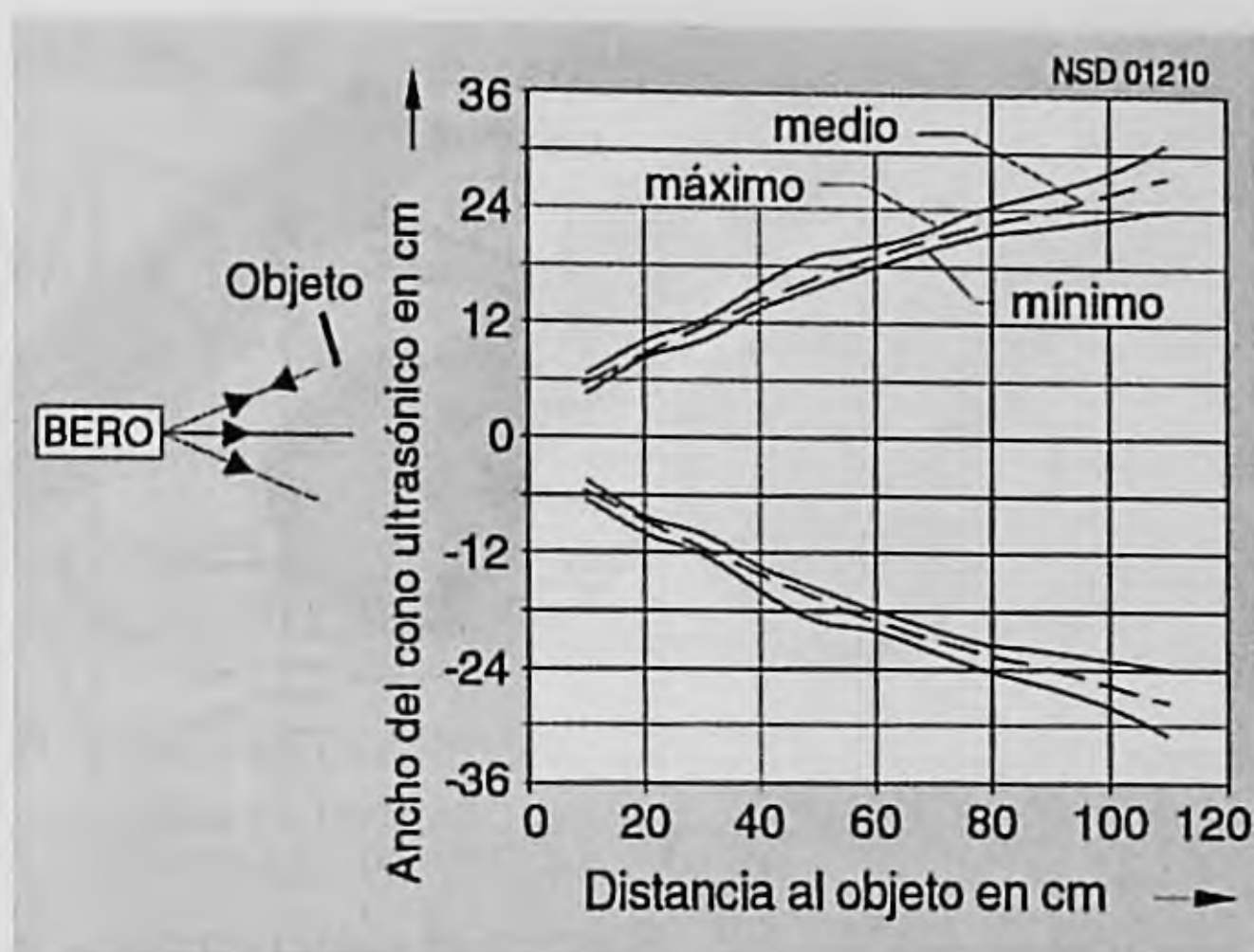


Figura 7.185 Campo de trabajo de un sensor ultrasónico de la familia Bero cuya distancia nominal de detección está comprendida entre 10 y 110 cm, utilizado en una aplicación en la que el objeto es plano y está colocado perpendicularmente a la línea que lo une con el sensor (cortesía de Siemens).

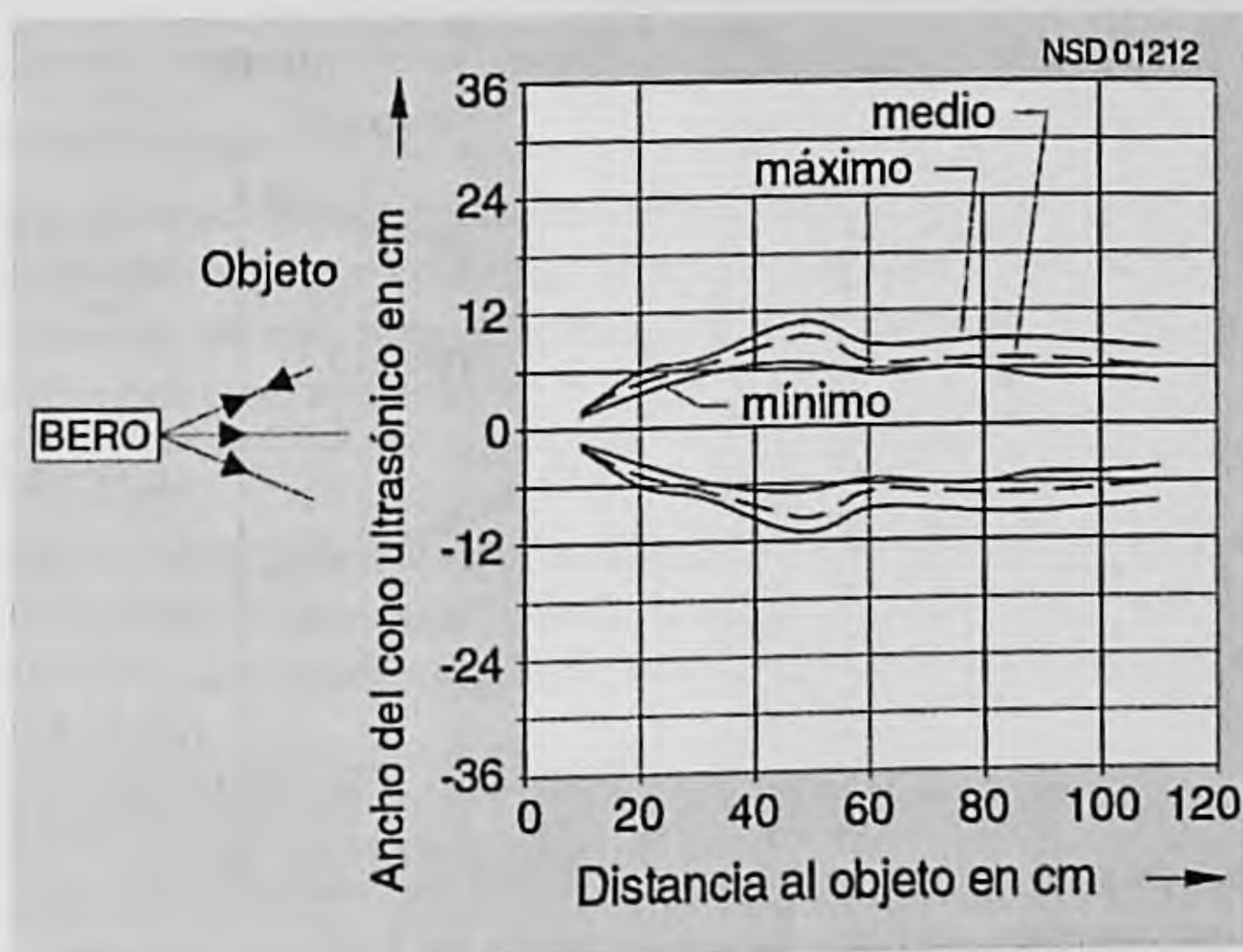


Figura 7.186 Campo de trabajo de un sensor ultrasónico de la familia Bero cuya distancia nominal de detección está comprendida entre 10 y 110 cm, utilizado en una aplicación en la que el objeto es plano y está colocado perpendicularmente al eje del cono de emisión del sensor (cortesía de Siemens).

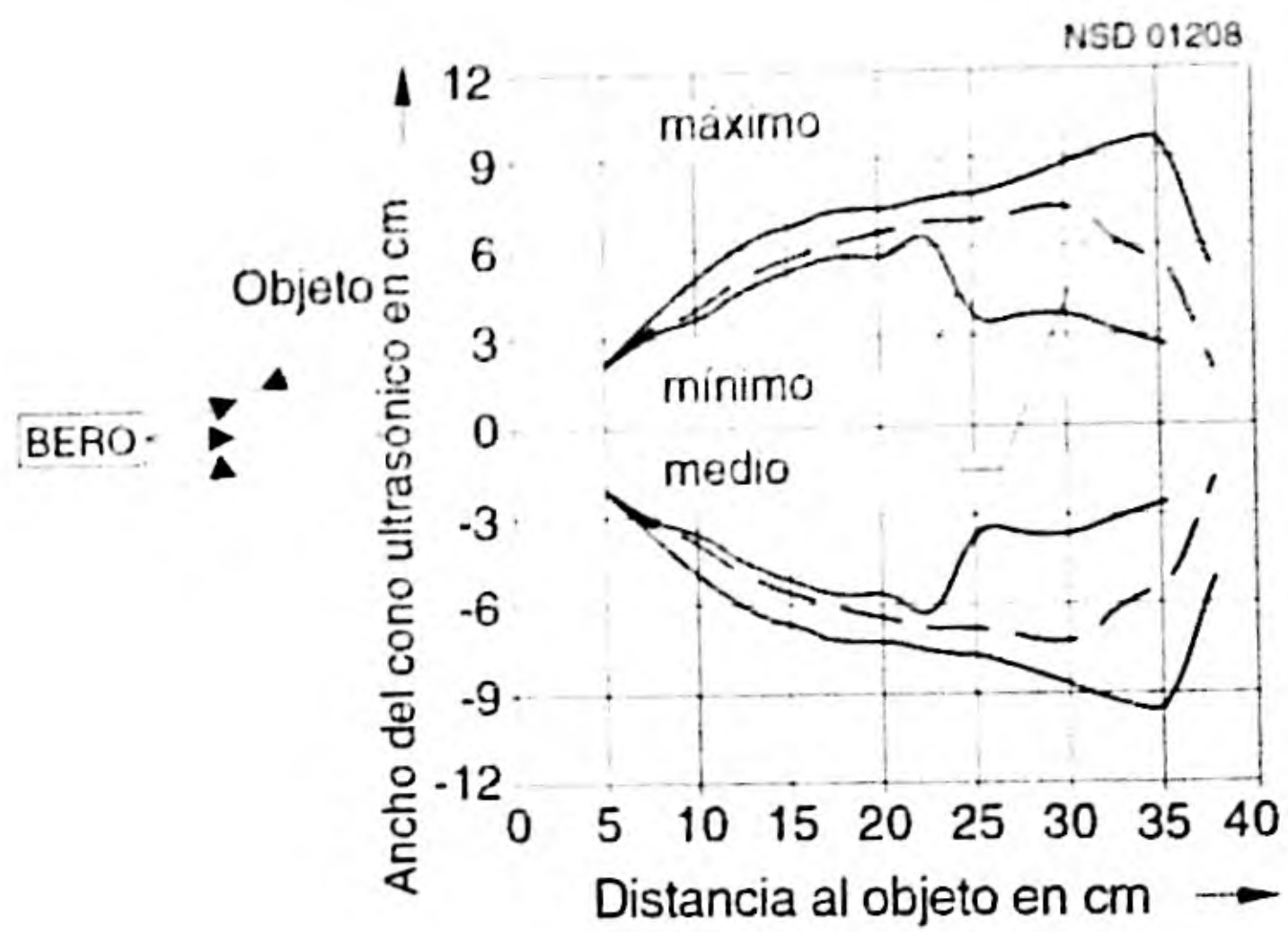


Figura 7.187 Campo de trabajo de un sensor ultrasónico de la familia Bero cuya distancia nominal de detección está comprendida entre 5 y 37 cm, utilizado en una aplicación en la que el objeto es cilíndrico (cortesía de Siemens).

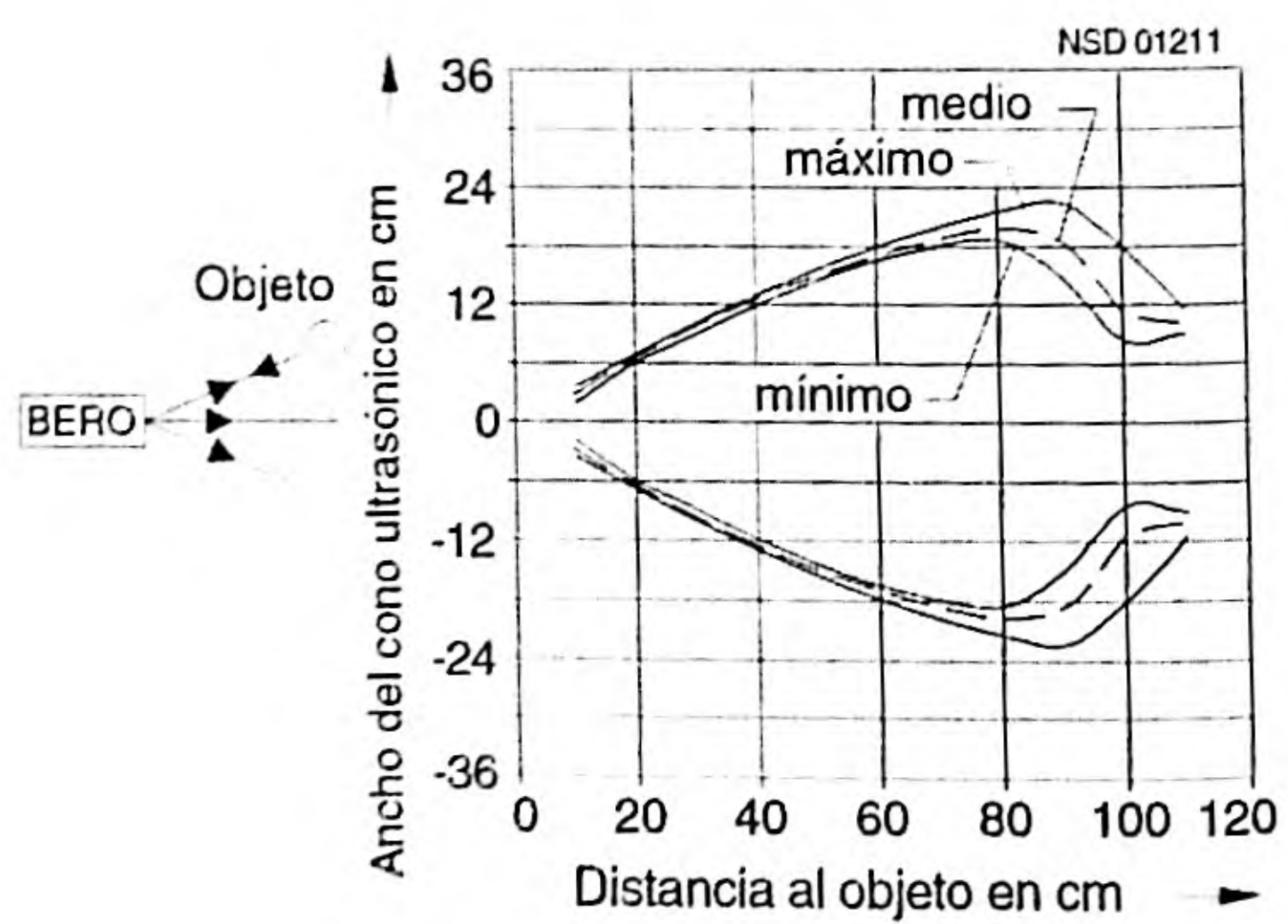


Figura 7.188 Campo de trabajo de un sensor ultrasónico de la familia Bero cuya distancia nominal de detección está comprendida entre 10 y 110 cm, utilizado en una aplicación en la que el objeto es cilíndrico (cortesía de Siemens).

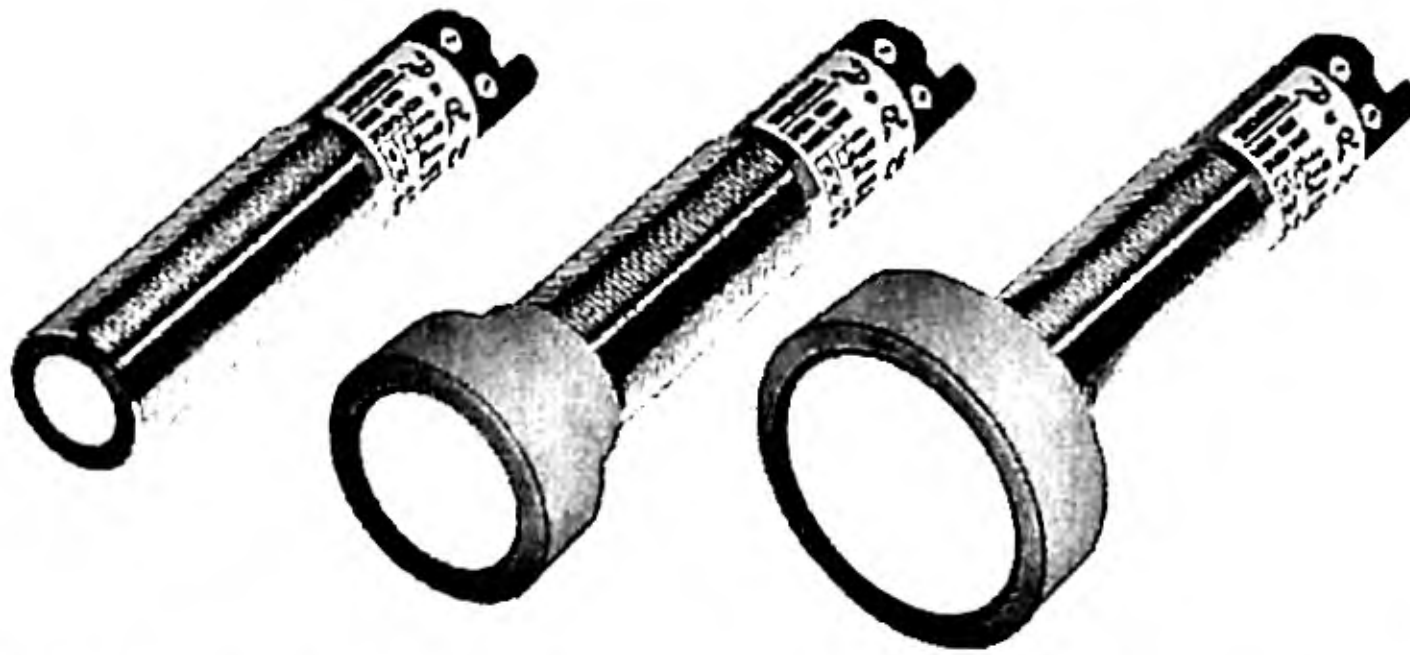


Figura 7.189 Sensores ultrasónicos con circuito de compensación de temperatura
(cortesía de Siemens)

Influencia de las condiciones ambientales

Al utilizar los sensores ultrasónicos de tipo eco se deben tener en cuenta los siguientes efectos:

- **Presión atmosférica**

Si se consideran las fluctuaciones habituales de la presión atmosférica en un punto determinado, se puede comprobar que su efecto sobre la velocidad del sonido es despreciable. Por ejemplo, entre el nivel del mar y 3000m de altitud la velocidad de propagación disminuye en menos del 1 %.

- **Humedad del aire**

A temperatura ambiente y con bajas temperaturas, la humedad ambiente tiene un efecto despreciable sobre la velocidad de propagación del ultrasonido. Cuanto mayor es la temperatura del aire, más aumenta la velocidad de propagación con la humedad, tal como se indica en el apartado 7.7.2.3.7.2.

- **Tipos de gases**

Los sensores ultrasónicos de proximidad funcionan generalmente en presencia de aire. Con otros gases se producen variaciones de la velocidad del sonido y del coeficiente de atenuación, que pueden producir errores importantes e incluso la imposibilidad de funcionamiento.

- **Corrientes de aire**

No es posible formular una ley de variación de validez general para el efecto que tienen sobre la velocidad de propagación los cambios permanentes de la dirección de circulación y la velocidad del aire. Los objetos que se encuentran a temperaturas elevadas, como es el caso de los metales al rojo, ocasionan turbulencias en el aire circundante que dan lugar a dispersiones o desviaciones de los ultrasonidos cuando el sensor se sitúa por encima de ellos y en su misma vertical, tal como se indica en la parte central de la figura 7.190.

- **La precisión del temporizador/contador**

Este temporizador está situado en el circuito de interfaz del microcontrolador y mediante él se mide el tiempo de ida y vuelta de la señal ultrasónica. Por ejemplo si el temporizador/contador es de 12 bits y el rango de operación del sensor está comprendido entre 40 y 300 cm la resolución es:

$$R = \frac{\text{Rango}}{2^{12}} = \frac{[3000 - 400]mm}{4096} = 0,73mm$$

- **Tipo de variables de salida**

Las variables de salida de los sensores ultrasónicos de proximidad de tipo eco, al igual que en los otros tipos de sensores de proximidad analizados anteriormente pueden ser:

- Todo-nada para indicar la presencia de objeto dentro del campo de trabajo establecido.
- Analógica, tanto de tensión como de corriente, que se debe conectar a una entrada analógica del procesador digital que debe tomar decisiones en función del resultado de la medida.
- Temporal, consistente en una secuencia de impulsos cuya frecuencia varía entre un mínimo y un máximo, a medida que el objeto se desplaza dentro del rango de operación del sensor. Este tipo de salida se puede conectar a una entrada de conteo de un procesador y hace innecesaria la utilización de una entrada analógica de corriente o de tensión. Un ejemplo de este tipo de salida es la que posee el sensor ultrasónico PXS200 K21 de Siemens que genera una señal cuadrada de frecuencia variable entre 70 y 500Hz cuando el objeto se desplaza en el rango de 35 a 250mm. Este tipo de salida permite la utilización de sensores ultrasónicos en combinación con autómatas programables compactos de bajo coste, como por ejemplo el 239RC de la familia LOGO (Figura 1.76 del capítulo 1) para controlar sistemas en los que es necesario medir distancias.

Al igual que los demás sensores detectores de objetos, al instalar los sensores ultrasónicos de proximidad de tipo eco es necesario tener en cuenta un conjunto de reglas prácticas relacionadas con el entorno de trabajo, como son las interferencias mutuas, los objetos circundantes, la existencia de obstáculos y la presencia de objetos perturbadores, que se describen seguidamente.

Normas de instalación

- **Interferencias mutuas**

La instalación próxima de dos o más sensores de ultrasonidos puede hacer que las emisiones de uno afecten a otro. Por ello cuando los sensores están situados de tal forma que sus respectivos haces son paralelos se deben separar una distancia mínima X , tal como se muestra en la figura 7.191a, en la que se indica que el valor de X depende de la distancia a la que está situado el objeto con relación al sensor (Figura 7.191b). En dicha figura se supone que el objeto es liso y está situado perpendicularmente al haz ultrasónico. Tal como se indica anteriormente es necesario también tener en cuenta el tipo de objeto y su inclinación (Figura 7.192).

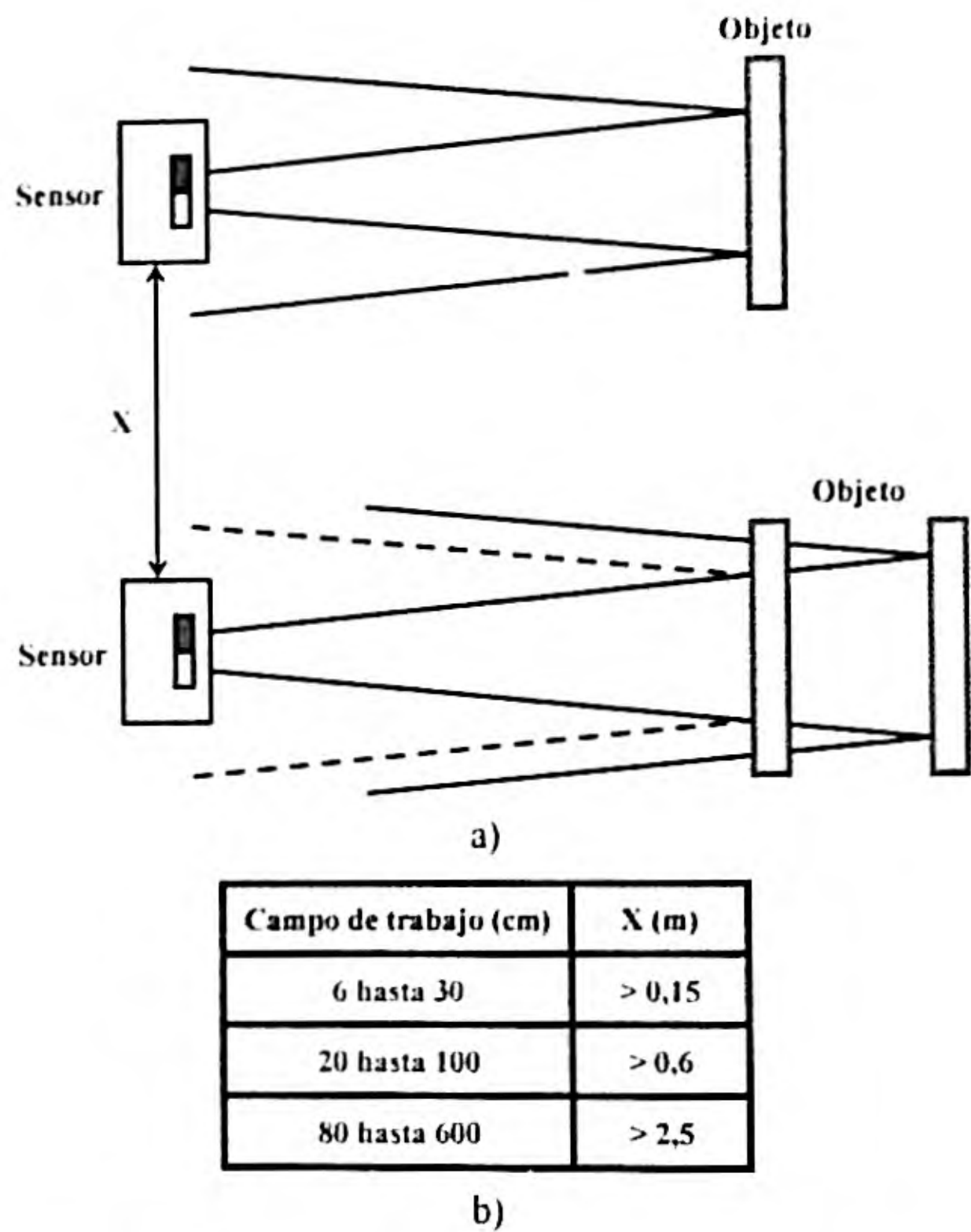


Figura 7.191 Distancia mínima X entre sensores ultrasónicos situados con sus respectivos haces paralelos: a) Representación gráfica; b) Valor de X en función del campo de trabajo.

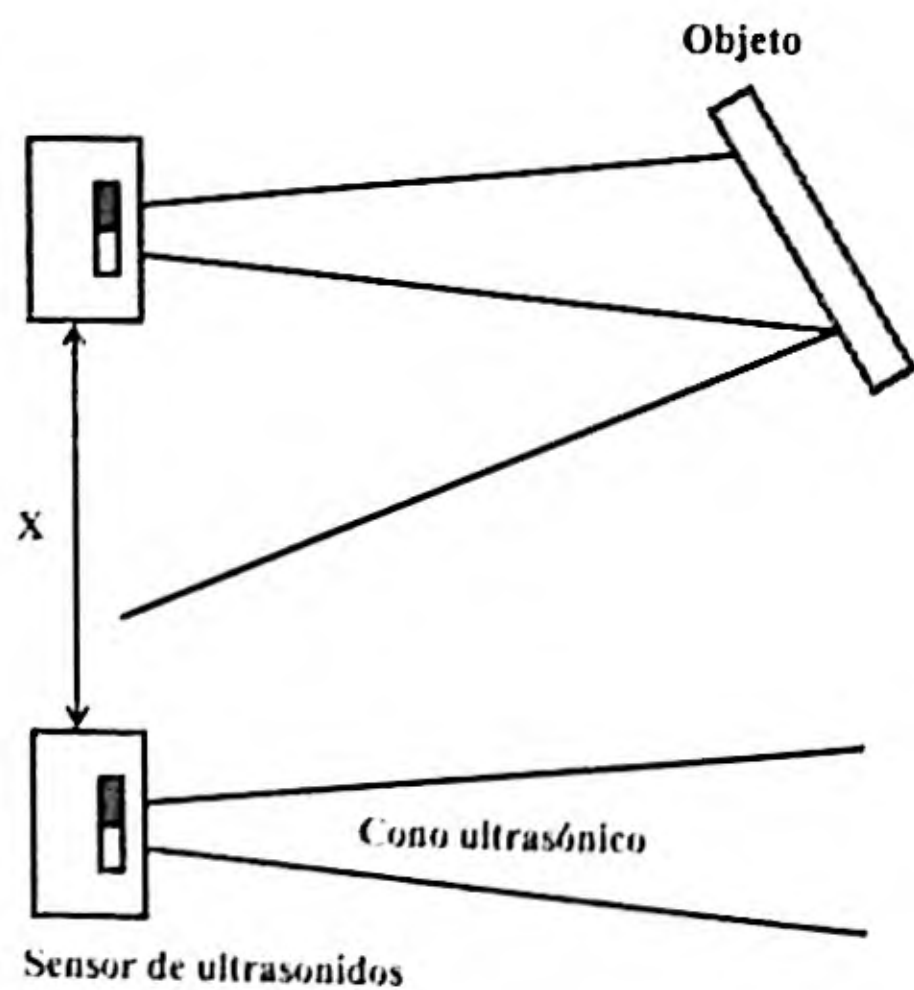


Figura 7.192 Influencia de la inclinación del objeto en la distancia X a la que hay que situar dos sensores ultrasónicos que tienen sus respectivos haces paralelos.

De igual forma, si los sensores están enfrentados, tal como se muestra en la figura 7.193a, puede ocurrir que uno detecte al otro y para evitarlo hay que colocarlos a una distancia mínima que también depende del campo de trabajo (Figura 7.193b).

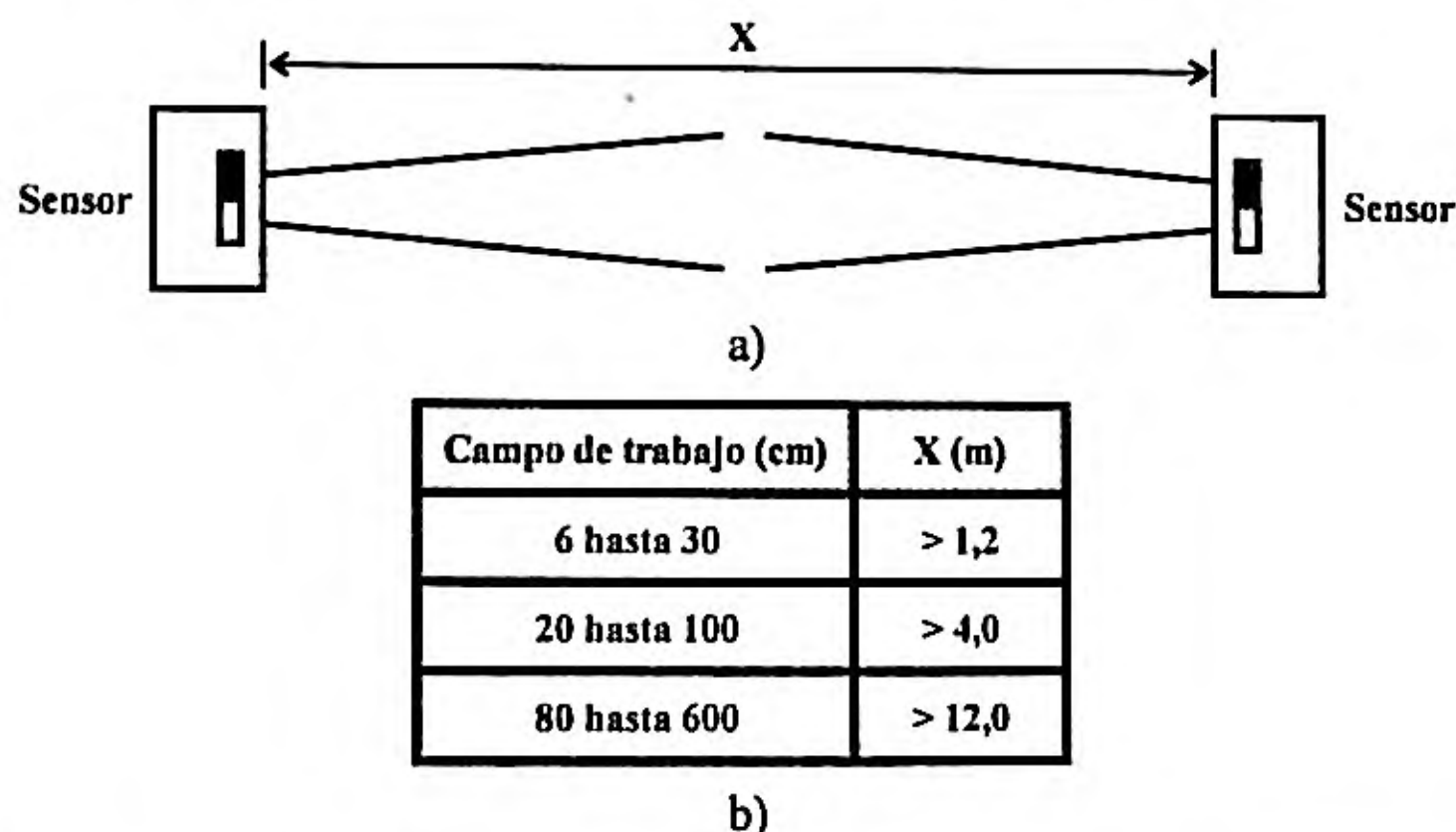


Figura 7.193 Distancia mínima X entre sensores ultrasónicos situados uno frente al otro: a) Representación gráfica; b) Valor de X en función del campo de trabajo.

- Presencia de objetos circundantes**

Los objetos que circundan al sensor, en general, y las superficies lisas próximas en particular, pueden tener una influencia importante en el funcionamiento del sensor. Debido a ello, el valor que debe tener la distancia mínima X (Figura 7.194a) a la que debe encontrarse una superficie lisa próxima al sensor, se indica en la figura 7.194b para diferentes valores del campo de trabajo. En la figura 7.195 se indican los mismos valores en el caso de que el sensor esté colocado entre dos superficies lisas.

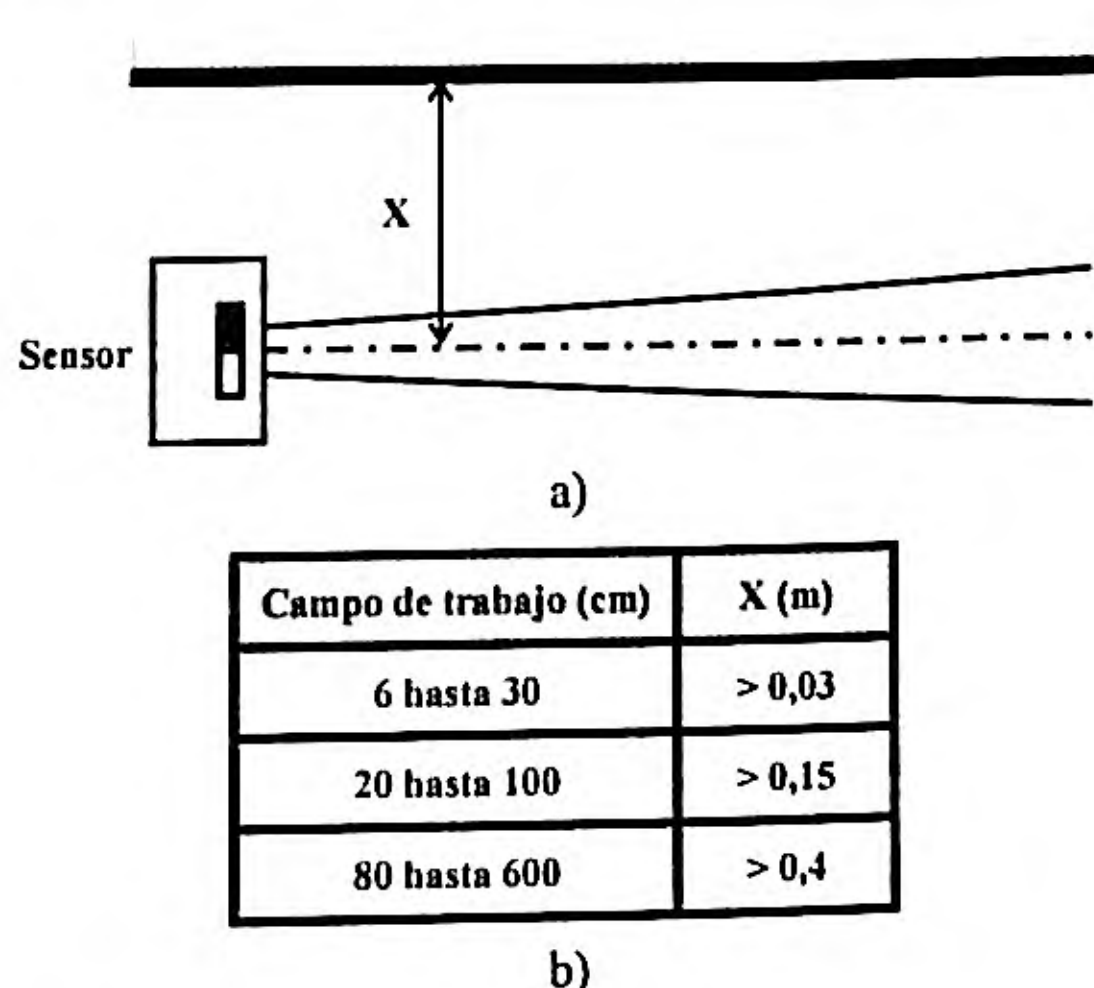
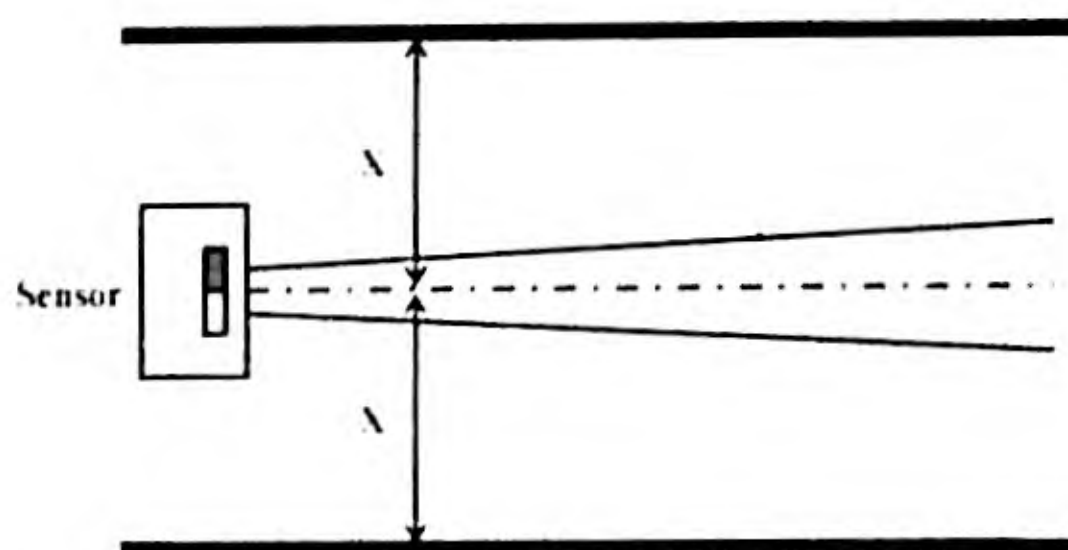


Figura 7.194 Distancia mínima X entre un sensor ultrasónico y una superficie lisa próxima al mismo: a) Representación gráfica; b) Valor de X en función del campo de trabajo.



a)

Campo de trabajo (cm)	X (m)
6 hasta 30	> 0,06
20 hasta 100	> 0,3
80 hasta 600	> 0,8

b)

Figura 7.195 Distancia mínima X entre un sensor ultrasónico y dos superficies lisas próximas al mismo: a) Representación gráfica; b) Valor de X en función del campo de trabajo.

- **Presencia de objetos perturbadores**

A veces es aconsejable reducir el cono ultrasónico para impedir la detección de objetos próximos, distintos del objeto a detectar, que alteran el correcto funcionamiento del sistema. La reducción se puede realizar mediante un diafragma perforado que absorbe la señal ultrasónica tal como se indica en la figura 7.196.

- **Existencia de obstáculos**

El haz de ultrasonidos se puede desviar por medio de reflectores planos o ligeramente cóncavos, para detectar objetos situados a la vuelta de una esquina tal como se indica en la figura 7.197. En general se recomienda no utilizar más de dos reflectores.

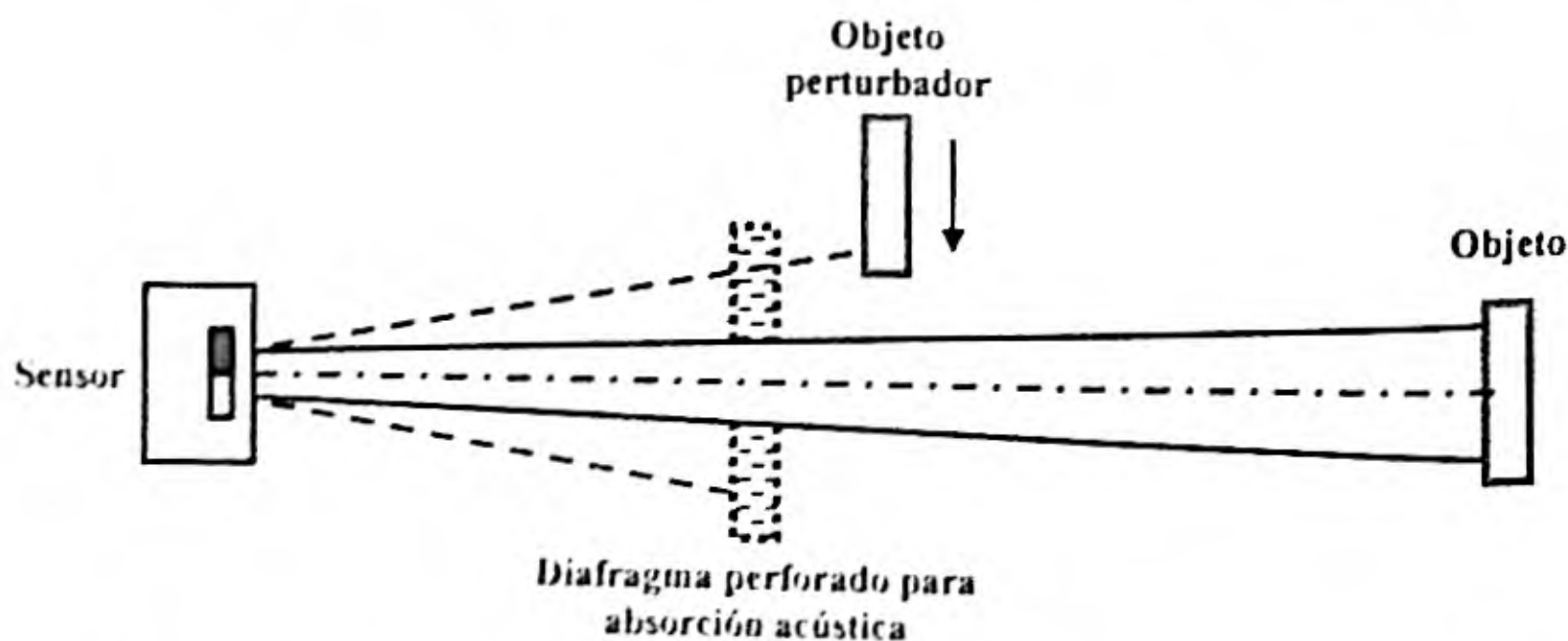


Figura 7.196 Utilización de un diafragma perforado para evitar la influencia de un objeto perturbador.

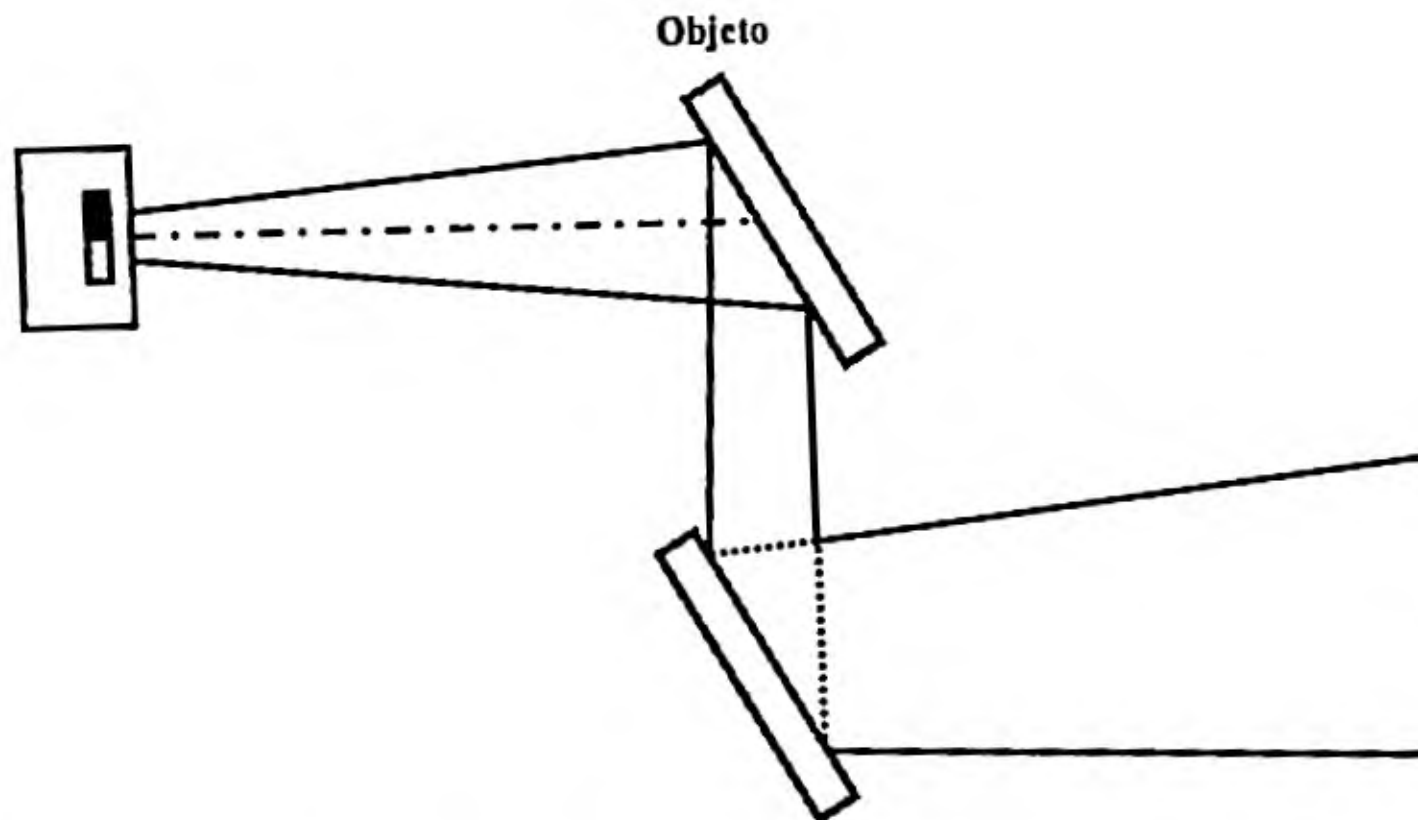


Figura 7.197 Desviación del haz ultrasónico mediante reflectores planos o ligeramente cóncavos, para detectar objetos situados a la vuelta de una esquina.

7.7.2.4 Sensores de medida de distancias

7.7.2.4.1 Introducción

Una de las variables físicas cuya medida es muy importante en los procesos de fabricación es el desplazamiento lineal o angular de un objeto.

En la práctica es necesario medir desplazamiento o conocer la posición de un objeto dentro de amplios márgenes, desde centésimas de milímetros hasta decenas de metros y, debido a ello, no existe ningún principio físico que permita realizar un sensor capaz de medir en un margen tan amplio.

Según el margen de medida se utilizan sensores basados en diferentes principios físicos. En el capítulo anterior se estudian los sensores ultrasónicos utilizados para medir distancias comprendidas, aproximadamente, entre algunos centímetros y algunos metros. A continuación se describen dos tipos de sensores muy utilizados en los procesos de fabricación como son los codificadores optoelectrónicos de posición y los transformadores diferenciales variables lineales.

7.7.2.4.2 Codificadores optoelectrónicos de posición

Los codificadores de posición (*Position encoders*) son sensores que proporcionan una información digital equivalente al desplazamiento lineal o angular de un elemento móvil.

Se utilizan para medir la posición, la velocidad y la aceleración tanto angular como lineal de un objeto y, de acuerdo con la forma en la que proporcionan la información a la salida, pueden ser incrementales o absolutos que se describen a continuación.

7.7.2.4.2.1 Codificadores absolutos de posición

Los codificadores absolutos de posición (*Absolute position encoders*) suelen ser ópticos y estar constituidos por un disco codificado que posee n pistas concéntricas que com-

binan zonas opacas y transparentes asignadas a los números “0” y “1”, tal como se muestra en la figura 7.198, intercaladas de tal manera que a cada posición le corresponde una combinación diferente en un cierto código binario. Cada pista tiene asignado un emisor (un diodo luminiscente) y un receptor sensible a la luz (generalmente un fotodiodo o un fototransistor), enfrentados entre sí y situados a uno y otro lado del disco respectivamente. El receptor de luz está asociado a un circuito cuya salida se encuentra en un nivel de tensión correspondiente a un determinado valor lógico cuando el citado receptor recibe luz y en el nivel opuesto cuando no la recibe. Si el número de pistas es n , el disco se puede dividir en 2^n sectores y, si a cada sector se le asigna una combinación diferente de zonas opacas y transparentes, el disco con sus emisores y receptores constituye un codificador óptico absoluto de posición. El código utilizado debe ser continuo y cíclico para evitar que alguna imperfección en la construcción de las zonas opacas y transparentes, una desalineación de los emisores y receptores o algún desajuste en el sistema mecánico cuyo eje se une al codificador, puedan hacer que aparezca una combinación intermedia errónea. El código mas utilizado es el Gray [MAND 08] y el codificador óptico absoluto de posición de la figura 7.198 posee cuatro pistas que dan lugar a dieciséis sectores codificados en el citado código. En la figura 7.199 se muestra una fotografía del sensor absoluto de posición P NC01 XX 00755P.

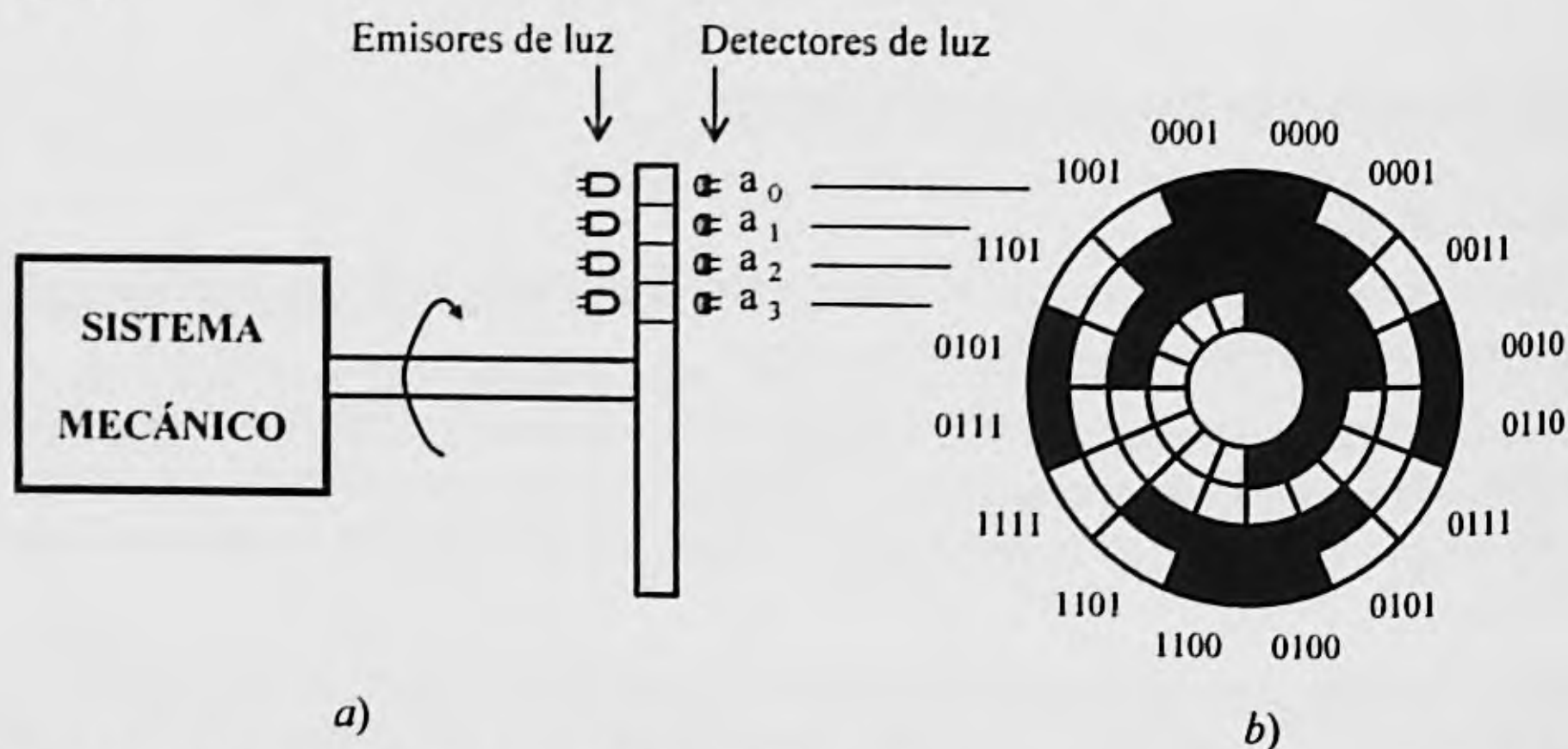


Figura 7.198 Codificador absoluto de posición: a) Acoplamiento a un sistema mecánico; b) Vista frontal del disco codificador.

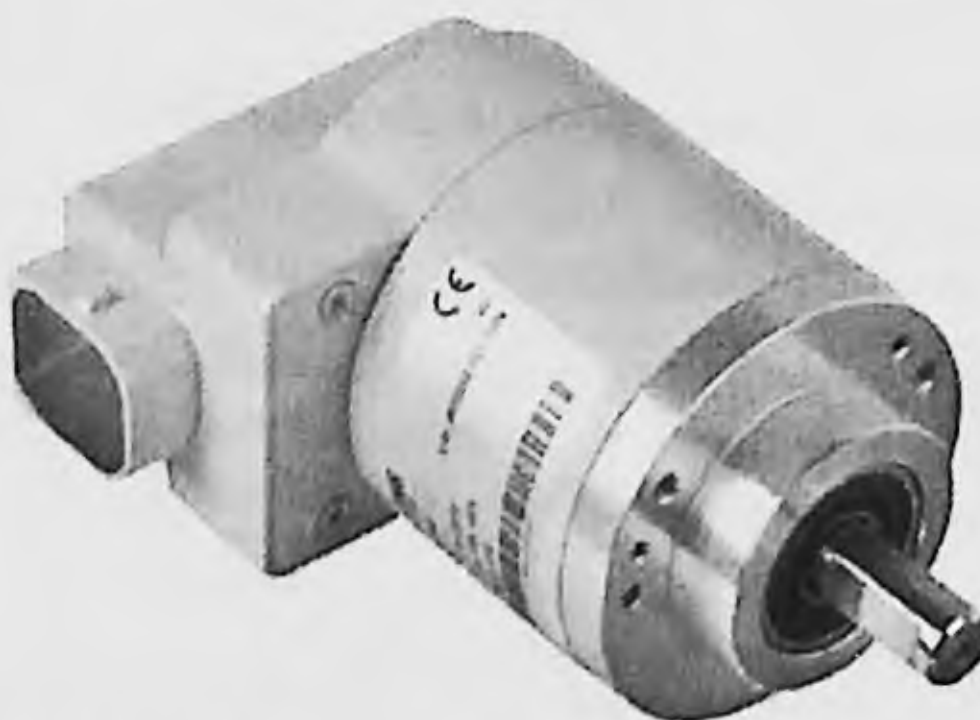


Figura 7.199 Codificador optoelectrónico absoluto de posición P NC01 XX 00755P (Cortesía de Siemens).

7.7.2.4.2 Codificadores incrementales de posición

Los codificadores incrementales de posición (*Incremental position encoders*) generan un número de impulsos proporcional al desplazamiento y pueden ser:

Unidireccionales

Poseen solamente una pista codificada mediante zonas opacas y transparentes iguales, dispuestas alternativamente tal como se indica en la figura 7.200a, y permiten medir desplazamientos o giros en un solo sentido. Poseen una ranura de índice que indica el paso del disco por la posición inicial (Figura 7.200b). Se utilizan en los ratones de ordenador y la frecuencia de los impulsos de salida del receptor es proporcional a la velocidad de giro. El desplazamiento angular es proporcional al número de impulsos que se generan a la salida del receptor a partir de la posición inicial. La resolución es $360^\circ / \text{número } n \text{ de ranuras}$.

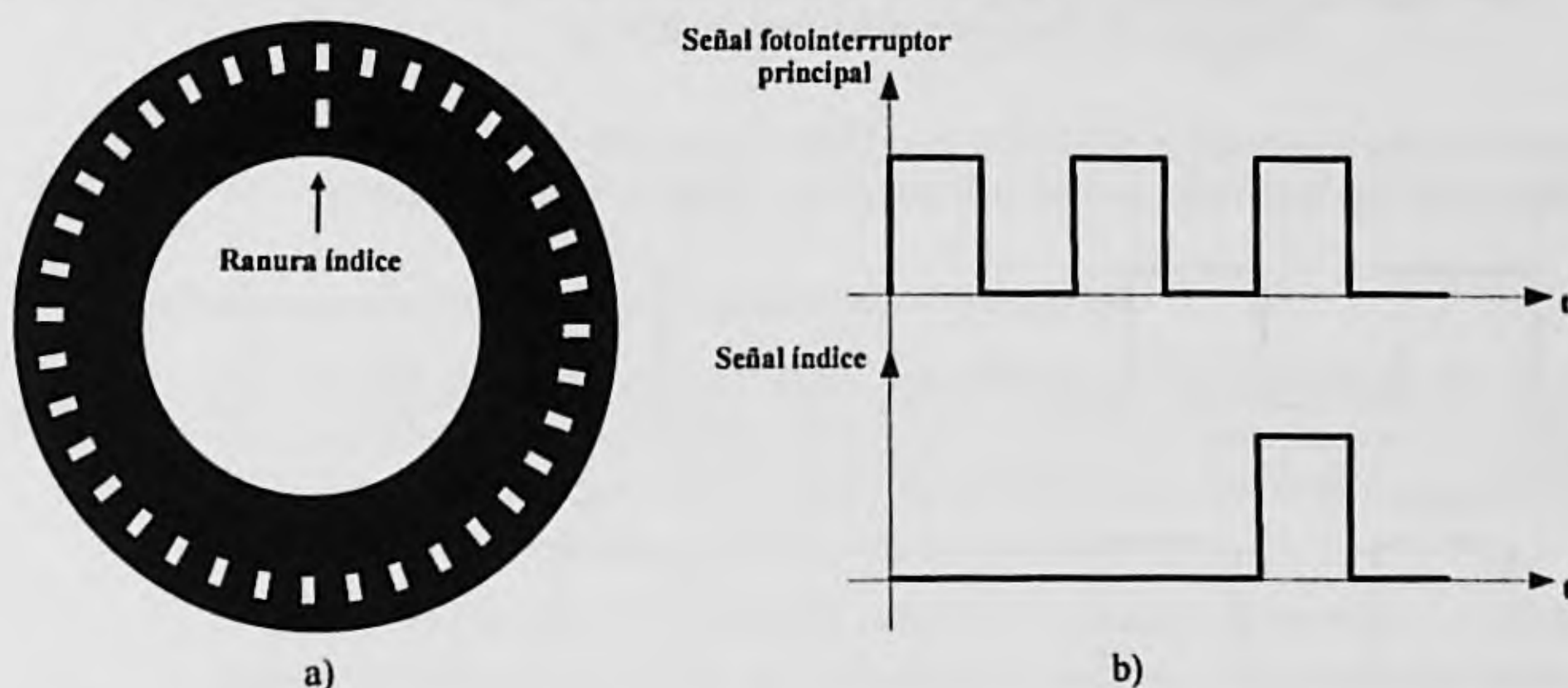


Figura 7.200 Codificador incremental de posición unidireccional: a) Disco codificador; b) Señales generadas por los detectores de luz.

Bidireccionales

Permiten la medida de giros en ambos sentidos y se pueden realizar mediante dos pistas codificadas, con las divisiones de una de las pistas desplazadas un cuarto de periodo (90°) con respecto a las de la otra, tal como se muestra en la figura 7.201, o mediante una sola pista con dos fotosensores desplazados. El sentido de giro se detecta mediante la fase de las señales generadas por ambos fotodetectores tal como se muestra en la figura 7.202. Si el eje gira en sentido horario, la señal B está siempre en nivel cero cuando se produce un flanco de subida de A, y en nivel uno cuando se produce un flanco de bajada de A. Si, por el contrario, el eje gira en sentido antihorario, la señal B está siempre en nivel uno cuando se produce un flanco de subida de A, y en nivel cero cuando se produce un flanco de bajada de A. La medida del desplazamiento se realiza mediante el conteo de los impulsos generados por el codificador a partir de un origen arbitrario. El conteo se puede realizar mediante un circuito digital contador o mediante un procesador digital como por ejemplo un autómatas programable o un microcontrolador [MAND 08], en una parte de cuya memoria se coloca un programa de conteo. En la medida de velocidad se puede duplicar la resolución contando los flancos de la señal obtenida mediante una puerta lógica O-exclusiva a la que se apliquen las señales A y B.

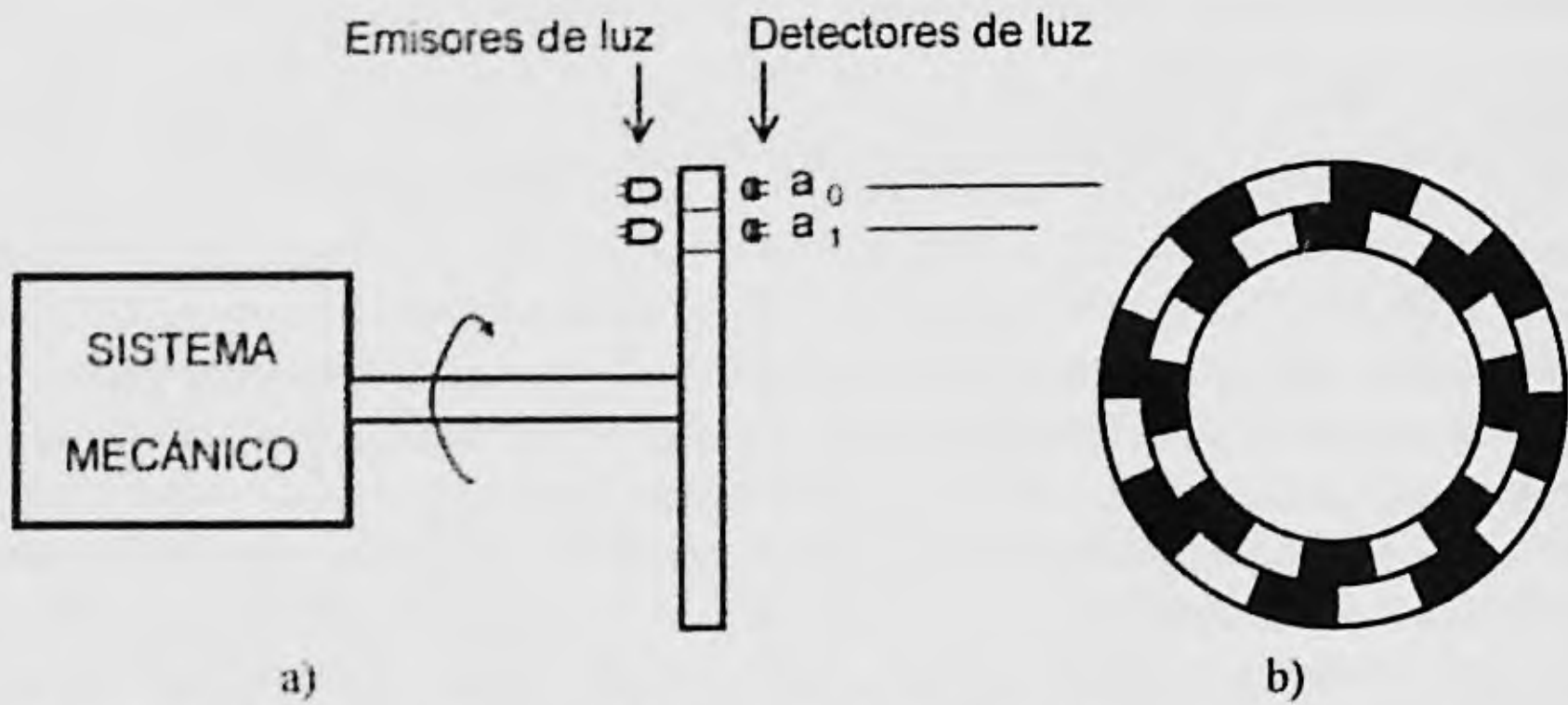


Figura 7.201 Codificador incremental de posición bidireccional: a) Acoplamiento a un sistema mecánico; b) Vista frontal del disco codificador.

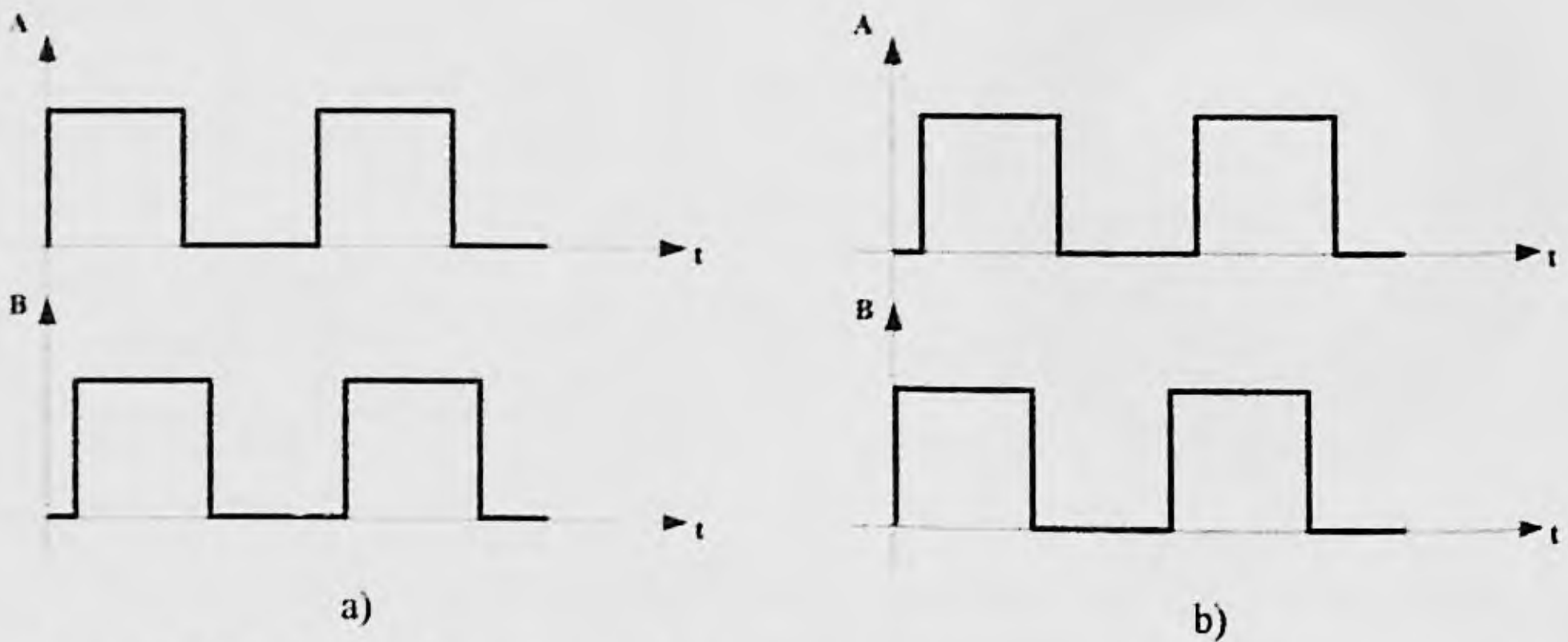


Figura 7.202 Codificador incremental de posición bidireccional: a) Señales generadas por los detectores de luz cuando gira hacia la izquierda; b) Señales generadas por los detectores de luz cuando gira hacia la derecha.

7.7.2.4.3 Transformador diferencial variable lineal

El transformador diferencial variable lineal conocido como LVDT (*Linear Voltage Differential Transformer*) es un elemento sensor o transductor basado en la variación de la inductancia mutua entre un primario y dos secundarios conectados en oposición ($M_2 - M_2'$). Dicha variación es debida al desplazamiento de un núcleo de material ferromagnético que está unido a la pieza cuya posición se quiere medir tal como se muestra en la figura 7.203a.

El primario se alimenta con una tensión alterna y los secundarios están situados simétricamente con respecto a él. En cada secundario se inducen tensiones iguales cuando el núcleo está situado en su posición central u origen con respecto a cada uno de los secundarios (posición de origen, $x = 0$) y, por estar conectados en oposición, la tensión de salida, que resulta de la diferencia entre las dos tensiones, es teóricamente nula. ($V_{AB} = 0$).

Al desplazarse el núcleo una de las tensiones crece y la otra se reduce en la misma magnitud. Debido a ello la curva característica de la tensión V_{AB} con respecto al desplazamiento x es la representada en la figura 7.203b.

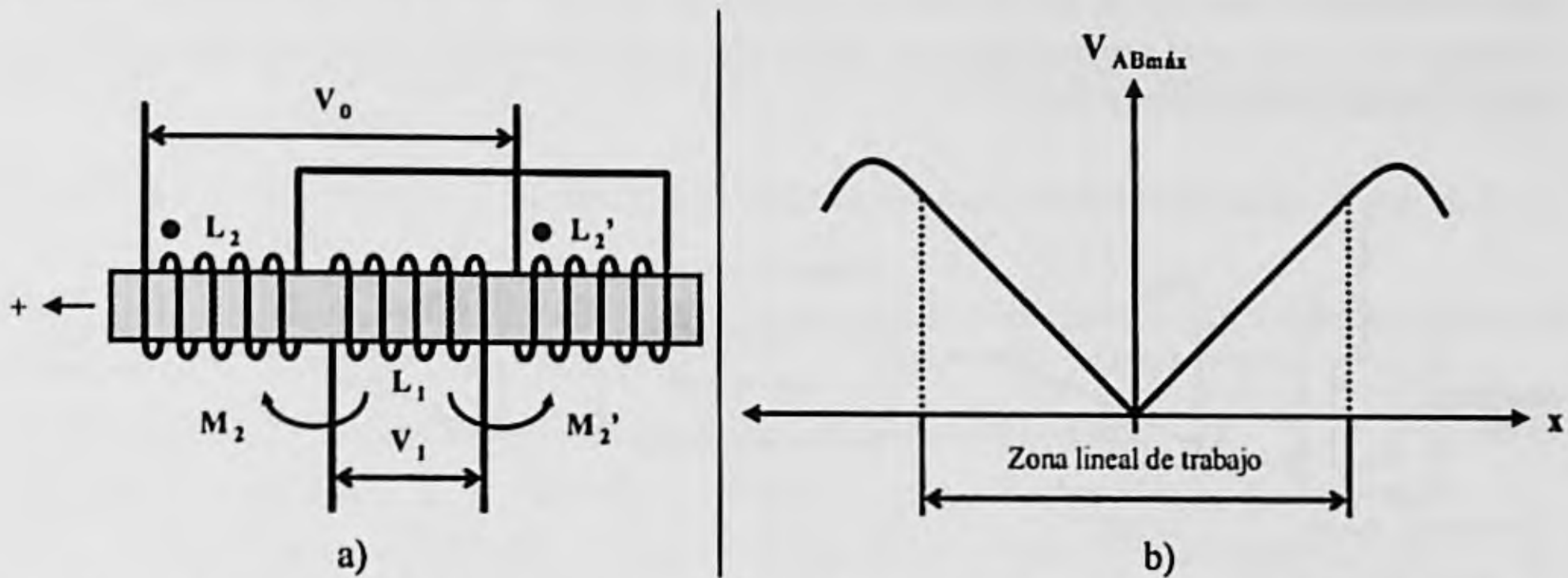


Figura 7.203 Transformador diferencial variable lineal (LVDT): a) Elementos que lo constituyen; b) Curva característica de la tensión V_{AB} con respecto al desplazamiento.

Las principales características del elemento sensor LVDT son:

- La ausencia de contacto físico entre el núcleo móvil y la estructura que aloja los devanados hace que sea un dispositivo sin rozamiento, lo que permite su utilización en medidas críticas que pueden tolerar la adición de una pequeña masa (el núcleo) pero que no admiten cargas que provoquen rozamiento
- Por no haber rozamiento ni contacto entre los devanados y el núcleo, no hay partes sometidas a desgaste. Esto le confiere un vida mecánica prácticamente ilimitada y lo hace apto para ser utilizado en mecanismos o sistemas en los que se precisa una gran fiabilidad
- El aislamiento entre el núcleo y los devanados permite fabricar modelos de elementos sensores LVDT capaces de ser utilizados en condiciones ambientales duras, tales como temperaturas extremas, alta presión, ambientes corrosivos, etc.
- Su principio de funcionamiento inductivo, junto con la ausencia de rozamiento, hacen que el sensor LVDT posea una resolución muy elevada y que sea sensible a desplazamientos muy pequeños del núcleo. El límite de la resolución lo establece el circuito electrónico conectado a él.

Como consecuencia de todo ello:

- Los sensores LVDT alcanzan valores de sensibilidad elevados que varían linealmente, tanto con la amplitud como con la frecuencia de la tensión aplicada al primario. Los valores de frecuencia que se utilizan están comprendidos generalmente entre 50 Hz y algunas decenas de KHz, mientras que los valores de tensión no superan los 50 V eficaces.
- Es frecuente encontrar en el mercado sensores LVDT con rangos de medida que van desde $\pm 0,1$ mm hasta ± 150 mm o superiores, con precisiones superiores al 0,1%.

La señal de salida de un elemento sensor LVDT contiene la información en la amplitud de una onda senoidal. Es por tanto una señal modulada en amplitud que es preciso demodular, tal como se indica en el diagrama de bloques de la figura 7.204 que representa el esquema de bloques completo de un sensor LVDT con el circuito electrónico de acondicionamiento asociado al elemento sensor, que proporciona a la salida una señal analógica continua cuya amplitud es proporcional al desplazamiento.

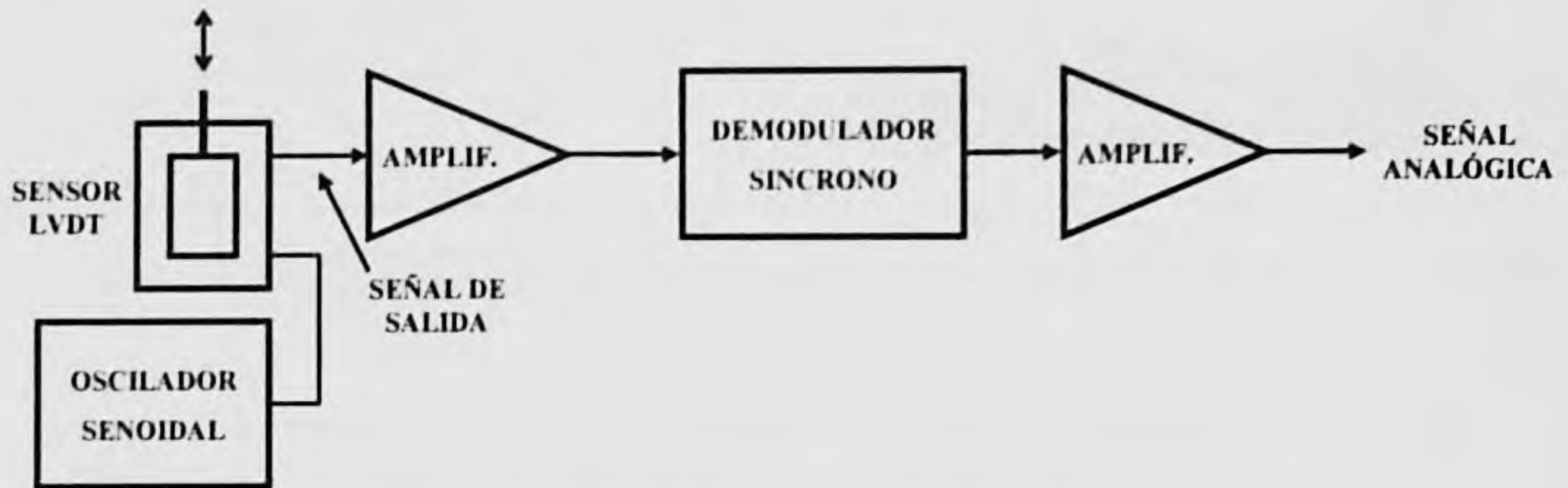


Figura 7.204 Diagrama de bloques de un sensor LVDT.

Bibliografía

- [AENO 05] AENOR. *UNE EN 60947-5-2: Aparata de baja tensión. Parte 5: Aparatos y elementos de conmutación para circuitos de mando. Sección 2: Detectores de proximidad*. AENOR. 2005.
- [AFNO 90] AFNOR. *Bus FIP pour échange d'information entre transmetteur, actionneur et automate. Architecture generale du système*. Norma NFC 46-601. 1990.
- [BART 03] T. Bartelt. *Industrial Control Electronics. Devices, Systems & Applications*. Thomson Learning, 2003.
- [BAXT 97] Larry K. Baxter. *Capacitive Sensors. Design and Applications*. IEEE Press Series on Electronics Technology. IEEE PRESS. 1997.
- [CRC 08] *Handbook of Chemistry & Physics*. 88th Edition, 2007-2008. CRC Press.
- [FRAD 96] J. Fraden. *Handbook of Modern Sensors*. Spriger-Verlag New York Inc. 1996.
- [GARC 97] G. García. *Interbus-S: Una herramienta de comunicación muy flexible*. Automática e Instrumentación. nº 234, pp56-59, mayo 1993.
- [MAND 95] E. Mandado, J. J. Rodríguez y L. J. Álvarez. *Manual de prácticas de Electrónica Digital. Cap. 6. Tecnologías electrónicas digitales*. 1995.
- [MAND 07] E. Mandado, L. M. Menéndez, E. L. Matos y L. F. Ferreira. *Microcontroladores PIC. Sistema integrado para el autoaprendizaje*. Editorial Marcombo. 2007.
- [MAND 08] E. Mandado & Y. Mandado. *Sistemas Electrónicos Digitales*. 9ª Edición. Capítulo 8. Editorial Marcombo. 2008.

- [MARS 97] D. Marsh. *CAN bus Ics marry mechanics with electronic supervision and control*. EDN, pp.122-125. March 3, 1997.
- [MORR 93] A. S. Morris. *Principles of Measurement and Instrumentation*. Prentice Hall. 1993.
- [PALL 98] R. Pallás. *Sensores y acondicionadores de señal*. Marcombo Boixareu Editores. 3ª edición. 1998.
- [PALL 01] R. Pallás & J. G. Webster. *Sensor and signal conditioning*. John Wiley & Sons. 2001.
- [PERE 03] M. A. Pérez García y otros. *Instrumentación electrónica*. Editorial Thomson Learning Paraninfo. 2003.
- [SALO 98] S. Salomon. *Sensors handbook*. Mc Graw Hill. 1998.
- [SIEM 96] Siemens AG. *PROFIBUS & AS-Interface. Fieldbus components*. Siemens Automation Group, Catalog ST PI. Order nº : E86060-K4660-A101-A1-7600. 1996.
- [SIEM 00] Siemens AG. *Basic of sensors*. Siemens. 2000.
- [SIEM 08] Siemens AG. Direccion de internet: <https://mall.automation.siemens.com>.
- [SZE 94] S. M. Sze. *Semiconductor sensors*. John Wiley & Sons, Inc. 1994.
- [WILS 05] J. Wilson, editor. *Sensor Technology handbook*. Newnes. 2005.
- [WOOD 95] G. Wood. *Field bus status 1995*. Computing & Control Engineering Journal, pp. 251-253. Diciembre 1995.

CAPÍTULO 8

Interfaces de entrada y salida

8.1 Introducción y clasificación

Como se indica en el capítulo 1, un autómata programable es en la actualidad un equipo electrónico especialmente diseñado y adaptado para realizar tareas de control de las máquinas que forman parte de un proceso industrial. Para ello debe relacionarse con este último a fin de:

- Conocer la situación de cada una de las máquinas que tiene que controlar, para lo cual debe recibir información del valor de las variables físicas (temperatura, presión, posición de objetos, etc.) adecuadas mediante un conjunto de elementos que las convierten en señales eléctricas. Dichos elementos, que reciben el nombre de sensores y se describen en el capítulo 7, se conectan al autómata programable a través de los adecuados circuitos de acoplamiento o interfaces de entrada.
- Actuar sobre cada una de las máquinas del proceso industrial como resultado de la ejecución del programa de control, para lo cual debe generar un conjunto de variables de salida conectadas al proceso a través de elementos que convierten las señales eléctricas en otras variables físicas no eléctricas o en otras eléctricas de mayor nivel de tensión o de corriente. Dichos elementos, que reciben el nombre de actuadores (*Actuators*), se deben conectar al autómata programable a través de los adecuados circuitos de acoplamiento o interfaces de salida.

El conjunto de circuitos electrónicos de acoplamiento o interfaces de entrada y de salida a través de los cuales el autómata programable se relaciona con el proceso controlado por él se denomina interfaz de conexión con el proceso.

Pero, además de relacionarse con el proceso controlado por él, el autómata programable debe comunicarse con los usuarios para que estos puedan:

- Desarrollar el programa de control en un sistema externo e introducirlo en la memoria del autómata programable.
- Modificar parámetros del programa de control.
- Recibir información del estado del proceso controlado por el autómata programable.
- Detectar fallos en el conjunto formado por el autómata programable y el proceso controlado por él.

Para desarrollar el programa de control se debe utilizar un computador en el que se ejecutan los programas que constituyen las herramientas de diseño asistido por computador (*CAD tools*) descritas en el apartado 5.2 del capítulo 5. El conjunto de elementos que llevan a cabo las demás tareas constituye el interfaz de conexión automática programable-usuario, conocido como HMI (acrónimo de *Human Machine Interface*) que se describe en el apartado 8.3.2.

En los apartados siguientes se analizan ambos tipos de interfaces.

8.2 Interfaz de conexión con el proceso

8.2.1 Introducción

La interfaz de conexión con el proceso está asociada a los dos conceptos principales siguientes (Tabla 8):

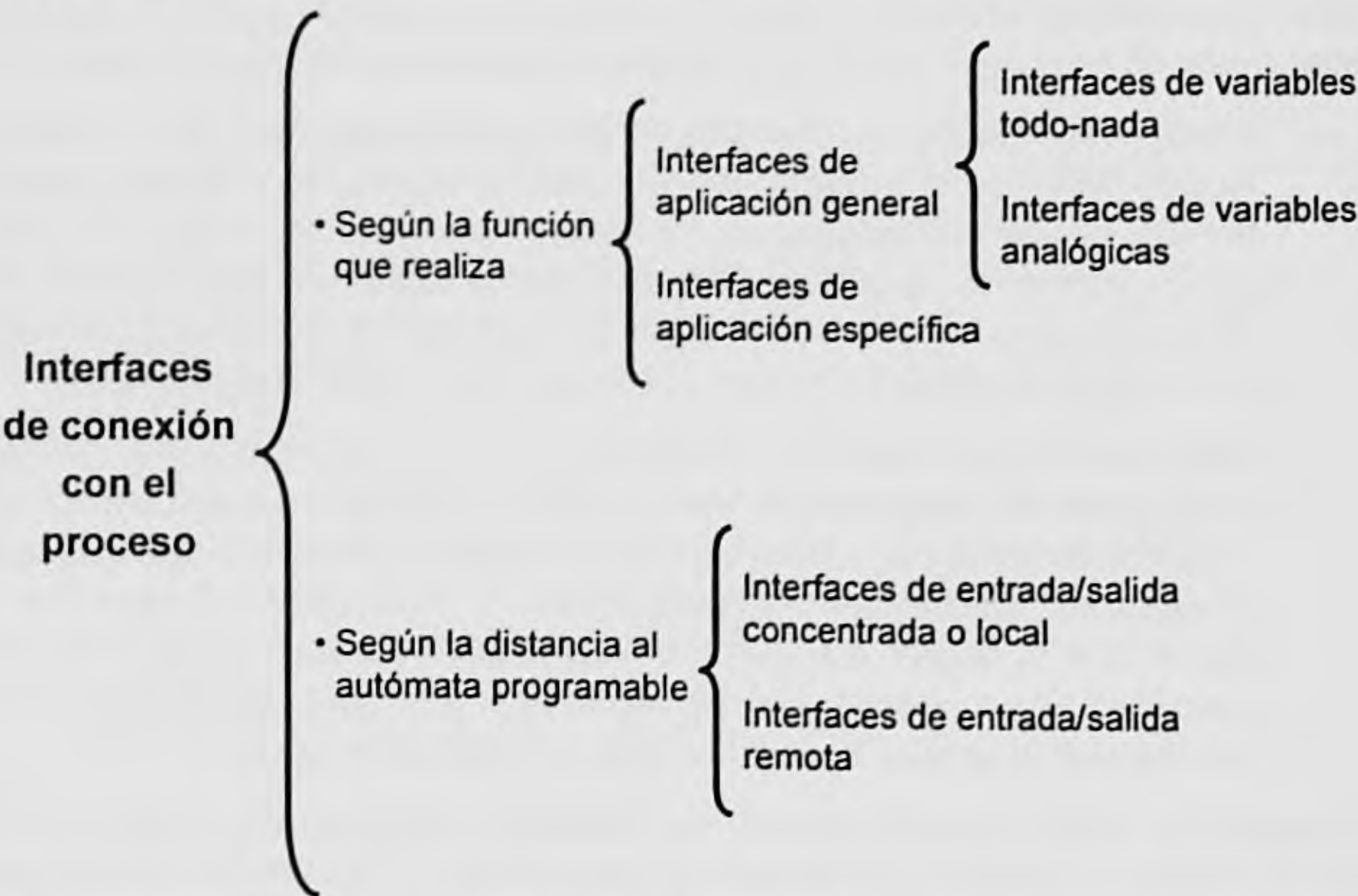


Tabla 8.1. Clasificación de las interfaces de conexión de un autómata programable con un proceso.

- La función que realiza la interfaz.
- La distancia entre el autómata programable y la interfaz a la que se conectan los sensores y actuadores del proceso.

Según la función que realizan las interfaces pueden ser de dos tipos:

- **Interfaces de aplicación general**

Las interfaces de aplicación general (*General purpose interfaces*) tienen como misión acoplar al autómata programable las variables de entrada o de salida, tanto digitales como analógicas, y a su estudio se dedica el apartado 8.2.2.

- **Interfaces de aplicación específica**

Las interfaces de aplicación específica (*Application specific interfaces*) realizan tareas concretas para el acoplamiento de determinadas variables o para llevar a cabo el control de ciertos sistemas. Su fabricación es el resultado del avance de la Microelectrónica que ha permitido el desarrollo de circuitos de entrada y salida que realizan el control de dispositivos específicos de uso muy extendido. Este es el caso de los circuitos de salida que realizan el control directo de motores paso a paso, los circuitos de entrada que realizan el conteo de impulsos de frecuencia elevada, la lectura de codificadores (*Encoders*) angulares, la conexión directa de sensores como por ejemplo termopares, termorresistencias, etc. A su estudio se dedica el apartado 8.2.3.

Según la distancia a la que se encuentran del autómatas programable, las interfaces pueden ser de dos tipos:

- **Interfaces de entrada/salida concentrada o local**

Se dice que la interfaz de entrada/salida de un autómatas programable es local cuando está situada en el interior de la unidad central o en módulos situados al lado de la misma (Figura 8.1). Este tipo de interfaz es adecuado cuando los sensores y actuadores están situados muy próximos al autómatas programable.



Figura 8.1. *Autómatas programable S7 314 de Siemens con módulos de entrada/salida concentrada o local.*

- **Interfaces de entrada/salida remota**

Se dice que la interfaz de entrada/salida de un autómatas programable es remota cuando está situada a una distancia elevada (superior a algunos metros) del mismo (Figura 8.2). En este caso, tanto el autómatas programable, como el módulo de interfaz de en-

trada/salida remoto, poseen un procesador de comunicaciones y se enlazan a través de una red de control que recibe el nombre de bus de campo (*Field bus*) y se estudia en apartado 9.3.2.3 del capítulo 9. Este tipo de interfaz es adecuado cuando los sensores y actuadores están situados a una distancia elevada del autómata programable porque reduce sensiblemente el coste del cableado, así como la influencia de las interferencias electromagnéticas.

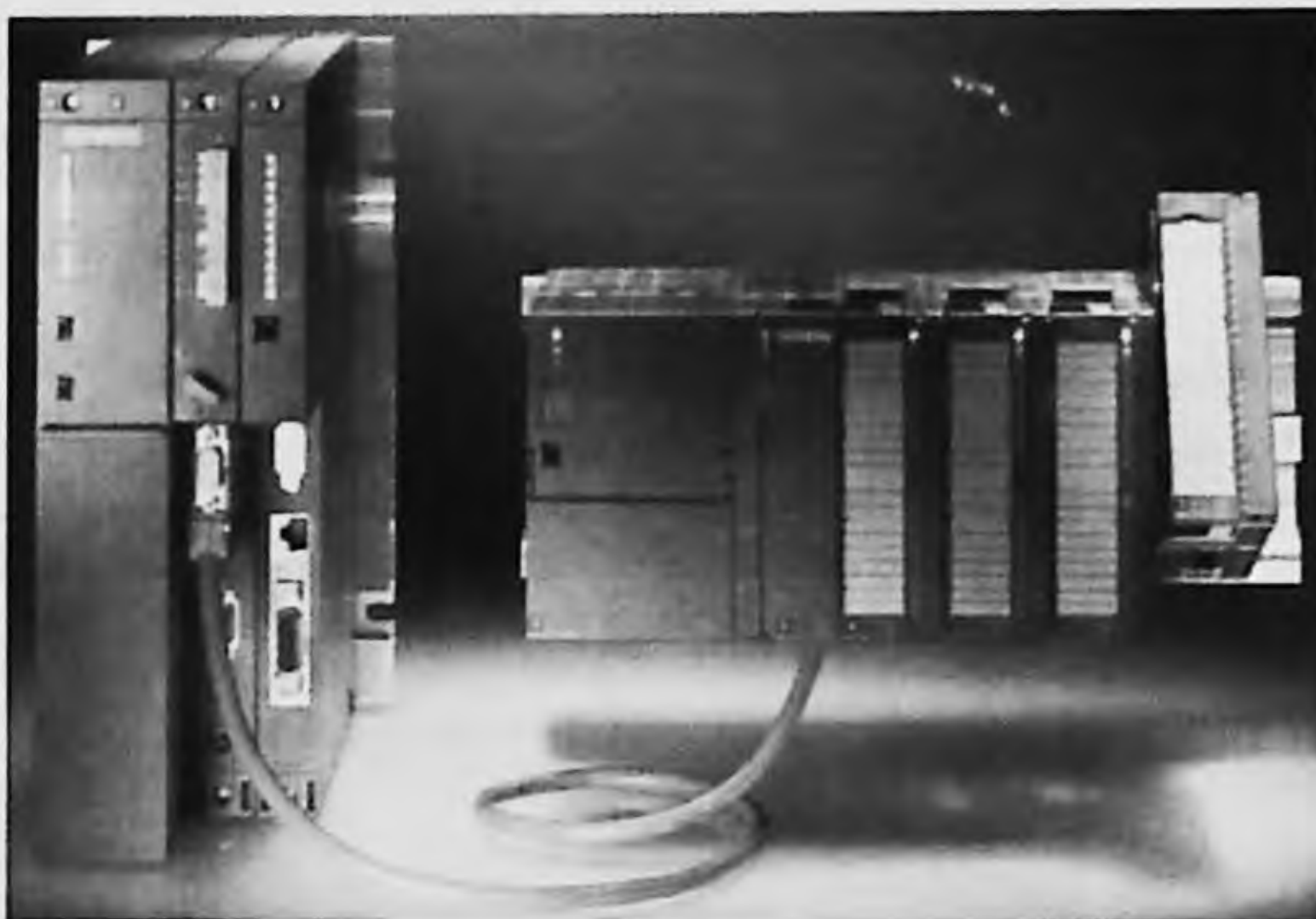


Figura 8.2. *Autómata programable de la familia S7-400 conectado a una unidad de entrada/salida remota.*

8.2.2 Interfaces de conexión con el proceso de aplicación general

8.2.2.1 Introducción

Tal como se indica en el apartado anterior, bajo la denominación de interfaces de aplicación general se conocen las interfaces cuya misión es simplemente transferir al autómata programable la información del proceso a través de sensores o aplicarle a éste las variables de salida mediante los correspondientes actuadores. Tanto las señales eléctricas de entrada como las de salida pueden ser todo-nada (*On-Off*) o analógicas (*Analog*), que se describen en el apartado 7.2.3 del capítulo 7, y por ello según el tipo de señales eléctricas que acoplan el autómata programable con el proceso, las interfaces de aplicación general pueden ser todo-nada o analógicas, cuyas características generales se describen a continuación.

8.2.2.1.1 Interfaces de variables todo-nada

Este tipo de interfaz conecta al autómata programable las variables todo-nada (descritas en el apartado 7.2.4 del capítulo 7), que son aquellas que sólo pueden tener dos valores y por ello algunos autores las denominan digitales.

Las interfaces de las variables todo-nada están ligadas a tres conceptos principales:

- El tipo de variable, que puede ser de entrada o de salida.
- El tipo de tensión de alimentación utilizado, que puede ser continua (CC) [*Direct current (DC)*] o alterna (CA) [*Alternate Current (AC)*].
- La forma de realizar el acoplamiento, que puede ser directo (sin aislamiento galvánico) o con aislamiento o separación galvánica.

En la tabla 8.2 se presenta la clasificación de las interfaces todo-nada de acuerdo con los tres conceptos citados. Se tiene así una visión panorámica de las diferentes interfaces de variables todo-nada, cuyas principales características se estudian en sucesivos apartados.

CLASIFICACIÓN DE LOS CIRCUITOS DE INTERFAZ TODO-NADA	
Variable	Entrada
	Salida
Alimentación	Continua
	Alterna
Acoplamiento	Con aislamiento galvánico
	Sin aislamiento galvánico

Tabla 8.2. Clasificación de los circuitos de interfaz todo-nada.

Las variables de entrada todo-nada (*On-Off*) las generan todos los sensores todo-nada (como por ejemplo termostatos, presostatos, interruptores de caudal, finales de carrera, pulsadores, interruptores, etc.). Dichos sensores cierran o abren un contacto libre de potencial, o hacen que un transistor PNP o NPN, un tiristor o un triac se encuentre en estado de saturación o de corte, tal como se indica en el apartado 7.4.1.3 del capítulo 7. Por ello, según el tipo de sensor y el tipo alimentación que se utilice, el circuito de interfaz puede estar alimentado en corriente continua o en corriente alterna.

Además, la conexión de un sensor a un autómata programable a través de su circuito de entrada se puede realizar de dos formas diferentes:

- **Sin aislamiento galvánico**
Se dice que la conexión entre un sensor que proporciona una variable todo-nada y un autómata programable se realiza sin aislamiento o separación galvánica cuando ambos tienen al menos dos puntos unidos eléctricamente, es decir, están al mismo potencial eléctrico.
- **Con aislamiento galvánico**
Se dice que el acoplamiento entre un sensor que proporciona una variable todo-nada y un autómata programable se realiza con aislamiento galvánico cuando el valor de la tensión a la salida influye en el de la tensión que se presenta en un terminal de entrada del autómata, sin que exista una conexión eléctrica entre ambos. Se logra así que, una sobretensión o una sobrecorriente en el circuito de entrada no afecte al autómata programable.

La tabla 8.3 muestra la clasificación de las interfaces de las variables de entrada todo-nada de acuerdo con el tipo de alimentación y el tipo de salida del sensor que se les puede conectar. En sucesivos apartados se analizan todos los tipos de interfaces de variables de entrada indicados en ella.

INTERFACES DE ENTRADA TODO-NADA		
Alimentación continua	Sensores con salida tipo relé	
	Sensores de dos hilos	
	Sensores de tres hilos	Salida tipo transistor PNP
		Salida tipo transistor NPN
Alimentación alterna	Sensores con salida tipo relé	
	Sensores de dos hilos	
	Sensores de tres hilos	

Tabla 8.3. Clasificación de las interfaces de los sensores todo-nada de acuerdo con el tipo de alimentación.

Las interfaces de salida todo-nada están también, en general, implementadas con aislamiento galvánico para evitar que las sobrecorrientes o sobretensiones que se produzcan en los circuitos de salida puedan afectar al autómata programable. La tabla 8.4 muestra la clasificación de las interfaces de salida según el tipo de corriente que conmutan, y el tipo dispositivo que lleva a cabo la conmutación que se suele denominar actuador. Según el tipo de dispositivo de conmutación utilizado, la interfaz de salida puede ser de tipo relé, de tipo transistor (PNP o NPN) o de tipo triac o tiristor. Si la alimentación de la carga que se conecta al circuito de salida del autómata programable es continua, la salida debe ser del tipo relé o transistor. Si dicha alimentación es alterna, la salida tiene que ser del tipo relé, triac o tiristor. En cualquier caso, la misión de un circuito de salida es hacer que el dispositivo de potencia (relé, transistor, triac o tiristor) se active o se desactive de acuerdo con las acciones que ejecuta el programa de control del autómata programable.

INTERFACES DE SALIDA TODO-NADA	
Alimentación en continua	Salida del tipo relé
	Salida del tipo transistor PNP
	Salida del tipo transistor NPN
Alimentación en alterna	Salida del tipo relé
	Salida del tipo relé de estado sólido (tiristor o triac)

Tabla 8.4. Clasificación de las interfaces de salida todo-nada.

En sucesivos apartados se analizan los distintos tipos de interfaces de variables todo-nada de salida.

8.2.2.1.2 Interfaces de variables analógicas

Este tipo de interfaz conecta al autómata programable los sensores que generan variables analógicas, que son aquellas que pueden tener cualquier valor dentro de un determinado rango.

Su utilización es frecuente en muchos procesos industriales. Por ejemplo, en un tanque puede ser necesario conocer el nivel del líquido que contiene y no solamente saber si está lleno o vacío.

Las variables analógicas pueden ser de tensión o de corriente:

- **Variables analógicas de tensión**

Su comportamiento se puede representar mediante el circuito de la figura 8.3, en la que el sensor es un generador de tensión ideal (V_G) en serie con su impedancia de salida (Z_o). Si la impedancia de entrada del receptor es Z_i , la tensión de entrada analógica V_R del autómata programable cumple la ecuación:

$$V_R = \frac{Z_i}{Z_o + Z_i} V_G$$

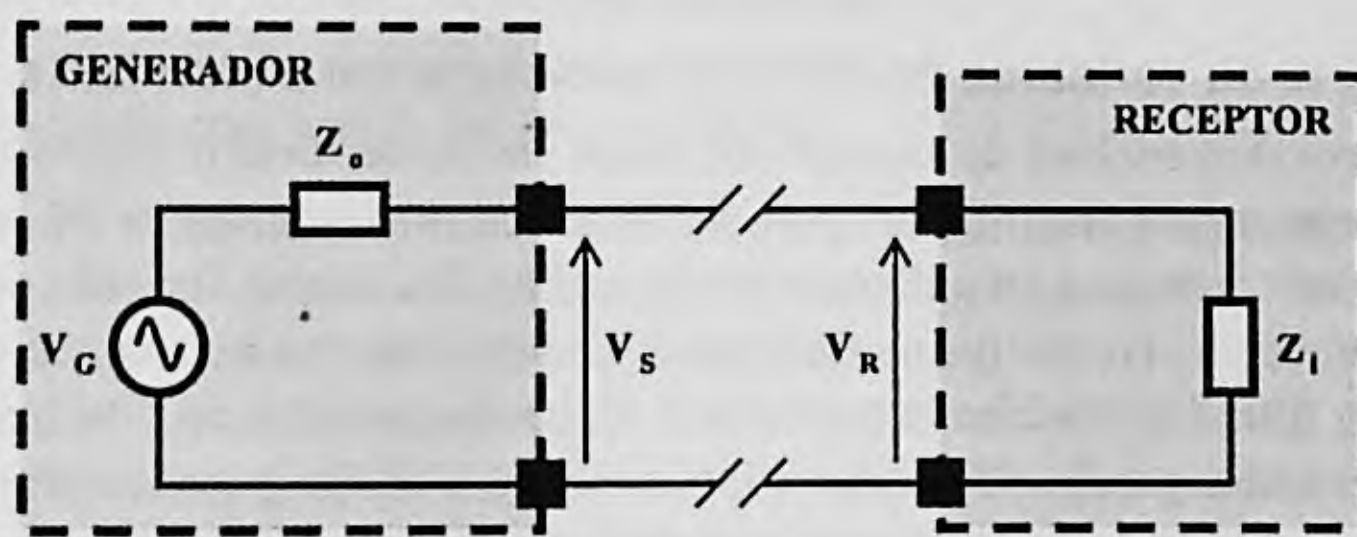


Figura 8.3. Generación de una variable analógica de tensión.

Si Z_i es mucho mayor que Z_o , V_R es aproximadamente igual a V_G y prácticamente toda la tensión que proporciona el generador se aplica al receptor. Por ello la impedancia de salida de este tipo de sensores debe ser pequeña (decenas de ohmios) y la impedancia de entrada del interfaz de entrada del autómata programable debe ser alta (superior a la decena de $K\Omega$).

- **Variables analógicas de corriente**

Su comportamiento se puede representar mediante el circuito de la figura 8.4, en la que el sensor es un generador de corriente ideal (I_G) en paralelo con su impedancia de salida (Z_o). Si la impedancia de entrada del receptor es Z_i , en este caso la entrada analógica de corriente del autómata programable, cumple la ecuación:

$$I_R = \frac{Z_o}{Z_o + Z_i} I_G$$

Si Z_i es mucho menor que Z_o , I_R es aproximadamente igual a I_G y prácticamente toda la corriente que proporciona el generador pasa a través del receptor. Por ello la impedancia de salida de este tipo de sensores debe ser muy elevada (superior a la decena de $K\Omega$) y la de entrada del autómata programable muy pequeña (del orden de la decena de ohmios).

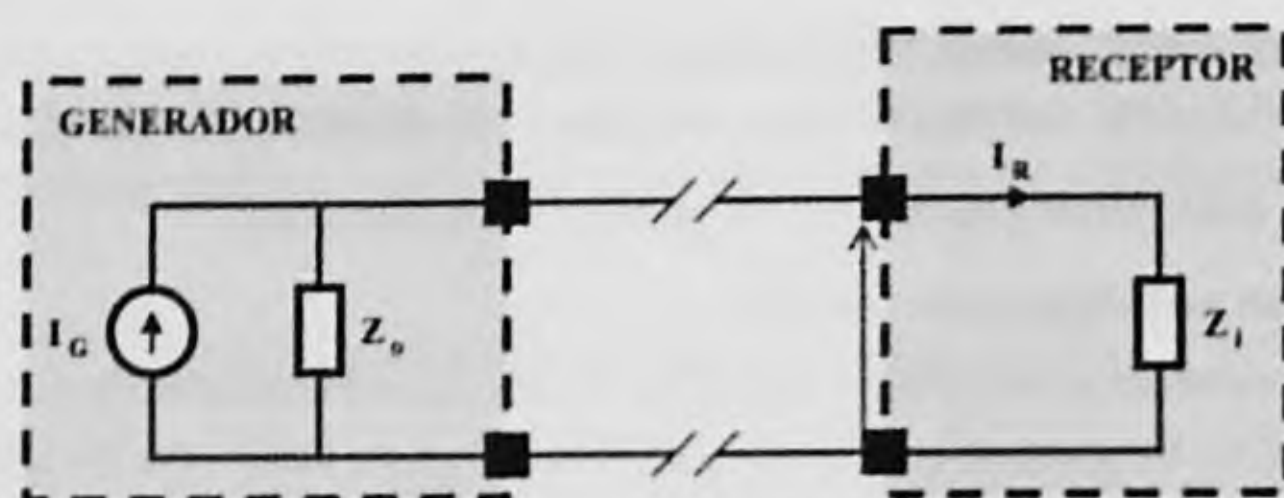


Figura 8.4. Generación de una variable analógica de corriente.

En sucesivos apartados se describen las interfaces de variables todo-nada y analógicas, tanto de entrada como de salida.

8.2.2.2 Interfaz de variables de entrada todo-nada sin aislamiento galvánico

Las interfaces de variables de entrada todo-nada sin aislamiento o separación galvánica se realizan solamente en continua. En la figura 8.5 se indica el circuito típico de este tipo de interfaz que acopla un contacto a un autómata programable. El circuito formado por la resistencia $R1$ y el condensador C , constituye un filtro paso-bajo que elimina los rebotes que se producen en el contacto y filtra las posibles interferencias electromagnéticas. Su constante de tiempo es del orden de los milisegundos. El diodo luminiscente (LED) D proporciona información visual del estado abierto o cerrado del contacto. La tensión de alimentación del circuito de la interfaz es continua y en algunos casos puede ser suministrada por el propio autómata programable.

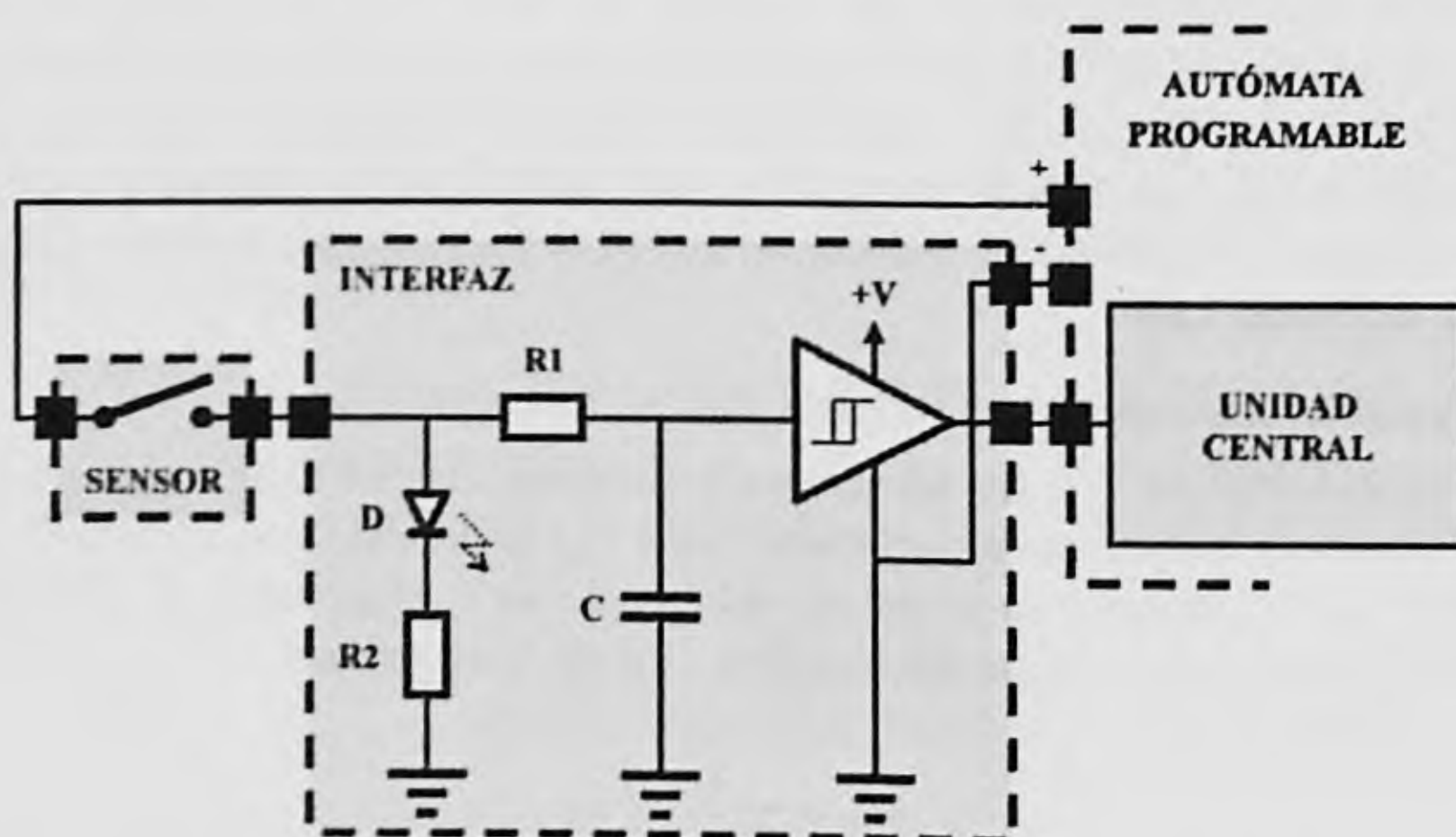


Figura 8.5. Circuito de interfaz típico alimentado en continua de una variable de entrada todo-nada del tipo contacto sin aislamiento galvánico de un autómata programable.

Este tipo de interfaz es adecuado cuando la longitud de los cables que conectan el contacto con el circuito es reducida y la posibilidad de que puedan aparecer sobretensiones, que provoquen una avería en el autómata programable, es prácticamente nula.

8.2.2.3 Interfaz de variables de entrada todo-nada con aislamiento galvánico

8.2.2.3.1 Introducción

Las interfaces de entrada de los sensores todo-nada se suelen suministrar en módulos como los de las figuras 8.1 y 8.2, que contienen varios circuitos de interfaz.

Las interfaces de este tipo se pueden alimentar en continua o en alterna y para analizar su funcionamiento se debe tener en cuenta que la salida de los sensores todo-nada puede ser un contacto libre de potencial o un dispositivo electrónico. En sucesivos apartados se analizan las diferentes variantes.

8.2.2.3.2 Interfaces de variables de entrada todo-nada con aislamiento galvánico y alimentación en continua

Sensores todo-nada con salida del tipo contacto

En la figura 8.6 se representa el circuito típico de la interfaz con aislamiento galvánico y alimentación en continua, entre un contacto y un autómata programable. Para lograr el aislamiento galvánico se utiliza un dispositivo optoacoplador (*Optocoupler*) formado por el diodo luminiscente (*LED*) D1 y un fototransistor. El diodo luminiscente D2 indica al usuario el estado de la variable de entrada y el diodo rectificador D3 protege a los diodos luminiscentes D1 y D2 contra una eventual inversión de la tensión de alimentación. El circuito formado por la resistencia R2 y el condensador C, conectado al fototransistor, constituye también un filtro paso-bajo, que elimina los rebotes que se producen en el contacto y filtra las posibles interferencias electromagnéticas. Su constante de tiempo es del orden de algunos milisegundos. Las señales de entrada deben estar activas durante un intervalo de tiempo mayor que la duración del ciclo del autómata programable para garantizar que éste detecta los cambios de nivel lógico de las mismas. Para detectar señales de elevada frecuencia se utilizan módulos especiales de contaje, que se describen en el apartado 8.2.3.2. La corriente que atraviesa el diodo luminiscente, cuando está cerrado el contacto, la proporciona una fuente de alimentación auxiliar que suele estar colocada en el mismo armario que el autómata programable.

En la figura 8.6 se supone que el sensor es un simple contacto (*Switch*), como por ejemplo un microrruptor (*Microswitch*) o un final de carrera (*Limit switch*), descritos en el apartado 7.7.2.2 del capítulo 7. Pero existen muchos sensores todo-nada, como por ejemplo los de proximidad, que están constituidos por un circuito electrónico que actúa sobre un relé. En este caso el circuito de interfaz se alimenta mediante una fuente auxiliar y se puede realizar de dos formas diferentes:

- La fuente de alimentación auxiliar proporciona energía al circuito electrónico del sensor y al diodo luminiscente del optoacoplador del circuito de interfaz (Figura 8.7).
- La fuente de alimentación auxiliar proporciona energía solamente al circuito electrónico del sensor y el diodo luminiscente del circuito de interfaz se conecta a la fuente de alimentación del autómata programable (Figura 8.8).

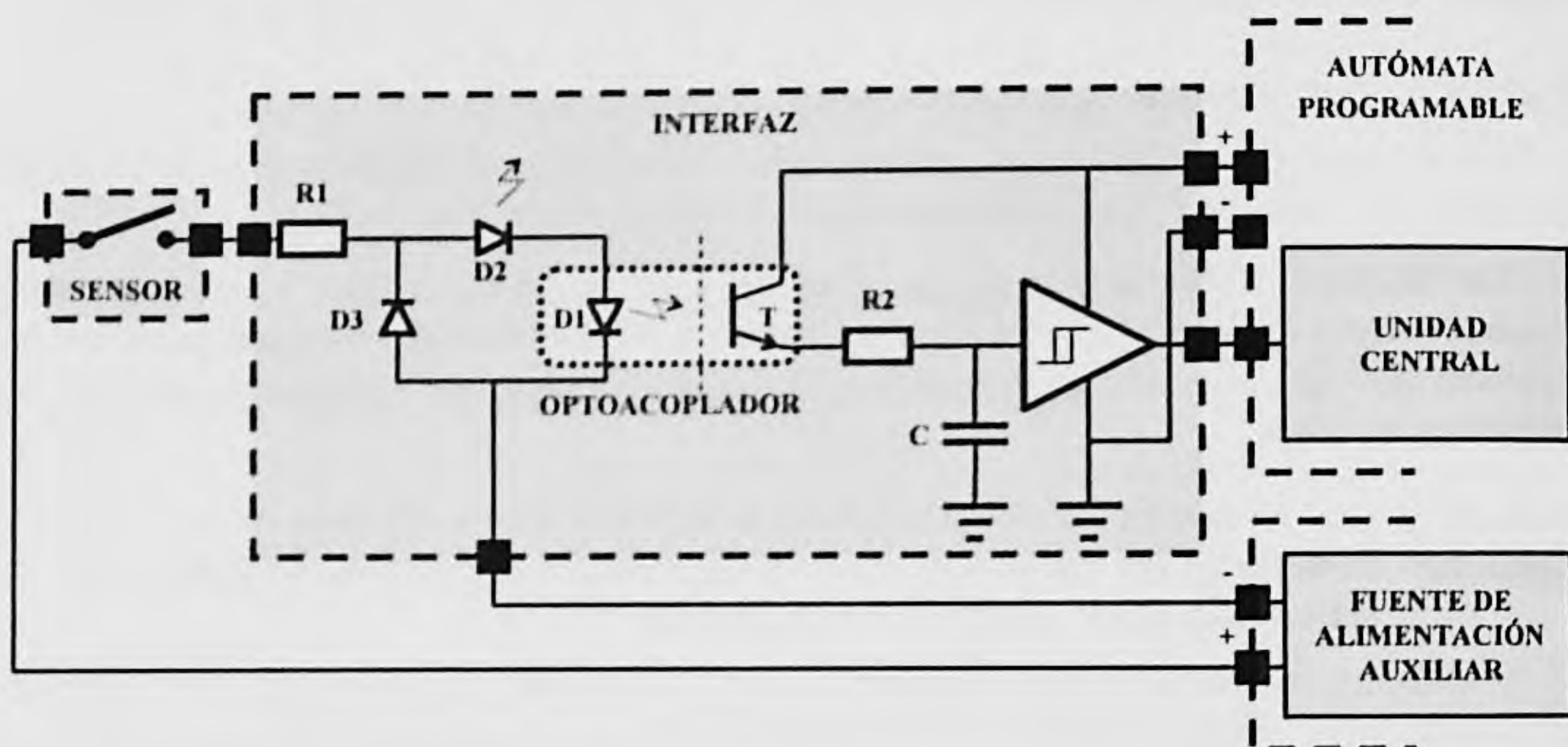


Figura 8.6. Circuito de interfaz típico alimentado en continua de una variable de entrada toda-nada del tipo contacto con aislamiento galvánico de un autómeta programable.

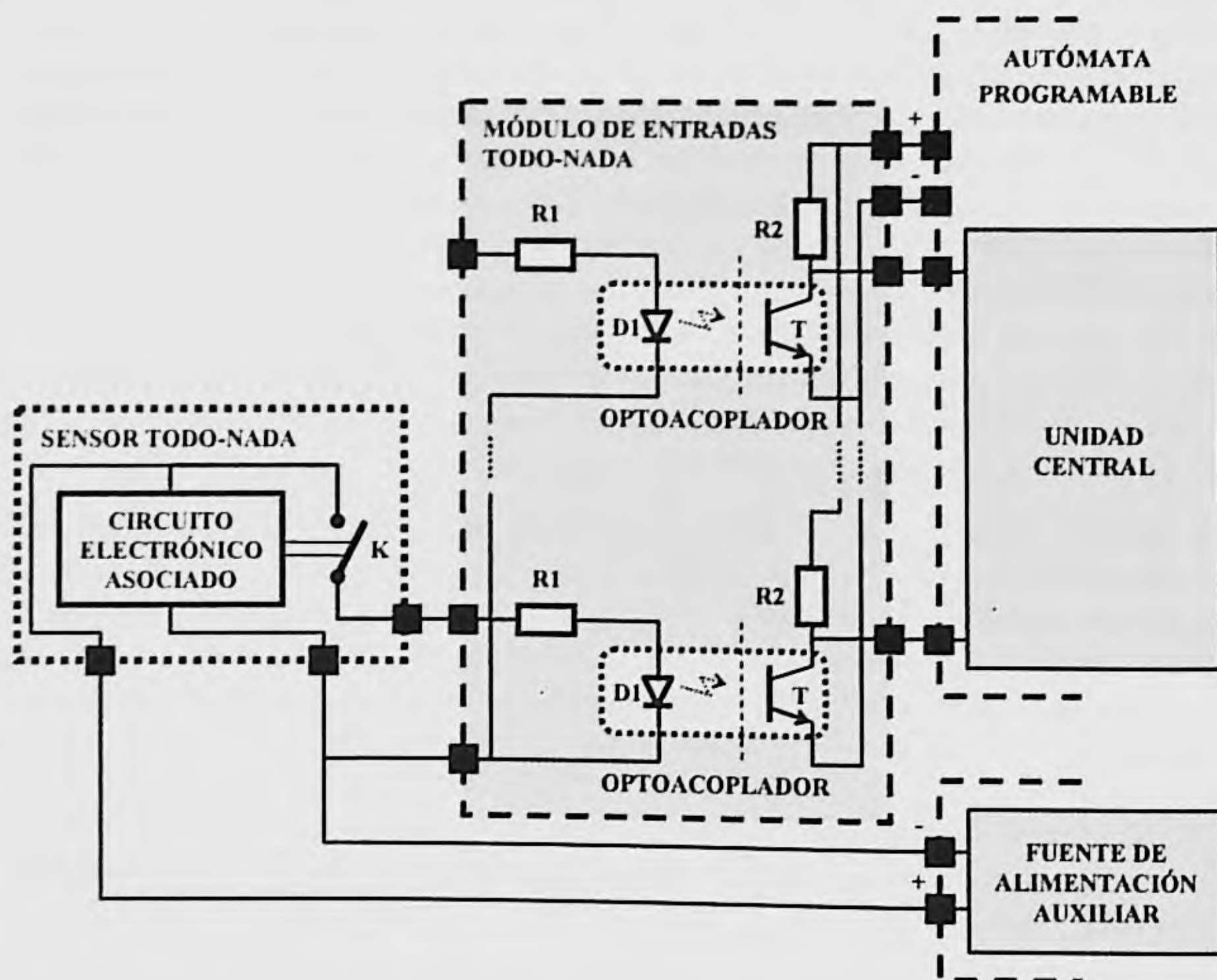


Figura 8.7. Circuito de interfaz típico alimentado en continua de un sensor electrónico toda-nada con salida del tipo relé con aislamiento galvánico en el que el diodo del optoacoplador se alimenta con la fuente de alimentación auxiliar.

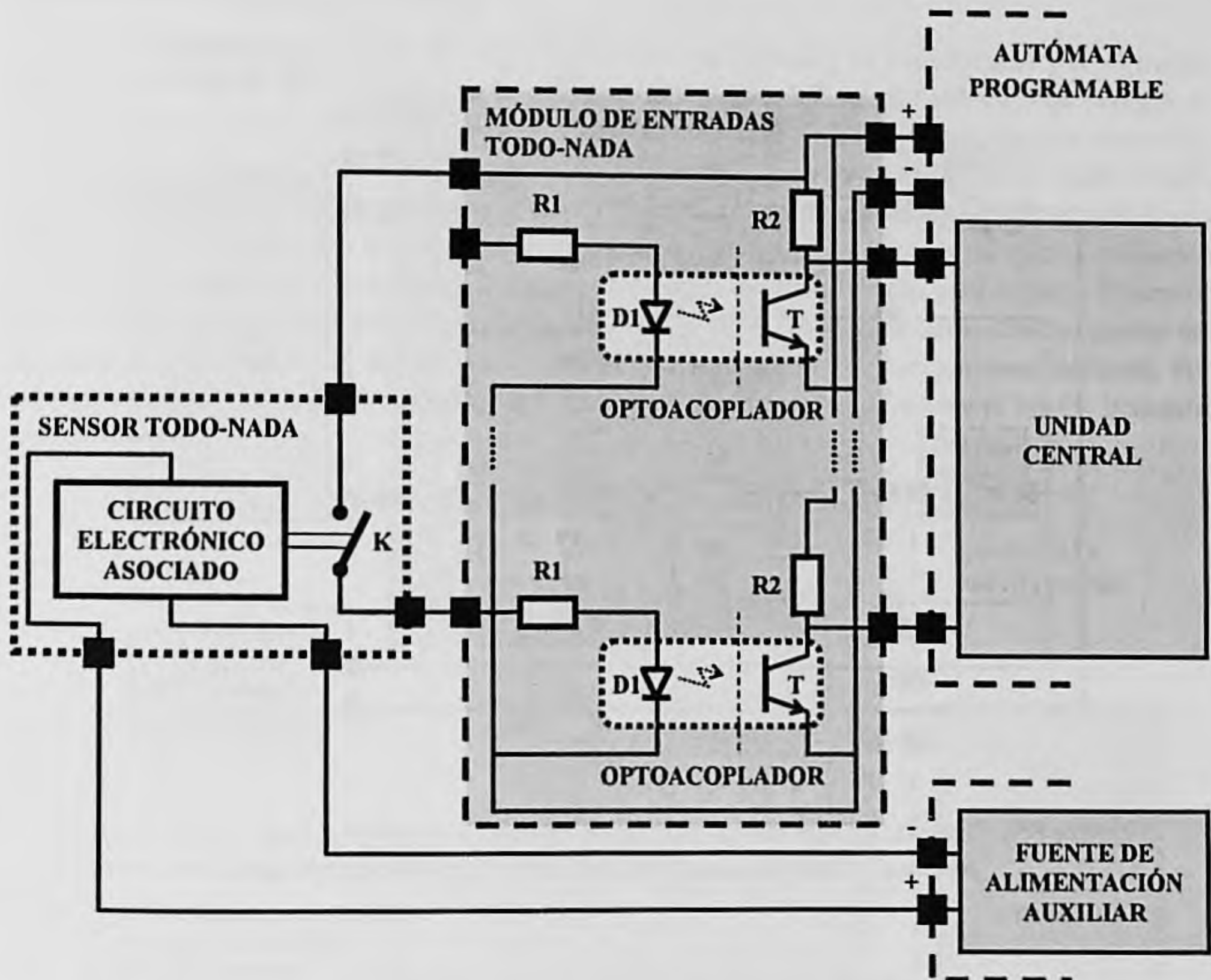


Figura 8.8. Circuito de interfaz típico alimentado en continua de un sensor electrónico todo-nada con salida del tipo relé con aislamiento galvánico en el que el diodo del optoacoplador se alimenta con la fuente de alimentación del autómata programable.

Tal como se indica en el apartado anterior, las interfaces todo-nada de variables de entrada se agrupan en módulos de entrada que poseen una alimentación única. Según que el terminal común a todas las interfaces de un módulo se conecte al positivo o al negativo de la fuente de alimentación de continua se pueden distinguir dos tipos de interfaces diferentes:

- Interfaces de entrada en las que el terminal común está conectado al negativo de la fuente de alimentación (Figura 8.9).
- Interfaces de entrada en las que el terminal común está conectado al positivo de la fuente de alimentación (Figura 8.10).

La tensión de alimentación más habitual de este tipo de interfaces es de 24 V.

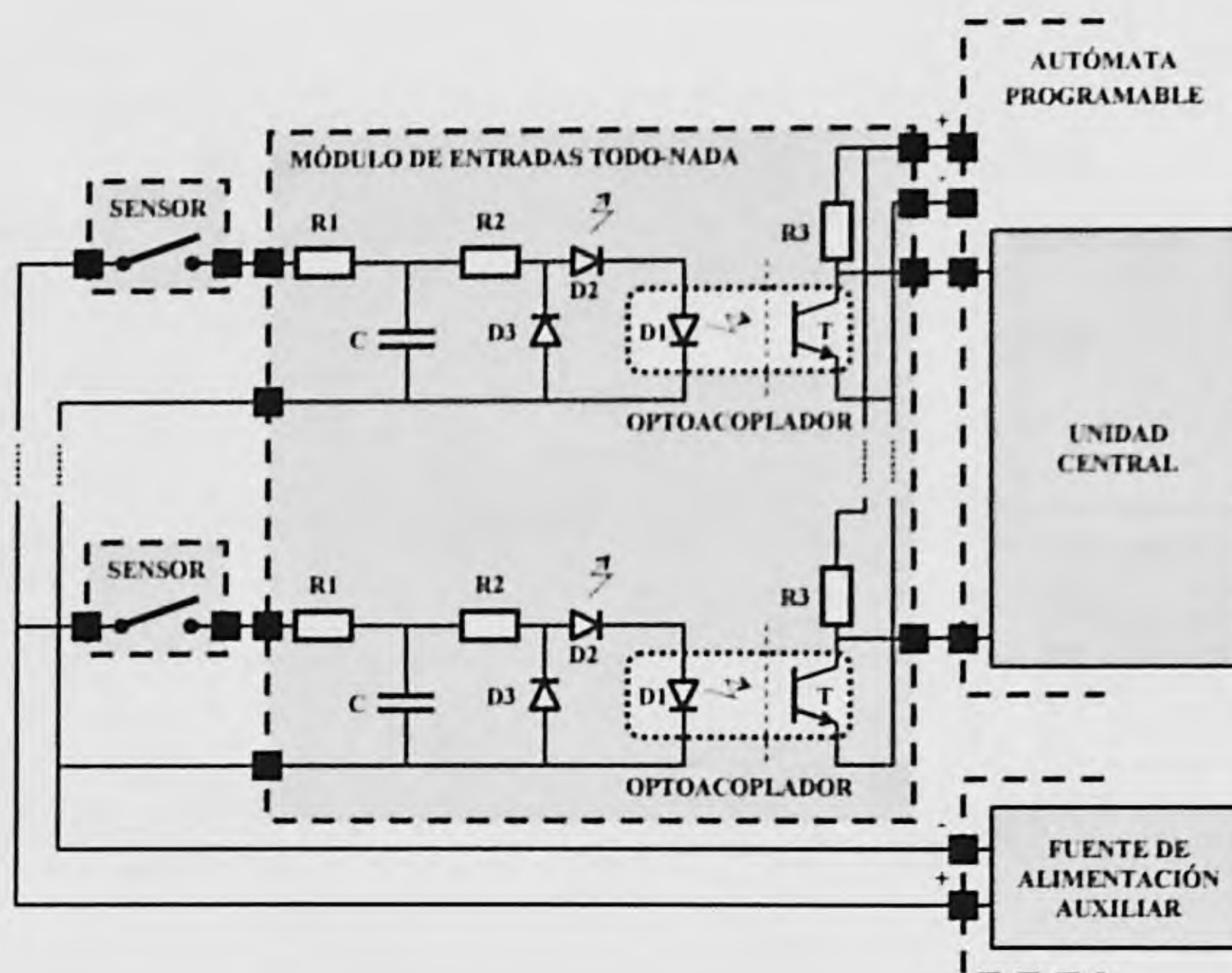


Figura 8.9. Interfaz de entrada en continua de sensores con salida del tipo relé, en la que el terminal común es el negativo de la fuente de alimentación auxiliar.

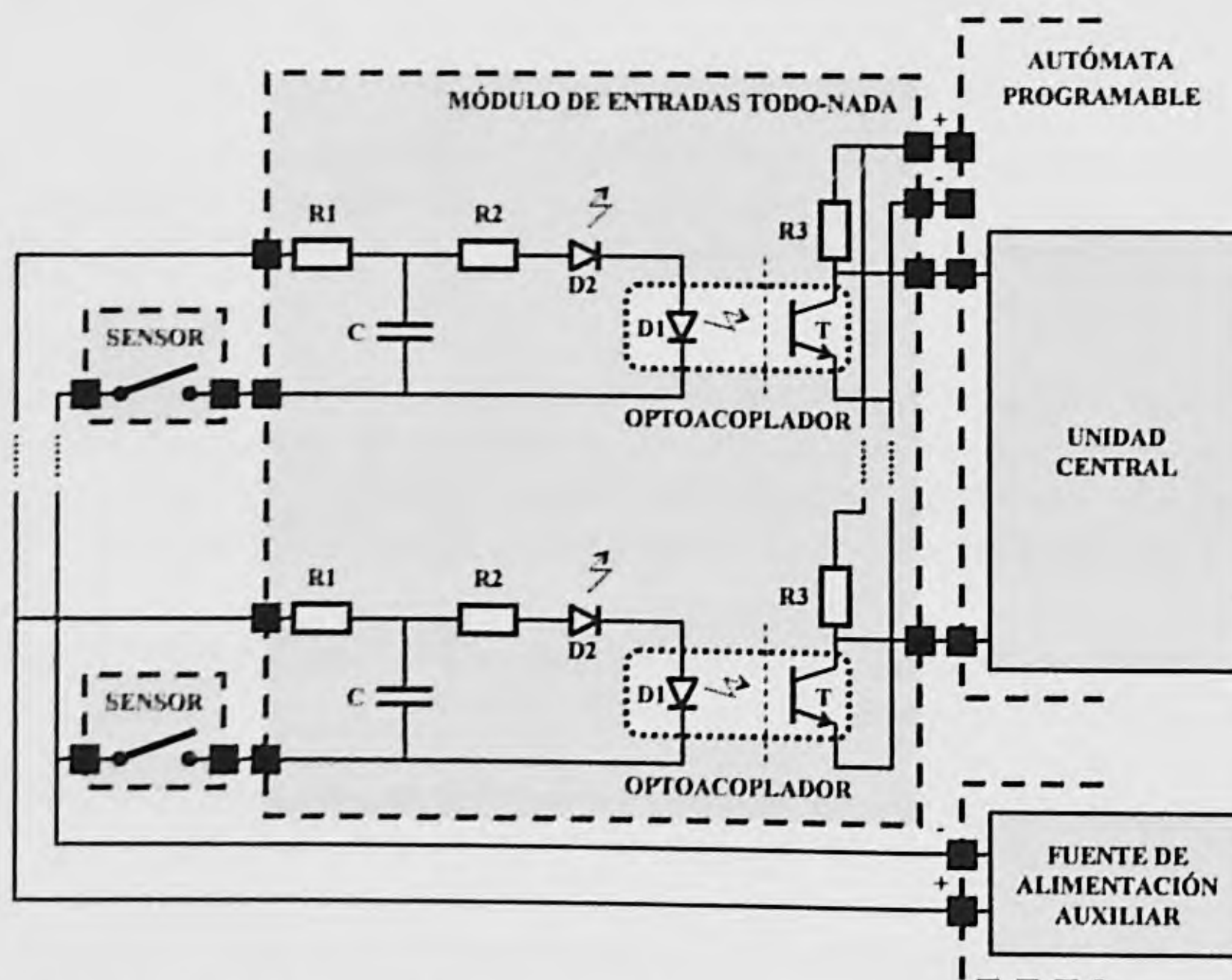


Figura 8.10. Interfaz de entrada en continua de sensores con salida del tipo relé, en la que el terminal común es el positivo de la fuente de alimentación auxiliar.

Sensores todo-nada de dos hilos

Los sensores todo-nada de dos hilos (*2-wire switches*) se caracterizan por el hecho de que, tal como se indica en el apartado 7.4.1.3 del capítulo 7, la alimentación de energía se realiza a través de uno de los terminales de salida conectado a la carga, que en este caso es el circuito de interfaz del autómata programable (Figura 8.11). Esto hace que la corriente de alimentación del sensor deba ser inferior a la corriente máxima que puede pasar a través del diodo luminoso D1 para que la salida del interfaz tenga un nivel de tensión tal que el autómata programable interprete que la salida del sensor no está activada. Por ello, para acoplar un sensor de dos hilos a una entrada todo-nada de un autómata programable, el usuario debe consultar las hojas de características de ambos para verificar que se cumple lo anteriormente indicado. Por tener solamente dos terminales, este tipo de sensor se utiliza frecuentemente en los procesos industriales.

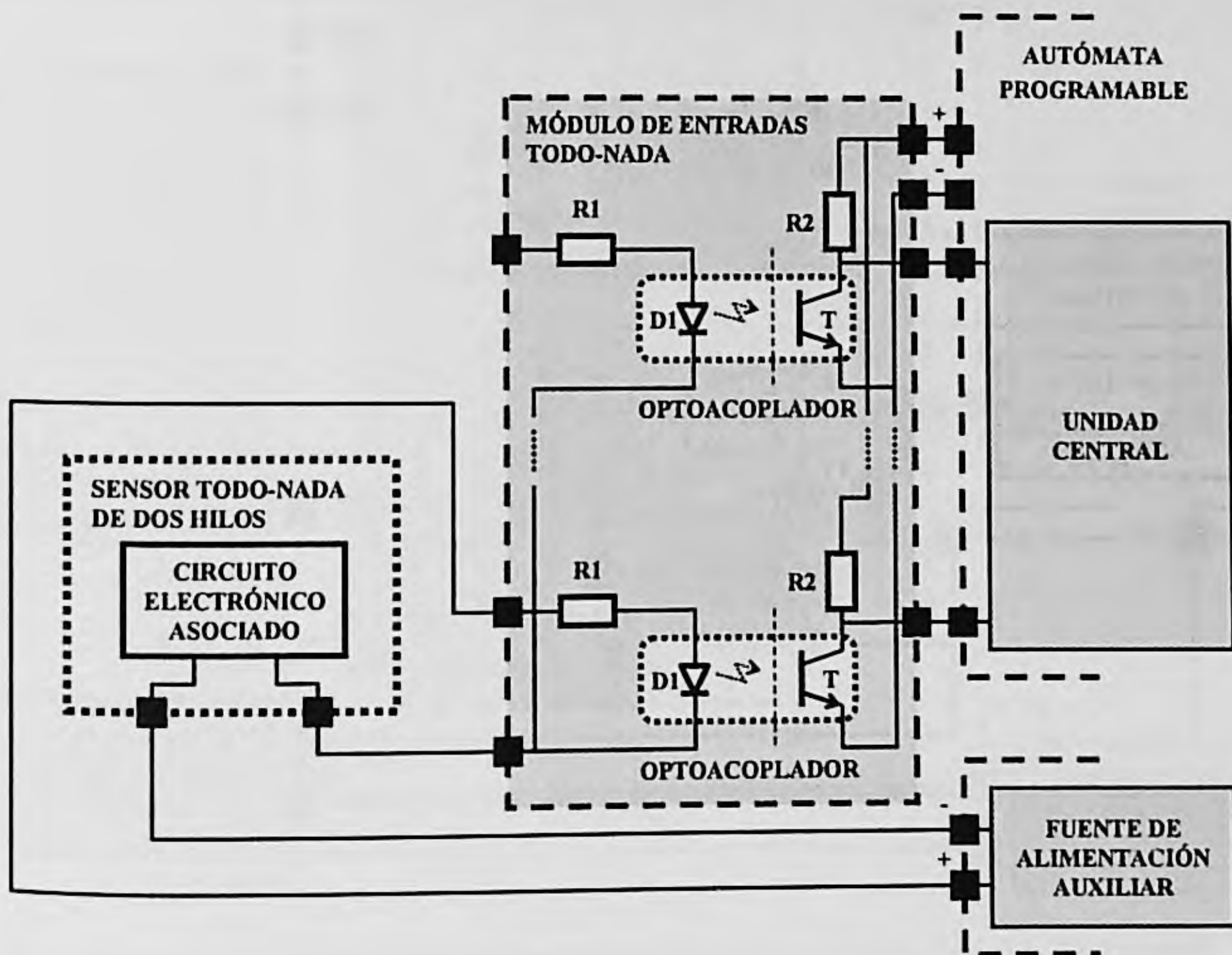


Figura 8.11. Circuito de interfaz típico con aislamiento galvánico y alimentado en continua de un sensor electrónico todo-nada de dos hilos.

Sensores todo-nada de tres hilos

Estos sensores se alimentan a través de un terminal distinto que el de salida y por ello la corriente de alimentación no pasa a través del circuito de entrada del autómata programable. Las figuras 8.12 y 8.13 indican la forma en que se conecta un sensor cuya salida está realizada

con un transistor PNP (sensor de salida PNP) y con un transistor NPN (sensor de salida NPN), respectivamente, a la entrada de un autómata programable. Los sensores de salida PNP se caracterizan por conmutar la corriente que sale (*Source current*) del circuito del módulo de entradas todo-nada del autómata programable, y por ello el negativo de la fuente de alimentación auxiliar es común a todas las entradas. Por el contrario, los sensores de salida NPN se caracterizan por conmutar la corriente que entra (*Sink current*) en el módulo de entradas todo-nada del autómata programable, y en este caso el positivo de la fuente de alimentación es el común a todas las entradas.

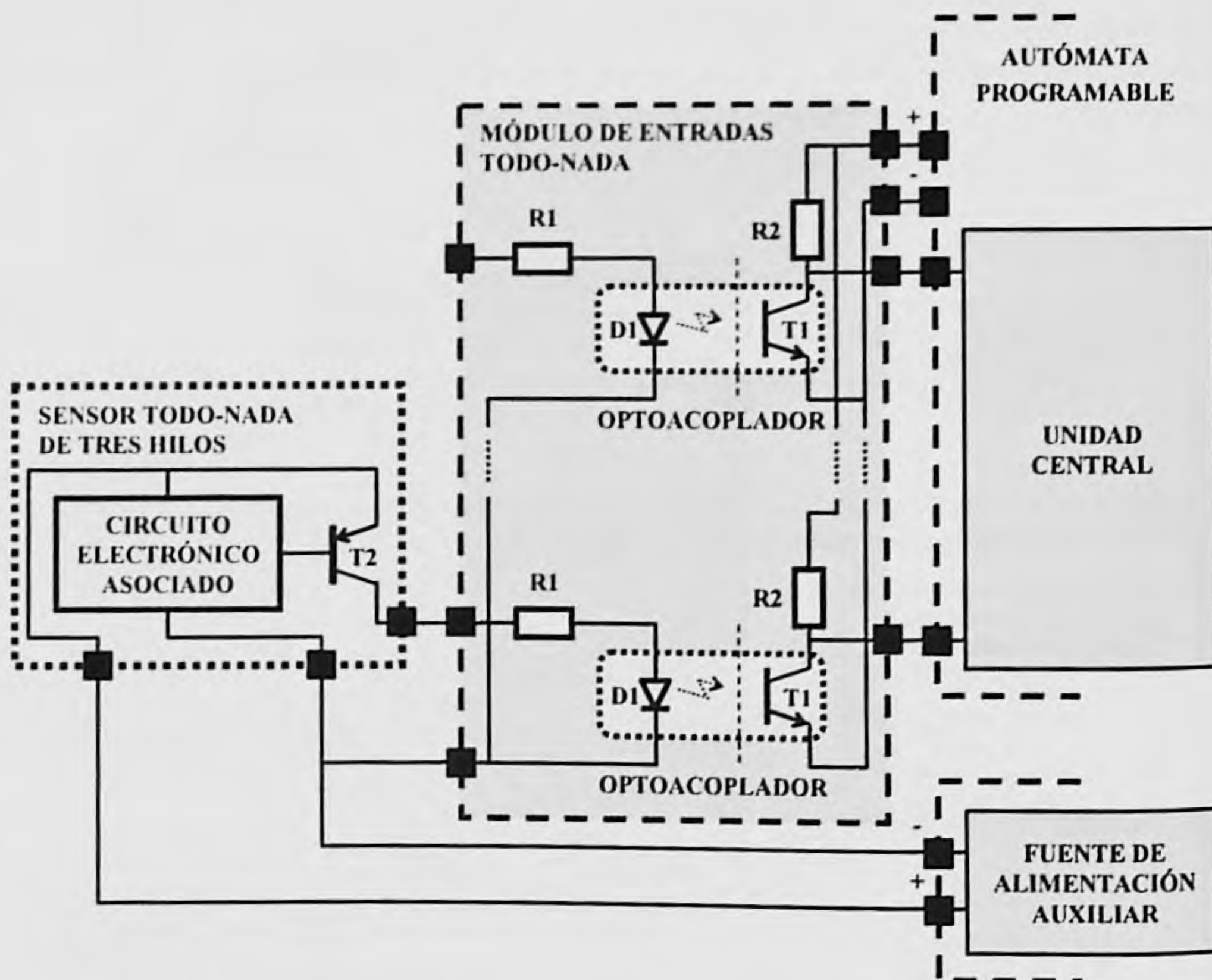


Figura 8.12. Circuito de interfaz típico con aislamiento galvánico y alimentado en continua de un sensor electrónico todo-nada de tres hilos con salida mediante un transistor PNP.

Los sensores de salida PNP y NPN se diferencian por su comportamiento cuando se produce un cortocircuito entre la salida del sensor y el negativo de la fuente de alimentación

En las figuras 8.14a y 8.14b se muestra, respectivamente, la conexión de un sensor de salida NPN y otro de salida PNP a un autómata programable. Si en el circuito de la figura 8.14a, se produce un cortocircuito entre los hilos 2 y 3, el circuito del módulo de interfaz todo-nada del autómata programable no puede diferenciarlo de la situación en la que el sensor está activado

y el fallo no se detecta. Por el contrario, si se produce un cortocircuito entre los hilos 2 y 3 del circuito de la figura 8.14b, en el instante que se activa el sensor el transistor PNP se satura y da lugar a un cortocircuito en la fuente de alimentación, lo que hace que su tensión se anule y por tanto que el fallo se convierta en avería. Ello hace que los sensores de salida PNP se utilicen más que los sensores de salida NPN, en especial cuando al no detectar el fallo se produce una situación peligrosa.

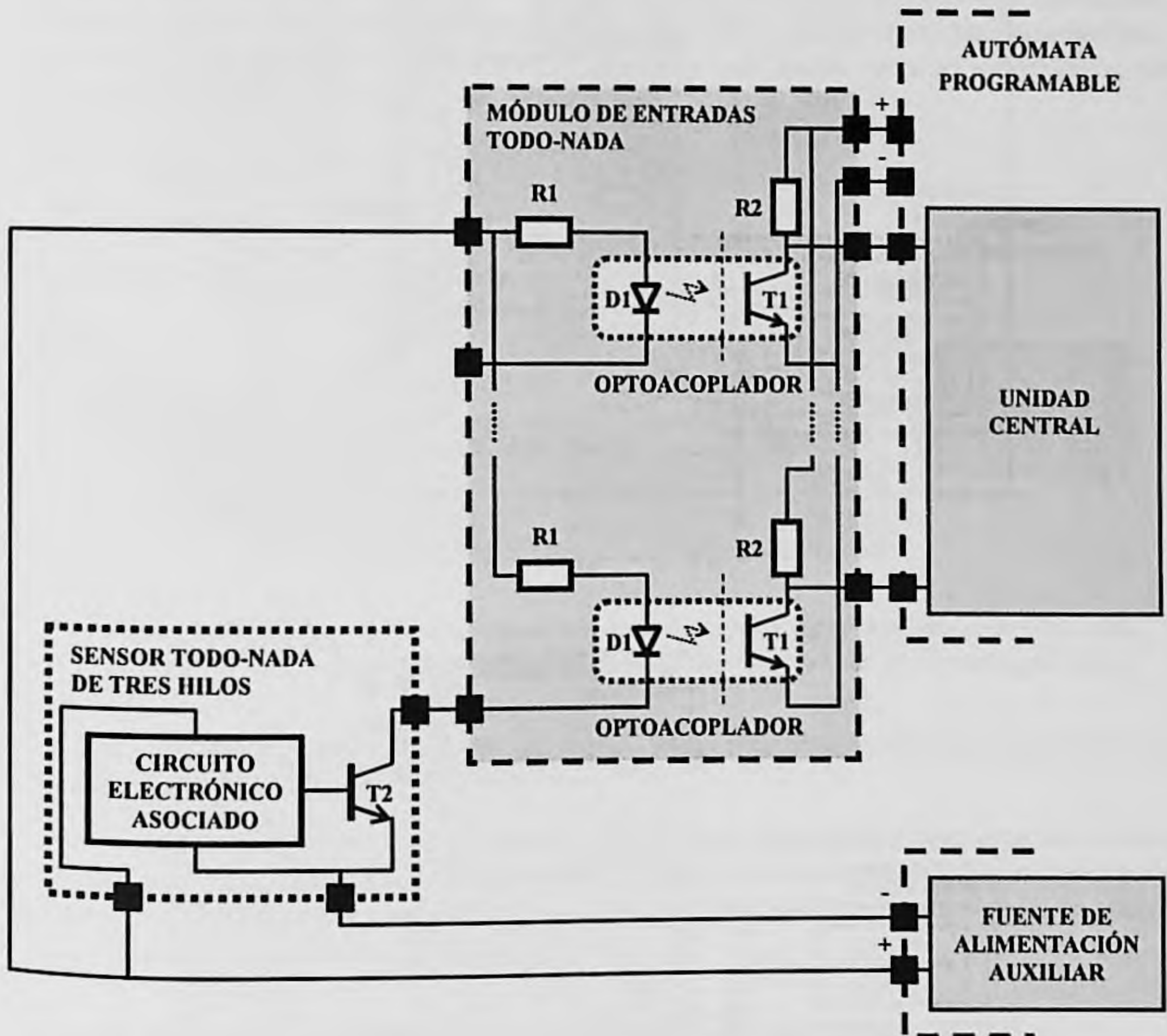
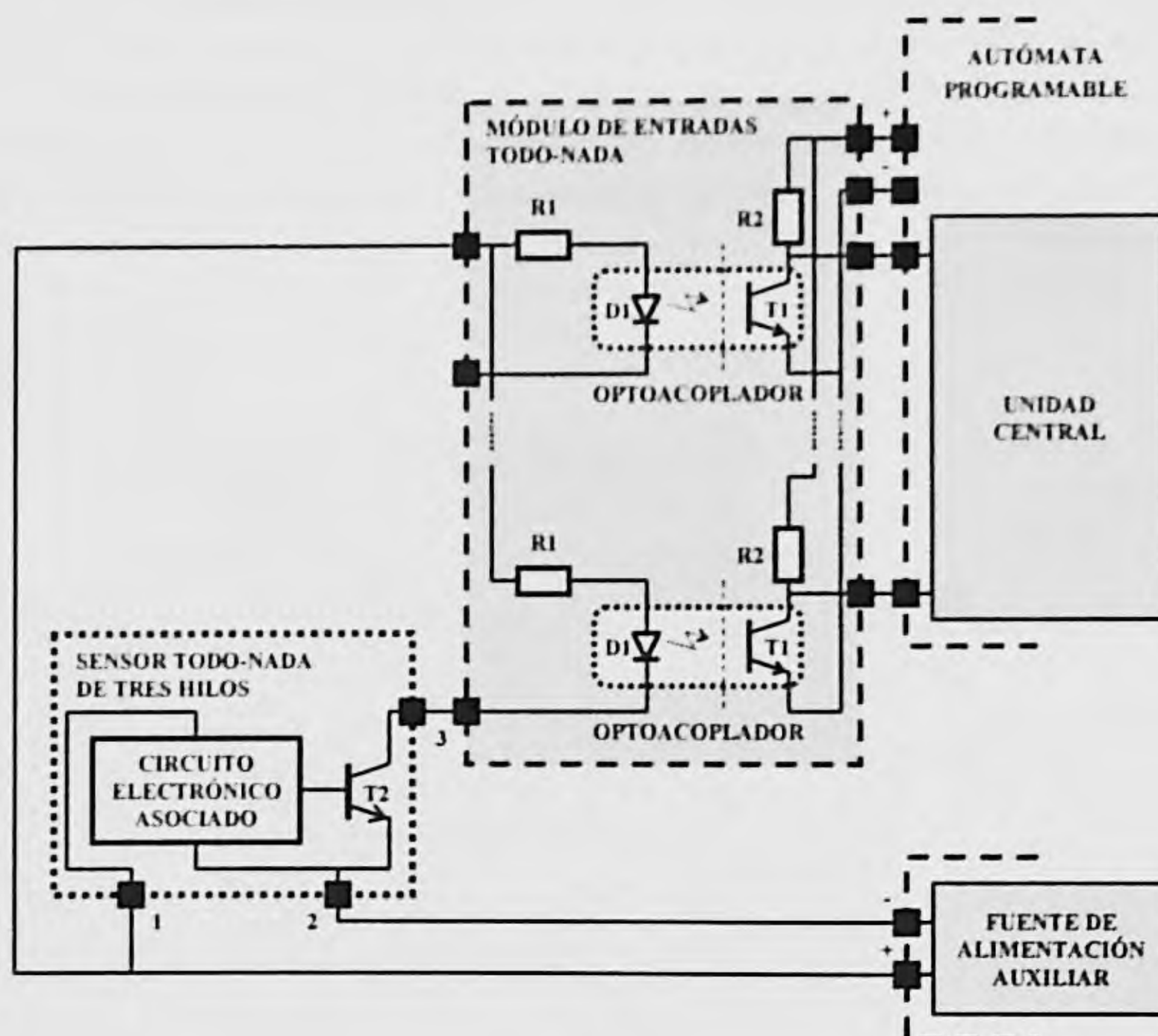
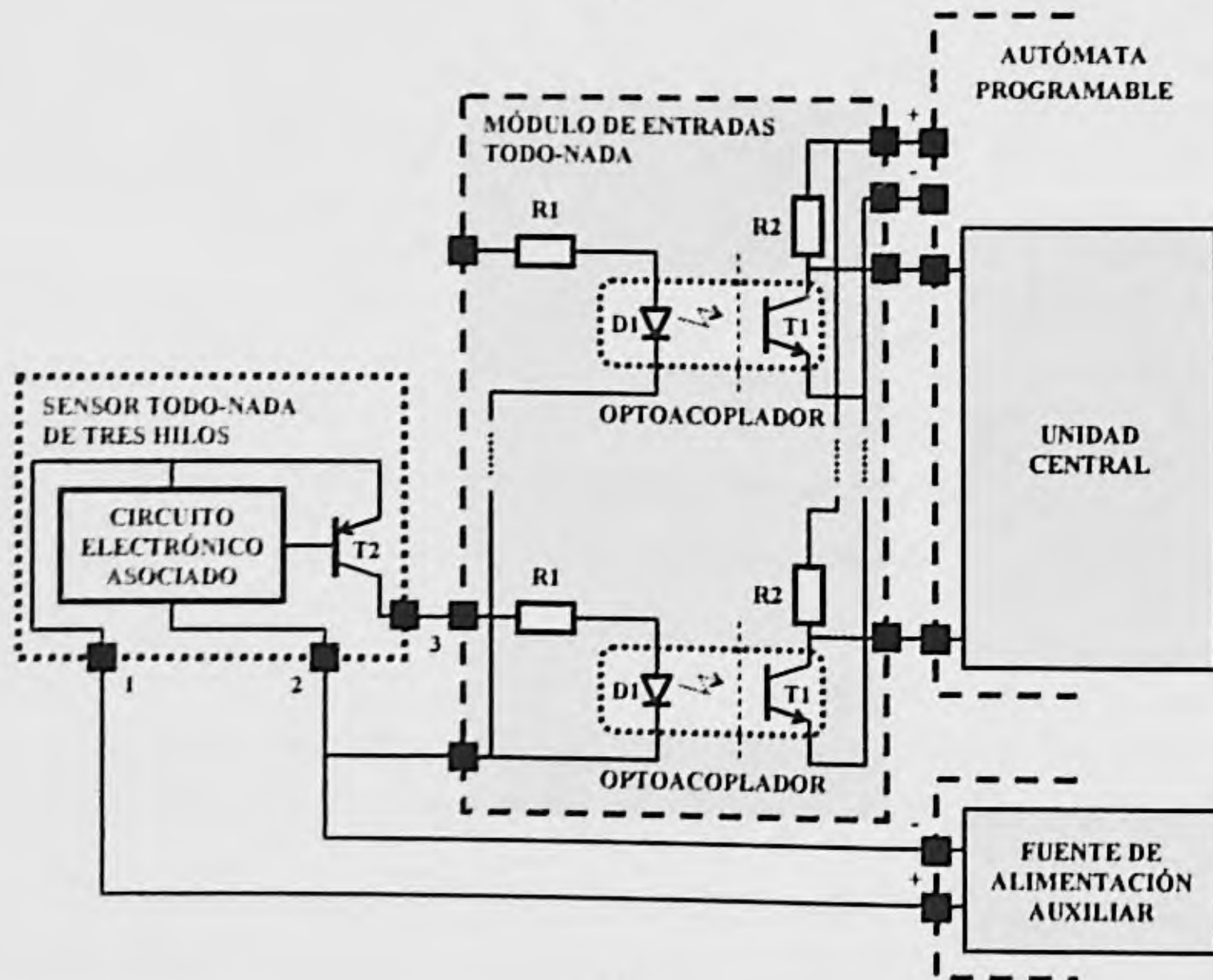


Figura 8.13. Circuito de interfaz típico con aislamiento galvánico y alimentado en continua de un sensor electrónico todo-nada de tres hilos con salida mediante un transistor NPN

Es conveniente resaltar que en los interfaces de las figuras 8.9 y 8.10, el diodo de señalización que indica la activación del sensor está colocado en la parte de circuito de interfaz conectado en serie con el sensor, pero también se puede colocar en la parte del circuito conectada eléctricamente al autómata programable. En ambos casos, un fallo del optoacoplador puede hacer que el diodo luminiscente emita luz sin que el sensor esté activado.



a)



b)

Figura 8.14. Análisis del efecto de un cortocircuito en la conexión entre un sensor electrónico todo-nada de tres hilos con salida mediante un transistor NPN (a) y otro con salida mediante transistor PNP (b).

8.2.2.3.3 Interfaz de variables de entrada todo-nada con aislamiento galvánico y alimentación en alterna

En la figura 8.15 se representa el circuito típico de la interfaz con aislamiento galvánico y alimentación en alterna, entre un final de carrera y un autómata programable. El puente rectificador proporciona una señal variable continua y la resistencia R_1 limita el valor máximo de la corriente que pasa a través del diodo luminoso D_1 del optoacoplador y del diodo luminoso D_2 de señalización. El circuito formado por R_2 y C constituye un filtro paso-bajo que elimina los rebotes que se producen en el contacto y filtra las posibles interferencias electromagnéticas, así como la componente alterna de la señal. Esta constante de tiempo suele ser mayor que en las entradas de continua y alcanza los 20 milisegundos.

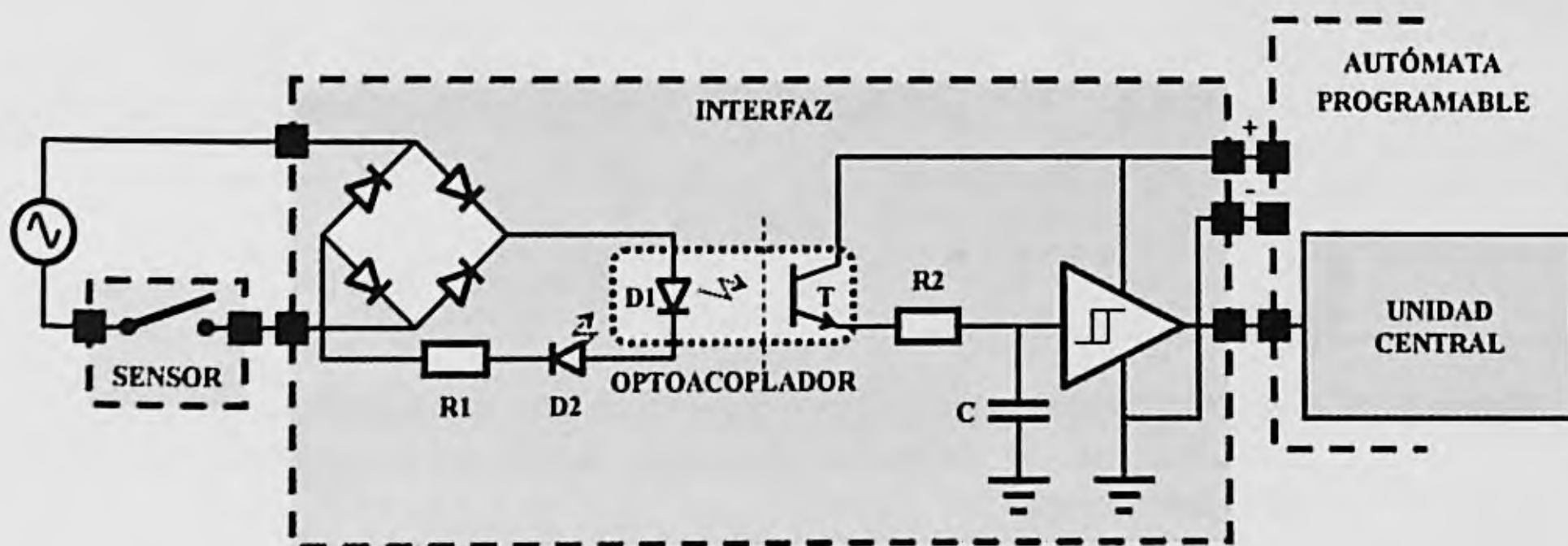


Figura 8.15. Circuito de interfaz típico con aislamiento galvánico y alimentado en alterna de una variable de entrada todo-nada del tipo contacto con aislamiento galvánico.

Las tensiones de alimentación más habituales en las interfaces todo-nada de alimentación en alterna son de 120 V ó 230 V.

Los circuitos de interfaz de este tipo también se agrupan en módulos que tienen un terminal común y el otro terminal se conecta a una fase de la alimentación a través del correspondiente sensor. En la figura 8.16 se representa el circuito de un módulo de interfaz alimentado en alterna y su conexión a sensores con salida del tipo relé.

8.2.2.4 Interfaces de variables de salida todo-nada con alimentación en continua

8.2.2.4.1 Clasificación y descripción

Tal como se indica en el apartado 8.2.2.1, estos interfaces son los encargados de activar los correspondientes actuadores para controlar el proceso. El dispositivo de conmutación utilizado (Tabla 8.4) en una interfaz de salida todo-nada puede ser un relé o un dispositivo electrónico. Las salidas de tipo relé suelen conmutar corrientes mayores pero su frecuencia máxima de operación es menor. A continuación se describen las diferentes interfaces de variables de salida todo-nada y la influencia en ellas del tipo de carga.

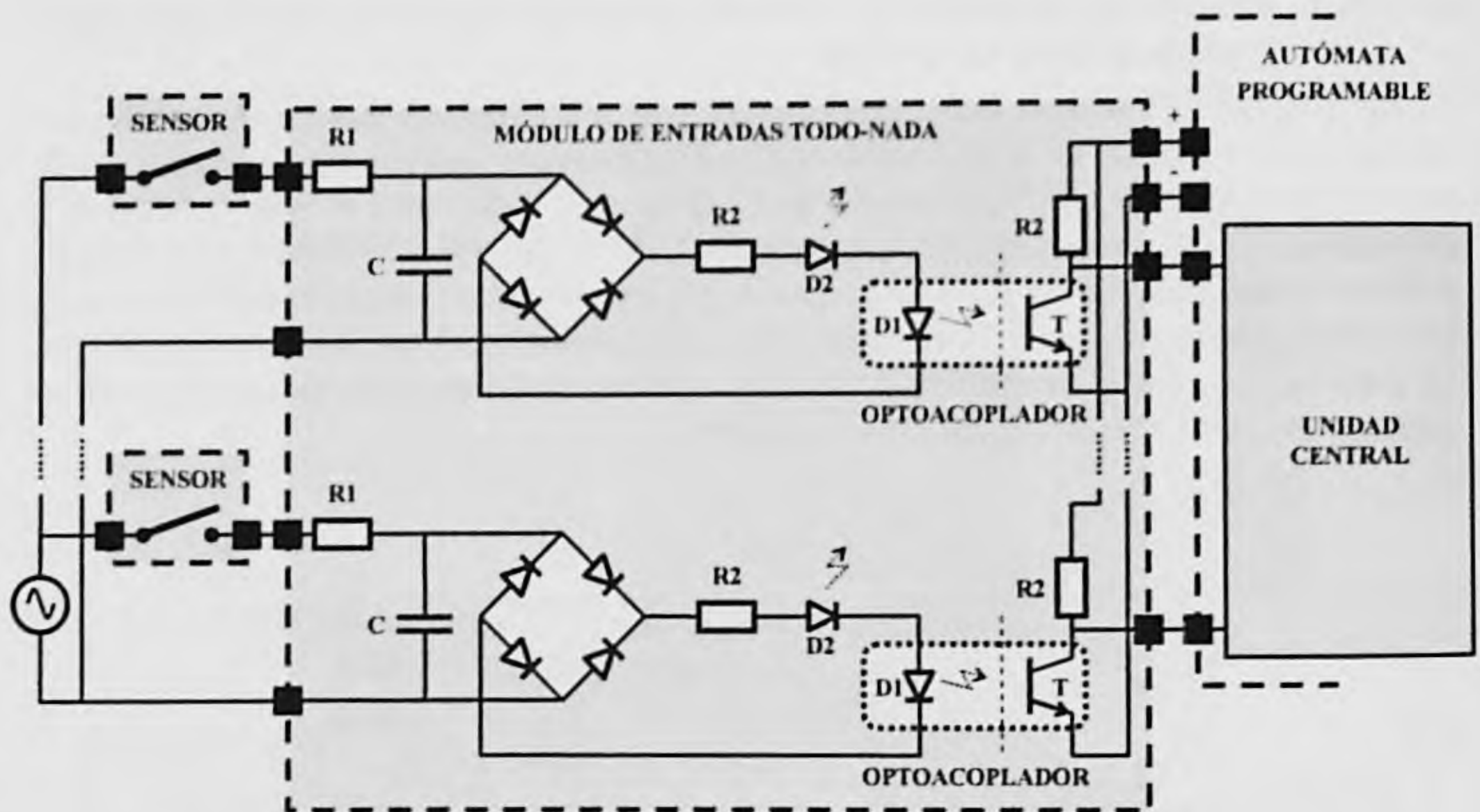


Figura 8.16. Circuito electrónico típico de un módulo de interfaz de variables de entrada todo-nada alimentado en alterna con aislamiento galvánico y su conexión a sensores de salida del tipo relé.

Interfaces de variables de salida todo-nada con relé

En la figura 8.17, se representa el circuito típico de un módulo de salidas todo-nada de un autómata programable, implementadas con relés. En este tipo de interfaz de salida se pueden utilizar contactos normalmente abiertos "NA" (en inglés "NO" de *Normally Open*), normalmente cerrados "NC" (en inglés "NC" de *Normally Close*) o conmutados del tipo SPDT (*Single Pole Single Through*), descritos en el apartado 7.4.1.3 del capítulo 7 (Figura 7.39). En todos ellos las salidas están libres de potencial, lo que permite que las cargas se puedan alimentar en continua o en alterna y que, además, cada salida pueda activar actuadores con distinto tipo de alimentación (alterna o continua) y con valores de tensión diferentes, tal como se indica en la figura 8.17.

En la práctica, los fabricantes de autómatas programables comercializan módulos de salida con relés, cuyos contactos tienen un terminal común para reducir el número de terminales del módulo y simplificar su conexionado. En este caso los actuadores se suelen alimentar con la misma fuente de alterna o de continua. También es frecuente que los fabricantes incorporen alguno de los tipos de circuitos de protección del contacto de cada relé que se describen en el apartado 8.2.2.4.2. La frecuencia de conmutación de este tipo de salidas es menor que la de las salidas realizadas con dispositivos electrónicos (también denominadas salidas de tipo semiconductor) que se estudian a continuación.

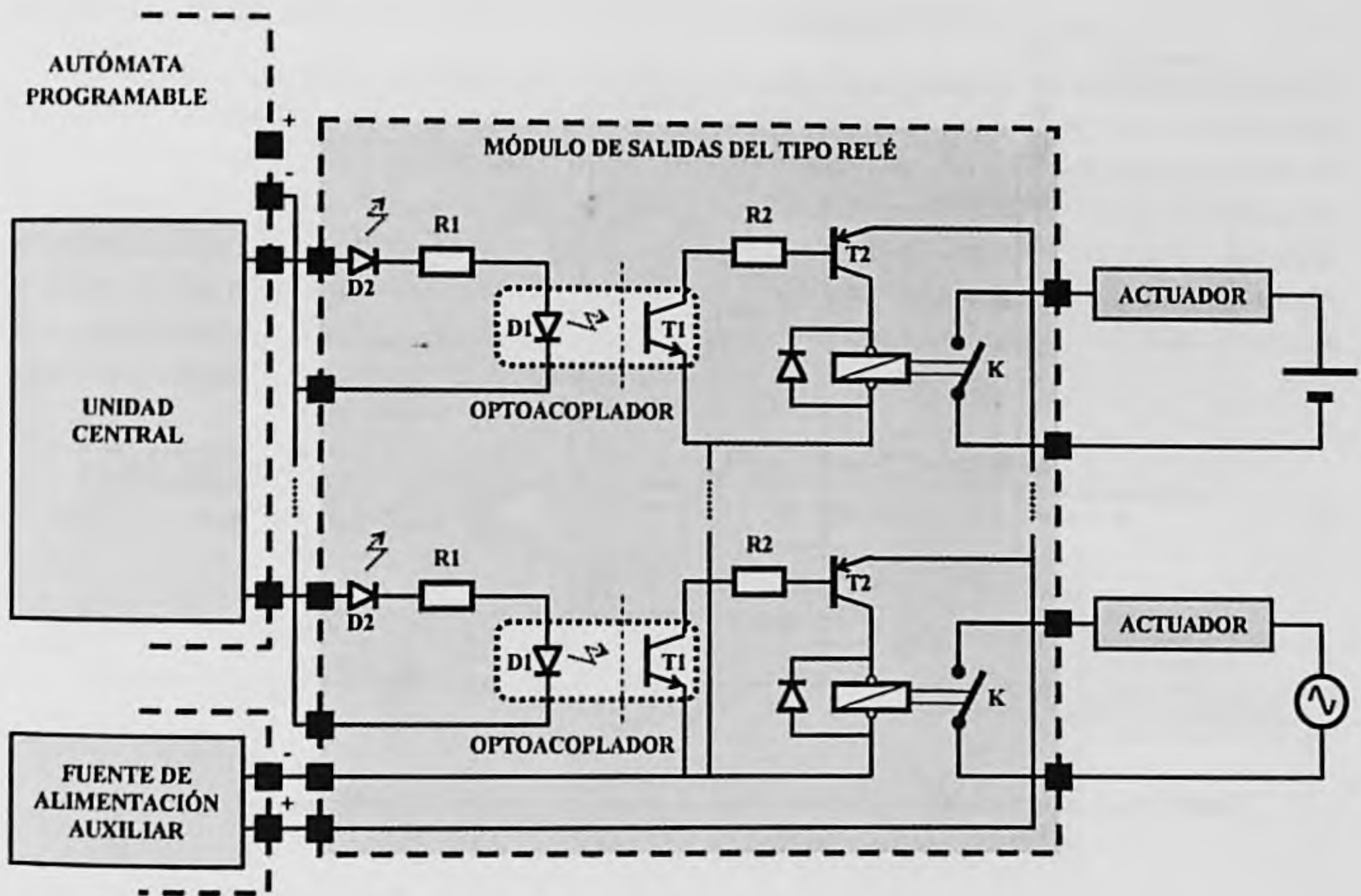


Figura 8.17. Circuito típico de un módulo de interfaz de salida implementado con relés.

Interfaces de variables de salida todo-nada con transistor NPN

En la figura 8.18 se muestra un módulo de interfaz de salidas todo-nada cuyos circuitos están implementados con transistores NPN controlados a través de un circuito comparador. En este tipo de interfaz el actuador se tiene que alimentar en continua y el diseñador de la instalación debe comprobar que la corriente que pasa a través del transistor NPN cuando está saturado es inferior a la corriente máxima que puede conmutar. Se debe tener en cuenta que, como consecuencia de la forma de funcionamiento de este dispositivo electrónico (salida de transistor NPN), la carga (actuador) se debe conectar entre el colector del transistor y el positivo de la fuente de alimentación auxiliar.

Existen diferentes variantes del circuito de este tipo de módulos. La figura 8.19 muestra otro circuito que se diferencia del de la figura 8.18 en que el diodo luminoso D2 de señalización de salida activada está situado en la parte del circuito alimentado por la fuente de alimentación auxiliar externa. Además el colector del transistor está conectado a un diodo D3 (diodo de libre circulación) que, tal como se indica en el apartado 8.2.2.4.2, elimina las sobretensiones que se producen al cortar la corriente que pasa a través de los actuadores de tipo inductivo (relés, contactores, electroválvulas, etc.).

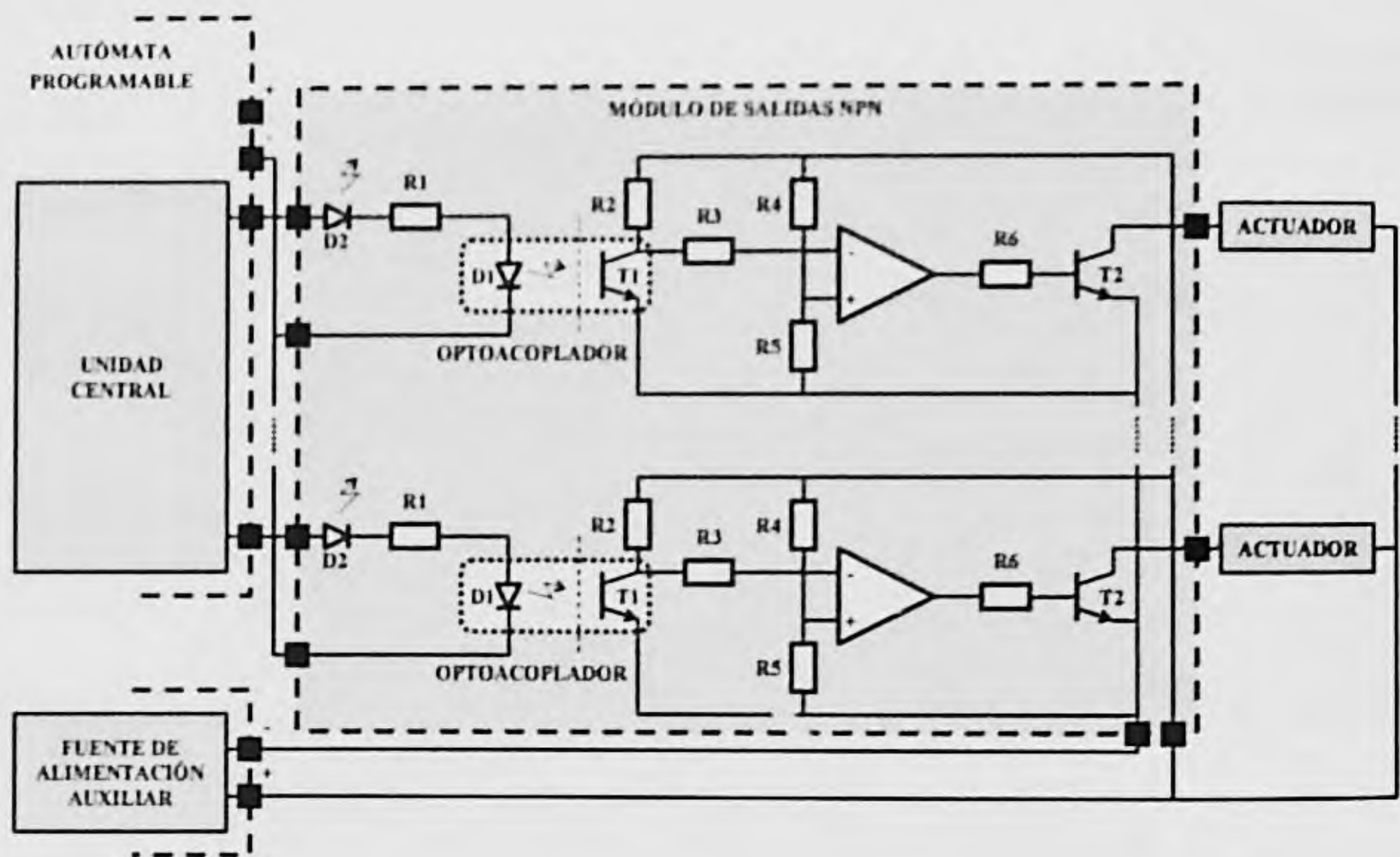


Figura 8.18. Circuito típico de un módulo de interfaz de salida implementado con transistores NPN en colector abierto.

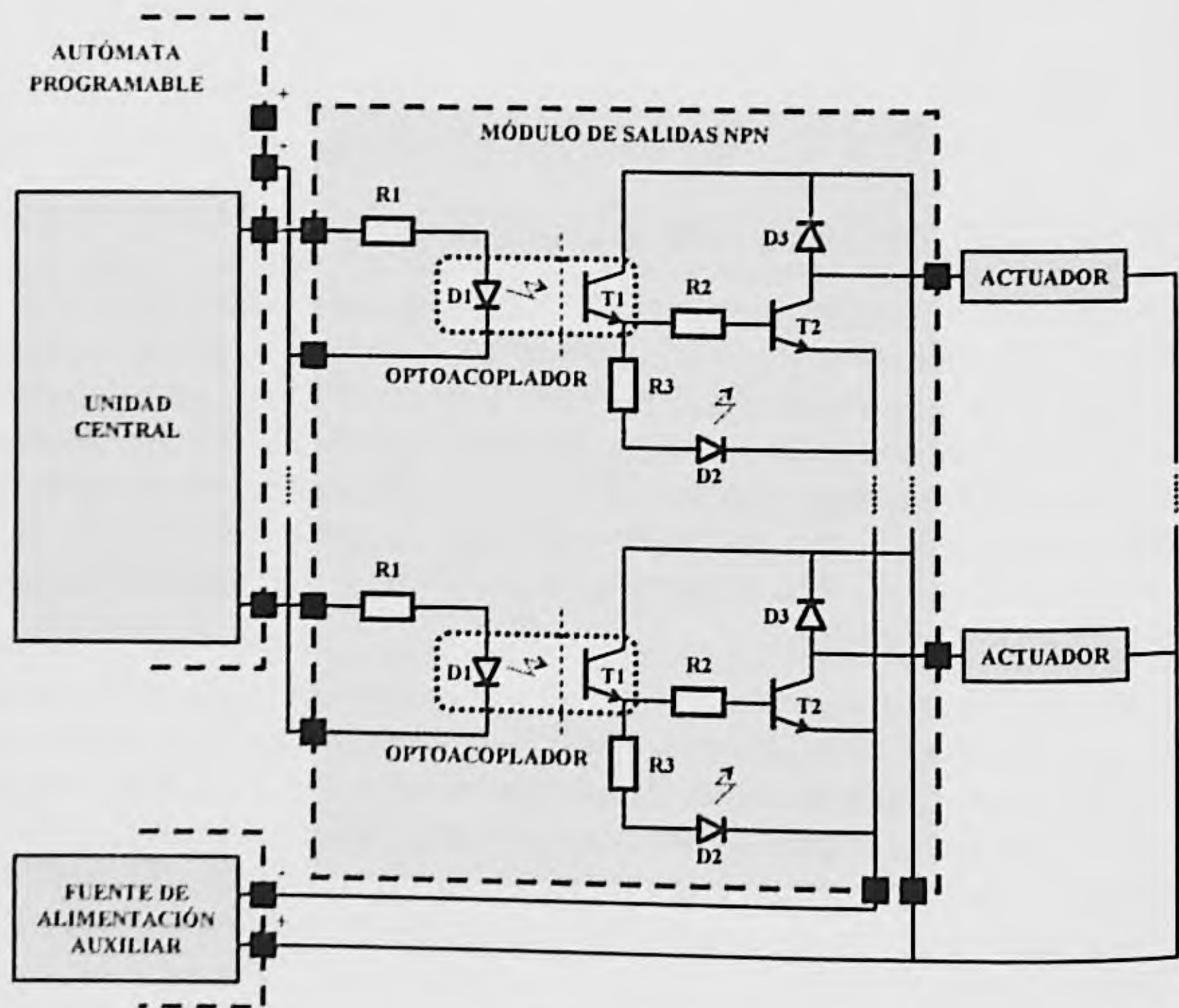


Figura 8.19. Circuito típico de un módulo de interfaz de salida implementado con transistores NPN y diodos de libre circulación.

Interfaces de variables de salida todo-nada con transistor PNP

En este tipo de interfaces el actuador también se tiene que alimentar en continua y la carga (actuador) se conecta entre el colector del transistor y el negativo de la fuente de alimentación auxiliar, tal como se indica en la figura 8.20 que muestra un circuito típico. El diodo D3 es, al igual que en la salida de transistor NPN, un diodo de libre circulación que, tal como se indica en el apartado 8.2.2.4.2, elimina las sobretensiones que se producen al cortar la corriente que pasa a través de los actuadores de tipo inductivo (relés, contactores, electroválvulas, etc.). El diodo de señalización D2 está situado, en este caso, en la parte del circuito alimentado por la fuente de alimentación del autómata programable.

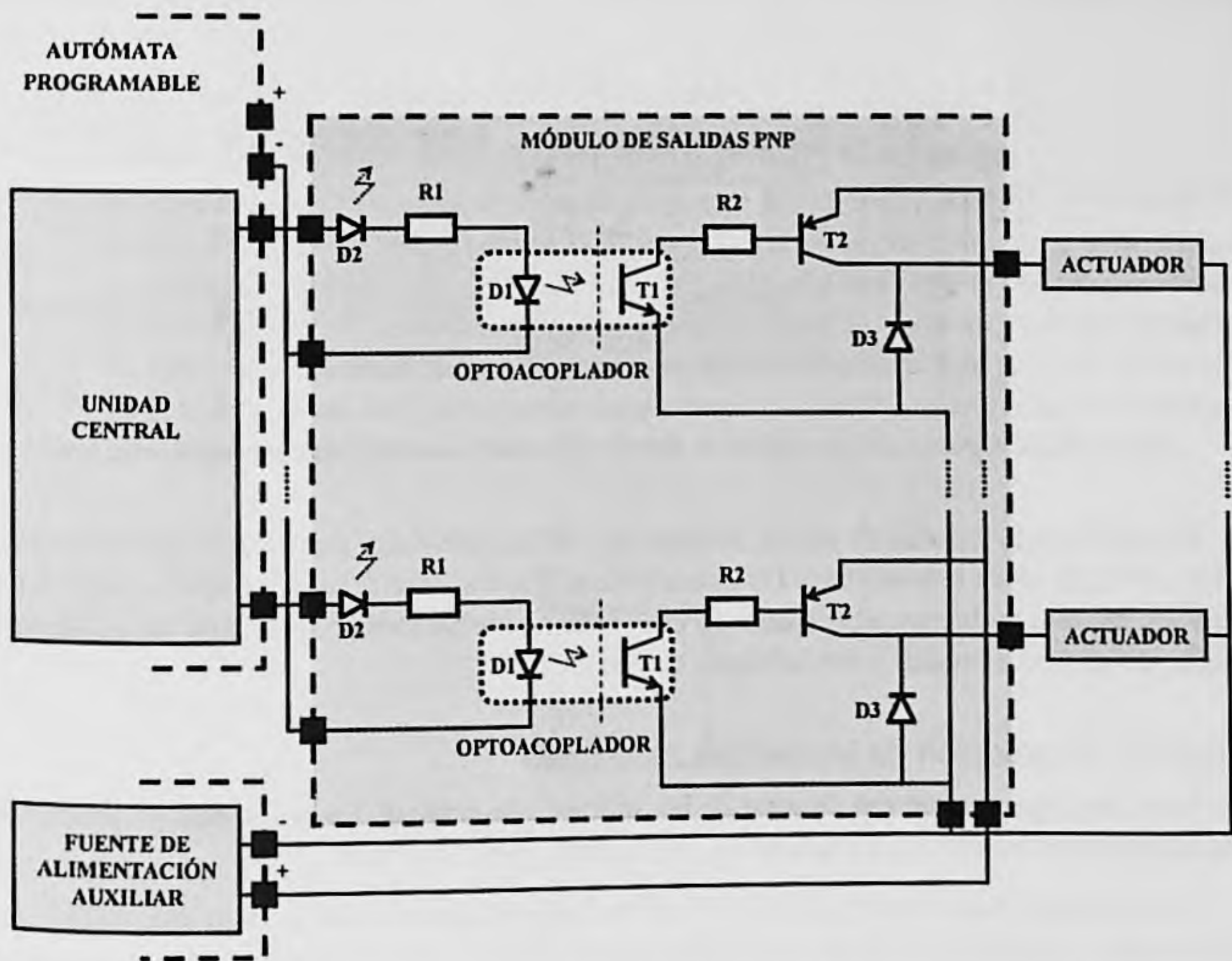


Figura 8.20. Circuito típico de un módulo de interfaz de salida implementado con transistores PNP y diodos de libre circulación.

Tanto en este tipo de salida como en la de tipo transistor NPN, analizada en el apartado anterior, existe una pequeña corriente de fuga cuando el transistor está en corte que no influye en el funcionamiento del actuador.

Interfaces de variables de salida todo-nada con tiristor o triac

En este caso el actuador se alimenta en alterna. La figura 8.21 muestra un circuito típico de variables de salida todo-nada en alterna en el que el dispositivo de conmutación es un triac. El optoacoplador utiliza un optodiac.

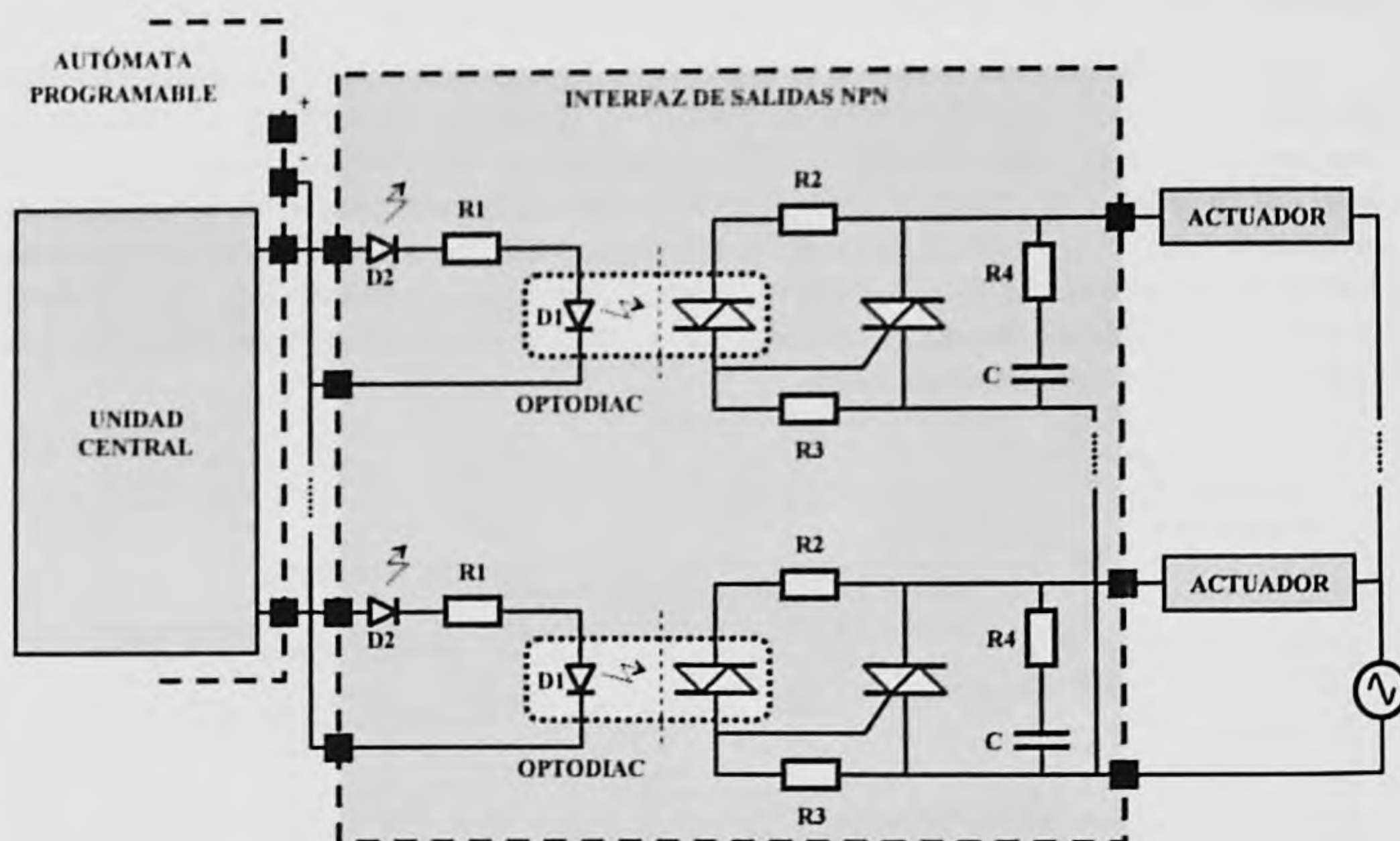


Figura 8.21. Circuito de un módulo de interfaz de salida en alterna implementado con triacs.

En este tipo de interfaz de salida, al igual que en los realizados con relés o con transistores NPN o PNP, el diodo luminoso D2 puede indicar que una determinada variable de salida está activada sin que realmente el actuador lo esté. Ello puede ser debido a un fallo del módulo de salida, del propio actuador o del cableado del mismo.

8.2.2.4.2 Protección de las salidas todo-nada

En la implementación del circuito de las salidas todo-nada de los autómatas programables hay que tener en cuenta las características de la carga conectada a ellas.

Si la carga es inductiva, el circuito de salida debe eliminar la sobretensión que, debido a la ley de Lenz, se produce al cortar la corriente que pasa por ella, para lo cual tiene que ser capaz de absorber la energía acumulada en el circuito magnético de la bobina. Dicha sobretensión no sólo puede provocar la destrucción del dispositivo de conmutación sino que también puede generar señales electromagnéticas que afecten a los circuitos electrónicos próximos. Además, el valor de la sobretensión depende del valor de la autoinducción de la bobina y de la velocidad de cambio de la corriente que circula por ella y, por lo tanto, es tanto mayor cuanto más rápido es el elemento de conmutación.

Si la carga está constituida por una lámpara de incandescencia se debe tener en cuenta que en el instante de activarla, la corriente es mucho más elevada que en régimen permanente.

En apartados sucesivos se analizan diferentes circuitos de protección que se pueden utilizar en cada caso. Los circuitos que se describen son válidos para proteger tanto los contactos de un relé (cuando la salida del autómata programable es de dicho tipo) como los dispositivos elec-

trónicos utilizados en las salidas de tipo semiconductor. Los fabricantes de autómatas programables suelen dar reglas prácticas para ayudar al usuario a seleccionar el circuito de protección más adecuado en función de las características de la carga [SIEM 02].

Salida con carga inductiva

El circuito de protección en el caso de cargas inductivas debe limitar la tensión máxima entre los extremos del dispositivo de conmutación. En la práctica existe un compromiso entre dicho valor máximo y el tiempo durante el cual circula corriente a través de la carga y por ello se pueden utilizar diferentes circuitos de protección que se caracterizan por la forma de resolver el citado compromiso. Además el circuito de protección depende de la alimentación en continua o en alterna de la carga.

Entre los diferentes circuitos de protección cabe citar:

- **Protección mediante diodo rectificador en antiparalelo**

El circuito de protección está constituido por un diodo D en paralelo con la carga (Figura 8.22). Sólo se puede utilizar con cargas alimentadas en continua y el diodo está polarizado inversamente. Al producirse la apertura del contacto, la bobina se descarga a través del diodo D durante un tiempo que teóricamente es infinito y que en la práctica es tanto mayor, cuanto más pequeña es la componente resistiva de la carga. Si la carga es la bobina de un relé, sus contactos permanecen activados mientras circula corriente a través de la misma.

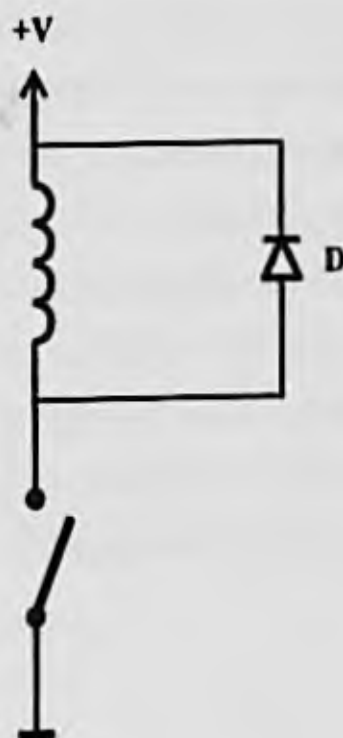


Figura 8.22. Circuito de protección de salidas todo-nada con carga inductiva, implementado con un diodo rectificador.

Este tipo de protección es el más utilizado para cargas en continua y los fabricantes de autómatas programables lo suelen incorporar en sus módulos de variables de salida en continua, como el representado en la figura 8.20 (diodo D_3).

- **Protección mediante diodo Zener**

El circuito de protección está constituido por un diodo zener D_z en paralelo con el elemento de conmutación (Figura 8.23). Este circuito sólo se puede utilizar con cargas

alimentadas en continua. Cuando se produce la apertura del contacto la bobina se descarga a través del diodo zener y de la fuente de alimentación. Con respecto al anterior, presenta la característica de que la corriente de descarga se anula cuando la tensión en bornes del diodo zener disminuye por debajo de la tensión de zener, lo que hace que disminuya el tiempo de recuperación de la carga a costa de alterar el nivel de masa transitoriamente.

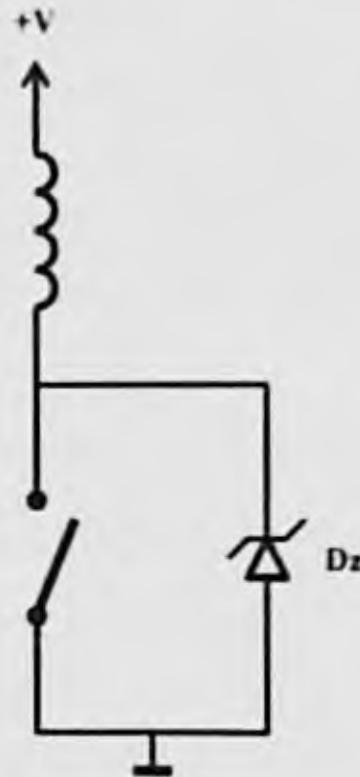


Figura 8.23. Circuito de protección de salidas todo-nada con carga inductiva alimentada en continua, implementado con un diodo zener.

- **Protección mediante diodo rectificador y diodo zener**

Este circuito de protección está constituido por un diodo zener D_z conectado en serie con un diodo D y ambos conectados en paralelo con la carga (Figura 8.24). En este caso, al producirse la apertura del contacto, la bobina se descarga a través del diodo D y del diodo zener D_z . Al igual que en el caso anterior, el diodo zener D_z hace que la anulación de la corriente en la carga se produzca antes. Este circuito sólo se puede utilizar con cargas alimentadas en continua, pero si el diodo D se sustituye por otro diodo zener se puede utilizar tanto con cargas alimentadas en continua como en alterna (Figura 8.25). Los diodos zener se deben elegir de forma que su tensión de zener sea superior al valor de pico de la tensión de alimentación en alterna.

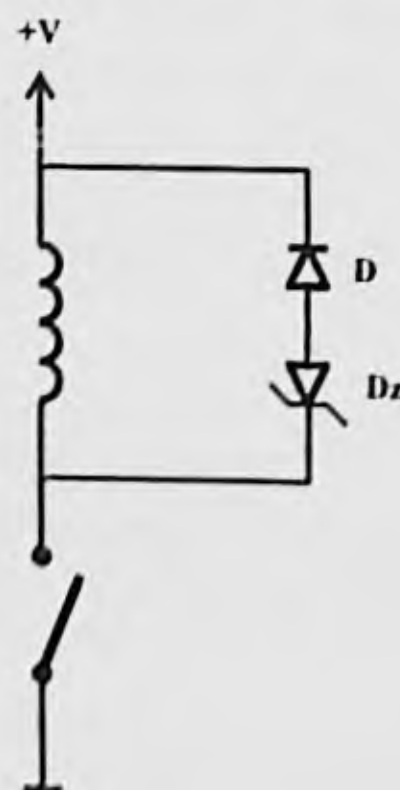


Figura 8.24. Circuito de protección de salidas todo-nada con carga inductiva alimentada en continua, implementado con un diodo rectificador y un diodo zener.

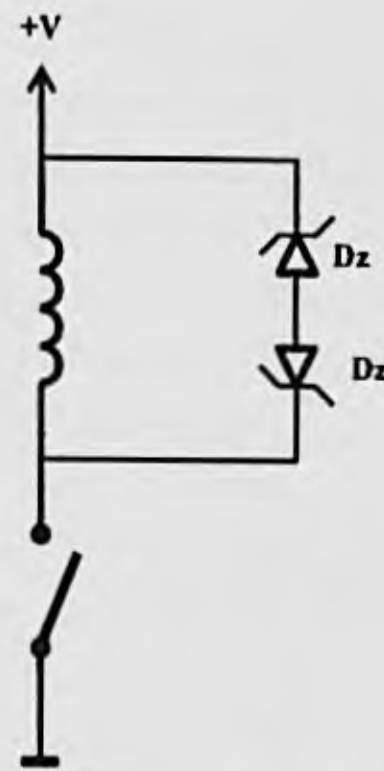


Figura 8.25. Circuito de protección de salidas todo-nada con carga inductiva alimentada en alterna, implementado con un diodo rectificador y dos diodos zener.

- **Protección mediante un circuito R-C**

El circuito de protección consiste en un circuito R-C conectado en paralelo con el elemento de conmutación (Figura 8.26) o con la carga (Figura 8.27) y se puede utilizar con cargas inductivas alimentadas tanto en alterna como en continua. El valor de R puede variar entre 10 y 100 Ω y el de C entre 0,07 y 0,1 μ F y dependen del máximo valor de la tensión que los elementos utilizados pueden soportar. Cuando el circuito R-C se conecta en paralelo con el contacto y la carga se alimenta en continua, puede aparecer un pico de corriente en el instante de la activación del dispositivo de conmutación, debido a la descarga del condensador C. La conexión del circuito R-C en paralelo con el elemento de conmutación, protege a este último contra parásitos en la red de alimentación de alterna. Esta protección está especialmente indicada cuando el elemento de conmutación es un dispositivo electrónico de cuatro capas, como por ejemplo un triac o un tiristor, que se puede disparar debido a variaciones imprevistas de elevado dV/dt de la alimentación alterna. Este tipo de protección lo suelen incorporar los fabricantes de autómatas programables en sus módulos de variables de salida en alterna como el representado en la figura 8.21 (Circuito R4-C).

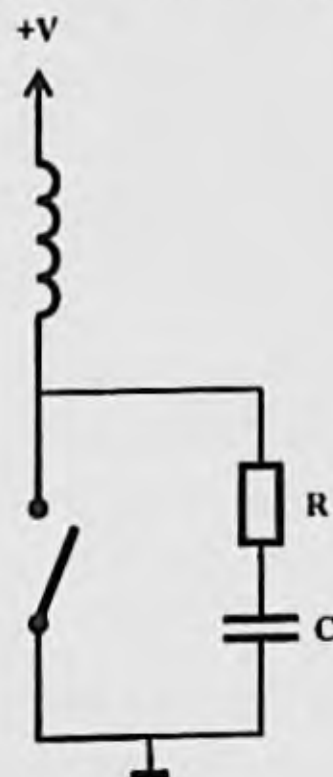


Figura 8.26. Circuito de protección de salidas todo-nada con carga inductiva alimentada en continua, implementado con un circuito R-C conectado en paralelo con el elemento de conmutación.

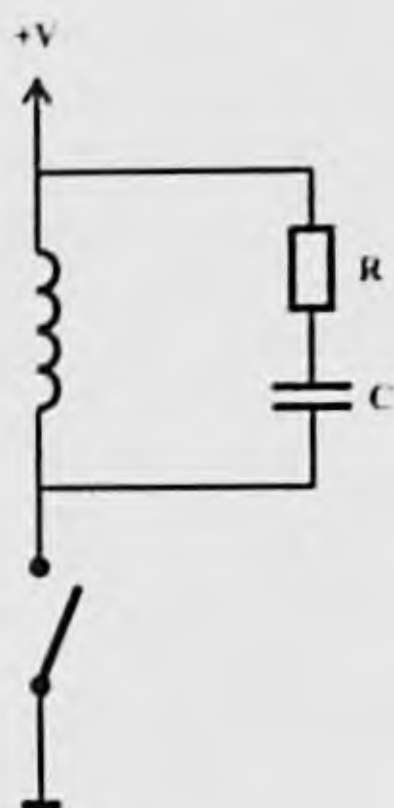


Figura 8.27. *Circuito de protección de salidas todo-nada con carga inductiva alimentada en continua, implementado con un circuito R-C conectado en paralelo con la carga.*

Salida con carga resistiva del tipo lámpara de incandescencia

Este tipo de cargas se caracteriza por tener un valor resistivo menor cuando no circula corriente a través de ella (está fría), lo cual provoca una corriente más elevada a través del elemento de conmutación en el instante en el que se cierra. El usuario debe conocer las características de la carga y del módulo de variables de salida del autómata programable y comprobar la compatibilidad entre ambas.

8.2.2.5 Interfaces de variables analógicas de entrada

Muchos sensores suministran variables analógicas de tensión o de corriente, cuyo valor puede variar dentro de un cierto rango. Por ejemplo, el nivel de líquido de un tanque se puede medir mediante un sensor analógico de nivel, cuya tensión de salida varíe entre 0 y 10 V. Para ello debe proporcionar 5V a su salida si el nivel del líquido contenido en el tanque alcanza el 50% y 7,5 V si alcanza el 75%.

Los fabricantes de autómatas programables comercializan interfaces de este tipo, con diferentes niveles de resolución y rangos de tensión y de corriente. Los rangos de tensión más habituales son de 0 a 5V, de 1 a 5V, de 0 a 10V, de -10V a +10V, de -1 a +1V, etc., y los de corriente son de 0 a 20mA, de -20mA a +20mA y de 4 a 20mA. A veces la señal que proporciona un sensor en el rango de corriente de 4 a 20mA se transforma en una tensión variable de 1 a 5V mediante una resistencia de 250Ω.

En la figura 8.28 se muestra el diagrama de bloques básico típico de una interfaz de variables de entrada analógicas que constituye un sistema electrónico de adquisición de datos [PERE 03]. El autómata programable selecciona a través de una unidad de control y mediante un multiplexor analógico [MAND 08a] las distintas variables analógicas. El circuito de muestreo y retención (*Sample & Hold*) mantiene estable el valor de la señal analógica mientras se realiza la conversión (del orden de los microsegundos). El convertidor analógico/digital (A/D) transforma la variable analógica en un conjunto de “n” bits que se introduce en la memoria del autómata programable. Cuanto mayor es el valor de “n”, mayor es la precisión con la que se

mide el valor de la variable analógica. Los fabricantes de autómatas programables, como por ejemplo Siemens, suministran distintas interfaces de este tipo en las que “n” varía entre 12 y 16 bits. Por ejemplo, si el interfaz analógico de entrada utiliza 12 bits (bit 0 a bit 11) de información en binario natural, además del bit de signo y otros bits de control, y al convertidor se le aplica una variable analógica que varía entre -1V y $+1\text{V}$, la resolución es de $2/4096 = 0,58 \text{ mV/bit}$. Esto supone que por cada $0,58 \text{ mV}$ que aumenta o disminuye el nivel de la señal analógica, la combinación digital equivalente se incrementa o decrementa respectivamente en una unidad.

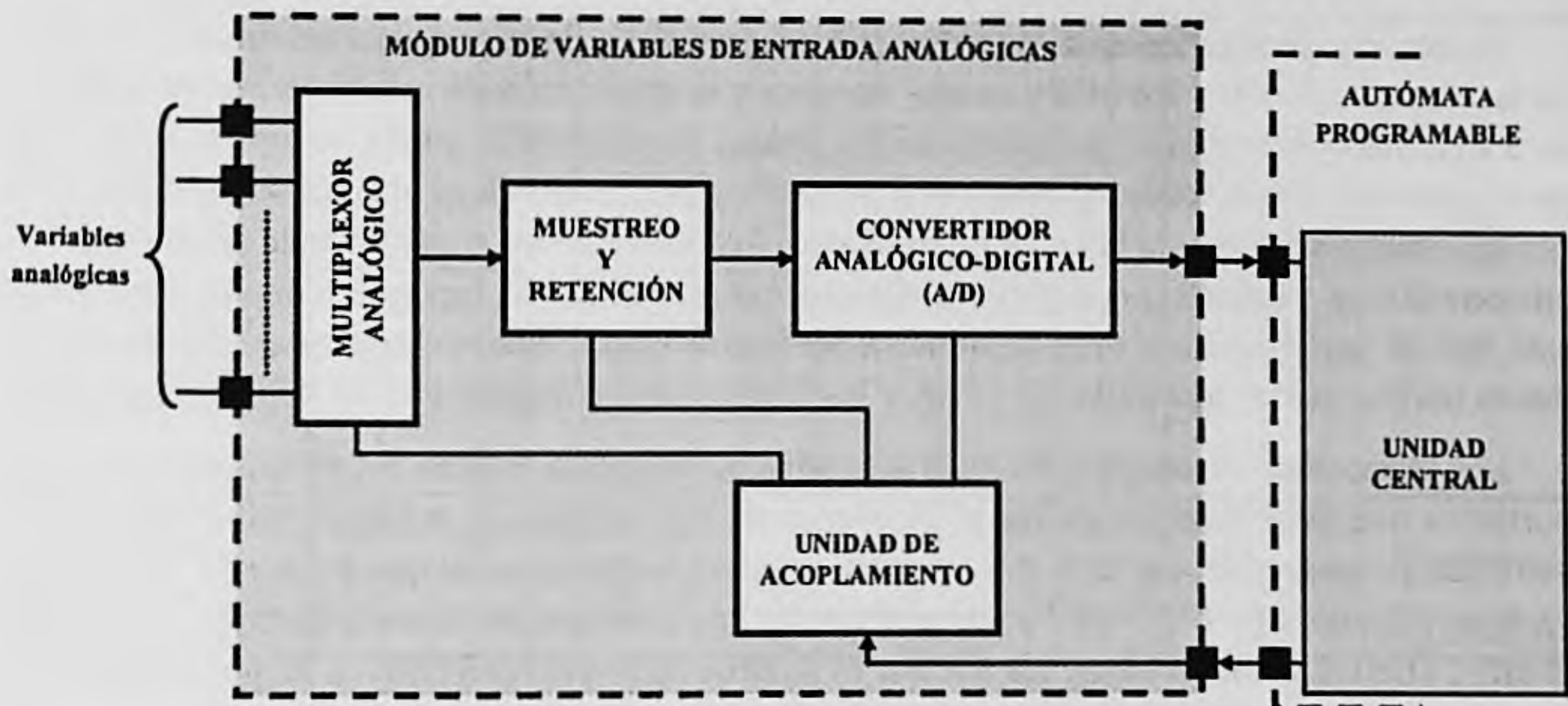


Figura 8.28. Diagrama de bloques típico de una interfaz de variables de entrada analógicas de un autómata programable.

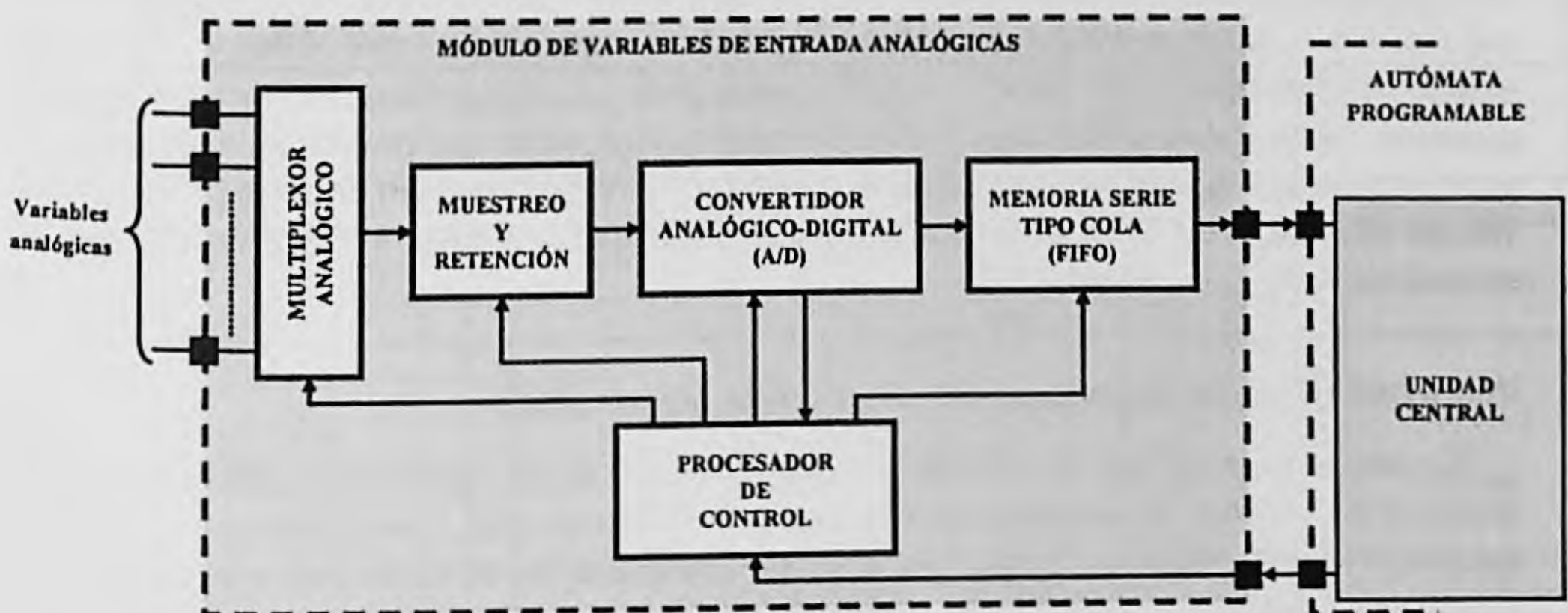


Figura 8.29. Diagrama de bloques de una unidad de variables analógicas de entrada que posee un procesador digital de control y una memoria serie del tipo cola (FIFO).

En la actualidad algunos fabricantes comercializan interfaces de variables analógicas de entrada que incorporan un procesador digital y una memoria asociada que suele ser de acceso serie del tipo cola (FIFO) (Figura 8.29) [MAND 08b]. De esta forma durante la fase de ejecución del

programa del autómata programable, el procesador digital escribe en la memoria serie las combinaciones binarias correspondientes a todos los sensores y a continuación el autómata programable la lee y transfiere su contenido a la memoria de imágenes de las variables de entrada.

A continuación se analizan las variables analógicas de entrada, tanto de tensión como de corriente.

Interfaces de variables analógicas de entrada por tensión

En este tipo de interfaces de los autómatas programables, la impedancia de entrada varía entre las decenas de $K\Omega$ y los $M\Omega$ y es muy superior a la impedancia de salida del sensor conectado a ella, que debe ser una fuente de tensión con una impedancia pequeña en serie (Figura 8.3), para garantizar que la totalidad de la señal generada por el sensor llega al autómata programable por ser prácticamente nula la caída de tensión en los hilos que los conectan cuando la distancia entre ambos es pequeña (no superior a algunos metros). Además, hay que tener en cuenta que este tipo de interfaces son muy sensibles a las interferencias electromagnéticas, por lo que se deben utilizar cables apantallados y seguir las instrucciones proporcionadas por el fabricante.

Los fabricantes suelen proporcionar una tabla en la que se indican los parámetros más importantes que debe tener en cuenta el diseñador de una instalación automatizada mediante un autómata programable. La tabla 8.5 muestra un ejemplo típico, en el que la variable analógica varía en el rango de -10 a $+10V$ y se convierte en una combinación binaria de 16 bits en la que el bit 15 (BS) es el bit de signo, los bits 3 a 14 (ambos inclusive) constituyen la medida en binario natural (en complemento a dos cuando el número es negativo), el bit 0 (D) se activa cuando se produce un desbordamiento (*Overflow*) (en este caso variable analógica $>$ de $+20V$ o menor de $-20V$), el bit 1 (E) se activa si se produce un error y el bit 2 no se utiliza. Como se puede ver en la tabla 8.5, cuando la señal de entrada varía entre -10 y $+10V$, el autómata programable recibe una señal en formato digital que varía entre 1000000000000000 y 0100000000000000, que a su vez se corresponde con los valores decimales comprendidos entre -2048 y $+2048$. Los valores comprendidos entre -4095 y -2049 , y entre 2049 y 4095 constituyen el margen de rebasamiento, en el cual se realiza una medida correcta con un sensor que proporcione una variable analógica cuyo rango esté comprendido entre -20 y $+20V$. Si la tensión de entrada supera los valores citados, el bit 1 (E) se pone a uno para indicar que la medida es incorrecta, lo que se suele denominar desbordamiento.

Interfaces de variables analógicas de entrada por corriente

La impedancia de entrada de este tipo de interfaces de un autómata programable es del orden de las decenas de ohmios (un valor típico es 30Ω) y es muy inferior a la impedancia de salida del sensor conectado a ella, que debe ser una fuente de corriente con una impedancia elevada en paralelo (Figura 8.4) para garantizar que el máximo de señal generada por el sensor llega al autómata programable.

Las variables analógicas de corriente, también denominadas bucle de corriente, presentan una inmunidad a las interferencias electromagnéticas muy superior a las de tensión y pueden transmitirse a mayores distancias.

VARIABLE DE ENTRADA ANALÓGICA DE Tensión EN FORMATO DIGITAL																		
Valor Decimal	Entrada (V)	15	14	13	12	11	10	9	8	8	6	5	5	3	2	1	0	Nº de Bit
		BS	2 ¹¹ 2048	2 ¹⁰ 1024	2 ⁹ 512	2 ⁸ 256	2 ⁷ 128	2 ⁶ 64	2 ⁵ 32	2 ⁴ 16	2 ³ 8	2 ² 4	2 ¹ 2	2 ⁰ 1		E	D	Significado
≥4096	20,000	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	Desbordamiento
4095	19,995	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	Rebasamiento
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
2049	10,0048	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	Rango De Operación
2048	10,000	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
2047	9,995	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
3	0,0156	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	
2	0,0097	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	
1	0,0058	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-1	-0,0058	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
-2	-0,0097	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
-3	-0,0156	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	
-5	-0,0195	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
-2048	-10,000	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-2049	-10,0048	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	Rebasamiento
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
-4095	-19,995	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	Desbordamiento
≤-4096	-20,000	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	

Tabla 8.5. Ejemplo de tabla en la que se indican las variables de salida de una interfaz analógica de entrada mediante tensión de -10 a +10V.

VARIABLE DE ENTRADA ANALÓGICA DE CORRIENTE EN FORMATO DIGITAL																		
Valor Decimal	Entrada (mA)	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	Nº de Bit
		BS	2 ¹¹ 2048	2 ¹⁰ 1024	2 ⁹ 512	2 ⁸ 256	2 ⁷ 128	2 ⁶ 64	2 ⁵ 32	2 ⁴ 16	2 ³ 8	2 ² 4	2 ¹ 2	2 ⁰ 1		E	D	Significado
≥4096	≥32,769	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	Desbordamiento
4095	31,992	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	Rebasamiento
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
2561	20,008	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	Rango De Operación
2560	20,000	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
2048	16	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
2047	15,992	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
512	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Rebasamiento
511	3,992	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	
384	3,0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	Rotura de hilo
383	2,992	0	0	0	0	1	0	1	1	1	1	1	1	1	0	0	0	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
3	0,0235	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	
2	0,0156	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	
1	0,0078	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Tabla 8.6. Ejemplo de tabla en la que se indican las variables de salida de una interfaz analógica de entrada mediante corriente de 4 a 20mA.

Los fabricantes también suelen proporcionar una tabla en la que se indican los parámetros más importantes que debe tener en cuenta el diseñador de una instalación automatizada mediante un autómata programable que utiliza este tipo de interfaz. Un ejemplo es la tabla 8.6 que representa una variable analógica ampliamente utilizada en la industria, como es la que varía entre 4 y 20mA. En este caso se tiene una resolución de $(20-4)/2048 = 0,0078\text{mA/bit}$. El rango comienza en 4mA, que corresponde a un valor decimal de 512 y finaliza en 20mA, que corresponde a un valor decimal de 2560. Cuando la corriente de entrada es inferior a 3 mA se considera que está roto el cable de conexión.

8.2.2.6 Interfaces de variables analógicas de salida

Los autómatas programables son procesadores digitales que proporcionan a su salida combinaciones binarias en un determinado código, pero a la mayoría de los actuadores se les deben aplicar señales analógicas, lo que hace que sea necesaria una interfaz que realice la conversión de un tipo de señal al otro, para lo cual debe incorporar un convertidor digital-analógico (D/A), tal como se indica en el diagrama de bloques de la figura 8.30. El circuito está formado por un convertidor D/A y una etapa de salida que puede proporcionar tensión o corriente, para lo cual debe tener la impedancia de salida adecuada, según lo indicado en el apartado 8.2.2.1.1.

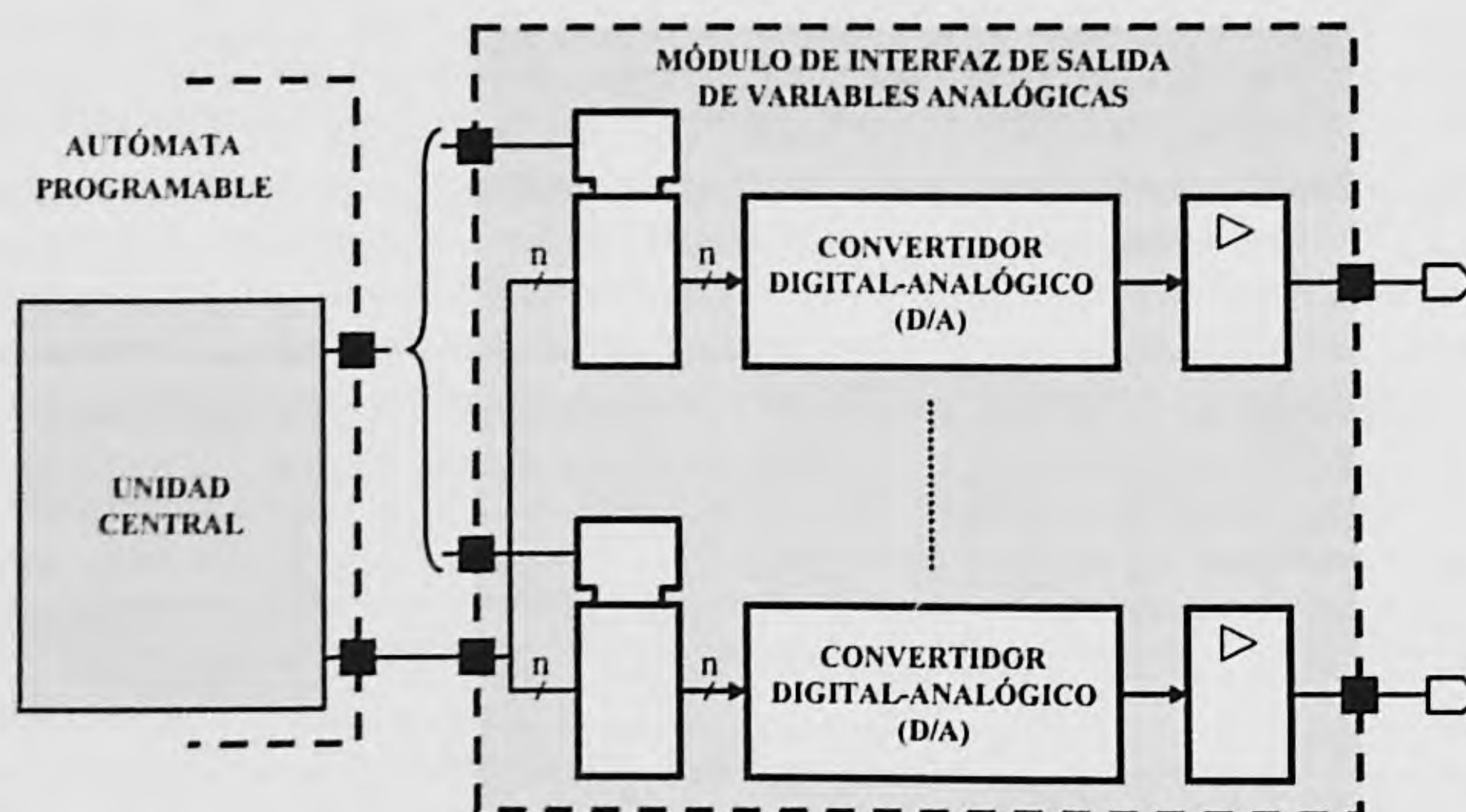


Figura 8.30. Diagrama de bloques típico de un módulo de variables de salida analógicas de un autómata programable.

El desarrollo de la Microelectrónica ha propiciado la comercialización de interfaces de variables analógicas de salida como la representada en la figura 8.31, que contiene un procesador digital, una memoria serie de tipo circular o una memoria de acceso aleatorio de doble puerto [MAND 08c], un único convertidor digital-analógico y un circuito de muestreo y retención en cada salida analógica. El autómata programable introduce en la memoria la información correspondiente a cada salida analógica y el procesador digital realiza cíclicamente una lectura de la memoria y la apertura secuencial de los circuitos de muestreo y retención, para mantener estable la tensión a su salida.

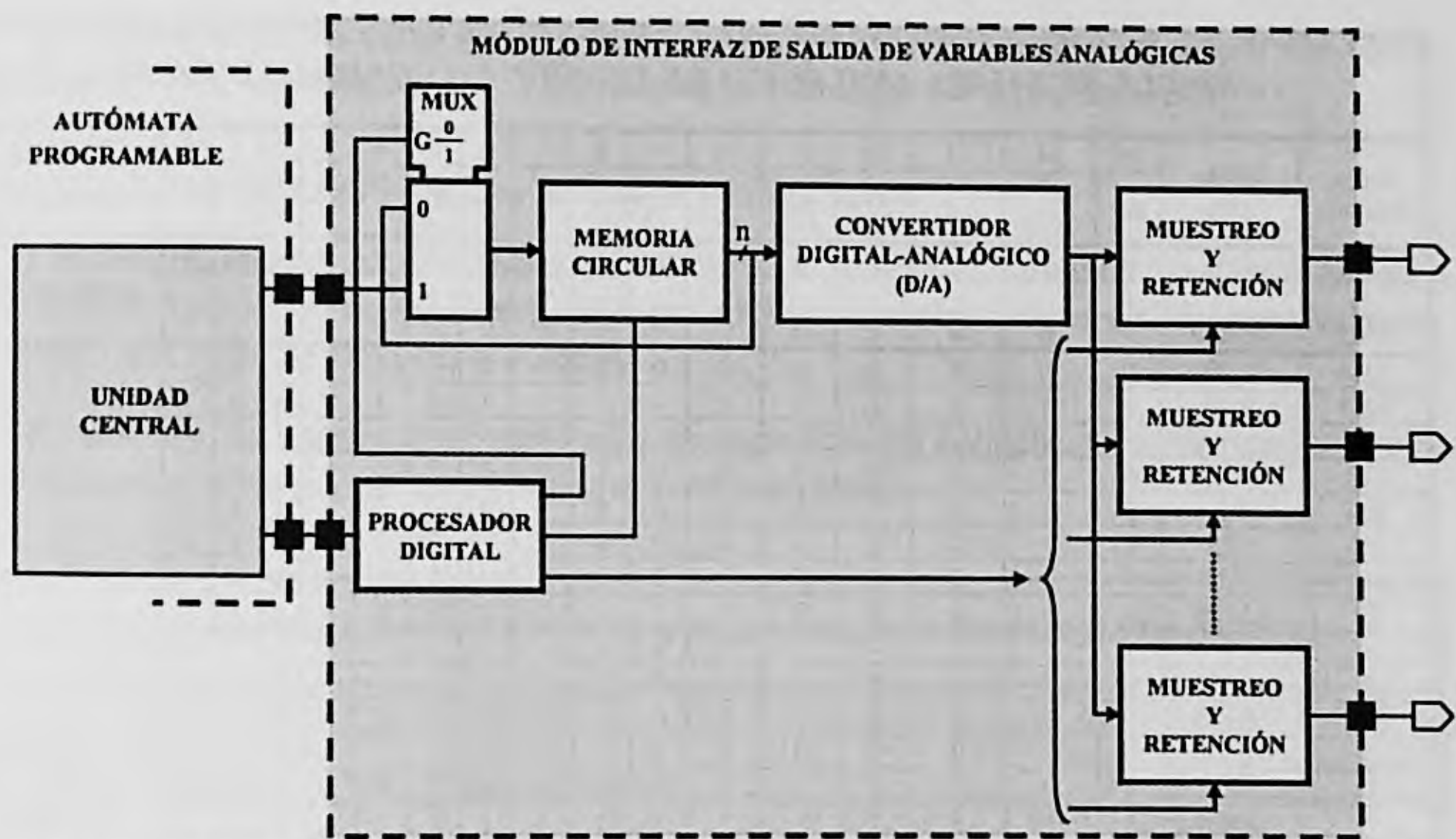


Figura 8.31. Diagrama de bloques típico de un módulo de variables de salida analógicas implementado con una memoria circular y un único convertidor analógico/digital.

En este tipo de interfaces es frecuente utilizar 10 bits de información ($2^{10} = 1024$ combinaciones) más bit de signo. Si la variable de salida es bipolar y se representa en el convenio del complemento a dos, la combinación correspondiente varía entre 0 y 1023 y entre -1 y -1024, y si es unipolar, varía entre 0 y 1023.

A continuación se analizan las variables analógicas de salida, tanto de tensión como de corriente.

Interfaz de variables analógicas de salida por tensión

Los diferentes fabricantes de autómatas programables comercializan módulos analógicos de salida que se diferencian por el rango de la tensión analógica que generan. Dos de los rangos más utilizados son los comprendidos entre -10 y +10V, y entre 1 y 5V. En la tabla 8.7 se muestran las combinaciones de una palabra de 16 bits que representa en binario natural en el convenio del complemento a dos, una variable analógica de salida que varía en el rango de -10 a +10 V. Se utilizan 11 bits de información y bit de signo (bit 4 al bit 15). Los bits 0 a 3 no se utilizan. La variación de 0 a +10V se representa mediante la combinación binaria 00000000000 a 01111111111, que corresponde a los valores decimales de 0 a 1023. De igual forma, los valores negativos, de -1 a -10V, se representan en complemento a dos mediante las combinaciones 11111111111 a 10000000000, que corresponden a los valores decimales -1 a -1024 respectivamente.

Si la salida es por tensión, la corriente que pasa a través de los hilos que conectan la interfaz con la carga puede producir un error debido a la caída de tensión en los mismos. Para evitarlo se suele dotar de entradas de alta impedancia de medida de la tensión en la carga (*Sense inputs*) $S+$ y $S-$, y el correspondiente circuito de compensación interno, a la interfaz de salida analógica de tensión (Figura 8.32). De esta forma el autómata programable ajusta el nivel de la señal analógica de salida para que en la carga se aplique el nivel de tensión adecuado en cada instante.

VARIABLE DE SALIDA ANALÓGICA DE TENSIÓN EN FORMATO DIGITAL																		
Valor Decimal	Salida (V)	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	Nº de Bit
		BS	2 ¹⁰ 1024	2 ⁹ 512	2 ⁸ 256	2 ⁷ 128	2 ⁶ 64	2 ⁵ 32	2 ⁴ 16	2 ³ 8	2 ² 4	2 ¹ 2	2 ⁰ 1	x	x	x	x	Significado
1280	12,5	0	1	0	1	0	0	0	0	0	0	0	0	x	x	x	x	Rebasamiento
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
1025	10,0098	0	1	0	0	0	0	0	0	0	0	0	0	x	x	x	x	
1024	10,000	0	1	0	0	0	0	0	0	0	0	0	0	x	x	x	x	Rango De Operación
1023	9,99	0	0	1	1	1	1	1	1	1	1	1	1	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
512	5	0	0	1	0	0	0	0	0	0	0	0	0	x	x	x	x	
256	2,5	0	0	0	1	0	0	0	0	0	0	0	0	x	x	x	x	
128	1,25	0	0	0	0	1	0	0	0	0	0	0	0	x	x	x	x	
64	0,625	0	0	0	0	0	1	0	0	0	0	0	0	x	x	x	x	
32	0,3125	0	0	0	0	0	0	1	0	0	0	0	0	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
3	0,0292	0	0	0	0	0	0	0	0	0	0	1	1	x	x	x	x	
2	0,0195	0	0	0	0	0	0	0	0	0	0	1	0	x	x	x	x	
1	0,0098	0	0	0	0	0	0	0	0	0	0	0	1	x	x	x	x	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	x	x	x	x	
-1	-0,0098	1	1	1	1	1	1	1	1	1	1	1	1	x	x	x	x	
-2	-0,0195	1	1	1	1	1	1	1	1	1	1	1	0	x	x	x	x	
-3	-0,0292	1	1	1	1	1	1	1	1	1	1	0	1	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
-32	-0,3125	1	1	1	1	1	1	1	0	0	0	0	0	x	x	x	x	
-65	-0,625	1	1	1	1	1	1	0	0	0	0	0	0	x	x	x	x	
-128	-1,25	1	1	1	1	1	0	0	0	0	0	0	0	x	x	x	x	
-256	-2,5	1	1	1	1	0	0	0	0	0	0	0	0	x	x	x	x	
-512	-5	1	1	1	0	0	0	0	0	0	0	0	0	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
-1023	-9,99	1	1	0	0	0	0	0	0	0	0	1	x	x	x	x	x	
-1024	-10	1	1	0	0	0	0	0	0	0	0	0	x	x	x	x	x	
-1025	-10,0098	1	0	1	1	1	1	1	1	1	1	1	x	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Rebasamiento
-1280	-12,5	1	0	1	1	0	0	0	0	0	0	0	x	x	x	x	x	

Tabla 8.7. Ejemplo de tabla en la que se indican las combinaciones binarias correspondientes a los valores de una variable de salida analógica por tensión de -10 a +10V de un autómata programable.

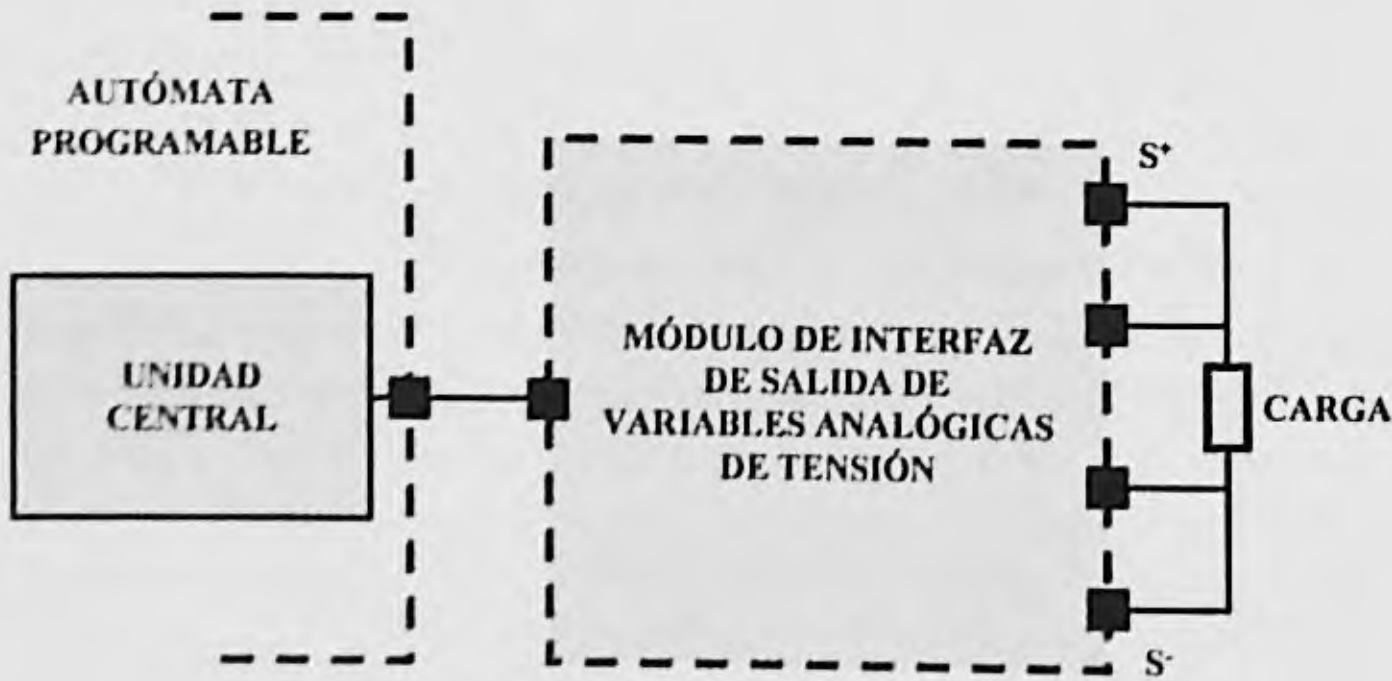


Figura 8.32. Diagrama de bloques típico de un módulo de variables de salida analógicas de tensión.

La impedancia de la carga que se conecta a este tipo de interfaz debe ser mayor que un determinado valor, especificado por el fabricante, que suele ser del orden de $1K\Omega$.

Interfaz de variables analógicas de salida por corriente

Las señales analógicas de salida por corriente más habituales varían en los rangos de -20 a $+20mA$, de 0 a $20mA$ y de 4 a $20mA$. En la tabla 8.8 se indica los valores analógicos correspondientes a una señal de salida de 4 a $20mA$.

En este tipo de interfaces la impedancia de carga debe ser pequeña, y los fabricantes suelen especificar un valor máximo, en general del orden de 500Ω .

VARIABLE DE SALIDA ANALÓGICA DE CORRIENTE EN FORMATO DIGITAL																		
Valor Decimal	Salida (mA)	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	Nº de Bit
		BS	2 ¹⁰ 1024	2 ⁹ 512	2 ⁸ 256	2 ⁷ 128	2 ⁶ 64	2 ⁵ 32	2 ⁴ 16	2 ³ 8	2 ² 4	2 ¹ 2	2 ⁰ 1	x	x	x	x	Significado
1280	24	0	1	0	1	0	0	0	0	0	0	0	0	x	x	x	x	Rebasamiento
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
1025	20,016	0	1	0	0	0	0	0	0	0	0	0	0	x	x	x	x	
1024	20	0	1	0	0	0	0	0	0	0	0	0	0	x	x	x	x	Rango De Operación
1023	19,98	0	0	1	1	1	1	1	1	1	1	1	1	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
512	12,0	0	0	1	0	0	0	0	0	0	0	0	0	x	x	x	x	
256	8	0	0	0	1	0	0	0	0	0	0	0	0	x	x	x	x	
128	6	0	0	0	0	1	0	0	0	0	0	0	0	x	x	x	x	
64	5	0	0	0	0	0	1	0	0	0	0	0	0	x	x	x	x	
32	4,5	0	0	0	0	0	0	1	0	0	0	0	0	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
3	4,0469	0	0	0	0	0	0	0	0	0	0	1	1	x	x	x	x	
2	4,0312	0	0	0	0	0	0	0	0	0	0	1	0	x	x	x	x	
1	4,016	0	0	0	0	0	0	0	0	0	0	0	1	x	x	x	x	
0	4	0	0	0	0	0	0	0	0	0	0	0	0	x	x	x	x	
-1	3,984	1	1	1	1	1	1	1	1	1	1	1	1	x	x	x	x	Rebasamiento
-2	3,9687	1	1	1	1	1	1	1	1	1	1	1	0	x	x	x	x	
-3	3,9531	1	1	1	1	1	1	1	1	1	1	0	1	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
-32	3,5	1	1	1	1	1	1	1	0	0	0	0	0	x	x	x	x	
-64	3	1	1	1	1	1	1	0	0	0	0	0	0	x	x	x	x	
-128	2	1	1	1	1	1	0	0	0	0	0	0	0	x	x	x	x	
-256	0	1	1	1	1	0	0	0	0	0	0	0	0	x	x	x	x	
-512	-4	1	1	1	0	0	0	0	0	0	0	0	0	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
-1023	-11,984	1	1	0	0	0	0	0	0	0	0	1	x	x	x	x	x	
-1024	-12	1	1	0	0	0	0	0	0	0	0	0	x	x	x	x	x	
-1025	-12,0156	1	0	1	1	1	1	1	1	1	1	1	x	x	x	x	x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
-1280	-16	1	0	1	1	0	0	0	0	0	0	0	0	x	x	x	x	

Tabla 8.8. Ejemplo de tabla en la que se indican las combinaciones binarias correspondientes a los valores de una variable de salida analógica por corriente de 4 a $20mA$ de un autómata programable.

8.2.3 Interfaces de conexión con el proceso de aplicación específica

Las interfaces de conexión con el proceso de aplicación específica, también denominadas módulos específicos de entrada/salida, son, tal como se indica en el capítulo 1, sistemas electrónicos comercializados por un fabricante de autómatas programables como elementos acoplables a los mismos para realizar un determinado tipo de proceso de forma más eficiente (en general en menor tiempo) que si se realiza por programa (*Software*), es decir, mediante la ejecución de una secuencia de instrucciones. Los tipos de módulos dependen de la capacidad de entrada/salida del autómata programable específico y varían de un fabricante a otro.

El desarrollo de la Microelectrónica ha hecho que cada vez sea mayor el número de aplicaciones de los autómatas programables, que exigen la medida de variables específicas, como por ejemplo las señales generadas por un codificador incremental, que no tienen un formato de tensión o corriente normalizados. De igual forma ocurre con las salidas, que en múltiples ocasiones deben generar las señales de control de ciertos tipos de actuadores, como por ejemplo los motores paso a paso. En apartados sucesivos se analizan algunos de estos tipos de interfaces.

8.2.3.1 Unidades de entrada de medida de temperatura

La temperatura es, junto con el caudal, una de las variables físicas más medidas en la industria. Además, según el rango de temperatura que se quiera medir, se utilizan distintos tipos de sensores (termopares, termorresistencias, de infrarrojos, termistores, etc.) que necesitan circuitos de acondicionamiento diferentes para convertir la señal que proporcionan en una señal analógica normalizada convertible, a su vez, en digital [PALL 98] [PERE 03]. Por ello, los fabricantes de autómatas programables comercializan, entre otros, módulos de interfaz para medir temperatura mediante termorresistencias Pt100, que se caracterizan porque su resistencia a 0° es $R_0 = 100\Omega$ y su variación con la temperatura θ se expresa mediante la ecuación $R_{Pt} = R_0 (1 + \alpha\theta)$, en la cual α es un coeficiente aproximadamente constante y θ es la temperatura en °C.

Los módulos de interfaz para acoplar termorresistencias Pt100 realizan las siguientes funciones:

- Utilizan como circuito básico de acondicionamiento un divisor de tensión o un puente de medida alimentado con tensión o con corriente constante.
- Amplifican la señal del acondicionador mediante un amplificador de instrumentación [PALL 98] [PERE 03].
- Si la resistencia está situada lejos del circuito acondicionador situado en el módulo, la conectan mediante tres o cuatro hilos en lugar de dos. En la figura 8.33 se representa el esquema de bloques de un módulo de medida de temperatura en el que la resistencia se conecta al puente de medida mediante tres hilos, de forma que la caída de tensión en el hilo 1 se compensa con la caída de tensión en el hilo 3, y el error en la medida sólo depende de la caída de tensión en el hilo 2 que es despreciable porque la señal V_i de salida del puente se amplifica mediante un amplificador de instrumentación AI que tiene una elevada impedancia de entrada.

Otras termorresistencias de uso habitual son la Pt200, la Pt500 y la Pt1000.

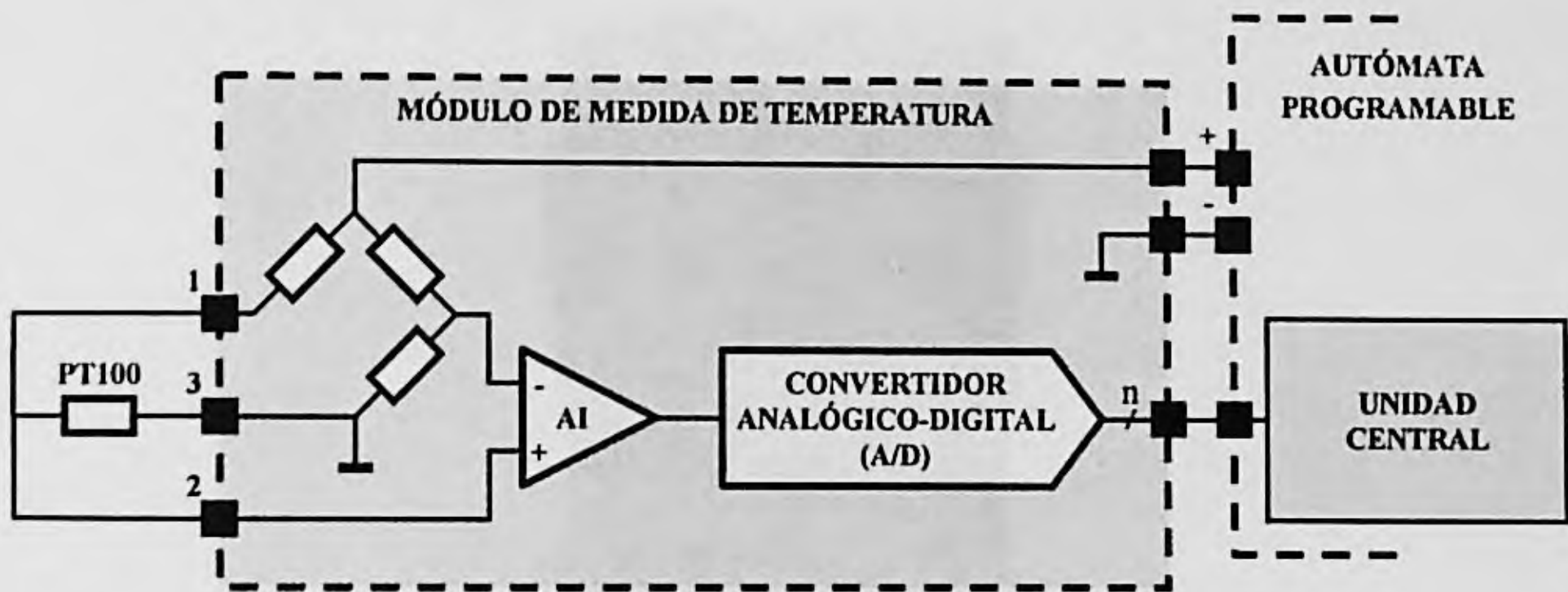


Figura 8.33. Módulo de medida de temperatura.

Existen otros sensores, ampliamente utilizados en aplicaciones industriales, que también necesitan un circuito de acondicionamiento, como por ejemplo los termopares que generan una pequeña tensión (efecto Seebeck), del orden de algunos microvoltios por °C, denominada fuerza termoelectromotriz. Algunos fabricantes de autómatas programables comercializan módulos de interfaz para este tipo de sensores, que acondicionan la señal mediante un amplificador de instrumentación AI, que también filtra la señal. Además tienen un circuito para la compensación de la unión fría, realizado con una resistencia R_T y su circuito acondicionador, un multiplexor analógico, y un convertidor analógico/digital. El control de los distintos elementos lo lleva a cabo un microcontrolador que proporciona al autómata programable la combinación binaria correspondiente a la temperatura medida (Figura 8.34).

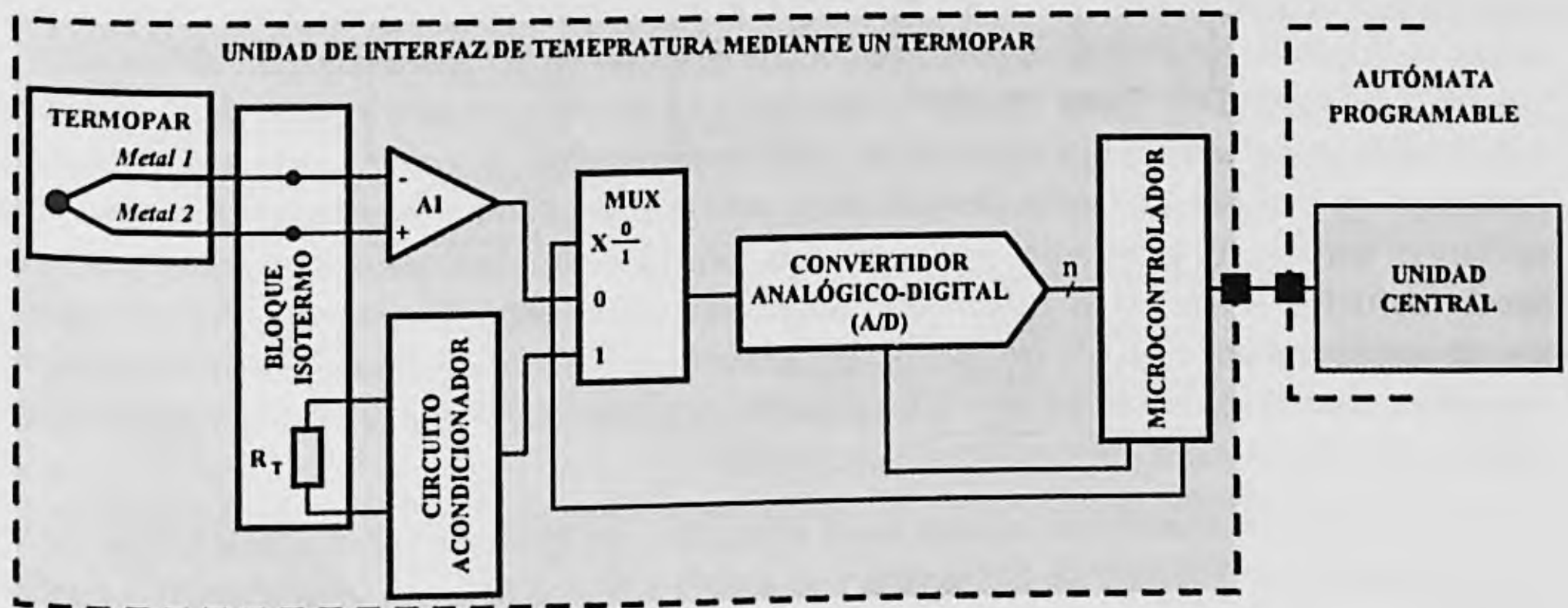


Figura 8.34. Diagrama de bloques del circuito de interfaz de un termopar.

Un ejemplo de módulo de interfaz para sensores de temperatura es el SM 331 de Siemens (Figura 8.35), que, además de entradas de tensión y corriente normalizadas para conectar diversos tipos de sensores, dispone de los circuitos de acondicionamiento adecuados para la conexión de termorresistencias Pt100 y Ni 100 y de termopares (tipos E, N, J y K).



Figura 8.35. Módulo de entradas analógicas SM 331 para el acoplamiento de termopares y termorresistencias a un autómata programable de la familia S7-300 de Siemens.

8.2.3.2 Unidades de entrada de conteo

Este tipo de módulo específico de entrada/salida se utiliza, como su nombre indica, para realizar el conteo de los impulsos generados por determinados sensores, como por ejemplo los codificadores angulares de tipo incremental, que, por su frecuencia, no pueden ser contados mediante la ejecución de las instrucciones de conteo descritas en los apartados 2.3.7 del capítulo 2 y 3.3.8.5 del capítulo 3. Estos impulsos, frecuentes en las aplicaciones industriales, no pueden ser contados por el autómata a través de un módulo de variables de entrada todo-nada porque su frecuencia es lo suficientemente elevada como para que el autómata programable no sea capaz de realizar el conteo por programa (*software*) de todos los impulsos generados por el sensor en sucesivos ciclos de operación. En este tipo de aplicaciones se debe utilizar un módulo de conteo que puede contar impulsos de frecuencia del orden de KHz. Periódicamente, la unidad central consulta el contenido del contador o bien el propio contador envía una señal a la unidad central cuando alcanza un valor previamente seleccionado.

Un ejemplo de módulo de conteo es el FM 350-1 de Siemens (Figura 8.36), que puede contar impulsos de 500KHz de frecuencia y se acopla a un autómata programable de la familia S7-300.

8.2.3.3 Unidades de entrada/salida remotas

En muchas aplicaciones, los sensores que generan las variables de entrada analógicas o todo-nada y los actuadores están situados a una cierta distancia del autómata programable y su conexión al mismo, a través de los módulos de entrada/salida analizados en el apartado 8.2, hace que sea necesario utilizar un cable distinto para cada uno de ellos, lo cual eleva el coste del cableado, así como la influencia de las interferencias electromagnéticas.



Figura 8.36. Módulo de contaje FM 350-1 (cortesía de Siemens).

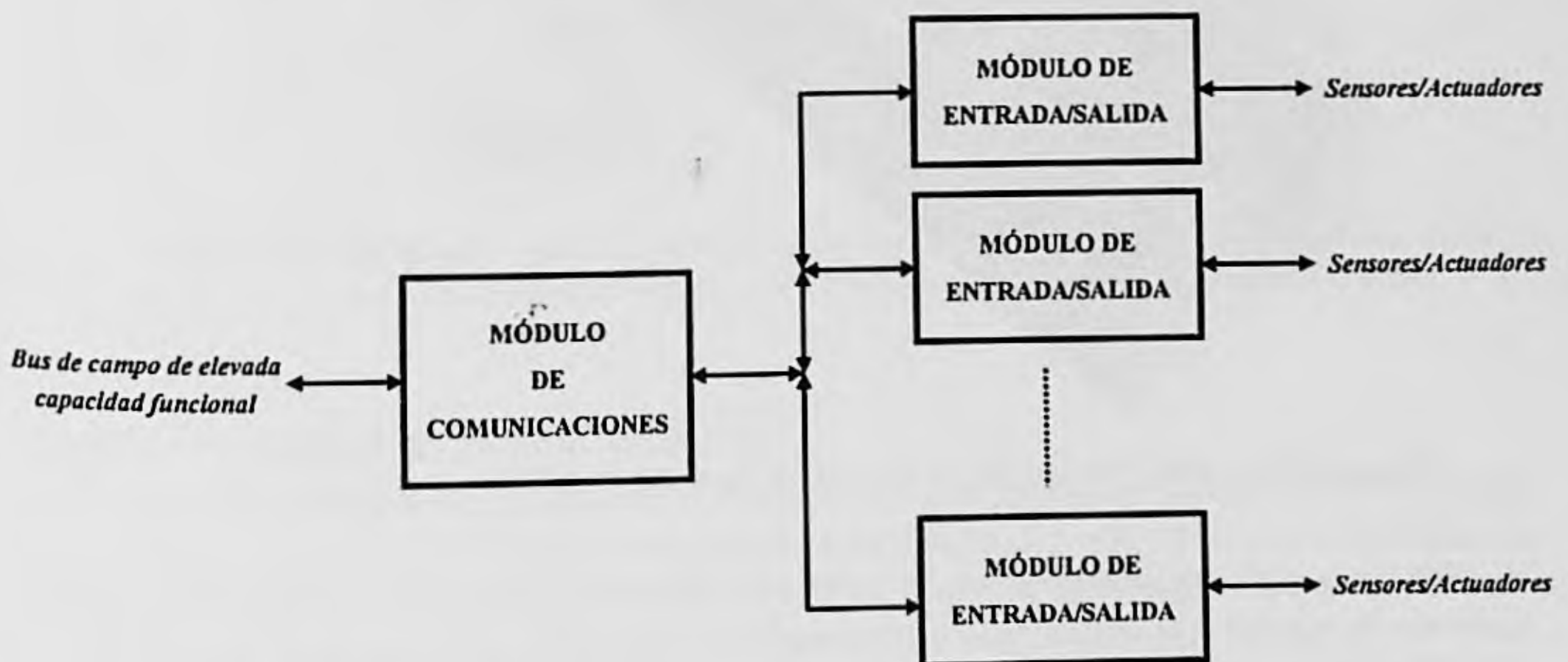


Figura 8.37. Diagrama de bloques de una estación de entrada/salida remota (Remote terminal units).

Por ello los fabricantes de autómatas programables comercializan unidades de entrada/salida remotas (*Remote Terminal Units*), denominadas también periferia descentralizada, que, tal como se indica en la figura 8.37, están formadas por:

- Un módulo de comunicaciones, que es un procesador digital con capacidad para enviar mensajes a un autómata programable y recibirlos procedentes de él, a través de un bus de campo (*Field bus*) normalizado de elevada capacidad funcional, que es una red de comunicaciones, denominada red de control, que tiene unas características adecuadas

para enviar información en un tiempo máximo preestablecido. La importancia de los buses de campo hace que a ellos se dedique el apartado 9.3.2.3 del capítulo 9.

- Un conjunto de módulos de entrada/salida de aplicación general o de aplicación específica, a los que se conectan sensores y actuadores todo-nada y analógicos. Algunos de estos módulos pueden a su vez incorporar un procesador de comunicaciones para conectar sensores a través de un bus de campo normalizado de capacidad funcional limitada, como por ejemplo el HART (*Highway Addressable Remote Transducer*) o el AS-i (*Actuator-Sensor-Interface*), descrito en el apéndice 2.



Figura 8.38. Unidad de entrada/salida remota Simatic ET 200 ECO de Siemens.

Son un ejemplo de este tipo de unidades de entrada/salida remotas los sistemas de periferia descentralizada ET 200 de Siemens que se conectan a los autómatas programables S7-300 y S7-400 a través de una red o bus de campo PROFIBUS DP (descrita en el apéndice 3) con una velocidad máxima de transmisión de 12 Mbits por segundo. Cada uno de los sistemas de periferia descentralizada ET 200 permite la conexión de módulos de variables de entrada todo-nada y analógicas y módulos de aplicación específica (de posicionamiento, de contaje, arrancador de motores, etc.). Algunos de los módulos incorporan un procesador de comunicaciones que les permite su conexión a distintos tipos de sensores a través del bus de campo normalizado HART o del AS-i descrito en el apéndice 2. Existen diversos tipos de sistemas de periferia descentralizada ET 200, como son los ET 200S, ET 200iS, ET 200X, ET 200L, ET 200M, ET 200B, ET 200ECO y ET200R que se distinguen unos de otros por lo que se refiere a su capacidad de ampliación, número y tipo de variables de entrada y de salida, diferentes grados de protección ambiental y sellado IP (apartado 7.6.1 del capítulo 7), etc. En la figura 8.38 se muestra como ejemplo la unidad periférica modular ET 200 ECO de Siemens, que se caracteriza por:

- Es un módulo de entrada/salida compacto, de bajo coste y no ampliable, de variables todo/nada (digitales).
- Posee un grado de protección ambiental o sellado IP67 que permite su instalación a pie de máquina de forma directa sin necesidad de colocarlo en el interior de un armario.
- Se puede conectar a la red de control PROFIBUS DP.

En la figura 8.39 se representa el diagrama de bloques de un sistema de control distribuido que conecta, a través de una red de control PROFIBUS DP, un autómata programable (PLC) con una unidad de interfaz máquina-usuario (HMI) (descrita en los apartados 1.3.3.3.2 y 8.3.2), un equipo o aparato de programación y dos unidades de entrada-salida remota ET200X.

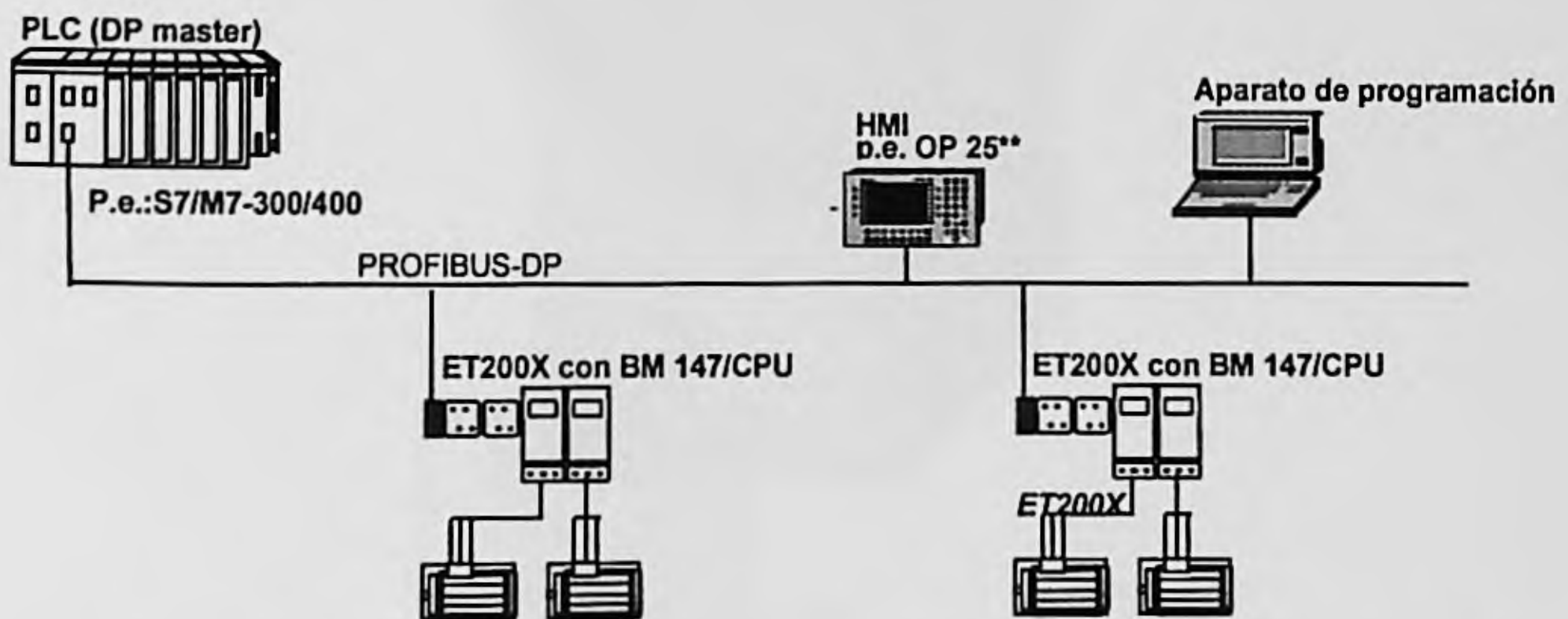


Figura 8.39. Sistema de control distribuido implementado con un autómata programable de la familia S7-300 o S7-400 y estaciones remotas ET 200X (cortesía de Siemens).

8.2.3.4 Unidades de posicionamiento

Otro tipo de módulos de aplicación industrial son los que controlan la posición de uno o más ejes. Su combinación con un autómata programable convierte a este último en un sistema de control numérico (descrito en el apartado 9.2.4.4.1 del capítulo 9). Existen distintos módulos de posicionamiento que se diferencian por sus prestaciones y su modo de operación.

Módulos de control en bucle cerrado

Un módulo de este tipo controla el movimiento del motor (avance, retroceso, movimiento rápido, movimiento lento) y mide la posición del eje mediante un codificador angular. Tal como se indica en el apartado 9.2.4.4.1 del capítulo 9, realizan el posicionamiento dinámico de los ejes con elevada precisión.

Un ejemplo de este tipo de módulo es el FM 351, que pertenece a la serie S7-300 y permite el control de dos ejes independientes. Para el control de cada eje dispone de 5 salidas todo/nada que controlan el motor y una entrada para el codificador angular. En la figura 8.40 se indica una aplicación de este tipo de módulo.

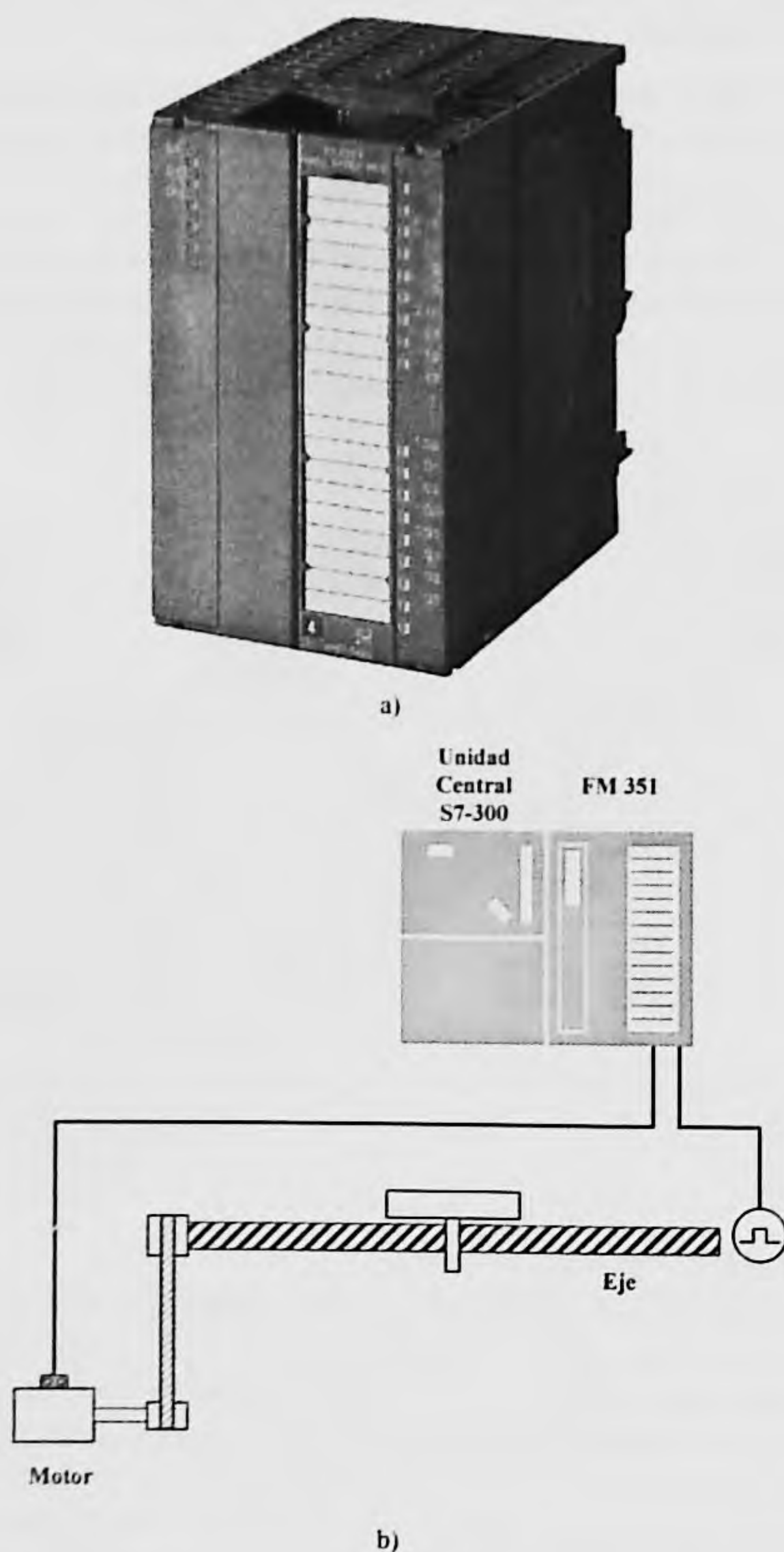


Figura 8.40. Módulo FM 351 de posicionamiento en bucle cerrado: a) Fotografía; b) Ejemplo de aplicación (cortesía de Siemens).

Módulos de control en bucle abierto

Utilizan motores paso a paso y se caracterizan porque la posición del eje sólo depende del número de pasos que gira el eje del motor, lo que hace innecesaria, en principio, la medición del giro con un codificador angular.

Esta forma de control se utiliza ampliamente también en máquinas herramienta, robots, etc. Un ejemplo de este tipo de módulos es la combinación del módulo FM353 y el SIMOSTEP o STEPDRIVE de la serie S7-300, que constituye la parte de potencia que actúa sobre un motor paso a paso de tres fases. En la figura 8.41 se representa un sistema de este tipo que puede controlar motores cuyo número de pasos puede variar entre 500 y 10.000 por revolución.

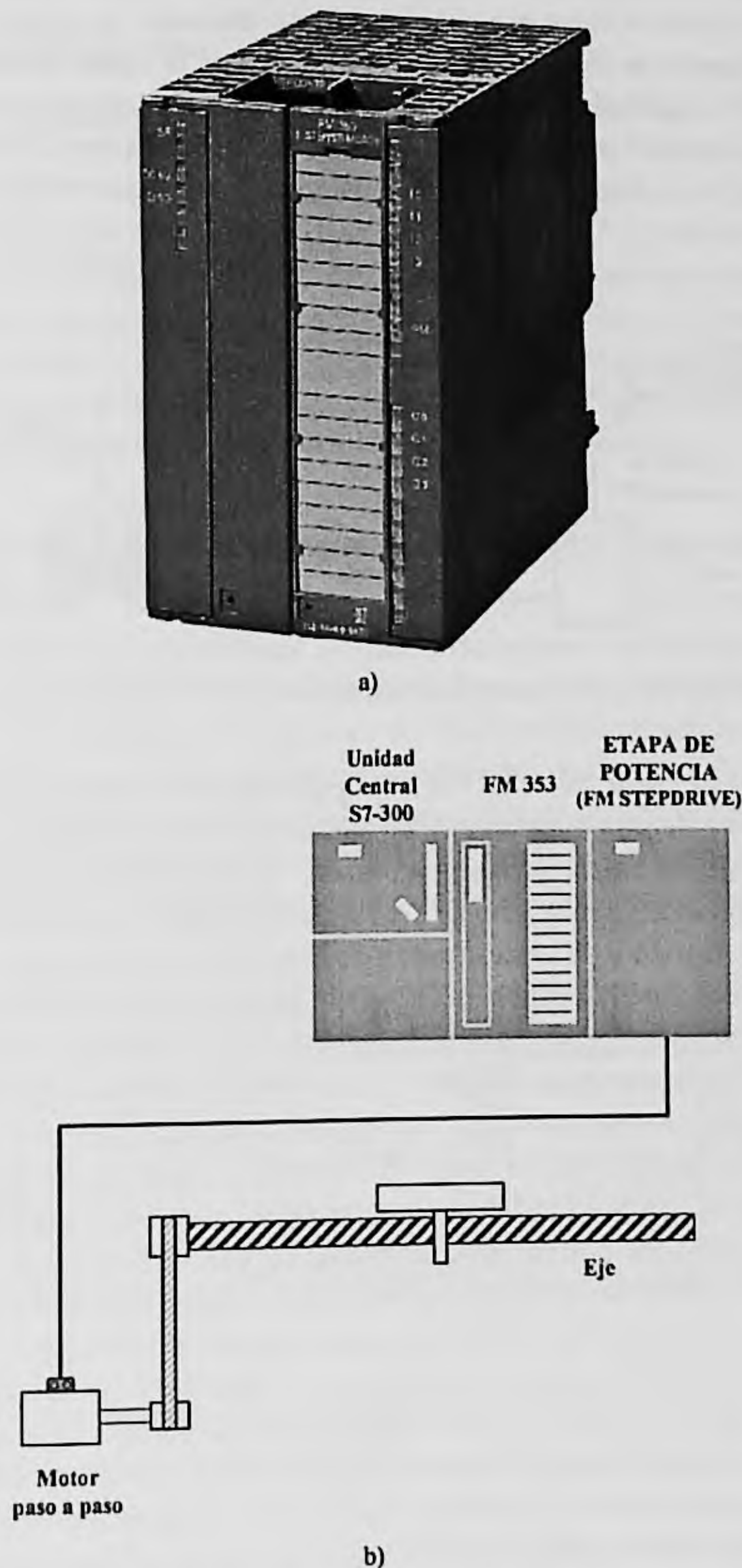


Figura 8.41. Módulo FM 353 de posicionamiento mediante motor paso a paso: a) Fotografía; b) Ejemplo de aplicación (cortesía de Siemens).

8.2.3.5 Unidades de regulación

Para llevar a cabo el control de muchos procesos industriales, es necesario regular (mantener constante o hacer que evolucione de una forma determinada) el valor de una variable física *PV* (*Process variable*), como por ejemplo el caudal del líquido que circula por una tubería. En la figura 8.42 se representa el esquema de bloques funcional de un sistema de este tipo, denominado regulador o controlador lineal continuo, al que se le proporciona el punto de consigna *SP* (*Set point*), que es el valor que debe alcanzar y mantener la variable controlada (en este caso, el caudal). Además se le proporciona el valor real de la citada variable *PV* por medio de un sensor de caudal o caudalímetro *C*. La diferencia entre el punto de consigna y el valor real constituye la señal de error *e* que puede ser positiva o negativa según que pase más o menos caudal del establecido mediante el punto de consigna. El regulador utiliza la señal de error para generar la señal de control *CV* (*control variable*) que actúa sobre la válvula de control *V* y la posiciona en el punto adecuado. Se tiene así un sistema de control en bucle cerrado.

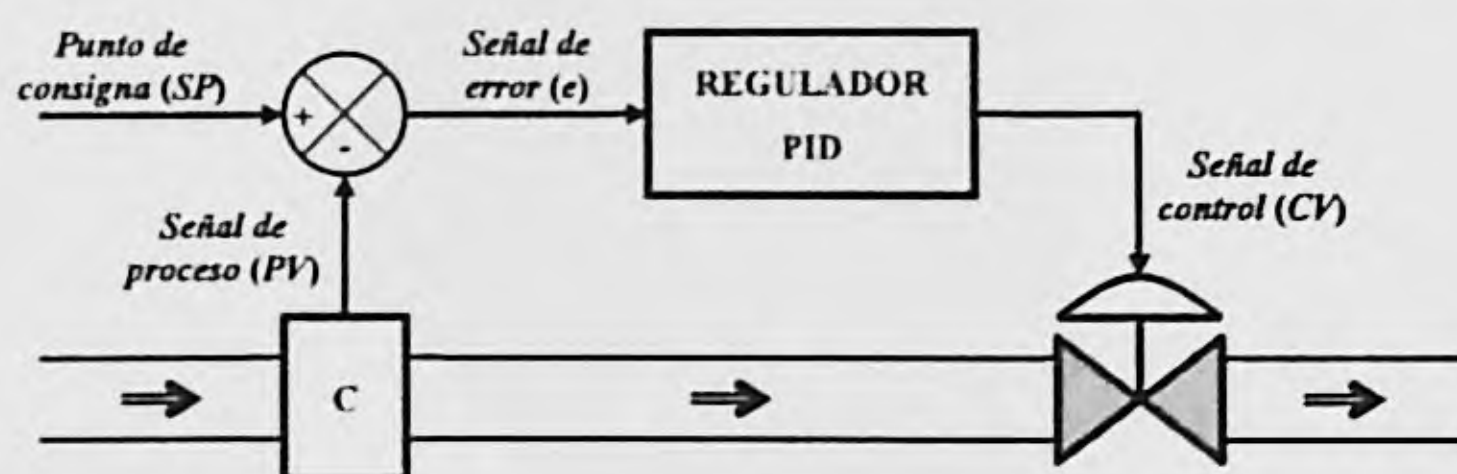


Figura 8.42. Diagrama de bloques de un regulador PID de caudal.

La señal de control puede tener un valor proporcional al de la señal de error, y cuanto más grande es la constante de proporcionalidad K_p , menor es el error en régimen permanente y más rápido actúa el regulador ante los cambios bruscos de la señal de consigna. Pero, en contrapartida, si dicha constante es muy grande y el sistema responde rápidamente (sistema de baja inercia) a un cambio brusco y de valor elevado de la variable de consigna, se puede convertir en inestable porque la señal de control se hace mucho mayor, lo que provoca que el error se haga negativo, y que el regulador actúe en sentido contrario y de nuevo vuelva a pasarse del punto de consigna, y así sucesivamente. En estas condiciones el sistema tarda en alcanzar el valor del punto de consigna, se hace inestable y en algunos casos puede que incluso se convierta en oscilante y que nunca alcance dicho valor. Por otra parte, si la constante de proporcionalidad es pequeña, la señal de control también lo es y se tiene un error permanente entre el punto de consigna y el valor real. Para evitar este problema se hace que la señal de control tenga una componente proporcional a la integral (I) del error en el tiempo (lo que equivale a disminuir la ganancia solamente en el instante en el que cambia la señal de consigna). La ejecución de estas dos operaciones por parte del regulador dan lugar a lo que se denomina un regulador PI, que es ampliamente utilizado en la industria. Pero en algunas aplicaciones industriales la variable que se quiere regular, como por ejemplo la temperatura de un horno, tiene mucha inercia al cambio. En este caso es necesario elevar la ganancia en el instante en que cambia la señal de consigna y para ello se debe hacer que la señal de salida tenga una componente proporcional a la derivada (D) de la señal de error respecto del tiempo. Por esta razón el regulador utilizado debe combinar la función P con la I y con la D y por ello se le denomina regulador o controlador PID. El algoritmo ejecutado por este regulador responde por lo tanto a la ecuación:

$$CV(t) = K_p e(t) + K_i \int e(t) dt + K_d \frac{de(t)}{dt}$$

en la cual:

- $CV(t)$: Señal de control
- $e(t)$: Señal de error
- K_p : Constante Proporcional
- K_i o T_i : Constante Integral
- K_d o T_d : Constante Derivativa

Inicialmente estos reguladores fueron mecánicos y neumáticos, y posteriormente, con el desarrollo de la Electrónica, y en concreto de los amplificadores operacionales, comenzaron a realizarse con circuitos electrónicos analógicos. En la actualidad dicha tarea puede ser realizada por un autómatas programable de acuerdo con dos estrategias diferentes:

- **Mediante la ejecución de un programa de control PID por parte del autómatas programable**

En la figura 8.43 se representa el diagrama de bloques correspondiente. El autómatas recibe la señal del sensor de caudal a través de un módulo entrada de variables analógicas y actúa sobre la electroválvula a través de un módulo de salidas de variables analógicas. El punto de consigna, así como las constantes de proporcionalidad, integración y derivación, las puede proporcionar el usuario a través de una interfaz máquina-usuario (HMI). Esta forma de realizar un regulador PID es la más adecuada cuando el autómatas programable tiene tiempo para llevar a cabo esta tarea por programa.

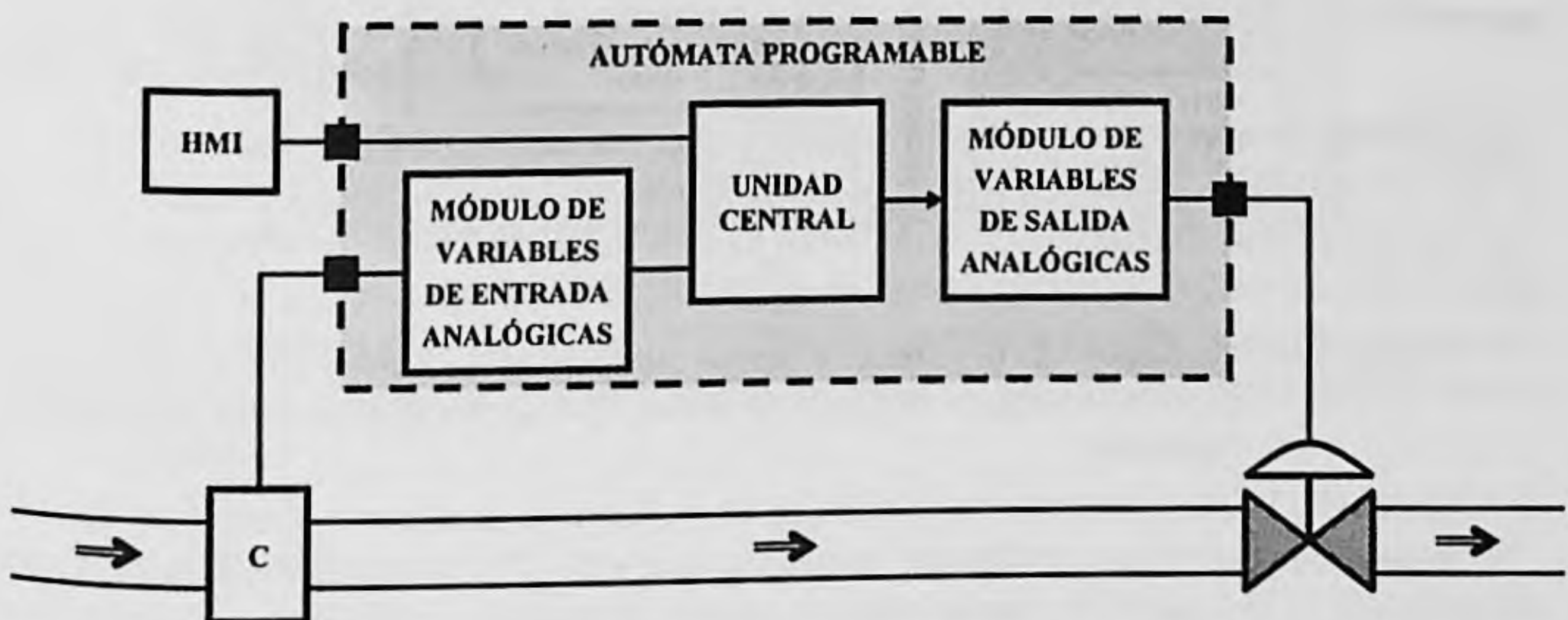


Figura 8.43. Diagrama de bloques de un regulador PID de caudal implementado con un autómatas programable mediante módulos de entrada/salida de aplicación general.

- **Mediante un módulo o unidad de regulación acoplado al autómata programable**

El diagrama de bloques correspondiente se representa en la figura 8.44a. La unidad central del autómata programable está conectada a un módulo de regulación, cuyo diagrama de bloques típico se indica en la figura 8.44b, que posee un procesador digital y un sistema de adquisición de datos al que se conecta el sensor de caudal. El procesador digital ejecuta el algoritmo con la información digital obtenida a partir del sensor y se acopla a un convertidor digital-analógico cuya salida se conecta a la electroválvula a través de un amplificador. El autómata programable proporciona al procesador del módulo de regulación la señal de consigna y las constantes de proporcionalidad, integración y derivación introducidas en el autómata programable, por ejemplo, a través de una unidad HMI.

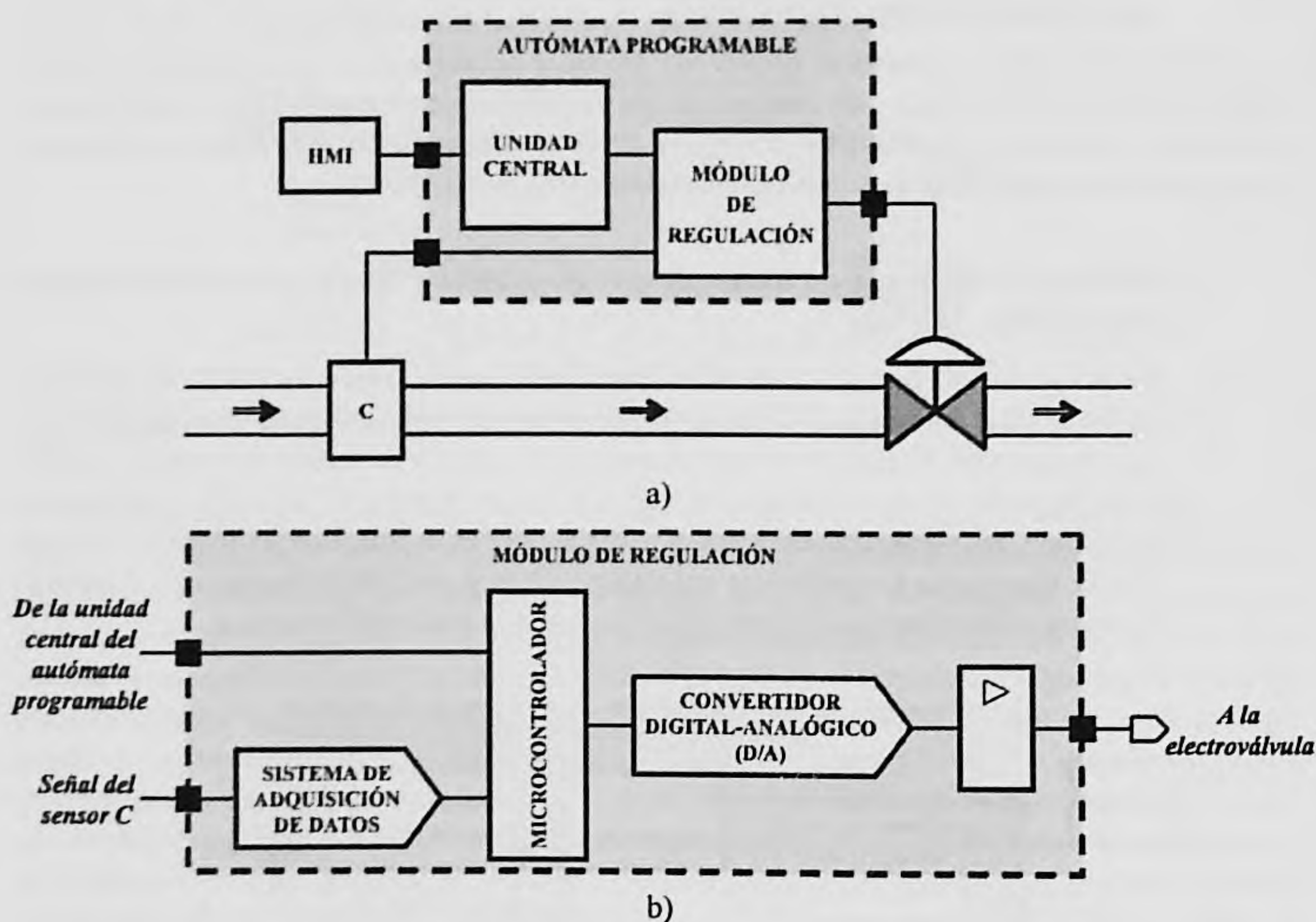


Figura 8.44. Regulador PID implementado con un autómata programable y un módulo de regulación: a) Diagrama de bloques general; b) Diagrama de bloques del módulo de regulación.

Un ejemplo de módulo específico de regulación es el FM 455 de la serie S7-400 de Siemens, representado en la figura 8.45, que es capaz de regular un máximo de 16 lazos de control. Este módulo admite señales analógicas de tensión y de corriente en los rangos normalizados y además se le pueden conectar directamente diferentes tipos de termopares y termorresistencias. De igual forma, a su salida proporciona diferentes valores normalizados de tensión y corriente.

En el capítulo 6 se estudia con mayor detalle el control de procesos mediante autómatas programables en general y los controladores lineales continuos en particular.



Figura 8.45. Módulo de regulación FM 455 de la familia S7-400 de Siemens.

8.3 Interfaces de conexión autómata-usuario

Los usuarios de los autómatas programables necesitan comunicarse con ellos para llevar a cabo dos funciones principales:

- Utilizar las herramientas de diseño asistido por computador (*CAD tools*) que el fabricante del autómata programable pone a su disposición, para lo cual se necesita una unidad de programación.
- Modificar parámetros (actuación) y observar el estado de determinadas variables (visualización), en especial cuando el proceso controlado es complejo, para lo cual es necesaria una unidad de acoplamiento (interfaz) entre el usuario y la máquina que suele recibir el nombre de HMI (acrónimo de *Human Machine Interface*) o MMI (acrónimo de *Man Machine Interface*). Además la interfaz máquina-usuario está asociada a un programa, que muchos fabricantes de equipos de automatización llaman SCADA.

A continuación se analizan las unidades de programación, los equipos HMI y los programas SCADA.

8.3.1 Unidades de programación

Inicialmente las unidades de programación eran procesadores digitales especializados, comercializados por cada fabricante de autómatas programables para llevar a cabo la tarea de

desarrollo del programa de control. El vertiginoso desarrollo de los computadores personales, y en especial de los portátiles, hizo que poco a poco se fuesen utilizando como estaciones de diseño de sistemas de control basados en autómatas programables.

La utilización de computadores personales proporciona grandes ventajas, debido a su gran capacidad de tratamiento y almacenamiento de información y su facilidad de comunicación. Además, los distintos fabricantes han elaborado diversas unidades de programación, que realmente son computadores personales portátiles especialmente diseñados para trabajar en ambiente industrial y en entornos agresivos. En la figura 8.46 se representan como ejemplo dos unidades de programación Field PG de Siemens.



Figura 8.46. Unidades de programación Field PG PIII de Siemens.

8.3.2 Equipos de interfaz máquina-usuario

8.3.2.1 Introducción

Los conceptos de visualización y actuación surgieron de forma natural cuando el ser humano comenzó, a principios del siglo XX, a realizar máquinas electromecánicas gobernadas mediante un sistema de control (tal como se indica en el apartado 4.1 del capítulo 4), que recibía órdenes o consignas de un operador y generaba señales que actuaban sobre la citada máquina. Los sistemas de control lógico se realizaban entonces con relés y los de proceso con válvulas y las órdenes o consignas las tenía que dar el usuario, para lo cual utilizaba elementos cuyo funcionamiento era fácil de entender. Se trataba de interruptores y pulsadores en el caso de variables lógicas o todo-nada y potenciómetros para generar variables analógicas. Para visualizar las variables analógicas se utilizaban, cuando era necesario, galvanómetros convertidos en amperímetros o voltímetros [MAND 95] [STOC 79] y para visualizar las variables todo-nada

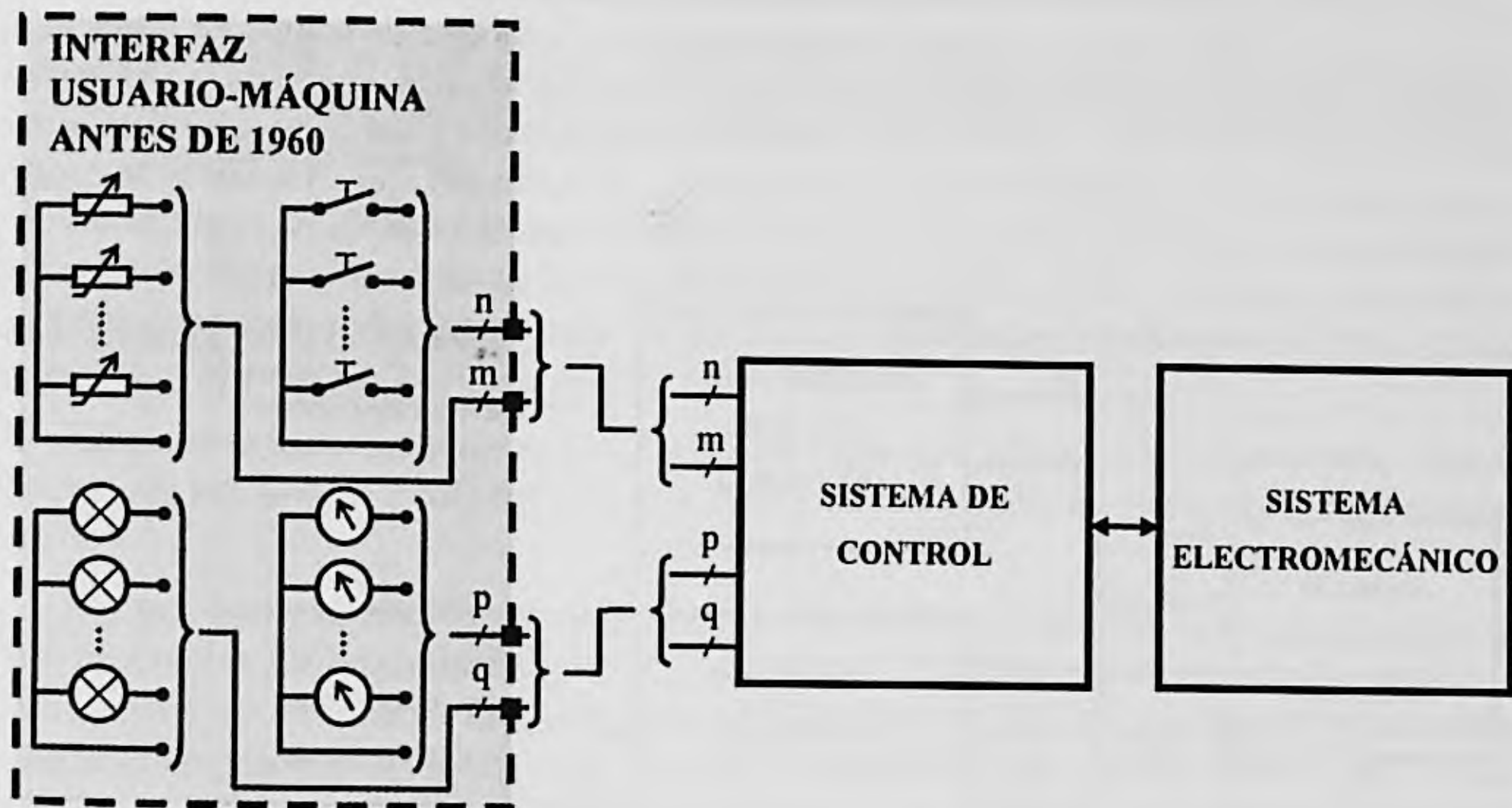


Figura 8.47. Diagrama de bloques de una unidad HMI anterior a 1960.

pilotos que eran lámparas de incandescencia o neón de reducido tamaño, gobernadas mediante relés. Por lo tanto en esa época la interfaz usuario-máquina estaba formada por un conjunto de los citados elementos tal como se indica en la figura 8.47.

El desarrollo de los transistores hizo que en la década de 1960 se utilizasen puertas lógicas con transistores para sustituir a los relés en los sistemas de control lógico y controladores PID con amplificadores operacionales para sustituir a los amplificadores implementados con válvulas en los de control de procesos continuos. La sustitución de los relés por transistores hizo que se tuviesen que utilizar indicadores luminosos para conocer si una determinada salida estaba en nivel cero o en uno, lo cual en los circuitos implementados con relés se comprobaba viendo los contactos. Esta fue una de las razones que impulsó la investigación básica para conocer las propiedades de emisión de luz de ciertos compuestos de semiconductores, como por ejemplo el arseniuro de galio, lo que trajo como resultado el desarrollo de los diodos luminiscentes conocidos por el acrónimo LED (*Light Emitting Diode*) y su incorporación como dispositivos de visualización de las variables de salida de los primeros controladores lógicos electrónicos digitales cableados (descritos en el apartado 1.2.2.2.3 del capítulo 1).

Además, el progreso de la Electrónica de estado sólido no sólo hizo cambiar la forma de construir los sistemas electrónicos de control sino que mejoró las prestaciones y redujo el coste de los computadores de aplicación general. Unido todo ello al aumento de la complejidad de algunas instalaciones de proceso continuo (como por ejemplo las plantas químicas) hizo que los computadores, que hasta ese momento se dedicaban solamente a la gestión (bancos, etc.), comenzasen a utilizarse en aquellos casos en los cuales su elevado coste no constituía un inconveniente insalvable. Nació así, tal como se indica en el apartado 4.2.2.3.2 del capítulo 4, el denominado Control Digital Directo conocido como DDC (acrónimo de *Direct Digital Control*) constituido por un computador digital en el que se utilizaba como interfaz máquina-usuario el subsistema formado por una pantalla y un teclado acoplados a la unidad central de

proceso (CPU) a través del correspondiente interfaz, tal como se representa en la figura 8.48. Fue en ese momento cuando comenzó a adquirir sentido el concepto de equipo de actuación y visualización que se denominó inicialmente interfaz hombre-máquina, conocida por el acrónimo MMI (*Man-Machine Interface*) y que pasó a denominarse posteriormente interfaz máquina-usuario o HMI (acrónimo de *Human-Machine Interface*), que es como se la conoce actualmente.

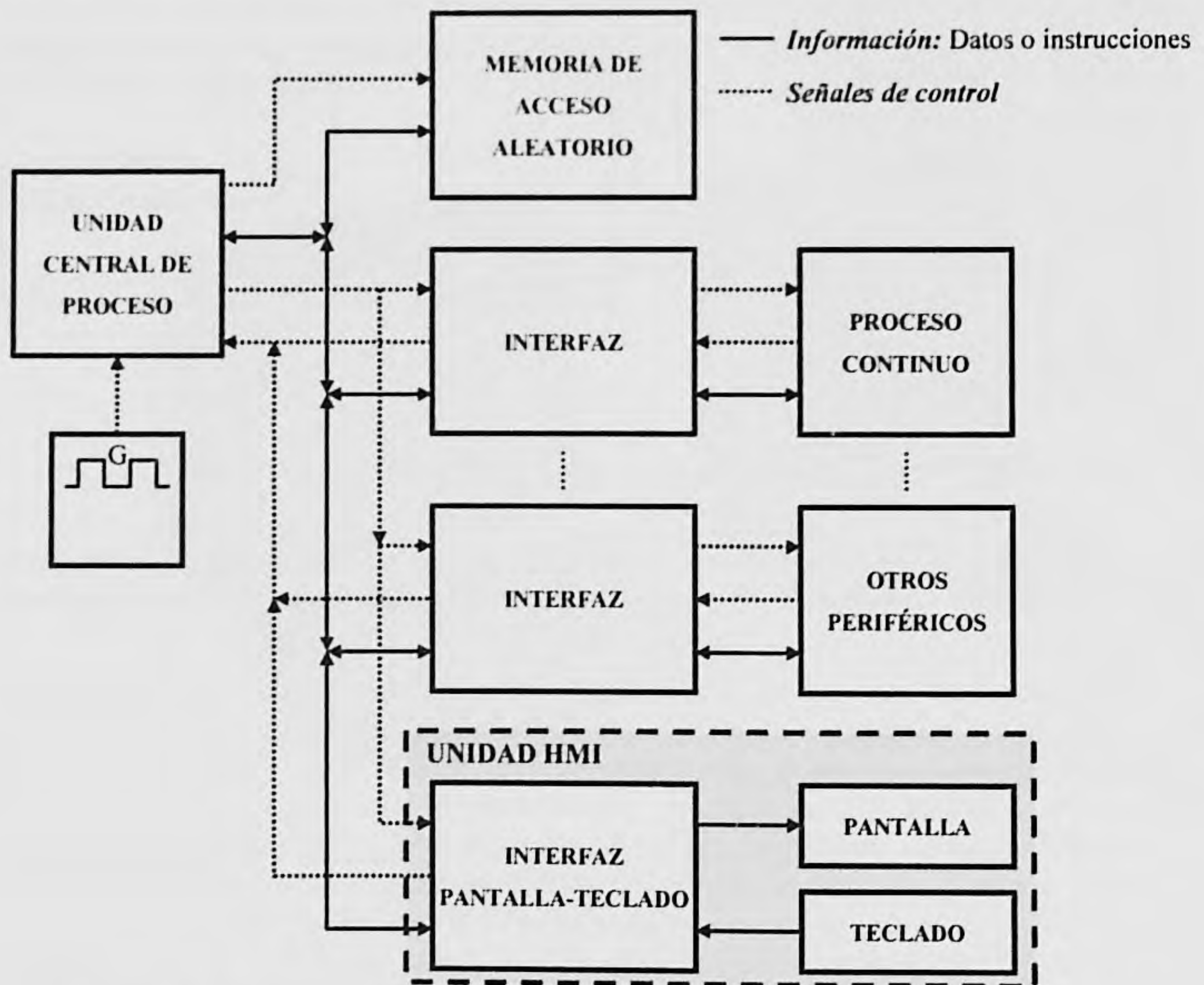


Figura 8.48. Diagrama de bloques de la unidad HMI de un computador de control digital directo.

8.3.2.2 Características de los equipos HMI

Los equipos de visualización y actuación reciben actualmente, tal como se indica en el apartado anterior, la denominación de interfaz máquina-usuario y se les suele conocer por el acrónimo HMI (*Human Machine Interface*).

Tal como se indica en el apartado 8.1, su utilización es necesaria cuando la aplicación exige que el usuario pueda:

- Modificar parámetros del programa de control y dar órdenes a los actuadores a través de él.

- Recibir información del estado del proceso controlado por el autómatas programable.
- Detectar fallos en el conjunto formado por el autómatas programable y el proceso controlado por él, que dan lugar a alarmas ante las cuales el operador debe realizar las acciones oportunas.

Constituyen un ejemplo de instalaciones en las que es imprescindible presentar al usuario la situación del proceso automatizado y facilitarle la actuación sobre el mismo a través de las órdenes oportunas, todos los sistemas de fabricación asistida por computador (CAM), que se estudian en el apartado 9.2.4 del capítulo 9.

Tal como se indica en el apartado 1.3.3.3.2 del capítulo 1, actualmente existen varias formas de realizar una unidad HMI, entre las que cabe citar los paneles de operación y las pantallas táctiles.

Los paneles de operación, conocidos por las siglas OP (acrónimo de *Operation Panel*), están formados por una pantalla gráfica y un conjunto de pulsadores de membrana asociados, que constituyen un teclado. Tal como se puede observar en la figura 8.49a, ambos se controlan mediante un procesador especializado que es un controlador de pantalla y teclado, que se acopla al autómatas programable a través de la correspondiente interfaz, y un ejemplo de panel de operación es la unidad OP 277B de Siemens (Figura 8.49b).

Los paneles o pantallas táctiles conocidos por las siglas TP (acrónimo de *Touch Panel*) utilizan una pantalla gráfica que posee elementos sensores sensibles al tacto. De esta forma,

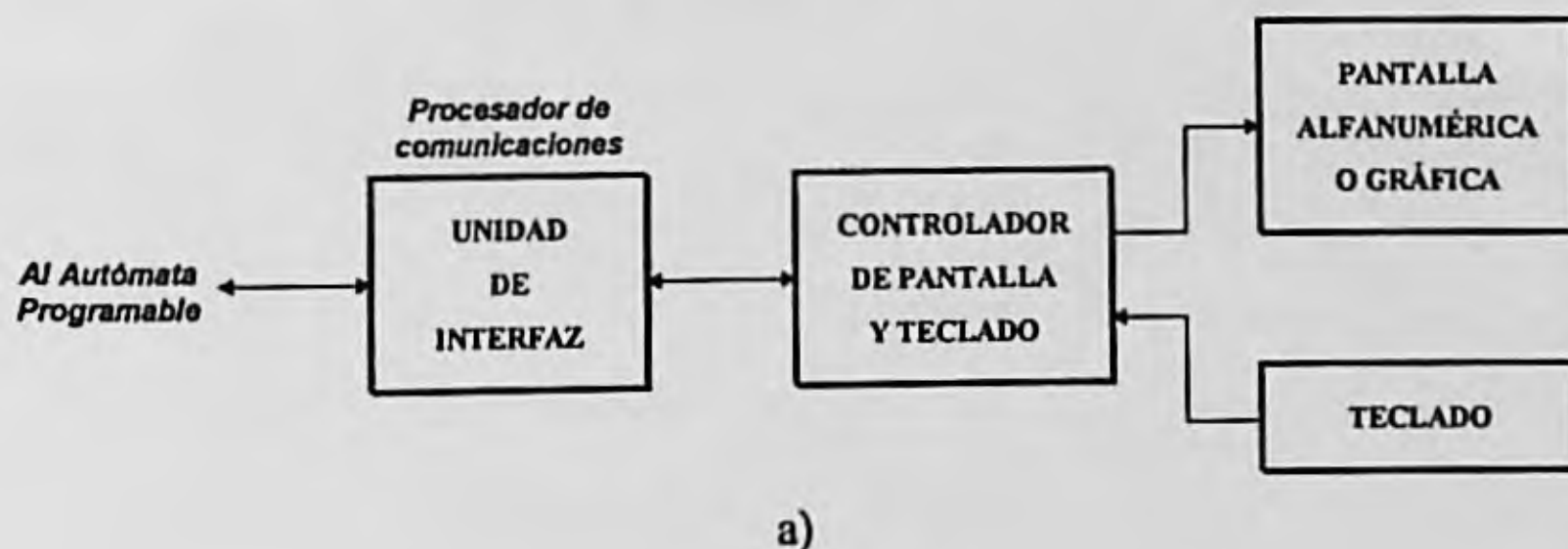


Figura 8.49. a) Diagrama de bloques de un panel de operación; b) Fotografía del panel de operación Simatic OP 277 de Siemens.

la pantalla realiza la función de entrada y de salida y se elimina el teclado. Su diagrama de bloques se representa en la figura 8.50a y es un ejemplo de panel táctil la unidad HMI TP277 de Siemens (Figura 8.50b). Tanto los paneles de operación como las pantallas táctiles pueden poseer diferentes características en lo referente a las dimensiones de la pantalla, capacidad gráfica de la misma, dimensiones del teclado, grado de protección ambiental o sellado IP (descrito en el apartado 7.7.2 del capítulo 7), etc. En cada aplicación se debe elegir el tipo de unidad HMI adecuada.

Por otra parte, tal como se indica también en el apartado 1.3.3.3.2 del capítulo 1 (Interfaz máquina-usuario), existen máquinas en las que es conveniente que el sistema electrónico de

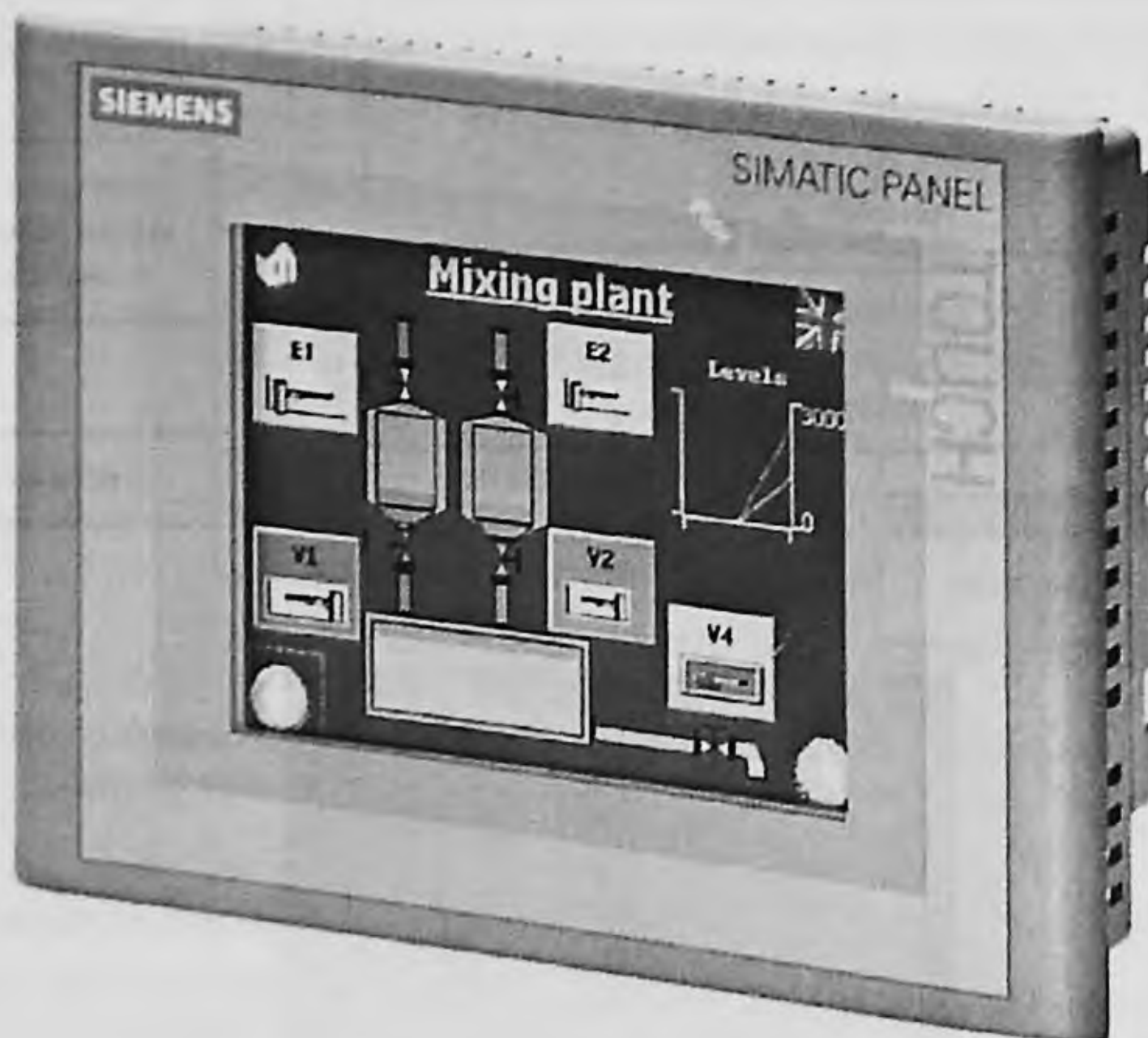
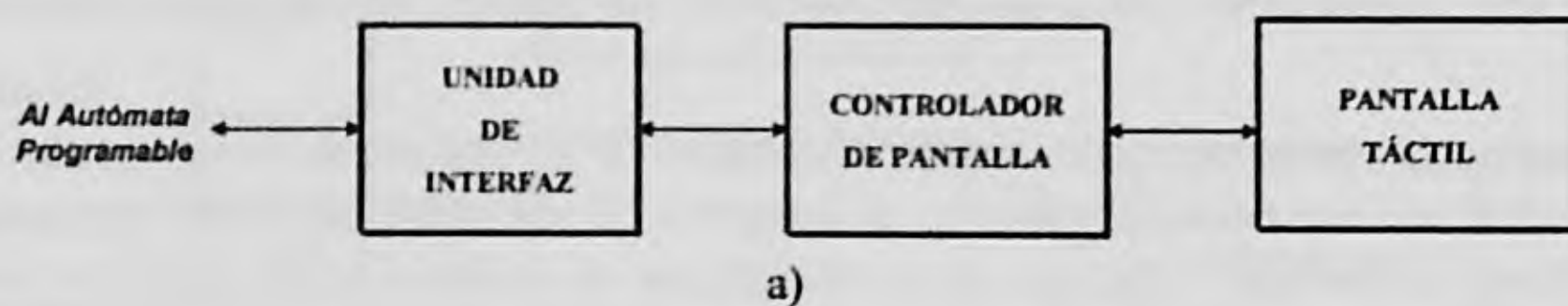


Figura 8.50. a) Esquema de bloques de un panel táctil; b) Fotografía del panel táctil Simatic TP 277 de Siemens.

control esté embebido o empotrado en las mismas (*Embedded Automation*), para lo cual la unidad HMI debe ocupar el mínimo espacio posible y proporcionar más funciones que un panel de operación o un panel táctil, como por ejemplo elevada capacidad de memoria, funciones de autómata programable, funciones de supervisión y adquisición de datos conocidas como SCADA (acrónimo de *Supervisory Control And Data Acquisition*), descritas en el apartado 8.3.3 a continuación, etc.

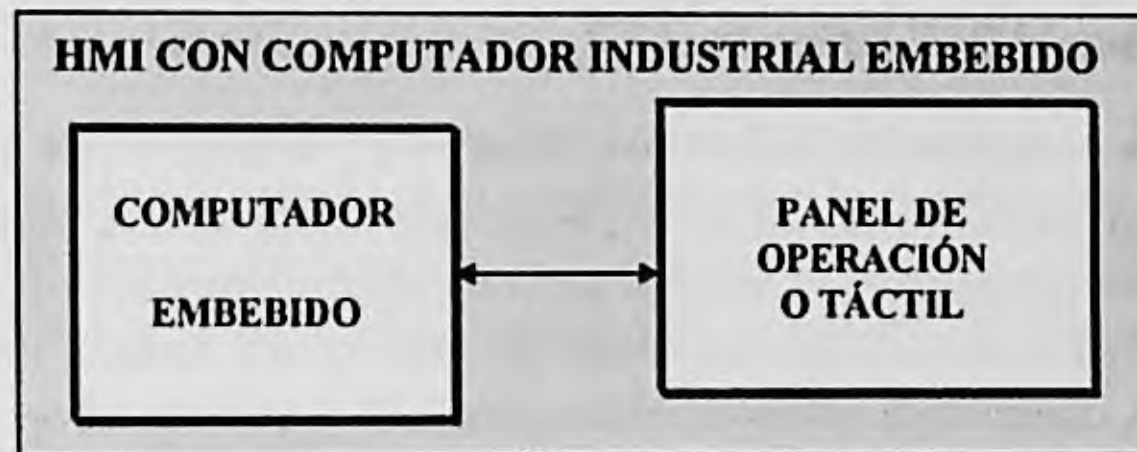


Figura 8.51. Diagrama de bloques de una unidad HMI con computador industrial embebido.

A fin de atender esta necesidad, los fabricantes de sistemas de automatización comercializan equipos, cuyo diagrama de bloques se representa en la figura 8.51, que incluyen en una sola carcasa una unidad HMI y un computador industrial embebido. De acuerdo con las características de este último, los paneles con computador industrial embebido se pueden dividir en dos grandes clases:



Figura 8.52. Fotografía del multipanel MP 277 de Siemens.

- **Paneles con computador industrial embebido y arquitectura cerrada**

Este tipo de paneles utiliza un computador embebido con un sistema físico no ampliable, un sistema operativo empotrado y recursos de programación, entre los que se puede incluir un emulador de autómatas programable (*Soft-PLC*), descrito en el apartado 9.2.4.4.3 del capítulo 9, y un programa SCADA de supervisión y adquisición de datos (descrito en el apartado 8.3.3 a continuación). Ejemplo de este tipo de unidad HMI es la familia MP de multipaneles (*Multipanels*) de Siemens, uno de cuyos miembros es el multipanel MP 277 (Figura 8.52).

- **Paneles con computador industrial embebido y arquitectura abierta**

Para mejorar las prestaciones de los paneles de arquitectura cerrada, los fabricantes de autómatas programables utilizan un computador embebido implementado con un microprocesador de elevada velocidad de cálculo, un sistema físico ampliable (*Expandable*), como el de un computador personal, y un sistema operativo de aplicación general (como por ejemplo Windows XP). El computador puede incluir una placa que implementa un autómata programable (*Slot-PLC*) así como un programa SCADA de supervisión y adquisición de datos (descrito en el apartado 8.3.3 a continuación). Ejemplo de este tipo de unidad HMI son los miembros de la familia Simatic Panel PC de Siemens (Figura 8.53).

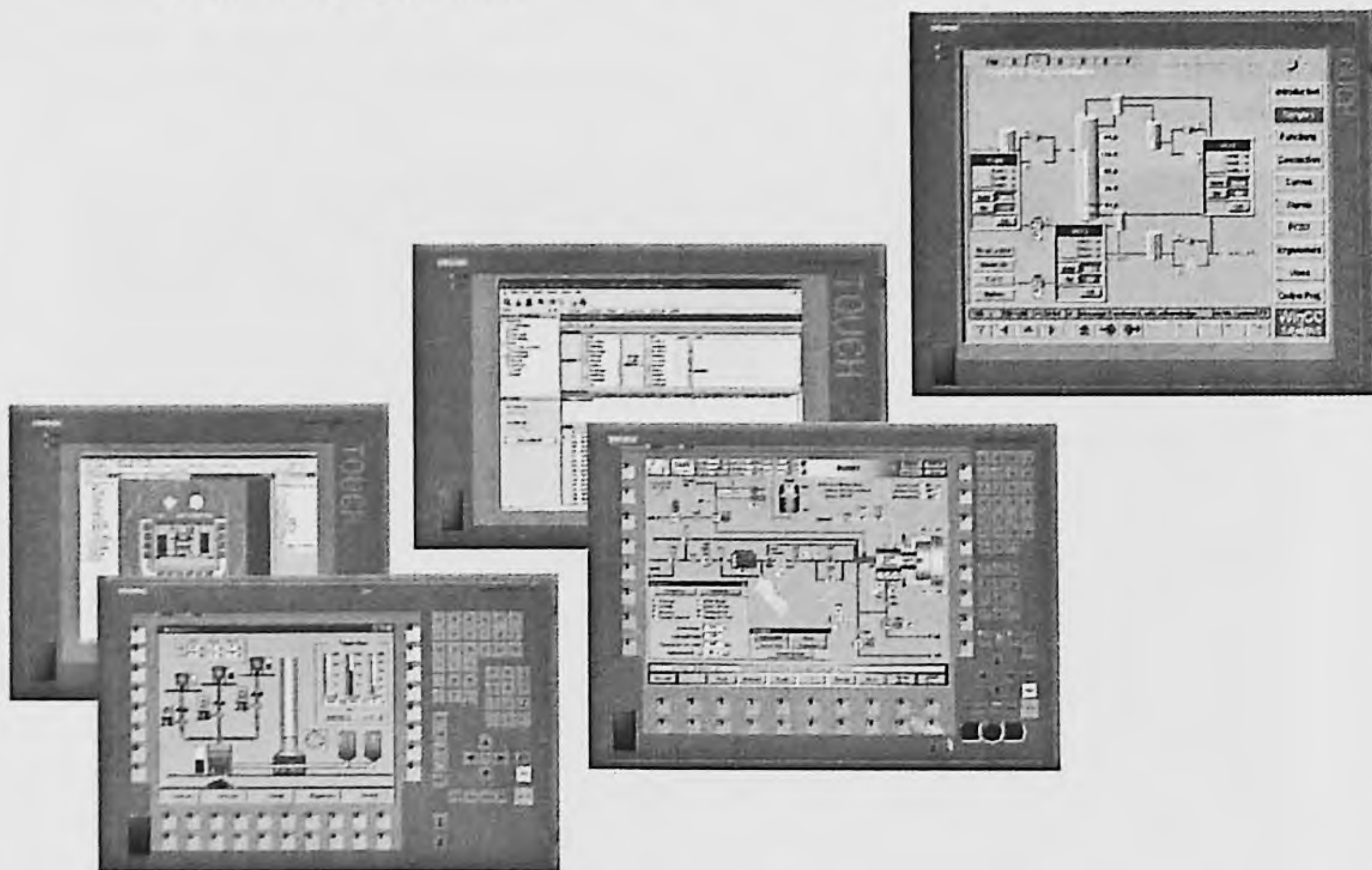


Figura 8.53. Familia Simatic Panel PC de Siemens.

Uno de las características más importantes de las interfaces HMI es la facilidad de utilización (*Usability*) [NIEL 93], también conocida como calidad de utilización (*Quality of use*), que de acuerdo con la norma ISO 9241 [ISO 08] está compuesta por tres factores:

- La eficacia (*Effectiveness*), que indica el grado de cumplimiento de las necesidades del usuario.
- La eficiencia (*Efficiency*), que indica la rapidez con la que el usuario aprende a utilizarla.
- La satisfacción (*Satisfaction*) del usuario, que indica si contribuye a facilitarle el trabajo que debe realizar.

Para lograrlo se ha de llevar a cabo el diseño de la interfaz HMI teniendo en cuenta los tres factores citados.

El desarrollo de la Microelectrónica ha propiciado que los fabricantes de unidades HMI las doten de capacidad de comunicación a través de una o más redes de control normalizadas, para facilitar así su ubicación en el lugar más adecuado de la instalación (Figura 8.39).

8.3.3 Sistemas de supervisión y adquisición de datos (SCADA)

Tal como se indica en el apartado anterior, un equipo HMI facilita la interacción entre el usuario y un sistema de control, generalmente distribuido, pero, en la mayoría de las instalaciones de fabricación o de control de procesos, es necesario además:

- Representar gráficamente instalaciones de fabricación o procesos muy complejos que exigen la visualización de un gran número de datos de manera clara y precisa.
- Almacenar en la memoria de un computador una secuencia de informaciones que incluya alarmas, evolución histórica (*Trends*) de diversas señales e informes (*Reports*), para que el usuario tenga constancia de lo que ha sucedido en diferentes instantes de tiempo durante el funcionamiento de la instalación.
- Facilitar la ampliación de la instalación o la modificación de las funciones que realizan las máquinas de la misma, para lo cual ha de ser modular.
- Tener una elevada disponibilidad para evitar, en algunas instalaciones, los inconvenientes de la puesta fuera de servicio del computador ligado al equipo HMI.
- Gestionar un sistema de comunicaciones que permita realizar la adquisición de datos y el envío de órdenes de forma distribuida.

Para que un sistema de automatización cumpla con los requisitos que se acaban de indicar no solo ha de tener un equipo HMI sino que además éste ha de estar asociado con un computador en el que se ejecute un programa que se encargue de adquirir la información de los sensores y la almacenada en la memoria de los controladores instalados en las máquinas, así como de gestionar la visualización. Son numerosos los fabricantes de equipos de automatización que le dan a este programa la denominación de SCADA (acrónimo de *Supervisory Control And Data Acquisition*).

Por lo tanto, se puede definir un SCADA como una herramienta informática que constituye un programa de computador que tiene como misión facilitar la adquisición de los datos generados por un conjunto de sensores y controladores de un proceso industrial, enviar órdenes a los actuadores de dicho proceso a través de los correspondientes controladores y gestionar la comunicación con el usuario a través de uno o más equipos HMI.

De lo expuesto se deduce que un programa SCADA está íntimamente ligado a una unidad HMI, especialmente a las basadas en un computador industrial embebido, de tal manera que en la práctica no se sabe donde acaba uno y empieza el otro.

Para proporcionar las prestaciones indicadas anteriormente el programa SCADA está compuesto por los siguientes módulos o subprogramas:

- **Módulo de configuración**
Permite al usuario definir el entorno de trabajo del programa SCADA para adaptarlo a la aplicación particular que se quiere desarrollar.
- **Módulo de interfaz gráfica del usuario**
Proporciona al usuario las funciones de control y supervisión de la planta. El proceso se representa mediante gráficos almacenados en el computador y generados mediante el editor incorporado en el SCADA o importados desde otro computador durante la configuración del SCADA.
- **Módulo de proceso**
Ejecuta las acciones de mando preprogramadas a partir de los valores actuales de las variables adquiridas.
- **Módulo de gestión y archivo de datos**
Se encarga de almacenar y procesar ordenadamente los datos, para que otra aplicación o dispositivo pueda tener acceso a ellos.
- **Módulo de comunicaciones**
Se encarga de la transferencia de información entre la planta y el computador en el que se ejecuta el programa SCADA, y entre éste y el resto de los elementos informáticos de gestión.

De lo expuesto se deduce también que una de las características más importantes de un programa SCADA es el acceso a los datos remotos de un proceso y la gestión de las comunicaciones necesarias para implementarlo [BAIE 03] [IEEE 93] [RODR 03]. Los primeros programas SCADA, que nacieron a finales de la década de 1980, se caracterizaron por contener controladores de comunicaciones (*Drivers*) diseñados a medida para cada dispositivo de campo diferente (autómata programable, sensor y actuador inteligente, equipo HMI, etc.), tal como se indica en la figura 8.54. Esta forma de actuar generó el problema de la incompatibilidad entre los programas SCADA y los diferentes controladores porque para conectar un nuevo controlador era necesario que el fabricante del SCADA desarrollase para él un subprograma específico de control de las comunicaciones (*Driver*).

Pero, al mismo tiempo que los programas SCADA, que poseían controladores de comunicaciones (*Drivers*) específicos, fueron entrando en el mundo de la automatización industrial los fabricantes de computadores de aplicación general que diseñaron una nueva forma de gestionar documentos compuestos por elementos heterogéneos. Dicha forma constituye una nueva tecnología de comunicación entre programas de gestión, basada en la programación orientada a objetos [BOOC 07] [JACO 99]. Un ejemplo muy difundido de esta nueva tecnología de gestión es OLE (acrónimo de *Object Linking and Embedding*) [ADLE 95] [BROC 95] que permite insertar (*Embedding*) o vincular (*Linking*) distintos objetos, como por ejemplo una hoja

de cálculo y una representación gráfica (*Chart*), en diversos programas de aplicación general, como por ejemplo las herramientas ofimáticas.

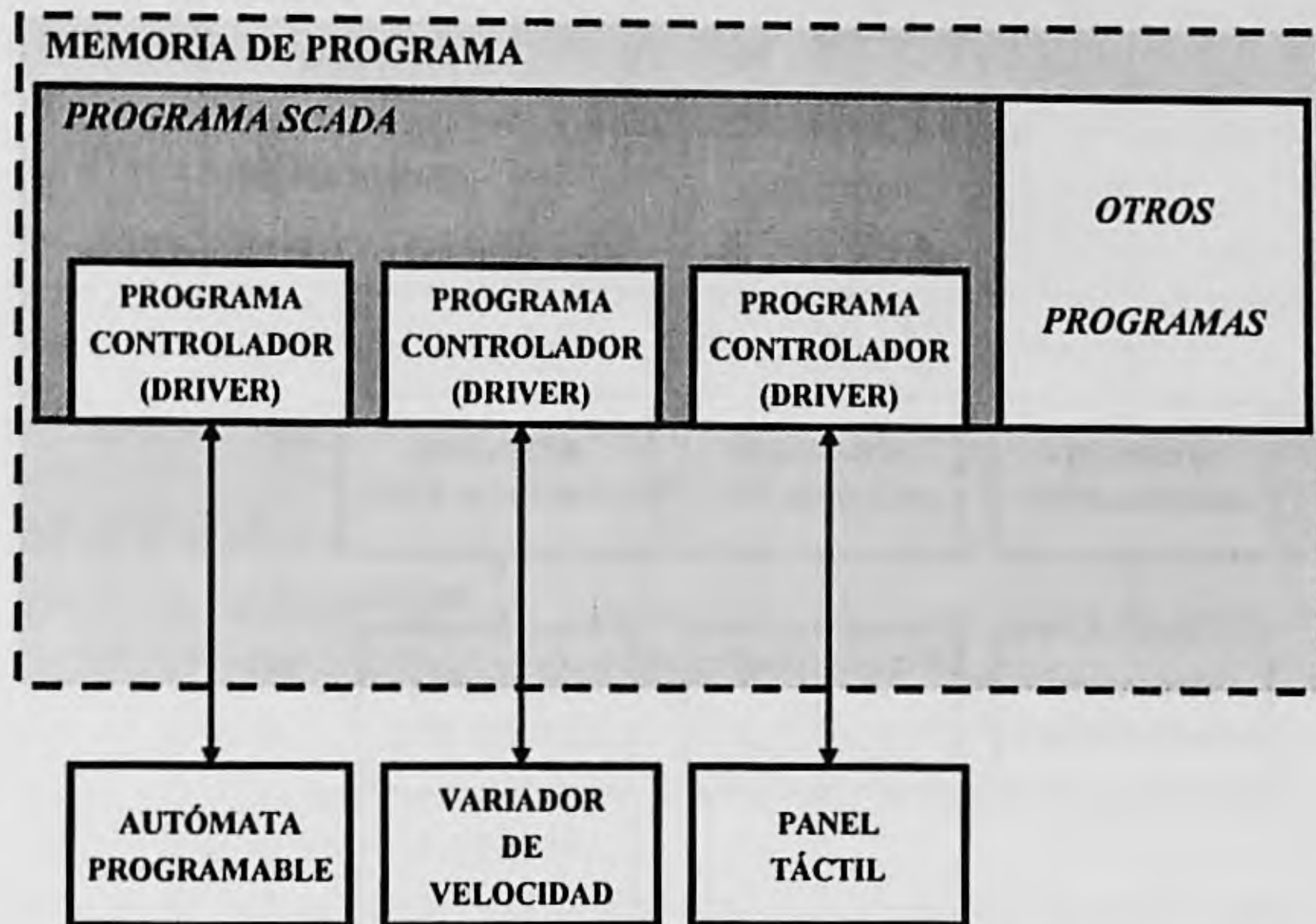


Figura 8.54. Intercomunicación sin OPC entre un programa SCADA y diferentes subsistemas de control de procesos.

El éxito que la tecnología OLE tuvo en el campo de la informática de gestión empresarial y la progresiva utilización del sistema operativo "Windows" de Microsoft como elemento básico de los programas SCADA, favoreció que, en 1994, los fabricantes de controladores y de equipos de automatización de procesos en general se pusieran de acuerdo para tratar de aplicar la tecnología OLE a la solución del problema de la incompatibilidad de las comunicaciones entre los programas SCADA y los controladores, indicada anteriormente. Para ello crearon un grupo de trabajo (*Task force*) que, en 1995, en el congreso internacional de ISA (*Instrumentation Systems and Automation Society*) publicó el primer borrador de la tecnología OPC (*OLE for Process Control*) que constituye una norma para facilitar la interconexión entre programas informáticos, especialmente programas SCADA, y distintos tipos de sistemas de control de procesos [OPCC 08]. OPC define una interfaz normalizada, que se utiliza para diseñar sistemas informáticos cliente-servidor [EVAN 95] y conseguir la comunicación entre elementos que cumplan la norma. En la figura 8.55 se representa el diagrama de bloques básico de un sistema basado en OPC. El fabricante del programa SCADA coloca en él un programa cliente OPC que puede trabajar con diversos servidores OPC, uno por cada sistema controlado por el SCADA. Los servidores OPC pueden estar situados en el mismo computador que el SCADA o en el sistema controlado por el SCADA. En el primer caso la conexión se establece a través de uno de los buses del computador (como por ejemplo el autómata programable y el variador de velocidad de la figura 8.55). En el segundo caso (pantalla táctil de la figura 8.55) la conexión se hace a través de un procesador de comunicaciones y un bus de comunicaciones, como por ejemplo Ethernet.



Figura 8.55. Intercomunicación con OPC entre un programa SCADA y diferentes subsistemas de control de Procesos.

En 1996 se creó la fundación OPC (*OPC Foundation*) [OPCF 08], constituida en la actualidad por más de 300 miembros, dedicada a hacer evolucionar OPC y asegurar la compatibilidad de las nuevas versiones.

Por otra parte, de lo que se acaba de exponer se deduce que un programa SCADA está en la práctica asociado siempre, como mínimo, con un procesador digital programable, como por ejemplo un computador industrial, un sistema de adquisición de datos que transmite al procesador digital la información procedente de un conjunto de sensores y un conjunto de actuadores. Además, en la mayoría de las instalaciones actuales, tanto los sensores como los actuadores están distribuidos espacialmente. Por ello, algunos autores [BAIE 03] [CLAR 04] [RODR 03] entienden por SCADA un sistema de control distribuido conocido por el acrónimo DCS (*Distributed Control System*) que está formado por un conjunto de elementos físicos (*Hardware*) y de programas (*Software*) interrelacionados.

El sistema físico está formado por:

- Un computador principal (*Central host*) que se denomina unidad terminal principal, a la que se conoce también por el acrónimo MTU (*Master Terminal Unit*).
- Un conjunto de unidades distribuidas de adquisición de datos, control y actuación, conocidas por el acrónimo RTU (*Remote Terminal Units*) conectadas a la MTU a través de los adecuados recursos de comunicaciones.

- Un conjunto de sensores y actuadores inteligentes conocidos por el acrónimo IED (*Intelligent Electronic Devices*) que están asociados a un procesador de comunicaciones para dotarlo de capacidad de diálogo con las unidades de adquisición de datos, control y actuación (RTU) y con el computador principal (MTU).
- Un sistema de comunicaciones a pequeñas o grandes distancias que enlaza la MTU con las RTU y los IED.

En el computador principal se ejecuta un programa al que, como se indica anteriormente, muchos fabricantes denominan SCADA, que monitoriza y controla las unidades de adquisición de datos y actuación (RTU) y los sensores y actuadores inteligentes (IED).

BIBLIOGRAFÍA

- [ADLE 95] R. M. Adler. *Emerging Standards for Component Software*. Computer, Volume 28, issue 3. IEEE Computer Society Press. 1995.
- [BAIE 03] D. Bailey & E. Wright. *Practical SCADA for Industry*. Elsevier. 2003.
- [BOOC 07] G. Booch. *Object oriented Analysis and Design with applications*. Addison Wisley. 2007.
- [BROC 95] K. Brockschmidt. *Inside OLE*. Microsoft Programming Series. 1995.
- [CLAR 04] G. Clarke, D. Reynders & E. Wright. *Practical Modern SCADA Protocols: DNP3, 60870.5 and related systems*. Elsevier. 2004.
- [EVAN 95] C. Evans, D. Lacey & D. Harvey. *Client-server: A Handbook of Modern Computer System Design*. Prentice Hall. 1995.
- [IEEE 93] IEEE Std 999-1992. *IEEE Recommended Practice for Master/Remote Supervisory Control and Data Acquisition (SCADA) Communications*. IEEE. 1993.
- [ISO 08] ISO 9241-400. *Ergonomics of human-system interaction. Part 400: Principle and requirements for physical input devices*. ISO. 2008. Incluida en Junio de 2008 en la página: www.iso.org
- [JACO 99] I. Jacobson & J. Rumbaugh. *The Unified Software Development Process*. Prentice Hall. 1999.
- [MAND 95] E. Mandado, P. Mariño, A. Lago. *Instrumentación Electrónica*. Editorial Marcombo. 1995.
- [MAND 08a] E. Mandado & Y. Mandado. *Sistemas electrónicos digitales. Apartado 3.7.4.3 del capítulo 3: Multiplexores realizados con interruptores*. 9ª edición. Editorial Marcombo. 2008.
- [MAND 08b] E. Mandado & Y. Mandado. *Sistemas electrónicos digitales. Apartado 7.7.3 del capítulo 7: Memorias cola (FIFO)*. 9ª edición. Editorial Marcombo. 2008.
- [MAND 08c] E. Mandado & Y. Mandado. *Sistemas electrónicos digitales. Apartado 7.6.1.2.5 del capítulo 7: Memorias activas volátiles de acceso múltiple*. 9ª edición. Editorial Marcombo. 2008.
- [NIEL 93] J. Nielsen. *Usability Engineering*. Academic Press. London. 1993

- [OPCC 08] OPCconnect. *History of OPC*. Incluida en Junio de 2008 en la página: www.opcconnect.com/history.com
- [OPCF 08] OPCfoundation. *Interoperability in automation*. Incluida en Junio de 2008 en la página: www.opcfoundation.org
- [PALL 98] R. Pallás. *Sensores y acondicionadores de señal*. Editorial Marcombo. 3ª edición. 1998.
- [PERE 03] M. A. Pérez García y otros. *Instrumentación electrónica*. Editorial Thomson Learning Paraninfo. 2003.
- [RODR 03] A. R. Penín. *Sistemas SCADA*. 2ª edición. Editorial Marcombo. 2007.
- [SIEM 02] Siemens. *Sistema de automatización S7-200, Manual del sistema*. C79000-G7078-C230-02.
- [STOC 79] M. Stockl & K. H. Winterling. *Técnica de las medidas eléctricas*. Editorial Labor. 1979.

CAPÍTULO 9

El autómata programable y las Comunicaciones Industriales

9.1 Introducción

Tal como se indica en el capítulo 1, los primeros autómatas programables, comercializados al final de la década de 1960, realizaban el control de una máquina o de un proceso productivo sencillo. Para ello, las variables todo/nada (*On/Off*) del proceso se convertían en variables eléctricas binarias que se conectaban al autómata programable a través de terminales independientes tal como se indica en la figura 9.1. Pero la elevación de la complejidad de los procesos productivos hizo que en ellos se tuviesen que utilizar varias máquinas, cada una de ellas especializada en la realización de una o más operaciones determinadas.

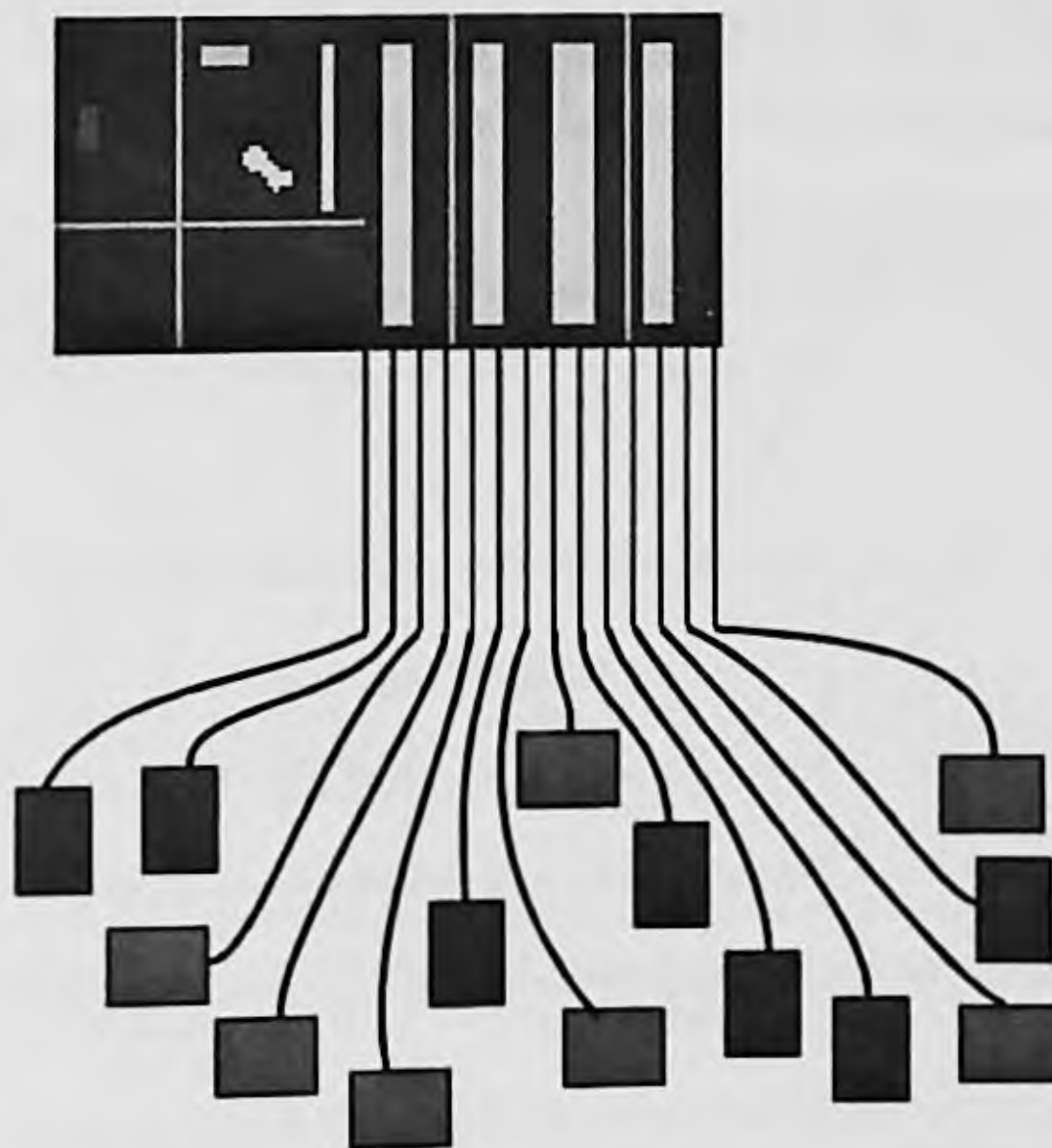


Figura 9.1. Autómata programable que se conecta con los diversos sensores y actuadores a través de terminales independientes (Cortesía de Siemens).

Además cada máquina estaba, en general, fabricada por una empresa diferente y llevaba incorporado un autómata programable distinto. Se generaron entonces las llamadas “Islas de Automatización”, denominadas así porque los sistemas electrónicos de control carecían de capacidad de diálogo entre ellos debido a lo cual las empresas industriales comenzaron a tener problemas para gestionar adecuadamente su proceso productivo. La situación así creada motivó el interés por dotar a los sistemas electrónicos de control de capacidad de comunicación entre ellos y con el entorno que les rodea y dio lugar a un área de las Telecomunicaciones dedicada a estudiar la transmisión de información entre circuitos y sistemas electrónicos utilizados para llevar a cabo las tareas de control de productos y procesos. Dicha área recibe en la actualidad la denominación de “Comunicaciones Industriales” (*Industrial Communications*) y era prácticamente inexistente antes de 1970. Su rápido desarrollo es el resultado (Figura 9.2) de la confluencia de dos circunstancias:

- La elevación de la complejidad de los procesos y los productos industriales que dejaron de poder ser controlados y gestionados de forma manual para pasar a serlo mediante un procesador digital.
- La evolución de las comunicaciones digitales (*Data Communications*) como resultado del avance de la Microelectrónica que permitió ampliar la capacidad de los procesadores digitales y crear mecanismos para establecer la comunicación a distancia entre los mismos [ELAH 01] [FORO 07] [STAL 07].

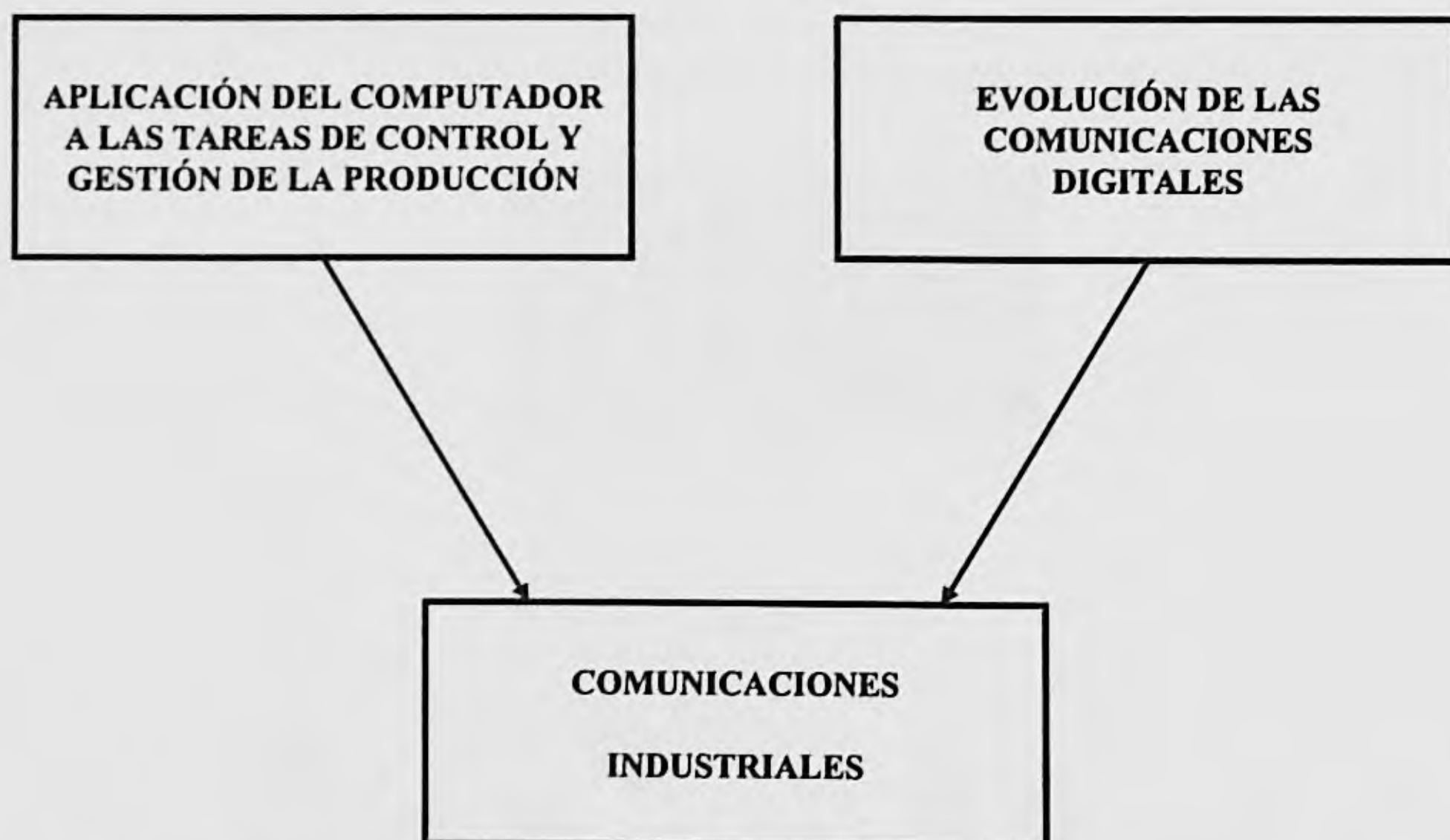


Figura 9.2. Descripción conceptual de las Comunicaciones Industriales.

Por ello, para comprender las Comunicaciones Industriales es necesario estudiar tanto los conceptos ligados a los procesos productivos y al papel que el computador desempeña en ellos en la actualidad, como las comunicaciones digitales que permiten que los computadores se comuniquen entre sí. Los primeros se estudian a continuación y las comunicaciones digitales se analizan en el apéndice 1.

9.2 El computador y el ciclo del proceso de un producto

9.2.1 Conceptos generales

La técnica, cuyo origen se pierde en la noche de los tiempos, ha tenido un desarrollo parejo al del ser humano sobre la tierra [MUMF 34]. Durante las primeras etapas de ese desarrollo, hasta finales del siglo XVIII, las diferentes áreas de la técnica evolucionaron individualmente, pero a partir de esa fecha comenzó un proceso de integración que ha llegado hasta nuestros días. En la figura 9.3 se representa gráficamente dicho proceso de integración que se puede considerar dividido en tres etapas:

- La integración mediante la Cinemática, que se inició en 1772 con la invención de la máquina de vapor y se desarrolló paulatinamente hasta 1871 mediante su utilización como elemento motriz de un eje central que accionaba un conjunto de máquinas-herramienta especializadas.
- La integración mediante la Electricidad, que se inició en 1871 con la invención de la dinamo y se desarrolló a lo largo del siglo XX mediante un conjunto de innovaciones basadas en la Electricidad que dieron lugar a las máquinas accionadas por motores eléctricos controlados por sistemas electrónicos y situadas en línea para efectuar una secuencia de operaciones.
- La integración mediante la Información que, aunque formalmente se inició en la década de 1960 con los primeros terminales remotos de los computadores, se puede considerar que se hizo realidad a partir de 1971 con la comercialización de los primeros microprocesadores y especialmente en la década de 1980 con los primeros computadores personales que permitieron implementar redes de sistemas electrónicos digitales programables (computadores, autómatas programables, sistemas de control numérico, etc).

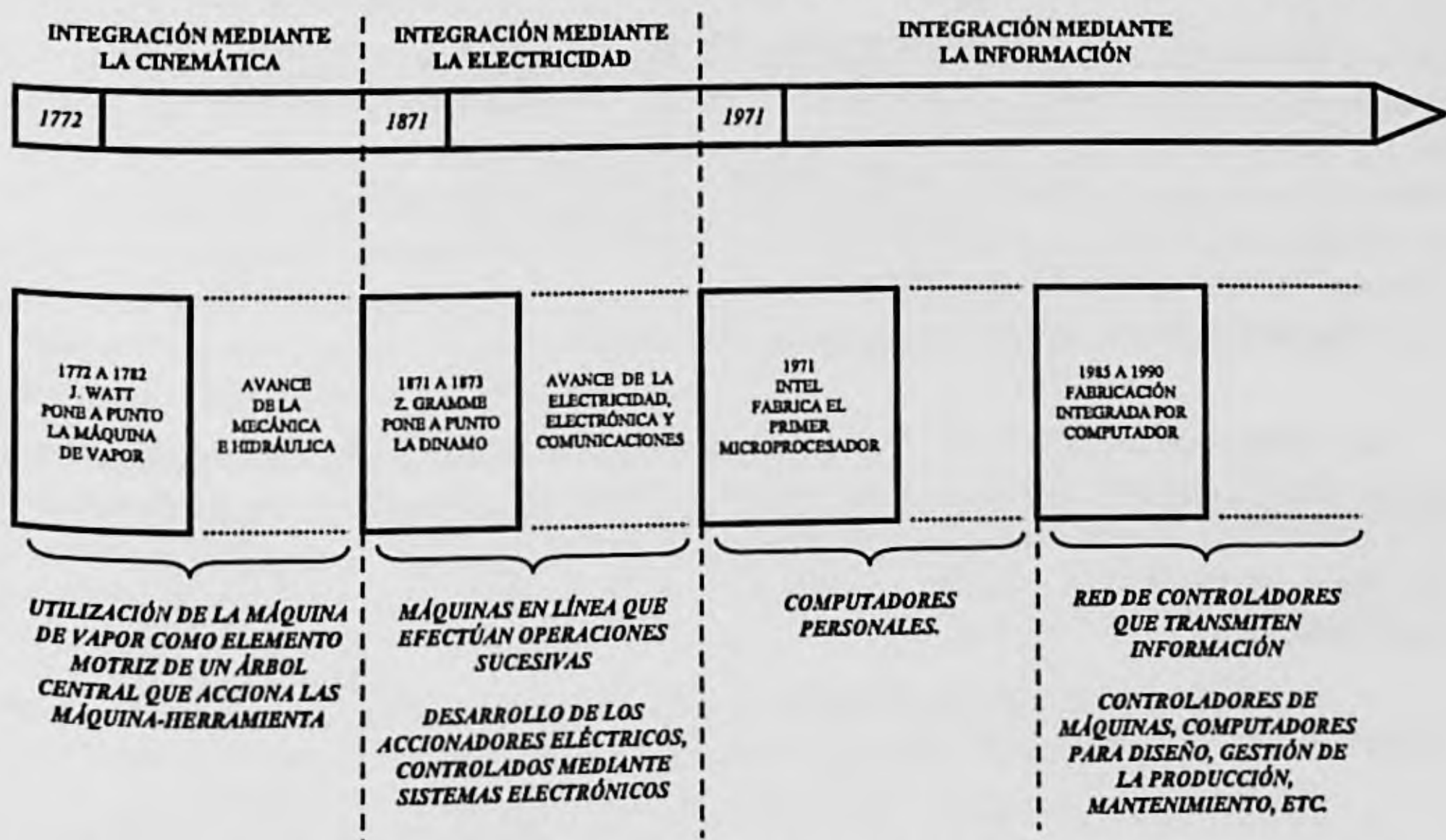


Figura 9.3. Evolución de la Mecanización: De la rueda a la producción integrada por computador.

Pero no solo los artefactos técnicos se fueron integrando entre sí, sino que, como dijo Des-sauer [DESS 58], los objetos técnicos, como por ejemplo un reloj o un microscopio, se convirtieron en una mercancía, palabra que no es una categoría técnica porque frente al valor técnico de un objeto, que es su valor específico no canjeable, está el valor de canje, el precio, a cuya determinación contribuye la oferta y la demanda en el mercado. Fue precisamente esa demanda a escala mundial la que hizo que decreciesen los lotes de fabricación y los plazos de industrialización, a la vez que se incrementaron la variedad y complejidad de los productos a desarrollar.

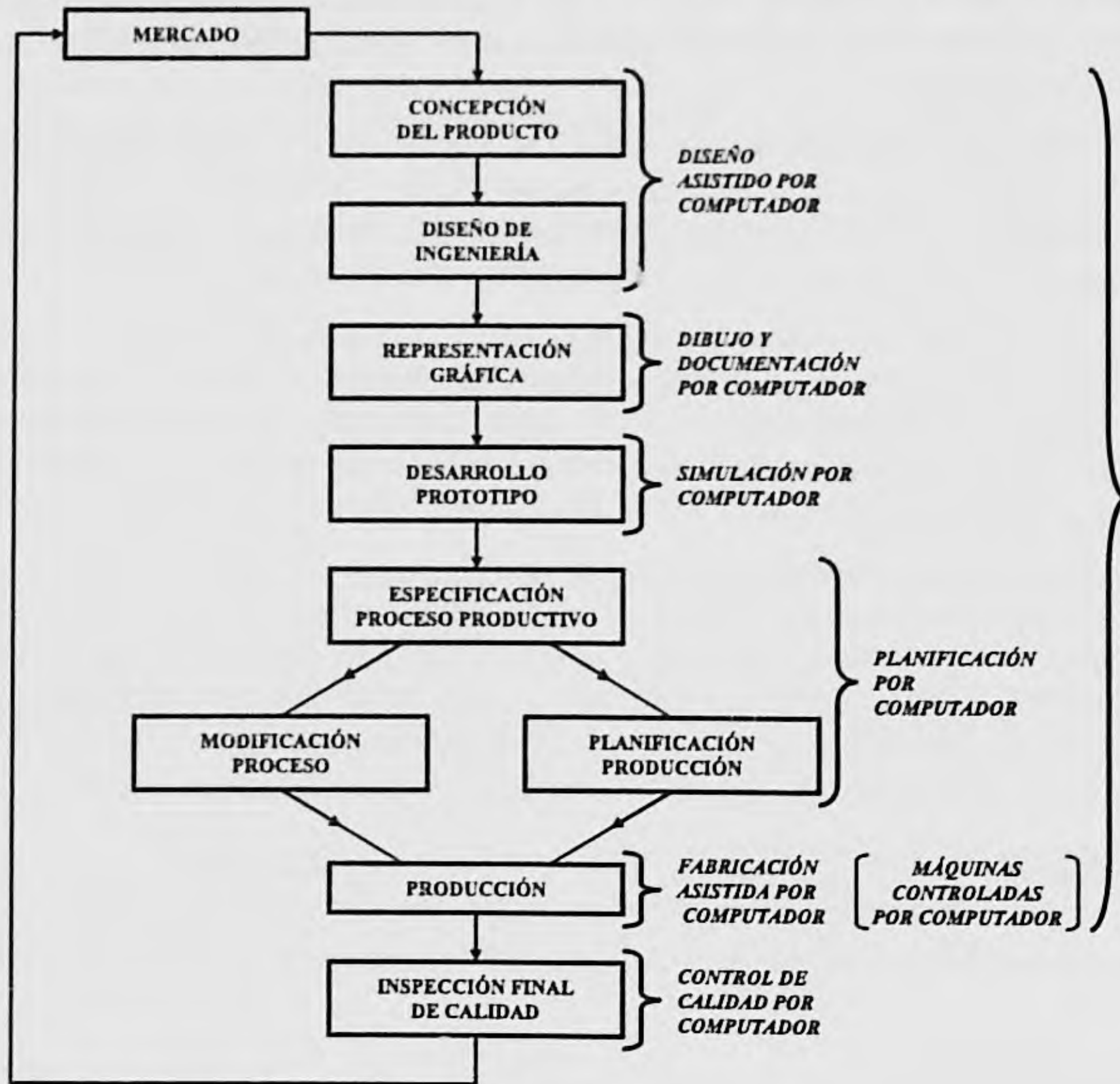


Figura 9.4. Ciclo completo del proceso de un producto.

Se produjo así la creación del ciclo de un producto, que se representa en la figura 9.4 en la que se observa que todas las tareas que forman parte del mismo se pueden llevar a cabo mediante la utilización del computador para automatizar el diseño y la fabricación y lograr dos grandes objetivos que forman parte de la estrategia de todas las empresas fabricantes de productos industriales:

- La utilización de métodos de diseño de sistemas complejos que garanticen el correcto funcionamiento del prototipo y su producción en serie.
- La ejecución de las tareas de producción mediante máquinas que, no solo sustituyen a veces a las personas, sino que permiten la ejecución de tareas que el ser humano, debido a sus limitaciones, es incapaz de acometer.

Mediante la automatización del diseño y la fabricación se pueden obtener las siguientes ventajas competitivas:

- Una mayor fiabilidad del proceso productivo con la consiguiente elevación de la calidad del producto.
- La reducción de los costes de desarrollo y fabricación.
- La reducción del tiempo de desarrollo (*Time to market*).
- Un menor consumo de todo tipo de recursos en el proceso productivo.
- La mejora de la calidad total de la empresa.

La automatización del diseño y la fabricación se llevan a cabo mediante el siguiente conjunto de técnicas:

- El diseño asistido por computador conocido como CAD (acrónimo de *Computer Aided Design*).
- La ingeniería asistida por computador o gestión del ciclo de vida del producto conocidas por el acrónimo CAE (*Computer Aided Engineering*) y también por el acrónimo PLM (*Product Life cycle Management*).
- La fabricación asistida por computador conocida como CAM (acrónimo de *Computer Aided Manufacturing*).

A continuación se describe brevemente cada una de ellas. Para un estudio más detallado se remite al lector a la bibliografía indicada al final de este capítulo [GROO 84] [RAPH 03] [REHG 05].

9.2.2 Diseño asistido por computador

El diseño asistido por computador (CAD) es un conjunto de técnicas que utilizan el computador con el objetivo de generar la información necesaria para fabricar un producto a partir de las especificaciones de sus características de funcionamiento [GROO 84] [REHG 05]. Su aplicación depende del área tecnológica, pero su importancia es cada vez más grande en todas ellas, y en algunas, como por ejemplo la Microelectrónica, resulta imprescindible debido a la imposibilidad de implementar un prototipo. La actividad básica del diseño asistido por computador es la descripción del sistema que implica la creación de lenguajes normalizados para facilitar su utilización por los técnicos.

Mediante el diseño asistido por computador se logra:

- Mejorar la calidad de los productos, porque permite desarrollar alternativas y solventar los problemas en las etapas iniciales del proceso de diseño.
- Reducir el tiempo de diseño, lo cual disminuye el coste y el tiempo que se tarda en lanzar un producto.
- Reducir los costes de fabricación, porque facilita los cambios y se puede, en algunos casos, combinar con la tecnología de grupos descrita en el apartado 9.2.4.3.
- Mejorar la gestión de las bases de datos.
- Facilitar la reutilización (*Reusability*) de los diseños.

Un ejemplo típico del diseño asistido por computador es el desarrollo del programa de control de un autómatas programable, llevado a cabo mediante una unidad de programación, que es un computador que permite editar el programa en, por ejemplo, uno de los lenguajes descritos en los capítulos 2 y 3 y enviarlo al autómatas programable, una vez traducido al lenguaje máquina del mismo, a través de un canal de comunicación.

9.2.3 Ingeniería asistida por computador

La ingeniería asistida por computador (*CAE or PLM*) es el conjunto de técnicas que utilizan el computador para analizar el resultado de un diseño, facilitar al máximo (optimizar) la fabricación (*Optimize Manufacturability*), optimizar las prestaciones y los costes totales del producto final y facilitar el diagnóstico de las averías y la reparación de las mismas, a lo largo del ciclo de vida útil del producto. Utiliza, entre otras, la información de la base de datos, generada por un programa de diseño asistido por computador, para analizar las características funcionales del sistema objeto de diseño, simular su comportamiento y proporcionar la información que necesita el sistema de fabricación [REHG 05] [SOUT 95].

La simulación por computador consiste en utilizar la descripción del sistema como entrada de un programa de computador que hace que éste último se comporte igual que él. Se considera que es una actividad que forma parte al mismo tiempo del diseño y de la ingeniería asistida por computador y permite llevar a cabo las pruebas necesarias para garantizar el correcto funcionamiento de un producto sin necesidad de implementar un prototipo. Para generalizarla es necesario normalizar la forma de describir los sistemas.

9.2.4 Fabricación asistida por computador

9.2.4.1 Introducción

La fabricación asistida por computador (CAM) consiste en un conjunto de técnicas que tienen como objetivo elevar la productividad de los procesos de fabricación mediante la sustitución de las manos del ser humano por sistemas físicos que combinan la Tecnología Electrónica con otras como la Teoría de Control, la Mecánica, las Máquinas Eléctricas, la Neumática, la Hidráulica, etc., cuyo conjunto suele conocerse bajo la denominación de “Automatización de la Producción o Automatización Industrial”.

Se trata, por lo tanto, de una técnica multidisciplinar que se caracteriza por:

- Necesita un conocimiento profundo del proceso productivo.
- Puede ser para la industria la mejor solución para elevar su rentabilidad y garantizar su competitividad.
- No siempre implica la mayor automatización posible, porque la automatización supone una inversión en activos fijos que si es elevada conduce a un considerable incremento de los costes fijos y también puede suponer un aumento de los costes de mantenimiento y una disminución de la flexibilidad de los recursos.

Aunque algunas de las técnicas que forman parte de la fabricación asistida por computador son específicas de cada tipo de producto concreto, existen técnicas comunes a todos los proce-

Los productivos que están basadas en la utilización de los sistemas electrónicos en general y del computador en particular. Las más importantes son:

- Las clases de automatización y sus características.
- La planificación de los productos a fabricar (Tecnología de grupos).
- Los sistemas electrónicos de control.
- Los sistemas de manipulación de elementos.
- Los sistemas de fabricación flexible.
- Los equipos de visualización y actuación (HMI) y los sistemas de supervisión y adquisición de datos (SCADA).

En sucesivos apartados se analiza cada una de estas técnicas, excepto los equipos de visualización y actuación (HMI) y los sistemas de supervisión y adquisición de datos (SCADA) que, por ser equipos de interfaz, se describen en los apartados 8.3.2 y 8.3.3 del capítulo 8 respectivamente.

9.2.4.2 Clases de automatización y sus características

Se suele definir la Automatización (*Automation*) como la aplicación de la tecnología para llevar a cabo procesos que se autocomprueban y se autocorrigien [SOUT 95]. Es, por lo tanto, una combinación de la Tecnología Mecánica, la Tecnología Eléctrica, la Teoría de Control, la Electrónica y los Computadores para implementar y controlar los sistemas de producción.

En relación con el control de los procesos de fabricación, la automatización se puede clasificar en cuatro grandes clases que se describen a continuación.

9.2.4.2.1 Automatización fija

Se entiende por automatización fija (*Fixed automation*) un sistema de fabricación en el que la secuencia de las operaciones está fijada por la configuración de los equipos que lo forman.

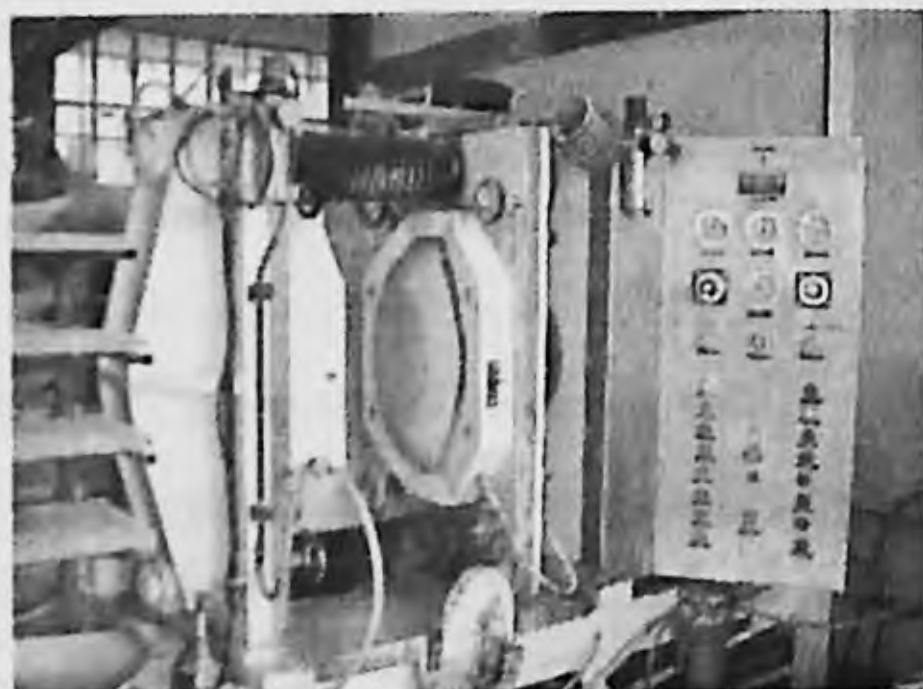
Se caracteriza por:

- Está constituida por una secuencia sencilla de operaciones.
- Necesita una inversión elevada en equipos especializados.
- Posee elevados ritmos de producción.
- Es muy inflexible, en general, para acomodarse a los cambios de los productos

La automatización fija se justifica económicamente tanto en los procesos de producción continuos como en los discretos (Figura 9.5). En los procesos continuos, denominados en general simplemente procesos (*Processes*), como por ejemplo las plantas químicas, es donde ha tenido una mayor implantación que continúa en la actualidad combinándola con la automatización programable mediante la utilización del control por computador que da lugar al denominado DCS (acrónimo de *Distributed Control System*) porque se lleva a cabo mediante un conjunto de sistemas electrónicos de control interrelacionados entre ellos a través de una red compleja de comunicaciones.



a)



b)

Figura 9.5. Ejemplos de automatización fija: a) Máquina de fabricación de platos; b) Máquina de colado a presión (Cortesía de Sargadelos, S.A.).

En los procesos discretos, conocidos como procesos de fabricación (*Manufacturing Processes*), la automatización fija se inició a principios del siglo XX y un ejemplo de ello son las líneas de montaje mecanizado formadas por un conjunto de puestos de trabajo operados manualmente y enlazados por una cinta transportadora. En la actualidad ha sido sustituida por la automatización flexible descrita más adelante.

9.2.4.2.2 Automatización programable

La automatización programable (*Programmable automation*) se identifica con los sistemas de fabricación en los que el equipo de producción está diseñado para ser capaz de cambiar la secuencia de operaciones a fin de adaptarse a la fabricación de productos diferentes (Figura 9.6). La secuencia de producción se controla mediante un programa, que es un conjunto de instrucciones que se pueden cambiar para fabricar un nuevo producto.

Se inició con las máquinas-herramienta con control numérico, conocidas por las siglas NC (acrónimo de *Numerical Control*), cuyo primer prototipo se realizó en 1952, y los robots industriales, aunque sus orígenes se remontan al telar de Jacquard en 1801.

Se caracteriza por:

- Una gran inversión en equipos de aplicación general como por ejemplo sistemas de control numérico.
- La necesidad de cambiar el programa y la disposición física de los elementos de las máquinas para cada lote de productos distintos.
- La existencia de un período de preparación previo a la fabricación de cada lote de productos distintos.

De lo expuesto se deduce que la automatización programable es adecuada para la fabricación por lotes (*Batch production*) y no proporciona suficiente flexibilidad para realizar cambios en la configuración del producto.

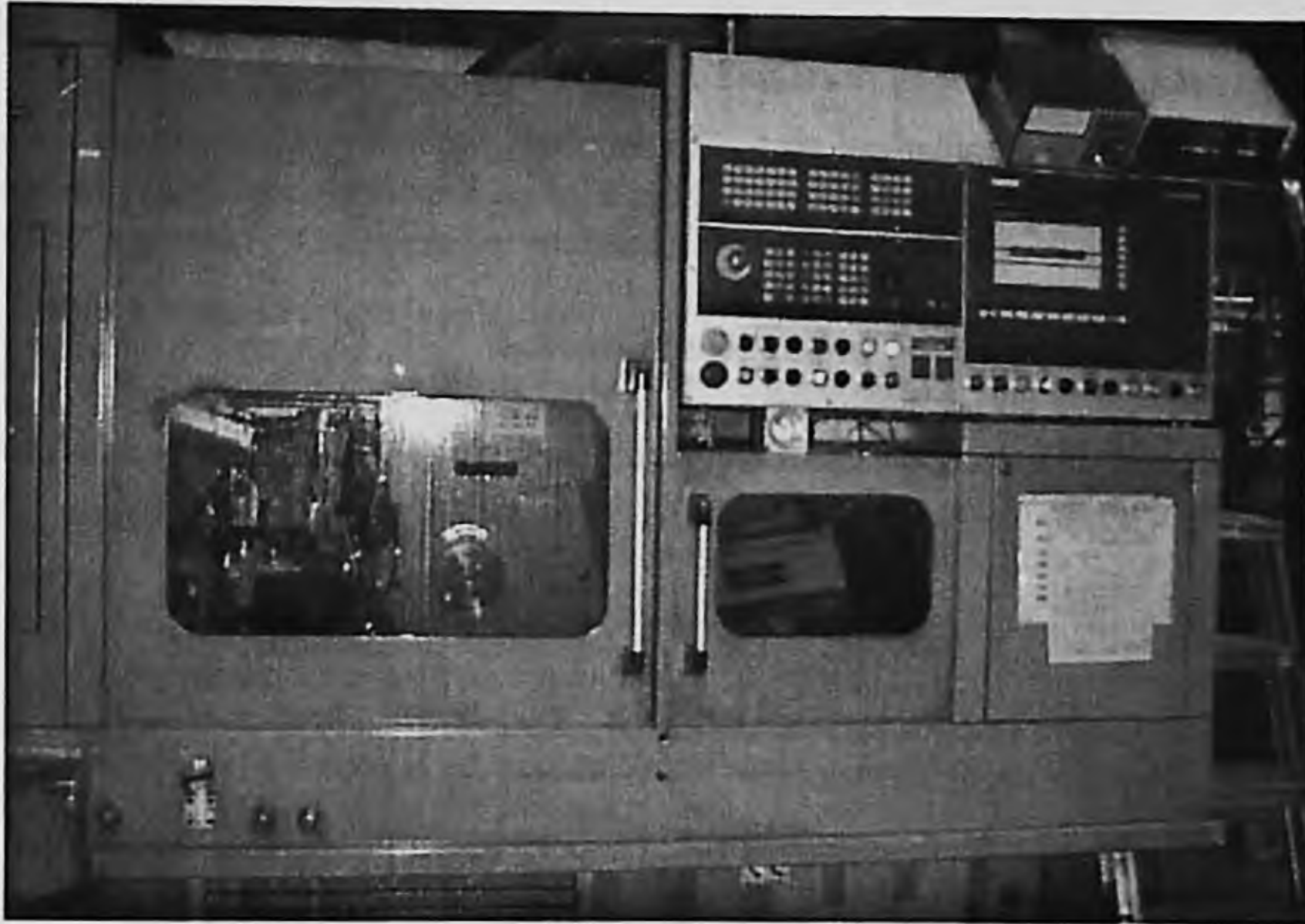


Figura 9.6. *Ejemplo de automatización programable. Máquina de fabricación de manguetas de transmisión controlada por un sistema de control numérico (Cortesía de GKN Driveline Vigo).*

9.2.4.2.3 Automatización flexible

La automatización flexible (*Flexible automation*) es una extensión de la automatización programable que da como resultado sistemas de fabricación en los que no solo se pueden cambiar los programas sino que además se puede cambiar la relación entre los diferentes elementos que los constituyen.

La automatización flexible ha dado lugar a los sistemas de fabricación flexibles que se estudian en el apartado 9.2.4.6.

9.2.4.2.4 Automatización integrada

Tal como se indica anteriormente, la automatización de la producción permitió inicialmente la fabricación en grandes lotes de un conjunto reducido de productos a lo largo de todo su ciclo de vida. Pero el aumento de la capacidad de producción propició nuevas estrategias comerciales que dieron como resultado la diversificación de los productos y el aumento de la variedad y complejidad de sus componentes, lo cual fue complicando paulatinamente la gestión de la empresa industrial (control de almacenes, planificación de la producción, etc). El resultado de todo ello fue la necesidad de disminuir el tamaño de los lotes de fabricación y de los plazos de industrialización, lo que impulsó la necesidad de disponer de sistemas productivos más versátiles y flexibles que los empleados anteriormente.

Surgió así la automatización integrada (*Integrated automation*), que es un sistema de fabricación que integra el diseño asistido por computador (CAD), la ingeniería asistida por computador (CAE) y la fabricación asistida por computador (CAM) con la verificación, la co-

mercialización y la distribución. Suele recibir el nombre de CIM (acrónimo de *Computer Integrated Manufacturing*) y se estudia en el apartado 9.2.5.

Dado que en ella se automatizan, de forma coordinada, todas las tareas que forman parte del ciclo completo del proceso de un producto (Figura 9.4) se la conoce también por las siglas TIA (acrónimo de *Totally Integrated Automation*).

9.2.4.3 Planificación de los productos a fabricar

La planificación de los productos a fabricar se lleva a cabo mediante la tecnología de grupos (*Group technology*) que consiste en agrupar los productos en familias a cada una de las cuales se asigna un grupo de máquinas. Su objetivo es conseguir alta repetibilidad con bajos volúmenes de producción y consta de las siguientes fases:

- Creación de familias en base al tamaño, forma, prestaciones, rutas de proceso, etc. Esta fase tiene como objetivo determinar un conjunto de productos cuyas necesidades de fabricación sean similares para minimizar el cambio de máquinas o el inicio de una nueva serie de productos diferentes.
- Organización de las máquinas necesarias para desarrollar los procesos básicos. Esta organización se realiza para llevar a cabo los procesos básicos en áreas separadas llamadas células o celdas de fabricación. Se logra así que las máquinas de cada célula necesiten tan sólo ajustes menores para adaptarse a las necesidades de los diferentes lotes de cada familia.
- Definición e implantación de un sistema de clasificación y codificación de los productos.

La tecnología de grupos ha sido favorecida por diversos factores entre los que cabe citar:

- La complejidad de muchas tecnologías o combinaciones de ellas, que han hecho proliferar el número de artículos con diferentes prestaciones, lo cual produce una reducción del tamaño de los lotes.
- La necesidad creciente de trabajar con una mayor variedad de materiales diferentes.
- El mayor peso del coste de los materiales en el coste total de un producto como resultado de la mayor eficacia del sistema productivo. Este hecho hace que se tengan que buscar formas de reducir la tasa de piezas defectuosas y los residuos.

La tecnología de grupos produce reducción de costes tanto en el diseño como en la fabricación de productos.

La utilización de un buen sistema de clasificación y codificación facilita el diseño porque proporciona:

- Información sobre piezas similares.
- La posibilidad de crear una base de datos eficiente que contenga información precisa para el diseño de productos, lo cual da lugar a una elevada productividad.
- La normalización de los diseños, lo cual evita la duplicación de tareas.
- La creación de familias de piezas.
- La incorporación de los cambios de diseño a los procesos de fabricación.

Por otra parte, la incorporación de la tecnología de grupos en la fabricación da como resultado:

- La simplificación de las rutas de fabricación, con lo cual se reduce el tiempo que cada lote de pedido está en el taller y se disminuyen, o incluso se eliminan, las colas de artículos que están a la espera de ser procesados.
- Mejor utilización de la maquinaria.
- Mejor distribución de la planta lo que reduce los costes de transporte de materiales.
- Mayor eficiencia en el aprovechamiento de las materias primas.

Además, al utilizar un buen sistema de clasificación y codificación se facilita:

- El desarrollo de un sistema de planificación asistida por computador conocido por las siglas CAPP (acrónimo de *Computer Aided Planning Process*), lo cual simplifica el proceso de planificación y hace que sea más sencillo de seguir y controlar.
- Una mayor eficacia de la gestión de operaciones.
- El desarrollo de rutinas de programación reutilizables para cada familia de piezas, lo cual proporciona la utilización de herramientas y de programas de control numérico.

Pero la tecnología de grupos no está exenta de inconvenientes, entre los que cabe citar:

- La instalación del sistema de clasificación y codificación consume mucho tiempo y tiene, en general, un coste elevado.
- La implementación suele resultar muy compleja debido a que no existen enfoques normalizados, ni tampoco unas normas sobre el propio concepto de tecnología de grupos.
- La agrupación de máquinas no siempre tiene como consecuencia que todas las de un grupo se utilicen eficientemente. Esto suele ser debido a que no se aprovecha al máximo el conjunto aunque los costes fijos totales de las distintas células de fabricación sean menores.
- En ocasiones, la redistribución en planta tiene un coste elevado.
- Debido a que hay que cambiar la forma y los métodos de trabajo, es posible que los empleados opongan cierta resistencia, como consecuencia de la adopción de la tecnología de grupos.

Por todo ello se puede afirmar que la implantación de la tecnología de grupos necesita un apoyo fuerte y continuado de la alta dirección de la empresa.

En este apartado se realiza una descripción somera de la tecnología de grupos. Para un estudio más profundo se remite al lector a la bibliografía [FERR 88] [HAM 85].

9.2.4.4 Sistemas electrónicos de control

El aumento de la complejidad de los sistemas productivos mecánicos y eléctricos, así como de los procesos químicos durante la primera mitad del siglo XX, generó la necesidad de automatizarlos mediante sistemas electrónicos programables cuyo comportamiento pudiese adaptarse fácilmente a las características cambiantes de aquellos. Esto dio lugar al desarrollo de los siguientes tipos de sistemas electrónicos de control:

- Los sistemas de control numérico.
- Los autómatas programables.
- Los computadores industriales.
- Los sistemas de control de procesos continuos.

Aunque cada uno de ellos tiene características específicas que los diferencian de los demás, todos poseen la característica común de ser sistemas de control en tiempo real (*Real time control systems*), que se definen (según el diccionario Oxford sobre computación) como “Cualquier sistema de control en el que el tiempo que tarda en producirse una determinada salida es significativo, lo cual ocurre habitualmente porque las variables de entrada son parámetros de algún movimiento o de los cambios en el valor de algunas variables de entrada del mundo físico, con el que están relacionadas también sus variables de salida”. Esto hace que el valor del intervalo de tiempo que transcurre entre la aparición de una determinada situación de una o más variables de entrada (como por ejemplo el cambio de estado de una variable lógica de entrada), y la generación de la correspondiente respuesta de salida deba ser inferior a un determinado valor para que el comportamiento del sistema pueda ser considerado correcto.

De lo expuesto se deduce que el principal aspecto que distingue a un sistema de control en tiempo real de otro en el que el tiempo no es especialmente significativo, es que su correcto funcionamiento no depende solo del resultado de sus cálculos sino también del tiempo que tarda en ejecutarlos.

A continuación se describen los diferentes sistemas electrónicos de control antes citados.

9.2.4.4.1 Sistemas de control numérico

Los sistemas de control numérico, conocidos por las siglas NC (acrónimo de *Numerical Control*) son sistemas electrónicos programables que controlan los movimientos de una máquina-herramienta. La unión de ambos da lugar a una unidad autónoma de mecanizado que es un conjunto mecánico con accionamiento propio que, por sí solo o en unión de otros conjuntos similares, realiza operaciones de mecanizado sin necesidad de intervención de un operario.

Los primeros sistemas de control numérico estaban formados por un sistema electrónico analógico de control en bucle cerrado en el que la señal de referencia la generaba un computador a través de un convertidor digital-analógico (Figura 9.7).

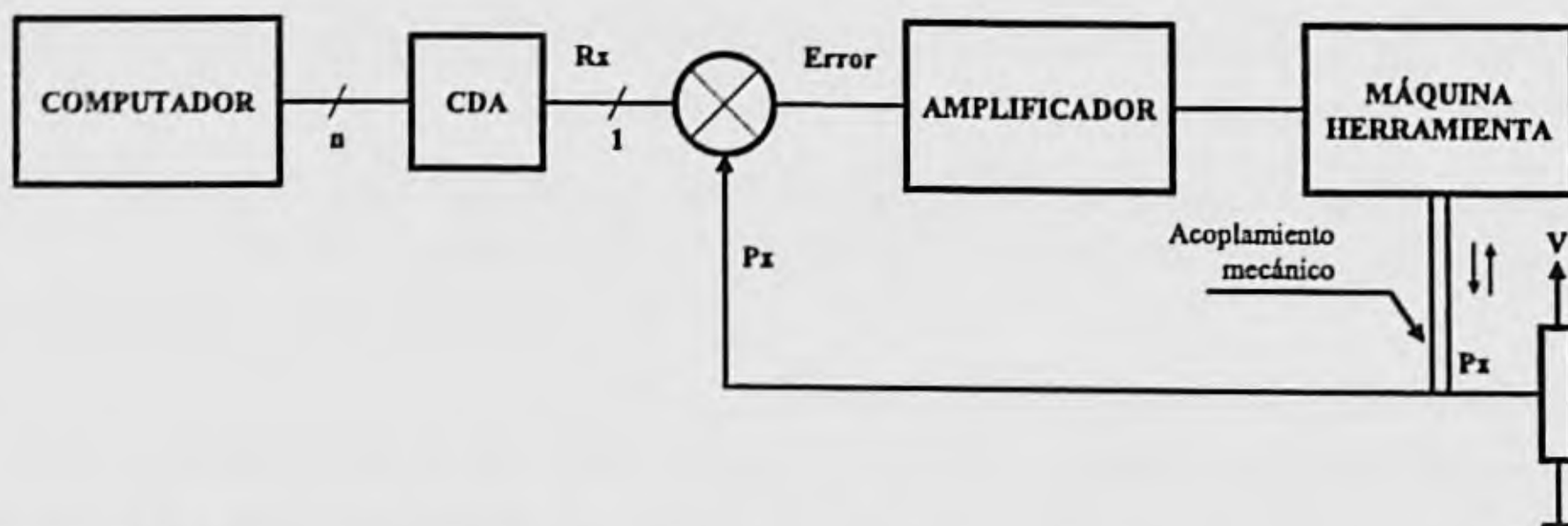


Figura 9.7. Esquema de bloques de un sistema electrónico de control numérico (NC) formado por un sistema de control de posición analógico cuya señal de referencia la genera un computador digital.

El progreso de la Electrónica propició el desarrollo en el Instituto Tecnológico de Massachusetts en la década de 1950 del primer sistema de control numérico en el que el computador cierra el bucle de control mediante la ejecución de un algoritmo que compara la información proporcionada por un codificador incremental de posición con la calculada mediante un programa de interpolación del propio computador (Figura 9.8). La utilización de un computador hizo que los sistemas de control numérico se denominasen CNC (acrónimo de *Computer Numerical Control*).

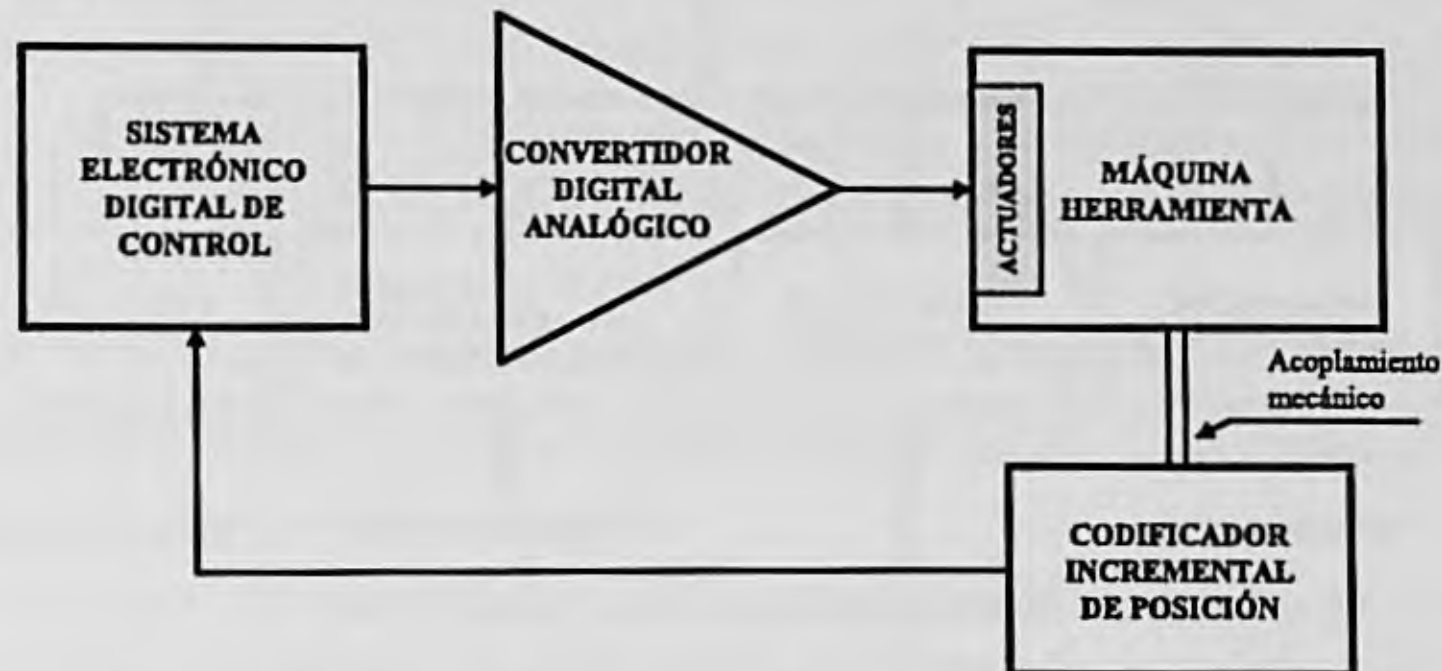


Figura 9.8. Esquema de bloques de un sistema electrónico de control numérico realizado mediante un computador (CNC) y un codificador incremental de posición.

De forma paulatina, los sistemas CNC fueron incorporando, además de las funciones de control de la máquina-herramienta, otras que mejoran las prestaciones como por ejemplo la adaptación a las condiciones de trabajo, el diagnóstico de averías, etc.

El desarrollo de los circuitos integrados a partir de 1980 propició que los sistemas de control numérico CNC evolucionasen siguiendo dos grandes líneas en función del campo de aplicación que se describen seguidamente.

Sistemas CNC de gran capacidad

Son sistemas CNC que realizan el posicionamiento dinámico de los ejes con elevada precisión y que están formados por dos módulos. (Figura 9.9):

- Un módulo de órdenes (*Command module*) que está implementado actualmente con un microprocesador de 32 o 64 bits y posee una interfaz gráfica fácil de utilizar.
- Un módulo de control numérico (*Numerical Control Module*) propiamente dicho, que controla la máquina-herramienta a partir de las informaciones que genera el módulo de órdenes.

Existe una gran variedad de sistemas CNC de gran capacidad que se caracterizan por poseer las siguientes características generales (Figura 9.10):

- Controlan tres ejes o más.
- Tienen un panel inteligente de operador asociado a una placa de computador que constituye un módulo de comunicación con el operador (*Human Machine Interface*) fácil de manejar.
- Se acoplan a la unidad central de un autómata programable de elevadas prestaciones.

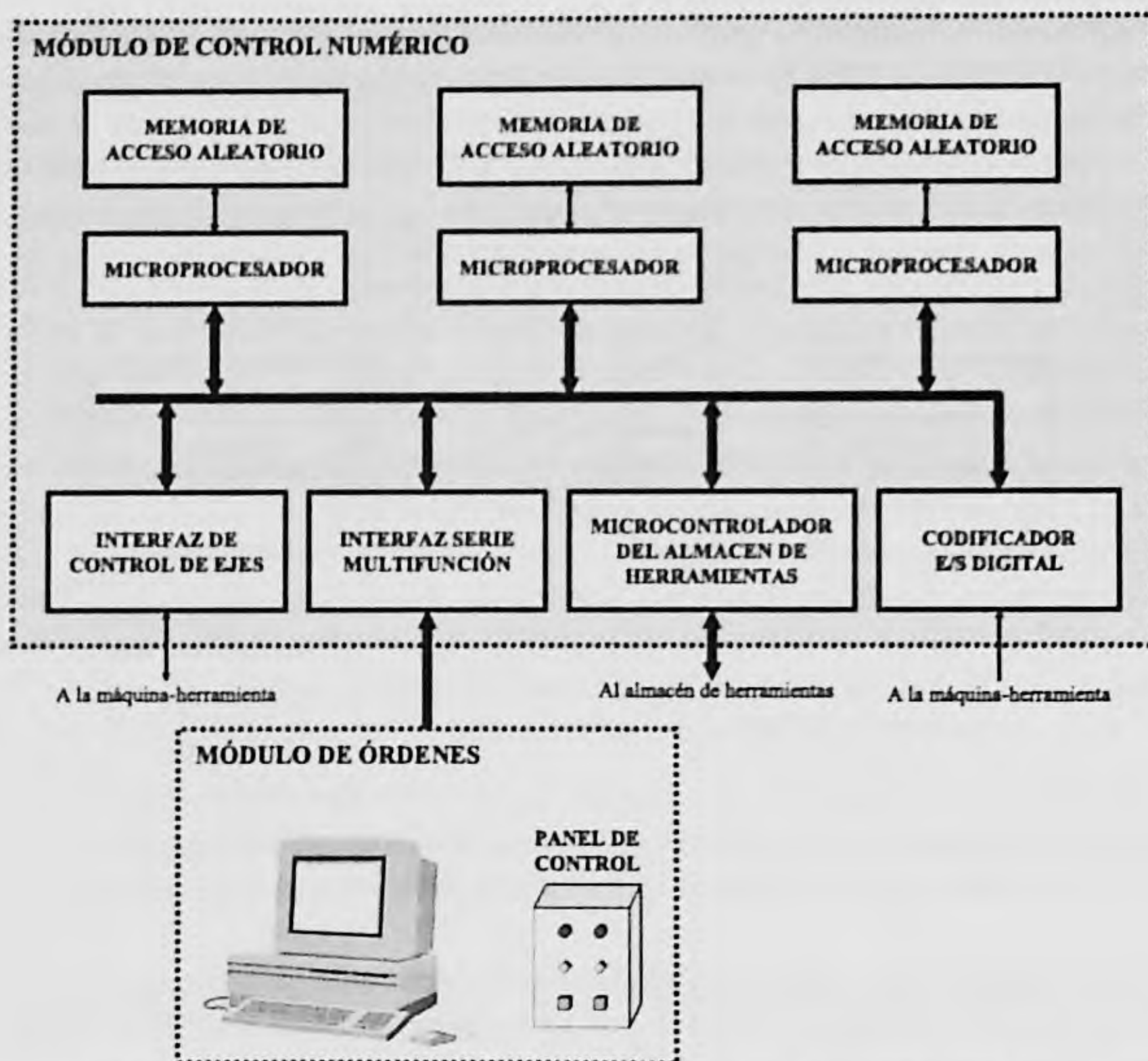


Figura 9.9. Esquema de bloques de un sistema electrónico de control numérico (CNC) de gran capacidad formado por un módulo de órdenes y un módulo de control.

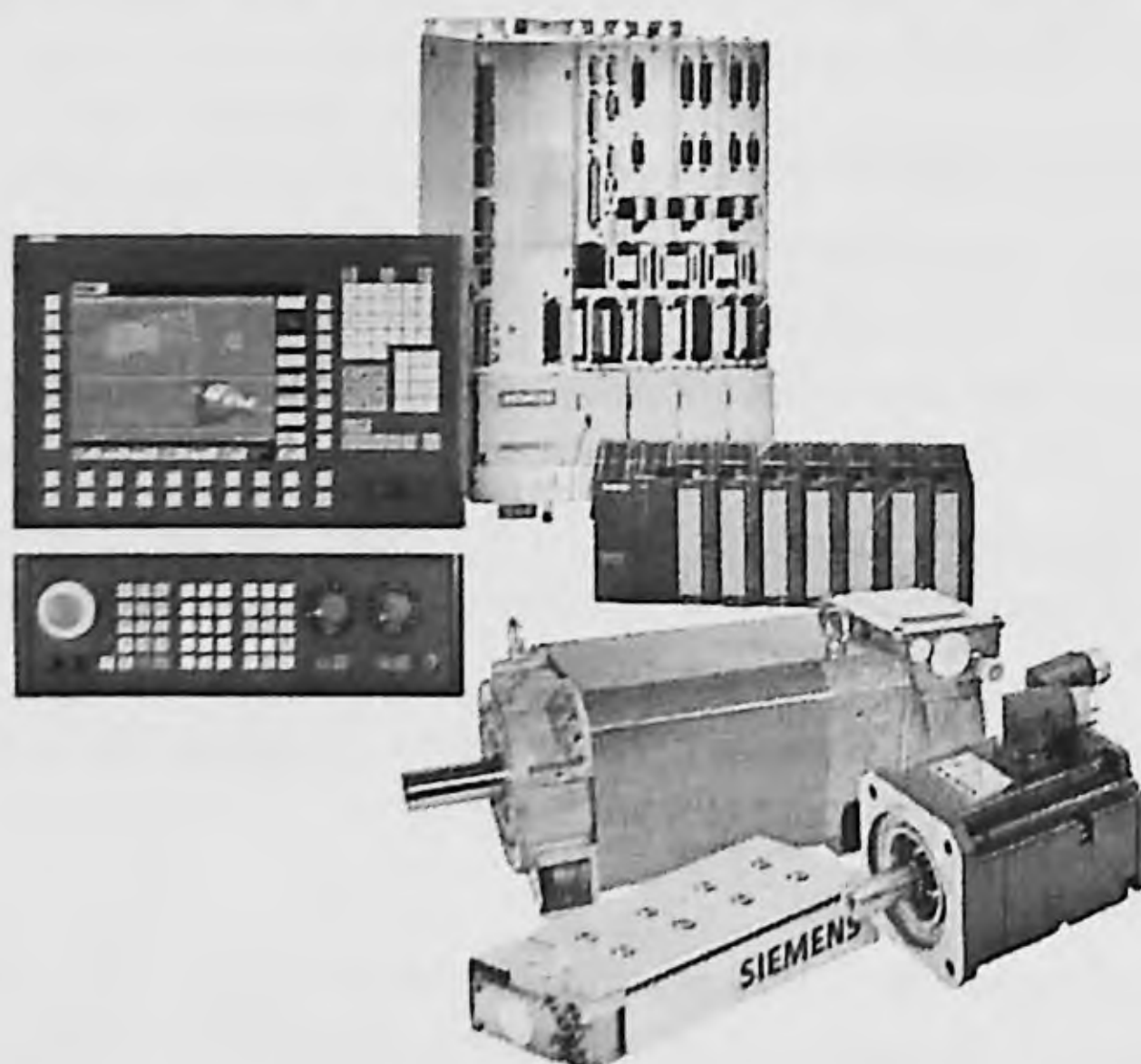


Figura 9.10. Elementos que forman parte de un CNC de gran capacidad (Cortesía de Siemens).

Sistemas CNC de capacidad reducida

Son sistemas CNC que controlan, por lo general, uno o dos ejes como máximo y que se realizan según dos arquitecturas diferentes:

- **CNC de control en bucle abierto**

Utilizan como actuador un motor paso a paso de mecánica sencilla y robusta (Figura 9.11a) adecuado para sistemas de posicionamiento de un par máximo del orden de la decena de Nm. Sus características más importantes son:

- Se realizan con computadores integrados denominados microcontroladores [ANGU 03] [MAND 07] [MART 93] o con dispositivos lógicos programables [MAND 02].
- Se pueden utilizar como módulos de salida de los autómatas programables (Figura 9.11b).

- **CNC de control en bucle cerrado**

Son sistemas CNC que utilizan como actuador un servomotor (Figura 9.12) y suelen estar basados en un procesador que ejecuta algoritmos de control de posición que permiten variar la velocidad y la aceleración durante el accionamiento y cuyo esquema de bloques es similar al de la figura 9.8. Sus características más importantes son:

- Funcionan en bucle cerrado y por ello tienen más precisión y alcanzan un par máximo mayor que los realizados con motores paso a paso.
- Se pueden utilizar como módulos de salida de los autómatas programables.

El desarrollo de la Microelectrónica favoreció el diseño y comercialización de sistemas de control numérico distribuido o directo denominados DNC (acrónimo de *Direct or Distributed Numerical Control*), consistentes en un computador central (*Host*), que se suele denominar computador DNC, que se conecta con varios sistemas CNC (Figura 9.13) a través de una red de controladores que se describe en el apartado 9.3.2.3.2.

El computador DNC proporciona, entre otras, las siguientes prestaciones:

- Memorización y gestión de los programas de los diferentes CNC conectados a él.
- Simulación gráfica del mecanizado.
- Eliminación de tiempos muertos.
- Memorización y gestión de datos de herramientas.
- Centralización de datos de las distintas máquinas de la fábrica.

Los sistemas DNC se combinan con los autómatas programables y otros sistemas electrónicos de control para dar lugar a los sistemas de fabricación flexible que se estudian en el apartado 9.2.4.6.

Al lector interesado en estudiar con más detalle los sistemas de control numérico CNC y DNC se le remite a la bibliografía [CRUZ 05] [GROO 80] [KIEF 98] [KRAR 00] [MATT 01].

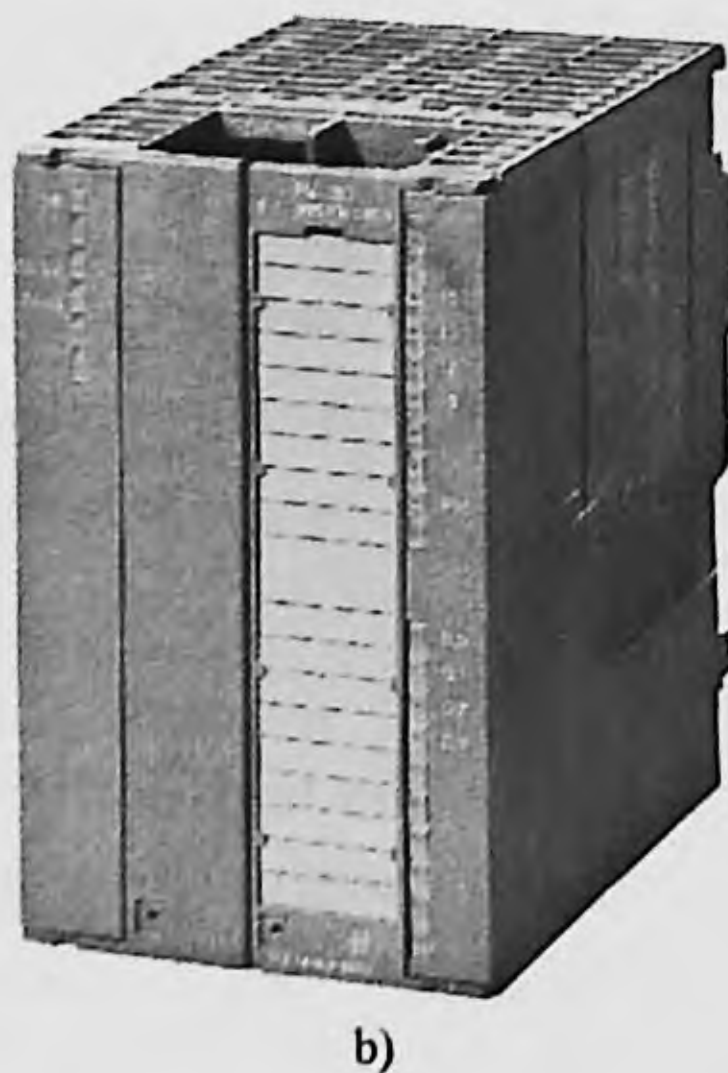
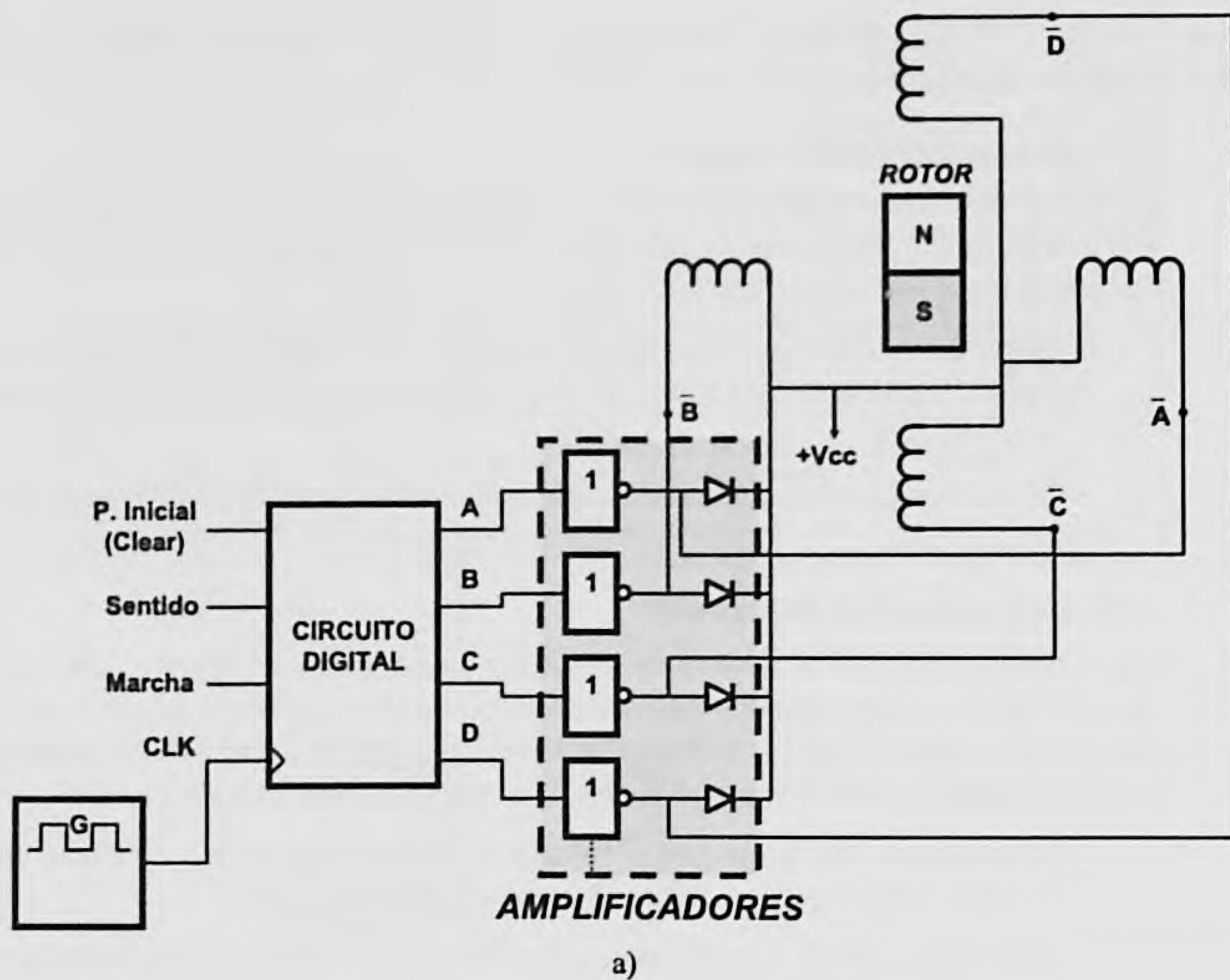
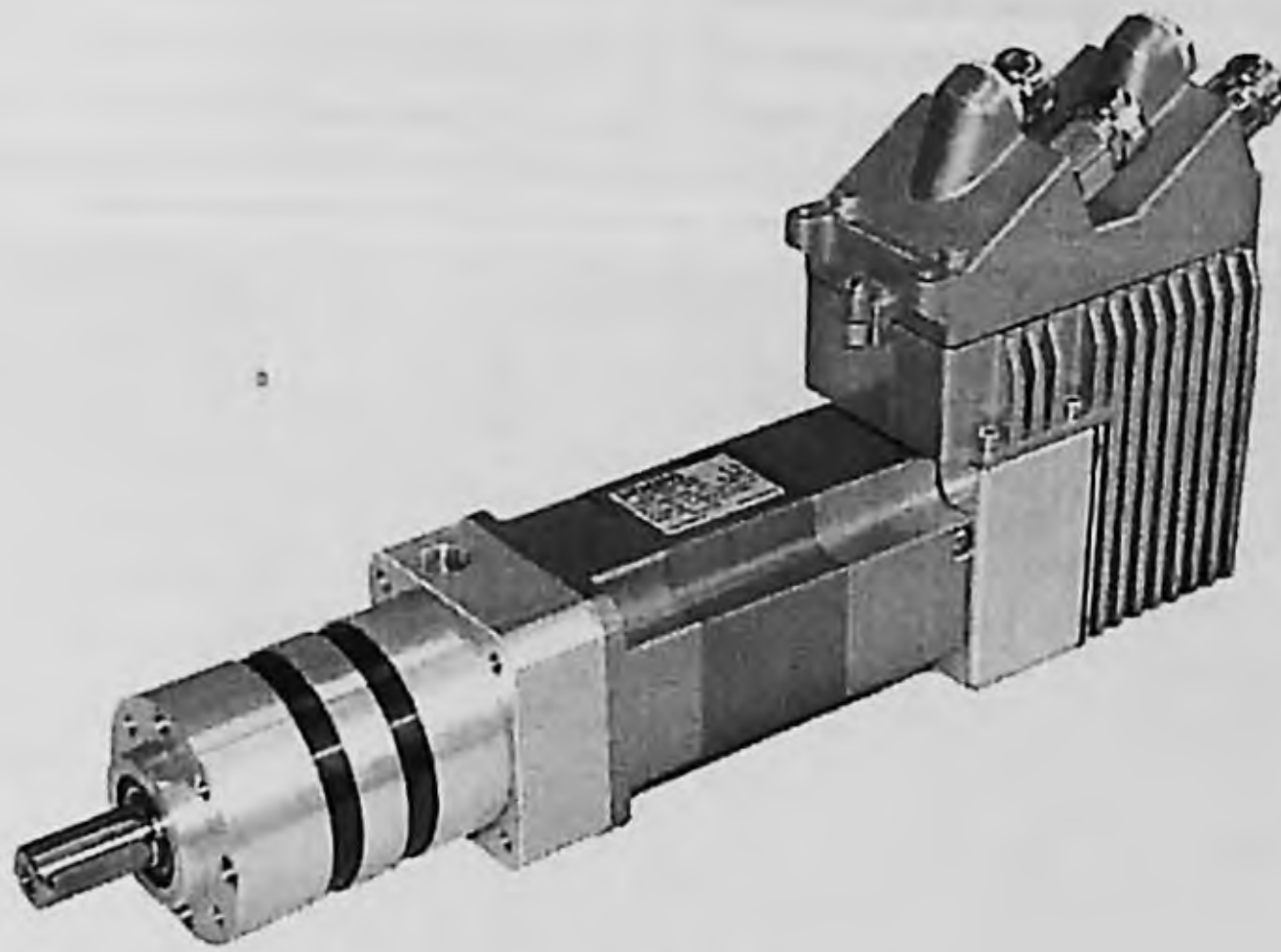
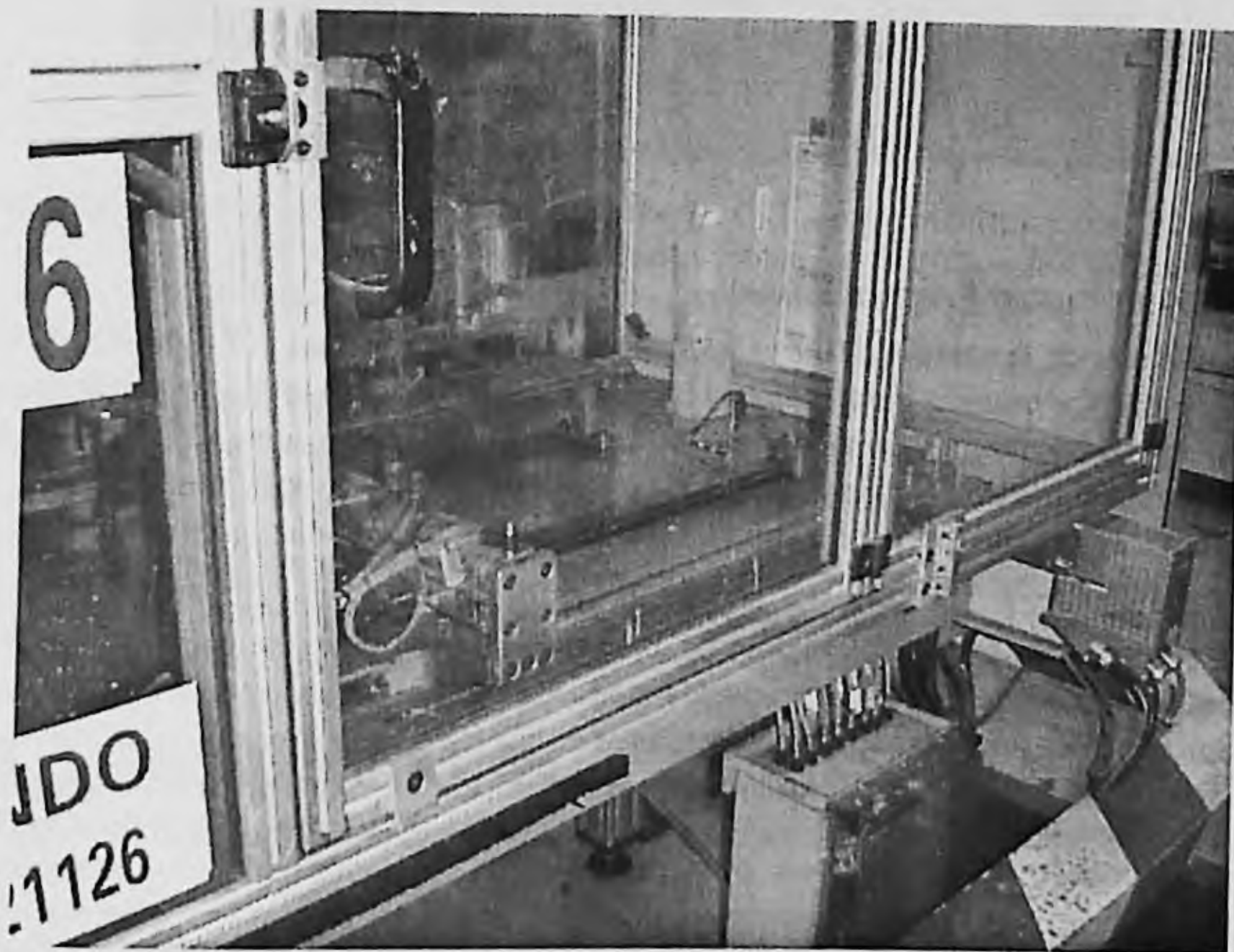


Figura 9.11. Sistema de control numérico realizado con un motor paso a paso: a) Esquema de bloques; b) Módulo FM353 de Siemens de posicionamiento de motores paso a paso acoplable a un autómata programable de la familia S7-300.



a)



b)

Figura 9.12. Sistema CNC que utiliza como actuador un servomotor: a) Módulo servomotor SIMODRIVE POSMO de Siemens; b) Sistema de control numérico para posicionar un eje realizado con el módulo SIMODRIVE POSMO (Cortesía de GKN Driveline Vigo).

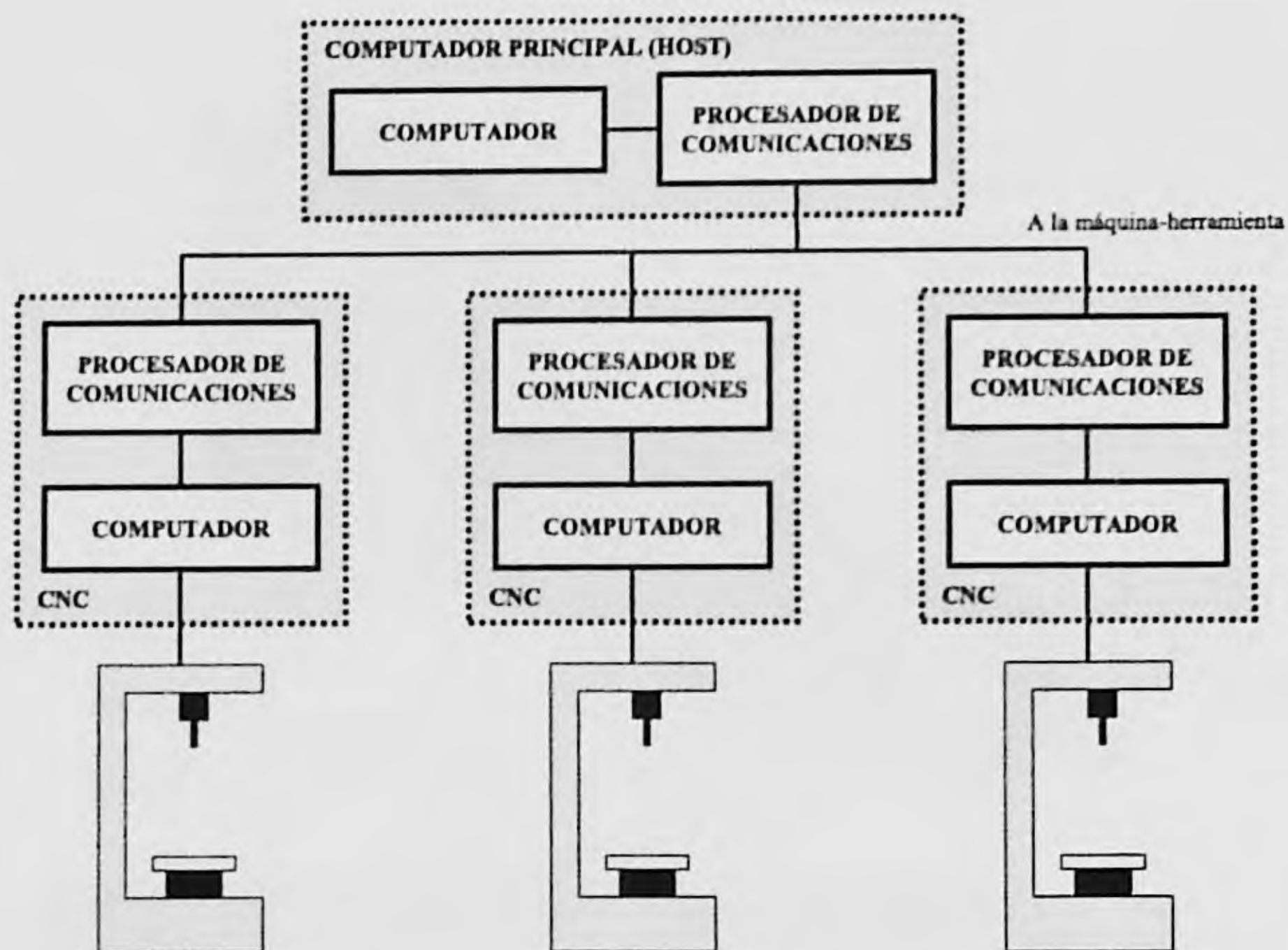


Figura 9.13. Sistema de control numérico distribuido o directo.

9.2.4.4.2 Autómatas Programables

El desarrollo de la energía eléctrica a lo largo de la primera mitad del siglo XX propició la introducción de los motores eléctricos en la práctica totalidad de los procesos productivos. Ello hizo que una de las partes más importantes de cualquier fábrica fuese su instalación eléctrica, formada por un conjunto de máquinas eléctricas estáticas y rotativas interconectadas por medio de elementos de control, un porcentaje elevado de los cuales eran dispositivos (tanto sensores como actuadores) todo-nada (descritos en los apartados 7.2.4 y 7.4.1.3 del capítulo 7).

Para controlar de forma automática las instalaciones eléctricas fue necesario realizar sistemas digitales secuenciales implementados mediante relés interconectados. Pero los sistemas realizados con relés presentaban el inconveniente de sus elevadas dimensiones, la reducida vida útil de sus contactos y la necesidad de modificar el cableado entre los dispositivos para adaptar la función del sistema de control a los cambios de la instalación eléctrica debidos a las modificaciones del proceso productivo.

Unido ello a la comercialización de los transistores hizo que se desarrollasen en Europa, en la década de 1960, los primeros controladores lógicos cableados realizados con componentes electrónicos discretos encapsulados en un bloque, que recibieron la denominación de “Relés estáticos”, de los que la serie SIMATIC de Siemens y los “NORBIT” de PHILIPS fueron ejemplos significativos (Figura 9.14).

La falta de flexibilidad de los relés estáticos impulsó la investigación aplicada que dio como resultado el desarrollo y posterior comercialización, a finales de la década de 1960, de un equi-

po electrónico programable por personal no informático, destinado a realizar funciones lógicas combinacionales y secuenciales en un ambiente industrial y en tiempo real. Dicho equipo, mucho más sencillo que un computador y con unas funciones mucho más limitadas, recibió en el mundo anglosajón la denominación de PLC (acrónimo de *Programmable Logic Controller*) que todavía se utiliza en la actualidad, junto con la de autómatas programable.

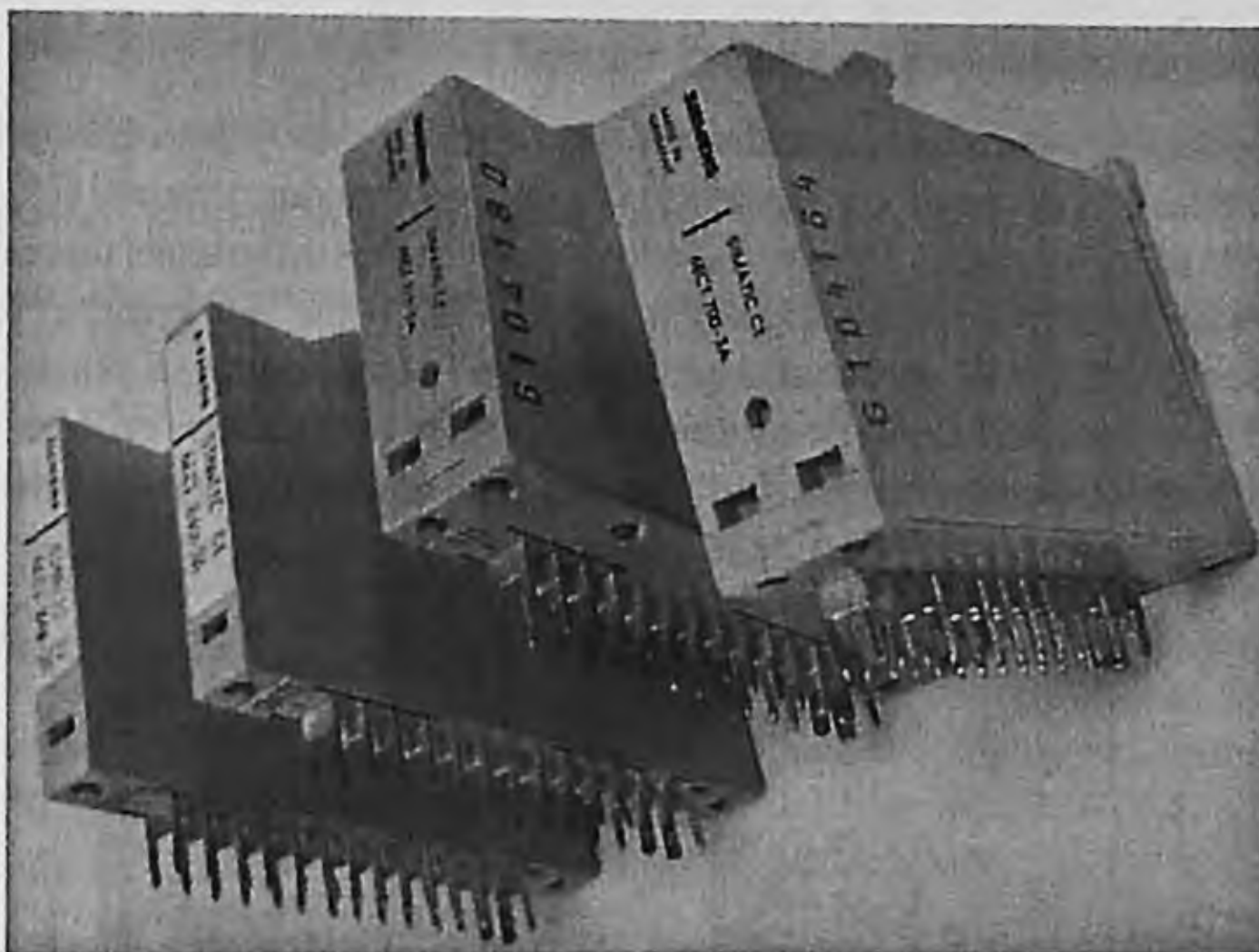


Figura 9.14. Relés estáticos de la familia SIMATIC C de Siemens realizados con circuitos integrados de tecnología TTL, utilizados a finales de la década de 1960 para realizar controladores lógicos cableados.

El rápido desarrollo de los microprocesadores a lo largo de la década de 1980 hizo que se utilizasen para realizar autómatas programables con capacidad de tratamiento de señales analógicas mediante su conversión en digitales y su proceso a continuación.

La importancia de los autómatas programables en las tareas de control de todo tipo de instalaciones industriales ha motivado la publicación de este libro, que en el resto de los capítulos describe los diferentes conceptos ligados actualmente con ellos.

9.2.4.4.3 Computadores Industriales

Los computadores industriales son sistemas electrónicos de control constituidos por un computador de aplicación general (*General purpose computer*) adecuadamente diseñado y montado para poder trabajar en el entorno de un proceso industrial continuo o discreto y soportar condiciones ambientales adversas (presencia de polvo, elevada temperatura, presencia de radiaciones electromagnéticas, movimientos bruscos, golpes, etc.).

Los computadores industriales, y en especial aquellos cuyo sistema físico (hardware) es equivalente al del computador personal, conocido popularmente como "PC" (acrónimo de *Personal Computer*), cuyo primer modelo, el PC XT, fue desarrollado inicialmente por la empresa IBM en 1981, han gozado en los últimos años de una gran aceptación en el ámbito específico de la planta industrial.

La entrada de estos sistemas electrónicos en el nivel de fábrica se hizo viable cuando su evolución técnica, que ha sido vertiginosa en los últimos veinte años, permitió el desarrollo de versiones cuyos elementos constitutivos (como por ejemplo la fuente de alimentación, el sistema de ventilación, los sistemas de almacenamiento masivo, etc.) están más adaptados para su utilización en el entorno industrial, lo cual eleva su fiabilidad y disponibilidad al mismo tiempo que mantiene la excelente relación precio-prestaciones funcionales que proporcionan actualmente los equipos utilizados en el entorno ofimático.

Históricamente, la utilización del computador en aplicaciones industriales se inició [de forma casi coincidente con la propia comercialización del computador personal (PC)] para llevar a cabo las tareas de edición, puesta en marcha y depuración de los programas de control asociados a los sistemas electrónicos de automatización de procesos como por ejemplo los sistemas de control numérico (CNC) o los autómatas programables. Actualmente se puede afirmar que la práctica totalidad de los fabricantes de este tipo de sistemas electrónicos utilizan el computador personal como plataforma que soporta la ejecución de este tipo de herramientas informáticas (Figura 9.15). Como ejemplo de ellas se pueden citar, entre otros, los sistemas de programación "STEP5" y "STEP7" de Siemens, "RSLogix" de ROCKWELL AUTOMATION, "PL7 AUTOMATION UNITY" de SCHNEIDER ELECTRIC, "CX-Programmer" de OMRON, etc.



Figura 9.15. Unidades de programación FIELD PG de Siemens (Cortesía de Siemens).

Pero el aumento de la capacidad de memoria y de la velocidad de operación de los computadores personales, propició también su utilización como sistemas de control. Se inició esta tendencia a principios de la década de 1990 en sectores industriales muy específicos y tecnológicamente punteros (como por ejemplo el aeroespacial, liderado por la NASA, y otras empresas internacionales) y en investigación (como por ejemplo en el entorno universitario) y se trasladó paulatinamente a las plantas industriales más tradicionales para llevar a cabo la integración de los procesos de gestión de la producción con las tareas de control realizadas por otros sistemas electrónicos utilizados tradicionalmente para ello (autómatas programables, controladores numéricos, robots industriales, etc.).

El progreso de la Microelectrónica hizo, además, que a partir de mediados de la década de 1990 se desarrollasen un conjunto de sistemas de control en los que el computador industrial juega un papel importante. Entre dichos sistemas cabe citar:

- Computador industrial y autómatas programable (*PC-PLC architecture*).
- Computador industrial y sistema operativo en tiempo real (*Real Time Operating System*).
- Computador industrial y software de autómatas programable (*Soft PLC*).
- Computador empotrado (*Embedded computer*).

Cada uno de estos sistemas de control tiene características propias, que se describen a continuación.

Computador industrial y autómatas programable (arquitectura "PC-PLC")

Este tipo de sistema de control nace cuando se desarrollan programas informáticos, que se pueden ejecutar en computadores industriales, denominados SCADA (acrónimo de *Supervisory Control And Data Acquisition*) por numerosos fabricantes de sistemas electrónicos de control [BAIE 03] [BOYE 04] [RODR 03]. Ejemplos de programas informáticos SCADA son "InTouch" de WONDERWARE, "RSView" de ROCKWELL AUTOMATION, "iFix Family" de INTELLUTION, "Factory Link" de US DATA, CIMPLICITY HMI de GE Fanuc y "WinCC" de SIEMENS (Figura 9.16).

Los programas SCADA permiten, mediante la conexión del computador a uno o varios equipos de automatización con los cuales se intercomunica, llevar a cabo tareas avanzadas de gestión (como por ejemplo la monitorización, supervisión, control estadístico y mantenimiento del proceso, la trazabilidad de la producción, etc.) que son, cada vez en mayor medida, imprescindibles en los procesos productivos modernos y forman parte del nivel de fábrica o del nivel de empresa de la pirámide CIM (descrita en el apartado 9.2.5). La importancia de los programas SCADA hace que a ellos se dedique el apartado 8.3.3 del capítulo 8.

Este sistema de control y gestión, fruto de la integración de los autómatas programables con los computadores industriales, se conoce popularmente con el nombre de arquitectura "PC-PLC" (acrónimo de *Personal Computer - Programmable Logic Controller*) y se realiza, actualmente, en dos versiones diferentes:

- Computador industrial y autómatas programables independientes enlazados mediante un canal de comunicaciones, que se representa en la figura 9.17.
- Computador industrial y placa de autómatas programable (*Slot-PLC*), que se representa en la figura 9.18.

La versión *Slot-PLC* es producto de la integración de los equipos de control de procesos y los computadores industriales y ha comenzado a ser desarrollada a finales del siglo XX. Consiste en realizar el autómatas programable en una placa de circuito impreso (con CPU y memoria propios) que se coloca en un conector o ranura (*Slot*) del bus principal (por ejemplo, los buses "PC-104" o "PCI" de un computador industrial) para llevar a cabo la comunicación con él. Por ello esta placa se suele denominar *Slot-PLC* (Figura 9.18a) y se alimenta, en muchos casos, de forma independiente del computador para seguir funcionando aunque se produzcan fallos

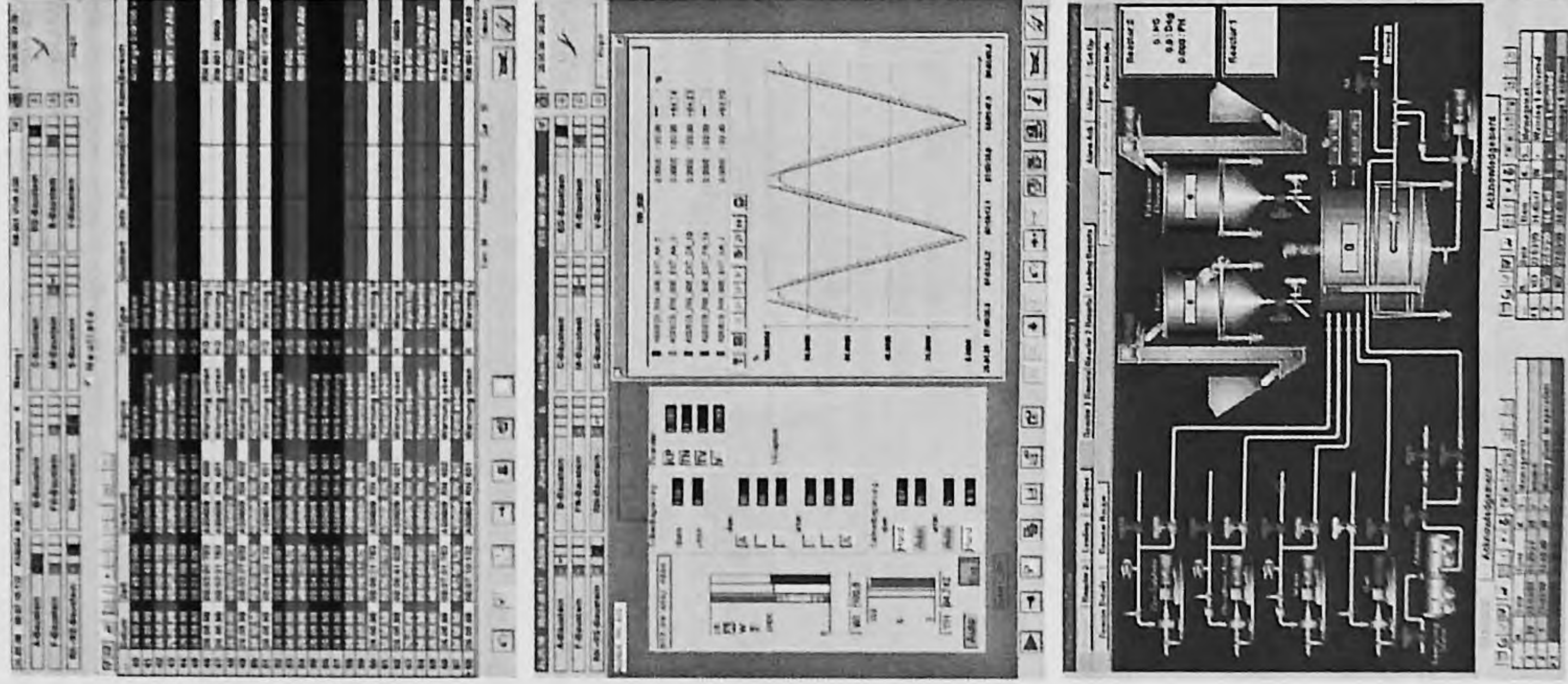


Figura 9.16. Pantallas de una aplicación del programa informático SCADA WinCC (Cortesía de Siemens).

(de funcionamiento, alimentación, etc.) del computador industrial en el que reside. Esta versión de la arquitectura PC-PLC es más económica que la que utiliza un autómata independiente de similares prestaciones. Además, la placa *Slot-PLC* suele interaccionar con los dispositivos de campo (sensores y actuadores) del proceso a través de un sistema de comunicaciones industriales (denominado usualmente bus de campo), que se describe en el apartado 9.3.2.3.

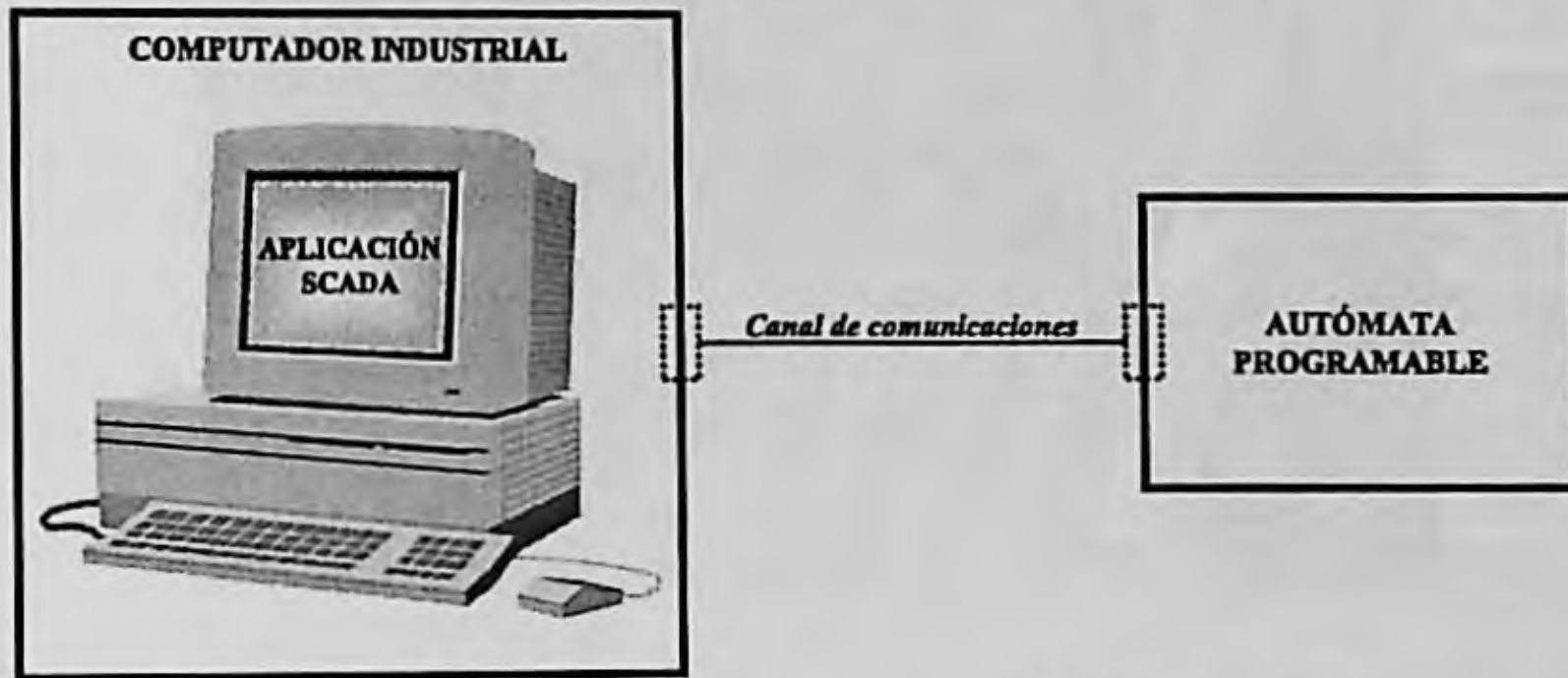
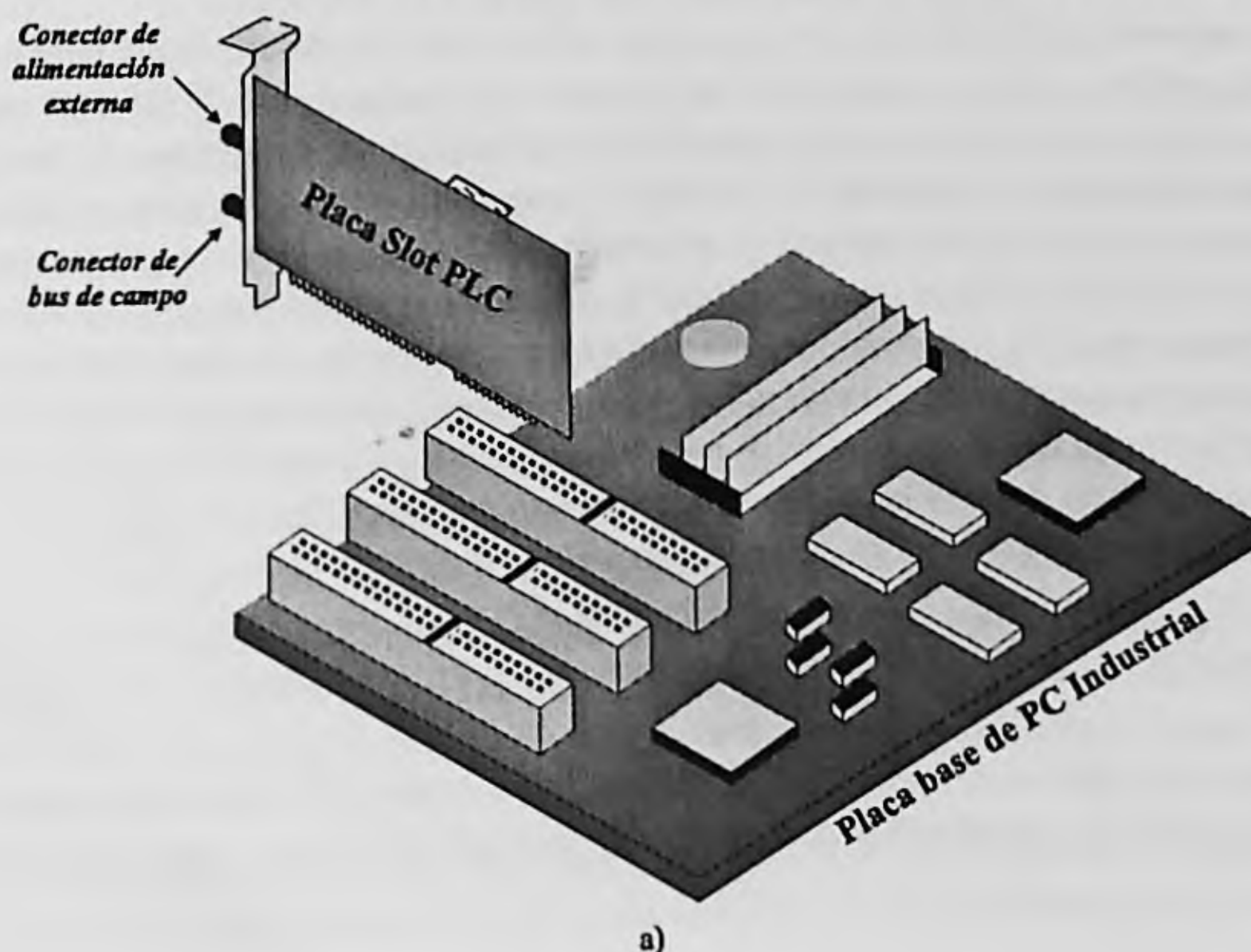


Figura 9.17. Esquema de bloques de la arquitectura "PC-PLC" formada por un computador industrial y un autómata programable independiente.

La conexión de la placa *Slot-PLC* con el bus principal del computador permite, fundamentalmente, que la velocidad y eficiencia del intercambio de información entre los procesos de control (que se ejecutan en ella) y los programas de gestión (que se ejecutan en el computador industrial en el que se aloja) sean lo más eficientes y fiables posibles.

Diversos fabricantes disponen ya de este tipo de sistemas en su catálogo. Son ejemplo de *Slot-PLC* el "WinAC Slot 412 o 416" de Siemens (Figura 9.18b), el "AS-i PCI Master" de BIHL&WIEDEMANN y el "IBS ISA FC/486 DX/I-T" de PHOENIX CONTACT.



a)



b)

Figura 9.18. *Arquitectura PC-PLC realizada con un computador industrial y una placa de autó-mata programable (Slot-PLC): a) Representación gráfica del concepto Slot-PLC; b) Computador industrial de Siemens que incorpora una placa "WinAC Slot 412".*

Computador industrial y sistema operativo en tiempo real

La aplicación del computador industrial en sectores tecnológicamente avanzados (como el aeroespacial, robótica, visión artificial, universitario, etc.) motivó el desarrollo de este tipo de sistema de control, que consiste en utilizar el propio computador industrial como un equipo en el que se implantan, simultáneamente, el control en tiempo real y la gestión de los procesos de fabricación.

Diversas empresas especializadas en este sector informático han desarrollado, para los computadores industriales, sistemas operativos en tiempo real denominados RTOS (acrónimo de *Real Time Operating System*), que posibilitan el desarrollo de programas de control que satisfacen las restricciones de determinismo temporal asociadas intrínsecamente con los procesos de producción. Estos sistemas operativos permiten, además, ejecutar otros programas como por ejemplo los de desarrollo y gestión, mediante la utilización de interfaces gráficos basados en ventanas y la comunicación con otros sistemas informáticos a través de redes ofimáticas. En la figura 9.19 se representa el entorno de desarrollo del sistema de manutención industrial basado en el Sistema Operativo de tiempo real RTX para Windows-XP de Interval Zero. Dicho entorno utiliza el lenguaje C para desarrollar los procesos de control y Visual Basic para las aplicaciones de interfaz máquina-usuario (HMI) y de monitorización.

Diversas empresas han desarrollado este tipo de sistemas operativos. Entre ellas cabe citar a Siemens que, a partir del sistema operativo "iRMX" de INTEL, desarrolló el RTOS denominado "RMOS" (que hoy en día es la base de sus equipos de control de procesos "STEP7"), QNX y LYNX que desarrollaron sendos RTOS que llevan su nombre, WINDRIVER, cuyo RTOS denominado "VxWorks" es utilizado con frecuencia por la NASA en sus proyectos aeroespaciales, etc.

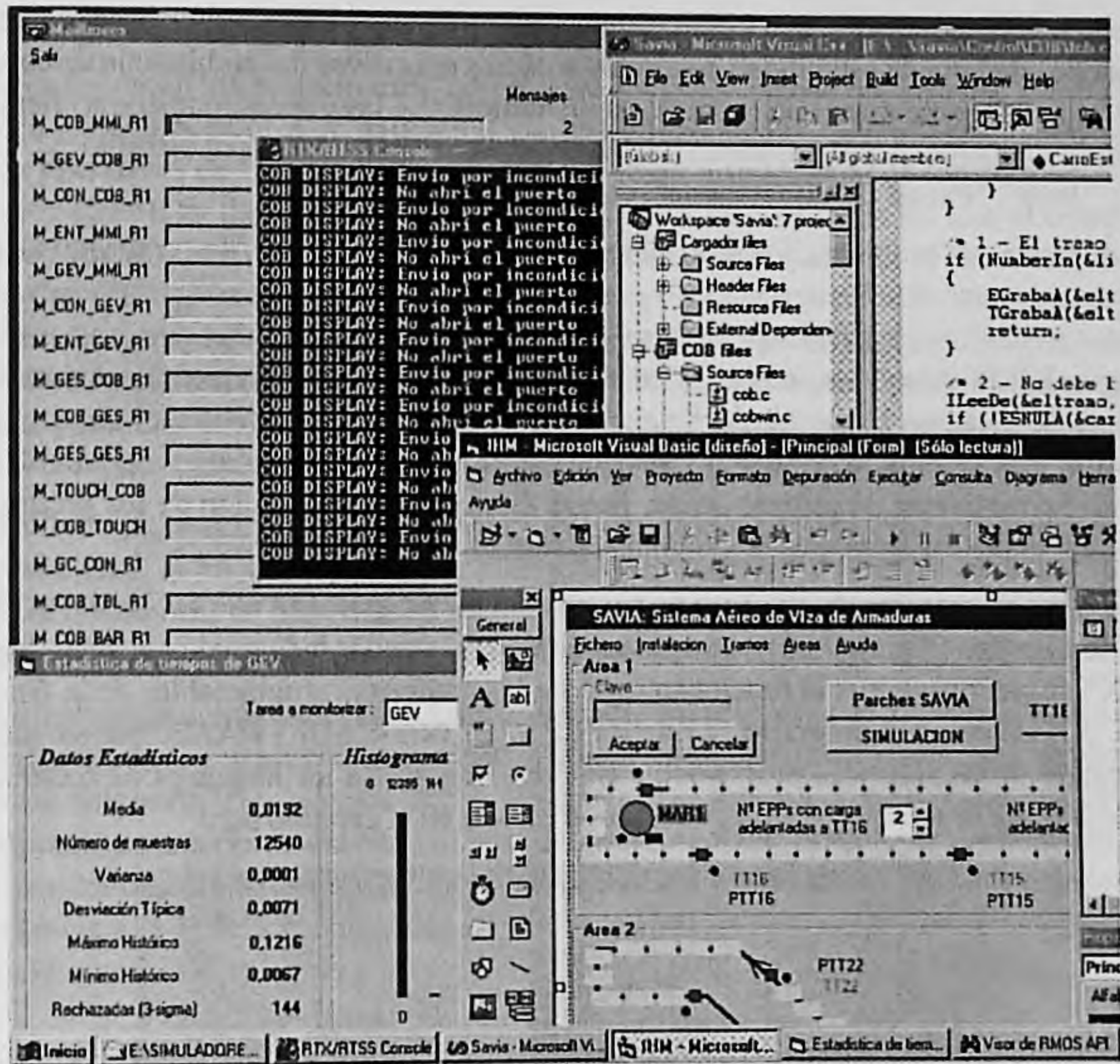


Figura 9.19. Ejemplo de entorno de desarrollo de aplicaciones de control en el sistema operativo de tiempo real RTX de Interval Zero (Cortesía del Departamento de Ingeniería de Sistemas y Automática de la Universidad de Vigo).

Una variante reciente de este tipo de sistema de control consiste en adaptar los sistemas operativos de aplicación general (es decir, que son una norma “de facto” en el entorno industrial actual) con el fin de utilizarlos para llevar a cabo el control en tiempo real de los procesos productivos. Para ello, en la mayor parte de los casos, se realizan modificaciones específicas del núcleo de dichos sistemas operativos a fin de que permitan la ejecución determinista de los programas de control en tiempo real de los procesos productivos para lo cual relegan, si es preciso, las tareas de gestión a un menor nivel de prioridad. Entre los productos disponibles en el mercado se pueden citar “RTX” (*RealTime eXtensión*) de Interval Zero (sistema utilizado por Siemens para una de sus soluciones *Soft-PLC* que se describen en el siguiente apartado), “iRMX for Windows” de TENASYS, e “Hyperkernel” de NEMATRON, desarrollados a partir de la familia de sistemas operativos “Windows” de MICROSOFT, así como “RTLinux Free” de Wind River, desarrollado a partir del sistema operativo de libre distribución “Linux”.

Computador industrial y software de autómatas programable

La limitada utilización actual del computador industrial como sistema de control de los procesos productivos está relacionada, en buena medida, con la tradicional resistencia al cambio

(o miedo a lo desconocido), con cierta “fama” (por otra parte bien merecida hasta no hace mucho tiempo) de falta de estabilidad de ciertos sistemas operativos del ámbito ofimático y con la falta de conocimiento que el personal de mantenimiento de las plantas industriales tiene de los entornos y lenguajes de programación utilizados en este ámbito, como por ejemplo “C”, “C++”, “Ada”, “Java”, “Pascal”, “Visual Basic”, etc.

Para paliar este problema, los fabricantes proponen, como solución de compromiso, programas conocidos con el sobrenombre de *Soft-PLC* (abreviatura de *Software Programmable Logic Controller*). Dichos programas son herramientas informáticas que, en combinación con algún RTOS (como los indicados en el apartado anterior), emulan el funcionamiento de un autómata programable, es decir, funcionan igual que él. Para el diseñador de aplicaciones de control, este sistema se programa y se comporta igual que un autómata programable debido a que las herramientas de programación, puesta en marcha y depuración de los programas de control, así como los lenguajes de programación utilizados, son los mismos.

Ejemplos de este tipo de sistemas de control son los programas “WinAC Basis” y “WinAC RTX” de Siemens (Figura 9.20) que emulan, en un computador industrial bajo “Windows XP” y “RTX”, respectivamente, el funcionamiento de los autómatas programables de la familia SIMATIC S7, así como el programa “CodeSys SP RTE” de 3-S SOFTWARE, que emula el funcionamiento de un autómata programable genérico que utiliza los lenguajes de programación establecidos por la norma IEC-61131-3, que se estudia en el capítulo 3.

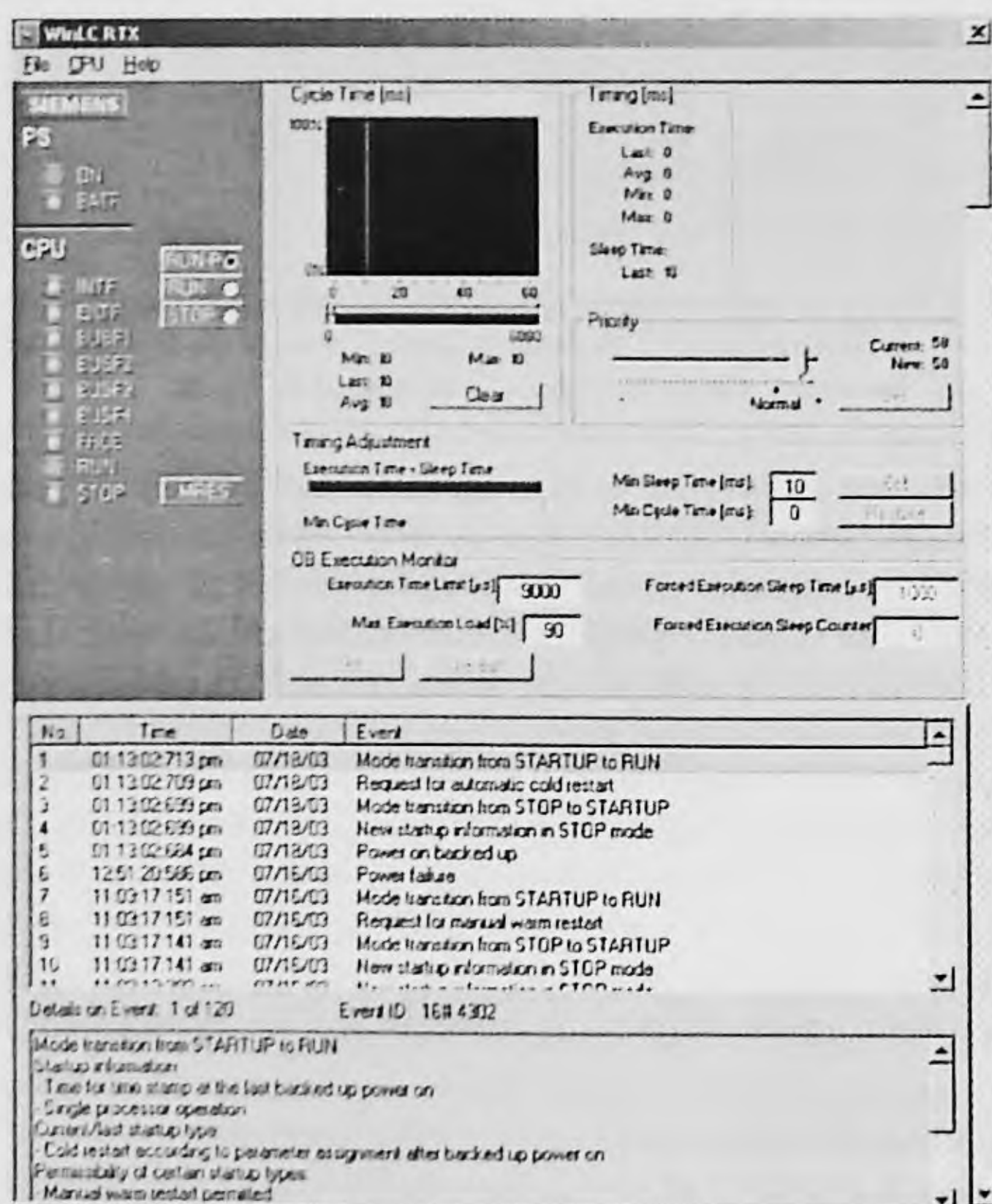


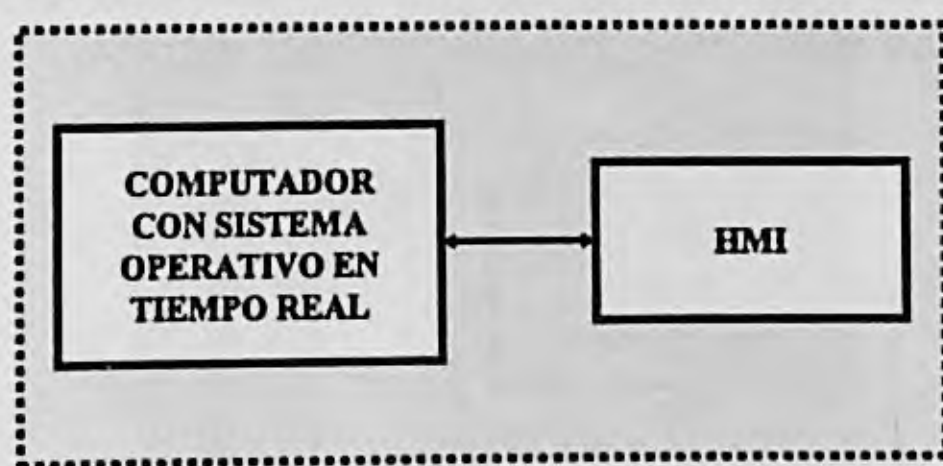
Figura 9.20. Interfaz de usuario del programa “WinAC RTX” de Siemens.

Computador empotrado

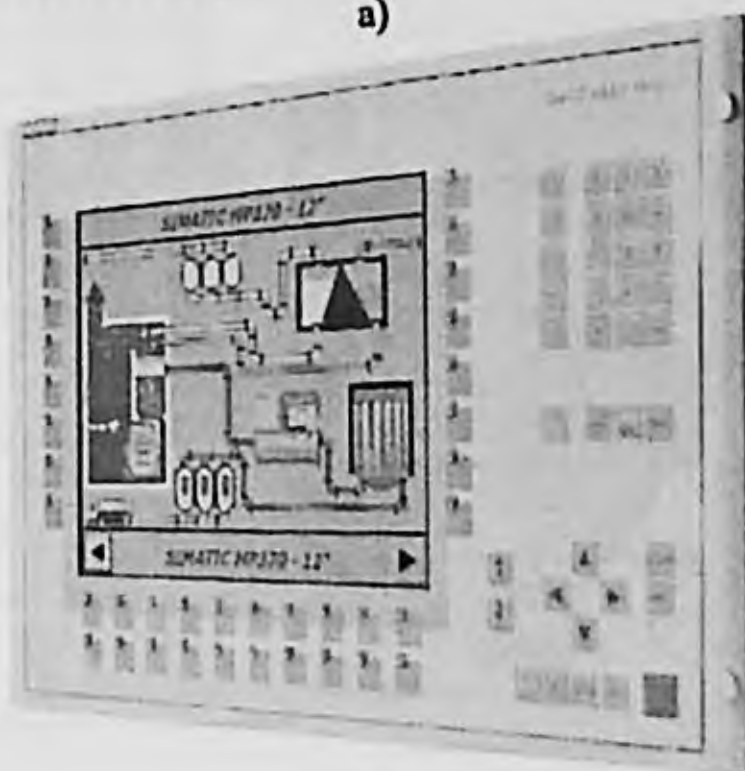
El avance de la Microelectrónica, que mejoró las prestaciones del sistema físico (*hardware*) de los computadores de aplicación general e hizo posible la implantación en ellos de sistemas operativos de tiempo real (cuya denominación más habitual es *Embedded Operating Systems*), ha dado lugar a otro tipo de sistemas de control en los que el computador queda empotrado dentro de otro tipo de sistemas, entre los cuales cabe citar los interfaces usuario-máquina (HMI), descritos en el apartado 8.3.2, las agendas personales denominadas PDA (acrónimo de *Personal Data Assistant*), los reproductores de audio y video, los sistemas electrónicos de asistencia al guiado y ocio en automóviles, etc. A este tipo de sistemas, que tienen en la actualidad una gran proyección en el ámbito industrial, se les suele dar en general la denominación de computadores empotrados (*Embedded computers*).

Entre los sistemas operativos de tiempo real desarrollados para los computadores empotrados cabe citar, entre otros, "Windows CE" y "XP Embedded" de MICROSOFT y "S60" de SYMBIAN. Además, se han desarrollado versiones para computadores empotrados de otros sistemas operativos de aplicación más general, como por ejemplo "QNX" y "Linux".

Un ejemplo de computador empotrado, utilizado recientemente en el ámbito industrial son los interfaces usuario-máquina, denominados plataformas multifuncionales, que combinan los paneles de operación y los paneles táctiles descritos en el apartado 1.3.3.2 con uno de los sistemas operativos anteriormente citados, para desarrollar aplicaciones de monitorización, supervisión, mantenimiento e incluso control de procesos industriales. Un ejemplo de estos sistemas son los equipos de la familia MP (acrónimo de *Multi Panel*) de Siemens. En la figura 9.21 se muestra el esquema de bloques y una solución comercial de esta clase de sistemas.



a)



b)

Figura 9.21. Equipo de la familia MP: a) Esquema de bloques; b) Fotografía del equipo MP370 (cortesía de Siemens).

9.2.4.4.4 Sistemas electrónicos de control de procesos continuos

Los procesos continuos (*Continuous Processes*) son aquellos cuyo producto final, en lugar de estar formado por un conjunto de elementos separados, está constituido por un material que fluye de forma continua (productos químicos en fase fluida, mezclas de sólidos como por ejemplo piensos, etc). Este tipo de procesos se caracterizan porque en ellos es necesario controlar elementos (como por ejemplo electroválvulas proporcionales, motores, etc.) mediante variables analógicas. Por lo tanto los sistemas electrónicos de control de procesos continuos deben recibir señales a partir de sensores de medida (descritos en el apartado 7.2.4 del capítulo 7), y generar variables analógicas que controlan los diferentes actuadores del proceso [KUO 95].

Los primeros sistemas de control de procesos continuos se realizaron con procesadores analógicos formados por amplificadores operacionales adecuadamente realimentados [FIOR 01]. En la figura 9.22 se representa, como ejemplo, el esquema de bloques de un sistema electrónico analógico de control de velocidad que utiliza como sensor una generatriz tacométrica.

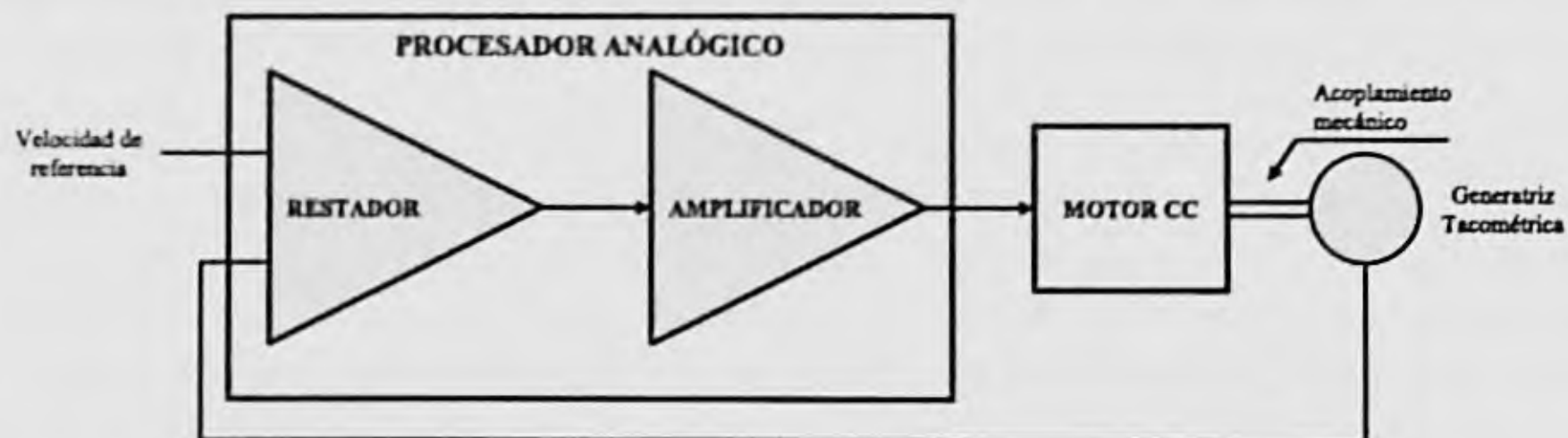


Figura 9.22. Esquema de bloques de un sistema electrónico analógico de control de velocidad que utiliza como sensor una generatriz tacométrica.

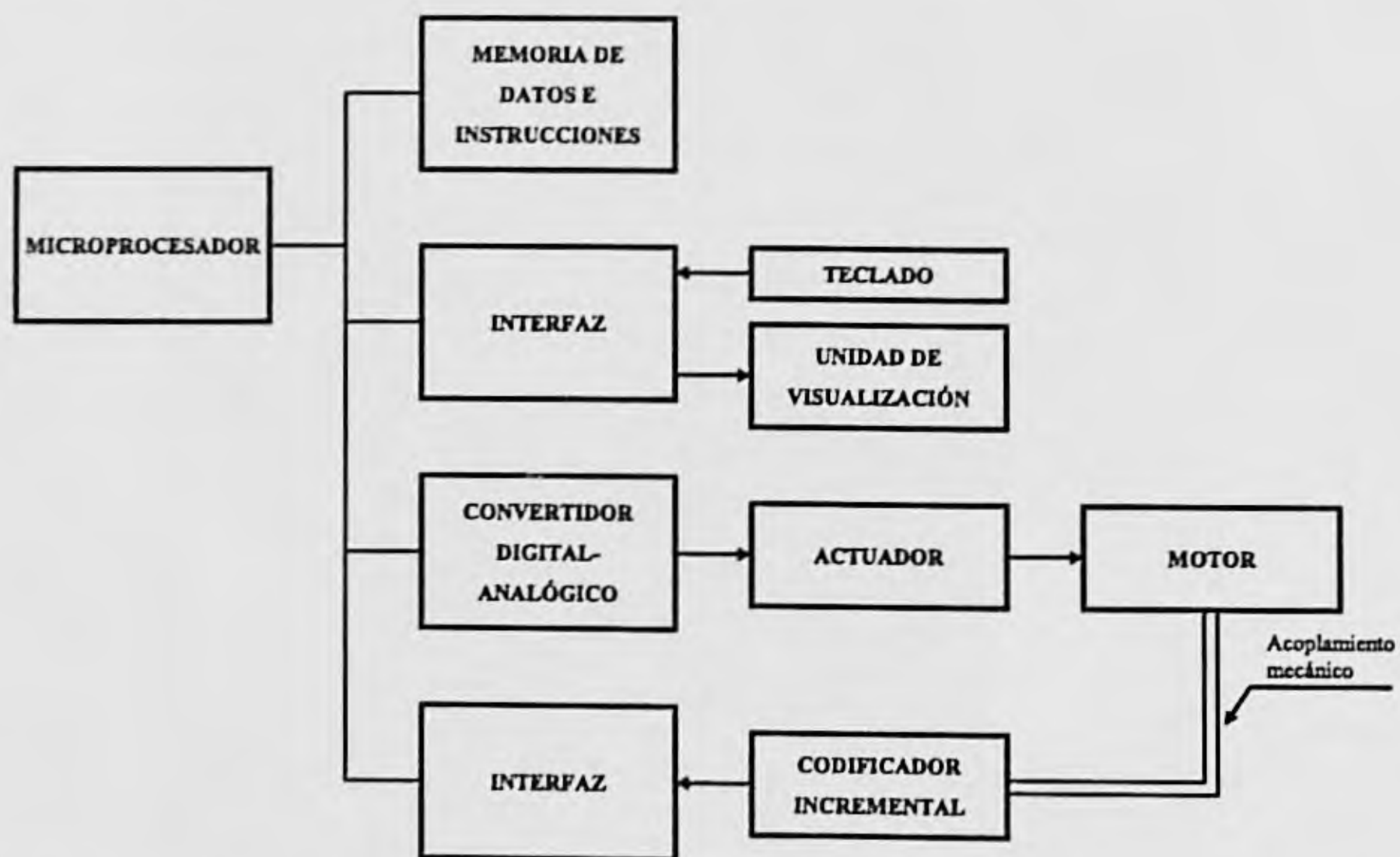


Figura 9.23. Esquema de bloques de un sistema de control de velocidad realizado mediante un microcomputador.

En la década de 1960 se inició la aplicación de los procesadores digitales al control de procesos continuos, tal como se indica en el apartado 4.2.2.3 del capítulo 4, pero su elevado coste, por estar realizados con componentes electrónicos discretos, restringió su campo de aplicación a aquellas instalaciones de gran complejidad cuyo coste era muy superior al del procesador. La comercialización de los microprocesadores primero y de los microcontroladores [ANGU 03] [MAND 07] [MART 93] posteriormente, consecuencia de la elevación de la capacidad de integración de los fabricantes de circuitos integrados, provocó una drástica disminución del coste y una elevación de la capacidad de los procesadores digitales, e hizo que sustituyesen paulatinamente a los procesadores analógicos en numerosas aplicaciones [LEIG 85] [SING 80] [KUO 95]. En la figura 9.23 se representa el esquema de bloques de un sistema de control de velocidad realizado mediante un microcomputador que utiliza como sensor un codificador incremental que genera un número de impulsos por unidad de tiempo proporcional a la velocidad de giro del motor al que está acoplado.

En la actualidad la práctica totalidad de los sistemas electrónicos de control de procesos continuos se implementa mediante procesadores digitales. En los procesos continuos sencillos en los que el número de variables a medir y controlar es reducido, se utilizan generalmente equipos basados en microcontroladores, que son computadores en los que la unidad central de proceso, la memoria volátil y no volátil así como diversas unidades de interfaz están incluidas en un único circuito integrado de muy elevada escala de integración [ANGU 03] [MAND 08] [MART 93]. Un ejemplo típico de regulador de procesos continuos sencillos es el regulador PID de la familia SIPART DR de Siemens (Figura 9.24).



Figura 9.24. Regulador PID industrial de la familia SIPART DR que está implementado con un microcontrolador (cortesía de Siemens).

En los procesos complejos, denominados en general simplemente procesos (*Processes*), en los que el número de variables a manipular puede llegar a ser de cientos e incluso de miles, se utilizan sistemas electrónicos digitales de elevada capacidad de cálculo, que combinan las prestaciones de los autómatas programables y los computadores industriales. Un ejemplo típico de este tipo de sistemas es el sistema SIMATIC PCS7 de Siemens, basado en los controladores de la familia S7-400 (Figura 9.25), en el que PCS es un acrónimo de *Process Controller System*. Este tipo de sistemas es conocido, en general por las siglas DCS (*Distributed Control System*) que es una denominación, que, tal como se indica el apartado 4.2.3 del capítulo 4, utilizan numerosos fabricantes de sistemas electrónicos de control para hacer referencia a los controladores de procesos continuos.

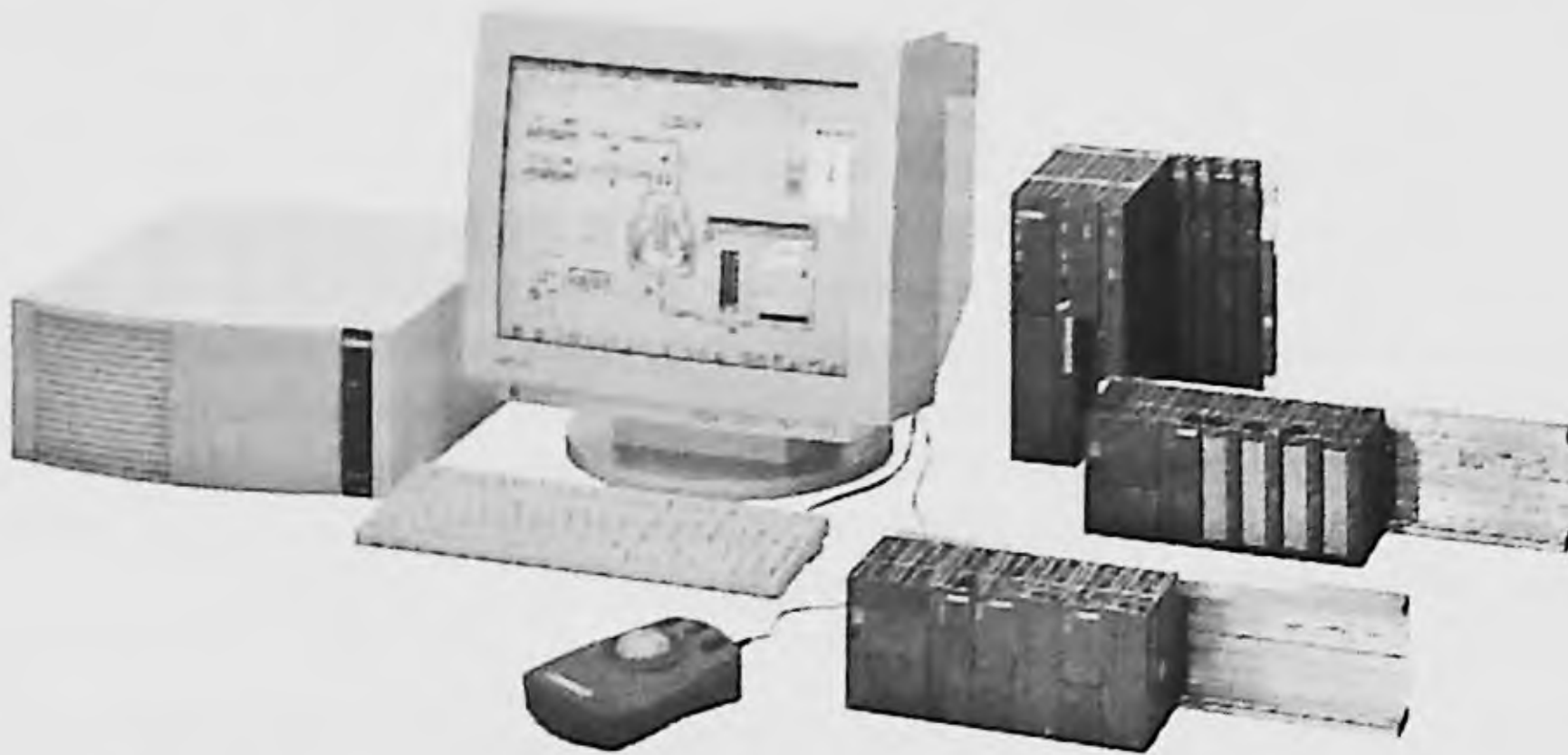


Figura 9.25. Sistema de control de procesos continuos SIMATIC PCS7 de Siemens basado en un autómata programable de la familia S7-400 (cortesía de Siemens).

9.2.4.4.5 Sistemas CAD-CAM

Suelen recibir esta denominación los sistemas electrónicos que ejecutan un conjunto de programas que automatizan y simplifican las tareas de diseño, simulación, y fabricación asistidas por computador.

Las principales funciones de un sistema CAD/CAM son:

- El diseño (eléctrico, mecánico, electrónico, de programas de control, etc.) asistido por computador.
- La simulación asistida por computador.
- La programación de sistemas de control numérico.
- La programación de robots.
- La planificación del proceso productivo.

Un sistema CAD/CAM es por lo tanto un computador que posee los recursos físicos (*Hardware*) y las herramientas de diseño y fabricación (*Software*) adecuadas para realizar las tareas antes indicadas. Suele formar parte del nivel de fábrica de la pirámide CIM (acrónimo de *Computer Integrated Manufacturing*) que se analiza en el apartado 9.2.5.

9.2.4.5 Sistemas de manipulación de elementos

9.2.4.5.1 Introducción

Reciben esta denominación los sistemas electromecánicos capaces de transportar elementos que constituyen productos o subproductos de un proceso de fabricación (*Manufacturing process*). Son sistemas que tienen una estructura mecánica fija que da lugar a un número reducido de trayectorias. Las máquinas de transferencia (*Transfer machines*) son uno de los ejemplos más característicos y los autómatas programables constituyen el tipo de sistema electrónico más utilizado para su control.

El aumento de la complejidad de determinados productos, como por ejemplo los automóviles, hizo necesaria la disponibilidad de manipuladores más flexibles e incentivó el desarrollo de los robots que la “*Robot Industry Association*” [AUDI 88] [FERR 86] define como manipuladores reprogramables multifuncionales diseñados para mover cargas, piezas, herramientas o dispositivos especiales (soldadores, pulverizadores, etc.) según diferentes trayectorias, lo cual les permite ser programados de forma automática para realizar tareas variables.

Los robots son sistemas complejos para cuyo diseño es necesaria la colaboración de expertos en diferentes áreas de la tecnología (mecánica, eléctrica, electrónica, neumática, teoría de control, etc.). Las diferentes formas de concebir las distintas partes de un robot hacen que se puedan realizar robots de prestaciones muy diferentes. El éxito de la robotización de una o más tareas de un proceso productivo solo es posible si se posee un conocimiento profundo de las mismas y se elige o se diseña el robot adecuado para llevarlas a cabo.

La utilización adecuada de robots en la producción genera los siguientes beneficios:

- Eleva la velocidad de la producción porque facilita la repetición automática de movimientos.
- Disminuye los tiempos muertos.
- Permite la realización de diferentes tareas sin más que cambiar el programa.
- Eleva la calidad de los productos.
- Disminuye la mano de obra directa.
- Facilita la realización de tareas que resultan penosas para el ser humano por:
 - Tener que ser realizadas en ambientes insalubres.
 - Ser duras, incómodas o altamente peligrosas.
 - Ser monótonas.

En el apartado siguiente se analizan los conceptos más importantes ligados a los robots.

9.2.4.5.2 Conceptos básicos de los robots y clasificación de los mismos

En la figura 9.26 se representa el esquema de bloques de un robot en el que se pueden distinguir las siguientes partes:

- Un sistema mecánico dinámico articulado que posee dispositivo de sujeción y agarre.

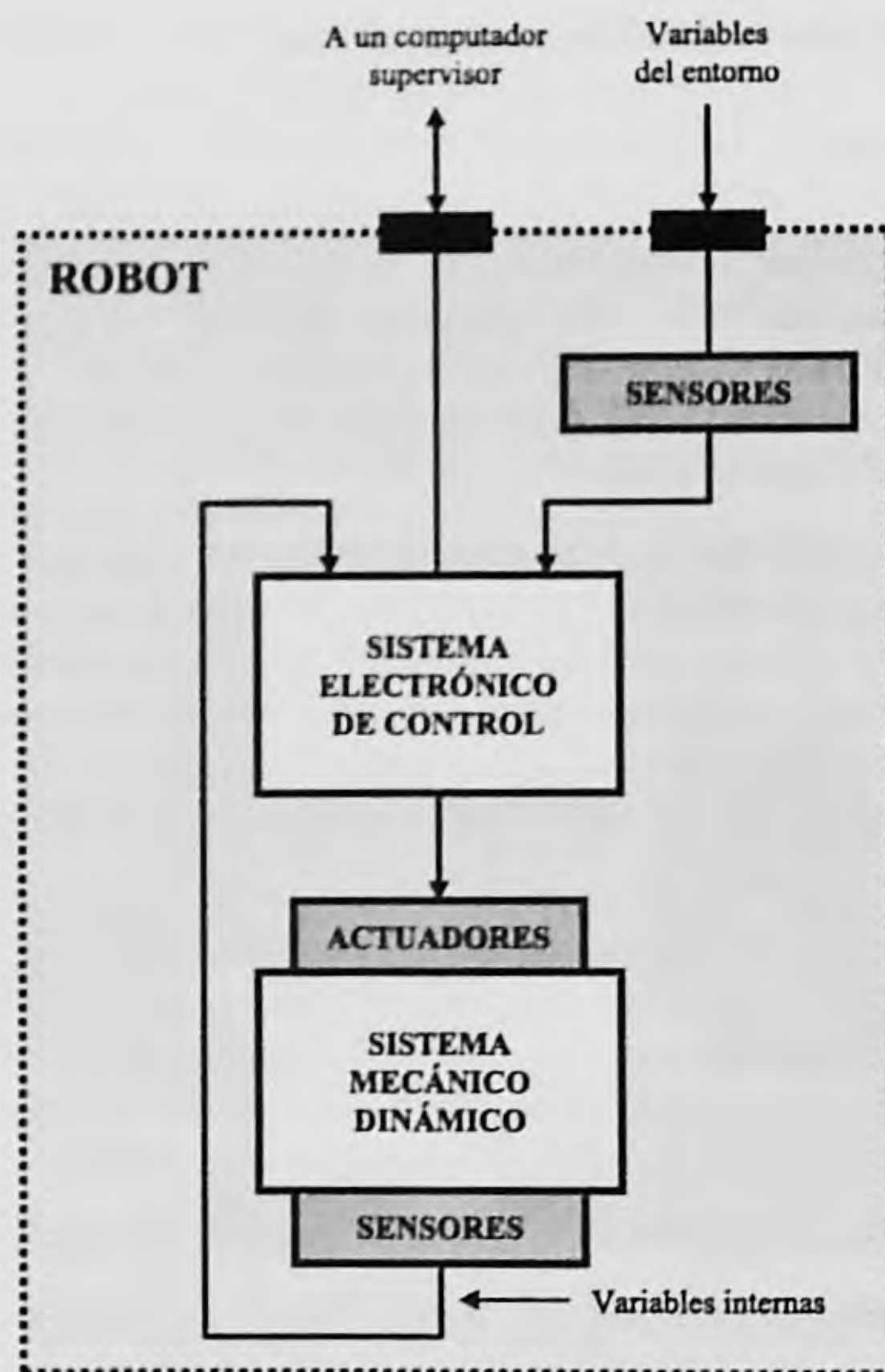


Figura 9.26. Esquema de bloques de un robot.

- El sistema motriz que puede ser electromecánico, neumático o hidráulico.
- Uno o mas sistemas electrónicos de control por computador.
- Un conjunto de sensores.

A continuación se analiza brevemente cada una de ellas. Para un estudio más profundo se remite al lector a la bibliografía [AUDI 88] [AYRE 83] [BARR 07] [FERR 86] [KIEF 98] [OLLE 01] [RENT 00] [SIMP 96] [TORR 02] [WEBB 93].

Sistema mecánico dinámico

Constituye el cuerpo del robot, que se encarga de ejecutar los desplazamientos dentro del área de trabajo. Como su nombre indica, está formado por un conjunto de elementos mecánicos relacionados entre sí, cuya posición relativa se puede modificar mediante unos elementos actuadores. El número de movimientos relativos da lugar a otros tantos grados de libertad. Dicho número se debe limitar al mínimo necesario para llevar a cabo un determinado trabajo.

De acuerdo con las características del sistema mecánico articulado, los robots se clasifican en tres grandes categorías [AUDI 88]:

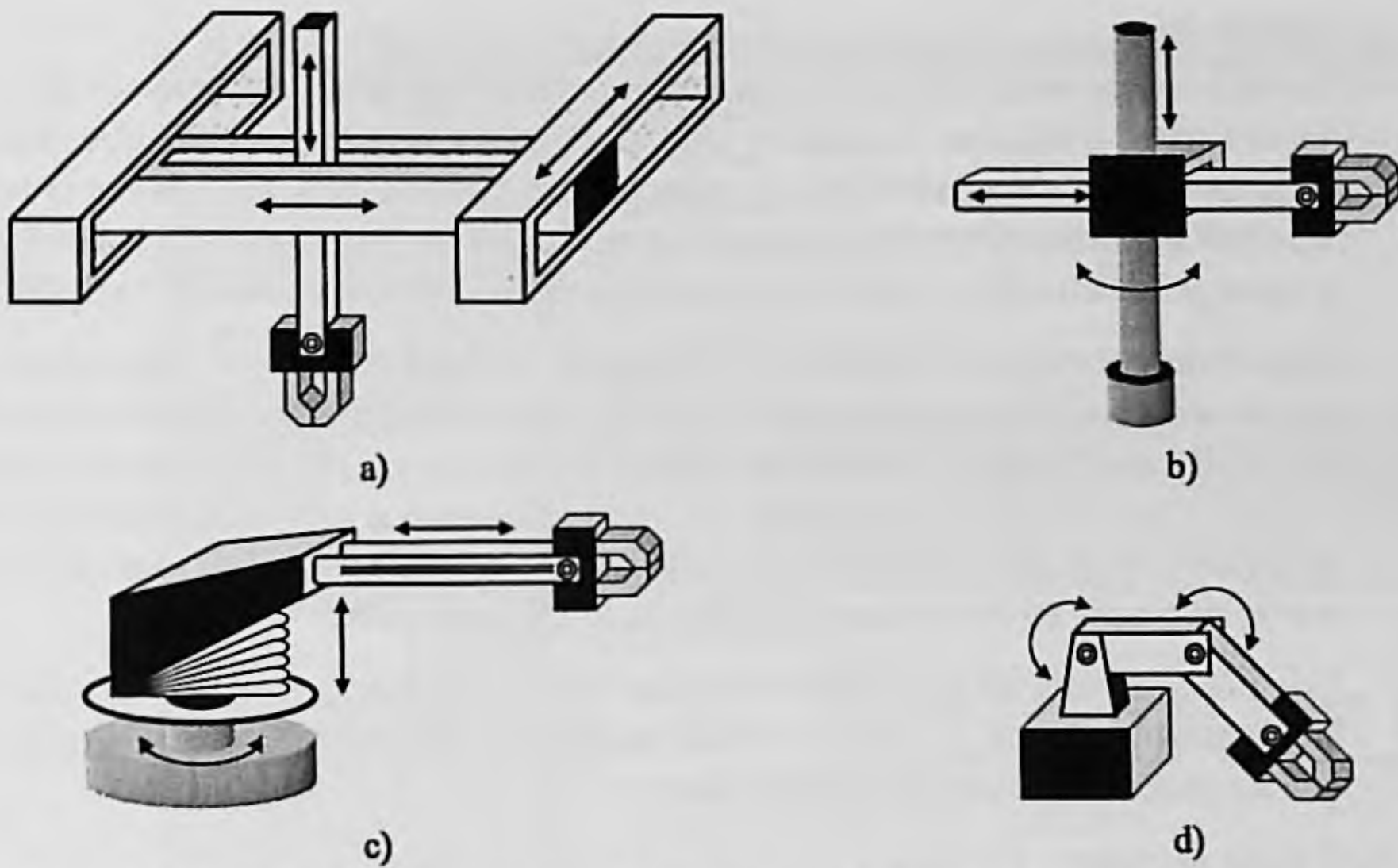


Figura 9.27. *Diferentes arquitecturas de los robots: a) Cartesiana; b) Cilíndrica; c) Esférica; d) Articulada.*

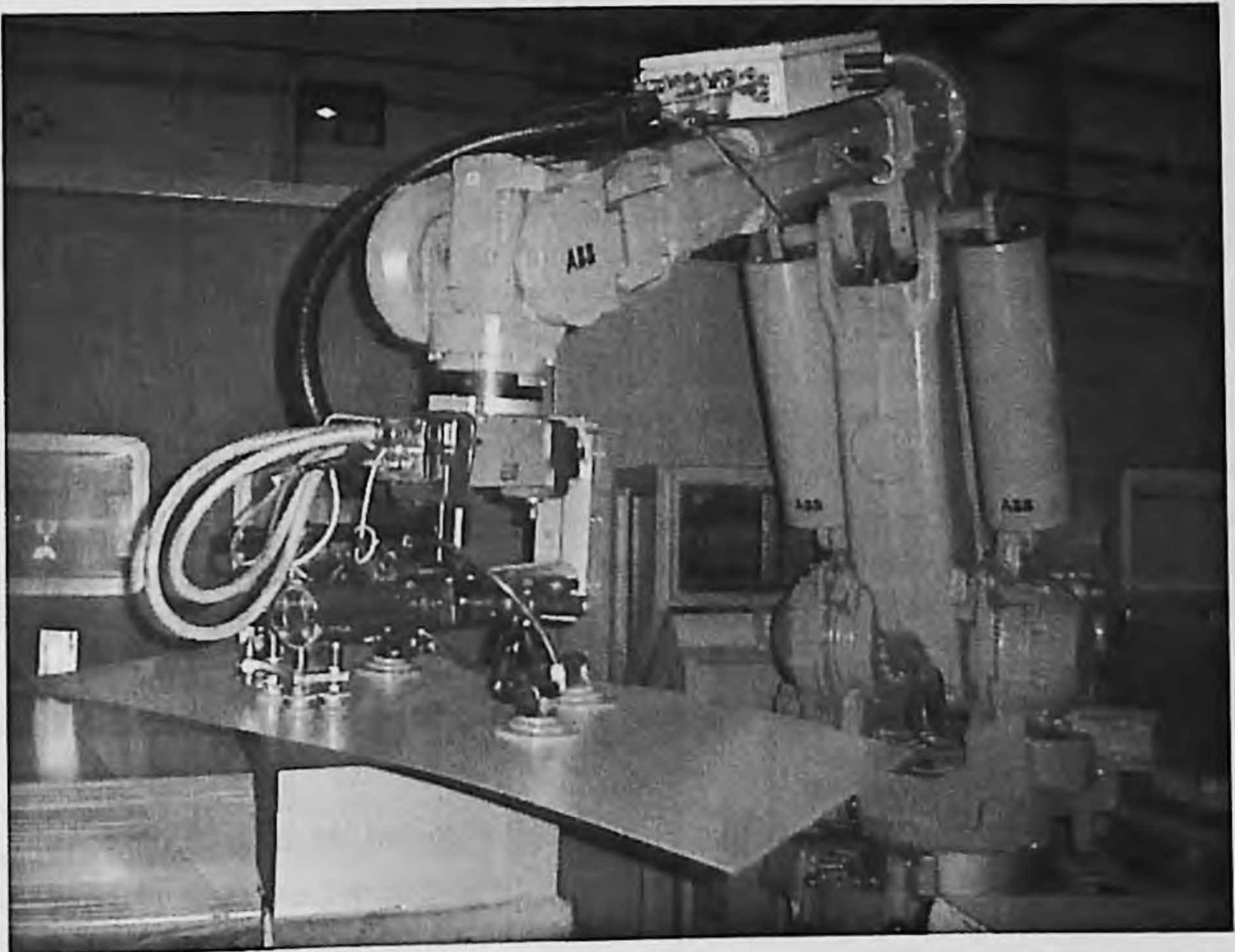


Figura 9.28. *Robot articulado para transferencia de piezas en una línea de prensas (cortesía de GESTAMP VIGO).*

- **Robots fijos**

Son robots que están rígidamente unidos a una superficie estática, que puede ser el suelo, una pared o el techo. Existe una gran variedad según la arquitectura que define los tipos de movimientos relativos que existen entre los elementos que forman el cuerpo. Los más importantes son los cartesianos o rectangulares, los cilíndricos, los esféricos o polares y los articulados cuyo principio de funcionamiento se indica en la figura 9.27.

Los robots cartesianos se mueven solamente en los tres ejes X , Y y Z . Los robots cilíndricos tienen solo dos ejes de movimiento que son el eje Z y el giro alrededor de él. El robot esférico tiene un movimiento rotativo y otro angular del elemento que sostiene el dispositivo de sujeción y agarre. Los robots articulados poseen al menos un brazo articulado que emula al hombro y se utilizan extensamente en aplicaciones de transferencia de piezas en los procesos de fabricación (Figura 9.28).

Los más utilizados en tareas de montaje en sistemas de fabricación son los esféricos y en especial los “*Scara*”, que combinan la arquitectura cilíndrica con la articulada lo que les proporciona una gran versatilidad.

- **Robots móviles**

Como su nombre indica son robots que se pueden desplazar. Pueden estar dotados de ruedas o de patas y por su complejidad sólo se utilizan, actualmente, en aplicaciones especiales.

- **Robots combinados**

Como su nombre indica, combinan las características de los fijos y de los móviles. Pueden estar montados sobre guías rectilíneas o poseer un mecanismo complementario que tiene sus propios ejes de movimiento como por ejemplo una mesa giratoria.

Dispositivo de sujeción y agarre

Forma parte del sistema mecánico articulado y su constitución depende de las características de los objetos a sujetar [AUDI 88]. Los robots más utilizados en las tareas de fabricación y soldadura poseen, por lo general, elementos terminales con forma de garra.

Sistema motriz

Está formado por un conjunto de elementos que proporcionan las fuerzas o pares necesarios para actuar sobre los elementos que constituyen el sistema mecánico articulado. Suelen recibir el nombre de accionamientos o actuadores y pueden ser eléctricos, neumáticos o hidráulicos.

Los actuadores más utilizados actualmente son motores eléctricos de alterna. En el pasado se utilizaron los hidráulicos en aplicaciones en las que se necesitaba una fuerza o un par elevados.

Sistema electrónico de control

El sistema electrónico de control de un robot es en general un sistema formado por varios computadores denominado multicomputador (*Multicomputer*). En la figura 9.29 se representa el esquema de bloques típico, en el que cada articulación está controlada por un computador y el conjunto de todos ellos se conecta a un computador central que a su vez está conectado a un procesador de comunicaciones que lo enlaza con los demás sistemas que forman parte de la planta de fabricación.

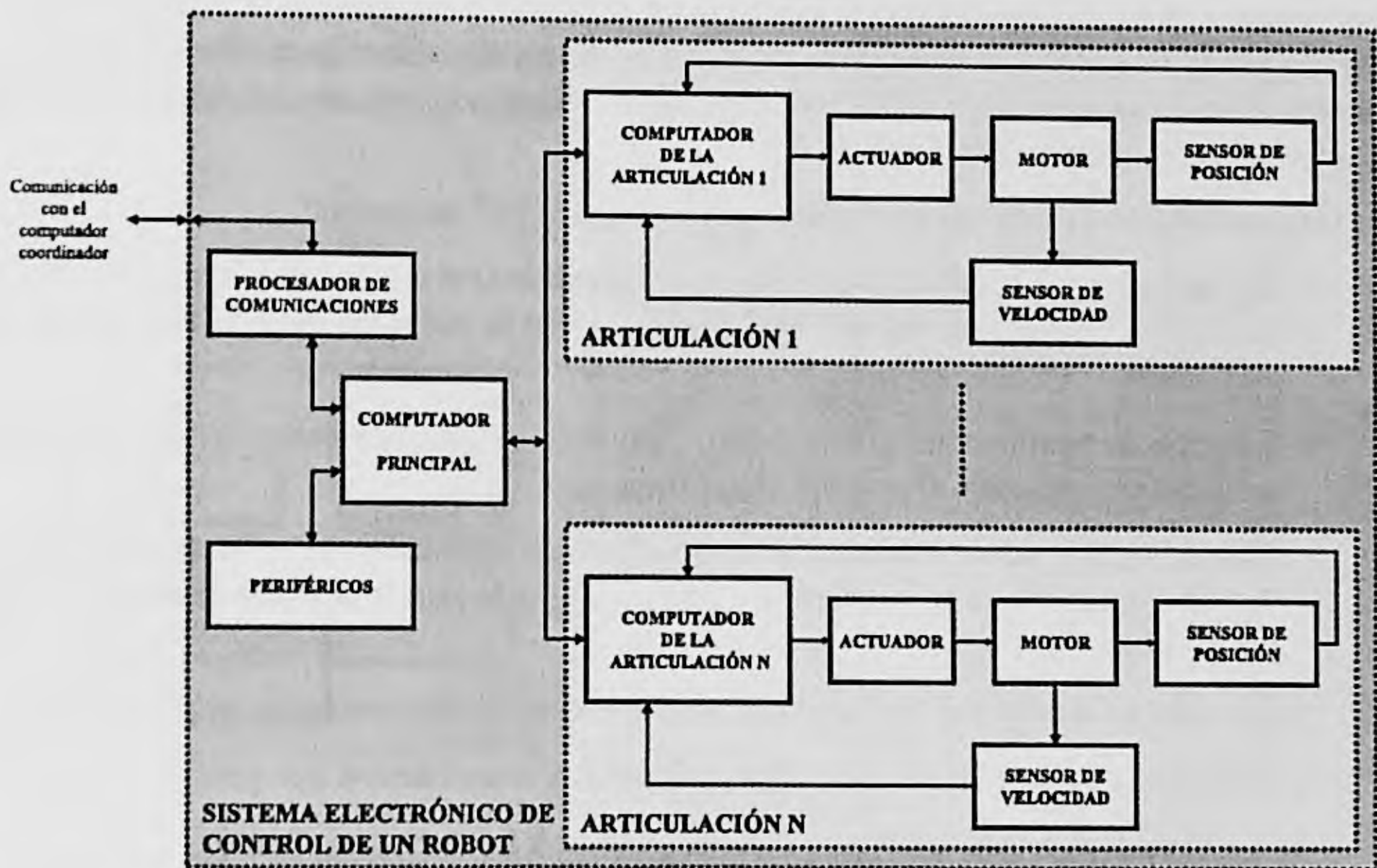


Figura 9.29. Esquema de bloques típico del sistema electrónico de control de un robot.

Sensores

Los sensores (descritos en el capítulo 7) son los elementos encargados de proporcionar al robot, en forma de señales eléctricas, la información de la evolución de un conjunto de variables físicas que se pueden dividir en dos grupos (Figura 9.26):

- Variables internas de las que las más importantes son la posición y velocidad de los distintos ejes, y la fuerza y el par de los dispositivos de sujeción.
- Variables externas que proporcionan al robot la capacidad de captación de las características de su entorno.

Los robots forman parte de la mayoría de los sistemas de fabricación flexible que se analizan a continuación. Debido a ello deben tener capacidad de comunicación con los diversos sistemas electrónicos de control (autómatas programables, sistemas de control numérico, computadores industriales, etc.) para lo cual su procesador de comunicaciones ha de utilizar el mismo tipo de red de control que ellos. Las redes de control forman parte de las Comunicaciones Industriales que se analizan en el apartado 9.3.

9.2.4.6 Sistemas de fabricación flexible

Los sistemas de fabricación flexible (*Flexible manufacturing system*) se pueden definir de dos formas diferentes según se ponga énfasis en la producción o en la automatización. En el primer caso se definen como un conjunto de máquinas e instalaciones, enlazadas entre sí mediante un sistema de transporte y control, que es capaz de producir una variedad de productos dentro de una gama sin necesidad de interrumpir el proceso de fabricación para realizar una

readaptación de los elementos que lo constituyen [FERR 88]. En el segundo caso un sistema de fabricación flexible se define como un sistema controlado por un computador central que conecta varios centros o estaciones de trabajo informatizadas mediante un sistema automático de manipulación de materiales [PARR 93].

Una instalación de fabricación flexible posee [MALE 91], en general:

- Equipos de producción que realizan automáticamente el cambio de piezas y herramientas, lo que les permite trabajar sin operarios a pié de máquina durante largos períodos de tiempo.
- Sistemas de manutención y transporte automáticos de piezas y herramientas, tanto entre máquinas como entre ellas y los almacenes.
- Entrada al azar de un conjunto de piezas distintas adecuadamente identificadas dentro de una gama, más o menos amplia, predeterminada asociada a la tecnología de grupos.
- Un sistema de monitorización y control informatizado que coordina todo el proceso.
- Un sistema de gestión de materiales, máquinas y herramientas que permite:
 - Fabricación justo a tiempo conocida como JIT (*Just In Time*).
 - Inspección de la producción.
 - Diagnóstico y mantenimiento preventivos.

Por todo ello el concepto de sistema de fabricación flexible presenta diversas variantes como son los módulos, las células o celdas, las líneas y el taller de fabricación flexible que se analizan a continuación.

9.2.4.6.1 Módulo de fabricación flexible

Un módulo de fabricación flexible conocido por las siglas FMM (*Flexible Manufacturing Module*) es una máquina controlada por computador que puede fabricar piezas diferentes sin necesidad de que un operario realice tareas de cambio de elementos. Un ejemplo típico es una máquina-herramienta con control numérico (CNC), que posee un almacén automático de herramientas. Algunos autores lo denominan célula de fabricación flexible.

9.2.4.6.2 Célula de fabricación flexible

Una célula de fabricación flexible conocida por las siglas FMC (*Flexible Manufacturing Cell*) es un conjunto de máquinas-herramienta capaz de mecanizar total o casi totalmente una cierta categoría de piezas y de realizar control de calidad sobre ellas. Está asociada a la tecnología de grupos y se caracteriza por:

- Cada máquina está dotada de un sistema de control numérico por computador (CNC) y posee un almacén automático de herramientas y piezas.
- Posee almacenes intermedios (*Buffers*) entre máquinas para garantizar la autonomía durante, por lo general, varias horas.
- Posee un computador que coordina los elementos de mecanizado, manutención y transporte entre las máquinas.

En la figura 9.30 se representa el esquema de bloques de una célula de fabricación flexible que posee dos máquinas de control numérico, un robot y un carrusel de piezas. La figura 9.31 es una célula de fabricación flexible realizada con un robot y dos sistemas de control numérico.

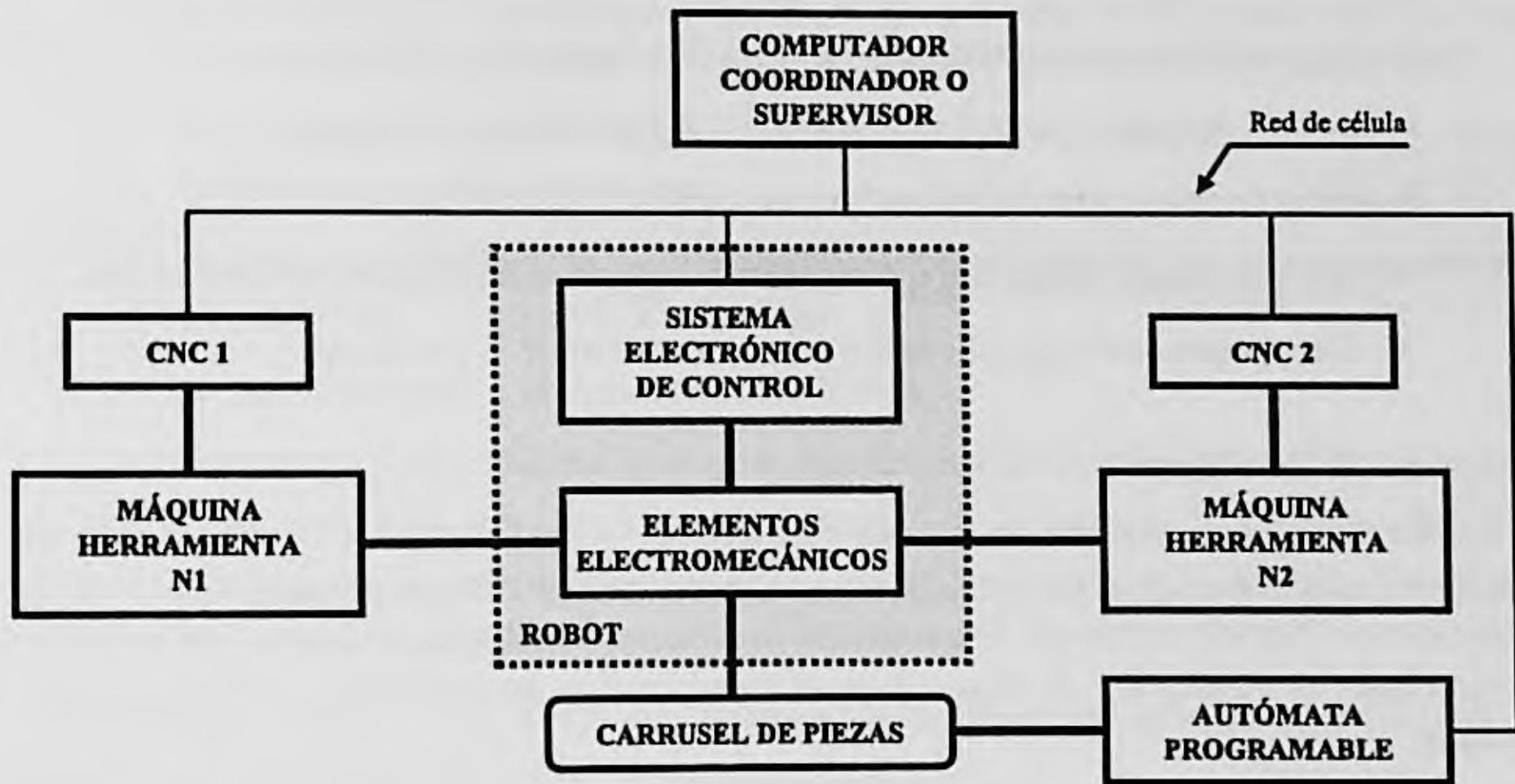


Figura 9.30. Esquema de bloques de una célula de fabricación que tiene dos sistemas CNC, un robot y un carrusel de piezas.

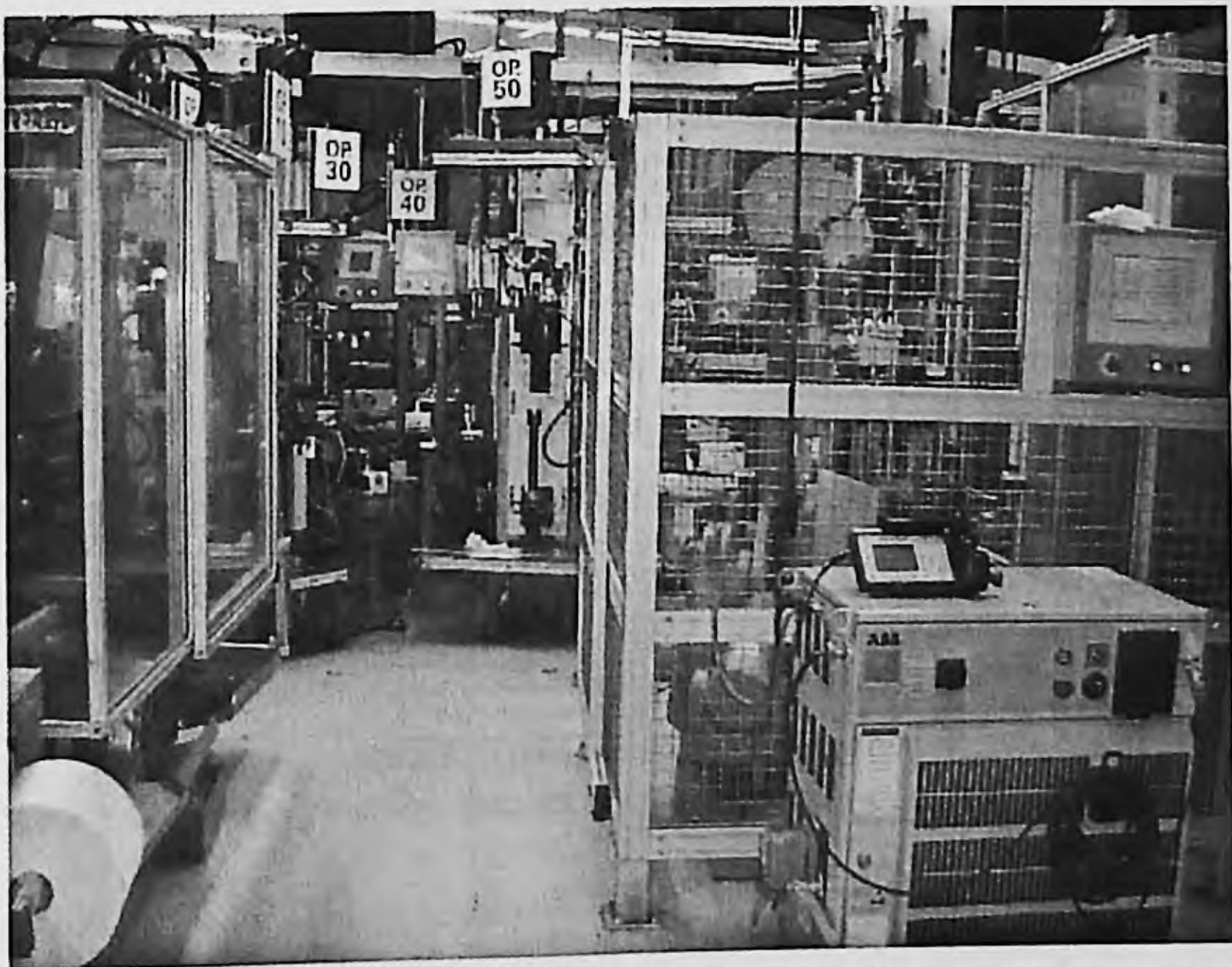


Figura 9.31. Célula de fabricación flexible realizada con un robot y dos sistemas de control numérico (cortesía de GKN Driveline Vigo).

9.2.4.6.3 Línea de fabricación flexible

Una línea de fabricación flexible (*Flexible Manufacturing Line*) está formada por un conjunto de células de fabricación flexible relacionadas entre sí mediante un sistema de transporte de piezas adecuadamente identificadas. En general se caracteriza por:

- Poseer un almacén automatizado en línea de piezas y herramientas.
- Admitir la llegada al azar a su entrada de un gran número de piezas.
- Poseer un computador coordinador que ejecuta:
 - Un programa de gestión que asigna cada pieza a la máquina más adecuada.
 - Un programa de planificación y programación de la producción.

9.2.4.6.4 Taller flexible

Recibe también el nombre de sistema de fabricación flexible, conocido como FMS (acrónimo de *Flexible Manufacturing System*) porque tiene integradas dentro de la filosofía de fabricación flexible todas las funciones de una planta de fabricación como son la recepción de materiales, la inspección, el almacenaje, el transporte, la mecanización, el montaje y la distribución.

Un sistema de fabricación flexible se caracteriza por producir cualquier tipo de pieza que se precise y para ello ejecuta programas de control y gestión de gran complejidad. Su implementación implica la normalización de los productos y de los medios de producción, lo que hace que solo esté al alcance de empresas líderes en su sector.

En realidad la fabricación flexible es una filosofía de producción que afecta a toda la empresa, porque hace necesario un cambio de estrategia empresarial en lo que se refiere a:

- La estrategia de mercado porque permite:
 - Truncar rápidamente el ciclo de vida de un producto.
 - Reforzar la obtención de innovaciones tecnológicas en los productos.
 - Aumentar la complejidad del producto.
 - Ampliar la gama en distintos segmentos del mercado.
 - Atender las peticiones de “personalización” de los productos por parte de los clientes.
 - Fragmentar el mercado.
- La ingeniería de producto porque permite:
 - Integrarla con la ingeniería de fabricación o proceso.
 - Implantar la tecnología de grupos.
- La estrategia de producción porque permite:
 - Trabajar contra pedido en lugar de la clásica producción por programa de previsión de ventas.
 - Implantar la gestión justo a tiempo.

- La ingeniería de proceso porque necesita integrar técnicos en proceso con técnicos en informática, comunicaciones, electrónica, circuitos neumáticos e hidráulicos, etc.
- Las finanzas porque:
 - Los sistemas de costes pasan de tener un componente variable muy alto a ser prácticamente fijos y, por lo tanto, a ser muy sensibles al volumen de producción.
 - Los procedimientos de análisis de inversiones y de su amortización experimentan también un cambio importante.
- Las relaciones industriales porque puede ocasionar problemas laborales importantes si no es asumida por el personal debido a que:
 - Reduce los puestos de mano de obra directa.
 - Aumenta las necesidades de formación, reconversión, y modificación de estructuras y funciones.
 - Necesita, en algunos casos, la contratación de expertos en condiciones más favorables de las usuales.

Por todo ello la fabricación flexible se debe implantar de forma progresiva y con una adecuada información, formación y adecuación del personal.

Se puede por lo tanto concluir que la fabricación flexible es un proceso que debe implicar a toda la organización, que debe estar motivada en todos sus estamentos y muy especialmente en la alta dirección.

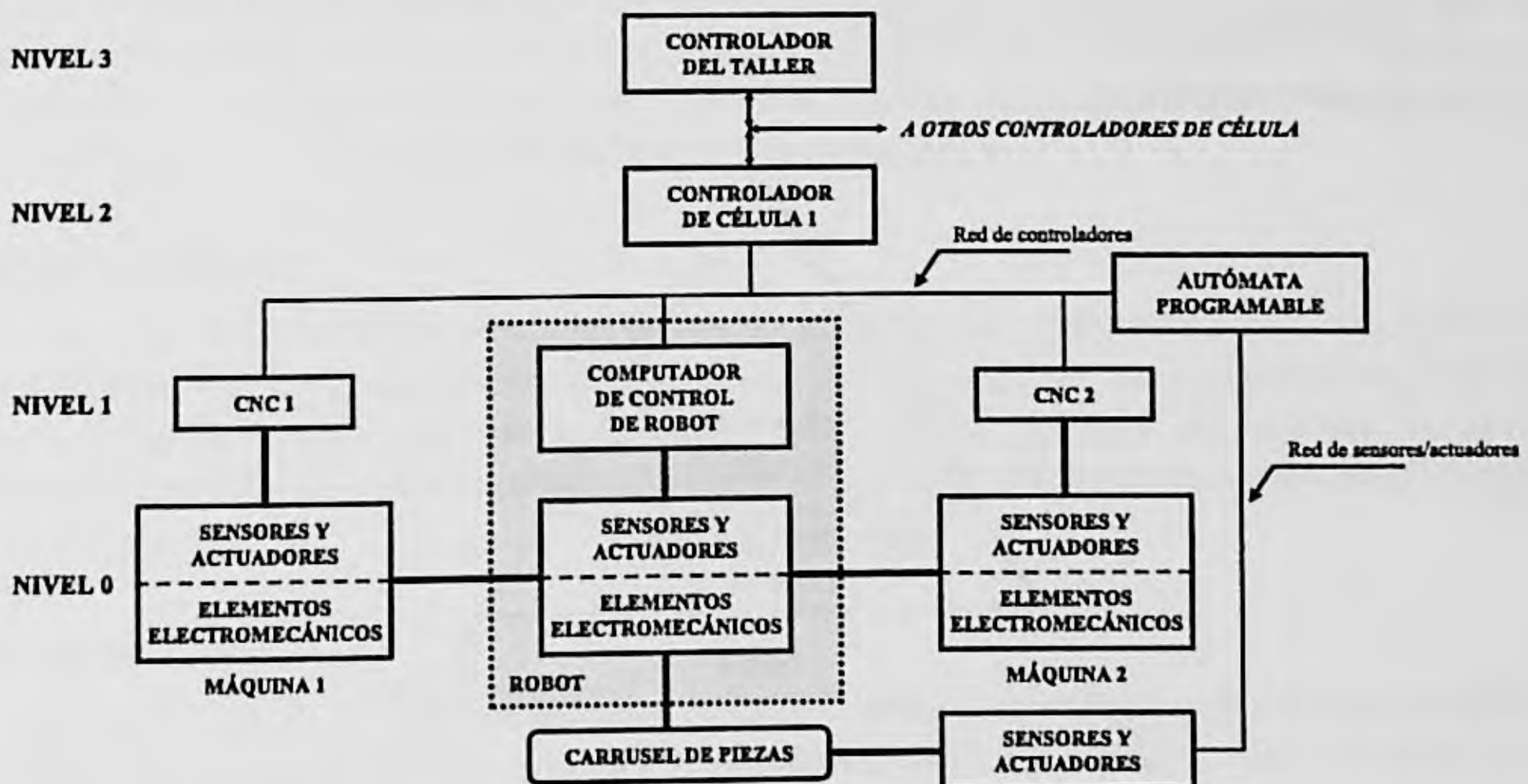


Figura 9.32. Ejemplo de niveles de la estructura funcional de un sistema de fabricación flexible.

La estructura funcional de un sistema de fabricación flexible se puede considerar dividida en niveles, tal como se indica en el ejemplo de la figura 9.32, que representa un sistema de fabricación flexible formado por varias células de fabricación coordinadas por un controlador de taller. El nivel 0 está formado por los sensores y actuadores de las diferentes máquinas. Cada

máquina posee un sistema electrónico de control y su conjunto constituye el nivel 1. Las diferentes máquinas constituyen una célula cuyo controlador forma parte del nivel 2. Finalmente un controlador de taller, que constituye el nivel 3, coordina entre sí las distintas células.

9.2.5 Fabricación integrada por computador. Pirámide CIM

9.2.5.1 Introducción

De lo expuesto en el apartado 9.2.4.6 se deduce la necesidad de integrar los procesos de producción (diseño, ingeniería y fabricación) con los de gestión de la empresa. Se obtiene así la fabricación integrada por computador, conocida como CIM (acrónimo de *Computer Integrated Manufacturing*) [SIEM 91] [REHG 05] [LEON 00], que forma parte en general de la estrategia de una empresa industrial que integra, en mayor o menor medida mediante la utilización adecuada de los computadores, todas las áreas de la empresa:

- Ordenes de entrada.
- Control de inventarios.
- Planificación de necesidades de materiales.
- Diseño del producto y del proceso.
- Simulación.
- Planificación de la fabricación.
- Automatización de la producción.
- Control de calidad.



Figura 9.33. Pirámide CIM.

- Ensamblado automático.
- Control de ventas.

La división en niveles de la estructura funcional de un sistema de fabricación flexible, descrita en el apartado 9.2.4.6, propicia la representación de un sistema de fabricación integrada por computador mediante la pirámide de la figura 9.33, que se denomina pirámide CIM y está formada por cinco o seis niveles que se describen seguidamente.

Nivel de Proceso

En este nivel se adquieren datos del proceso mediante sensores situados en él y se actúa sobre él mediante actuadores. Los primeros se transfieren a los sistemas que forman parte del nivel de estación inmediatamente superior para que ejecuten los algoritmos de control y que, teniendo en cuenta los resultados obtenidos, envíen las órdenes oportunas a los actuadores. Por lo tanto este nivel es el encargado de la comunicación de los diferentes controladores del nivel inmediatamente superior de estación con los dispositivos de campo (*Field devices*), que es el nombre utilizado para hacer referencia tanto a los sensores como a los actuadores porque interaccionan, de forma directa, con el proceso productivo [que suele ser denominado campo (*Field*)].

Nivel de Estación

En este nivel se elabora la información procedente del nivel de proceso inferior y se informa al usuario de la situación de las variables y alarmas. Forman parte de él los diferentes sistemas electrónicos de control utilizados en cada máquina como son los autómatas programables, los sistemas de control numérico (CNC), los robots, los computadores industriales etc., descritos en el apartado 9.2.4.4, que reciben por ello el nombre genérico de controladores de máquinas. De ahí que a este nivel se le denomine también nivel de máquina.

Nivel de Célula

En este nivel se realiza la coordinación de las máquinas pertenecientes a una célula de fabricación. Las tareas generadas en el nivel superior de área o de fábrica se descomponen en un conjunto de operaciones más sencillas que se trasladan, de forma sincronizada, hacia los subprocesos del nivel inferior (almacenamiento y transporte, fabricación, ensamblado, control de calidad, etc.).

Nivel de Área

En este nivel se coordinan entre sí las diferentes células que constituyen una línea de fabricación. Solo existe en instalaciones de una cierta complejidad y por ello a veces no se incluye en la pirámide CIM.

Nivel de Fábrica

En este nivel se realiza el secuenciamiento de las tareas y la administración de los recursos. Suele ser el responsable de la gestión de una planta o fábrica concreta. Las principales actividades se centran en la planificación y control de la producción. En él se diseñan y definen los

procesos de fabricación y su secuencia concreta, se gestiona el material y los recursos (máquinas, programas, etc.) necesarios para la obtención del producto final, se planifican las labores de mantenimiento, etc.

Nivel de Empresa

En este nivel se lleva a cabo la gestión e integración de los niveles inferiores. En él se consideran principalmente los aspectos de la empresa desde el punto de vista de su gestión global (compras, ventas, comercialización, investigación, objetivos estratégicos, planificación a medio y largo plazo, etc.).

9.2.5.2 Implantación del modelo CIM

El modelo CIM, que se acaba de describir, se puede implantar en la realidad tal como se indica en la figura 9.34.

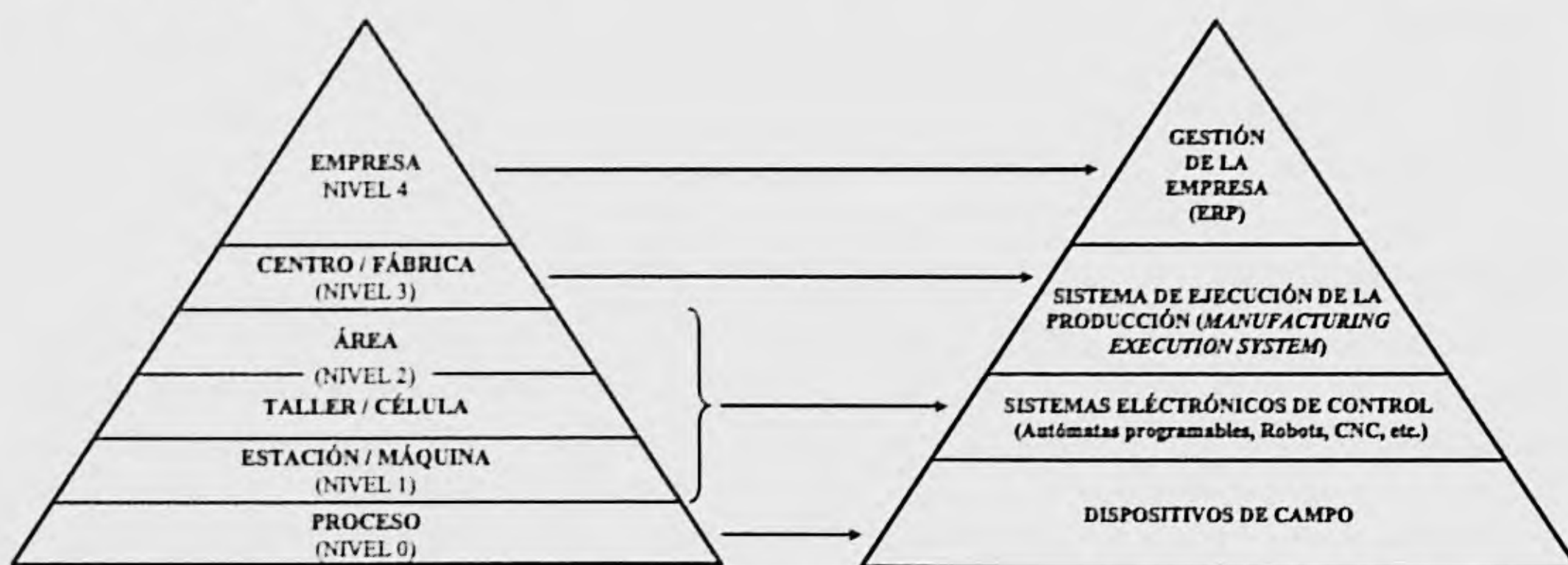


Figura 9.34. Relación entre el modelo teórico de la pirámide CIM y su implantación real en la empresa.

El nivel de proceso, tal como se indica anteriormente, está formado por los dispositivos de campo (sensores y actuadores) que interactúan de forma directa con el proceso productivo. A su estudio se dedica el capítulo 7, que se combina con el capítulo 8 dedicado a las unidades de interfaz entre los autómatas programables y el proceso.

El conjunto formado por el nivel de estación/máquina, taller/celda y área está constituido por un conjunto de sistemas electrónicos de control que se describen someramente en el apartado 9.2.4.4, excepto los autómatas programables a los cuales se dedica principalmente el resto de los capítulos de este libro.

El nivel de fábrica, en el que se realiza el secuenciamiento de las tareas y la administración de los recursos, es el último en el que se han comenzado a utilizar, en los últimos años, programas que tienen como objetivo contribuir a ejecutar eficientemente el plan de fabricación de una planta. Por eso, los suministradores de equipos tanto hardware como software orientados a la fabricación integrada por computador, como Siemens, desarrollaron programas específicos que reciben el nombre de MES (acrónimo de *Manufacturing Execution System*). Los sistemas MES son principalmente sistemas informáticos en línea que proporcionan herramientas para llevar

a cabo las distintas actividades de la administración de la producción. La importancia de los sistemas MES ha hecho que se haya constituido una asociación para impulsarlos, denominada *Manufacturing Execution Systems Association* (MESA), que ha elaborado una lista descriptiva de las tareas que pueden ser incluidas en un sistema MES. Entre ellas cabe citar:

- Ubicación y estado de los recursos.
- Calendario detallado de operaciones.
- Unidades de despacho de producción.
- Control de documentos.
- Adquisición de datos.
- Administración del trabajo.
- Administración de la calidad.
- Gestión del mantenimiento.
- Administración de procesos.
- Trazabilidad de los productos.
- Análisis de rendimiento.

La asociación MESA divide las funciones de un sistemas MES en básicas (directamente asociadas con la gestión de la producción) y complementarias. Las principales funciones básicas son:

- Interfaz con el sistema de planificación de la producción.
- Gestión de las órdenes de trabajo.
- Gestión de las estaciones de trabajo.
- Trazabilidad y gestión de inventarios.
- Gestión de movimiento de materiales.
- Adquisición de datos.
- Administración de sucesos imprevistos.

Las principales funciones complementarias son:

- Gestión del mantenimiento.
- Sistemas de control del personal.
- Control estadístico de procesos.
- Gestión integral de la calidad (aseguramiento de la calidad).
- Análisis de rendimientos.
- Gestión documental.
- Trazabilidad de los productos.
- Gestión de proveedores.

En la figura 9.35 se muestra de forma gráfica el modelo de la asociación MESA que se acaba de describir.

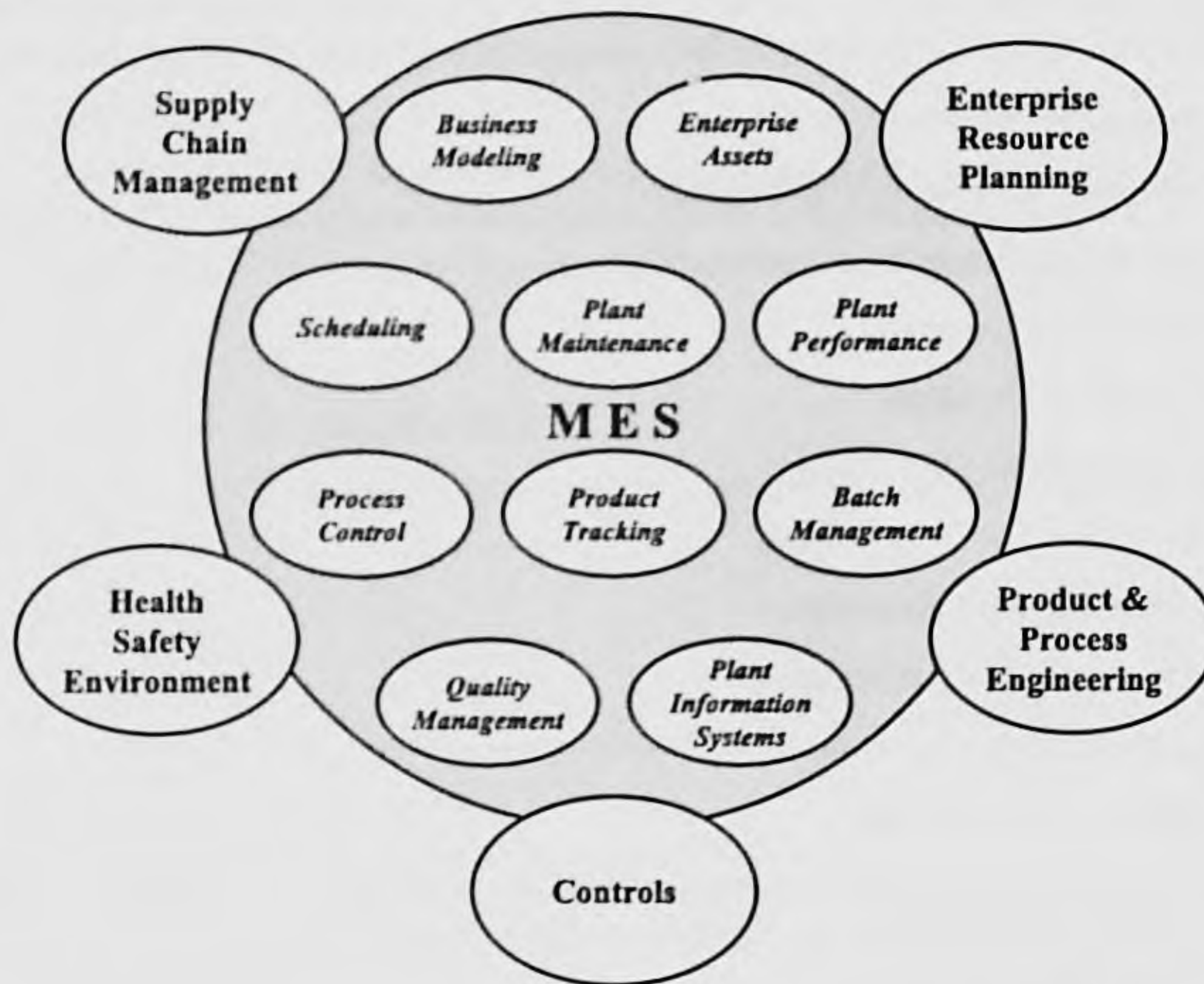


Figura 9.35. Modelo de un sistema MES desarrollado por la asociación MESA.

Un ejemplo de sistema MES es el conjunto de programas SIMATIC IT de Siemens, que se representa gráficamente en la figura 9.36. SIMATIC IT está formado por el siguiente subconjunto de programas denominados componentes de SIMATIC IT:

- **Conjunto de componentes de producción (*Production Suite Components*)**
 - **Gestor de órdenes de fabricación (*Production Order Manager*)**
Edita, transfiere y monitoriza las órdenes de fabricación.
 - **Gestor de materiales (*Material Manager*)**
Define los materiales según la norma S95. Recopila en tiempo real toda la información relativa a los materiales y lleva a cabo la trazabilidad del producto a través de todo el proceso productivo.
 - **Gestor de personal (*Personnel Manager*)**
Proporciona las funciones necesarias para la gestión del personal involucrado en el proceso de fabricación. Permite la división de los operarios en grupos y la asignación de turnos de trabajo a los mismos.
- **Histórico (*Historian*)**
Captura los datos de la planta de producción proporcionados por el laboratorio (*Unilab*) y los almacena para realizar informes, certificaciones, estadísticas, monitorización de rendimientos, etc.

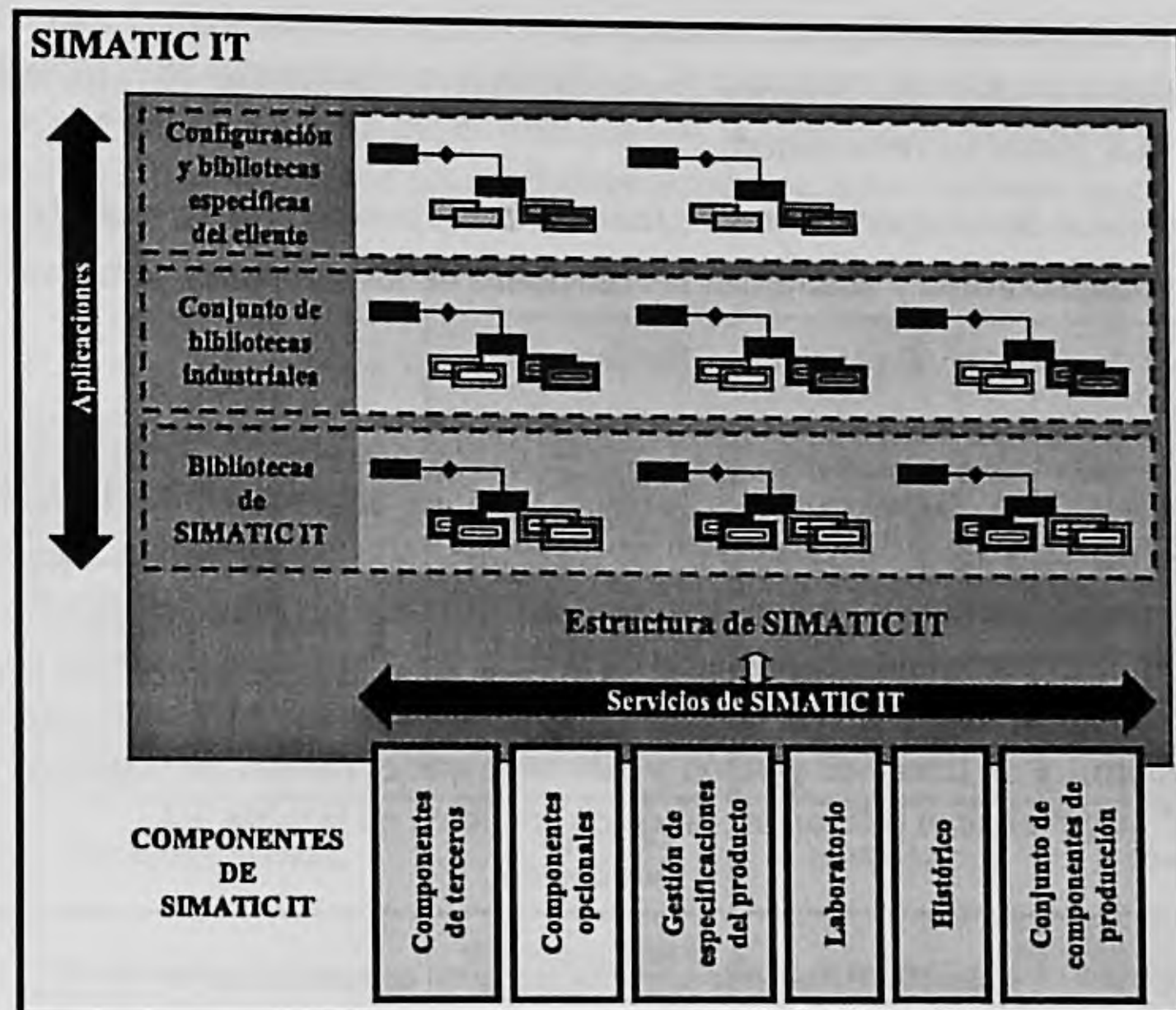


Figura 9.36. Representación gráfica del sistema MES de SIEMENS y su relación con el nivel de sistemas de control y el nivel de planificación de recursos de la empresa (ERP).

- **Laboratorio (Unilab)**

Recoge las muestras de los productos para determinar su calidad u en función del resultado toma decisiones en relación con su aceptación, rechazo o reprocesamiento.

- **Gestión de especificaciones (Interspec)**

Gestiona las especificaciones del producto establecidas por la empresa en un entorno que puede tener varias plantas de fabricación y varios idiomas.

- **Componentes opcionales**

- **Planificador de la producción (Detailed Production Schedule)**

Planifica y sincroniza simultáneamente las máquinas, la mano de obra y los materiales y proporciona un plan optimizado de fabricación.

- **Servidor (Server)**

Gestiona el modelo de integración de datos para realizar la certificación del punto de acceso del servicio (SAP).

- **Gestor de informes (Report Manager)**

Facilita la realización de todos los informes necesarios.

Además SIMATIC IT contiene un conjunto de programas que enlaza los componentes antes citados con las aplicaciones. Dicho conjunto, denominado “Estructura de SIMATIC IT” (*SIMATIC IT Framework*) esta formado por los siguientes programas:

- **Gestor de informes (Report Manager)**
Realiza el modelo que describe la capacidad de las máquinas, del sistema y del personal de la planta de producción.
- **Registrador de las operaciones de fabricación (Production Operation Recorder)**
Combina, coordina y sincroniza las funciones de los diferentes componentes.
- **Gestor de equipos (Equipment Manager)**
- **Servicios (Services)**

De la pirámide CIM, representada en la figura 9.33, así como de su implantación real en la empresa, puesta de manifiesto en la figura 9.34, se deduce que la automatización integrada de la producción se realiza mediante un conjunto de dispositivos y sistemas electrónicos de proceso de datos que realizan las funciones indicadas en la tabla 9.1. Pero las actividades correspondientes a cada uno de los niveles 0 a 3 de la tabla 9.1 se diferencian por la relación entre la cantidad de tareas de control y de tareas de gestión y ello hace que el tiempo de respuesta mínimo y la disponibilidad mínimas sean diferentes, tal como se indica en la tabla 9.2.

Nivel 4	Gestión global de la empresa mediante computadores
Nivel 3	Gestión automatizada de la producción de la fábrica mediante computadores
Nivel 2	Gestión automatizada de las células/áreas de fabricación mediante autómatas programables y computadores industriales
Nivel 1	Control de máquinas mediante autómatas programables, sistemas de control numérico (CNC), etc.
Nivel 0	Medidas de variables (mediante sensores) y acciones sobre el proceso (mediante actuadores)

Tabla 9.1. Actividades realizadas por los sistemas electrónicos de control utilizados en cada uno de los niveles de la pirámide CIM.

En los niveles superiores de la pirámide CIM se trabaja frecuentemente con grandes volúmenes de datos, aunque el tiempo de respuesta no es en general crítico y se sitúa entre pocos segundos hasta minutos e incluso horas o días. Por el contrario, los sistemas electrónicos de control utilizados en los niveles más cercanos a las fases de producción trabajan en tiempo real (descrito en el apartado 9.2.4.4) y debido a ello se exigen tiempos de transmisión mucho menores y, sobre todo, un comportamiento determinista de las comunicaciones, aunque los volúmenes de información a transmitir son en general menos elevados.

De la pirámide CIM se deduce la necesidad de disponer de un sistema de comunicaciones en cada uno de los niveles y entre los diferentes niveles para lograr la integración de los procesos mediante la información. Dicho sistema ha evolucionado progresivamente a partir de 1990 y ha dado lugar al área de las Telecomunicaciones denominada “Comunicaciones Industriales” que se estudia en el apartado 9.3 a continuación.

Los autómatas programables son, por sus características, descritas en el capítulo 1, especialmente idóneos para actuar como controladores de máquinas, de células e incluso de un área o fábrica. Por ello deben tener un papel relevante en la fabricación integrada por computador (CIM) y tienen que disponer de los adecuados recursos de comunicaciones que se describen en el apartado 9.4.

NIVEL	TIPO DE SISTEMA ELECTRÓNICO DE CONTROL	PÁRAMETRO		
		<i>Tiempo de respuesta</i>	<i>Relación (%) de tareas Gestión/Control</i>	<i>Disponibilidad (Availability) exigible (%)</i>
4	Computador de planta	De días a segundos	95-100/0-5	> 60
3	Controlador de área	De minutos a segundos	90-95/5-10	70-80
2	Controlador de célula	De segundos a milisegundos	80-90/10-20	80-90
1	Controlador de proceso	De milisegundos a microsegundos	5-10/90-95	90-95

Tabla 9.2. *Valor de la relación entre la cantidad de tareas de control y las de gestión, del tiempo de respuesta y de la disponibilidad exigible a los sistemas electrónicos de control utilizados en cada uno de los niveles de la pirámide CIM.*

9.3 Comunicaciones Industriales

9.3.1 Introducción

Se pueden definir las Comunicaciones Industriales (*Industrial Communications*) como el área de la tecnología que estudia la transmisión de información entre circuitos y sistemas electrónicos utilizados para llevar a cabo tareas de control y gestión del ciclo de vida de los productos industriales.

Las Comunicaciones Industriales deben, por lo tanto, resolver la problemática de la transferencia de información entre los equipos de control del mismo nivel y entre los correspondientes a los niveles contiguos de la pirámide CIM (Figura 9.33). En la década de 1980, las Comunicaciones Industriales comenzaron a realizarse mediante conexiones punto a punto (descritas en el apartado A1.2 del apéndice 1) para facilitar la utilización de los computadores como sistemas de diseño asistido por computador del programa de control. Pero las conexiones punto a punto dan lugar a múltiples canales de comunicación con la consiguiente complejidad. Por ello para llevar a cabo dicha transferencia de información de la manera más eficaz posible, los equipos deben compartir canales de comunicación, lo cual incentivó el desarrollo de redes de comunicaciones que tienen que poseer las características adecuadas.

Para estudiar las redes utilizadas en las Comunicaciones Industriales (*Industrial Networks*), que en adelante se denominan redes industriales, el lector debe conocer el modelo de la conexión de sistemas informáticos abiertos denominado OSI (acrónimo de *Open Systems Interconnection*), desarrollado por ISO (acrónimo de *International Standard Organization*). En este modelo se basan la mayor parte de dichas redes, que se describen en sucesivos apartados.

En la figura 9.37 se representa gráficamente el modelo OSI, que pone en evidencia que los nodos de una red se relacionan a través de siete niveles de diálogo, cada uno de los cuales tiene asociado un protocolo, que es un conjunto de reglas que gobiernan el intercambio de datos entre las entidades del mismo nivel de cada nodo. Por ejemplo, en el nivel de aplicación son entidades los programas que, en un sistema ofimático, gestionan el correo electrónico y los que controlan, en un entorno industrial, la transferencia de información entre un autómatas programable y sus módulos de interfaz de dispositivos de campo situados a distancia.

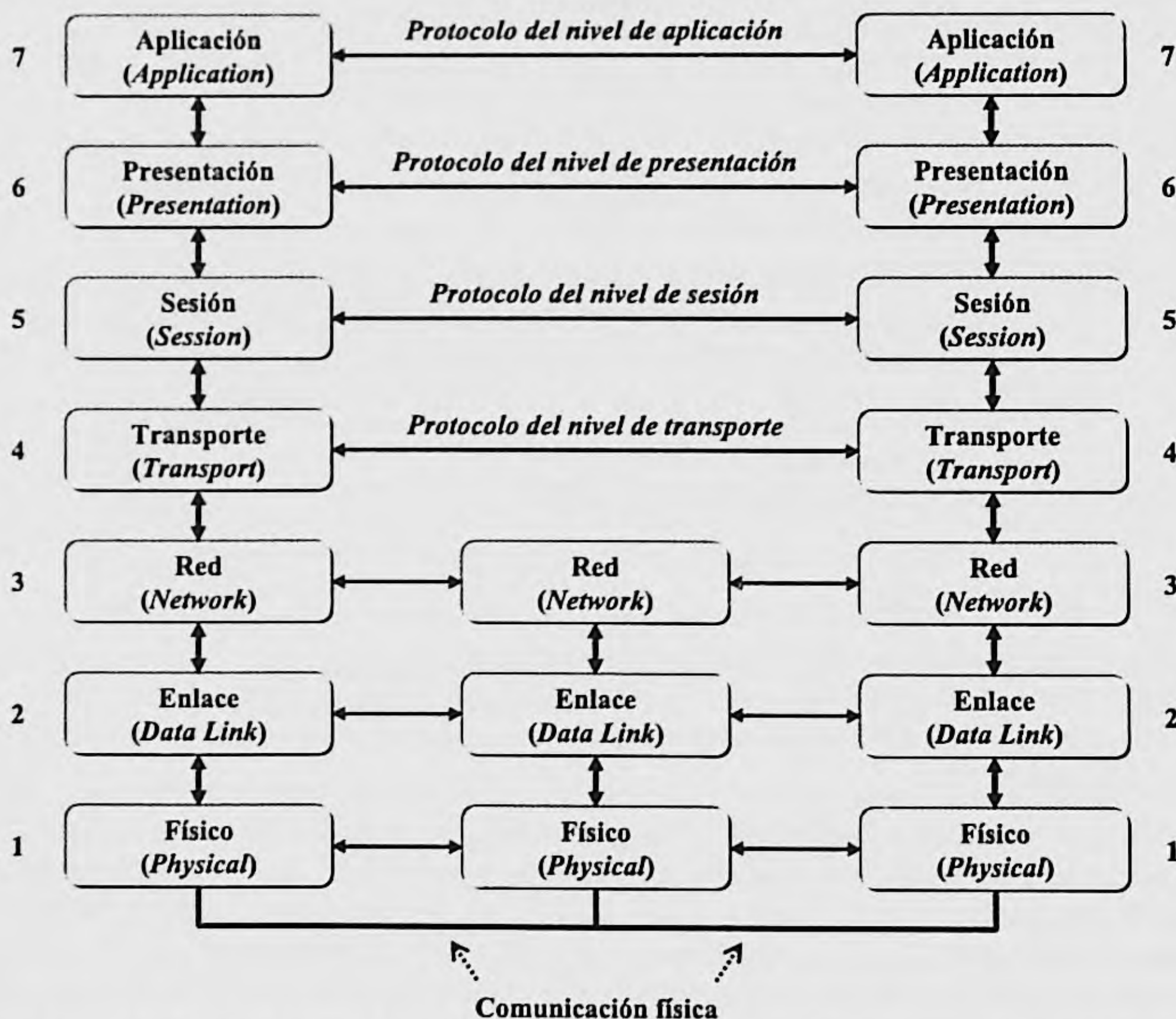


Figura 9.37. Modelo OSI desarrollado por ISO para la conexión de sistemas informáticos abiertos.

Al lector que no conozca los conceptos básicos de las redes de comunicaciones y en particular el modelo OSI, se le recomienda la lectura del apéndice 1 y la consulta de la bibliografía indicada en él.

9.3.2 Redes de Comunicaciones Industriales

9.3.2.1 Clasificación de las redes de Comunicaciones Industriales

Los diferentes tiempos de respuesta exigidos al sistema de comunicaciones de cada uno de los niveles de la pirámide CIM, indicados en la tabla 9.2, hacen que la red utilizada para implementar la comunicación en cada uno de ellos deba tener unas características específicas. Por ello, para solucionar las comunicaciones en una planta de producción se pueden adoptar teóricamente dos estrategias diferentes:

- La utilización de redes distintas, específicas para cada uno de los niveles de la pirámide CIM.
- La utilización de una red universal o una red integrada que atienda de forma diferente a los distintos tipos de transferencias de información que se deben realizar en la pirámide CIM.

El estado de la Microelectrónica a finales de la década de 1970 y principios de la de 1980, hizo que diversos fabricantes de autómatas programables comenzasen a desarrollar redes industriales de aplicación específica para los niveles bajos de la pirámide CIM y posteriormente a aplicar las redes ofimáticas en los niveles altos. Esto justifica que, en una primera aproximación y de acuerdo con la opinión de diversos expertos en el tema, las redes industriales se clasificasen en redes de datos y redes de control, tal como se indica en la tabla 9.3.

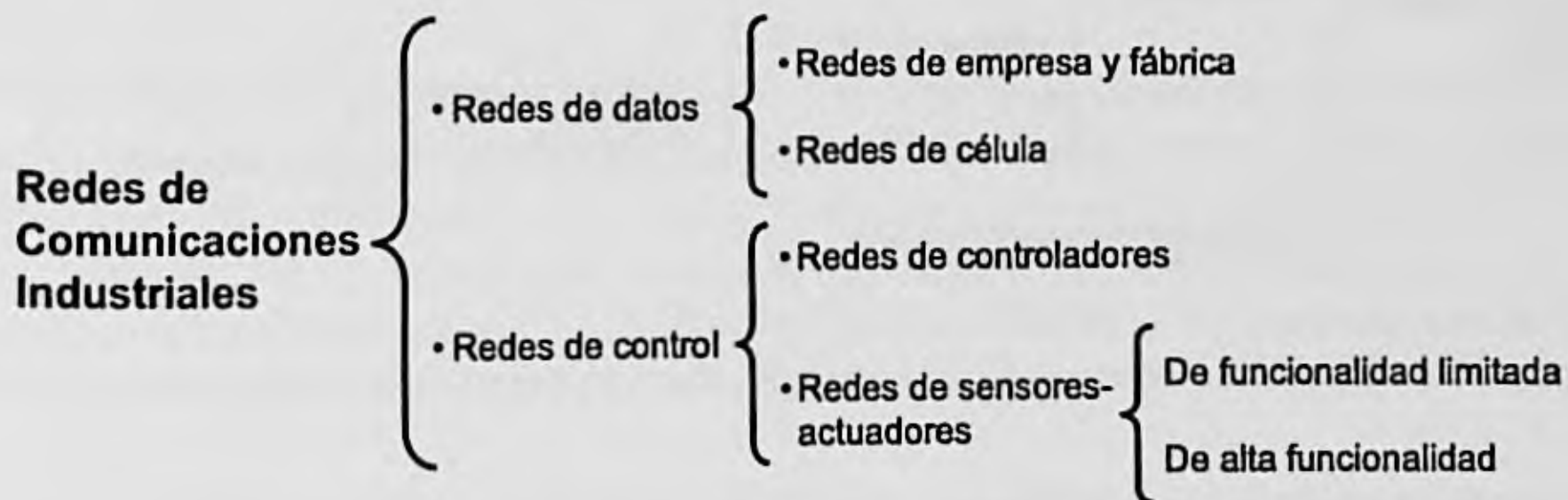


Tabla 9.3. Clasificación de las redes de Comunicaciones Industriales a finales del siglo XX.

Las redes de datos, derivadas de las redes ofimáticas, están ligadas a los niveles altos de la pirámide CIM y por ello tienen como principal objetivo transportar grandes paquetes de información de forma esporádica (baja carga), pero a elevada velocidad (gran ancho de banda) para permitir el envío rápido, a través de ellas, de una gran cantidad de datos entre un volumen potencialmente elevado de estaciones interconectadas.

Las redes de control están ligadas a los niveles bajos de la pirámide CIM y, por ello, deben ser capaces de soportar, en tiempo real (concepto descrito en el apartado 9.2.4.4), un tráfico de información formado por un gran número de pequeños paquetes procedentes de un número de estaciones proporcionalmente menos elevado que en las redes de datos.

Es conveniente resaltar que las redes de datos y las de control no son conjuntos claramente disjuntos, debido a que cuando las redes de datos se utilizan en el nivel de área o de célula de la

pirámide CIM, hay una tendencia a dotarlas de características que son más propias de las redes de control.

Por otra parte, las redes de datos y las de control pueden compartir las especificaciones de las capas física y de enlace de datos del modelo OSI. Surge así el concepto de “familia” de redes industriales, descritas en el apartado 9.3.2.4, en la que cada uno de sus elementos resuelve las transferencias de información en un determinado nivel de la pirámide CIM.

El extraordinario avance de la Microelectrónica hizo que a partir de finales de la década de 1990 diversos fabricantes de equipos de automatización industrial se planteasen el desarrollo de una red única que diese respuesta a las necesidades de comunicación de los diferentes niveles de la pirámide CIM (Tabla 9.2). Nacieron así las redes industriales universales que, por estar basadas en la red de datos Ethernet (véanse los apartados A1.3.3.3.2 y A1.3.6 del apéndice 1) reciben en general la denominación de red Ethernet Industrial (*Industrial Ethernet*). Este hecho hace necesario realizar una nueva clasificación de las redes de Comunicaciones Industriales, representada en la tabla 9.4, que las clasifica en redes de aplicación específica y redes universales.

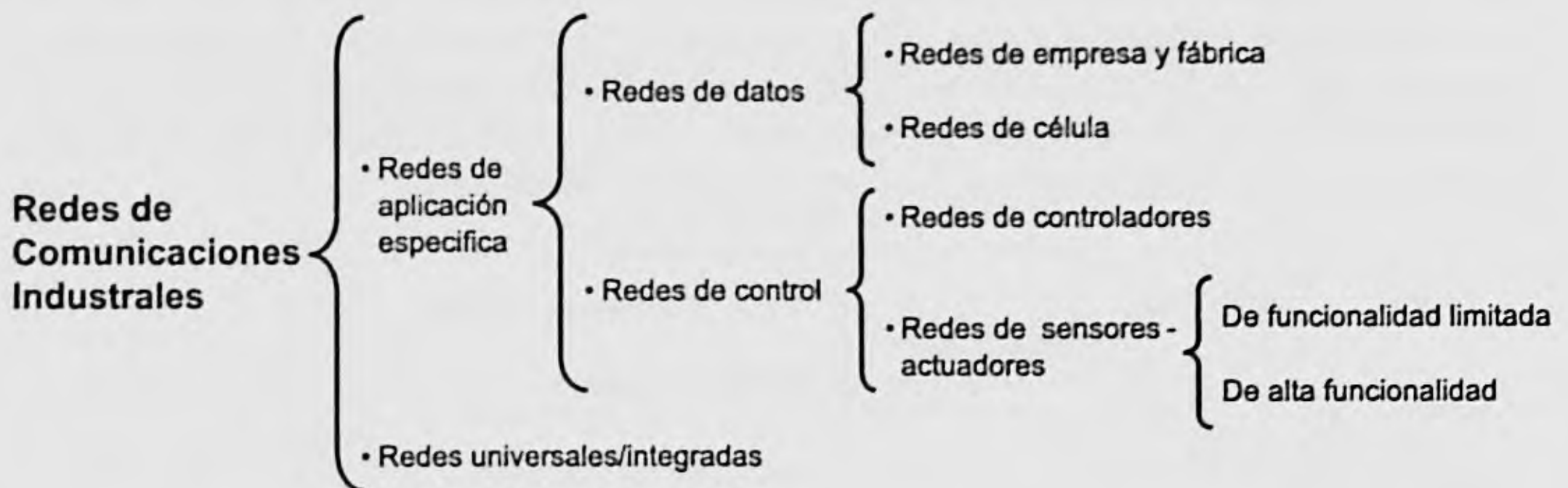


Tabla 9.4. Clasificación de las redes de Comunicaciones Industriales en el siglo XXI.

En sucesivos apartados se analizan en primer lugar las diferentes redes de datos y de control desarrolladas a finales del siglo XX, así como las familias de redes utilizadas para implementar las Comunicaciones Industriales modernas. Finalmente se describen las redes Ethernet Industrial que iniciaron su comercialización a partir de comienzos del siglo XXI.

9.3.2.2 Redes de datos

Como se indica en el apartado anterior, se suelen considerar redes de datos (*Data networks*) las dedicadas al establecimiento de las comunicaciones entre los equipos informáticos que conforman los niveles de empresa, fábrica, área y, en ocasiones, de célula de la pirámide CIM. En este apartado se analizan las principales características de esta clase de redes industriales.

9.3.2.2.1 Redes de empresa y de fábrica

Los sistemas de control de los niveles de empresa y de fábrica ejecutan, entre otras, las siguientes herramientas informáticas:

- Programas dedicados a la planificación de recursos de la empresa, conocidos como ERP (acrónimo de *Enterprise Resource Planning*).
- Programas dedicados a la gestión de los sistemas de ejecución de la fabricación, que se suelen conocer como MES (acrónimo de *Manufacturing Execution Systems*) (véase apartado 9.2.5.2).
- Programas de diseño, simulación, ingeniería y fabricación asistidos por computador denominados CAD/CAM/CAE (acrónimos de *Computer Aided Design/ Manufacturing/ Engineering*).
- Herramientas de aplicación general que permiten el trabajo en grupo (*Groupware*) del personal de todas las áreas de la empresa.

Cuando los sistemas enlazados mediante una red de empresa y fábrica están situados en la misma planta, o en emplazamientos relativamente próximos, se utiliza una red de área local que se suele denominar LAN (acrónimo de *Local Area Network*). La red de área local más utilizada en empresas industriales es la red “*Ethernet*” conmutada que combina la técnica de acceso al medio basada en la detección de colisiones (CSMA/CD) (véase el apartado A1.3.3.3.1 del apéndice 1), con la topología en estrella mediante conmutadores (*Switches*) (véase apartado A1.3.4 del apéndice 1) y el conjunto de protocolos de red y transporte TCP/IP (acrónimo de *Transmission Control Protocol/Internet Protocol*). Se estima que actualmente lo utilizan el 80% de las comunicaciones en este sector. Para estudiar las características básicas de las redes de área local, se remite al lector al apartado A1.3.3.3 del apéndice 1 y a la bibliografía indicada en él.

Para comunicar entre sí las distintas sedes de una empresa, situadas por lo general en emplazamientos distantes, se utilizan redes de área metropolitana y extensa, denominadas respectivamente, MAN (acrónimo de *Metropolitan Area Network*) y WAN (acrónimo de *Wide Area Network*). Un ejemplo de ellas es la red mundial de comunicaciones conocida como *Internet*. En el apartado A1.3.3.2 del apéndice 1 se describen los fundamentos de las redes de área metropolitana y extensa.

9.3.2.2.2 Redes de célula

Las redes de área local y de área extensa antes citadas no fueron diseñadas inicialmente, para satisfacer determinados requisitos que son propios o característicos del ambiente industrial, entre los que se pueden destacar:

- Funcionamiento en ambientes hostiles (presencia de fuertes perturbaciones electromagnéticas, temperaturas extremas, polvo y suciedad, etc.).
- Gran seguridad en el intercambio de datos en un intervalo de tiempo cuyo límite superior se fija con exactitud (lo que se denomina comportamiento determinista) para poder trabajar correctamente en tiempo real.
- Elevada fiabilidad y disponibilidad de las redes de comunicación, mediante la utilización de dispositivos electrónicos y/o medios físicos redundantes y protocolos de comunicación que dispongan de mecanismos avanzados para la detección y corrección de errores en la comunicación, etc.

Por ello, en las dos últimas décadas del siglo XX se acrecentó el interés por desarrollar redes de comunicación específicamente diseñadas para ambientes industriales, que integrasen sistemas heterogéneos de diferentes fabricantes, y que diesen soporte a la intercomunicación entre las operaciones del nivel de fábrica y las funciones de apoyo y gestión de la producción, correspondientes al nivel de empresa.

Fue la empresa GENERAL MOTORS la que, a mediados de los años 80, decidió desarrollar una red de comunicaciones adecuada a las necesidades de transferencia de información entre los sistemas de control de una fábrica. Dicha red fue denominada MAP (acrónimo de *Manufacturing Automation Protocol*) y su principal objetivo era uniformizar los métodos aplicados para la comunicación de los sistemas de control utilizados a nivel de planta o fábrica. Posteriormente, en el año 1986, surgió en el seno de la empresa BOEING la red de comunicaciones TOP (acrónimo de *Technical and Office Protocol*), para desarrollar un sistema orientado hacia la comunicación de datos en el área técnico/administrativa de la empresa. La red TOP fue incorporada a la red MAP, y se generó así el nuevo proyecto de red MAP/TOP, que contemplaba de esta forma también la capacidad de interconexión de los sistemas de control utilizados en el nivel de fábrica con los equipos informáticos presentes en los niveles superiores de la empresa.

La red MAP se desarrolló de acuerdo con el modelo de interconexión OSI de sistemas abiertos que se describe en el apartado A1.3.2 del apéndice 1. En su capa física y en la de control de acceso al medio se utilizaba, en especial por su comportamiento determinista, el protocolo correspondiente a la red de área local denominada bus con paso de testigo (*Token Bus*), definida en la norma IEEE 802.49 (ver apartado A1.3.3.3.2 del apéndice 1), en dos de sus variantes para cable coaxial:

- Transmisión en banda ancha, con una tasa de 10 Mbaudios, adecuada para los niveles altos de la pirámide CIM.
- Transmisión en banda base digital con una tasa de 5 Mbaudios para los niveles inferiores de la misma.

Por su parte la red TOP, dada su orientación más próxima a las redes de datos que la red MAP, permitía el uso, para la capa física y de acceso al medio, tanto de las redes *Ethernet* como la de la norma IEEE 802.5, más conocida como red de paso de testigo en anillo (*Token Ring*) que se describe en el apartado A1.3.3.3.2 del apéndice 1.

El protocolo más importante de la capa de aplicación de una red MAP, y de todas las derivadas de ella, es el conocido como MMS (acrónimo de *Manufacturing Message Specification*). Este protocolo fue diseñado con el objetivo de facilitar la monitorización y gestión de los sistemas electrónicos de control de procesos de fabricación como los sistemas de control numérico (CNC), autómatas programables, células robotizadas, computadores industriales, etc. Para ello, MMS posee un conjunto de servicios, orientados al establecimiento de relaciones de comunicación con estos sistemas, que permiten la manipulación, entre otras, de entidades conceptuales (denominadas “objetos” de forma genérica) como por ejemplo variables (que pueden ser independientes o estar estructuradas en forma de registros o matrices), programas (que pueden ser enviados, activados y finalizados de forma remota), mecanismos de comunicación y sincronización entre procesos como por ejemplo semáforos y sucesos (*events*), etc.

El protocolo MMS está especificado en dos documentos normalizados:

- En ISO/IEC 9506-1 se describen los servicios que es posible invocar, de forma remota, para actuar sobre los "objetos" MMS asociados a un determinado dispositivo.
- En ISO/IEC 9506-2 se describe la sintaxis del protocolo MMS, en términos de unidades de datos de protocolo denominadas PDU (acrónimo de *Protocol Data Unit*), por medio de las reglas de codificación ASN.1 (*Abstract Syntax Notation - One*) de la norma ISO 8824.

A pesar de sus características, adecuadas para este segmento de las Comunicaciones Industriales, la red MAP casi no se utiliza actualmente a nivel industrial, entre otros, por los siguientes motivos:

- Cubre adecuadamente los requisitos de comunicación del nivel de empresa de la pirámide CIM pero, dada la robustez de los protocolos en los que se soporta para ello, proporciona en la práctica, tiempos de respuesta en la comunicación que son relativamente elevados para los exigidos habitualmente en el nivel de planta o fábrica.
- Su especificación es tan vaga en algunos aspectos (de hecho no existe un único documento que la defina claramente) que se ha hecho muy complejo y difícil el desarrollo de interfaces, tanto a nivel físico (*hardware*) como de programación (*software*) que le den soporte, con lo que ello implica a nivel de coste en términos comerciales.

Para superar estos problemas surgieron posteriormente, en la década de 1990, diferentes proyectos de redes basadas en la red MAP que tratan de solucionar los problemas de comunicación en tiempo real (determinista) de los niveles inferiores de la pirámide CIM. En dichos proyectos, se limitaron diversas características de la red MAP como por ejemplo la complejidad de los modelos y normas seleccionadas en las diferentes capas y protocolos de comunicación, la funcionalidad de las interfaces de cada capa, y el tamaño máximo de las tramas de comunicación o bloques de información (*Frames*), descritos en el apartado A1.2.3 del apéndice 1, etc. Entre ellos, se pueden destacar las redes denominadas MAP/EPA y Mini-MAP [BEND 93] que constituyeron una solución transitoria en el camino hacia las redes industriales universales, que como se indica en el apartado 9.3.2.1, dan respuesta a las necesidades de comunicación de los diferentes niveles de la pirámide CIM (Tabla 9.2) y, por estar basadas en la red de datos Ethernet (véanse los apartados A1.3.3.3.2 y A1.3.5.2 del apéndice 1), reciben en general la denominación de red Ethernet Industrial a cuyo estudio se dedica el apartado 9.3.2.5.

9.3.2.3 Redes de control

9.3.2.3.1 Conceptos generales

Las redes de control resuelven los problemas de comunicación en los niveles de estación y proceso de la pirámide CIM, aunque también se utilizan, en ocasiones, en el nivel de célula. Suelen recibir también el nombre genérico de redes o buses de campo (*Field buses*) porque utilizan la topología de bus de las redes de área local (descrita en el apartado A1.3.3 del apéndice 1) y se utilizan en una planta (*Field*) de fabricación.

Históricamente, el desarrollo de esta clase de redes (que se produjo en la década de 1980) fue debido a la elevación de la complejidad de la automatización de los sistemas industriales, que incrementó desmesuradamente el volumen de cableado que era preciso realizar para conectar un elevado número de dispositivos sensores y actuadores (denominados en adelante disposi-

tivos de campo) a un controlador de estación o de célula (autómata programable, robot, sistema de control numérico (CNC), etc.), mediante hilos independientes (Figura 9.1).

Para resolver el problema, surgió la idea de conectar cada dispositivo o grupo de dispositivos de campo a un procesador de comunicaciones. Conectando entre sí los distintos procesadores de comunicaciones, se logra que los sensores y/o actuadores compartan un canal de comunicaciones serie, digital, bidireccional y multipunto (denominado bus de campo) a través del cual intercambian información con los sistemas electrónicos de control (Figura 9.38). Además, el procesador de comunicaciones suele tener capacidad para ejecutar funciones de autodiagnóstico que le proporcionan un cierto nivel de seguridad ante averías. Esta clase de redes son las que se conocen en la actualidad con el nombre de *redes de sensores-actuadores*.

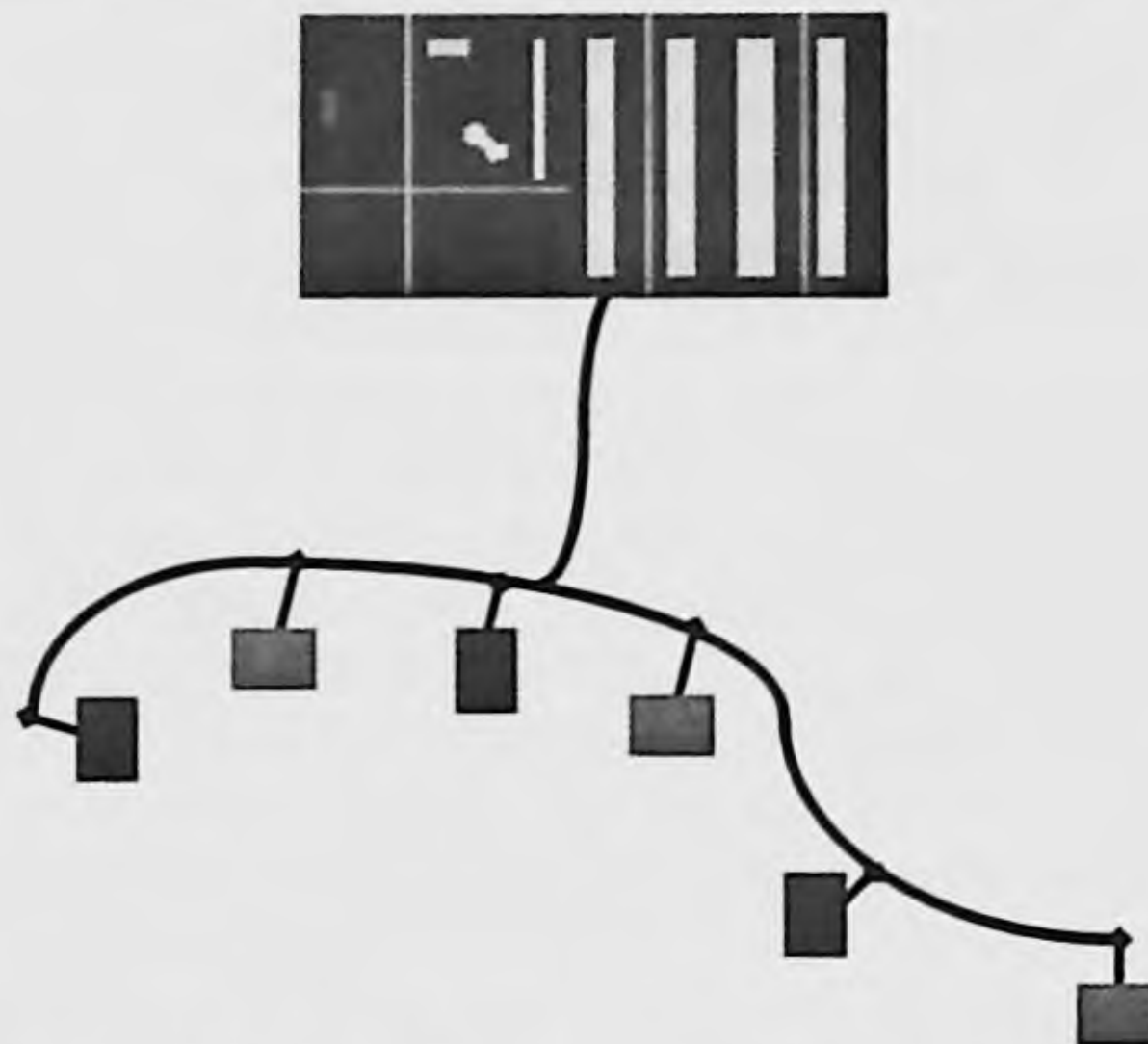


Figura 9.38. Conexión de los dispositivos de campo con un autómata programable a través de una red o bus de campo (Cortesía de Siemens).

Además, las redes de campo no tienen por qué limitarse a la conexión de sensores y actuadores con un sistema de control, sino que pueden también servir para enlazar entre sí todos los sistemas electrónicos de control del nivel de fábrica (autómatas programables de diferentes fabricantes, computadores industriales, controladores de robots, etc). Este tipo de redes se denominan *redes de controladores*.

Ambos tipos de buses de campo (de sensores-actuadores y de controladores) se diferencian de las redes de datos en que permiten la transmisión de información en pequeños paquetes (con unos tamaños que varían entre las decenas y centenas de octetos) y con unos requisitos temporales que exigen su intercambio en un intervalo de tiempo muy reducido (comprendido entre los milisegundos y las décimas de segundo).

Las redes o buses de sensores-actuadores se caracterizan, en general, por estar formados por un único sistema electrónico de control (por ejemplo un único autómata programable) y un conjunto de módulos de sensores/actuadores que deben enviarle o recibir información de él en intervalos de tiempo cuyo límite superior está acotado para poder trabajar correctamente

en tiempo real (lo que se denomina comportamiento "determinista"). Para ello utilizan diversos mecanismos de acceso al medio, entre los que cabe citar el acceso priorizado denominado principal-subordinado (*master-slave*), descrito en el apartado A1.3.3.3.1 del apéndice 1, que asegura que, como máximo, sólo un nodo transmite información en un instante determinado. El nodo o estación principal (*master*) se encarga de coordinar y distribuir los procesos de comunicación en la red, para lo cual interroga o consulta de forma cíclica (*polling*), a los nodos subordinados (*slaves*).

Los buses de controladores se caracterizan, en general, por estar formados por varios sistemas electrónicos de control que deben tener la misma capacidad de acceso a un medio de comunicación compartido, típico de las redes ofimáticas de igual a igual (*peer-to-peer*), descritas en el apartado A1.3.3.3.2 del apéndice 1. Por ello, este tipo de buses implementan en general un mecanismo de reparto del tiempo de acceso al medio mediante paso de testigo (*Token Bus*), descrito en el apartado A1.3.3.3.2 del apéndice 1), que permiten que varios nodos de la red puedan intercambiar, en el tiempo, el papel de nodo principal. Dicho mecanismo se combina con el de acceso al medio denominado principal-subordinado (*master-slave*) (apartado A1.3.3.3.1 del apéndice 1) que es el que regula la relación entre el nodo activo (el que posee el testigo en un instante dado) y el resto de los nodos de la red.

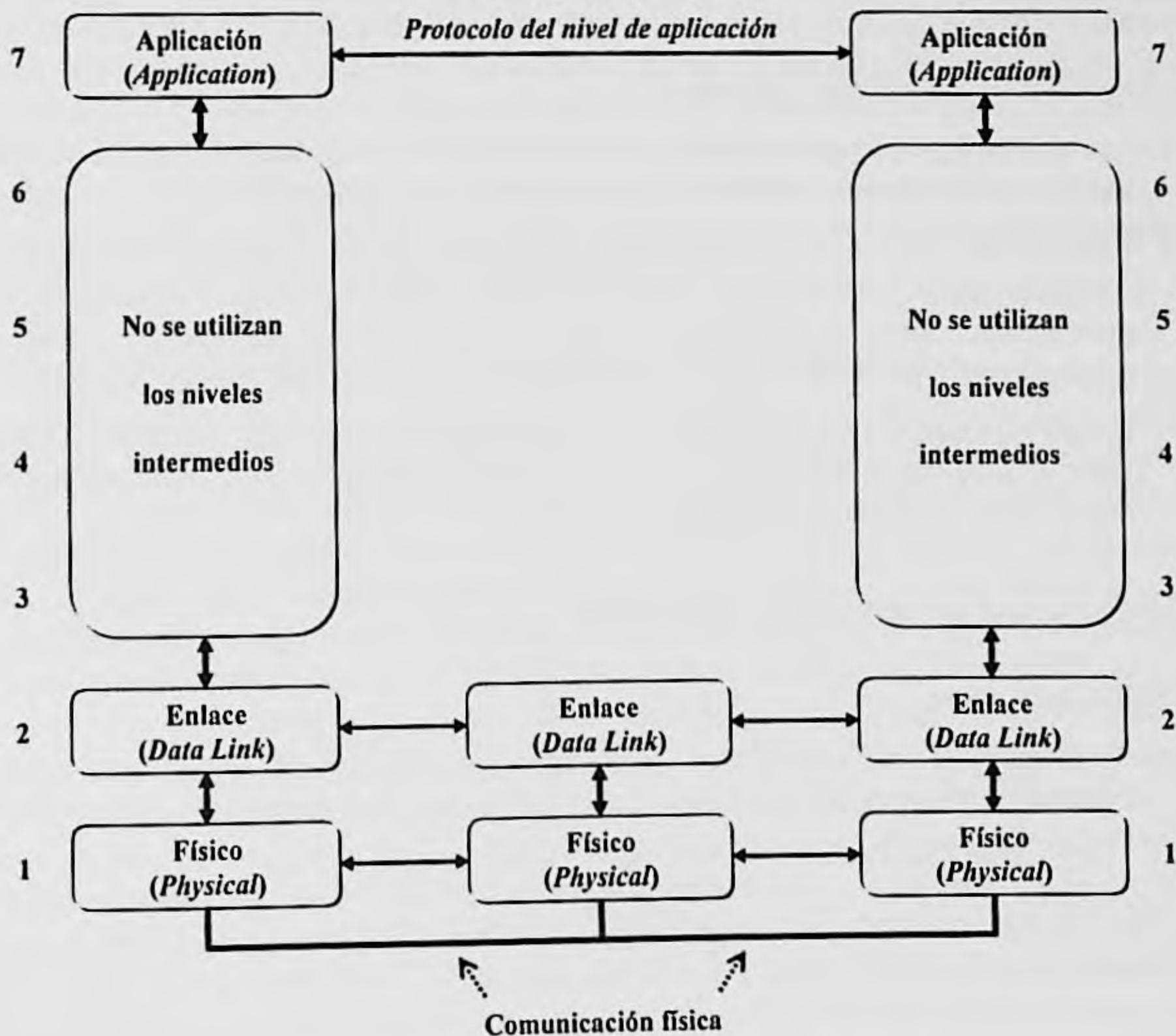


Figura 9.39. Modelo OSI simplificado utilizado en las redes de control.

Debido a sus características peculiares, citadas anteriormente, en las redes de control no es necesario emplear protocolos para las capas de red, transporte, sesión y presentación del modelo OSI. Por ello, el modelo seguido por la práctica totalidad de las redes de control, tanto propietarias como normalizadas, es el indicado en la figura 9.39.

En sucesivos apartados se analizan las características principales de las redes de controladores y de las redes de sensores-actuadores.

9.3.2.3.2 Redes de controladores

Las redes de controladores (*Controller networks*) están diseñadas para realizar la comunicación de varios sistemas electrónicos de control (autómatas programables, robots, controladores numéricos, terminales de operador (HMI), computadores industriales, etc.) entre sí.

Son por lo general redes de área local de tipo principal-subordinado (*Master-slave*) descritas en el apartado A1.3.3.3.1 del apéndice 1 que poseen varios nodos principales (*Multi-master networks*) o redes productor-consumidor (*Producer-consumer*). Los servicios de comunicación que proporcionan permiten, no sólo el intercambio más estructurado de la información entre los diferentes sistemas de control, sino, también, llevar a cabo tareas de diagnóstico, programación, carga, descarga, ejecución y depuración de los programas ejecutados en ellos.

Son redes de controladores, entre otras:

- La red MAP que, como se indica en el apartado 9.3.2.2.2, actualmente no se utiliza, aunque la capa de aplicación de algunas redes está basada en ella.
- La red PROFIBUS-FMS, perteneciente a la familia de redes PROFIBUS, que ha sido diseñada específicamente para cubrir este tipo de redes de control. Precisamente, gran parte de sus conceptos funcionales se derivan de las especificaciones desarrolladas en torno a la red MAP (entre las que cabe destacar el protocolo MMS) [BEND 93].
- La red ControlNet, desarrollada inicialmente por ROCKWELL AUTOMATION en 1995. Esta red de control está especialmente diseñada para el intercambio de información entre los sistemas de control de procesos.

9.3.2.3.3 Redes de sensores-actuadores

Conceptos generales

Bajo la denominación de redes de sensores-actuadores se agrupan las redes o buses de campo (*Field buses*) diseñados con el objetivo específico de intercomunicar los sistemas electrónicos de control con los dispositivos de campo (*Field devices*) conectados al proceso. Este tipo de redes funcionan en aplicaciones de tiempo real estricto en una pequeña zona de la planta (típicamente una máquina o una célula) y se suelen utilizar, por ejemplo, para comunicar los autómatas programables con los dispositivos sensores y/o actuadores del sistema. Por ello, los fabricantes de estas redes las suelen denominar de periferia distribuida (*Distributed periphery*).

Han sido numerosos los fabricantes de sistemas de control que han desarrollado redes de sensores y actuadores, que se diferencian en algunos aspectos funcionales tales como:

- La posibilidad de disponer de uno o más nodos principales (*master*) en la red.
- La comunicación de datos de sensores y actuadores todo/nada (*On/off*) o analógicos.
- La capacidad de diagnóstico y/o parametrización de los sensores y actuadores.

Es frecuente, además, que los fabricantes de este tipo de redes traten de normalizar el intercambio de información con los dispositivos de uso más frecuente en las aplicaciones industriales mediante los denominados perfiles de comunicación (*Communication Profiles*). Un perfil de comunicación es una información digital que define las características funcionales de cada tipo de módulo conectable a la red. El objetivo del perfil es facilitar el diseño asistido por computador de las instalaciones automatizadas mediante redes de sensores-actuadores. En el apartado A2.4.2.1 del apéndice 2 se describen los perfiles de la red AS-i.

De acuerdo con ello, y aunque resulta difícil determinar con exactitud los límites, los autores proponen clasificar las redes de sensores y actuadores, de acuerdo con su capacidad funcional, en dos subgrupos que se estudian a continuación.

Redes de sensores-actuadores de capacidad funcional limitada

Se entiende por redes de sensores-actuadores de capacidad funcional limitada las que están diseñadas para integrar principalmente dispositivos todo/nada como por ejemplo finales de carrera, fotocélulas, relés, contactores, electroválvulas, etc. Se caracterizan también por tener, en general, un único nodo principal. Como ejemplo de este tipo de redes se puede citar la red AS-i (acrónimo de *Actuator Sensor Interface*), que es un bus de campo que esencialmente permite la conexión de sensores y actuadores del tipo todo/nada y en su última versión también de sensores analógicos de medida de determinadas magnitudes como por ejemplo temperatura, distancia, etc. A su estudio se dedica el apéndice 2.

Redes de sensores-actuadores de capacidad funcional elevada

Se entiende por redes de sensores-actuadores de capacidad funcional elevada las que disponen de una capa de enlace adecuada para el envío eficiente de bloques de datos de mayor tamaño que en el caso anterior. Los mensajes que circulan por estas redes permiten que, mediante ellas, se puedan configurar, calibrar e incluso programar sensores de medida (como por ejemplo codificadores absolutos, sensores de temperatura, presión, caudal, etc.) y/o actuadores inteligentes (como por ejemplo variadores de velocidad de motores de corriente alterna, servoválvulas, etc.), además de los todo/nada. Son redes capaces por lo tanto de comunicar, de forma eficiente y con bajo coste, los sistemas electrónicos de control, como por ejemplo los autómatas programables, con dispositivos de campo (*Field Devices*) más complejos e inteligentes que los sensores todo/nada. Son ejemplos de redes de sensores-actuadores de alta capacidad funcional:

- La red InterBus-S, desarrollada por la empresa alemana PHOENIX-CONTACT.
- La red DeviceNet, desarrollada inicialmente por ROCKWELL AUTOMATION, que utiliza la capa física y de enlace de la red CAN y les añade una capa de aplicación orientada a objetos.
- La red PROFIBUS-DP, que es una red perteneciente a la familia PROFIBUS (descrita en el apéndice 3), que está específicamente orientada a la comunicación de los sistemas electrónicos de control con todos los distintos tipos de dispositivos de campo, tanto sensores como actuadores.

- La red CANopen, desarrollada en 1993 a partir del bus de campo CAN, diseñado por Bosch en 1986, añadiéndole la capa de aplicación del mismo nombre. Esta red, está gestionada por la asociación de usuarios y fabricantes denominada CiA (*CAN in Automation*).

9.3.2.4 Familias de redes industriales

Tal como se indica en el apartado 9.3.2.1, una familia de redes industriales es un conjunto de redes de datos y de control que comparten la capa física y de enlace de datos del modelo OSI. Su objetivo es utilizar una base común y, a partir de ella, implementar un conjunto de protocolos de aplicación que satisfagan los requisitos funcionales de los diferentes niveles de la pirámide CIM. Un ejemplo de familia de redes industriales es la familia de redes PROFIBUS, que se describe en el apéndice 3.

9.3.2.5 Redes Ethernet Industrial

9.3.2.5.1 Introducción

Los trabajos de investigación aplicada y desarrollo llevados a cabo por diferentes fabricantes de equipos de automatización a partir de finales de la década de 1990 han promovido la adaptación y redefinición de las tecnologías que son normas “de facto” en el ámbito de las redes de datos, para poder utilizarlas en ámbitos más cercanos a los de las redes de control [SCHE 02] [PIGA 06]. Como resultado de ello se definió en los últimos años el conjunto de redes denominadas “*Industrial Ethernet*” cuya capa de enlace está basada en la técnica de acceso al medio normalizada denominada *Ethernet* y los protocolos básicos de comunicación de las capas de red y transporte TCP/IP del modelo OSI, utilizados por las redes ofimáticas y descritos en el apéndice 1.

A grandes rasgos, las redes *Industrial Ethernet* tratan fundamentalmente de rediseñar el *hardware* y el *software* asociado a las redes *Ethernet* con el objetivo principal de adecuarlas para que tengan un tiempo de respuesta que permita utilizarlas como redes de control y redes de sensores-actuadores. Para ello se han llevado a cabo actividades de investigación aplicada y desarrollo tecnológico, tanto en el sistema físico (*Hardware*) como los recursos de programación (*Software*), para conseguir nuevas redes industriales que proporcionen comunicaciones redundantes y que presenten una mayor fiabilidad y vida útil que las utilizadas tradicionalmente en el entorno de las redes de datos ofimáticas. Se logra así que estos sistemas electrónicos de comunicación puedan funcionar en las condiciones ambientales que se presentan habitualmente en una fábrica (alto nivel de interferencias electromagnéticas, temperatura elevada, vibraciones, humedad, elevada velocidad de comunicación y mínimo tiempo de respuesta en determinadas aplicaciones, etc.).

En lo que respecta al protocolo de la capa de aplicación del modelo OSI, que se debe utilizar en combinación con los protocolos de las capas inferiores del mismo, no existe actualmente una única red *Industrial Ethernet* normalizada y están planteadas diferentes soluciones tecnológicas como por ejemplo “Modbus/TCP”, “EtherNet/IP” y “PROFINet”. Se simplifica al máximo, mediante ellas, la implantación de sistemas electrónicos de intercomunicación entre las redes de datos de los niveles superiores y las redes de control de los niveles inferiores de la pirámide CIM, lo que hace que deje de tener vigencia la clasificación de las redes de Comuni-

caciones Industriales de la tabla 9.3 y pase a tenerla la de la tabla 9.4, que indica que además de las redes de aplicación específica, en el siglo XXI existen redes universales y redes integradas, estudiadas a continuación, que satisfacen las necesidades de los niveles altos y bajos de la pirámide CIM.

9.3.2.5.2 Tipos de redes Ethernet Industrial

Para implementar una red Ethernet Industrial existen dos opciones:

- **Modificar Ethernet para utilizarla en todos los niveles de la pirámide CIM**

Se obtiene así una red industrial universal (*Universal industrial network*) que permite la implantación de las comunicaciones en los diversos niveles de la pirámide CIM y se representa gráficamente de forma simplificada en la figura 9.40, en la que se puede observar que a través de una red única se conectan los diferentes sistemas que forman parte de la pirámide CIM:

- Los controladores (PLC) y los sistemas SCADA utilizados en aplicaciones en las que el tiempo de respuesta T_R (*Response time*) está comprendido, en general, entre 10 y 100ms.
- Los controladores (PLC), las interfaces máquina-usuario (HMI) y las estaciones de entrada-salida remotas (*Distributed periphery*) en las que el tiempo de respuesta T_R está comprendido, en general, entre 1 y 10ms.
- Las unidades de control de movimiento (*Motion control*) y las estaciones de entrada-salida remotas en las que el tiempo de respuesta T_R es, en general, inferior a 1 ms.

Es un ejemplo de una red industrial universal la red Profinet que se describe en el apéndice 4.

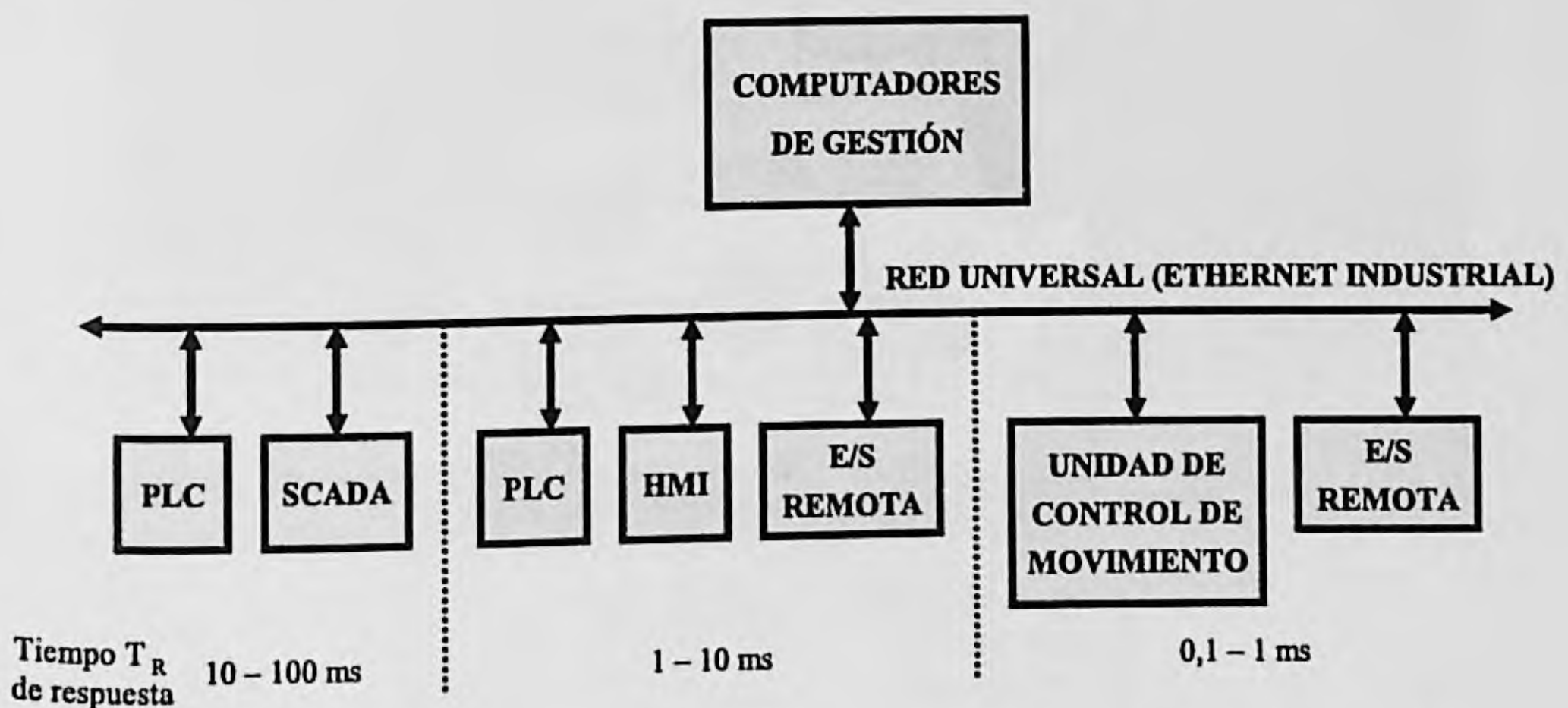


Figura 9.40. Diagrama de bloques del principio de funcionamiento de una red universal Ethernet Industrial.

- **Combinar la red Ethernet con una red de controladores y otra de sensores/actuadores**

Integrar la red Ethernet con una red de controladores y otra de sensores/actuadores que utilizan el mismo protocolo de la capa de aplicación que ella. La red Ethernet, que en este caso se suele denominar también "Ethernet Industrial", se utiliza hasta que el nivel de sus prestaciones de velocidad y determinismo lo permiten, teniendo en cuenta las características de la planta. Las redes de controladores y de sensores/actuadores implementan el resto de las comunicaciones de la planta. En la figura 9.41 se representa el diagrama de bloques del principio de funcionamiento de una red Ethernet Industrial de este tipo de red industrial integrada (*Integrated industrial network*). Las unidades de control de movimiento y las estaciones de entrada-salida remotas se conectan, en general, a un bus de sensores-actuadores que tenga el tiempo de respuesta necesario para este tipo de sistemas. Las unidades de interfaz máquina-usuario y los controladores se conectan a un bus de controladores y además estos últimos hacen de puente (*Bridge*) entre las dos redes citadas que se diferencian en el protocolo de enlace y poseen un protocolo idéntico en la capa de aplicación. Por último a la red Ethernet Industrial se conectan controladores, sistemas SCADA y las estaciones de entrada-salida remotas que exigen un tiempo de respuesta mayor que las conectadas a los buses de sensores-actuadores y de controladores.

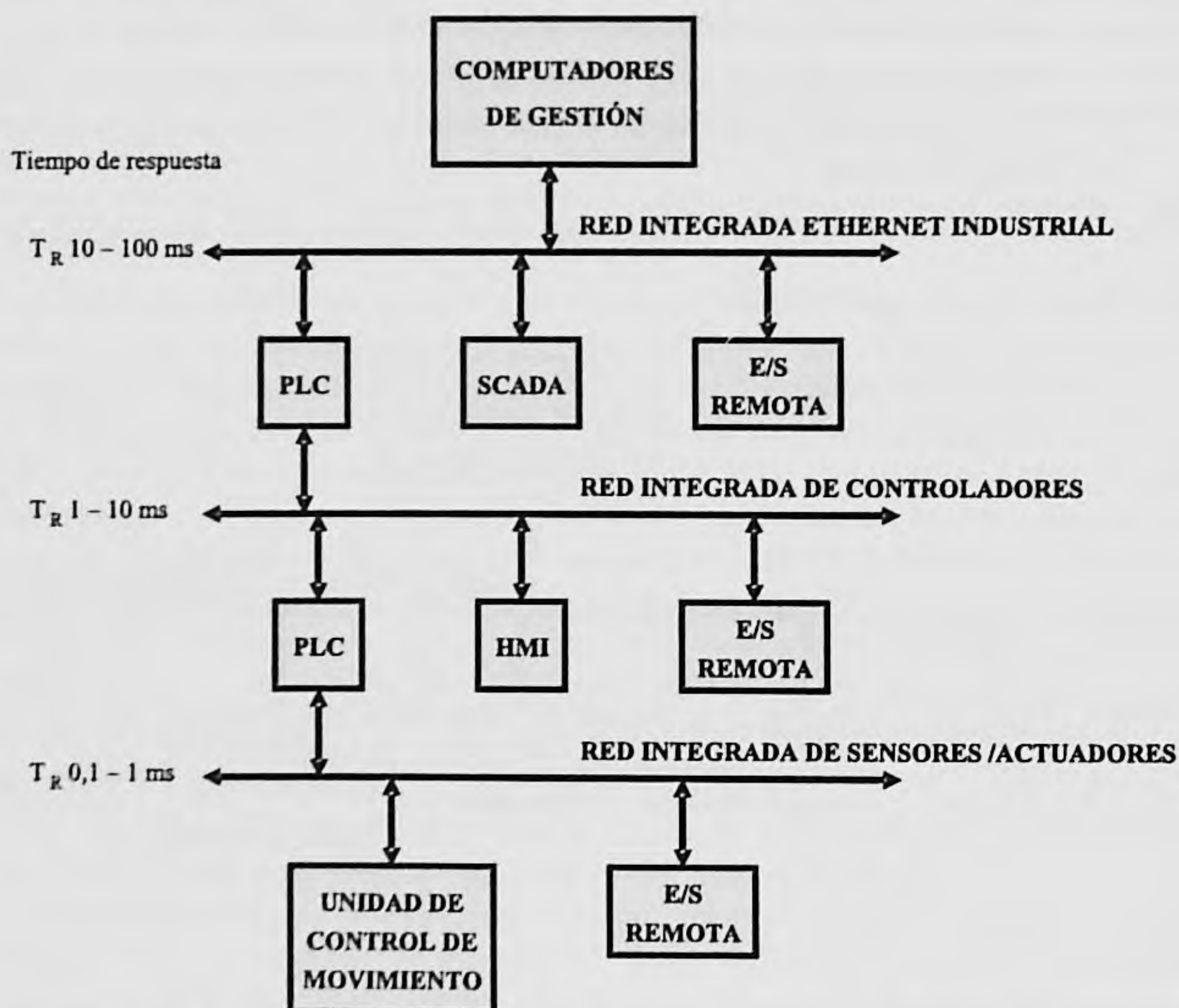


Figura 9.41. Diagrama de bloques del principio de funcionamiento de una red integrada Ethernet Industrial.

Es un ejemplo de una red integrada Ethernet Industrial la combinación de la red Ethernet/IP (denominada simplemente Ethernet Industrial), la red de controladores ControlNet y la de sensores-actuadores DeviceNet, que utilizan todas ellas en la capa de aplicación el protocolo CIP (acrónimo de *Common Interface Protocol*).

9.4 El autómata programable y las Comunicaciones Industriales

9.4.1 Introducción

Por ser el autómata programable, tal como se indica en el capítulo 1, uno de los sistemas electrónicos más utilizados en la automatización de los procesos productivos como parte integrante de la pirámide CIM (Figura 9.33) de fabricación integrada por computador, también conocida como TIA (acrónimo del inglés *Totally Integrated Automation*), en especial en los niveles de control de máquinas, células y áreas de fabricación (Figura 9.42), se fueron incorporando a él, de forma paulatina, las Comunicaciones Industriales. Para ello se les fue dotando de capacidad de comunicación tanto con los dispositivos de campo como con los demás sistemas electrónicos de control situados en su entorno (otros autómatas programables, sistemas de control numérico, computadores industriales, etc).

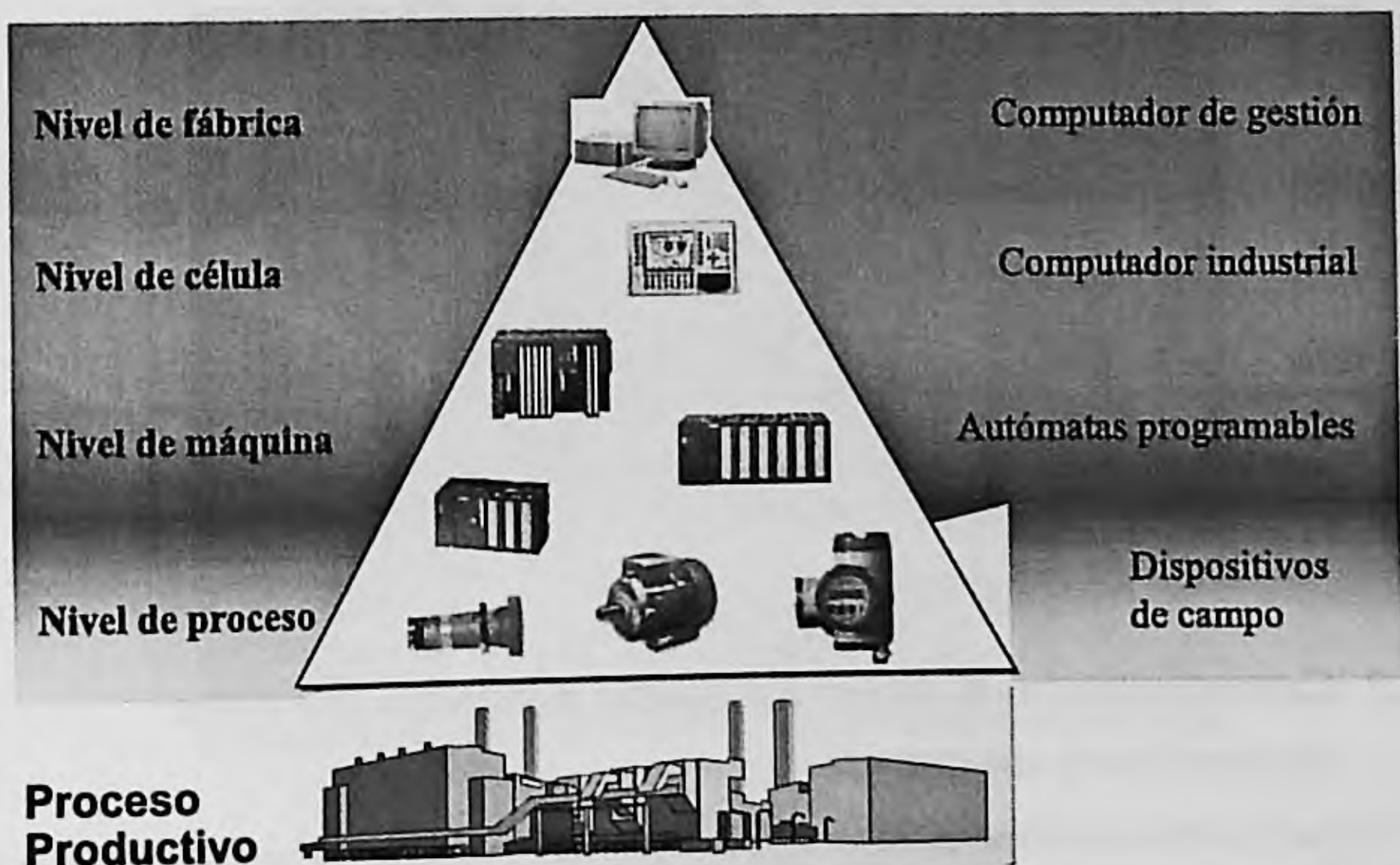


Figura 9.42. Utilización del autómata programable en la pirámide CIM (cortesía de Siemens).

En la figura 9.43 se representa gráficamente la evolución en Europa de los autómatas programables, en la que se pueden distinguir las siguientes etapas a partir de 1975:

- 1975: Comercialización de los autómatas programables basados en un microprocesador (por ejemplo, la familia SIMATIC S3 de Siemens).
- 1980: Comienzo de la conexión de los autómatas programables entre sí y con los sensores y actuadores (dispositivos de campo), mediante redes de control.
- 1994: Descentralización generalizada de los procesos de control: Sistemas de control distribuido conectados mediante redes de control, algunas de ellas basadas en Ethernet (por ejemplo la red SINEC H1 de Siemens).
- 2000: Distribución generalizada de los dispositivos de campo inteligentes mediante redes de control.

En los siguientes apartados se describen las relaciones que se han ido produciendo entre los autómatas programables y las redes de control, y se indican algunas de las soluciones normalizadas que se utilizan actualmente en el entorno europeo.

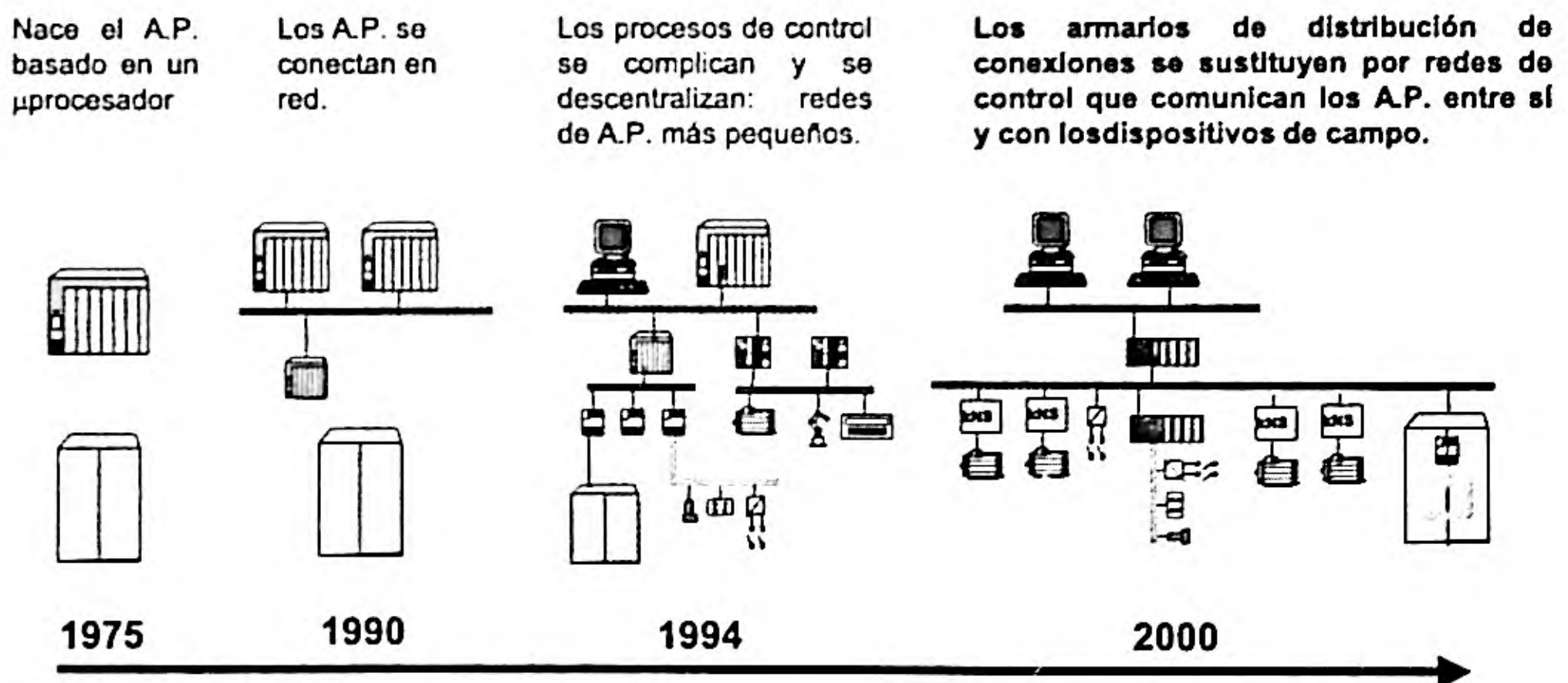


Figura 9.43. Evolución en Europa de los autómatas programables basados en un microprocesador (cortesía de Siemens).

9.4.2 Comunicación entre el autómata programable y los dispositivos de campo

Al igual que otros sistemas electrónicos de control que se utilizan en los niveles de estación, célula y fábrica de la planta industrial, los autómatas programables modernos se han adaptado, a lo largo de la última década del siglo XX (Figura 9.42), para disponer de capacidad de conexión, mediante una red de control, con los dispositivos de campo (*Field devices*) (tanto sensores como actuadores) presentes en el proceso productivo.

Las redes que conectan los autómatas programables con los dispositivos de campo (Figura 9.44) tienen, en general, las siguientes características:

- Garantizan un tiempo máximo de lectura de los sensores y de actualización de la señal de control de los actuadores (modo de operación determinista).
- Facilitan la instalación y conexión de todos los elementos de la red.
- Proporcionan capacidad de ampliación (*Scalability*) del sistema a lo largo de toda su vida útil (o ciclo de vida).
- Proporcionan funciones complementarias de diagnóstico y seguridad.

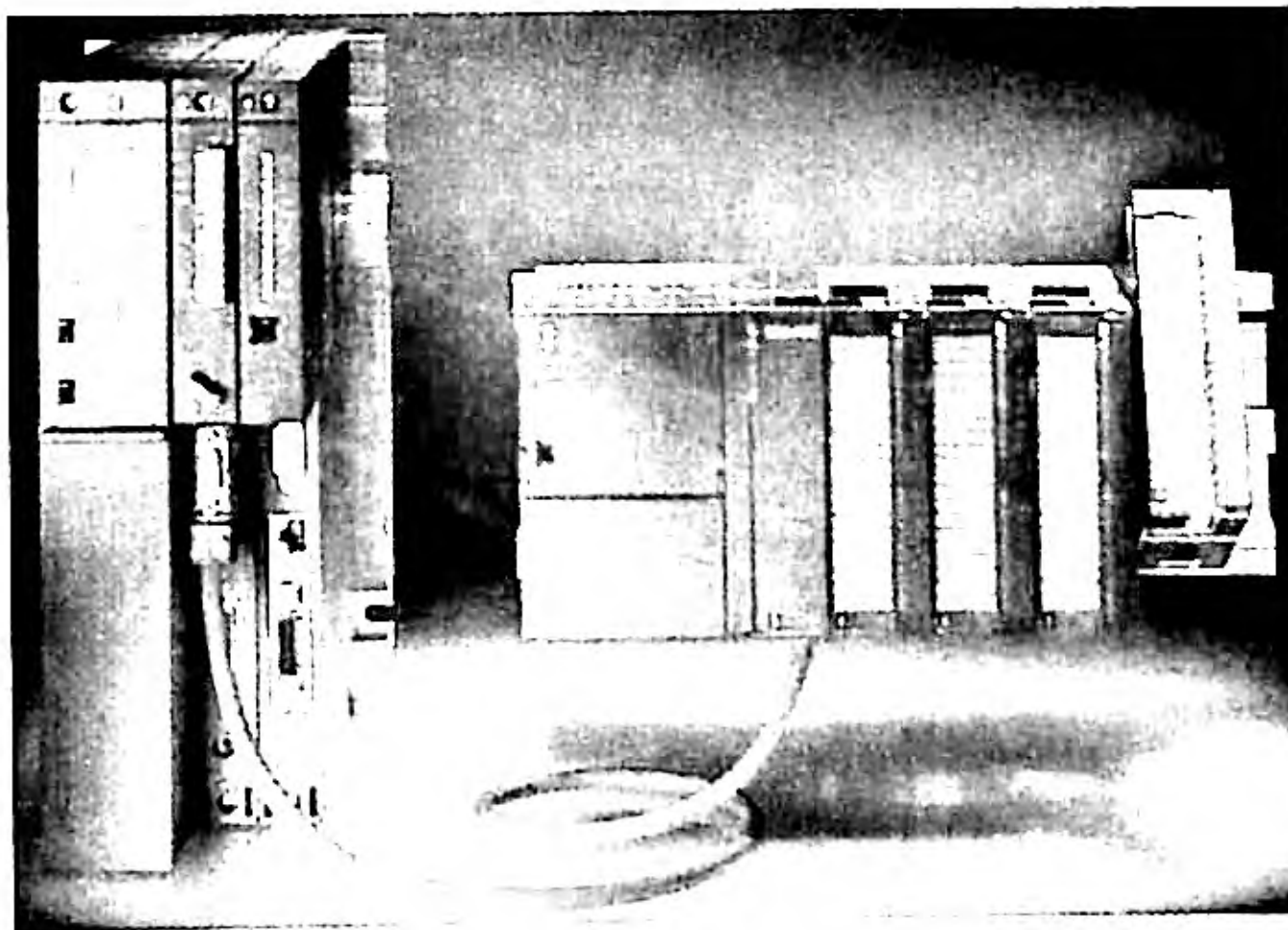
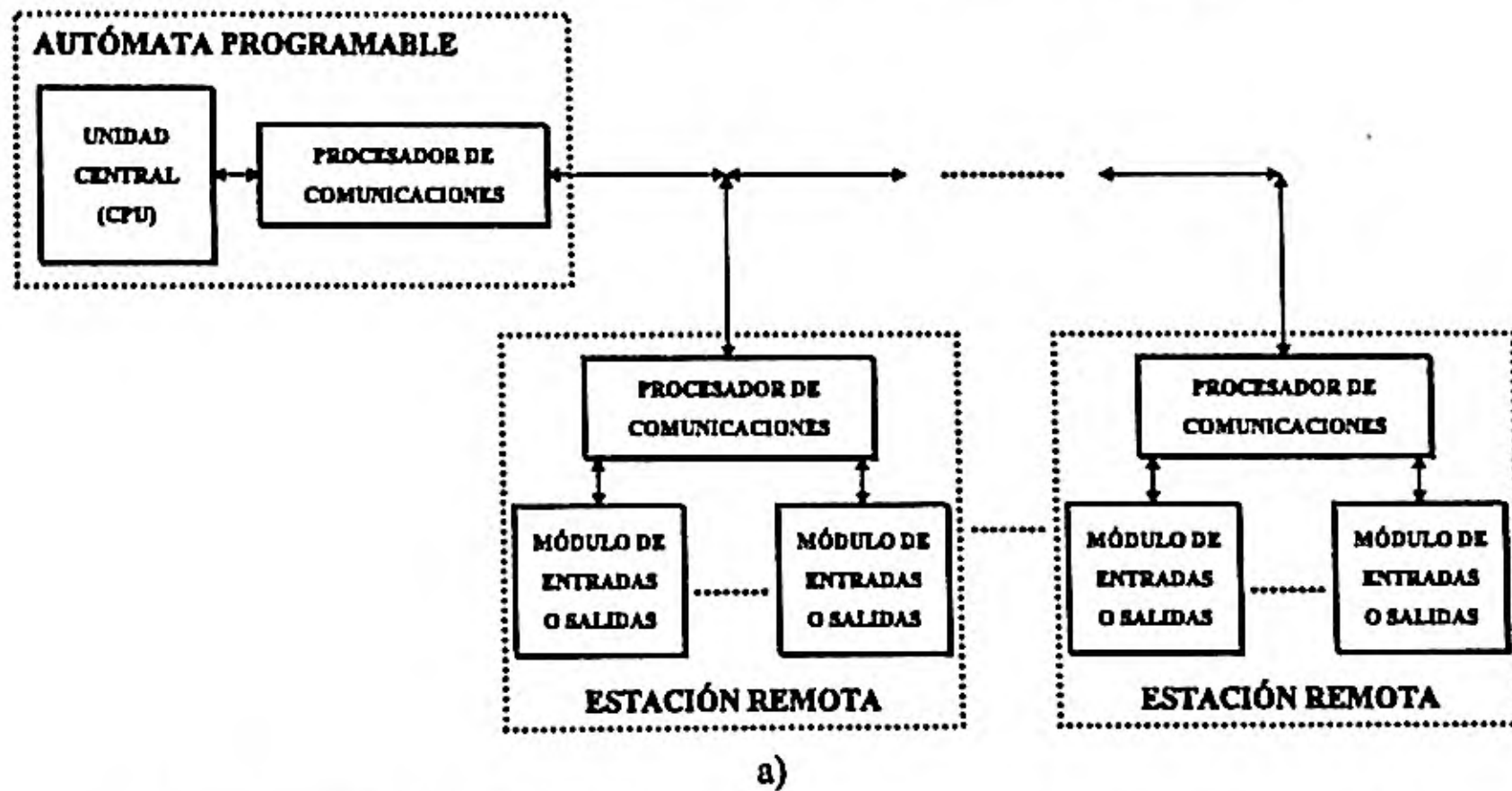


Figura 9.44. Red de sensores-actuadores: a) Esquema de bloques; b) Red PROFIBUS-DP de Siemens que enlaza un autómatas S7-400 con una estación remota de entrada/salida.

Inicialmente, numerosos fabricantes de sistemas electrónicos de control desarrollaron redes de sensores y actuadores propietarias. Posteriormente, y debido principalmente al interés por facilitar la interoperabilidad entre los elementos de automatización comercializados por distintos fabricantes, se normalizaron algunas de las redes propietarias (para convertirlas en no

propietarias) y se desarrollaron otras redes normalizadas basadas en general en el modelo OSI simplificado (Figura 9.39).

Tal como se indica en el apartado 9.3.2.3.3, las redes de sensores-actuadores pueden ser de capacidad funcional limitada o de alta capacidad funcional.

En el apéndice 2 se describen las características más significativas de la red de sensores-actuadores de capacidad funcional limitada denominada "AS-i" (acrónimo de *Actuator Sensor Interface*). Esta red está especialmente diseñada para satisfacer las necesidades de comunicaciones entre un autómata programable y los sensores todo/nada (*On-off*) y la utilizan numerosos fabricantes de autómatas programables.

En el apéndice 3 se describen las redes de sensores-actuadores de elevada funcionalidad PROFIBUS-DP y PROFIBUS-PA.

En la actualidad se utilizan las redes Ethernet Industrial en sus versiones universal o integrada, como por ejemplo la red universal Profinet y su servicio Profinet IO, para sustituir a las redes de sensores-actuadores. En la figura 1.87 se muestra un autómata programable Simatic S7 315-2PN/DP de Siemens que posee capacidad de comunicaciones a través de un bus de campo PROFIBUS DP y dos redes universales Profinet.

9.4.3 Comunicación entre autómatas programables y otros sistemas electrónicos de control

La primera técnica utilizada para el establecimiento de la comunicación entre los autómatas programables y otros sistemas electrónicos de control de procesos (como por ejemplo otros autómatas programables de igual o distinto fabricante, robots industriales, sistemas de control numérico, etc.) consistió en conectar entre sí sus módulos analógicos y digitales de entrada/salida, tal como se indica en la figura 9.45. Pese a sus limitadas prestaciones funcionales y su coste relativamente elevado, este método aún se utiliza hoy en día en muchas instalaciones industriales porque es la única forma posible de comunicar dispositivos de control de características muy dispares. No obstante, la situación va cambiando a medida que aumenta la normalización y disponibilidad de sistemas de comunicación abiertos, más avanzados y versátiles.

Otro método de comunicación, utilizado en los sistemas electrónicos de control industrial casi desde sus orígenes, es la comunicación punto a punto basada en la norma "RS-232" (denominada usualmente puerto serie) normalizada por la EIA (acrónimo de *Electronics Industry Association*) a principio de la década de 1980, que utiliza, a nivel de enlace, el protocolo de transmisión serie asíncrona (descrita en el apartado A1.2.3.1 del apéndice 1).

Las interfaces de comunicación "RS-232" o puerto serie (*Serial port*) se hicieron muy populares en el ámbito de los computadores domésticos con la normalización y difusión de dispositivos, como por ejemplo el vulgarmente conocido como "ratón", y hasta hace un par de años se incluían en la mayor parte de dichos sistemas informáticos. Rápidamente, los fabricantes de equipos electrónicos de control industrial apreciaron sus virtudes y comenzaron a incorporar esta y otras variantes de transmisión de información serie asíncrona en sus aparatos para poder intercomunicarlos, por ejemplo, con computadores encargados de su configuración, programación, depuración y puesta en marcha, equipamiento auxiliar (lectores de códigos de barras, impresoras, paneles de operador, pantallas táctiles, convertidores de frecuencia, etc.), otros sis-

temas de control, etc. Hoy en día se puede afirmar que todavía una gran parte de los sistemas electrónicos de control industrial disponen de alguna clase de comunicación serie asíncrona como la "RS-232" o sus derivadas, aunque están siendo sustituidas paulatinamente por nuevas interfaces más versátiles como "USB", "FireWire", etc.

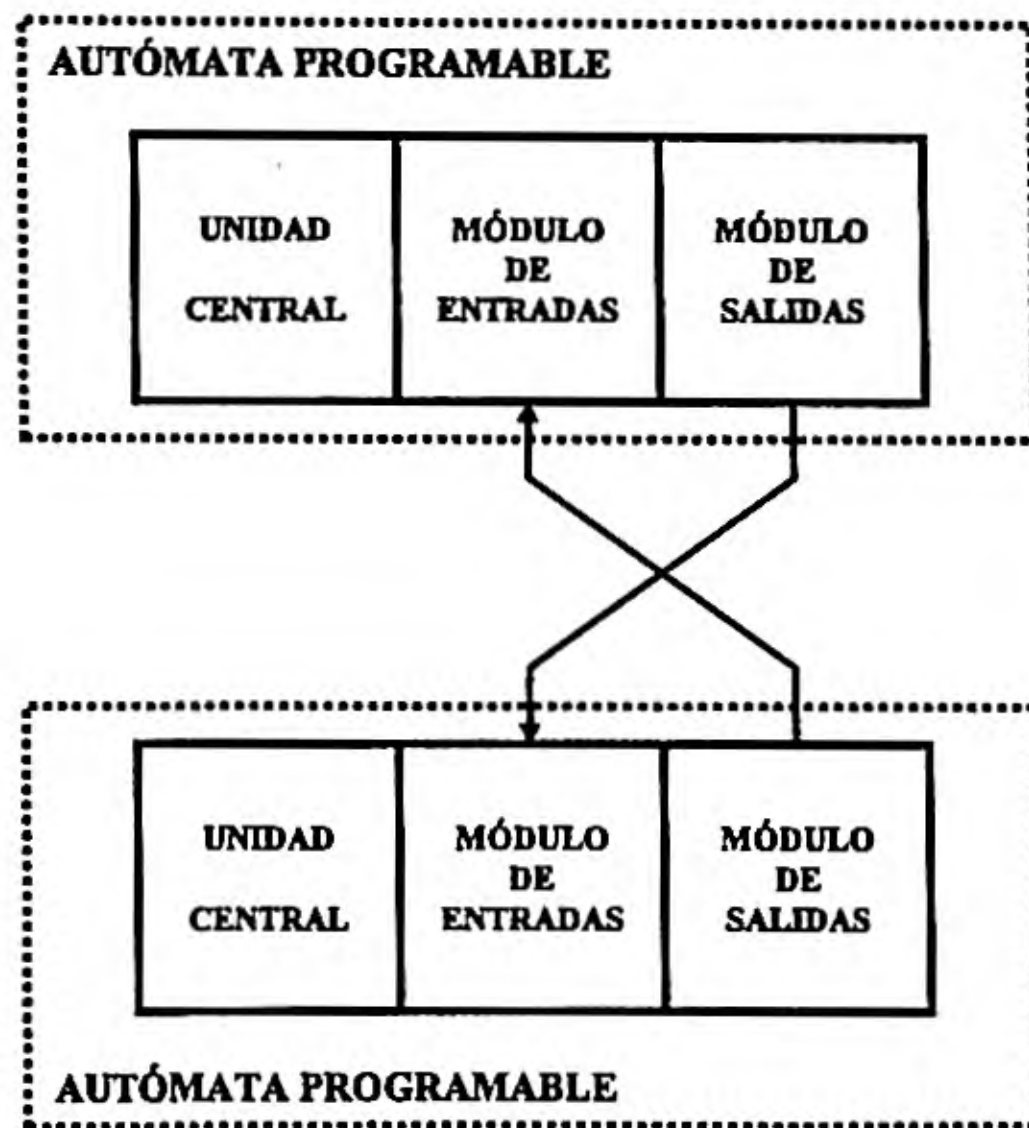


Figura 9.45. Comunicación entre dos autómatas programables a través de sus módulos de entrada/salida.

Uno de los problemas que surgieron a raíz de la difusión industrial de la norma "RS-232" fue la ausencia, en su contexto, de especificaciones relativas a los protocolos de aplicación a implantar sobre la misma, debido a que solamente establece las capas física y de enlace de datos del modelo OSI. Esta situación hizo que se generase una auténtica "Torre de Babel" de protocolos, muchos de ellos propietarios, que comprometió sin duda la interoperabilidad de los equipos industriales (y contribuyó a que se produjesen las islas de automatización). Algunos fabricantes, conscientes de este problema, trataron de normalizar en el mercado protocolos de aplicación y entre ellos destacó el protocolo conocido como "ModBus", definido inicialmente por la empresa MODICON (actualmente SCHNEIDER ELECTRIC), y que aún hoy en día goza de gran aceptación en ciertos mercados internacionales.

El posterior desarrollo de las tecnologías de transmisión serie asíncrona de información derivó en la creación de nuevas normas de comunicación con mayores prestaciones funcionales, como las denominadas "RS-422" y "RS-485" (descritas en el apartado A1.2.7 del apéndice 1), que son la base de muchos de los sistemas de comunicación industrial que existen actualmente en el mercado.

Sobre la base de estas tecnologías, los fabricantes comercializan en la actualidad protocolos de aplicación propietarios y no divulgados (como los denominados "MPI" y "PPI" de Siemens), divulgados pero no normalizados (como el "USS" de Siemens, "MODBUS" de SCHNEIDER ELECTRIC, etc.) y normalizados (como "PROFIBUS", "ControlNet", "InterBus-S",

“WorldFIP”, etc). En el apéndice 3 se describen, como ejemplo, las características más significativas de la familia de redes industriales normalizadas PROFIBUS (abreviatura de *PROcess Field BUS*).

Hoy en día, la tendencia en la evolución de las Comunicaciones Industriales continua siguiendo, a una cierta distancia, la implantación de técnicas de transmisión de información sin hilos (infrarrojas, ondas electromagnéticas, etc.), que constituyen normas *de facto* en el mundo de las comunicaciones como “*BlueTooth*”, “*Wi-Fi*”, *Zigbee*, etc.

Otro ámbito en el que las Comunicaciones Industriales han supuesto una gran revolución tecnológica es el relacionado con la realización, especialmente en la última década del siglo XX, de sistemas de control distribuido formados por un conjunto de sistemas electrónicos de control que intercambian la información necesaria para coordinar y secuenciar las operaciones de un determinado proceso productivo. En este entorno, se ha desarrollado un conjunto de redes industriales de control capaces de cubrir satisfactoriamente los requisitos específicos en relación con el determinismo, el tiempo de respuesta, la fiabilidad, la tolerancia a fallos, etc. Algunas de ellas, las más ambiciosas, pretenden también, en ciertos casos, integrarse con el área propia de las redes de datos u ofimáticas de la empresa. Un ejemplo de ello es la red Profinet que se describe en el apéndice 4.

9.4.4 Diseño e implantación de sistemas de control distribuido basados en autómatas programables

9.4.4.1 Introducción

Tal como se indica en los apartados anteriores, la implantación de un sistema electrónico de control distribuido implica la utilización de un conjunto de redes de comunicaciones o una red universal que resuelvan las transferencias de información entre los sistemas utilizados en los diferentes niveles de la pirámide CIM.

Para el diseño de dichas redes, el fabricante suministra elementos físicos (*Hardware*) y herramientas de programación (*Software*):

- **Elementos físicos**

Resuelven el nivel físico (1) y la mayor parte del de enlace de datos (nivel 2) del modelo OSI mediante un procesador de comunicaciones (*Communication processor*) que implementa el protocolo de enlace a través de un programa situado en una memoria no volátil (*Firmware*).

- **Herramientas de programación**

Son programas de diseño integrado que se ejecutan en un computador personal y permiten configurar de forma sencilla cualquier sistema de control distribuido que combina redes de datos y redes de control o que utiliza una red universal.

9.4.4.2 Metodología de diseño de sistemas de control distribuido

Las principales tareas a realizar para implementar una instalación de control distribuido son las siguientes:

- Elección de los sistemas electrónicos de control.
- Elección de las diferentes redes que comunican entre sí los sistemas electrónicos de control.
- Elección de la red que comunica los sistemas electrónicos de control y los sensores asociados con cada uno de ellos.
- Configuración del sistema.

A continuación se describe brevemente cada una de ellas.

Elección de los sistemas electrónicos de control

Se ha de tener en cuenta el número de variables de entrada y salida, la velocidad de cálculo necesaria y las especificaciones de comunicación con los sensores y con otros sistemas de control y supervisión.

En esta elección se deben considerar otros tipos de parámetros como por ejemplo, económicos, de estrategia de mantenimiento, etc., y por ello es difícil dar reglas generales.

Elección de las redes de comunicaciones

La elección de las redes de comunicaciones está supeditada al resultado de la tarea anterior, porque cada uno de los fabricantes con mayor implantación en el mercado utiliza principalmente un determinado conjunto de redes de comunicaciones normalizadas.

Además, algunos fabricantes de equipos de Comunicaciones Industriales comercializan procesadores de comunicaciones que permiten conectar autómatas programables de diferentes fabricantes a redes normalizadas distintas de las del propio fabricante de autómatas programables.

Configuración del sistema

El proceso de configuración (*Configuration phase*) es llevado a cabo mediante un programa de diseño integrado tal como se indica en el apartado anterior. Dicho programa permite, de forma gráfica e interactiva, establecer las topologías y los parámetros detallados de cada elemento de la red, y una vez finalizada la descripción del sistema de control distribuido se transfiere a través de un puerto de comunicaciones al autómata programable y al procesador de comunicaciones, directamente o a través del primero. En la figura 9.46 se representa el esquema de bloques de un sistema de control formado por un computador personal, un autómata programable y un procesador de comunicaciones (CP).

Ejemplos de sistemas de configuración de una red de control distribuido son las herramientas SIMATIC NET y NetPro de Siemens. Esta última se representa en la figura 9.47.

Como resumen de todo lo expuesto es conveniente resaltar que, mediante los elementos físicos (*Hardware*) y las herramientas de programación (*Software*) y a pesar de que las redes industriales son muy complejas, su configuración resulta extremadamente sencilla, es decir dicha complejidad es transparente al usuario.

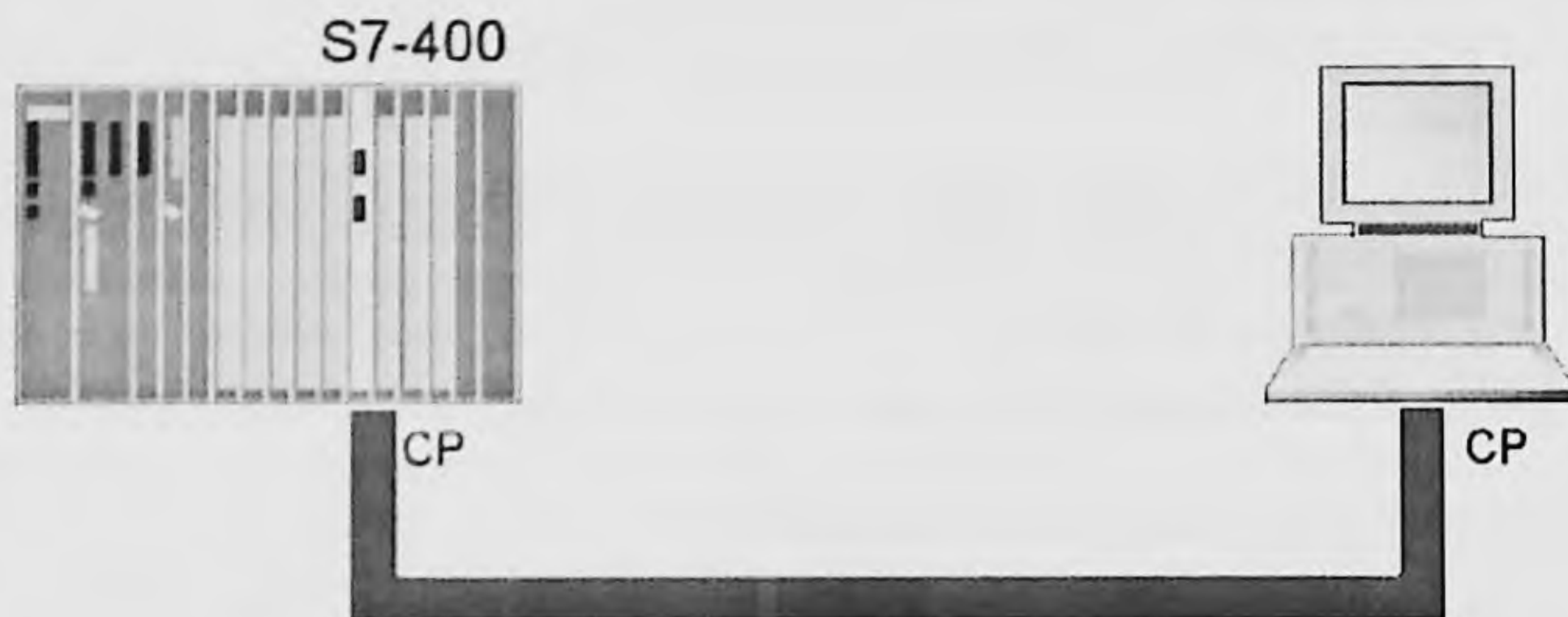


Figura 9.46. Comunicación entre un ordenador personal, un autómata programable y un procesador de comunicaciones (CP) que constituye un módulo del autómata programable.

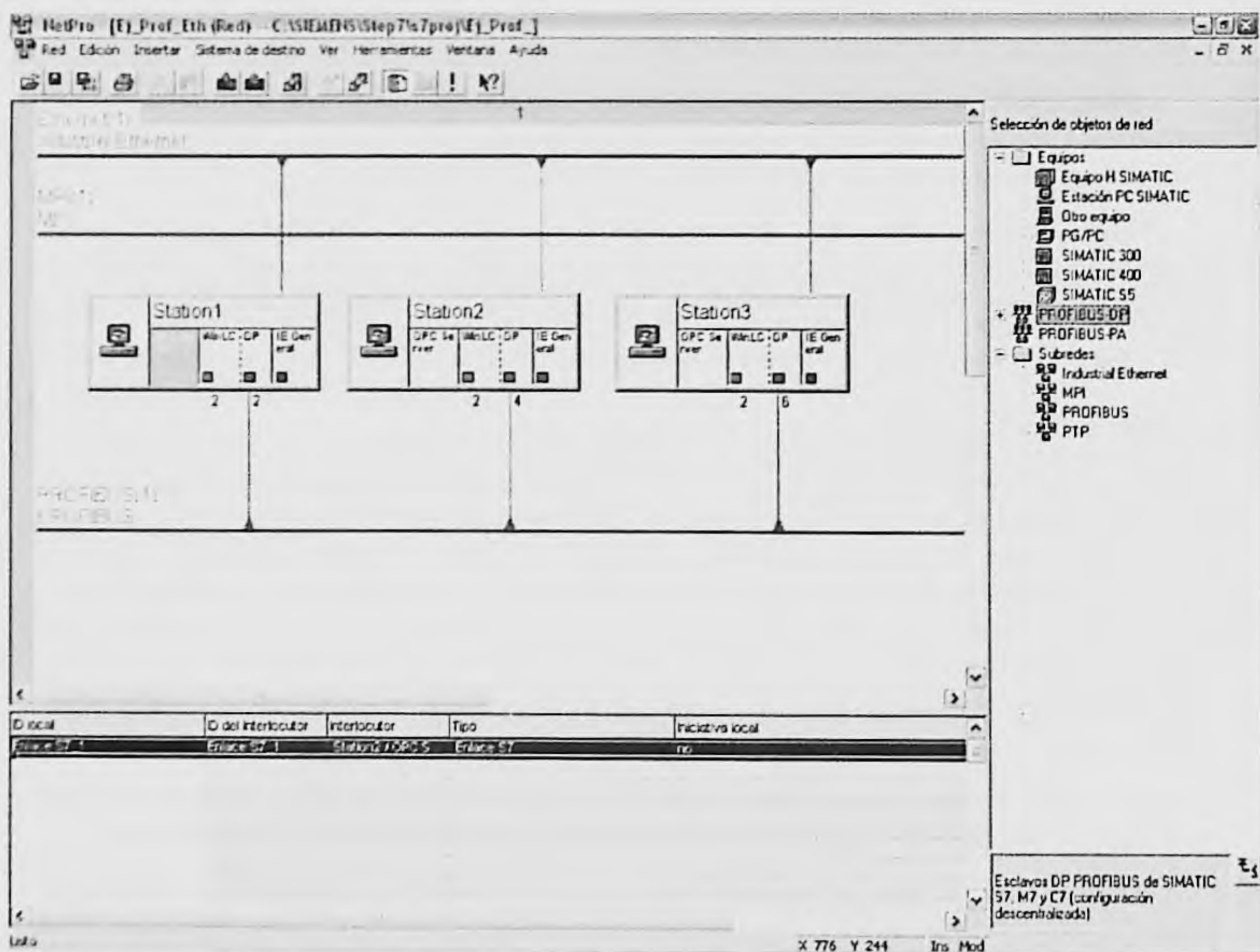


Figura 9.47. Herramienta NetPro de configuración de una red de control distribuido (Cortesía de Siemens).

Bibliografía

- [AUDÍ 88] D. Audí. *Como y cuando aplicar un robot industrial*. Colección Productiva, número 6. Editorial Marcombo. 1988.
- [AYRE 83] R. U. Ayres, Steven M. Miller. *Robotics: Applications and social implications*. Ballinger Publishing Company. 1983.
- [BAIE 03] D. Bailey, E. Wright. *Practical SCADA for Industry*. Newnes. 2003
- [BARR 07] A. Barrientos, L. Peñín, C. Balaguer y R. Aracil. *Fundamentos de Robótica*. Mc Graw Hill. 2007.
- [BEND 93] K. Bender, M. Katz, et al. *PROFIBUS. The Fieldbus for Industrial Automation*. Prentice-Hall. 1993.
- [BOYE 04] S. A. Boyer. *SCADA: Supervisory Control And Data Acquisition*. ISA (Instrumentation Systems and Automation Society). 2004.
- [CRUZ 05] F. Cruz. *Control numérico y programación: Sistemas de fabricación de máquinas automatizadas*. Marcombo. 2005
- [DESS 58] F. Dessauer. *Streit um die Technik*. Editorial Verlag Josef Knecht, 1958. Versión española: *Discusión sobre la técnica*. Ediciones Rialp S.A. 1964.
- [ELAH 01] A. Elahi. *Network communications technology*. Delmar Thomson Learning. 2001
- [FIOR 01] James M. Fiore. *Amplificadores operacionales y circuitos integrados lineales*. Thomson Learning Paraninfo. 2001.
- [FERR 86] G. Ferraté y otros. *Robótica Industrial*. Editorial Marcombo, 1986.
- [FERR 88] Rafael Ferré. *La fábrica flexible*. Colección Productiva, número 9. Editorial Marcombo. 1988.
- [FORO 07] B. A. Forouzan. *Data communications and networking*. Mc Graw Hill. 2007.
- [HAM 85] I. Ham, K. Hitomi & T. Yoshida. *Group Technology: Applications to Production Management*. Kluwer-Nijhoff. 1985.
- [GROO 80] M.P. Groover. *Automation, Production Systems, and Computer Aided Manufacturing*. Prentice Hall, 1980.
- [GROO 84] M. P. Groover, E. W. Zimmers. *CAD/CAM: Computer Aided Design and Manufacturing*. Prentice Hall. 1984.
- [IEC 00] International Electrotechnical Commission standard IEC 61158-2. *Fieldbus standard for use in industrial control systems - Part 2: Physical Layer specification and service definition*. 2000.
- [KIEF 98] H. B. Kief. *Manual CN/CNC*. Editorial Gran Duc S.L. 1998.
- [KRAR 00] S. F. Krar, A. Gill & P. Smid. *Computer numerical control simplified*. Industrial Press Inc. 2000.
- [KUO 95] B. C. Kuo. *Automatic Control Systems*. Prentice Hall. 1995.
- [LEIG 85] J. R. Leigh. *Applied digital control*. Prentice Hall International. 1995.
- [LEON 00] C. T. Leondes, editor. *Computer aided design, engineering and manufacturing: Systems technics and applications*. CRC Press. 2000.

- [MALE 91] R.A. Maleki. *Flexible manufacturing systems: The technology and management*. Prentice Hall. 1991.
- [MAND 07] E. Mandado, L. M. Menéndez, L. F. Ferreira y E. L. Matos. *Microcontroladores PIC. Sistema integrado para el autoaprendizaje*. Editorial Marcombo. 2007.
- [MART 93] J. Martínez y M. Barrón. *Prácticas con microcontroladores de 8 bits. Aplicaciones industriales*. Editorial Mc Graw-Hill. 1993.
- [MART 03] E. Martín Cuenca, J. M. Angulo, I. Angulo. *Microcontroladores PIC*. Editorial Thomson Learning Paraninfo. 2003
- [MATT 01] M. Mattson. *CNC programming: Principles and applications*. Thomson Delmar Learning. 2001.
- [MUMF 34] Lewis Mumford. *Technics and civilization*. 1ª edición. Editorial Harcourt, Brace & World Inc., 1934. Versión española: *Técnica y civilización*. Alianza Universitaria (2ª edición), 1988.
- [OLLE 01] A. Ollero. *Robótica: manipuladores y robots móviles*. Editorial Marcombo. 2001.
- [PARR 93] D. Parrish. *Flexible Manufacturing*. Editorial Butterworth- Heinemann Ltd., 1993.
- [PIGA 08] R. Pigan & M. Metter. *Automating with Profinet. Industrial communication based on industrial Ethernet*. Wiley-VCH. 2008.
- [PNO 91] Profibus Nutzerorganisation. *PROFIBUS standard: DIN 19245*. Translation of the German standards. 1991.
- [PNO 94] Profibus Nutzerorganisation. *Process Field Bus - Decentralized Periphery (DP)*. Translation of the German Standard DIN 19245, Part 3. 1994.
- [PNO 95] Profibus Nutzerorganisation. *PROFIBUS-PA: Protocol Specification. Technical Guideline*. 1995.
- [PNO 99] Profibus Nutzerorganisation. *PROFIBUS Technical Description*. 1999.
Disponible en:
http://www.hta-be.bfh.ch/~wwwfbus/download/pno/tbe_0999.pdf
- [RAPH 03] B. Raphael & I. F. C. Smith. *Fundamentals of computer aided engineering*. John Wiley. 2003.
- [REHG 05] J. A. Rehg, H. W. Kraebber. *Computer-Integrated Manufacturing*. Prentice Hall. 2005.
- [RENT 00] A. Rentarí y M. Rivas. *Robótica Industrial*. McGraw Hill 2000.
- [RODR 03] A. Rodríguez Penín. *Sistemas SCADA*. 2ª edición. Editorial Marcombo. 2007.
- [SCHN 02] G. Schneider. *Ethernet Industrial va en camino de ser una auténtica norma abierta*. Revista Automática e Instrumentación. Octubre 2002.
- [SEMP 02] Sempere Payá, V. et al. *PROFIBUS, un bus de campo industrial*. Cuadernos CEA-IFAC, 2002.
- [SIEM 91] H. Baumgartner, K. Knischewski, H. Wieding. *CIM: consideraciones básicas*. Siemens. Editorial Marcombo. 1991.

- [SIMP 96] C. D. Simpson. *Industrial Electronics*. Prentice Hall Inc., 1996.
- [SING 80] M. G. Singh, J. P. Elloy, R. Mezencev y N. Munro. *Applied Industrial Control*. Pergamon Press, 1980.
- [SOUT 95] D. W. South. *Encyclopedic dictionary of industrial automation and computer control*. Editorial Prentice Hall, 1995.
- [STAL 07] W. Stallings. *Data & computer communications*. Sixth edition. Prentice Hall. 2000
- [TORR 02] F. Torres. *Robots y sistemas sensoriales*. Prentice Hall 2002.
- [WEEB 93] J. Webb, K. Greshock. *Industrial control electronics*. Second edition". MacMillan Publishing Company, 1993.
- [WEIG 01] J. Weigmann, G. Kilian. *Decentralization with PROFIBus-DP: Architecture and Fundamentals. Configuration and use with SIMATIC S7*. John Wiley & Sons. 2001.
- [ZURA 04] R. Zurawsky. *The Industrial Information Technology Handbook*. CRC Press. 2004.
- [ZURA 05] R. Zurawsky. *The Industrial Communication Technology Handbook*. CRC Press. 2005.

5

GARANTÍA DE FUNCIONAMIENTO DE LOS SISTEMAS ELECTRÓNICOS DE CONTROL

En esta parte, que está constituida por el capítulo 10, se estudian los principales conceptos relacionados con la garantía de funcionamiento o confiabilidad (*Dependability*) de los sistemas electrónicos de control.

Para ayudar al lector a comprender el capítulo 10, en el apéndice 5 se describen los principales conceptos asociados a la confiabilidad y se hace especial hincapié en los sistemas seguros ante averías así como en las normas EN e IEC que los definen y clasifican.

CAPÍTULO 10

Confiabilidad de los Sistemas Electrónicos de Control

10.1 Introducción

El aumento de la capacidad de integración de los fabricantes de circuitos integrados, debido al desarrollo de la Microelectrónica, ha disminuido drásticamente el coste del hardware de los sistemas electrónicos y ha hecho que, tal como se indica en los apartados 1.1 y 1.3.3.3.2 del capítulo 1, los sistemas electrónicos en general y los autómatas programables en particular se utilicen para llevar a cabo el control de numerosos procesos industriales en los que es muy importante tener en cuenta la garantía de funcionamiento, también denominada confiabilidad (*Dependability*), que se define como la propiedad que permite al usuario de un sistema tener una confianza justificada en el servicio que proporciona. La confiabilidad es un concepto complejo que está relacionado con diversos atributos de los sistemas como son la fiabilidad (*Reliability*), la disponibilidad (*Availability*), la seguridad ante averías (*Safety*), la protección de la propiedad intelectual o seguridad ante sabotajes (*Security*), la verificabilidad (*Testability*) [BS 99], etc. Para alcanzar una adecuada confiabilidad de una máquina o proceso se debe:

- Realizar la máquina o el proceso con el objetivo de que su confiabilidad tenga el nivel adecuado. La estrategia que se debe seguir al realizar el diseño es, en general, específica del área correspondiente de la tecnología o tecnologías utilizadas.
- Utilizar un sistema electrónico de control con un nivel adecuado de confiabilidad. Al lector que no esté familiarizado con los conceptos relativos a la confiabilidad de los sistemas electrónicos se le recomienda la lectura del apéndice 5.

Pero, además, el aumento de la actuación automática de los procesos productivos ha hecho que cada vez sean más importantes los aspectos de la confiabilidad relacionados con la seguridad ante averías (*Safety*) de las instalaciones, de las personas que trabajan en ellas y del medio ambiente que las rodea. Por ello las sociedades modernas exigen que dichas instalaciones cumplan determinadas condiciones que garanticen su seguridad de operación cuando se produce una avería, lo cual ha impulsado el desarrollo de normas de seguridad ante averías, por parte de los organismos de normalización. Por otra parte, la legislación de muchos países tiende a endurecer las penas por negligencias que den lugar a accidentes de consecuencias graves para las personas o el medio ambiente, lo que hace que las empresas deban estar cada vez más preocupadas por los temas de seguridad ante averías.

En general las normas relativas a la seguridad ante averías en las máquinas establecen un nivel de seguridad que no es posible conseguir actuando solamente sobre los elementos que

las constituyen. Por ello se han elaborado normas como la EN 954-1, descrita en el apartado A5.5.2.5.2 del apéndice 5, en las que se establecen diferentes categorías de sistemas electrónicos de control seguros ante averías para facilitar la elección del más adecuado en cada aplicación concreta.

La gran variedad de instalaciones automatizadas da lugar a diferentes situaciones que presentan exigencias distintas desde el punto de vista de su confiabilidad:

- Máquinas o procesos en las que la automatización es fija (descrita en el apartado 9.2.4.2.1 del capítulo 9) y además se lleva a cabo mediante elementos mecánicos, neumáticos, hidráulicos, etc., en los que se puede generar una situación peligrosa debido a una avería de los citados elementos o a una maniobra indebida por parte del operario que utiliza la máquina. En este caso se debe utilizar un sistema independiente de seguridad ante averías, descrito en el apartado 10.2 a continuación.
- Máquinas o procesos que, por su complejidad, están controladas por un sistema electrónico. En ellas, además de la situación anterior, se pueden producir, en función de las características del proceso productivo, diversas situaciones relacionadas con la confiabilidad:
 - Una situación peligrosa debida a una avería de la máquina o del sistema electrónico de control. En este caso se pueden utilizar dos estrategias diferentes para garantizar la seguridad ante averías:
 - La combinación de un sistema electrónico de control de aplicación general para realizar las tareas de control y un sistema independiente de seguridad para llevar a cabo las correspondientes a la seguridad ante averías. A la confiabilidad de los autómatas programables de aplicación general se dedica el apartado 10.3.2.
 - Un sistema electrónico de control seguro ante averías (*Safety or Fail-safe electronic system*) que ejecuta tanto las funciones de control como las de seguridad ante averías. Dicho sistema debe disponer de recursos que le permitan detectar los fallos que se producen en él y, en esa situación, debe pasar a un estado que garantice que no se producen daños en el proceso del que forma parte, en el entorno del mismo, o a las personas que lo utilizan. Al estudio de los autómatas programables seguros ante averías, denominados autómatas programables de seguridad, se dedica el apartado 10.3.3.2.
 - Una situación peligrosa debida a una acción incorrecta, intencionada o por falta de conocimiento, que haga que el sistema electrónico de control realice una operación incorrecta. En este caso se necesita que el sistema electrónico tenga una elevada seguridad ante sabotajes (*Security*). Al análisis de la seguridad ante sabotajes (*Security*) de los autómatas programables se dedica el apartado 10.3.2.4.
 - Graves pérdidas económicas, debidas a la puesta fuera de servicio del proceso productivo, por causa de una avería del sistema electrónico que lo controla. En este caso el sistema electrónico debe tener una elevada disponibilidad (*Availability*), es decir, debe estar adecuadamente implementado para garantizar que la probabilidad de que preste servicio en cualquier instante sea muy alta. Al estudio de los autómatas programables de elevada disponibilidad (*High availability*) se dedica el apartado 10.3.3.3.

10.2 Sistemas electrónicos independientes de seguridad

10.2.1 Definición y descripción

Se define un sistema independiente de seguridad (SIS) como aquél sistema que es capaz de detectar una situación crítica en un proceso o en una máquina y de tomar las medidas adecuadas para evitar que se convierta en peligrosa, es decir, que se produzca un accidente. Este tipo de sistemas debe ser intrínsecamente seguro o seguro ante averías propias (*Safety or Fail-safe electronic system*), es decir, una avería del mismo que haga que no pueda realizar la tarea que tiene encomendada debe dar lugar a una parada de la máquina o bien hacer que la máquina o el proceso pasen a un estado seguro. Al lector que no conozca los fundamentos de los sistemas electrónicos seguros ante averías se le remite al apartado A5.5.2.3 del apéndice 5.

Numerosas máquinas deben ser puestas fuera de servicio cuando se avería alguno de los elementos que las componen o se actúa incorrectamente sobre ellas, por ejemplo abriendo una puerta, invadiendo una zona con alguna parte del cuerpo, etc. En este tipo de máquinas, no es posible utilizar solamente un sistema electrónico de control de aplicación general, como los autómatas programables descritos en el capítulo 1, porque un fallo de alguno de sus componentes (sensores, módulos de entradas, unidad central, módulos de salidas o actuadores) puede hacer que se produzca un accidente. En este caso se pueden utilizar diferentes estrategias para elevar el nivel de seguridad ante averías de la máquina, una de las cuales consiste en utilizar un sistema complementario o auxiliar que actúe ante una situación de emergencia y lleve la máquina a un estado seguro que implica en general la parada de la misma. Tradicionalmente este tipo de sistemas se realizó mediante circuitos digitales implementados con relés y contactores y recibió el nombre de sistema enclavado de seguridad ante averías (*Safety Interlock system*) o sistema independiente de seguridad. Pero el avance de la Microelectrónica ha hecho que en la actualidad se implementen mediante circuitos electrónicos y que se definan como un dispositivo o grupo de dispositivos electrónicos diseñado para detectar una condición límite o fuera de límite, o una secuencia incorrecta de sucesos y llevar la instalación a una situación segura.

Diversos organismos internacionales de normalización han elaborado normas relativas a los SIS, como por ejemplo la IEC 61508 y la EN 954-1, que establecen que su estructura está formada por un sistema electrónico que, tal como se indica en la figura 10.1, tiene asociados un conjunto de sensores y actuadores y por ello los denominan "*Safety Instrumented Systems*" (SIS) [ANSI 96]. En el apartado A5.5.2 del apéndice 5 se analizan las características técnicas de estos sistemas, así como las normas antes citadas que les son aplicables.

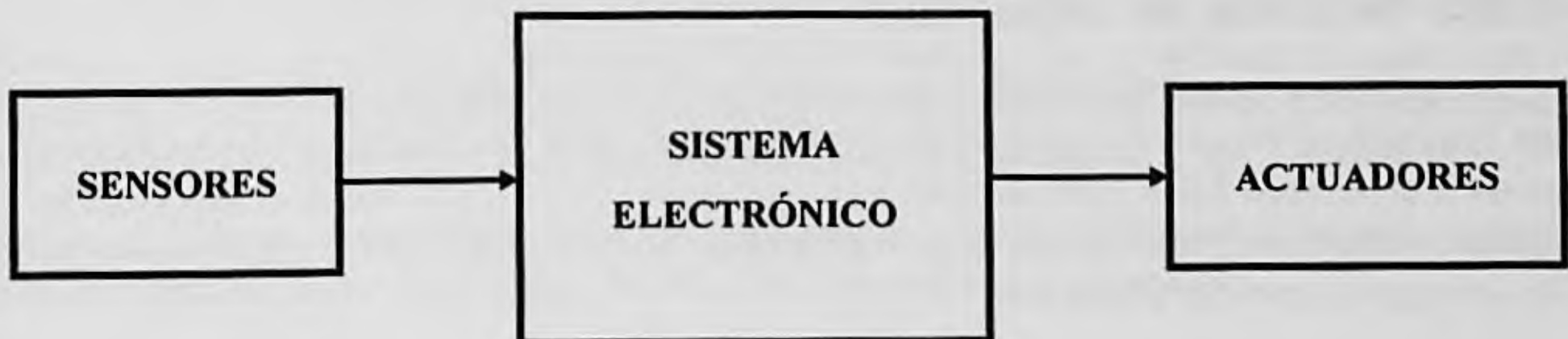


Figura 10.1. Diagrama de bloques de un sistema independiente de seguridad (SIS).

De lo expuesto se deduce que un SIS vigila la operación de un proceso o de una máquina para asegurar que no se produce en él una situación peligrosa debida a una avería de cualquiera de los elementos que la componen o a una utilización inadecuada de los mismos, y por lo tanto debe actuar en las situaciones siguientes:

- Una variable crítica rebasa un valor límite.
- Un operador realiza una acción incorrecta.
- Un operador activa un pulsador de parada de emergencia.
- El propio SIS sufre una avería que le impide llevar a cabo la misión para la que fue diseñado.

Las máquinas o procesos en los que es necesario utilizar un SIS se pueden dividir en dos grupos:

- **Máquinas o procesos sencillos desde el punto de vista de la seguridad ante averías**

Son máquinas o procesos en los que el número de variables cuyo valor es necesario conocer para alcanzar un grado elevado de seguridad ante averías (variables críticas) es reducido. Un ejemplo típico es la detección de la presencia de un operario en un área peligrosa. La misión de un SIS es, en estos casos, la de provocar la parada de la máquina, y por ello se le suele denominar “Dispositivo de paro de emergencia”, “Módulo de seguridad para paradas de emergencia” o “Relé de seguridad”. Un ejemplo típico de aplicación de un SIS de este tipo es la protección del área de trabajo de un robot para evitar un accidente. En este caso, el SIS hace que el robot se pare cuando uno o más sensores de presencia detectan que una persona penetra en el área de trabajo del robot. Al estudio de los relés de seguridad se dedica el apartado 10.2.2 a continuación.

- **Máquinas o procesos complejos desde el punto de vista de la seguridad ante averías**

Son máquinas o procesos en los que el SIS debe ser capaz de observar muchas variables de distinta naturaleza. Aunque algunos fabricantes han desarrollado en el pasado SIS de elevada complejidad, la solución más utilizada en la actualidad, debido a que diversas variables del proceso deben ser controladas automáticamente, consiste en un sistema electrónico seguro ante averías que además de las tareas de seguridad ante averías lleva a cabo las funciones de control. Por ello se les suele denominar sistemas electrónicos de control seguros ante averías y al estudio de los autómatas programables seguros ante averías, también denominados autómatas programables de seguridad, se dedica el apartado 10.3.3.2.

10.2.2 Módulos de seguridad

Los módulos de seguridad que, como se indica en el apartado anterior, también se denominan dispositivos de paro de emergencia o relés de seguridad (en la bibliografía anglosajona se les suele denominar MCR (acrónimo de *Master Control Relay*), son sistemas electrónicos digitales seguros ante averías que poseen un número reducido de variables de entrada y salida. Sus principales características son:

- Utilizan redundancia masiva (descrita en el apartado A5.4.2.2 del apéndice 5).
- Suelen disponer de un número reducido de variables de entrada del tipo todo-nada (*On-off*) y variables de salida de tipo relé con contactos normalmente abiertos y contactos

normalmente cerrados, así como elementos de salida del tipo semiconductor [por ejemplo diodos luminiscentes (*LED*)] para señalización.

- Pueden alcanzar distintas categorías de seguridad ante averías, pero siempre son sistemas redundantes y por lo menos deben disponer de una entrada redundante y de una salida también redundante.
- Las variables de salida de tipo relé utilizan contactos de elevada fiabilidad que poseen un mecanismo que garantiza su apertura incluso aunque se queden pegados por efecto de un pico excesivo de corriente.

Los primeros módulos de seguridad se implementaron con dispositivos electrónicos discretos que constituían sistemas digitales cableados. En la actualidad están implementados con dos microcontroladores [MAND 07] [MART 93] [MART 03] que trabajan en paralelo y se vigilan mutuamente (Figura 10.2).

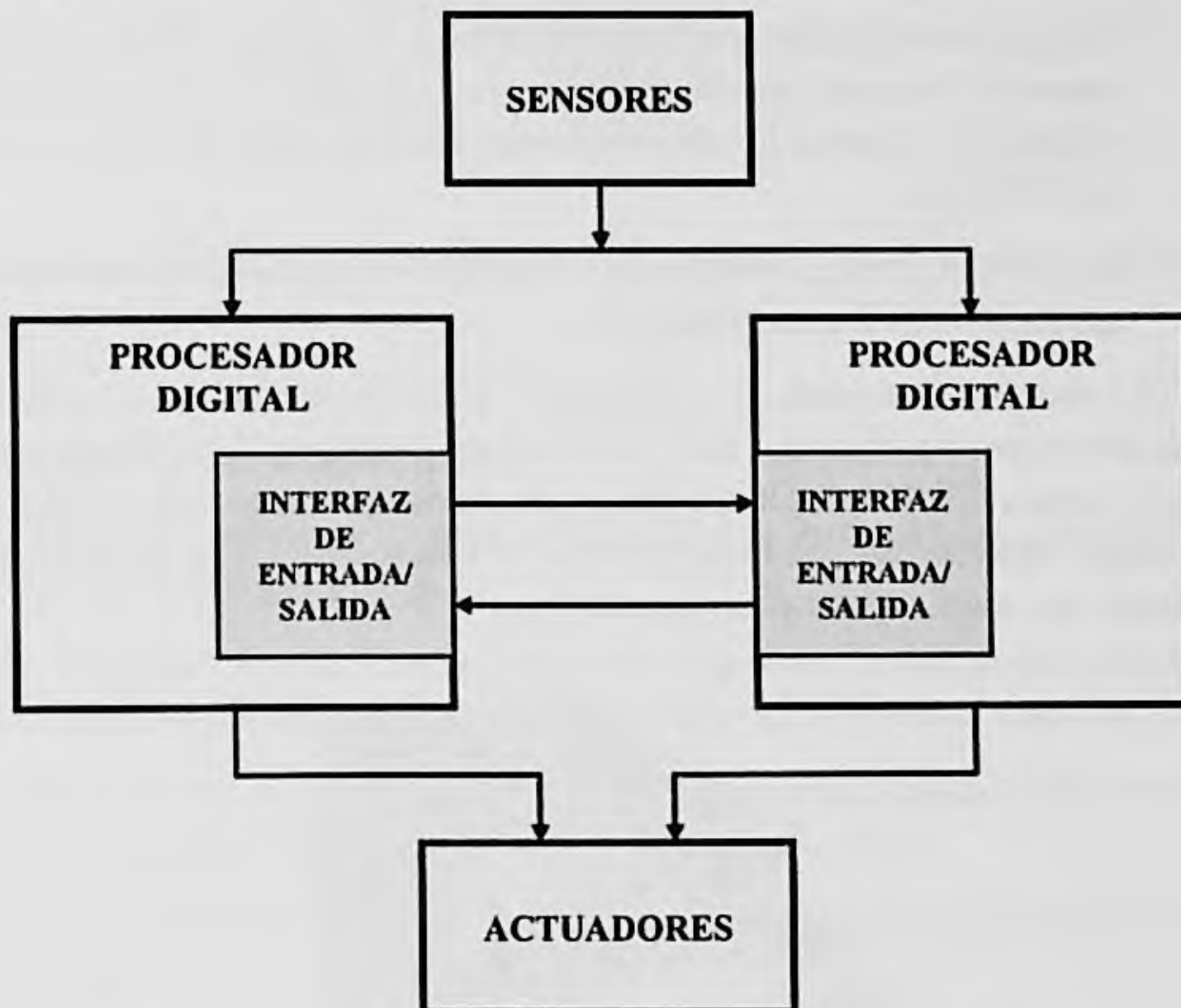


Figura 10.2. Diagrama de bloques de un módulo o relé de seguridad.

Un ejemplo de módulos de seguridad para paradas de emergencia o relés de seguridad son los elementos pertenecientes a la familia 3TK28 de Siemens, que se diferencian entre sí en el número de variables de entrada y salida y en el nivel de seguridad ante averías que proporcionan de acuerdo con la norma EN 954-1, que se describe en el apartado A5.5.2.5.2 del apéndice 5.

Los módulos de seguridad de la familia 3KT28 tienen un comportamiento seguro, que consiste en la desactivación de la carga, debido a las circunstancias siguientes:

- **Protección contra circuito abierto**

Detectan el circuito abierto, al romperse uno de los cables de conexión del sensor de emergencia (SE), mediante la medida de la caída de tensión.

- **Apertura positiva de los contactos**

Poseen un mecanismo de accionamiento de los contactos que garantiza su apertura aunque se queden pegados. Los contactos son de elevada fiabilidad y alcanzan tasas de fallo del orden de un fallo por cada 100 millones de operaciones.

- Detectan la conexión a masa de uno de los cables del sensor de emergencia SE o el cortocircuito entre ellos.
- Se alimentan en corriente continua para asegurar que la bobina del relé no se activa debido a las capacidades parásitas del cableado.
- Se diseñan de forma que la falta de tensión de alimentación o su disminución por debajo de un cierto valor supone el paso al estado seguro de la máquina. Para ello utilizan contactos normalmente abiertos que están cerrados durante el funcionamiento normal de la máquina.
- Tienen incorporados los relés de salida y se deben utilizar conjuntamente con sensores de seguridad (descritos en el apartado 10.2.3 a continuación) para que el conjunto alcance las distintas categorías de seguridad ante averías de la norma EN954-1.
- Se utilizan principalmente con sensores de proximidad, que se analizan en el apartado 10.2.3 a continuación.

La figura 10.3 muestra el módulo de seguridad 3TK28 25 de Siemens que dispone de 3 contactos de salida normalmente abiertos (NA) y 2 normalmente cerrados (NC) y puede alcanzar la categoría 4 de la norma EN 954-1. Otro ejemplo de módulo de seguridad es el modelo 3RG78 47-4BJ de Siemens, representado en la figura 10.4, al que solamente se pueden conectar sensores de proximidad optoelectrónicos de seguridad.

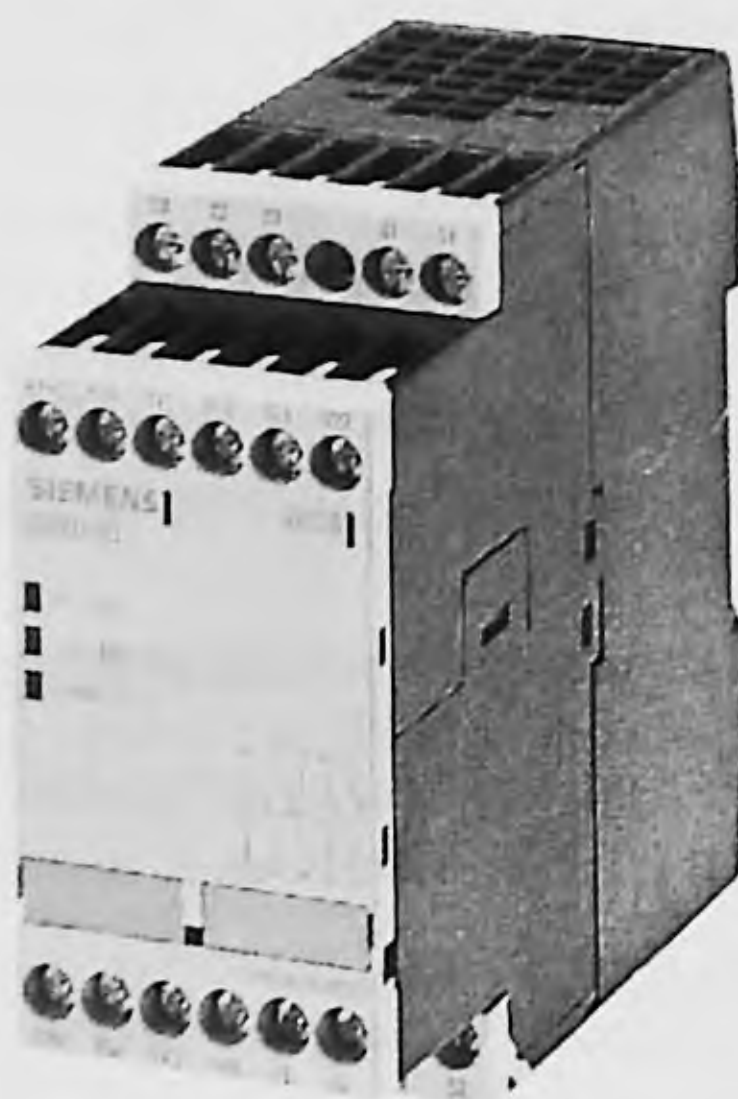


Figura 10.3. Módulo de seguridad 3TK28 25 de Siemens.

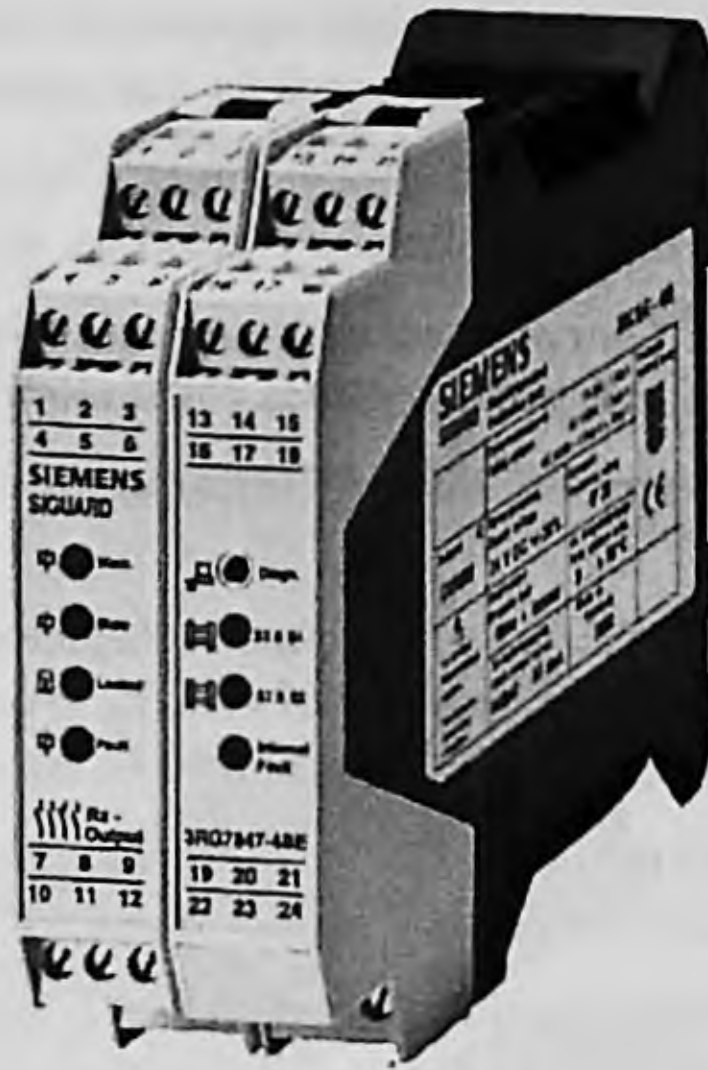


Figura 10.4. Módulo de seguridad 3RG78 47 de Siemens.

10.2.3 Sensores de seguridad

Se suelen denominar sensores de seguridad los que se utilizan para detectar la presencia de personas y objetos en áreas peligrosas.

Aunque las características de los sensores de seguridad dependen del principio físico que utilizan, poseen un conjunto de características comunes a todos ellos, entre las que cabe citar:

- Cumplen uno de los niveles de seguridad ante averías, especificados en la norma EN 954-1 descrita en el apartado A5.5.2.5.2 del apéndice 5.
- Cumplen la norma EN 60204-1 relativa a la seguridad de equipos eléctricos utilizados en máquinas.
- Están específicamente diseñados para ser utilizados con sistemas electrónicos seguros ante averías.

A continuación se estudian brevemente los principales sensores de proximidad de seguridad.

10.2.3.1 Sensores de proximidad de seguridad de actuación mecánica

Estos sensores se denominan también “Interruptores de posición” o “Finales de carrera de seguridad” y poseen las siguientes características generales:

- Se colocan dentro de una caja de material aislante no inflamable.
- Poseen un grado de protección IP elevado (descrito en el apartado A1.5.2 / apéndice 1).

- Suelen disponer de contactos dobles, lo que por un lado facilita la conmutación de elevadas corrientes y por el otro proporciona una resistencia del contacto reducida, imprescindible cuando controlan pequeñas corrientes.
- Los contactos tienen separación galvánica entre ellos, lo que hace que se puedan utilizar en circuitos diferentes.
- Tienen una elevada vida útil, tanto mecánica como eléctrica.

Son ejemplo de este tipo los sensores SIGUARD de la serie 3SE 100 (Figura 10.5), de Siemens, que se caracterizan por:

- Su sistema mecánico está diseñado para que los contactos normalmente cerrados (NC) realicen “apertura positiva” (norma IEC 60947-5-1), lo que garantiza su apertura incluso aunque los contactos se queden pegados debido a un pico de corriente.
- Su vida útil mecánica es del orden de 30 millones de operaciones y la eléctrica del orden de los 10 millones de operaciones.
- Están homologados como dispositivos de enclavamiento de acuerdo con las exigencias de las normas EN 1088 y EN 292.
- Su estado de funcionamiento se indica mediante dos diodos luminiscentes (LED).

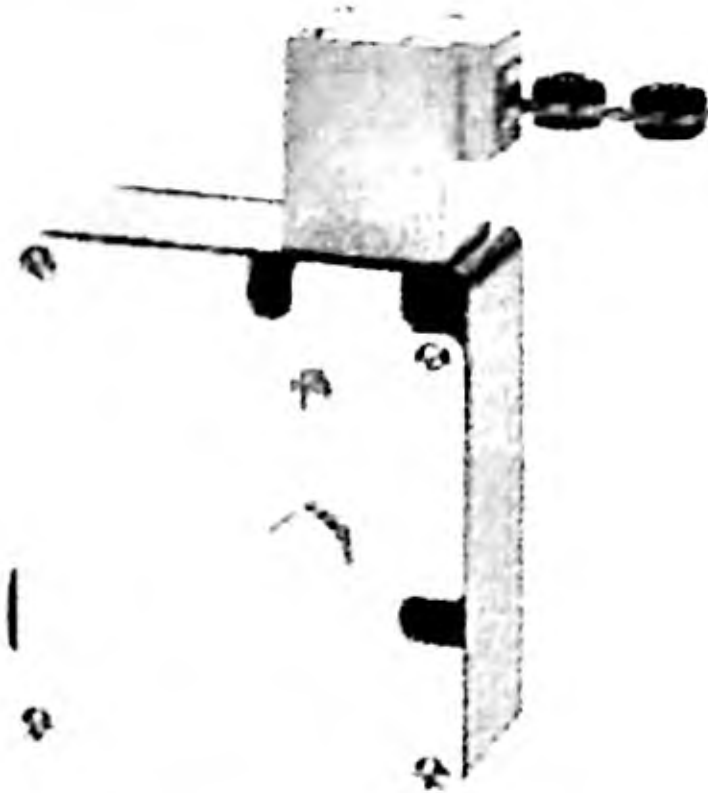


Figura 10.5. *Final de carrera de seguridad SIGUARD de la serie 3SE 100 de Siemens.*

También se pueden considerar interruptores mecánicos de seguridad los sensores de mando a dos manos, como por ejemplo los sensores SIGUARD de la serie 3SB38 de Siemens y los interruptores de pedal, como por ejemplo los sensores SIGUARD de la serie 3SE3 de Siemens.

10.2.3.2 Sensores magnéticos de posición de seguridad

Este tipo de sensores está formado por un imán, un sensor de campo magnético y el circuito electrónico de control asociado. Presentan las siguientes características:

- Son idóneos para ser utilizados en las instalaciones en las que en el entorno del sensor hay agentes que pueden afectar a su funcionamiento (líquidos, polvo, etc.).

- Su salida es del tipo relé de seguridad con contactos normalmente cerrados (NC) y normalmente abiertos (NA).
- Suelen tener salidas del tipo transistor para conectarlos a la entrada de un autómata programable.
- Alcanzan las categorías más exigentes de la norma EN 954-1 (apartado A5.5.2.5.2 del apéndice 5).

Son ejemplo de este tipo los sensores SIGUARD de la serie 3SE6, de Siemens (Figura 10.6).



Figura 10.6. Sensor magnético de posición de seguridad SIGUARD de la serie 3SE6 de Siemens.

10.2.3.3 Sensores optoelectrónicos de seguridad detectores de objetos

Este tipo de sensores se basa en la emisión y recepción de luz para detectar la presencia de personas y/o objetos en zonas peligrosas [CEDE 02]. Se implementan de varias formas diferentes que se describen a continuación.

Cortinas fotoeléctricas

Están constituidas por un número elevado de dispositivos optoelectrónicos emisores de luz y receptores de luz. Los rayos de luz que van del emisor al receptor forman una cortina de luz, que es interrumpida por la sombra que provoca un objeto o una persona al cruzarla. En general disponen de salidas de tipo transistor para conectarlos a un sistema electrónico de control. Son ejemplo de este tipo los sensores SIGUARD de la serie 3RG78 de Siemens (Figura 10.7), que cumplen la norma EN 61496-1 y alcanzan la categoría de seguridad 4 de la norma EN 954-1 (apartado A5.5.2.5.2 del apéndice 5).

En la figura 10.8 se representa un ejemplo de aplicación de una cortina fotoeléctrica, cuyas principales parámetros son:

- La resolución, que establece las dimensiones mínimas que debe tener un objeto para ser detectado.
- La altura de campo, que indica la altura de la zona de detección.
- El alcance, que define la separación que puede haber entre los emisores y los receptores de luz.

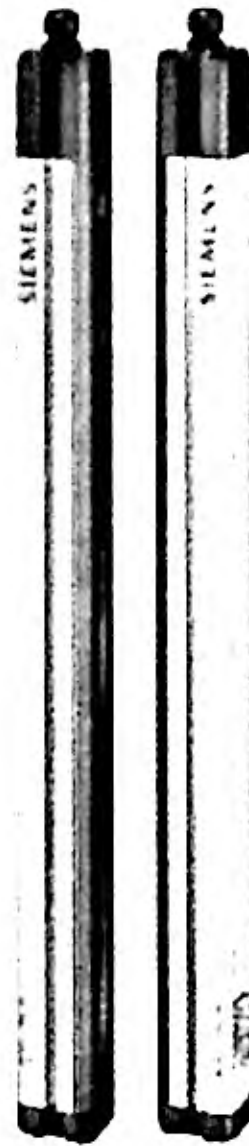


Figura 10.7. Cortina fotoeléctrica SIGUARD de la serie 3RG78 de Siemens.

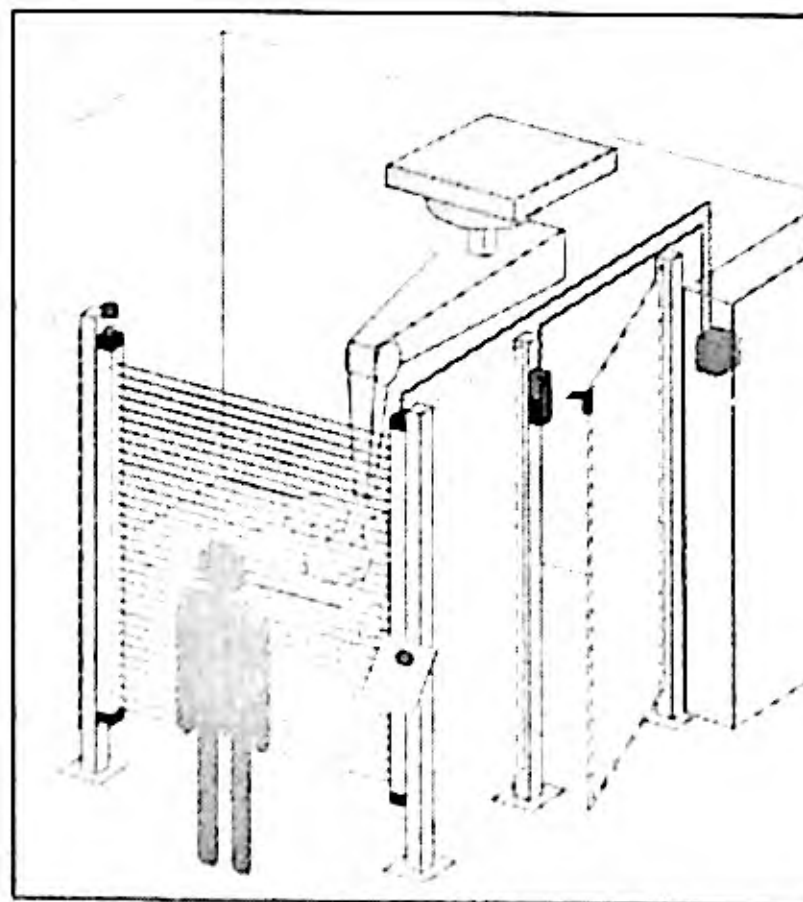


Figura 10.8. Ejemplo de aplicación de una cortina fotoeléctrica SIGUARD de la serie 3RG78 de Siemens.

Rejillas fotoeléctricas

Disponen de un número reducido de dispositivos optoelectrónicos emisores y detectores de luz y sus principales características son:

- La resolución, que representa el número de rayos de que dispone la rejilla. Son valores típicos 2, 3 y 4 rayos.
- La altura de campo, que indica la altura de la zona de detección.
- El alcance, que define la separación que puede haber entre los emisores y los receptores de luz.

En general disponen de salidas realizadas con dispositivos electrónicos (de tipo semiconductor) para conectarlos a un sistema electrónico de control. Son ejemplo de este tipo los sensores de rejilla fotoeléctrica SIGUARD de la serie 3RG78 de Siemens, que cumplen la norma EN 61496-1 y pueden alcanzar la categoría de seguridad 4 de la norma EN 954-1 (apartado A5.5.2.5.2 del apéndice 5).

Escáner (explorador) láser

Son sistemas optoelectrónicos que emiten rayos láser en un ángulo amplio del orden de 180° y lo barren mediante pequeños desplazamientos de un espejo giratorio. Cada barrido se realiza en un intervalo de tiempo reducido del orden de las decenas de milisegundos. Detectan la presencia de objetos mediante la medida del tiempo que tarda la luz en alcanzar el detector situado próximo al emisor.

Son ejemplo de este tipo los sensores SIGUARD de la serie LS4, de Siemens (Figura 10.9) que cumplen la norma de protección contra radiación láser EN 60825-1, además de la EN 61496-1 y EN 61496-3 y alcanzan la categoría 3 de la norma EN 954-1 (apartado A5.5.2.5.2 del apéndice 5).



Figura 10.9. Escáner láser SIGUARD de la serie LS4 de Siemens.

Barrera fotoeléctrica de seguridad

La barrera de luz combina un solo emisor y un receptor sólo de luz. Si un objeto o una persona entran en la zona de seguridad se interrumpe el haz de luz y se produce la detección. Un ejemplo de este tipo son los sensores SIGUARD de la serie 3RG78 de Siemens, que alcanzan la categoría 4, según la norma EN 954-1 (apartado A5.5.2.5.2 del apéndice 5).

10.3 Confiabilidad de los autómatas programables

10.3.1 Introducción

Tal como se indica en el apartado 10.1, la utilización de los sistemas electrónicos de control para automatizar numerosos procesos industriales ha hecho que los usuarios deban tener cada vez más en cuenta los aspectos relacionados con la confiabilidad de los mismos. Por ser el autómata programable el principal sistema electrónico utilizado para automatizar los procesos industriales, al estudio de su confiabilidad se dedica este apartado.

De acuerdo con el nivel de confiabilidad los autómatas programables se pueden clasificar en tres clases:

- **Autómatas programables de aplicación general**

Se consideran autómatas programables de aplicación general aquellos que no están especialmente concebidos para ser utilizados en aplicaciones en las que se necesita una elevada confiabilidad, y por ello corresponden a la categoría B de la norma EN954-1 y a la SIL 1 de la norma IEC 61508, descritas respectivamente en los apartados A5.5.2.5.2 y A5.5.2.5.3 del apéndice 5. Son ejemplos de este tipo de autómatas programables los pertenecientes a las familias S-200, S7-300 y S7-400 de Siemens que se describen en el apartado 1.3.3.3 del capítulo 1.

- **Autómatas programables de seguridad**

Los autómatas programables de seguridad son sistemas electrónicos de control cuya estructura interna coincide con alguna de las establecidas en la norma IEC 61508, descritas en el apartado A5.5.2.5.3 del apéndice 5. Se trata de sistemas redundantes con diagnóstico que junto con los sensores y los actuadores conectados a ellos reciben la denominación de sistemas electrónicos de control seguros ante averías, porque son capaces de realizar funciones de control y de seguridad ante averías, tal como se indica en el apartado A5.5.2.3 del apéndice A5. Son ejemplo de este tipo de autómatas programables los pertenecientes a las familias S7-300F y S7-400F de Siemens que se describen en el apartado 10.3.3.2.

- **Autómatas programables de elevada disponibilidad**

Son autómatas programables que se caracterizan por la elevada probabilidad de que presen servicio en cualquier instante y para ello utilizan redundancia masiva en una de las configuraciones 2003 ó 2004D establecidas en la norma IEC 61508 descrita en el apartado A5.5.2.5.3. Son ejemplo de este tipo de autómatas los pertenecientes a las familias S7-300H y S7-400H de Siemens que se describen en el apartado 10.3.3.3.

En sucesivos apartados se analiza cada una de estas tres clases.

10.3.2 Confiabilidad de los autómatas programables de aplicación general

10.3.2.1 Introducción

Tanto los autómatas programables básicos como los primeros autómatas programables basados en microprocesadores eran de aplicación general porque se diseñaron y comercializaron entre 1970 y 1980, cuando las limitaciones de la capacidad de integración obligaban a reducir al mínimo la complejidad del sistema físico para minimizar su coste y no permitían, por lo tanto, desarrollar sistemas electrónicos programables de elevada confiabilidad (*Dependability*), que, tal como se indica en el apéndice 5, son sistemas redundantes, es decir, que poseen elementos que tienen como misión la detección de fallos. Pero la situación cambió con el progreso de la Microelectrónica, que hizo que los fabricantes de autómatas programables de aplicación general fuesen incorporando paulatinamente los recursos adecuados, tanto físicos (*Hardware*) como de programación (*Software*), que permiten elevar la confiabilidad mediante la verificación (*Checking*) del correcto funcionamiento del sistema, al mismo tiempo que lleva a cabo las tareas de control que tiene encomendadas.

Por otra parte, los usuarios pueden también adoptar estrategias tendentes a elevar la confiabilidad de la instalación en la que se utiliza el autómata programable, mediante la incorporación de recursos físicos externos a los mismos.

Los recursos dedicados a la comprobación del funcionamiento de un sistema reciben por ello la denominación de circuitos de diagnóstico. Al análisis de los recursos internos y externos que se pueden utilizar para elevar la confiabilidad de los autómatas programables de aplicación general se dedican los dos apartados siguientes.

10.3.2.2 Recursos internos de diagnóstico

Son numerosos, y de muy diversa índole, los recursos que un fabricante de autómatas programables de aplicación general puede incorporar para elevar su confiabilidad. A continuación se describen los de uso más generalizado.

10.3.2.2.1 Circuito de vigilancia

El “Circuito de Vigilancia o Perro Guardián” (*Watch Dog*) [SCHW 84] fue uno de los primeros recursos internos de diagnóstico que incorporaron los autómatas programables. En su versión más básica consiste en un monoestable o temporizador redisparable [MAND 08] cuya entrada se activa al comienzo de cada ciclo de trabajo del autómata. Esto hace que su salida permanezca siempre activada excepto cuando el ciclo de trabajo del autómata excede un determinado valor, lo cual indica que se ha producido alguna anomalía en su hardware o en su software. La totalidad de los autómatas programables actuales poseen un circuito de vigilancia que es un temporizador digital programable cuya salida actúa sobre el microprocesador (Figura 10.10) que controla el funcionamiento del autómata (normalmente a través de su sistema de interrupción), y hace que éste se pare (*Stop*). La eficacia de los circuitos de vigilancia, en la comprobación del funcionamiento de los procesadores digitales secuenciales, ha sido probada experimentalmente en aplicaciones de computadores tolerantes a fallos en diversos centros de investigación [GIL 92] [LU 82] [MADE 91] [MAHM 85a] [MAHM 85b] [MAHM 88].

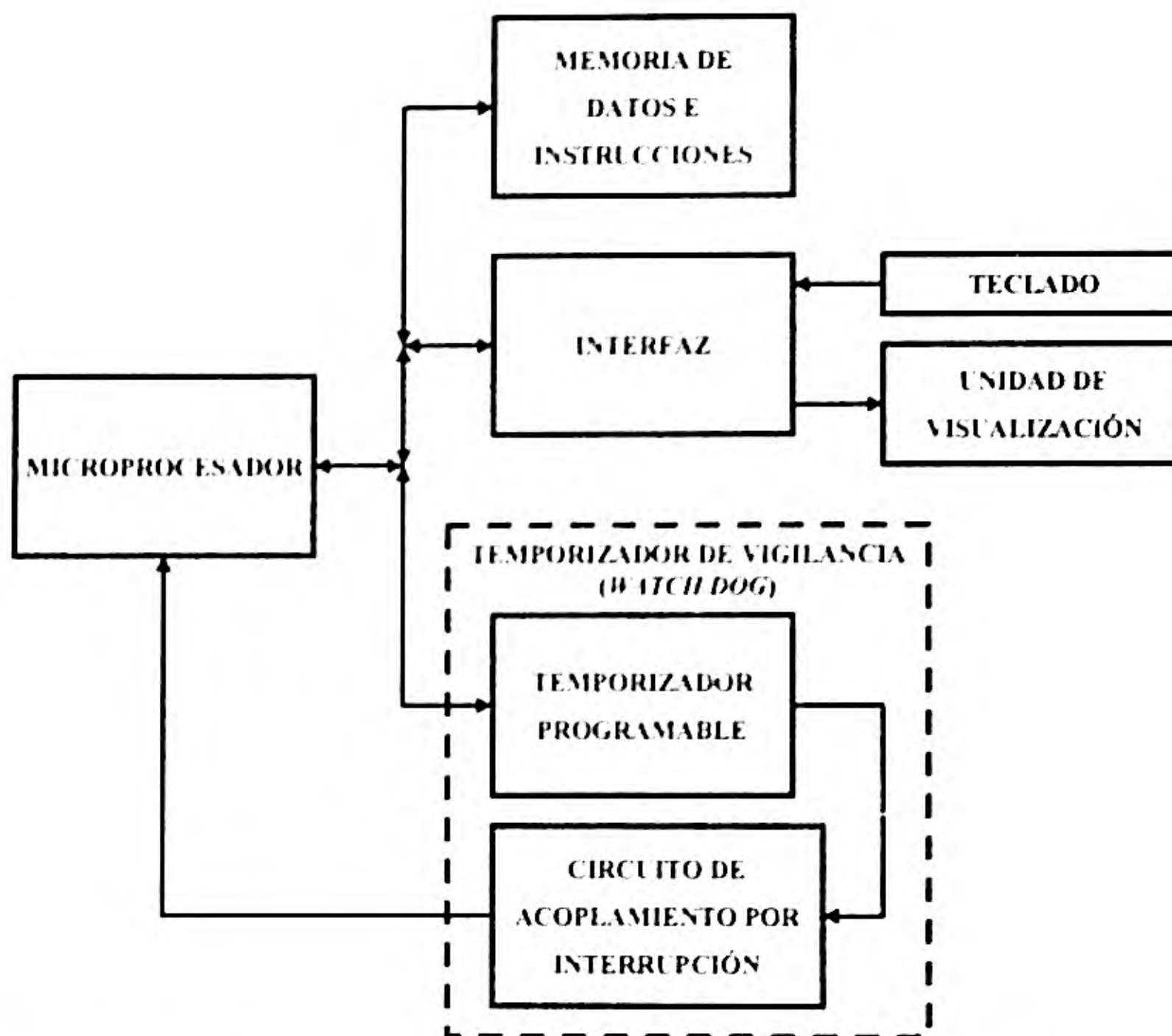


Figura 10.10. Diagrama de bloques de un procesador digital que posee un circuito de vigilancia.

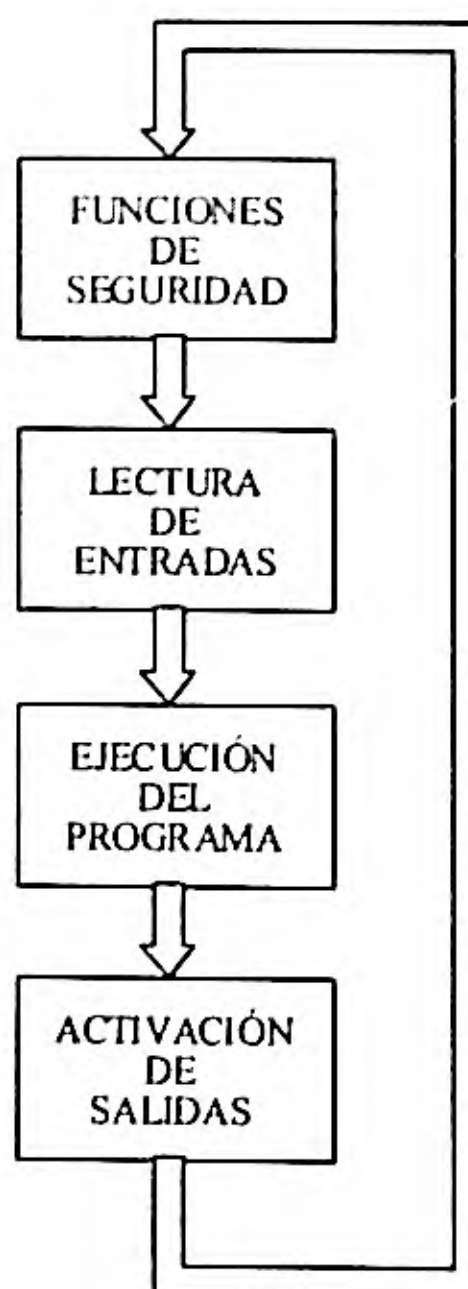


Figura 10.11. Ciclo de funcionamiento de un autómata programable de aplicación general que posee un perro guardián.

El temporizador que realiza la función del perro guardián, junto con otras comprobaciones que el autómata programable lleva a cabo en cada ciclo, da lugar al bloque de funciones de seguridad ante averías que se ejecutan en cada ciclo de trabajo del autómata programable, tal como se indica en la figura 10.11.

10.3.2.2.2 Circuito de comprobación de la batería

Otra comprobación que suelen hacer los autómatas programables de aplicación general es la del nivel de carga de las baterías que alimentan a la memoria que almacena el estado del proceso y el punto de ejecución del ciclo de trabajo cuando se produce un corte de la tensión alterna de alimentación.

10.3.2.2.3 Entradas de seguridad

Las entradas de seguridad son especialmente adecuadas para hacer que el autómata programable atienda de forma prioritaria determinadas variables críticas del proceso bajo control, cuya activación supone una indicación de que existe una anomalía que puede provocar una situación peligrosa. La activación de una entrada de este tipo genera una interrupción (descrita en el apartado 1.3.3.3.3 del capítulo 1) de la unidad central de proceso de la unidad central del autómata programable y hace que deje de ejecutar la tarea que estaba realizando en ese instante y pase a ejecutar una tarea específica, programada por el usuario. Dicha tarea ejecuta el programa adecuado para llevar la instalación a un estado seguro.

10.3.2.2.4 Instrucciones del tipo MCR

Las instrucciones de control de la ejecución de un grupo de instrucciones denominadas MCR (acrónimo de *Master Control Relay*), descritas en el apartado 2.3.8.3 del capítulo 2, hacen que se desactiven en el instante oportuno las salidas comprendidas dentro de un bloque de programa.

10.3.2.2.5 Instrucciones de acceso a la periferia

Las instrucciones de acceso a la periferia del autómata se utilizan para leer el estado de una o varias variables críticas en un determinado instante. Se evita así que se produzca un retraso excesivo en la detección del cambio de una de estas variables, que, en el caso más desfavorable, es igual al ciclo de trabajo del autómata programable. También hay que resaltar que la utilización de estas instrucciones puede hacer que una determinada variable tenga dos valores distintos en un mismo ciclo de proceso, lo cual se debe tener en cuenta al diseñar el programa de control.

10.3.2.3 Recursos externos de diagnóstico

Los recursos internos de diagnóstico descritos en el apartado anterior están incorporados en la unidad central de los autómatas programables de aplicación general actuales de la mayoría de los fabricantes, pero es un hecho, verificado experimentalmente, que más del 90% de los fallos de los autómatas programables se producen en la periferia de los mismos, es decir, en los dispositivos de campo (sensores y actuadores), en los módulos de entrada/salida o en el cableado entre éstos y aquéllos.

Por ello, mediante la utilización de técnicas de redundancia selectiva (descritas en el apartado A5.4.2.2 del apéndice 5) en la periferia del autómata programable es posible detectar muchos de estos fallos e incluso se puede lograr que el sistema sea tolerante a algunos de ellos. Desde este punto de vista, a continuación se describen diversas soluciones tecnológicas que incrementan la confiabilidad de los autómatas programables de aplicación general. Es conveniente indicar que estas soluciones son aplicables también a los autómatas programables de seguridad y a los de elevada disponibilidad que se analizan en los apartados 10.3.3.2 y 10.3.3.3 respectivamente.

10.3.2.3.1 Seguridad de las entradas

Una de las averías típicas que se producen en los sistemas electrónicos de control basados en autómatas programables es debida a que uno de los circuitos de un módulo de variables de entrada todo-nada del autómata programable proporciona una información incorrecta debido a un fallo del propio circuito, del sensor asociado con él o de la conexión entre ambos. Por ello, en los casos en los que es necesario elevar el nivel de seguridad ante averías de un autómata programable de aplicación general, se debe:

- Seleccionar las variables de entrada críticas, es decir, aquellas que proporcionan una información fundamental para la seguridad ante averías del sistema porque un fallo de las mismas produce una disminución del nivel de seguridad ante averías del sistema controlado por el autómata programable.
- Hacer redundantes los elementos sensores y sus terminales de conexión con el autómata programable (junto con el circuito interno correspondiente), asignados a las citadas variables.

A continuación se describen diversas alternativas. Todas ellas se caracterizan por combinar la redundancia espacial (*Hardware redundancy*) con la redundancia por programa (*Software redundancy*), que constituye una redundancia temporal. Los programas correspondientes se indican en el lenguaje de esquema de contactos.

Duplicación de sensores y unidades de entrada

Esta solución consiste en:

- Utilizar dos sensores que detecten la misma variable física y conectarlos a través de terminales de entrada diferentes que proporcionen dos variables de entrada distintas, como por ejemplo E 0.0 y E 0.1.
- Realizar un programa (Figura 10.12a) que genere una variable de estado interno M 0.0, que se active solamente cuando lo hacen los dos sensores. La variable de estado interno M 0.0 se utiliza como variable de entrada del programa de control de forma que su activación indica que la variable crítica está activada (situación peligrosa).
- Generar una variable de estado interno M 0.1 (Figura 10.12b), que se utiliza para detectar que las dos variables de entrada E 0.0 y E 0.1 tienen valores lógicos diferentes. La activación de M 0.1 puede ser utilizada por el programa del autómata programable para hacer que la instalación alcance un estado seguro mediante la activación de una variable de salida través de un temporizador. El temporizador tiene como misión evitar

el paro imprevisto debido a que las dos variables de entrada E 0.0 y E 0.1 no se activen simultáneamente. El valor de la temporización se debe ajustar para cada aplicación.

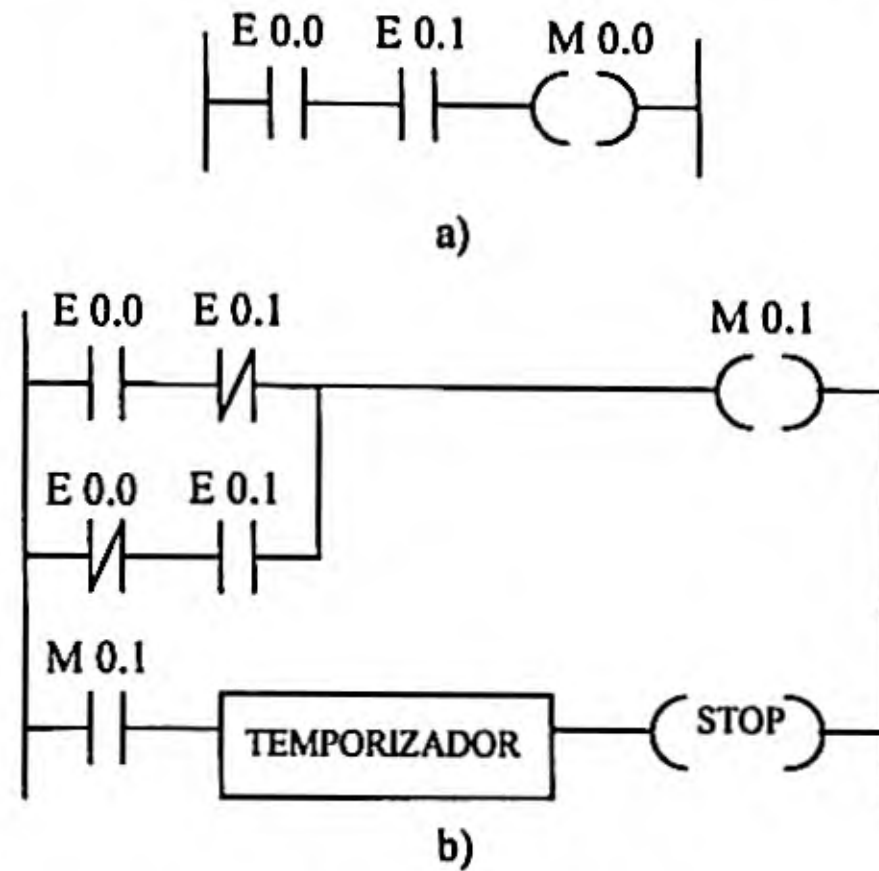


Figura 10.12. Programa, realizado en el lenguaje de esquema de contactos, asociado a la duplicación de sensores y unidades de entrada.

En el método descrito, cualquier discrepancia entre el estado de las variables de entrada E 0.0 y E 0.1 determina la existencia de un fallo, pero no permite detectar cuál es la entrada defectuosa y en consecuencia es imposible conocer si la variable crítica está en realidad activada o no. Debido a ello, y en el supuesto de que se pueda producir una situación peligrosa, es necesario detener el sistema para preservar la seguridad ante averías y averiguar dónde se ha producido el fallo. El sistema no tiene necesariamente que pasar a STOP, sino que puede pasar a cualquier otro estado seguro que se haya previsto a tal efecto.

Triplicación de sensores y unidades de entrada

Esta variante consiste en:

- Utilizar tres sensores que detecten la misma variable física y conectarlos a través de terminales de entrada diferentes que proporcionen tres variables de entrada distintas, como por ejemplo E 0.0, E 0.1 y E 0.2.
- Generar una variable de estado interno M 0.2 que se utiliza como variable de entrada del programa de control, tal como se indica en la figura 10.13, en la que dicha variable se activa solamente cuando lo hacen dos cualesquiera de los tres sensores o los tres. En la aplicación del método se supone que no se averían dos sensores o su correspondiente circuito de entrada, simultáneamente. Si la información de un sensor difiere de la de los otros dos, se supone que éste es el sensor estropeado y la información proporcionada por los otros dos es la correcta.
- Generar tres variables de estado interno M 0.3, M 0.4 y M 0.5, que se utilizan para detectar cual es el sensor que ha fallado. Tal como se indica en la figura 10.14, la variable de estado interno M 0.3 se activa cuando la información de E 0.0 difiere de la de E 0.1 y E 0.2. Por lo tanto las variables M 0.2 y M 0.3 se activan simultáneamente cuando falla

el sensor E 0.0. De forma similar M 0.2 y M 0.4 se activan simultáneamente cuando falla el sensor E 0.1 y M 0.2 y M 0.5 se activan simultáneamente cuando falla E 0.2.

- Generar una variable de salida, A 1.0, para indicar al usuario que existe una avería aunque la instalación continúe funcionando correctamente (Figura 10.15). A 1.0 se activa un cierto tiempo después de activarse alguna de las tres variables M 0.3, M 0.4, o M 0.5, para evitar una activación incorrecta debida a que los tres sensores E 0.0, E 0.1 y E 0.2 no se activan exactamente en el mismo instante.

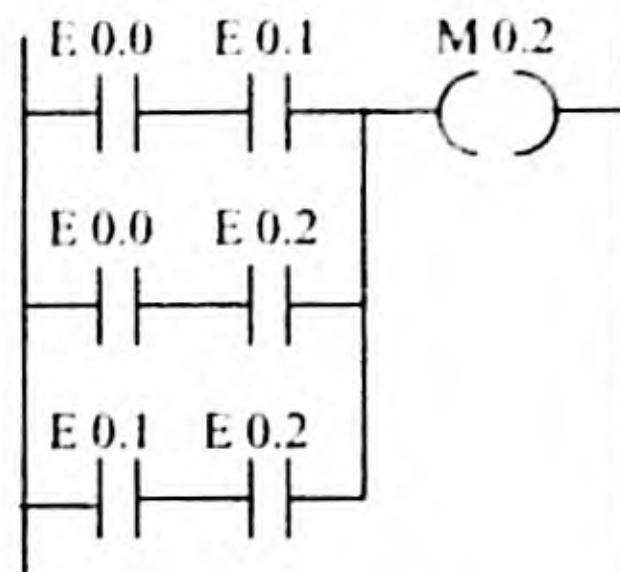


Figura 10.13. Programa, realizado en el lenguaje de esquema de contactos, asociado a la triplicación de sensores y unidades de entrada.

Pero, dado que el circuito del módulo de salida conectado a la variable A 1.0 también puede fallar, es posible que el servicio de mantenimiento no detecte la avería. Si esto sucede y falla otro sensor antes de que el averiado sea sustituido, se produce una situación en la que el programa de la figura 10.14 da un resultado erróneo. Por ello, para evitar una pérdida de seguridad ante averías hay que aumentar la redundancia y entre las estrategias posibles cabe citar:

- Autocomprobar periódicamente que la salida A 1.0, que está normalmente desactivada, funciona correctamente. Esta opción implica la utilización de un hardware complementario que se describe en el apartado 10.3.2.3.2 (Figura 10.18).
- Pasar al sistema de duplicación cuando se detecta un fallo en uno de los sensores (se activa A 1.0). Para ello, cuando se activa A 1.0, se dejan de ejecutar los programas de las figuras 10.14 y 10.15 y se ejecuta el programa de la figura 10.16, que consiste en un sistema de duplicación. Conocido el sensor que ha fallado, se comprueba que la información proporcionada por los otros dos restantes coincide. En caso de que no sea así, se activa una variable interna M 0.6. Si después del tiempo programado en el temporizador, la variable M 0.6 continúa activada, el sistema se detiene. Es muy recomendable que el sistema detenido pase a un estado seguro para la instalación, en el que sea posible acceder a las variables internas del autómata programable para comprobar que es lo que ha fallado y subsanarlo antes de arrancarlo de nuevo. Si el sistema se pone en marcha de nuevo sin comprobar que es lo que ha fallado, las rutinas de comprobación anteriores no funcionan correctamente. Dado que el sistema produce una salida correcta a pesar de la existencia de un solo fallo en las entradas, se consigue la tolerancia del sistema contra fallos en los sensores y sus circuitos de entrada.

Autocomprobación de las interfaces de entrada

Esta forma de elevar el nivel de seguridad ante averías de las entradas consiste en que el autómata programable se autocompruebe periódicamente para detectar la avería de uno de los interfaces de entrada y dar la correspondiente señal de alarma. Se puede implementar mediante un relé activado por una variable de salida del propio autómata programable, haciendo que uno de sus contactos active la variable de entrada cuyo funcionamiento se quiere comprobar. En la figura 10.17 se representa el esquema correspondiente, en el que el sensor conectado a la entrada se puentea mediante el contacto K del relé en el instante en el que se quiere comprobar el circuito de la interfaz de entrada. Mediante esta estrategia se pueden comprobar varias entradas del autómata programable utilizando un único contacto auxiliar del relé. Evidentemente, para llevar a cabo las comprobaciones el autómata programable, ha de disponer de tiempo para ejecutar las instrucciones necesarias porque esta estrategia es un ejemplo de redundancia temporal (descrita en el apartado A5.4.2).

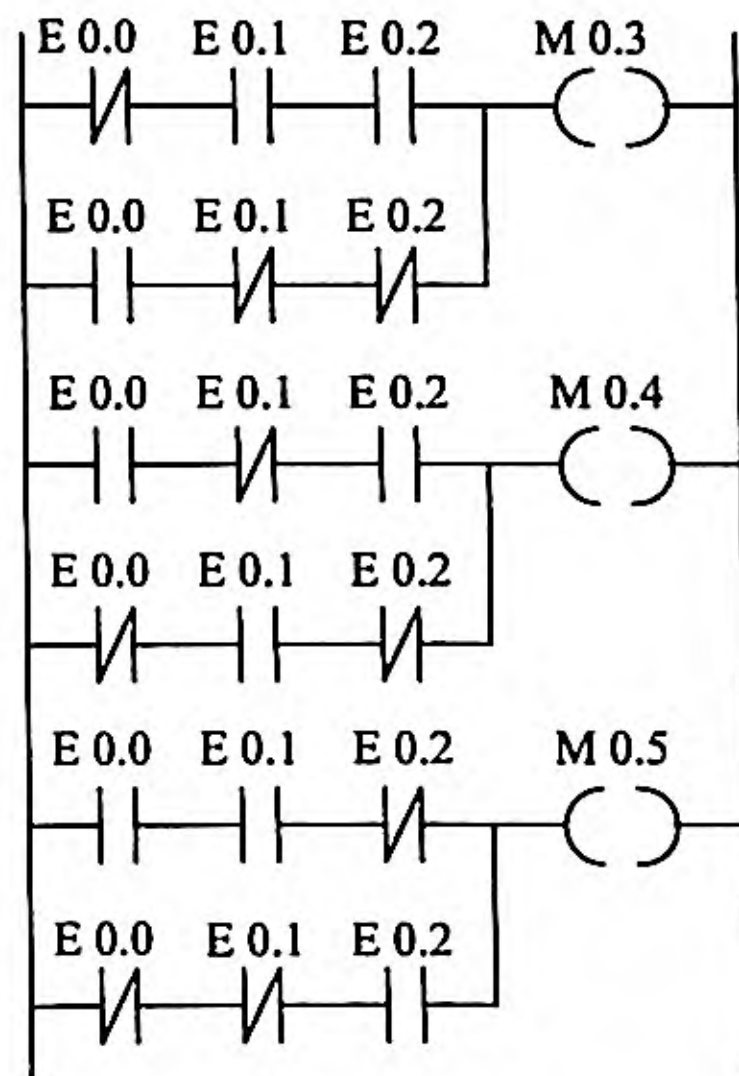


Figura 10.14. Programa, realizado en el lenguaje de esquema de contactos, de generación de tres variables de estado interno M 0.3, M 0.4 y M 0.5, utilizadas para detectar cual es el sensor que ha fallado.

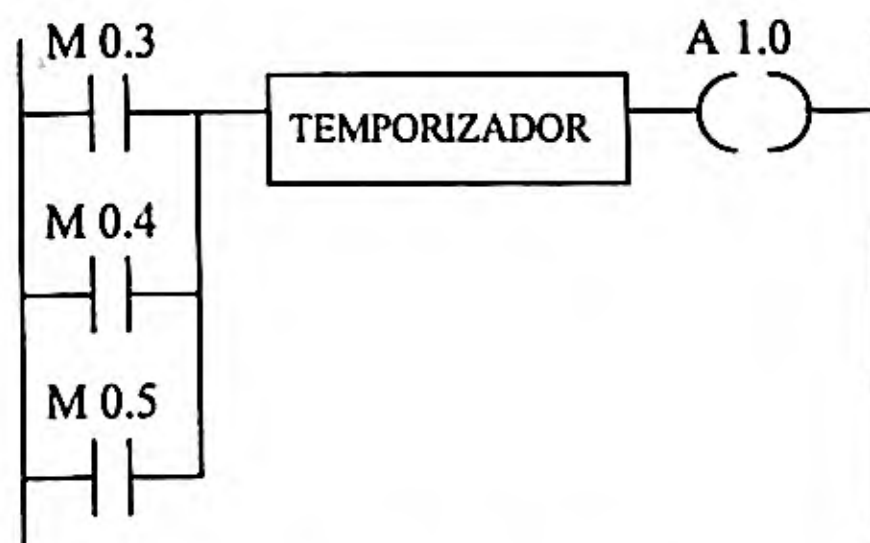


Figura 10.15. Programa de duplicación, realizado en el lenguaje de esquema de contactos, que genera la variable de salida A 1.0 para indicar al usuario que existe una avería.

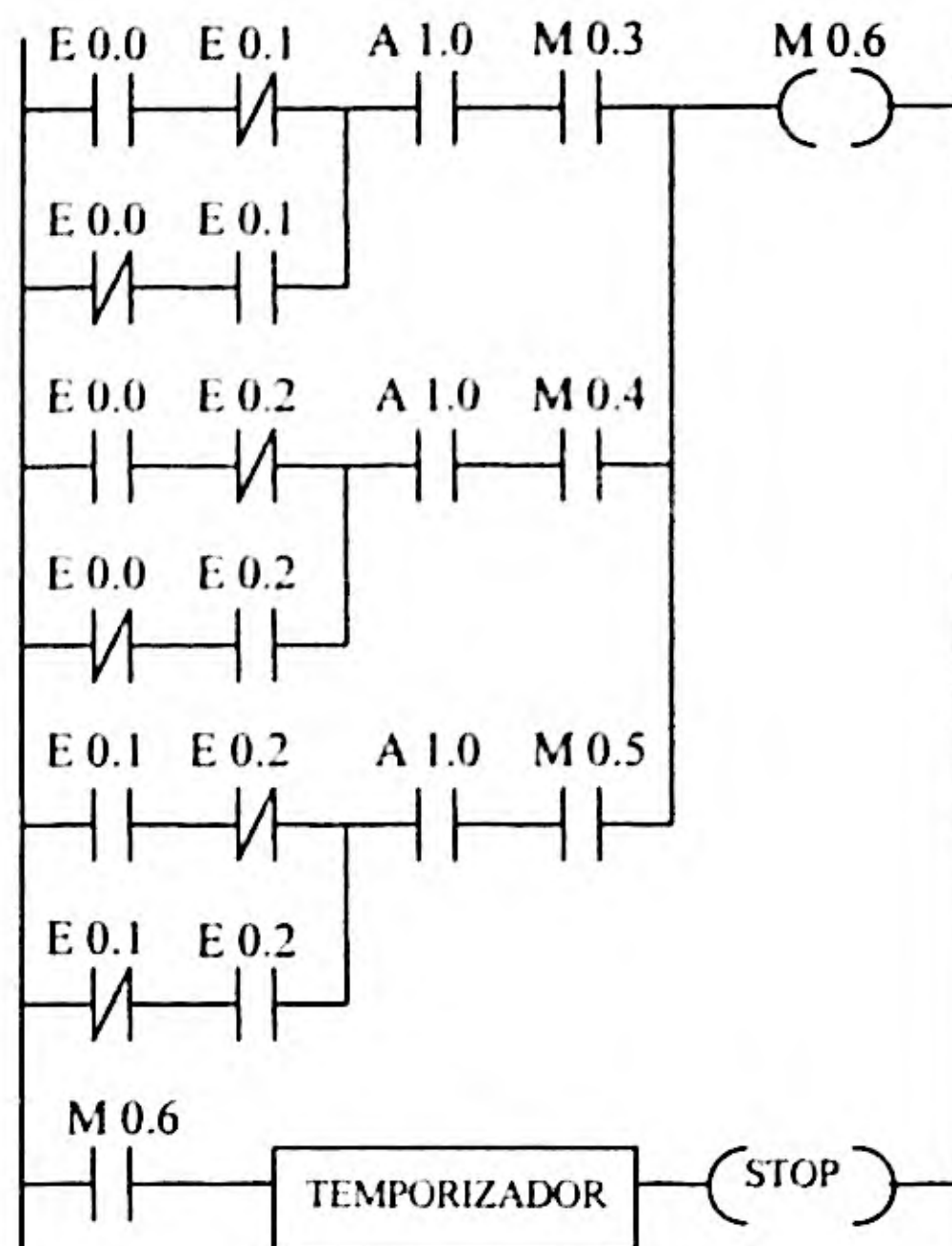


Figura 10.16. Programa de duplicación, realizado en el lenguaje de esquema de contactos, que se ejecuta cuando se activa la variable A 1.0, generada por el programa de las figuras 10.14 y 10.15.

10.3.2.3.2 Seguridad de las salidas

Desde el instante en que se comercializaron los primeros autómatas programables, a finales de la década de 1960, se han utilizado diversas formas de conseguir un funcionamiento seguro de las variables de salida todo-nada. En la figura 10.18 se representa una solución simple y muy eficaz consistente en hacer que el autómata programable detecte si la salida correspondiente está realmente activada o desactivada.

Se supone que la variable de salida todo-nada A1.0 del tipo relé, cuyo funcionamiento correcto se desea comprobar, está realizada con un transistor PNP. A dicha salida se conecta un relé (K) que además de activar el correspondiente actuador a través de su contacto K1, posee un contacto auxiliar K2 que se utiliza para activar la entrada E 0.0 del autómata programable. Para conocer si la activación o desactivación de la salida A 1.0 se corresponde realmente con la activación o desactivación del relé K se ejecuta el programa representado en el lenguaje de esquema de contactos en la figura 10.19. La activación de la salida A1.1 indica que hay un fallo en el circuito de la interfaz de salida o en el propio relé. La seguridad ante averías se puede además incrementar mediante la utilización de un relé de seguridad.

Otras soluciones muy efectivas, aunque más complejas, se basan en la utilización de la lógica dinámica [VAUT 83a] [VAUT 83b].

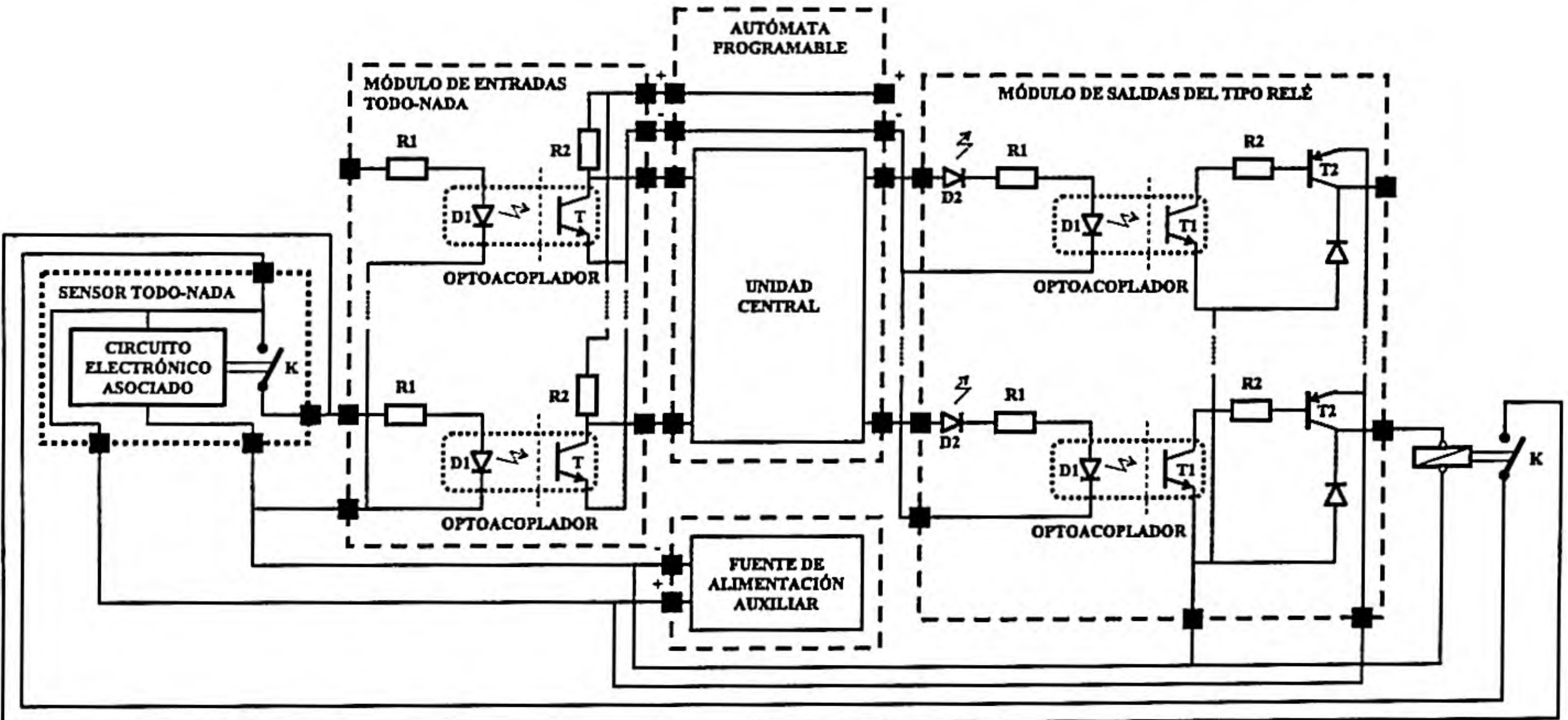


Figura 10.17. Circuito para comprobar una avería en un módulo de variables de entrada todo-nada.

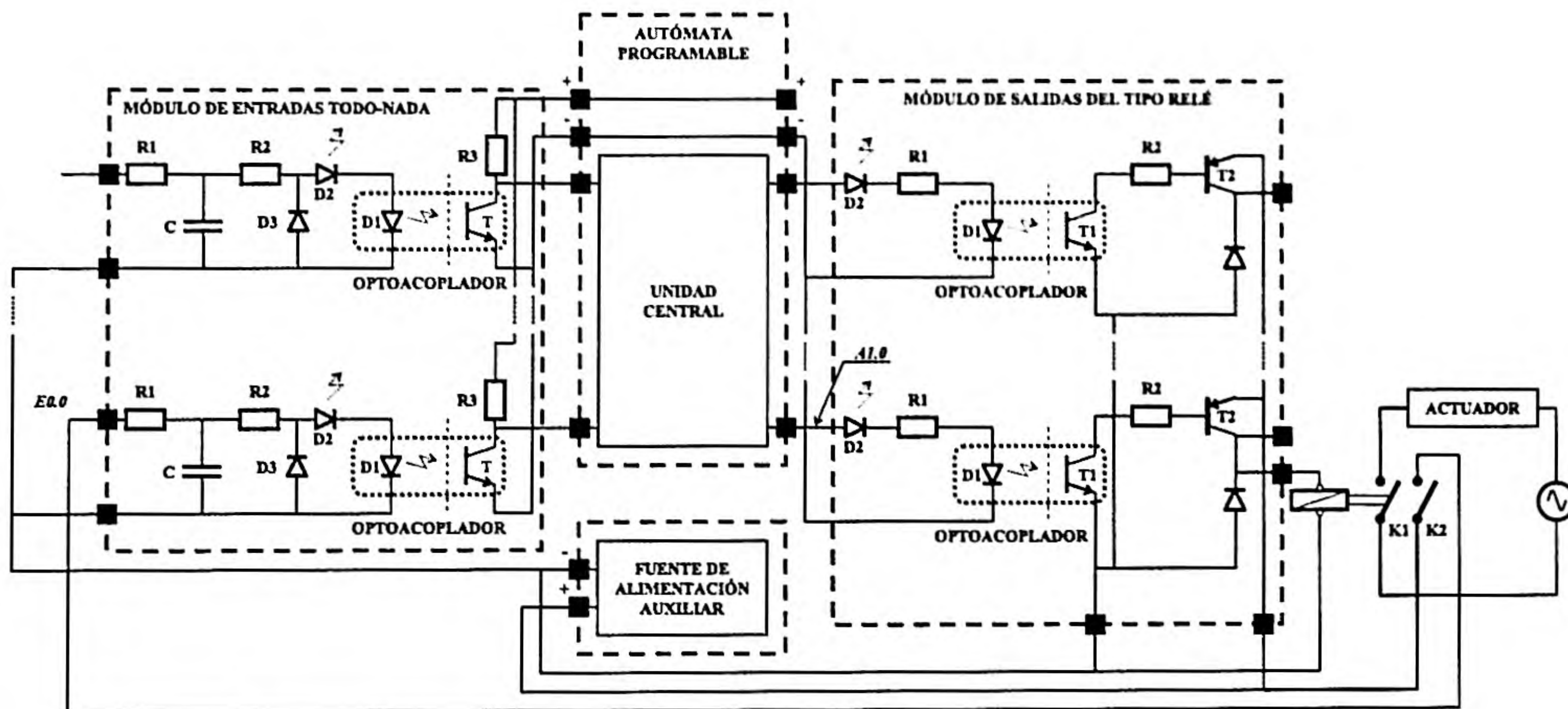


Figura 10.18. Utilización de una variable de entrada todo-nada para elevar la seguridad ante averías de una salida del tipo relé.

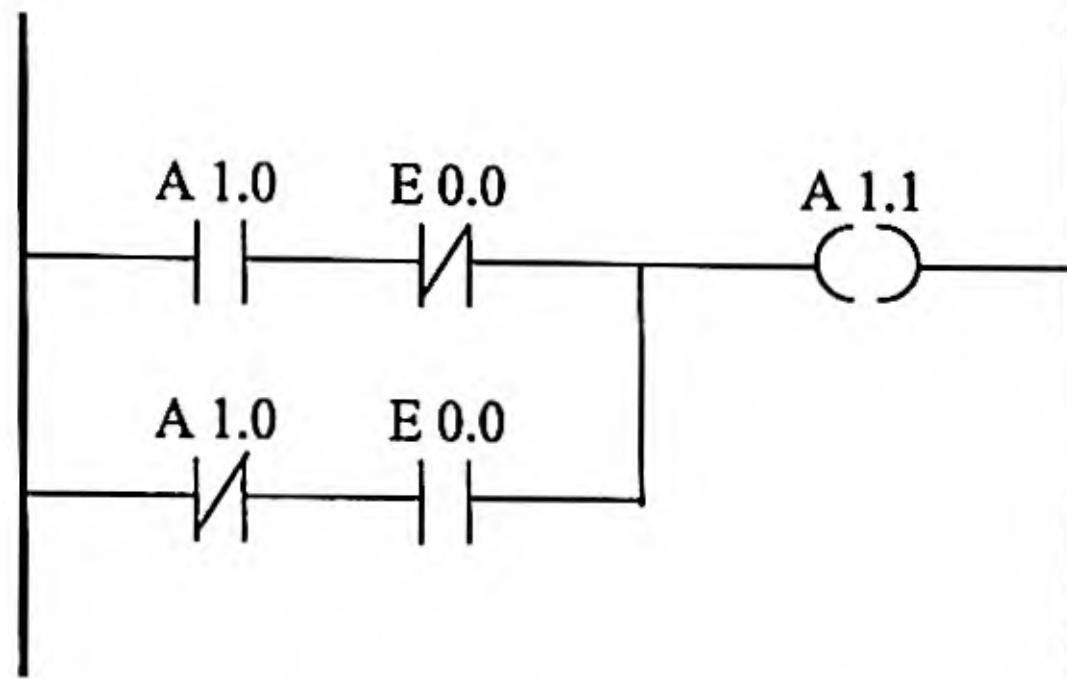


Figura 10.19. Programa, realizado en el lenguaje de esquema de contactos, que comprueba el fallo de una salida todo-nada.

10.3.2.4 Seguridad ante sabotajes (*Security*)

Otro aspecto de la seguridad que tiene gran importancia, no sólo en los autómatas programables de aplicación general sino también en los de seguridad y en los de elevada disponibilidad, que se estudian en sucesivos apartados, es el relativo a la capacidad que tienen para impedir que agentes externos no autorizados puedan producir averías en la instalación controlada por él mediante la modificación del programa de control sin tener autorización para ello.

Por ello los fabricantes de autómatas programables proporcionan al usuario la posibilidad de establecer una clave con varios niveles de acceso, como por ejemplo:

- Acceso libre en el que cualquier operador puede leer y modificar el programa.
- Acceso en lectura en el que se puede leer el programa pero no se puede modificar.
- Bloqueo total en el que no se puede leer ni modificar el programa y sólo es posible modificar parámetros.

10.3.3 Autómatas programables de elevada confiabilidad

10.3.3.1 Conceptos generales

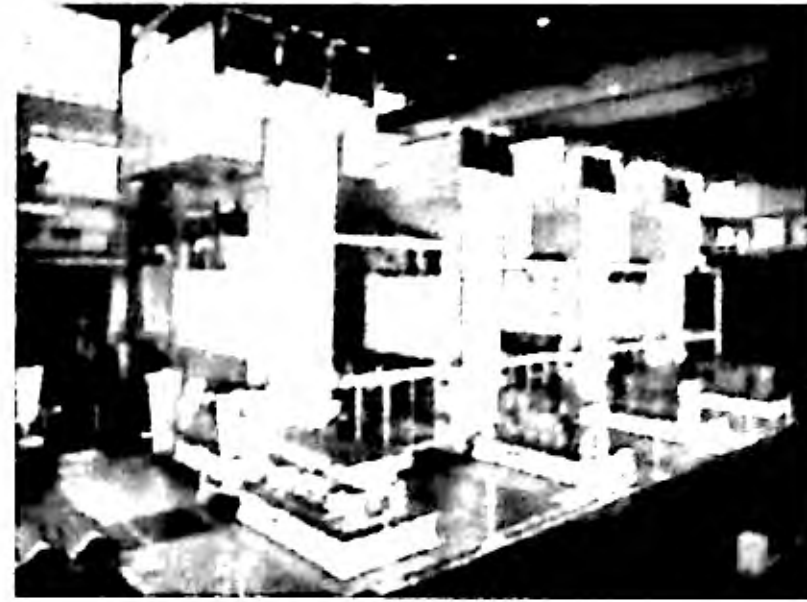
Son numerosos los procesos industriales complejos (Figura 10.20) que exigen la utilización de sistemas electrónicos de control de elevada confiabilidad y entre ellos cabe citar los siguientes:

- Los procesos de generación y distribución de energía eléctrica.
- Las plantas químicas y farmacéuticas.
- Las instalaciones industriales que incorporan calderas, hornos, etc.
- Las instalaciones que poseen grandes máquinas rotativas, como por ejemplo turbinas, compresores, etc.

- Los sistemas de transporte de material y de personas, tanto en instalaciones industriales como en la navegación aérea.
- Las líneas de producción en las que es necesario prever paradas de emergencia, impedir accesos a determinados elementos de las máquinas, etc.
- Las instalaciones petroquímicas y de gas.



a)



b)



c)



d)



e)

Figura 10.20. Ejemplos de instalaciones y procesos industriales que exigen la utilización de sistemas electrónicos de control de elevada confiabilidad.

Pero un análisis detallado de los procesos que se acaban de citar permite clasificarlos en dos clases:

- **Procesos en los que el atributo más importante en relación con su funcionamiento es la seguridad ante averías**

Se entiende por seguridad la probabilidad de que el sistema lleve la instalación a un estado seguro cuando se produce una situación de peligro debido a que una variable del proceso se sale de rango. La norma IEC 61508 denomina a esta característica seguridad integral ante averías (*Safety Integrity*) y representa la probabilidad de que el sistema realice satisfactoriamente las funciones de seguridad ante averías exigidas, bajo las condiciones especificadas para las que fue diseñado y durante un determinado período de tiempo. Al nivel de fiabilidad exigido al sistema, la norma IEC 61508 lo denomina SIL (*Safety Integrity Level*) (Apartado A5.5.2.5.3 del apéndice 5).

- **Procesos en los que el atributo más importante en relación con su funcionamiento es la disponibilidad**

Se entiende por disponibilidad de un sistema la probabilidad de que el equipo esté operativo en un instante determinado, bien porque no se avería o porque si se avería se repara. También se puede entender como la probabilidad de que el sistema no ejecute la función de seguridad ante averías, y de que por lo tanto sólo lleve la instalación a un estado seguro debido a que una variable crítica esté fuera de rango, y que no lo haga en el caso de que se produzca un fallo del propio sistema. Cuanto mayor sea la disponibilidad menor es la probabilidad de que el sistema pare la instalación por un fallo del propio sistema. Esta característica está muy relacionada con la fiabilidad, si bien las normas no establecen especificaciones para la misma.

Mediante la combinación de los recursos de diagnóstico internos y externos descritos en los apartados anteriores se eleva el nivel de seguridad ante averías de los autómatas programables de aplicación general y se obtiene un sistema cuyo diagrama de bloques se representa en la figura 10.21. Este sistema posee una estructura que se denomina 1oolD (*1 out of 1 with diagnostic*) porque posee una sola unidad central y el circuito de diagnóstico se puede considerar dividido en tres partes:

- Un circuito de diagnóstico de la interfaz de entrada que comprueba el funcionamiento de los circuitos de la misma, para lo cual puede, por ejemplo en el caso de los sensores todo-nada, puentear periódicamente el contacto de salida de los mismos de forma similar a la indicada en la figura 10.17.
- Un circuito de diagnóstico de la unidad central de proceso, basado en el funcionamiento de los circuitos de vigilancia descritos en el apartado 10.3.2.
- Un circuito de diagnóstico de la interfaz de salida que utiliza, por ejemplo, contactos auxiliares de las salidas todo-nada para detectar su correcto funcionamiento.

El conjunto formado por los tres circuitos de diagnóstico interconectados actúa sobre una variable de salida todo-nada (Figura 10.21) y hace que se deje de proporcionar tensión a los actuadores adecuados para colocar la instalación en un estado seguro, cuando se detecta una situación anómala.

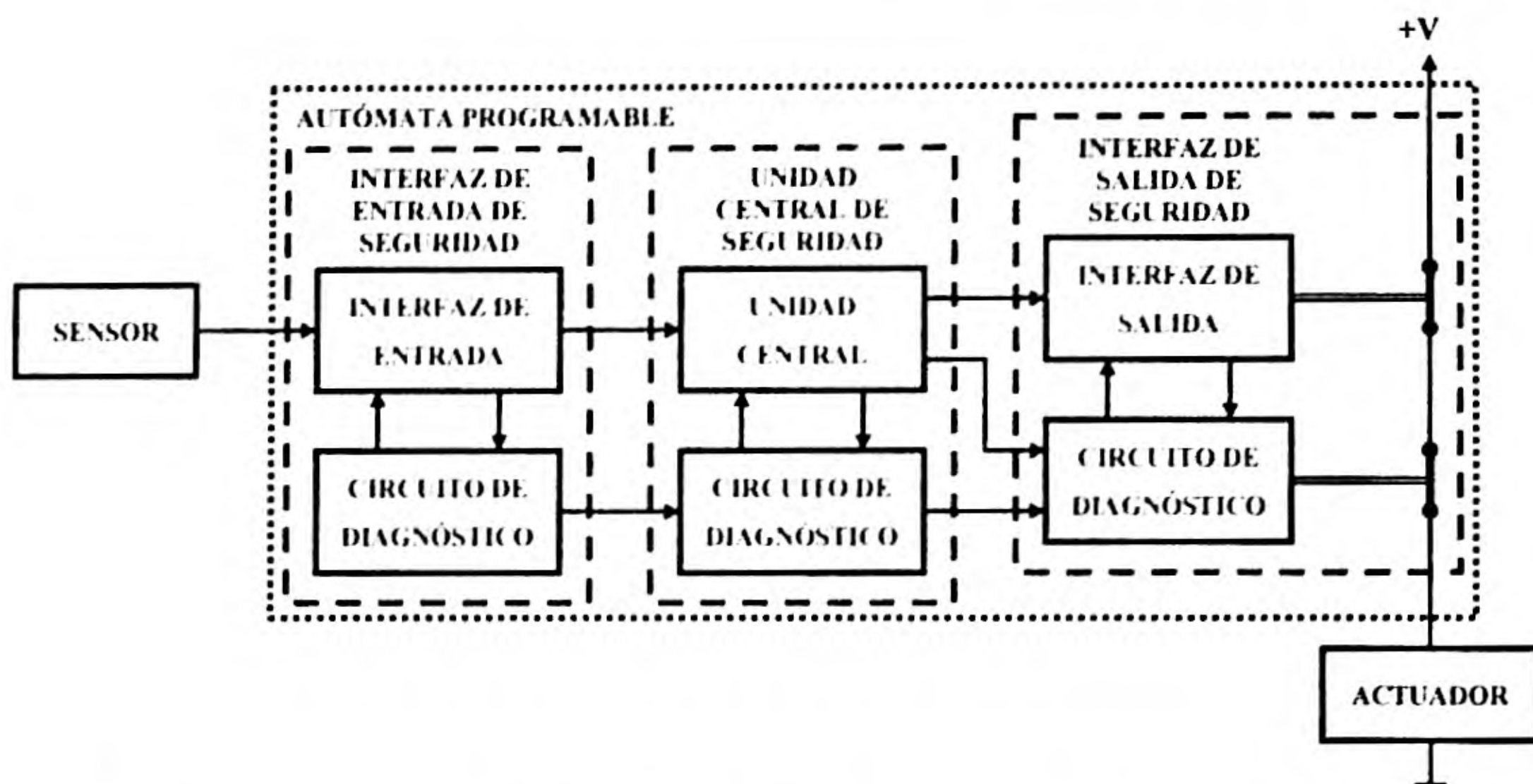


Figura 10.21. *Autómata programable de estructura denominada 1 out of 1 with diagnostic).*

Utilizando los recursos que se acaban de citar se obtiene un sistema electrónico de control con capacidad de autodiagnóstico de forma concurrente con la operación del mismo, pero el progreso de la Microelectrónica permite que se puedan realizar circuitos de autodiagnóstico de gran complejidad que elevan dicha capacidad. Las técnicas de autocomprobación de sistemas digitales complejos se denominan BIST (acrónimo de *Built In Self Test*) y consisten en añadir elementos en el interior del propio circuito que se encargan de la generación aleatoria de vectores de prueba y el análisis de la respuesta del circuito [KOEN 79] [McCl 81]. Estas técnicas se han combinado con los recursos de exploración periférica (*Boundary Scan Techniques*) [MAND 02] y permiten la implementación de sistemas digitales de elevada confiabilidad cuyo estudio se sale de los límites de este libro [GLOS 89] [NAGV 91] [ZORI 94]. No obstante y aunque teóricamente se puede conseguir que el sistema de la figura 10.21 presente un nivel alto de seguridad ante averías, es conveniente indicar que en cuanto falla la única unidad central que posee, el sistema deja de estar operativo.

10.3.3.2 Autómatas programables de seguridad

Los autómatas programables de seguridad son sistemas electrónicos de control seguros ante averías (*Safety or fail-safe systems*) que, tal como se indica en el apartado A5.5.1 del apéndice 5, además de ejecutar la tarea de control de un proceso, disponen de recursos que les permiten detectar sus propios fallos y los del proceso que controlan, así como acciones incorrectas de los usuarios del mismo y, al hacerlo, pasan a un estado que garantiza que no se producen daños en el proceso del que forman parte, al entorno del mismo, o a las personas que los utilizan.

Las principales características de los autómatas programables de seguridad son:

- **Alta fiabilidad del hardware**

Utilizan componentes electrónicos de elevada calidad y se diseñan de acuerdo con procedimientos y criterios de elevada fiabilidad, que proporcionan, entre otras características, una gran inmunidad ante interferencias electromagnéticas. Poseen una estructura redundante y un elevado nivel de autodiagnóstico como los establecidos en la norma IEC61508, que se describen en el apartado A5.5.2.5.3 del apéndice 5. Esto hace que sean tolerantes a un fallo o a un número determinado de fallos.

- **Elevada fiabilidad del software**

Tanto el software del propio sistema como el desarrollado por el usuario son fiables y seguros. Para ello, el software del sistema está dotado de una elevada capacidad de diagnóstico que le permite detectar el mayor número posible de fallos propios y proporcionar información sobre ellos, así como tomar las medidas oportunas. Por otra parte, el software del usuario debe ser realizado con una determinada metodología y modularidad que evite los posibles errores y facilite su documentación.

- **Elevada seguridad ante averías**

Detectan sus propios fallos y en esa circunstancia hacen que el proceso pase a un estado seguro. Aunque esta característica resulta fundamental en los tipos de aplicaciones indicadas en el apartado anterior, en algunas de ellas la parada del proceso debida a un fallo del propio sistema electrónico de control (no debida a que algún parámetro del proceso se salga del rango de seguridad ante averías) provoca pérdidas de producción que son muy costosas. Por ello la disponibilidad de estos equipos también resulta ser de vital importancia y a los autómatas programables de elevada disponibilidad se dedica el apartado 10.3.3.3 a continuación.

Existen múltiples formas de implementar un autómata programable de seguridad, debido a que, tal como se indica en el apartado A5.4.2 del apéndice 5, la redundancia necesaria para lograr la función de seguridad ante averías se puede implementar de diversas maneras.

En la figura 10.22 se representa el diagrama de bloques de una de las estructuras utilizadas para implementar autómatas programables de seguridad, que está constituida por dos subsistemas como el de la figura 10.21, cuyas variables de salida todo-nada se conectan en serie. Este sistema es, desde el punto de vista de la seguridad ante averías, del tipo 1oo2D, descrito en el apartado A5.5.2.5.3 del apéndice 5, porque si falla uno de las dos subsistemas que contiene, pone la instalación en un estado seguro.

Un ejemplo de este tipo de autómatas programables son los autómatas programables S7-300F y S7-400F de Siemens, en cuya denominación la letra F indica que son seguros ante averías propias (*Fail-safe*). Sus principales características son:

- Utilizan módulos de entrada y salida de seguridad que se caracterizan por poseer circuitos de autodiagnóstico y por ello se denominan periferia de seguridad.
- Utilizan un programa de seguridad ante averías (Figura 10.23) que se caracteriza por:
 - Procesa las variables críticas con su valor directo e inverso.
 - Compara los resultados obtenidos y, en caso de discrepancia, pone la instalación en el estado seguro (*Stop*).

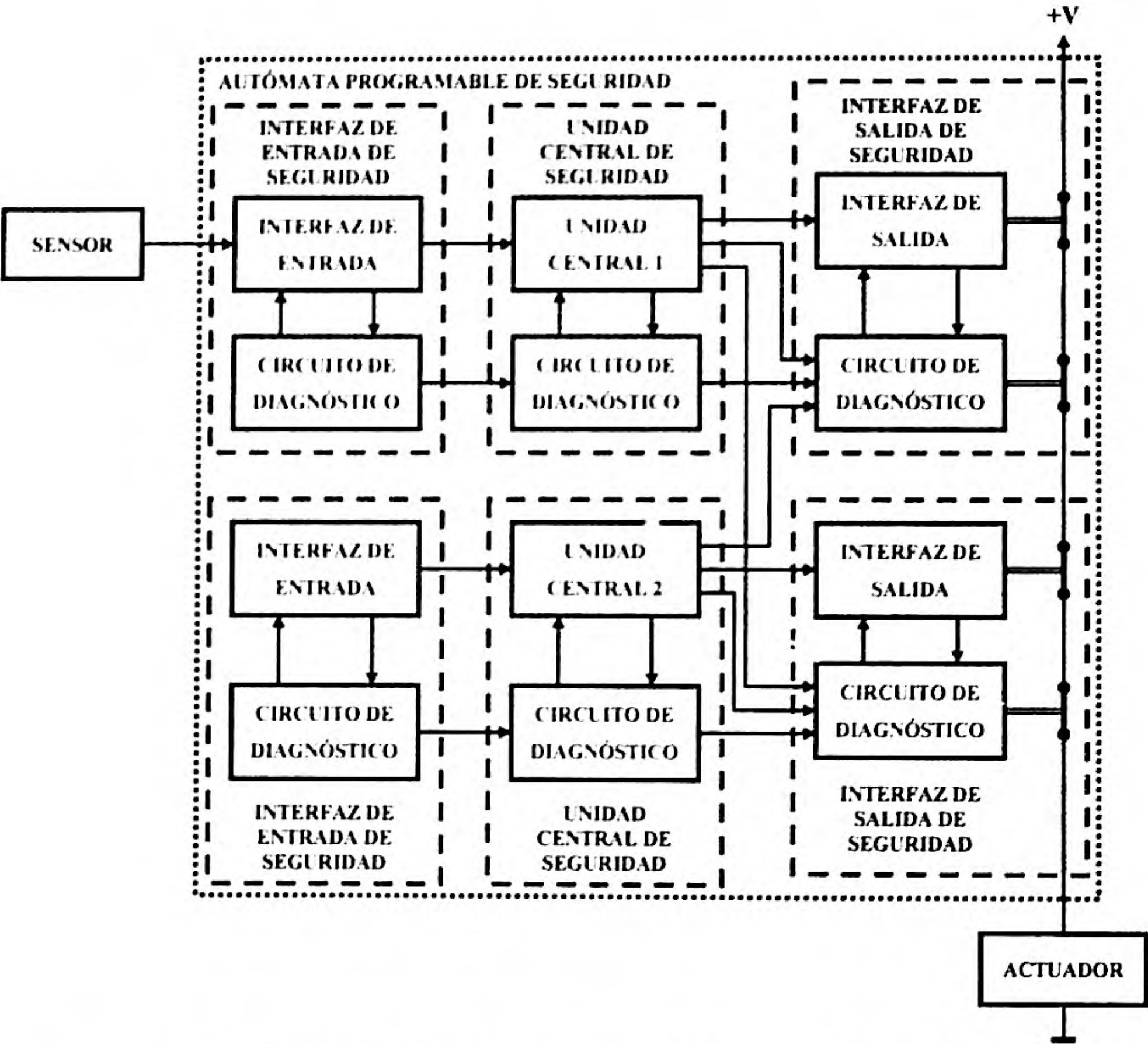


Figura 10.22. Diagrama de bloques de una estructura 1oo2D, utilizada para implementar autómatas programables de seguridad ante averías.

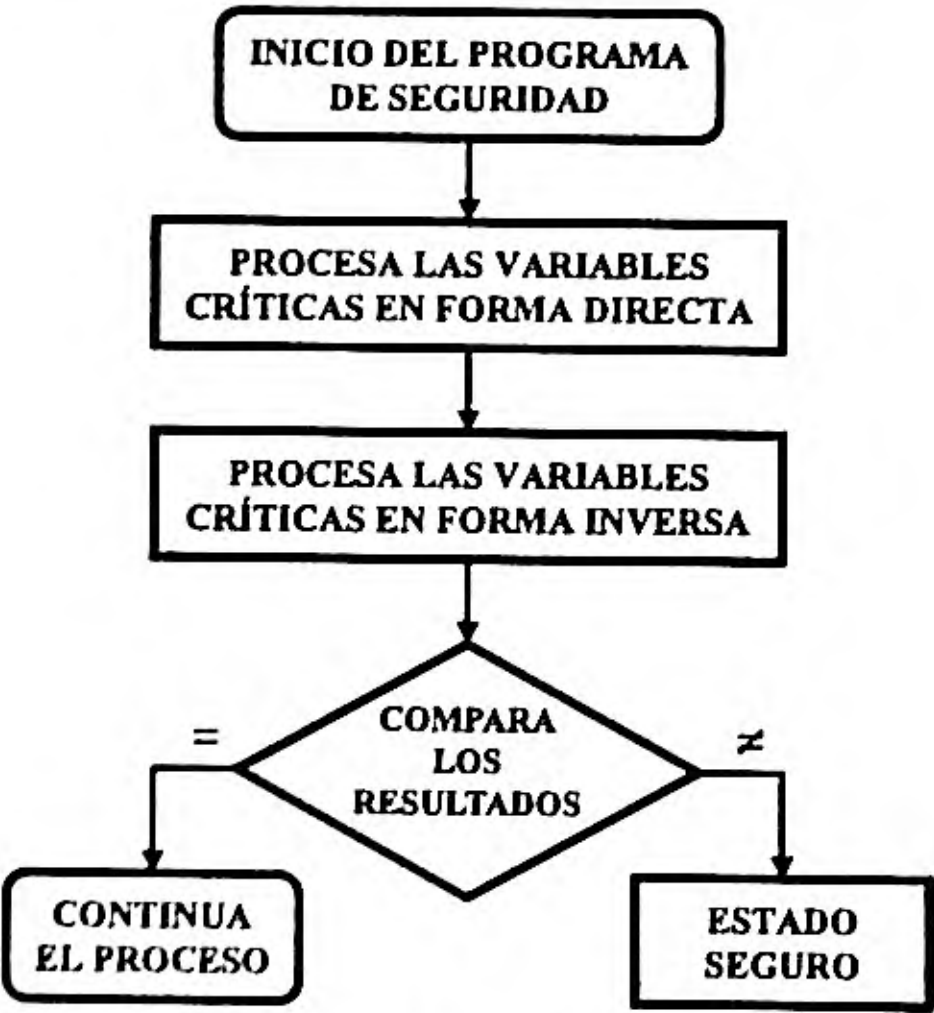


Figura 10.23. Algoritmo del programa de seguridad ante averías de un autómata programable de seguridad.

- Poseen un procesador de comunicaciones que permite la transferencia de información entre el autómata programable y módulos de entrada y salida remotos (periferia descentralizada) a través de un bus de campo.
- Pueden ejecutar las tareas de control y las de seguridad ante averías. Para ello se les pueden conectar módulos de entrada/salida remota de aplicación general y de seguridad, tal como se indica en la figura 10.24 en la que se representa un autómata programable S7-400F de Siemens, conectado, a través de un bus de campo PROFIBUS-DP (descrito en el apéndice 4), a un módulo de entrada/salida de aplicación general y otro de seguridad. En la figura 10.25 se muestra la fotografía de un autómata de seguridad S7-300F que posee módulos de entrada/salida local y de entrada/salida remota, tanto de aplicación general como de seguridad ante averías.

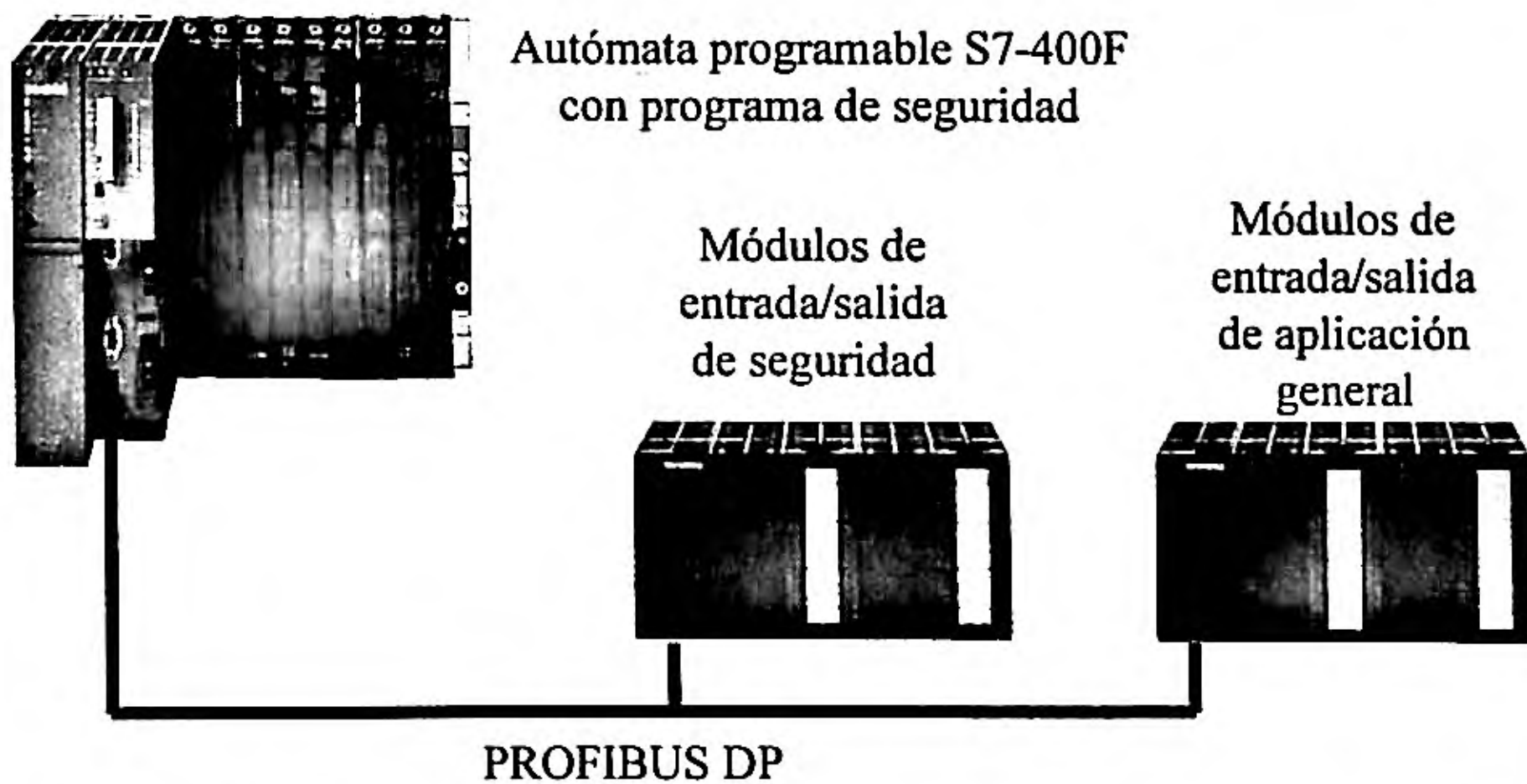


Figura 10.24 Autómata programable S7-400F de Siemens conectado, a través de un bus de campo PROFIBUS-DP, a un módulo de entrada/salida de aplicación general y otro de seguridad ante averías.

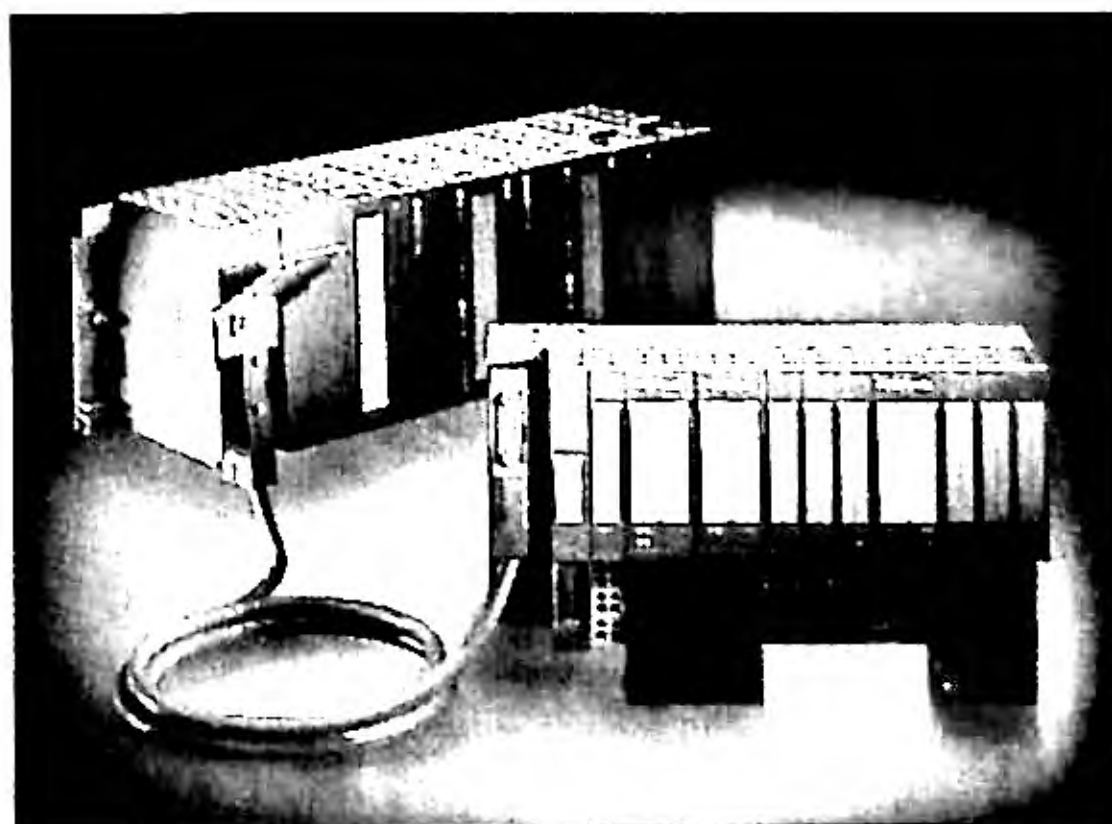


Figura 10.25. Fotografía de un autómata de seguridad ante averías S7-300F que posee módulos de entrada/salida local y de entrada/salida remota.

- Su diagrama de bloques es el representado en la figura 10.22 y se comportan como un sistema electrónico de estructura 1oo2D, es decir, la instalación se para y se pone en un estado seguro en cuanto uno de ellos toma la decisión de hacerlo.
- Utilizan tanto sensores como actuadores que tienen también individualmente el nivel de seguridad ante averías adecuado para que el conjunto alcance el nivel de seguridad ante averías exigido por la aplicación. En muchas aplicaciones se duplican los sensores de las variables críticas. En este caso si uno de los sensores se avería el sistema toma la decisión de ejecutar la función de seguridad ante averías y pone la instalación en un estado seguro. Si se pretende que el sistema sea tolerante al fallo de uno de los sensores y que por lo tanto tenga elevada disponibilidad ante ese tipo de fallo, es necesario instalar tres sensores y realizar un sistema 2oo3. Este tipo de aplicaciones no son habituales por su elevado coste.

Para elevar la confianza de los usuarios en los autómatas programables de seguridad, los fabricantes validan sus equipos a través de determinados organismos de reconocida solvencia (TÜV en Europa y FMRC en USA). Dichos organismos se encargan de comprobar los métodos de diseño utilizados y de certificar el nivel de seguridad ante averías del sistema.

10.3.3.3 Autómatas programables de elevada disponibilidad

Se denominan sistemas electrónicos de elevada disponibilidad (*High availability*) los que están adecuadamente implementados para proporcionar una elevada probabilidad de que presten servicio en cualquier instante. Su utilización está motivada por el hecho de que en muchas ocasiones es muy elevado el coste que supone la puesta fuera de servicio de un proceso productivo debido a una avería del sistema electrónico que lo controla. De ello se deduce que en estos autómatas programables debe ser muy baja la probabilidad de que dejen de funcionar, y por lo tanto paren la instalación, debido a una avería interna.

El gran incremento de la utilización de los autómatas programables en el control de las instalaciones industriales ha hecho que en muchas de las aplicaciones la disponibilidad sea un parámetro añadido a la seguridad ante averías, porque la falta de disponibilidad del sistema electrónico de control no sólo pone el proceso controlado fuera de servicio sino que genera una situación peligrosa que es necesario evitar. Por ello los fabricantes de autómatas programables suelen ofrecer a los diseñadores de sistemas de control industrial configuraciones que garantizan una elevada disponibilidad mediante la utilización de la redundancia masiva, que se estudia en el apartado A5.4.2.2 del apéndice 5.

Para aclarar la diferencia entre los sistemas electrónicos de seguridad y los de elevada disponibilidad se puede analizar un sistema redundante que está formado por cuatro sistemas electrónicos o canales idénticos y un sistema de diagnóstico de elevada fiabilidad [estructura 2oo4D (dos de cuatro con diagnóstico) de la norma IEC 61508 descrita en el apartado A5.5.2.5.3 del apéndice 5]. Al analizar este sistema se concluye que:

- Desde el punto de vista de la seguridad ante averías, ejecuta la función de seguridad y para la máquina o pone la instalación en un estado seguro en cuanto dos de los cuatro sistemas que lo forman detectan que la variable crítica está fuera de rango. Por tanto, incluso con dos de los cuatro sistemas averiados se puede ejecutar la función de seguridad y el sistema en cuestión es tolerante a dos fallos.

- Desde el punto de vista de la disponibilidad, este sistema es tolerante a un solo fallo porque, en cuanto dos de los sistemas fallan, pueden tomar la decisión de parar la máquina o instalación, cuando realmente no existe ninguna variable crítica fuera de rango y lo único que sucede es que dos de los cuatro sistemas existentes tienen al menos un fallo.

De lo expuesto se pueden sacar las siguientes conclusiones:

- Un autómata programable de elevada disponibilidad es un sistema que no puede dejar de funcionar si se produce un solo fallo interno. Por ello, aunque la estructura 1001D representada en la figura 10.21 puede teóricamente servir para implementar un autómata programable seguro ante averías propias, no puede ser utilizada para implementar un autómata programable de elevada disponibilidad porque en cuanto uno de los elementos que lo forman tiene una avería, el autómata programable deja de realizar la tarea que tiene asignada y pasa a un estado seguro.
- Para que la estructura 1002D constituya un autómata programable de alta disponibilidad se deben conectar en paralelo los contactos de salidas correspondientes a cada uno de los dos subsistemas. De esta forma el fallo de uno de ellos no pone al sistema fuera de servicio sino que sigue funcionando y avisa de la anomalía al usuario (Figura 10.26).

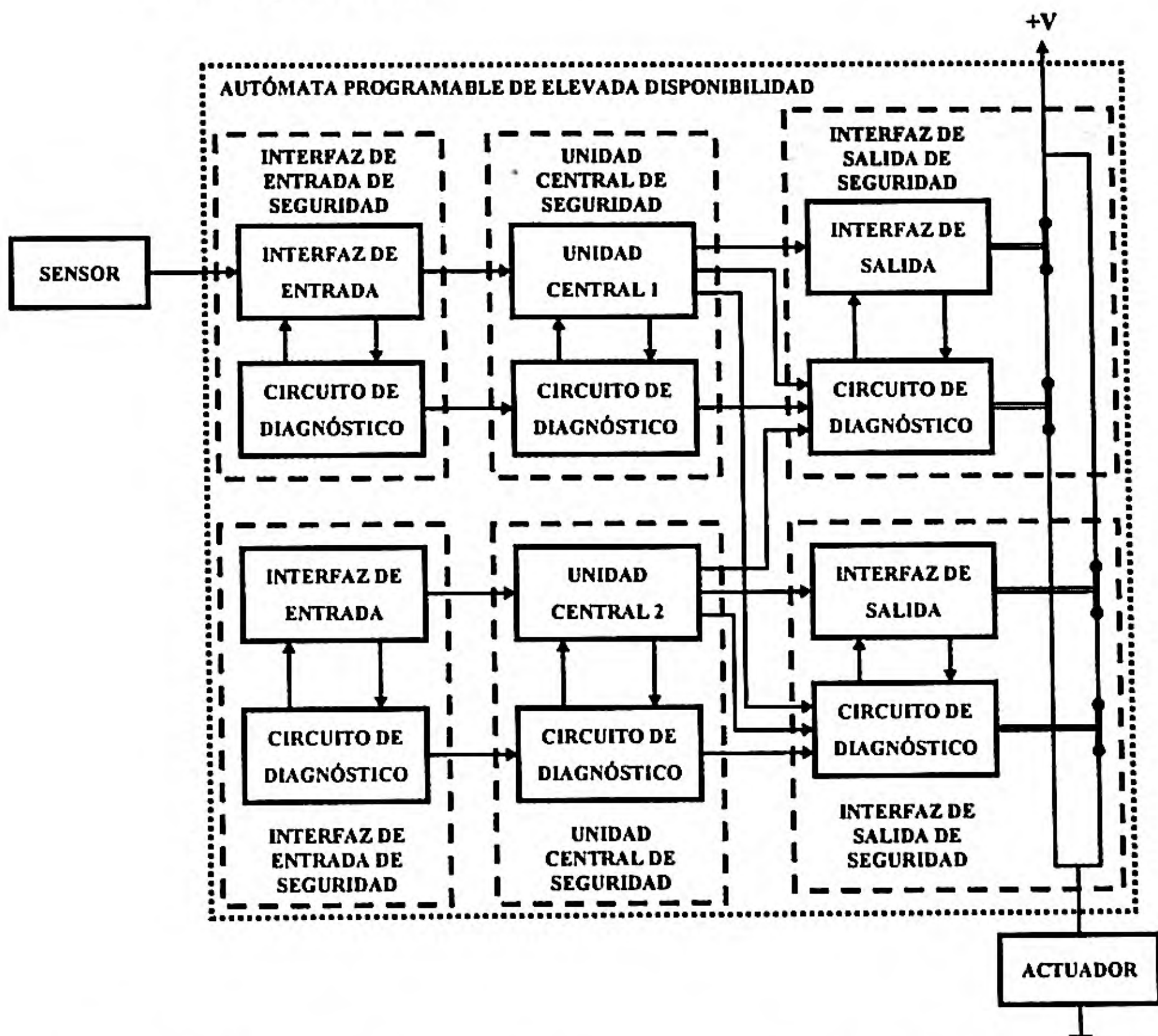


Figura 10.26. Diagrama de bloques de una estructura 1002D, utilizada para implementar un autómata programable de elevada disponibilidad.

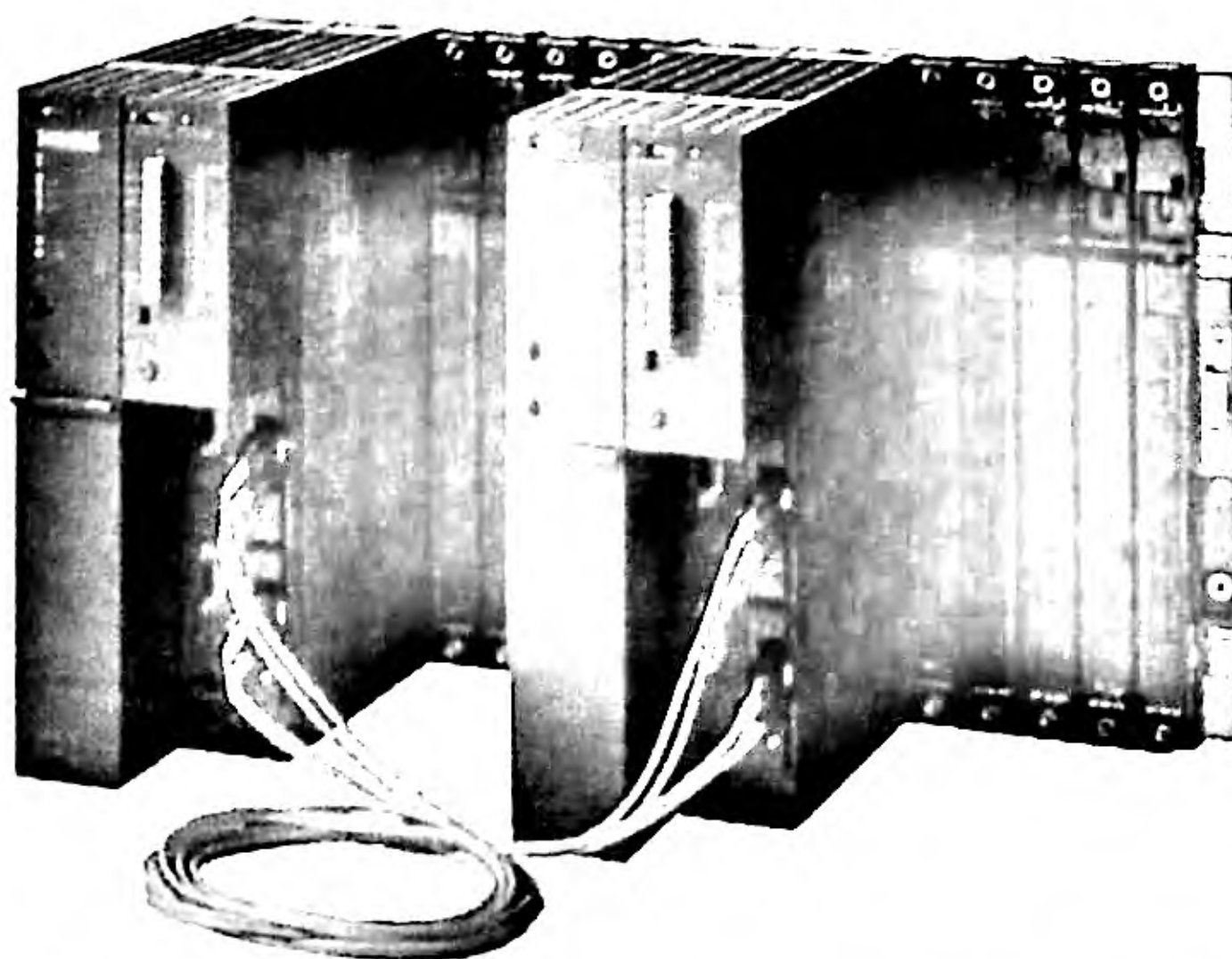


Figura 10.27. Fotografía de un autómata programable de elevada disponibilidad S7-400H de Siemens.

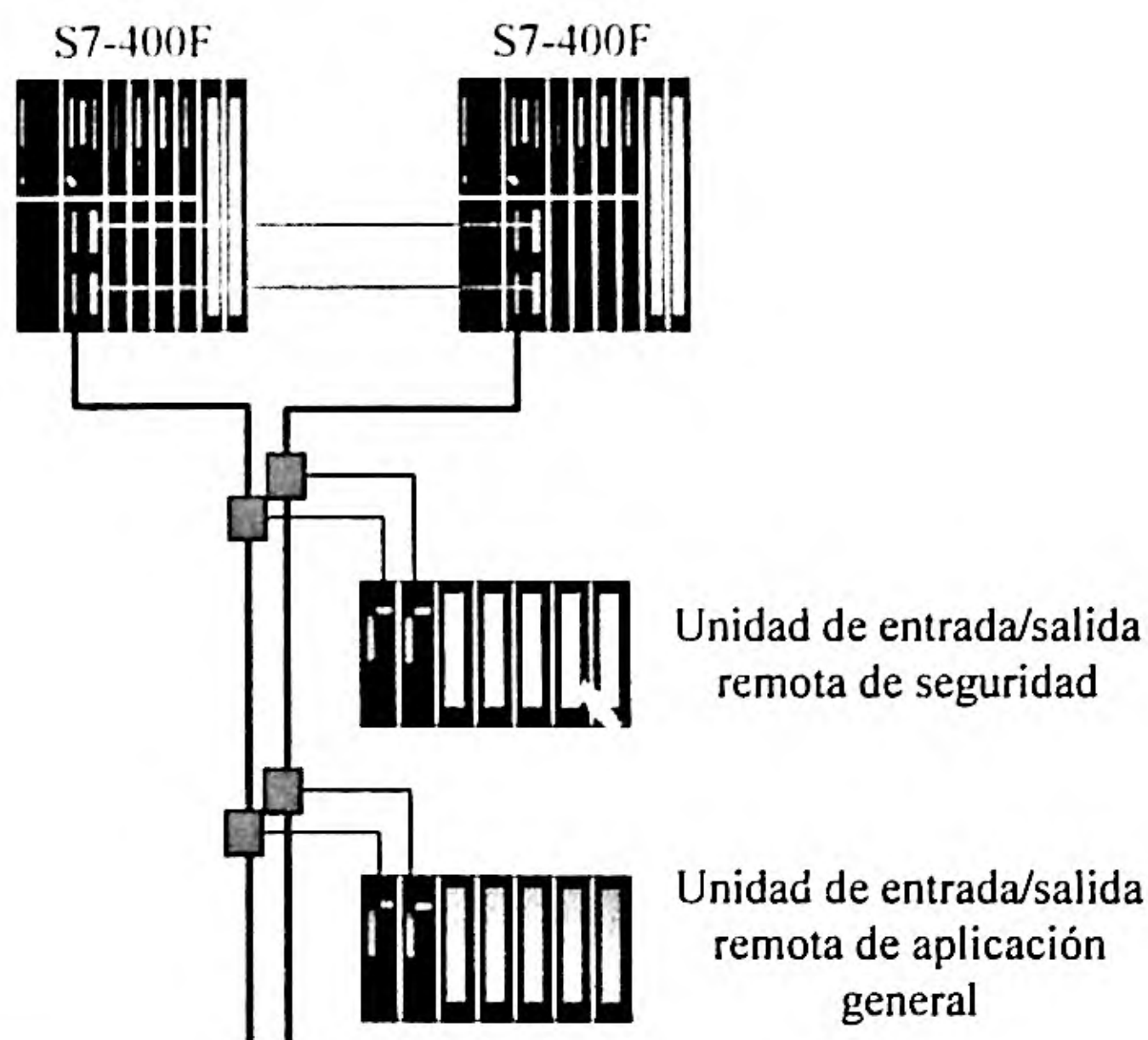


Figura 10.28. Autómata programable S7-400FH implementado con dos autómatas programables S7-400F que comparten, a través de dos buses de campo independientes, una unidad de entrada/salida remota de aplicación general y otra de seguridad ante averías.

En la figura 10.27 se muestra la fotografía de un autómata programable de elevada disponibilidad S7-400H (la H indica *High availability*) realizado con dos unidades centrales

S7-400 interconectadas en una configuración principal-subordinado (*Master-slave*) en la que el principal ejerce el control en situación normal y el subordinado pasa a ejercerlo cuando falla el principal.

La combinación de dos autómatas programables de seguridad que tengan el diagrama de bloques de la figura 10.22 da lugar a un sistema que es al mismo tiempo de seguridad y de elevada disponibilidad. Un ejemplo de sistema electrónico de control de estas características es el S7-400FH de Siemens que está realizado con dos autómatas programables S7-400F que ejecutan un programa de seguridad ante averías, además del programa de control, y comparten a través de dos buses de campo independientes una unidad de entrada/salida remota de aplicación general y otra de seguridad ante averías (Figura 10.28).

Bibliografía

- [ANSI 96] American National Standard. ANSI/ISA-84.01-1996. *Application of Safety Instrumented Systems for the Process Industries*. 1996.
- [BS 99] British Standard. *Railway Applications – The Specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS)*. BS EN 50126. 1999.
- [CEDE 02] Cedes. *Safety & Automation*. CEDES AG, Switzerland, 2002.
- [GIL 92] Pedro J. Gil. *Sistema tolerante a fallos con procesador de guardia: Validación mediante inyección física de fallos*. Tesis doctoral. Universidad Politécnica de Valencia. 1992.
- [GLOS 89] C.S. Gloster, F. Brglez. *Boundary scan with Built in self test*. IEEE Design & Test of Computers. February 1989.
- [KOEN 79] B. Koenemann, J. Mucha, G. Zwiehoff. *Built-in logic block observation techniques*. Dig. Papers, 1979 Test Conference. IEEE Pub. 79CH1509-9C, pp. 37-41, Oct. 1979.
- [LU 82] D.J. Lu. *Watchdog processors and structural integrity checking*. IEEE Transactions on Computer, vol. 31, nº 7, pp. 681-685. Julio 1982.
- [MADE 91] H. Madeira, J. Camoes y J.G. Silva. *A watchdog processor for concurrent error detection in multiple processor system*. Microprocessors and Microsystems, vol. 15, nº 3, pp. 123-130. Abril 1991.
- [MAHM 85a] A. Mahmood y E.J. McCluskey. *Watchdog processor: Error coverage and overhead*. Proceedings 15th IEEE International Symposium on Fault Tolerant Computing, FTCS-15, pp. 214-219. 1985.
- [MAHM 85b] A. Mahmood, A. Ersoz y E.J. McCluskey. *Concurrent system-level error detection using a watchdog processor*. Proceedings 15th International Test Conference, pp. 145-152. 1985.
- [MAHM 88] A. Mahmood y E.J. McCluskey. *Concurrent error detection using a watchdog processor. A survey*. IEEE Transaction on Computers, Vol. 37, nº 2, pp. 160-174. Febrero 1988.

- [MAND 02] E. Mandado, L. J. Alvarez y M. D. Valdés. *Dispositivos lógicos programables y sus aplicaciones*. Editorial Thomson Learning Paraninfo. 2002.
- [MAND 07] E. Mandado, L. M. Menéndez, L. F. Ferreira y E. L. Matos. *Microcontroladores PIC. Sistema integrado para el autoaprendizaje*. Editorial Marcombo. 2007.
- [MAND 08] E. Mandado y Y. Mandado. *Sistemas electrónicos digitales. Apartado 4.2.4.2 del capítulo 4*. 9ª edición. Editorial Marcombo. 2008.
- [MART 93] J. Martínez y M. Barrón. *Prácticas con microcontroladores de 8 bits. Aplicaciones industriales*. Editorial Mc Graw Hill. 1993.
- [MART 03] E. Martín, J. M. Angulo & I. Angulo. *Microcontroladores PIC*. Editorial Thomson Learning Paraninfo. 2003.
- [McCl 81] E.J. McCluskey, S. Bozorgui-Nesbat. *Design for autonomous test*. IEEE Trans. on Computers, vol.C-30. Pp.866-875, Nov.1981.
- [NAGV 91] P.Nagvajara, M.G. Karpovsky, L.B. Levitin. *Pseudorandom testing for Boundary-Scan design with Built-In self Test*. IEEE Design & Test of Computers. September 1991.
- [SCHW 84] A. Schweitzer, J. P. Girardin. *Méthodes permettant d'améliorer le niveau de sécurité des systèmes à logique programmée. Logique programmée et sécurité: Enoncé du problème*. Electronique, Techniques et Industries. (París). N° 12, Novembre 1984.
- [VAUT 83a] J. P.Vautrin, M.Collier. *Les fonctions dynamisme autocontrôle et sécurité des systèmes logiques*. Revue Electronique Industrielle, N° 60, Novembre 1983.
- [VAUT 83b] J.P.Vautrin, M.Collier. *Application of dynamic and self-checking functions in the improvement of the safety of logic systems*. Electronique Industrielle N° 60, Nov. 1983.
- [ZORI 94] Y. Zorian. *A structured testability approach for multi-chip modules based on BIST and Boundary-Scan*. IEEE Transactions on Components, Packaging, and Manufacturing Technology. August 1994.

APÉNDICE 1

Comunicaciones digitales

A1.1 Introducción

Las Comunicaciones Digitales (*Digital or Data Communications*), conocidas actualmente como Telemática, (combinación abreviada de las palabras “Telecomunicación” e “Informática”), nacieron como resultado del progreso de la Informática que, a comienzos de la década de 1960, necesitó, para ampliar su campo de aplicación, dotar a los computadores de terminales remotos y enviarles o recibir la información procedente de ellos a través del mínimo número de hilos posible. Para ello fue necesario convertir del formato paralelo al formato serie (que se describen en el apartado 9.2.3.2 del capítulo 9), tanto la información generada en paralelo por la unidad central de proceso del computador como por los terminales. Nació así el concepto de transmisión o comunicación de datos, utilizado tradicionalmente para referirse a la comunicación de información digital en el formato serie a una cierta distancia superior a decenas de centímetros y se inició el Teleproceso o proceso de datos a distancia.

Para llevar a cabo la transmisión de datos es necesario que al menos una de las unidades de interfaz del computador, cuyo diagrama de bloques se representa en la figura 1.62 del capítulo 1, lleve a cabo la tarea antes citada y por ello dicha unidad recibe en la actualidad el nombre genérico de procesador de comunicaciones (*Communication processor*) también denominado por algunos autores controlador de comunicaciones (*Communication controller*) [HELD 03]. Se tiene así el esquema básico de bloques de un computador (Figura A1.1) que posee capacidad de comunicación a distancia con un periférico.

Pero el Teleproceso no permite distribuir entre varios computadores un proceso complejo y llevar a cabo la toma de datos en puntos separados espacialmente. Fue el avance de la Microelectrónica, que elevó vertiginosamente la complejidad de los circuitos integrados, el que dio lugar a las redes de computadores que permiten la ejecución de un proceso de información complejo entre varios computadores situados en lugares distantes entre sí. Para ello se tuvo que elevar la complejidad de los procesadores de comunicaciones y de esa forma nació el Proceso Distribuido [PALM 00][SLOM 87] y surgió el concepto de Sistema Informático definido, por la Organización Internacional de Normalización conocida como ISO (acrónimo de *International Standard Organisation*), de la forma siguiente:

Uno o más computadores, el conjunto asociado de programas (Software), los terminales, los operadores humanos, los procesos físicos y los medios de transmisión que constituyen un todo autónomo capaz de realizar un tratamiento de la información.

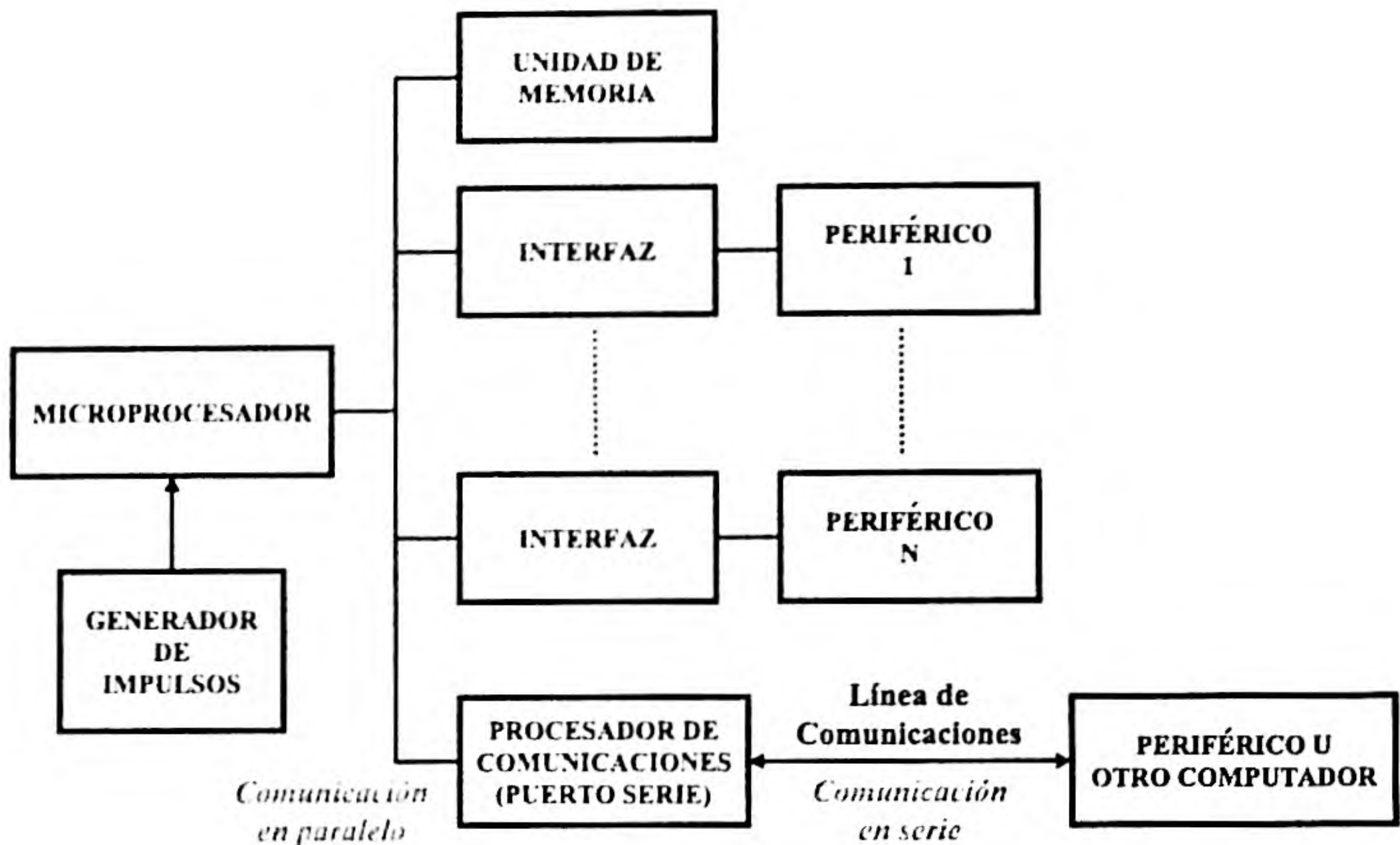


Figura A1.1. Computador con capacidad de comunicación a distancia.

Por otra parte, la gran cantidad de conceptos asociados con las comunicaciones digitales hace que no resulte fácil su análisis y que sea necesario ir de lo particular a lo general, y analizar primero cada concepto por separado. Por ello es conveniente estudiar primero la comunicación entre un computador y un periférico que recibe la denominación de comunicación punto a punto (*Point to point*) y permite comprender los conceptos básicos relativos a las Comunicaciones Digitales. A partir de ellos se procede al estudio de las redes de comunicaciones digitales, también denominadas redes de datos o redes informáticas (*Data Network Communications*) [FORO 07a] [STAL 07] que constituyen conexiones multipunto (*Multipoint or Multi-drop*).

En este apéndice se describen los conceptos más importantes ligados a las comunicaciones digitales y, en especial, los que, en opinión de los autores de este libro, debe conocer cualquier técnico que tenga que utilizar las Comunicaciones Industriales en general y las asociadas con los autómatas programables en particular, que se describen en el capítulo 9. Para una descripción más amplia se remite al lector a la bibliografía [BLAK 02] [ELAH 01] [FORO 07a] [HALL 05] [MILL 00] [STAL 07] [WALD 91].

A1.2 Conceptos básicos de las comunicaciones digitales punto a punto

La comunicación punto a punto (*Point to point*) se puede definir como la transmisión de información en serie entre un computador y un periférico u otro computador, que no está

compartida con ningún otro elemento de proceso de datos. En la figura A1.2. se representa su esquema básico en banda base digital (descrita en el apartado A1.2.1 a continuación) en el que cada computador está asociado a una interfaz serie que se denomina procesador de comunicaciones, que algunos fabricantes de autómatas programables, como por ejemplo Siemens, denominan CP (acrónimo de *Communication Processor*) porque realiza las acciones adecuadas para llevar a cabo la comunicación. Tal como se indica en ella, además de comunicarse entre sí los procesadores de comunicaciones, lo hacen también entre sí los usuarios de ambos computadores.

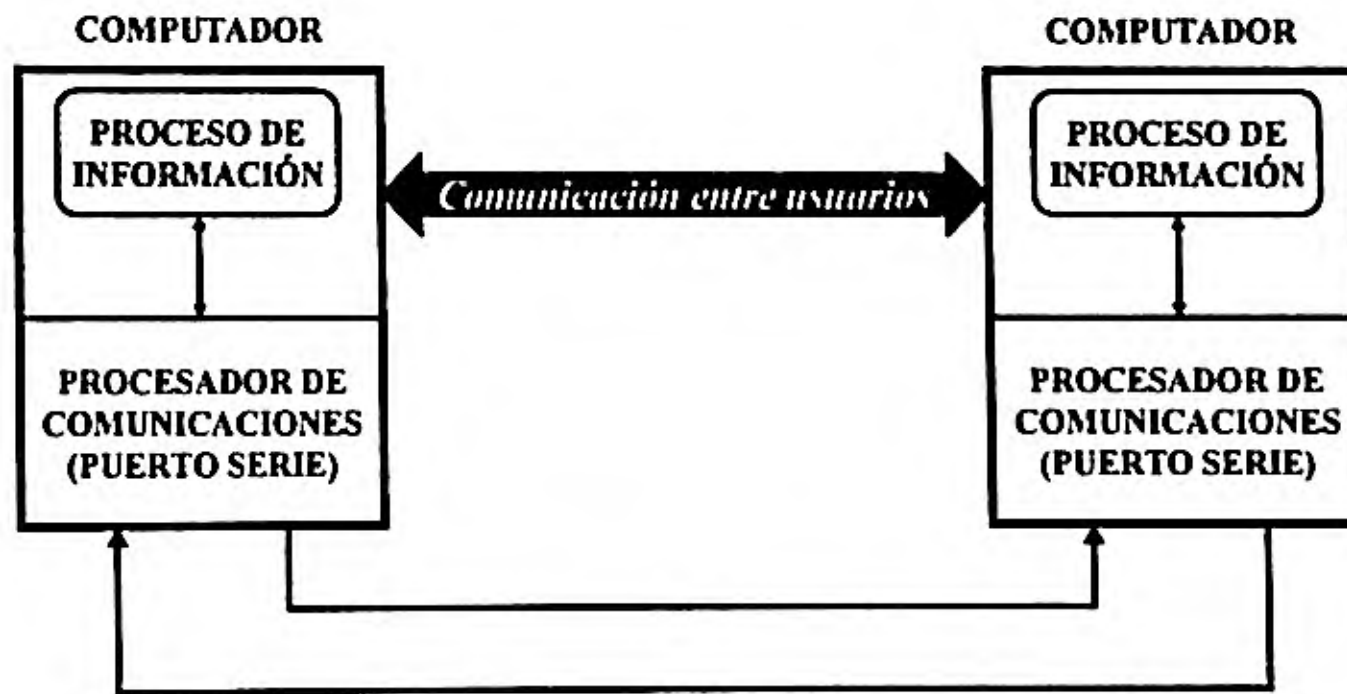


Figura A1.2. Esquema básico de la comunicación punto a punto en banda base digital entre dos computadores.

La transmisión punto a punto está asociada al siguiente conjunto de conceptos básicos interrelacionados, que son además comunes a las comunicaciones multipunto:

- Los modos o métodos de transmisión de las señales (*Signaling modes*).
- Los modos de comunicación (*Data Flow modes*).
- Los modos de sincronización (*Synchronization modes*).
- Los métodos de detección de errores (*Error detection methods*).
- Los medios de transmisión (*Transmission media*).
- Las características mecánicas y eléctricas (*Electrical and mechanical characteristics*).

A continuación se analiza cada uno de ellos.

A1.2.1 Modos o métodos de transmisión de las señales

Las señales digitales se pueden transmitir de dos modos:

- En banda base (*Baseband*).
- En banda ancha (*Broadband*).

A continuación se describe cada una de ellas.

Transmisión en banda base

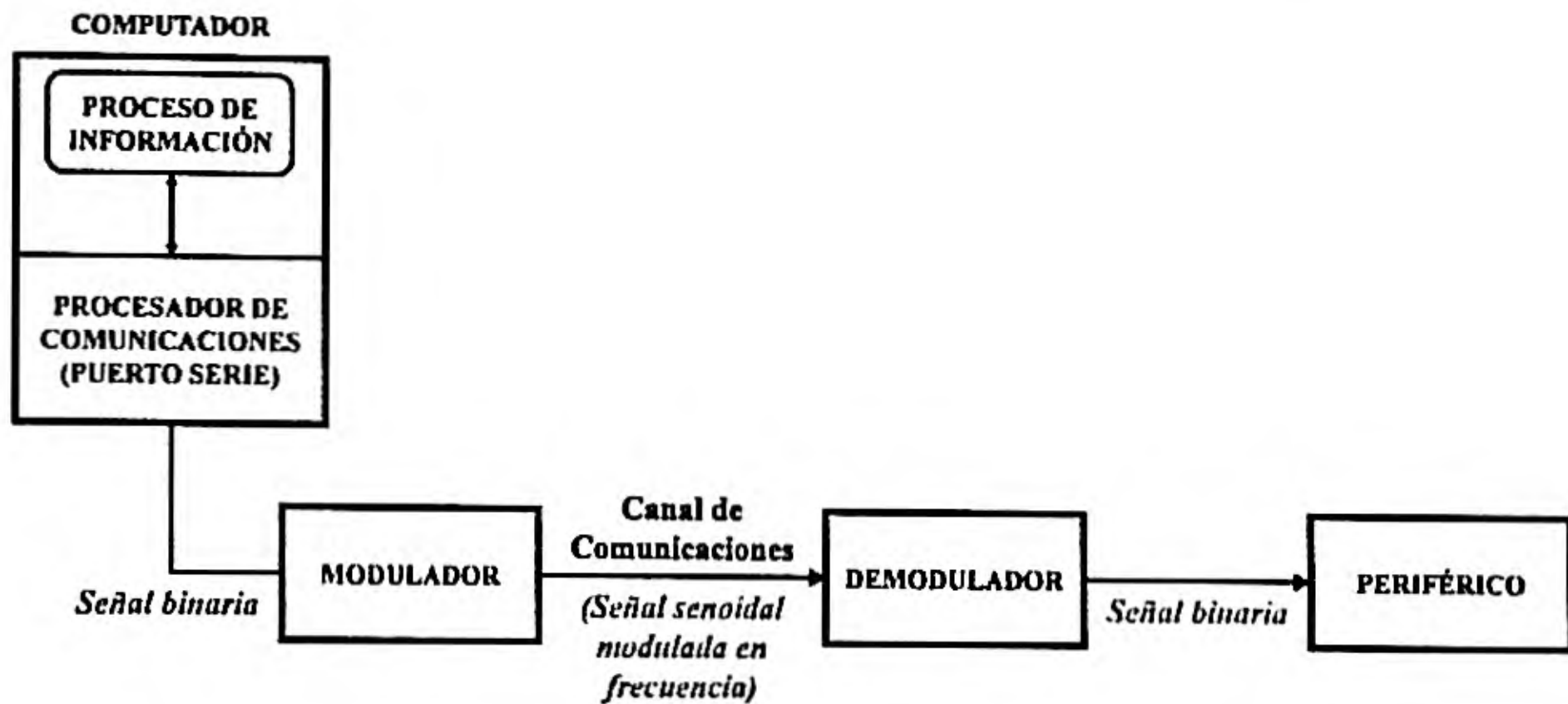
Se suele denominar transmisión en banda base (*Baseband transmission*) al modo de transmisión en el que a través del medio de comunicación utilizado sólo se transmite una señal de información en cada instante en uno o en ambos sentidos. La transmisión de información digital en banda base se puede realizar de dos formas:

- Transmisión directa en formato digital, que según los autores, se conoce por transmisión en banda base o en banda base digital. Los sistemas cuyos diagramas de bloques se representan en la figuras A1.1 y A1.2 constituyen un ejemplo de comunicación digital en banda base digital.
- Transmisión mediante la modificación de alguno de los parámetros (modulación) de una señal senoidal denominada portadora. Este modo de transmisión se suele denominar banda base analógica [BLAK 02]. A partir de la figura A1.1 se obtiene un sistema de comunicación en banda base analógica colocando un circuito modulador (*Modulator*) entre el procesador de comunicaciones y la línea de comunicaciones y un circuito demodulador (*Demodulator*), que elimina la portadora, entre ésta y el periférico tal como se indica en la figura A1.3a. Cuando se quiere transmitir la información en ambos sentidos entre el computador y el periférico se utiliza en cada extremo de la línea de comunicaciones un circuito denominado MODEM (*Modulator/Demodulator*) que combina un modulador y un demodulador (Figura A1.3b). Son varios los parámetros de una señal moduladora que se pueden modificar mediante una señal digital. Como ejemplo en la figura A1.3c se representa una señal modulada en frecuencia que se suele denominar FSK (acrónimo de *Frequency Shift Keying*). Al lector interesado en analizar con más detalle el proceso de modulación/demodulación se le remite a la bibliografía [BLAK 02] [FORO 07a] [HALL 05] [STAL 07].

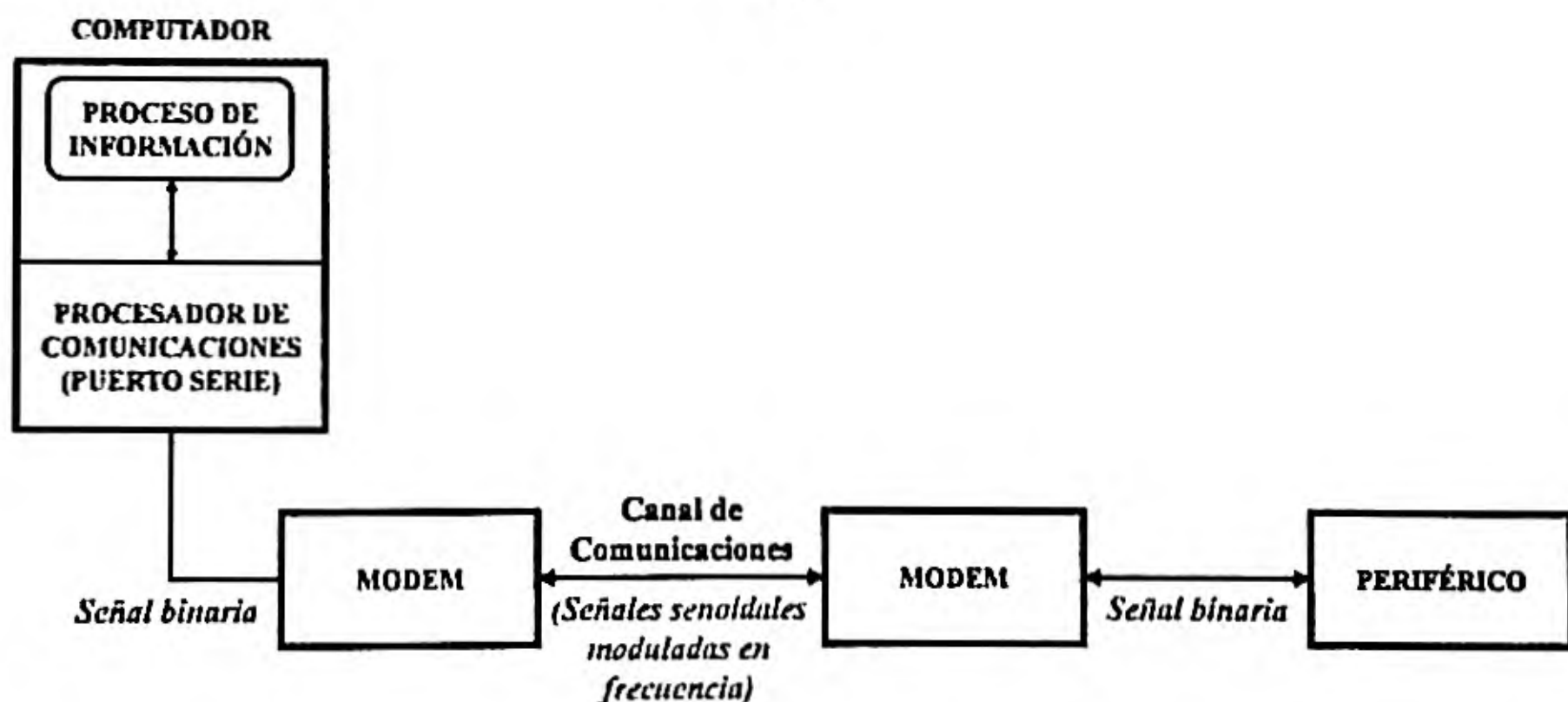
Para transmitir varios datos en banda base es necesario secuenciarlos (multiplexarlos en el tiempo) o utilizar tantos medios de transmisión (canales de comunicación) como datos se quieren transmitir simultáneamente. A la transmisión en secuencia de varios datos a través de un único medio de transmisión se la conoce como TDM (acrónimo de *Time Division Multiplexing*) y propicia que varios emisores y varios receptores de información puedan compartir un medio de transmisión, aunque sólo uno de ellos lo utilice en cada instante. La figura A1.4a constituye el diagrama de bloques de un sistema de transmisión punto a punto mediante TDM en el que un computador envía información secuenciada en el tiempo a dos periféricos diferentes a través un único canal. La figura A1.4b expresa gráficamente el comportamiento de la transmisión de datos TDM.

Transmisión en banda ancha

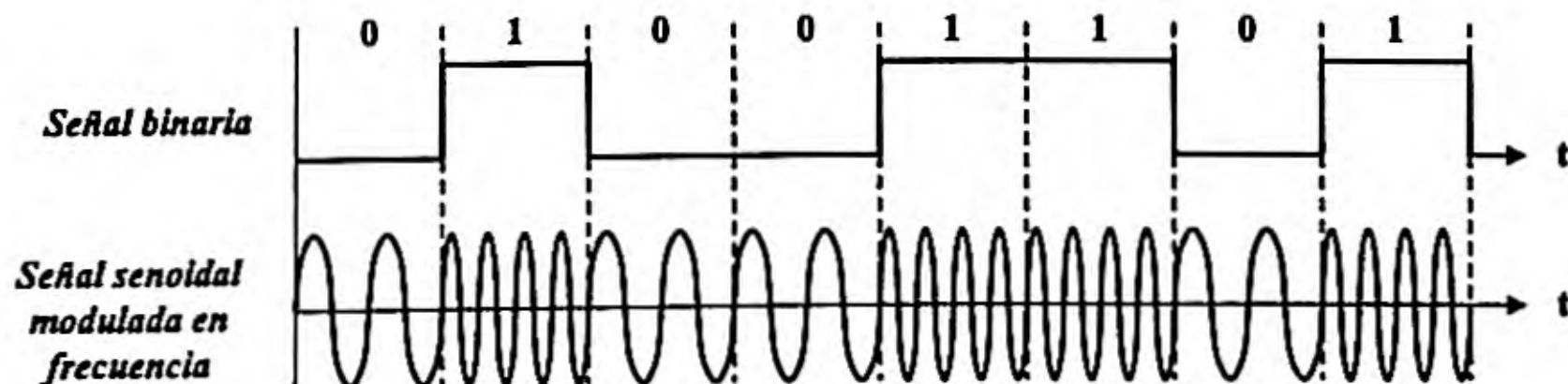
Cuando la señal que soporta la información modula una señal senoidal de elevada frecuencia se tiene un sistema de transmisión en banda ancha (*Broadband transmission*) que se caracteriza por permitir la transmisión simultáneamente, a través de un único medio de transmisión, de varias señales digitales, haciendo que cada una de ellas module una portadora senoidal de una frecuencia diferente, lo cual recibe la denominación de multiplexado en la frecuencia y se conoce como FDM (acrónimo de *Frequency División Multiplexing*) [FORO 07a] [STAL 07] (Figura A1.5). La transmisión en banda ancha es la única que permite la comunicación a distancia a través de la atmósfera, tal como se indica en el apartado A1.2.5.5 y además



a)



b)



c)

Figura A1.3. *Computador con capacidad de comunicación en banda base analógica:*
 a) Diagrama de bloques de la comunicación entre un computador y un periférico que recibe información de él; b) Diagrama de bloques de la comunicación entre un computador y un periférico en ambos sentidos utilizando un MODEM; c) Ejemplo de señal senoidal modulada en frecuencia.

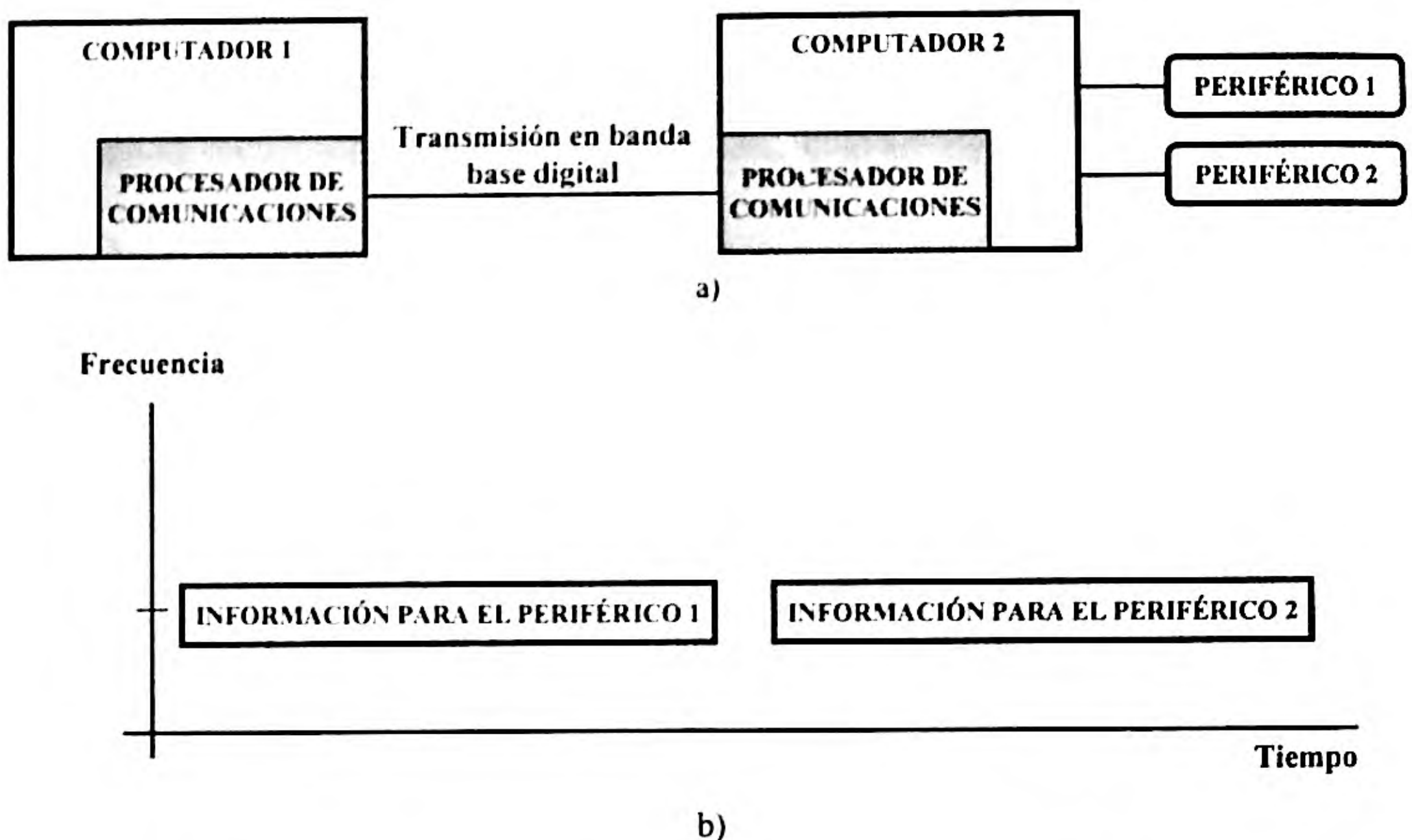


Figura A1.4. Concepto de multiplexación en el tiempo en una conexión punto a punto: a) Diagrama de bloques; b) Representación gráfica de la transmisión de información.

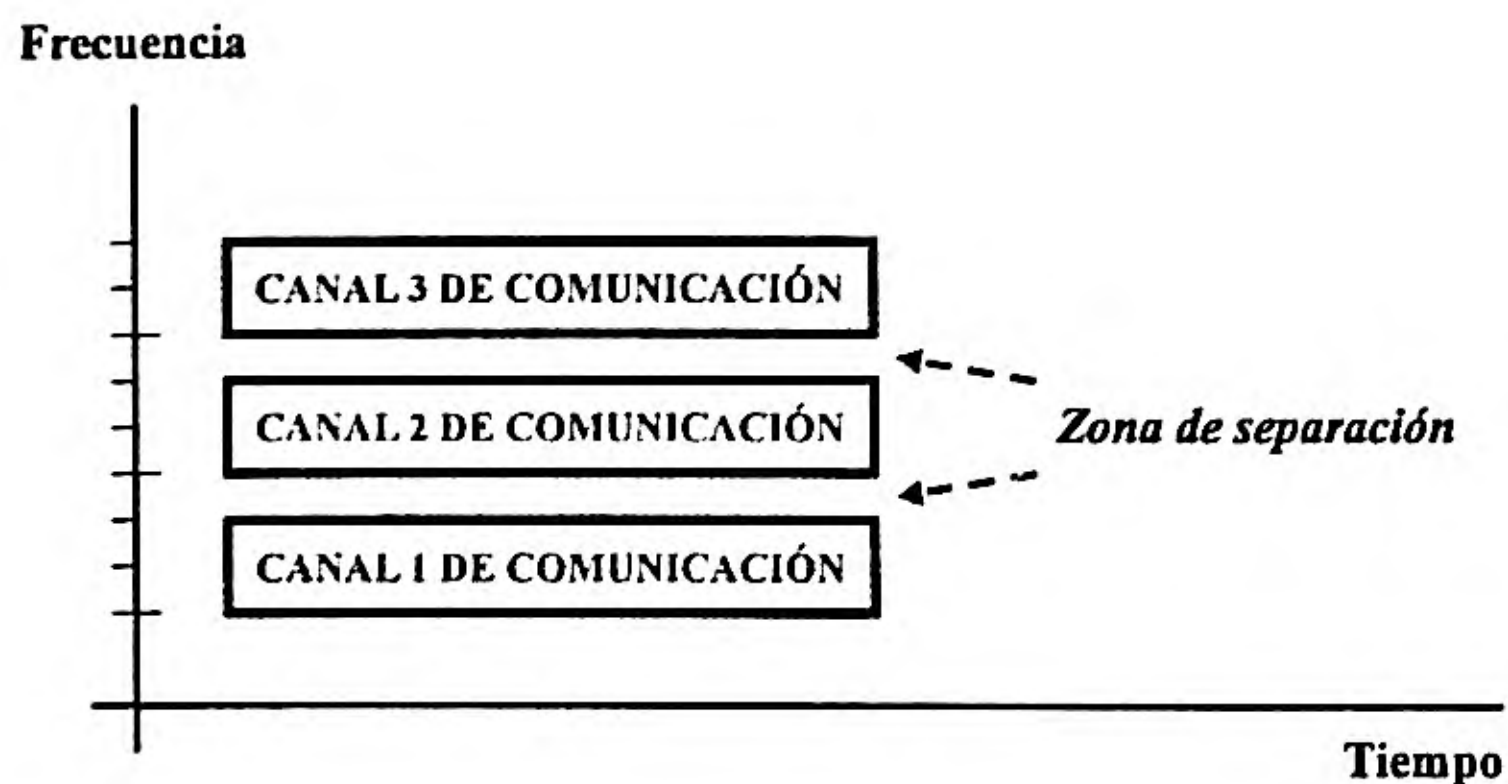


Figura A1.5. Representación gráfica de la relación entre la frecuencia de transmisión y el instante en que se realiza la misma en un sistema de multiplexación en la frecuencia.

la transmisión en banda ancha de señales analógicas mediante FDM permitió la radiodifusión inalámbrica mediante la cual diversas fuentes de información comparten el aire y pueden ser detectadas simultáneamente (Figura A1.6). Otro ejemplo de utilización del multiplexado en la frecuencia es la tecnología ADSL (acrónimo de *Asymmetric Digital Subscriber Line*)

desarrollada para facilitar el acceso a Internet a través de la línea telefónica tradicional. En la figura A1.7 se representa el diagrama de bloques en el que se puede observar que a la línea telefónica de abonado (*Local loop*) se conecta el teléfono a través de un filtro paso bajo (*Low pass filter*) que elimina las señales de frecuencia superior a 4 kHz. Además el computador se conecta también a la línea telefónica a través de un filtro paso alto (*High pass filter*) y un MODEM que hace que pueda enviar y recibir datos a través de la misma a una frecuencia superior a la de la voz.

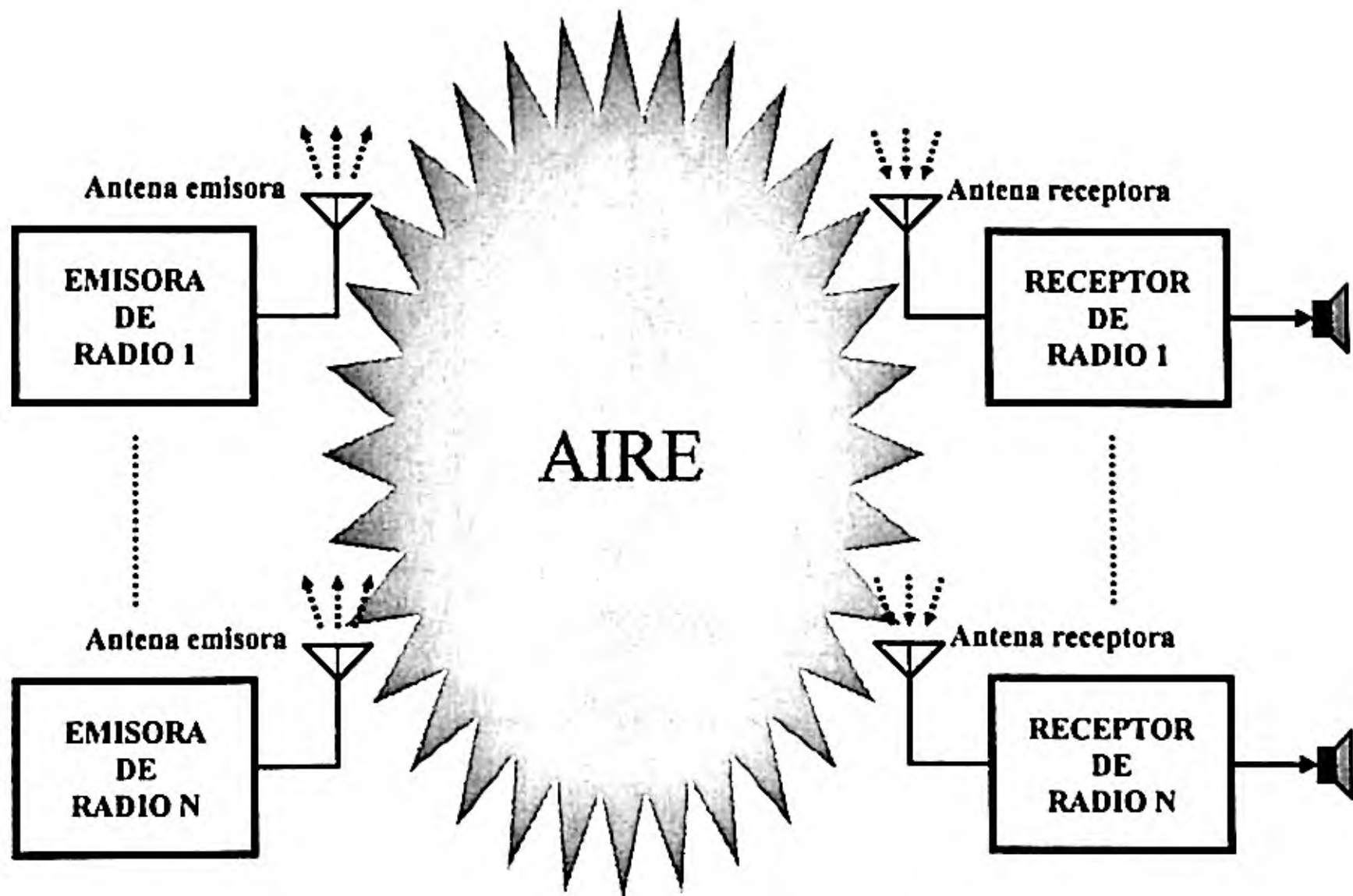


Figura A1.6. Utilización de la multiplexación en frecuencia para realizar transmisiones de radio.

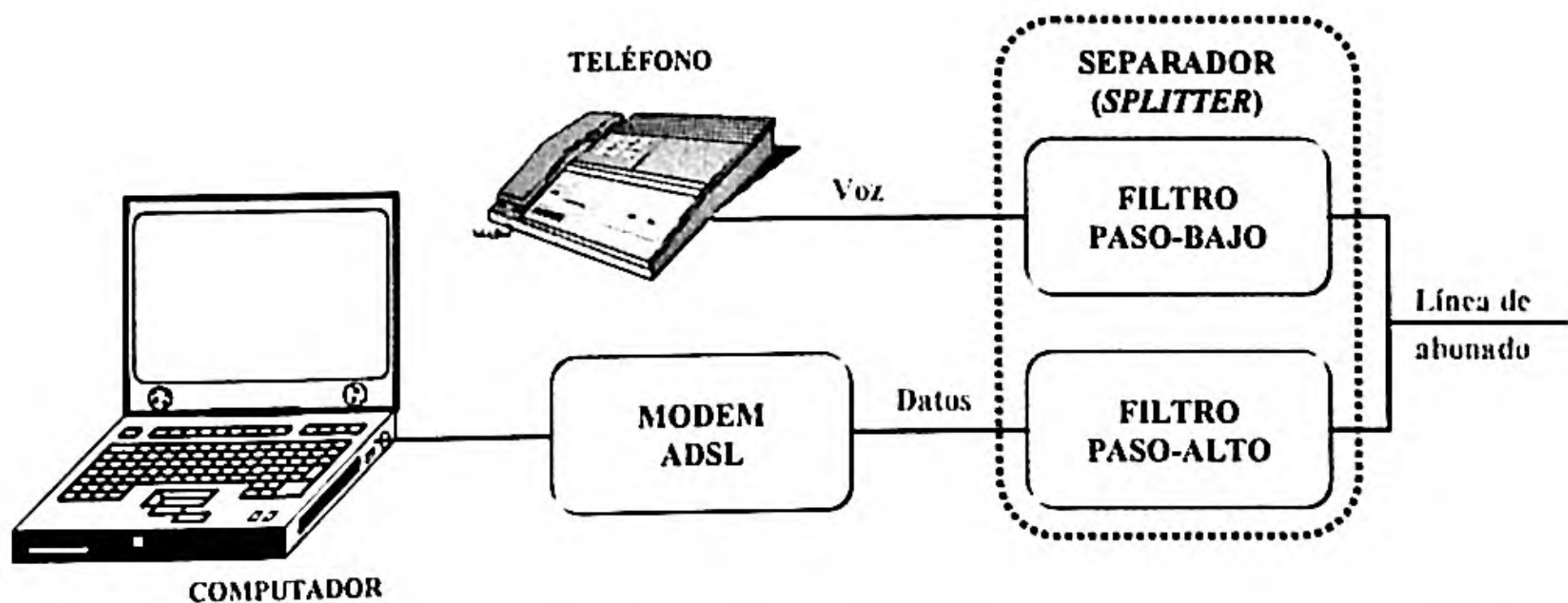


Figura A1.7. Diagrama de bloques de la comunicación de datos en banda ancha a través de la línea telefónica.

Un estudio en profundidad de la transmisión en banda ancha se sale fuera de los límites de este libro y por ello se remite al lector a la bibliografía [BLAK 02] [FORO 07a].

La combinación de la transmisión TDM con la FDM hace que se pueda enviar gran cantidad de información a través de un único medio de transmisión (cable, fibra óptica, aire, etc.) en un tiempo mínimo y ha permitido la realización de redes de datos a nivel mundial mediante la utilización de satélites de comunicaciones (Figura A1.8).

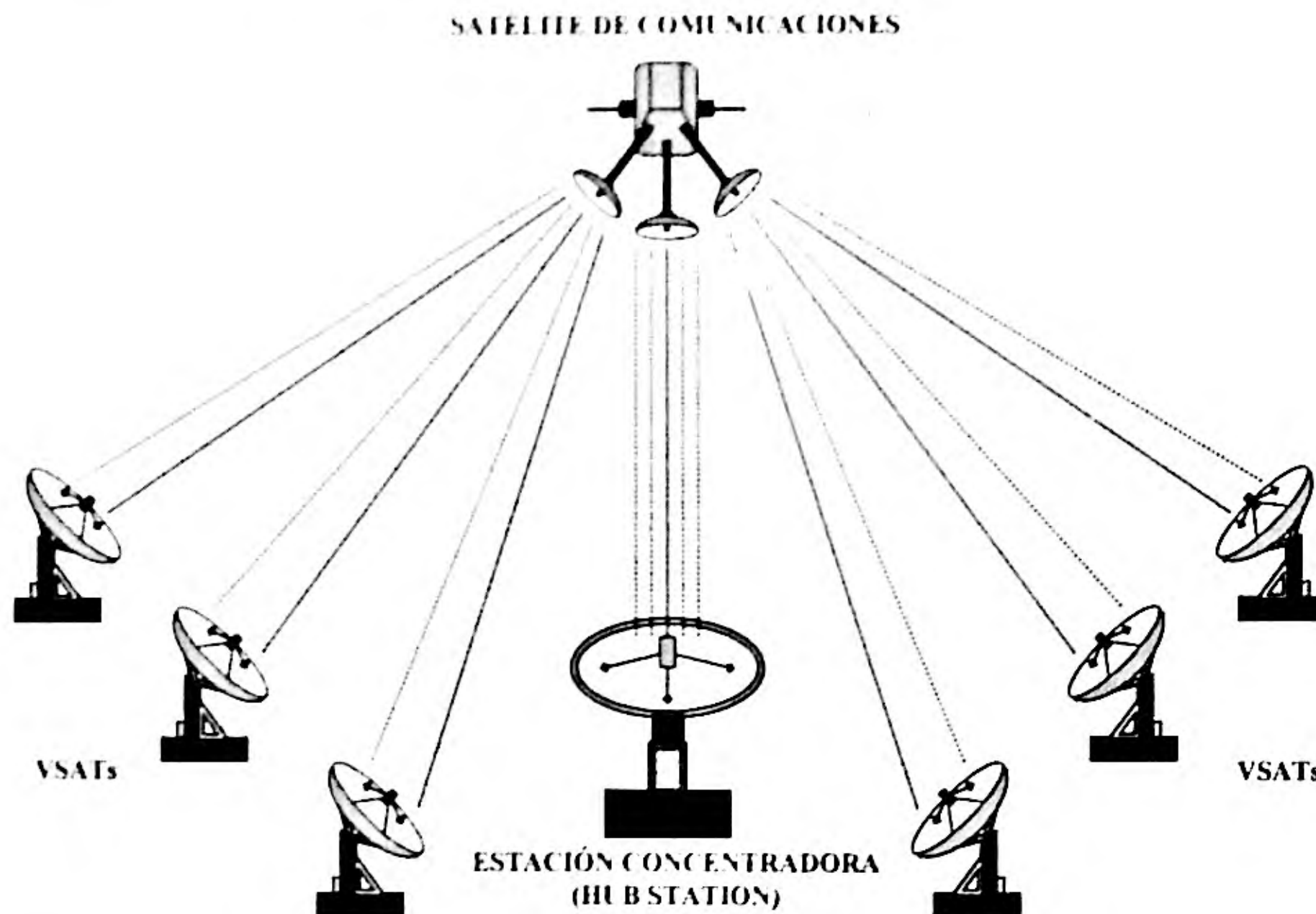


Figura A1.8. Red de datos mediante la combinación de TDM Y FDM en una red de comunicaciones via satélite.

A1.2.2 Modos de comunicación

El modo de comunicación (*Data flow mode*) indica la capacidad de dos sistemas digitales enlazados por una conexión punto a punto para intercambiar información entre ellos. Según el modo, la comunicación digital puede ser simple (*Simplex*), semidoble (*Half duplex*) o doble simultánea (*Full duplex*). Estos tres modos de comunicación se combinan con los modos de transmisión de las señales analizados en el apartado anterior, lo cual amplía el abanico de formas de implementar las comunicaciones digitales punto a punto.

Se dice que la comunicación se realiza en modo simple cuando sólo puede tener lugar en una dirección. Se utiliza para transferir información desde un procesador digital a un periférico de salida o para transferir información desde un periférico de entrada a un procesador digital. La comunicación simple sólo necesita un canal de comunicación o medio de transmisión cuando se realiza en banda base y por ello constituye una solución interesante en la práctica cuando sólo

se tiene que transmitir información en una dirección. Un ejemplo de comunicación en modo simple es la representada en la figura A1.1 si el puerto serie sólo tiene capacidad de transmisión o de recepción de información. También es un sistema de comunicación en modo simple el representado en la figura A1.3a.

La transmisión doble no simultánea, semidoble o semisimultánea es aquella en la que la transmisión de información se puede realizar en ambos sentidos pero no simultáneamente. Cuando se lleva a cabo en banda base, sólo necesita un único canal de comunicación. El sistema de la figura A1.1 puede corresponder tanto a un sistema de comunicación simple como semidoble. En este último caso el puerto serie del computador ha de poseer capacidad de transmisión y de recepción de información.

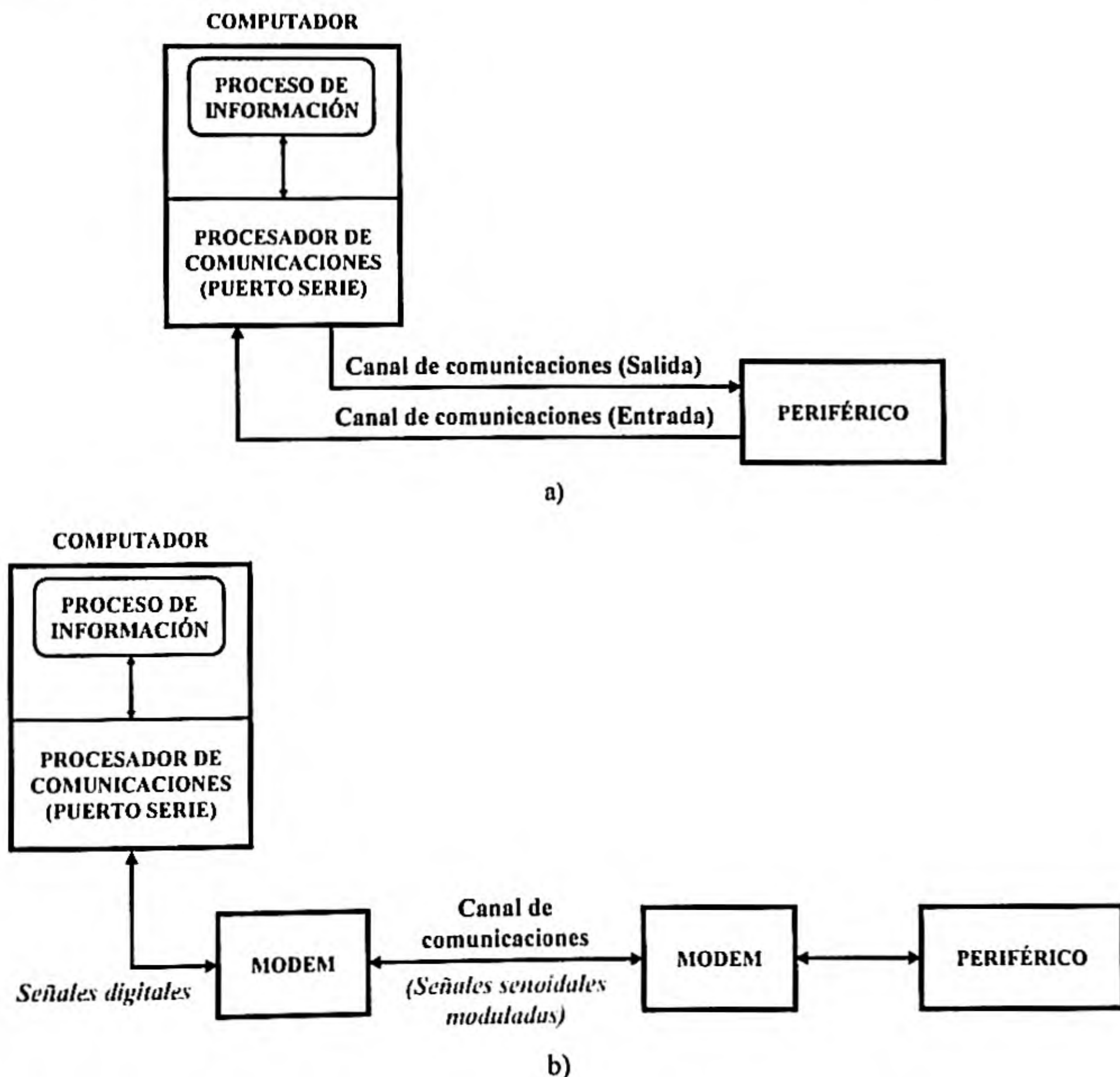


Figura A1.9. Sistema con capacidad de comunicación doble simultánea (Full dúplex): a) En banda base digital; b) En banda base analógica.

La transmisión doble simultánea se caracteriza por permitir la transmisión de información en ambos sentidos simultáneamente. No es posible realizarla en banda base digital a través de

un único canal de comunicación sino que necesita dos, uno para cada sentido (Figura A1.9a). Por el contrario, utiliza un único canal de comunicación en banda base analógica, haciendo que la información que se transmite en cada sentido module una portadora diferente. La figura A1.9b representa un sistema con capacidad de transmisión doble simultánea en banda base analógica para lo cual tanto el computador como el periférico están conectados a un MODEM.

A1.2.3 Modos de sincronización

La comunicación digital se realiza siempre entre sistemas digitales síncronos en los que la duración de los bits está fijada por un generador de impulsos o reloj (*Clock*) de una determinada frecuencia. Por otra parte la transmisión de la información en serie implica que los sucesivos bits estén separados sólo por la variable tiempo. Por ello, para recomponer adecuadamente la información, el sistema receptor ha de ser capaz de determinar la duración de cada bit y de detectar el instante en el que se inicia la transmisión, lo cual implica la sincronización de los instantes en los que capta la información con los establecidos por el transmisor. En la figura A1.10 se representa un bloque de información o trama (*Frame*) de una transmisión serie, que está en general formado por una información de inicio, la información transmitida propiamente dicha y una información de final.



Figura A1.10. Trama (*Frame*) de una transmisión en serie.

Según la forma en que se lleva cabo la sincronización (*Synchronization mode*), la transmisión puede ser asíncrona, síncrona o isócrona. Las tres constituyen un protocolo que es un conjunto de reglas que permiten la comunicación entre dos sistemas informáticos, tanto en banda base como en banda ancha. A continuación se describe el fundamento de cada una de ellas. Al lector interesado en estudiar con más detalle los modos de sincronización se le remite a la bibliografía [BLAK 02] [ELAH 01] [FORO 07] [HALL 05] [MILL 00] [STAL 07] [WALD 91].

A1.2.3.1 Transmisión asíncrona

La denominación de transmisión asíncrona (*Asynchronous transmission*) proviene del hecho de que el receptor utiliza su propio generador de impulsos o reloj, independiente del reloj del transmisor, para extraer la información contenida en las señales transmitidas.

La unión entre ambos sistemas digitales se realiza en banda base digital o analógica mediante un único canal de comunicación a través del cual se envían en serie la información y las señales de control que sincronizan el reloj del transmisor y el del receptor que oscilan a la misma frecuencia (Figura A1.11a). En la figura A1.11b se representa el formato de la información transmisión asíncrona. El receptor debe observar el canal y continuar realizando dicha observación, mientras éste se encuentre en el nivel correspondiente al uno lógico. El comienzo de la

transmisión de información lo indica el transmisor mediante la bajada a cero del canal durante un intervalo de tiempo, igual a la duración de un bit, que se denomina de “Comienzo o inicio de transmisión” (*Start*). A continuación de este bit de inicio, el receptor recibe, procedentes del transmisor, n bits de información (que suelen ser igual a ocho, siete de los cuales corresponden a un carácter del código alfanumérico ASCII y uno es de paridad) seguidos de uno o dos bits de “Fin de transmisión” (*Stop*) cuya función se entiende fácilmente al analizar como capta la información el receptor, lo cual se indica seguidamente.

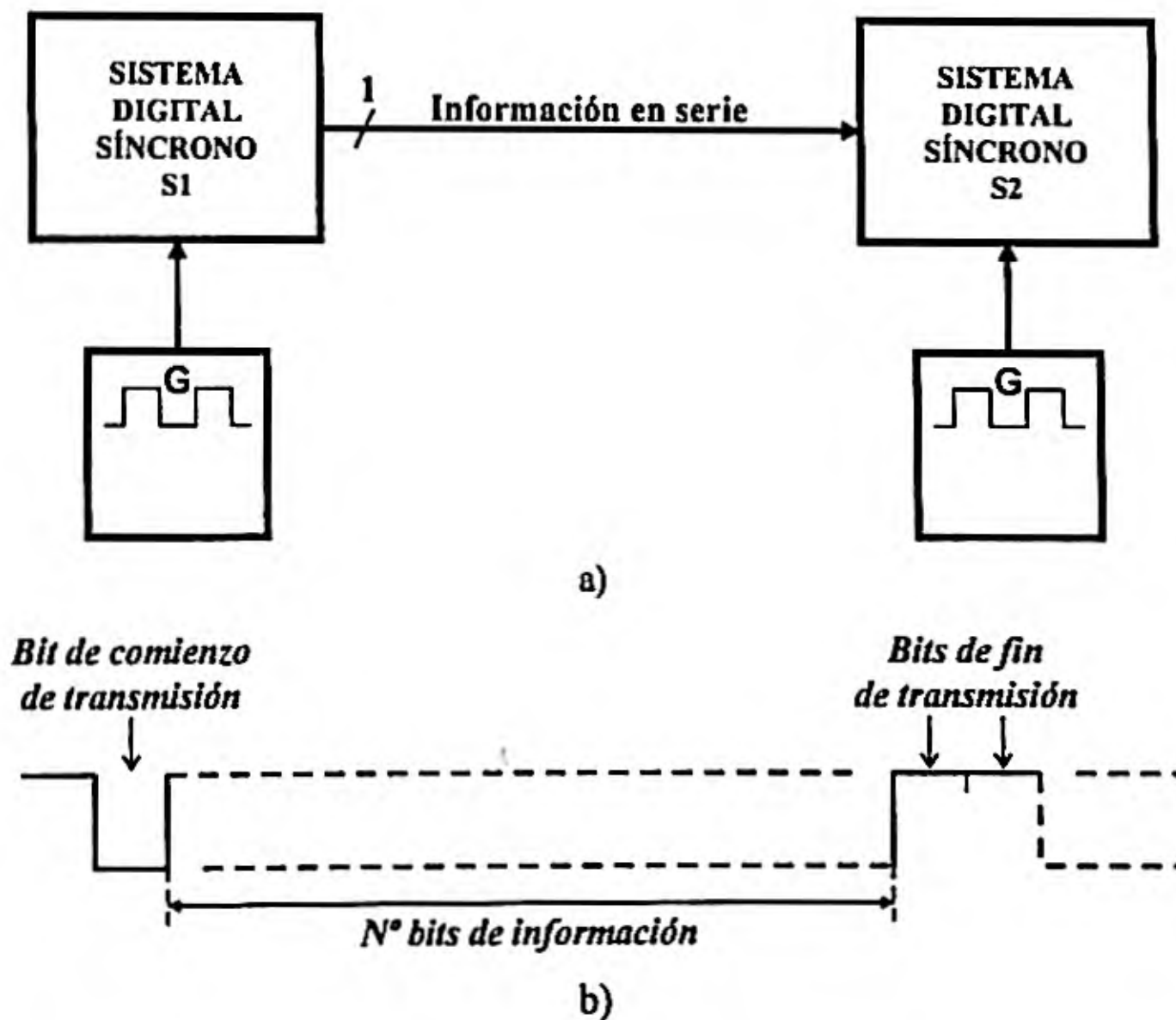


Figura A1.11. Transmisión asincrónica: a) Diagrama de bloques básico; b) Protocolo de comunicación a nivel de enlace.

El receptor es un sistema secuencial síncrono cuyo generador de impulsos posee un período que tiene un valor que es un submúltiplo de la duración de los bits de información (se suelen utilizar valores 16 ó 64 veces inferiores a la duración del bit). Esto permite realizar la sincronización con una precisión elevada en el medio de cada bit, tal como se indica gráficamente en la figura A1.12, en la que se supone que el período del generador de impulsos del receptor es 16 veces inferior a la duración del bit. A partir del instante en el que el canal pasa al nivel cero, la unidad de control cuenta 8 impulsos, capta el bit de comienzo y a continuación cada 16 impulsos del generador capta un nuevo bit. La duración de los bits recibidos depende del valor del período del generador situado en el transmisor y por ello puede diferir ligeramente de 16 períodos del generador de impulsos del receptor. Debido a ello, el número de bits enviados después del de inicio está limitado a un máximo de 8, y a continuación se transmiten uno o dos bits de fin de transmisión que permiten al receptor volver a sincronizarse. Por ello, si la diferencia entre la frecuencia de los generadores del emisor y del receptor no es elevada, el desplazamiento de la captación de la información con respecto al punto medio es insuficiente para producir una re-

cepción incorrecta de alguno de los 8 bits de información transmitidos entre el de inicio y los de fin de transmisión. Queda, por tanto, justificada la existencia de los bits de fin de transmisión.

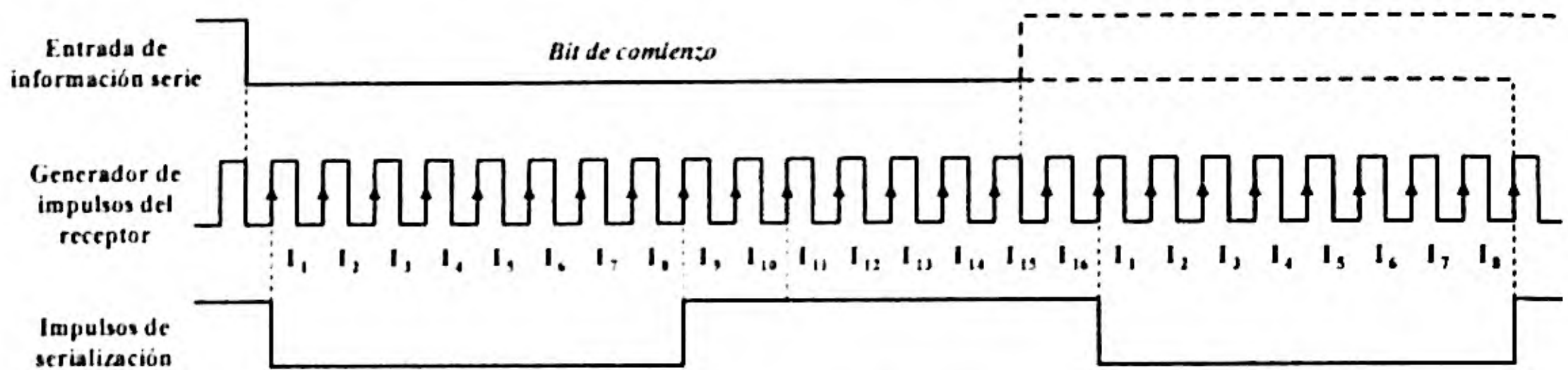


Figura A1.12. Diagrama de secuencia de la sincronización entre la información y el generador de impulsos del receptor.

Las funciones de añadir los bits de inicio y fin en el proceso de transmisión, así como de eliminarlos en el proceso de recepción, se suelen realizar mediante un sistema digital que, además, pasa la información del formato paralelo al serie. Por ello constituye un procesador de comunicaciones que se suele denominar UART (acrónimo de *Universal Asynchronous Receiver-Transmitter*). Diversos fabricantes de microprocesadores y microcontroladores le han dado a la UART nombres diferentes, como por ejemplo Motorola que la denomina ACIA (acrónimo de *Asynchronous Communication Interface Adapter*). El conjunto formado por el bit de inicio, los 8 bits de información (un carácter ASCII y el bit de paridad) y los dos bits de fin constituyen una trama (*Frame*). Para que el lector se haga una idea de la complejidad de una UART, en la figura A1.13 se representa el diagrama de bloques del transmisor basado en un registro de desplazamiento y un registro de modo de operación que permite programar varios parámetros, como por ejemplo la velocidad de transmisión, el número de bits de la trama, etc.

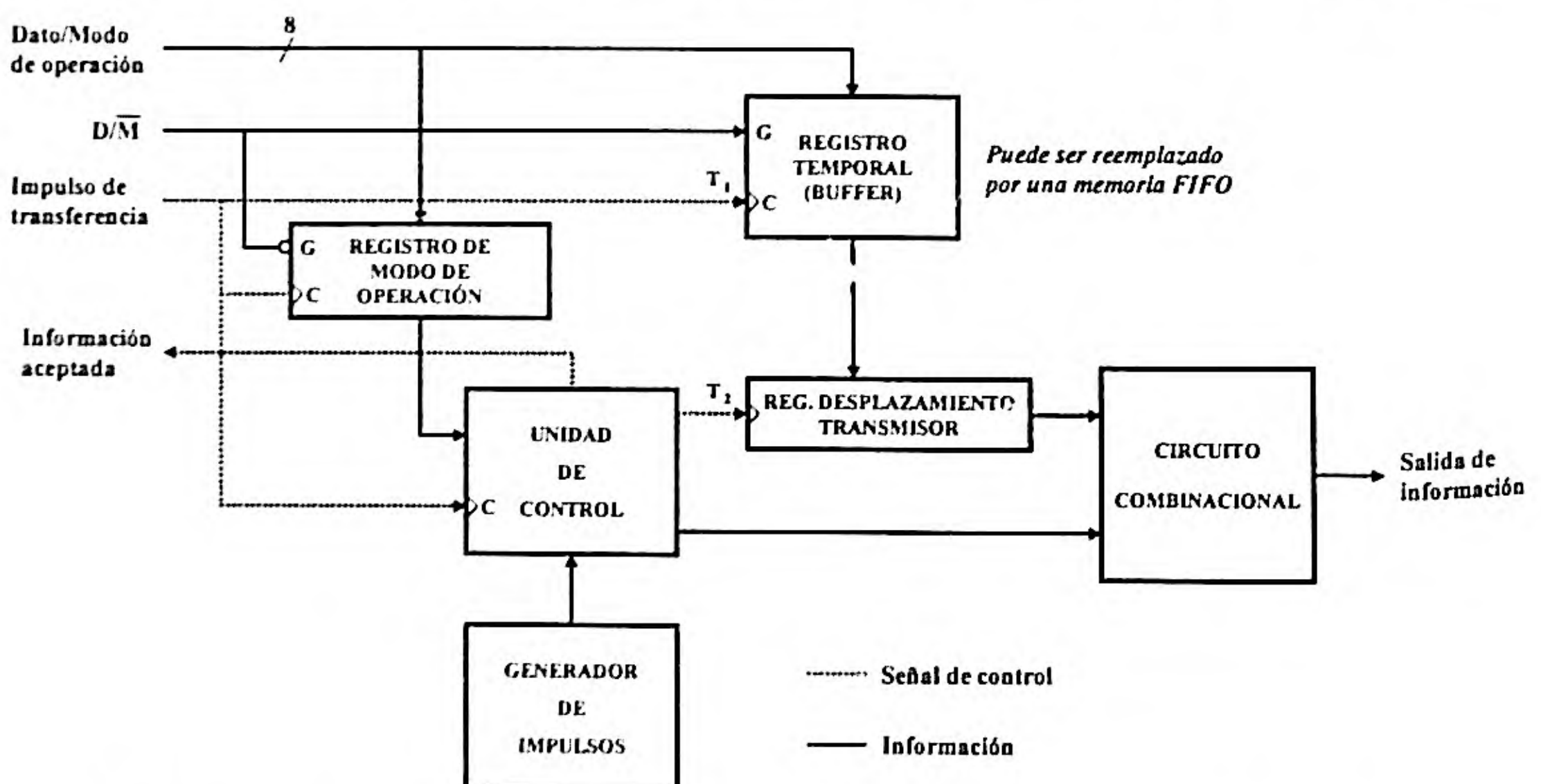


Figura A1.13. Diagrama de bloques de un transmisor asíncrono con registro de modo de operación.

A1.2.3.2 Transmisión síncrona

Tal como se acaba de indicar, la transmisión asíncrona no permite la transmisión de un número elevado de bits entre el de inicio y el de final para garantizar que la sincronización entre el transmisor y el receptor se realiza correctamente. Por ello su rendimiento es muy bajo, dado que el dato que se transmite es de ocho bits y por ello tres de cada once bits que se transmiten no constituyen información propiamente dicha. Además la longitud de la trama es muy reducida. Esto generó un gran interés por desarrollar un modo de transmisión más eficaz y de ello surgió la transmisión síncrona (*Synchronous transmission*) en la que el receptor no sólo recibe la información procedente del transmisor sino que también recibe una onda cuadrada que establece la duración de los sucesivos bits que la componen.

En la figura A1.14a se representa el fundamento de la transmisión síncrona, en la que la interconexión entre el procesador y el periférico se realiza mediante dos canales:

- Un canal de información a través del cual se envía la información propiamente dicha y la indicación del inicio de la transmisión.
- Un canal de control a través del cual se envían los impulsos de sincronización que indican los instantes en que se pasa de un bit al siguiente.

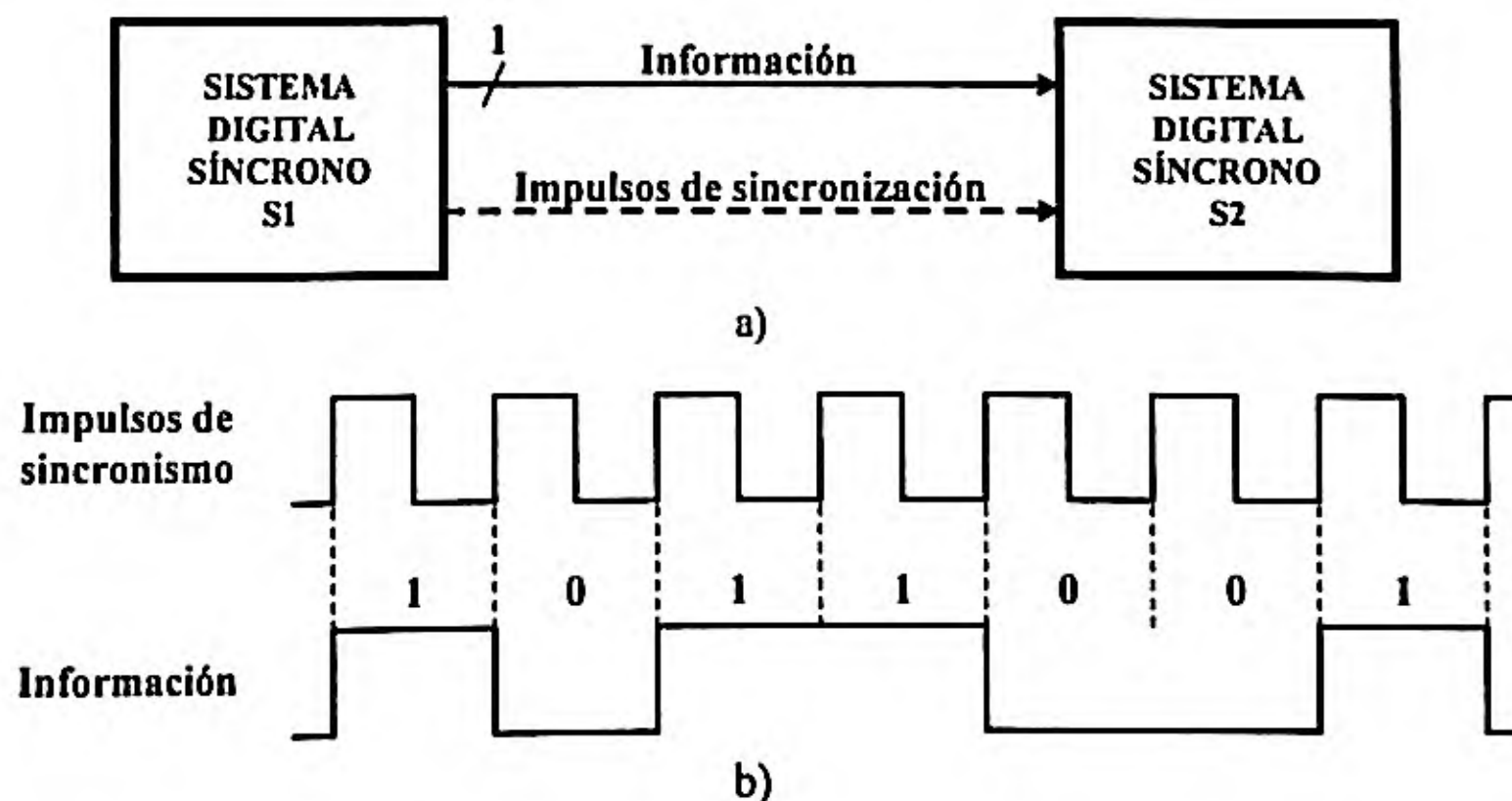


Figura A1.14. Transmisión síncrona a distancias cortas: a) Diagrama de bloques básico; b) Relación temporal entre el canal de información y los impulsos de sincronismo.

En la figura A1.14b se representa el diagrama de secuencia de las señales de la figura A1.14a en la que se supone que los instantes de conmutación de los sucesivos bits de información coinciden con los flancos de subida de los impulsos de sincronismo. Por lo tanto en los flancos de bajada de estos impulsos la información permanece estable en el canal correspondiente y puede ser captada por el receptor. La necesidad de enviar los impulsos de sincronismo a través de un canal independiente es debida a que el envío de dos o más ceros o unos seguidos no puede ser detectado por el receptor sin ayuda de los mismos. El hecho de que la información permanezca fija en el mismo nivel cuando se transmiten consecutivamente varios bits idénticos hace que se diga que está codificada (*Encoding*) en un código NRZ (acrónimo de *Non Return to Zero*). Es conveniente indicar que la palabra código hace referencia a la forma de transmitir los bits en

serie y no al código binario (binario natural, BCD natural, ASCII, etc.) utilizado. En la figura A1.15 se muestra como ejemplo el diagrama de bloques de una transmisión síncrona doble simultánea (*Full duplex*) a distancias cortas entre dos microcontroladores o computadores integrados (*Embedded microcomputers*) basada en la norma SPI (*Synchronous Peripheral Interface*) de Motorola [VALV 00]. Uno de ellos actúa como principal (*Master*) y el otro como subordinado (*Slave*) que responde a las órdenes que le envía el primero.

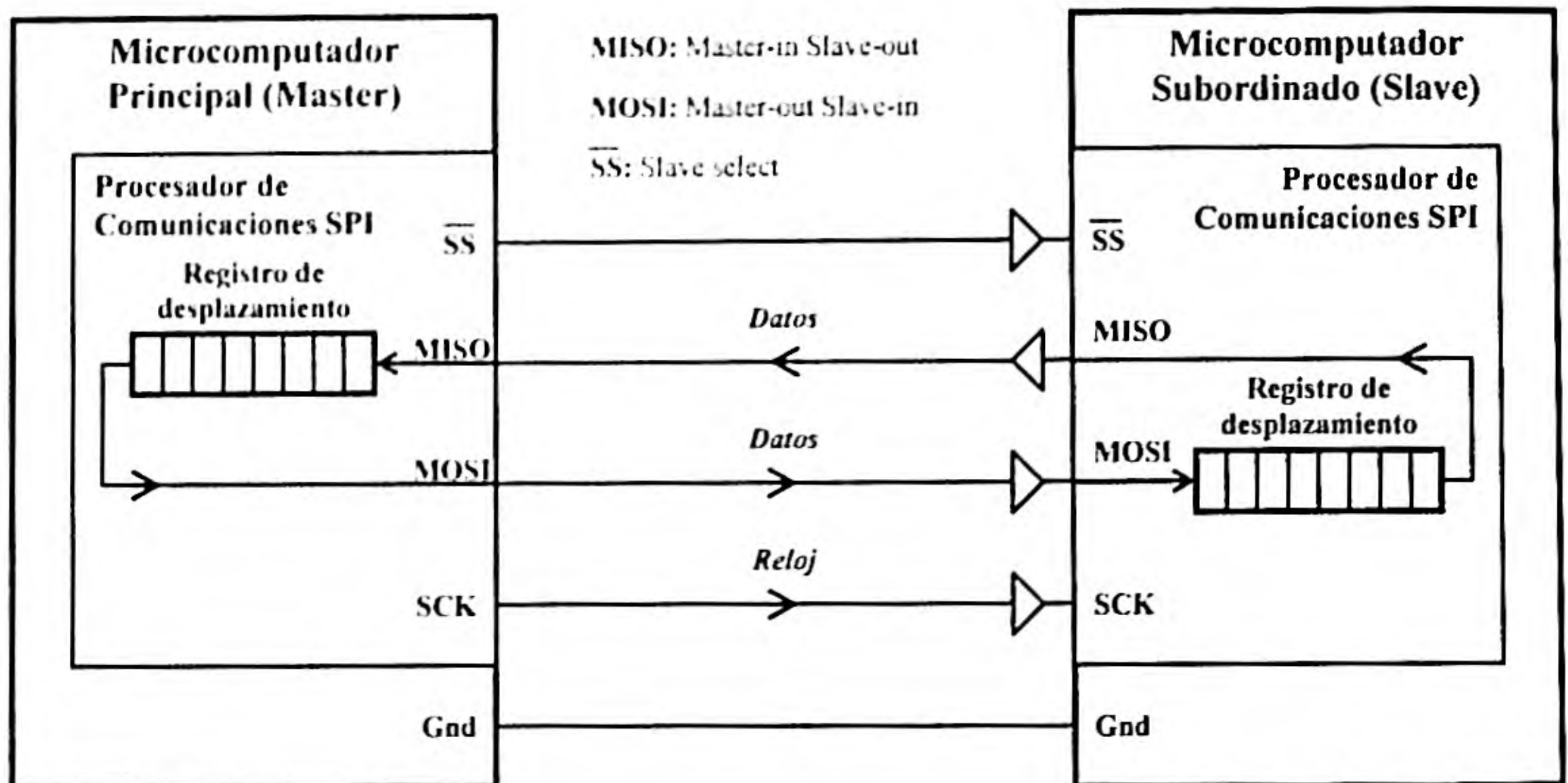


Figura A1.15. Ejemplo de diagrama de bloques de una transmisión síncrona doble simultánea (*Full duplex*) a distancias cortas basada en la norma SPI (*Synchronous Peripheral Interface*) de Motorola.

Pero cuando la distancia entre el transmisor y el receptor es elevada, no es posible enviar los impulsos de sincronismo y la información por canales diferentes porque la diferencia entre los retardos de propagación de ambos no garantiza una correcta recepción. Esto hace que sea necesario enviar la información combinada con los impulsos de sincronismo de tal forma que en el receptor se puedan separar a fin de utilizar dichos impulsos para captar la información en el receptor (Figura A1.16). Para ello, la información que se transmite debe representar los unos y los ceros de forma tal que en cada uno de ellos se produzca una transición de uno a cero o viceversa, lo que hace que se pueda obtener el reloj a partir de ellas. No existe unanimidad en la denominación de este tipo de códigos pero el hecho de que la información no permanezca fija en un nivel, tanto cuando representa un uno como un cero, hace que diversos autores los denominen códigos autosincronizados (*Self-clocking codes*) [STAL 07].

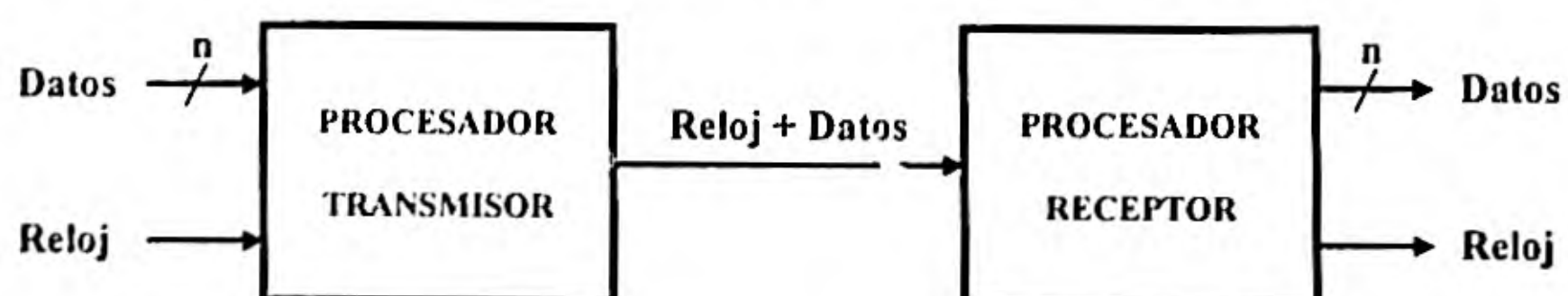


Figura A1.16. Fundamentos de la transmisión síncrona a distancias largas.

Son numerosos los códigos autosincronizados. Dos de los más utilizados en las redes informáticas son el código Manchester y el Manchester diferencial. En la figura A1.17 se representan junto con el NRZ para que el lector pueda compararlos. El código Manchester se caracteriza por dividir el bit en dos partes, de forma que el nivel de la primera es igual al inverso del nivel del código NRZ y el de la segunda coincide con él. De esta forma en el medio de cada bit se tiene una transición a partir de la cual se obtiene el reloj. Por su parte el código Manchester diferencial se caracteriza porque además de cambiar de estado en el medio de cada bit lo hace al principio cuando el código NRZ cambia de nivel y no lo hace en el caso contrario.

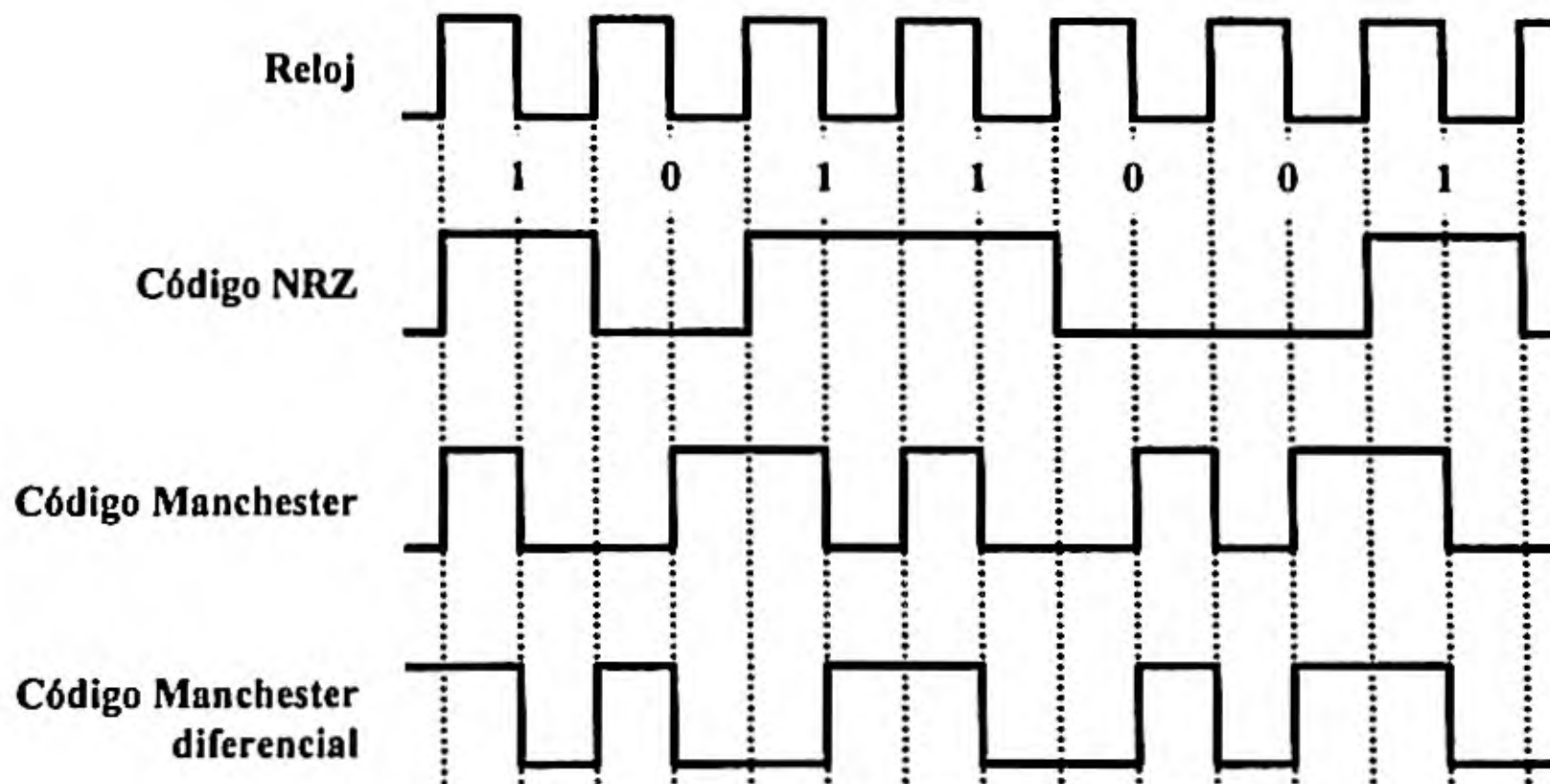


Figura A1.17 Representación de la información transmitida en serie en los códigos NRZ, Manchester y Manchester diferencial.

Los códigos autosincronizados se obtienen fácilmente a partir del NRZ y viceversa mediante un circuito digital, tal como se indica en la figura A1.18. En la figura A1.19 se representa el circuito convertidor del código NRZ en el Manchester que está constituido por una puerta lógica que implementa la función O-exclusiva [MAND 08a]. En la figura A1.20 se muestra el esquema de bloques de un receptor síncrono.

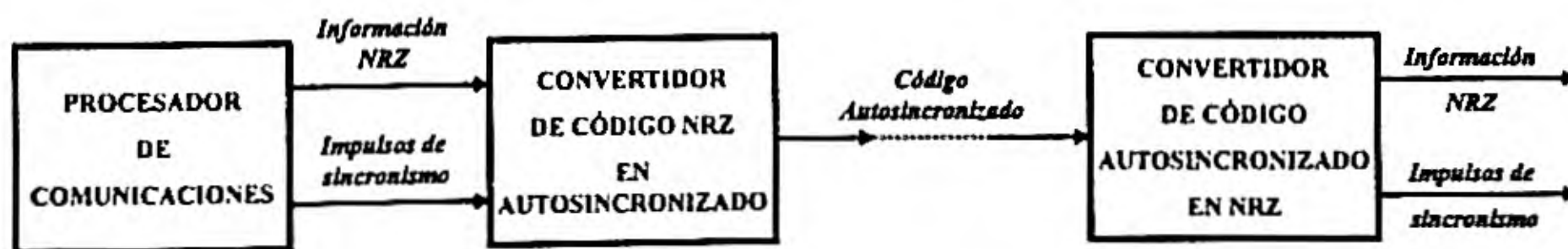


Figura A1.18. Diagrama de bloques de un circuito de transmisión en un código autosincronizado.

Pero, además de recuperar el reloj, en el receptor es necesario detectar el inicio y el fin de la transmisión, a fin de obtener un protocolo síncrono (*Synchronous Data-Link Protocol*). Existen infinitas opciones para establecer un protocolo síncrono, de las que se han utilizado un

número reducido impulsadas por los fabricantes más importantes de equipos informáticos. Según sus características los protocolos síncronos se pueden clasificar en tres tipos:

- Protocolos síncronos orientados al carácter.
- Protocolos síncronos orientados al conteo de objetos.
- Protocolos síncronos orientados al bit.

A continuación se describe brevemente cada uno de ellos.

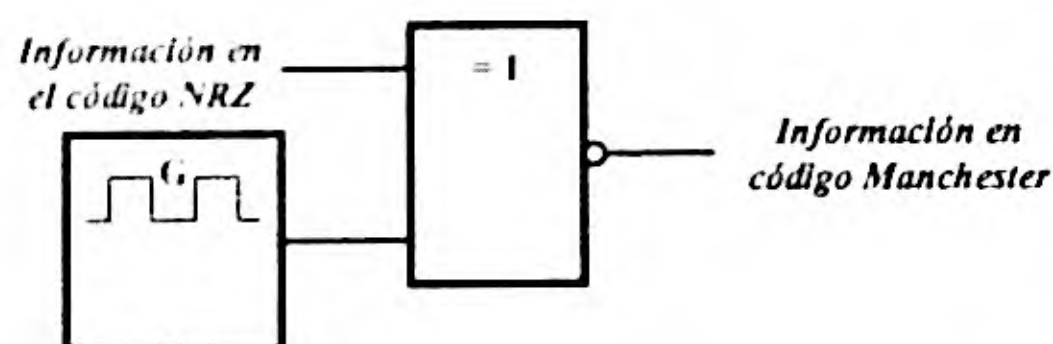


Figura A1.19. Fundamentos de la transmisión síncrona a distancias largas.

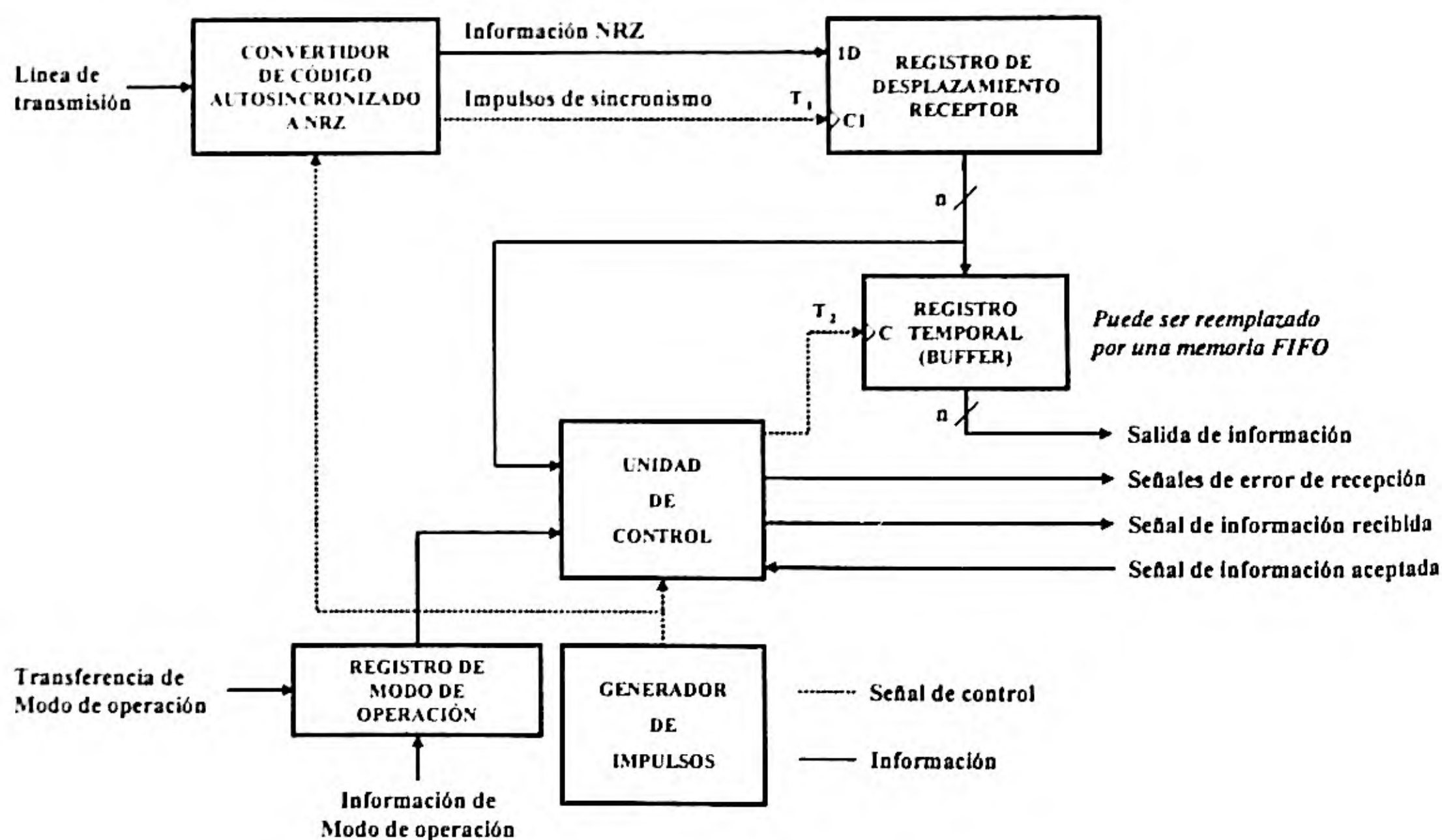


Figura A1.20. Diagrama de bloques de un receptor síncrono.

Protocolos síncronos orientados al carácter

Son protocolos en los que los datos o texto están formados por un conjunto de caracteres y se utilizan caracteres especiales, como por ejemplo el de inicio de la transmisión SYNC (abreviatura de *Synchronism*), el de inicio de texto STX (abreviatura de *Start of Text*) para indicar el comienzo del mensaje y el de fin del bloque de transmisión ETB (abreviatura de

End of Transmission Block). En la figura A1.21 se representa el formato general de un protocolo síncrono orientado al carácter. El primer protocolo síncrono fue el BISYNC o BSC (acrónimo de *Binary Synchronous Control*) cuya trama de información se indica en la figura A1.22. El protocolo BYSINC tiene problemas de transparencia porque existe la posibilidad de que algún carácter del texto sea confundido con los caracteres ETX o ETB de fin de texto o fin de bloque respectivamente. Para lograr transparencia BISYNC utiliza el carácter DEL (acrónimo de *Data Link Escape*) que precede siempre a los símbolos de control que deben ser tratados como datos.

SINCRONIZACIÓN Y CONTROL	DATOS O TEXTO	FIN DE TRANSMISIÓN	INFORMACIÓN REDUNDANTE
-----------------------------	------------------	-----------------------	---------------------------

Figura A1.21. Formato general de la trama (Frame) de un protocolo orientado al carácter.

SYN	SYN	SOH	Cabecera	STX	Datos	ETX	BCC
-----	-----	-----	----------	-----	-------	-----	-----

Figura A1.22. Representación de una trama (Frame) de información en el protocolo BISYNC.

Protocolos síncronos orientados al conteo de octetos

Para superar las dificultades del protocolo BYSINC, a fin de lograr una transmisión y recepción transparentes, se concibieron protocolos que consiguen la transparencia mediante el conteo de octetos sin necesidad de utilizar el carácter DEL. Un ejemplo típico es el código DDCMP (acrónimo de *Digital Data Communication Message Protocol*), desarrollado por Digital Equipment Corporation, cuya trama de información se indica en la figura A1.23, que pone en evidencia su complejidad.

SYN	SYN	CLASS	CONTAJE 14 bits	INDICADOR (FLAG) 2 bits	RESPUESTA (RESPONSE) 8 bits	SECUENCIA (SEQUENCE) 8 bits	DIRECCIÓN (ADDRESS) 8 bits	COMPROBACIÓN CRC 1 (16 bits)	INFORMACIÓN hasta 16K de caracteres (8 bits)	COMPROBACIÓN CRC 2 (16 bits)
-----	-----	-------	-----------------	----------------------------	--------------------------------	--------------------------------	-------------------------------	---------------------------------	--	---------------------------------

Figura A1.23. Formato general de una trama (Frame) del protocolo DDCMP.

Protocolos síncronos orientados al bit

Las desventajas de los protocolos orientados al carácter y la complejidad de los orientados al conteo de octetos promovieron actividades de investigación aplicada para desarrollar protocolos orientados al bit de los que los más típicos son el SDLC (acrónimo de *Synchronous Data Link Control*) de IBM, utilizado por ella en sus redes, el ADCCP (acrónimo de *Ad-*

vanced Data Communications Control Procedure) normalizado por ANSI a partir del SDLC, y el HDLC (acrónimo de *High-level Data Link Control*) normalizado por ISO a partir también del SDLC.

El protocolo HDLC, cuya trama de información se indica en la figura A1.24, utiliza la combinación (*Flag*) 0 1111110 para indicar tanto el inicio como el final de la transmisión de una trama (*Frame*) de información. Para que en el medio de la información no pueda estar incluida la combinación antes citada, el transmisor añade un cero cada vez que envía cinco unos seguidos y el receptor lo elimina. El campo de dirección (*Address*) indica cuál es la estación receptora, entre las diferentes existentes, a la que va destinado el mensaje. El campo de control facilita la retransmisión de mensajes recibidos con algún error, para lo cual posee varios formatos, como son el de supervisión, el de información y el de fuera de secuencia. Este último indica que la información está fuera de la secuencia normal, por ser, por ejemplo, fruto de una retransmisión por error.

Combinación de inicio	Dirección	Control	Información	Comprobación	Combinación de fin
0 1 1 1 1 1 1 0	1 octeto	1 o 2 octetos	Longitud variable	CRC 16 bits	0 1 1 1 1 1 1 0
<i>Opening flag</i>	<i>Address</i>	<i>Control</i>	<i>Information</i>	<i>Checking</i>	<i>Closing flag</i>

Figura A1.24. Formato general de la trama (*Frame*) del protocolo HDLC.

El protocolo HDLC fue concebido inicialmente para facilitar el enlace de varios procesadores, sin jerarquía establecida entre ellos (*Peer to peer*), a través de un canal único de comunicaciones en forma de bucle que recibe la denominación de red de área local de topología en anillo, que se describe en el apartado A1.3.3.1.5.

De lo expuesto en este apartado y en el anterior se deduce que, para llevar a cabo tanto una transmisión asíncrona como una síncrona, es necesario realizar un conjunto de operaciones con la información que se desea enviar. Esto justifica la afirmación realizada en el apartado A1.1 de que la transmisión de datos implica la necesidad de que los equipos de procesos de datos, como por ejemplo los autómatas programables, posean una unidad de interfaz que ejecute las tareas correspondientes, por lo que recibe en la actualidad el nombre genérico de procesador de comunicaciones (*Communication processor*). Las figuras A1.13 y A1.20 son ejemplos de circuitos digitales que forman parte de un procesador de comunicaciones.

A1.2.3.3 Transmisión isócrona

La transmisión síncrona permite transmitir gran cantidad de información con un mínimo de información redundante y por ello supera a la asíncrona cuando se debe transmitir una gran cantidad de datos. Pero existen aplicaciones, como por ejemplo la transmisión de datos en una instalación industrial, en las que a través de un único canal hay que enviar:

- Un número reducido de datos que deben llegar al receptor en un tiempo que no puede superar un límite determinado para que la información recibida sea útil. Un ejemplo de ello es la transmisión de información de máquinas que tienen partes en movimiento

(*Motion control*), que se deben controlar a distancia con tiempos de respuesta inferiores a 1 ms. En general se suele decir que la comunicación en este caso se debe hacer en tiempo real (*Real Time transmission*).

- Muchos datos que no presentan una exigencia elevada en lo que se refiere al tiempo máximo que pueden tardar en transmitirse. Un ejemplo de ello es la transmisión de información de muchas variables, como por ejemplo la temperatura, que exigen tiempos de respuesta mínimos superiores a 10 ms e incluso a 100 ms.

Para combinar ambas situaciones se desarrolló la transmisión isócrona (*Isochronous transmission*) que divide la transmisión en ciclos, tal como se indica en la figura A1.25, en cada uno de los cuales se transmite una parte de la información que tiene una menor exigencia temporal junto con la que sí la tiene. El valor máximo del tiempo de ciclo debe ser elegido adecuadamente para garantizar la transmisión correcta de la información en tiempo real.

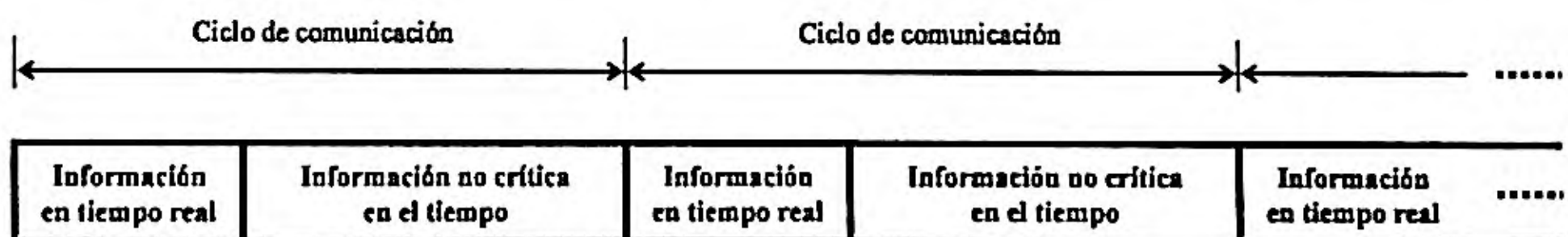


Figura A1.25. *Formato básico de la transmisión isócrona.*

La transmisión isócrona es, por lo tanto, un caso especial de la transmisión síncrona, que se utiliza en las redes Ethernet Industrial que forman parte de las Comunicaciones Industriales, que se estudian en el apartado 9.3.2.5 del capítulo 9, y en particular en la red Profinet descrita en el apéndice 4.

A1.2.4 Métodos de detección de errores

La posibilidad de detectar errores (*Error detection*) en la información recibida, debidos, por ejemplo, a las interferencias electromagnéticas a que pueden estar sometidas las señales eléctricas portadoras de la información es una de las grandes ventajas del formato digital sobre el analógico. Para ello es necesario, tal como se indica gráficamente en la figura A1.26, añadir información redundante (*Redundant*) antes de proceder a la transmisión y verificar la información recibida para comprobar que no se han producido cambios en la misma.

La generación de información redundante se puede llevar a cabo de múltiples formas y por ello ha sido necesario elegir unas cuantas y proceder a su normalización. Por otra parte, las diferencias entre los modos de sincronización asíncrono y síncrono, descritos en los apartados anteriores, hace que los métodos para generar la información difieran notablemente entre ellos.

A continuación se describen los fundamentos de ambos y al lector interesado en estudiarlos con más detalle se le remite a la bibliografía al final de este apéndice [BLAK 02] [ELAH 01] [FORO 07a] [MILL 00] [STAL 07].

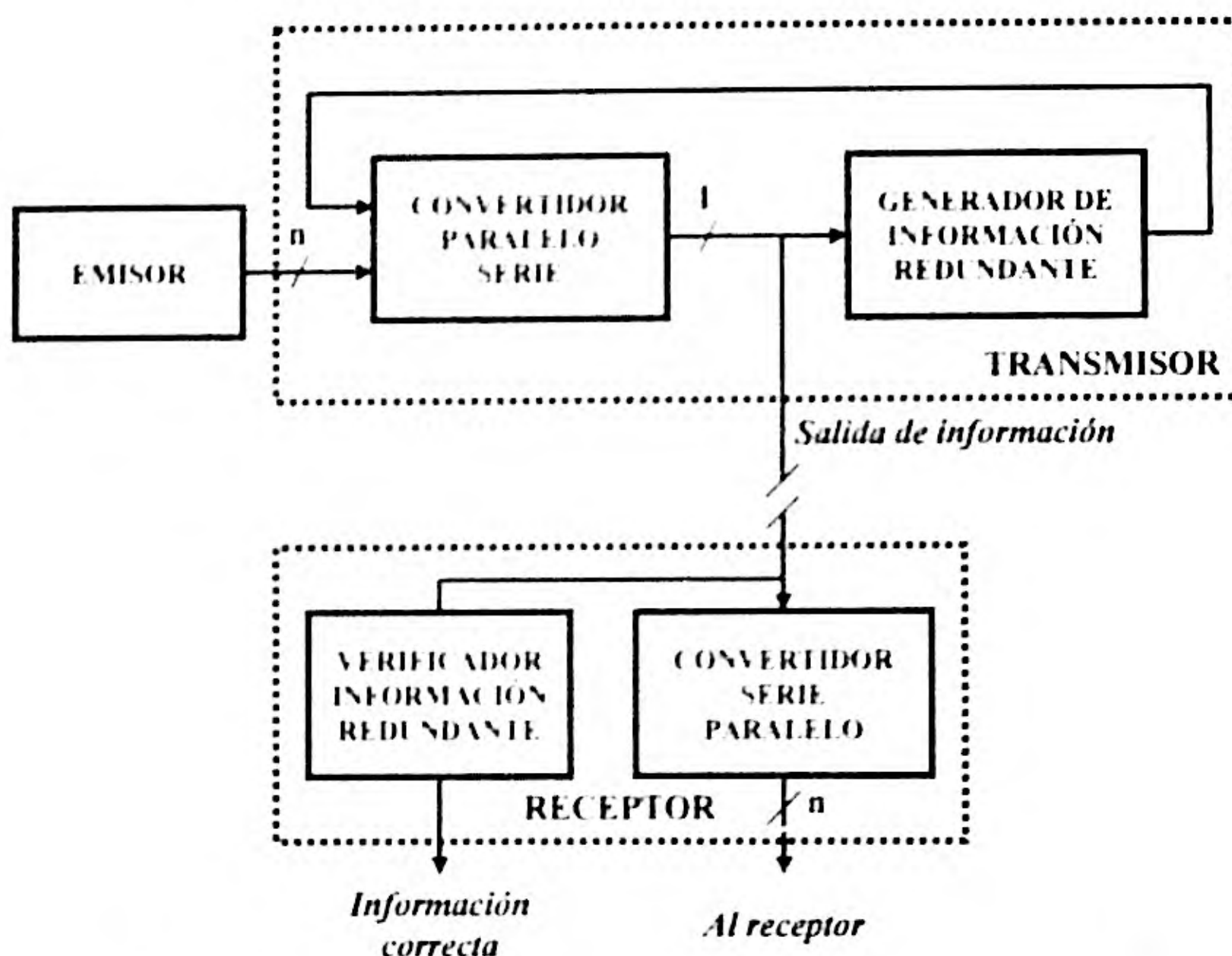


Figura A1.26. Diagrama de bloques básico de un sistema de transmisión digital que tiene un circuito generador de información redundante.

A1.2.4.1 Detección de errores en la transmisión asíncrona

Tal como se indica en el apartado A1.2.3.1, en la transmisión asíncrona se transmiten en general caracteres ASCII de 7 bits entre los bits de inicio y fin de transmisión. Por ello la detección de errores se realiza añadiendo un bit de paridad impar que da lugar a un carácter de ocho bits [MAND 08b]. Mediante este proceso, denominado de paridad vertical, (*Vertical parity*) se detectan errores en un solo bit. Para mejorar la capacidad de detección de errores se suele además añadir un carácter de paridad cada 2^8 (256) caracteres, de forma que la secuencia de bits del mismo peso tenga la paridad constante. Esta paridad recibe el nombre de longitudinal u horizontal (*Horizontal parity*) para distinguirla de la paridad de cada carácter que se denomina vertical. Estas denominaciones fueron adoptadas cuando los computadores utilizaban las cintas perforadas como memoria externa, que fue sustituida posteriormente por los discos.

A1.2.4.2 Detección de errores en la transmisión síncrona

Dado que en la transmisión síncrona el campo de la información propiamente dicha tiene una longitud variable (Figura A1.24), la capacidad de detección de errores se consigue añadiendo una combinación binaria obtenida a partir de aquélla. En el receptor se vuelve a obtener dicha combinación a partir del contenido del campo de información y sólo si ambas coinciden se considera que la comunicación es correcta. La forma más utilizada de generar dicha combinación binaria es la verificación mediante redundancia cíclica denominada CRC (acrónimo de *Cyclic Redundancy Check*) que consiste en añadir a un bloque de k bits una combinación de n bits tal que el bloque resultante de $n+k$ bits sea divisible exactamente por un determinado número [FORO 07a] [MAND 08].

La forma más usual de generar el CRC es la constituida por un conjunto de funciones O-exclusiva combinadas con desplazamientos y puede ser realizada en el procesador de comunicaciones mediante puertas O-exclusiva y un registro de desplazamiento. Existen varios tipos de CRC según el divisor utilizado. En la figura A1.27 se representa como ejemplo el circuito correspondiente al CRC-16.

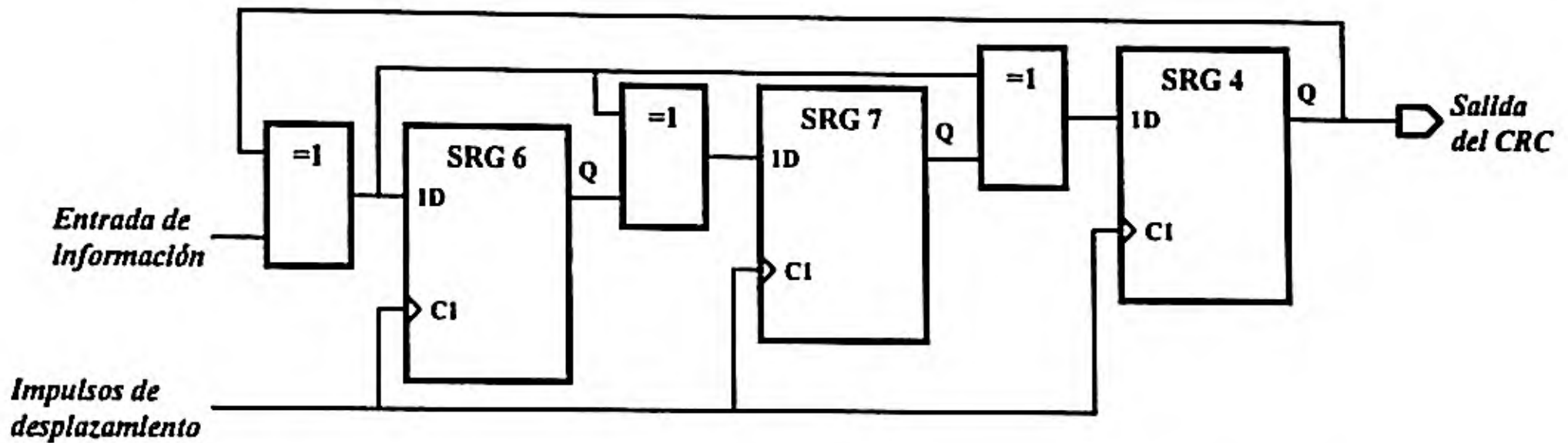


Figura A1.27. Circuito de generación de la combinación de comprobación mediante redundancia cíclica (CRC).

A1.2.5 Medios de transmisión

A1.2.5.1 Introducción

Las señales eléctricas portadoras de información constituyen radiaciones electromagnéticas que deben propagarse a través de un medio físico que enlaza el transmisor con el receptor. Los medios de transmisión se pueden clasificar en dos grupos:

- **Medios guiados (*Guided or wired transmission media*)**

Los medios guiados son aquellos en los que las señales eléctricas se transmiten a través de una región del espacio perfectamente delimitada. Los más importantes son los conductores y especialmente el par de hilos, el cable coaxial y la fibra óptica.

- **Medios no guiados (*Unguided transmission media*)**

El medio no guiado por excelencia es la atmósfera y la transmisión realizada a través de él recibe el nombre de transmisión inalámbrica (*Wireless transmission*).

Para analizar los diferentes medios de transmisión es conveniente tener en cuenta el espectro de las señales eléctricas que pueden soportar información en alguno de sus parámetros. Se entiende por espectro de una señal la gama de frecuencias de las señales senoidales cuya suma es igual a la misma. La frecuencia es el número de veces que se repite la señal en la unidad de tiempo y se mide en ciclos por segundo o hercios. Se define la longitud de onda λ de una onda electromagnética como el espacio que recorre en un tiempo igual a la duración de un periodo y por lo tanto se obtiene mediante la ecuación:

$$\lambda = C \cdot T = \frac{C}{f}$$

en la cual C es la velocidad de propagación de una onda electromagnética en el vacío, que es aproximadamente igual a 300.000 Km/seg.

Por ejemplo, si la frecuencia de una onda electromagnética es de 30 MHz su longitud de onda es:

$$\lambda = \frac{3 \times 10^8}{3 \times 10^7} = 10 \text{ m}$$

El espectro de comunicaciones es la gama completa de frecuencias utilizadas en la transmisión de información a distancia. En la tabla A1.1 se representa el espectro total de las radiaciones electromagnéticas dividido en gamas. Las señales de frecuencia comprendidas entre 1 y 40GHz reciben la denominación de microondas, debido al valor reducido de su longitud de onda.

ONDAS ELECTROMAGNETICAS	FRECUENCIA	DENOMINACIÓN	SIGLAS EN INGLÉS	LONGITUD DE ONDA (λ)
	30 - 300 Hz	Frecuencias Ultrabajas [Extremely Low Frequencies]	ELF	10 a 1Mm
	300 - 3000 Hz	Frecuencias de Voz [Voice Frequencies]	VF	1000 a 100Km
	3 - 30 KHz	Frecuencias muy Bajas [Very Low Frequencies]	VLF	100 a 10Km
	30 - 300 KHz	Frecuencias Bajas [Low Frequencies]	LF	10 a 1Km
	300 - 3000 KHz	Frecuencias Medias [Medium Frequencies]	MF	1000 a 100m
	3 - 30 MHz	Frecuencias altas [High Frequencies]	HF	100 a 10m
	30 - 300 MHz	Frecuencias muy altas [Very High Frequencies]	VHF	10 a 1m
	300 - 3000 MHz	Frecuencias ultra altas [Ultra High Frequencies]	UHF	100 a 10cm
	3 - 30 GHz	Frecuencias super altas [Super High Frequencies]	SHF	10 a 1cm
	30 - 300 GHz	Frecuencias extra altas [Extremely High Frequencies]	EHF	10 a 1mm
	300 - 3000 GHz			1000 a 100 μ m
LUZ		Infrarrojo		1 a 0,75 μ m
		Luz Visible		750 a 380nm
		Ultravioleta		380 a 1nm

Tabla A1.1. Espectro de radiaciones electromagnéticas.

Los principales parámetros que se deben tener en cuenta al seleccionar el medio de transmisión, debido a su influencia en la distancia y la velocidad de transmisión, son:

- **La anchura de banda (*Bandwidth*)**

La anchura o ancho de banda establece la gama de frecuencias de las señales senoidales que pueden transmitirse sin deformación o distorsión. La distorsión se produce en las señales compuestas, que son suma de diferentes componentes senoidales que sufren un desfase y una atenuación diferentes al desplazarse a través del medio.

- **La atenuación**

La atenuación de una señal es la disminución que se produce en su amplitud al desplazarse a lo largo del medio. Para evitar su efecto negativo es necesario utilizar amplificadores (repetidores) que devuelven a la señal su amplitud original.

- **La influencia de las interferencias electromagnéticas**

Se definen las interferencias electromagnéticas como señales electromagnéticas generadas por fuentes no previstas que alteran la señal portadora de información.

- **El número de receptores**

En los medios guiados utilizados para establecer conexiones punto a punto compartidas por varios emisores y receptores de información, la atenuación y la distorsión de la señal y por lo tanto la máxima distancia y la velocidad de la transmisión, dependen del número de los mismos.

En sucesivos apartados se describen las principales características de los diferentes medios de transmisión citados. Para un estudio más detallado se remite al lector a la bibliografía indicada al final del capítulo [FORO 07a] [MIL 00] [STAL 07].

A1.2.5.2 Conductores

El par trenzado (*Twisted pair*) (Figura A1.28) constituye el tipo de medio guiado más utilizado para realizar las comunicaciones digitales a distancias cortas, tanto en banda base como en banda ancha debido a su coste reducido. Consiste en dos hilos de cobre aislados y entrelazados de forma regular para reducir al máximo la influencia de las interferencias electromagnéticas. Frecuentemente, bajo una cubierta única se dispone un cierto número de pares trenzados. Se puede realizar de dos formas distintas:

- **Par trenzado sin blindar**

El par trenzado sin blindar denominado UTP (acrónimo de *Unshielded Twisted Pair*) (Figura A1.28a) es el medio de comunicación más utilizado por su bajo coste y fácil instalación. Mediante él se realizan diferentes tipos de redes de área local descritas en el apartado A1.3.3.3.

- **Par trenzado blindado**

El par trenzado blindado denominado STP (acrónimo de *Shielded Twisted Pair*) tiene una capa metálica que rodea a cada par de cables entrelazados (Figura A1.28b). Además de reducir las interferencias de las señales electromagnéticas externas, reduce la influencia de un par sobre otro. Su coste es mayor, así como su instalación más complicada, porque el blindaje se debe conectar a tierra en un extremo.

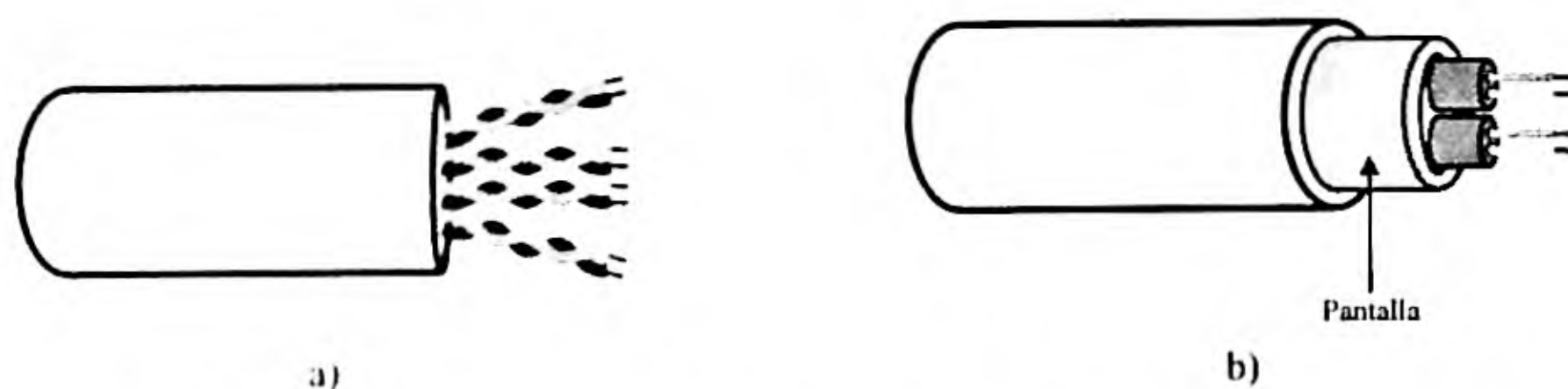


Figura A1.28. Conductores. a) Par trenzado sin apantallar; b) Par trenzado apantallado.

A1.2.5.3 Cable coaxial

El cable coaxial (*Coaxial cable*), al igual que el par trenzado, está formado por dos conductores pero su forma constructiva es diferente y consiste en un conductor cilíndrico que rodea a un único conductor central (Figura A1.29). Debido a su construcción concéntrica es menos susceptible a las interferencias electromagnéticas que el cable de par trenzado y puede transmitir señales de mayor frecuencia y a una distancia superior.

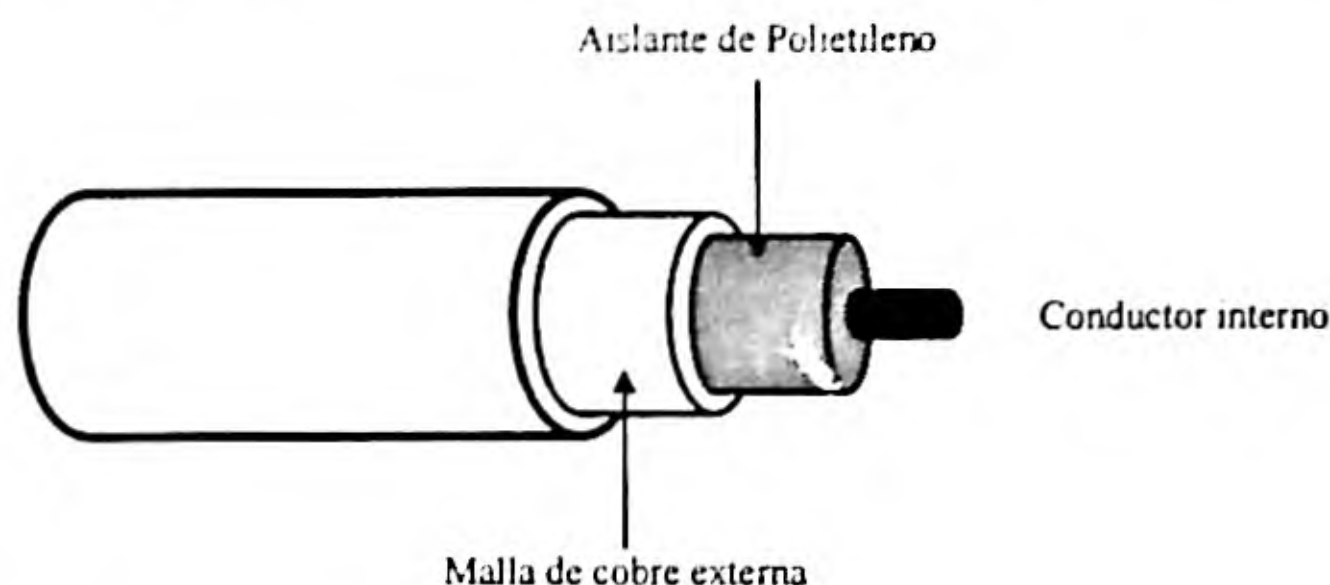


Figura A1.29. Cable coaxial.

A1.2.5.4 Fibra óptica

La fibra óptica (*Optical fiber or fiber-optic cable*) es un medio guiado utilizado para transmitir información mediante radiaciones ópticas confinadas en una región del espacio. Está constituida por un núcleo (*Core*) compuesto por un material de cristal o plástico (en general SiO_2) rodeado de una cubierta (*Cladding*) de un material en el que se refleja la luz, debido a que posee un índice de refracción diferente. Además tienen un tercer cilindro externo, denominado funda o recubrimiento (*Jacket*), que es una protección primaria a base de barniz de acrilato o resina de silicona que protege a la fibra contra la abrasión, rotura y agresiones químicas.

Las principales características de la fibra óptica son:

- Atenúa muy poco las señales, lo que hace que se puedan realizar líneas de transmisión de gran longitud sin utilizar amplificadores intermedios.
- Es flexible y tiene un tamaño reducido, lo cual facilita su instalación.
- Es muy estable frente a variaciones de las condiciones ambientales, como por ejemplo la temperatura, la humedad, etc., lo que le proporciona una gran fiabilidad.

- Posee un elevado ancho de banda, lo que permite la comunicación a frecuencias muy superiores a las de cualquier otro tipo de medio físico y le confiere una gran capacidad de transmisión de información.
- Utiliza material dieléctrico, inmune a las radiaciones electromagnéticas, lo que le proporciona una gran seguridad de funcionamiento y hace que sea especialmente idónea para ser utilizada en ambientes con elevados niveles de ruido eléctrico, como el que existe en muchas plantas industriales y en algunos sistemas como los automóviles.
- No emite radiaciones electromagnéticas, por lo que no necesita apantallamiento.
- La distorsión de las señales que se transmiten a través de ella es reducida e independiente de la frecuencia.

La fibra óptica puede trabajar en monomodo (*Monomode fiber*) o en multimodo (*Multi-mode fiber*). La fibra monomodo tiene un núcleo de diámetro reducido (algunas micras) y minimiza la distorsión y la atenuación de las señales. La fibra multimodo posee un núcleo de diámetro superior a las decenas de micras y es más fácil de acoplar a una fuente de luz, aunque presenta una mayor atenuación y distorsión que la hace idónea para distancias menores.

Dado que los computadores representan la información mediante señales eléctricas es necesario convertirlas en radiaciones ópticas mediante un diodo luminiscente conocido también como LED (acrónimo de *Light Emitting Diode*) o un diodo láser, palabra castellanizada a partir del término inglés LASER (acrónimo de *Light Amplification by Stimulated Emission of Radiation*).

En la figura A1.30 se representa el fundamento de la transmisión mediante fibra óptica de las señales digitales generadas por un procesador de comunicaciones. A la salida de este último se conecta un elemento conversor de una señal eléctrica digital en una señal óptica que se acopla a la fibra óptica. Al final de la misma, un dispositivo electrónico sensible a la luz la convierte de nuevo en una señal eléctrica que se conecta al procesador de comunicaciones receptor.

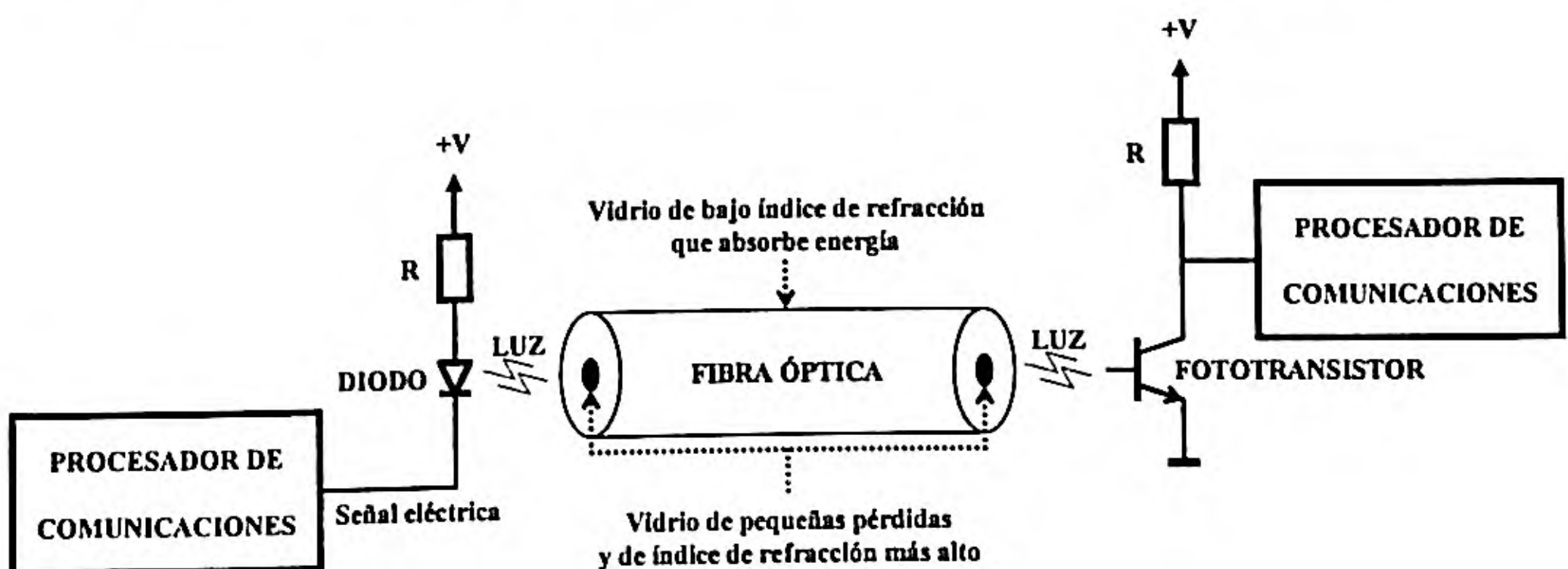


Figura A1.30. Fibra óptica como medio de transmisión.

Al lector interesado en estudiar con mayor detalle las comunicaciones a través de las fibras ópticas se le remite a la bibliografía [MILL 00] [PERE 03] [FORO 07a].

A1.2.5.5 La atmósfera

La atmósfera (*Atmosphere*) es el medio típico no guiado con el que se realizó la transmisión a grandes distancias sin necesidad de utilizar cables, primero de la voz (Radio) y de la imagen posteriormente (Televisión) mediante señales analógicas (descritas en el apartado 7.2.3 del capítulo 7) combinadas con señales temporales de sincronización (descritas en el apartado 7.4.1.4 del capítulo 7), para lo cual las señales eléctricas que soportan la información (señales en banda base) modulan una portadora de mayor frecuencia y se envían al aire mediante una antena transmisora. Las ondas electromagnéticas así generadas (*Radio waves*) se propagan a través de la atmósfera hasta alcanzar una antena receptora que lleva a cabo la función opuesta.

La transmisión de ondas electromagnéticas a través del aire recibe la denominación de inalámbrica (*Wireless communication*) y puede ser omnidireccional o direccional. Las transmisiones omnidireccionales son aquellas en las que las ondas electromagnéticas se propagan en todas las direcciones y pueden ser recibidas en múltiples antenas. Se utilizan para transmitir las señales de radio y de televisión en frecuencias comprendidas entre 300 KHz (límite inferior de la MF) y 3000 MHz (límite superior de la UHF) aproximadamente. Las señales de las bandas MF y HF (Tabla A1.1) se reflejan en la ionosfera (capa de partículas cargadas que se encuentra en la parte más alta de la atmósfera terrestre) y por ello se pueden utilizar para transmitir la voz de un punto de la tierra a otro sin que exista enlace visual entre ambos. Las señales de las bandas VHF y UHF no se reflejan en la ionosfera y su alcance es óptico, debido a lo cual necesitan reemisores o repetidores fijos o móviles y se utilizan en la práctica totalidad de las comunicaciones, es decir, de la voz (telefonía), televisión, datos, localización de móviles, etc.

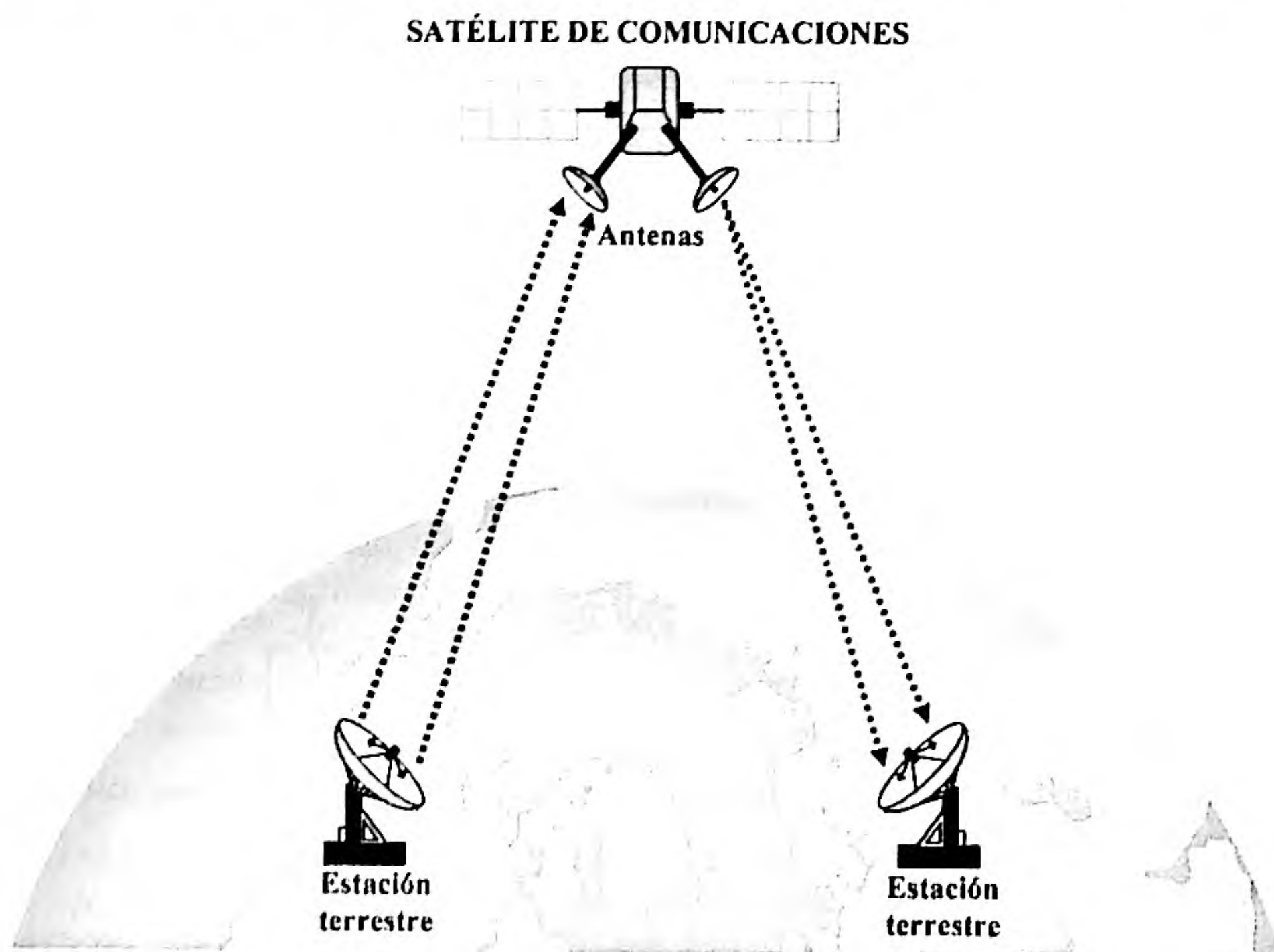


Figura A1.31. Comunicación punto a punto vía satélite entre dos estaciones terrestres a la frecuencia de las microondas.

La transmisión inalámbrica direccional se realiza mediante microondas que son ondas electromagnéticas cuya frecuencia está comprendida entre 1 GHz y 40 GHz.

En relación con las comunicaciones digitales, la atmósfera es el medio que se utiliza para llevarlas a cabo en banda ancha mediante portadoras. Es el medio que permite realizar redes de datos a nivel mundial.

Las microondas se utilizan para realizar la comunicación punto a punto vía satélite entre dos estaciones terrestres, tal como se muestra en la figura A1.31.

A través del aire se realiza también la comunicación a distancias cortas mediante rayos infrarrojos (*Infrared waves*), que son radiaciones electromagnéticas inocuas para el ser humano. El principio básico es similar al de la transmisión por fibra óptica con la diferencia de que el dispositivo convertidor de señales eléctricas en ópticas genera rayos infrarrojos que se transmiten a través del aire, tal como se indica en la figura A1.32. En el receptor se coloca un transistor sensible a los rayos infrarrojos que los convierte en una señal eléctrica. La aplicación más típica de la comunicación por infrarrojos es el telemando de diferentes electrodomésticos, puertas de garajes, etc., pero la ausencia de cables facilita su utilización también para implementar redes de área local, descritas en el apartado A1.3.3.3.

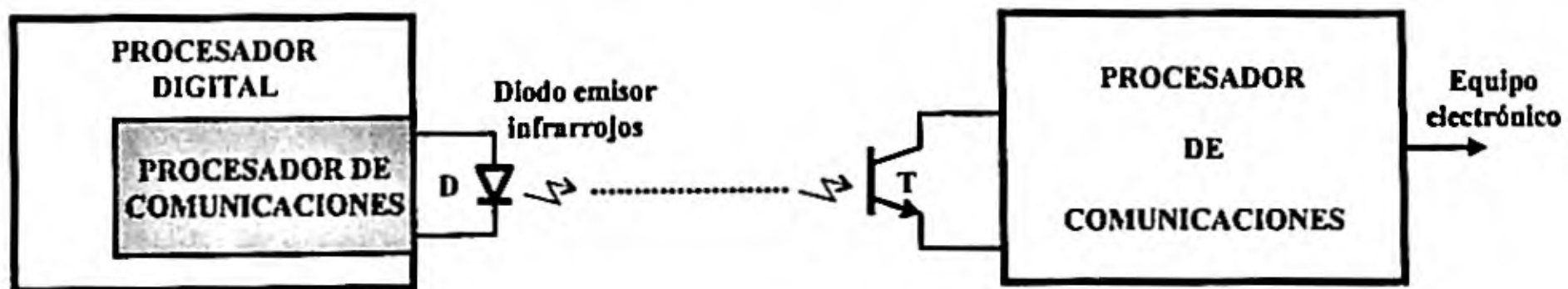


Figura A1.32. Transmisión inalámbrica mediante rayos infrarrojos.

A1.2.6 Características mecánicas y eléctricas de las conexiones

Para hacer realidad el concepto de sistema informático abierto, descrito en el apartado A1.3.1, y poder conectar diferentes periféricos a un procesador digital, es necesario establecer las características de las conexiones mecánicas y eléctricas entre ellos.

Características mecánicas

Los periféricos de un computador se caracterizan por ser sistemas:

- Mecánicamente independientes del computador.
- Fabricados, en general, por empresas distintas de las que fabrican los computadores.

Esto hace que la conexión entre el computador y el periférico tenga que:

- Ser removible.
- Tener elevada fiabilidad.

- Facilitar la identificación de las distintas señales.
- Estar normalizada, para que se puedan intercambiar periféricos de diferentes fabricantes.

Por ello, la conexión entre los medios de transmisión guiados y un procesador digital se realiza mediante conectores, uno macho y otro hembra, que tienen un número de terminales y unas dimensiones establecidas por algún organismo de normalización. En la figura A1.33a se representa como ejemplo el conector utilizado para efectuar la conexión de un computador con un periférico mediante la norma RS-232 (descrita en el apartado A1.2.7.2) y en la A1.33b, el conector RJ-45 utilizado con cables de cuatro pares trenzados.

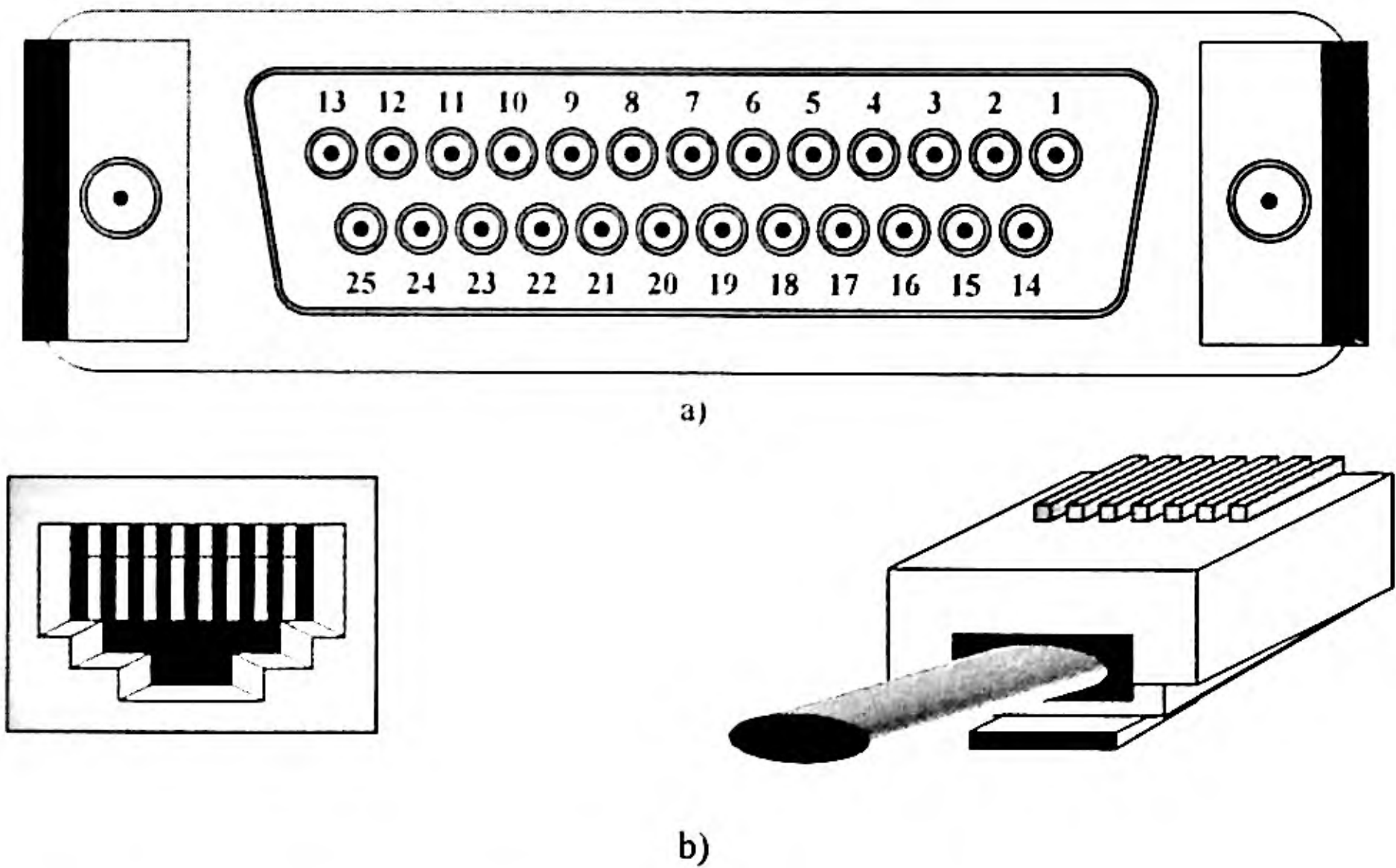


Figura A1.33. Ejemplo de conectores utilizados en diferentes normas de comunicación:
a) RS - 232 de 25 terminales; b) RJ - 45 de cuatro pares trenzados.

Características eléctricas

Las señales eléctricas que soportan la información deben tener unos determinados niveles de tensión y de corriente que se deben establecer teniendo en cuenta:

- La inmunidad ante interferencias electromagnéticas.
- La velocidad de transmisión.
- La protección de los usuarios.

Según el modo de transmisión y el modo de sincronización utilizado se han desarrollado diferentes formas de representar la información digital mediante señales eléctricas. A continuación se indican las más utilizadas en banda base digital que es, a su vez, la más utilizada en las Comunicaciones Industriales descritas en el apartado 9.3 del capítulo 9.

A1.2.7 Normalización de las comunicaciones punto a punto

De lo expuesto en los apartados anteriores se deduce el interés por normalizar la forma de implementar los diferentes conceptos ligados a las conexiones punto a punto.

Las normas deben establecer el modo de transmisión de las señales, el modo de sincronización, el método de detección de errores y las características mecánicas y eléctricas.

A1.2.7.1 Bucle de corriente

El bucle de corriente asigna, como su nombre indica, el cero y el uno lógicos a dos corrientes distintas como son la ausencia de corriente al cero y la corriente de 20 mA al uno. Nació para acoplar a un computador digital los periféricos electromecánicos, como por ejemplo el teletipo.

Se caracteriza por:

- Transmitir la información en banda base digital.
- Tener una velocidad de transmisión baja.
- Se puede realizar con aislamiento galvánico.

En la figura A1.34 se representa su esquema básico. El transmisor actúa sobre un dispositivo conmutador que conecta o desconecta de la línea una fuente de corriente de 20 mA. A su vez el receptor detecta la corriente que pasa por la línea y la convierte en una tensión eléctrica que se acopla al receptor. En la figura A1.35 se representa el esquema típico de un bucle de corriente realizado entre el transmisor de una UART y el receptor de otra, con aislamiento galvánico obtenido mediante un optoacoplador.

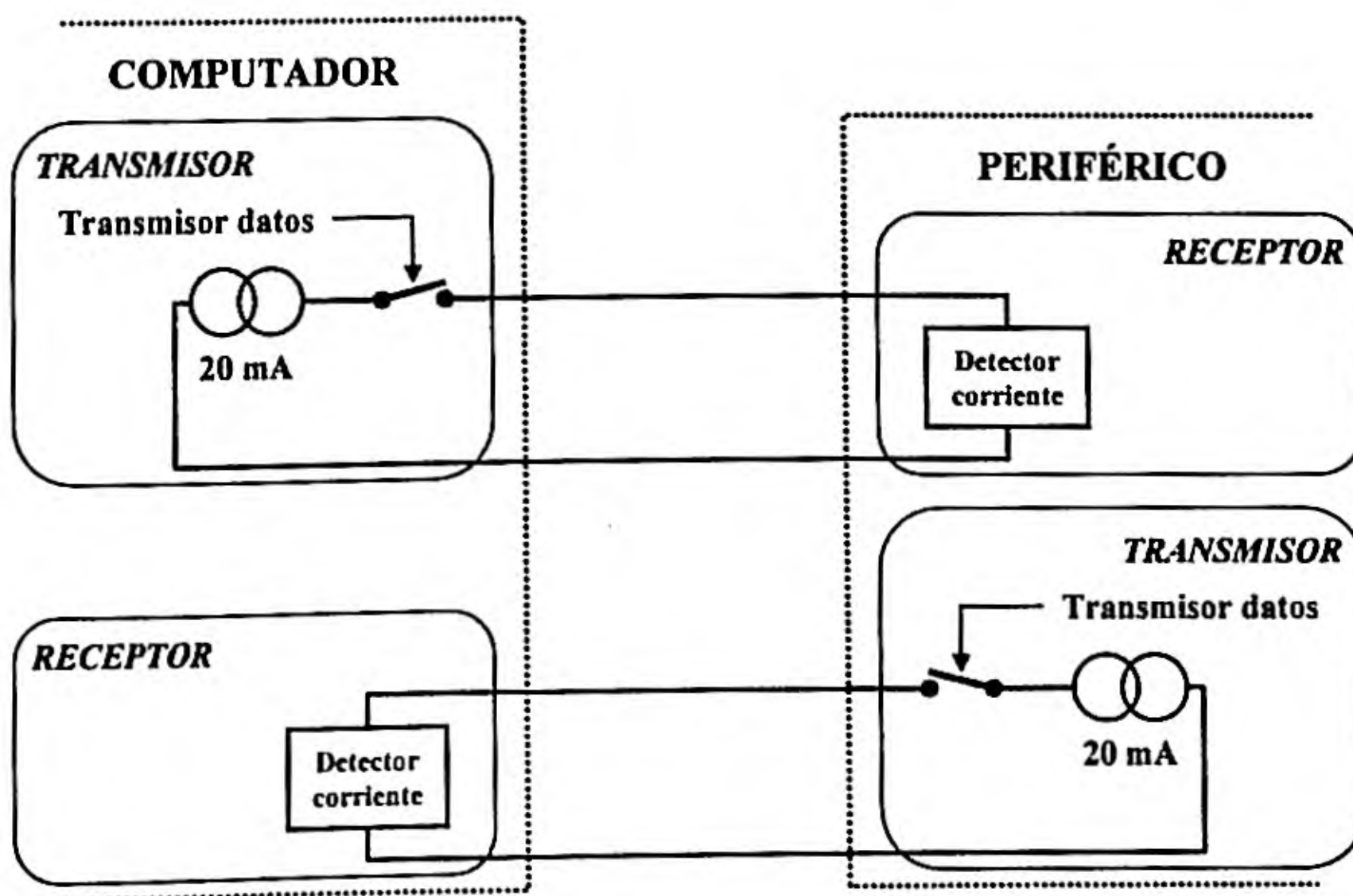


Figura A1.34. Bucle de corriente.

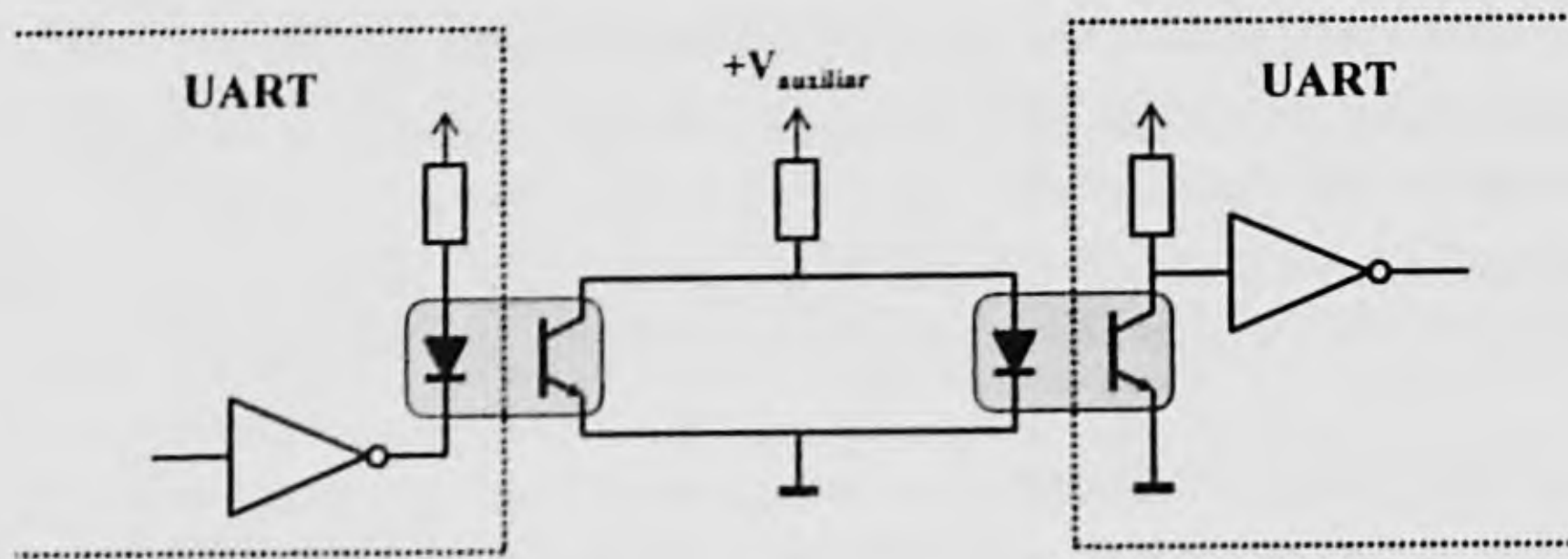


Figura A1.35. Bucle de corriente.

Especificaciones	RS-232	RS-423 A	RS-422	RS-485
<i>Modo de enlace físico</i> (Mode of operation)	Terminación única (Single-ended)	Terminación única (Single-ended)	Diferencial (Differential)	Diferencial (Differential)
<i>Máximo número de transmisores</i> (Drivers)	1	1	1	32
<i>Máximo número de receptores</i> (Receivers)	1	10	10	32
<i>Distancia máxima en pies</i> [Maximum distance (ft)]	50	4000	4000	4000
<i>Velocidad máxima</i> (Maximum data rate)	20 Kbits/sec	100 Kbits/sec	10 Mbits/sec	10 Mbits/sec
<i>Máxima tensión de salida del amplificador de salida</i> (Maximum driver output)	$\pm 25\text{ V}$	$\pm 6\text{ V}$	-0,25 a +6 V	-7 a +12 V
<i>Tensión del amplificador de salida con carga</i> [Driver output (Loaded)]	$\pm 5\text{ V}$	$\pm 3,6\text{ V}$	$\pm 2\text{ V}$	$\pm 1,5\text{ V}$
<i>Tensión del amplificador de salida sin carga</i> [Driver output (Unloaded)]	$\pm 15\text{ V}$	$\pm 6\text{ V}$	$\pm 5\text{ V}$	$\pm 5\text{ V}$
<i>Impedancia de los amplificadores de salida</i> (Driver load impedance)	3 K Ω a 7 K Ω	450 Ω min	100 Ω	54 Ω
<i>Tensión de entrada del receptor</i> (Receiver input voltage)	$\pm 15\text{ V}$	$\pm 12\text{ V}$	$\pm 7\text{ V}$	-7 a +12 V
<i>Sensibilidad de entrada del receptor</i> (Receiver input sensitivity)	$\pm 3\text{ V}$	$\pm 200\text{ mV}$	$\pm 200\text{ mV}$	$\pm 200\text{ mV}$
<i>Impedancia de entrada del receptor</i> (Receiver input resistance)	3 K Ω a 7 K Ω	4 K Ω min	4 K Ω min	12 K Ω min

Tabla A1.2. Resumen de las principales características de las normas RS-232, RS-422, RS-423 y RS-485.

En la actualidad el bucle de corriente no se utiliza en el diseño de las comunicaciones serie punto a punto porque ha sido sustituido por los bucles de tensión correspondientes a las normas RS 232, RS 423, RS 422 y RS 485 cuyos parámetros principales se indican en la tabla A1.2 y se describen sucintamente en sucesivos apartados a continuación. Al lector interesado en ampliar la información se le remite a la bibliografía al final del este apéndice [MILL 00] [ELAH 01].

A1.2.7.2 Norma RS-232

La norma RS-232 define los conceptos asociados con toda comunicación serie, como son el protocolo de la capa de enlace, que establece la forma de sincronizar el transmisor y el receptor, la velocidad de transmisión, las características eléctricas y las características mecánicas. A continuación se describe brevemente cada una de ellas.

Protocolo de la comunicación y velocidad de transmisión

El protocolo de la norma RS-232 define una comunicación que utiliza el modo de sincronización denominado asíncrono que se describe en el apartado A1.2.3.1 (Figura A1.11) en el que se indica que el transmisor y el receptor utilizan señales de reloj independientes de igual frecuencia y, para sincronizarlos, la información se envía precedida de una señal de inicio y otra de fin de transmisión. Fue la más utilizada en los puertos de salida serie de los computadores personales y actualmente ha sido sustituida por la USB.

Los diferentes parámetros de la transmisión son programables, como por ejemplo el número de bits de datos que puede variar entre 5 y 8 y la velocidad de transmisión que se mide en baudios (bits/seg.) y puede variar entre 50 y 19.200 baudios. Por lo tanto, para garantizar una comunicación correcta entre dos procesadores digitales, realizada con la norma RS-232, es imprescindible que ambos utilicen el mismo valor de cada parámetro y la misma velocidad de transmisión en baudios.

Características mecánicas y eléctricas

La comunicación serie mediante la norma RS-232 se puede realizar en banda base digital (transmisión directa) o en banda base analógica modulando una portadora (mediante un MODEM).

Para transmitir a través de un MODEM, la norma RS-232 define un conjunto de 22 señales divididas en señales de datos y señales de control, que se distribuyen en un conector de tipo D de 25 terminales como el representado en la figura A1.33a. No todas las señales de control definidas por la norma son imprescindibles para establecer una comunicación entre dos procesadores digitales y, por ello, en muchas aplicaciones se utiliza un conector macho tipo D de 9 terminales, tal como se indica en la figura A1.36.

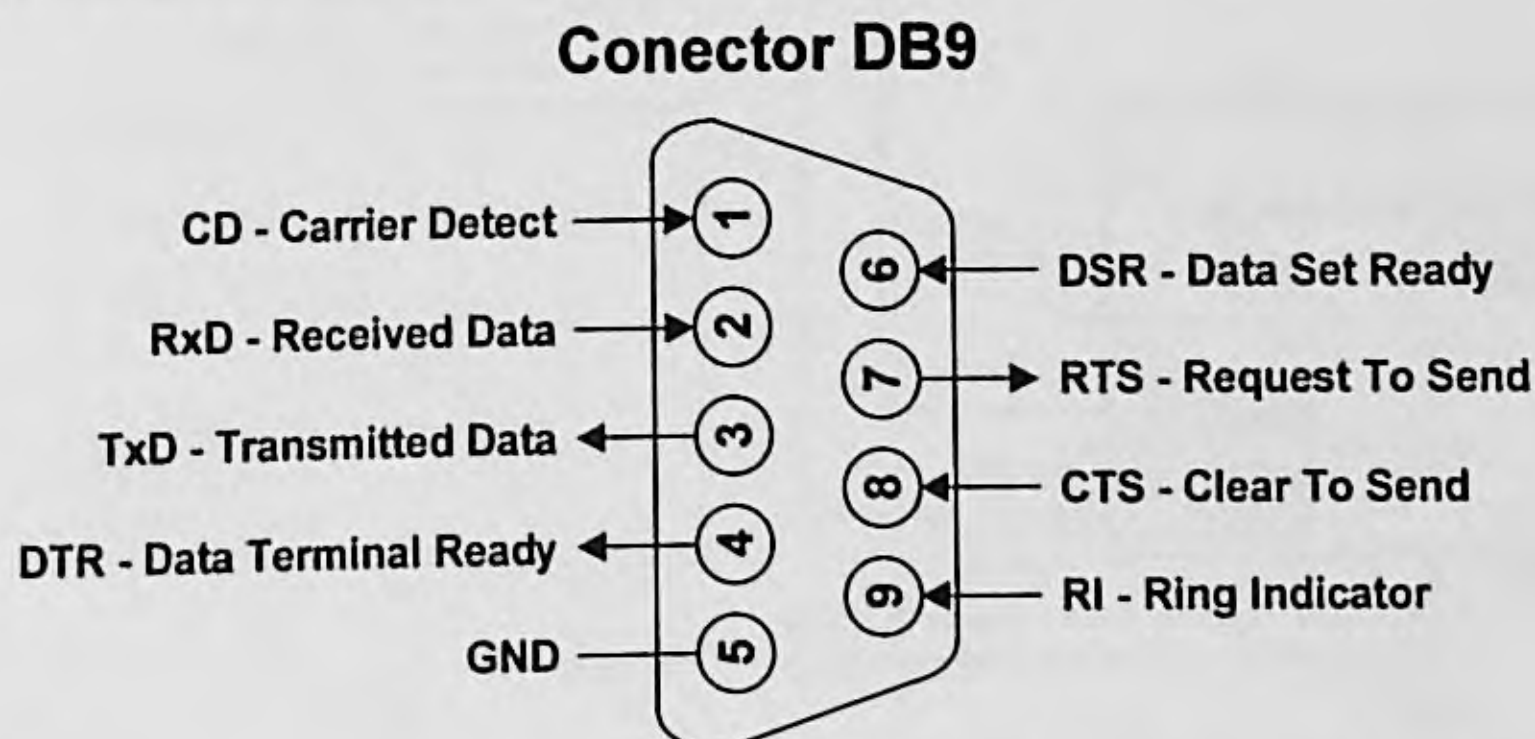


Figura A1.36. Norma RS-232 con doble transmisión simultánea en banda base digital.

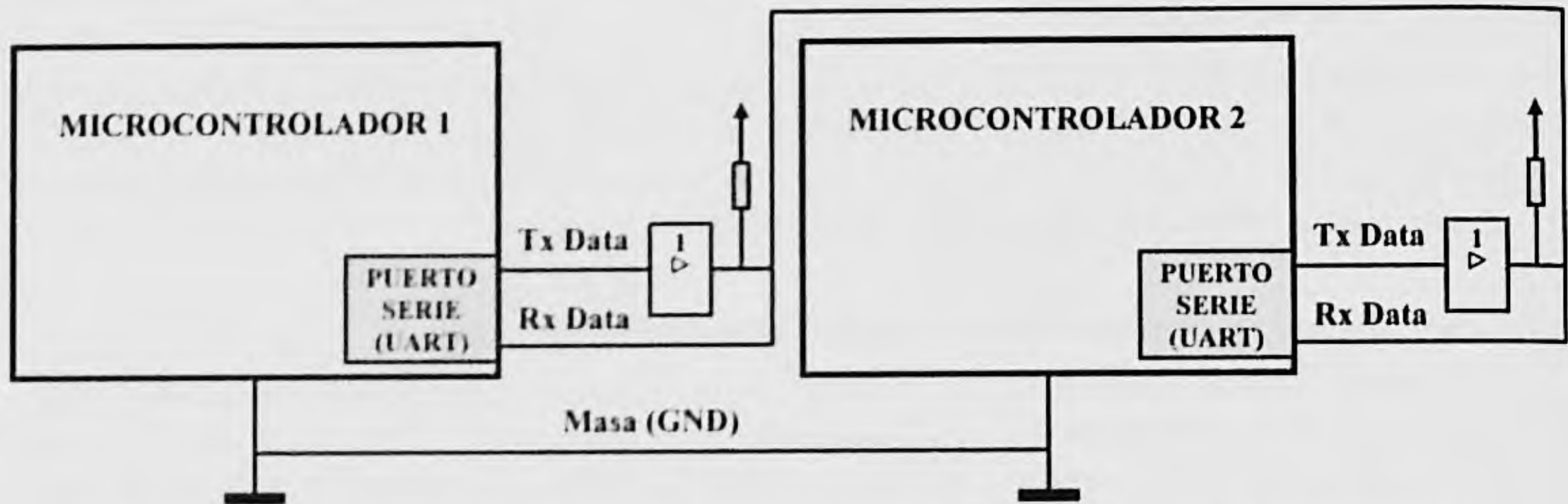


Figura A1.37. Diagrama de bloques de una comunicación asíncrona semidoble simultánea (*Half duplex*) en banda base digital entre dos microcontroladores.

Cuando se transmite en banda base digital se puede establecer una configuración mínima en la que se prescinde de las señales de control. Si se realiza la transmisión en un único sentido simultáneamente (*Half Duplex*) se utilizan sólo dos líneas (la de masa y una de las líneas de datos) tal como se indica en la figura A1.37 que representa el diagrama de bloques de una comunicación asíncrona semidoble en banda base digital entre dos microcontroladores. Si se realiza la transmisión en los dos sentidos simultáneamente (*Full Duplex*) se utilizan tres líneas (la de masa y las dos líneas de datos) tal como se indica en la figura A1.38. Pero si sólo se utilizan las líneas de datos para establecer la comunicación puede ocurrir que un sistema envíe información al otro sin que éste último esté preparado para recibirla, lo cual daría lugar a un error. Para evitar esta situación se pueden utilizar las señales de control CTS/RTS (*Clear To Send*/"*Ready To Send*) y DTR/DSR (*Data Terminal Ready/Data Set Ready*). Existen varias formas de hacerlo, una de las cuales se indica en la citada figura A1.39. En este caso, el sistema 1 antes de transmitir información comprueba a través de su terminal DSR si el estado del terminal DTR del sistema 2 es un uno lógico, lo cual indica que está preparado para recibir la información. Por su parte, cuando el sistema 2 desea transmitir información al sistema 1, comprueba a través de su terminal CTS si el estado del terminal RTS del sistema 1 es un uno lógico, lo cual indica también que está preparado para recibir la información.

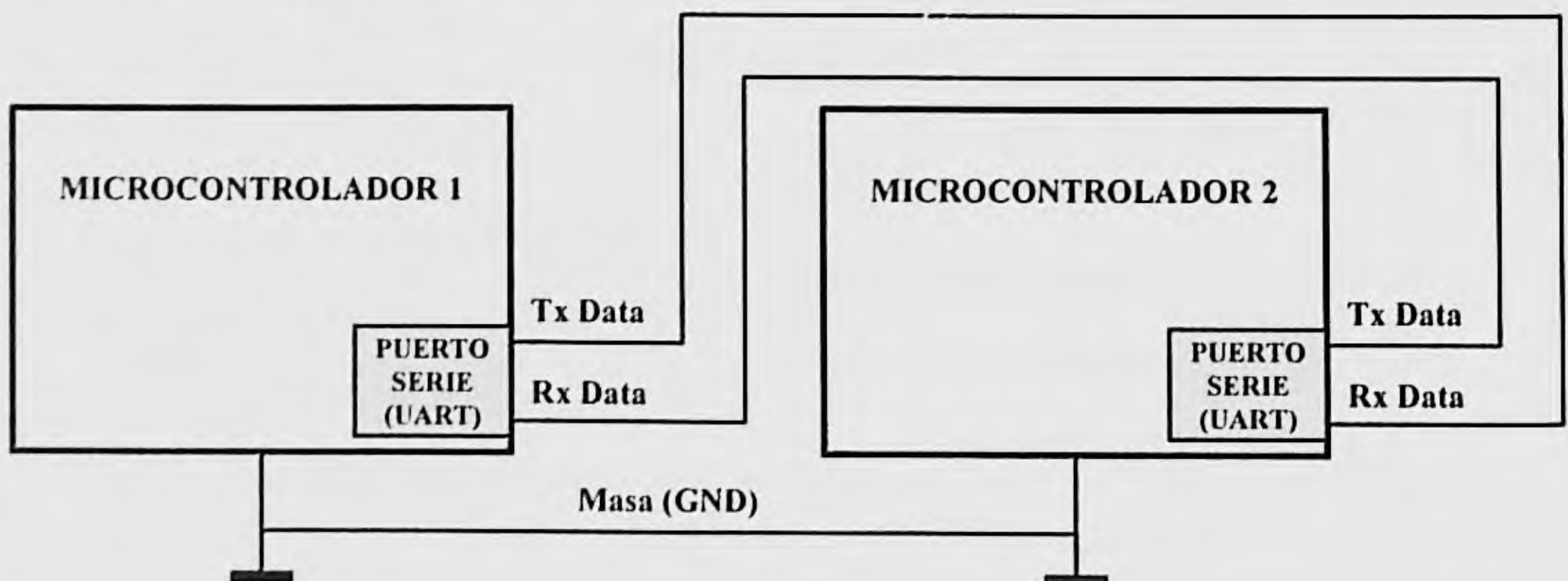


Figura A1.38. Diagrama de bloques de una comunicación asíncrona doble simultánea (*Full duplex*) en banda base digital entre dos microcontroladores.

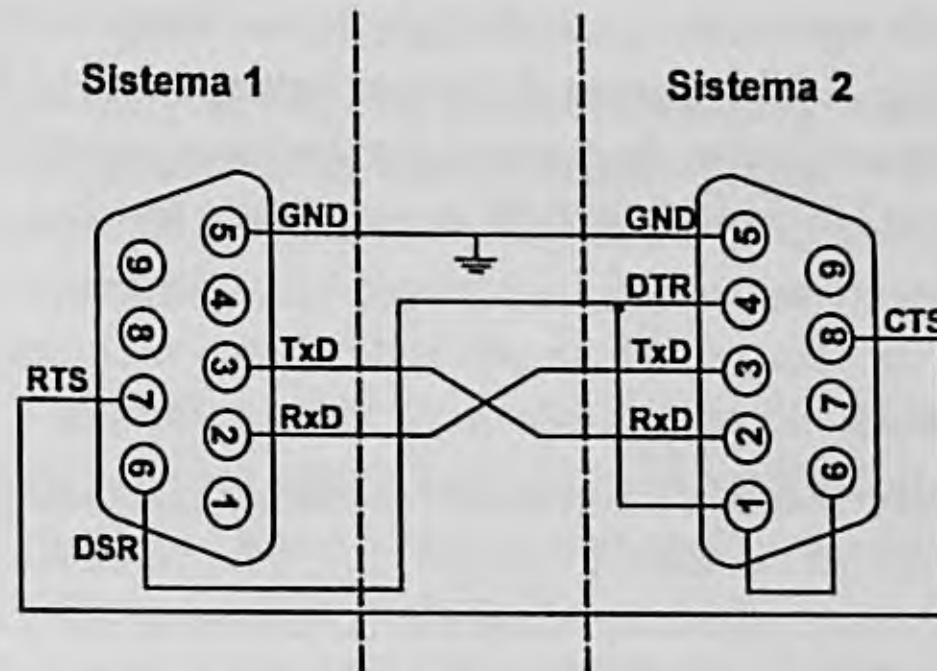


Figura A1.39. Norma RS-232: a) Conector DB9; b) Conexión de dos sistemas.

Tal como se indica en el apartado A1.2.3.1, el proceso de comunicación según la norma RS-232 implica la conversión paralelo-serie en la transmisión (Figura A1.13) y serie-paralelo en la recepción. Además, en la transmisión hay que añadir la información de los bits de inicio y de fin de la transmisión así como el de paridad (opcional) y en la recepción hay que detectar dichos bits para recuperar la información original. Estas tareas pueden ser realizadas mediante un programa, pero esta solución no es en general eficiente porque impide que el procesador pueda realizar otras tareas específicas de la aplicación. Por ello lo más adecuado es utilizar un circuito de interfaz dedicado específicamente a la comunicación, al que se suele dar el nombre genérico de procesador de comunicaciones y en el caso de la transmisión asíncrona recibe, tal como se indica en el apartado A1.2.3.1, la denominación de UART (acrónimo de *Universal Asynchronous Receiver Transmitter*).

En relación con las características eléctricas hay que destacar que la norma RS-232 define un margen de tensión de +3V a +15V para el cero lógico y -3V a -15V para el uno lógico. Esto hace que sea necesario utilizar, tanto en la transmisión como en la recepción, un circuito de adaptación que transforme los niveles de tensión utilizados en los circuitos digitales en los que se acaban de indicar y viceversa, tal como se indica en la figura A1.40. Para un análisis más detallado de la norma RS-232 se remite al lector a la bibliografía [TISC 94].

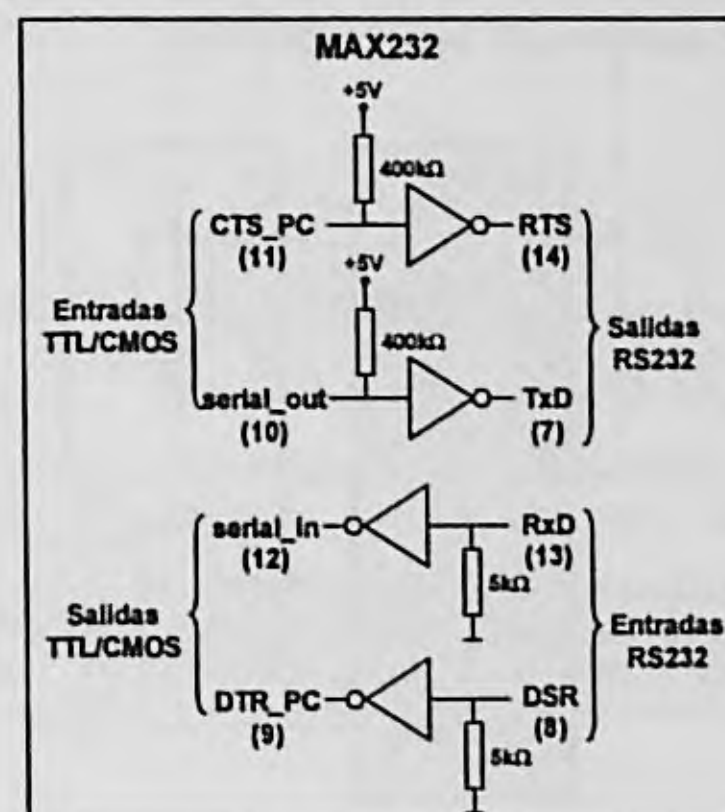
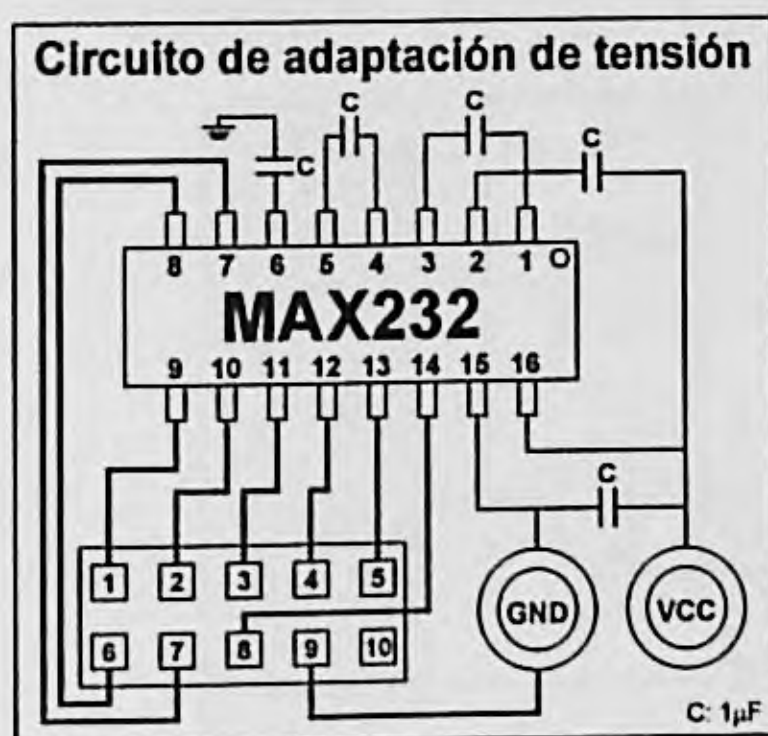


Figura A1.40. Circuito de adaptación de los niveles de tensión de TTL/CMOS a los de la norma RS-232.

Los computadores de aplicación general dispusieron siempre de puertos de salida serie normalizados como el RS-232. El progreso de la microelectrónica ha hecho que se desarrollasen unidades de interfaz normalizadas de mejores prestaciones, entre las que destaca el bus serie universal conocido como USB (acrónimo de *Universal Serial bus*) que se describe en el apartado A1.2.7.6.

A1.2.7.3 Norma RS-422

La norma RS-232 representa la información mediante el estado cero o uno de un terminal y por ello se dice que es de salida única (*Single ended*) (Tabla A1.2). La norma RS-422, que en la figura A1.41 implementa una comunicación asincrónica doble simultánea en banda base digital, utiliza en el transmisor un amplificador (*Driver*) que representa la información mediante la diferencia de nivel entre sus dos salidas (*Balanced circuit or differential*) y en el receptor un amplificador diferencial. Se consigue así una mayor inmunidad al ruido, al igual que una mayor distancia y velocidad de transmisión, que mediante la norma RS-232.

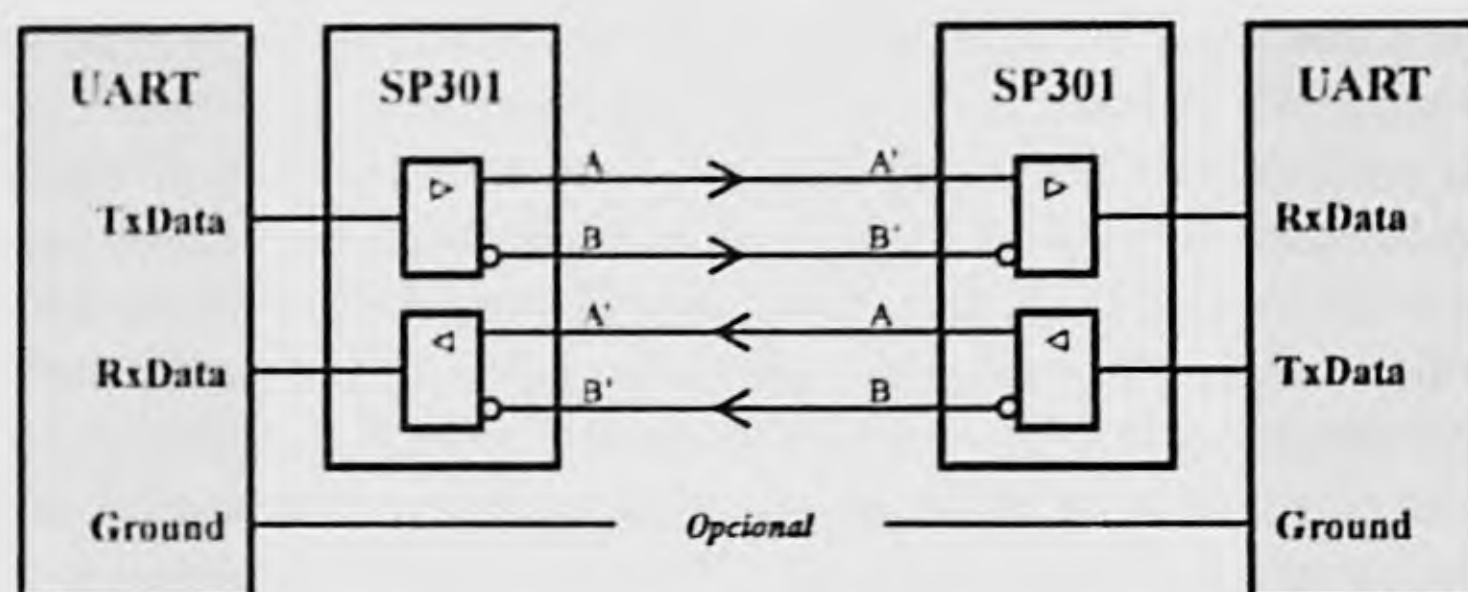


Figura A1.41. Norma RS 422 con doble transmisión simultánea en banda base.

A1.2.7.4 Norma RS-423

La norma RS-423, cuyo circuito básico se representa en la figura A1.42, es una solución intermedia entre la RS-232 y la RS-422, porque utiliza una salida única en el transmisor y una entrada diferencial en el receptor.

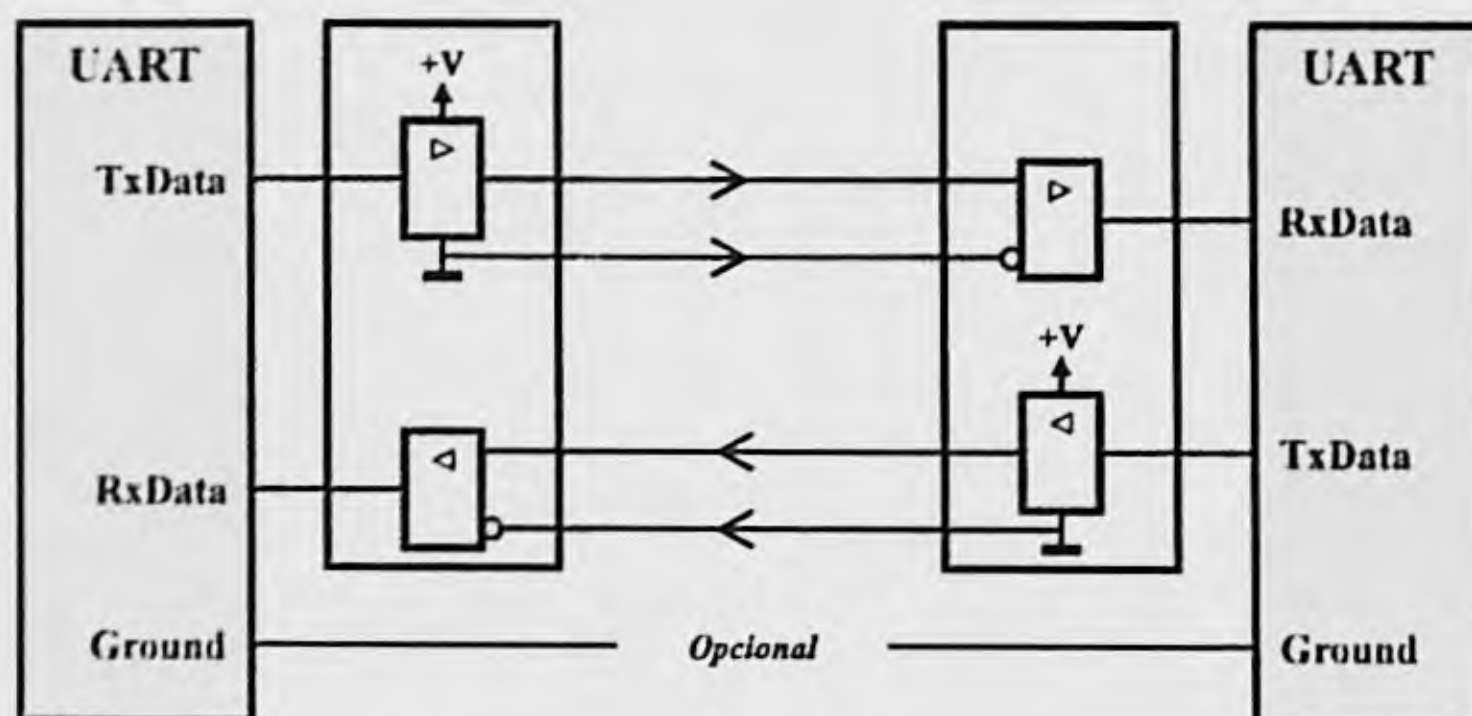


Figura A1.42. Norma RS-423.

A1.2.7.5 Norma RS-485

Las normas RS-232, RS-422 y RS-423 se desarrollaron para realizar conexiones punto a punto entre dos procesadores digitales o entre un procesador digital y un periférico. Pero el desarrollo de la Microelectrónica, que disminuyó el coste de los procesadores digitales, propició el desarrollo de una norma que permitiese conectar a dos hilos un número elevado de procesadores para realizar una comunicación serie entre ellos que recibe, en general, el nombre de bus serie, y en las Comunicaciones Industriales recibe la denominación de bus de campo descrito en el apartado 9.3.2.3 del capítulo 9. Dicha norma es la RS-485, que facilita la implementación de un bus serie con un máximo de 32 transmisores y otros tantos receptores (Tabla A1.2). En la figura A1.43a y b se representa la implementación de una transmisión doble simultánea (*Full duplex*) y semidoble (*Half duplex*) respectivamente, que enlaza un procesador de comunicaciones principal (*Master*) y un conjunto de procesadores subordinados (*Slaves*) cuyo funcionamiento (acceso priorizado) se describe en el apartado A1.3.3.3.

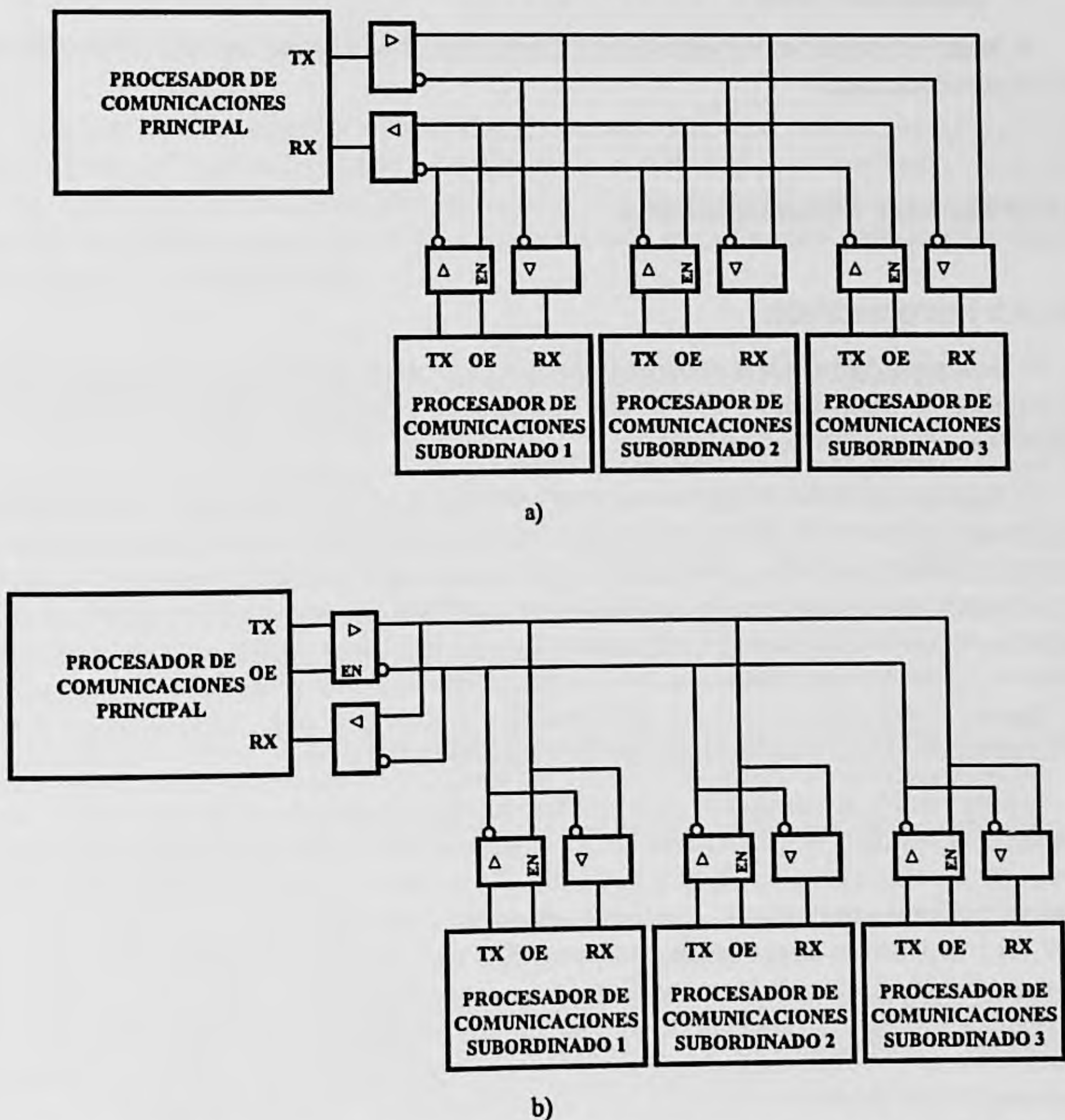


Figura A1.43. Comunicación en banda base digital basada en la norma RS-485: a) Doble simultánea (*Full duplex*); b) Semidoble (*Half duplex*).

A1.2.7.6 Norma USB

El bus serie universal, conocido como USB (acrónimo de *Universal Serial Bus*) sustituye con ventaja a los descritos en los apartados anteriores y sus principales características son:

- Es una conexión punto a punto.
Sustituye a la norma RS-232 y por ello ha sido diseñado para conectar periféricos a un computador principal (*Host*) o conectar éste último a otros procesadores subordinados del mismo. Para dotar a un procesador de varias conexiones USB es necesario utilizar un concentrador (*Hub*) descrito en el apartado A1.3.4.
- Combina la transmisión sincrónica y la asíncrona.
- Permite el envío de paquetes de pequeña y gran longitud.
- Utiliza cuatro hilos.

Al lector interesado en estudiar el bus USB se le remite a la norma: <http://www.usb.org/developers/docs.html>.

A1.3 Redes Informáticas

A1.3.1 Introducción

Se define una red de datos o red informática (*Data Network Communications*) como un conjunto de computadores que pueden intercambiar información entre ellos a través de un sistema de comunicaciones digitales.

En la figura A1.44 se representa esquemáticamente la comunicación entre dos computadores a través de una red de datos o informática. El lector puede observar que esta figura constituye una generalización de la comunicación punto a punto representada en la figura A1.2. En ella se indica que los procesadores de comunicaciones dialogan de forma efectiva a través de la red de datos, pero existe una comunicación entre ellos que se obtiene mediante la eliminación de la información que hay que añadir para que los datos enviados por un procesador de comunicaciones lleguen al otro a través de la red. De forma similar tiene que haber una comunicación entre los usuarios de los dos computadores que intercambian información.

Es innumerable la cantidad de formas diferentes de concebir una red de datos, así como de establecer el diálogo entre procesadores de comunicaciones y entre usuarios de los computadores que se relacionan a través de ella. Unido ello al interés por poder interconectar entre sí equipos de diferentes fabricantes, provocó la necesidad de definir el intercambio de información entre los componentes de un sistema informático con el objetivo de establecer un conjunto de normas que les permitiesen cooperar. Esto hizo que la Organización Internacional de Normalización (ISO) crease el concepto de sistema informático abierto como aquél que es capaz de comunicarse con otros de acuerdo con unas normas determinadas y que estableciese un modelo denominado OSI (acrónimo de *Open System Interconnection*) para definirlo.

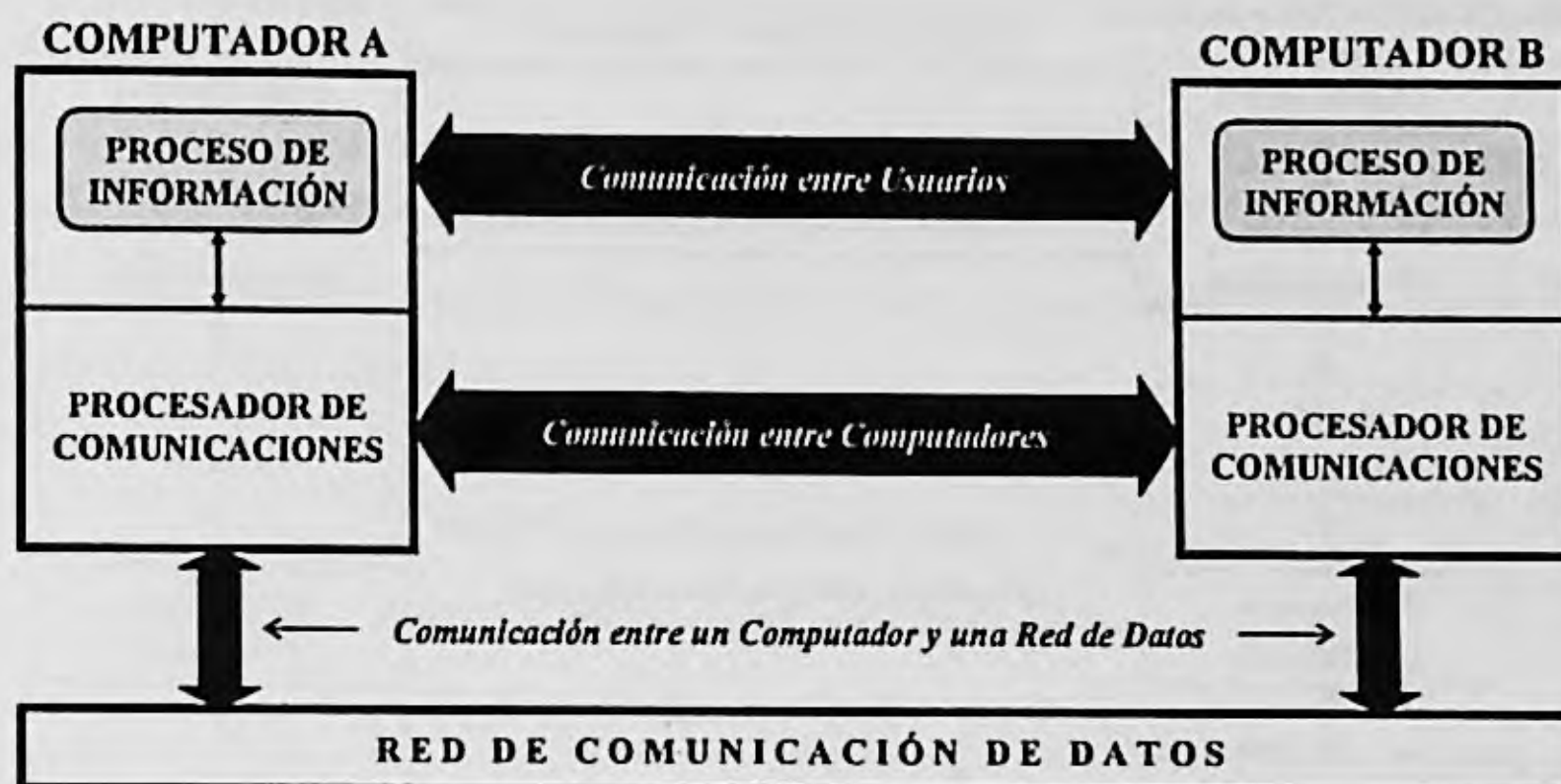


Figura A1.44. Esquema de la comunicación entre dos computadores a través de una red.

Según las características de la red de datos, el modelo OSI (*OSI model*) se aplica de una u otra forma. Además, antes de que se normalizase el modelo OSI se desarrollaron otros, entre los que destaca el DoD (acrónimo de *Department Of Defense*). Por ello a continuación se estudian en primer lugar ambos modelos y a continuación se analizan los principales conceptos específicos de las redes de datos.

A1.3.2 Modelos de interconexión de sistemas informáticos

Tal como se indica en el apartado anterior, los sistemas de comunicación entre computadores deben resolver la transferencia de información entre ellos de forma transparente para los usuarios, de tal manera que el diálogo entre los mismos se realice de forma sencilla, como si dialogasen directamente. Para ello, y dado que tanto los usuarios como los computadores pueden dialogar entre sí de formas muy diferentes, el modelo OSI establece siete niveles de diálogo, que constituyen otros tantos protocolos, y representa una manera estructurada de abordar el problema de la interconexión de los sistemas informáticos. En la figura A1.45 se representa gráficamente.

Aunque el modelo OSI no propone soluciones tecnológicas concretas, aporta procedimientos normalizados de intercambio de información, para lo cual tiene en cuenta tres aspectos:

- El punto de vista del usuario para el cual el carácter distribuido del sistema debe ser transparente.
- El hecho de que el sistema puede estar formado por varios sistemas informáticos independientes entre los que debe ser transportada la información.
- El hecho de que la interconexión puede utilizar una red pública de transmisión de datos u otro medio de transmisión, lo cual implica que la transmisión debe ser considerada como una parte de las funciones que constituyen el transporte.

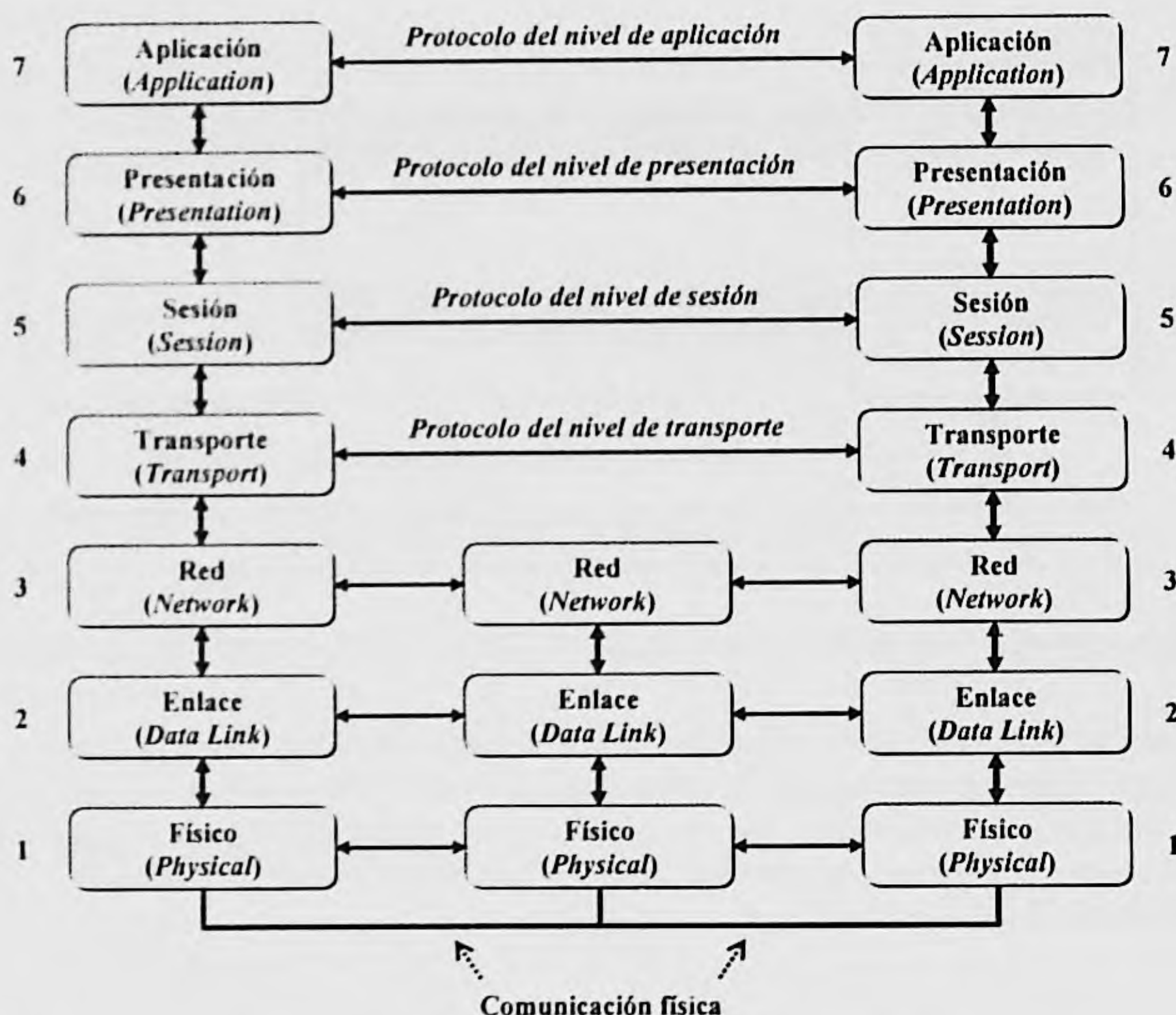


Figura A1.45. Modelo OSI de intercambio de información entre dos computadores a través de una red de datos.

Para ello los siete niveles o capas (*Layers*) se asocian en tres bloques, tal como se indica en la tabla A1.3:

- Un bloque de transmisión orientado a las comunicaciones que comprende la capa física (*Physical Layer*), la capa de enlace (*Data-link Control layer*) y la capa de red (*Network Layer*). Estas capas se suelen implementar mediante elementos físicos y circuitos electrónicos combinados con los programas adecuados. Dichos circuitos reciben el nombre de procesadores de comunicaciones.
- Un bloque de transporte que tiene asignada la capa de su mismo nombre (*Transport Layer*). Esta capa se implementa en general mediante un programa ejecutado por el procesador que intercambia información con otro a través de la red.
- Un bloque de usuarios del bloque de transporte que está formado por la capa de sesión (*Session Layer*), la de presentación (*Presentation Layer*) y la de aplicación (*Application Layer*). Las dos primeras, junto con la de transporte, están orientadas al sistema de proceso de datos, mientras que la de aplicación es la que resuelve la comunicación entre los usuarios que deben ser expertos en la aplicación pero no tienen que tener conocimientos de telemática. Estas tres capas también se implementan mediante otros tantos programas.

La transferencia de información se realiza desde la capa 7 hacia la 1 en el nodo emisor y de la 1 a la 7 en el receptor.

La relación entre las diferentes capas (*Interface between layers*) se basa en los siguientes principios, que se indican gráficamente en la figura A1.46:

- Las capas del mismo nivel “i” cooperan entre sí, y solo entre sí, de acuerdo con un determinado protocolo “i”.

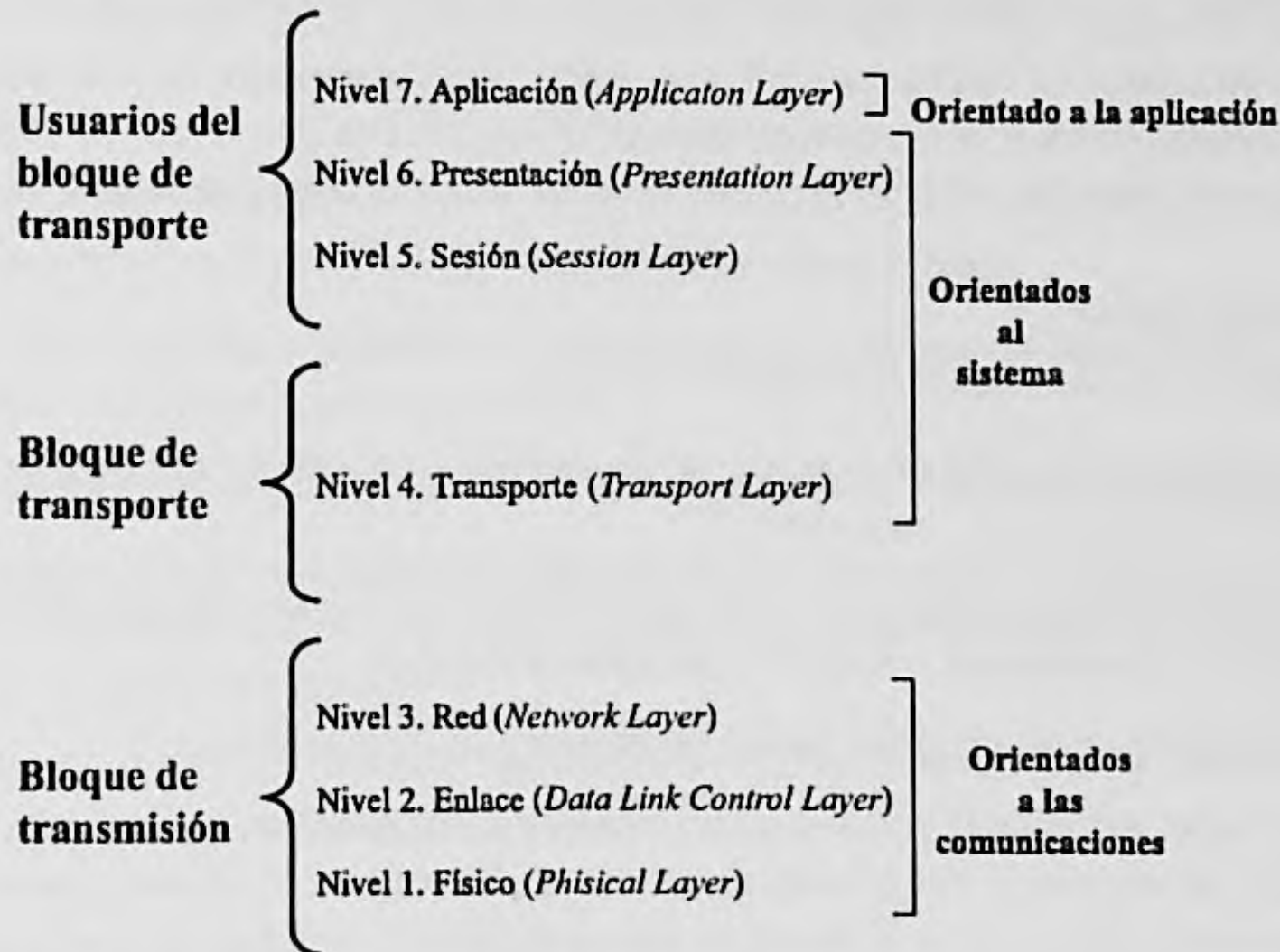


Tabla A1.3. Esquema básico del modelo de referencia OSI.

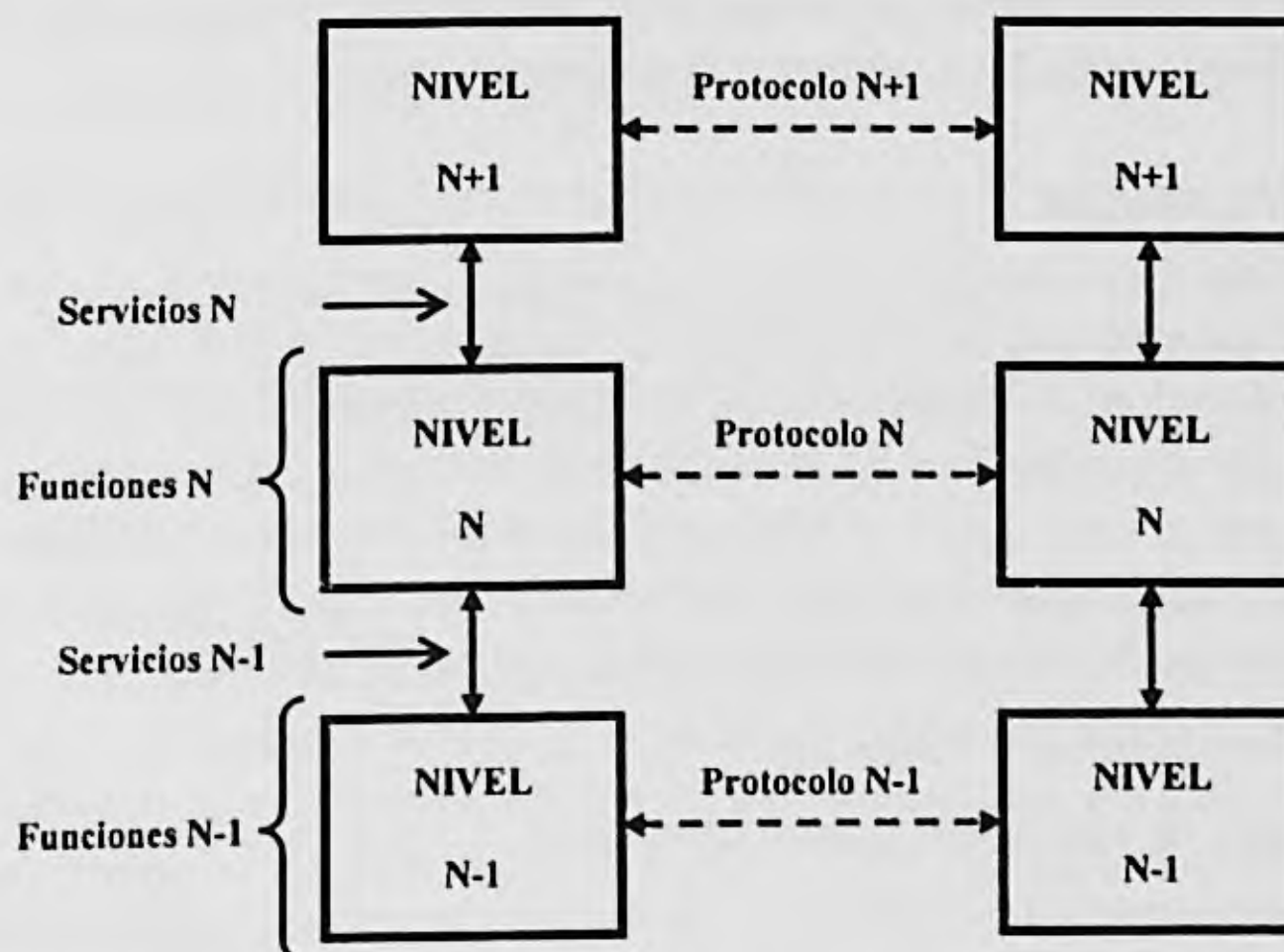


Figura A1.46. Esquema de configuración del modelo de referencia para interconexión de sistemas abiertos [*Open Systems Interconnection (OSI)*].

- Cada una de las capas del mismo nivel “i” utiliza los servicios proporcionados por la capa adyacente que le transfiere información sin necesidad de conocer la estructura de las capas que preceden a esta última.
- Las capas del nivel “i” realizan unas determinadas funciones “i”, para lo cual utilizan los servicios de las capas “i - 1”, y proporcionan a su vez los servicios a las capas de nivel “i + 1”.
- Las capas 7 y 1 interactúan con el usuario y con el medio físico, respectivamente.

Los datos utilizados en cada capa reciben la denominación genérica de unidades de datos del protocolo correspondiente y se les conoce como PDU (acrónimo de *Protocol Data Unit*). A continuación se indican las principales características de cada una de las capas.

Capa 1 o nivel físico

Esta capa (*Physical Layer*) se caracteriza por:

- Garantiza el transporte de la información (en bits) a través del medio físico de transmisión (*Transmission medium*).
- Ha sido la primera en normalizarse, mediante el establecimiento del adecuado interfaz mecánico (conectores), eléctrico y de señalización.
- Incluye reglas para elegir el medio de transmisión y el cableado.
- Es la responsable de la codificación/decodificación de los datos mediante señales eléctricas y de establecer los niveles eléctricos de las señales utilizadas, las velocidades de transmisión, el tamaño y la forma de los conectores utilizados, las funciones y procedimientos para establecer y desactivar las conexiones físicas, etc.

Por tanto esta capa establece el modo de transmisión de las señales, el modo de sincronización y las características eléctricas y mecánicas descritas en los apartados A1.2.1, A1.2.3 y A1.2.6 respectivamente, para las comunicaciones punto a punto.

Capa 2 o nivel de enlace

La capa de enlace (*Data-link layer* o *Data-link Control layer*) proporciona las funciones precisas para establecer, mantener y liberar conexiones fiables de enlace de datos entre los nodos de la red, para lo cual define la forma en que se produce la transmisión es decir el tamaño, la estructura, el contenido y la secuencia de los bloques de información o tramas (*Frames*). Además garantiza la transmisión sin errores de las tramas, para lo cual utiliza los métodos de detección de errores descritos en el apartado A1.2.4, que le permiten detectar y corregir los errores introducidos por la falta de fiabilidad de los circuitos de datos.

Se divide en dos subcapas o subniveles:

- **La subcapa de control de acceso al medio**

Se la conoce como MAC (acrónimo de *Medium Access Control*) y soporta las funciones de gestión del direccionamiento de los mensajes, así como el acceso a la red.

- **La subcapa de control del enlace lógico**

Se la conoce como LCC (acrónimo de *Logical Link Control*) y establece el formato de las tramas (*Frames*), gestiona la detección de errores, los métodos de pe-

tición de reenvío automático, los avisos de recepción de mensajes (*Messages acknowledgment*) y los procesos de relación entre usuarios (*Handshakes*).

Capa 3 o nivel de red

La capa de red (*Network Layer*) realiza funciones de conmutación y encaminamiento, es decir, de selección del camino físico de los bloques de datos o paquetes, a través de los posibles nodos intermedios de la red. Se encarga de establecer, mantener y liberar conexiones de red, para lo cual proporciona los procedimientos precisos para el intercambio de datos entre un origen y un destino.

Capa 4 o nivel de transporte

La capa de transporte (*Transport Layer*) tiene como misión:

- Controlar el transporte fiable de información (en bloques denominados mensajes) de extremo a extremo a través de la red.
- Facilitar un enlace transparente (independiente de la red) entre el usuario y la red.
- Garantizar que los mensajes procedentes de los usuarios de la red llegan correctamente a sus destinatarios. Para ello, debe ensamblar los mensajes que hayan sido divididos en paquetes para su transporte.
- Corregir, en su caso, posibles deficiencias de niveles inferiores, recuperar datos perdidos, reinicializar la comunicación, etc.

Capa 5 o nivel de sesión

La capa de sesión (*Session Layer*) gestiona el establecimiento, sincronización y control del diálogo entre procesos de aplicación remotos, así como su finalización, para lo cual utiliza la transacción como unidad de información intercambiada.

Las normas establecidas en ella incluyen:

- El reconocimiento de palabras de paso (*Passwords*).
- El proceso de conexión/desconexión de usuarios (*Log on/Log off*).
- Las funciones de gestión de red (*Network management*).

Capa 6 o nivel de presentación

La capa de presentación (*Presentation Layer*) es la responsable de la interpretación y presentación de la información intercambiada entre las aplicaciones. Garantiza la compatibilidad sintáctica entre los sistemas que se comunican, es decir la estructura, lenguaje y formato de los datos. Incluye la transferencia de ficheros, la compresión de los datos (*Data compression*), la encriptación de los datos (*Data encryption*) la seguridad de la red (*Network security*), las funciones de formato, etc.

Capa 7 o nivel de aplicación

La capa de aplicación (*Application Layer*) se caracteriza por:

- Constituye el punto de acceso de todos los usuarios al modelo a través de aplicaciones que son procesos realizados mediante la ejecución de un determinado programa (*software*).
- Define la semántica de intercambio de datos entre procesos de aplicación que se comunican para la resolución de una tarea.
- Incluye procesos que dan servicio al usuario y están bajo su control directo, como por ejemplo la compartición de ficheros, el establecimiento de colas de trabajo de impresión, el correo electrónico, la gestión de bases de datos, etc.

En esta capa se pueden considerar dos tipos de aplicaciones:

- **Aplicaciones en modo conectado o en línea (*On line*)**

Son aplicaciones en las que las dos entidades que dialogan están activas y participan simultáneamente en la comunicación. Son ejemplos característicos de aplicación en línea la transferencia de ficheros (*File transfer*) y la charla electrónica (*Chat*).

- **Aplicaciones en modo no conectado o fuera de línea (*Off line*)**

Son aplicaciones en las que una de las dos entidades puede no estar activa y es sustituida por un buzón electrónico. El ejemplo más típico de este tipo de aplicación es la mensajería no instantánea (SMS) o el correo electrónico (*Electronic mail*).

A partir de su creación en 1979, en el modelo OSI se han basado numerosas normas de redes de comunicaciones desarrolladas por diversos organismos, como por ejemplo las redes de control o buses de campo, que se describen en el apartado 9.3.2.3 del capítulo 9. Pero antes de esa fecha ya existía el concepto de división en capas del protocolo (*Protocol layering*) de diálogo entre dos equipos informáticos y se desarrollaron otros modelos diferentes de OSI. El ejemplo más significativo de modelo de capas diferente de OSI es el modelo en el que se basa el conjunto de protocolos conocido como TCP/IP utilizado por la red Internet descrita en el apartado A1.3.3.4.2. Dicho modelo, que recibe el nombre de DoD porque su desarrollo estuvo ligado al Departamento de Defensa de E.E.U.U., posee sólo cuatro capas tal como se indica en la figura A1.47.

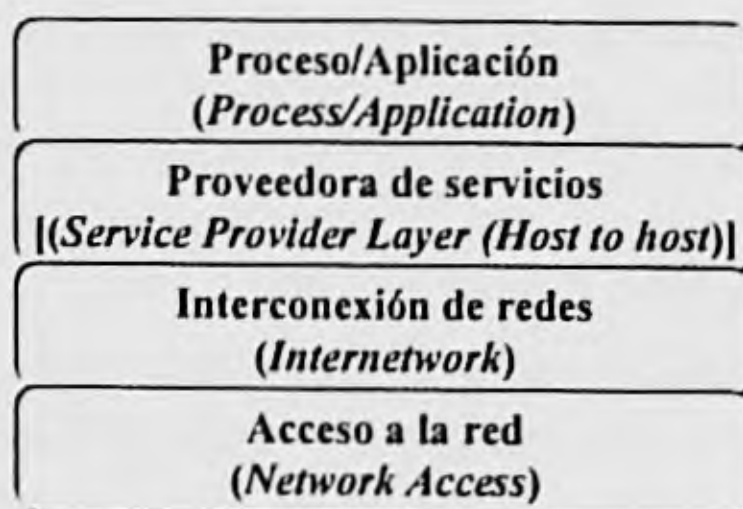


Figura A1.47. Modelo DoD de intercambio de información entre dos computadores a través de una red de datos.

La capa inferior del modelo DoD es la de acceso a red (*Network Access Layer*) que incluye los elementos físicos de conexión tales como los cables, los transceptores (*Transceivers*), las interfaces de red (*Network boards*), los protocolos de enlace (*Link proto-*

cols), los protocolos de acceso de las redes de área local (*LAN access protocols*), como por ejemplo el protocolo CSMA/CD de la red Ethernet y el paso de testigo de la red en anillo (*Token ring*) que se describen en el apartado A1.3.3.2. Los servicios proporcionados por la capa de acceso a la red los utiliza la capa de red (*Internetwork layer*), que se responsabiliza de darle una dirección lógica a la dirección física proporcionada por la primera. La capa de red se encarga también del enrutado de los paquetes de información entre los diversos computadores conectados a la red y sus servicios son utilizados por la capa de enlace entre computadores (*Host to Host layer*) que implementa el diálogo entre computadores y se encarga de gestionar el reenvío de los paquetes en los que se detecta la presencia de errores. La capa de aplicación (*Process/Application layer*) utiliza los servicios de la capa inferior de diálogo entre computadores y en ella se incluyen los distintos programas utilizados por el usuario, como por ejemplo el de correo electrónico, denominado SMTP (acrónimo de *Simple Mail Transfer Protocol*) y el de transferencia de ficheros denominado FTP (acrónimo de *File Transfer Protocol*).

En la figura A1.48 se representa la relación entre el modelo DoD y el OSI y el lector puede comprobar su similitud. La capa de acceso a red (*Network Access layer*) de DoD corresponde a las capas de nivel físico y nivel de enlace del modelo OSI. La capa de red de DoD (*Internet layer*) corresponde a la capa de red de OSI y la de diálogo entre computadores (*Host to Host layer*) de DoD a la de transporte (*Transport layer*) de OSI. Finalmente la capa de aplicación de DoD equivale a las de sesión, presentación y aplicación de OSI.

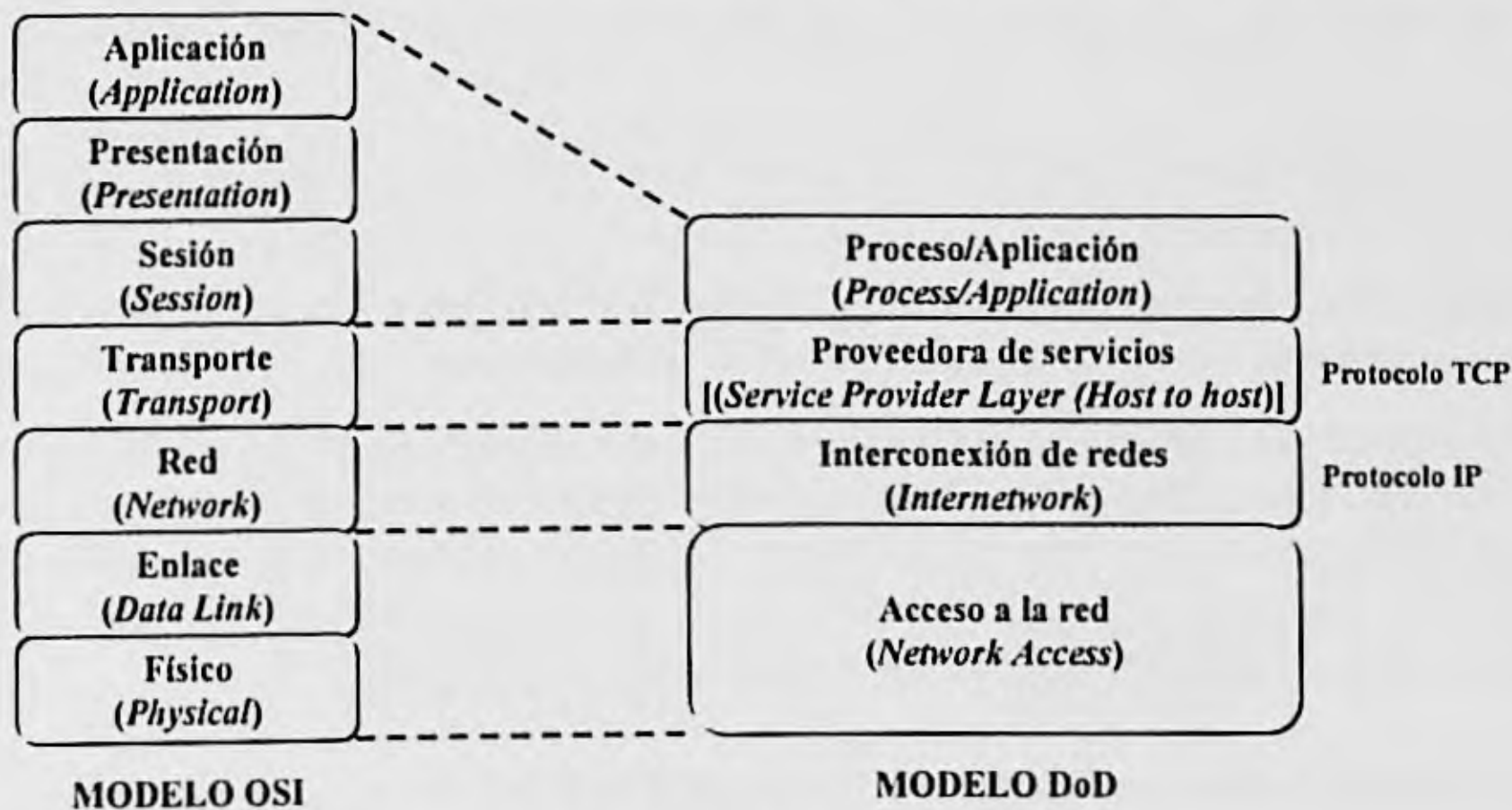


Figura A1.48. Relación entre el modelo DoD y el OSI.

Como resumen del análisis de los modelos OSI y DoD de interconexión de sistemas informáticos, se puede decir que del mismo se deduce que la normalización de las comunicaciones en una red informática no es una tarea sencilla porque existen infinitas maneras de implementarlas que no presentan ventajas sustanciales unas con respecto a otras. Por ello, las decisiones se basan, en muchos casos, en intereses y además están influenciados por la cuota de mercado que tienen las empresas, porque, cuanto más alta es esta última, mayor es la disponibilidad de los recursos necesarios para llevar a cabo las tareas de investigación aplicada y desarrollo que una empresa tiene que realizar para lograr innovaciones en este campo de la tecnología.

A1.3.3 Conceptos específicos de las redes de datos

Tal como se indica en el apartado A1.1, en las redes de datos, que son conexiones multi-punto, hay que considerar todos los conceptos ligados a las conexiones punto a punto. Por ello existen redes de datos en las que la transmisión se hace en banda base y otras en las que se hace en banda ancha. Además, las redes de datos pueden teóricamente utilizar el modo de sincronización asíncrono, el síncrono o el isócrono, así como cualquiera de los medios de transmisión indicados en el apartado A1.2.5.

Pero además hay dos conceptos específicos de las redes de datos, que también deben ser tenidos en cuenta:

- **La topología de la red**

La topología de una red de comunicaciones (*Communication network topology*) establece la forma en que se interconectan los diferentes procesadores que la constituyen.

De acuerdo con la topología, las redes de comunicaciones se dividen en diferentes tipos que se analizan a continuación en el apartado A1.3.3.1.

- **La extensión de la red**

Según su extensión, las redes informáticas se pueden clasificar en:

- Redes de área local, conocidas como LAN (acrónimo de *Local Area Network*).
- Redes metropolitanas, conocidas como MAN (acrónimo de *Metropolitan Area Network*).
- Redes de área extensa o de área amplia, conocidas como WAN (acrónimo de *Wide Area Network*).

La extensión no es un parámetro independiente de la topología, porque de ella depende el tipo de topología más adecuado para una aplicación determinada.

De lo expuesto se deduce que el estudio de las redes de datos no es tarea fácil porque son muchos los conceptos interrelacionados a tener en cuenta. Por ello primero se analizan las topologías utilizables para realizar una red de datos y a continuación se describen las redes de datos de acuerdo con su extensión y la forma en la que la misma ha influido en el pasado e influye actualmente en la elección de la topología más adecuada.

A1.3.3.1 Clasificación de las redes de datos de acuerdo con la topología

A1.3.3.1.1 Topología de canales independientes

La topología de canales de comunicación independientes, también llamada topología de malla (*Mesh topology*), se caracteriza porque cada pareja de procesadores se conecta a través de un canal independiente tal como se indica en la figura A1.49. Se trata, por lo tanto, de un conjunto de comunicaciones punto a punto y por ello es discutible su consideración como una auténtica red. Tuvo un cierto interés en la década de 1980 como evolución natural de las comunicaciones punto a punto, pero el elevado número de canales que necesita cuando el número de procesadores supera la media docena ha hecho que haya sido sustituida por las topologías en las que los diferentes elementos de proceso de datos comparten canales de comunicación entre ellos.

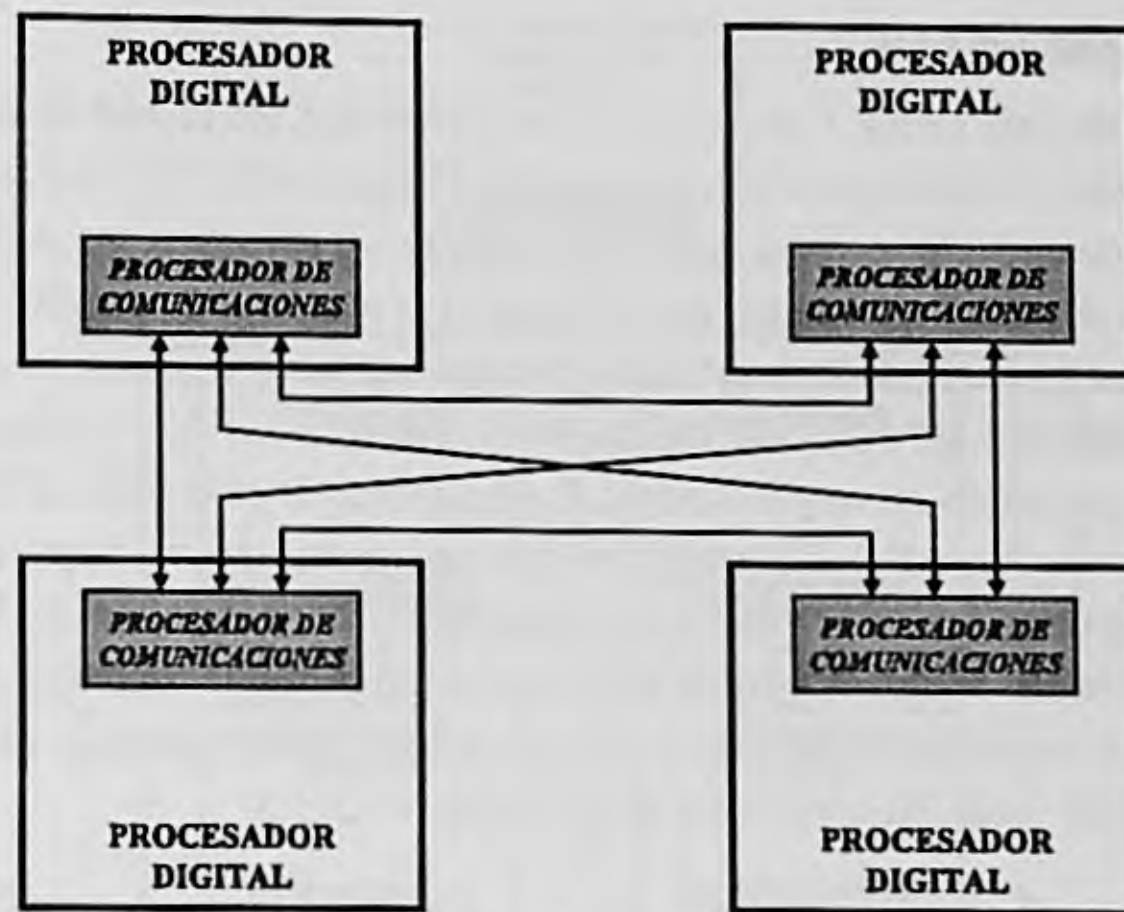


Figura A1.49. Topología de red de canales independientes (malla).

A1.3.3.1.2 Topología en estrella

La topología en estrella (*Star topology*) se caracteriza por la utilización de un elemento central a través del cual pasan indefectiblemente todas las transferencias de información entre los distintos procesadores de la red (Figura A1.50). Dicho elemento recibe distintas denominaciones según se ponga el énfasis en su capacidad de concentración de las comunicaciones (Concentrador), de distribución de la información (Distribuidor) o de repetición de la información que un computador envía a otro (Repetidor). Se le solía conocer por su denominación inglesa de *Hub*, pero el avance de la Microelectrónica permitió que se elevase su complejidad y como consecuencia de ello sus prestaciones, y por ello se le denomina conmutador (*Switch*). El conmutador es un procesador que posee los recursos físicos y el programa adecuados para facilitar las transferencias entre los distintos procesadores de la red.

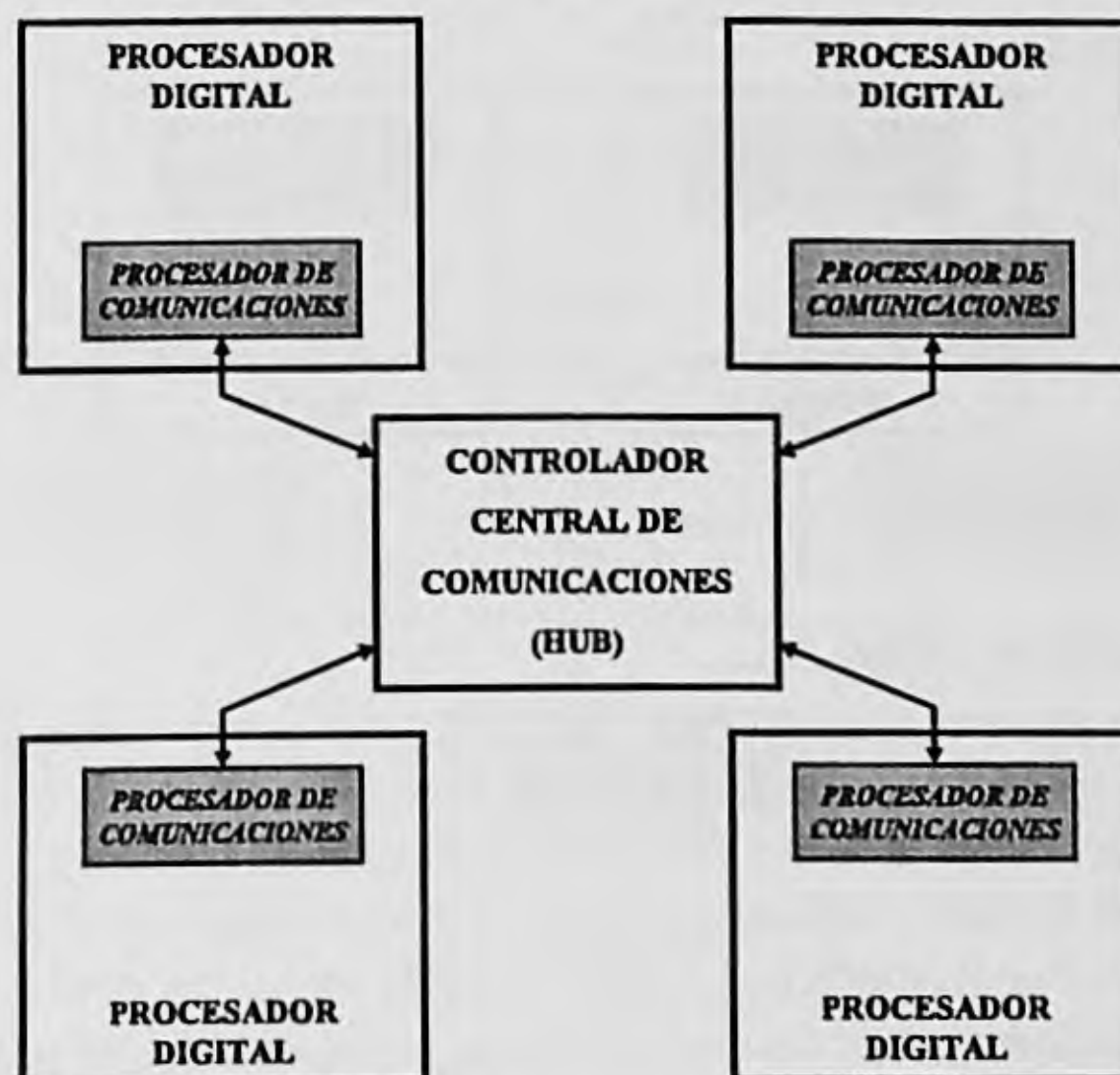


Figura A1.50. Topología de red en estrella.

A1.3.3.1.3 Topología en bus

En la topología de bus (*Bus topology*) los elementos de la red se conectan a un único canal de comunicación (*Backbone*) común a todos (Figura A1.51). Mediante ella un procesador de comunicaciones puede realizar una transmisión de información a un único procesador (*Unicast*), a varios al mismo tiempo (*Multicast*) o a todos los demás conectados a la red (*Broadcast*). Por ser el canal de comunicación único, es necesario resolver el acceso al medio de los diferentes elementos, tal como se indica en el apartado A1.3.3.3.1 dedicado al análisis de los mecanismos que permiten su implantación. Esta topología es la más utilizada en las redes de área local de aplicación general o informáticas que se describen en el apartado A1.3.5 y en las redes de control, denominadas buses de campo (apartado 9.3.2.3 del capítulo 9), que se utilizan para llevar a cabo la fabricación integrada por computador (CIM) descrita en el apartado 9.2.5 del capítulo 9. En la actualidad se combina con la topología estrella para mejorar las prestaciones de las redes de área local, descritas en el apartado A1.3.3.3.

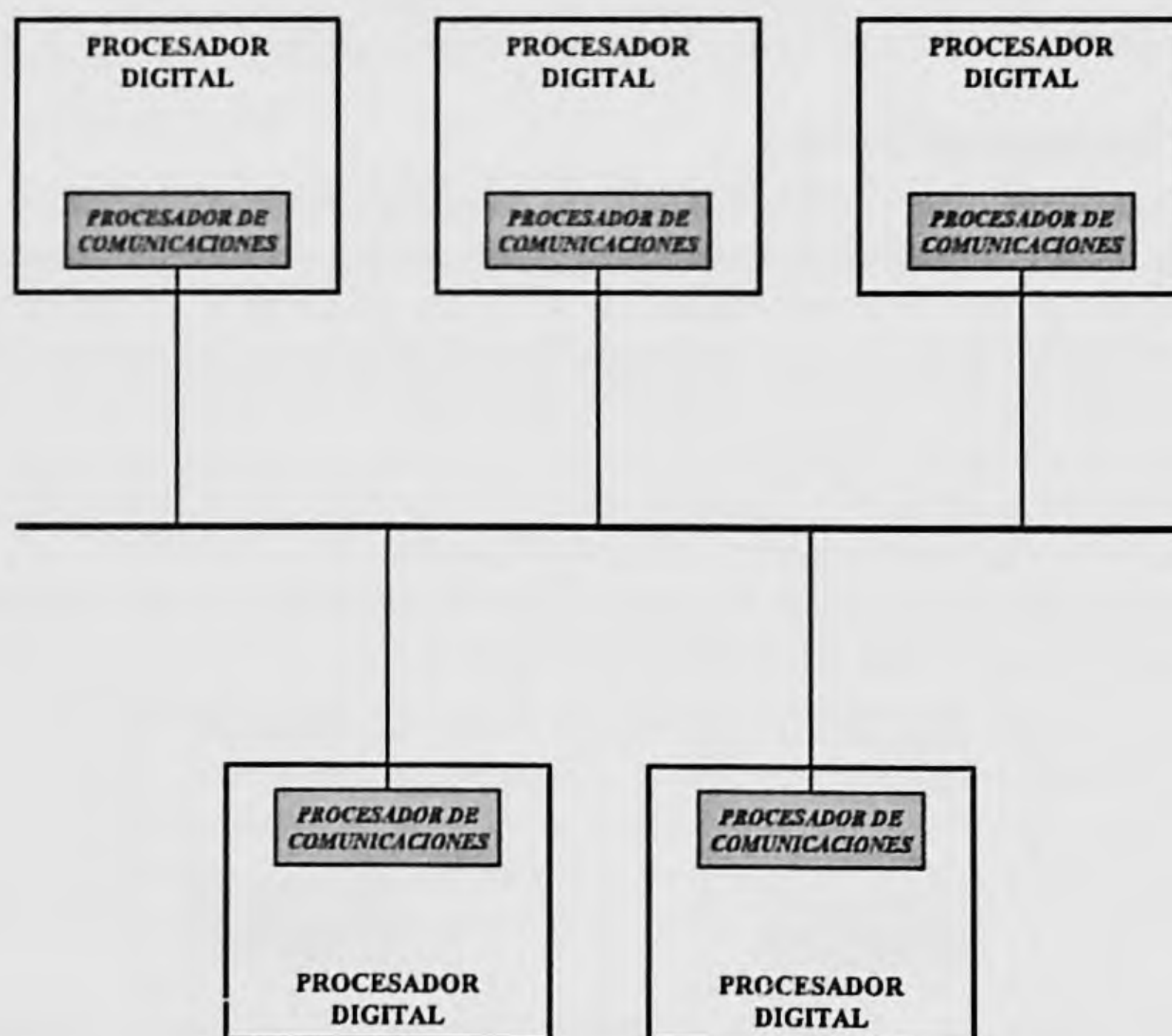


Figura A1.51. Topología de red en bus.

A1.3.3.1.4 Topología en árbol

La denominación de topología en árbol (*Tree topology*) presenta varias acepciones según los autores. Por ejemplo, Stallings [STAL 07] denomina árbol a la topología de bus en la que se conectan varias ramas a un terminal único (*Headend*), tal como se indica en la figura A1.52. Para disminuir la carga conectada al punto común se asigna en ocasiones un repetidor a cada una de las ramas. Por el contrario, Elahi [ELAH 01] denomina árbol a la arquitectura en la que los computadores se asocian con concentradores o repetidores (*Hubs*) conectados entre sí, tal como se indica en la figura A1.53.

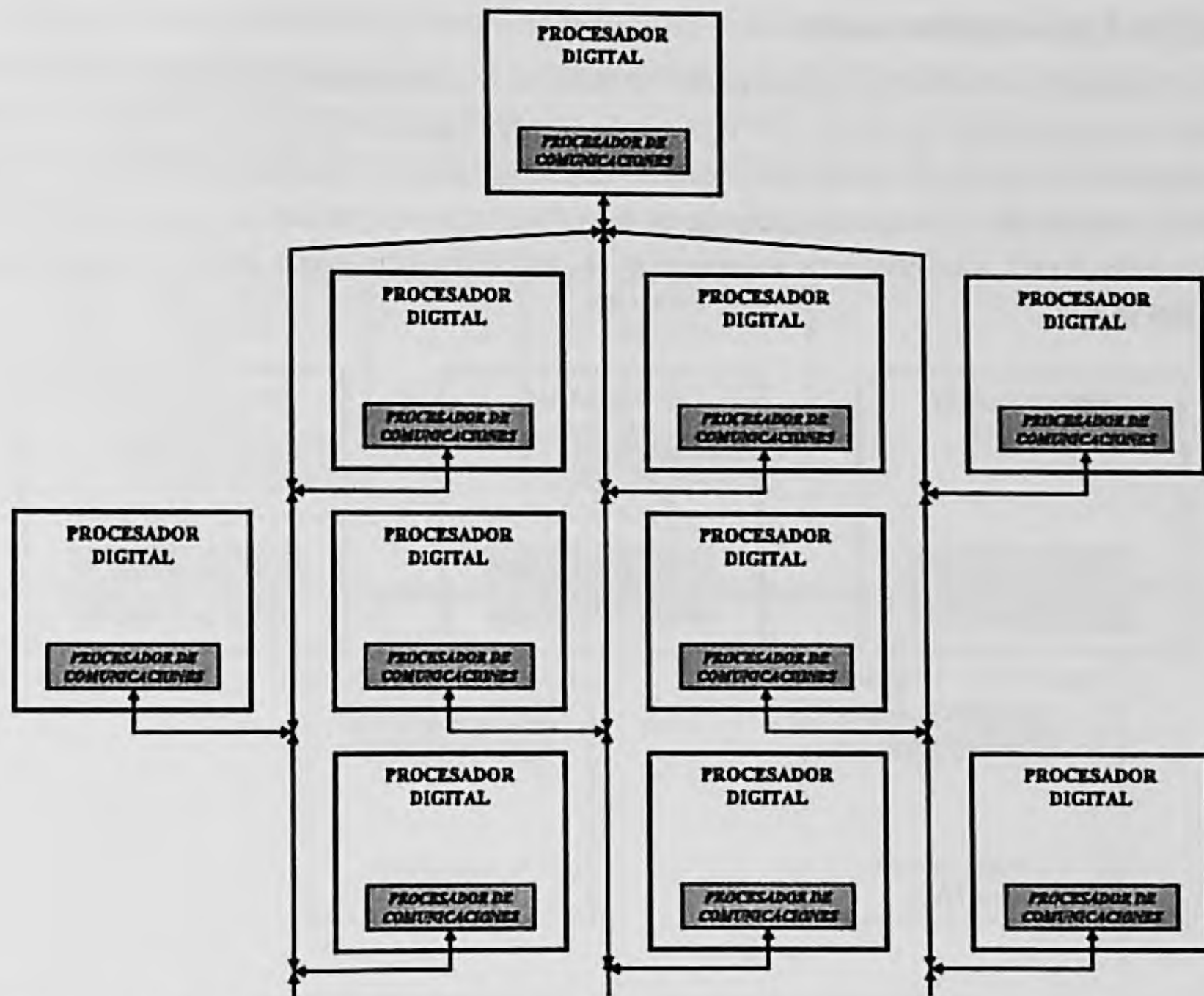


Figura A1.52. Topología de red en árbol.

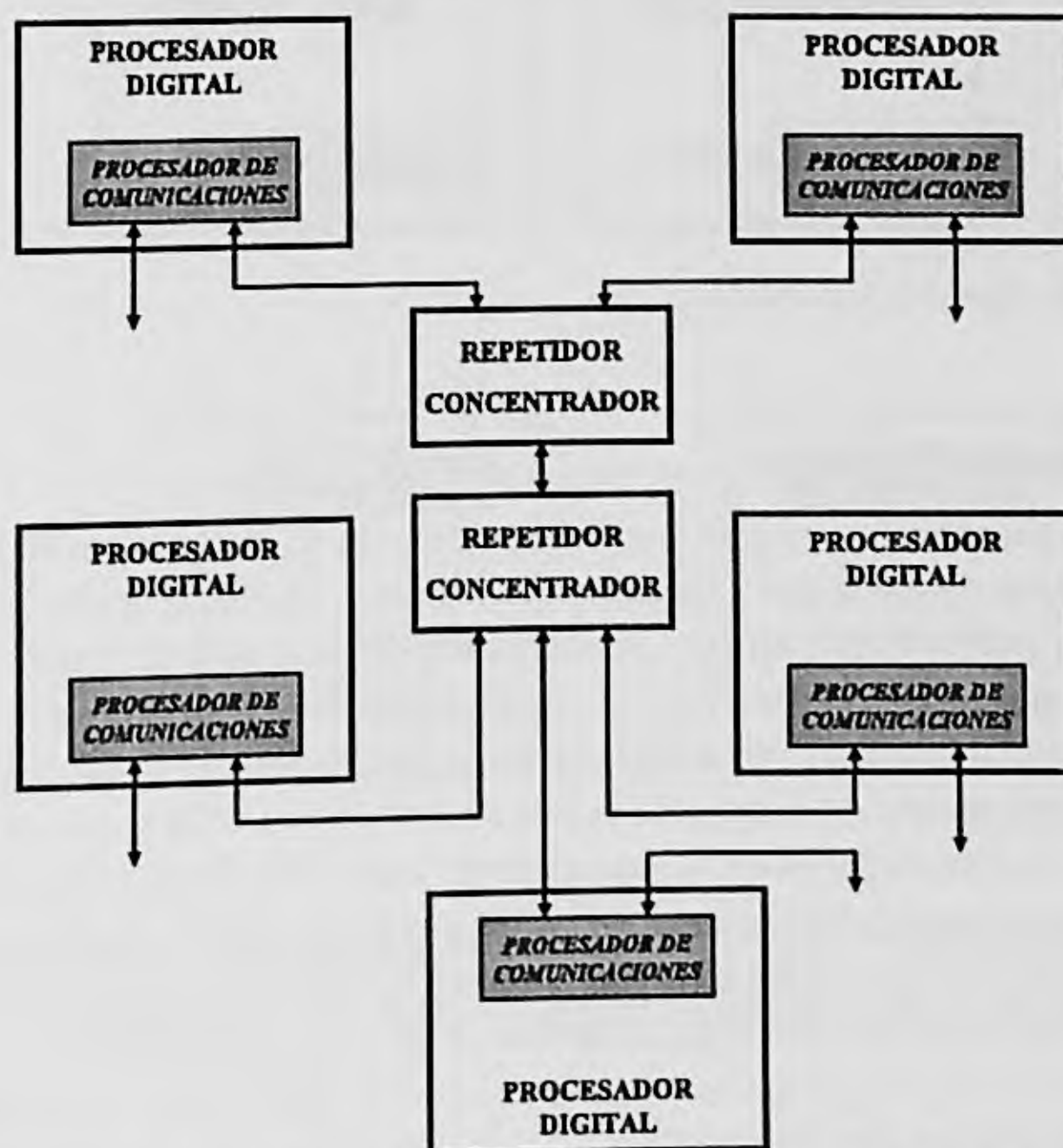


Figura A1.53. Topología de red en árbol que incluye la utilización de repetidores.

A1.3.3.1.5 Topología en anillo

En la topología en anillo los diferentes elementos se comunican mediante un bucle cerrado, de ahí su denominación de anillo (*Ring topology*) (Figura A1.54). Los datos se transmiten de un elemento al siguiente en un solo sentido desde su origen a su destino a través de todos los elementos intermedios. Un inconveniente de esta topología es que una avería del medio situado entre dos estaciones cualesquiera interrumpe la comunicación entre todas las estaciones que conforman la red.

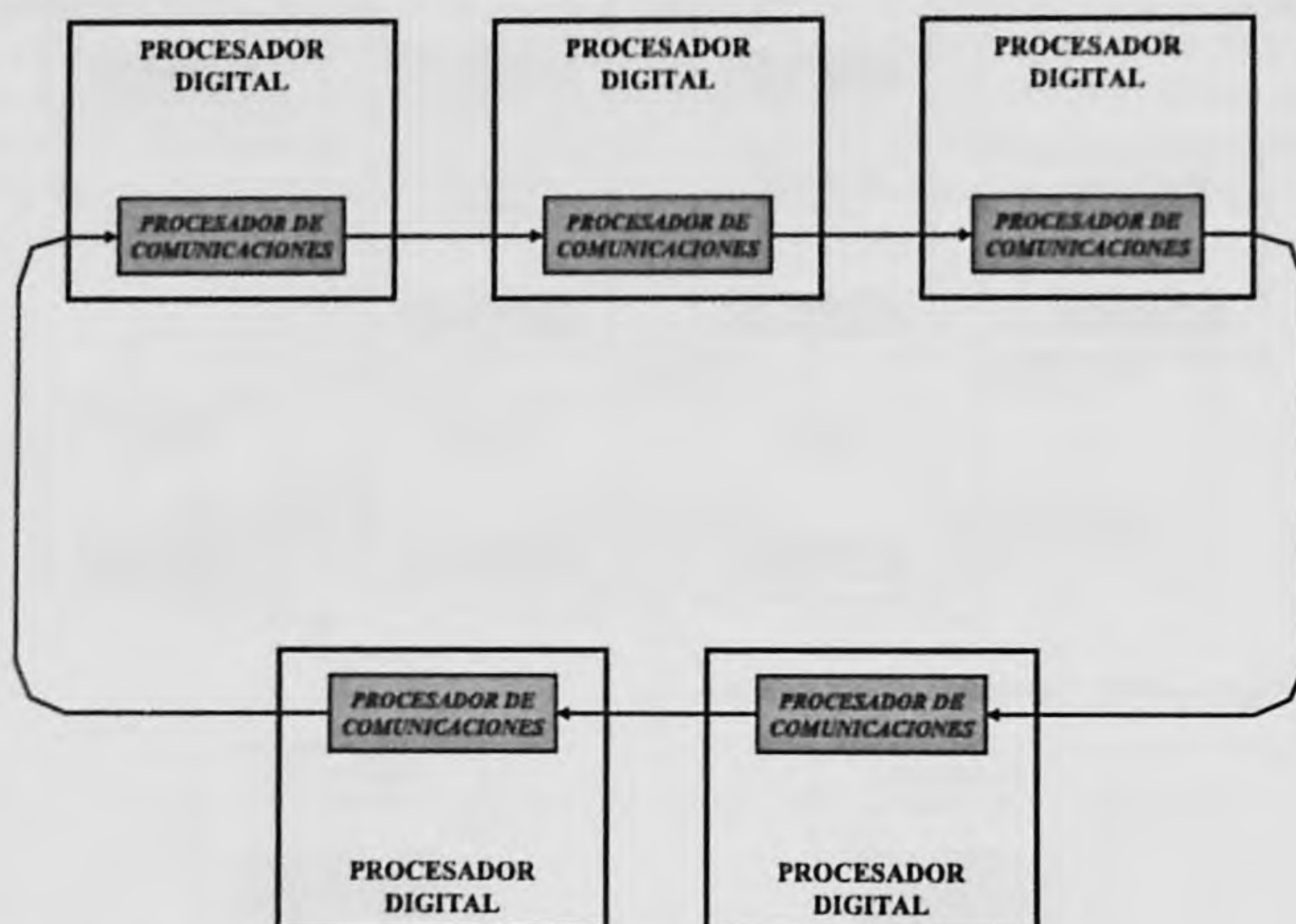


Figura A1.54. Topología de red en anillo.

A1.3.3.1.6 Topologías híbridas

Se denominan híbridas (*Mixed or hybrid topologies*) las topologías que combinan dos o más de los tipos de redes que se acaban de analizar o incluso redes de la misma topología que se diferencian por la forma en que implementan el protocolo de alguna de las capas del modelo OSI descrito en el apartado A1.3.2. En la figura A1.55 se representa una red híbrida formada por una red en bus y otra en anillo, en la que se supone que las tareas de conversión de la información del protocolo de la capa de enlace de una red a la otra y viceversa la realizan los dos procesadores de comunicaciones pertenecientes a cada una de las redes que están conectados entre sí. Es usual que el enlace entre redes diferentes se realice mediante un equipo dedicado a tal fin que, en función de las diferencias existentes entre los protocolos de la capa física y de enlace de las mismas, se denomina repetidor (*Repeater*), puente (*Bridge*), conmutador (*Switch*) o pasarela (*Gateway*), que se describen en el apartado A1.3.4, dedicado al análisis de los elementos de interconexión entre redes (*Internetworking*).

Un ejemplo real típico de red híbrida es la que resuelve las comunicaciones entre los diferentes elementos que forman parte de un proceso industrial automatizado representado mediante la pirámide CIM descrita en el apartado 9.2.5 del capítulo 9. En la figura A3.1 del apéndice 3 se muestra la red híbrida desarrollada por Siemens basada en la familia de buses de campo PROFIBUS.

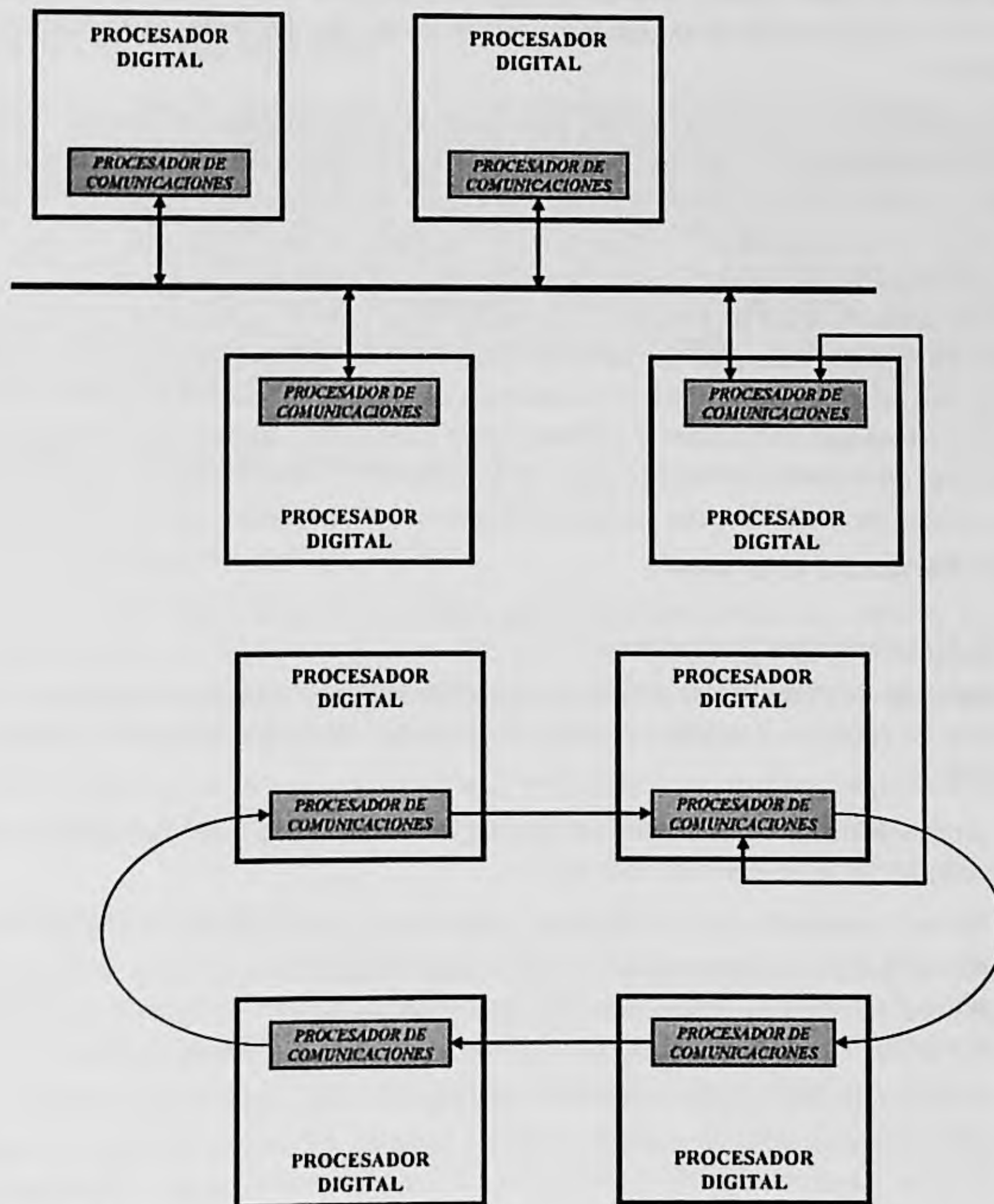


Figura A1.55. Topología de red híbrida.

A1.3.3.2 Clasificación de las redes informáticas de acuerdo con la extensión

Según su extensión las redes informáticas se pueden clasificar en locales, metropolitanas y de área extensa.

Las redes de área local conocidas como LAN (acrónimo de *Local Area Network*) se caracterizan por estar limitadas a un edificio, ya sea una oficina, una fábrica, etc., y las distancias

entre los equipos que forman parte de ellas no superan los centenares de metros.

Las redes metropolitanas, conocidas como MAN (acrónimo de *Metropolitan Area Network*), se ciñen a una ciudad y su entorno próximo o a un conjunto de edificios próximos. Las distancias típicas entre los equipos no superan las decenas de kilómetros.

Las redes de área extensa o de área amplia, conocidas como WAN (acrónimo de *Wide Area Network*), son redes de comunicaciones en las que las distancias entre equipos no tienen limitaciones.

Las redes de área local son idóneas para resolver los problemas de proceso distribuido de cualquier entidad pública o privada. Además, las redes de área local se pueden interconectar entre sí para comunicar las diversas sedes de una misma entidad a través de una red de área extensa. Por ello, a continuación se estudian primero las redes de área local, que son muy utilizadas en las Comunicaciones Industriales y a continuación: las redes de área extensa en general, y en especial, la normalización de los protocolos del nivel de red (IP) y del nivel de transporte (TCP) que han dado lugar al concepto de “Internet” primero y de “Intranet” después. Estos protocolos, contribuyen a interrelacionar entre sí los diversos campos de aplicación de las redes de comunicaciones, y hacen que comiencen a utilizarse para supervisar a distancia el comportamiento de los diferentes elementos que forman parte de la pirámide CIM de un proceso productivo.

A1.3.3.3 Redes de área local

A1.3.3.3.1 Conceptos generales

Las redes de área local se caracterizan porque en ellas sólo hay que definir los niveles físico y de enlace de datos [con sus dos subniveles de control de acceso al medio (MAC) y control lógico (LLC)].

Los procesadores de comunicaciones que implementan el protocolo de una red de área local suelen implementarse de diversas formas:

- En los computadores de aplicación general suele colocarse en su interior en forma de placa de circuito impreso, conectada al bus. Dicha placa suele denominarse “Tarjeta o placa de interfaz de red”, y algunos autores anglosajones la conocen como NIC (acrónimo de *Network Interface Card*) [ELAH 01] [BLAK 02] [MILL 01].
- En los autómatas programables de elevadas prestaciones suelen colocarse en un módulo que se acopla a la unidad central en aquellas aplicaciones en las que se tienen que realizar transferencias de información a los módulos de entrada/salida remota o a otros autómatas programables, tal como se indica en el apartado 9.4 del capítulo 9.
- En los computadores empotrados y en los autómatas programables de capacidad de operación limitada forman parte de la misma placa de circuito impreso que la unidad central.

Las redes de área local presentan un conjunto de conceptos interrelacionados que hacen difícil su aprendizaje debido a la gran cantidad de variantes de cada uno de ellos, así como de la forma de combinarlos. Los más importantes son:

- La topología

- El modo de transmisión de las señales.
- El medio físico de transmisión.
- La forma de compartir el medio físico entre varios procesadores
- El mecanismo de control del acceso al medio.

Topología de una red de área local

Tal como se indica en el apartado A1.3.3, la topología de una red es la forma en que los computadores se conectan entre sí. De las diversas topologías descritas en el citado apartado, las más utilizadas en las redes de área local son el bus, el anillo, el árbol y la estrella. En el pasado se utilizaron principalmente las dos primeras descritas en los apartados A1.3.3.1.3 y A1.3.3.1.5. En la actualidad se combina la topología bus con la estrella y el árbol para implementar redes de gran velocidad y mínimo tiempo de respuesta, utilizadas para transmitir gran cantidad de información multimedia o información entre procesadores de control de procesos, que necesitan tiempos de respuesta inferiores a 1 ms.

Modo de transmisión de las señales

La transmisión de las señales en una red de área local se puede realizar en banda base o en banda ancha, descritas en el apartado A1.2.1.

La transmisión en banda base se caracteriza por su sencillez porque no necesita moduladores ni demoduladores. Se utiliza en redes en las que sólo se transmite un determinado tipo de información y no se necesita realizar multiplexado en la frecuencia. Un ejemplo característico son los buses de campo descritos en el apartado 9.3.2.3 del capítulo 9.

Las redes en banda ancha son más complejas y utilizan técnicas de multiplexado en la frecuencia, mediante las cuales se dedica un canal distinto a cada tipo de servicio.

Medio físico de transmisión

Cualquiera de los medios de transmisión descritos en el apartado A1.2.5 puede ser utilizado para implementar una red de área local. Para elegir el más adecuado en una determinada aplicación se deben tener en cuenta diversos parámetros, como son la anchura de banda, la longitud, la fiabilidad ante interferencias electromagnéticas, la facilidad de instalación y el coste.

Un estudio detallado se sale fuera de los límites de este libro, pero se pueden hacer las siguientes puntualizaciones:

- El par trenzado se caracteriza por su bajo coste y su fácil instalación y con él se pueden alcanzar velocidades elevadas del orden de 150 Megabits por segundo, a distancias del orden de los cien metros.
- El cable coaxial permite alcanzar distancias y velocidades de transmisión mayores y se suele utilizar en redes de banda ancha.
- La fibra óptica constituye, debido a sus características indicadas en el apartado A1.2.5.4, un medio guiado de transmisión utilizado en muchos campos de aplicación de las redes de área local en los que se necesita una gran velocidad. Es el medio que permite alcanzar mayor velocidad y seguridad en la comunicación.

- La atmósfera se utiliza para evitar los inconvenientes del tendido de cables en algunos entornos, como por ejemplo el doméstico. Operan a distancias que, por lo general, nunca superan los cientos de metros y utilizan las técnicas de modulación de una portadora que puede ser:
 - Una señal de UHF de 300 MHz a 3 GHz.
 - Microondas de la banda de 3 GHz a 30 GHz.
 - Rayos infrarrojos.

Para un estudio más detallado de las redes de área local inalámbricas se remite al lector a la bibliografía [ELA 01] [FORO 07a] [STAL 07].

Forma de compartir el medio físico entre varios procesadores

Es la forma en la que se logra que varios procesadores compartan el ancho de banda de un medio común. Se suele denominar división en canales (*Channelization*) y puede realizarse de tres maneras diferentes:

- **Acceso múltiple mediante división en la frecuencia**
Se la conoce por el acrónimo FDMA (*Frequency-Division Multiple Access*) y, como su nombre indica, consiste en asignar a cada procesador una banda de frecuencia para transmitir sus datos. Es en realidad una aplicación del multiplexado en la frecuencia (FDM), descrito en el apartado A1.2.1.
- **Acceso múltiple mediante división en el tiempo**
Se la conoce por el acrónimo TDMA (*Time-Division Multiple Access*) y, como su nombre indica, consiste en asignar a cada procesador un intervalo de tiempo (*Time slot*) durante el cual envía sus datos. Es en realidad una aplicación de la multiplexación en el tiempo (TDM), descrita también en el apartado A1.2.1.
- **Acceso múltiple mediante división de código**
Se la conoce por el acrónimo CDMA (*Code-Division Multiple Access*) y, como su nombre indica, consiste en asignar a cada procesador un código. Se superan así los inconvenientes del FDMA que hace necesaria la utilización de moduladores/demoduladores y no facilita la transferencia de información entre varios procesadores, así como del TDMA que tiene problemas de sincronización entre procesadores independientes que acceden a un medio común [STAL 07].

Estrategia de acceso al medio

Las redes de área local de topología bus y anillo se caracterizan por poseer un medio de comunicación compartido por varios procesadores que utilizan el método CDMA de acceso múltiple mediante división de código. Esto hace necesario el establecimiento de una estrategia o mecanismo que regule el acceso al mismo por parte de cada uno de los procesadores conectados a la red. Dicha estrategia, a la que se denomina “Control de acceso al medio” y se la conoce como MAC (acrónimo de *Medium Access Control*), forma parte de la capa de enlace (*Data Link*) del modelo OSI (Figura A1.45).

Es posible idear un gran número de estrategias o formas de controlar el acceso a un medio común mediante CDMA por parte de varios procesadores digitales, de las que a continuación se indican las más utilizadas:

- **Acceso aleatorio**

Recibe la denominación de CSMA/CD (acrónimo de *Carrier Sense Multiple Access with Collision Detection*) y el procesador de comunicaciones que lo implementa ejecuta un programa que realiza las acciones indicadas en la figura A1.56.

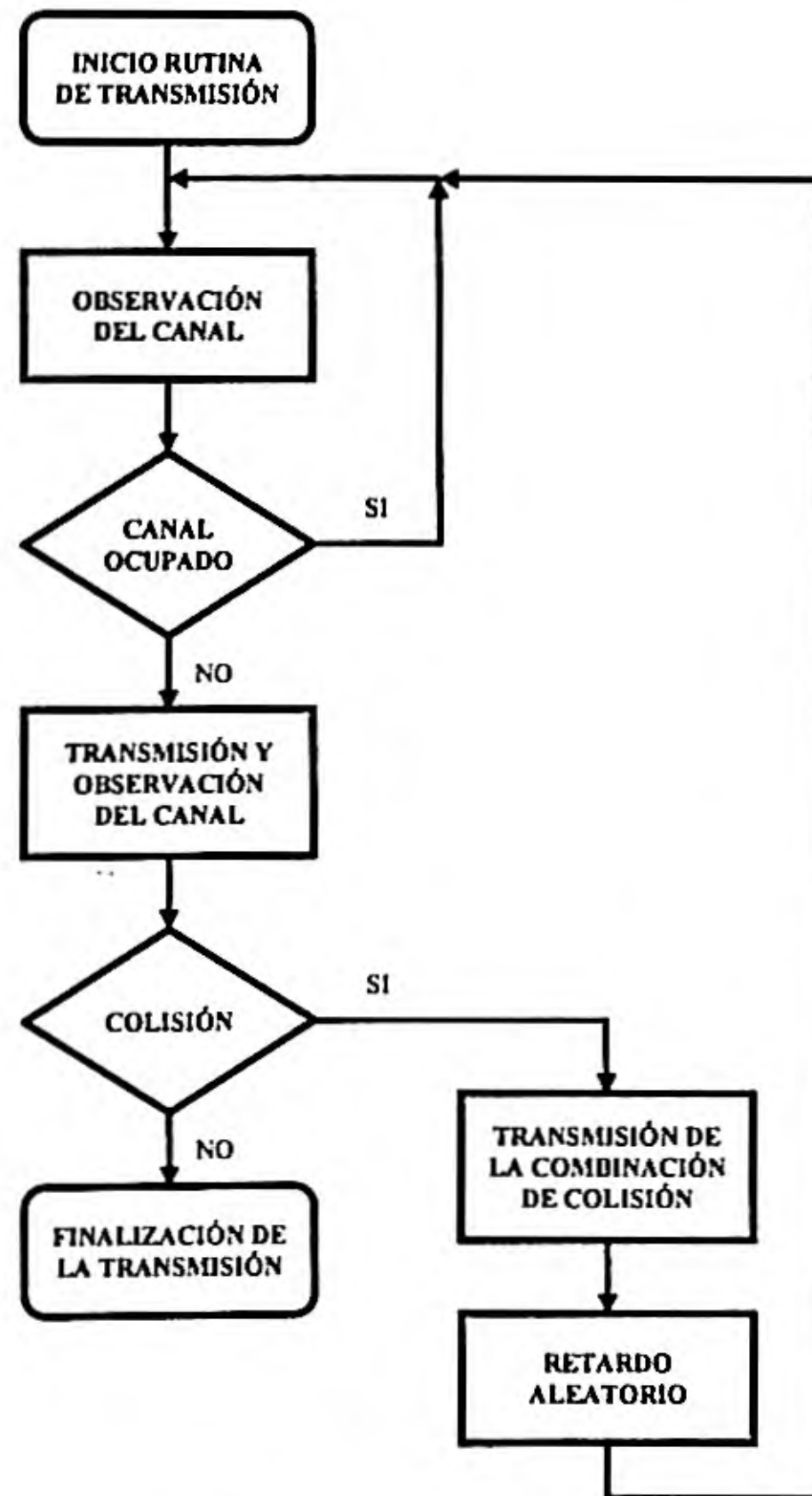


Figura A1.56. Mecanismo de acceso al medio CSMA/CD.

Cuando un procesador quiere transmitir una información al medio, observa si está o no ocupado y no inicia la transmisión hasta que el medio está desocupado. En ese instante realiza la transmisión y observa permanente el canal hasta que la misma finaliza, para comparar la información presente en el canal con la transmitida por él. Si ambas informaciones coinciden, considera que la transmisión se ha realizado correctamente y da por finalizada la rutina. En caso contrario, detecta que se ha producido una colisión

con la información enviada por otro procesador y envía al medio una combinación de colisión (*Jam*) si antes no la recibe del otro. Seguidamente genera un retardo de duración aleatoria e inicia de nuevo la ejecución de la rutina de transmisión. Todos los procesadores conectados a la red poseen la misma prioridad para transmitir información a través de la misma.

- **Acceso controlado con paso de testigo (*Token*)**

En la estrategia de acceso controlado con paso de testigo los procesadores sólo pueden transmitir cuando están en posesión del testigo, que es una trama corta que especifica el procesador que puede utilizar, en un determinado instante, el canal y el que puede transmitir cuando acabe el que está accediendo al canal. Para evitar que un procesador que tiene que transmitir mucha información monopolice el canal se limita el tiempo durante el cual puede retener el testigo ininterrumpidamente.

- **Acceso priorizado**

Se denominan de acceso priorizado las estrategias que hacen que los procesadores conectados a una red tengan diferente nivel de prioridad para acceder al medio común. Se pueden concebir múltiples estrategias de este tipo entre las que cabe citar:

- **Estrategia Principal/Subordinado**

La estrategia Principal/Subordinado (*Master/Slave*) hace que sólo uno de los procesadores (Principal) pueda acceder al medio en cualquier instante y que los demás (Subordinados) sólo lo hagan cuando aquél se lo permite y que en ese instante envíen el mensaje adecuado al procesador principal. Este mecanismo se puede además combinar con otros, como el de paso de testigo, para que sean varios, en lugar de uno sólo, los procesadores que pueden acceder al canal en cualquier instante. Este tipo de acceso es el utilizado por el bus de campo PROFIBUS que se describe en el apéndice 3.

- **Estrategia Proveedor-Consumidor**

La estrategia Proveedor/Consumidor (*Provider-Consumer*), también denominada Productor/Consumidor (*Producer-Consumer*), se puede concebir de diversas formas, entre las que cabe citar:

- La estrategia productor distribuidor-consumidor, en la que hay un procesador que actúa como árbitro y decide cual es el procesador que debe colocar una determinada información en el bus (*Producer or Provider*) y cuales deben ser los que la reciban (*Consumers*). Para ello el árbitro envía un mensaje en el que indica la información a transmitir, cual es el procesador que debe actuar como productor y cuales son los que deben actuar como consumidores. Una vez que tanto los consumidores como el proveedor reciben el citado mensaje se produce la transferencia de información entre ellos.
- La estrategia productor-consumidor propiamente dicha, en la que el productor y los consumidores de una determinada información se definen durante el proceso de configuración del sistema de control distribuido para que las transferencias se produzcan en los instantes adecuados.

A1.3.3.3.2 Redes de área local descentralizadas

Las redes de área local de igual a igual (*Peer to peer*) [ELAH 01], se denominan también descentralizadas porque carecen de un procesador central que controle el proceso global realizado por el conjunto de computadores que la conforman. Son conocidas simplemente como LAN y se utilizan en entornos en los que se interconectan un conjunto de computadores que deben tener capacidad para intercambiar información entre ellos sin que ninguno prevalezca sobre los demás.

Red Ethernet

Conceptos básicos

La norma IEEE 802.3 (más conocida como "Ethernet"), desarrollada y registrada inicialmente por las empresas Xerox Corporation, Digital Equipment e Intel se caracteriza por:

- Utiliza la topología de tipo bus (descrita en el apartado A1.3.3.1.3), tal como se indica en la figura A1.57.
- Posee una trama como la representada en la figura A1.58, constituida por el siguiente conjunto de campos:
 - Un preámbulo (*Preamble*) que contiene 7 octetos de ceros y unos alternados para advertir del envío de una trama a los posibles destinatarios y facilitar su sincronización.
 - Un delimitador de inicio de la trama SFD (acrónimo de *Start Frame Delimiter*) constituido por el octeto 10101011 que indica el verdadero inicio de la trama. Los dos unos seguidos al final indican a los posibles destinatarios, que a continuación van a recibir la dirección de aquél al que va dirigida la información.
 - La dirección del destinatario al que va dirigida la información o dirección de destino DA (acrónimo de *Destination Address*), formada por 6 octetos.

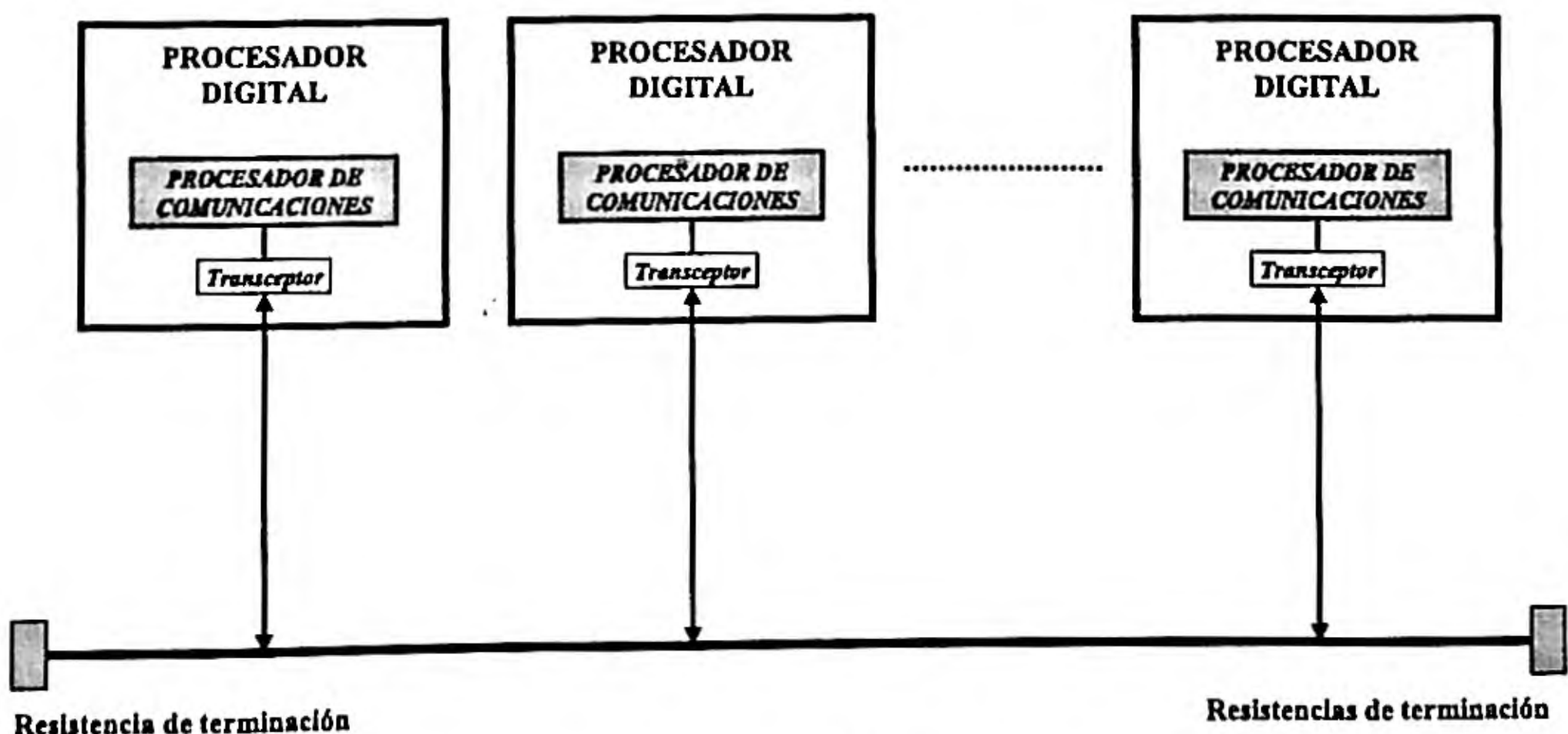


Figura A1.57. Diagrama de bloques de una red Ethernet.



Figura A1.58. Trama de la subcapa MAC de acceso al medio de la norma IEEE 802.3 de la red Ethernet.

- La dirección del emisor SA (acrónimo de *Source Address*) que envía la información a través de la red.
- La longitud (*Length*) que indica el número de octetos que constituyen el campo de datos.
- Los datos (*Data*) que tienen un número de octetos comprendidos entre 46 y 1500.
- La combinación de detección de errores CRC-32 (acrónimo de *Cyclic Redundancy Check*).
- Utiliza el mecanismo de acceso al medio denominado “Detección de portadora con acceso múltiple y detección de colisiones” conocido como CSMA/CD (acrónimo de *Carrier Sense Multiple Access with Collision Detection*) descrito en el apartado A1.3.3.3.1. Para ello, cada elemento de la red está dotado de un procesador de comunicaciones que tiene un receptor (*Receiver*) y un emisor (*Transmitter*) que constituyen conjuntamente un transceptor (*Transceiver*) conectado al medio físico de transporte de la información, tal como se indica en la figura A1.57. En general, el transceptor y su conexión al medio reciben la denominación de unidad de enlace con el medio, conocida como MAU (*Medium Attachment Unit*).
- Debido a la aleatoriedad del acceso al medio, en esta clase de red no se puede garantizar el valor máximo del intervalo de tiempo que transcurre desde que un procesador inicia la rutina de transmisión hasta que la finaliza con éxito. Se dice que la transmisión no es determinista, lo que puede resultar inadmisibles en aquellas aplicaciones en las que el proceso distribuido debe realizarse en tiempo real (descrito en el apartado 9.2.4.4 del capítulo 9).

Para facilitar la transmisión de grandes cantidades de información, típicas de la transmisión de información multimedia, y disminuir el tiempo de respuesta en las aplicaciones de control industrial, la red Ethernet ha evolucionado desde la velocidad inicial de 10 Mbps, mediante la modificación del protocolo de la capa de enlace y del medio de transmisión, para lograr el desarrollo de tres redes Ethernet avanzadas:

- La red Ethernet rápida (*Fast Ethernet*) [FORO 07a] [HELD 03] [STAL 07].
- La red Ethernet ultrarrápida (*Gigabit Ethernet*) [FORO 07a] [HELD 03] [STAL 07].
- La red Ethernet isócrona (*Isochronous Ethernet*) [HELD 03].

Estos tres tipos de redes contribuyen a disminuir el tiempo máximo de acceso al medio y por ello hacen que la red Ethernet sea más determinista, de tal forma que pueda ser utilizada en aplicaciones en las que se necesita un tiempo de respuesta inferior a 1 ms. A continuación se describen brevemente.

Red Ethernet rápida

La red Ethernet rápida (*Fast Ethernet*) se describe en la norma 802-3i del IEEE y alcanza los 100 Mbps. Para lograr esa velocidad de transmisión, se caracteriza por:

- Tiene el mismo protocolo de la capa de enlace que la red Ethernet de la norma 802.3.
- Utiliza códigos distintos del Manchester de la norma IEEE 802.3 en función del tipo de medio [FORO 07a].
- Utiliza elementos de interconexión (*Connecting devices*) y por ello es una red Ethernet conmutada (*Switched Ethernet*) con topología distinta del bus, descrita en el apartado A1.3.5.2.
- Puede utilizar el modo de comunicación (*Data flow mode*) doble simultáneo (*Full duplex*) para lograr la máxima velocidad de transmisión, con la consiguiente disminución del tiempo de respuesta.

Red Ethernet ultrarrápida

La red Ethernet ultrarrápida (*Gigabit Ethernet*) se describe en la norma 802-3z del IEEE y alcanza 1 Gbps. Para lograr esta velocidad:

- Tiene el mismo protocolo de la capa de enlace que la red Ethernet de la norma 802.3.
- Utiliza también códigos distintos del Manchester de la norma IEEE 802.3 [FORO 07a].
- Utiliza elementos de interconexión y por ello también es una red Ethernet conmutada (*Switched Ethernet*) descrita en el apartado A1.3.5.2.
- Puede utilizar también el modo de comunicación (*Data flow mode*) doble simultáneo (*Full duplex*) para lograr la máxima velocidad de transmisión con la consiguiente disminución del tiempo de respuesta.

Red Ethernet isócrona

La red Ethernet isócrona (*Isochronous Ethernet*) está basada en el modo de transmisión isócrono descrito en el apartado A1.2.3.3 y es una extensión de la red Ethernet 802.3i que operó inicialmente a 10 Mbps y utilizó como medio un par trenzado no apantallado. Fue pensada inicialmente a principios de la década de 1990 para transmitir información sensible al tiempo de transmisión, como por ejemplo información multimedia, pero el desarrollo de las redes Ethernet rápida y ultrarrápida pareció que la relegaba a un segundo plano [HELD 03]. La red Ethernet isócrona se utiliza actualmente combinada con la rápida y la ultrarrápida en las redes denominadas Ethernet Industrial que se describen en el apartado 9.3.2.5.2 del capítulo 9.

Para mejorar el determinismo, y lograr tiempos máximos de transmisión inferiores a 1ms, con una desviación máxima (*Jitter*) de algunos microsegundos, necesarios en algunas de las aplicaciones de control industrial, además de elevar la velocidad de transmisión, se ha convertido la topología bus en una topología estrella o árbol mediante la utilización de elementos de interconexión (*Connecting devices*) denominados conmutadores (*Switches*), que se describen en el apartado A1.3.4. Por tratarse de redes que poseen dichos elementos se las puede denominar redes Ethernet conmutadas (*Switched Ethernet networks*) y a su estudio se

dedica el apartado A1.3.5.2. Mediante ellas se logra que la red Ethernet pueda ser utilizada para implementar las comunicaciones en todos los niveles de la pirámide CIM, tal como se indica en el apartado 9.3.2.5 del capítulo 9 y en el apéndice 4 en el que se describe la red Profinet.

Al lector interesado en estudiar con mayor detalle las diferentes redes Ethernet se le recomienda la bibliografía indicada al final del capítulo [FORO 07a] [HELD 03] [STAL 07] [YOUL 02].

Red en anillo con paso de testigo

Este tipo de red (denominada en inglés *Token Ring*) utiliza la topología en anillo (descrita en el apartado A1.3.3.1.5) y la combina con el protocolo de acceso al medio que utiliza un mensaje que especifica cual es el nodo de la red que está utilizándola en cada instante y el próximo que podrá utilizarla en el momento en que el actual deje de hacerlo, que se describe, junto con otros protocolos de acceso, en el apartado A1.3.3.3.1.

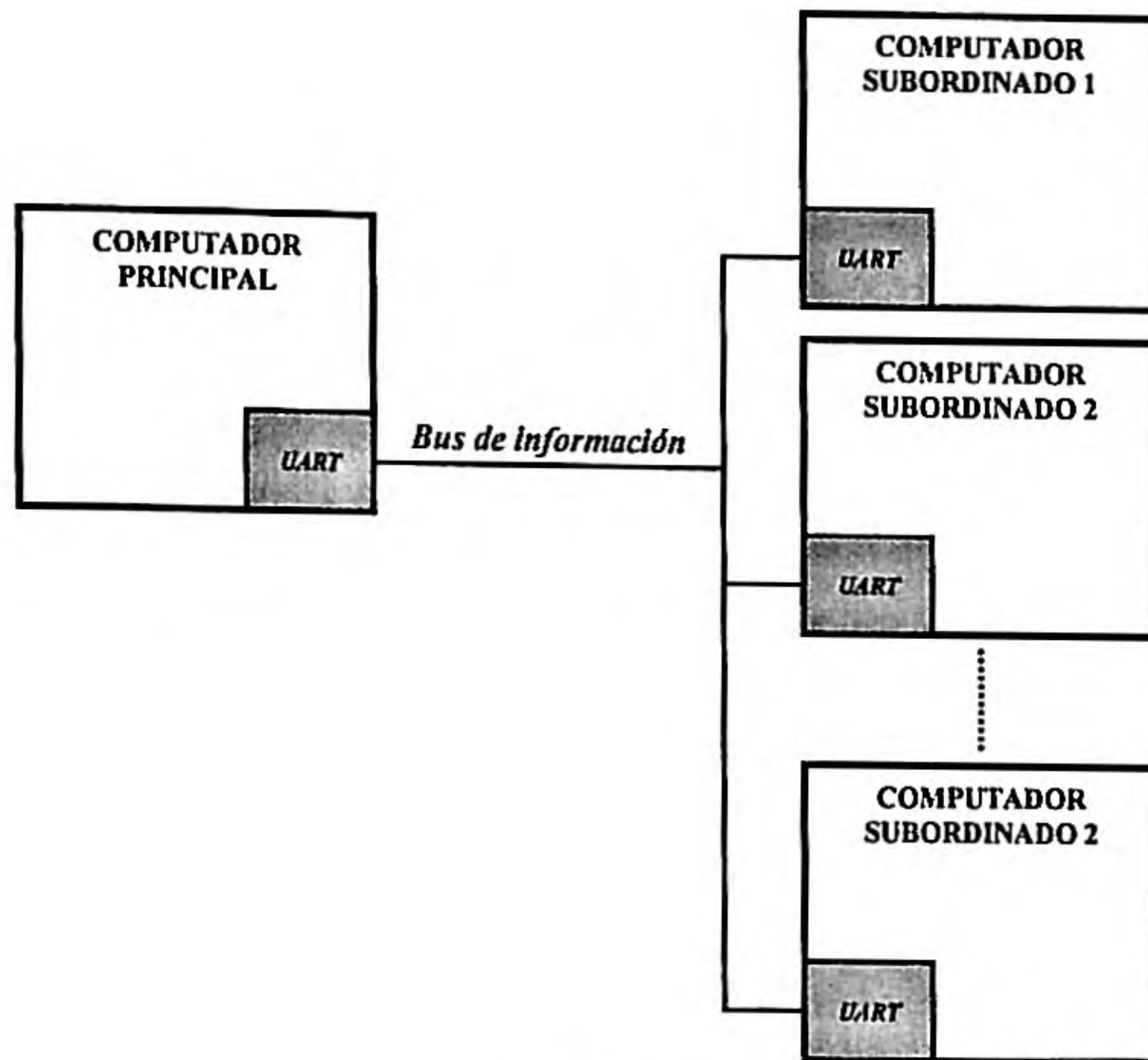
Red en bus con paso de testigo

Este tipo de red (denominada en inglés *Token Bus*) utiliza la topología del tipo bus (descrita en el apartado A1.3.3.1.3) combinándola con un protocolo de acceso al medio que utiliza un mensaje como el de la red en anillo descrita en el apartado anterior, por lo que se dice que genera un anillo lógico. El anillo lógico presenta la ventaja de que el fallo de uno de los procesadores puede ser detectado comprobando que no responde al envío del testigo por parte del procesador que le precede y procediendo, en ese caso, a su eliminación. Además, la red en bus con paso de testigo presenta la ventaja de que mediante los programas adecuados se pueden combinar los mecanismos de paso de testigo y el principal/subordinado que se describe a continuación. Se obtiene así una red que supera las limitaciones de las redes principal/subordinado, en las que hay un solo procesador principal, sin elevar excesivamente la complejidad. Un ejemplo característico de ello es la red PROFIBUS que se describe en el apéndice 3.

A1.3.3.3.3 Redes de área local principal/subordinado

Las redes de área local principal/subordinado o principal/secundario (*Master/slave*) tienen topología bus y en ellas existe al menos un procesador denominado principal (*Master*) que determina cuando pueden acceder a la red los demás denominados subordinados (*Slaves*). Utilizan por lo tanto el mecanismo de acceso priorizado al medio descrito en el apartado A1.3.3.3.1. En la figura A1.59 se representa el esquema de una red realizada con una topología bus en banda base digital con modo de sincronización asíncrono y mecanismo de acceso al medio principal/subordinado. El computador principal pone el noveno bit en uno y coloca en el bus la dirección del computador. A continuación pone el citado bit en cero y coloca en el bus el dato.

Este tipo de redes de área local utiliza el sondeo periódico que se caracteriza por tener como mínimo un procesador principal que solicita periódicamente información a los subordinados. Por ello tiene especial aplicación en la resolución de las comunicaciones en los niveles bajos de la pirámide CIM que se estudian en el apartado 9.3.2.3 del capítulo 9. Es posible concebir de muchas formas diferentes el protocolo de la subcapa de control lógico (LLC) de una red de este tipo. En los apéndices 2 y 3 se describen los utilizados en las redes o buses normalizados AS-i y PROFIBUS, respectivamente.



a)



b)

Figura A1.59. Red en bus que utiliza el concepto de TDM para realizar una comunicación en banda base digital simple con modo de sincronización asíncrono: a) Diagrama de bloques; b) Protocolo de la capa de enlace.

A1.3.3.3.4 Redes de área local proveedor/consumidor

Las redes de área local proveedor/consumidor utilizan el mecanismo de acceso prioritario del mismo nombre descrito también en el apartado A1.3.3.3.1. Son típicas de la solución que algunos fabricantes de sistemas electrónicos de control distribuido utilizan para controlar el acceso a un bus de los diferentes controladores o sensores inteligentes, para diferenciarse de los que utilizan el mecanismo principal/subordinado.

A1.3.3.4 Redes de área extensa

A1.3.3.4.1 Introducción

En el apartado anterior se pone en evidencia que existen múltiples redes de área local que se diferencian por su topología, su mecanismo de acceso al medio, el modo de transmisión de las señales y el medio físico de transmisión que utilizan, pero todas ellas tienen en común la distancia reducida y el número limitado de nodos.

En este apartado se estudian las redes de área extensa (WAN) que están compuestas por un conjunto de nodos de conmutación interconectados. La información enviada por un procesador a otro puede atravesar diferentes nodos y, por ello, se dice que es enrutada o encaminada a través de ellos, de forma transparente para el usuario, hasta llegar a su destino.

De acuerdo con la forma en que se enrutan o encaminan los paquetes de información, las redes de área extensa pueden ser de conmutación de circuitos o de paquetes, que se describen seguidamente.

Redes de conmutación de circuitos

Una red de conmutación de circuitos (*Circuit switching network*) establece un canal (*Path*) de transmisión dedicado entre los dos procesadores que se comunican. El ejemplo más típico de este tipo de procesadores es la Red Telefónica Básica (RTB) que se desarrolló para implementar la transmisión de la voz mediante señales analógicas. Su esquema básico se representa en la figura A1.60 y se caracteriza por:

- Los elementos de comunicación se unen físicamente entre sí mediante un canal que se mantiene dedicado a la comunicación desde que comienza hasta que finaliza, instante en el que queda libre.
- Es una alternativa adecuada para comunicar elementos que generan o reciben información sensible al retardo, es decir, que debe llegar en un orden determinado y con retardo uniforme, como es el caso de la voz y la imagen.
- Es poco rentable en el caso de que en la comunicación se produzcan pausas, como por ejemplo las que se producen en las comunicaciones digitales cuando un elemento solicita a otro una información que tarda un cierto tiempo en elaborarse.

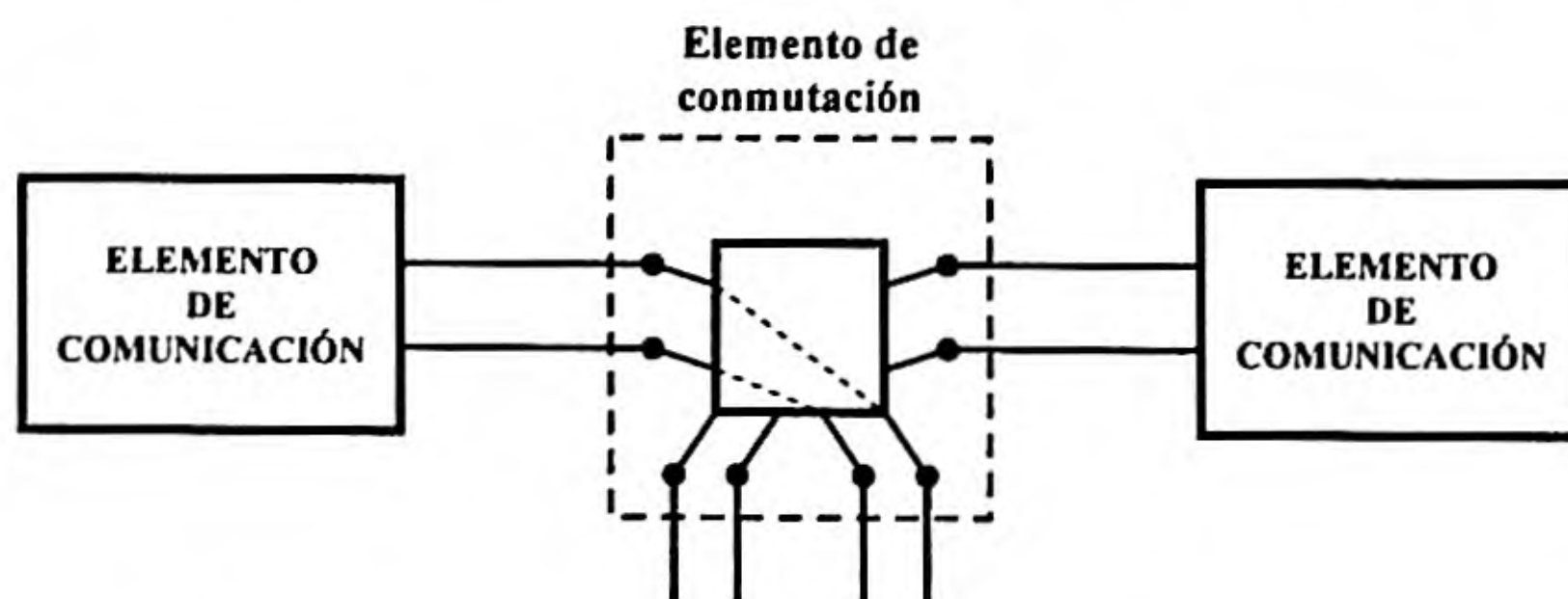


Figura A1.60. Esquema básico de una red de datos de área extensa de conmutación de circuitos.

Las grandes inversiones realizadas en las redes telefónicas hacen que su utilización continúe vigente en la actualidad e incluso ha dado lugar a nuevos tipos de redes como la RDSI (acrónimo de Red Digital de Servicios Integrados) conocida también como ISDN (acrónimo de *Integrated Services Digital Network*) [MILL 00].

Redes de conmutación de paquetes

Una red de conmutación de paquetes (*Packet switching network*) se caracteriza por no establecer canales dedicados sino que divide los mensajes en una secuencia de combinaciones binarias denominadas paquetes (*Packets*) que contienen un determinado número de bits. Cada paquete pasa de un nodo a otro hasta llegar a su destino final. Para ello, cada nodo tiene que tener capacidad para recibir paquetes, almacenarlos temporalmente y transmitirlos.

Sus principales características son:

- Los paquetes se propagan a través de la red desde un emisor hasta un receptor sin necesidad de que haya un enlace físico directo entre ellos.
- Se utiliza fundamentalmente en la transmisión de datos que no son sensibles al retardo.
- Define la segmentación del paquete en un conjunto de partes denominadas campos que permiten identificar al receptor, detectar errores, etc.
- Está basada en la multiplexación en el tiempo (TDM) y es muy flexible porque rentabiliza los canales de transmisión, dado que para realizar una comunicación se pueden utilizar diferentes canales.

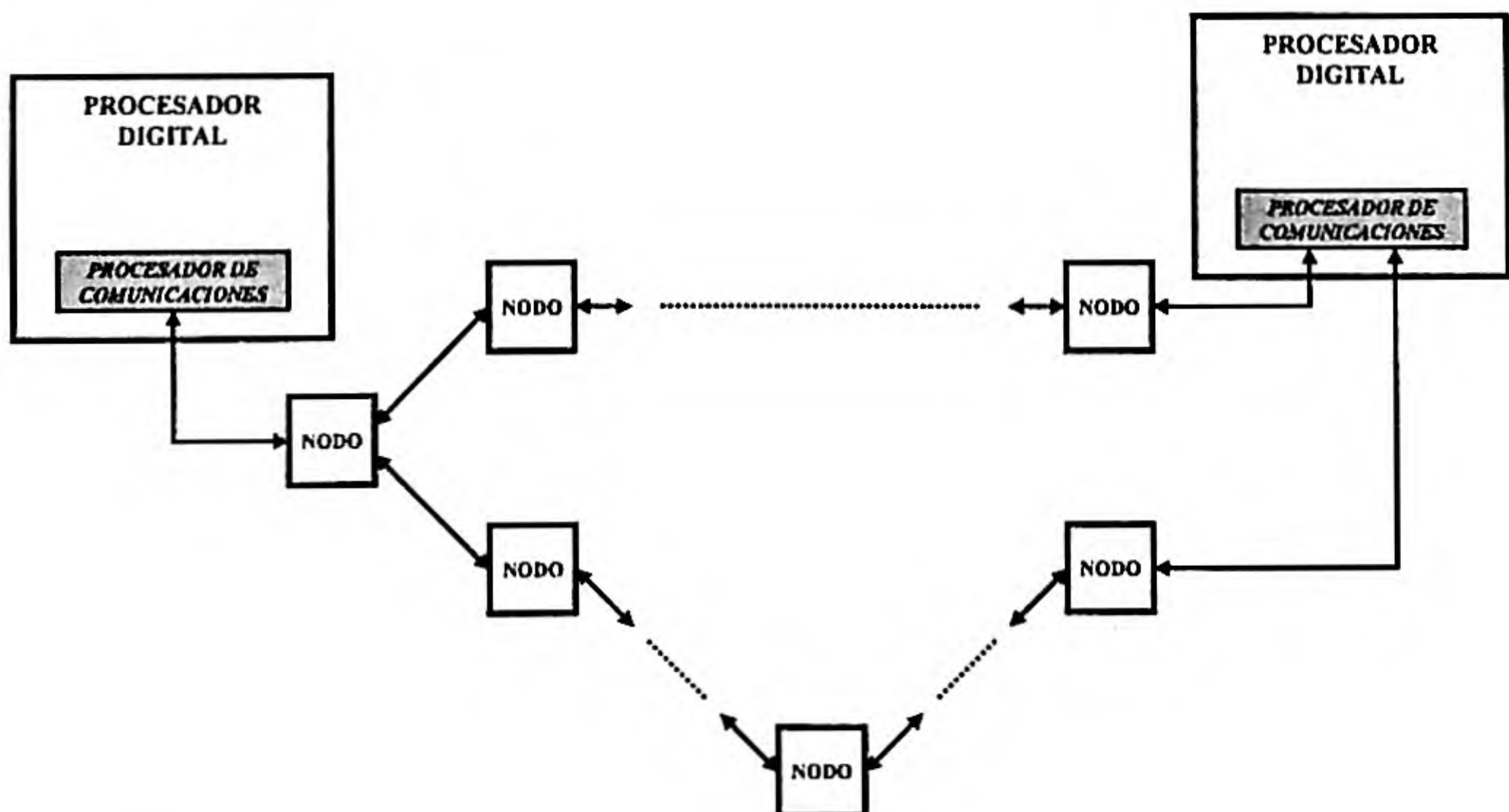


Figura A1.61. Esquema básico de una red de datos de área extensa de conmutación de paquetes.

Son realmente redes de conmutación de paquetes las diferentes redes de área local descritas en el apartado A1.3.3.3.3. En la figura A1.61 se representa gráficamente una red de área extensa de conmutación de paquetes que enlaza dos procesadores digitales a través de un conjunto de nodos que proporcionan múltiples canales de comunicación para lograr la transferencia de información entre ambos.

Un concepto básico en este tipo de redes de área extensa es la elección de la ruta que siguen los paquetes, conocida como enrutado (*Routing*). La evolución, tanto de los recursos físicos (*Hardware*) como de programación (*Software*) de los computadores, ha dado lugar a dos nuevas formas de implementar el enrutado denominadas "Retransmisión de tramas" (*Frame Relay*) y "Modo de transferencia asincrónica" conocida como ATM (acrónimo de *Asynchronous Transfer Mode*) cuya descripción se sale fuera de los límites de este libro y por ello se remite al lector a la bibliografía [FORO 07a] [SMIT 93] [STAL 07].

Las redes de área extensa, y en particular el conjunto de redes que constituyen Internet, utilizan los diferentes medios de transmisión indicados en el apartado A1.2.5. Entre ellos hay que destacar las comunicaciones inalámbricas, que han tenido un desarrollo espectacular durante la última década del siglo XX como resultado del desarrollo de la combinación de los procesadores empotrados con los circuitos transmisores y receptores de radiofrecuencia que han dado lugar a los servicios de comunicaciones móviles, cuyo estudio se sale de los límites de este libro, por lo que se remite al lector a la bibliografía [BLAK 02] [RAPP 02].

A1.3.3.4.2 Redes TCP/IP

Conceptos generales

Se denominan redes TCP/IP a un conjunto de redes interconectadas de conmutación de paquetes que comparten el protocolo de la capa de red denominado IP y el de la capa de transporte denominado TCP. Dado que las redes TCP/IP pueden utilizar protocolos de nivel físico y de nivel de enlace diferentes se interconectan mediante sistemas de comunicaciones denominados pasarelas (*Gateways*), descritas en el apartado A1.3.4.

Las redes TCP/IP están organizadas de acuerdo con el modelo de interconexión de sistemas informáticos abiertos denominado DoD, que tiene cuatro niveles o capas de diálogo, que constituyen otros tantos protocolos (Figura A1.47). El conjunto de redes TCP/IP interconectadas recibe la denominación de Internet.

En la figura A1.62 se representa el modelo de las comunicaciones entre dos computadores a través de un red TCP/IP. El nivel físico puede utilizar cualquier tipo de medio, tanto guiado como no guiado. El nivel de enlace de datos también puede ser cualquiera de los analizados en apartados anteriores, como por ejemplo Ethernet, etc.

La implementación de la capa de red (*Internetwork layer*) de TCP/IP se realiza mediante el protocolo denominado IP (*Internetworking Protocol*) y en la capa de transporte se utilizan los protocolos TCP (acrónimo de *Transmission Control Protocol*), UDP (acrónimo de *User Datagram Protocol*) y SCTP (acrónimo de *Stream Control Transmission Protocol*).

El nivel de aplicación (*Application Process*) está constituido por los diferentes programas que proporcionan servicios a un usuario, como por ejemplo el correo electrónico (*Electronic mail*).

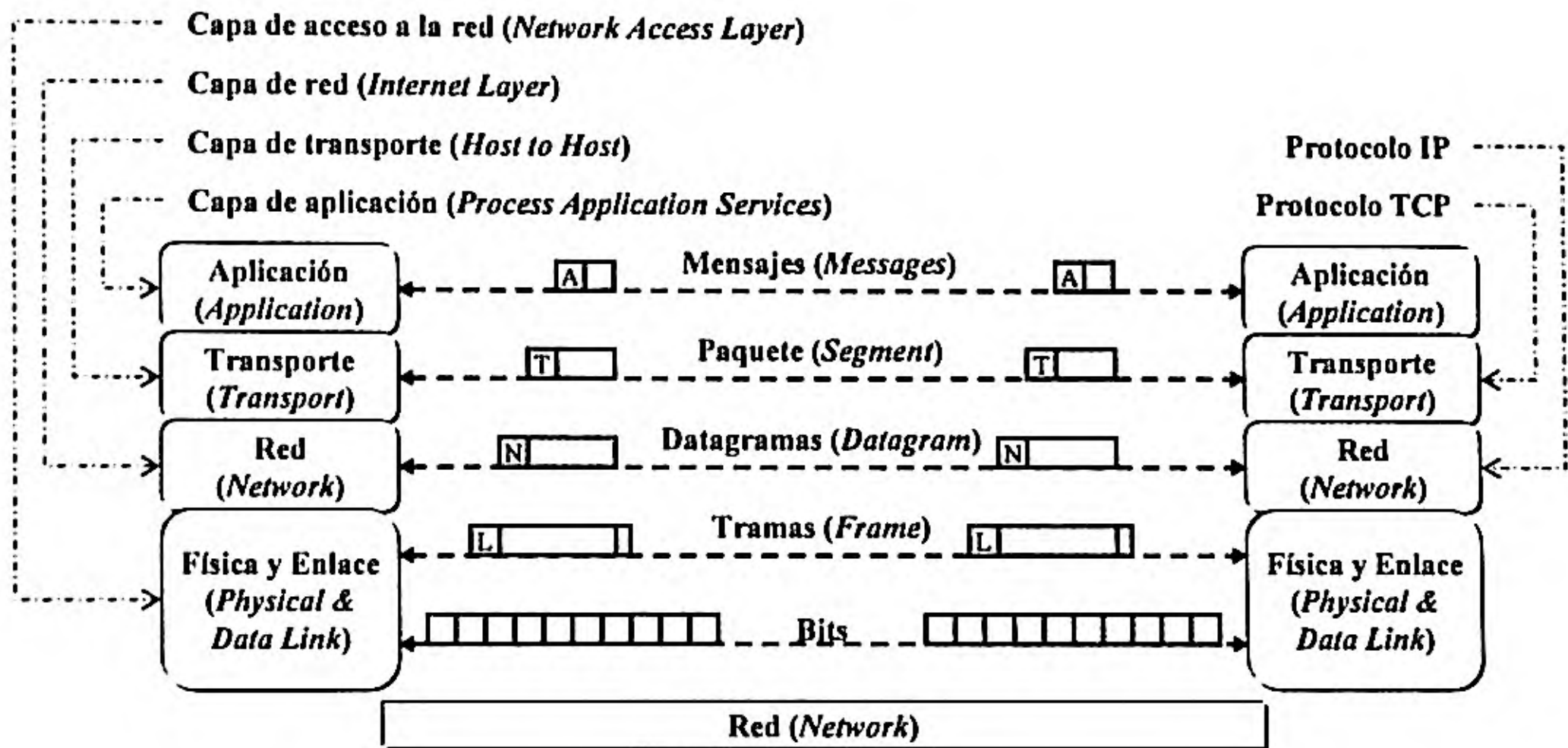


Figura A1.62. Comunicaciones en internet que utilizan el modelo DoD.

A continuación se describen brevemente los protocolos IP, TCP, UDP y SCTP, así como los principales conceptos asociados a Internet. Al lector interesado en profundizar más en las aplicaciones de Internet se le remite a la bibliografía [COME 95] [ELAH 01] [FORO 07a] [STAL 07] [MILL 01].

Protocolo IP

Su función es el suministro de paquetes de información mediante mensajes llamados datagramas (*Datagrams*), por lo que se dice que es un protocolo que no está orientado a la conexión (*Connectionless*). En este protocolo se especifica, mediante un conjunto de campos, la red (*Net ID*) y el computador (*Host ID*) al que se envía información [SIYA 96].

Protocolos TCP, UDP y SCTP

El protocolo TCP es el responsable de enlazar los datagramas para constituir un paquete (*Segment*) y viceversa (Figura A1.62). También se encarga de detectar los paquetes perdidos y de reenviarlos. Opera en los nodos extremos de un canal de comunicación de una red de datos y complementa al protocolo IP que selecciona la ruta más adecuada para cada uno de los paquetes que forman parte de una transacción.

El protocolo UDP es un protocolo de transporte que sólo proporciona comunicación entre procesos (*Process to process communication*), en lugar de entre computadores (*Host to Host communication*) y realiza una verificación de errores más sencilla que el protocolo TCP. Es más útil que TCP para enviar mensajes de tamaño reducido, en los que no es necesario obtener la máxima fiabilidad (*Reliability*).

El protocolo SCTP es un protocolo de transporte fiable y orientado al mensaje, que combina las ventajas del TCP y el UDP.

Para estudiar con mayor profundidad los protocolos de transporte TCP, UDP y SCTP se remite al lector a la bibliografía [FORO 07b].

Intranet

De lo expuesto anteriormente se deduce que Internet no es más que un conjunto de redes que utilizan los protocolos TCP/IP y que se interconectan entre sí.

Pero son numerosas las entidades que poseen una red privada y desean a su vez comunicarse con otra mediante los protocolos de Internet. Esto hizo que surgiese el concepto de "Intranet" que es una red privada que utiliza el mismo protocolo TCP/IP que Internet para resolver la comunicación entre los diversos usuarios en el nivel de aplicación [BLAK 02] [CIMI 97].

La utilización de una intranet por parte de una entidad facilita el acceso controlado de los computadores de la misma a los otros computadores conectados a Internet, tal como se indica en la figura A1.63. Dicho acceso se realiza a través de sistemas pasarela (descritos en el apartado A1.3.4 a continuación) que, además, deben hacer la función de cortafuegos (*Firewall*). Un cortafuegos es una combinación de hardware (un sistema de interconexión entre redes) con un programa que analiza, de acuerdo con determinados criterios de seguridad, todos los paquetes de información que entran y salen de una red privada, para impedir el acceso a la misma de usuarios no autorizados, así como la salida de información confidencial contenida en los procesadores de la misma [BLAK 02] [ELAH 01].

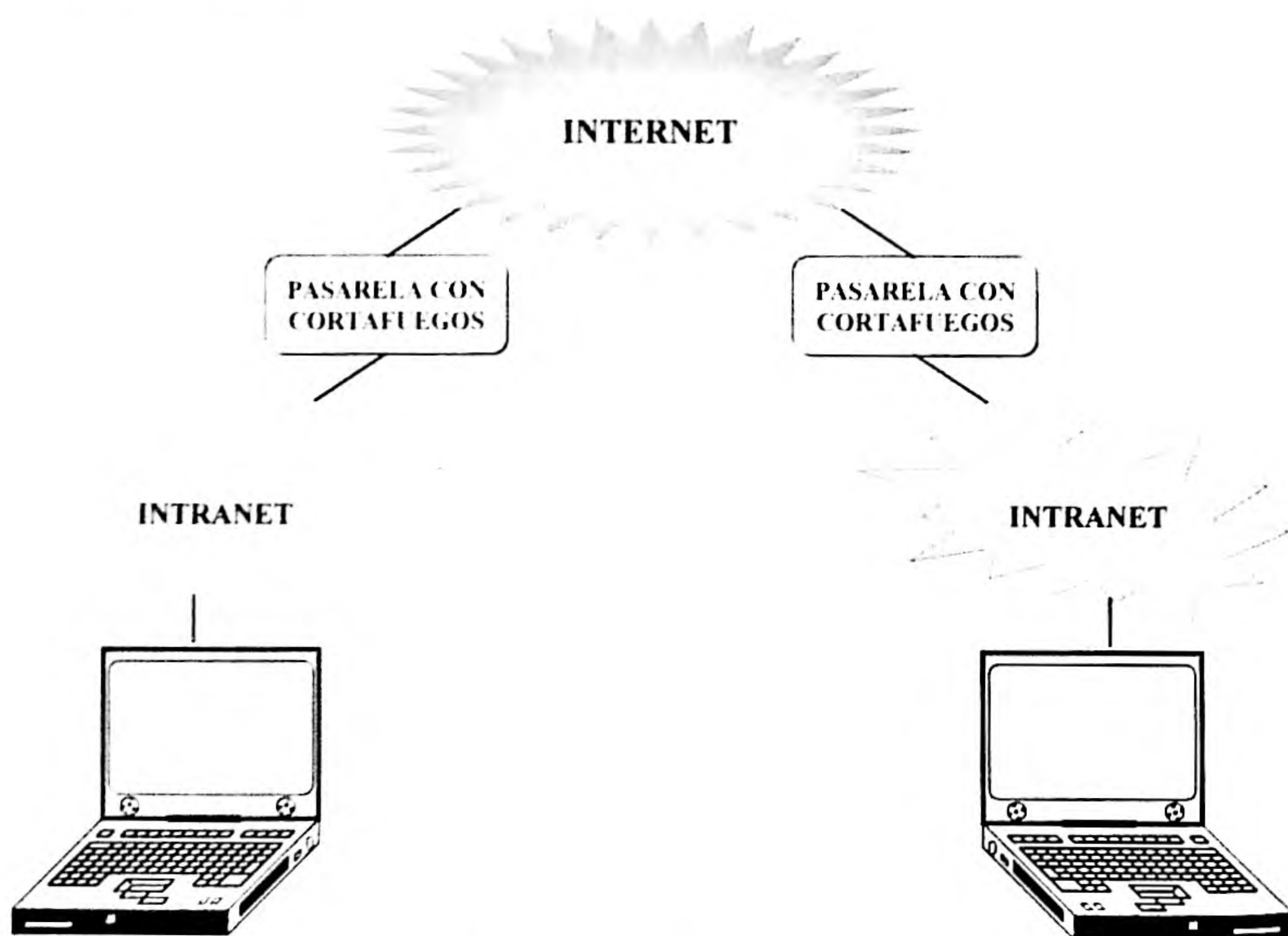


Figura A1.63. Transferencia de datos entre dos intranets a través de Internet.

Conexión de usuarios individuales a Internet

Cada computador que se conecta a Internet debe tener asignada una dirección de 32 bits (*IP address*) dividida en tres campos: el número del propio computador, el de la red a la que pertenece y el de la posible subred dentro de ella. Las direcciones numéricas se convierten en combinaciones de letras mediante computadores denominados “Servidores de nombres de dominio” (*Domain Name Servers*) que tienen asociados un conjunto de computadores dedicados a este fin. Las entidades que tienen una red interna de computadores disponen, en general, de un servidor de nombres de dominio que hace la función de una pasarela a través de la cual se conectan a Internet. Los usuarios individuales que quieren conectarse a Internet deben hacerlo a través de un suministrador de servicios de Internet (*Internet Services Provider*), tal como se indica en la figura A1.64.

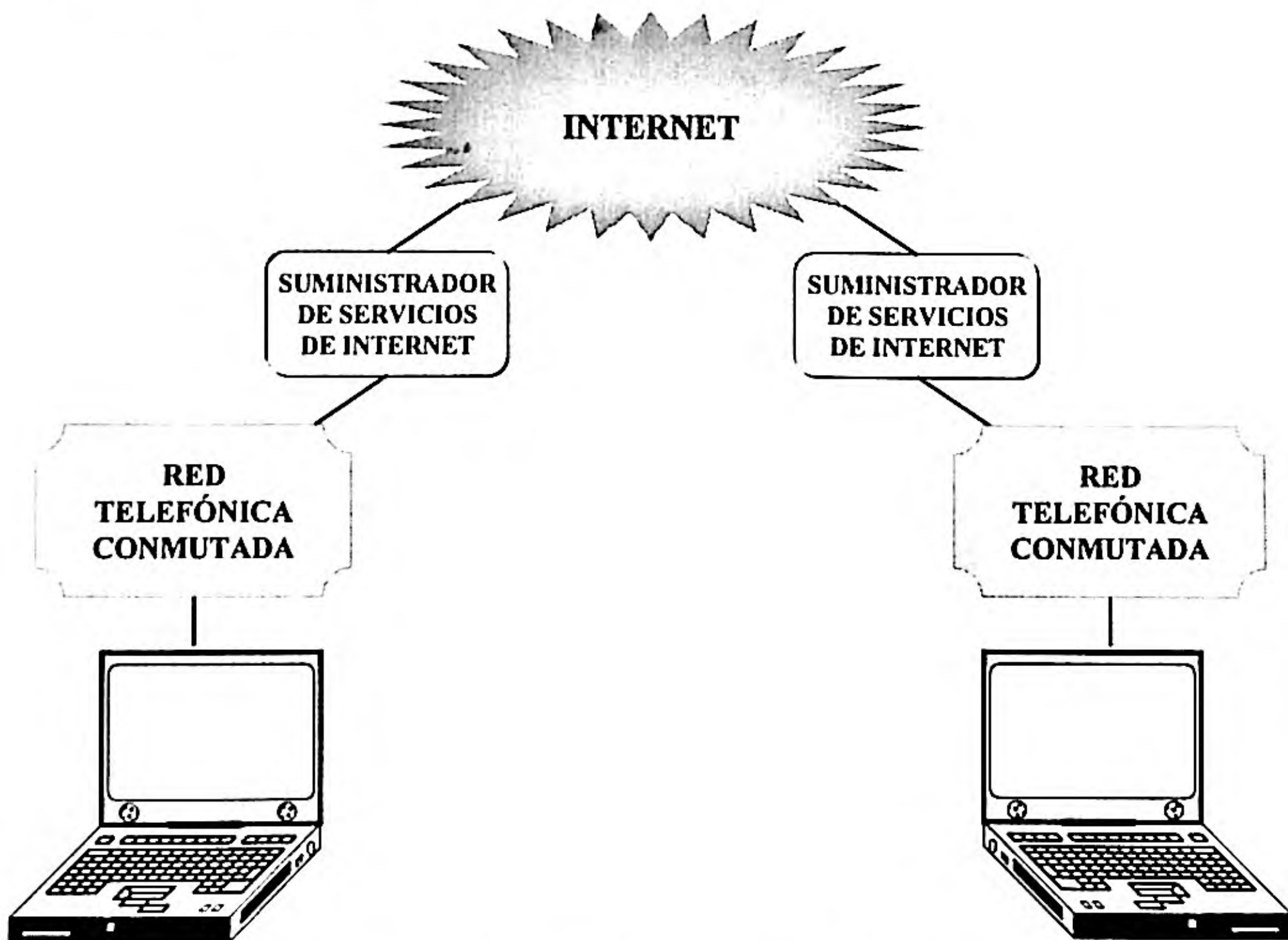


Figura A1.64. Diagrama de bloques de la transferencia de datos entre dos usuarios a través de Internet y de la Red Telefónica Conmutada, por medio de un subministrador de servicios de Internet.

Integración de la voz en Internet

Mediante pasarelas, descritas en el apartado A1.3.4 a continuación, es posible facilitar la comunicación entre redes de voz, como por ejemplo las telefónicas conmutadas, a través de una red de datos, tal como se indica en la figura A1.65. Para ello, la pasarela que conecta un generador de voz a la red de datos convierte las señales analógicas en digitales y la pasarela que recibe

la información procedente de la red de datos la convierte de nuevo en una señal analógica que se aplica a un altavoz.

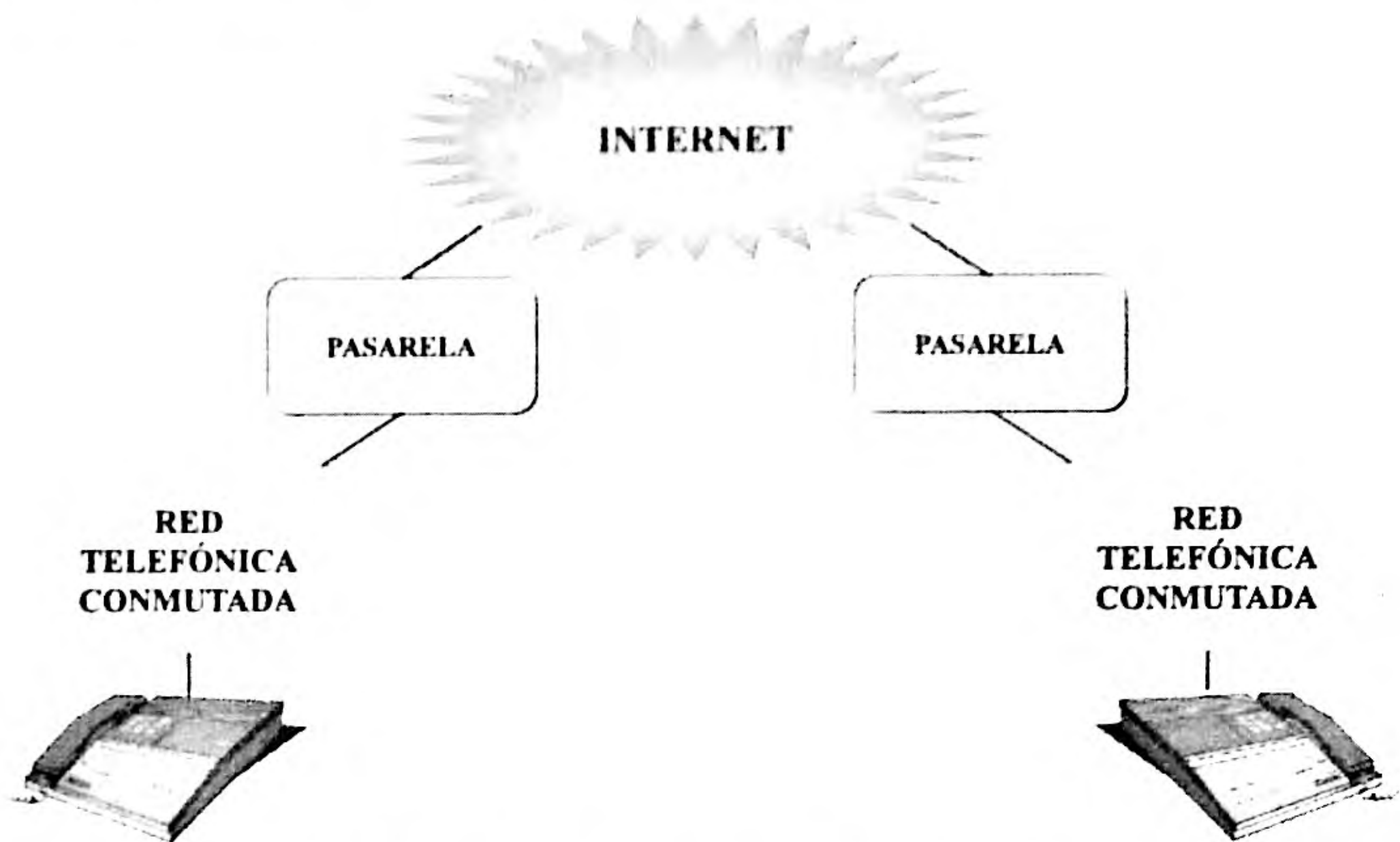


Figura A1.65. Diagrama de bloques de la transferencia de voz a través de Internet.

A1.3.4 Interconexión de redes

De lo expuesto en el apartado anterior se deduce que es habitual tener que interconectar entre sí redes de datos (*Data networks Interconnection or Internetworking*) que poseen características distintas en cuanto al tipo de datos que generan, la velocidad con la que los transmiten, así como los protocolos de las distintas capas del modelo OSI utilizados, etc.

En particular, en muchas aplicaciones, se tienen que conectar entre sí redes de área local diferentes y el conjunto a su vez a una red de área extensa mediante un conjunto de elementos o subsistemas de interconexión (*Connecting devices*) cuya complejidad depende de las diferencias entre las redes que se interconectan. De acuerdo con el nivel más alto del modelo OSI que utilizan, los elementos de interconexión se dividen en cuatro categorías (Figura A1.66):

- Elementos de la capa física.
- Elementos de la capa de enlace de datos.
- Elementos de la capa de red.
- Elementos de las capas de transporte y aplicación.

A continuación se analizan los más importantes y el lector puede observar que la complejidad de estos elementos y la gran cantidad de formas diferentes de implementarlos lleva a que no existan denominaciones normalizadas aceptadas de forma generalizada. Un ejemplo de ello son las palabras inglesas *Hub* y *Switch*.

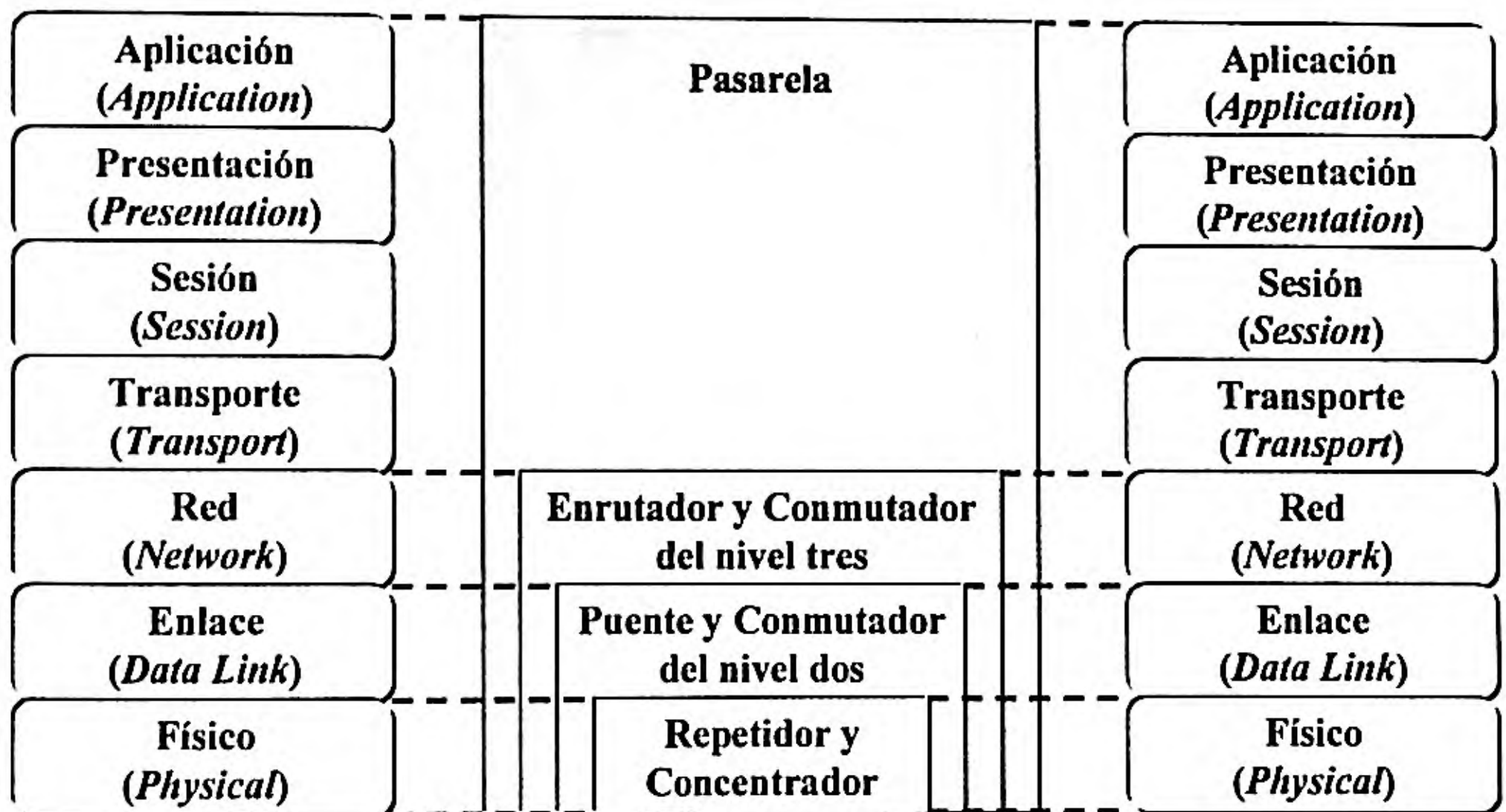


Figura A1.66. Categorías de los elementos de interconexión.

Elementos de interconexión de la capa física

Repetidor

Un repetidor (*Repeater*) es un elemento de interconexión que enlaza dos redes idénticas. Tiene como misión, en general, regenerar los niveles de tensión y la duración de los bits cuando se eleva la distancia entre los equipos que deben formar parte de una red de área local. Por ello, tal como se indica en la figura A1.66, el enlace se realiza al nivel de la capa física del modelo OSI.

Los repetidores sólo realizan la función de regeneración de las señales eléctricas y por ello carecen de capacidad de almacenamiento de información. Se utilizan en las redes de control de las Comunicaciones Industriales, como por ejemplo las redes PROFIBUS descritas en el apéndice 3 (véase el apartado A3.6.5).

Concentrador

Un concentrador (*Hub*) es un repetidor multipuerto, es decir, que realiza el enlace entre sistemas informáticos en el nivel físico. Por ello sus principales aplicaciones son:

- Enlazar entre sí varios computadores de una red de área local en estrella en lugar de a través de un bus.
- Enlazar un computador con varios periféricos que utilizan la misma norma de comunicación, como por ejemplo USB (Figura A1.67).

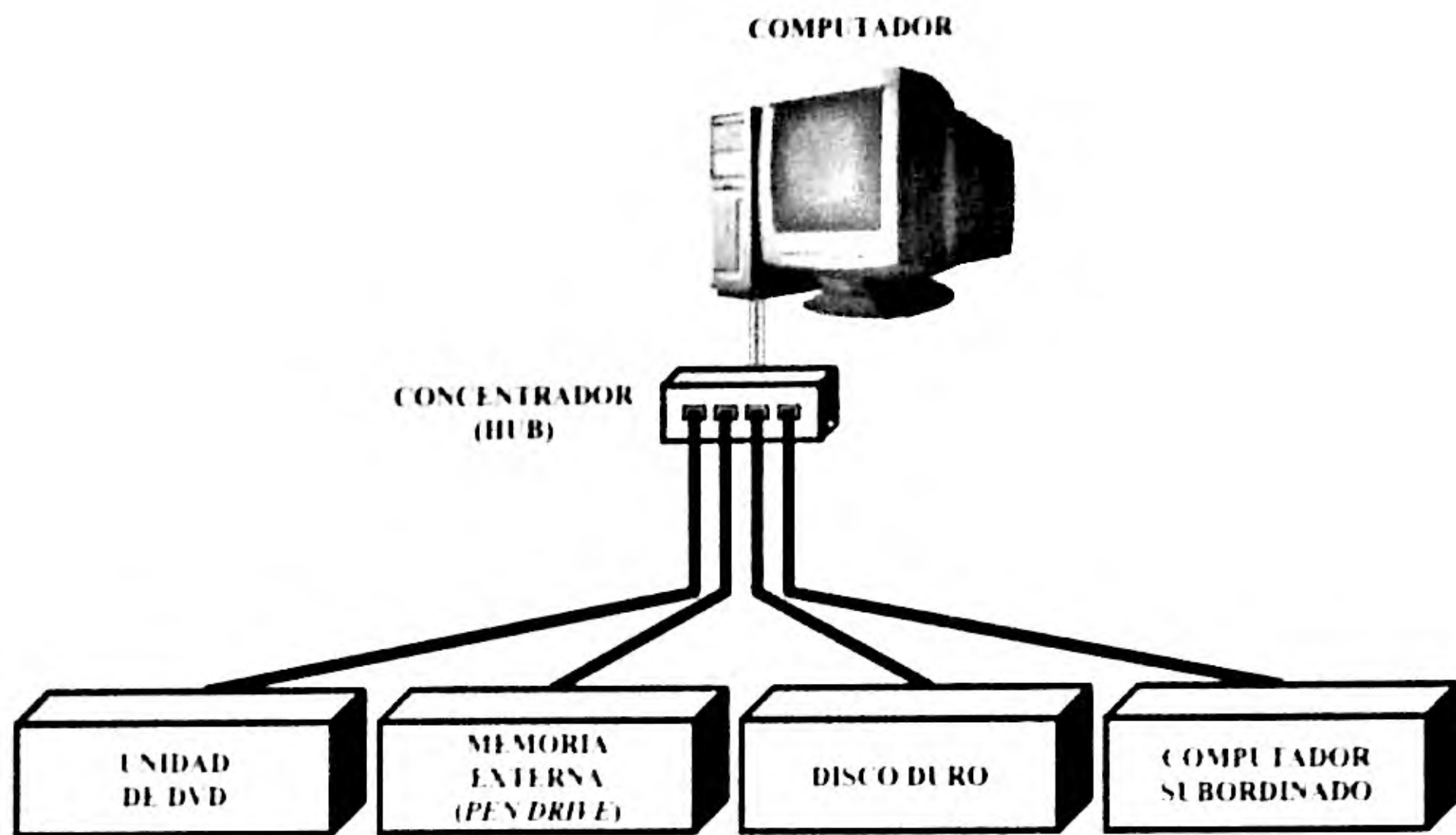


Figura A1.67. Diagrama de bloques de un sistema informático que utiliza un concentrador (Hub) para enlazar varios periféricos con un procesador.

Elementos de interconexión de la capa de enlace de datos

Puente

Un puente (*Bridge*) es un elemento de interconexión que realiza las funciones necesarias para unir entre sí dos redes no sólo a nivel físico sino también a nivel de enlace de datos. Por lo tanto, un puente adapta, tal como se indica en la figura A1.66, además de la capa física, las capas de enlace de datos de dos redes que utilizan el mismo protocolo o protocolos diferentes.

Cuando conectan redes que poseen el mismo protocolo de enlace de datos, el objetivo es dividir una red de topología bus en dos redes separadas y se suelen denominar puentes transparentes (*Transparent bridges*) [FORO 07a]. Un ejemplo típico de aplicación de este tipo de puentes es la red Ethernet, tal como se indica en el apartado A1.3.5.1.

Cuando conectan redes que poseen protocolos de enlace de datos diferentes se denominan puentes traductores (*Translating bridges*).

Conmutador de la capa de enlace de datos

Recibe la denominación de conmutador de la capa de datos o conmutador de la capa dos (*Two-layer switch*) un elemento de interconexión que posee múltiples puertos de entrada/salida, cada uno de los cuales puede estar conectado a un único sistema informático o a una red. En ocasiones se le denomina concentrador inteligente (*Intelligent hub*) [HELD 03] porque está basado en un procesador que lleva a cabo la función de conectar un único computador a varios procesadores que no intercambian información directamente entre ellos. Un ejemplo típico de conmutador de la capa de datos es el sistema que se conecta a la salida de un puerto serie de un computador para ampliar el número de periféricos que se pueden conectar al mismo y por ello se le suele llamar también concentrador (*Hub*) [STAL 07].

Elementos de Interconexión de la capa de red

Conmutador de la capa de red

Recibe el nombre de conmutador de la capa de red o conmutador de la capa tres (*Three-layer switch*) un elemento de interconexión que enlaza dos o más redes idénticas o diferentes a nivel de la capa de red, es decir, permite la transferencia de información entre procesadores pertenecientes a redes separadas. Algunos autores los denominan concentradores conmutadores o concentradores con capacidad de conmutación (*Switching hubs*) [HALL 05] [HELD 03] y cuando se utilizan con redes de área local se denominan conmutadores de red de área local (*LAN switches*) (en el caso de Ethernet, *Ethernet Switches*) [HELD 03]. En la actualidad suelen recibir el nombre simplemente de conmutadores (*Switches*) y poseen mecanismos para conmutar varios canales de comunicación simultáneamente.

Se puede, por lo tanto, expresar la diferencia entre un concentrador (*Hub*) y un conmutador (*Switch*) de la forma siguiente:

- Un concentrador (*Hub*) es un elemento de interconexión que recibe información a través de un puerto y la retransmite a todos los demás.
- Un conmutador (*Switch*), por el contrario, permite la transferencia directa de información entre todos los sistemas informáticos conectados a él, e incluso las transferencias entre pares de ellos simultáneamente.

Existe una multiplicidad de formas de implementar los mecanismos de conmutación de un conmutador, debido a que se pueden realizar mediante componentes físicos (*Hardware*), mediante un programa (*Software*) o mediante combinación de ambos. El progreso de la Microelectrónica ha hecho que unos de los más utilizados actualmente sean los componentes físicos denominados matrices de interconexión programables, utilizadas en los circuitos digitales configurables del tipo dispositivo lógico programable (PLD) y conjuntos configurables de puertas (FPGA) [MAND 08c], combinándolos con un microprocesador o con un microcontrolador. Los conmutadores especializados en algunos tipos de redes industriales se implementan con circuitos integrados de aplicación específica (ASIC) o con matrices de puertas configurables (FPGA). Un ejemplo de ello son los descritos en el apéndice 4, dedicados a implementar la red industrial universal Profinet.

Enrutador

Un enrutador o encaminador (*Router*) es un sistema informático conmutador de paquetes que trabaja en el nivel de red y, como su nombre indica, determina la mejor ruta que deben seguir los paquetes en una red de área extensa para alcanzar su destino. Utiliza la información del campo de dirección de los paquetes para pasar mensajes de una parte de una red a otra. En la práctica su función es similar a la de un conmutador de la capa de red y por ello también se le denomina, a veces, simplemente, conmutador (*Switch*) [FORO 07a].

Suelen recibir la denominación de enrutadores o encaminadores, las pasarelas, descritas a continuación, que realizan las funciones de intercambio de información entre un usuario o una red privada e Internet.

Elementos de interconexión de la capa de transporte y aplicación

Pasarela

Se reserva el nombre de pasarela (*Gateway*) para denominar a los elementos de interconexión que enlazan dos redes a nivel de las capas de transporte y aplicación, tal como se indica en la figura A1.66. Una mención especial merecen las pasarelas cliente-servidor conocidas bajo el nombre inglés de *Proxy*, que son recursos informáticos que según la aplicación están constituidos por un programa [STAL 07] o por una combinación sistema físico-programa (*Hardware-Software*). Un ejemplo de este tipo de elementos de interconexión son los sistemas que enlazan un bus de campo, como por ejemplo PROFIBUS, con una red industrial universal, como por ejemplo Profinet, que se estudia en el apéndice 4.

A los lectores interesados en estudiar con mayor profundidad los diferentes elementos de interconexión se les remite a la bibliografía al final del capítulo [FORO 07a] [HALL 05] [HELD 03] [STAL 07].

A1.3.5 Redes de área local conmutadas

A1.3.5.1 Introducción

Las redes de área local, descritas en el apartado A1.3.3.2, se caracterizan por compartir un único canal de comunicaciones y por ello no son adecuadas, especialmente la red Ethernet original correspondiente a la norma 802.3 [conocida como Ethernet normalizada (*Standard Ethernet*)], cuando se tiene que intercomunicar un conjunto elevado de procesadores repartidos en un área de cierta extensión, como sucede en una instalación industrial o en un conjunto de edificios, porque el tiempo de respuesta a una petición de transmisión no es determinista y se eleva hasta alcanzar valores inaceptables. Por ello, en las redes de área local modernas la topología de bus ha sido sustituida, en parte o totalmente, en función de la aplicación, por la arquitectura en estrella mediante la utilización de los elementos de interconexión adecuados.

Una de las formas de disminuir el tiempo de respuesta de una red Ethernet es la disminución del tiempo de colisión mediante la utilización de puentes. En la figura A1.68 se representa una red Ethernet formada por dos subredes enlazadas a través de un puente (*Bridged Ethernet*).



Figura A1.68. División de una red Ethernet en dos mediante un puente (*Bridge*).

Pero, tal como se indica en el apartado A1.3.4, un puente o un concentrador (*Hub*) operan en las capas físicas y de enlace de datos y no gestionan la comunicación a nivel de red. Esto provocó el interés por desarrollar redes Ethernet conmutadas, que se describen a continuación.

A1.3.5.2 Red Ethernet conmutada

La red Ethernet conmutada (*Switched Ethernet*) utiliza elementos de interconexión del tipo conmutador (*Switches*) de la capa tres, descritos en el apartado A1.3.4, para sustituir al bus utilizado en la red Ethernet estándar. Se tiene de esta forma una topología estrella como la representada en la figura A1.69, que evita las colisiones de la topología tipo bus y garantiza un tiempo de respuesta máximo a las solicitudes de transferencia de información.

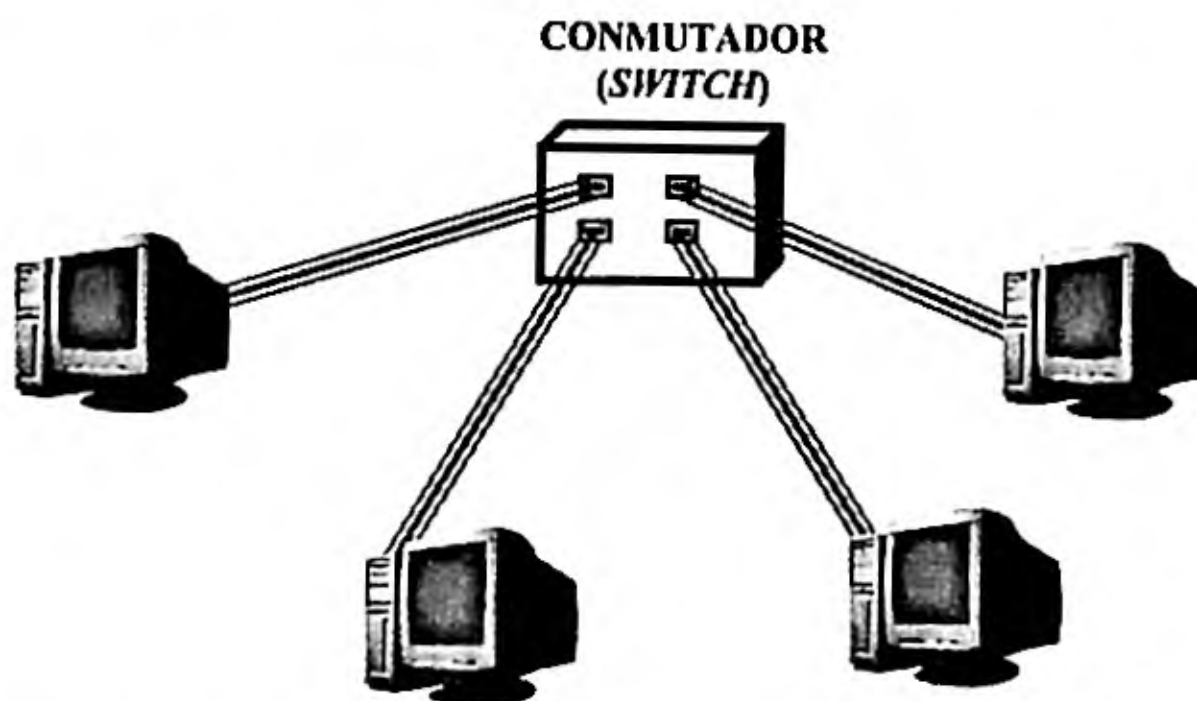


Figura A1.69. Red Ethernet conmutada de topología estrella implementada con un conmutador (Switch).

Por otra parte, en la red Ethernet estándar de la norma 802.3, las transferencias a través del bus constituyen una comunicación que se realiza en el modo semidoble (*Half duplex*), descrito en el apartado A1.2.2, porque cualquier computador de la red puede recibir o transmitir una trama en cualquier instante pero no puede realizar ambas acciones simultáneamente. La red Ethernet de la figura A1.69 permite, utilizando un conmutador que posea los recursos físicos y el programa adecuados, realizar una comunicación en modo doble simultáneo (*Full duplex*) e incluso comunicaciones simultáneas entre más de dos computadores, de acuerdo con la norma IEEE - 802.3X. Además la norma IEEE - 802.3Q establece un máximo de siete niveles de prioridad de las tramas, que facilita la implementación de conmutadores que gestionan la transmisión de tramas de forma priorizada. Un ejemplo de ello son los conmutadores de la red Profinet descrita en el apéndice 4.

Mediante conmutadores se puede también implementar una red de topología en árbol de varios niveles. En la figura A1.70 se representa, como ejemplo, una red de topología árbol de dos niveles en la que la capacidad de un conmutador se eleva mediante su conexión a dos conmutadores, a los que se conectan a su vez los computadores que forman parte de la red.

Es posible también implementar una red híbrida como la representada en la figura A1.71, en la que el conmutador está conectado a dos o más redes de tipo bus. La función del conmutador es superior a la del puente de la figura A1.68.

La combinación de Ethernet con los protocolos TCP e IP, correspondientes respectivamente a las capas de transporte y de red del modelo OSI, junto con los conmutadores, da lugar a las redes Ethernet conmutadas utilizadas en aplicaciones de control industrial y conocidas bajo la denominación Ethernet Industrial (*Industrial Ethernet*), que se describen en el apartado 9.3.2.5 del capítulo 9.

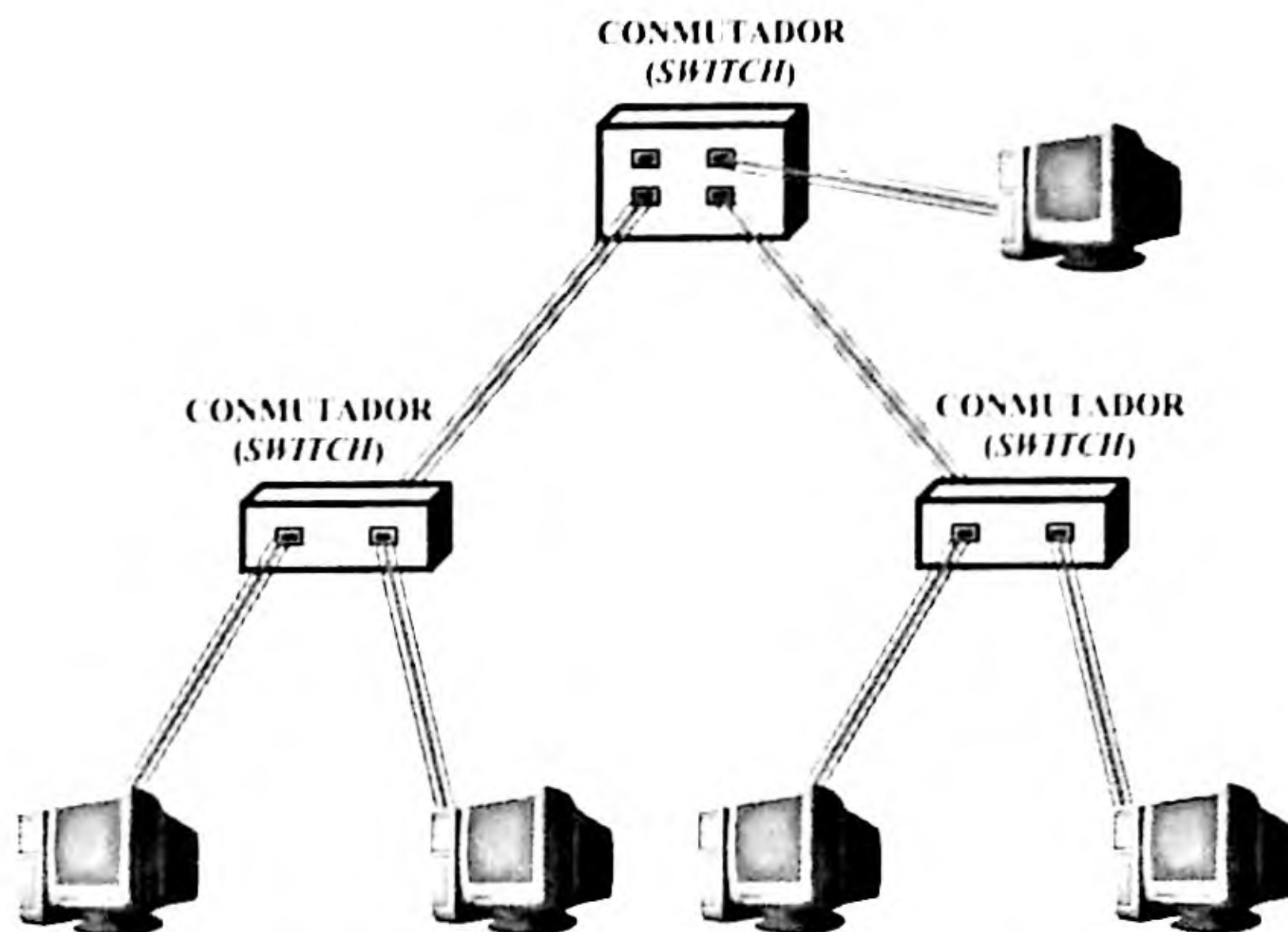


Figura A1.70. Red Ethernet conmutada de topología árbol de dos niveles implementada con conmutadores (Switches).

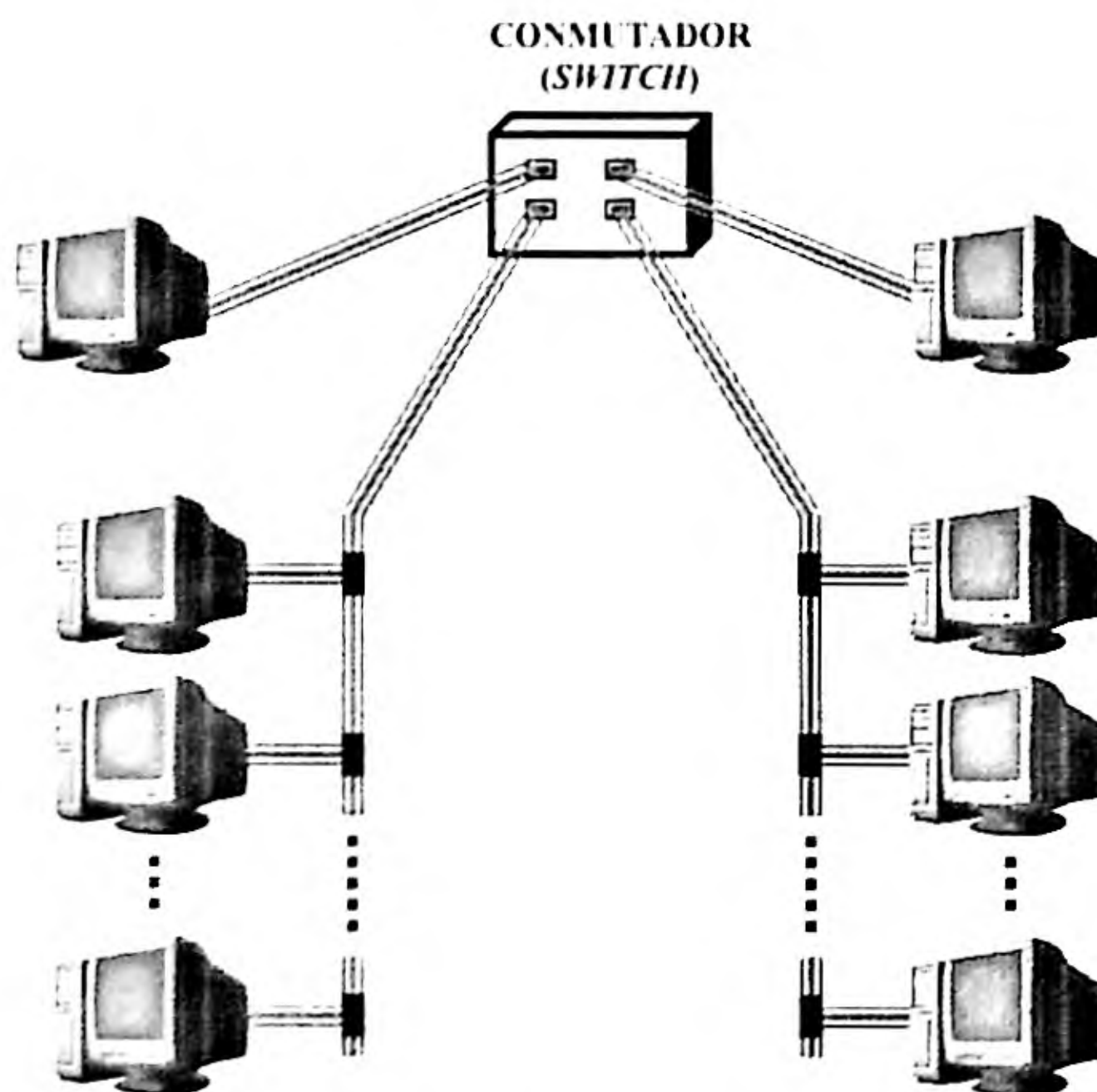


Figura A1.71. Red Ethernet conmutada híbrida que combina la topología estrella implementada con un conmutador con la topología bus.

A1.4 Infocomunicaciones

A1.4.1 Introducción

A partir de la comercialización de los primeros microprocesadores, a principios de la década de 1970, se produjo la explosión de la informática, en especial a partir de la comercialización del computador personal, vulgarmente conocido como PC (acrónimo de *Personal Computer*), por parte de IBM en 1981.

En los apartados anteriores se estudian las redes de datos y sus principales características, pero mucho antes de que la Microelectrónica hiciese posibles las comunicaciones digitales se desarrollaron la telefonía, la radio y la televisión, que se caracterizaron por transmitir señales analógicas (descritas en el apartado A1.2.3 del apéndice 1), correspondientes a la voz las dos primeras y a la imagen la tercera. La conversión de las señales analógicas en digitales, previamente a su transmisión, está produciendo una integración de todos los sistemas de comunicación a distancia y da lugar a una nueva área de la tecnología, que surge como resultado de la integración de las anteriores y se puede denominar Infocomunicaciones.

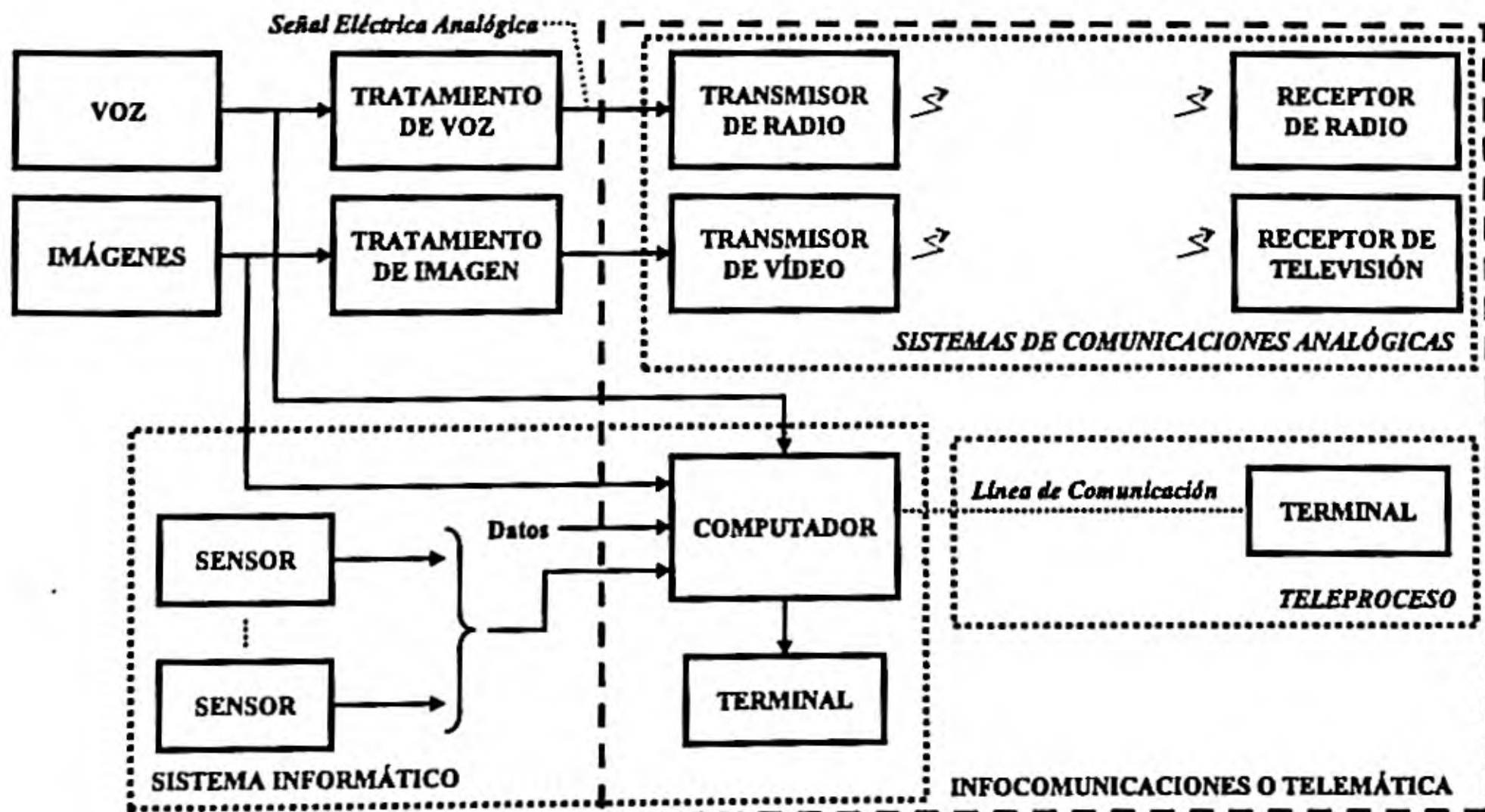


Figura A1.72. Modelo inicial de integración de las Infocomunicaciones.

Las Infocomunicaciones se pueden definir como un conjunto de técnicas que relacionan entre sí la Informática, las Telecomunicaciones y los Medios Audiovisuales, para dar lugar a un conjunto de servicios que permiten implantar un elevado número de aplicaciones. En la figura A1.72 se indica el concepto inicial de Infocomunicaciones, nacido como resultado de la integración de los sistemas de comunicaciones analógicas y del teleproceso, que implementa la conexión a distancia de los periféricos de un sistema informático. En la actualidad las Infocomunicaciones integran las señales analógicas de voz e imágenes convertidas en digitales

A1.4 Infocomunicaciones

A1.4.1 Introducción

A partir de la comercialización de los primeros microprocesadores, a principios de la década de 1970, se produjo la explosión de la informática, en especial a partir de la comercialización del computador personal, vulgarmente conocido como PC (acrónimo de *Personal Computer*), por parte de IBM en 1981.

En los apartados anteriores se estudian las redes de datos y sus principales características, pero mucho antes de que la Microelectrónica hiciese posibles las comunicaciones digitales se desarrollaron la telefonía, la radio y la televisión, que se caracterizaron por transmitir señales analógicas (descritas en el apartado A1.2.3 del apéndice 1), correspondientes a la voz las dos primeras y a la imagen la tercera. La conversión de las señales analógicas en digitales, previamente a su transmisión, está produciendo una integración de todos los sistemas de comunicación a distancia y da lugar a una nueva área de la tecnología, que surge como resultado de la integración de las anteriores y se puede denominar Infocomunicaciones.

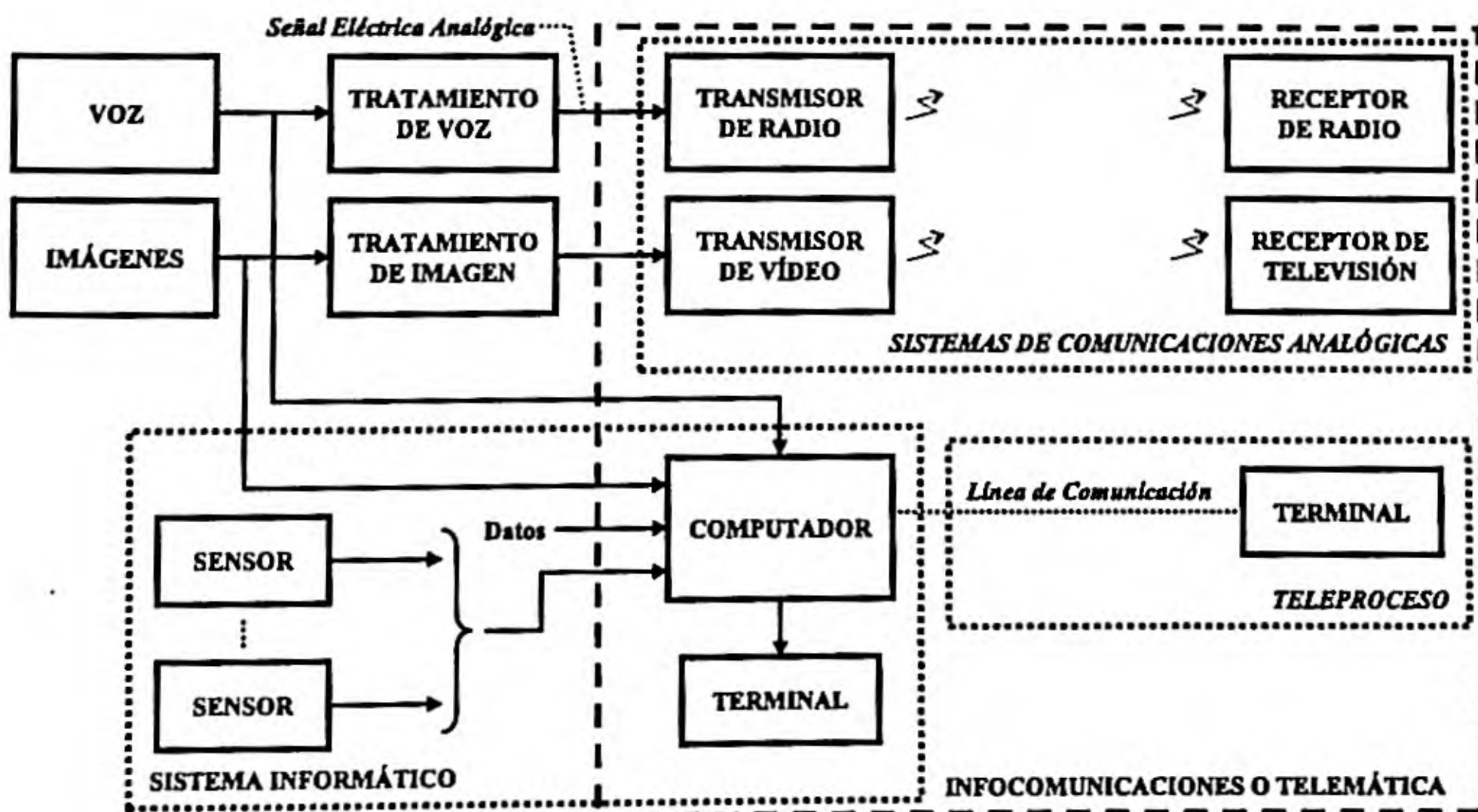


Figura A1.72. Modelo inicial de integración de las Infocomunicaciones.

Las Infocomunicaciones se pueden definir como un conjunto de técnicas que relacionan entre sí la Informática, las Telecomunicaciones y los Medios Audiovisuales, para dar lugar a un conjunto de servicios que permiten implantar un elevado número de aplicaciones. En la figura A1.72 se indica el concepto inicial de Infocomunicaciones, nacido como resultado de la integración de los sistemas de comunicaciones analógicas y del teleproceso, que implementa la conexión a distancia de los periféricos de un sistema informático. En la actualidad las Infocomunicaciones integran las señales analógicas de voz e imágenes convertidas en digitales

con las digitales que constituyen datos procedentes de diversas fuentes, como por ejemplo un teclado, un conjunto de sensores todo/nada de un proceso productivo u otro tipo de actividades empresariales. En la figura A1.73 se representa un ejemplo de sistema distribuido de Infocomunicaciones.

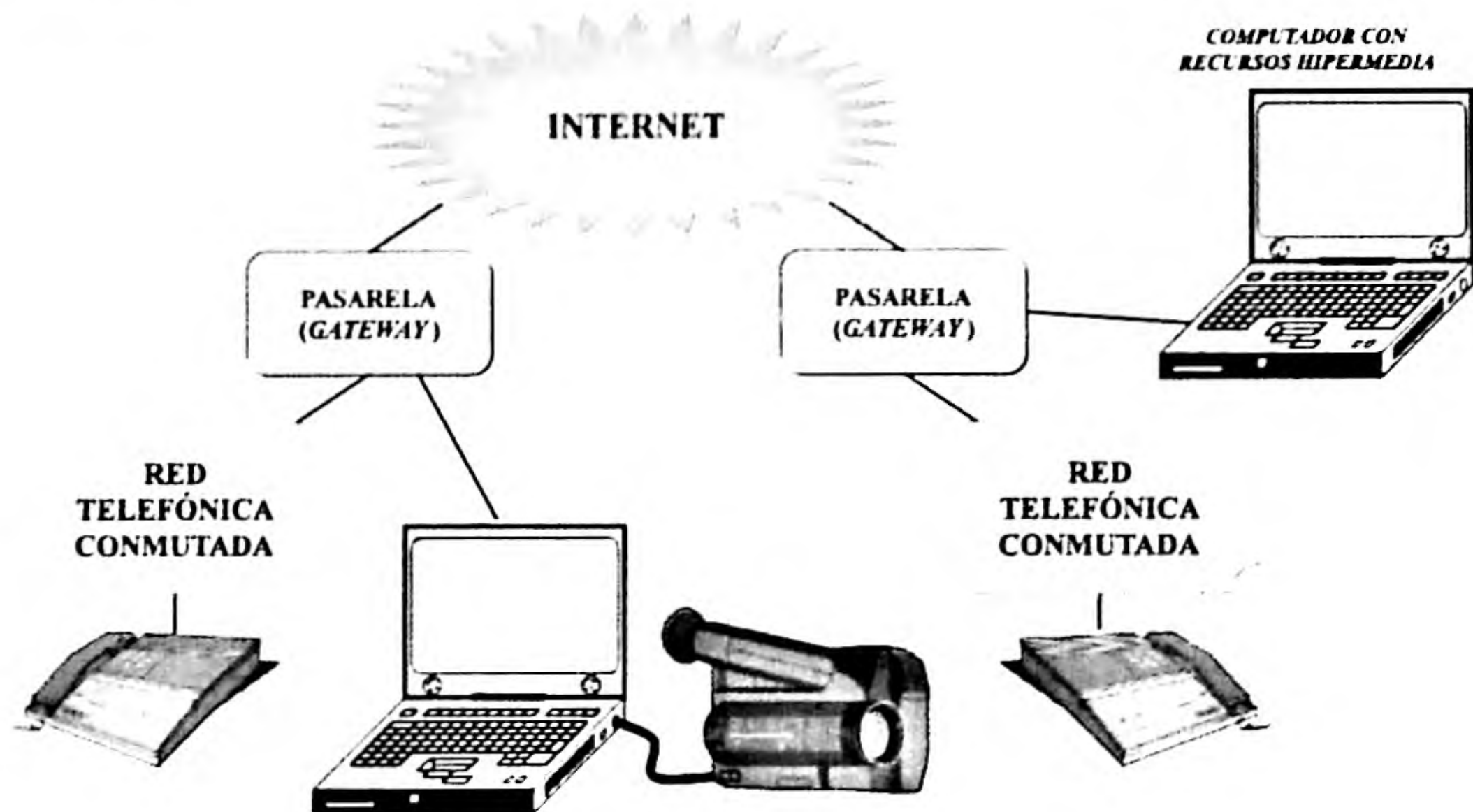


Figura A1.73. *Ejemplo de sistema de Infocomunicaciones.*

En apartados anteriores se analizan las técnicas de las Infocomunicaciones directamente relacionadas con las comunicaciones digitales, especialmente en banda base, más utilizadas en las Comunicaciones Industriales. Un estudio más profundo de todas las técnicas utilizadas en las Infocomunicaciones, como por ejemplo los sistemas de comunicación inalámbrica (*Wireless Communication Systems*), se sale fuera de los límites de este apéndice y se remite al lector a la bibliografía indicada al final del mismo [FORO 07a] [STAL 07].

Las Infocomunicaciones se caracterizan por:

- Su complejidad es elevada, debido a que combinan gran cantidad de conceptos interrelacionados que presentan múltiples variantes: medios de transmisión, topologías, tipo de información, etc.
- Su evolución es vertiginosa, lo que hace que la implantación de soluciones deba realizarse rápidamente para garantizar su rentabilidad.
- Su utilización es estratégica. Los costes asociados con ellas pueden constituir uno de los capítulos importantes del presupuesto de una empresa.

A1.4.2 Aplicaciones de las Infocomunicaciones

El número de aplicaciones de las Infocomunicaciones no es cuantificable y se pueden clasificar en generales y sectoriales, que se estudian en sucesivos apartados.

A1.4.2.1 Aplicaciones generales de las Infocomunicaciones

Se consideran aplicaciones generales las que se pueden utilizar en la gran mayoría de los sectores industriales. A continuación se analizan algunas de las más importantes.

Almacenes de datos

Un almacén de datos (*Data Warehouse*) es un conjunto integrado de bases de datos diseñado para facilitar la toma de decisiones mediante el acceso a información relevante adecuadamente clasificada. Un almacén de datos se caracteriza por memorizar la información mediante hipercubos en los que cada una de las dimensiones constituye una variable relacionada con las operaciones de la empresa. La explotación de un almacén de datos se denomina proceso analítico de datos en tiempo real y se conoce por las siglas OLAP (acrónimo de *On Line Analytical Processing*).

Un estudio más detallado de los almacenes de datos se sale de los límites de este libro y se remite al lector a la bibliografía [ALAR 03].

Minería de datos

La minería de datos (*Data Mining*) es un proceso automático de detección y cuantificación de las relaciones entre los datos de un almacén de datos, para describir y predecir conductas y patrones que enriquezcan y faciliten la toma de decisiones. Para ello, una vez identificadas las variables que pueden explicar el funcionamiento de determinada área de actividad de una empresa, se toman muestras aleatorias, se analizan y se sacan conclusiones. Tiene gran importancia en aquellas actividades en la que influyen numerosos factores y no es posible encontrar una ley matemática que establezca la relación entre ellos. Un ejemplo típico es la relación entre la probabilidad de devolución de un crédito, diversos datos de sus clientes como la edad, el sexo, el tipo de estudios, el área geográfica en la que residen, etc. [ALAR 03].

Correo electrónico

El correo electrónico (*e-mail*) es una herramienta informática que permite a un usuario de un computador conectado a Internet elaborar un mensaje y enviarlo a otro usuario identificado mediante una combinación determinada de caracteres.

Comercio electrónico

El comercio electrónico (*e-commerce*) se puede definir como el conjunto de herramientas informáticas que propician las transacciones comerciales a través de Internet. Tiene sus antecedentes en el intercambio electrónico de datos, denominado EDI (*Electronic Data Interchange*), entre clientes y proveedores, en especial entre empresas, iniciado a finales de la década de 1980.

Para su implementación es necesario disponer de un servidor de Internet, propio o proporcionado por un proveedor de servicios en Internet.

Hipermedia

El término hipermedia es el resultado de la combinación de los términos hipertexto y multimedia. El hipertexto es un sistema constituido exclusivamente por textos que se enlazan a

través de palabras activas (*Hot words*) sobre las que puede actuar un usuario por medio de un periférico. La multimedia es un sistema formado por información de sonidos, video, gráficos estáticos y animados, y textos.

Se puede, por tanto, definir un sistema hipermedia como un programa y un conjunto de datos asociados con él que hacen que un computador pueda mostrar gráficos, videos, sonidos, textos y animaciones e integrarlos en un único entorno sobre el que puede actuar un usuario para obtener un resultado visible, audible o ambas cosas.

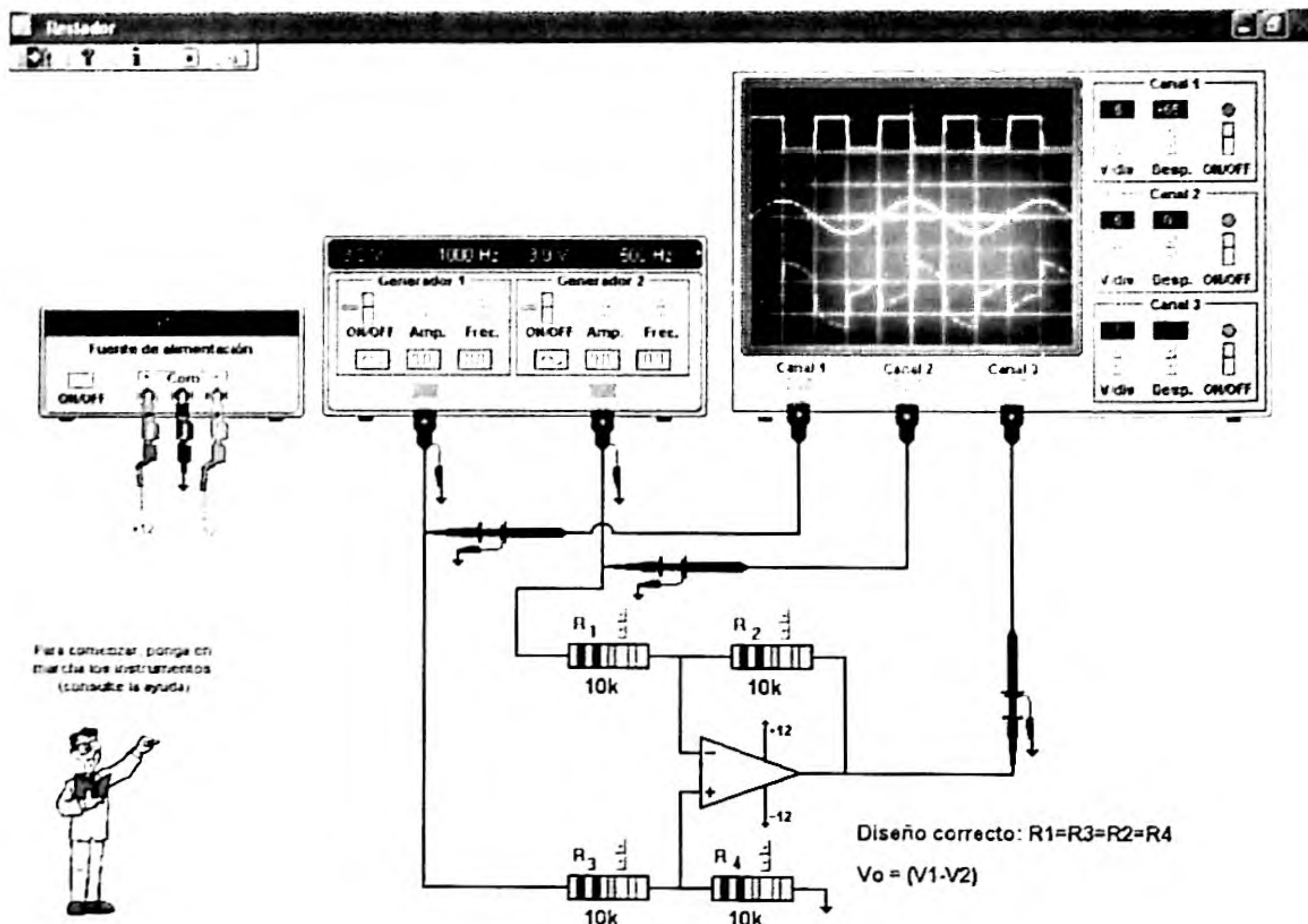


Figura A1.74. *Simulador básico de Electrónica Analógica que representa un circuito restador en el que se utilizan dos generadores de señales y un osciloscopio (cortesía del Departamento de Electrónica y Comunicaciones de la Universidad del País Vasco y del Departamento de Tecnología Electrónica y del Instituto de Electrónica Aplicada de la Universidad de Vigo).*

La hipermedia constituye una herramienta que está llamada a revolucionar los métodos de enseñanza y aprendizaje utilizados hasta el presente basados en la clase magistral presencial. En la figura A1.74 se representa, como ejemplo, la interfaz de usuario de un simulador básico de Electrónica Analógica, que permite comprobar el funcionamiento de los distintos circuitos realizados con amplificadores operacionales y comprobar la influencia en su funcionamiento de la modificación de las resistencias utilizadas en ellos. En la figura A1.75 se representa un laboratorio virtual que facilita al alumno el aprendizaje de las técnicas de conexión de los circuitos electrónicos para facilitar su experimentación en el laboratorio.

Un análisis en profundidad de los sistemas hipermedia y de los recursos de computación que permiten desarrollarlos se sale de los límites de este libro y por ello se remite al lector a la bibliografía [SALA 03].

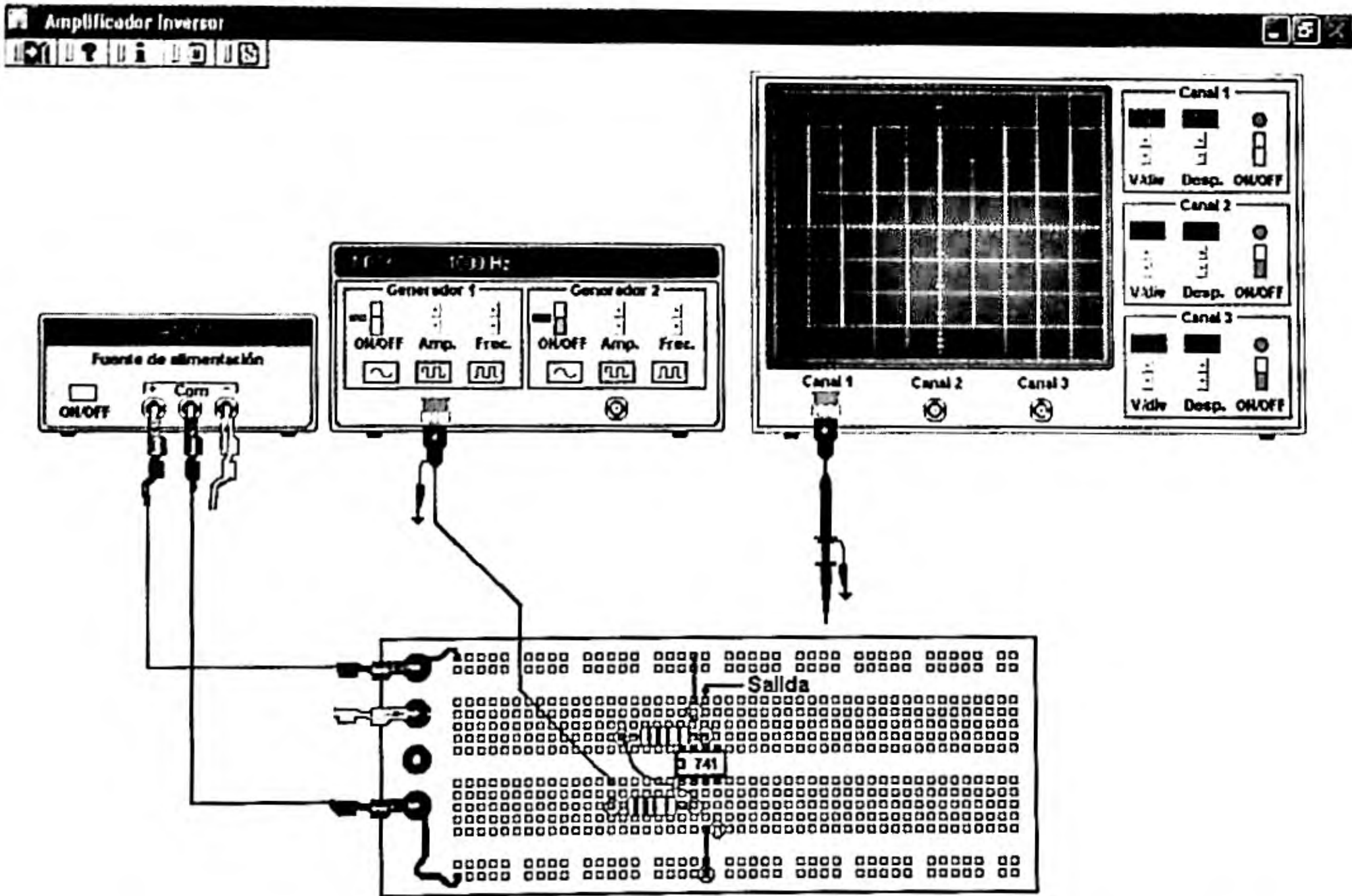


Figura A1.75. Laboratorio virtual de Electrónica Analógica que representa un amplificador inversor en el que se utiliza un generador de señales y un osciloscopio (cortesía del Departamento de Electrónica y Comunicaciones de la Universidad del País Vasco y del Departamento de Tecnología Electrónica y del Instituto de Electrónica Aplicada de la Universidad de Vigo).

A1.4.2.2 Aplicaciones sectoriales de las Infocomunicaciones

Se consideran aplicaciones sectoriales las que se pueden utilizar en determinados tipos de empresas o sectores industriales. A continuación se indican algunas de las más importantes.

Comunicaciones Industriales

Las Comunicaciones Industriales (*Industrial Communications*) constituyen el área de las Comunicaciones Digitales, que estudia la transmisión de información entre circuitos y sistemas electrónicos digitales, utilizados para llevar a cabo tareas de control y gestión del ciclo de vida de los productos industriales. Su combinación con los almacenes de datos y con la minería de datos forma parte de la estrategia de fabricación integrada por computador de numerosas empresas industriales y por ello a su estudio se dedica el apartado 9.3 del capítulo 9.

Control de edificios

El control de edificios, que se suele denominar “Domótica” (término obtenido a partir de las palabras “Doméstico” e “Informática”) combina las comunicaciones digitales con el control por computador de los diferentes elementos utilizados en un edificio como son los interruptores eléctricos, los elementos de iluminación, calefacción, etc.

Bibliografía

- [ALAR 03] J. M. Alarcón. *Innovación y Tecnologías de la Información. Capítulo 6 del libro “La innovación tecnológica en las organizaciones”* Editorial Thomson Learning Paraninfo. 2003.
- [BLAK 02] R. Blake. *Electronic Communication Systems. Second edition.* Delmar Thomson Learning. 2002.
- [CIMI 97] J. D. Cimino. *Intranets.* Editorial Thomson Learning Paraninfo. 1997.
- [COME 95] D. E. Comer. *El libro de Internet.* Prentice Hall Hispanoamericana S.A. 1995.
- [ELAH 01] A. Elahi. *Network Communications Technology.* Delmar Thomson Learning. 2001.
- [FORO 07a] B. A. Forouzan. *Data Communications and Networking.* Fourth edition. Editorial Mc Graw Hill. 2007.
- [FORO 07b] B. A. Forouzan. *Data Communications and Networking. Chapter 23.* Fourth edition. Editorial Mc Graw Hill. 2007.
- [HALL 05] F. Hallsal. *Computer networking and the Internet.* 5th edition. Pearson 2005.
- [HELD03] G. Held. *Ethernet networks.* Fourth edition. John Wiley & Sons. 2003.
- [HUID 01] J. M. Huidobro y A. Blanco. *Redes de área local.* Editorial Thomson Learning Paraninfo. 2001.
- [MAND 08a] E. Mandado & Y. Mandado. *Sistemas electrónicos digitales. Capítulo 2. Álgebra de Boole.* Editorial Marcombo. 2008.
- [MAND 08b] E. Mandado & Y. Mandado. *Sistemas electrónicos digitales. Capítulo 1: Sistemas y códigos de numeración.* Editorial Marcombo. 2008.
- [MAND 08c] E. Mandado & Y. Mandado. *Sistemas electrónicos digitales. Capítulo 6: Tecnologías de implementación de los circuitos digitales.* Editorial Marcombo. 2008.
- [MILL 01] M.A. Miller. *Data and Network Communications.* Delmar Thomson Learning. 2001.
- [PALM 00] M. J. Palmer. *Redes Informáticas.* Paraninfo Thomson Learning. 2000.
- [PERE 02] M. A. Pérez, J. A. Álvarez, J. C. Campo, F. J. Ferrero y G. J. Grillo. *Instrumentación Electrónica.* Editorial Thomson Learning Paraninfo. 2003.
- [RAPP 02] T. Rappaport. *Wireless Communications.* Prentice Hall PTR. 2002.

- [SALA 03] A. Salaverría. *Tesis doctoral: Nueva metodología para la enseñanza asistida por computador de la Electrónica Aplicada*. Universidad del País Vasco. 2003.
- [SIYA 96] K. S. Siyan. *Windows NT Server 4 Professional Reference*. New Riders Publishing. 1996.
- [SLOM 87] M. Sloman & J. Kramer. *Distributed systems and computer networks*. Prentice Hall International. 1987.
- [SMIT 93] P. Smith. *Frame Relay*. Addison-Wesley Publishers. 1993.
- [STAL 07] W. Stallings. *Data and Computer Communications*. Eighth edition. Pearson Prentice Hall International Inc. 2007.
- [TISC 94] M. Tischer. *PC interno*. Editorial Marcombo. 1994.
- [VALV 00] J. W. Valvano. *Embedded microcomputers systems*. Brooks/Cole. 2000.
- [WALD 91] J. Waldrand. *Communications Networks: A first course*. Asksen associates Inc. 1991.
- [YOUL 02] Z. Youlu and S. Akhtar. *Networks for computers scientists and engineers*. Oxford University Press. 2002.