Appendix B

AUTOMATIC TRAIN CONTROL TECHNOLOGY"

The technology of automatic train control embraces many kinds of equipment and engineering techniques. All aspects of this technology cannot be adequately presented in a brief appendix such as this. Therefore, the discussion is confined to two major elements of train control technology: the track circuit and methods for speed command and control. The technology forms the basis for almost all automation of train protection, train operation, and train supervision functions. Of the two, the track circuit is the more basic. It was the first to be developed, and it underlies the operation of speed command and control systems. It is the fundamental method of train detection and, while there has been experimentation with other methods over the years, none has proven to be as effective and reliable as this electrical technique for determining the presence and location of transit vehicles. From this basic positional information, signal systems are operated, train protection is accomplished, train operation is controlled, and supervisory functions are carried out.

TRACK CIRCUITS

The track circuit is an electrical circuit which includes a length of running rails (or special rails) and permits detection of the presence of a train. A track circuit may also be used to communicate commands, instructions, or indications between the wayside and a train. Track circuits provide information on the location of the trains, and this information is used to command train speeds so that the trains operate safely. For instance, if a train attempts to approach too close to the rear of another train, information on the locations of the two trains, provided by the track circuits, is used to command a slowdown or stop of the following train before there is danger of a rear-end collision.

The basic d.c. track circuit was invented by Dr. William Robinson and first used in a railway application is 1872. Although the equipment and technology have changed considerably in their detail since that time, the basic principle has remained the same. An electrical signal of some kind is impressed between the two running rails, and the presence of a train is detected by the electrical connection between the two running rails provided by the wheels and axles of the train (wheel-to-rail shunting).[98]

Before proceeding to a discussion of the various types of track circuits, it must first be considered how track circuits are used in the operation of a transit system. A track circuit provides information on whether a train occupies a given length of track (a block). The occupancy information for a particular block and for contiguous upstream of blocks is used to control the operation of all trains within the given area, For instance, when a train is detected in a block, that occupancy information is used to cause a zero-speed command for the block immediately behind the train. Depending upon the block lengths, the line speeds involved, and the number of available speed commands, the second block behind the train may have a command speed between zero and full line speed. The third block behind the train may have a commanded speed greater than or equal to the second block, and so on. In all cases, the blocks behind a train are signaled so that a train entering a block has sufficient braking distance to leave the block at a speed not greater than the commanded speed. In the case of a zero-speed command, the

[97]This appendix is based on material originally prepared by Battelle Columbus Laboratories in support of the OTA study. The editors gratefully acknowledge the contribution of Battelle Columbus Laboratories but accept full responsibility for the version presented here and for any alteration of content that may have been introduced in condensing and editing the material for publication,

[98]In transit systems with rubber-tired vehicles, special rails are mounted beside the guideway and brushes or shoes on the vehicle contact these rails as the vehicle moves along. The special rails replace the running rails of a conventional steel-wheel, steel-rail system, and the brushes or shoes replace the wheels and axles of a train in the operation of the track circuits, This is only a difference in detail; the principle is the same as in conventional track circuits.

train must be able to stop before it comes to the end of the block.

D.C. Track Circuits

In all track circuits an electrical signal of some kind is impressed between the running rails, and the presence of a train is detected by the electrical connection that the wheels and axles of the train make between the two running rails. In d.c. track circuits, the electrical signal is direct current, usually supplied by batteries, The detector for the electrical signal is a relay.

Figure B-1 shows a simple d.c. track circuit. The track circuit consists of a block or length of track which is defined at each end by insulated joints in the running rails. The insulated joints provide electrical insulation between a given track circuit and the abutting tracks which comprise other track circuits. The signal source, in this case a battery, is connected to the rails at one end of the track circuit while the receiver (a relay) is connected to the other end, When no train is present, the track circuit is said to be unoccupied, and the direct current supplied by the battery is transmitted by the running rails to the relay and energizes it or "picks it up. " When the relay is energized, the upper set of relay contacts is connected causing the green signal light to be turned on. When a train enters the track circuit its wheels and axles connect the two running rails together, shorting the battery and thereby reducing the current through the relay. This causes the relay to "drop," as shown by the dashed line in the diagram. This action connects the bottom set of relay contacts, turning off the green signal light and turning on the red light to indicate that the block is occupied by a train. The resistor in series with the battery protects the battery by limiting the current the battery must provide when a train is present,

The terms "pick up" and "drop" refer to the position of the special "fail-safe" relays used for train detection. These relays are constructed from specifications approved by the Association of American Railroads and are designed so that their normally open "front" contacts will be closed only when sufficient electrical energy is being supplied to the coil, One or both of the normally open contact members are made of carbon or carbon impregnated with silver, which cannot be welded. The relays use gravity rather than spring return and are mounted vertically so that the relay armature, to which the contacts are attached, is returned to the dropped position when the current through the coil is reduced below some critical value. The failure rate of these relays for the mode in which the normally open contacts would be closed with no power applied to the relay coil is so low that for all practical purposes it is considered to be zero.

The track circuit shown here has been simplified for the purpose of illustration. In actual practice, the relay would have several sets of contacts connected in combination with the contacts from other relays in nearby track circuits to form logic circuits for the control of the signaling devices (the red and green lights). Even in the simple form shown in Figure B-1, however, it can be seen that the breaking of any conductor or the loss of power in the circuit will cause either a red signal or no signal at all to be displayed. A red or "dark" signal is always to be in-

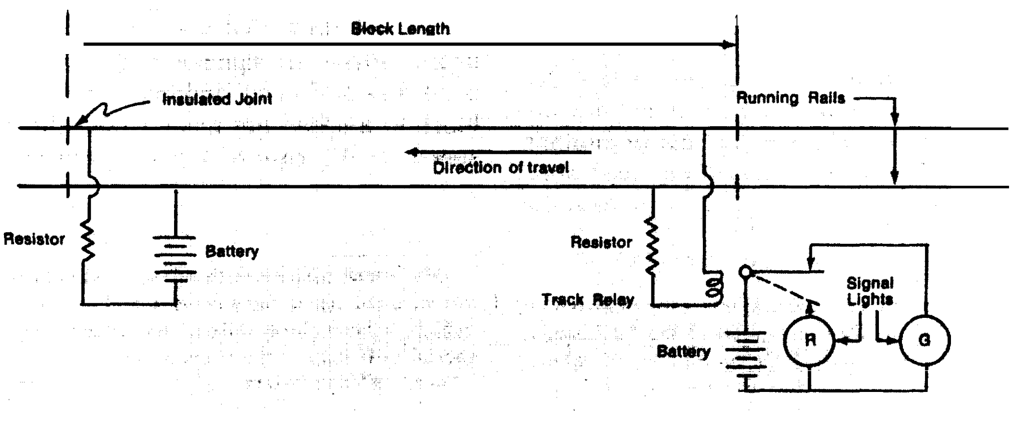

FIGURE B-1.—Simple D.C. Track Circuit

terpreted as a command to stop. If, for instance, the green light burned out or the relay coil open-circuited so that the relay could not be "picked up, " it would be impossible to have a "green" signal. In that case a train would be required to stop when it arrived at that signal. To put it another way, all signaling systems are designed so that a green signal (meaning proceed) is presented only when the track circuits provide positive information that it is safe to do so.

The double-rail d.c. track circuit is susceptible to interference when the running rails are also used as the return for d.c. electric propulsion current. For this reason, d.c. circuits are not used in rail rapid transit.[99] Single-rail d.c. track circuits could be used, but in fact all modern rail rapid transit systems use some form of a.c. track circuit.

Power-Frequency A.C. Track Circuits

The power-frequency a.c. track circuit is energized by an alternating electrical current with a frequency in the range of 50 to 150 hertz.[100] Except for the type of current and apparatus used, the a.c. track circuit is similar in operation to the d.c. track circuit described above.

[99]The principal modern application of the double-rail d.c. track circuit is in railroads with diesel-powered locomotives,

[100]This type of circuit is often called simply an a.c. track circuit.

Figure B-2 shows a simple power-frequency a.c. track circuit. As with the d.c. circuit, the a.c. track circuit consists of a block or length of track which is defined at each end by insulated joints in one or both of the running rails. Figure B-2 shows a double rail circuit with insulating joints in both rails. The a.c. signal source (usually a transformer) is connected to the rails at one end of the track circuit while the receiver (a relay) is connected to the other end, In addition to the signal source and the receiver, the a.c. track circuit contains a pair of impedance bonds at each pair of insulated joints. An impedance bond is a center-tapped inductance which is connected across the rails on both sides of the insulated joints. The center taps of the pair of impedance bonds are connected together as shown. The purpose of the impedance bonds is to provide continuity between the track circuits for the d.c. propulsion power and to distribute the propulsion current between the two running rails. The impedance bonds do this while still maintaining a relatively high impedance at the signaling frequencies between the two rails and between adjacent track circuits.

When no train is present, the alternating current supplied by the transformer at the left side of the diagram in Figure B-2 is transmitted by the running rails to the relay and "picks it up. " The energized relay turns on the green signal light, exactly as in a d.c. track circuit. The wheels and axles of a train en-

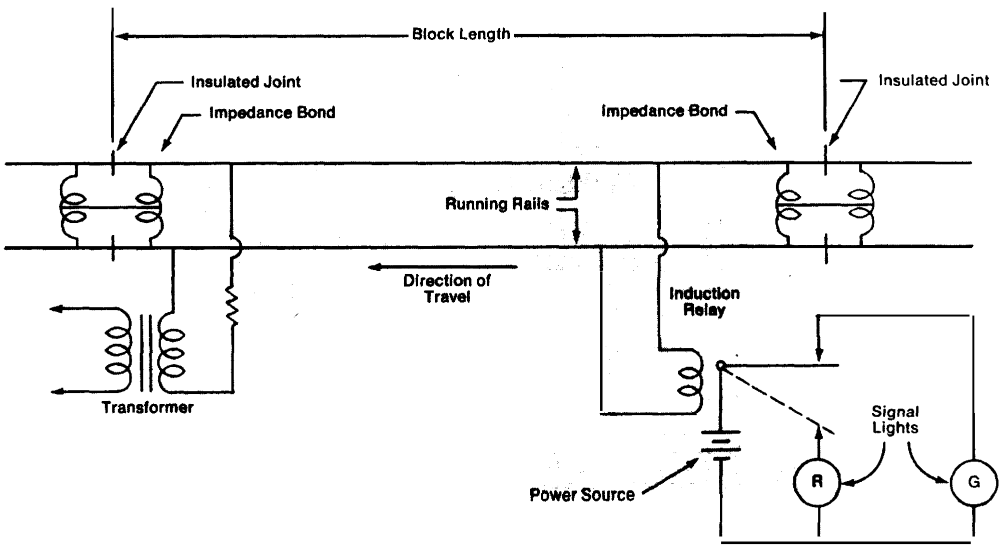

FIGURE B–2.—Simple Power-Frequency A.C. Track Circuit

tering the track circuit connect the two running rails together; and the current through the relay is reduced, causing the relay to "drop." This connects the bottom set or relay contacts, turning off the green light and turning on the red light to show that the block is occupied. The resistor in series with the transformer (at the left in the diagram) protects the transformer by limiting the current that the transformer must provide when a train is present.

High-Frequency A.C. Track Circuits

Some a.c. track circuits use a current that alternates at a frequency in the range of hundreds or thousands of hertz. Because this frequency range corresponds roughly to the spectrum of audible sound, such circuits are sometimes called audiofrequency track circuits.

High-frequency a.c. track circuits eliminate the need for insulated joints in the running rails. Because insulated joints are expensive to install and to maintain, eliminating them leads to a significant cost reduction. Eliminating insulated joints also allows the track circuit to operate with the continuous welded rails being used in some newer installations.

Figure B-3 shows a simple high-frequency a.c. track circuit. Since no insulated joints are used in the running rails, the ends of the block established by special transformers are connected to the rails. The transformer winding attached to the rails is usually a single turn of heavy copper bar stock. The transformer core is often a toroid. The other transformer winding is tuned to resonate at the operating frequency by a capacitor. The transmitter is the a.c. signal source and provides electrical energy at the operating frequency in the audiofrequency range. The receiver in this case is not simply a relay, as with the d.c. and power-frequency a.c. track circuits, but an electronic circuit which responds to the electrical signal provided by the transmitter. The receiver may be used to actuate a relay which performs functions like those in the d.c. or power-frequency a.c. track circuits. Thus, when no train is present the high-frequency a.c. potential supplied by the transmitter is connected to the running rails by the transformer and transmitted along the running rails to the other transformer and its associated receiver. When the receiver detects the high-frequency a.c. signal, the relay is energized and the green signal light is turned on. When a train enters the block, the circuit behaves much as it would with the a.c. or power-frequency a.c. track circuits, That is, the train wheels and axles connect the two running rails together, and the current to the receiving transformer and its associated receiver causes the track circuit relay to drop, turning off the green light and turning on the red light.

The circuit illustrated in Figure B-3 is highly simplified. In practice, it is necessary to accommo-

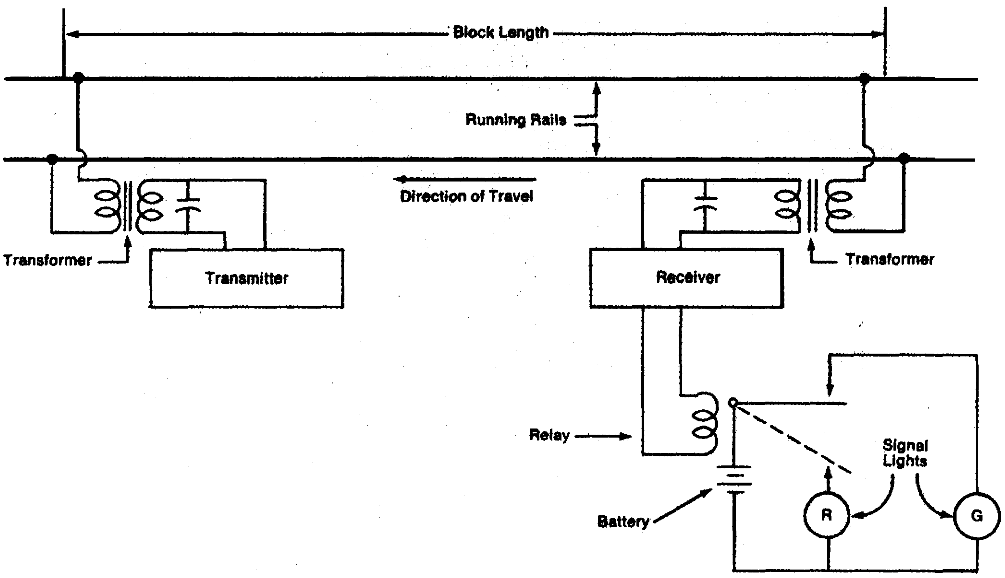

FIGURE B-3.—Simple High-Frequency A.C. Track Circuit

date the adjacent track circuits on either side, Rather than install two separate transformers for each track circuit, a second resonant winding can be included in each transformer or a heavy primary winding can be passed through more than one transformer core. Thus, a single transformer assembly is used at the boundary between adjacent track circuits and serves each. Although part of the same transformer assembly, the resonant windings are effectively isolated from each other because they are tuned to and operate on different frequencies.

Figure B-4 shows another type of high-frequency a.c. track circuit. No insulated joints are used in the running rails. The ends of the block or track circuit are established by shunts which are heavy copper cables or bars attached to the rails, The transmitter is the a.c. signal source and supplies electrical energy to a loop which is placed between the rails as shown. The loop is the primary of a transformer of which the rails and the shunts form the secondary, At the other end of the track circuit, a pickup directs the high-frequency a.c. energy to the receiver, which in turn actuates the track circuit relay. When no train is in the track circuit, the high-frequency a.c. potential supplied by the transmitter is directed into the loop and thence into the running rails to the pickup associated with the track shunt shown at the right portion of the diagram, When the receiver detects the high-frequency a.c. signal, the relay is energized and the green light is turned on. When a train enters the track circuit, the wheels and axles connect the two running rails together, and the current to the receiver is reduced. The reduced current to the receiver causes the track circuit relay to drop, turning off the green light and turning on the red light, as in other types of track circuits. The relay in a practical circuit would have several sets of contacts which would be connected in combination with the contacts from relays in nearby track circuits to form logic circuits for the control of the signaling devices.

In practice a transmitting loop and a pickup are associated with each track shunt. Adjacent track circuits are operated on different frequencies, and the receivers have frequency selectivity so they only respond to their intended frequencies. This is the type of track circuit used in the BART system.

Check-In/Check-Out Circuits

All of the track circuits described up to this point operate on the closed-circuit principle. Any disruption of the circuit by a train passing along the rails or by power or component failure, "opens" the circuit and causes a red (stop) indication to be displayed by the signal system.

An alternate approach to track circuit design uses "check-in/check-out" logic, Simply stated, this cir-

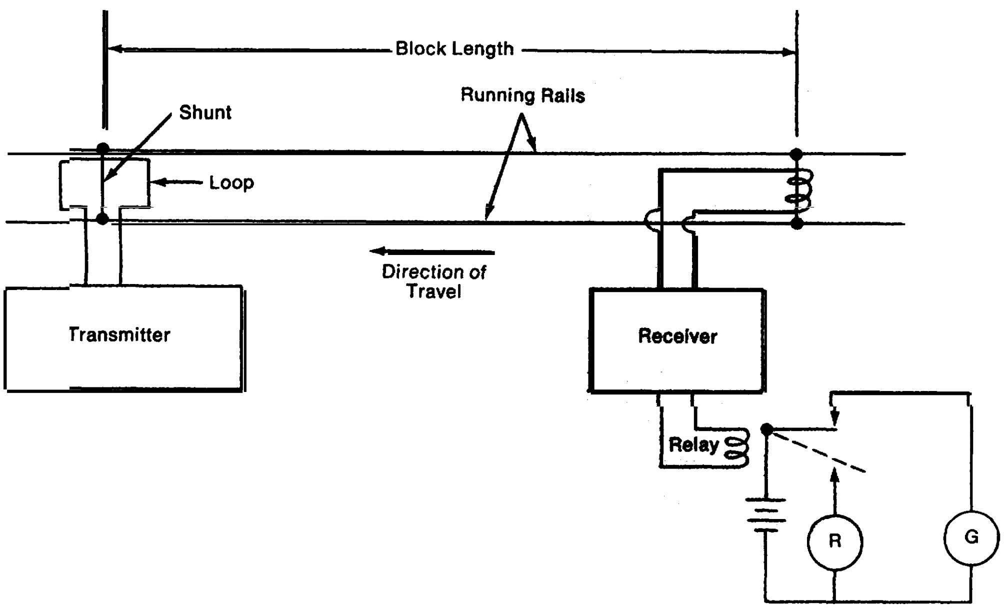

FIGURE B-4.—Alternate High-Frequency A.C. Track Circuit

'2-683 O - 76 - 14

cuit is based on the principle that once a train is detected or "checked in" to a block, it is assumed to be there until it is "checked out" by being detected in an adjacent block. The presence of a train may be detected only intermittently at the time when it enters a new block. This is in contrast to the conventional track circuits described above in which the presence of a train is detected continuously. In some check-in/check-out systems the first and the last cars of a train are checked in and out of the blocks as the train moves through the system, A transmitter of some kind located on the train can be used with a receiver at a fixed wayside location to indicate that a train has entered the block associated with the wayside receiver, In some cases, two transmitters are used, one at the head end of a train and the other at the rear. When the head-end transmitter enters a new block and is checked in, the block remains in the occupied condition until the rear-end transmitter also indicates that the rear end of the train has entered the new block, At that time, the train is checked out of the block behind.

Check-in/check-out has some operational disadvantages, For instance, consider the effects of a temporary loss of power to the signal system. With conventional track circuits, the loss of signal power will cause all track circuits to indicate occupancy, but when the signal power is restored, the true occupancy situation is again indicated. With a check-in/check-out system, the loss of signal power may destroy the "memory" circuits charged with "remembering" that a train has entered a block. The memory often consists of electrical relays which are energized (or deenergized) to indicate the presence of a train in a block. The loss of electrical power can destroy the information stored in such a memory. (A memory whose information can be lost by a loss of electrical power is termed "volatile.") Thus, when the signal power is restored, the information on track circuits which are occupied may have been lost. In this case the identity and location of each train in the affected portion of the system must be established before the entire transit system can be operated again safely. In a small transit system the identification and location of each train may not be difficult to establish. However, in a large, complex system even a short-term interruption of a portion of the system can create a bottleneck which makes it very difficult to restore the system to full operation. Thus, check-in/check-out systems do not find application as the primary train detection system in rail rapid transit systems.

A special case of a check-in/check-out system is the SOR (sequential occupancy release) system recently installed at BART, which uses the check-in/check-out principle as a logical back-up to the primary train detection system which uses high-frequency a.c. track circuits. The purpose of SOR is to protect against the loss of train detection in the event the primary system fails and to prevent service interruptions due to false occupancy indications,

The SOR system provides a latch such that an occupied track circuit continues to indicate occupancy until it is reset by the detected occupancy of the second downstream track circuit. Thus, with the loss of shunt or failure to detect the presence of a train, the latched-up track circuit still indicates occupancy and prevents a following train from colliding with the rear of the leading train. A series of computers is used in the SOR system, and the logic is such that the computers can recognize false occupancies, i.e., a track circuit which shows occupancy without a prior occupancy of the preceding track circuit is considered by the computer to be falsely occupied,

SPEED COMMAND AND CONTROL

In considering how the speed of transit vehicles is controlled by automatic devices, it is important to understand the principle of closed-loop control before proceeding to a discussion of the means by which speed commands are transmitted and received, A closed-loop control system is one in which some feedback of information on the status of the system (or its response to command inputs) is used to modify the control of the system. As a minimum, feedback verifies that the command was received, Feedback may also be used to modulate subsequent command inputs so as to smooth out irregularities of response, to make increasingly more precise adjustments of the state of the system, or to compensate for external perturbations. Thus, the basic purpose of closed-loop control is to assure continuity of control by confirming that command inputs have been received and that the commanded state of the system has, in fact, been achieved.

The alternative to closed-loop control is open-loop control, where commands are transmitted from the controlling to the controlled element without any feedback or acknowledgment that the command signal has been received and interpreted properly. The traditional wayside signaling of rail rapid transit is an open-loop system. So, too, is a manually operated train with cab signals, although the automatic overspeed and stop enforcing mechanisms of cab signals represent the beginning of a closed-loop system. Systems with ATO are true closed-loop systems. Feedback is used to monitor the response to propulsion and braking commands and to regulate the performance of the system on a continuous, real-time basis. Thus, a closed-loop system, in contrast to an open-loop system, is characterized by continuous control and self-adjusting commands conditioned by observation of system response.

The technology for controlling the speed of transit vehicles is based on the track circuit. The signals used for train detection can also be used for the transmission of speed commands to wayside signaling devices and to the trains. Two general methods are used for the transmission of such commands. In one method, the track circuit signal is turned on and off at a specific rate, which is interpreted as a speed command. This rate modulation scheme is called a coded track circuit. The second method is called binary message coding. With either method, equipment on the wayside or on the train senses the signals in the rails and decodes the speed command.

Coded Track Circuits

This technique is applicable to either d.c. or a.c. track circuits. The track circuit signal is switched on and off (modulated) at a rate which is related to the speed command. The switching rates are in the range from about 50 to 500 times per minute. In a d.c. track circuit, the direct current applied to the running rails at one end of the track circuit is simply turned on and off at the desired rate. Wayside equipment at the other end of the track circuit receives and decodes the signals. A code-following track relay is used in the track circuit and codes continuously when the circuit is not occupied. The relay is energized when the current is allowed to flow and is deenergized or "drops" when the current stops. The decoding equipment is actuated by the contacts of the code-following relay. When a given code (rate of transmission) is received, a particular relay in the decoding equipment is energized and remains energized as long as that code is being received. The relay, in turn, controls the appropriate wayside signal. When another code is received, another relay is energized as long as that code is being received. When a train enters the track circuit, the code-following relay is deenergized, and this fact is used to indicate the presence of a train. Typical interruption rates for these circuits are 75, 120, and 180 times per minute,

In a.c. track circuits, either power-frequency or audiofrequency, the a.c. signal is turned on and off at a selected rate. Since the switching rates for the coded signals are so much slower (I-3 per second) than the frequencies of the a.c. signals applied to the track circuit (50-150 per second), many cycles of the a.c. signal occur during the time that the code signal is switched on. The coded track signal can be received by wayside equipment at the far end of the track circuit and used to control wayside signals or it can be received on board a train and used to control the speed of the train. The presence of a train stops the operation of the code-following relay and indicates occupancy of the track circuit. The coded track signals are received on board a train by a pair of coils mounted near the front of the leading car just a few inches above each of the two running rails and in front of the first set of wheels and axle, The magnetic field from the electric current carried in the rails produces a signal in these coils (sometimes called antennas), and this signal is processed or decoded to determine the switching rate and hence the speed command. The decoded speed command is used in an automatic system to control the speed of the train. In a semiautomatic system, the decoded speed command is displayed to the train operator who then regulates train speed manually.

Binary Coded Track Circuits

This technique is sometimes used with audiofrequency a.c. track circuits. It could also be used with power-frequency a.c. circuits, but customarily it is not. Instead of turning the track circuit signal on and off (rate modulation), the frequency of the track signal is changed from one to the other of two discrete frequencies. This technique is particularly adaptable to digital systems in which one frequency corresponds to the transmission of a "l" and the other frequency corresponds to the transmission of a "0." The track circuit receiver responds to both of the signaling frequencies that are used. When a train enters the track circuit, the amplitude of the signals at the track circuit receiver is reduced below some threshold and this information is used as an indication of the presence of the train. On board the train two antennas or coils are mounted near the front of the lead car close to the running rails and in front of the first set of wheels and axle. As with the coded track circuits, the magnetic field from the electric current carried in the rails produces a signal in these coils, and this signal is decoded to determine the speed command. In an automatic system, the decoded speed command is used to control the train speed, In a semiautomatic system, the decoded speed code information is displayed to the train operator who controls train speed manually,