NASA SP-159

12

RECENT ADVANCES IN DISPLAY MEDIA

A symposium held in

Cambridge, Massachusetts

September 19–20, 1967

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA SP-159

RECENT ADVANCES IN DISPLAY MEDIA

A symposium held in Cambridge, Massachusetts, September 19–20, 1967, and sponsored by the Electronics Research Center, NASA

Scientific and Technical Information Division
OFFICE OF TECHNOLOGY UTILIZATION 1968
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C.

For sale by the Superintendent of Documents,
U.S. Government Printing Office, Washington, D.C. 20402
Price $1 (Paper Cover)
Library of Congress Catalog Card Number 67–62659

FOREWORD

The symposium on Advanced Display Media was the eighth in a series of technical meetings sponsored by the Electronics Research Center, NASA. It was the first NASA symposium held on display media. The attendance of more than 350 engineers and scientists from universities, industrial organizations, and Government agencies reflects the increasing interest in improving displays in aeronautical and space vehicles and the continuing university-industry-Government relationship in aerospace technology.

Many of the techniques in display media discussed at the symposium and documented herein have been made possible by the tremendous advances in electronic technology in recent years. The application of digital computers to drive new displays has opened new vistas for engineers and scientists to improve the information available for astronauts and pilots to make decisions.

A unique feature of this seminar was the combination of theoretical papers and demonstrations of experimental or breadboard display devices by 31 aerospace research organizations.

The Advanced Display Media Symposium was one of a continuing program of conferences and seminars planned by the Electronics Research Center, NASA, to expedite interchange of electronics research and technology.

JAMES C. ELMS
Director, Electronics Research Center, NASA

RICHARD J. HAYES
Assistant Director for Guidance and Control Research

PRECEDING PAGE BLANK NOT FILMED.

CONTENTS

INTRODUCTION

This volume documents the proceedings of the symposium, "Recent Advances in Display Media," sponsored by the Electronics Research Center, NASA, Cambridge, Mass. It was held at the Kresge Auditorium of the Massachusetts Institute of Technology on September 19–20, 1967.

Recent trends toward the display of computer-generated information are rendering obsolete the use of traditional mechanical movements for information display. While there is a wealth of information as to precisely *what* information requires display in a given system, no previous seminar has concerned itself primarily at the display media level with the problem of *how* this information may best be displayed. It was the purpose of this seminar to examine the present and projected state of the art of the various display media applicable to computerized systems and to assist the system designer in the selection of media most appropriate for his particular needs.

The first three technical papers were written to highlight areas where further developments are required by citing the deficiencies of existing media for advanced spacecraft, advanced commercial aircraft, and general aviation. To some extent, these papers reflect the Electronics Research Center's emphasis on cockpit display. The remaining papers provide state-of-the-art reviews of display media.

The papers include technology equally applicable to nonaerospace systems such as command and control centers. All of the developments reported here were not sponsored by NASA. Rather, the results of Department of Defense, university, and industry-sponsored research are also included. Every effort has been made to provide research-sponsorship acknowledgment in the reference lists at the end of the papers.

GEORGE KOVATCH
Chairman

EDWIN H. HILBORN
Program Chairman

PRECEDING PAGE BLANK NOT FILMED.

SESSION I

CHAIRMAN, *Lloyd Stevenson*

COMPUTER-MANAGED DISPLAY SYSTEM FOR ADVANCED COMMERCIAL TRANSPORTS

W. C. DERSCH AND ROBERT T. JOHNSON
The Boeing Co.

The computer-managed display system is part of an automatic flight management system study designed to evaluate applications for automation in commercial air transport operations. The display system employs cathode ray tubes (CRT's) and appropriate solid-state displays to provide the flight crew with automatically sequenced, time-varying information that is not only more accurate and complete than that provided by conventional flight deck instrumentation, but is more pertinent to each flight regime and requires less manual control and adjustment. The prime purpose of this paper is to stimulate the development of components needed for such display systems.

The function of this display system is to inform the flight crew of the overall operation of advanced commercial transports. Development of advanced display systems and display components are, of course, mutually dependent. In order to exploit the inherent flexibility of modern CRT's and other active display components, computer management is necessary to avoid an otherwise hopeless clutter of controls.

With respect to the display system, a central flight management computer (ref. 1) is assumed to—

(1) Provide adequate storage to refresh and update the displays

(2) Manage the display system controls to reduce clutter and to supply the pilot with maximum useful information

(3) Process data into pictographs and improved informative display configurations

(4) Manage all the inputs to the display system, such as navigation and flight control computations, data link instructions, and air data Updating the navigation system via a moving map display is also included as a direct input function.

Figure 1 is a block diagram of the Automatic Flight Management (AFM) system (ref. 1) showing the relationship of the advanced display system to the overall airplane management system. Loops 5, 6, and 7 typify today's manually controlled airplane. Loops 1 to 4 show, respectively, the man-to-computer control, the computer-to-airplane control, the airplane-to-computer feedback, and the display system man-machine interface.

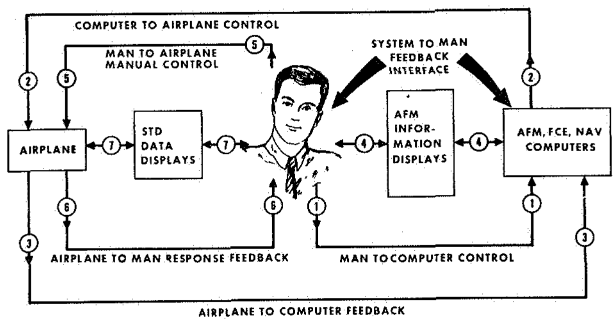

FIGURE 1.—Automatic flight management system.

The standard displays (loop 7) show unprocessed data, such as the number of pounds of fuel remaining and the airplane's altitude. The information displays in loop 4 show processed data, such as a pictograph of the relationship of fuel on board to the flight plan and the airplane's position relative to its destination, and its altitude in relation to the Mach number, air

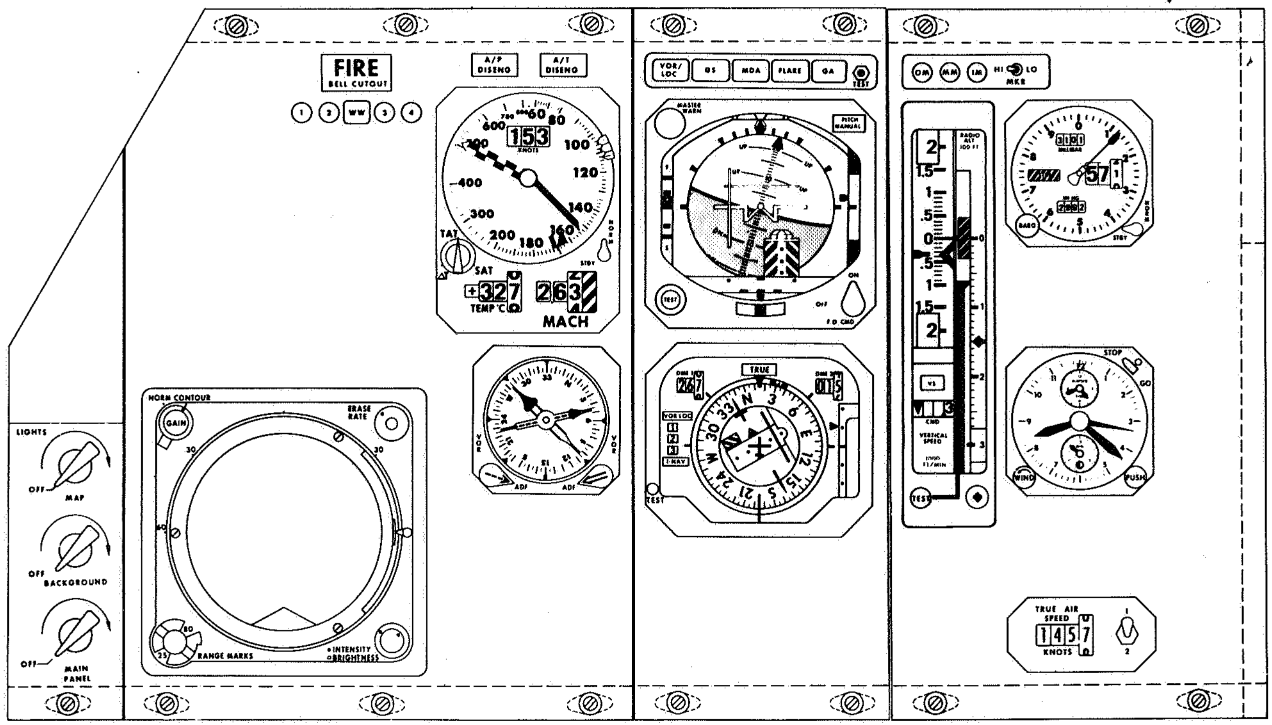

FIGURE 2.—Conventional pilot's panel.

temperature, and overpressure limits. The remainder of this discussion will be confined to the computer-managed display system in loop 4; the evolution of this system is shown in figures 2 to 4.

Figure 2 is a conventional pilot's panel showing some of the major instruments. These displays are in loop 7 of figure 1. Figure 3, an advanced pilot's panel, shows the substitution of cathode-ray-tube displays in place of the electromechanical displays. In addition, a Mach-altitude vertical profile display is included. Note the improved readability and pitch scale sensitivity on the attitude director display and the additional navigational information presented on the moving-map display. The advanced pilot's panel (ref. 2) is a bridge between the present standard display system and the computer-managed display system in loops 7 and 4, respectively. Figure 4 is a sample of the panel configuration of the display system under study; as is apparent, this panel borrows liberally from the panel reported in reference 2. The major difference is that the AFM display system represents a more extensive utilization of the cathode-ray-tube displays made possible by the management capability of the previously mentioned AFM computer facility.

The nature of the computer-managed display panel is illustrated in figure 5, which shows four different configurations on the same panel corresponding to four different points on the flight profile. When shown side by side, these four sample panels show how the CRT displays are configured in an effort to increase the display format efficiency as the airplane progresses to its destination along the flight profile.

Thus, with the display management concept, the pilot is furnished only usable information. With the pilot's prior permission, the display panel is automatically configured to the most efficient layout; however, the pilot is provided with the controls to take exceptions to the automatic configuration program. For example, in the case of suspected nonnormal operation of an engine, the pilot and flight engineer can examine engine readings in detail on their respective multimode displays. This permits the pilot to remain in full command of the airplane both through his other chosen essential displays and his copilot, the latter's displays being in the normal flight mode configuration.

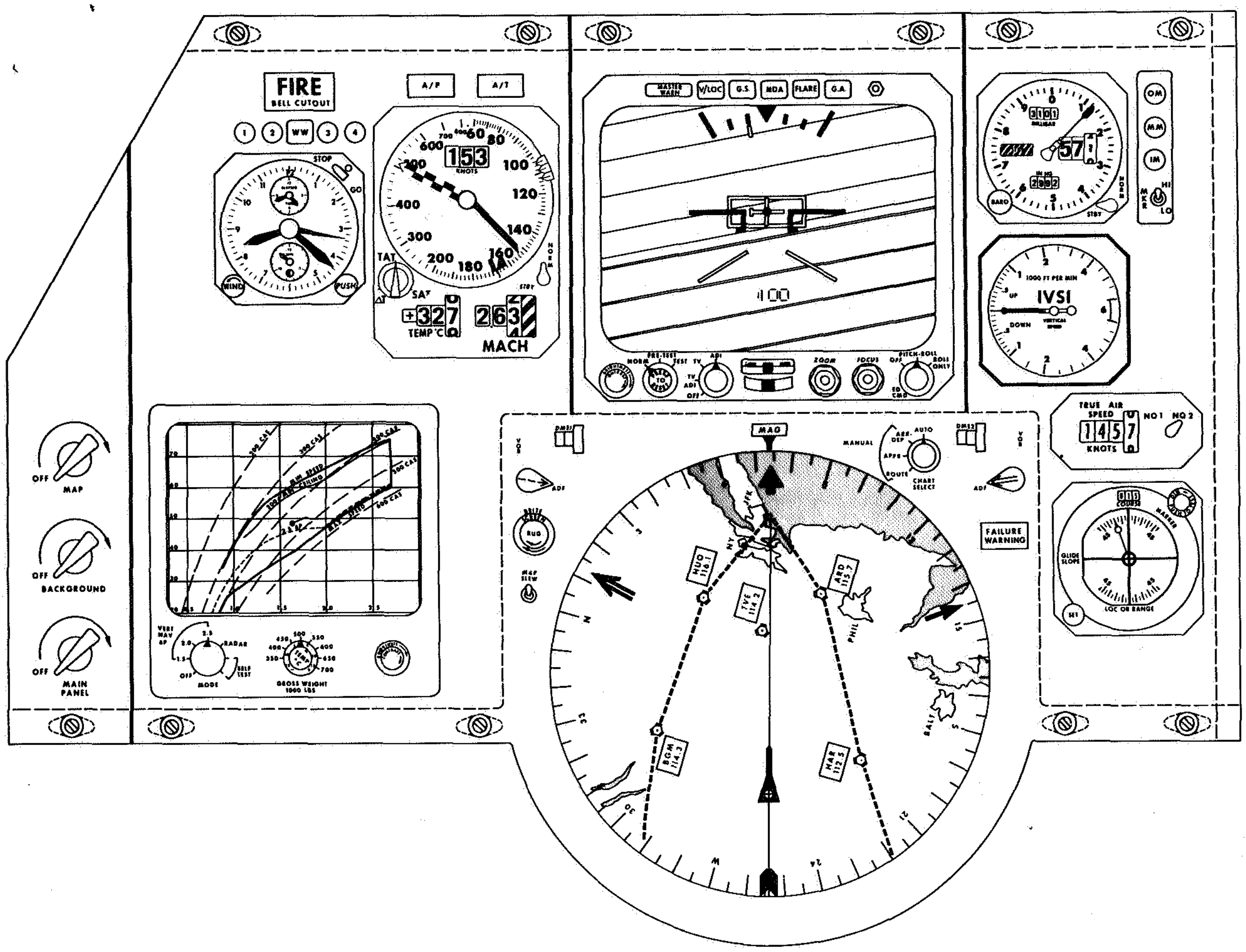

Figure 3.—Advanced pilot's panel based on conventional system organization.

The goal of the display management concept is to control by computer those routine functions that can be preplanned and that do not require judgment. This reduces the workload on the flight crew and conserves their finite human capacity for decisions requiring skillful and experienced judgment. In addition, the computer can sometimes sense a subtle trend before even the most attentive human observer, and thus can alert the flight crew in advance of situations which may require decisions.

Inherent in the managed display concept is the built-in program causing the display system to return to the automatically programed configuration as the flight progresses. Also, a simple arrangement is provided so that the pilot can reset the display system to the automatic mode. This prevents him from becoming lost in his display system, particularly after having changed his displays to a highly nonroutine configuration, as in the engine example just given.

Figure 6 shows a flight profile organized around the functional requirements of the managed display system rather than the usual navigationally oriented plan. For example, the display system requirements are the same when the airplane is taxiing at departure as when it is taxiing at arrival. Eight flight regimes, typical of supersonic flight and requiring different data display, are as follows:

1. Pre-post-flight checklist
2. Engine autostart
3. Taxi in-out terminal
4. Roll, liftoff, and subsonic noise abatement
5. Climb-cruise, subsonic
6. Climb-cruise, supersonic, and overpressure limits (sonic boom)

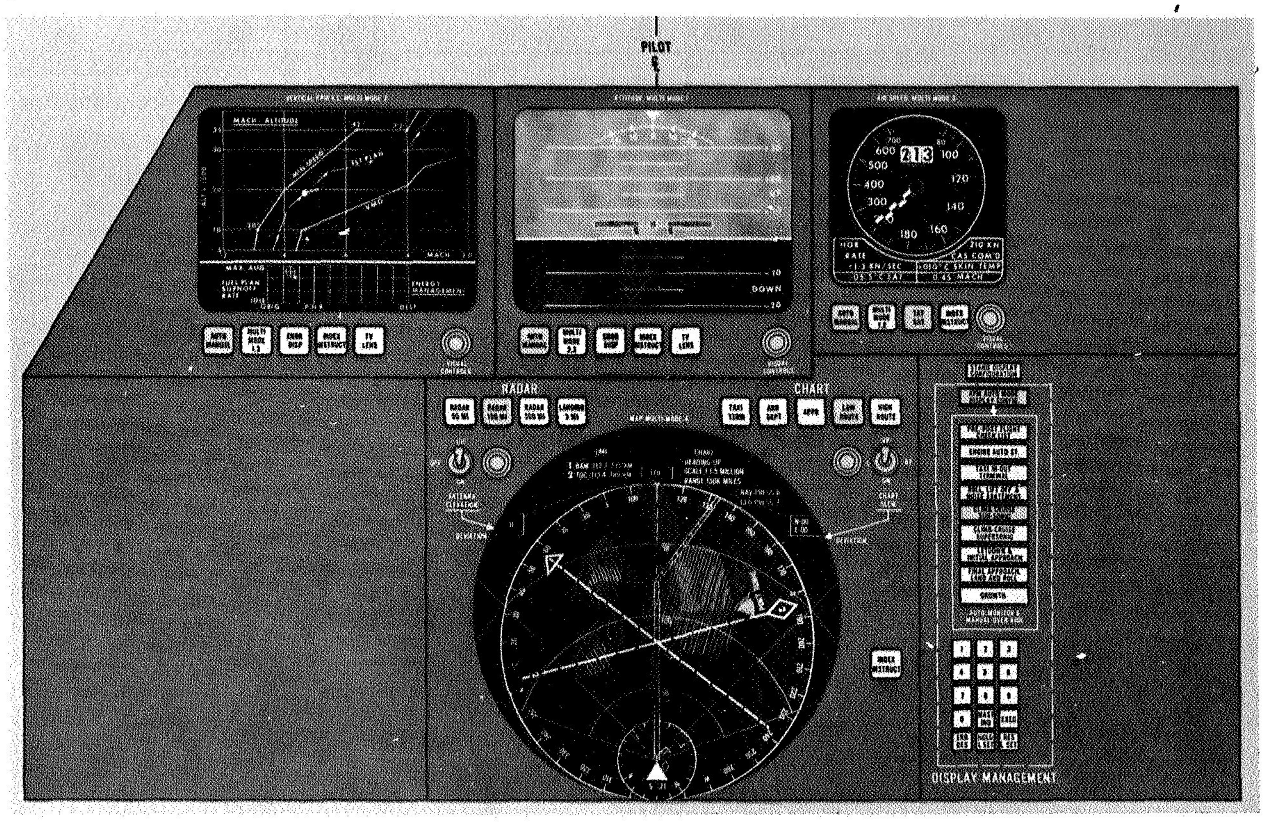

FIGURE 4.—Automatic flight management panel, climb-cruise, subsonic.

7. Letdown and initial approach
8. Final approach, land, and rollout

Panel layouts are being made of each mode (four of which were shown in fig. 5) to study what and how information should be displayed and managed. Development hardware and associated simulator designs are in process as the next step to eventual flight tests.

Let us take two points in time on the flight plan and examine the pilot's display panel. By the time the pilot gets to "Engine Autostart," he has approved the appropriate checklist items with the assistance of an auto checklist. The central management computer has been instructed to proceed in its program with the engine-start operation, and the display panel is automatically configured to this flight-plan mode. (See fig. 7.) The panel shows that engine 1 has already been started and is idling. The remaining three engines are being started in parallel on the assumption that the one running engine can provide adequate start services. This approach saves a few minutes which become significant when compared to a desired 30-minute turnaround time.

Since the airplane is standing still on a level surface, the attitude director display is not needed. Instead, its CRT is used to display the picture from a split-field TV camera observing the engine exhausts. Certain types of malfunctions show up immediately. Along with the engine-exhaust view, the engine status is quantitatively summarized for the pilot with the four sets of gages showing percent rpm, exhaust gas temperature (egt), and fuel flow. These gages are monitors, and the AFM computer is monitoring and controlling these and other transient data (such as oil-pressure buildup) against stored templates. In the event of an out-of-tolerance reading, preplanned action would be taken immediately and the pilot notified accordingly.

The pilot must be careful that the exhaust blast from his engines does not cause damage from jet heat or velocity. Thus he is shown a TV view of the underside and to the rear of

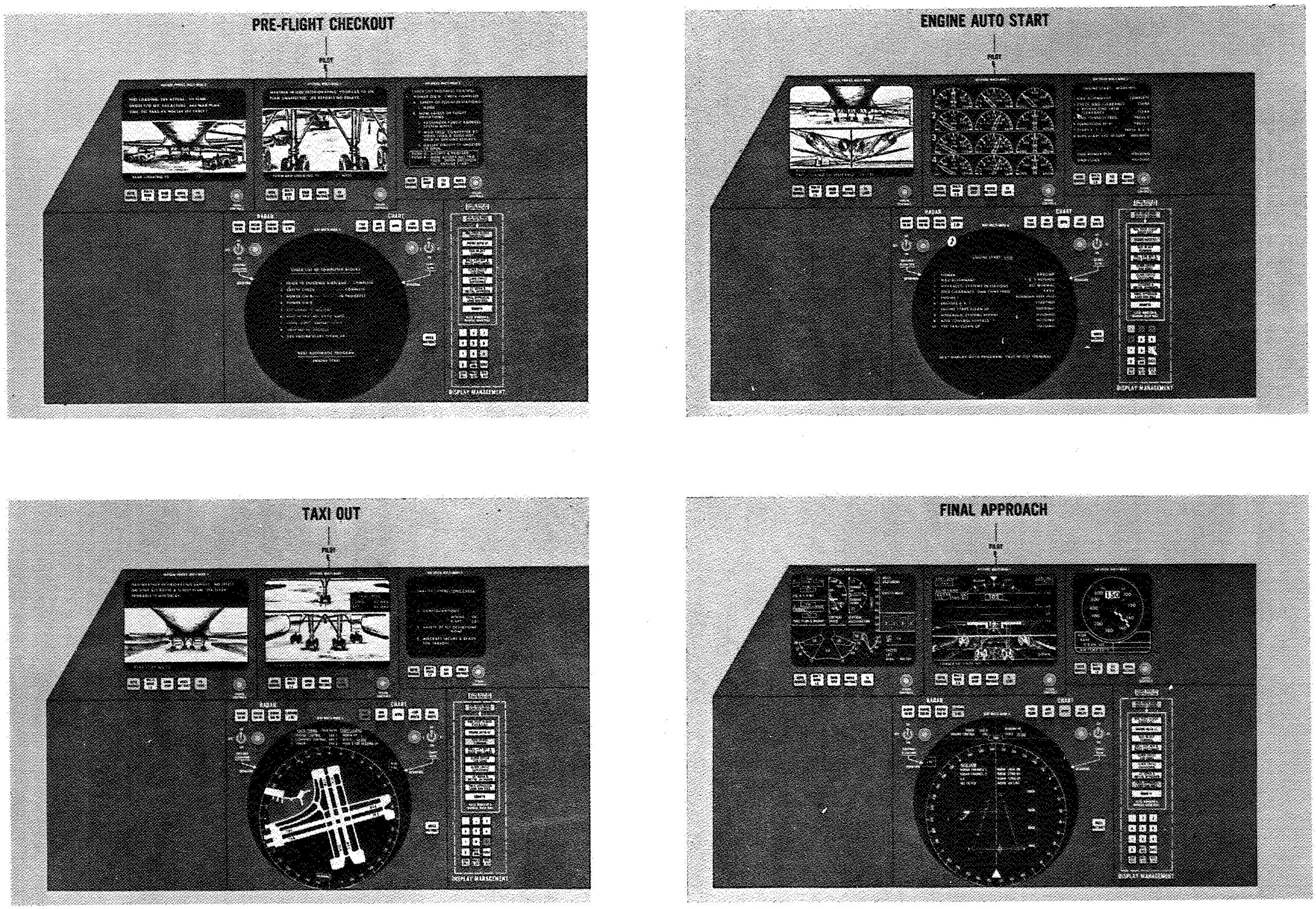

FIGURE 5.—Automatic flight management panel configuration comparison.

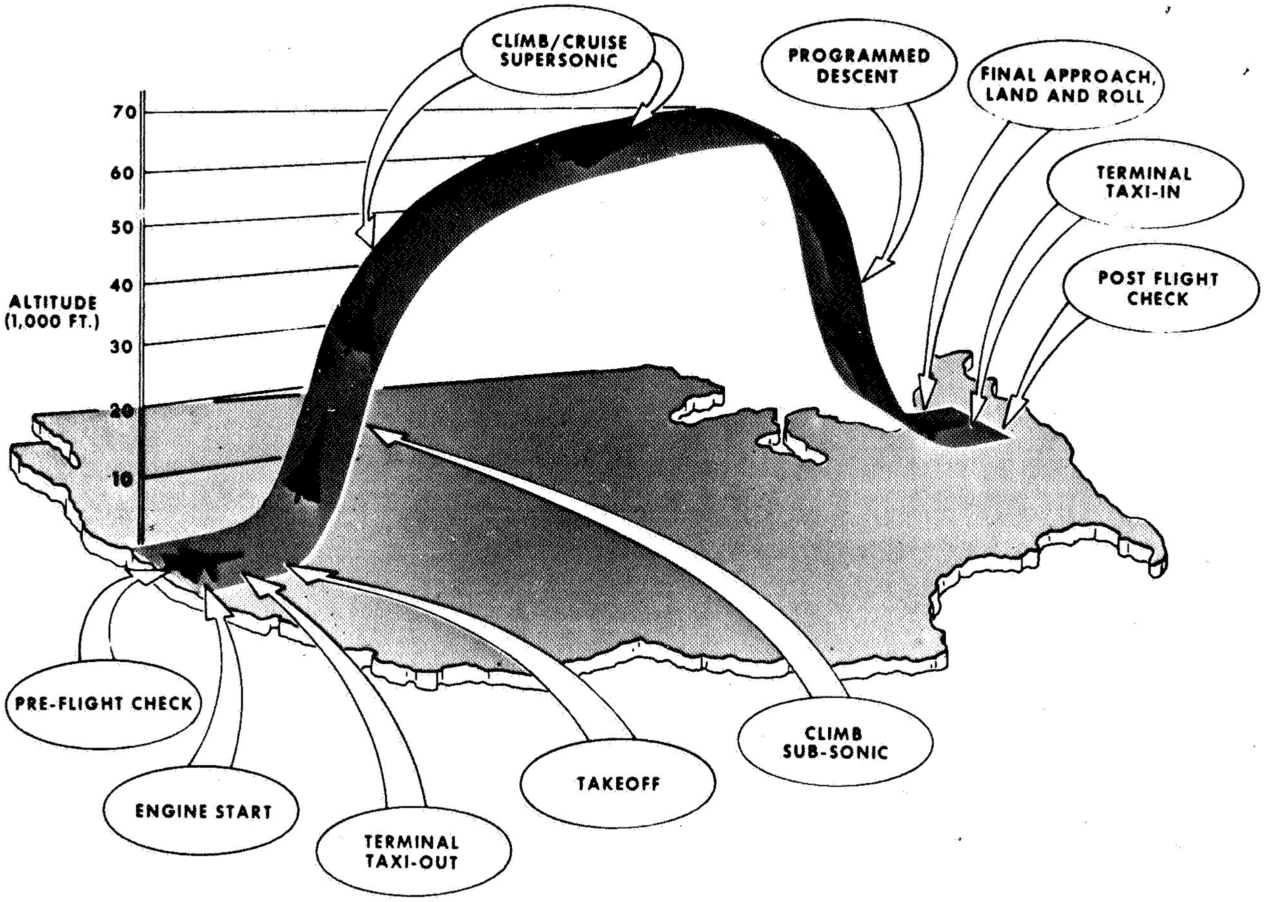

FIGURE 6.—Automatic display management flight profile.

the airplane, which enables him to verify the all-clear message from the ground crew.

As the auto checklist is proceeding, only the deviations or judgment-requiring decisions are displayed on the fourth multimode display. A log of completed items is shown on the third display, which is normally used for map and navigation information. High-priority messages, such as changing weather conditions would usually be displayed on the attitude director CRT or vertical profile multimode CRT.

Figure 8 shows what the pilot's panel might look like on final approach. The center of attention is the attitude director display, which shows a TV picture of the landing gear and its surroundings. The camera is located on the leading edge of the rear underside of the ventral fin, many feet behind the pilot. Tests are being run on the supersonic transport (SST) configuration on the Boeing space-flight simulator, in which both pilot and nonpilot engineers "land" the SST, using only the attitude director with the simulated "real world" TV background.

The rearward location of the camera gives a pictorial indication of the pitch attitude and yaw angle, in addition to quantitative pitch-and-yaw indicators. Vertical speed and vertical acceleration are shown, but in a less prominent location than the radio altimeter. A high-priority message area is also provided.

The vertical profile display is now used to show engine thrust and its relation to noise-abatement limits. The gross weight and the center of gravity, with respect to the aircraft flight plan, are computed and shown; fuel weight remaining and its relationship to the flight plan is computed and shown; and wing flap and wing sweep are shown.

On the moving map, multimode, CRT display, the airplane heading is shown along with a high-resolution radar picture of the runway.

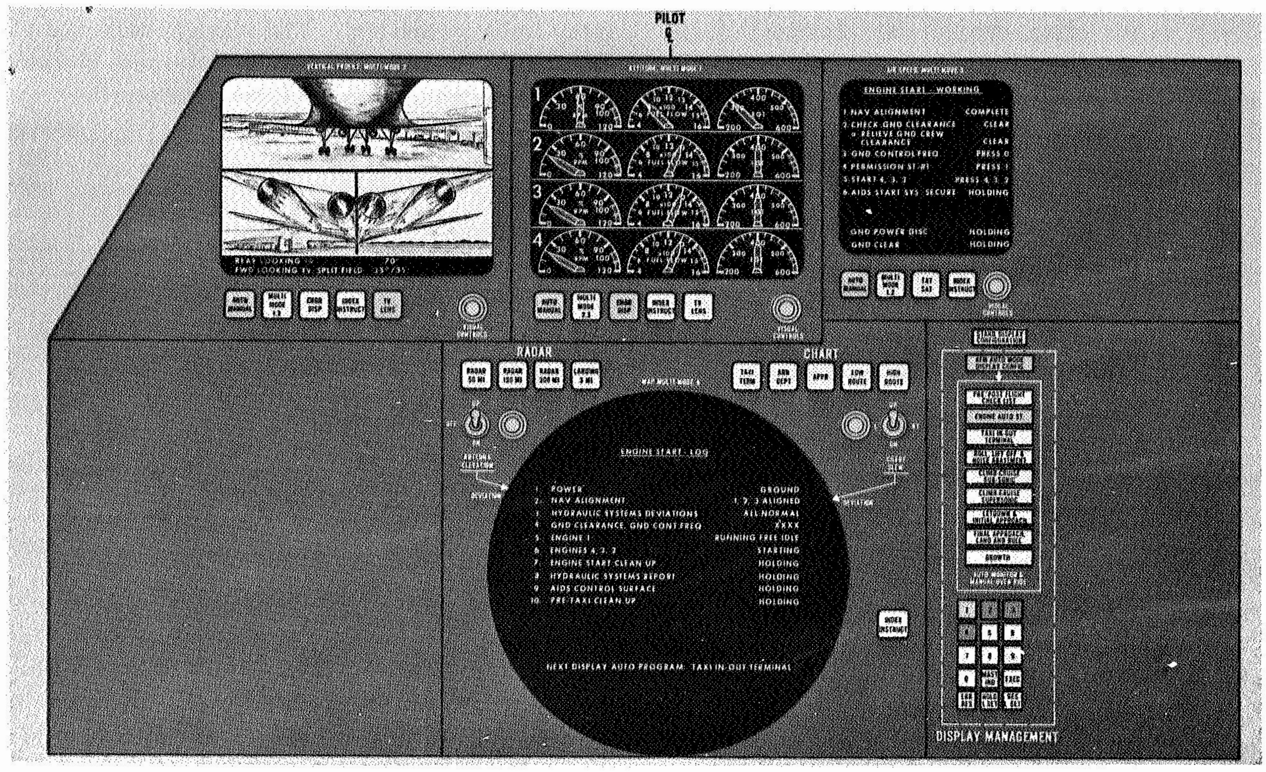

FIGURE 7.—Pilot's AFM display panel, engine autostart.

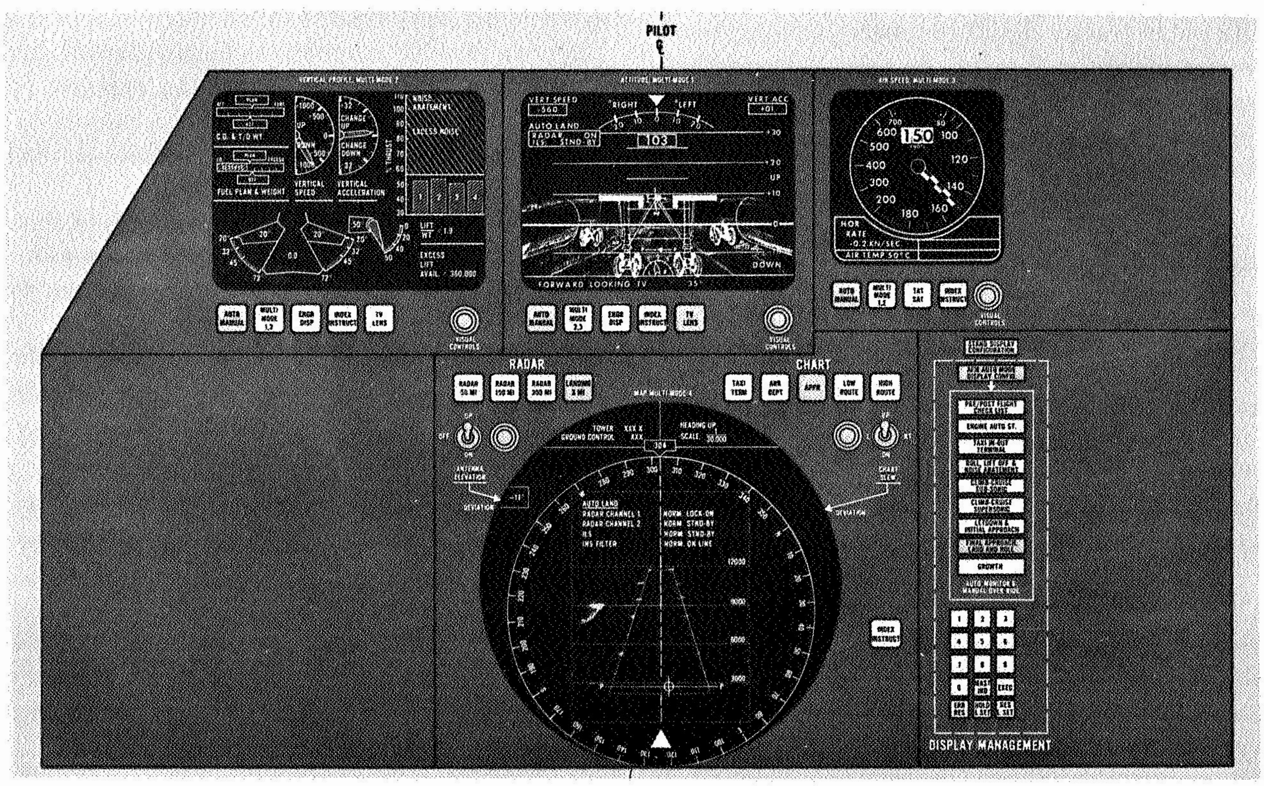

FIGURE 8.—Pilot's AFM display panel, final approach.

These radar data would be obtained from ground-located points which are fed into the computer to calculate the relative position and the change in position of the airplane with respect to the runway. In an autoland system, these data, along with data from other sources such as the standard Instrument Landing System (ILS) or inertial platforms, would be used to generate appropriate signals to maneuver the airplane safely onto the runway.

Under conditions of impaired visibility, the runway would not be visible either on the TV or through the window. In this circumstance, the pilot is shown, in natural pictorial form, the data and resulting computer control signals used to autoland the airplane. Thus he is informed and exerts overall management control of the airplane's safe flightpath.

The computer storage capability needed to service this display system is quite modest. Four display data sources are anticipated. The first and second—standard TV and weather or high-resolution radar—operate in real time and require no storage other than a few addresses to locate the several simultaneous pictures that can occur (radar signal processing is considered separately from display storage). Computer-generated curves and facsimile symbol-pictorial generators are the other two data sources which do require storage facilities.

Twenty computer-generated curves, with 40 segments per curve and vector line drawing capability, located within one-sixteenth inch on any one of four CRT displays, require about 1600 words. This type of display signal generator will be used for flightpaths, trending, and special symbols. Only full white-black contrast (no gray scale) is planned.

The facsimile generators are monoscopes and flying spot scanners. Thirty-three messages averaging 75 alphanumeric characters each can be located within one-sixteenth inch on the four CRT's. Some of these messages are canned groups of characters and thus require only one source and one display location. In addition, 36 symbols are provided, such as the airplane symbol, horizon lines, and pointers, with full gray scale as needed. The storage capacity of the display system, as now understood, is summarized as follows:

	Words
1. Pilot's panel	2000
2. Copilot's panel (additional)	1000
Total	3000

The following have not yet been detailed, but for the present have been sized in relationship to the complexity of the preceding panels:

	Words (estimated)
3. Flight engineer panel	2000
4. Center panel and miscellaneous	1000
Total	3000

The total of 6000 words assumes that the display capabilities of the various panels are highly independent. The pilot's and copilot's panels will have nearly duplicate displayed data at any one time, and there will be some commonality between the engineer's and center panels.

In summary, a computer-managed display system has been described, through which the pilot can give his attention to the airplane flying a safe path, rather than his being engaged in doing it. The flight crew will be as important, but "they will be managers rather than control pushers" (ref. 3).

This display system requires the highest flexibility in display components, and a computer is necessary to manage the display system to avoid an otherwise prohibitive complex of knobs. This system probably could be implemented with available components, but to realize maximum efficiency, improved components are required. Medium-sized screen, compact CRT's with multiple independent guns, high contrast and/or brightness, and multiple colors exemplify the functional requirements of desirable components. Satisfactory solid-state displays with their usually more limited flexibility, but higher space and weight efficiency factors, would most certainly be blended into the display system.

References

1. Petrie, D.: Automatic Flight Management. 4th annual AIAA meeting, Anaheim, Calif., October 1967.
2. Gannett, J. R.: New Flight Instrumentation for Future Transport Aircraft. SAE meeting, New York, N.Y., Apr. 26, 1967.
3. Schaiber, G. S.: How the Airplane Designer Can Help the People in the Cockpit. SAE meeting, New York, N.Y., Apr. 26, 1967.

f

N68-22304

DISPLAY REQUIREMENTS FOR ADVANCED MANNED SPACECRAFT

ROBERT C. DUNCAN

Assistant Director, Electronics Research Center, National Aeronautics and Space Administration

INTRODUCTION

A fundamental purpose of displays and controls in manned vehicles is to integrate man into the vehicle system. Displays provide information from which the pilot can control the vehicle. Displays take the form of—

(1) Dials to give status of subsystems.

(2) Attitude indicators; that is, state vector and rate of change of state-vector indicators. (Note that both items (1) and (2) are common both to aircraft and to the first three generations of manned spacecraft.)

(3) Computer-derived status and maneuver data, which were first used in second- and third-generation manned spacecraft. The third-generation manned spacecraft, the Apollo Command and Lunar Modules, use the computer to give data in numerical form to the pilot concerning (*a*) future state-vector data (that is, trajectory data projected to some future time or future orbit), and (*b*) emergency maneuver data (if required during the launch phase or other phases of the missions).

A significant man-machine integration factor for both the Gemini and Apollo spacecraft was the use of a powerful computer with provisions for communication between pilot and computer through proper input and output equipment. Another significant factor was the development of software during the mission-planning phases to give an important capability to the crew that has not been exploited in aircraft and space vehicles. This capability may also be of value to pilots of aircraft, particularly of high-performance aircraft such as the supersonic transport.

Heads-up displays are not used in the first three generations of manned spacecraft except for optical reticles provided in Gemini to aid in rendezvous and a reticle system in the Lunar Module, the landing-point designator, to assist in landing efficiently at a predetermined point on the surface of the Moon. The landing-point designator is a pattern on the left window of the Lunar Module that uses the parallax of the inner and outer windows to show the pilot where the computer is taking him. The computer provides the pilot with the window coordinates of the landing site by displaying this information on the computer keyboard. The pilot can then visually locate the site on the window and redesignate to a new target if desired.

MAN'S CAPABILITIES

What are man's unique capabilities as a controller of systems? Man has a unique facility for exercising judgment. He can reason inductively and has the ability to draw inferences from isolated elements in one situation and apply them to another. The trained human being has the capacity to analyze problems never before encountered and to make decisions on the basis of general, rather than specific, experience. He is a valuable technical troubleshooter, and he can assure reliable operation of all equipment on board his spacecraft by continuously checking and choosing alternate control modes.

Man can perform an almost unlimited diversity of tasks, such as the following:

(1) Correlate many nonrelated observations.

(2) Readily perform diverse physical tasks in any order—moving from place to place, replacing or repairing a faulty electronic component, controlling a moving spacecraft, collecting minerals and plantlife, and comparing observations with data previously accumulated.

(3) Make rapid decisions based on seemingly independent events.

(4) Have intuition, imagination, and initiative.

It is interesting to compare some of the debits and credits which enter the engineering ledger as a result of designing a spacecraft system to include the human crew. The man excels over the machine in the following:

(1) Detecting minimal changes in visual or auditory stimuli.

(2) Perceiving in the presence of noise meaningful patterns of information.

(3) Choosing new courses of action with great flexibility and adaptability when circumstances change unexpectedly.

(4) Storing tremendous quantities of data for long periods of time and recalling the required relevant information rapidly.

Engineering liabilities incurred by incorporating man into the system include the following:

(1) Much engineering effort is required to make the spacecraft habitable and safe because of the inherent vulnerability of man to the space environment.

(2) Engineering provisions beyond a utilitarian minimum must be made for living and rest because more is required of man than mere survival. He must perform at top efficiency, a requirement closely coupled with morale and physiological considerations. Equipment required (such as displays, controls, and data-processing facilities to make man effective in his duties) is costly in weight, size, complexity, and dollars.

Evolution of Control Displays in Manned Spacecraft

Let us summarize now in a series of drawings the evolution of pilot displays and controls in the first three generations of U.S. manned spacecraft.

Figure 1 shows the main display panel for the Mercury spacecraft. This spacecraft had no gyro horizon; it used aircraft-type instruments with some miniaturization and simplification. A left panel (not shown) provided status lights and controls for such things as jettisoning the abort tower. The top instruments in the main (center) panel provided attitude data in the form of needles indicating roll, pitch, and yaw. These were used in lieu of a gyro horizon (commonly called an eight-ball or FDAI, flight director attitude indicator). Also shown on the center panel is the altimeter and the rate of descent indicator. The right panel contains status lights and controls for the environmental system and other subsystems.

If the spacecraft shown by the U.S.S.R. at such international exhibits as Expo 67 are accurate, the Mercury display and control panel is complex in comparison with that of the Vostok. The Vostok had a small number of switches in one panel to the left of the pilot. In front of the pilot was a small number of dials and an interesting navigation indicator showing the geographic position of the satellite over the Earth. In lieu of a gyro horizon, the astronaut had a small window, approximately 8 or 10 inches in diameter, between his feet through which he could observe the Earth.

Medical data released by the Soviet Union and by the United States at various international meetings on the problems of manned space flight have been somewhat contradictory. The Soviet astronauts have apparently had greater physiological difficulty in accommodating to space flight than have the U.S. astronauts. One of the primary differences between Soviet manned spacecraft and U.S. manned spacecraft is the greater data on spacecraft status provided to the U.S. astronaut. Another significant difference is that the U.S. astronaut has been provided with considerably larger windows, giving him a much larger solid-angle view of the Earth. Perhaps these factors have contributed to the easy accommodation of U.S. astronauts to space flight. None has experienced illness or disorientation. No flight has

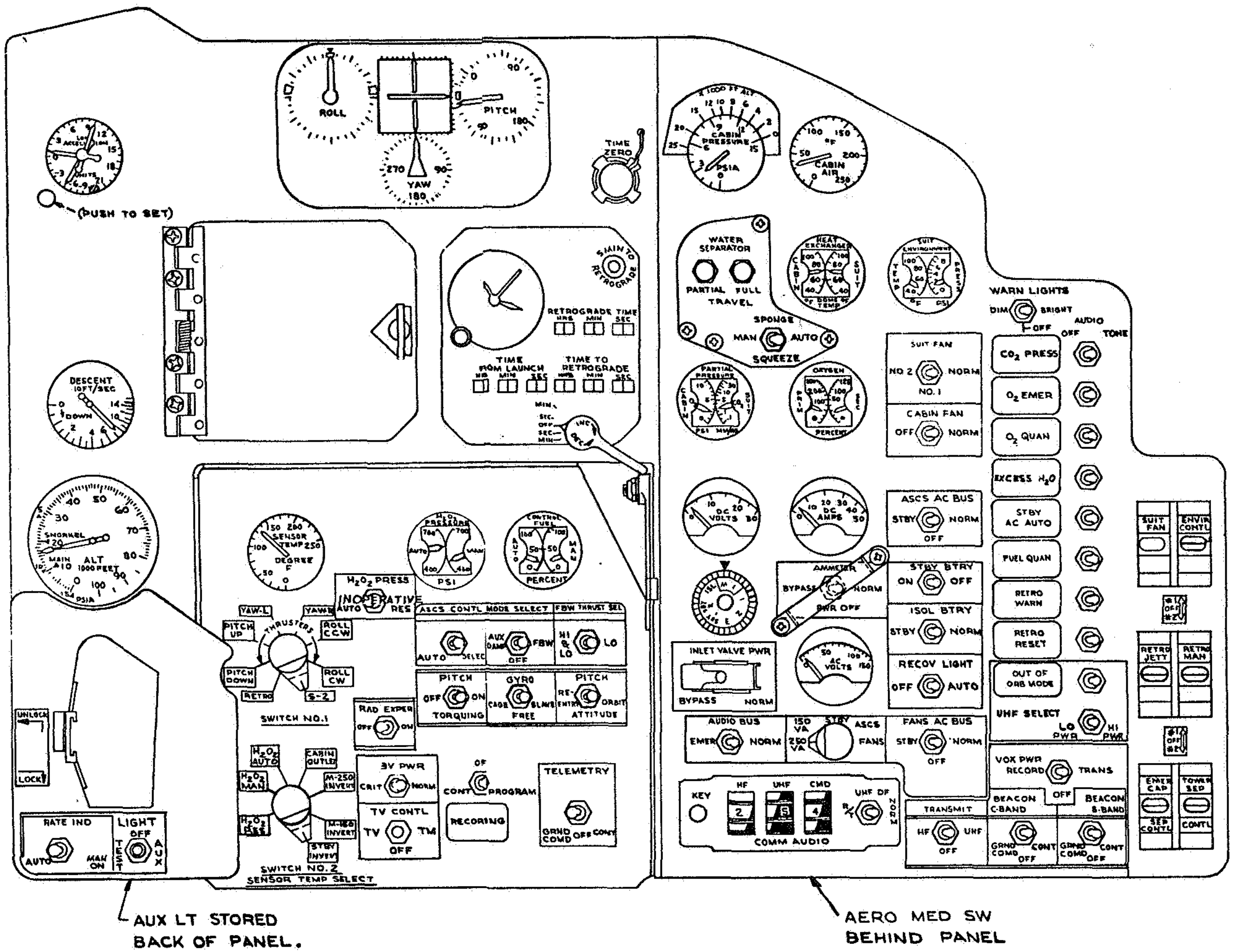

FIGURE 1.—Mercury spacecraft main display panel.

been prematurely terminated because of the astronaut's physiological reactions. This is not true of the Soviet astronauts. These factors are certainly not inconsequential in designing displays and evaluating the importance of them.

Figure 2 shows the main display and control panel for the Gemini spacecraft. This panel is much more complex and complete than that of Mercury. The Gemini spacecraft had an inertial guidance system and could maneuver in orbit for rendezvous and docking and during reentry for landing at a specified point on Earth. This maneuverability required that considerably more information be provided the pilots. For example, in the lower left-hand corner of figure 2 the incremental velocity indicators show Delta V maneuvers in all three axes. Also required is information on the status of the fuel, as shown by the propellant quantity indicators. To control the vehicle during maneuvers, the pilot was required to have an FDAI, and these were provided on both panels as shown. Many of the indicators were duplicated to increase the overall reliability of the spacecraft system. Gemini was the first manned spacecraft to have a high-speed digital computer, and methods were required for the astronaut and the computer to communicate with each other. This is shown in the lower right-hand side of the control panel.

Figure 3 shows the Gemini panel in more detail. In the lower right-hand panel are the mechanisms through which the astronaut can enter numbers into the computer and through which the address and message are indicated to

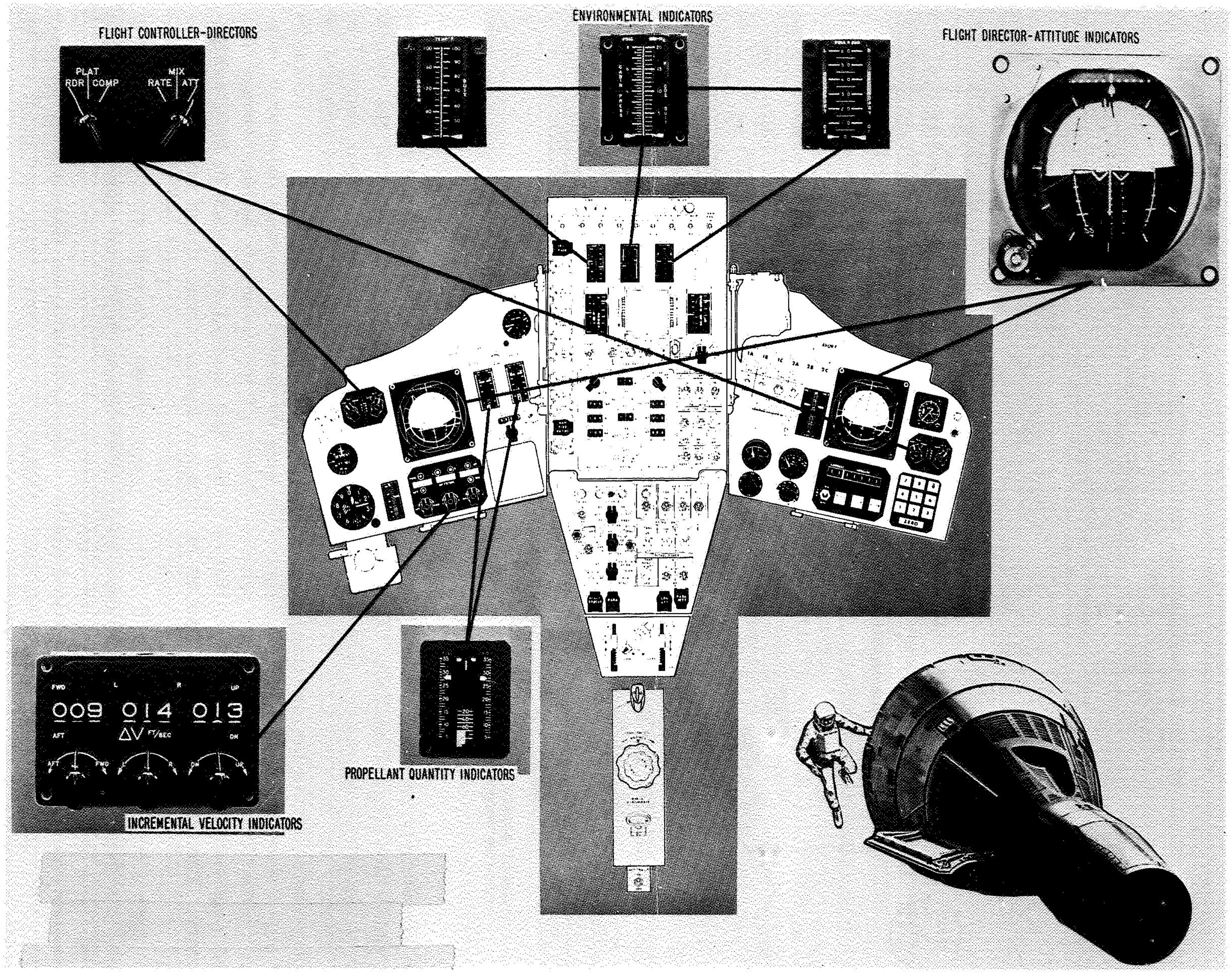

FIGURE 2.—Gemini spacecraft main display and control panel.

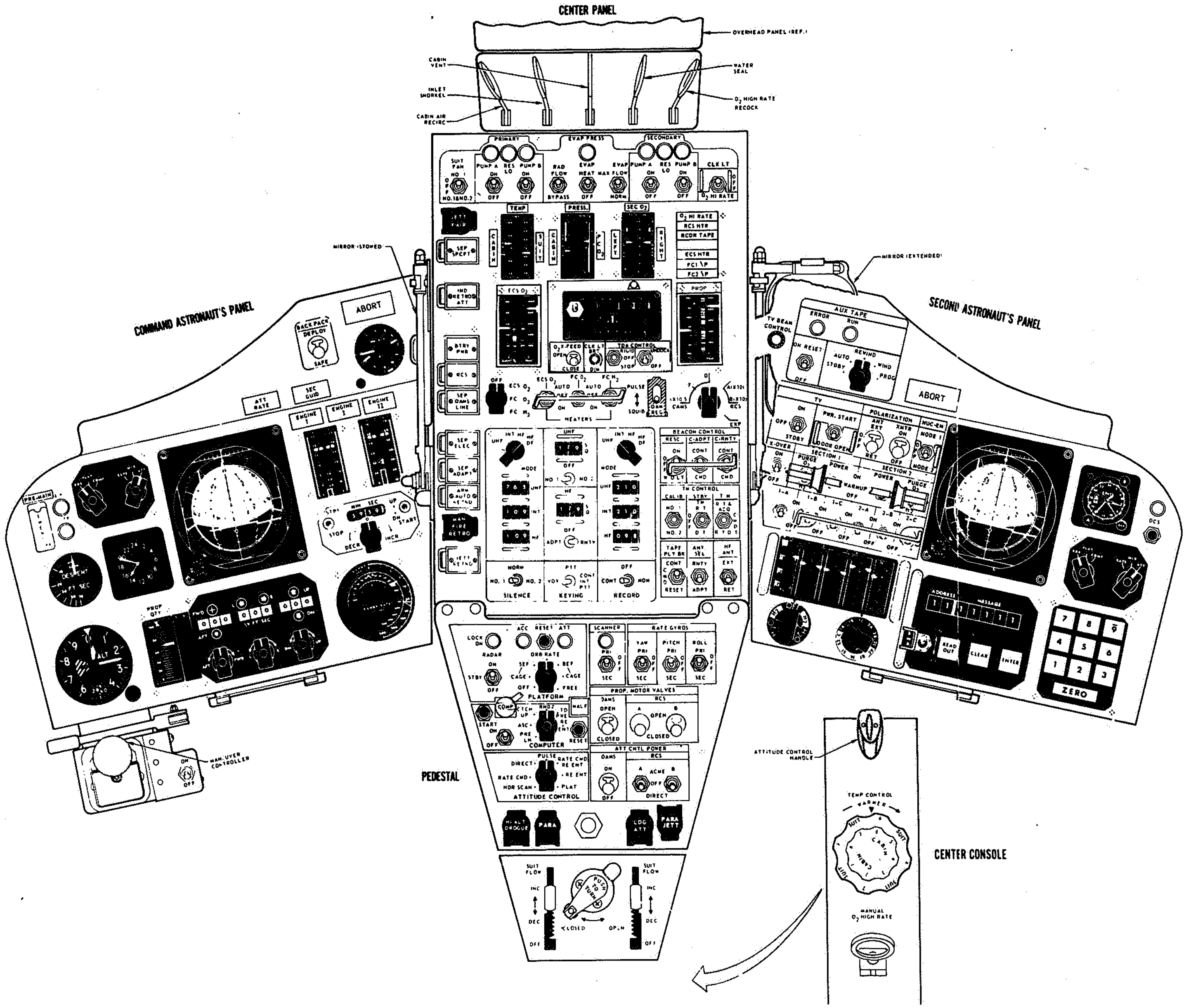

Figure 3.—Detail of Gemini spacecraft main display and control panel.

him. The center panel is largely a panel of subsystem status and controls. The command astronaut's panel (left) and the second astronaut's panel (right) contain attitude and flight data for both pilots. The second astronaut was integrated with the computer system.

Figure 4 shows the third generation of manned spacecraft. This is the main display console for the Command Module in Apollo. Some of the important displays on this panel are as follows.

Panel	Display
1	The altimeter
2	FDAI switches and an independent accelerometer output along the main fore-and-aft axis of the spacecraft
3	Engine data for both the launch vehicle and the propulsion system
4	The flight director attitude indicator
5	Launch vehicle status lights
6	Attitude-set group
7	Direct ullage information
8	Various switches
9	VHF antenna controls
10 and 11	Various status lights
12	Reaction control system indicators
13	Communications switches and data
14	Display and keyboard assembly through which the astronaut communicates with the computer in Apollo (this will be covered in more detail later)
15 to 20	Various indicators and a number of switches for controlling various subsystems comprising this relatively complex spacecraft

Figure 5 shows another important display and control panel in the Command Module—the guidance and navigation panel located in the lower equipment bay. Apollo provides the astronaut with the capability of navigating his spacecraft completely independent of ground control. This capability is exercised through this panel. Here, we can see the two major optical elements: the scanning telescope and the sextant shown in panel 4. Panel 2 shows the instantaneous angular positions of the various gimbals on the inertial measurement unit and the optical subsystem, and panel 6 is the display and keyboard assembly through which the astronaut communicates with the computer. With this system, the astronaut measures the angle between a star and a landmark on either the Earth or the Moon or measures the angle of the star with the horizon of these planetary bodies; this information is entered automatically into the computer, and the position of the spacecraft is determined automatically. Panel 5 contains various controls for carrying out navigation operations.

Figure 6 shows the relative location in the spacecraft of the guidance and navigation display and control panel and of the main display and control panel.

Figure 7 is the display and control system in the Lunar Module. The Lunar Module is the first manned spacecraft designed specifically to operate solely in conjunction with a second spacecraft and not to have an Earth-return capability. The Lunar Module has been designed without benefit of precedence and it contains enough duplication and crew-machine interaction to assure mission success. In this figure we can see that the right and left panels, that is, 10, 11, 12, and 1, 2, and 3, are made up largely of circuit breakers and switches for control of various subsystems. The triangular units toward the middle are the windows for the command pilot and the pilot of the Lunar Module. Panels 7 and 8 are the flight instrument panels. Panel 5 is the display and keyboard assembly, through which the crew communicates with the computer of the primary guidance and navigation system.

Figure 8 shows panels 5, 6, 7, and 8 of the Lunar Module in greater detail. Note that each pilot is provided a flight-director attitude indicator (eight-ball). There is also a series of instruments showing the temperature, pressure, and throttle setting for the main propulsion system and similar data for the reaction control system. The dials or the pointers show both fore-and-aft velocity and right-and-left velocity. To the right of the FDAI, on the left-hand panel, can be seen the abort buttons.

Figure 9 shows the face of the Apollo guidance computer, the display and keyboard assembly (DSKY). This display and control unit is a very significant element in the third-generation spacecraft, and may be of further significance if applied to high-performance aircraft. It is the means by which the pilot and

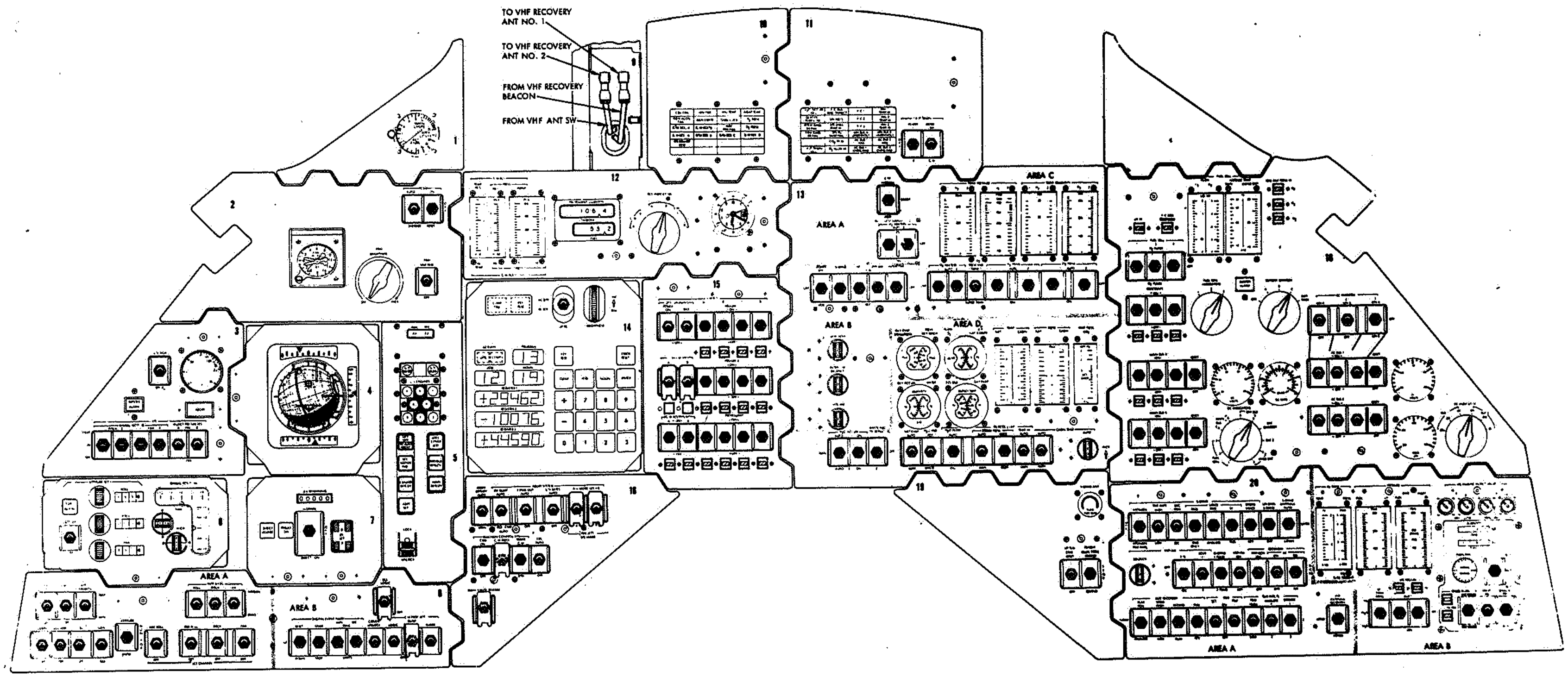

FIGURE 4.—Apollo spacecraft main display console (Command Module).

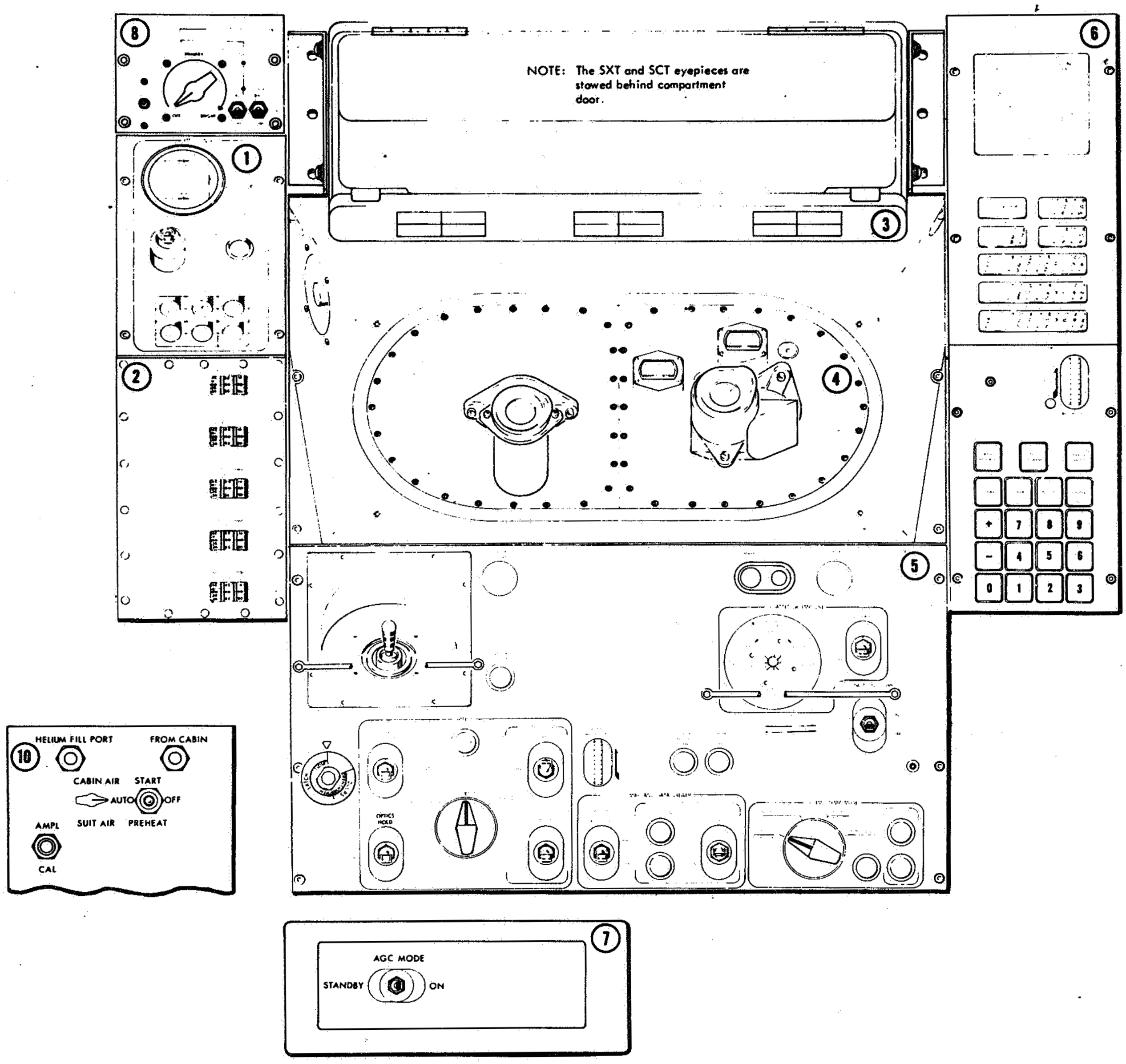

FIGURE 5.—Apollo spacecraft guidance and navigation panel (Command Module).

the computer are integrated into a single, unified system. This unit is on the main display panel of the Command Module, at the main guidance and navigation station in the lower equipment bay of the Command Module, and at the lower part of the main display panel in the Lunar Module. The unit is identical in all cases. It was made identical for ease of crew training and also for economy in production. The unit for each spacecraft is produced from the same production line at the factory. The DSKY contains certain lights to show the status of the computer. It also contains some switches, as indicated in figure 9. It contains a number of buttons such as the "key release," "error reset," "clear," "verb," "noun," "enter," numbers "0" through "9," "plus," and "minus." It contains an electroluminescent display of "program," "verb," and "noun," and three five-digit displays below these indicators. The "program," "verb," and "noun" each contain two-digit displays. The crew can select any one of 100 programs which are associated with the various mission phases, such as midcourse cor-

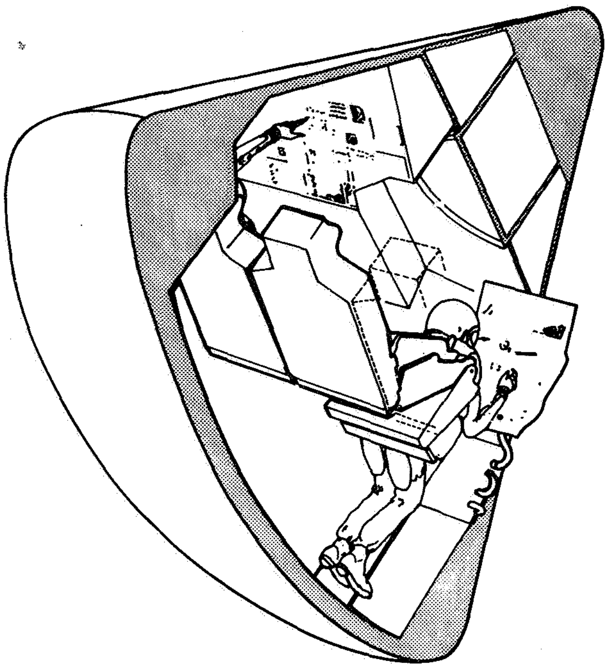

FIGURE 6.—Location within the spacecraft of the main guidance and navigation display and control panels.

rection program or lunar landing program. The crew can select for each of these programs any one of 100 verbs and 100 nouns. Hence, in each program the crew may choose 10^4 sentences, each consisting of a verb and a noun. For example, in the program corresponding to the lunar orbital phase, the crew may select a verb such as "compute" and a noun such as "abort velocity." The orthogonal components of abort velocity would be provided on the three five-digit displays below the "verb" and "noun" displays. These numbers would be constantly changing and would give the crew the abort velocity for safe return to the Earth if an abort were chosen at that precise moment while in lunar orbit. With 10^4 sentences for each program and 10^2 programs, 10^6 types of communication are available between the computer and the astronaut. This communications capability is limited only by the memory capacity of the computer, by the software developed for the computer prior to the mission, and by the crew's

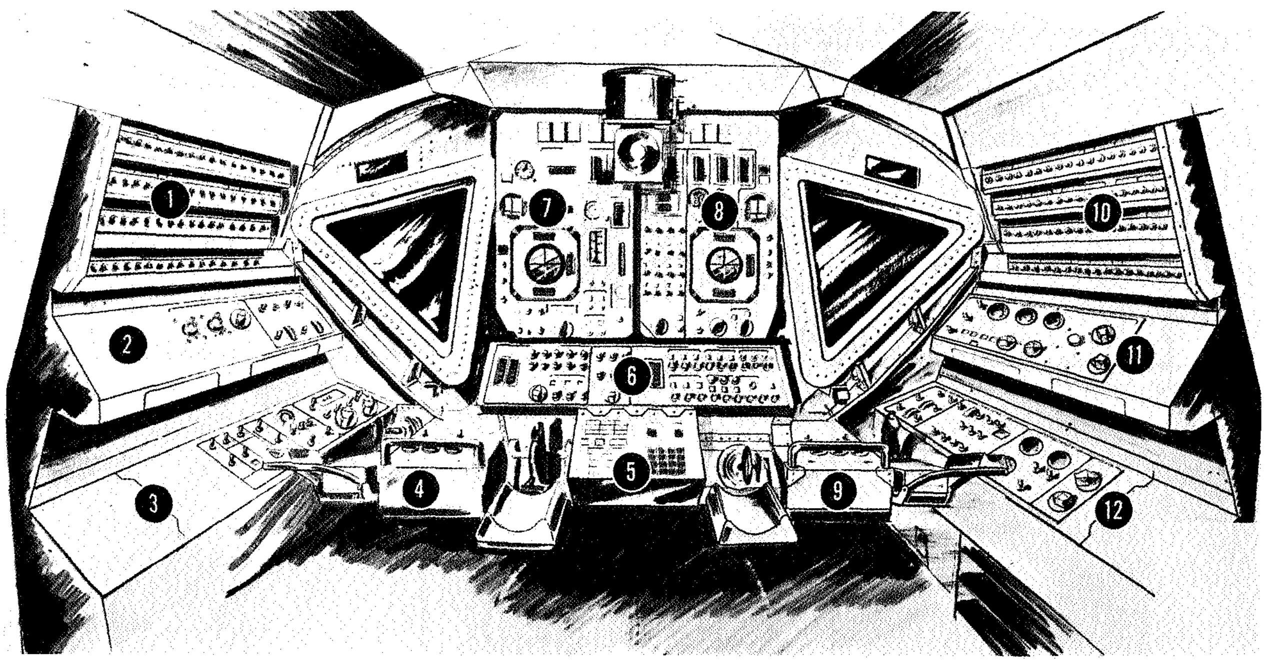

FIGURE 7.—Display and control system (Lunar Module). (1) Commander's upper side console: circuit-breaker panel. (2) Commander's center side console: power distribution panel; audio control panel. (3) Commander's lower side console: pyrotechnics console panel; radar panel; abort guidance panel. (4) Commander's lighting control panel. (5) Primary guidance and navigation panel. (6) Lower center panel: cryogenic storage panel; stabilization and control panel; power generation panel. (7) Commander's center panel. (8) Systems engineer's center panel. (9) Systems engineer's lighting control panel. (10) Systems engineer's upper side console: circuit-breaker panel. (11) Systems engineer's center side console: power distribution panel. (12) Systems engineer's lower side console: audio control panel; communications antenna panel.

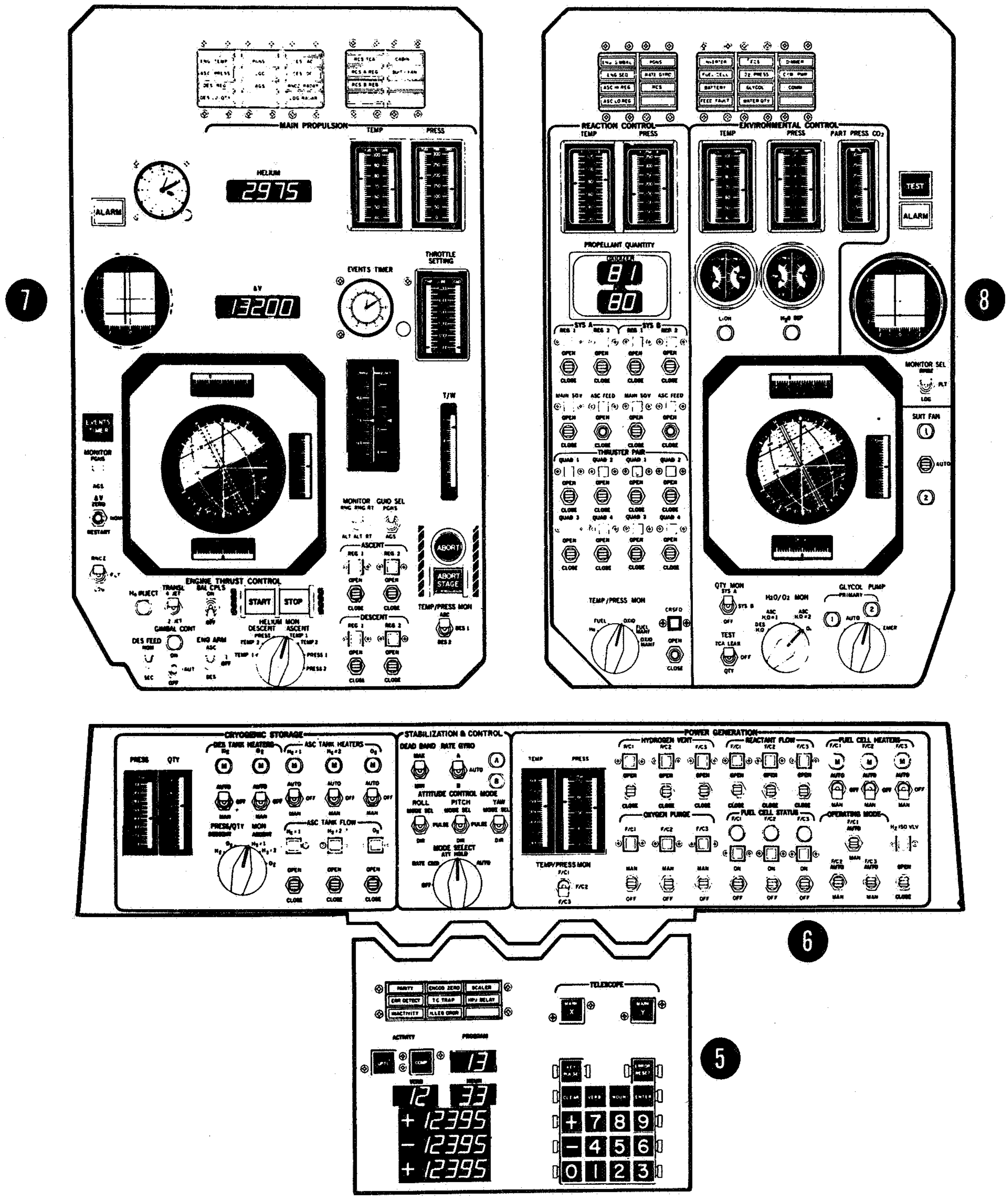

FIGURE 8.—Detail of panels (5), (6), (7), and (8).

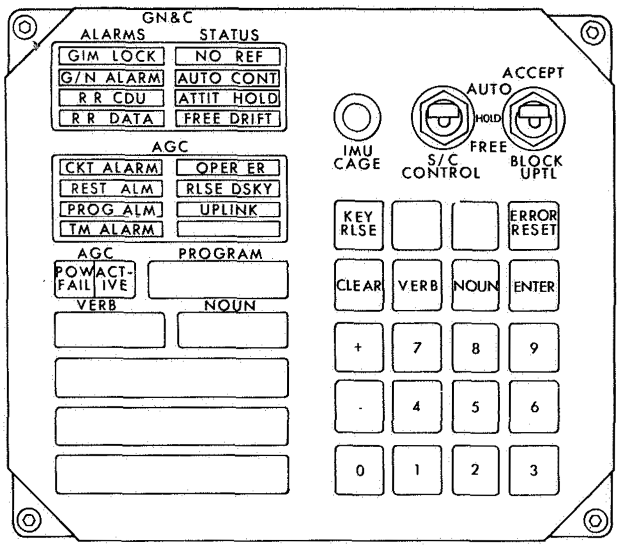

FIGURE 9.—Apollo guidance computer; display and keyboard assembly. (Representative only of display data and control functions, not physical configuration and arrangement.)

ability to use this system under all conceivable situations. Probably the latter is the limiting factor in the full exploitation of this capability.

MAN-MACHINE RELATIONS

What is the relative role of man and automated systems for control during the Apollo mission? Man will be the primary controller of the vehicle during the terminal phases of rendezvous, during the transportation and docking maneuver, and during descent below 1000 feet during the lunar landing maneuver. The transposition and docking maneuver is performed soon after injection from Earth orbit toward the Moon, at which time the Command and Service Modules are pulled away from the Saturn IV–B and Lunar Module. The Command Module is docked with the Lunar Module enabling the astronauts to enter the Lunar Module via the docking tunnel. At all other times throughout a nominal lunar mission, the vehicles will be controlled automatically. The astronaut will be monitoring the performance of the vehicle and will take over and control it manually in the event of equipment malfunctions.

The astronaut will compare the performance of redundant systems and will choose the solution that compares most closely with his own simplified solutions made from charts and nomograms. During Gemini rendezvous, for example, the astronaut compared the ground-calculated solution and the onboard closed-loop solutions calculated by his guidance system with a simplified solution made from charts prepared prior to the mission. The time of making burns and the magnitude of burns were chosen by the astronaut after comparing these three sources of data. During reentry, as another example, the astronaut had a predetermined nominal trajectory which he compared with the onboard derived closed-loop solution. He could choose to fly either automatically or manually with the onboard solution, or he could fly manually with the precalculated nominal solution.

The issue of man versus automatic guidance in spacecraft has too often been stated as a disagreement between the humanist and the technologist—the pilot and the engineer. This is a complex engineering situation which is amenable to solution by conventional engineering criteria. Man's role in guidance and control tasks may be evaluated on a technical basis by considering such factors as size, weight, power, bandwidth, versatility, reliability, and cost in relation to mission requirements. To perform such an engineering evaluation, the realistic aspects of the underlying engineering characteristics of a man as well as the automatic components must be evaluated.

The appropriate role of man in stabilization, control, and guidance of spacecraft is not as a permanent component of any of the normal action patterns, except in specific portions of the missions such as rendezvous or landing. He performs primarily as an off-line parallel or complementary observer. He must be provided with information in a form which is usable. If he can be provided with redundant information from dual sources, and if he is well trained and is carrying with him the proper aids, the pilot can act intelligently to enhance greatly the reliability and capability of his spacecraft guidance and control system.

N68-22305

MEDIA REQUIREMENTS FOR GENERAL AVIATION

Charles H. Gould and Roger L. Winblade
National Aeronautics and Space Administration, Washington, D.C.

General aviation includes all civil flying except that performed by the interstate and intrastate air carriers operating large aircraft. It embraces a multitude of aircraft uses ranging from pleasure flying and the transportation of people and cargo in privately owned aircraft by business firms, to such specialized uses as crop dusting, pipeline patrol, and aerial advertising.

In recent years there has been a strong uptrend in all areas of general aviation activity. All current indications are that this trend will continue and will even accelerate. There are around 110 000 general aviation aircraft in the current fleet, and about 15 000 are produced a year. The figures show that while the whole fleet is increasing, the preponderance of aircraft will be of the single-engine passenger type.

The general aviation fleet which emerged over the past 20 years has adopted such concepts as all-metal construction, tricycle landing gear, and navigation/communication avionics. In general, though, today's aircraft differ from their counterparts of 20 years ago, primarily because of increased engine power and payload with the corollary benefits of increased cruise speed and range.

Current light aircraft can make a 400- to 800-mile trip in 4 hours; this has established these aircraft as fine transportation for that relatively small segment of the population able to cope with the cost, pilot skill, and knowledge requirement. However, the use of these longer range aircraft for transportation greatly increased the probability of encountering poor weather on at least some segment of a trip. For effective utilization, the general aviation pilot is faced with virtually the same task as the professional military and commercial pilots. The training and proficiency are usually lower, resulting in much higher relative workload. Unfortunately, the displays with which the general aviation pilot must function, for the most part, have not been designed with total pilot workload as a major consideration.

It is true that many techniques exist that could reduce the pilot workload in general aviation aircraft such as stability augmentation, procedures, and displays. In general, these display advances have not been applied because of cost. The difficulties encountered under IFR flight in general aviation aircraft are not unique, and, for the most part, similar problems have been encountered and solved in military and commercial systems. While technically compatible, very few of these solutions can be used by general aviation because of the high cost.

The direct result of the pilot workload reaching total pilot capacity, coupled with the lack of workload-relieving displays and controls, is an increasing rate of clearance deviations, near-misses, and eventually accidents involving general aviation. The potential market growth of general aviation may never be achieved due to this situation. Displays, and especially display media and hardware, offer the promise of solution to these problems.

An immediate need in general aviation is the development and production of equivalents to

291-351 O—68——3

those workload-relieving devices available to commercial and military aircraft. The resulting user cost of direct duplication is prohibitive; consequently, considerable effort must be directed toward achieving comparable functions, accuracy, and reliability through new techniques. The techniques of display format design are in hand, but could probably be greatly improved for the general aviation pilot. The techniques of display media, in order to display the required information for general aviation, are not now sufficient. The results of this symposium should be viewed with this in mind.

One thing is clear. The workload advantages of an integrated display, whether for navigation, flight control, or engine status, indicate the desirability of the integrated approach format in general aviation. Electromechanical displays, the mainstay of general aviation display media, are seldom suited to complex, two-dimensional integrated formats. The media papers scheduled during this symposium describe, in general, more suitable approaches. The economics, reliability, and compatibility with existing avionic equipment must be stressed.

It seems possible that in the future we can consider digital computer-driven displays for general aviation, for the following reasons:

(1) Routine flight management requires many computations which, though relatively simple individually, are a major workload factor in total.

(2) Small multipurpose digital computers, comparable in complexity and cost to a good desk calculator, can greatly reduce this workload, and probably will appear in general aviation aircraft.

(3) At present, many display devices include low-order computation, which, in total, justifies consideration of centralized computation.

Coupling these factors, computer-driven integrated displays are a possibility for general aviation, provided the total installed system cost can be made competitive.

A major factor in general aviation displays, almost completely aside from the technical factors we have discussed, is the marketing side of the picture. General aviation manufacturers will produce what will sell, and display development programers should keep in mind that eventual market reduction will be necessary. Also, the range of capability demanded by the customers implies a system modularity. In other words, some customers may want complete and dual navigation equipment, while others may want a minimum. In fact, each aircraft is a personal vehicle, and will reflect the individual's desires, biases, and financial capability.

If these general aviation requirements are considered in display media developments and adequate displays are made available, a major contribution to this important and expanding field will be possible. With increased safety will come increased utilization and a larger market potential.

N68-22306

RECENT ADVANCES IN CATHODE-RAY-TUBE DISPLAY DEVICES

JOSEPH A. DAVIS

Purdue University, Fort Wayne Regional Campus

The present state of the art in cathode-ray-tube (CRT) displays is reviewed, and current research and development in CRT components is described, along with the most recent application of these materials. Future prospects regarding the CRT are discussed. The major points discussed in this paper include the following: methods for generating electrons; developments in luminescent materials; resolution, brightness, and contrast techniques; new display tubes; developments in color CRT technology; developments in bandwidth reduction; and developments in projection display techniques.

INTRODUCTION

The CRT is a well-known electro-optical conversion device which, in its simplest form, consists of a source of electrons, called the cathode, and a luminescent material, or phosphor, deposited on the inner surface of an optically transparent container that is evacuated to a pressure less than 10^{-6} torr. The electrons emitted by the cathode can be collimated into a beam and deflected by means of electric and/or magnetic fields and, hence, be directed to impinge on any area of the phosphor layer (or, as it is normally referred to, the phosphor screen). The electron beam can be made to trace a regular pattern or raster on the phosphor screen electronically. Modulation of the beam is accomplished by inserting a control grid near the cathode and applying a variable voltage to the grid.

Within the past year, a number of articles have appeared in the literature describing a variety of applications of CRT's. The design factors to consider in choosing a CRT for information display have been recently described (ref. 1). This paper describes operating voltage, power, brightness, sensitivity, size of the display, deflection and focusing methods, and resolution capability and their interrelationships. The information presented should provide a convenient guide for the system designer.

An earlier article (ref. 2) describing CRT's for display systems should prove useful in conjunction with reference 1. More attention is given to electron-gun types, resolution measurement, and tube life.

Graham (ref. 3) has reported the use of a standard CRT monitor in alphanumeric displays. DuBois (ref. 4) discussed a miniature CRT for the same purpose, approximately ½ by 1 inch, which operates under low-power conditions, providing a display easily readable in a moderately lighted room.

The use of the CRT as a microscope has been applied to the examination of missile case-wall materials by converting the X-ray image to an electrical signal through an X-ray vidicon (ref. 5). There are many applications of CRT's as the interface between man and machine for the purpose of displaying stored information. Winfield (ref. 6) reported on a computer-controlled, multichannel CRT symbol generator. Stotz and Cheek (ref. 7) described the use of a direct-view storage tube with an alphanumeric keyboard console as a replacement for the remote teletypewriter connected to a central computer. The use of a CRT-computer combination in the educational field has also been reported (ref. 8). In this system, which has as many as 32 stations connected to a central processing unit, the student uses a light pen device to indicate answers to questions by covering the correct figure on the CRT face. A discussion of the sensitivity of the light pen is given in reference 9, while reference 10 describes the operation and use of the beam pen.

The light pen is used to detect light produced on a CRT screen employing a photomultiplier or similar light-sensitive device. The beam pen, however, actually detects the electron beam producing the luminescence on the phosphor screen. It is a conducting probe capacitively coupled to the beam, and, as the beam approaches the pen, the capacitance decreases and the signal increases.

To complete this survey of CRT uses, 35 models of display consoles, described in reference 11, are also included, along with suggestions for proper selection.

The remainder of this discussion will be devoted primarily to developments recently reported and also to the few new discoveries that have not been described in the easily accessible literature. A description of the more recent research studies will be given from which new and improved materials for CRT fabrication result, because the capabilities of the next generation of CRT's are determined by the more fundamental studies of today.

METHODS FOR GENERATING ELECTRONS

There are four significant areas of study—research on the "cold cathode," the nonstoichiometric compound thermionic cathode, the well-known mixed alkaline earth oxide cathode, and the pure metal or alloy cathode.

The effort in cold cathodes is now taking an approach familiar to the semiconductor scientist; namely, the application of III–V compound semiconductors such as gallium phosphide with platinum, tungsten, or other metals evaporated onto the surface, and barium oxide evaporated onto such metals as tungsten (ref. 12). These studies are being done in an effort to lower the surface barrier sufficiently to permit electrons to escape into the vacuum region by application of an electric field. At present, the most promising combination appears to be GaAs/Cs, in which the electron affinity apparently approaches zero. Band bending occurs with very thin GaAs layers, and emission efficiencies of 50 percent are claimed (refs. 13 and 14) based on photoemission studies. The combination of evaporated BaO on GaAs is also under study. It is felt that negative values of electron affinity will be obtained, giving still higher emission efficiencies.

The nonstoichiometric compound-type thermionic cathode research has evolved from the initial field emission studies (ref. 15) and reports on cathodes (refs. 16 and 17). The emission studies have shown that a pronounced minimum in the work function occurs, which appears to be attributable to a mixture of Sr_3WO_6 and $SrWO_4$ on tungsten. The preparative conditions are such that a high degree of nonstoichiometry exists in the final surface after thermal activation. Lowering the work function of a surface permits greater ease of removal of electrons; hence, one can expect to obtain a more copious electron emission and/or much lower operating voltage current conditions.

An investigation (ref. 18) of alkaline earth tungstates for thermionic cathodes is presently being conducted, based primarily on the work of Cape and Coomes. The very low work function of these cathodes promises high current densities, possibly as much as 100 A/cm^2.

The studies of high-current, high-resolution CRT's have led to the development of high-current-density cathodes in which current densities of 10 to 20 A/cm^2 have been obtained. However, some attendant serious problems exist regarding ease of contamination by gaseous substances. These newer cathodes are composed of a mixed alkaline earth basic tungstate, with tungsten metal powder and zirconium hydride (ZrH_2). They are formed at high pressure and activated at elevated temperatures. Under these conditions, a nonstoichiometric alkaline earth tungstate compound is, in all probability, formed. Under ideal conditions of preparation, including careful cathode design and exclusion of gases, these cathodes exhibited extremely long life (9000 hours) and high current densities.

The conventional triple-oxide-on-nickel thermionic cathode (BaO, SrO, CaO) up to now was capable of producing 1 A/cm^2 under the most ideal conditions (ref. 19). A recent paper (ref. 20) on the high current density operation of oxide-coated cathodes describes the achievement of high current density output from diode- and triode-structure cathodes under dc and pulsed conditions. In diode structures, current densities as high as 6 A/cm^2 (dc operation) were realized, while in triode structures, 2 A/cm^2 current densities were obtained. Lifetimes on the order of 1500 hours at current levels greater than 1 A/cm^2 were obtained through improved tube processing and design techniques with triodes. Interestingly, the limitations on tube operating conditions and life were found to be attributable to tube elements other than the oxide cathode itself.

The efforts of Albert, Atta, and Gabor (ref. 21) in thermionic converter research have led to an improved all-metal electron emitter. Mixtures of tungsten, thorium, and a small amount of zirconium, coated on tungsten wire, produced emission densities up to 30 A/cm^2. When tungsten carbide was substituted for tungsten powder in the mixture, even higher emission densities were obtained, with operating temperatures as high as 1600° C. Lanthanum may be substituted for thorium, yielding equivalent emission densities at temperatures about 200° C lower. In these substitutions, the emission density for safe operation is about 10 A/cm^2. If hafnium is substituted for zirconium, significantly higher current densities are obtained (about 100 A/cm^2). The lifetimes of these cathodes are comparable to pure tungsten cathodes. The role of the small concentrations of hafnium or zirconium in these cathodes appears to be one of a "mobilizer," which increases the rate of transport of activator atoms, lanthanum or thorium, to the surface of the cathodes. A summary of the various types of cathodes is given in table I.

Developments in Luminescent Materials

Factors that should be considered when discussing new phosphors include the measured characteristics of the phosphor powder and the measured characteristics of the screen in a tube. Each year many new phosphor powders are synthesized and reported, but few ever find application in a device. The measurement of phosphor parameters in tubes is therefore an important factor. The various ways in which measurements can be carried out and the relative merits of these methods have been reported and discussed (ref. 22).

TABLE I.—*Cathode Characteristics*

Type	Composition	Maximum emission density, A/cm^2	Operating temperature, °C	Life, hr[a]
Cold cathode	GaAs/Cs		Room	Very long
	GaP/W			
	BaO/Pt(W)			
	GaAs/BaO			
Nonstoichiometric	Ba-Sr-tungstate/W	20	1000	9000
	Sr_3WO_6-$SrWO_4$/W	100	1000	
Pure metal	W-Th(La)-Zr(Hf)	100	1400	1000
Triple oxide	BaO-SrO-CaO/Ni	6	727	1500

[a] Life measurements not at maximum emission density.

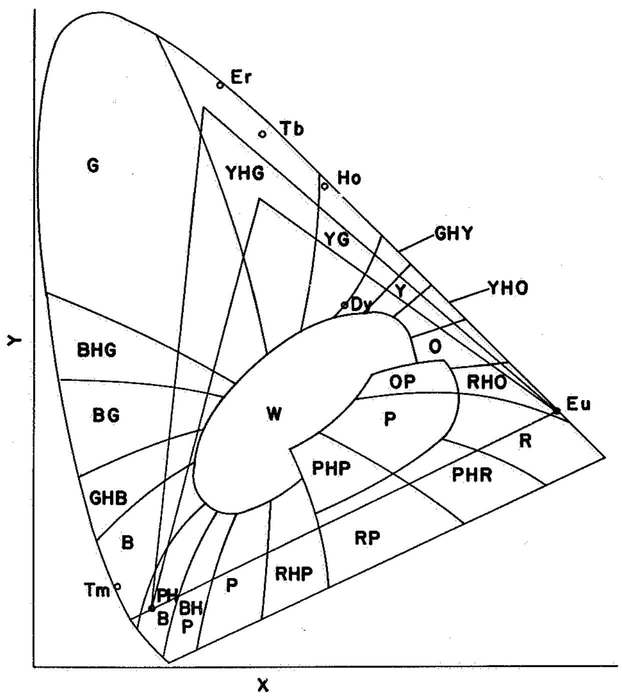

FIGURE 1.—Attainable color gamut in CRT's and location of lanthanide-activated phosphor coordinates.

New Visible-Emitting Phosphors

The most recent compositions are based almost exclusively on rare-earth-activated host crystals. Specifically, the europium-activated oxides, or oxysulfides such as yttrium oxide (ref. 23), yttrium vanadate (refs. 24 and 25), gadolinium oxide (refs. 26 and 27), and yttrium oxysulfide (refs. 24 and 25), are among the most efficient yet synthesized. All emit narrowband luminescence in the red region of the visible spectrum. Zinc sulfide activated with thulium has a narrowband emission in the blue region and is comparable to zinc sulfide-silver in visual brightness, but has a lighter blue emission and is therefore not considered suitable for color television. Indium orthoborate activated with terbium has an efficient green emission comparable to P–1, zinc silicate-manganese (refs. 28 and 29). With regard to improvement in the color gamut (the range of chromaticity coordinates) of full-color displays, the most obvious need is to increase the y chromaticity coordinate of the phosphor as illustrated in figure 1. The newer visible-emitting phosphors and their characteristics are shown in table II.

The energy conversion efficiency of phosphors under electron excitation has been calculated for several selected materials (refs. 30, 31, and 32). The efficiencies for some of these phosphors are shown in table III.

It is interesting to note that only the sulfides exhibit efficiencies higher than 15 percent, while most oxides are less than 10 percent.

To illustrate the relative intensities of luminescence, the spectral energy distribution for several new phosphors and two well-known phosphors is shown in figure 2. All of the emission curves were recorded under identical conditions of current density and accelerating voltage on powder samples for each phosphor and are uncorrected for monochromator and photomultiplier wavelength sensitivity varia-

TABLE II.—*Visible-Emitting Phosphors*

Composition	CIE coordinates		Peak wavelength, nanometers	Radiant efficiency, percent
	x	y		
$ZnS:Tm$	~0.11	~0.10	478	----------
$InBO_3:Tb$	.31	.66	542 550	8
$Y_2O_3:Eu$	.64	.36	611	7
$YVO_4:Eu$	.675	.325	615 619	7
$Y_2O_2S:Eu$	.66	.34	615 630	7
$Gd_2O_3:Eu$	.64	.36	611	8

TABLE III.—*Conversion Efficiency of Selected Phosphors*

Composition	Code No.[a]	Peak wavelength, nanometers	Measured conversion efficiency, percent
ZnS:Ag	NBS 1020	452	21
ZnS:Cu	NBS 1022	525	11
ZnCdS:Ag	NBS 1023	570	19
ZnCdS:Ag	NBS 1024	590	12
Zn_2SiO_4:Mn	NBS 1021	525	8
ZnS:Mn	----------	591	4
ZnO:Zn	P–15	505	7
$CaWO_4$	----------	430	3
$CdSiO_3$:Mn	----------	590	5
$Zn_3(PO_4)_2$:Mn	NBS 1025	643	6

[a] The National Bureau of Standards (NBS) phosphors correspond to the following P-number phosphors:

1020	Blue component of P–4 all-sulfide phosphor
1021	P–1
1022	P–2
1023	Yellow component of P–4 all-sulfide phosphor
1024	Orange component of P–14 phosphor
1025	Red component of old P–22 phosphor

tions. A corrected set of curves would show about a twofold improvement in the output of the red phosphor.

New Ultraviolet-Emitting Phosphors

The P–16 type phosphor, calcium magnesium pyrosilicate-cerium-lithium (ref. 33) and P–15, a zinc oxide phosphor, have been the only useful uv phosphors employed for over a decade. The main application for such phosphors has been in flying spot scanning tubes, in which the primary attribute of the phosphors has been their rapid decay properties. In recent years, however, increased interest in the use of phosphors having high-intensity emission in the ultraviolet has been demonstrated for the purpose of dry-film recording. The P–16 phosphor is employed for this purpose, but considerable difficulty has been encountered because of the rather high power output required for available dry-film materials and the fact that the P–16 phosphor experiences rapid aging or deterioration under high-current-density operation in CRT's.

Within the past 3 years, two significant studies have been conducted to develop new ultraviolet phosphors (refs. 34 and 35). While some discrepancies have been found (ref. 34), a number of promising materials were discovered, and several new and improved phosphors were reported (ref. 35). The spectral energy distribution for several of these phosphors is shown in figure 3. Also shown is a new ultraviolet phosphor which exhibits an energy emission intensity nearly twice that of P–16 and at a peak wavelength of 300 nanometers.

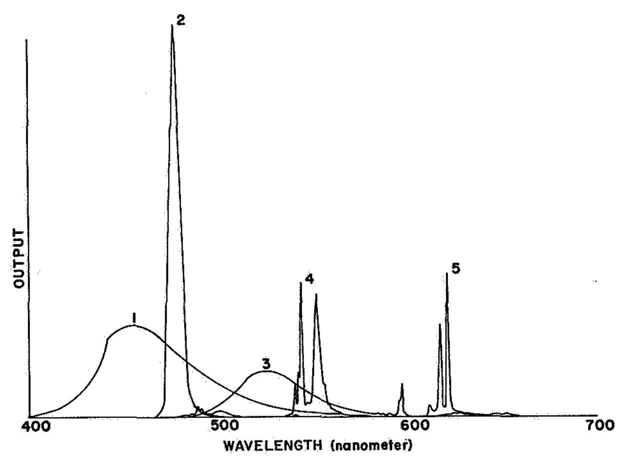

FIGURE 2.—Spectral energy distribution for visible-emitting phosphors. 1. ZnS:Ag:Cl. 2. ZnS:Tm. 3. Zn_2SiO_4:Mn. 4. $InBO_3$:Tb. 5. YVO_4:Eu.

TABLE IV.—*UV-Emitting Phosphors*

Composition or code designation	Peak, nanometers	Accumulated charge density to 50% point, C/cm²	Peak intensity, % relative P–16	Decay time, μsec
$Ca_2MgSi_2O_7$:Ce	385	0.1	100	0.06
$CaSiO_2$:Pb	345		20	.23
$Ca_3Si_2O_7$:Pb	355	>0.1	40	.19
$BaSi_2O_5$:Pb	355	0.1 to 1.0	150	.25
JJ13K	300	>0.1	180	.25
$SrSi_2O_5$:Pb	280		40	.25

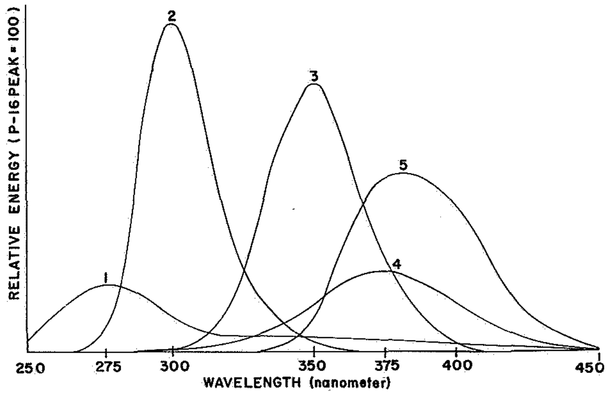

FIGURE 3.—Spectral energy distribution for uv-emitting phosphors. 1. $SrSi_2O_5$:Pb. 2. ITT phosphor JJ–13K. 3. $BaSi_2O_5$:Pb. 4. $Ca_3Si_2O_7$:Pb. 5. P–16.

Table IV contains selected ultraviolet phosphors and their pertinent properties.

It is noteworthy that all the newer uv-emitting phosphors contain lead as an activator, and most are silicate compounds. In addition, table IV shows that all of the phosphors reported have longer decay times than the P–16. These decay times have been determined by the method of Sisneros (ref. 36).

Infrared-Emitting Phosphors

While there is relatively little interest in phosphors emitting in the infrared spectrum, it is appropriate to point out that such phosphors do exist. Table V lists several representative phosphors and their characteristics. It is interesting that, from a theoretical standpoint, preferential emission from the phosphors ZnS:Ag:V or ZnS:Cu:V is entirely from the 2.0-μ band of vanadium (ref. 37). The normal silver- or copper-activated visible emission is "quenched." Cobalt and iron, normally considered as luminescence poisons, actually have emission bands in the infrared; while nickel, another efficient "poison," apparently does not have an infrared emission band, at least below 5.0 μ (ref. 38).

Registered Phosphors

At the request of the Electronic Industries Association (EIA), the Joint Electron Device Engineering Council (JEDEC) Committee 6.3 has recently revised and expanded their publication (formerly no. 16, now no. 16A) (ref. 39) on the optical characteristics of CRT screens. In order to register phosphors, certain requirements or rules must be fulfilled, as specified in reference 39. Some representative phosphor registrations are given in table VI.

JEDEC publication 16A is quite useful to the CRT display engineer in that a special CRT incorporating any of the listed phosphors can readily be obtained from most CRT producers.

The most recent registration of P-number

TABLE V.—*Infrared-Emitting Phosphors*

Composition	Peak wavelength, μ	Conversion efficiency, percent
ZnS:Ag:V	2.0	2.5
ZnS:Cu:V	2.0	.5
CdSe:Cu:Ga	1.2	.7
CdSe:Au:In	1.2	.6
ZnS:Fe	.66	
ZnS:Co	3.2	

phosphors is listed in table VII. All appear to be nearly identical with earlier P-number phosphors registered, except the persistence (decay) is made shorter, or longer, than their earlier counterparts by variation in synthesis methods. The use of a given P number to represent several completely different phosphors, or phosphor combinations, is confusing to the design engineer, and there is an obvious need for clarification of the P-number designation method.

RESOLUTION, BRIGHTNESS, AND CONTRAST TECHNIQUES

Fundamentally, resolution and brightness are limited by a number of fixed conditions in a CRT. Resolution is determined by the beam-spot size and the phosphor particle (or, more correctly, aggregate) size, disregarding other contributing factors such as screen thickness and type of viewing surface. Brightness is limited by the screen current attainable and phosphor luminous efficiency, in addition to the limited controllable conditions of light absorption and reflection in the phosphor and viewing surface. Achieving high contrast is one problem area that has been solved in the "optical diode" CRT, to be described later.

The relation between these and other quantities has been discussed quite recently (ref. 40), with particular attention given to very-high-resolution CRT's exceeding the resolving power of the human eye. The spot size of the electron beam is described, along with the effects of tube geometry on resolution. In addition, phosphor light output, blemishes, noise, and uniformity are discussed, giving special attention to the associated electronic equipment.

Image-recognition experiments (ref. 41) have been conducted wherein a raster scan display was used to examine the effects of variations in visual angle, active to inactive raster line ratio, number of raster lines, contrast, image classes, dimension ratios of images, and image orientation.

Several observations may be made regarding brightness and resolution, which have not been described extensively. Because simultaneous high resolution and brightness are incompatible, a compromise is always necessary, and is determined by the particular requirement for the CRT display. This incompatibility was considerably lessened by the phosphor scientist, in that the old idea that phosphor efficiency decreased with particle size has been proved false. It is now possible by suitable synthesis techniques to prepare small-particle phosphors with light output equal to the more familiar large-particle compositions. Phosphors with average particle size of 0.5 μ can be synthesized. This achievement, coupled with the anticipated developments in very high current density cathodes, should make possible a significant improvement in high-resolution, high-brightness CRT's. Full realization of this improvement will require fiber optic faceplates on the CRT. Currently, fiber optic faceplates are available with 85 percent transmission in the visible spectrum and a fiber diameter of 4 μ.

The "optical diode" approach (refs. 42 and 43) to high-contrast CRT's is one of the more unique and interesting developments in recent years. Basically, the system incorporates two filters: a uv-transmitting, visible-absorbing filter, and a visible-transmitting, uv-absorbing filter. A diagram of the faceplate structure is shown in figure 4.

A uv cathodoluminescent phosphor is excited by the electron beam. This uv phosphor is settled on a uv-transmitting, visible-absorbing bandpass filter. The uv luminescence excites a visible emitting phosphor sandwiched between the uv-transmitting, visible-absorbing layer and a visible-transmitting, uv-absorbing bandpass filter. Thus, only the luminescence emission in the visible is transmitted through the filters to the viewer, while ambient light is absorbed in these filters, thus rendering an extremely high-contrast image.

Palilla (ref. 25) has discussed the possible application of narrowband phosphors to high-contrast color CRT's. Employing three rare-earth-activated phosphors emitting narrowband luminescence in the blue, green, and red spectral regions, and a narrow-bandpass filter corresponding to these narrowband phosphors, one can effectively absorb ambient illumination and realize a reasonably high-contrast display.

TABLE VI.—*Characteristics of Selected P-Number Phosphors*

P No.	Composition	CIE coordinates		Peak wave-length, nano-meters	Decay time to 10% level	Color		Uses
		x	y			Fl [a]	Phos [b]	
P-1	Zn_2SiO_4:Mn	0.218	0.712	525	24 msec	YG	YG	Radar, oscilloscopes.
P-2	ZnS:Cu	.279	.534	543	35 to 100 μsec	YG	YG	Oscilloscopes.
P-4	ZnS:Ag+	.270	.300	440	25 μsec	W	W	Black and white standard.
	ZnS(50) CdS(50) Ag	------	------	565	60 μsec	------	------	
P-5	$CaWO_4$	.169	.132	430	25 μsec	B	B	Photography.
P-7	ZnS:Ag+	.151	.032	440	40 to 60 μsec	W	YG	Radar and oscilloscopes.
	ZnS-CdS:Cu	.357	.537	560	0.4 sec	------	------	
P-11	ZnS:Ag	.139	.148	460	25 to 80 μsec	B	B	Photography.
P-16	$Ca_2MgSi_2O_7$:Ce	.175	.003	385	~0.12 μsec	UV	UV	Flying spot scanning systems and photography.
P-20	ZnS-CdS:Ag	[c].444	[c].536	520–560	0.05 to 2 msec	Y to YG	Y to YG	Image-conversion devices. TV–P–4 yellow comp.
P-22B	ZnS:Ag	.155	.060	440	~25 μsec	B	B	Color TV.
P-22G	ZnS:CdS:Ag	.285	.600	520	60 μsec	YG	YG	Color TV.
P-22R	YVO_4:Eu	.675	.325	615–619	0.9 msec	OR	OR	Color TV.
P-25	$CaSiO_3$:Pb:Mn	.557	.430	610	45 msec	O	O	Long-persistence displays up to 10 sec.
P-27	$Zn_3(PO_4)_2$:Mn	.674	.326	635	27 msec	R–O	R–O	Color TV monitors.
P-28	ZnS:Ag:Cu	.370	.540	550	~0.5 sec	YG	YG	Radar.
P-31	ZnS:Cu	.193	.420	522	4 μsec	G	G	Oscilloscopes—high efficiency phosphor.
P-33	MgF_2:Mn	.559	.440	585	------	O	O	Radar—very long decay.
P-34	ZnS:Pb:Cu	.235	.364	490	~40 sec	BG	YG	Oscilloscopes, radar, visual.
		.409	.564	535	------	------	------	Information storage.
P-35	ZnS-ZnSe:Ag	.286	.420	540	1 msec	G	B	Oscilloscopes.
		.200	.245	485				

[a] Fluorescence-luminescence during excitation. [b] Phosphorescence-luminescence after excitation is discontinued. [c] Average.

Notes

P–1 General reference standard for relative peak efficiency measurements.

P–2 Decay decreases as beam current increases.

P–4 Sulfide version of P–4; there are two others referred to as sulfide-silicate and all-silicate.

P–11 May contain nickel to decrease persistence.

P–20 A range of CIE coordinates and, hence, colors are permitted. Coordinates stated are for a representative P–20 employed in preparation of P–4.

P–22 Data for sulfide blue and green, and vanadate red. The vanadate red was described earlier.

P–28 Curve has shoulder on long-wavelength side at 580 nanometers.

P–31 Curve also has blue peak at 450 nanometers.

P–33 Decay decreases as beam current decreases, burning very rapidly when used with stationary or slow-moving electron beam.

P–34 Infrared stimulable phosphor. Y-phosphor.

P–35 Resists burning compared to P–11.

TABLE VII.—*New Registered Phosphors*

P No.	Composition	CIE coordinates x	CIE coordinates y	Peak wave-length, nano-meters	Decay time to 10% level	Color Fl	Color Phos	Uses
P-36	(ZnCd)S:Ag:Ni	0.400	0.543	550	0.25 μsec	YG	YG	Flying spot scanning tubes.
P-37	ZnS:Ag:Ni	.143	.208	470	0.155 μsec	B	B	Flying spot scanning tubes.
P-38	(ZnMg)F_2:Mn	.561	.437	600	1040 msec	O	O	Integrating phosphor for low repetition rate displays and radar.
P-39	Zn_2SiO_4:Mn:As	.223	.698	525	150 msec	YG	YG	Integrating phosphor for low repetition rate displays and radar.
P-40	ZnS:Ag	.276	.3117	440	150 μsec	W	YG	Integrating phosphor for low repetition rate displays and radar.
	(ZnCd)S:Cu	----	----	555	0.55 sec	----	----	Integrating phosphor for low repetition rate displays and radar.

Notes:
P-39 similar to P-1, but with longer decay.
P-40 similar to P-7, but with longer yellow decay.
P-37 similar to P-11, but with shorter decay.
P-36 similar to P-20, but with shorter decay.
See also notes to table VI.

The tube face itself could conceivably be the filter, and considerably lower luminous output would be required for an acceptable display.

A method has been reported in which the picture quality of a CRT display is much improved (ref. 44). Referred to as the "image enhancer," this system compares and enhances the contrast by using a short-term storage or memory bank to store picture information to allow comparison of one point to all points surrounding it and then enhancing the contrast of these points.

Recently, Hilborn and Stevenson (ref. 45) reported a method for improving apparent resolution that enhances character recognition as much as 75 percent for the smallest characters generated. It is based on the ability of the human visual system to perform a short-term integration of information in successive TV frames.

The method involves changing the normal 0° phase relation between horizontal and vertical oscillators from one frame to the next by a switchable time-delay network in the vertical oscillators of both the camera and the monitor.

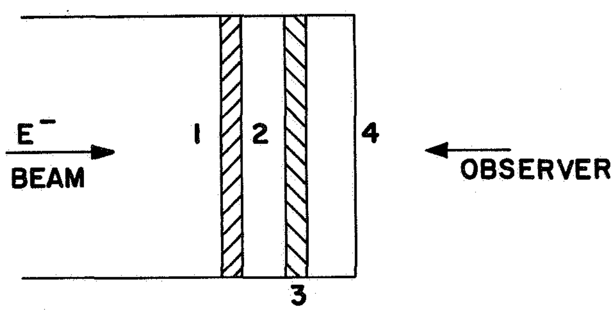

FIGURE 4.—Optical diode faceplate. 1. uv-emitting phosphor. 2. uv-transmitting, visible absorbing filter. 3. Visible-emitting, uv-excited phosphor. 4. Visible-transmitting, uv-absorbing filter.

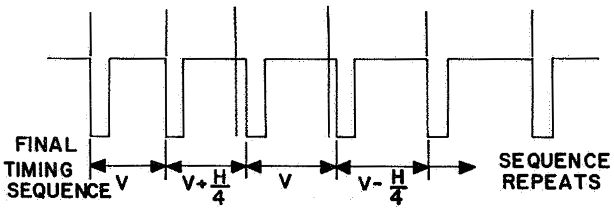

FIGURE 5.—Field-frame interlace waveform sequence.

Thus, by delaying two successive frames and then repeating the sequence, one obtains twice the normal number of raster lines and an interlace of frames as well as fields, provided one maintains the correct timing relationship. This is shown in figure 5 where V is the duration of one field (16.689 milliseconds) and H is the duration of one scan line (63.492 microseconds).

NEW DISPLAY TUBES

The report by Winningstad (ref. 46) discusses the relative advantages and limitations of the use of storage CRT's in computer output applications.

The paper by Engel and Eldridge (ref. 47) describes the use of an electrostatic display storage tube, originally developed by Clauer and Kuehler (ref. 48) with a new electron-gun design and an independent optical system consisting of a dual-chamber CRT. These chambers are separated by a thin dielectric membrane upon which recording and storage is done. The outer chamber contains an electrographic developer consisting of a mixture of carrier beads and a pigment that yields a net positive charge on the pigment when triboelectrically excited. This mixture then adheres to the negatively charged areas on the dielectric membrane, producing a high-resolution image that can be projected on a screen. Erasure is accomplished with a low-voltage floodbeam on the dielectric membrane, which emits secondary electrons more copiously under these conditions, and, thus, discharge is effected.

Tilton (ref. 49) has described the principles of three-dimensional CRT displays, which are too involved for discussion here. Strikingly real simulations of three-dimensional display have been realized.

An "anaglyph" (stereoscopic CRT display system) has also been reported by Wolvin (ref. 50). The stereo effect is obtained by the use of parallax in the image display, with the left eye viewing only the left field and the right eye viewing only the right field through the use of colored displays and special viewers. Both the left and right pictures are displayed on the same CRT and recorded on film for projection viewing.

DEVELOPMENTS IN COLOR CRT TECHNOLOGY

There are perhaps a dozen different "conventional" approaches to the realization of color CRT display in which variations in tube shape, phosphor screen structure, electron-gun number, position, and accelerating voltage are employed. The most recent developments, however, involve the first attempt to affect color variation in a display by means of the beam current variation in a single phosphor layer (ref. 51). Almost all phosphors exhibit a linear variation with normal beam current and accelerating voltage levels. There are at the present time a number of phosphors emitting broadband cathodoluminescence, in which the output versus current density is nonlinear. Several of these phosphors are illustrated in figure 6.

If one of these nonlinear phosphors is mixed with a linear phosphor of a different color and lower conversion efficiency, then a change of color results as the beam current is increased. If a green nonlinear phosphor and a red linear phosphor are employed at low current densities, the primary emission is from the red phosphor and one obtains a low brightness red emission. As the current density is increased, the green phosphor begins to emit more strongly, and resultant colors in the orange, orange-yellow, yellow, and green-yellow are obtained, as shown in figure 7. A chromaticity diagram showing the locations of the primaries and the color gamut for an experimental CRT is shown in figure 8.

It is obvious that a much more acceptable arrangement is obtained if the red phosphor exhibits nonlinearity with beam current density because of the greater eye sensitivity to green as compared with red. This system is ideally suited to high-density computer data readout, wherein it is necessary to rapidly detect and/or correct low or high values of parameters. The

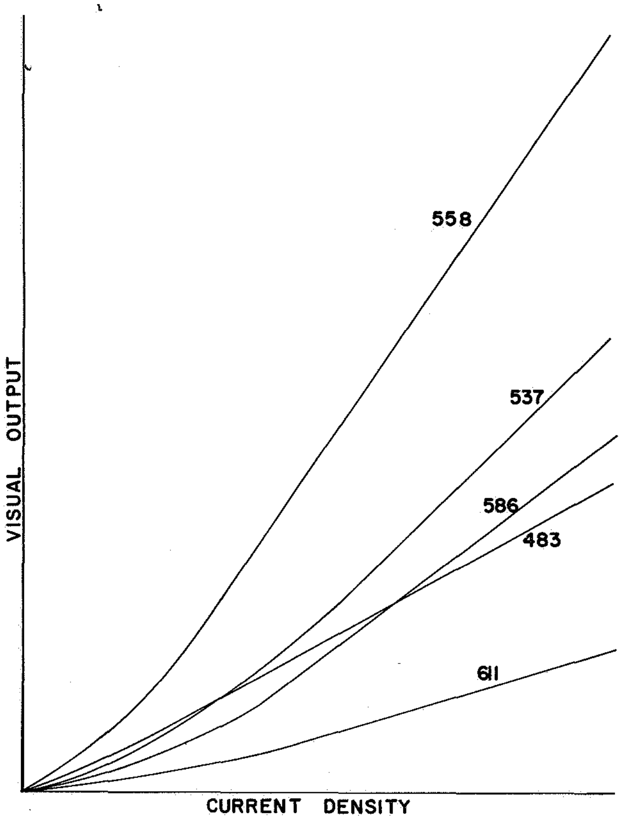

FIGURE 6.—Nonlinear phosphors, visual output versus current density.

number of easily discriminable colors obtained thus far is three; however, with anticipated CRT improvement, this number should increase to at least five.

Another two-color phosphor tube has been recently advertised.[1] Data for this tube are unavailable as of this writing, except that the CIE color coordinates are red $x=0.659$, $y=0.337$; green $x=0.376$, $y=0.573$. It is a penetration type of CRT in which the single-gun accelerating voltage is switched several kilovolts to obtain a color change from red to green.

The patent recently granted to Pritchard (ref. 52) for penetration color screen, color tube, and color television receiver is one of the more comprehensive found in the CRT field.

Several processes are described for phosphor screen fabrication, including details of the Kell process (ref. 53) employing gelatin as an adsorbing layer for phosphor particles. Some of the ideas covered in the Pritchard patent are single, double, triple, and multilayer screens deposited by evaporation or settling successive phosphor layers on a faceplate.

Also described is the process of constructing multilayer phosphor particles in which a central blue-phosphor particle is coated successively with layers of smaller particle green and, lastly, red phosphors, and then treated as a normal phosphor powder and settled on faceplates to form a phosphor screen.

To obtain a color variation, the method requires variation of accelerating voltage from either single gun, or operation of multiple electron guns at fixed voltages, to obtain penetration through, and selective excitation of, the phosphor layers.

The salient factors requisite to optimum color gamut achievement are thoroughly described, including the optimum thickness and efficiency of the phosphor layers, and the thickness of the spacer or inert separating layers. Finally,

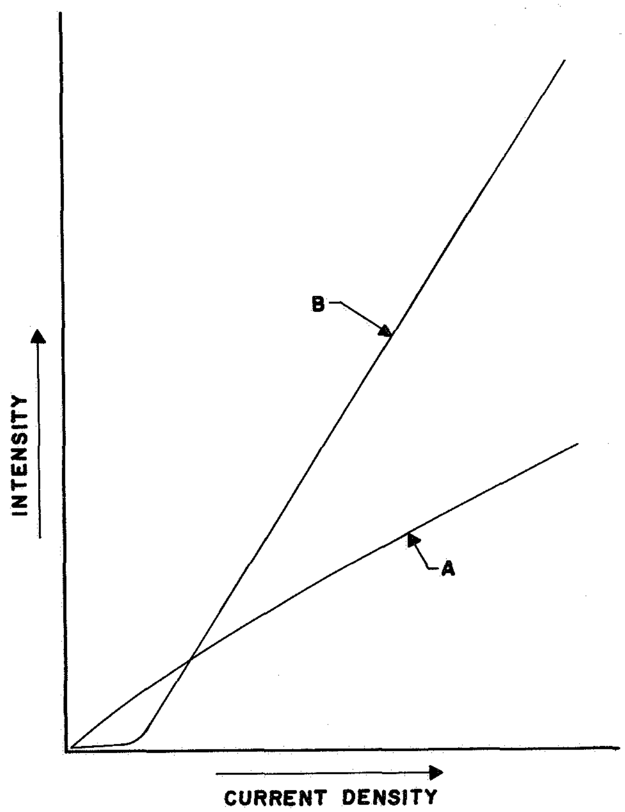

FIGURE 7.—A. Red phosphor linear. B. Green phosphor nonlinear.

[1] Sylvania Division GT&E advertisement appearing in various electronics and display periodicals.

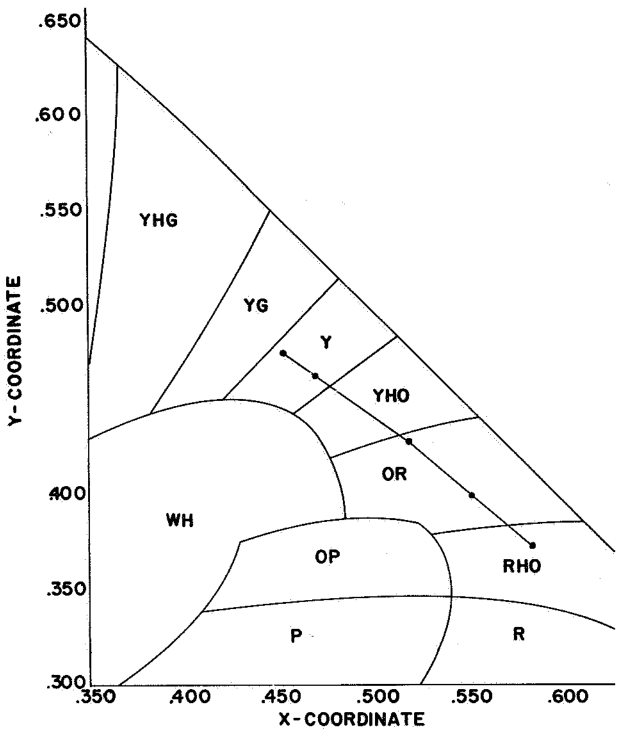

FIGURE 8.—Chromaticity coordinates versus current density. ($\mu a/cm^2$) for linear-nonlinear phosphor combination.

the electronic circuitry required for the various types of gun-phosphor combinations is described.

A radar display tube has been reported (ref. 54) which is referred to as an electron-beam-barrier tube in which a uv phosphor and a red phosphor, excited by an electron beam, are deposited on an electron-beam-opaque substrate which is transparent to uv and visible radiation. On the opposite side of this substrate, the viewing side, is a P–2 phosphor which has a long green phosphorescence. This display yields red moving target information and white stationary targets.

Damon (ref. 55) has recently described a high-resolution, multicolor storage CRT. This tube is similar to standard storage types, with the exception of a special phosphor, which is apparently layered red and green phosphors. At low voltage (6 kilovolts) the red phosphor layer is excited, and at high voltage (12 kilovolts) the green layer is excited. Of course, all colors lying along the line connecting these two points on the chromaticity diagram are attainable by varying the beam voltage from 6 to 12 kilovolts. References to layered and evaporated screens include the patent of Fonda (ref. 56) and Feldman's work (ref. 57) described in his original patent for the formation of multilayer screens by his well-known evaporation technique.

Of the many other methods described in the literature for the fabrication of color CRT's, relatively few have actually succeeded in reaching the consumer market as a finished product. Several transparent CRT display tubes are available.[2] These CRT's have evaporated multilayer phosphor screens of extremely high resolution but relatively low brightness, although the latter is not as important for transparent phosphor layers in view of the much lower ambient light reflection compared with polycrystalline-phosphor aluminized CRT screens.

Only two commercial TV systems have been reported (ref. 58). These sets contain a CRT based on the Lawrence tube, called the colornetron and the chromatron, both of which, manufactured by Japanese firms, contain vertical stripes of red, green, and blue phosphors and are scanned by a single gun.

A survey of various color systems has been published by Solomon (ref. 59), while a state-of-the-art examination was conducted by Burdick et al. (ref. 60). A comparison of the three-gun, shadow-mask CRT and single-gun, beam-index CRT was made by T. Seke et al. (ref. 61).

Only a single reference to the Gabor tube is known (ref. 62). This appears to be one of the more practically applicable CRT's for wall mounting.

A deflection-modulation CRT display has been described recently (ref. 63) in which the striking three-dimensional qualities are discussed. The first report on the method by which these displays are obtained was given by Johnson (ref. 64).

---

[2] Literature on high-resolution, transparent phosphor CRT's may be obtained from the Transylume Tube Division of Panaura Corp.

Developments in Bandwidth Reduction

As existing communication channels became more crowded, people began to look for ways to put them to use more efficiently. This has led to two principal philosophies of bandwidth reduction. The first, and probably the one that has received the most attention (ref. 65), requires an analysis of the information to eliminate redundancies and some type of encoding of the remaining pertinent information to communicate it from point A to point B. At the receiving point B, the transmitted information must be decoded and the redundancies, or some type of quasi-redundancies, created to put the information into a usable form. These systems are generally quite complex and expensive because of the many functions they must perform.

The second philosophy of bandwidth reduction is to format the image information without paying any attention to particular redundancies in a manner that takes advantage of the fact that a human viewer with specific visual capacities and definite subjective attitudes is, in the end, the decisive factor on whether there is enough information in the image for a given situation. An example of the latter would be a dot interlace television system (ref. 66); of the former, run length, or block length coding or delta modulation. A discussion of bandwidth reduction is beyond the scope of this review, and the reader is referred to "References," particularly reference 67.

Developments in Projection Display Techniques

In a report on an automated display chart for program management, Hilt (ref. 68) refers to a technique employed to obtain images on Kalvar film from a P–16 phosphor CRT screen, in which the exposure required from the CRT for Kalvar is reduced by a factor of 10. This is accomplished by a proprietary film-developing technique.

Morgan, Werner, and Libby (ref. 69) have recently reported on the use of CRT exposure of dry silver recording materials. Peak spectral sensitivity can be varied, and high-resolution, high-contrast images are obtained.

The recent development of a number of organic, uv-sensitive materials, referred to as ultraviolet imaging, or uvi materials (ref. 70), with significantly lower power requirements for exposure than materials previously available, has opened a new and exciting area for the application of uv phosphors. In addition, the wavelength sensitivity of these materials is variable, but the most promising and useful materials have a spectral sensitivity matching the newer high-output uv phosphors described in table IV. The potential applications for these uv phosphor-uvi material combinations in projection display (and other areas of application for rapid dry-film processing) are very encouraging indeed.

Conclusions

From the wealth of recent literature on CRT's, it is apparent that new applications are still being discovered and existing applications are being improved. The most promising of these are the optical-diode, high-contrast tube; the two-color CRT wherein color change is obtained by screen current variation; the exciting developments in the uv phosphor and imaging materials field, which should make possible essentially real-time projection display and hardcopy transfer; the improved character recognition techniques; the bandwidth reduction achievements; the multilayer phosphor screen fabrication techniques; the new visible emitting phosphor discoveries; and the significant improvements attainable in electron-gun current densities.

References

1. Johnson, A. C.; and Cowden, D. G.: Information Display, vol. 4, May–June 1967, p. 46.
2. Gleichauf, P. H.; and Niklas, W. F.: Electrotechnology, Apr. 1966, p. 75.
3. Graham, S.: Information Display, vol. 4, May–June 1967, p. 59.
4. DuBois, R.: 8th Natl. Symp. Society for Information Display, 1967, p. 9.
5. McMaster, R. C.; Rhoten, M. L.; and Mitchell, J. P.: Materials Evaluation, Mar. 1967, p. 46.
6. Winfield, R.: 8th Natl. Symp. Society for Information Display, 1967, p. 273.
7. Stotz, R. H.; and Cheek, T. B.: 8th Natl. Symp. Society for Information Display, 1967, p. 91.

8. Aziz, R. A.: 8th Natl. Symp. Society for Information Display, 1967, p. 83.
9. Stupar, T.: Information Display, vol. 4, May–June 1967, p. 67.
10. Haring, D. R.: Proc. FJCC, 1965, p. 847.
11. Corbin, H. S.: Control Engineering, vol. 12, Dec. 1965, p. 77.
12. Geppart, D.; and Dore, B.: Research on Cold Cathodes. 2d Quarterly Report, Stanford Research Laboratories, Contract NAS 5–9581, Nov. 1965.
13. Scheer, J.; and VanLaar, J.: Solid State Comm., vol. 3, 1965, p. 189.
14. Eden, R. C.; Moll, J. L.; and Spicer, W. E.: Phys. Rev. Letters, vol. 18, 1967, p. 597.
15. Cape, J. A.; and Coomes, E. A.: J. Chem. Phys., vol. 32, 1960, p. 210.
16. Rate, E. T.: Design, Development and Fabrication of an Ultra High Resolution CRT. General Electric Co., Syracuse, N.Y., Contract DA28–043 (AMC)00476(CE), Mar. 1966.
17. Tube Source. General Electric Co., Syracuse, N.Y., Contract DA28–043(AMC)00021(E), Dec. 1966.
18. ITT Industrial Laboratories: Strontium Tungstate Cathodes. Contract DAAB07–67–C–0381, Apr. 1967.
19. Afflech, J. H.: Tech. Information, General Electric Series, RG2ETI–19, Aug. 1962.
20. Dominguez, R.; Doolittle, H. D.; and Varadi, P. F.: Electron Devices, ED14. Trans. IEEE, no. 5, May 1967, p. 267.
21. Albert, M. J.; Atta, M. A.; and Gabor, D.: Brit. J. Appl. Phys., vol. 18, 1967, p. 627.
22. Bryden, J. E.: 7th Natl. Symp. Society for Information Display, 1966, p. 83.
23. Buchanan, R. A., et al.: Proc. 6th Rare Earth Res. Conf., May 1967, p. 271.
24. Levine, A.; and Palilla, F. C.: Appl. Phys. Letters, vol. 5, no. 6, 1964, p. 118.
25. Palilla, F. C.: The Trivalent Rare Earth in Inorganic Phosphor Systems. GT&E Tech. Rept., T.R. 67,020–3, June 1967.
26. Bril, A.; and Wanamaker, W. L.: Philips Tech. Rev., vol. 27, 1966, p. 23.
27. Bril, A.; Wanamaker, W. L.; and de Laat, C.: J. Electrochem. Soc., vol. 112, 1965, p. 111.
28. Avella, F. J.: J. Electrochem. Soc., vol. 113, 1966, p. 1225.
29. Avella, F. J.; Sovers, O. J.; and Wiggins, C. S.: J. Electrochem. Soc., vol. 114, 1967, p. 613.
30. Bril, A.; and Klasens, H. A.: Philips Res. Rept., vol. 7, 1952, p. 401.
31. Bril, A.: Absolute Efficiencies of Phosphors With Ultraviolet and Cathode Ray Excitation. *In:* Luminescence of Organic and Inorganic Materials (H. P. Kallman and G. M. Spruch), John Wiley & Sons, Inc., 1961, p. 479.
32. Eberhardt, E. H.: ITT Phosphor Chart and Research Memo 362.
33. Crosby, G.; and McCreary, H.: U.S. Patent 2,821,508.
34. Westinghouse Electric Corp., Aerospace Division, Baltimore, Md.: Ultraviolet Phosphor Development for an Active Optical Imaging System. Final Rept., Contract NObs–90449, June 1964.
35. Sisneros, T. E.; and Davis, J. A.: Ultraviolet Generator Study. Final Rept., ITT Industrial Laboratories, Fort Wayne, Ind., Contract AF30 (602)–3879, July 1966.
36. Sisneros, T. E.: Appl. Optics, vol. 6, 1967, p. 417.
37. Avinor, M.; and Meijer, G.: Phys. Chem. Solids, vol. 12, 1960, p. 211.
38. Shinoya, S.: Luminescence of Lattices of the ZnS Type. *In:* Luminescence of Inorganic Solids (P. Goldberg), Academic Press, 1966, p. 205.
39. JEDEC: Optical Characteristics of Cathode Ray Tube Screens. Publication No. 16A, Jan. 1966.
40. Wurtz, J. E.: Information Display, vol. 4, May–June 1967.
41. Baker, C. A.; and Nicholson, R. M.: Proc. 19th NAECON, 1967, p. 285.
42. Amberger, D.: Final Rept., Hartman Systems Inc., Contract NAS12–97, 1967.
43. Lally, K. P.: 7th Natl. Symp. Society for Information Display, 1966, p. 47.
44. CBS Laboratories: Electronic Design, vol. 15, no. 12, June 7, 1967, p. 14.
45. Hilborn, E. H.; and Stevenson, L. F.: Information Display. (To be published.)
46. Winningstad, C. N.: 8th Natl. Symp. Society for Information Display, 1967, p. 129.
47. Engel, J. M.; and Eldridge, A. J.: 7th Natl. Symp. Society for Information Display, 1966, p. 31.
48. Clauer, C. K.; and Kuehler, J. D.: U.S. Patent 3,109,062, Oct. 29, 1963.
49. Tilton, H. B.: Control Engineering, Feb. 1966, p. 74.
50. Wolvin, J.: 7th Natl. Symp. Society for Information Display, 1967, p. 169.
51. Sisneros, T. E.: Single Gun Color CRT. ITT Industrial Laboratories, Fort Wayne, Ind., Contract NAS12–534, Apr. 1967.
52. Pritchard, D. H.: U.S. Patent 3,204,143, Aug. 31, 1965.
53. Kell, R. D.: Patent Application No. 108,535, May 8, 1961.
54. Sylvania Division GT&E: Research and Development of Cathodoluminescent Phosphor With Excitation Time Dependent Color, Contract DA36–039(AMC)03228(E), 1966.
55. Damon, P. P.: Information Display, vol. 3, Nov.–Dec. 1966, p. 31.
56. Fonda, G. R.: U.S. Patent 2,504,674, Apr. 18, 1950.
57. Feldman, C.: U.S. Patent 3,046,154, July 24, 1962.
58. Sugehara, Y.; Ito, H.; and Horaguchi, A.: Electronics, vol. 38, May 31, 1965, p. 81.
59. Solomon, L.: Electronics World, Jan. 1964, p. 34.

60. Burdick, D. C., et al.: NRL Report 6348 (ASTIA No. AD 623960), Oct. 1965.
61. Seke, T., et al.: 6th Natl. Symp. Society for Information Display, 1965, p. 1.
62. Gabor, D.; Stuart, P. R.; and Kalmon, P. G.: Proc. IEE, vol. 105, 1958, p. 581.
63. Everhart, T. E.: Proc. IEEE, 1966, p. 1480.
64. Johnson, T. W.: Proc. IEE (London), vol. 105B, supp. 12, 1958.
65. Deutsch, S.: 7th Natl. Symp. Society for Information Display, 1966, p. 203.
66. Gouriet, G. G.: Electronic Engineering, Apr. 1962, p. 166.
67. Kaufman, L., et al.: Evaluation of Pseudo Random Dot Scanning T.V. Systems, Contract AF19 (628)–5672, 1966.
68. Hilt, D. J.: 8th Natl. Symp. Society for Information Display, 1967, p. 137.
69. Morgan, D. A.; Werner, T. J.; and Libby, W. H.: 8th Natl. Symp. Society for Information Display, 1967, p. 57.
70. Du Pont, E. I.: Instant Access Ultraviolet Imaging Materials. Data Sheets 1–4 and Brochure (available from Organic Chemicals Dept., UVI Marketing Division).

291–351 O—68——4

Page intentionally left blank

N68-22307

CARRIER INJECTION ELECTROLUMINESCENCE

JACQUES HANLET AND RALPH W. HAAS
Marquardt Corp., Van Nuys, Calif.

During the past decade, interest in carrier injection electroluminescence as a display medium has steadily increased, as evidenced by the scope of research in this area. This research has been directed toward a wide range of materials and a large number of device implementations for display purposes. This interest is attributed to the number of desirable features promised by the carrier injection phenomenon. Operation is possible at a few volts; switching is possible at megacycle rates. A relatively long life is promised with chemical stability.

In a survey of carrier injection electroluminescence, it is desirable to consider the relationship of the materials, the device structure, and the fabrication processes to the display system requirements. Desirable features for a visual display include the following:

(1) A light emission that matches the spectral response of the human eye

(2) A low operating voltage and current that minimize the power necessary for continuous excitation

(3) Materials and a simple device structure that permit an economical method of fabrication

Before these desirable properties can be evaluated, a review of the mechanisms involved in producing electroluminescence by carrier injection should be discussed. These mechanisms are basic semiconductor phenomena and include such aspects as intrinsic conduction, extrinsic conduction, and photon generation.

SPECTRAL RESPONSE

The range of possible materials that look promising for use with carrier injection in a visible display can be considerably narrowed by first considering only those materials having a spectral response that provides a reasonable match to the peak-wavelength sensitivity of the human eye. This relationship is established by the energy gap of the material, which approximately determines the maximum possible energy of the emitted photons. The response of the eye dictates the following requirements:

Visual threshold (at blue)	0.40 μ, or 3.06 eV
Visual maximum (at green)	0.55 μ, or 2.25 eV
Visual threshold (at red)	0.70 μ, or 1.75 eV

Because processing of the material during fabrication can alter the requirements slightly, materials with a band gap between 1.75 and 3.6 to 4.0 eV might be selected for use. A review indicates that the I–V and II–V groups have an average band gap of only 1 eV and thus may be excluded from consideration.

As shown in table I, the I–VII and IV–IV groups provide only a few compounds that can be classified as satisfactory. However, both the III–V and II–VI groups contain many binary compounds of sufficient band gap.

INTRINSIC CONDUCTION

If we wish to examine some of the electron processes that occur in the crystal structure, it is best to consider the various energy relationships by means of a simple model describing the position of the various energy levels. Normally, the electrons in a crystalline material are restricted to defined ranges of energy that can be represented by bands of energy. For an ideal, or perfect, semiconductor without impurities, the structure can best be shown by three bands: the valence band, the forbidden energy band, and the conduction band (fig. 1).

The valence band is that range of energies exhibited by the valence electrons of the element which is least tightly bound to the atomic nuclei. The forbidden band of energies represents the range that electrons normally cannot occupy in an ideal semiconductor. The next higher energy level above the valence band that an electron can occupy is represented by the conduction band. The energy gap E_g of a solid-state material is used to describe the separation between the highest energy level of the valence band E_v and the lowest energy level of the conduction band E_c. If an energy, at least equal to E_g, is introduced into the material, some of the valence electrons will be raised from their normal energy levels in the valence band to the others in the conduction band. Energies less than E_g will not remove electrons from the valence band. The density of electrons is given by equation (1):

$$N=N_C e^{-\frac{E_g}{2kT}} \quad (1)$$

Electrons raised to states in the conduction band are free to respond to the influence of an applied electrical field. Some of the electrons will gain a net momentum in the direction of the field, and this energy represents a current flow. Raising electrons to the conduction band leaves both the valence and conduction bands partially filled; thus, current effectively flows in both bands—the free electrons in the conduction band and holes in the valence band.

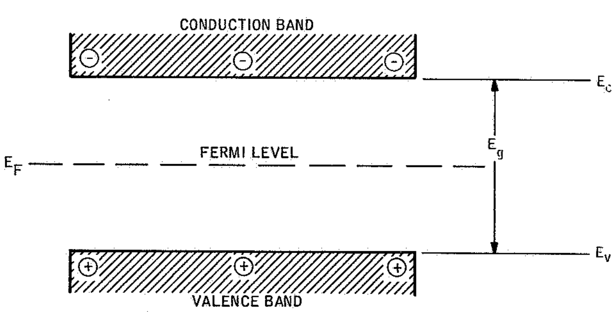

FIGURE 1.—Energy band model.

TABLE I.—*Band-Gap Comparison*

General group	Material	Band gap
I–VII	CuBr	2.9
	Agl	2.8
II–VI	CaS	5.4
	CaSe	5.0
	CaTe	4.3
	MgSe	5.6
	MgTe	4.7
	ZnO	3.2
	ZnS	3.7
	ZnSe	2.6
	ZnTe	2.1
	SrO	5.8
	SrS	4.8
	SrSe	4.6
	SrTe	4.0
	CdS	2.4
	CdSe	1.7
	CdTe	1.5
	BaO	4.2
	BaS	4.0
	BaSe	3.7
	BaTe	3.4
	HgS	2.0
	HgSe	.6
	HgTe	.02
III–V	BN	----------
	AIN	5.0
	AIP	2.5
	AIAs	2.4
	AISb	1.5
	GaN	3.4
	GaP	2.24
	GaAs	1.4
	GaSb	.67
	InP	1.25
	InAs	.33
	InSb	.18
	BP	5.0
IV–IV	SiC	[a] 2.3
		[b] 2.9

[a] Cubic. [b] Hexagonal.

The distribution of electronic states over the band structure is described by the fermi distribution function, which indicates the probability that a state at a given energy level E will be occupied by an electron. In an intrinsic semiconductor, the number of electrons occupying states in the conduction band is equal to the number of holes in the valence band. The fermi level E_F lies approximately midway between the valence and conduction band edges.

Extrinsic Conduction

Nominally, the number of electrons available for conduction in an intrinsic material is an inverse exponential function of the energy band gap. Because of this fundamental relationship, most of the wider band-gap materials are classified as insulators. For example, at room temperature, a semiconductor such as germanium, with a band gap of 0.68 eV, has an electron density of 6×10^{13} cm^3. An insulator such as cadmium sulfide, with a band gap of 2.42 eV, has an electron density of 3 per cm^3. Therefore, it is necessary to incorporate impurity atoms into the crystal structure to increase the conductivity. The properties of the semiconductor will then depend strongly on the impurity content. This incorporation may take a number of forms. A silicon atom, for example, has four valence electrons with which it forms four electron-pair bonds with four silicon neighborhood atoms. If an atom with five valence electrons, such as phosphorus, is substituted in the crystal lattice for a silicon atom, the structure will remain unchanged and partially covalent. The fifth electron will remain in the vicinity of the impurity atom, because of the extra positive charge in the atomic nucleus. A very small amount of energy can overcome this binding and excite the extra electron to the conduction band. Thus, impurity atoms, with more valence electrons than atoms of the host crystal, can donate excess electrons to the conduction band and are called donor impurities. Because the density of conduction electrons is increased in the crystal, such a semiconductor is called an n-type extrinsic semiconductor.

If an atom with three valence electrons, such as gallium, is substituted in the host crystal, one of the covalent bonds cannot be satisfied. This loss can be compensated by the transfer of a valence electron from a neighboring silicon atom. A hole in the valence band can be used to represent the loss of a valence electron. This type of impurity is called an acceptor because it accepts a valence electron from the host crystal. A p-type of extrinsic semiconductor is created by the incorporation of acceptor impurities which form holes in the valence band (fig. 2).

Referring again to the energy band model, it can be shown that the impurities described above will occupy energy levels in the forbidden energy band. The donor impurities will occupy electronic states very close to the conduction band, and the acceptor impurities will occupy states at an energy level close to the valence band. An approximation of the position of these levels can be obtained from a relationship which indicates that the energy is proportional to the effective mass of an electron in a material and inversely proportional to the square of the dielectric constant of the material.

In an n-type semiconductor, the fermi level will be situated close to the conduction band and, for a p-type semiconductor, the fermi level will be situated close to the valence band. The position will be determined by the type and concentration of impurities at a given temperature. If both donors and acceptors have been added, the level is located such that the crystal is neutral; there are equal numbers of negative and positive charges. (See eq. (2).) The density of electrons and ionized acceptors must equal the density of holes and ionized donors.

$$n+N_A^-=p+N_D^+ \tag{2}$$

The powerful effect of impurity concentration on conductivity can be illustrated with a classic example. Consider a germanium crystal containing 10^{16} donors per cm^3 and no acceptors. A solution of the equation for charge neutrality indicates that the fermi level lies 0.52 eV above the valence band at room temperature. From

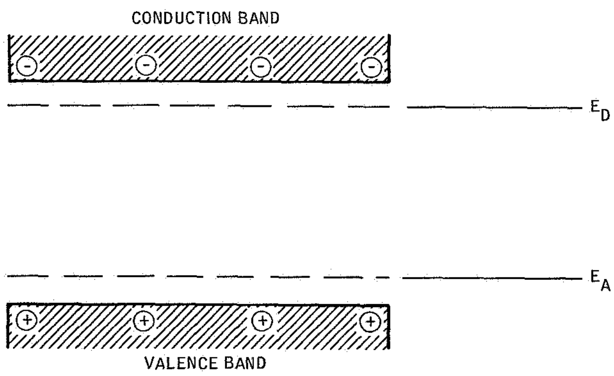

Figure 2.—Impurity levels.

equations (3) and (4), the density of electrons is 10^{16} per cm^3 and the density of holes is 10^{10} per cm^3. This differential was accomplished with an impurity incorporation of approximately 1 ppm.

$$n=N_C e^{-\frac{(E_C-E_F)}{kT}} \tag{3}$$

$$n=N_V e^{-\frac{(E_F-E_V)}{kT}} \tag{4}$$

Photon Generation

The mechanism of carrier injection electroluminescence is based on the production of light by an excited semiconductor returning to a state of equilibrium concentration of holes and electrons by means of recombination transitions. These transitions involve the recombinations of electrons and holes accompanied by the emission of photons. Other recombination processes which are nonradiative must be suppressed or minimized. A transition that takes place between a minimum in the conduction band and a maximum in the valence band is called a direct transition if it does not involve a change in direction of motion of the electron and hole. For this condition, both the electron and hole must have the same momentum, the transition can occur directly with the emission of a photon, and both energy and momentum are conserved. (See fig. 3(*a*).) However, if the above minimum and maximum do not occur for the same direction of motion, an indirect transition will be the result of the recombination. (See fig. 3(*b*).) The lowest energy conduction electron will have a momentum that is different from the momentum of the hole in the valence band, and a phonon is emitted to conserve momentum, thus producing heat instead of light. In an indirect transition, a recombination can occur with the electron and hole having the same momentum resulting in a radiative recombination, but this case is statistically less likely to happen. With direct transitions, the rate of recombination is both fast and competitive with other recombination processes, resulting in a high internal quantum efficiency. Indirect transitions have much slower recombination rates with an attendant lower internal quantum efficiency.

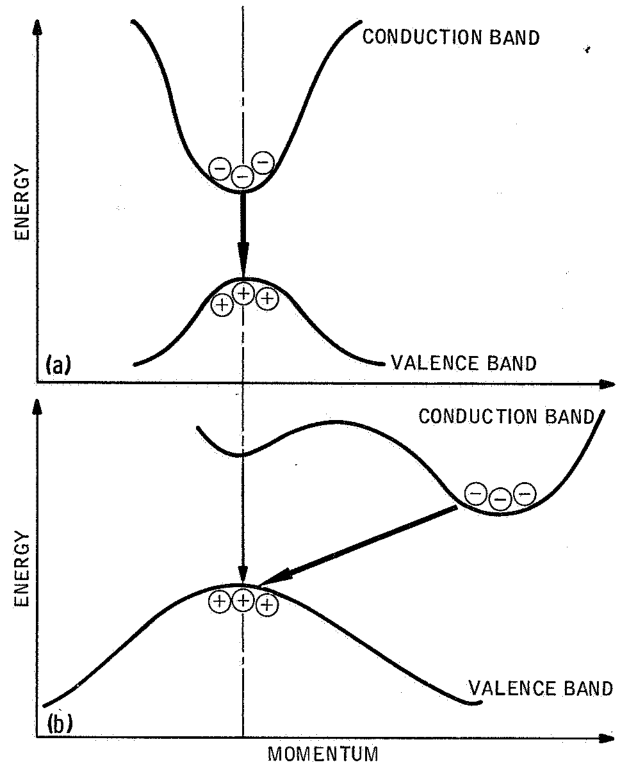

FIGURE 3.—(*a*) Direct transition. (*b*) Indirect transition.

In materials that exhibit indirect transitions which are likely to produce nonradiative recombinations, a higher current density is necessary to increase the probability of occurrence of radiative recombinations and the production of useful light. The two factors, low internal quantum efficiency and high operating current, make these materials poor candidates for use in visual displays.

Table II lists the materials that meet the spectral response requirements discussed above and indicates the type of transition. Aluminum phosphide has been excluded because it is chemically unstable. Copper bromide and silver iodide are both purely ionic crystals, so they have been deleted.

It can be seen that, in general, the III–V compounds exhibit an indirect type of transition, while the II–VI compounds provide a direct transition, thereby permitting a higher internal efficiency. Of the II–VI compounds, ZnSe, ZnTe, and CdS most closely match the peak of the visual region (2.25 eV). In addition, these materials have a relatively high value of mo-

TABLE II.—*Transition Comparison*

Group	Material	Band gap	Type transition	Mobility	
				Electron	Hole
II–VI	ZnO	3.2	Direct	180	
	ZnS	3.7		110 to 140	
	ZnSe	2.6	Direct	100 to 530	
	ZnTe	2.1	Direct		425 to 1000
	CdS	2.4	Direct	200	15
	BaSe	3.7	Direct		
	BaTe	3.4	Direct		
	HgS	2.0	Direct		
III–V	AlAs	2.4	Indirect	1200	200
	GaN	3.4			
	GaP	2.24	Indirect	80	100

bility. Because the requirements for high conductivity in a material are related to carrier mobility and depth of donors and acceptors, this becomes a very desirable property.

Implementation of Carrier Injection

The selection of a device structure to provide the necessary excitation is just as important as the selection of a material based on spectral response, efficiency, and other factors. At least three basic methods will be discussed: the injection of electrons into a *p*-region, where a large concentration of holes exists, the injection of holes into an *n*-region where a large concentration of electrons exists, or the injection of both electrons and holes into an insulating region. Normally, a *p-n* junction is used to accomplish the first two methods, and a form of *p-i-n* structure permits the implementation of the third method.

If an *n*-type material is placed in contact with a *p*-type material, a junction is formed. Initially, electrons will tend to diffuse from the *n*-type into the *p*-region and holes will diffuse from the *p*-type material into the *n*-region. The removal of electrons from the donor sites in the *n*-type material provides a collection of positively charged donor ions. Similarly, the flow of holes from the acceptor sites leaves a number of negatively charged acceptor ions in the *p*-region. This condition will cause an electrostatic layer to be formed at the junction with an electric field whose polarity will oppose the current flow. The magnitude of this potential will be determined by the position of the fermi level in the *n*- and *p*-type materials. As described above, the position of the level in each region is determined by the temperature and by the concentration of the impurity incorporation by doping (fig. 4).

If an external potential is applied to the junction with a polarity such that the height of the barrier is reduced, the junction is said to be

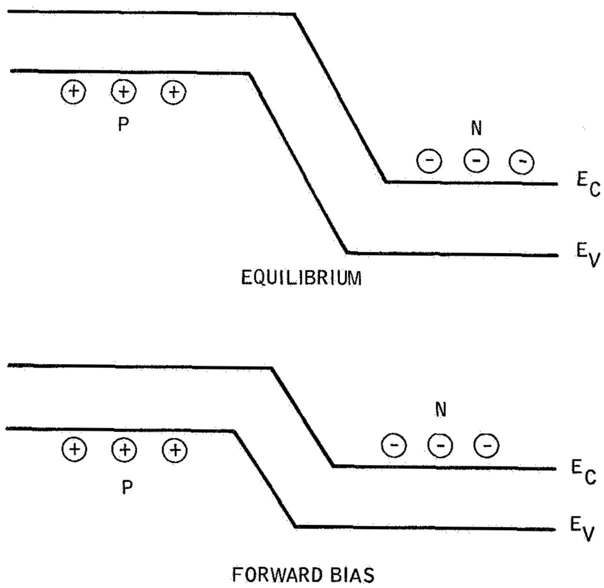

FIGURE 4.—*pn* junction.

forward biased. A large hole current will flow from the p-region to the n-region, and a large and equal electron current will flow from the n-region to the p-region. If, however, the n-region is very heavily doped and the p-region is lightly doped, a condition of minority carrier injection will occur. The junction current will be carried mainly by electrons, and a large excess of electrons will be injected into the p-region. These carriers will diffuse approximately one diffusion length away from the junction and recombine. For opposite conditions of doping, hole injection into the n-region is possible. The selection of the type of minority carrier injection is usually based on the material, since, in many compounds, one carrier will exhibit a much higher mobility.

Another implementation for carrier injection can be provided by placing a metal in contact with an n-type material. It is then possible to inject holes under a condition of forward bias similar to the p-n junction mechanism. The efficiency is low because the injection current represents only a part of the total current.

A tunneling injection technique is possible by using an insulating film in contact with a semiconductor material. This approach requires a high-voltage drop across the insulator.

Another form of junction has been used for injection with those materials that tend to form unipolar extrinsic conduction states. Some materials are strongly n-type and others are strongly p-type and resist changing from this type of conductivity, even with heavy impurity doping. For example, ZnTe has a tendency to remain p-type, while ZnSe has a tendency to remain n-type. If the two materials are used to form junctions, called heterojunctions, these junctions will behave similarly to a simple p-n junction.

The recombination of an electron and a hole can never produce more than one photon; thus, the internal quantum efficiency discussed above can never exceed a value of 1. A more important criterion for a carrier injection device is the external quantum efficiency, which relates the number of photons produced at the surface of the device and the current required to produce these photons during operation. This measure of efficiency accounts for losses in the device due to absorption and transmission, in addition to internal quantum efficiency. The ratio of photons emitted to photons created will certainly depend on the critical angle beyond which the light is internally reflected. The light emerging from the junction is limited to the photons emitted in a direction within a small cone having an aperture which is defined as a function of the index of refraction. Within this angle, light will still be reflected at the interface and reduced by this factor. For example, in a material with an index of refraction of 3, 25 percent of the light will be reflected. In a material with an index of 6.7, 55 percent of the light will be reflected. In addition to this loss, photons may be destroyed by free carrier absorption. It can be shown that the absorption coefficient is directly proportional to the carrier density and the square of the wavelength. These device parameters place an additional burden on the material by requiring a low index of refraction.

Certain problems dominate for each method of injection discussed above, whether the device implementation is a simple p-n junction, a heterojunction, or a hybrid configuration, such as a p-i-n structure.

If the semiconductor material is amphoteric, or can be made n-type or p-type, abrupt p-n junctions can be made from now-familiar technologies. With the direct transition materials, a high internal efficiency can be expected. The major problem pertaining to this structure is the photon extraction from inside the junction in order to achieve an equally high external efficiency.

The problem with the heterojunction approach is that of finding another semiconductor of opposite conductivity type with equal band gap, having an identical lattice constant, similar coefficient of expansion, and similar growth habits. If this condition is met, the internal efficiency could be comparable to that of the p-n junction. The fact that the electrodes must be made purely ohmic for both electron and hole injection poses another requirement.

With the p-i-n junction structure, the problem is that of material compatibility. A high degree of physical and chemical compatibility must be exhibited by the five material layers that form the device; the thin layer of intrinsic material in which the recombination occurs, the wide band-gap material of low work function and the wide band-gap material of high work function that are deposited on either side of the intrinsic layer, and the outer injecting electrodes. (One must be transparent.) The lattice and band-gap requirements are somewhat relaxed with this approach as compared with those mentioned in the previous structures. However, the requirements for similar coefficients of expansion and a high conductivity in the transparent conductor are quite stringent.

In all of the structures mentioned above, a high current requires high mobilities, a low forward resistance, and a junction width that is shorter than the diffusion length of the carriers. The diffusion is controlled by the carrier lifetime which depends on the various capture and recombination processes.

Material Relationships

Because no simple element can satisfy the desired band-gap requirements, with the exception of diamond, the choice of a host crystal material is limited to compounds. The most likely candidates seem to be from the II–VI group of elements. These compounds, by Goldschmidt's rule, which determines lattice organization as a function of atomic radii, crystallize in the blende or wurtzite structures. It can be inferred from Welker's rule, which defines the degree of ionicity in the atomic band, that these partially ionic, high-bond-strength materials should provide high mobilities with small scattering of the mobile carriers.

However, as they chemically form, all II–VI compounds have a tendency to form vacancies, or point defects, resulting from missing atoms in the lattice structure, and have a tendency to form various spatial arrangements of the atoms forming the elementary lattice cells. The vacancies are formed because of the polymeric form assumed by the chalcogene (an element that is listed in col. VI of the periodic table), and result in low conductivity. The mixtures of crystalline structure occur, in spite of the fact that the length of the atomic bonds remains unchanged, and the differences in structure appear only when other than nearest-neighbor atoms are considered. The differences result in complex scattering modes between optical and acoustical vibrations in the lattice. A combination of these vibration modes can produce an undesirable shift in band gap and a lower mobility because of an increase in the effective mass of the carrier.

Defects, vacancies, and structures are all strongly dependent on the processing temperature, because, as they are heated, the compounds dissociate into elements of dissimilar vapor pressures. The effects of the temperature, however, vary in opposite directions with respect to vacancies and structures. For example, at moderate temperature of preparation, the chalcogenes do not appear in a monoatomic form, and hardly at all even as diatomic species, but mainly as molecules of four, six, or eight atoms. On the other hand, the crystalline orientation preferentially assumes only the blende structure, which presents a single optical mode of vibration from its spatial arrangement. However, at high temperature, the contribution from diatomic molecules increases, though not to the extent of excluding puckered rings of six or eight chalcogene atoms. The structure then assumes a hexagonal packing and, as a consequence, it has an acoustical mode or low frequency of vibration ensuing from its spatial arrangement of atoms.

The problem of polymerism is about the same for any chalcogene; however, the structural problem is mostly one of degree; that is, the extremes of the vertical series of the periodic table assume only one crystalline form, varying from hexagonal zincite for the oxygen compounds to cubic sphalerite for the telluride compounds.

Besides the polymerism, responsible for the vacancy formation, the atomic sizes of the constituent elements define the levels where native donors and acceptors will be found in the gap, as well as the type of conductivity these native defects will produce. Hence, the increasing size of the Group II elements will promote negative conductivity, whereas elements from Group VI will give opposite results. In other words, going down column II, from beryllium to mercury, with the atomic diameter increasing from 2.25 Å to 3.10 Å, will secure an n-type conductivity; conversely, going down column VI from oxygen to tellurium, with the atomic diameter increasing from 1.32 Å to 2.9 Å, will secure a p-type conductivity.

With respect to the location of the various levels in the band donor or acceptor of native origin, it can be said that the biggest of the two atoms in a band will always be close to the band to which it belongs when a vacancy of the other atom occurs. Hence, because the metallic element is associated with the conduction band and the nonmetallic one with the valence when, for example, in CdS a sulfur vacancy occurs, the excess cadmium atom will be close to the conduction band, where it may easily ionize and contribute to the conduction. In ZnTe, if a tellurium vacancy is produced, the excess zinc atom of smaller radius will be far below the conduction band and contribute nothing to the conductivity.

Conversely, if cadmium or zinc vacancies occur, the sulfur atom in excess will be high above the valence band, so high as to be unable to contribute to hole conduction, whereas the much bigger tellurium atom would be close enough to the valence band to be easily ionized and give hole conductivity. This follows from the cation/anion radii ratio which takes a value greater than unity for CdS, and smaller than unity for ZnTe; this ratio, therefore, gives an indication on the compensation tendency in these II–VI compounds. This compensation depends on the energy gained in a transition of electron from donor to acceptor. This gain in energy can be considered high only when both are shallow, in which case the energy gained by the transition may exceed the energy needed to form the vacancy.

Compensation between vacancies means electrical neutralization and should be prevented by proper doping; this is possible when the cation and anion have equally deep ionization levels and an adequate solubility; that is, when they can be introduced into the host crystal at the desired concentration.

Numerous studies have shown that incorporation of Group III elements in substitution for Group VI provides n-type conductivity, and Group V elements in substitution for Group VI provide p-type conductivity in II–VI compounds, although sufficient concentration has not been achieved equally well for p- and n-dopants. Thus, the carrier concentration achieved by impurity incorporation is still somewhat short of its goals, and there is limited hope in that direction.

Considerable work in industry has shown that carrier injection can be excited in many materials with a wide range of device structures. The only decent efficiencies have been achieved with the III–V compounds such as GaAs, which can emit in the visible red. The following efforts, however, are representative of efficiencies achieved in the visible region:

(1) External quantum efficiencies of 2 percent were obtained with injection through an insulating film into P-doped ZnTe at 5340 Å with the temperature at 77° K.

(2) With a metal-insulator-p-type tunnel diode of ZnTe, an efficiency of 0.01 percent was observed at 5375 Å and 77° K.

(3) With an avalanche structure, an external efficiency of 2 percent at 5380 Å and 77° K was observed with P-doped ZnTe.

(4) An 18-percent external quantum efficiency was obtained from p-n junctions at 70° K with a 6300-Å emission in $ZnSe_xTe_{1-x}$.

In summary, on the basis of spectral response and internal efficiency, the II–VI compounds appear most desirable for use as carrier injection host crystal materials. ZnSe, ZnTe, and CdS most closely match the peak of visual sensitivity. From a structure standpoint, the most encouraging results to date have been achieved with the heterojunction structure. It is believed that, with a more efficient technique in preparing the materials and in producing junctions, a combination of II–VI compounds can be considered to be the most promising to give efficient electroluminescence at room temperature.

BIBLIOGRAPHY

AVEN, M.: Efficient Visible Injection Electroluminescence From p-n Junctions in $ZnSe_xTe_{1-x}$. Appl. Phys. Letters, vol. 7, 1965, p. 146.

AZAROFF, L.; AND BROPHY, J. J.: Electronic Processes in Materials. McGraw-Hill Book Co., Inc., 1963.

CROWDER, B. L.; MOREHEAD, F. F.; AND WAGNER, P. R.: Efficient Injection Electroluminescence in ZnTe by Avalanche Breakdown. Appl. Phys. Letters, vol. 8, 1966, p. 148.

EASTMAN, P. C.; HAERING, R. R.; AND BARNES, P. A.: Injection Electroluminescence in Metal-Semiconductor Diodes. Solid State Electronics, vol. 7, 1964, p. 879.

KITTEL, C.: Introduction to Solid State Physics. John Wiley & Sons, Inc., 1956.

MIKSIC, M. G., ET AL.: Injection Electroluminescence in p-type ZnTe. Phys. Letters, vol. II, 1964, p. 202.

Page intentionally left blank

SESSION II

CHAIRMAN, *Hans Bullinger*

N68-22308

THERMOCHROMIC DISPLAYS[1]

D. GRAFSTEIN, R. P. BURKOWSKI, M. KORNBLAU, AND W. L. FLINT

Aerospace Research Center, General Precision Systems Inc., Little Falls, N.J.

The feasibility of using thermochromic materials in display devices is under investigation with particular emphasis on their use in cathode-ray tubes (CRT's). Reversible color changes have been observed when Ag_2HgI_4 or Cu_2HgI_4 target films were irradiated with electron beams. Likewise, simultaneous thermochromism and cathodoluminescence have been demonstrated for mixtures of Ag_2HgI_4 and conventional CRT phosphors. Some limitations of thermochromics for device use are discussed. Studies directed at improving the existing ternary mercuric iodides, as well as other potential thermochromic materials, are also described.

An increasing need exists to view information presented by display devices under conditions of high ambient illumination, such as often prevail in both aircraft and space vehicles and in central command rooms. In aerospace vehicles particularly, visual displays may be subject to variations in illumination levels encompassing a range of several orders of magnitude. The more traditional displays that incorporate mechanical movements, such as the dials of meters, are usable under these extremes of ambient, but they suffer from the disadvantages of moving parts and catastrophic failure modes.

More modern nonmechanical systems, such as electroluminescent (EL) displays or CRT's, saturate at brightness levels far too low to allow their use in high ambients. In these devices, a phosphor material emits light (luminescence) as a result of the application of an electric field, in the case of electroluminescent materials, and bombardment by an electron beam, in the case of cathodoluminescent materials. One of their main limitations is the relatively low brightness levels and the concomitant low contrasts that have been achieved. Attempts to increase the brightness of EL displays by increasing the applied electric field and its frequency result in a severe lowering of the lifetime of the material.

Scotophors are materials whose change in opacity under electron beam bombardment is viewed by reflected light. The alkali halides, such as potassium chloride, are darkened under excitation by high-energy particles such as X-rays or beta rays. This darkening is the result of the formation of color centers (trapped electrons) in the material. Alkali halides thus provide better contrast when viewed under high ambient illumination. The image will remain on the screen until the trapped electrons are freed by heating or other means. Several devices utilizing an alkali halide are commercially available; among these are the recorditron tube and dark trace tubes. While these produce images useful under high ambient illumination levels, erasure to allow rapid updating of information is extremely difficult.

[1] This work was sponsored by the National Aeronautics and Space Administration, Electronics Research Center, under Contract NAS 12-89.

Photochromic materials are also viewed by reflected light. These change color by absorption of light of a particular wavelength, usually in the ultraviolet range. The color can be erased by heating or by illumination of the colored form of the material by light of longer wavelength than was used to accomplish the initial color change.

The concept developed in our laboratories is to display information using thermochromic materials that will change color upon activation by heat and revert to their original color when cooled. The observation of the color change in these materials is due to a differential reflection of the ambient light by the two different-colored forms.

General Considerations

Thermochromics are materials that undergo a color change when they are heated above a certain temperature, called the transition temperature, and revert back to their original color when cooled below the transition temperature. This phenomenon has been known for some time, but the application of thermochromic materials to practical display devices is new (refs. 1 to 6).

The response to the temperature change may be slow or fast, sharp or gradual, depending on the mechanism that causes the color change. Thermochromism has been observed in a variety of organic and inorganic materials, and the mechanism has been found to vary with the molecular structure of the material. The type of mechanism invoked to explain the thermochromic activity has included equilibrium between two molecular species, broadening of a near-ultraviolet absorption curve, ring opening, thermal achievement of a triplet state configuration, formation of free radicals and order-disorder phenomena. This list by no means exhausts the mechanisms by which thermochromism can occur. Many of these mechanisms, however, involve the breaking of chemical bonds and/or the migration of chemical species. In such systems, the rate of the color change with temperature of the material is relatively slow.

For information display applications, the color changes must be fast and result in a sharp contrast of one color on another. Because both maintenance and power requirements are important factors in device considerations, especially in aircraft and spacecraft, the materials also should show good thermal stability and require low power when they are incorporated into the device. The most interesting thermochromic materials, which satisfy the above conditions, are those that involve a rearrangement of cations in a close-packed anion crystal lattice, where no breaking of chemical bonds occurs. Theoretically, there appears to be no upper limit to the rate at which such transitions, often called order-disorder reactions, can occur. Thus, these transitions are intrinsically rapid and take place at a definite temperature.

Before actually discussing devices for which thermochromics are being considered, some of the properties that make them attractive for such devices will be described. In principle, it is feasible to design repetitively driven thermochromic displays which are flicker free, even with low repetition rates. A standard phosphor, even one having so-called long persistence, reaches peak brightness almost immediately after being pulsed, and then gradually decreases in brightness. Thus, fairly rapid repetition rates are necessary if the flicker is to be avoided. For a repetitively driven display using thermochromics, however, it is only necessary to repeat the signal before the material has cooled below the transition temperature to obtain a completely flicker-free display, because of the fact that thermochromic materials have only two distinct states. This behavior becomes important, particularly when the displayed information is being derived from a computer, because it permits the use of a slower computer and reduces the size of the computer memory that is required.

The rise time and persistence of the thermochromic display is a function of a number of material and experimental parameters. These include the heat and temperature of transition, heat capacities, and thermal conductivities of both the thermochromics and the various substrates employed. Persistence can be controlled by a suitable selection and trimming of these parameters.

The mass and thermal conductivities of the substrates are particularly easy parameters to adjust for the generation of a set of desired rise times and persistence values. As will be discussed later, a wide range of transition temperatures is also available, but, of course, the selection of any particular thermochromic fixes the transition temperature and the heat of transition. Persistence values and rise times are also a function of the ambient thermal level as well as the rate and magnitude of the energy input.

Because thermochromic materials do not emit light, but differentially reflect incident ambients, the operation of thermochromic devices is not limited by the usual energy transfer relationships, where emission occurs at longer wavelengths than the energy which produced it, as is the case with phosphors. Thus, with thermochromics, a visible color change can be produced by infrared stimulation.

Another property, namely the large difference in electrical conductivity between the high- and low-temperature forms, provides an added storage mechanism and control element which may be useful when, for example, thermochromics are combined with electroluminescent materials. It should be mentioned, however, that this property is specific to those thermochromic materials that result from an order-disorder mechanism. Our research efforts, thus far, have emphasized this type of thermochromic material.

Since the thermochromic change is accomplished by heating, a variety of display devices have suggested themselves depending on the mechanism used to heat the material. These mechanisms include resistance heating, heating with a laser and electron beam, or high-energy particle bombardment. Numeric prototype display devices that utilize the well-known thermochromics Ag_2HgI_4 and Cu_2HgI_4 and depend on resistance heating for their operation have been constructed. Similarly, in other experiments, the feasibility of writing with a laser beam on these same materials was also demonstrated. A detailed description of the prototype display device and laser "writing" on thermochromics was given at the October 1966 meeting of the Society for Information Display (ref. 7).

THERMOCHROMICS IN A CATHODE-RAY TUBE

Recently, we have been investigating the use of thermochromics in a CRT, both as a substitute for the phosphor and in conjunction with the phosphor. Earlier calculations indicated that the use of thermochromics in a CRT was feasible.

For example, with Ag_2HgI_4 as the target material, an energy of approximately 7.32 joules per gram is required to heat the material to 60° C from room temperature, assuming an average specific heat of 0.21 joule per degree-gram for this temperature range. With a surface density of the target material of 10 mg/cm^2 and an electron beam diameter of 3×10^{-3} centimeters, at least 0.52×10^{-6} joules would be required to heat this area to 60° C. This would be the actual energy required, if there was 100 percent efficient transfer of the kinetic energy from the electron beam to the thermochromic. The dwell time necessary to accomplish this change then would depend on the values of the accelerating voltage and electron beam current. For an accelerating voltage of 2000 volts and a beam current of 0.1 milliampere, a dwell time of approximately 3 microseconds would be required. This corresponds to a sweep rate of 10^3 cm/sec. These beam energies and sweep rates are available in conventional CRT's.

The wide range of beam energies, energy densities, and sweep rates available in CRT's makes it possible to achieve a wide range of temperature variations in the target material. The objective of the present work has been to demonstrate experimentally and evaluate the use of thermochromics in a CRT. The results of these experiments are now described.

In one series of experiments, the electron beam in a Hitachi electron microscope, Type HU-11, was utilized. The thermochromic is put onto a glass surface by spraying a suspension of the thermochromic in xylene using an artist's airbrush. After the solvent has evaporated, the sample is tested.

291-351 O—68——5

Several sets of experiments were conducted using the electron microscope. In the first set, samples of Ag_2HgI_4 and Cu_2HgI_4 painted onto a ⅛-inch-thick conducting surface of the plate were electrically connected to the chassis of the instrument. The target chamber was evacuated. Each of these samples was then bombarded with a focused beam of 75-kilovolt electrons at an indicated beam current of 30 microamperes. Both materials underwent their characteristic reversible thermochromic color change: Ag_2HgI_4 went from yellow to orange, and Cu_2HgI_4 went from red to black. Upon moving the beam across the surface of the plate, the transition of the previously irradiated spot to the low temperature form was observed to be very rapid.

A second set of experiments was conducted in much the same way as described above, except that a very thin (0.0033 to 0.0051 inch thick) conducting glass plate was used, and thermal contact with the microscope was minimized by supporting the plate on alumina thimbles. Under these conditions, the thermochromic transition was observed to be more rapidly induced, and the persistence was increased so that it was possible to write on the target with a moving beam. The transition to the low-temperature phase was slower than in the first experiment.

Several sets of experiments were conducted on 50 percent mixtures of each of the two thermochromics with each of five standard CRT phosphors (P–1, P–2, P–7, P–14, and Sylvania Red 1120). Each of the mixtures was shown to be thermochromic, with an observed transition temperature within at most 2° of that of the parent thermochromic component. Generally, as was expected, the contrast change was diminished by the presence of the phosphor, which acts as a diluent for the color. However, in one case, that of Ag_2HgI_4 and P–14 (a lemon yellow phosphor), the contrast was undiluted and very good. This is because the ground color of the phosphor is similar to the color of the low-temperature phase of the thermochromic.

It was possible to induce the thermochromic color change in the Ag_2HgI_4:P–14 mixture, but not in the Cu_2HgI_4:P–14 mixture. Both mixtures showed cathodoluminescence but at a lower output intensity than the pure P–14 phosphor. In addition, the color of the emitted light was different from that of the yellow-emitting P–14. The Ag_2HgI_4:P–14 mixture emitted an orange light and the Cu_2HgI_4:P–14 mixture emitted a reddish light. This change in output is probably due to some filtering of the phosphor light output by the colored thermochromics.

Experiments on mixtures of the green-emitting phosphor P–2 and Ag_2HgI_4 were performed in the electron microscope. Mixtures containing 50, 25, and 10 percent by weight were milled in a mortar and pestle, dispersed in methyl alcohol as a slurry, and painted on thick conducting glass plates. Experiments were performed at both 50 and 75 kilovolts with currents nominally at 30 microamperes. All of the mixtures fluoresced with a blue-green light and the fluorescence could be observed in both light and dark ambients. Comparison with the pure phosphor emission indicated a slight shift toward the blue and decreased intensity of emission with decreased P–2 content. The change in hue is a result of filtering by the yellow-colored thermochromic. Reversible thermochromic color changes were observed only in the 10-percent P–2 mixture when viewed in a normal room level. The viewer is conscious of seeing both the thermochromic and phosphor transition together, but the thermochromic transition has less visual impact. With the other mixtures, either burning of the composition or absence of thermochromic transition was noted, depending on the current level.

Experiments on mixtures of the red-emitting Sylvania phosphor and Ag_2HgI_4 were also performed in the electron microscope. Mixtures containing 50 and 25 percent by weight of the phosphor were prepared and coated as in the preceding experiment. The red fluorescence was quite weak in both the mixtures and in the pure phosphor at 50 kilovolts and 20 microamperes. At 75 kilovolts and 30 microamperes the fluorescence was much stronger and no color shift was noticed as compared with the pure phosphor. Because of the red body color of this phosphor, it was extremely difficult to determine whether a thermochromic transition occurred. In one case (that of the 25-percent phosphor mixture in a 75-kilovolt beam), some faint reversible color shifts were observed. The use of a strong red-colored phosphor is not recommended with Ag_2HgI_4. It might be more useful with Cu_2HgI_4.

Figures 1 and 2 show a schematic diagram and a photograph, respectively, of the demountable CRT utilized in subsequent experiments. This is a 5FP–A tube, which has both magnetic focus and deflection and is suitable for radar applications. The maximum operating conditions, as specified by the manufacturer, are as follows:

Heater voltage	6.3 V.
Heater current at 6.3 V	0.6 ± 10% A.
Accelerator voltage	8800 V dc.
Accelerator input	6 maximum W.
Grid 2 voltage	770 maximum V dc.
Grid 1 voltage:	
Negative bias value	180 maximum V dc.
Positive bias value	0 maximum V dc.
Positive peak value	2 maximum V.

The demountable CRT was attached to a Veeco, Type C quick connect coupling through a flange in an MRC vacuum collar, which is located on the baseplate of a Veeco vacuum system. A bell jar was then placed on top of the collar. A vacuum gage was present in the bell jar of the vacuum system. Because of the presence of relatively long, small-diameter tubing between the bell jar and the CRT, the indicated pressure of about 2×10^{-6} torr was probably about an order of magnitude higher in the CRT.

The target materials were applied to various glass substrates rather than directly on the faceplate of the CRT. The glass substrates were in turn supported on a wire gauze which was located below the faceplate inside the tube. It was possible to irradiate the thermochromics directly, or indirectly, through the glass by inverting the target. Thin coatings of Cu_2HgI_4 or Ag_2HgI_4 were applied to the glass substrates in a manner similar to that described before for the experiments with the electron microscope. A small area of the slide was painted with P–14 phosphor, making it possible to observe the beam for focusing purposes.

After the cathode was properly activated, samples of Ag_2HgI_4 and Cu_2HgI_4, sprayed on thin (3 to 5 mils) conducting glass, were placed in the demountable tube. A well-focused beam was moved across the surface at rates of up to

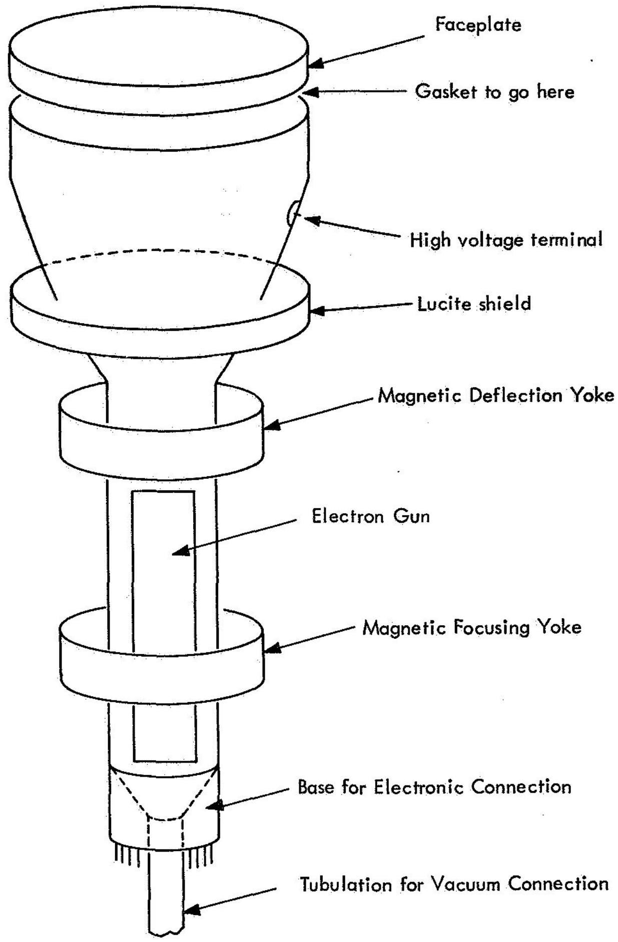

FIGURE 1.—Schematic of demountable CRT.

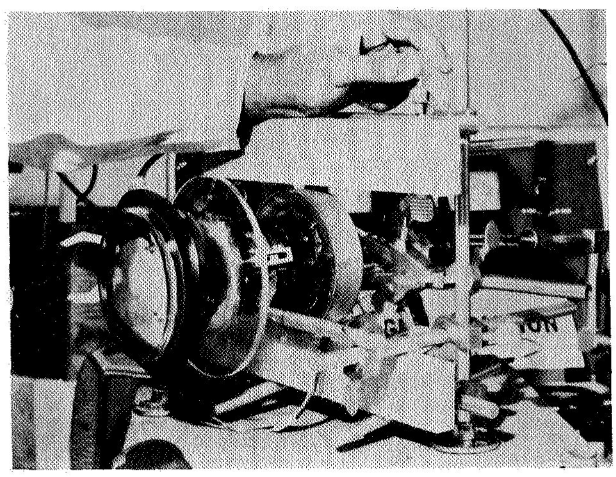

FIGURE 2.—Experimental demountable CRT.

a few centimeters per second. At voltages between about 3 and 10 kilovolts, it was possible to induce some irreversible color changes in both samples. Very faint reversible changes were also present, but the contrast in these cases was very poor, indicating only partial conversion to the high-temperature phase. There did not seem to be any setting of the tube using voltages between 0 and 10 kilovolts and currents between 0 and 70 microamperes that caused a good reversible transition. Currents were measured with a microammeter. Since beam area is not known for these initial experiments, the beam current density is not known. However, with the glass surface facing the electron beam, reversible thermochromic transitions in Ag_2HgI_4 were observed above about 5 kilovolts and at 30 microamperes. (See figs. 2, 3, and 4.) Under these conditions, it was also possible to decompose the specimen by holding the beam stationary. Similar results were observed with Cu_2HgI_4. The transition was reversible, but it was important not to allow the material to overheat. If overexposure occurred, there was partial decomposition.

Finally, a sample of Ag_2HgI_4 on thin conducting glass was placed in the demountable CRT with the thermochromic facing the beam. Working at beam currents of about 1 microampere and systematically varying the voltage, only irreversible color changes were observed as the voltage was increased from 10 kilovolts. At about 30 kilovolts the appearance of the re-

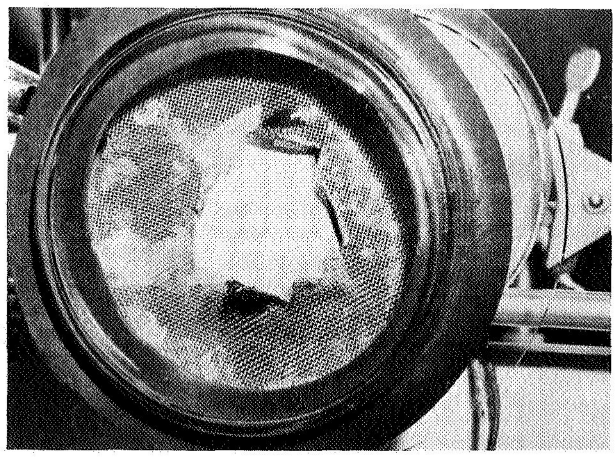

FIGURE 3.—Writing on thermochromic in demountable CRT.

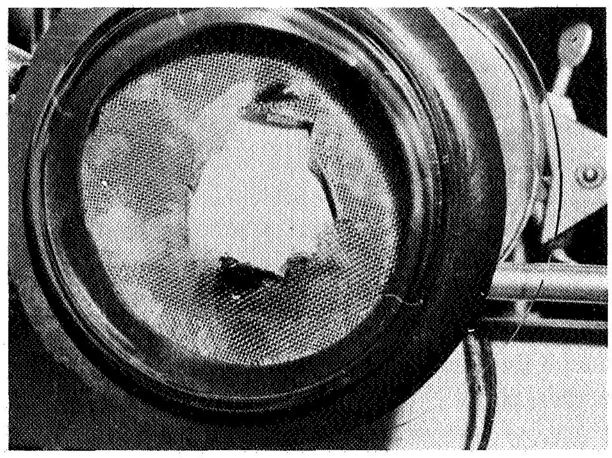

FIGURE 4.—Second view of writing on thermochromic in demountable CRT.

versible transition was noted. It thus appears from this experiment and from the electron-microscope experiments that voltages higher than 30 kilovolts are desirable if the electron beam is to impinge directly on the thermochromic.

The work thus far has demonstrated that thermochromic behavior can be induced by electron bombardment, either with a thermochromic alone or with a thermochromic-phosphor mixture as the target material. However, much remains to be done before a practical CRT using a thermochromic or a thermochromic-phosphor mixture is a reality. The effect of the substrate on the "writing" and "erasure" times was emphasized by the results of the experiments in which substrates of different thicknesses were used. The unexpected result that thermochromic behavior was observed under high accelerating voltages, while decomposition of the thermochromic occurred apparently under less energetic conditions, points to the fact that a quantitative knowledge of the beam energy, the dwell time of the electron beam, and other operating parameters of the CRT is necessary before a display device can be properly designed.

In addition to the applications already mentioned, at least two more possibilities are being considered. One involves the use of a thermochromic computer input board similar to a Rand Tablet, and the other involves a combined electroluminescent-thermochromic device.

A Rand Tablet is a graphical input device for a digital computer using a crossed grid of wires separated by a dielectric. One "writes" using a capacitance probe, and the junctions are sampled sequentially to sense the position of the probe. With a matrix of reasonable size, for example, 1000 by 1000 wires, this requires extremely high sampling rates to sense the position of the probe unless it is moved very slowly. The use of thermochromics and a heated probe would allow one to see what he is writing and also greatly reduce the requirements for high sampling rates, since the conductivity of individual junctions remains high until the material has reverted to its low temperature form.

Display devices combining thermochromics and electroluminescence should yield a display suitable for operation over a wider range of ambient illumination than either a pure thermochromic or electroluminescent device alone. The electrical conductivity of the high-temperature form of the thermochromic could provide a storage mechanism for the EL.

Materials

It is appropriate to emphasize some of the limitations of thermochromic materials. Their main disadvantage is that the color change depends on heat. This implies possible high power consumption and thermal instability. Before either of these factors can be fully assessed, an actual device configuration is necessary. In an actual device, the heat capacity and thermal conductivity of the substrate must be taken into account.

A slow decomposition is observed when either Ag_2HgI_4 or Cu_2HgI_4 is maintained for extended periods of time about 10° C above their respective transition temperatures. Static weight loss studies are shown in figure 5. Thermal decomposition has also been observed in long-term cycling tests. In the cycling tests, Ag_2HgI_4 still showed some thermochromic activity after 475 000 cycles (~11 months), while the Cu_2HgI_4 was almost totally decomposed after 153 000 cycles (~3½ months). Because decomposition involves the formation of the relatively volatile mercuric iodide with subsequent loss of thermochromic activity, some preliminary encapsulation techniques were attempted

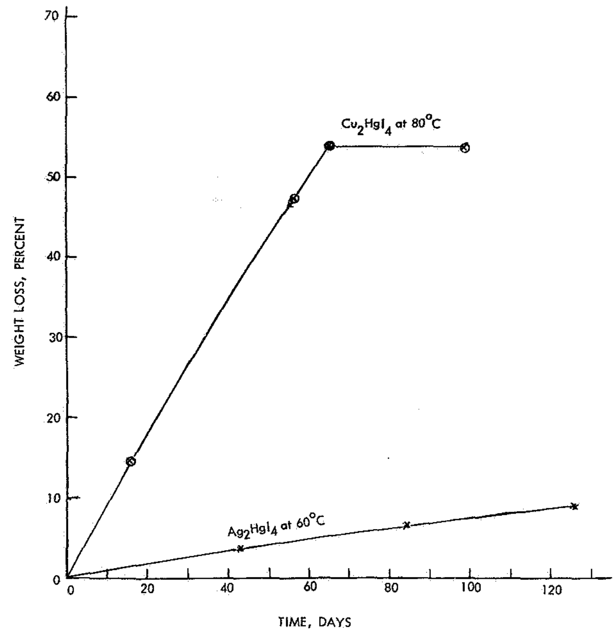

FIGURE 5.—Percent weight loss of Ag_2HgI_4 and Cu_2HgI_4 after 3 months at constant temperature.

to impede this loss. The results (fig. 6) indicate that encapsulation improves lifetime, and further improvement can be expected by varying the nature of the encapsulant.

A series of ternary chalcogenides of the general formula $M_2M'X_4$, where M was either Al^{3+}, Ga^{3+}, or In^{3+}, M′ was Zn^{2+}, Ca^{2+}, or Hg^{2+}, and X was either S^{2-}, Se^{2-}, or Te^{2-}, have been investigated. These compounds were chosen for study because many of them were reported to have crystalline structures similar to the ternary mercuric iodides mentioned above and might be expected to undergo an order-disorder transformation (ref. 8). In addition, the chalcogenides are known to be more stable than the iodides. Most of these compounds were found to be black and as such could not be used. Most of the colored ternary chalcogenides showed only very small reflectivity shifts. One exception, $CdIn_2S_4$, gave a reflectivity shift at approximately 500 Å on heating (fig. 7). A visual color change from dullish orange at room temperature to a reddish orange color at 95° to 100° C was observed for this compound. Both static thermal stability tests at 110° C and weight loss studies at room temperature under continuous evacuation (10^{-7} torr) showed that $CdIn_2S_4$ was very stable. This material, however, has

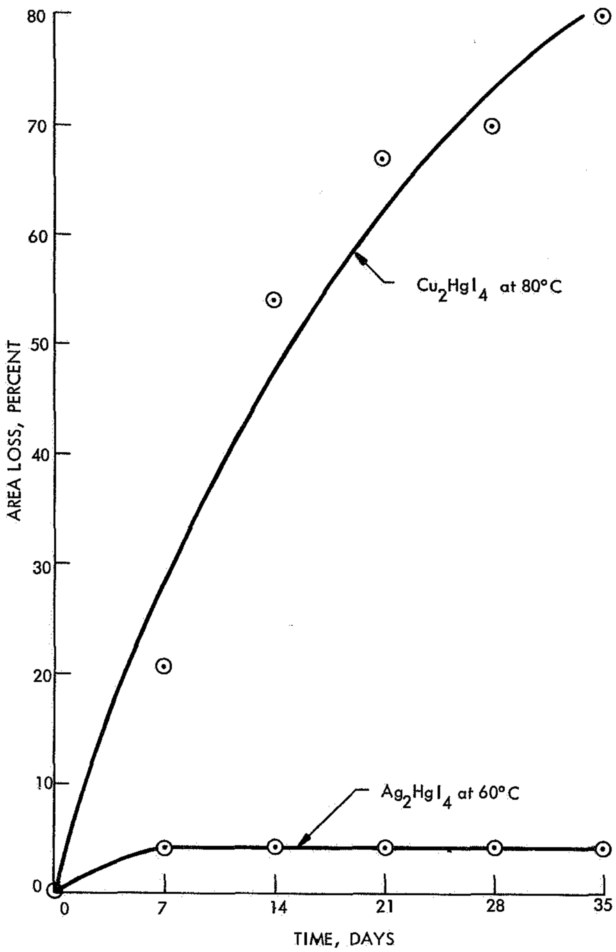

FIGURE 6.—Percent area loss of plastic encapsulated Ag_2HgI_4 and Cu_2HgI_4 as a function of time at a constant temperature.

limitations; the color change takes place over a relatively broad temperature range and does not afford good contrast in white light. Specific heat data as a function of temperature have shown that no latent heat accompanies this color change (fig. 8). This is not the case with the ternary mercuric iodides.[2]

In addition to the ternary chalcogenides, the preparation of more useful thermochromics has involved partial substitution in the ternary mercuric iodides. Materials have been prepared in which Cd^{2+} was substituted for Hg^{2+} and Br^- or Cl^- ion was partially substituted for iodide ion in Ag_2HgI_4. In the case of partial cationic substitution, the transition temperature in-

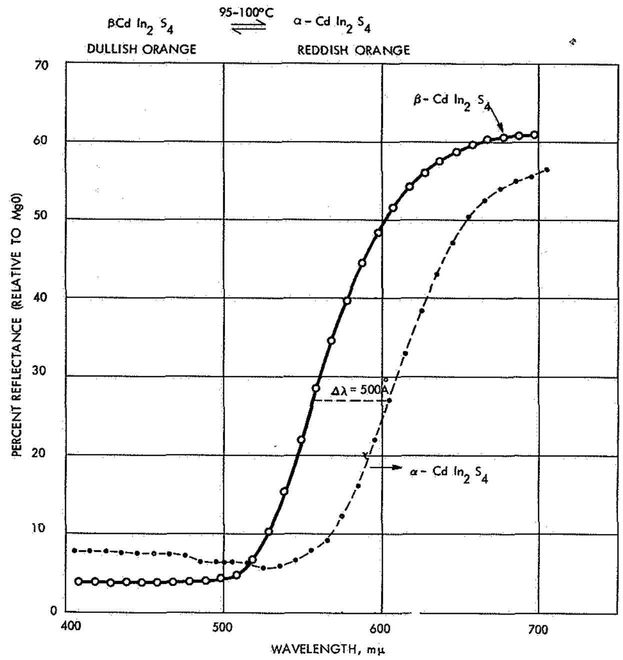

FIGURE 7.—Reflectance curves for β-$CdIn_2S_4$ and α-$CdIn_2S_4$.

creased with increasing amounts of Cd^{2+}, reaching a range near 88° C for 40 percent Cd^{2+}. Rather broad transition temperature ranges were noted for these formulations. Partial anionic substitution by either Br^- or Cl^- caused a lowering in the transition temperature. No linear relation between the amount of bromide ion substituted and the transition temperature was noted. The optimum thermochromic properties for the bromine-substituted formulations occurred at 4 percent bromide ion and the transition temperature changed from 50.5° C for the Ag_2HgI_4 to 42±2° C and from 69° C to 50° C for Cu_2HgI_4.

Measurement of specific heat as a function of temperature on the optimum bromide-substituted formulation $Ag_2HgI_{3.84}Br_{0.16}$ showed a highly endothermic phenomenon starting at

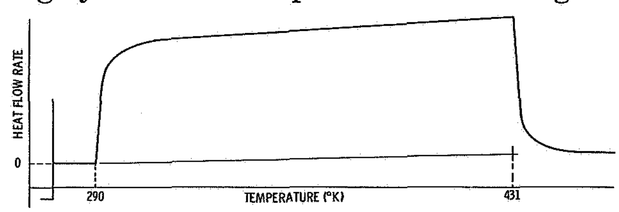

FIGURE 8.—Thermogram for $CdIn_2S_4$. Range: 2 millicalories per second, full scale. Scan speed: 10° per minute. Paper speed: 1 in. per minute. N_2 purge: 20 cc per minute.

[2] The authors are indebted to the Perkin Elmer Corp., Norwalk, Conn., for these measurements.

about 44° C. The results are shown in figure 9. Endothermic behavior was observed in earlier studies on Ag_2HgI_4 (ref. 9).

Efforts to produce more stable and useful thermochromics are continuing. One system that shows promise at this time is the AgI-PbI_2 system. The thermochromic compositions $2AgI:PbI_2$, $3AgI:PbI_2$, $AgI:PbI_2$, and $2AgI:3PbI_2$ have been prepared. The thermochromic behavior of such compositions is described as follows:

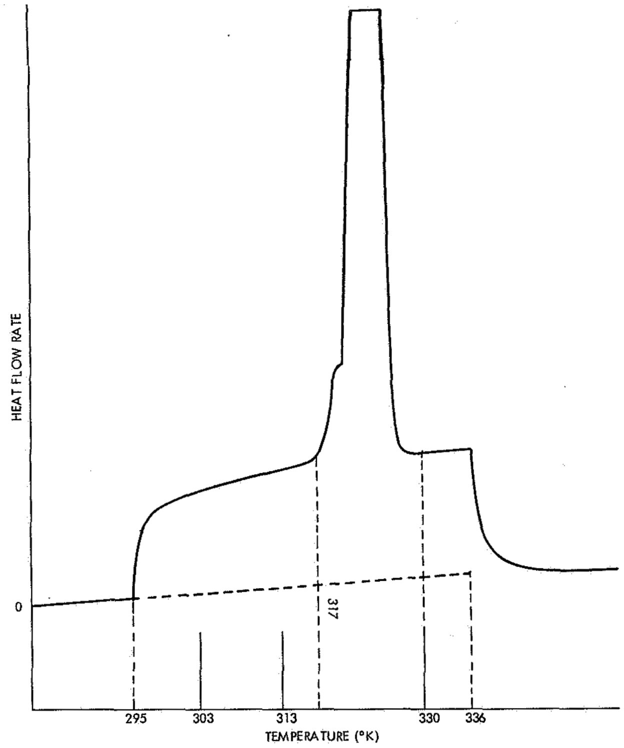

FIGURE 9.—Thermogram for $Ag_2HgI_{3.84}Br_{0.16}$. Range: 2 millicalories per second, full scale. Scan speed: 10° per minute. Paper speed: 1 in. per minute. N_2 purge: 20 cc per minute.

$3AgI:PbI_2$—broad transition, yellow to red, 99° to 137° C, pronounced change at 124° C

$2AgI:PbI_2$—broad transition, lemon yellow to reddish orange, 110° to 137° C

$AgI:PbI_2$—broad transition, yellow to orange red, 99° to 135° C

$2AgI:3PbI_2$—broad transition, yellow to red orange, 97° to 137° C

X-ray studies indicate that these are all solid solutions rather than stoichiometric compounds.

Static weight-loss tests were conducted on $2AgI:PbI_2$ at 150° C and atmospheric pressure for 2 months. No change was observed in that time. Weight-loss tests were also conducted on $2AgI:PbI_2$ under continuous evacuation at 10^{-7} torr and room temperature for 79 days. Initially, there was a weight gain of 2 percent after 5 days. No further weight change was noted for the remaining time, but the sample did show some darkening of the surface color. These systems definitely show promise as possible display materials. They have good contrast, thermal stability, and apparent fast response.

REFERENCES

1. DAY, J. H.: Chem. Rev., vol. 63, 1963, p. 65.
2. KETELAAR, J. A. A.: Z. Krist., vol. 80, 1931, p. 190.
3. KETELAAR, J. A. A.: Z. Krist., vol. 87, 1934, p. 436.
4. TAMMANN AND VESZI: J. Anorg. Chem., vol. 168, 1927, p. 46.
5. SUCHOW, L.; AND KECK, P. H.: J. Am. Chem. Soc., vol. 75, 1953, p. 518.
6. ASMUSSEN, R. W.; AND ANDERSEN, P.: Acta Chem. Scand., vol. 12, 1958, p. 939.
7. BURKOWSKI, R. P.; FINNIE, L. N.; KORNBLAU, M.; GRAFSTEIN, D.; AND HILBORN, E. H.: Applications of Thermochromic Materials in Display Devices. Proc. of the 7th Natl. Symp. of the Soc. for Information Display, Boston, Mass., Oct. 1966.
8. HAHN, H.; FRANK, G.; KLINGLER, W.; AND STORGER, A. AND G.: Z. Anorg. Allgem. Chem., vol. 279, 1955, p. 241.
9. KETELAAR, J. A. A.: Z. Physical Chem., vol. 26B, 1934, p. 327.

PROPERTIES AND APPLICATIONS OF PHOTOCHROMIC GLASSES

R. J. Araujo
Research and Development Laboratories, Corning Glass Works, Corning, N.Y.

The behavior of photochromic glasses results from reaction of light with silver halide crystals deliberately formed in the glass during its manufacture. The mechanism of the photolytic reactions is postulated and compared with the theory of the photographic process for silver halides. Reversibility of photochromism in these glasses is explained by prevention of diffusion of the products of photolysis from the original crystal site within the glass structure and their subsequent recombination when the activating light is removed. These glasses show large ranges in all their photochromic properties resulting from ranges in composition and in size and number of the included crystals. The behavior of typical glasses is reported in this paper, and some applications for these glasses are suggested.

Introduction

A photochromic material is one that undergoes a change of color when exposed to light and reverts to its original color when the light is removed. Articles carrying extensive bibliographies have been written by Brown and Shaw (ref. 1) and by Schwab and Bertelson (ref. 2).

Organics

Schwab and Bertelson divide the reactions in different organic photochromic materials into six categories; Windsor (ref. 3) reduces these to the following three main classes of general interest, based on how they work.

Stereoisomers.—Absorption of light breaks one of the chemical bonds in a ring molecule, thus allowing the molecule to unwind and form a different geometrical arrangement. The reverse process is a re-forming of the bond. Examples of this class are the spiropyrans and the anils.

Dyes.—A triphenylmethane dye, for example, is oxidized by energetic light; the absorption characteristics of the positive ion so formed are different from those of the original electrically neutral benzene rings.

Triplet states.—In the class of polynuclear aromatic hydrocarbons, ground-state molecules are excited first to a singlet state by irradiation, and then, via the lowest triplet state, go to an excited triplet state. Visible light is absorbed in the triplet-triplet transition.

Inorganics

To the major classes of organic photochromics must be added several kinds of inorganics, also listed and described by Brown and Shaw.

Alkaline earth sulfides.—Traces of a metal such as manganese or bismuth appear to be necessary for photochromism.

Zinc sulfide.—Lithopone, observed as early as 1881, is a compound of zinc sulfide and barium sulfate. The zinc sulfide appears responsible for the compound's sensitivity to light.

Titania and alkaline earth titanates.—In the titanates, a contaminant, such as iron or any of several other metals, also appears necessary for darkening to occur.

Mercury compounds.—Many of the mercury compounds, particularly those containing a halogen, have been observed to be photochromic.

In all of these materials, the photochromic response will depend on the intensity and spectral character of the incident light; on environmental parameters such as temperature, supporting matrix, or solvent; and, in most cases, on previous history. Most of the systems so far reported are only partially reversible, reversible with difficulty, or subject to "fatigue"—a change in behavior due to long use or lengthy storage. If the photochromic reaction is to be truly reversible, the quantum yield generally will be equal to, or less than, unity. When we compare this with a yield several orders of magnitude higher (in extreme cases, as high as 10^8) for ordinary silver halide photography, in which energy is added to the system chemically during development, we realize that photochromic processes are very "slow" in the photographic sense (ref. 4). However, these light-sensitive materials are unique in that the image is formed directly, and chemical processing to develop a latent image formed during exposure is unnecessary. In general, these inorganic materials are both reversible and reusable.

Glasses

Three general classes of photochromic glasses have been reported in the literature and these are described in the following sections.

Hackmanite types.—Hackmanite is a naturally occurring mineral of the soda alumina silicate-sodalite group; it has the stoichiometric composition 18 $(Na_2O \cdot Al_2O_3 \cdot 2SiO_2) \cdot 3NaCl \cdot Na_2SO_4$ and is, supposedly, a cubic crystal. It is usually opaque, white or blue, but can be melted to a glassy state, translucent to reasonably transparent, if a flux such as B_2O_3 is added. The minimum reported haze is 30 percent. With proper amounts of flux, the material darkens with exposure to ultraviolet light and can be bleached with longer wavelength (visible) light (ref. 5). Addition of other halides, such as bromide and iodide, can shift the absorption spectrum (color) of the resultant glassy material when it is darkened.

Cerium or europium.—Cohen and Smith (ref. 6) report that in suitably purified base glasses, either of pure silica or soda-silica, the addition of small amounts of cerium or europium, typically 100 ppm, has produced photochromic materials. Ultraviolet irradiation is absorbed by bands of cerium III or europium II centered in the ultraviolet, and it transfers photoelectrons to nearby traps that absorb in the visible region, producing an amethyst color. Decay times are typically a few seconds. Although the coloring and fading processes may be cycled repeatedly, the absorption band in the visible region (which produces the color) decreases in intensity with usage. This is believed to result from the photo-oxidation of the europium II to europium III. However, the band may be reduced, and the glass therefore resensitized, by exposure to short-wavelength ultraviolet light. These glasses, and their fatigue after exposure, have been studied in detail by Swarts and Pressau (ref. 7).

Silver Halide Photochromic Glass

Photochromic glasses that are truly reversible have been reported by Armistead and Stookey (ref. 8) of Corning Glass Works. The interesting properties of these glasses are a result of the minute particles of silver halide that are suspended in the host glass. A wide range of base glasses has been found to be suitable—alkali-metal borosilicates are perhaps the best, considering general glass qualities (clarity, durability, ease of melting, and forming) and photochromic behavior. Both the composition and the thermal history of the glass have large parts in determining its resultant photochromic properties. At high concentrations of the suspended colloid or following heat treatments which produce large average particle size, the glasses are translucent or opaque. The upper limit of silver for the transparent glasses is usually about 0.7 weight percent. Other metals, in the form of polyvalent oxides, including arsenic and antimony, tin and lead, and copper, increase the sensitivity and the photochromic absorbance. Electron-microscopic examination of the photochromic glasses shows small, dense particles, which are not seen in glasses that, because of composition or improper heat treatment, are not photochromic. The average particle size and number can be determined (with limitations) by counting from such photographs, or, with more precision, by small-angle X-ray scattering measurements. In general, glasses with particles less than about 50 Å in diameter are not photochromic. Above about 300 Å, the particles scatter light, and the resultant glass is opal. For particles of average diameter (100 Å), present in a concentration of, say, 0.2 percent in the glass, there will be about 4×10^{15} particles/cm^3, with an average spacing of 600 Å between particles.

General Properties

Large possible ranges and variations in composition, coupled with variations in time, temperature, and schedule of any subsequent heat treatment, provide a wide range of photochromic properties, greatly different rates of darkening and of recovery, and similarly large ranges of dependence of reaction rates and equilibrium states on temperature. The glasses are darkened by absorption of high-energy photons, in the near-ultraviolet or shorter wavelength visible region of the spectrum. The long-wavelength limit of the spectral sensitivity

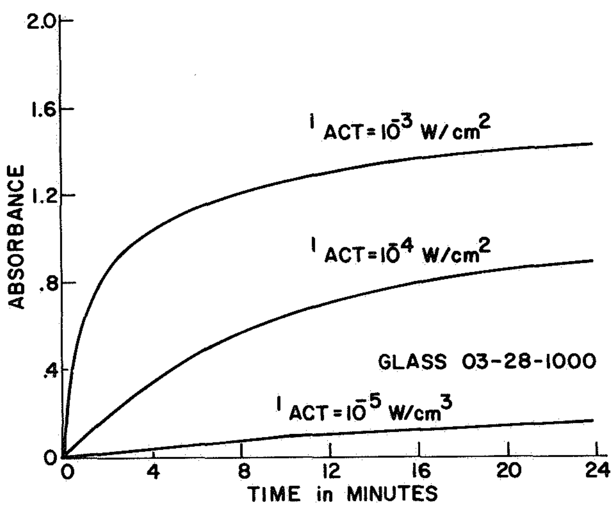

FIGURE 1.—The approach to equilibrium absorbance at different levels of incident energy. Wavelength of the activating light was 4000 Å.

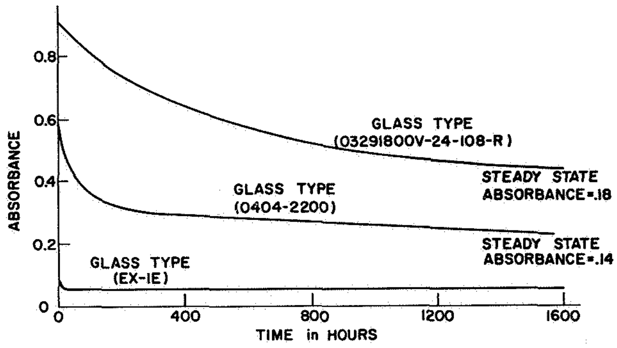

FIGURE 2.—Recovery of three glasses of different fading rates, after activation. The glass labeled EX–IE has a relatively very high thermal-fading constant.

for darkening is higher for glasses that contain the heavier halogens.

The rate of darkening is primarily dependent on the intensity of the light (in the proper spectral region). The approach to equilibrium absorbance for a glass illuminated at three different intensities is shown in figure 1. The rate of natural recovery is primarily determined by glass composition and heat treatment. Recovery in normal room light after irradiation is shown for three different glasses in figure 2. Recovery (to half-maximum absorbance) in the dark at room temperature is measured at times

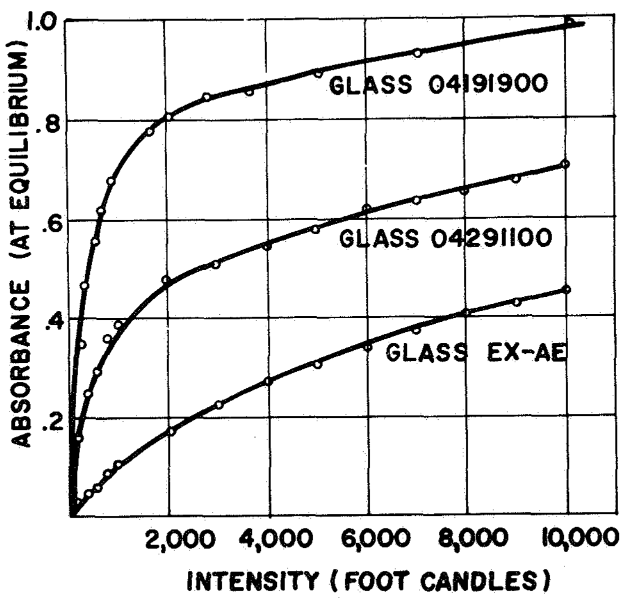

FIGURE 3.—Steady-state optical density versus light intensity, at constant temperature. The thermal-fading rate constant of glass EX–AE is greater than that of glass 04291100. Glass 04191900 is a relatively very slowly fading glass.

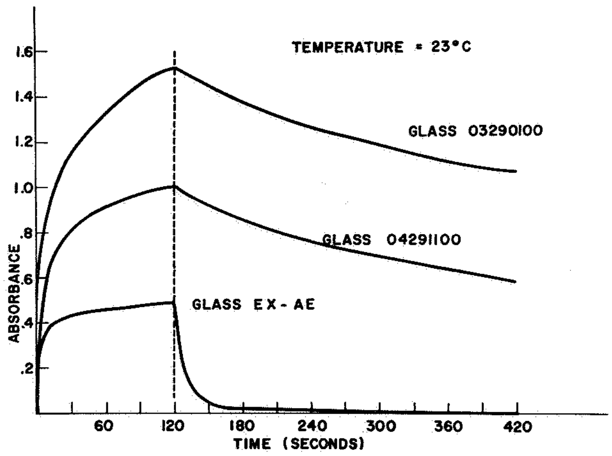

FIGURE 4.—Photochromic darkening and fading in three representative glasses, at room temperature (23° C), with light of constant intensity.

ranging from a few seconds to hundreds of hours. Absorption coefficients at equilibrium range up to about 20 cm^{-1} for the initially transparent glasses; that is, a resultant transmittance of about 15 percent for 1-millimeter thickness. The photochromic darkening for three glasses, selected to have a wide range of darkening and fading rates, is depicted in figure 3. The relative linearity of absorbance with intensity can be seen to depend on the fading rate of the glass, which is rationalized later in this paper. The short-time approach to equilibrium of three selected glasses, under constant illumination at room temperature (23° C), is seen in figure 4;

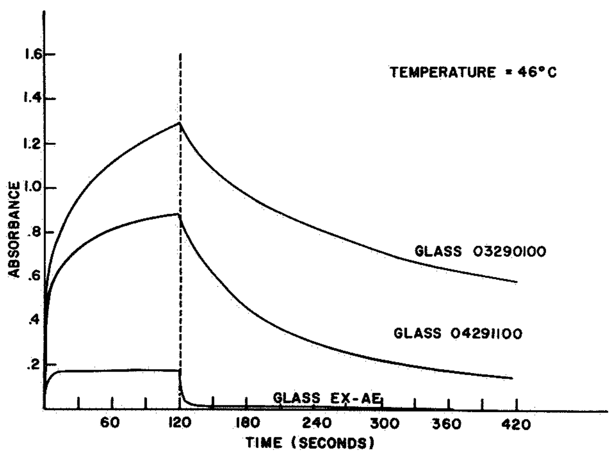

FIGURE 5.—Photochromic darkening and fading in three representative glasses at 46° C. Comparison with the curves of figure 4 shows the interdependence of equilibrium absorbances, rates of approach to equilibrium, and temperature.

figure 5 shows the behavior of the same three glasses at 46° C, with the same light source. The light from the arc was turned on at times zero and turned off after 120 seconds. The glasses were chosen to show differences of fading rates and equilibrium absorbance. The samples were maintained at the stated fixed temperature so that the glass temperature was not appreciably raised by the energy absorbed. The increased effect of change in ambient temperature on a faster clearing EX–AE is illustrated by these curves.

THEORY

It is expected, of course, that silver-halide crystals will be decomposed by light to form a silver image, as in conventional silver-halide photographic film. An essential difference, however, is that in conventional silver-halide photography, the incident photons decompose the atoms within the crystal into elemental silver and halogen; the silver may be subsequently chemically developed, and the halogen diffuses away from the original crystal site. In the silver-halide glasses, the halogen is held within the surrounding glass matrix, and is available for recombination with the silver, permitting recovery of the glass to its original, colorless state after the light is removed. The recombination occurs by two independent processes: a natural, thermal recovery, or by interaction with light of longer wavelength (lower energy) than that which darkens the glass. For these silver-halide crystals, the behavior is determined both by states at their surfaces and in their interior. But the simplifying assumption of a single species and first-order reactions permits generalizations about the behavior of the glasses. Under illumination, the change of concentration of absorbing color centers for this case will be given simply by

$$\frac{dc}{dt}=k_d I_d A-(k_f I_f+k_d I_d+k_t)c \qquad (1)$$

where c is the concentration of color centers; k_d, k_f, and k_t are generalized rate constants for darkening, optical bleaching, and thermal fading, respectively; I_d *and* I_f are the integrated intensities of the light, darkening and fading, over the respective wavelength ranges; and A is the initial concentration of sensitizable centers in the glass. When equilibrium is attained, $dc/dt=0$, and the equilibrium concentration of color centers will be

$$c_s=\frac{k_d I_d A}{k_f I_f+k_t+k_d I_d} \qquad (2)$$

The photochromic behavior of any glass will be determined, therefore, by the relative magnitude of the rate constants describing it; these constants are determined by the composition of the glass and by the state of the crystals produced from it; that is, by its heat treatment. If we assume the thermal-fading rate constant(s) to be vanishingly small, then c_s is independent of the light intensity (for constancy of the ratio of darkening to bleaching light intensities); if the total thermal-fading rate is large and becomes the determining rate constant, then c_s is proportional to the intensity. Figure 6 schematically displays these relationships at an assumed fixed temperature. (Note that the curves in fig. 3 are similar to the intermediate case shown in fig. 6.)

Although equation (1) is useful for qualitatively explaining some properties of the glass, it fails to quantitatively describe actual kinetic data. The kinetics of darkening is most conveniently studied (separately from the kinetics of fading) in glasses having extremely slow fading rates. Best attempts to fit curves to carefully measured darkening and fading data lead to the following conclusions.

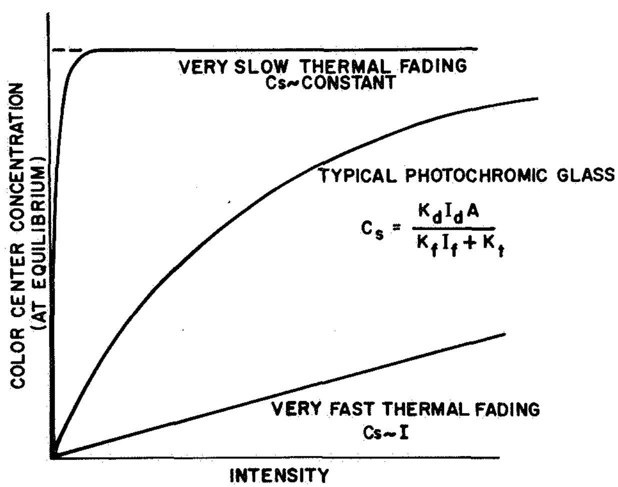

FIGURE 6.—Color center concentration at equilibrium versus intensity of illumination. The dependence on intensity is determined by the relative values of the rate constants for darkening and bleaching by light and for thermal fading.

The darkening is described by two first-order darkening terms:

$$\frac{d(x_1+x_2)}{dt}=k_1 I_d(A_1-x_1)+k_2 I_d(A_2-x_2) \qquad (3)$$

where the x's are related to the concentration of color centers in such a way that x_1+x_2 is equal to the absorbance, and A_1+A_2 is the maximum concentration of color centers attainable; therefore, $A_1+A_2=x_1+x_2$ at $t=\infty$. This suggests strongly that surface darkening and volume darkening are two separate modes, which determine the behavior of these crystals.

Data from a single thermal fading experiment are very well described by equation (4),

$$\ln(x)=-k\ln(t+t_o)+k' \qquad (4)$$

where (x) represents the absorbance, k and k' are adjustable constants, and t_o depends on k, k' and the initial value of (x).

The constants in equation (4) change when one changes the darkening conditions, such as length of time illuminated and intensity of activating light. In other words, the fading rate is not a unique function of the absorbance.

One can construct a model of carriers diffusing through the crystal and recombining with oppositely charged carriers when a critical separation is reached. Such a model can be shown to be completely consistent with equation (4). Furthermore, it qualitatively explains why the fading rate is not a unique function of the absorbance.

A continuous density function of electron concentration exists, which depends on the separation of the electrons from the holes with which they will recombine. The fading rate depends only on the gradient of the density function at the critical separation, while the absorbance depends on the integral of the density function over all separations. Hence, these two properties are not rigidly correlated.

Applications

The properties of the photochromic glasses discussed previously suggest some possibilities for their use in optical systems including information storage and display, photography, and glazing in buildings and in vehicles. Prescription lenses of photochromic glass are commercially available. Megla (ref. 4) suggested a display system that can be used to intensify the image on a screen without increasing the power density of the projection light source (fig. 7). The projected light beam is first filtered to permit only the bleaching wavelength (λ_{BL} around 600 mμ) to pass via a dichroic mirror to a photochromic glass plate acting as a display screen. There, the beam forms a bleached image on the preactivated photochromic glass plate. The rest of the glass is continually darkened by a flood illumination source that peaks at around 380 mμ. This uv irradiation is adjusted so that it is weak enough to be overruled by the information contained in the bleaching beam. As shown in figure 7,

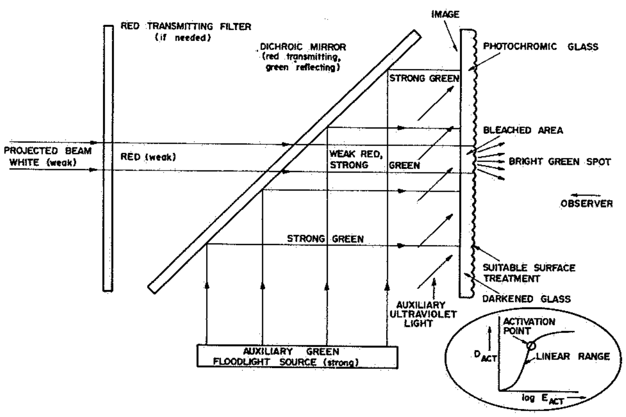

FIGURE 7.—Photochromic glass as display screen.

this can be achieved by activating the glass to a level corresponding to the upper part of the linear range of the D versus log E_{ACT} curve. A green floodlight of high intensity illuminates the photochromic screen after being reflected by the dichroic mirror. As pointed out previously, the green light of $\lambda=550$-mμ wavelength can be considered as a neutral wavelength and, therefore, does not affect the photochromic glass. However, it will pass through the bleached parts of the photochromic glass plate and will be absorbed by the activated ones. Thus, an amplified image can be viewed from the observer side of the screen.

Depending on the fading rate of the particular photochromic glass used, the image can also be stored for a longer period. Several variations of this principle are possible. For instance, during the display process, information can be added onto the photochromic glass screen with a separate collimated bleaching light source. Another variation would be to coat the back side of the photochromic glass with a green-light-transmitting and red-light-reflecting dichroic layer, permitting the front side of the screen to be viewed. An additional advantage of this technique is that the bleaching light energy would be reflected from the dichroic layer and would pass through the photochromic glass twice, thereby increasing the bleaching action.

Although the detailed requirements for each application may vary widely, silver-halide-containing photochromic glasses belong to such a large family that it is believed many of these requirements can be satisfied.

References

1. Brown, G. H.; and Shaw, W. G.: Phototropism (Photochromism). Rev. Pure Appl. Chem., vol. 11, no. 2, 1961, pp. 2–32.
2. Schwab, H.; and Bertelson, R. C.: Photochromism—State-of-the-Art Review. Presented at the 1966 Symp. on Unconventional Photographic Systems, Soc. of Photographic Scientists and Engineers, Washington, D.C., pp. 94–106.
3. Windsor, M. W.: Photochromism, *In:* Encyclopedia of Chemistry (G. L. Clark, ed.), Reinhold, New York, 1966, pp. 816–818.
4. Megla, G. K.: Optical Properties and Applications of Photochromic Glass. Appl. Opt., vol. 5, June 1966, pp. 945–960.
5. Radler, R.; and Chenot, D.: Synthesis of Inorganic Phototropic Materials for High Density Computer Memory Applications. Tech. Doc. Rept. AL TDR 64–170, Oct. 1964.
6. Cohen, A. J.; and Smith, H. L.: Variable Transmission Glasses Sensitive to Sunlight. Science, vol. 137, Sept. 21, 1962, p. 981; Ceram. Abstr., Feb. 1963, p. 41i.
7. Swarts, E. L.; and Pressau, J. P.: Phototropy of Reduced Silicate Glasses Containing the 570 mμ Color Center. J. Am. Ceram. Soc., vol. 48, July 1965, pp. 333–338.
8. Armistead, W. H.; and Stookey, S. D.: Science, vol. 144, p. 15, Apr. 10, 1964; U.S. Patent 3,208,860, Sept. 8, 1965.

FLUIDIC DISPLAYS

J. VAN DER HEYDEN
Martin Marietta Corp., Orlando, Fla.

This paper discusses recent developments in the field of fluidic displays and illustrates some of the areas in which fluidic display may be used.

Some of the basic advantages of fluidic systems are discussed in the first section of this paper. The treatment of fluidic displays is subdivided into a section on conventional fluid display systems, which can be used in conjunction with fluidic systems, and a section that covers the fluidic displays especially developed for use with fluidic computers and control systems.

In conclusion, some ideas that up to now have only been investigated in a cursory manner are presented as possible solutions to some of the problems associated with fluidic system displays.

INTRODUCTION

Fluidics is the field of technology of no-moving-parts fluidic components and systems for sensing, control, and logic (refs. 1 and 2). Fluidic devices operate faster than conventional pneumatic or hydraulic components. It has been established that, for certain applications, fluidic systems will exhibit high reliability, low cost, simplicity, freedom of maintenance problems, and an unparalleled hardness to nuclear radiation.

The significant contribution of fluidics has been that of partially closing the gap between conventional fluid and hydromechanical controls and electronic controls. Fluidic systems will be used extensively where new emphasis on automation and increased demands on cost and reliability are encountered.

Fluidics technology has grown rapidly from the basic inventions of 1959. Predictions are that the gross market in fluidics will be on the order of $200 million to $300 million by 1970, and part of this expenditure will be applied to fluidic display systems.

THE NEED FOR FLUIDIC DISPLAYS

Generally, any fluidic system application will have to be considered on an overall system basis rather than on an individual part-for-part substitution for control components. For instance, in cases where a fluidic system is considered as a replacement for an electronic system, it is easier to use fluidic sensing equipment than to use electrical-to-fluid signal transducers.

Interfaces between fluidic and electrical systems must be kept to a minimum or, if possible, eliminated completely to preserve the basic advantages of fluidic systems.

An example of a rather complex fluidic device (by present-day standards) using a fluidic display is shown in the model of a digital integrator (fig. 1). A digital integrator, called a digital differential analyzer (DDA), is an essential part of a digital computer. It performs a numerical integration of an area under a curve by successively summing all small area increments under a curve, as shown in figure 2(*a*).

291-351 O—68——6

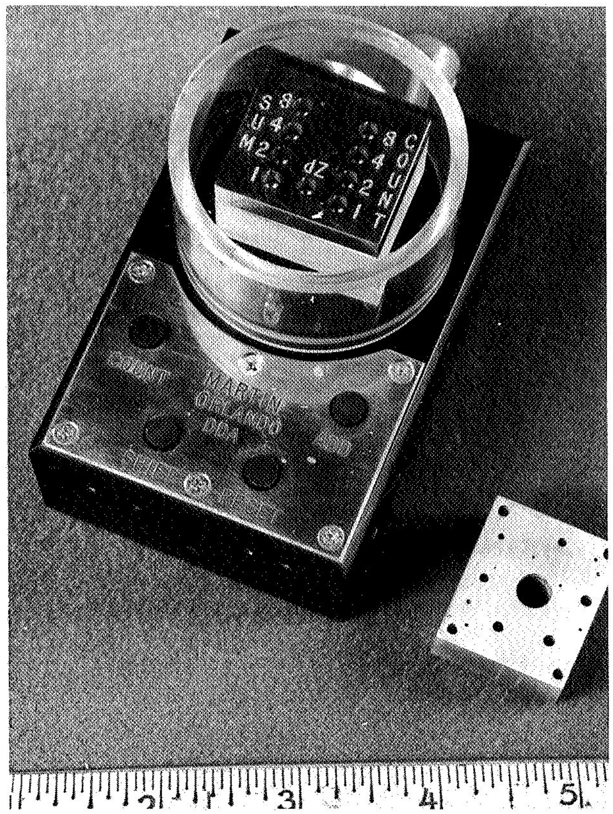

FIGURE 1.—Fluidic DDA.

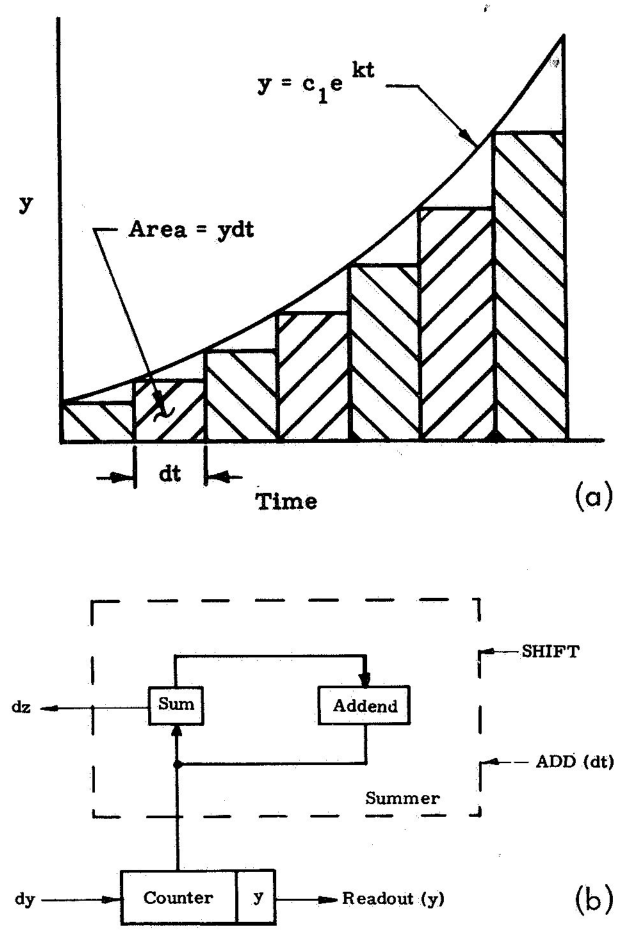

FIGURE 2.—DDA operation. (*a*) Numerical integration; (*b*) logic diagram.

The DDA consists of two binary circuits—a counter and a summer. The initial value of a dependent variable y is stored in the counter and can be changed by supplying pulses to the dy input shown. The summer contains a sum register and an addend register and the necessary carry logic to complete requirements for the addition process. In operation, the summer is supplied with two periodic signals, 180° out of phase. These are the add, or dt, and shift inputs shown in figure 2(*b*). In the proof-of-principle model, these signals were supplied through manually operated valves rather than by a fluidic signal generator.

The dt input essentially divides the time scale into small increments, each dt units wide. When it is momentarily interrupted, it commands the system to add the number in the counter (the y value) to the number that exists in the addend register. The number in the addend register is the value of the previous sum that was stored in the sum register. This value is transferred to the addend register by a momentary interruption of the shift input.

Assume that some number (y) is stored in the counter and that the sum and addend registers are set to zero. The first dt command adds this value (y) to the value in the addend register (0) and stores the answer in the sum register. This value represents the area of the first y dt area block under the curve of figure 2(*a*). Now the shift input is pulsed, which transfers this value to the addend register. The second dt command then adds the value in the counter to the value in the addend register (the previous sum) and stores this new value in the sum register. This value corresponds to the area under the curve represented by the first two y dt area blocks.

The sum register of a DDA has a finite capacity that is determined by the number of binary bits making up the system. When enough of these y dt area blocks are added into it, the register will overflow and generate a dz output. This dz output is proportional to a known amount of area under the y time curve and is equal to ky dt. If the value of y is decreased, a greater number of additions will be required to generate a dz output. If the value of y is increased, fewer additions will be required. Each dz pulse represents an area under the curve of ky dt; to obtain the total area, or the integral of ky dt, it is only necessary to count the number of dz pulses that appear.

Consider what happens when the dz output is connected back to the dy input. Since dz= ky dt, and in this case dy=dz, by substitution dy=ky dt is obtained. Separating the variables produces d$y/y=k$ dt, and integrating produces $\log y=kt+C$. Solving for y gives $y=e^{(kt+C)}=C_1e^{kt}$. So the number y that appears in the counter is the solution to the simple differential equation d$y/y=k$ dt. One DDA can solve this equation only; however, by suitably interconnecting a number of DDA's, many complex equations can be solved.

The DDA uses two distinct fluidic logic elements—the flip-flop and the OR–NOR gate. The logic gates are schematically shown in figure 3. The logic gates shown work on the wall attachment principle, also called the "Coanda effect." The fluid supply tends to attach itself to one sidewall as shown. A control signal applied through the control ports will force the output fluid jet to the output leg opposite the control port. The Coanda effect will, due to the

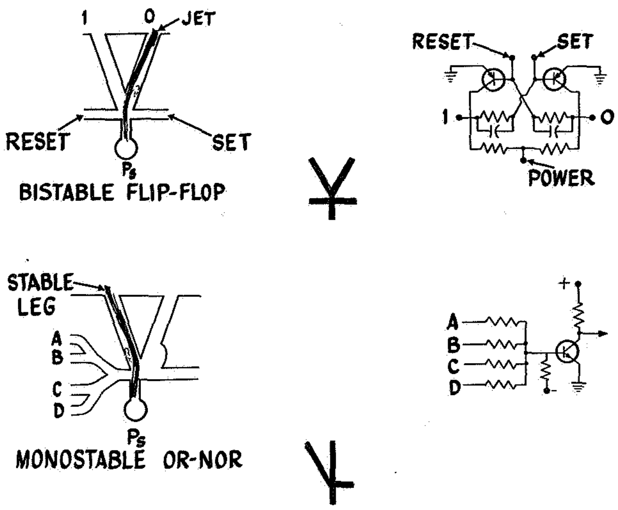

FIGURE 3.—Logic gate schematics.

geometry of the logic gate, cause the output jet to remain in the last commanded output leg, even after the control signal is removed. In order to obtain the monostable action of the OR–NOR gate, the sidewall is cut back sufficiently on one side to obtain stable operation in the desired direction when control signals are not present. Figure 3 also shows the electronic counterparts of the particular logic elements and the simplified schematic representations used throughout this presentation. Figure 4 shows the complete logic schematic of the four-bit DDA.

As shown in the schematic, a total of 132 logic elements were used to mechanize the circuit. The volume of the logic module is less than 0.6 cubic inch. The logic elements are chemically

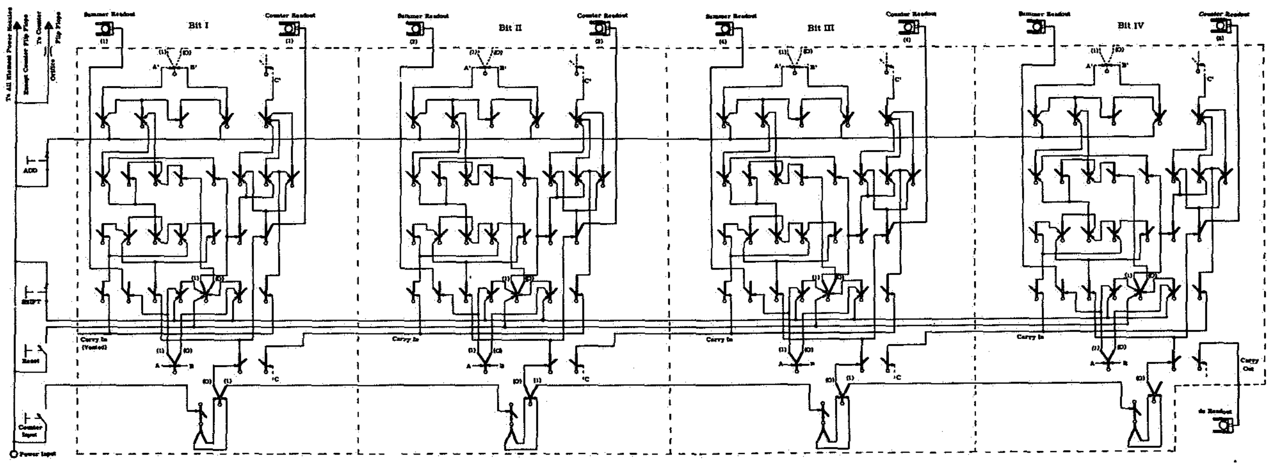

FIGURE 4.—Four-bit DDA schematic.

etched into copper laminate with a thickness of 0.004 inch. The interconnecting channels are also etched into 0.004-inch-thick metal planes. All logic elements and interconnecting planes are metallurgically bonded together. The complete assembly of the 134 logic planes (shown in fig. 5) is approximately 0.54 inch high and each is approximately 1 by 1 inch in area. The package density of the logic gates is then approximately 260 logic gates per cubic inch, which compares favorably with conventional electronic logic circuits. Power consumption for each of the logic gates is less than 50 milliwatts in equivalent pneumatic power.

The display part of the DDA consists of a readout block mounted on the column of planes (fig. 1). The readouts consist of tiny flowmeters made by inserting precision steel balls in bored tubes. Because of the limited flow output of the miniature elements, clearance between the ball and tube had to be held to 1 mil to insure proper operation. A readout is provided for each counter and summer stage and for the overflow indicator. A more specific description of the mechanization and construction of the fluidic DDA can be found in references 3 and 4.

Obviously, if fluidic computers are going to be part of our aeronautic and space technology, new display methods must be found that will be

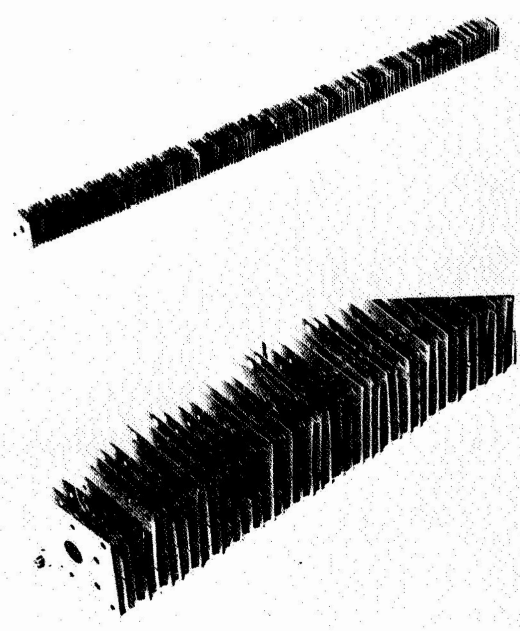

FIGURE 5.—Logic planes assembly.

compatible with the specific requirements of the fluidic systems. For instance, moving steel balls suffice for the demonstration unit, but are of limited use in airborne application. New and more advanced display techniques will have to be developed.

FLUIDIC DISPLAYS

For the purpose of this paper, a fluidic display can be defined as any device that converts fluidic signals into information that can be sensed by human beings. Most of the fluidic displays covered are visual displays; however, aural and other displays can be easily generated with fluidic techniques. Obviously, the more attractive fluidic displays are governed by the same ground rule that covers the fluidic technology as a whole: no moving parts.

Like electronic display systems, one or more parameters important to the controlled process are sensed and then displayed as monitoring or instructional aids to human operators. The sensed and displayed parameter is some measurement of one or more of the properties or energy levels of the fluid, such as pressure, density, flow, and chemical affinity.

Presently, most fluidic systems are used as control systems in an analog as well as digital mode, or as computers. Fluidic display hardware developments have, up to now, been restricted to the readouts on one or more of the controlled parameters of control systems and as readouts for fluidic computers. The field of fluidic displays can be conveniently subdivided into what may be called conventional fluidic displays, which are mainly direct applications of fluid instrumentation already available before the development of the fluidic technology, and the more advanced fluidic display techniques developed because of, and in conjunction with, requirements specifically generated by the application of fluidic technology.

CONVENTIONAL FLUIDIC DISPLAYS

A considerable amount of versatile instrumentation for measuring fluid properties is available and can be used without major modifications as a fluidic display system. In keeping with the distinction that separates fluidic systems from other fluid-operated systems, the ing with the distinction that separates fluidic systems from other fluid-operated systems, the conventional displays will be discussed in the following order:

(1) No-moving-parts displays
(2) Moving-parts displays
(3) Fluidic-electronic displays

No-Moving-Parts Displays

Examples of what are considered conventional no-moving-parts displays are fluid manometers, such as the familiar mercury barometer used to measure air pressure, and fluid-level indicators, such as steam-boiler-level indicators. The manometer is used to obtain information regarding fluid pressure, and can be used as an analog display as well as for indications of the on-off type.

Moving-Parts Displays

The design of most conventional moving-parts, analog-type fluidic displays is based on either the bourdon-type pressure gage or the fluid flow rotameter. In a typical bourdon-type pressure gage, the change in position of a coiled hollow tube or bellows resulting from a varied internal pressure is translated into a needle movement along a scale via some mechanical linkage. The basic design of a rotameter consists of a tapered hollow tube with a float inside the tube. If fluid flow is induced from the bottom of the tube, the float will move a distance proportional to the amount of flow, in cases where analog display is required. Two practical fluidic applications of these ideas are the demonstration model of a fluidic DDA, previously discussed, and an on-off indicator as shown in figure 6 (ref. 5).

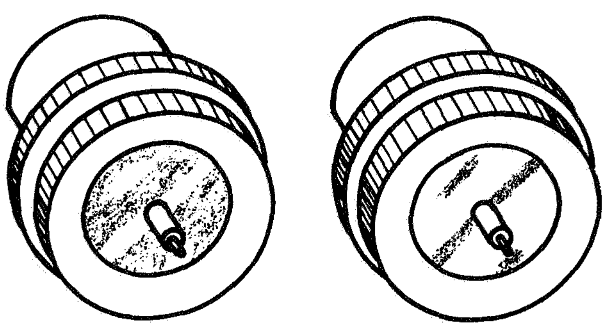

FIGURE 6.—ON-OFF indicator.

The on-off indicator uses a movable multicolored disk, which is partly exposed through a window. A change in pressure, applied to a pressure-sensitive mechanism, will move the disk behind the window. Several versions are commercially available and have been used extensively in similar form in hydraulic and pneumatic equipment.

Fluid-Electronic Displays

The advantages and availability of electronic display systems, such as CRT's and recorders, are generally utilized to the fullest in measurement of fluid phenomena. Suitable transducers that translate fluid properties into electrical signals are available.

Instrumentation, such as pressure transducers and hot-wire anemometers, is used extensively in the research and development stage of fluidic technology. The pressure transducers generally employ strain gages attached to a diaphragm. When fluid pressure is applied to the diaphragm, the strain-gage output will, with proper design, be proportional to the amount of pressure applied to the diaphragm.

Hot-wire anemometers are employed to measure flow in gases. The hot-wire anemometers consist of a small wire suspended in a fluid stream. The wire is heated to a predetermined value by an electrical current. The gas flow will cool the wire. The amount of heat extracted from the wire is proportional to the amount of gas flowing around the wire. An indication of the amount of gas flowing is obtained by monitoring the amount of current necessary to maintain the wire at the predetermined temperature level.

Special instrumentation for fluidic work has been developed and is commercially available. From a research scientist's standpoint, the versatility of electronic display equipment makes it almost ideal for use with fluidic equipment. In most cases, however, where display systems are part of an actual fluidic control system or computing device, the electronic equipment will be completely useless.

The advantages of choosing a fluidic system instead of another system (particularly, an electronic system) are low cost, resistance to nuclear radiation, and reliability. Adding an electronic display system to a fluidic system defeats these advantages.

Advanced Fluidic Displays

This category includes the display units specifically developed for use with fluidic control systems or computers. In most devices classified as advanced displays, the concept of no moving mechanical parts and no electronics is stringently followed to preserve the advantages of fluidic systems. Most major efforts, up to now, have been concentrated on obtaining an acceptable fluidic alphanumeric display driven by fluidic logic circuits used as a decoder.

Fluid Bead Display

One device proposed recently (ref. 6) consists of a matrix of hollow glass cavities, as shown in figure 7. The cavity has a narrow section in the center of the tube. When a drop of fluid having sufficient surface tension is introduced in the cavity, the surface tension forces of the fluid will make the droplet attain a shape that exhibits the smallest surface area for the given volume—a sphere. Because the droplet of fluid is just large enough to fit in the end cavities, the narrow center section is an unstable state and the fluid bead will be located in either one of the two end cavities. Introducing pressurized gas in either one of the two end cavities will force the fluid bead to the opposite side. The fluid bead will remain at the last position even after the gas pressure is turned off and thus acts as a memory device.

A drawback of this display system is its physical size and its sensitivity to shock and vibration. At higher g-levels, the fluid bead may disintegrate and will not reunite spontaneously. Some of the fluids used as memory and display beads are mercury and various combinations of water and glycerine with coloring agents.

Mercury has an extremely high surface tension; however, two objectionable characteristics (namely, the g-sensitivity and the toxicity of the fluid) make other liquids more desirable (ref. 6). The most advantageous system thus far reported is a drop of glycerine coated with a thin layer of silica powder. The surfaces touched by the fluid bead are coated with a water repellant to increase the surface-tension effect.

Thermochromic Digital Display

A more versatile approach to the fluidic alphanumeric display problem is the use of thermochromic material (ref. 7). A thermochromic material will change color at a certain temperature. The change from one color at the low-temperature level to the high-temperature color state can be repeated fast enough for use in a fluidic alphanumeric display unit. Several materials exhibit thermochromic properties and investigation of new materials is continuing.

Figure 8 shows a model of a fluidic decoder and display device being built as part of a NASA-sponsored study contract on fluidic displays (ref. 8). It is a good example of what can be obtained with state-of-the-art miniature fluidic components. The unit is only 6.25 cubic inches in volume and will be capable of decoding 16 bits of binary information into a five-digit octal display. The unit consists of the following components:

(1) Electrofluid interface
(2) Register
(3) Binary-to-octal converter
(4) Display decoder logic
(5) Display

The function of the electrofluid converter is to translate the incoming electrical signals into pneumatic signals of sufficient power to operate the fluidic logic elements in the remaining part

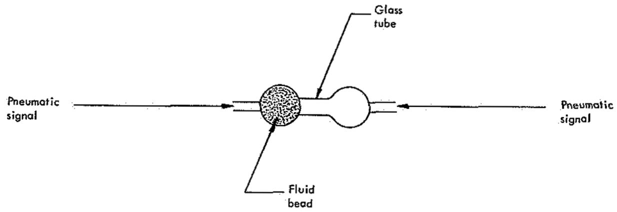

Figure 7.—Glass cavity display.

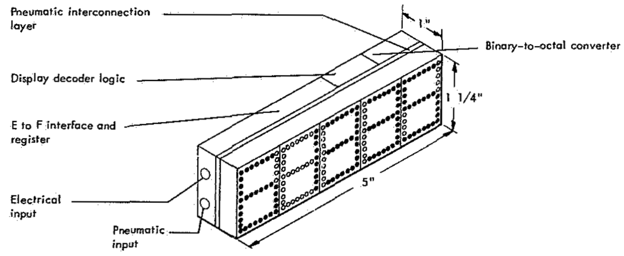

Figure 8.—Fluidic decoder and display module.

of the decoder and display device. The available electrical signal is a 20-volt square wave of 15-millisecond duration. It is postulated that a current of only 10 milliamperes is available because of computer limitations. The amount of pneumatic power required to obtain switching flow for the miniature fluidic elements is small. These low-power requirements make it possible to consider several concepts of electrofluid interface devices. The most promising device is a pneumatic flapper valve in which a piezoelectric crystal is used to open and close a small orifice. The pressure changes obtainable with this miniature device are sufficient to switch a miniature fluidic gate. The latest model of this device has been operated with an input as low as 10 volts dc. Because the amount of power required to switch the fluidic elements is extremely small, the necessary movement of the piezoelectric crystal can be obtained without using electronic oscillators and step-up transformers, which is a significant improvement over prior practice.

The function of the register is to store the binary information, supplied by the computer via the electrofluid interface, for further processing. The register consists of fluidic flip-flops that perform the function of a memory. A schematic of the register is shown in figure 9. The register input is obtained from two parallel pneumatic inputs: one input is the clock pulse; the second input contains the binary information. The register is actually a 16-bit fluidic shift register.

The decoder logic module transforms the binary information into corresponding signals to the readouts. Figure 10 shows the logic design used to obtain the proper activation signals A through G from the four binary bits of information available from the register.

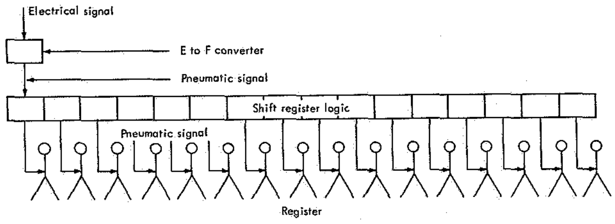

FIGURE 9.—Serial register schematic.

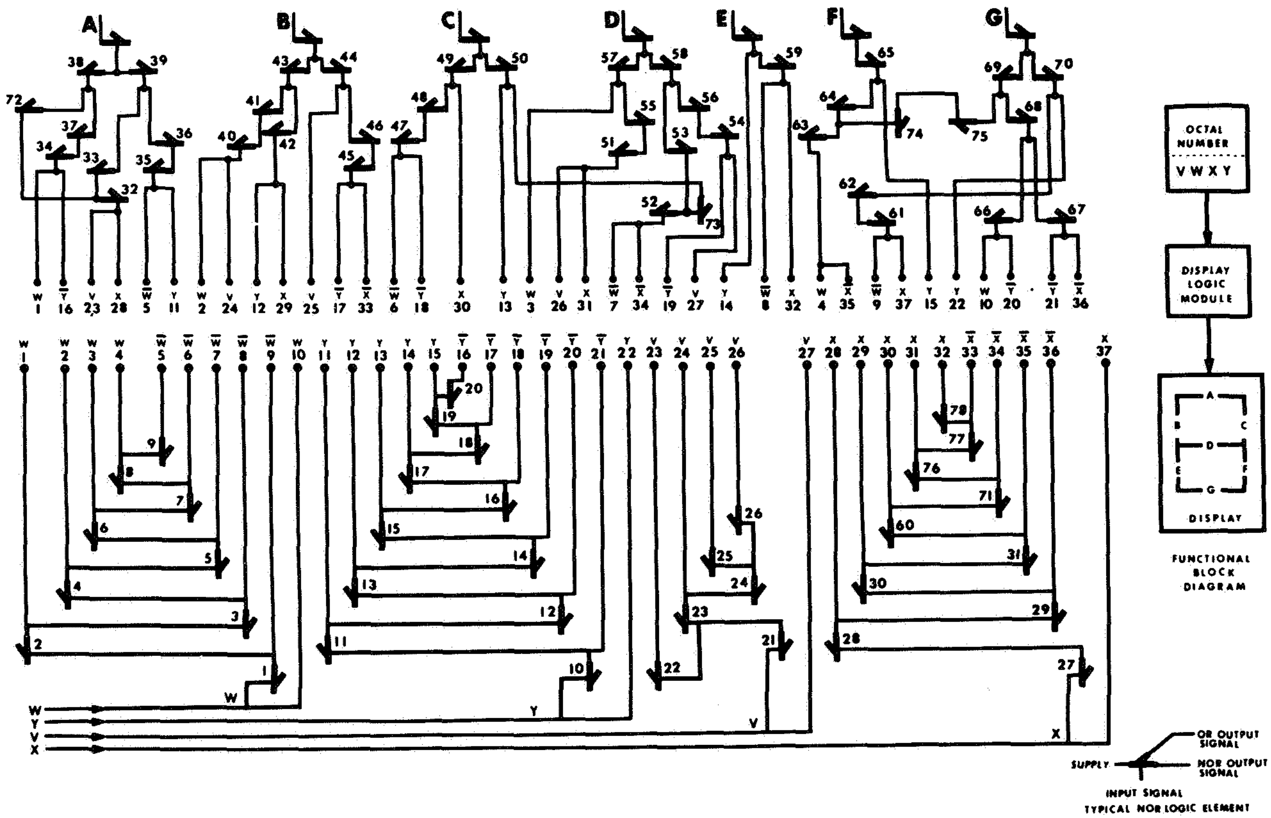

FIGURE 10.—Display logic module schematic.

The particular logic design used is by no means the most economical, from the standpoint of the number of logic gates used, but experiments showed that when certain limitations on fan in and fan out were observed, the logic gates could be operated at lower gas supply pressures which, in turn, reduce the total power consumption of the decoder.

The decoder logic gates are formed in sheets of 0.004-inch-thick copper by a chemical etching process. Each etched plane contains two logic gates, as illustrated in figure 11 (which shows the actual size of the planes used in the decoder logic module). Also shown is one of the planes containing interconnections between the logic element planes. The complete decoding logic block is formed by diffusion bonding of interconnecting planes and element planes into one integrated assembly measuring approximately 1.2 by 1.0 by 0.4 inches. The integrated bonded assembly is free from leakage paths, which are common causes of malfunctions in conventional fluidic circuitry, and the small interconnections have resonant frequencies far above the response time of even the comparatively fast miniature fluid gates, thus eliminating another problem so often experienced with larger fluidic logic element assemblies.

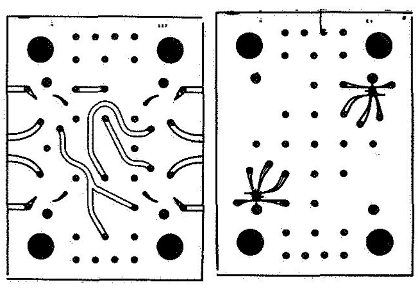

FIGURE 11.—Decoder planes.

Figure 12 shows a fluidic numeric display digit which uses the thermochromic material to obtain color contrasts. The conversion temperature is around 50.5° C for the material used in this demonstration model. Other thermochromic materials with different conversion temperatures and different color characteristics may be used, however, when so desired.

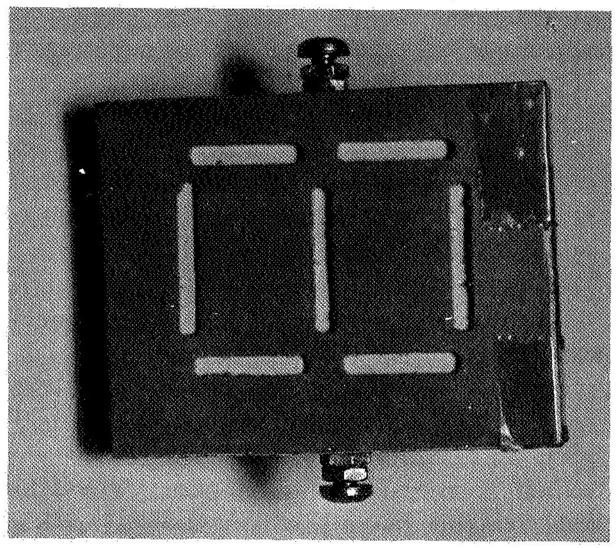

FIGURE 12.—Thermochromic display.

The model shown in figure 12 uses a Nichrome wire heater to keep the material in the "hot" color state. When room temperature gas is applied to the display module, the gas will cool the appropriate areas sufficiently to obtain a readable display. It is possible to eliminate the heater and use hot gas to effect the color conversion. Since experiments show that the obtainable conversion cycle time is increased when the hot-gas system is used, the technique employed in the model, wherein the thermochromic material is normally kept in the "hot" color state, is preferred.

Because cooled gas can be obtained fluidically, by either expansion of a gas through an orifice or with the use of a Hilsch tube, it is possible to obtain two gas supplies at two different temperatures. When two gas supplies are available, the cycle time is again increased considerably. Gas consumption will, of course, increase.

Figure 13 shows a cross section of the thermo-

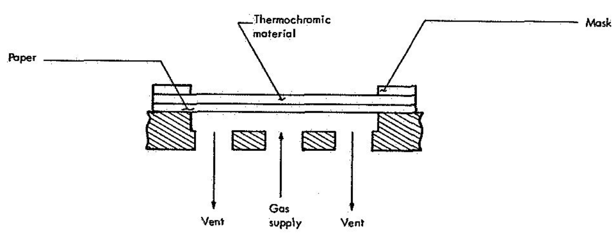

FIGURE 13.—Cross section of the thermochromic display models.

chromic display module. The thermochromic material is deposited on a paper-base film. The film is fastened to a suitable base material, such as aluminum or plastic. A recessed cavity underneath the thermochromic material is used to house the heating wire and is also used as a channel for the fluids.

This particular configuration of the display model was developed with two somewhat contradictory requirements in mind: low gas consumption and relatively severe updating requirements. The aimed-for updating cycle was once per second, which is extremely slow for electronic alphanumerical displays, but proved to be somewhat of a problem for fluidic thermochromic displays.

As far as the construction of the display unit is concerned, the best results were obtained with a hard-coated aluminum base block, where the hard coat serves as an insulation for the uninsulated Nichrome wire, a paper base, and a Mylar-base film mask. (See fig. 13.)

Considerable difference in time cycles can be obtained by changing some of the dimensions of internal channels, etc., since the obtainable minimum time intervals are largely dependent on the thermal capacity of the immediate environment of the thermochromic material.

The design goal set for the construction details was to obtain a miniature fluidic-thermochromic display module approximately 1¼ by 1 by ½ inches, which included a complete fluidic binary-to-decimal decoder.

Other Fluidic Display Devices

Some rather simple devices that can be used as fluidic display units are described in the following paragraphs. Exploratory work has been done on some of these ideas, but to date no known display actually uses the characteristics described.

The first series of relatively unexplored devices consists of small, hollow geometrical shapes that are built from resilient material. Application of air pressure to the hollow device will change the shape sufficiently to obtain visual effects. Figure 14(*a*) shows one of the hollow

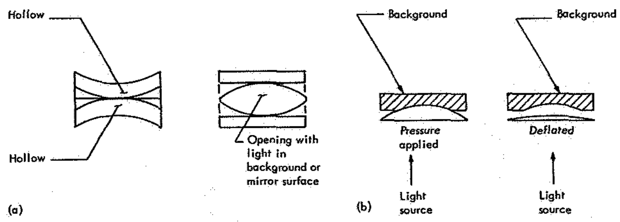

FIGURE 14.—Geometric shapes. (*a*) Hollow; (*b*) deflatable.

devices that may be constructed in such a way that the applied fluid pressure will expose a light or a mirror. Another device that works on a similar actuation principle is shown in figure 14(*b*). This device consists of a small, deflatable, transparent hemisphere. In the inflated position, the hemisphere is in close contact with a suitable contoured background. This close contact is lost in the deflated position.

The amount of light that will be returned from the device when used as a reflector will be a function of the refractive indices of the material of the hemisphere and the material with which it is in contact. Proper selection of the background material and the hemisphere material may result in a usable contrast between the inflated and deflated states.

Another device, which has only been explored briefly and insufficiently, is a fluidic analogy of a CRT. A two-dimensional fluid amplifier having four orthogonal control ports may be used in the neck of the tube to bend a fluid stream. The fluid stream could be made to impinge upon a sensitized plate. Thermochromic materials can be used as sensitizing agents, and temperature differences in the fluid can be created by means of expansion or, where more severe temperature cycles are necessary, by means of a Hilsch tube. The response time for a fluidic equivalent of the CRT will be slow compared with a regular CRT, but is certainly sufficient to provide displays for tracking tasks to be performed by man, or other similar slow-response systems.

An example of the capability of moving fluid stream display units in which techniques similar to the one described are used is shown in figure 15. These photographs show a stream-flow visualization technique used to investigate

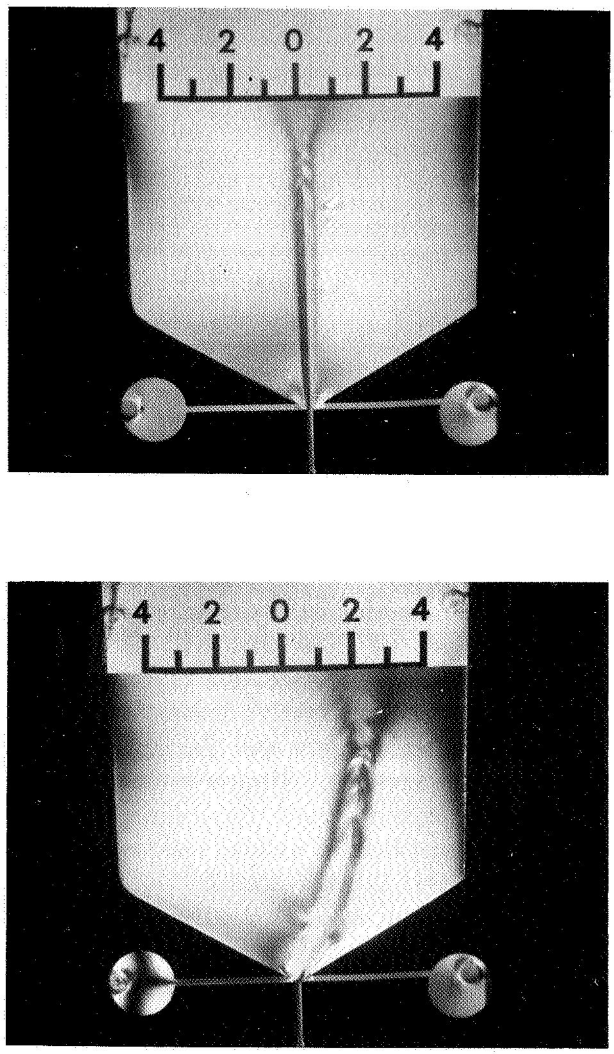

FIGURE 15.—Birefringent patterns.

fluidic phenomena. The stream of liquid was moved by applying a differential pressure across the two opposing control ports located perpendicular to the main-stream direction. The flow was made visible by using a special dye dissolved in water. The solution exhibits birefringent properties similar to those encountered in photoelastic materials. Pressure gradients in the flow patterns are shown as color gradients when polarizing filters are used to change the rays emitted from a light source placed in the back of the transparent assembly (ref. 9).

CONCLUSIONS

Many functions governing aeronautics and astronautics in general are related to properties of fluids. Propulsion and aerodynamic behavior of airborne vehicles are controlled by it. It is obvious that working directly with these properties in control systems may be advantageous from the standpoints of simplification and reliability. Recent developments of fluidic flight instrumentation, such as gyros, accelerometers, angle-of-attack sensors, and onboard computers, made it possible to obtain low-cost navigational equipment. Fluidic display systems will extend this low-cost capability to man-vehicle relationships in the future, not only in space application but also in the commercial and general aviation fields.

REFERENCES

1. KIRSHNER, JOSEPH M., ED.: Fluid Amplifiers. McGraw-Hill Book Co., Inc.
2. Fluidonics Division of Imperial Eastman Corporation: Fluidic System Design Guide.
3. DEPPERMAN, W. B.: Miniaturized High Speed Fluidic Circuits. Martin Marietta Corp., Orlando, Fla. Presented at the SAE Fluid Power Conference, Miami Beach, Fla., Apr. 1967.
4. DEPPERMAN, W. B.: Miniature Fluidic Digital Differential Analyzer. Final Rept. (ASTIA No. AD 484334), Martin Marietta Corp., Orlando, Fla.
5. C. A. Norgren Co.: Catalog. Littleton, Colo.
6. BELSTERLING, C. A.: A Fluidic Digital Computer With Non-Destructive Memory. Giannini Controls Corp. Presented at the SAE Fluidic Symposium, San Francisco, Calif., 1966.
7. BURKOWSKI, R. P.; FINNIE, L. N.; KORNBLAU, M.; GRAFSTEIN, D.; AND HILBORN, E.: Application of Thermochromic Materials in Display Devices. General Precision, Inc., Little Falls, N.J., and Electronics Research Center, NASA, Cambridge, Mass.
8. VAN DER HEYDEN, J.: Fluidic Decoder and Display Device. First Quarterly Rept., Martin Marietta Corp., Orlando, Fla., Contract NAS 12-532.
9. MCKENZIE, C. P.; AND DORSEY, W. P.: Transition to Turbulence. Martin Marietta Corp., Orlando, Fla. Presented at the SAE Aerospace Systems Conference, 1967.

MAGNETIC DISPLAY DEVICES

RICHARD C. SINNOTT
The Sinnott Corp., San Mateo, Calif.

INTRODUCTION

The role of a display device is to act as an interface between the man and the machine. In the following discussion the display device will be considered as a separate entity that must relate to both the needs of the man as well as the information processor (machine).

In accordance with this definition, the ideal display device may be described as follows:

Resolution	500 to 5000 line (or 0.25 to 25 million bits).
Contrast	Constant, as in the printed page.
Monochrome	Yes.
Color	Yes.
Memory	Nonvolatile.
Address time	Less than 100 nsec (preferably 1 to 10) per dot.
Image formation time.	Less than 50 msec.
Cost	Potentially low (consumer use).
Life	Greater than 5000 hr (20 000 or more ideal).
Drive	Matches well with solid-state devices. (Preferably current- rather than voltage-sensitive.)
Hardcopy	Upon command only. (Image is independent of hardcopy production.)
Size	From tens of feet to tenths of inches per side in a flat geometry.
Mechanical characteristics.	No moving parts (except when making hardcopy), no vacuum, and fabricated of impact- and temperature-resistant material.
Interface	Bilateral or bidirectional (machine-display-man or man-display-machine).
Power	1 to 10 W ideal, 100 W acceptable.

A device meeting these qualifications would generate images approaching photographic resolution appearing as a printed page with color. The ramifications of hardcopy upon command, as well as a bilateral interface, will be left to the imagination of the reader.

The closest approximation to these idealized characteristics is accomplished by the CRT. The areas in which a CRT fails to meet the qualifications are not within the scope of this paper; however, an excellent discussion of the topic is given in a recent publication (ref. 1).

MAGNETOMECHANICAL DISPLAYS

Many of the earliest numeric and alphanumeric displays were of a magnetomechanical nature and to this day are popular in many applications. Most of them are well known and are of interest here only as a basis of comparison to newer developments.

Probably one of the oldest devices is the d'Arsonval moving coil galvanometer. The advantages of this display device include sensitivity, constant contrast, analog response, low cost, and easily attained accuracies in the 1- to 10-percent region. However, in addition to being difficult to read more accurately than, say, 3 percent, its disadvantages are that it is both fragile and large, and has a maximum practical accuracy on the order of one-quarter percent. This type of indicator can possess memory, but usually does not.

A device of both historical and contemporary interest is the ratcheting drum display. The method of operation typically involves a magnetically actuated pawl that ratchets a drum with printed numbers or letters. Displays of this nature allow geometrics for in-line readout and were, therefore, a step forward. They can be linked mechanically and behave as a counter, or be addressed on a character-by-character basis for more rapid readout. This class of display will generate hardcopy (print) by appropriate modification of the drums.

Memory, constant contrast, optimum character shape, and planar readout are the inherent advantageous characteristics of the ratcheting drum display. The disadvantages are mechanical wear, bulk, and a long access time, particularly if the drum contains many characters.

These displays are found in wide and often very clever variations, but have in common a drum, tape, etc., containing printed characters each in a separate position that must be addressed for display.

In the ratcheting drum display it should also be noted that if the drum, etc., were to slip, or not properly respond to a write command, the viewer would have no way of knowing that a malfunction had occurred, unless the device has some kind of position feedback to the driving source for error check. Many of the mechanical displays have this feedback, or at least return to zero as a test before displaying the proper number. The mechanical display with feedback is more complicated, but with complete feedback the driving electronics now has a memory it can use in its logic. This is often an important advantage.

Another class of electromechanical displays circumvents the problem of the specific address for each character by using bar or dot matrices. A sphere [1] (or bar) is permanently magnetized and placed in proximity to an electromagnet. Both sides of the sphere are colored: one side is colored to blend with the display panel and the opposite side to contrast with the panel (fig. 1). The attraction of the permanent magnet to the core of the nonenergized electromagnet maintains the position of the sphere and gives the device memory. When the electromagnet is energized, the sphere rotates or remains in position, depending on the polarity of the electromagnetic field.

[1] A device of this class is marketed by Ferranti, Ltd., England.

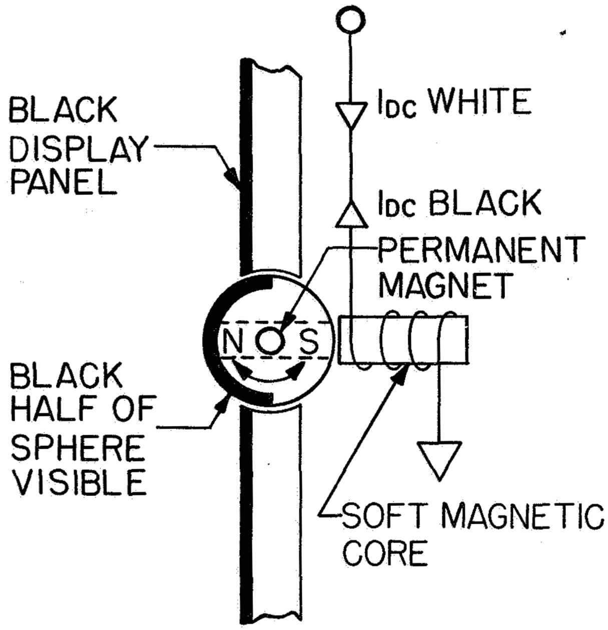

FIGURE 1.—Schematic cross section of an electromechanical dot matrix display.

The advantages of such displays are memory, constant contrast, increased speed (relative to drum displays), and, as mentioned previously, simplified mechanical address. The approximation to character shape by these matrices is a disadvantage, although a five-by-seven-dot matrix generates a reasonably readable character.

In a bar or dot matrix display, the viewer usually knows if a malfunction has occurred after very few characters are displayed.

Magneto-Optic Displays

A unique and different method of creating a display utilizing the formation of "stripe" domains in thin NiFe films in conjunction with a Bitter solution has been reported (refs. 2 and 3).

Early studies of magnetic domain structure resulted in the development of a technique to render domains visible by the application of a colloidal suspension of ferromagnetic particles. When placed on a magnetic surface, the particles in this solution will agglomerate on magnetic gradients occurring along domain boundaries. The use of sufficiently small particles will resolve boundary separations in the submicron region, thus causing the detailed magnetic state of the film to become visible.

A characteristic of negative magnetostrictive, isotropic tensile stress NiFe films having a thickness of 10 000 to 40 000 Å is the formation of domains of a very regular nature. These domains are long and slender with parallel edges, the long dimension alined with the applied field.

Combining a film of this nature with the Bitter solution results in the agglomeration of particles along the edges of the domains in a regular, parallel pattern thus forming an optical diffraction grating (fig. 2). By illuminating this grating appropriately, the zero-order reflection is discarded and the first-order diffraction is made visible to the viewer.

The characteristics of the magnetic film are variable enough for the domain size to be controlled by film thickness, modifications of the plating solution, etc. Selective plating of two or three different strips in a repeating pattern

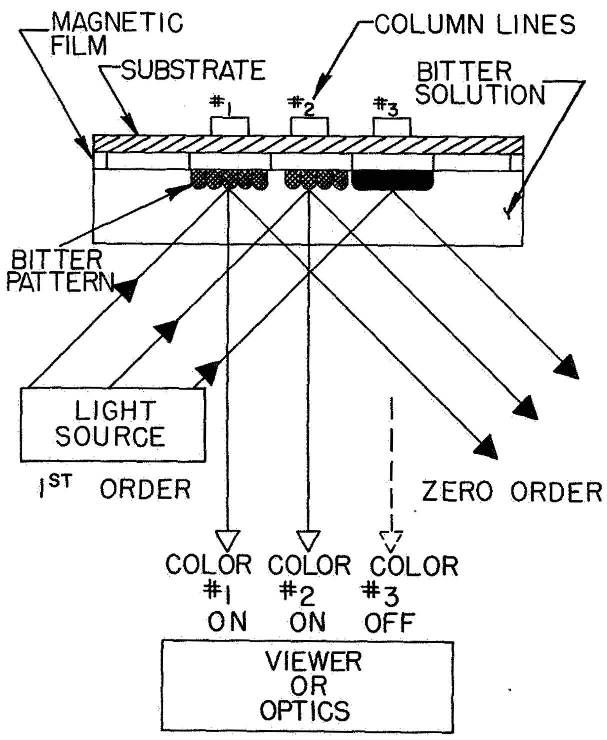

FIGURE 2.—Simplified schematic of magneto-optical display.

results in the formation of gratings with different periods thus diffracting two (or three) colors toward the viewer. Once the domains have formed they will remain, unless re-formed or moved by application of a magnetic field. This display, therefore, has nonvolatile memory.

The magnetic film may be magnetized (addressed) by any method, with sufficient field to modify the domains. Such methods include a hand-held permanent magnet ("magnetic pencil"), a recording head, or an x–y matrix imparts field to the film via the coupling between the conductor and the film. When dc is used as the driving current, a field of 1.4 times that around each conductor is realized at the intersections. When the current in one conductor is reversed, the vector of the field rotates 90°, which causes the grating to rotate and the image to disappear. Although the pulsed dc technique will switch the film in 20 nanoseconds, sufficient field exists around the conductors to cause partial rotation of the domains at positions other than the intersection. This crosstalk problem reduces the quality of the display.

A clever technique is used to avoid this problem (fig. 3) by driving the row line (write) with a dc pulse and the column line with a slightly shorter burst of ac. The interaction of the two fields at the intersection of interest apparently increases the sensitivity of the film to the dc field allowing a reduction of this current (relative to dc-dc switching) and thereby reduces the effect of the row line current at other places on the film. Apparently the column line ac burst also does not influence the film in other areas. Address time with ac-dc switching is 200 nanoseconds, the limit of experiments performed to date.

Although the film may be addressed rapidly, it takes somewhat longer for the Bitter particles to redistribute themselves on the displaced domain boundaries. The time for this to occur is considerably longer than the switching time, because the particles move by Brownian motion. Although this display time is unknown, it is certainly less than 1 second and very possibly in the 100-millisecond region.

Resolution of this system has been demonstrated at 100 lines per inch (single color) with

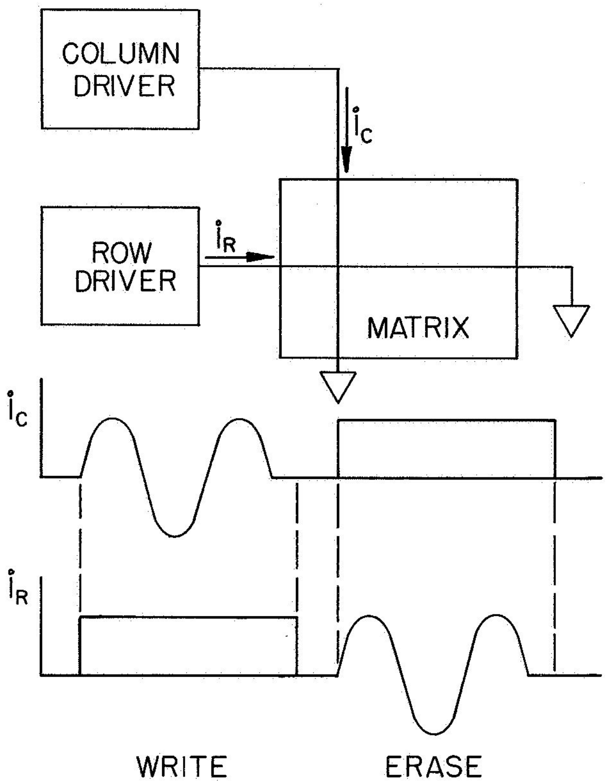

FIGURE 3.—Write and erase drive current for magneto-optical display.

contrasts varying from 5 to 50:1, brightness of 10 000 ft-L with tungsten illumination, and write currents of about 5 amperes. Two-color displays have also been constructed with 100-horizontal and 50-vertical-line-per-inch resolution.

The problem of color change as the relative position of the observer and display surface changes will probably be solved by some optical techniques for color displays. A monochrome display is formed if a lens and diffuse, rear-viewing screen is placed between the observer and the device (also solves the position problem).

The ability of the observer to alter the magnetic state of an element (or elements) via a magnetic probe, and the present capability of the device to read its state optically, indicates some very exciting possibilities in man-machine and machine-man communications. Use of the inherent magnetic memory characteristics to give electronic readout of the state of any element would be a very worthwhile refinement.

Although this device is in an early stage of development, a potential ability to fulfill an astonishing number of the idealized objectives stated earlier has been demonstrated.

Another recent development in the field of magnetic displays is the SIMAD principle (ref. 4). This method utilizes small magnetic particles as a display medium by propelling them from a chamber below the display surface. Magnets (dc) of appropriate design behind this surface are energized and attract some of the display particles from the "cloud" formed by the particle gun, onto the display surface, thus forming a planar, constant, high-contrast display (fig. 4).

The display particles, in the version described here, are permanently magnetized ferrites between one and two one-thousandths of an inch in diameter. The display surface is thin enough that the attraction between these particles and a nonenergized display magnet is sufficient to keep them in position, giving the device nonvolatile memory.

Erasure of an image is accomplished by an alternating magnetic field roughly perpendicular to the display surface and generated by the erase coil. This field is of a magnitude great enough to overcome the attraction between the display particles and a deenergized display magnet, but not an energized one (fig. 5).

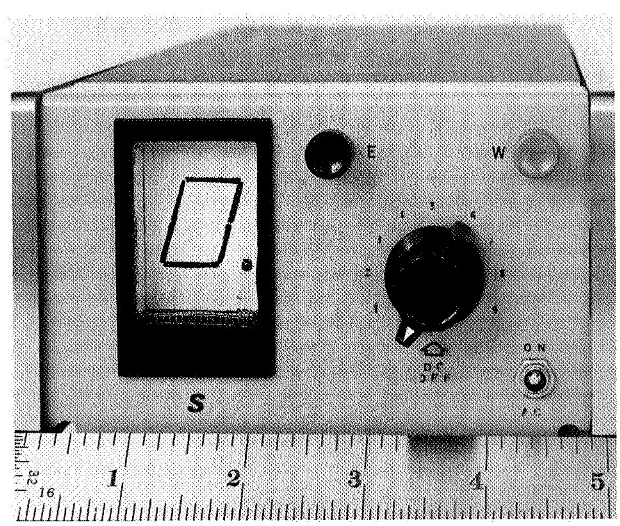

FIGURE 4.—Magnetic display unit employing SIMAD principle.

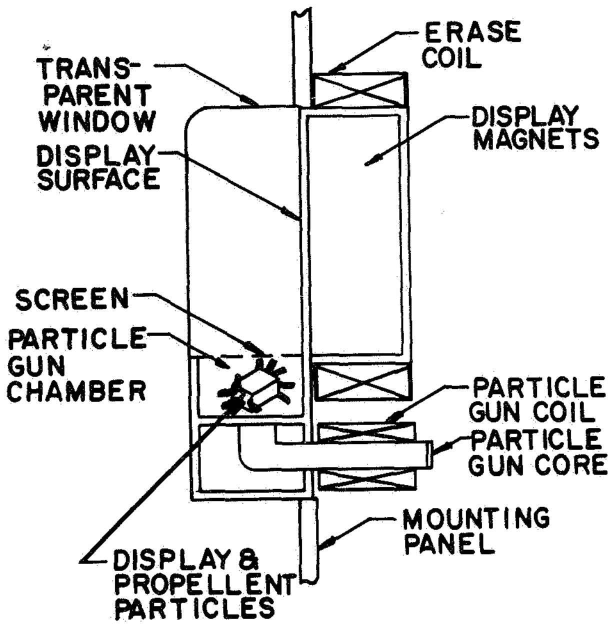

FIGURE 5.—Image erasure.

The particle gun is a transducer that imparts the energy in a magnetic field into kinetic energy of the display particles. This is accomplished by utilizing the interaction between a permanent magnet and a high gradient alternating field. Disposed in the particle-gun chamber are several cubical permanent magnets magnetized as bar magnets. When the particle-gun coil is energized with alternating current, the field generated penetrates the particle-gun chamber, interacting with the cubes or propellent particles, causing them to agitate violently. The screen separating the particle-gun chamber from the display area prevents the propellent particles from entering, but allows passage of the display particles. While in the particle-gun chamber the display particles are attached to the propellent particles and thus gain considerable kinetic energy. When the propellent particles strike the screen and rapidly decelerate, the display particles dislodge themselves and "fly" into the display chamber with high velocity.

The particle-gun structure can be extended horizontally with ease, and it appears feasible to construct guns capable of propelling particles 12 to 15 inches vertically. If the device is used to display characters, one particle gun can be used with many matrices, or each character may have its own gun.

The interaction of the particles, display surface, and erase field in the present device indicates line resolution is limited only by the display magnet geometry. Future development of the magnet designs utilizing "bulk" manufacturing techniques will hopefully yield 30 to 100 lines per inch.

The display logic may be reversed by masking the display surface with a material of the same color and texture as the display particles in a dot or bar matrix format. This mask would be alined over the display magnets so that when all the particles are removed from the display surface, a series of white dots (or bars) would be visible against a black background. When the display magnets are energized and the particle gun is activated, the particles would fill areas where the information was not wanted, thus creating a white image on a black background.

Image formation time with the present device is 100 milliseconds, and the display magnets require 35 milliamperes at 4 volts dc. The erase coil and particle gun operate at 60 cycles, 12 volts and 24 volts, respectively. The display is planar, has nonvolatile memory, and is of constant contrast. The basic nature of the device should allow designs that would survive in environments where shock, vibration, and temperature are severe; and since the structure can be completely encapsulated (with the exception of the display and particle-gun chambers), no vacuums are required, and the particles already use vibratory motion to operate the display.

A Combination of Magnetic and Electrostatic Techniques

The display particles used in the SIMAD device are insulators and, as such, may be charged triboelectrically. When so charged they behave in much the same manner as toners, which are used in electrostatic devices. The magnetic characteristics of the particle remain unchanged and thus allow the particles to interact with both electric and magnetic fields.

A CRT has been developed (ref. 5) that writes a negative charge on a thin dielectric membrane. After charging, a toner is applied to the opposite side of the film, thus forming a high-resolution, high-contrast display with memory. With an appropriate lens system, such a display may also be projected.

In the embodiment of the tube mentioned above (ref. 5), the toner is applied by slow, cumbersome mechanical means. By adding a SIMAD particle gun (fig. 6) and an "erase" coil, an interesting combination of electrostatic and magnetic devices results.

The display particles will be triboelectrically charged positive if all the elements of the particle-gun chambers are of a lower dielectric constant. The vibratory nature of the particle gun will insure complete charging of the particles, causing them to adhere to the charged areas of the display membrane.

The alternating magnetic field generated by the erase coil is adjusted to a level where it will remove particles adhering to the display membrane in areas where stray charge has accumulated, but will not remove particles attracted by the written charge. The slight vibratory motion imparted to the display particles by the erase magnetic field tends to sharpen the displayed image.

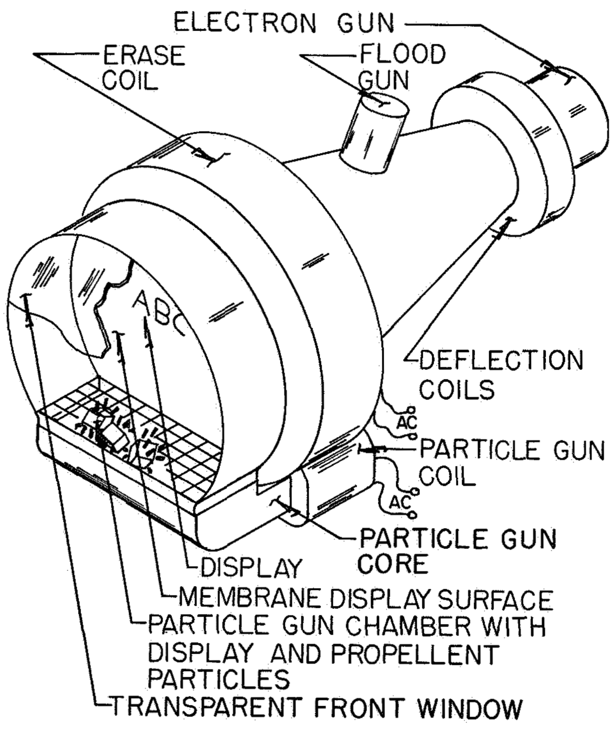

Figure 6.—Electrostatic image storage device combined with the SIMAD particle gun.

The marriage of the two techniques has not been tested in the embodiment shown in figure 6, but tests have been conducted with the display illustrated in figure 4 by applying charge to the Lucite viewing window. The particles remained in the same position for 10 hours after application by the particle gun.

The availability of a microscopic (or macroscopic) particle with capabilities of interacting with both electric and magnetic fields adds another dimension to display device technology, and hopefully will suggest many applications to the reader.

Ferromagnetography

It is possible to imprint magnetic fields on a magnetic medium at high speed and high resolution. With the application of an appropriate toner, a visible image is produced that can be transferred to paper, or other suitable media, by pressure, heat, etc.

It is within the state of the art in the tape-recording industry to record wavelengths of 50 to 100 microinches at tape-head velocities of 1500 inches per second. Bit densities between 1000 and 2000 per inch are accomplished every day with noncontact techniques in computer disk and drum memories.

Work reported 15 years ago (refs. 6 and 7) more than proved the feasibility of the method for character and facsimile copy generation. Work described in a later report (ref. 8) demonstrated more sophisticated techniques, utilizing higher resolution magnetic heads in facsimile recording and bar matrix recording directly on a magnetic medium (fig. 7).

Recording of characters with a wire-bar matrix is accomplished by pulsing a 3-mil-diameter wire to a peak current of 25 amperes with a very short pulse. The field around the wire (700 oersteds peak) is sufficient to magnetize the medium. Subsequent application of magnetic toner developed the latent image.

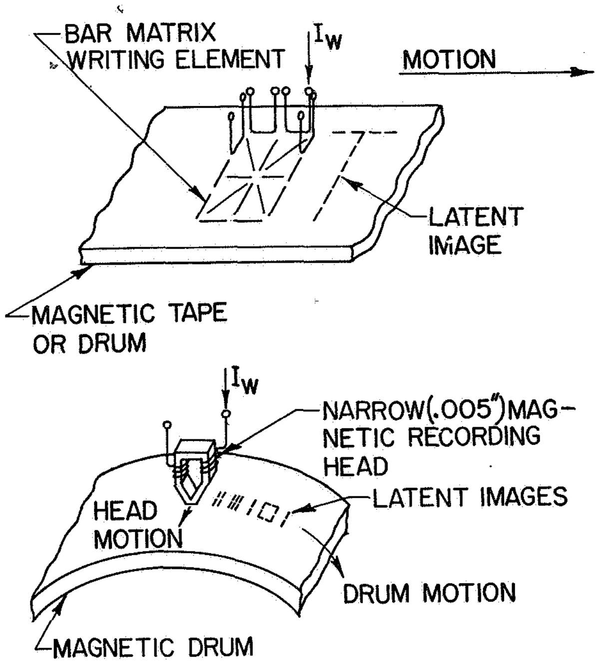

FIGURE 7.—Character and facsimile copy generation on magnetic medium. Top: Character writing; bottom: facsimile writing.

The same basic technique is used in the facsimile recording method (fig. 8). The magnetic head is placed in contact with a drum coated with suitable magnetic material (Ni-Co in this case) and the drum and material to be reproduced are rotated in synchronism. The head scans the surface of the drum appropriately magnetizing it. In the system described, the head width is 0.005 inch, and the shortest wavelength printed was 0.005 inch, or 40 000 elements per square inch. Present technology would easily allow packing densities of 1 million elements per square inch.

Magnetographic and electrostatic printing techniques are comparable in speed, resolution, etc. The magnetic technique is superior from the standpoint of memory; once a drum is magnetized, toner may be applied many times without degradation in resolution, making large-scale multiple copying practical. Electrostatic printing has dominated, however, because the writing elements can be addressed with a CRT and consist of a simple wire rather than a more complicated magnetic structure.

291-351 O—68——7

CONCLUSIONS

One of the oldest practical forms of electrical readout is a magnetic device—the galvanometer. The magnetically actuated drum display, capable not only of readout but hardcopy production, was next developed. Progress of magnetic display devices after this period has trailed far behind other technologies in the field.

In terms of a fit with the so-called idealized display, not one comes close. However, many of the necessary elements are present—constant contrast, rapid switching, hardcopy capability, high resolution, color capability, bilateral interface, and memory.

The advent of solid-state devices, bulk processing techniques, and the stringent demands of computer displays may well cause a resurgence of interest in the field. Magnetic devices are compatible with the solid-state devices (current sensitive); are fabricated of materials well suited to mass-production methods of printed circuitry, chemical milling etc.; and they have memory.

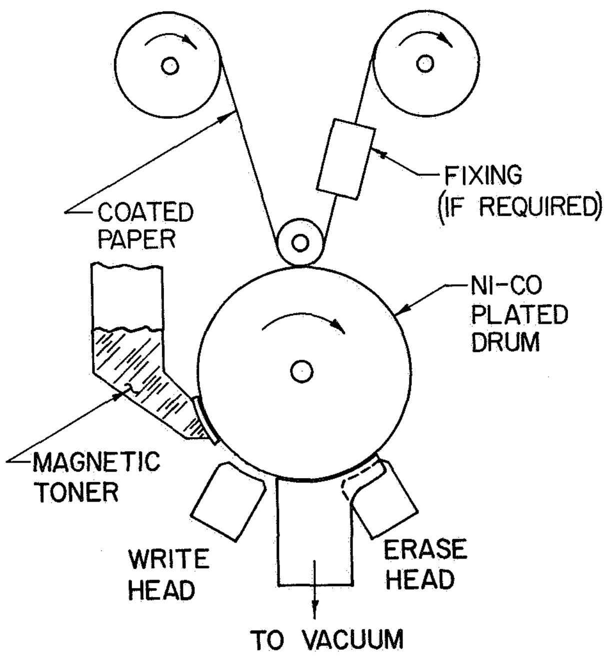

FIGURE 8.—Hardcopy printer utilizing magnetic techniques.

REFERENCES

1. HOBBS, L. C.: Display Applications and Technology. Proc. IEEE, vol. 54, no. 12, Dec. 1966, pp. 1870–1884.
2. EIDE, J. E.; GLUSICK, R. E.; SOMERS, L. E.; AND WINDSOR, H. F.: Magneto-Optical Display Panel With Memory. General Electric Co., Syracuse, N.Y., Tech. Rept. Ecom–0144–1 (ASTIA No. AD 629586), Jan. 1966, 1st Quarterly Rept. prepared for U.S. Army Electronics Command, Fort Monmouth, N.J., Contract DA–28–043–AMC–01442(E).
3. SOMERS, L. E.: Magneto-Optic Display. General Electric Co., Syracuse, N.Y., Nov. 1966, prepared under sponsorship of U.S. Army Electronics Command, Fort Monmouth, N.J., Contract DA–28–043–AMC–01442(E).
4. SINNOTT, RICHARD S.: A Novel Application of Magnetic Technology to Electronic Display Devices. Proc. of the 8th Natl. Symp. of the Society for Information Display, San Francisco, Calif., 1967, pp. 15–23.
5. CLAUER, C. K., ET AL.: Electrostatic Printing Device. U.S. Patent 3,109,062, issued Oct. 29, 1963.
6. ATKINSON, R. B.; AND ELLIS, S. G.: Ferrography. Journal of the Franklin Inst., Philadelphia, Pa., Nov. 1951, p. 373.
7. HANNA, J. P.: Ferromagnetography. Proc. of the Fifth Annual Technical Meeting, Tech. Assoc. of the Graphic Arts, Washington, D.C., Apr. 27–29, 1953, pp. 22–27.
8. BEGUN, S. J.: Theory of Magnetography. IRE Natl. Conv. Record, pt. 5, 1958, pp. 190–197.

BIBLIOGRAPHY

ATKINSON, R. B.; AND ELLIS, S. G.: Ferrography. Journal of the Franklin Inst., vol. 252, no. 5, 1951.

HANNA, JOHN P.: Ferromagnetography. General Electric Laboratories.

SINNOTT, R. C.: A Novel Application of Magnetic Technology to Electronic Display Devices. 8th Natl. Symp., Society for Information Display, May 1967.

SOMERS, LEWIS E.: Final Rept., Contract DA–28–043–AMC–01442(E), U.S. Army Electronics Command.

TRIMBLE, L. S.: U.S. Patent 3,320,523, May 1967.

N68-22312

ELECTROSTATIC DISPLAYS

PHILIP RICE
Stanford Research Institute, Menlo Park, Calif.

Electrostatic forces can be used in at least three ways to generate a volatile display. The feasibility of all three methods has been established, but displays based on these electrostatic effects are not yet in widespread use.

The first, and perhaps simplest, effect that has implications in the display field is the tendency for fine particles to collect in areas of high static-charge concentration. This effect is the basic principle underlying the field of electrostatic printing.

A pattern of charges can be established on a surface in several ways, among them direct charge deposition from an array of electrodes and selective discharging of a photoconductive surface by light images. Once the pattern has been established, pigmented particles, either dry or suspended in a liquid, are brought in contact with the charge pattern. They will be attracted to those areas having a high concentration of surface charge of the proper polarity and will cling tenaciously to the surface. In the normal printing application, the particles are fixed permanently to the surface, usually by a drying or heating sequence.

Electrostatic printing is in widespread use today mainly for document reproduction and for printing address labels for magazines. The familiar office copier makes use of the light-addressed photoconductive surface mentioned previously. An image of the original document is projected onto a precharged selenium drum, draining charge away in the illuminated areas. A mixture of pigmented fine particles and coarse plastic beads is cascaded over the drum. The particles cling to the drum in regions of high field gradient. As a final step, the particles are transferred to a sheet of paper by contact (aided by an electric field).

The label printer relies on direct charge deposition from the ends of fine wires onto a dielectric surface. A special CRT performs the function of a rapid switch that permits selection of the desired wires from among a large array.

The faceplate of the tube consists of a close-packed mass of very fine wires running through the faceplate parallel to the electron beam (fig. 1). As the beam is scanned across the inside ends of these wires, they charge rapidly toward the potential of the cathode of the tube. The outside ends of the wires contact the dielectric surface, which is usually a coated paper. Charge is transferred to the paper, and the pattern is made visible in a later step by the application of the pigmented particles. Unlike the case of the office copier, the pattern to be printed

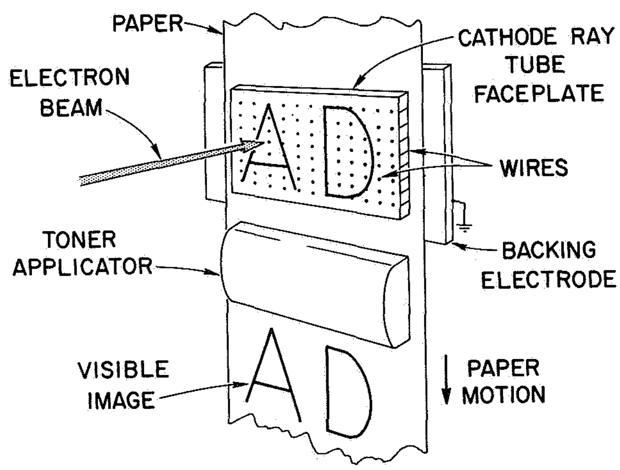

FIGURE 1.—Cathode-ray tube electrostatic display.

need never exist in hardcopy form. For the label-printing application, appropriate beam deflection signals are produced by a separate character generator under the direction of a central computer.

The techniques described above produce hardcopy, which is a form of display, although a rather specialized one. The fact that the hardcopy can be generated at a very rapid rate suggests that this form of "printing" could have application in the display field. The CRT used in the label printer is capable of recording standard television pictures, frame by frame, in real time.

The recorded information can be displayed by any standard projector if the printing is done on a transparent tape rather than on paper. Because the particles are readily removed before the fixing operation, it is possible to design an endless belt recorder-projector.

There are a number of other ways of recording and displaying information by using various combinations of charge deposition techniques and "inking" or "toning" steps. Some of these techniques allow the reproduction of color images (refs. 1, 2, and 3). An important feature of these systems is their very high sensitivity; only a minute amount of charge (about 10^{-9} C/cm^2) is needed to make a visible mark and, in this respect, these systems rival photography. Also, under proper conditions, the visible mark can be made in about 20 nanoseconds.

Another application of electrostatic effects to the display field is the use of static forces to deform the surface of a transparent film. One of the first uses of this principle was in a projection television system. The system used a special CRT containing an electron gun directed at a conducting faceplate that was covered with a thin film of oil. The tube was constructed so that a beam of light could be projected through the faceplate onto a screen. A schlieren mask and a projection lens were placed between the faceplate and the screen. The beam was scanned across the faceplate in a normal TV raster, using video signals to intensity-modulate the beam. In those areas where electrons struck the oil film, the electrostatic forces across the oil film caused the film thickness to decrease. As the beam scanned, the video information appeared as variations in thickness of the film. Because the film acts as a phase grating, the ripples thus produced cause the light passing through the film to be diffracted. The schlieren mask blocks all but diffracted light. Smooth areas on the oil film appear black on the screen, and the amount of light corresponding to the rippled areas of the film is a function of the depth of the ripples. As the charge leaks off the film, surface tension restores its original smooth surface.

This process was used extensively for theater television projection. One major problem has been contamination of the cathode of the electron gun by vapors produced by bombardment of the film.

A number of variations on this same order have been developed in recent years. Thermoplastic recording (ref. 4) allows permanent recordings to be made by "freezing" the surface deformations in a thermoplastic film. Charge is deposited by a beam on the surface of a plastic film while the surface is cold. It is then heated, allowing the electrostatic forces to deform the film. When cooled, the ripples are fixed on the surface, and the image can then be projected in the same manner as in the oil-film tube (fig. 2). The film can be reheated, allowing the surface tension to smooth the surface again. Very high resolution can be obtained with this system permitting frequencies in the range of 100 MHz to be recorded. Because the film is exposed in the vacuum chamber containing the electron gun, some of the same cathode poisoning problems mentioned previously are present.

A CRT with a wire-embedded faceplate, of

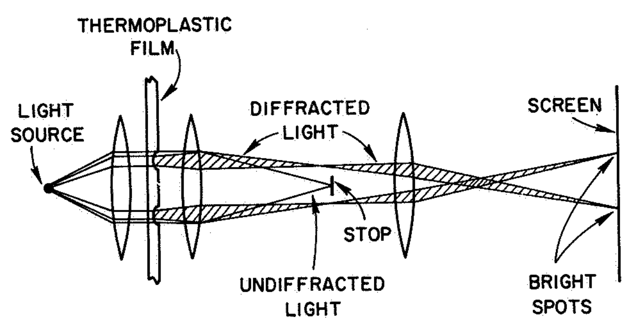

FIGURE 2.—Schlieren optical system for thermoplastic film.

the kind described earlier, can serve as a source of electrons for thermoplastic recording (ref. 5). The recording medium is no longer required to be contained in the vacuum chamber with the electron gun. Another "in air" thermoplastic recording system consists of an *xy* array of grids producing a discharge at the intersection of selected wires (ref. 6). In this system, unlike the others described, the schlieren optical method is not used. The thickness variations are made visible by making the deformable medium serve as one face of a prism (fig. 3). Light is incident on that face at such an angle as to suffer total internal reflection. The reflected light is then projected onto a screen. In the absence of any surface deformation, the screen is uniformly illuminated. Voltage is applied to selected *x* and *y* wires causing a discharge to take place that deposits charge in the area of the intersections. When heated, the film contracts at those points. In contracting, the film distorts the prism surface so that the conditions for total internal reflection are no longer satisfied at those points. The corresponding areas on the screen go dark. Grids with spatial frequencies as high as 500 per inch have been made and tested.

Another electrostatic-thermoplastic process has been developed that permits the charge pattern to be controlled by light (ref. 7). A transparent photoconductor and a thermoplastic film are used in combination with a transparent conducting layer. The photoconductor receives a uniform charge while in the dark. When the photoconductor is illuminated, a charge pattern corresponding to the optical image is established producing stress across the thermoplastic layer. Heating develops the thickness variations in the thermoplastic film as before.

All the foregoing electrostatic display techniques produce images that are best viewed by projection—either conventional or schlieren. One electrostatic display should be mentioned in which the images can be seen by direct viewing under ambient light. An *xy* array of lightweight plates hinged about their horizontal axes in such a way as to present either a black or white appearance, depending on their state (fig. 4). For example, the plates may be painted black on the side facing the observer and, when open, reveal a white background, or the opposite condition may be desired. In conjunction with the *xy* array of plates, there is a corresponding *xy* array of address wires. The system is designed so that when voltage is applied to a selected pair of *x* and *y* wires, the plate at the intersection swings upward, reversing the appearance of that cell from black to white, or vice versa. The plate remains in this condition until voltage is applied again, causing the cell to return to its original appearance. Thus, the cells have memory and remain in one state until directed to return to the other state. The forces responsible for the movement of the plates are the same as those that cause the plates of the familiar goldleaf electroscope to diverge.

Writing time per element varies from milliseconds, for plates one-sixteenth inch on a side, to tenths of seconds for 5-inch-square elements. The entire array is enclosed in a weatherproof box. Prototype displays have been in operation for some time in exposed locations.

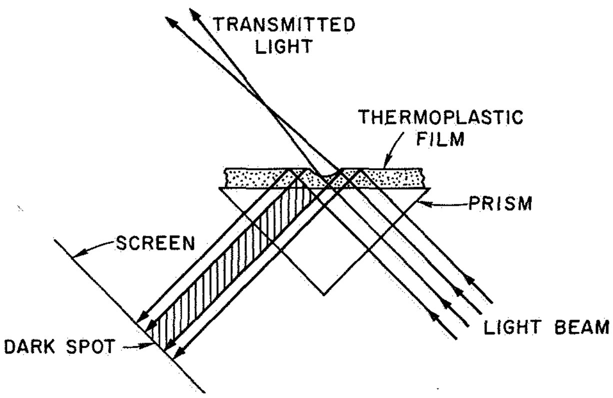

FIGURE 3.—Total internal reflection system for thermoplastic film.

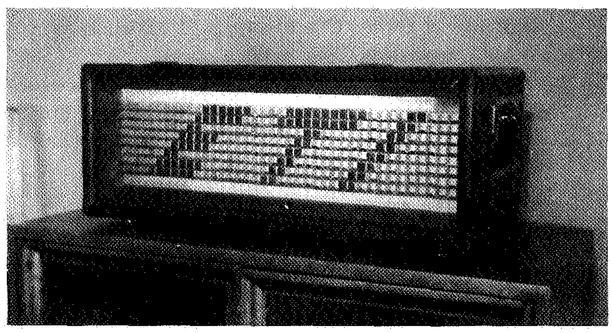

FIGURE 4.—*xy* array of lightweight plates for a direct viewing display.

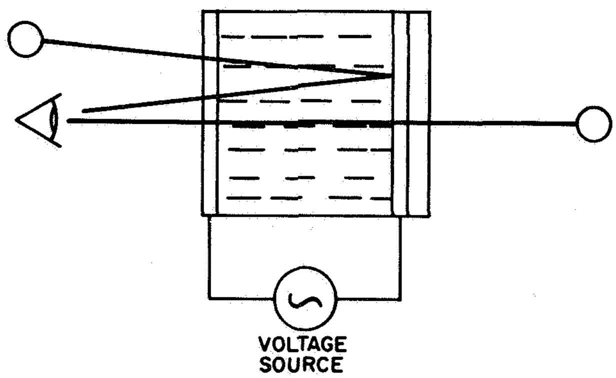

FIGURE 5.—Energized electrostatic display (oriented particle state).

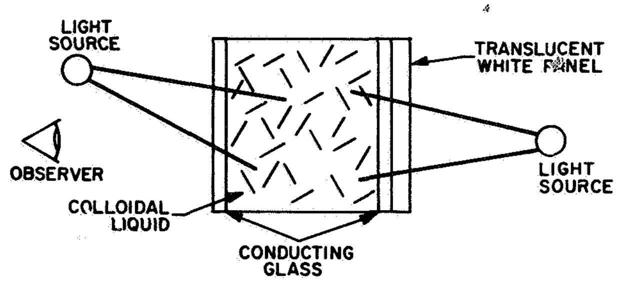

FIGURE 6.—Deenergized electrostatic display (random particle state).

A variation of the reflective display described uses the hinged plates as mechanical light valves. In this mode of operation, the display is lighted from the rear by a uniform, distributed light source. A given cell will either pass light or not, depending on its state.

The final example of electrostatic display makes use of the fact that elongated particles colloidally suspended in a nonconducting fluid will tend to aline their long axis parallel to the direction of an applied electric field (fig. 5). In the absence of the field (fig. 6), Brownian motion causes them to assume a random orientation (ref. 8). Because the particles are opaque, the amount of light transmitted through the cell will be greater in the oriented state than in the random state. For particles with a side-to-end area ratio of 5 to 1, transmission varies by a factor of 3. The particles can be oriented from the random state in about 100 microseconds.

The display can be operated either to control the amount of light passing through to the observer (transmission mode) or as a reflective display using only ambient light.

Phenomena that tend to disrupt the operation of the display are electrophoretic deposition of the particles on the conducting electrodes forming the walls of the cell, coagulation due to the loss of the dispersive forces between particles that are present in a normal colloidal suspension, and particle settling caused by gravity. Electrophoretic deposition is minimized by using ac rather than dc fields to orient the particles. Coagulation is prevented by restricting the magnitude of the applied field. The terminal settling velocity of a sphere in a fluid is proportional to the square of the radius of the sphere. The settling time of the particles in the display can be made to last many days if the particle size is limited to the micron range.

REFERENCES

1. DISSAUER, J. H.; AND CLARK, H. E.: Xerography and Related Processes. Focal Press, 1965.
2. WEINER, J.; AND ROTH, L.: Electrostatic Printing. Institute of Paper Chemistry Bibliographic Series No. 212, Appleton, Wis., 1964.
3. RICE, P.: Electrostatic Printing. *In:* Discovery (England), vol. 25, no. 7, July 1964, pp. 30–35.
4. GLENN, W. E.; AND WOLFE, J. E.: Thermoplastic Recording. *In:* International Science and Technology, June 1962, pp. 28–35.
5. POOLE, H. H.: Fundamentals of Display Systems, Spartan, Washington, D.C., 1966.
6. ORIMENKO, A. P.: A Matrix Controlled Display. 4th Natl. Symp. on Information Display, Society for Information Display, 1964, pp. 113–125.
7. GAYNOR, J.; AND AFTERGUT, S.: Photoplastic Recording. *In:* Photographic Science and Engineering, vol. 7, July to August 1963, pp. 209–213.
8. Simmonds Reflective Display, Simmonds Precision Products, Inc., Tarrytown, N.Y.

N68-22313

A SURVEY OF LASER DISPLAY

D. W. Kennedy, C. R. Grauling, A. J. Devaney, and R. D. Wright
Computer and Display Branch, Control Laboratory, Electronics Research Center, NASA, Cambridge, Mass.

Introduction

Present laser display systems may be classified into one of two distinct categories: those systems based on moving (scanned) laser beams, and those systems utilizing the concepts of holography. Of these two classes of laser display, those devices of the former category reflect the more advanced state of development. At this time, there are at least three complete display systems utilizing a scanned laser beam. This advanced state of development is due, in part, to the fact that flying spot display techniques are well known from industrial CRT experience.

The field of holography, on the other hand, is relatively new and not as well developed. Holographic displays offer more potential for realistic three-dimensional imaging than any other display technique developed to date. In addition to display applications, holography is being found to have uses in many other fields including pattern recognition, data processing, and stress analysis.

Scanning Laser Displays

Existing scanning laser displays use basic techniques from the CRT field, but a laser beam and appropriate scanning devices are substituted for the CRT and its deflection apparatus. The major problems in development are in refining techniques to manipulate the laser beam and to modulate its intensity.

Deflection Techniques

There are presently three basic methods of deflecting a laser beam:

(1) Mechanically, by deflecting a mirror with a motor or piezoelectric crystal

(2) Refractively, by changing the index of refraction of a transmitting fluid or crystal through an electro-optic effect

(3) Diffractively, by rapidly altering the distance between grating lines, usually by means of ultrasonic vibrations in a medium.

The most obvious form of mechanical beam deflector is a motor-driven rotating mirror. A system in which mirrors are attached to the faces of a polygon-shaped mount, which is rotated about its center by a motor, is easily capable of producing scan *rates* sufficient for high-resolution television horizontal scanning. For a 1000-line raster, scanned 30 times per second, a horizontal scan rate of 30 000 scans per second is required; however, a serious problem arises with the stability of such a system. Even when precise synchronous motors are used to drive a rotating mirror, slight speed variations make synchronism with autonomous signals very difficult (ref. 1).

A rotating-mirror deflector has been used successfully as a horizontal scanner for a laser digital display system (ref. 2) by synchronizing the signal flow with the rotating mirror, rather than the reverse. This system uses 32 mirrors arranged around the edges of a 32-faced polygonal prism, with each face tilted at a slight angle with respect to its neighbors. Each mirror produces a horizontal scan as the angle between the mirror and beam is changed by rotation. The slight tilt from mirror to mirror produces the vertical scan. The mirror and information storage device are driven by the same motor and are rigidly coupled.

A galvanometer and mirror system has been designed to produce a relatively low-frequency sawtooth scan (ref. 3). This device produces a reasonably faithful sawtooth scan at 60 Hz with a low dc voltage level for the scan, followed by a large restoring voltage, and then a large, but opposite, polarity stopping pulse.

An interesting method for scanning a beam at high frequency utilizes torsionally resonant crystalline fibers (ref. 4). This system uses a large-diameter single crystal that tapers gradually to a small fiber. The shape of the fiber is designed to match Fourier transform predictions for load, stress, and taper angle so that a slight vibration of the large base produces large vibration of the tip. The base is excited by coils through the magnetostrictive effect (figs. 1 and 2). Two of these fibers (each bearing a small mirror) set at right angles to one another and excited to vibrate in a sinusoidal manner, but 90° out of phase, will produce a perfect circular scan. This circular scan can be converted to a horizontal sawtooth scan by appropriate fiber optics. This method of scanning produces clean waveforms up to several tens of kHz. A single crystal of this type can also be stimulated to produce a circular scan if the proper field is applied at its base (fig. 3). This device is called a "nutating mirror," and is effective to about 23 kHz, after which low-

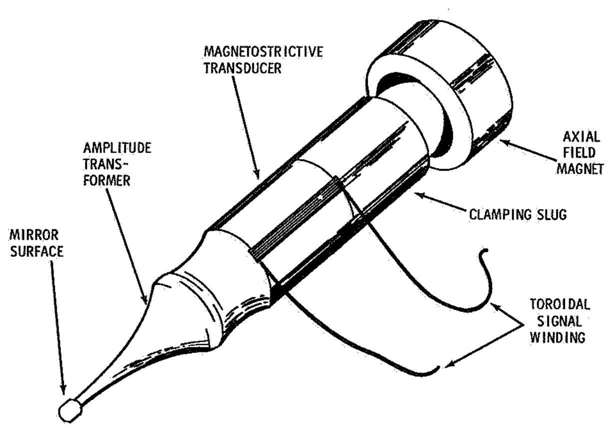

FIGURE 1.—Torsional scanner construction.

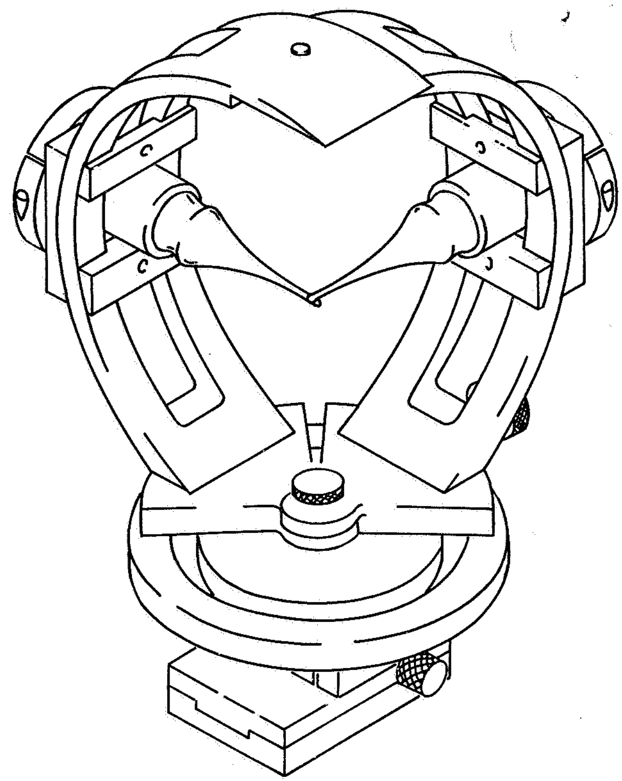

FIGURE 2.—Two-axis torsional scanner assembly.

frequency harmonics disrupt the scan (ref. 3). Both these systems use the natural resonance of a fiber to achieve a stable scan at the desired frequency.

Substantial work has been done with another interesting type of mirror deflection device.

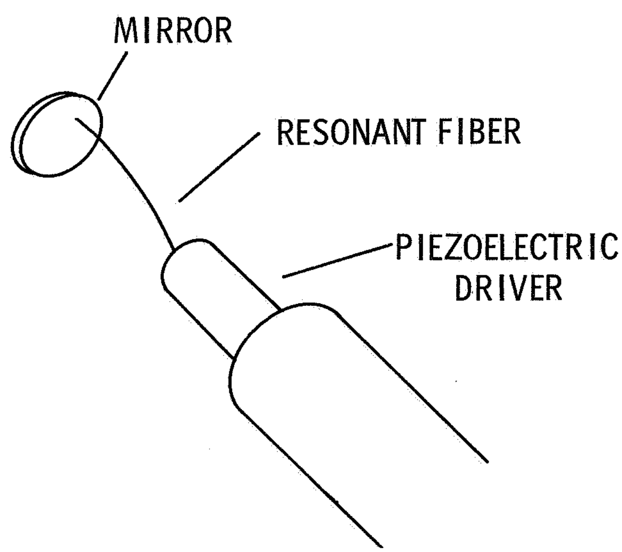

FIGURE 3.—Nutating mirror, horizontal scanning device.

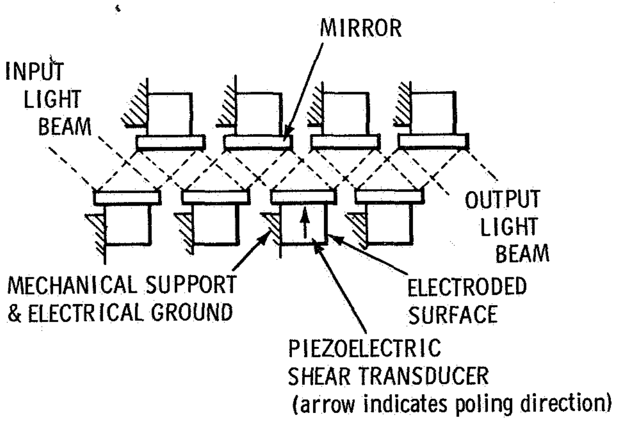

FIGURE 4.—Multiple-mirror deflector using shear-mode transducers.

This device (ref. 5) consists of a train of many mirrors, each attached to a piezoelectric crystal (fig. 4). Each mirror produces a small angular deflection of the beam when its crystal is distorted by an electric field. The proper input to the crystals produces a net deflection of the laser beam at the output of the device.

Using index-of-refraction principles, a digital deflection device has been developed. This device consists of electro-optic crystals to rotate the polarization vector and birefringent crystals to discriminate the direction of polarization (ref. 6). Each pair of crystals—a polarizer crystal and a birefringent crystal—can channel the transmitted light to one of two positions, as shown in figure 5. By cascading n such crystal pairs so that the output of one pair serves as the input to the next pair, a total of 2^n discrete output beam positions may be generated. The scheme described is conceptually one dimensional; however, the extension to two-

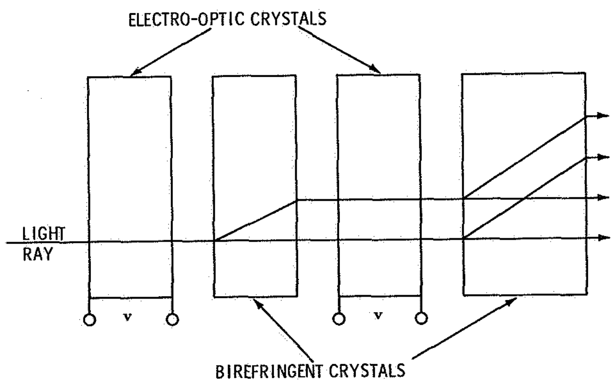

FIGURE 5.—Digital lateral displacement deflector.

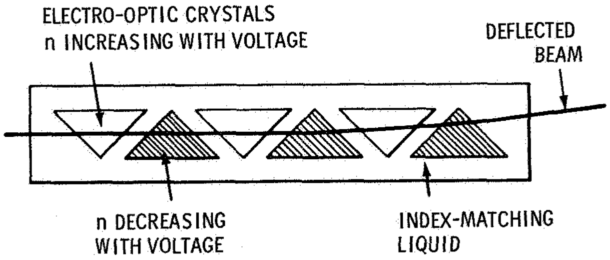

FIGURE 6.—Cascading of crystals in electro-optic beam deflector.

dimensional deflection is obvious. The X- and Y-axis deflectors could be arranged alternately or in cascade.

Another type of index-of-refraction deflector (ref. 7) is a series of alternately opposing prisms (fig. 6). An electric field is applied to alter the index of refraction of each prism and, thus, to deflect the transmitted beam. An index-matching fluid surrounds the prisms to reduce light loss from surface reflections. The device can produce a maximum net angular deflection of about 5°. This type of deflector generally has good frequency response—up to several hundreds of kHz—but overall transmissivity tends to be poor because of contact with many surfaces. Deflection angles are a function of wavelength and, thus, problems arise with nonmonochromatic light.

The third type of deflector uses an ultrasonic source to set up acoustic waves inside a crystal (ref. 8). Because this wave alters the index of refraction of the crystal from the node to the

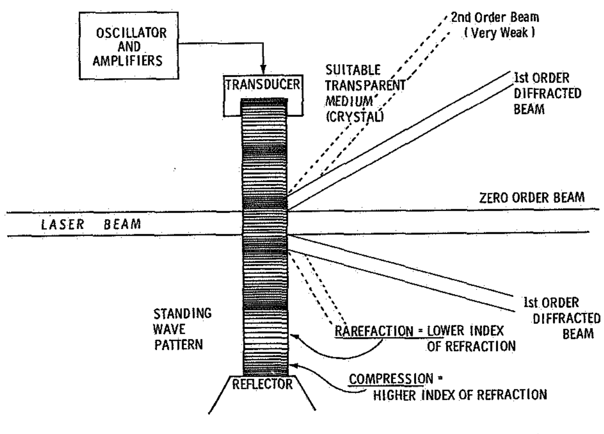

FIGURE 7.—Phase diffraction grating based on index-of-refraction pattern caused by audio or ultrasonic formation of standing waves within the light transmitting medium.

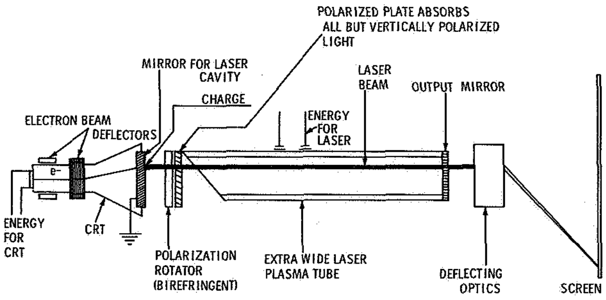

FIGURE 8.—CRT-stimulated modal operation wide-tube laser, producing lateral deflection in x and y directions.

antinode, the crystal becomes a phase-diffraction grating. The beam is split into three parts: two first-order images and one zero-order image (fig. 7). The two first-order beams are deflected at some angle to the original beam and, as the ultrasonic source alters the wavelength of the standing (or traveling) wave, the spacing of the grating is altered. This change causes the beam to be deflected at a different angle (ref. 9).

This method of deflection is impractical because serious problems due to transients arise in changing from one wavelength to another inside the crystal and, at best, only one-third of the light entering the crystal is transmitted after deflection. As a result, its use as a deflection method has been generally limited.

Another interesting laser scanning device employs an electron beam and a laser (ref. 10). The laser is pumped to lasing power but is prevented from lasing by polarization spoiling within the laser cavity. (See fig. 8.) A Brewster's angle window is built into one end of the plasma tube and has the property of transmitting 100 percent of the light polarized in the plane of incidence (vertically in the figure) and only about 85 percent of light polarized perpendicular to that plane. The remainder of the light is reflected from the window surface. Thus, very efficient light transmission is achieved for light of the proper (vertical) polarization, while some reflection losses occur otherwise. A vertical polarizer and a birefringent polarization rotator are included in the laser cavity to produce further polarization selectivity and about 10° of polarization rotation, respectively. Light is then reflected from a mirror at the end of the cavity and, again, passes through the polarization rotator where it is rotated another 10° and then passes on through the vertical polarizer. Enough light loss takes place at the vertical polarizer, where the polarization mismatch is now about 20°, to inhibit lasing.

If a negative charge is deposited on the mirror surface by the electron beam, the polarization is rotated in the opposite direction and the cavity gain becomes sufficient to support lasing action. Because the laser resonator is built with parallel mirrors, lasing will occur within a small area opposite the charged portion of the mirror. As the electron beam is scanned, lasing will occur along a corresponding axis. The plasma tube has a large cross-sectional area to accommodate the range of lasing axes.

In this system, the output of the laser is controlled by deflecting an electron beam in a conventional manner, which has the advantage of easily controlled, high-intensity output. The scan thus produced can be expanded and projected by appropriate optics. At present, the spot size is about 30 μ in diameter. Some difficulty has been encountered in modulating the beam, because a residual charge tends to remain on the mirror after the electron beam has swept by.

Of the systems mentioned, the mechanical ones lose less light and seem to be the most efficient low-frequency deflectors. Although index-of-refraction devices have very good response time, their poor transmittance makes them impractical for many applications (refs. 11 and 12).

Modulators

Scanning laser display systems employ, almost universally, a type of electro-optic polarization device to modulate the intensity of the light beam (refs. 7 and 13). Commonly known as the "Pockels cell," a potassium dihydrogen phosphate (KDP) crystal is equipped with electrodes for applying an electric field, as shown in figure 9. The field changes the index of refraction of the crystal to light polarized in one direction normal to the propagation direction of the light. In this way, the polarization of light passing through the crystal can be

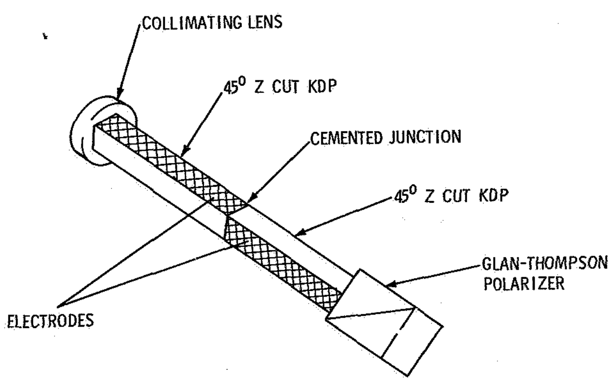

FIGURE 9.—Electro-optic light modulator design.

altered and, by means of an output polarizer, the change in polarization of the light can be converted to a change of intensity. The bandwidth of these intensity modulators extends well beyond typical video frequencies and, in some cases, to microwave frequencies.

Two crystals are used in series to cancel out temperature effects that would cause a birefringent delay in just one crystal (fig. 10). It should be noted that the axes of the crystals are oriented orthogonally so that the birefringent effect of one crystal cancels that of the other (ref. 2).

A piezoelectric modulator has been developed that employs the principles of interferometry (ref. 14). The device splits an input beam into two parts, each of which is reflected from a high-quality mirror driven by a piezoelectric crystal. The two reflected beams are then recombined to form a modulated output beam (fig. 11).

The two control mirrors are mounted to piezoelectric crystals so that they can be driven in the direction normal to their plane. This movement allows the path length of the two light beams to be changed by as much as one-quarter of a wavelength. When the two beams are re-

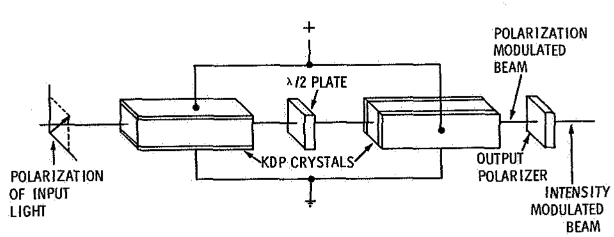

FIGURE 10.—Compensated-birefringence electro-optic modulator.

combined, they may have a phase difference of up to 180°. Intensity changes resulting from this phase modulation and interferometric recombination range from full off to full on.

Because each mirror needs to move only one-eighth of a wavelength, less than 100 volts is required to drive the modulator. The frequency response of the modulator is quite flat to about 2 MHz; however, more precise methods of construction are expected to extend the response of the system to about 5 MHz.

Lasers

There are many types of lasers available today, and the choice of a laser for a display system requires some consideration. Of the three basic categories of lasers—crystal, gas ion, and solid state—gas ion lasers are very attractive for display applications because they are capable of producing continuous emission.

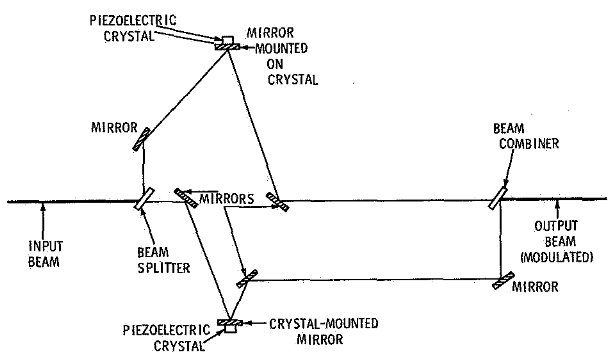

FIGURE 11.—Piezoelectric laser beam modulator.

Typically, gas ion lasers produce comparatively low power output. The most common gas ion laser, which uses a helium-neon gas mixture, seems inherently limited to less than 1 watt of light output power. Other types of gas ion lasers, such as argon ion and krypton ion lasers, are capable of higher power output, but are only about one-hundredth of 1 percent efficient in electrical to optical power conversion.

One watt of optical power, which is on the high end for He-Ne lasers and easily attainable with the other gas ion lasers, is sufficient to build a moderate-size laser display. An argon laser, having primary output in the blue and green

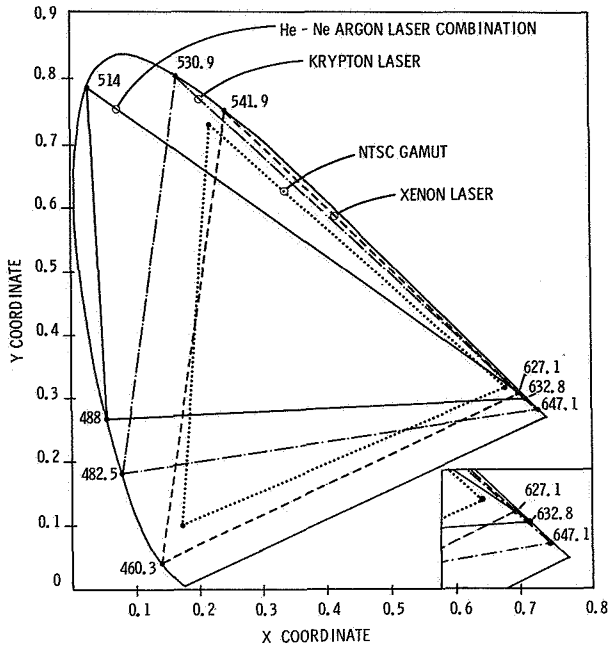

FIGURE 12.—Chromaticity diagram.

wavelengths, can illuminate a 1-meter-square screen to produce a brightness of about 30 ft-L.

Red, green, and blue beams are necessary to produce a color display. Because He-Ne and argon lasers are readily available today, most experimental systems derive a red beam from the He-Ne laser and the green and blue portions of the beam from the argon ion laser (fig. 12). This gives primary colors at 6328 Å, 5145 Å, and 4880 Å, respectively (ref. 2). Although rich in reds and greens, this combination cuts off many of the deep blues.

Both krypton and xenon lasers (ref. 2) have been operated successfully, and each produces a red, green, and blue component. The krypton laser produces red and green fairly close to the red and green components of the He-Ne-argon system (5208 Å, 6471 Å), but its blue (4762 Å) is also fairly long in wavelength, and the deeper blues would not be included in the combination. The xenon laser provides a good red and green (6271 Å and 5419 Å, respectively) and also a blue (4603 Å) for the best combination of the three, but this laser has had less developmental work of all the lasers mentioned and is, therefore, not likely to be available for some time.

Frequency Modulation

An interesting effect in a lithium niobate crystal has been observed (refs. 15 and 16). Temperature variation in the crystal causes it to fluoresce at different wavelengths when excited by the green light from an argon laser. Although this fluorescence is noncoherent, it does produce output power about 50 percent of that of the input beam. A suitable cavity will allow this output to lase (ref. 17), but the power of the coherent output beam is only about one-tenth of 1 percent of that of the input laser beam.

This development suggests a possible means of frequency modulating a laser beam. However, the difficulty of modulating at high rates by temperature change and the power loss involved may prove to be insurmountably awkward for a video system.

Complete Scanning Laser Display Systems

The devices that are necessary to build a scanning laser display are as follows:

(1) Beam deflectors and modulators to manipulate the beam

(2) Electronics to drive the deflectors and modulators

(3) Laser(s) of sufficient power output

As previously mentioned, components of sufficient quality with which to construct a laser display exist today, and further research is likely to produce even better components for the future.

At this time there are at least three complete scanning laser display systems in a developmental stage, and at least one is very close to being a marketable system. For purposes of this report, the three systems to be discussed are designated display 1, display 2, and display 3.

Display 1

Display 1 (refs. 18 and 19) is a large-screen color television (fig. 13). The system produces a standard 525-line television raster, utilizes conventional driving electronics from a commercial color television receiver, and can display commercial broadcast television.

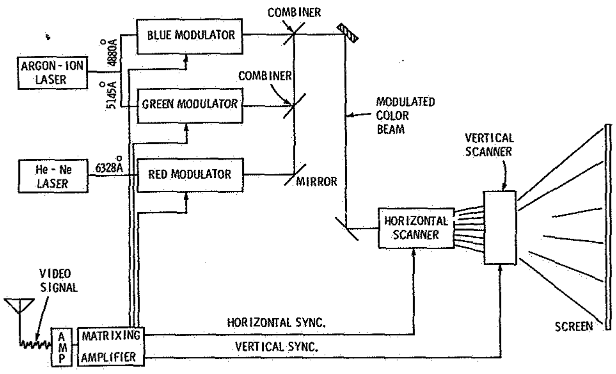

FIGURE 13.—Display 1—laser color television.

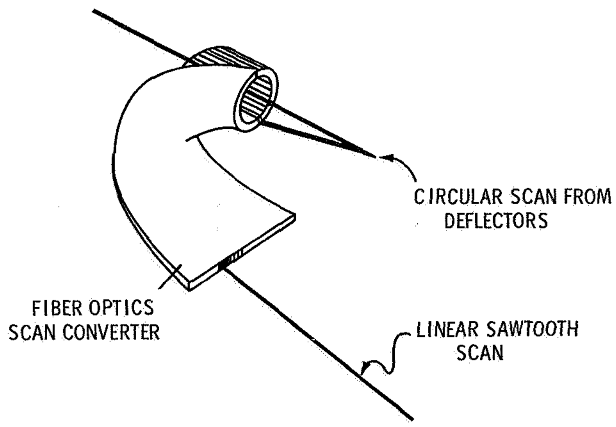

FIGURE 14.—Fiber optics scan converter.

The system uses an argon laser for green and blue wavelengths, and a He-Ne laser for red. The green and blue components of the argon laser are separated by a prism, and the three beams from the two lasers are intensity modulated individually. The modulated beams are then combined in a prism and become a properly coordinated color beam.

The four standard television signals (red, green, blue, and luminescence) are derived in a matrixing amplifier to form three signals for the red, green, and blue modulators of the display system. These three video signals are then amplified and applied to the modulators (ref. 19).

The horizontal scanner is a single quartz crystal that tapers from a large cylindrical base to a fine fiber. The fiber is driven in a circular manner by applying torsional stress. The system is resonant at 15 734 Hz and is, therefore, a tuned oscillator. It is driven by coils that are situated at 90° to one another around the cylindrical base of the crystal. These coils drive the scanner at its resonant frequency, and the system is synchronized by pulses amplified and extracted from the video signal. This deflector is described in more detail in the section titled "Deflection Techniques."

The circular scan is converted to a horizontal sawtooth scan by a fiber optic bundle, which accepts the circular scan on its circular face and transmits light down the fibers to its horizontal tail (fig. 14).

The vertical scanner is a d'Arsonval type of galvanometer. It deflects a 0.8- by 1.5-inch mirror, which is large enough to ensure that the entire horizontal scan is transmitted.

The system is completely compatible with the National Television System Committee (NTSC) color system. It has a 4.5-MHz bandwidth, and the display has a contrast ratio of 100 to 1. The display output has a scan angle of about 7°, and it can be projected from between 5 and 30 feet. Sweep linearity is 2.5 percent, and intensity is accurate to 10 percent. The lasers were operated to produce 100 milliwatts of red, 130 milliwatts of blue, and 23 milliwatts of green light, for a total flux of 44 lumens.

The fiber optic element presents the greatest difficulty, because even though transmission is very high—on the order of 90 percent—flaws in the bundle produce vertical lines in the picture. The system as a whole will transmit not less than 35 percent of the light entering the modulator. The modulator itself transmits only about 50 percent of the incident light. Work is continuing to perfect the horizontal scanner and to increase the transmission and overall brightness of the system.

Display 2

Display 2 (ref. 2) is a seven-color, digital data display system. It is a developmental model for a much more complicated display system. The ultimate objective is to produce a 1024- by 1024-element display which will illuminate a 50-ft² screen at a brightness of 50 ft-L.

The present system uses a 5-ft² screen and produces a display with 512 by 512 resolution elements. It operates with a brightness of about 10 ft-L. The system produces seven colors: red,

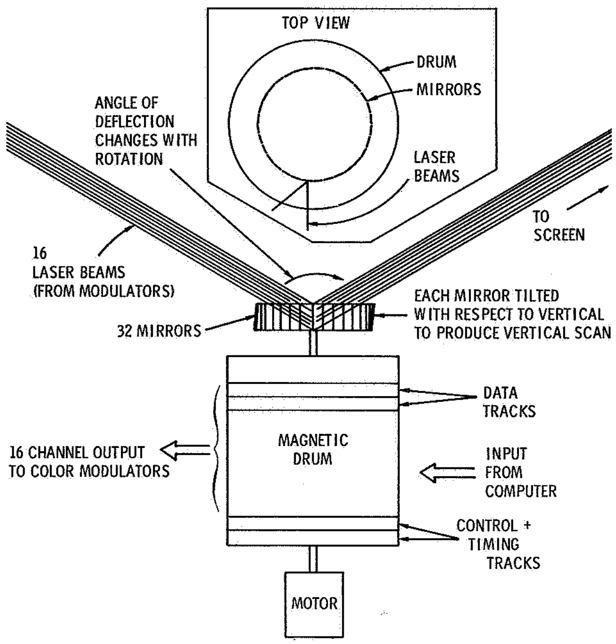

FIGURE 15.—Digital scanner (one for each color).

green, blue, yellow, cyan, magenta, and white, which are obtained by mixing the three primary colors (red, blue, and green) from two lasers in a digital fashion. The system uses an argon laser and a He-Ne laser.

The unique method of scanning used in this system requires that the data be in digital form. The scanner consists of 32 mirrors that are arranged in a polygon, as described earlier. Each face is slightly tilted with respect to the neighboring mirrors, producing a vertical scan (fig. 15). This mirror arrangement is rotated on the same axis and by the same motor as the magnetic drum on which the display data are recorded. Each primary-color laser source has been split into 16 beams, each of which is modulated in a digital fashion by one of the 16 tracks on the drum. In this way, as the data from the drum are read into the modulators, the beam is scanned across the screen. Each primary color has a 16-track data storage drum and a rigidly coupled polygonal mirror deflector associated with it. The system has a frame rate of 60 Hz. The contrast ratio is 100 to 1 in a darkened room, and linearity and registration are better than 0.25 percent (fig. 16).

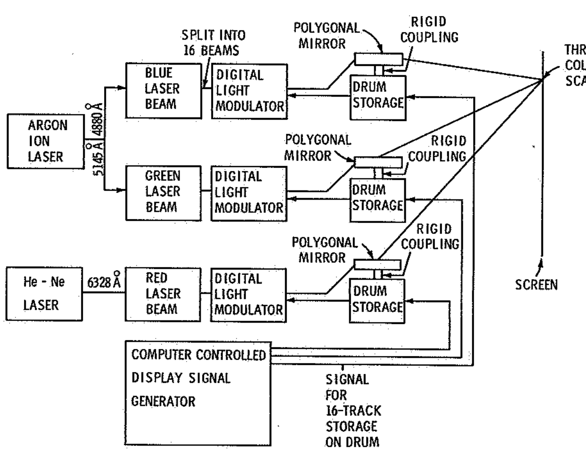

FIGURE 16.—Display 2—schematic diagram.

Display 3

Display 3 (refs. 5, 7, and 20) is a scanned, color laser display system using a flying spot both to illuminate the scene and to illuminate the display screen. A three-color beam is scanned in a raster pattern by appropriate electronics. Part of the moving beam is diverted from the screen to scan an object of interest. A sensing device that consists of three photodetectors, each sensitive to only one of the color components of the scanning beam, monitors the light reflected (or transmitted) from the scene and thus produces three color-separated video signals for modulating the display beams. Each signal is amplified and drives the corresponding modulator. (See fig. 17.)

The three colors are produced from argon and He-Ne lasers. Each beam is polarization modulated so the intensity of that portion of the beam scanning the object does not change. The beams directed toward the screen are passed through a polarizer to produce intensity change in response to the polarization change.

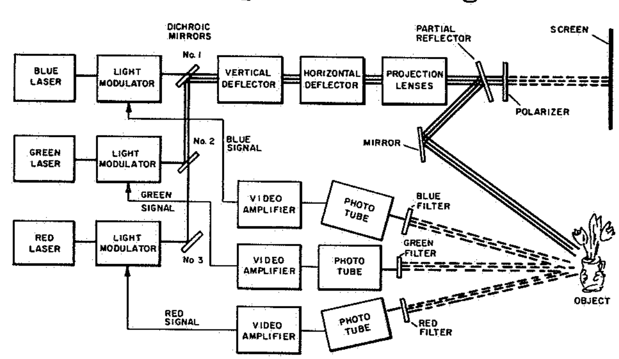

FIGURE 17.—Experimental three-color scanner laser display system.

This technique of simultaneously scanning the scene and the display screen was employed because synchronization of the video signal with the scan pattern is automatic. These researchers were unable to develop a satisfactory sawtooth scanning device, and the present system was developed to avoid the need for one.

The objective of this program is to develop a scanning laser display that is compatible with standard television signals, and work is continuing along this line. The resolution of the present system is comparable to standard television, and its contrast ratio is 100 to 1 in a dark environment.

This laboratory model is also being used to study computer-driven scanning systems. Such a device would choose an optimal scan for the objects being viewed, and would eliminate the necessity of recording and displaying useless information. It is hoped that computation equipment to accomplish this could be included in the display computer interface.

In addition to the displays mentioned and described, there are many other single-color scanning laser display systems that utilize the techniques of scanning and modulation described previously.

Although bulky and expensive in the research stage, scanning laser displays have the potential for substantial improvement over conventional CRT's. Ultimately, color-scanning laser displays will be developed to the point where they will replace the CRT in many flying spot applications. They offer greater brightness and greater resolution, and the display beam need not be in a vacuum. In addition, a scanning laser system can provide a very large screen display for multiple viewing with a practical degree of brightness and resolution (ref. 21).

Holographic Displays

In 1947, Gabor proposed that, if two wavefronts were allowed to interfere and the pattern were recorded on film, and if one of the wavefronts were passed through the resulting transparency, the other would appear as the first-order diffraction pattern. He called this technique "wavefront reconstruction" and observed that spatially coherent light would be required to make his technique work.

Not long after lasers were developed (particularly the class of continuous emission gas lasers), Gabor's techniques were applied, and wavefront reconstruction became a practical reality. Stroke proposed that the name "holography" be adopted for the new technique; *holos*, Greek for "everything," and *graphein*, Greek for "to depict," imply that, in a hologram, everything necessary to reconstruct a picture is contained in its wavefront.

A conventional photograph records the intensity pattern of a scene by "mapping" individual points of the scene onto corresponding points of film by use of a lens. A hologram also retains intensity information from the scene and, additionally, records the phase of light originating from each point. This phase information is related to the distance of points from the hologram and, therefore, depth information is contained in the holographic record. The hologram pattern is not related point for point to the recorded scene, but, instead, light from each point of the scene is distributed over the entire hologram. So much information is recorded in the two-dimensional plane of the hologram that a hologram has some unique properties. (See fig. 18(*a*) and 18(*b*).)

Properties of Holograms

Because the wavefront is entirely reconstructed, there is, substantially, no difference between the view of the reconstructed wavefront and that of the original wavefront. Full three-dimensional aspect is apparent. Also, each small section of the hologram plane has received light from every point of the recorded object; each small segment of the hologram will reproduce its view of the entire scene. Of course, different segments produce views of the scene from slightly different angles, further demonstrating that the reconstructed image of the entire plane is three dimensional.

A hologram is produced by recording the interference pattern between two wavefronts. As previously stated, the pattern can diffract one wavefront into the other. Normally, one of the wavefronts used to produce the hologram is very uniform; for example, a plane wavefront or a spherical wavefront. The other

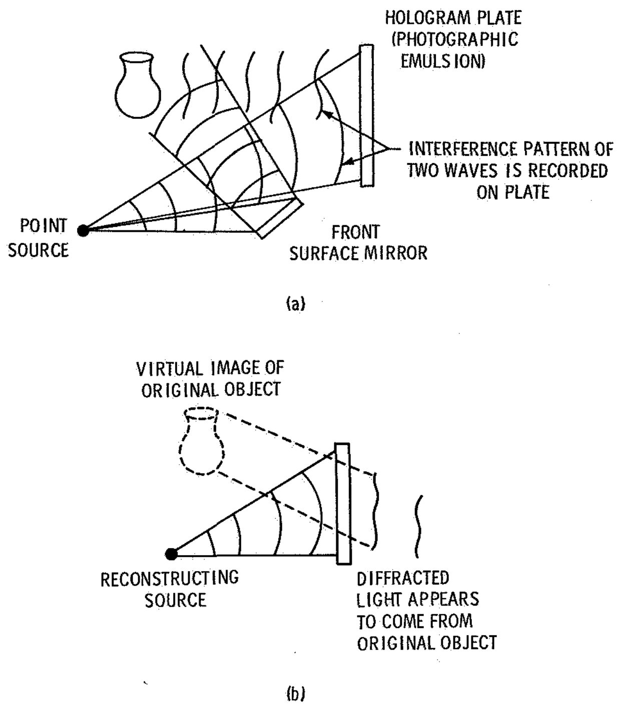

FIGURE 18.—Holographic displays. (*a*) Reference beam is part of wavefront from point source. Rest of beam is used to illuminate object. (*b*) First-order diffraction pattern from hologram produces wavefront identical with that which would come from object (as shown by dotted lines).

wavefront would be complicated, such as that from a scattering object. When the uniform wavefront is projected through the hologram plate, the irregular wavefront is produced, and one sees the scattering object as if through a window.

Several holograms can be stored on one plate by using another technique, namely, the "Bragg effect." The photosensitive material in which the interference pattern is stored is not strictly a two-dimensional transparency. The emulsion thickness, after development, for Kodak 649F spectroscopic plates is typically 15 μ. This is equal to approximately 25 or 30 wavelengths. Thus, the "interference lines" are, in reality, dark bands existing throughout the thickness of the emulsion. The result of this characteristic is a sensitivity to the wavelength and direction of the reconstructing light. The image intensity as a function of reconstructing beam angle for planar holograms is as shown in figure 19, the principal maximum occurring at the recording reference beam angle.

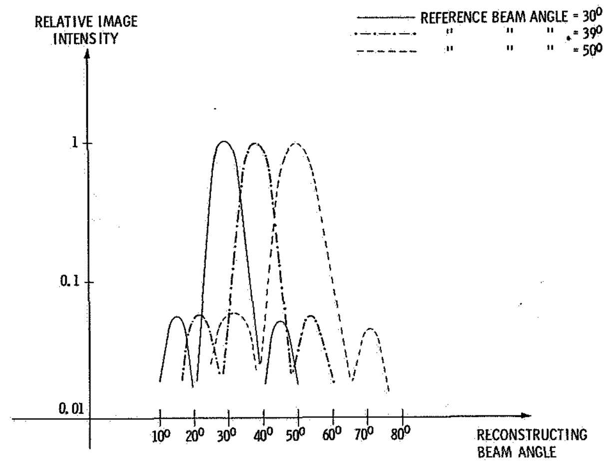

FIGURE 19.—Image intensity as a function of reconstructing beam angle.

Several holograms can be stored on the same plate by choosing reference beam angles so that the image intensity maximum for any one hologram corresponds to an intensity minimum for all the other holograms. As many as seven holograms have been stored on one plate by using the technique of Leith et al. (refs. 22 and 23). This technique can also be used to produce color holograms. By using reference beams of different colors and by choosing reference beam directions so that crosstalk between images is minimized, it is possible to reconstruct the images associated with the several colors superimposed in space, thereby generating a full-color hologram (refs. 24, 25, and 26). Color holograms that reconstruct with white light have also been made using the Lippmann effect (ref. 27).

A constraint that threatened to keep holograms in the laboratory was the necessity to produce them with coherent light. This limitation prevented making holograms of outdoor, or large, scenes. Subjects that could not be illuminated by laser light, and only laser light, were immune to holography. A technique, known as the fly's-eye method of photography, was developed early in 1967 (ref. 28) and eliminates the difficulty created by the need for coherent light. In this procedure, a composite, many-faceted lens, much like that of a fly's eye, is used to produce many images of the scene of interest (fig. 20). If a fly's-eye lens of about 2 inches in diameter is used, the scenes recorded

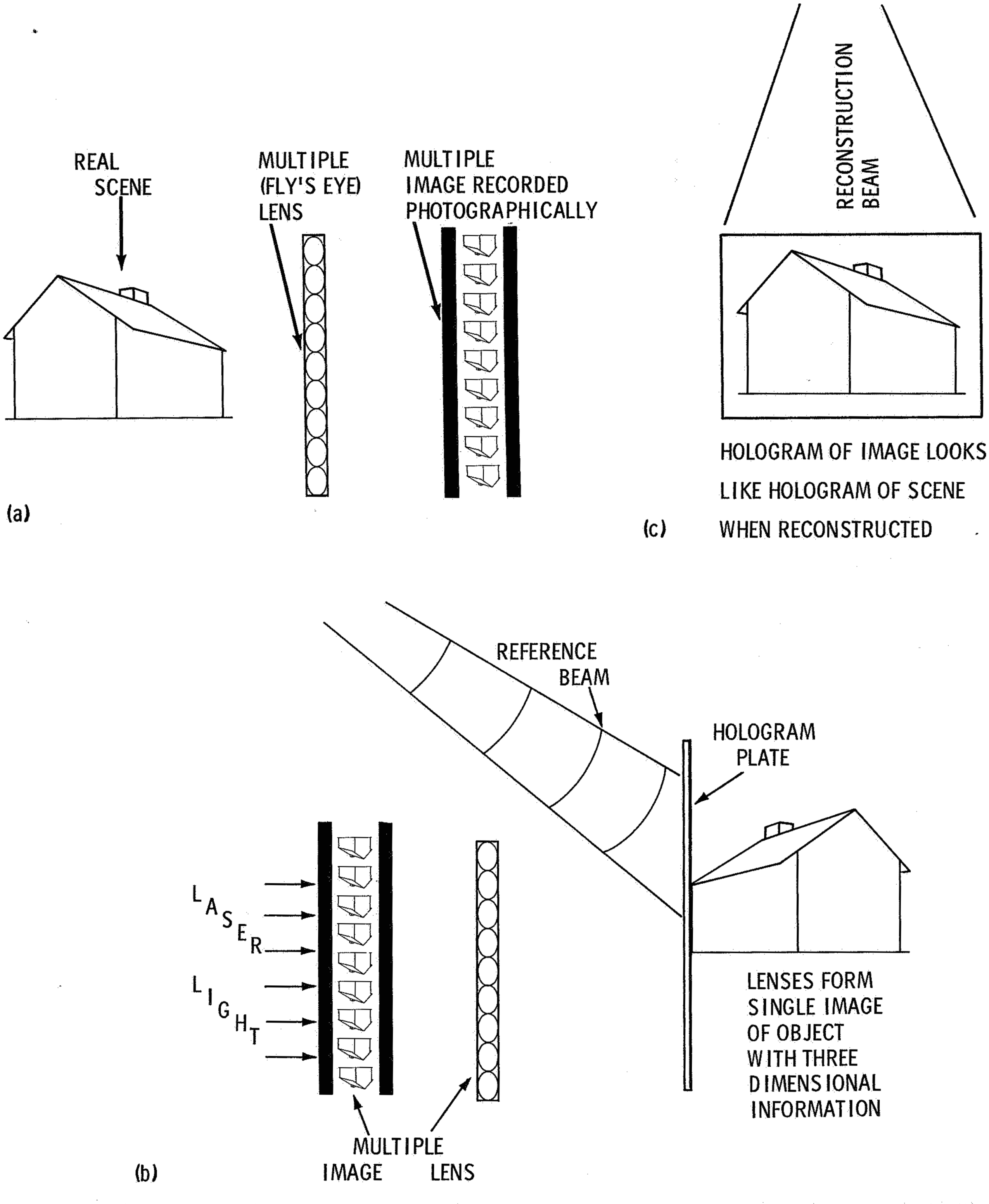

FIGURE 20.—Fly's-eye method of photography. (*a*) Fly's-eye lens makes multiple recording of scene. (*b*) Multiple recording is illuminated by laser and imaged by fly's-eye lens. Hologram of image is recorded. (*c*) Three-dimensional reconstruction of original scene.

by opposite edges of the lens will show as much angular variation as the scenes observed by the two eyes of a human. This indicates that the fly's-eye lens can record three-dimensional data as complete as human eyes can detect. The images from a fly's-eye lens are recorded in the manner of conventional photography, and a color transparency is made. If this transparency is projected through another fly's-eye lens of exactly the same size containing the same number of elements, and oriented in the same way as the original lens that made the picture, it produces a single real image of the scene of interest. This image contains three-dimensional information. If a laser is used to project this image and to make a hologram of it, the hologram has all the characteristics that one of the scene would have. In this way, a hologram of any scene, illuminated by any light, can be made if a photograph of the scene can be made.

Another property of holograms having application to displays is the ability to alter the virtual output image by reconstructing with a wavefront that is not exactly the same as the reference beam wavefront. Consider the case of a hologram constructed with plane waves and reconstructed with a point source located at some distance f from the plate and in the same direction as the reference beam (fig. 21). One can think of this configuration as a reconstruction with the original plane wave reference beam through a negative lens of focal length f, located in the plane of the hologram. The virtual image in this case is a closer, "minified" image. It is the virtual image of the original scene as viewed through a negative lens. Similarly, positive lenses can be simulated by using converging reconstructing bundles. Images produced in this way are enlarged virtual images if the effective focal length is greater than the object distance, or inverted real images if the effective focal length is less than the object distance. Lenses of any focal length can be simulated in this way. Mechanical manipulation of the position of the point source can provide a continuously variable focal length.

Prisms can be simulated in the same way. If a hologram, recorded with a plane reference wave, is reconstructed with plane waves originating from a slightly different direction than that of the original reference beam, the resulting virtual image is of the original scene viewed through a prism. The prism is that which would redirect the reference beam to be alined with the reconstructing beam. The resulting image appears somewhat like the original scene rotated about an axis in the plane of the hologram (ref. 29). The image intensity is somewhat lower than maximum when reconstructing with plane waves originating from other than the preferred direction. This is due to the sensitivity of intensity to reconstructing beam angle discussed earlier. This technique is effective for apparent rotations of $\pm 5°$ with Kodak 649F plates. Rotated versions of a scene can be stored at extinction points, as discussed earlier, thereby extending the range of reconstructing beam angles over which the hologram is useful (ref. 30).

A prototype display system using this method to obtain image rotations is being designed at the Electronics Research Center of NASA. Combined with mechanical manipulation of the hologram plate (translation within the hologram plane and rotation of the hologram about an axis normal to its plane), this technique will display a three-dimensional scene that can be controlled in five degrees of freedom.

One serious problem in the field of holographic display is that conventional holograms cannot be generated in real time. A hologram must be carefully created over a period of time ranging from a few minutes to a few hours. A carefully controlled environment is required during hologram construction; therefore, the hologram must be produced for a display situation ahead of time. One possible way to bypass this problem is through computer generation of holograms. Generation and manipulation of holograms by computer is interesting for many other applications as well. The realization of real-time holographic display must remain for the future, because even the problem of real-time hologram generation by computer has not been completely solved. The general method is to use mathematics that correctly describes the phenomena of interference at a plane, and allow a computer to calculate the degree of in-

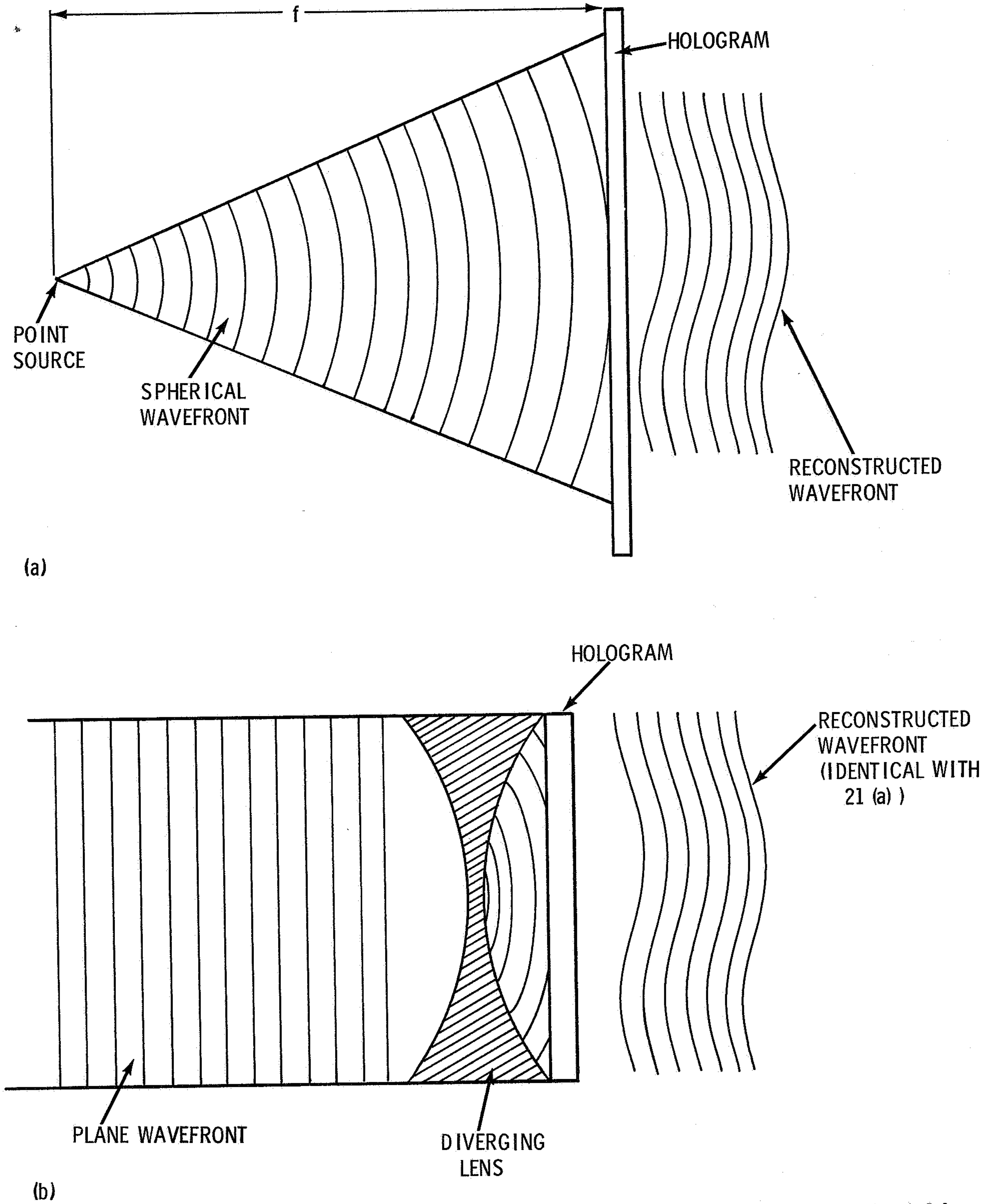

FIGURE 21.—Reconstructed holograms. (*a*) Hologram made with a plane reference beam reconstructed by a spherical beam. (*b*) Planar hologram reconstructed with plane waves and a diverging lens (physical equivalent of (*a*)).

terference at points in the hologram plane. Because the wavefront from any object can be approximated as the superposition of spherical wavefronts from points on the object, the hologram can be generated by consideration of the reference wave interfered with the summation of the many spherical wavefronts. A method has been devised for reducing the number of calculations necessary to produce a hologram, and the number of calculations necessary to perform operations on a created hologram to produce another one (ref. 31). This work makes the computer generation of complex holograms a somewhat more practical approach for the near future. At present, holograms of simple objects have been generated by computer (ref. 32).

Another property of holograms applicable to display is that any form of energy propagated by a wave will produce a hologram. Holograms have been produced by using ultrasonic vibrations as the "illuminating" source (refs. 33 and 34). Because these waves have a wavelength two or three orders of magnitude greater than that of light, the hologram so produced was proportionately larger in size. A conventional hologram was then made by reducing the original pattern to optical proportions. The resulting hologram, when illuminated by light, produced a visual reconstruction of the object illuminated by the sound. A similar experiment was successfully conducted using microwave radiation (ref. 35).

Work has also been performed to develop materials, other than photographic emulsions, that will respond to laser light. Some work has been completed in the use of alkali-halide single crystals for storage of hologram (and other) patterns. These crystals can be induced to form electron "holes" in the lattice where an atom has been displaced. The electron that fills this position is capable of responding to light in unique ways. For instance, a technique by which a laser can be made to write in a crystal in a specific manner and then erase what is written has been proposed. Theory predicts other dramatic properties for this type of crystal (ref. 36).

Many applications (ref. 37) for holography are immediately apparent. The concept has been used extensively in the science of interferometry because a hologram is basically an interference pattern. Especially valuable to interferometry is a method of taking "flash" holograms with a Q-switched laser (refs. 38, 39, and 40).

The capability of a hologram to magnify an object by large proportions makes it very useful in measuring. "Flash" holograms have been made of dynamic microscopic subjects; they can be studied in full three-dimensional reconstruction and at the researcher's leisure (ref. 41). Techniques for recording bubble-chamber data (information requiring detection of three-dimensional spatial relations) have been perfected (ref. 42). The reconstructed image of a hologram can be examined in a natural form or the interference pattern itself can be measured to extract the recorded data.

Extensive work has been performed to evaluate the use of holograms as data storage devices (ref. 43) for computers. The hologram acts as a storage device when any measurable information is recorded on it. The experiments with alkali-halide crystals, cited previously, indicate the potential for recording digital data, with the option to erase and re-record. In a different research effort (ref. 44), holograms were used to record digital data in a manner which allowed associative retrieval of the stored information; that is, information was retrieved from the hologram on the basis of content, rather than location (as in conventional memories). This research suggests the possibility for pattern recognition systems based on holograms as the storage medium.

The application of holography in the field of display seems to surpass its possible use in other fields. Here is a technique that can, potentially, provide a dynamic, full-dimensional, full-color, high-resolution display of any visual scene. This potential is realizable in the near future.

Any holographic display system must include a light source of suitable intensity for the display; and a mechanism to manipulate the hologram, either a hologram changing device if a hologram movie is to be displayed, or a multidimensional manipulator if the hologram is to be reoriented to provide different views.

Because of the availability of very bright lasers, a suitable light source for the display of a hologram does not pose a problem. Actually, a laser is not necessary; reconstruction is possible in any reasonably uniform illumination with a wavefront shape similar to the reference beam. Manipulation of the hologram(s) is by lateral motion of the hologram plate, by rotation of the hologram plate, or by replacement of the hologram plate with another one. A complete display might utilize a combination of these manipulations.

A hologram can be used to display any type of three-dimensional data. With a perfected method of computer generation of holograms, output data from a computer could take on visual, three-dimensional form.

A holographic display could be used as a head-up, all-weather landing and navigation display for aircraft and spacecraft pilots. In the case of an aircraft, it would provide a visual representation of the airport at which the pilot is landing. This visual representation would be driven by a small onboard computer, which would obtain driving data from the aircraft's instruments and ground control. The view would be reflected from the windshield, or installed as a window in the instrument panel. The update rate could be rapid enough to give the pilot an effectively continuous display (see fig. 22). A similar display could be used in flight-simulation devices to give a realistic feel of flight to the simulator operator.

Presently, only a few types of holographic displays are at the prototype stage. One is a holographic movie system which displays a 1½-hour sequence of holograms at a rate of 20 frames per second. It displays an apparent 2-meter cube of space, and can be viewed effectively from a floor area of about 30 square feet. Ultimately, the designers hope to produce a system displaying 24 frames per second with 2 views per frame. This would provide 48 pictures per second, which is sufficient for the eye to see a continuous image. Work is being done to develop a method for viewing color holograms by reflection to improve the quality of the hologram movie.

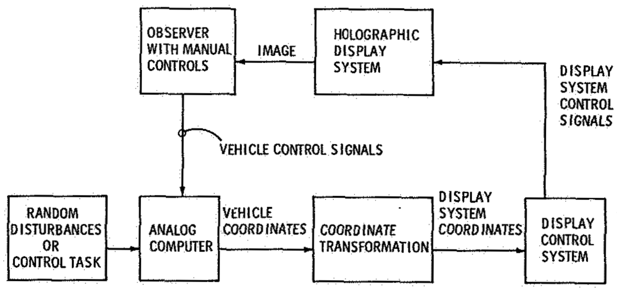

FIGURE 22.—Flight simulator utilizing a holographic display.

Several industries and NASA are working on the development of holographic display devices to simulate flight. These displays will have characteristics similar to the theoretical pilot display discussed above.

Holography, as a new science, has generated wide interest. The many applications of holography have been recognized for some time (refs. 45 and 46). The literature explains the concept of holography and describes, in greater detail than this report, many of its applications.

REFERENCES

1. NAGY, G. T.: Scanning Techniques With Light Beams. Information Display, May–June 1966.
2. BAKER, C. E.; AND PIPKEN, L. L.: A Multi-Color Laser Digital Display System. Texas Instruments, Inc., Dallas, Tex.
3. BAKER, C. E.; AND ALASBROOK, C. M.: A Large Screen Color TV Laser Display. Texas Instruments, Inc., Dallas, Tex.
4. BAKER, C. E.: Large Angle Deflection Techniques for Laser Display, Final Report. RADC–TR–66–722 (DDC No. AD 651301), Texas Instruments, Inc., Dallas, Tex., Feb. 1967.
5. FOWLER, V. J.: Investigation of Electro-Optic Techniques for Controlling the Direction of a Laser Beam. NASA TR–66–722.17, Jan. 29, 1966.
6. SENF, H. R.: Laser Display. Preprint no. 100–2, Society of Motion Picture and Television Engineering, 100th Technical Conference, Oct. 2, 1966.
7. FOWLER, V. J.: Scanned Laser Display Techniques. Symposium to Discuss a Display System for a Command and Control Center, North Atlantic Treaty Organization, Munich, Germany, Nov. 10, 1965.

8. AAS, H. G.; AND ERF, R. K.: Application of Ultrasonic Sound Waves to the Generation of Optical Beam Scanning. J. Acoust. Soc. Am., vol. 36, no. 10, Oct. 1964, p. 1906.

9. ERF, R. K.: Application of Laser Beam Scanning Systems Employing Ultrasonic Techniques. Proceedings of the 10th Annual Symposium of the Society of Photo-Optic Instrumentation Engineering, San Francisco, Calif., Aug. 20, 1965.

10. MYERS, R. A.: Scanning With Lasers. International Science and Technology, May 1967, p. 40.

11. EVERETT, S. L., JR., ET AL.: Broadband Laser Digital Deflector. TR–ECOM–2795 (DDC No. AD 651220), U.S. Army Electronics Command, Fort Monmouth, N.J., Jan. 1967.

12. LOTSPEICH, J. F., ET AL.: Selective Access Laser Display Beam Positioner. TR no. RADC–TR–66–447 (DDC No. AD 639968), Hughes Research Laboratories, Malibu, Calif., Aug. 1966.

13. FOLEY, W. L.: A Study of Light Modulation and Scanning Techniques for Application to Simulation Display Generation. AMRL–TR–66–9 (DDC No. AD 637307), U.S. Air Force Systems Command, Wright-Patterson AFB, Dayton, Ohio, Mar. 1966.

14. FRIED, D. L.; READ, W. S.; AND POLLOCK, D. B.: An Interferometric Optical Modulator. Appl. Opt., vol. 3, no. 6, June 1964, p. 697.

15. HARRIS, S. E.; OSHMAN, M. K.; AND BYER, R. L.: Observation of Tunable Optical Parametric Fluorescence. Phys. Rev. Letters, vol. 18, no. 18, May 1, 1967, p. 732.

16. GIORDMAINE, J. A.; AND MILLER, R. C.: Optical Parametric Oscillation in the Visible Spectrum. Appl. Phys. Letters, vol. 9, no. 8, Oct. 15, 1966, p. 298.

17. OSTERINK, L. M.: Single Frequency Light From an Argon FM Laser. Appl. Phys. Letters, vol. 10, no. 4, Feb. 15, 1967, p. 115.

18. BAKER, C. E.: A Large Screen Real-Time Display Technique. Information Display, Mar.–Apr. 1966, p. 37.

19. BAKER, C. E., ET AL.: A Multi-Color Laser Display. RADC–TR–66–711 (DDC No. AD 650203), Texas Instruments, Inc., Dallas, Tex., Feb. 1967.

20. STONE, S. M.: Experimental, Multi-Color, Real-Time, Laser Display System. Technical Report, General Telephone & Electronics, Bayside Research Labs., Bayside, N.Y., Jan. 1, 1967.

21. FOWLER, V. J.; AND KOHN, E. S.: A Photochromic Laser Display. Proc. IEEE, Mar. 1967, p. 424.

22. LEITH, E. N., ET AL.: Holographic Data Storage in Three-Dimensional Media. Appl. Optics, vol. 6, no. 8, Aug. 1966.

23. GRAULING, C.: Holograms Which Reconstruct With Widely Varying Reconstructing Beam Angle for Display Applications. NASA/ERC/GCD. Unpublished technical paper, 1966.

24. PENNINGTON, K. S.; AND LINN, L. H.: Multi-Color Wavefront Reconstruction. Appl. Phys. Letters, Aug. 1, 1965, p. 56.

25. MAROM, E.: Color Imagery by Wavefront Reconstruction. J. Opt. Soc. Am., vol. 57, Jan. 1967, p. 101.

26. FRIESEM, A. A.; AND FEDUROWITZ, R. J.: Multi-Color Wavefront Reconstruction. Appl. Optics, vol. 6, no. 3, Mar. 1967, p. 298.

27. STROKE, G. W.; AND LABEYRIE, A. E.: White-light Reconstruction of Holographic Images Using the Lippmann-Bragg Diffraction Effect. Phys. Letters, vol. 20, no. 4, Mar. 1, 1966, p. 368.

28. POLE, R. V.: 3–D Imagery and Holograms of Objects Illuminated in White Light. Appl. Phys. Letters, vol. 10, no. 1, Jan. 1, 1967, p. 20.

29. DEVANEY, A. J.; BARON, S.; AND GRAULING, C.: Holographic Display Systems. NASA/USC Ann. Conf. on Man Control, USC, Mar. 1967.

30. ROSEN, L.: Apparent Rotation of Hologram Virtual Images. J. Opt. Soc. Am., vol. 57, no. 2, Feb. 1967, p. 278.

31. COOLEY, J. W.; AND TUKEY, J. W.: An Algorithm for the Machine Calculation of Complex Fourier Series. Mathematics of Computations, vol. 19, Apr. 1965, p. 297.

32. LESEM, L. B., ET AL.: Computer Generation and Reconstruction of Holograms. IBM Pub. 37.021, International Business Machines, Inc., 1966.

33. DREHER, J. J.; EL-SUM, H. M. A.; ET AL.: Optical Reconstruction From Sampled Holograms Made With Sound Waves. Phys. Letters, vol. 24A, no. 10, May 8, 1967, p. 67.

34. PRESTON, K., ET AL.: Ultrasonic Imaging Using a Synthetic Holographic Technique. Appl. Phys. Letters, vol. 10, no. 5, Mar. 1, 1967, p. 150.

35. TRICOLES, G.; AND ROPE, E.: Reconstruction of Visible Images From Reduced Scale Replicas of Micro-wave Holograms. J. Opt. Soc. Am., vol. 57, no. 1, Jan. 1967, p. 97.

36. ANON.: An Introduction to Alkali-Halide Holography. Carson Laboratories, Inc., Bristol, Conn., 1966.

37. BROOKS, R. E.: Scientific Applications of Holography. TRW/Systems Technical Publication, Redondo Beach, Calif., Apr. 26, 1967.

38. BURCH, J. M., ET AL.: Dual and Multiple Beam Interferometry by Wavefront Reconstruction. Nature, Mar. 5, 1966, p. 105.

39. BROOKS, R. E., ET AL.: Holographic Interferometry. J. Appl. Phys., Feb. 1962, p. 642.

40. BROOKS, R. E., ET AL.: Holographic Photography of High Speed Phenomena With Conventional and Q-Switched Ruby Lasers. Appl. Phys. Letters, Aug. 15, 1965, p. 92.

41. KNOX, C.: Holographic Microscopy as a Technique for Recording Dynamic Microscopic Subjects. Science, vol. 153, no. 3739, Aug. 26, 1966, p. 989.
42. WARD, J. H.; AND THOMPSON, B. J.: In-Line Hologram System for Bubble-Chamber Recording. J. Opt. Soc. Am., Feb. 1967, p. 275.
43. CORSON, R. C.: Display Data Recording by Laser, Final Report. TR-ECOM-01398F (NASA No. 489374), Nortronics Division, Northrop Corp., Palos Verdes Peninsula, Calif., Sept. 1966.
44. KENNEDY, D. W.: Feasibility Study of a Holographic Associative Memory. Sc.D. thesis, Dept. of Mechanical Engineering, MIT, Cambridge, Mass., Feb. 1967.
45. ANON.: Some Applications of Lasers to Information Handling. Optical Group, Perkin-Elmer, Inc., Norwalk, Conn., 1966.
46. STROKE, G. W.: Recent Advances in Holography. MIT Technology Review, May 1967, p. 17.

Page intentionally left blank

SESSION III

CHAIRMAN, *Edwin H. Hilborn*

THE PLASMA DISPLAY—A DIGITALLY CONTROLLABLE, HIGH BRIGHTNESS DISPLAY WITH AN INHERENT MEMORY

R. H. WILLSON
Defense and Space Center, Surface Division, Westinghouse Electric Corp.

INTRODUCTION

Although display technology has progressed rapidly during the past few years, there is at present no satisfactory large tactical display nor is there a satisfactory digitally controllable display for airborne systems or for ground portable van systems (refs. 1 and 2). Recently, an advance in the state of the art was made with the invention of the plasma display (refs. 3 and 4). Although the plasma display does not satisfy as many of the conditions as one would like, it does more closely satisfy these conditions than any other technique (ref. 5). This paper discusses in detail the characteristics of the plasma display. Also covered are the history of the development, cell construction, and the experimental setup used to measure the cell characteristics, a few of the experimental results which shed light on the mechanisms responsible for the cell's behavior, an explanation of the bistable characteristic, techniques for writing and erasing, the current status of the display, and, finally, problem areas and possible solutions.

HISTORY

To minimize the amount of electronics necessary to control an array of discrete light-emitting elements, crossed grid arrays have been made. For instance, for a crossed grid array of $n \times n$ elements, instead of requiring n^2 control circuits, only $2n$ control circuits are needed if the array is arranged in a crossed grid. However, if the light-emitting elements of the array are not bistable only a single element can be selected at a time, an external memory device must be used, and the display must be cyclically refreshed at a 30-Hz rate. Because only one element is on at a time, the average cell on-time is low, so the average cell brightness is much lower than it would be if the cell were on continually. If the elements have a bistable characteristic, the display does not have to be continually refreshed, no external memory is required, and on cells are on all the time thereby yielding the full brightness of the cell.

Under appropriate conditions, a larger voltage is necessary to ignite a gas discharge cell than is required to sustain the discharge (ref. 6). At intermediate voltages the gas cell is bistable and the properties of a light source and memory are combined. Thus, arrays of gas cells are ideally suited for arrays of self-luminous elements because gas cells combine the functions of the light-emitting element and memory element into the same device.

One of the first arrays of gas cells was made by Skellet in 1954 (ref. 7). This design had two orthogonal sets of wires, which were separated by a small distance and sealed in a neon atmosphere. There were 100 horizontal wires and 20 vertical wires. Skellet reported that a discharge could be confined between any crossed pair of electrodes in the array. He did not report on experiments in which more than one cell was fired concurrently. However, Harris (ref. 8), in a memorandum on the Skellet display, pointed out that when a number of cells were ignited, additional unwanted cells would fire.

Moore (ref. 9) made a rectangular array in 1963 similar to that shown in figure 1(*a*). A honeycombed glass panel was placed between two outer glass panels, and parallel transparent thin-film electrodes were deposited on the inside surfaces of the outer glass panels. Air was evacuated from the array and the array was filled with neon. Initially, Moore was interested in igniting only a single spot of light, and in this he was successful. Later, when he tried to ignite a number of cells, he discovered that cells having electrodes in common with the fired cells would also fire.

In 1964, Thompson (ref. 10) added a resistance in series with each gas cell to reduce the coupling between the electrode. He made a matrix of 10 by 10 cells, and demonstrated that any combination of cells could be fired without firing unwanted cells. One set of parallel electrodes was on the inside surface of the outer plate. The other set was isolated from the discharge cells by resistors, which were connected to individual electrodes in each cell. This was accomplished by using one electrical feed-through for each cell. However, fabrication difficulties limit both the cell density and the cell size with this technique, and cathode sputtering and local impurities cause an undesirable variation in the cell's characteristics.

In 1964, Bitzer, Slottow, and Willson (ref. 11) placed electrodes on the outside surfaces of the outer plates. Their first experiments were on a single cell that was filled with neon. They observed that the cell had a large bistable characteristic and that the characteristic was time dependent. The time dependence was attributed to impurities in the cell which were introduced, in part, by a dirty gas-filling vacuum system, so a much cleaner all-glass, bakable, gas-filling vacuum system was made. Using the new system, they observed that with neon alone the cell had almost no bistable characteristic, but with small percentage of a molecular gas added to the neon, 5 percent N_2 for instance, the cell again exhibited a large bistable characteristic. The dependency of bistable characteristics on gas mixture will be explained in "Bistable Characteristics."

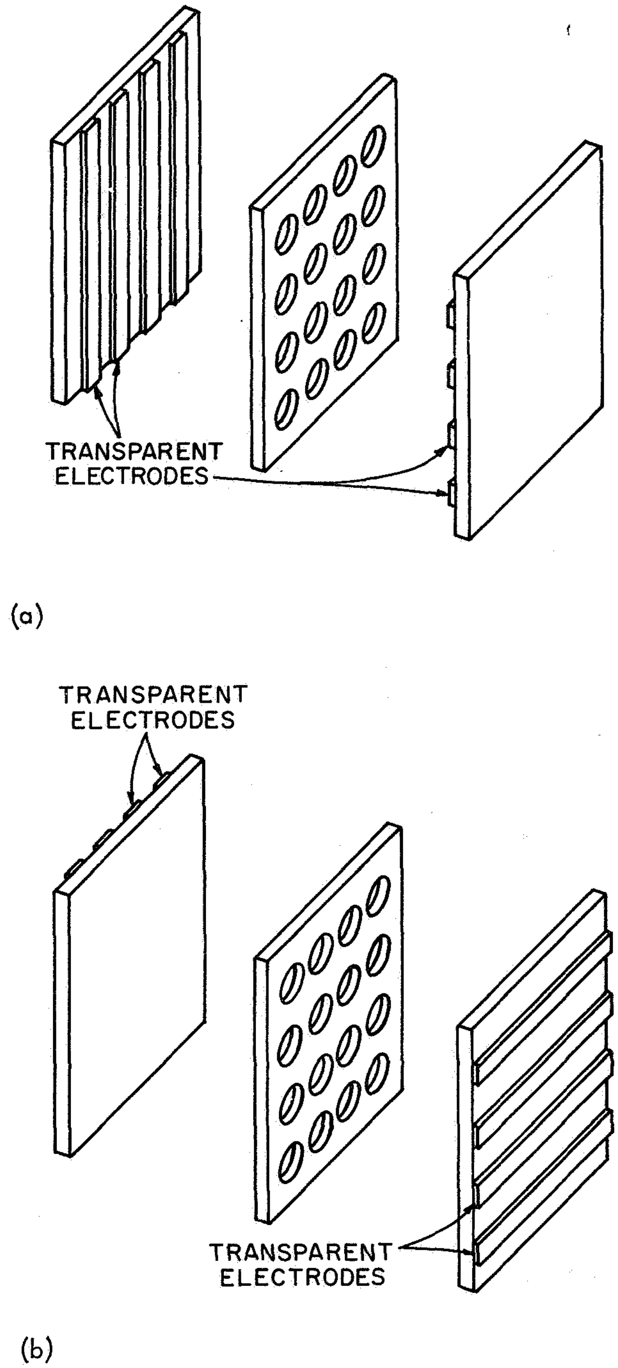

FIGURE 1.—Rectangular array of gas cells. (*a*) Internal electrodes; (*b*) external electrodes.

The first successful array of cells consisting of external electrodes was made by Bitzer, Slottow, and Arora in 1966 (ref. 12). The array consisted of 64 cells (8 rows and 8 columns), 16

(4 by 4) of which were used to demonstrate selective writing of cells; the cell density was 40 cells per inch. (See fig. 1(*b*).) Later that year plasma displays also were made in a number of different industrial laboratories. Work on plasma displays is continuing, some of which is discussed in "Current Status."

Cell Construction

A plasma display is constructed from three pieces of thin, flat glass panels that are sandwiched together. The center panel is honeycombed with holes, and transparent electrodes are vapor deposited on the outside surfaces of the outer glass panels. The two sets of electrodes are orthogonally positioned. (See fig. 1(*b*).)

Each cell forms a small volume which is completely surrounded by glass. Voltages applied to the appropriate row and column electrodes are capacitively coupled into the cell. The capacitive reactance isolates the cells from the row and column electrodes, and any combination of cells can be on at one time.

Usually, the holes in the honeycombed panel are etched; however, early arrays were made by ultrasonic drilling. Typically, the holes are 0.015 in^2, although both larger and smaller holes have been made. The panels are typically 0.006 inch thick.

The sandwich of three plates is sealed to a tube with epoxy. The panel is evacuated and then filled with a neon-nitrogen mixture to a pressure of approximately 700 torr.

Figure 2 is a photograph of a 16- by 16-element plasma display with a symbol written into the display. The cell density is 40 cells per inch, and only sustaining voltages are being applied to the array. The cell brightness was 600 ft-L and the excitation frequency was 300 kHz.

Experimental Results

Measurements of the current and the radiated light of the capacitively coupled plasma discharge cell show that the characteristics of the discharge are dependent on the gas mixture used. For instance, discharges in neon, in helium, and, to a lesser extent, in argon, are relatively slow in forming and are generally characterized by long light pulses that are nearly as long as a half cycle of the sustaining voltage. On the other hand, discharges in nitrogen, and in neon-nitrogen mixtures, are formed very rapidly and are characterized by extremely short light pulses, which can be as short as 40×10^{-9} second in duration. Furthermore, it was found that a significant bistable characteristic occurs only for the rapid discharges (refs. 5 and 13).

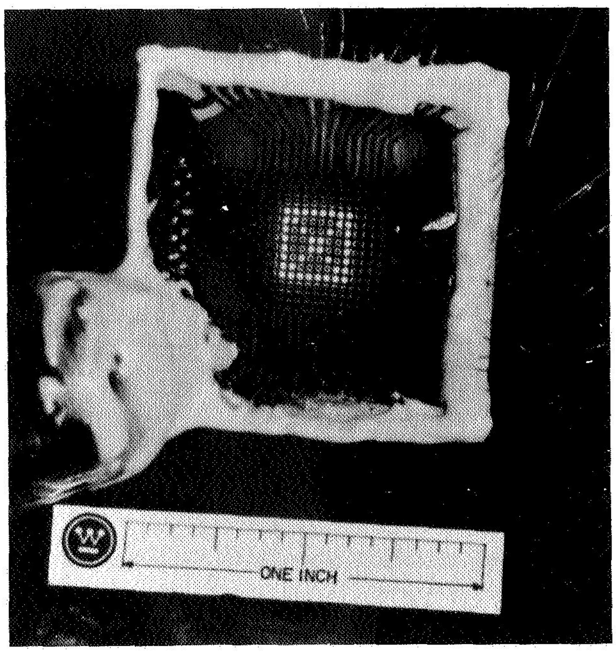

FIGURE 2.—Experimental 16- by 16-element plasma display panel.

Figure 3 shows oscillograms of the sustaining voltage and the light pulse, as measured with a photomultiplier. The light from a number of cells (about 12 cells) was channeled via a light pipe onto the cathode of the photomultiplier. A 400-element plasma display was used; the cells are 0.015 inch on a side and 0.006 inch thick. In figure 3(*a*), the cell is filled with neon only, and the light pulse is rather broad and persists for a large fraction of a half cycle. With 4.3 percent of nitrogen added to the neon, the light pulse is considerably narrower and larger (fig. 3(*f*)). Figure 3(*b*) to (*f*) shows the waveforms for a number of different nitrogen concentrations. The narrower, larger light pulse corresponds to a more rapid, more intense discharge.

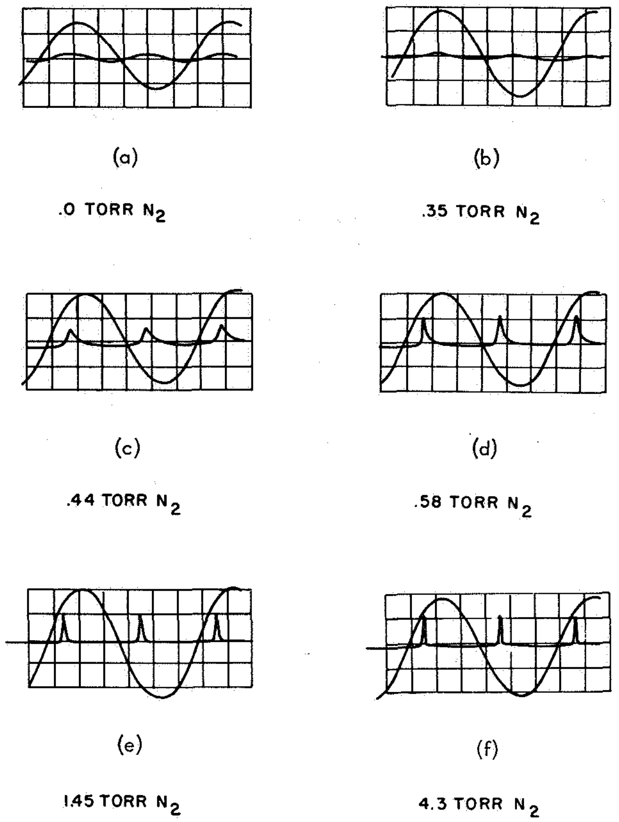

FIGURE 3.—Light pulse and voltage N_e+N_2. Neon pressure=650 torr, square cell r=0.038 cm, thickness=0.015 cm, time scale: 0.5 μsec/div.

A much more rapid discharge is also observed when other gases are added to neon; that is, neon plus carbon monoxide and neon plus water vapor (ref. 5). Likewise, helium plus nitrogen and argon plus nitrogen (ref. 13) exhibit the sharp discharge and a bistable characteristic.

BISTABLE CHARACTERISTICS

It is well known that the bistable characteristics of a dc gas discharge is due to the space charge of the discharge (ref. 6). If the space charge is removed, the full firing voltage must be reapplied to reignite the discharge. On the other hand, for the plasma display with a proper gas mixture, the discharge persists for only a short period of time. However, once the discharge is initially ignited, subsequent discharges occurred even though the applied voltage is reduced. Calculations show that practically no particles remain in the volume between discharges, so a space charge mechanism cannot account for the bistable characteristic. A new mechanism is necessary to explain the bistable characteristic of the plasma display.

Figure 4(a) is a sectional view of a single gas cell and its equivalent circuit. Capacitors C_1 represent the coupling capacity between the external electrodes and the adjacent cell walls, capacitor C_2 represents the capacity of the unfired cell, and G represents the gas discharge.

Let us assume that the cell is filled with a proper gas mixture so that the discharge is very rapid. Let a time-varying voltage be applied

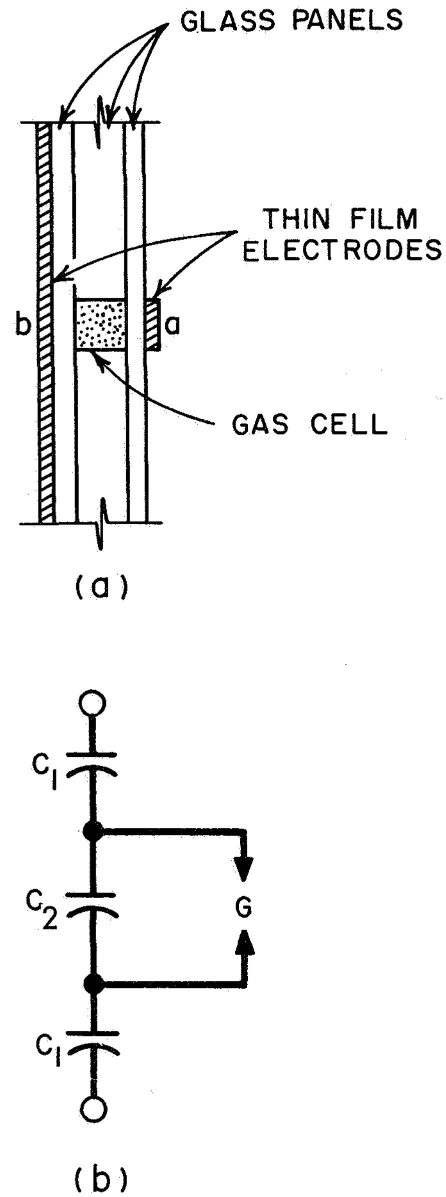

FIGURE 4.—Sectional view and equivalent circuit.

across terminals a–b, and assume that at time t_1 the firing voltage is applied across the cell (capacitor C_2 in our model). Shortly, after time t_1, the discharge starts and rapidly creates a large number of electrons and ions. The charges flow in a direction to minimize the energy of the system, and shortly after the discharge is started, enough changes are present to reduce the cell voltage below the minimum dc sustaining voltage and the discharge is extinguished. Charge continues to be removed from the volume, and due to the charge, a voltage V_0 appears across the cell C_c. If the total voltage difference across the cell does not change sign while charge is in the volume, V_0 will be larger than it would be if there were a sign change. The charge leakage time is long compared with the times between firings, and V_0 is essentially constant between discharges. During the next half cycle, the component of the external voltage which is across the cell adds to the charge voltage V_0 and only V_f–V_0 need be supplied to fire the cell. At any voltage between V_f, and V_f–V_0, the cell has a bistable characteristic, and the state of the cell is determined by the presence or absence of the charge voltage V_0.

The time dependence of the current buildup is strongly influenced by a small increase in the applied voltage above the firing voltage; a small percentage increase in the applied voltage above the firing voltage causes a large increase in the current. Thus, if the slope of the applied voltage at time t_1 is large, V_0 will be larger than it would be if the slope were smaller. This slope dependence is useful in changing the state of cells. (See "Drive and Selection Circuitry.")

It appears most likely that in the rapid discharges a fast electron multiplication mechanism dominates (photon bombardment of the negative surface), while in the slower discharge, a slower electron multiplication mechanism dominates (ion bombardment of the negative surface). A more detailed explanation of this mechanism may be found in reference 3.

Drive and Selection Circuitry

Many techniques are possible for inputting data into the plasma display. The most straightforward method and, indeed, the first method used is to control the voltage on each row and column electrode independently. A second method is to drive the lines in parallel with a single sustaining voltage source and to add pulses to appropriate lines (ref. 13). Each method has its advantages, as indicated in the following discussion.

Fast Write and Erase

The first write-erase scheme, called fast write and erase, makes use of three voltage levels: V_w a write voltage, V_s the sustaining voltage, and V_e the erase voltage. Of course, $V_w > V_F$ and $V_e < V_E$, where V_F is the cell firing voltage and V_E is the cell extinguishing voltage. Through control circuits, either $V_w/2$, $V_s/2$, or $V_e/2$ is applied to a given row electrode, and the corresponding voltage of opposite polarity is applied to the column electrode. Thus, the cell has either V_w, V_s, or V_e across it and is written, sustained, or erased. Of the two techniques, this one is the fastest, but relatively high voltage control circuits are required.

An interesting result has been derived (ref. 12) which relates the variation in the cell parameters and in the control voltages to the memory margin, α. The symbols V_F, V_E, V_e, V_w, and V_s are taken to be mean values, and the voltage range is assumed to be $V_j \pm \Delta V_j$ for each voltage. In general, the voltage range is not symmetrical with respect to the mean voltage but the assumption that it is symmetrical is not too much of a simplification. Using the condition that unwanted cells are not fired or erased, one obtains

$$V_w \geq V_F + \Delta V_F + \Delta V_w$$

$$V_e \leq V_E - \Delta V_E - \Delta V_e$$

$$V_E + 3\Delta V_E + 2\Delta V_e + \Delta V_s \leq V_s \leq V_F - 3\Delta V_w - 2\Delta V_w - \Delta V_s \quad (1)$$

The memory margin α is defined as

$$\alpha = \frac{(V_F - V_E)}{1/2 V_F} \quad (2)$$

Solving for V_E in equation (2) and substituting for V_E into equation (1), one obtains

$$\frac{2-\alpha+(3\Delta V_E+2\Delta V_e+\Delta V_s)}{1/2V_F} \leq \frac{V_s}{1/2V_F} \leq \frac{2+(-3\Delta V_F-2\Delta V_w-\Delta V_s)}{1/2V_F} \quad (3)$$

The minimal acceptable α occurs with equality of equation (3) and one has

$$\alpha_{\min}=\frac{3\Delta V_E+3\Delta V_F+2\Delta V_e+2\Delta V_w+2\Delta V_s}{1/2V_F} \quad (4)$$

If, for simplicity, one further assumes that $\Delta V_F=\Delta V_E=\Delta V_s=\Delta V_w=\Delta V_e$, then one has

$$\alpha_{\min}=\left(\frac{24\Delta V_F}{V_F}\right) \quad (5)$$

Actually, α could be somewhat smaller than that indicated by equation (5) because the numerator of equation (4) is usually less than $24\Delta V_F$. When V_E is the smallest, α is the largest. Now $V_E \leq V_F/2$, so $\alpha \leq 1$. If we assume that ΔV_w, ΔV_e, and ΔV_s are all very much less than ΔV_F or ΔV_E, and letting $\Delta V_E=\Delta V_F=\Delta V$, we have

$$\frac{\Delta V}{V_F}<\frac{1}{12} \quad \text{or 8.2 percent} \quad (6)$$

Thus, with near-perfect regulation of the applied voltages and with complete cell charging, the cell's electrical characteristics must not vary by more than about 8 percent. Even with perfect regulation of V_w, V_e, and V_s, the fraction of these voltages which are applied across the cells is a function of the plate thickness and the local dielectric constant. If one is to minimize the variations in the applied voltages, both uniform plates and regulated voltages must be used.

Slow Write and Erase

A diagram of the control circuitry for the second control technique, designated slow write and erase, is shown in figure 5. A single sustaining voltage source supplies the sustaining voltage to the whole array and all of the lines are driven in parallel. The switching networks

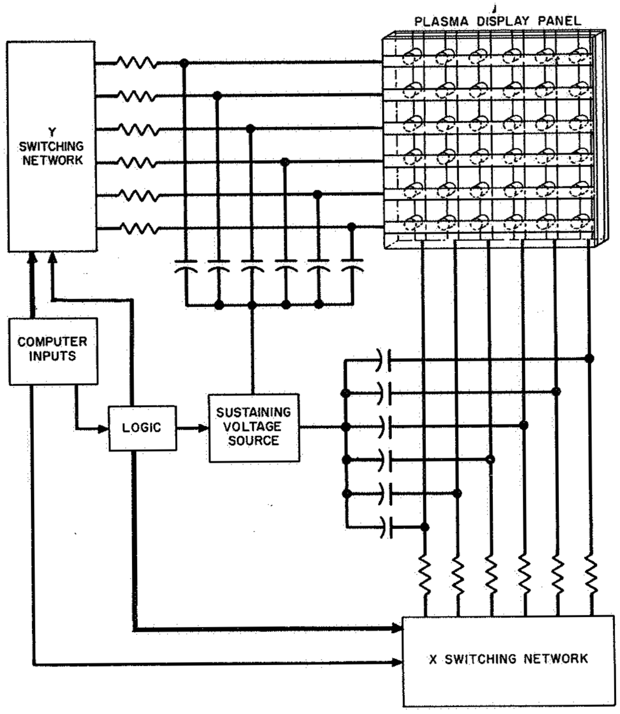

FIGURE 5.—Drive circuitry.

apply voltages to selected lines through resistors R. The R–C network isolates the switching network from the sustaining voltage source, and relatively low-voltage switching transistors can be used. Logic circuitry controls the application of both the write and erase pulses and the sustaining voltage.

Usually, only the sustaining voltage is applied to the array, and cells that are in the one state continue to refire each half cycle, but cells in the zero state do not. To fire a cell, the sustaining voltage for the whole array is removed at a zero crossing of the sustaining voltage and voltage pulses are added to the selected row and column. The sustaining voltage is then reapplied and, with a proper choice of pulse amplitude, the total voltage across the cell exceeds the firing voltage and the cell fires. The cell is now in the one state and will refire each half cycle. The pulses are now removed and the cell "tracks" the voltage because of the slope dependence mentioned in "Bistable Characteristics." This sequence of operations is illustrated in figure 6 (top).

To erase a cell, the pulses are added before the sustaining voltage is removed and the cell

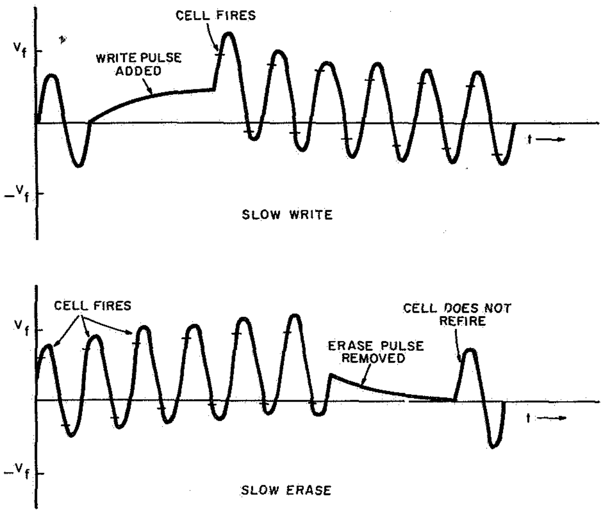

FIGURE 6.—Slow-write erase.

tracks the resultant voltage, as shown in figure 6 (bottom). Notice that just before the erase voltage is removed, the total voltage across the cell is nearly zero. The sustaining voltage is now removed. The sustaining voltage is removed just after the wall is positively charged, and if the amplitude of the voltage pulse is correct, no net voltage will appear across the cell. With the resumption of the sustaining voltage, since the net wall voltage V_0 is zero, the cell will not fire so the cell has been erased.

Measurements show that about 4 half cycles of the sustaining voltage must be allowed for the tracking of the pulse voltages (ref. 12). Thus, at least 2 periods must be allowed for writing or erasing the cell. Also if the sustaining voltage is off too long, some cells might not refire (ref. 13). The permissible off time is strongly determined by the cleanliness of the cells. For relatively "dirty" cells, the off time must be less than 15 microseconds, while for rather "clean" cells the off time can be as large as 1000 microseconds. Typical times are 15 microseconds for the total rise and fall times and 30 microseconds for the switching voltage off time, which gives a 60-microsecond write or erase cycle (ref. 14). If the cells are written one at a time, 16 000 cells can be written in 1 second. For most applications this speed is adequate; if the cells are written a row at a time, 16 000 rows can be written in 1 second.

CURRENT STATUS

Considerable work is being done in the plasma display area. As of August 1967, at least six different industrial and university laboratories were actively doing research and development work on the plasma display. The feasibility of the concept has been established and now the major goal is to make a reliable array of reasonable size with appropriate control circuitry. A number of problem areas must be solved, and some of the more important problem areas will be discussed later in this paper. The status of some of the research and development work is summarized in the following discussion.

Bitzer, Slottow, and their colleagues have confined their work to a 16- by 16-element array. They invented the two write-erase techniques and are presently perfecting them. In addition, they are investigating computer memory aspects of the plasma display and are looking into some aspects of the physics of electrodeless discharges that are germane to plasma displays (ref. 15).

Harris and his colleagues demonstrated their plasma display at a plasma display symposium in June 1967 (ref. 2). This display was 32 by 32 cells with 32 cells per inch. An ultraviolet light was placed behind the display to enhance the firing probabilities of the cells. The display could be written or erased under control of a computer but this was not demonstrated at that time. They used the fast-write technique. Beryllium copper electrodes were used for the transparent electrodes.

Mayer and his colleagues demonstrated their plasma display at the conference also (ref. 2). The array had 96 by 96 cells and had a cell density of 32 cells per inch. It was controlled by a computer (for writing only). They also used an ultraviolet source to illuminate the array. Because of the low firing probabilities associated with the cells it was necessary to cyclically rewrite the display a number of times. The fast-write technique was used. Opaque beryllium copper electrodes, with etched holes filled with highly transparent tin oxide, were used. The display was very dim and was viewed without room lighting.

291-351 O - 68 - 9

Demetrick and his colleagues (ref. 2) have been involved in evaluation of the plasma display work and in the feasibilities of the plasma display for tactical airborne systems.

Our work (ref. 14) has been along a number of lines. We have made various arrays, the largest being 60 by 30 cells, with 40 cells per inch, although most were 20 by 20 cells. Control electronics that are digitally controllable were designed and fabricated. The slow write-erase technique was used. During the write cycle the write pulse is left on for a period of time after the sustaining voltage has been applied so that the cell firing probability is increased. An ultraviolet light is also used. Our first objective was to show feasibility and to develop techniques for making both reliable arrays and less expensive electronics. Uniform arrays have been made, and the first step in developing less expensive electronics has been taken. Our second objective is to consider problems associated with making larger arrays from basic module building blocks, and we are considering various interconnection techniques for doing this. Also, techniques for enhancing the firing probabilities of the cells are being investigated with a view toward the eventual elimination of the uv source. Our third objective is to investigate the possibilities of using the plasma display as an associative memory. This application looks very promising.

Work is currently being done on plasma displays with internal electrodes and on some new and improved gas mixtures. The arrays have operating voltages of 200 volts, a cell density of 32 lines per inch, and a brightness of 300 ft-L.

Work is also being done in the area of color plasma displays and gray-scale capabilities.

Problem Areas

Many problems are associated with making a larger plasma display, such as making uniform arrays, supplying a source of initial electrons, and making inexpensive electronics.

The thickness of the three plates and the holes in the honeycombed plate must all be uniform. The electrodes must be placed over the cells consistently. These constraints are not severe when one makes small arrays, but for larger arrays they pose an increasingly difficult task. It appears that for arrays larger than a few square inches (perhaps 6 or 8 in.^{2}), a modular approach should be used. Modularization is especially attractive for displays in which the cell density can be smaller (tactical large area displays for instance), and a wider space between lines is permitted.

Plates can be made highly parallel, but this will increase the cost of the array. The honeycombed hole structure can be etched fairly uniformly, but as the area gets larger, inhomogeneities of the glass start to influence the structure. Tradeoffs between the difficulties of making larger modules and the difficulties of interconnecting smaller modules should be carefully considered.

Another problem is in supplying a source of initial electrons for cells that are off so that the cells can be fired at any specified time. A uv light is presently used to illuminate the array, and photoelectrons are ejected from the glass surfaces by the photoelectric effect. Work should be done in areas of coatings for the cell to enhance this mechanism. An electrode material should be chosen that is transparent to uv. Of course, the glass surface next to the uv source should also be transparent to uv, and the surface on the front side of the array should be opaque to uv, but transparent to visible light.

It has been mentioned that metastable atoms most likely account for the initial source of electrons between discharges for "on" cells (ref. 3). It might be possible, with the application of an appropriate excitation waveform, to fire off cells often enough to guarantee a supply of excited atoms (metastables), but at a low enough rate so that the discharge is not visible. This firing must be done so that it leaves the walls of the cells in the zero state but does not charge the state of cells in the one state.

The easiest method for guaranteeing reliable writing of the array is to write negatively. The whole array would normally be on, and writing would correspond with turning off selected cells. Cells can be erased very reliably because there is no problem of available initial electrons.

Much work remains to be done in making inexpensive drive and selection electronics. Indeed, it appears that the major expense of the display will be the control electronics. It should be possible to make the line drivers from microelectronics using thin film-discrete chip components, which will significantly reduce the cost of the display.

Conclusion

Some of the more important aspects of the plasma display have been presented. The significant events in the early development of a plasma display were sketched and the cell construction was shown. A typical set of experimental results was presented which illustrates the influence of the cell's characteristics on the gas mixture used. Furthermore, it was shown that with a proper gas mixture a three-capacitor model could be used to explain the bistable characteristic. Two techniques were mentioned which are being used to control the display. The current status of plasma display research and various problem areas was also discussed.

Acknowledgments

Many people have contributed to the development of the state of the art of plasma displays and as far as was practical we have referenced their work. A special thanks is due W. Hoff and D. Bartling who contributed greatly in developing the author's solid-state electronics, and to E. Heitman who developed etching techniques.

References

1. Thomas, P.: Large-Screen Displays. Space/Aeronauts, 82, 1967.
2. Graham, J.: Syracuse University Research Corporation, conference on plasma displays, June 1967.
3. Willson, R. H.: A Capacitively Coupled Bistable Gas Discharge Gas Cell for Computer Controlled Displays. Report R–303, Coordinated Science Laboratory, University of Illinois, June 1966.
4. Arora, B. M.; Bitzer, D. L.; Slottow, H. G.; and Willson, R. H.: The Plasma Display Panel—A New Device for Information Display and Storage. 8th Natl. Symp., Society for Information Display, May 1967.
5. Willson, R. H.: The Applicability of Plasma Displays to Large Tactical Displays and to Digitally Controlled Van Displays. Westinghouse Defense and Space Center, Baltimore, Md., internal report (to be published).
6. Cobine, J. D.: Gaseous Conductors. Dover Publications, Inc., New York, 1958.
7. National Union Electric Corp.: Lattice Videation. Engineering Bulletin, Orange, N.J., Apr. 1955.
8. Harris, F. H.: Private communication concerning work on matrix displays done by Skellet, National Union Electric Corp., Nov. 1954.
9. Moore, D. W.: Gas Discharge X–Y Display Panel. Winter Convention on Military Electronics, PGME of the IRE, vol. 8, 1963.
10. Lear Siegler, Inc.: Development of Experimental Gas Discharge Display. Q.P.R., Nos. 2, 3, 4, 5, 6, and 7, Contract No. Nobsr 89201, BuShips, August 1963–June 1965.
11. Coordinated Science Laboratory, Q.P.R. 31, University of Illinois, Sept.–Nov. 1964.
12. Coordinated Science Laboratory, Q.P.R. 155, University of Illinois, Mar.–Aug. 1966.
13. Coordinated Science Laboratory, Q.P.R. 157, University of Illinois, Sept. 1966–Feb. 1967.
14. Willson, R. H.; Hoff, W. J.; and Bartling, D. A.: Westinghouse Plasma Display Research. Westinghouse Defense and Space Center, Baltimore, Md., internal report (to be published).
15. Bitzer, D. L.; and Slottow, H. G.: Private communication, Coordinated Science Laboratories, University of Illinois.

N68-22315

MULTICOLOR ELECTROLUMINESCENT DISPLAYS

KENNETH P. LALLY
Hartman Systems Co.,[1] *Huntington Station, N.Y.*

By gaining a better understanding of the behavior of a point-source emitter when immersed in a transparent solid of high refractive index, we have recently been able to design electroluminescent (EL) displays capable of generating all three primary colors in a single compact element and mixing them at will to provide seven possible color states.

The resulting displays provide notable features not generally available in the past, the most salient of which are:

(1) A single EL phosphor system is used throughout, which operates at a constant voltage and frequency

(2) Each display element is capable of intrinsically generating or mixing the color primaries, and producing high color saturation

(3) Intrinsic contrast is higher than conventional EL displays

(4) Element-to-element isolation is higher, yet the percentage of display area that can be illuminated is nearly 100 percent

(5) The unusual optical configuration employed permits a number of additional special effects and design options that normally are unavailable in conventional displays.

The theoretical principles governing the design of these multicolor elements are presented, and application to an x–y display structure is discussed.

INTRODUCTION

Electroluminescent displays have always possessed valuable characteristics that are potentially suitable to general aerospace, tactical, and group viewing requirements. By providing a true multicolor characteristic, display flexibility is further increased. This has aroused again the attention and interest of many people concerned with display system implementation.

Two basic principles are exploited in the display element to be described:[2] fluorescent stimulation and dielectric reflection. A conventional electroluminescent lamp is employed as a compact source of short-wavelength visible radiation, which is then used not for direct viewing but to stimulate any of several fluorescent compounds into emission at longer visible wavelengths. The resulting emissions are collected in a common optical channel and then transmitted to the viewing screen for direct observation. A diagram of the display element is shown in figure 1. Three identical blue EL emitters stimulate three different secondary emitters, colocated on an adjacent dielectric slab. The resulting output from these secondary sources is then viewed in a plane normal to the original emission plane of the lamp.

The following paragraphs describe fundamental design considerations and their contributions in determining optimum element parameters.

PROPERTIES OF FLUORESCENT BODIES

Efficient organic fluorescent materials exist which are responsive to short-wavelength visible radiation in addition to the fluorescence nor-

[1] Formerly Huyck Systems Co.
[2] Patents applied for.

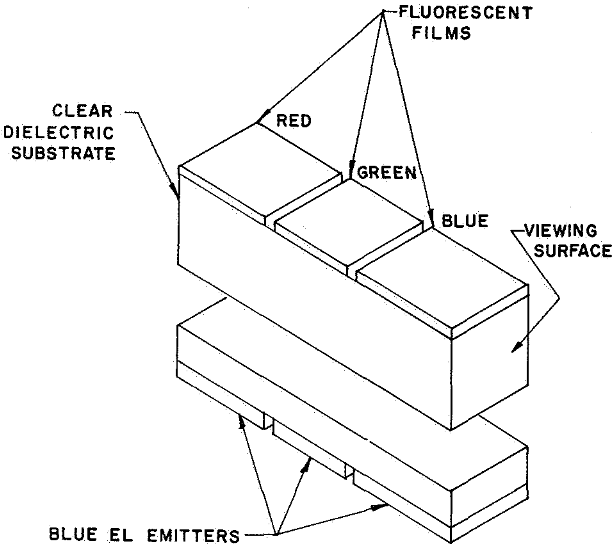

FIGURE 1.—Multicolor element seven-color output.

mally associated with ultraviolet excitation. They possess intense color values, strongly absorb radiation in the green, blue, or ultraviolet, and reemit at longer wavelengths.

Suitable fluorescent materials are found among the aromatic hydrocarbons (ref. 1) and the rare-earth chelates (ref. 2). Because of their high quantum efficiencies, they have been the subject of recent intensive studies for liquid laser applications. They are available in sufficient variety so that samples can be chosen with absorption spectra that are reasonably matched with the emission spectra of the more efficient electroluminescent phosphors.

If incorporated in a suitable plastic vehicle, these materials can often be cast in a thin, specular film (that is, nondiffuse) upon a transparent substrate. In this form a substantial amount of fluorescent light can be seen emitted from the substrate edge, when it is exposed to a stimulating source (fig. 1). Such an effect cannot be obtained if either the coating or the substrate is optically diffuse because the observed light is actually the cumulative flux of all emitters in the coating over a very wide angle, conserved within the substrate by repeated dielectric reflections at the major surfaces until an edge is reached. Obviously, this type of emitter has flux distribution characteristics that are quite different from conventional light sources, and must be understood to be employed properly.

Examination shows that fluorescent molecules behave as isotropic emitters in the plastic medium; that is, their flux is emitted over a solid 360° with even distribution. Thus, at any optical interface with a lower refractive index than the film, a critical angle is established beyond which total internal reflection takes place. Reflections will continue to be sustained between the two parallel major surfaces until a surface is encountered outside the critical angle, permitting escape to the outside world and the observer.

After the light is generated within the film, the governing physics are the same as those for fiber optics, which sustain reflections at the interface between a high-index core and a low-index cladding. In this instance, however, air is simply employed as the low-index cladding. Normally, optical fibers would suffer severe losses and cross-coupling without any solid low-index cladding. However, in this instance the deliberate omission of a solid cladding is only possible because the size range of interest for these display elements is generally not finer than 30 per linear inch, and depth is usually limited to a few inches: this is quite gross for an optical fiber. Thus, there can never be any extensive optical contact with adjacent structures, in comparison to the fiber size, nor will dust or handling have any significant effect on total reflection efficiencies. Well-made optical fibers exhibit transmission efficiencies of 99.5 percent per inch for thousands of sustained internal reflections, so the mechanism is quite efficient.

If the emission properties of this isotropic radiator are compared with a conventional lambertian radiator, it is found that an EL lamp, for instance, cannot trap any significant amount of light within its faceplate, because the faceplate is optically immersed in a scattering diffuse surface (the phosphor itself) which spoils any potential internal reflection. This, incidentally, gives rise to the familiar problem of halation, wherein a ray is critically reflected at the air interface of the polished faceplate, but scattered on the next bounce from the phosphor plane.

Thus, a transparent fluorescent emitter in a high-index medium can efficiently trap, or pipe,

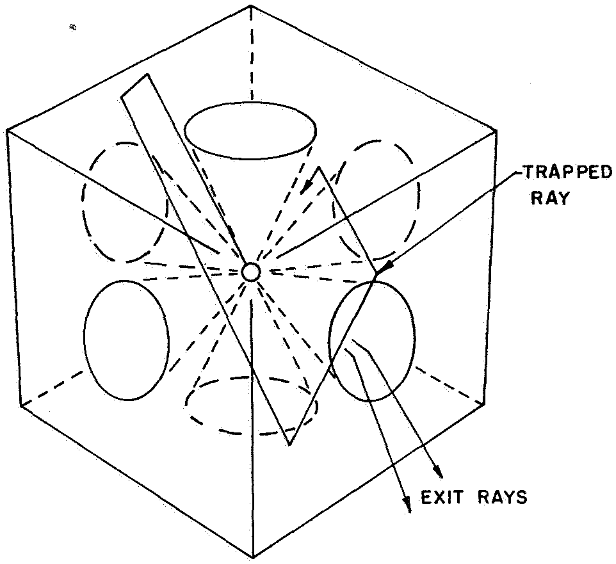

FIGURE 2.—Trapping and loss patterns for a point source. Isotropic emitter in high-index media, bounded by a rectangular parallelepiped.

light and an EL lamp cannot. One may conclude that a transparent fluorescent body could be employed as a radiation pattern converter for a lambertian source. A few simple calculations will show that for a typical optical grade plastic with a refractive index of 1.6, the amount of fluorescent flux which naturally escapes through each face (that is, outside the critical angle for internal reflection) of a rectangular parallelepiped bounding the emitter and coupled to air, is about 11 percent of the total stimulated flux. Thus, 66 percent escapes through the six cube faces and 34 percent remains trapped, reflecting endlessly within the bounding space (fig. 2).

If the internal reflection on any one surface is spoiled by diffusing it, typically 70 percent of the flux incident on that face will be transmitted, yielding 0.7 (0.11+0.34) = 0.315. If any one of these cube surfaces is chosen as a viewing area, approximately 10 to 30 percent of the total fluorescent flux will be available for display use, depending on how the viewing screen is finished.

If the cube is extended in any dimension (as shown in fig. 2), the preceding figures still apply: 11 percent of the flux will still be emitted from each face. Obviously, if the flux per face is constant but the face dimension can be changed, the flux per unit area, or brightness, is a design variable that may now be controlled by dimensions. If the isotropic source is viewed from the smallest face of a bounding rectangular parallelepiped, this face will be the brightest, and the brightness is completely within design control, providing design freedom rarely possible. This feature is so unusual that it is worth restating in more formal terms.

Given an isotropic source of constant power immersed in a medium of higher refractive index than the surrounding space, there is ideally no limit to the brightness one can obtain. In this extreme, for increasing length and decreasing exit area, diffraction effects limit the ultimate brightness. For practical purposes, the real design limits encountered are due entirely to the familiar, nonideal properties of optical materials, that is, bulk absorption and scattering in transparent media, optical surface finish, and the efficiency of dielectric reflection, etc. Some of the properties of isotropic emitters are being exploited for real value in ruby laser rods and rare-earth doped glass rods, which, however, further constrain fluorescent emission to an on-axis mode only, via the resonant cavity.

DESIGN CONFIGURATIONS

If a transparent substrate is coated with films of several different fluorescent compounds (as in fig. 3), each particular film will emit at a characteristic wavelength in the substrate. If, for simplicity, it is assumed that films and sub-

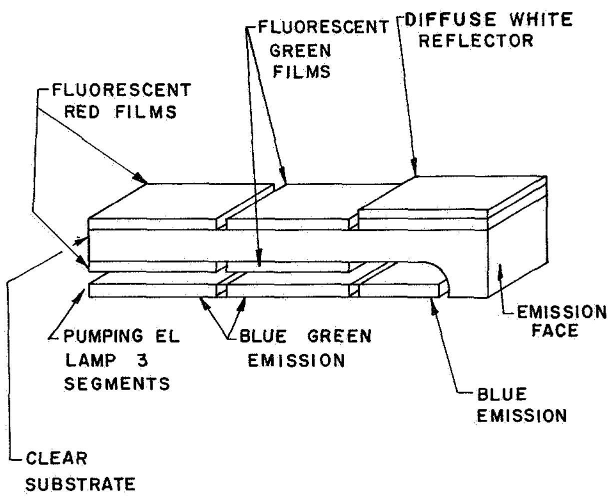

FIGURE 3.—Flared seven-color element.

strates are matched for refractive index, they behave optically as one; therefore, all radiation generated in the surface film is free to transmit through the supporting substrate without interference. If the films are very thin compared with the thickness of the substrate, the greatest part of the path length for totally reflected rays is through the clear substrate, while only a small fraction of the path is through the actual films. Therefore, many different wavelengths can be efficiently reflected down the length of the substrate without suffering significant absorption in the cladding films, regardless of their transmission characteristics, as long as they are not diffuse. This would not be true if the fluorescent materials were incorporated *in* the transparent substrate, because each has its own characteristic transmission spectrum that would severely attenuate other wavelengths.

The major advantages of employing a fluorescent film as a radiation pattern converter for EL illumination are as follows:

(1) Design control of brightness is possible, independent of source EL brightness

(2) Control of output wavelength is possible by choice of fluorescent materials

(3) Multiple wavelengths can be obtained as an output with high efficiency.

Because all the fluorescent materials obey Stokes' law by absorbing short-wavelength radiation, the arrangement shown in figure 3 is deliberate. If the viewing surface is located at the right, then the longest wavelength emitter is placed to the left so that it will not be absorbed to any significant degree by shorter wavelength emitters, or stimulate fluorescence in them when it is activated. For the sake of simplicity in the present analysis, it is chosen to allow the 55-percent loss of light out of the five nonviewed faces. With regard to the light lost from the sides of the element, once the ray couples to air, by the basic laws of refraction, it cannot be subsequently trapped in any adjacent element with coparallel surfaces and, therefore, contrast is not significantly degraded. In certain design configurations it is possible to recover some of these losses by employing either dielectric reflection from nonparallel surfaces or metallic reflection. The present analysis will be confined to the simple case.

A segmented electroluminescent lamp whose phosphor emits at some short wavelength is placed adjacent to the fluorescent-clad substrate in figure 3. Other types of sources, such as gas-discharge elements, can also be employed to some advantage, depending on design goals. Currently available EL phosphors based on ZnS generally will not emit below 4600 Å. The lamp is segmented to illuminate the individual films clad on the substrate, stimulating their characteristic fluorescence, which is then viewed at the right-hand end of the assembly. Illuminating more than one lamp results in direct mixing of the stimulated wavelengths within the common channel. As a result, by simply keying the three lamps on or off, one can generate red, green, or blue primaries; yellow, cyan, or magenta secondaries; and white.

DESIGN FEATURES

Brightness

As discussed previously, about 30 percent of the generated flux will couple to the viewer through a diffuse face. Keeping the exit face dimensions constant, one can increase the total flux in the substrate by increasing the length of the free dimension: doubling the length will nearly double the brightness. The brightness of an EL lamp B_E is generally a function of voltage and frequency. The brightness of a fluorescent/EL combination B_F is, additionally, directly proportional to the area over which the film is optically pumped A_P, inversely proportional to the exit area A_E, directly proportional to the conversion efficiency K (quantum efficiency × absorption), of stimulating light to the fluorescent light, and directly proportional to the coupling coefficient X through the viewed surface to the observer. Then

$$B_F = f\left(v,\ f,\ \frac{A_P}{A_E},\ K,\ X\right)$$

For the materials of interest, K is quite high, in the range of 0.5 to 0.8. As we have seen, X will range between 10 to 30 percent, depending on screen design; thus

$$B_F=(B_E)\ (0.5)\ (0.3)\ \frac{A_P}{A_E}$$

for a diffuse screen, in which case the units of radiation of the EL lamp and the fluorescent element can be compared directly in foot-lamberts (because both are now lambertian). If the pump area is the same in length as the exit area is in height, then $B_F=0.15\ B_E$. If the pumped film is seven times as long as the exit area is high (both have the same coextensive width), the brightness $B_F=B_E$. Typically, if a display element were 0.033 inch high, an element 1 inch deep would provide a brightness 4.5 times greater than the illuminating EL source (with suitable adjustment for any change in sensitivity for the output wavelength).

Color Balance

Different fluorescent materials exhibit different quantum efficiencies, absorption coefficients, etc., and, furthermore, their characteristic emission, depending on wavelength, does not stimulate the eye to the same degree. For instance, a unit radiant in red will stimulate the eye only about 30 percent as much as a unit radiant in green. Even if the two films were equally efficient, how can they be brought into balance on the display face? Because brightness can be controlled as a design variable, three times the area on the substrate is allocated to the red film (and its coextensive lamp) as compared with the green film.

Contrast

Ambient.—The design configuration automatically removes the electroluminescent phosphor from the direct field of view, particularly if an expanding substrate is employed (as shown in fig. 3), which gives a convenient shoulder behind which the lamp and any associated control elements may be concealed. If a transparent view surface is chosen rather than a diffuse one, there are no other diffusely reflecting surfaces in the element, and thus contrast under ambient lighting is much better than for conventional electroluminescent displays. The expanded end of the substrate further permits the introduction of materials that will absorb all or selected portions of the incident ambient to avoid the possibility of stimulating fluorescence in the element from outside sources, without significantly affecting the characteristics of emitted display light.

Element to Element.—The mechanism of dielectric reflection permits these substrates to be stacked extremely close to one another (within a few wavelengths of light) if they are not in extended optical contact. Any ray so disposed as to leave one element will merely transmit through the next element at the same entrance/exit and refraction angles, without being trapped. Extramural cladding can additionally be employed, with proper precautions, to yield the same contrast/crosstalk benefits available to fiber optic bundles.

Active Display Area

The expanded substrate end, plus the isolation provided by dielectric reflection, yields a display surface whose active area is considerably higher than that available for conventional EL displays (no interelectrode gaps). Because the element dimensions are large compared with normal optical fibers, little space is devoted to the low-index cladding of the substrate, which may often be just air.

Color Saturation

Fortunately, most fluorescent materials emit over a narrow spectral range and thus their apparent color saturation is high, approaching 80 to 90 percent more, in most instances. Figure 4 shows some of the values that have been obtained in multicolor elements in the past on a CIE chart. The uppermost point plotted represents a green whose saturation level has been increased by the addition of a passive filtering dye.

Radiation Pattern

It has already been seen that an isotropic radiation pattern permits a specular display surface to be employed for high contrast, or permits rediffusion of the display surface to reestablish lambertian emission. Certain other effects can be obtained with isotropic radiation that are not possible in a conventional display. The substrate end may be capped with a small molded lens for certain special effects, or it may be cut on a bias in several different fashions,

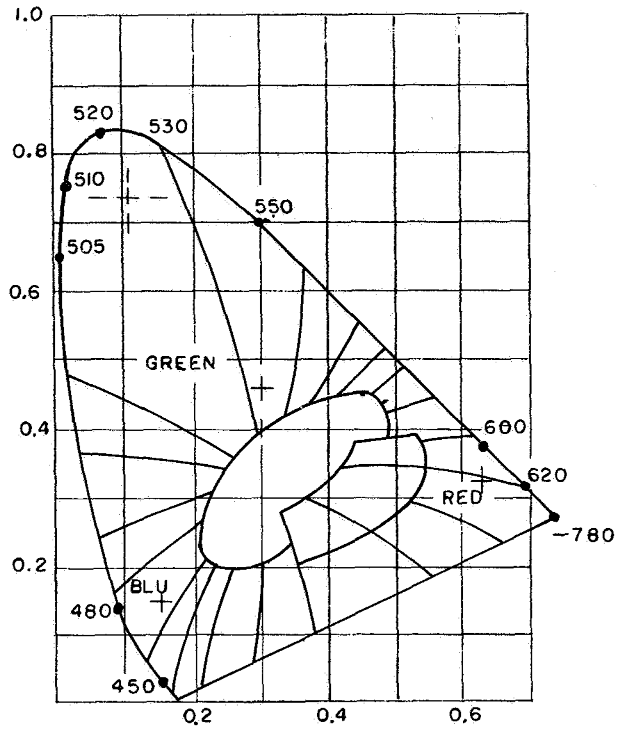

FIGURE 4.—C.I.E. chromaticity diagram showing multicolor primary colors.

	x	y
Red	0.63	0.33
Green	.30	.45
Blue	.16	.15

nonnormal to the other two substrate axes. Such a bias cut, which has been dealt with in fiber optics literature, will tend to make most of the exiting flux bend toward the tip of the cut, thus preferentially controlling the angular distribution of light. Changing the substrate design to one with converging instead of parallel surfaces, stepped surfaces containing parallel or nonparallel segments, etc., will also alter the flux distribution. Combining special surfaces in the substrate in one axis with bias cuts on the substrate end in another axis will provide a measure of control over both vertical and horizontal emission distribution. All these results may be simply analyzed by ray tracing, reflection, and refraction, and will provide design options not available in conventional lambertian displays.

Experimental Results

To date, linear arrays (column indicators) 0.75 inch wide employing two primary colors have been experimentally fabricated at a resolution of 30 lines per inch. With a 35-ft-L EL source and a 4-inch assembly depth, the brightness obtained was 110 ft-L in the green and 35 ft-L in the red. More recently, x–y arrays in modular form at a resolution of 10 elements per inch have been investigated in 2-inch-deep assemblies, yielding an integrated white output of 20 ft-L. Investigation is also proceeding into x–y displays having 30 elements per inch.

Certain design compromises must be employed to obtain blue as a display primary because blue excitation by ZnS of a blue-emitting fluorescent agent is not very efficient. Ideally, it would be most advantageous to employ a uv-emitting EL phosphor, but because they have only been observed experimentally, a blue-emitting phosphor is used, and a diffusely reflecting surface is carefully introduced into the element at this point. This serves to efficiently couple blue emission directly from the lamp into the substrate, and if proper design precautions are taken, the loss of previously trapped red and green light at the new diffuse surface can be held to an acceptable level. In such designs, blue intensity cannot be increased beyond the intensity of the source lamp, but it can be nearly equaled. If a balanced-primary display is desired, the blue brightness level of present EL phosphors becomes the limiting factor.

Conclusions

By combining fluorescent stimulation with dielectric reflection, true multicolor displays can be fabricated, providing brightness levels generally superior to conventional electroluminescent displays at higher intrinsic contrast. The design requires modest structural extension into the depth dimension to obtain these properties. Major benefits are foreseen in the additional design options presented, in the degree of freedom obtained from EL phosphor brightness limits, and in the ability to obtain balanced primaries with a single phosphor system, thus providing a single brightness/life characteristic for the entire display. Organizational characteristics still permit addressing by all relevant x–y address techniques, with the addition of z-axis modulation for color (or alternatively x–$3y$ for two-axis address). Structural design still permits sufficient space for the addition of various control elements to the matrix for continuing memory/persistence studies.

Acknowledgments

Acknowledgment is given to John N. McGuire (now of TRG, Inc.) for first recognizing the potential contributions of fluorescence to displays, and for many hours of patient listening in the development of theoretical models; and to Joseph F. Lidoski and James V. Masi for their persistence and enthusiasm in reducing theory to practical constructions. The work described herein was supported, in part, by Rome Air Development Center under contracts AF30602–4137 and AF30602–67–C–0307, by Wright-Patterson Air Force Base and NASA Manned Spacecraft Center jointly under contract AF33615–67–C–1842.

References

1. Pringsheim, P.; and Vogel, A.: Luminescence of Liquids and Solids. Interscience, New York, 1955.
2. Lempicki, A., et al.: Chelate Lasers. U.S. Government Printing Office (ASTIA No. AD 632178), Apr. 1966.

Page intentionally left blank

CONCLUDING SESSION

N68-22316

CONSIDERATIONS FROM ENGINEERING PSYCHOLOGY

R. C. Casperson
Dunlap & Associates, Inc.

The engineering psychologist is concerned with two classes of requirements in the development and implementation of display media. The first considers the information required by the human operator to perform effectively in a system. The second defines the human factors that are important for the effective transfer of this information via the display to the human operator.

Several studies are reviewed and specific visual phenomena are discussed which point out some of the limitations of the available data concerning human visual performance and the faulty conclusions that can derive from the uncritical use of certain "cookbook" data.

Some general guidelines are suggested for use by media researchers and display designers to assist them in meeting their ultimate goal: the transfer of useful information to man in a form that is compatible with his sensory-perceptual capabilities.

Introduction

Ten or fifteen years ago an engineering psychologist could acquit himself as the guardian of the human user, before a group of engineers and physical scientists engaged in the development and design of displays, by simply summarizing and interpreting the available textbook material concerned with the traditional parameters of visual performance such as brightness, contrast, color, and duration as they determine display effectiveness. This is no longer the case for two reasons:

1. Today's research-and-development engineers working in the field of information display are generally more sophisticated concerning the human factors related to visual performance than they were 10 or 15 years ago.
2. Despite some of the deceptively simple relationships described in handbooks and source materials, the visual mechanism is an intricate system which cannot yet be fully understood in terms of the physics and chemistry of the eyeball when stimulated by a complex signal, such as we find in most of our advanced display concepts.

The increased sophistication of today's display researchers and engineers in the human factors concerned with their ultimate objectives is illustrated by the contents of the papers that have been presented in this seminar. Although this is the only paper scheduled for discussion of engineering psychology considerations, the human-factors problems attending the use of each of the media discussed have been touched on in some fashion in the foregoing presentations. When we look at the broad spectrum of professional backgrounds represented in the membership of the Society for Information Display and the subject matter dealt with in its publication, this should not surprise us. The amalgamation of physical, chemical, and behavioral sciences that has been accomplished by the Society for Information Display has been a significant factor in the integration of engineering psychology into the basic process of designing displays for human use. Poole's book (ref. 1) not only confirms this fact but demonstrates that a single author can deal comprehensively and understandably with both the engineering and the human factors concerned with information display.

The remainder of this presentation will consider some of the specific reasons why a simple review of the basic principles of the visual process no longer satisfies the needs of the designer of information displays.

General Considerations

The general area of concern for the engineering psychologist must derive from the requirements of the human user. However, there are two kinds of user requirements that must be considered, and they stem from different origins. The information requirements posed by a specific system come from its mission objectives, while the human factors requirements derive from man's visual and perceptual capabilities in performing specific tasks.

Information Requirements

The first three presentations of this symposium dealt with the media requirements for a select segment of the possible user population; for example, aerospace systems. The accepted procedure for establishing these requirements is to begin with a mission description, which can be analyzed into flight phases or functions. These phases can then be broken down into operations or procedures, which imply specific requirements for information, decisions, and activities on the part of the human operator in the total system operation in the accomplishment of its mission. The engineering psychologist should be an integral member of the team which prepares this set of requirements. Given the responsibility and authority, he would prepare a document often referred to as an operational sequence diagram (OSD), which is a form of flowchart of man-related activities, over time. The OSD is a tool that first organizes the analyst's thinking, much the same as a PERT chart might do for the production manager, and thus provides him with a systematic method for understanding and describing the details of system operation. Second, if carefully done, the OSD assists significantly in specifying the requirements for information displays and, equally importantly perhaps, the constraints and conditions under which they must be used.

Engineers often rely on a simple inquiry of the operational user concerning his need for information, or they rely totally on their own insights concerning requirements for display. Engineering psychologists have found that the first alternative can be quite frustrating and often misleading. Operational people often find it very difficult to express their needs in terms of the basic elements of information. Pilots, for instance, often cannot tell us directly what information they need in order to fly. Rather, they will resort to answers such as "a better altimeter," or "a more sensitive gyro horizon." Unwittingly, such responses may constrain and even mislead the display designer into making an "improved" version of what is already there, rather than encouraging him to look at the problem of keeping the pilot informed of his position and motion relative to three-dimensional space.

On the other hand, the engineer who is creating or working on a system may feel that he has sufficient understanding of a system operation to set the requirements, and perhaps even design the displays necessary for the human operators to perform their functions. This approach can also lead to faulty design decisions, for at least two diametrically opposite reasons. First, unless the engineer goes through a rigorous analysis such as that typified by the OSD mentioned above, or has performance experience of the kind ultimately required of the operators (in the total operational situation), it is almost impossible for him to anticipate the conditions which ultimately will prevail. On the other hand, suppose the engineer is qualified to make the judgments concerning the conditions under which the system operators will perform and thus can establish the "logical" needs for information display. He will then implement these display requirements based on this same supersophistication concerning his system. As a result, there is an exceedingly good chance that the poor operator, without his background and perhaps with a much lower native intelligence, will *still* be shortchanged.

There are many examples of the first kind of problem in military aircraft instrumentation today. Many have experienced the results of the latter situation: because of unfamiliar equipment you were at a loss to interpret the readouts, but if you were as familiar with it as the man who designed it, the displays might have been adequate.

Human Factors

The second set of user requirements of concern to the engineering psychologist are the human factors, which deal with the interface between the man and the display, and therefore directly involve the media and techniques that are used. After the detailed requirements for information have been stated by the user, whether they be derived by analytical techniques, such as the one suggested above, or by arbitrary decision, the display designer must decide how best to provide them within the state of the art of the media he has at his disposal. If he is unable to satisfy the requirements with the media and techniques available, he must do his best through research and development to upgrade the state of the art, or he must make tradeoffs among the possible characteristics of his display concept and the stated needs for information. Regardless of the alternative he chooses, he should be constantly aware of the implications of his design decisions for the ultimate human user.

The following considerations should be implicit in the design decisions required in the development of a display concept, regardless of the medium used for its implementation.

(1) A valid statement of information requirements should be established by or for the user, including specifications for data rate, legibility, format, coding, etc.

(2) These information requirements should be translated into appropriate values for the parameters which define the media or techniques available to the designer for implementation.

(3) The values of these parameters available to carry the required information should be assessed against the visual capabilities of the human operator to determine how well his information needs are met by the medium and/or techniques available.

(4) Where shortcomings are apparent, optimum tradeoffs or compromises should be made to provide the best overall solution to the information deficit, and research and development efforts should be undertaken to satisfy the specific requirements still unsatisfied.

These four statements are obvious oversimplifications of the problem. It is not a simple task to obtain firm and valid information requirements. It is much easier for the user to simply ask for "everything." Silver and Cruikshank (ref. 2) have pointed out that the tendency is to ask for whatever the state of the art will give rather than what is actually needed. They cite the frequent requirement for fast data rates, when no objective requirement appears to exist. Their example is a good one, because in computer-generated systems, as in any other, excessively high data rates will create penalties in other aspects of the display output. The point to be made is that conservatism, if not austerity, should be encouraged in those responsible for stating the information requirements. Caution at the outset may result in less chance of dissatisfaction with the final product.

In addition to a simple statement of the requirements, a description of the conditions of use should be obtained. Display engineers are sufficiently sophisticated these days to ask the user what the ambient light level will be in the area surrounding a display that he has been asked to design. What he may neglect to find out is the kind of light that will be provided. A CRT or electroluminescent display with a refresh rate of 50 cycles would, under most conditions, be sufficiently above the critical flicker frequency (CFF) to preclude the appearance of flicker. If the display is installed in a room lighted by a single 60-cycle fluorescent lamp, we might find that the interaction of the two frequencies produces a visual beat that is well below the flicker threshold and can prove to be quite annoying. Often such problems can be solved simply if they are anticipated.

When we consider the next step of converting the information requirements into values of the parameters that define our display concept, another problem may arise. Because any useful display system ultimately will be produced in

291-351 O - 68 - 10

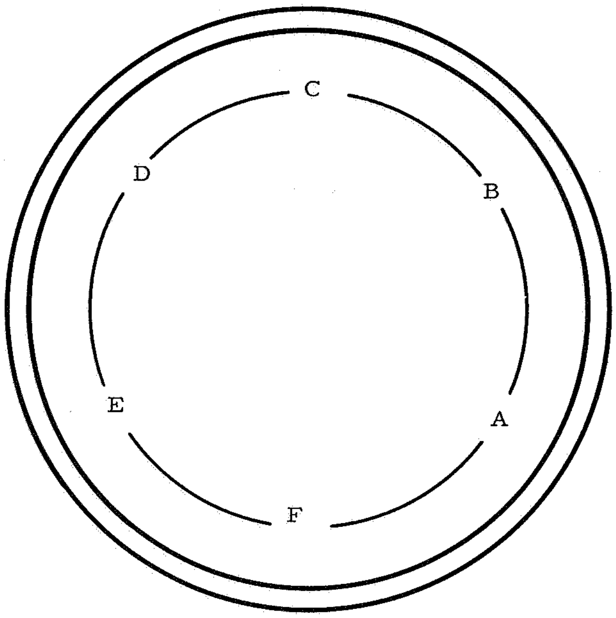

FIGURE 1.—Brightness measurements made at six locations on a commercially available CRT with a P–7A phosphor. Values represent average of 12 measures taken at each point (in ft-L): $A=11.40$, $B=9.10$, $C=8.85$, $D=10.80$, $E=7.60$, $F=7.00$.

quantity, the media and materials used in their implementation will be fabricated by some production process. When we define the capabilities of a display in terms of manufacturers' specifications for a particular phosphor, pigment, or filter, we must consider what happens in the real world of manufacture. In the study described later, Turnage (ref. 3) found that the P–28 phosphor on a CRT he planned to use for experimentation deviated so much from the manufacturer's specifications that he could not use it. In a study currently being conducted for the Navy by Dunlap & Associates, Inc., a commercially available oscilloscope was to be used as a display for some visual performance measures on a simulated sonar problem. A circular sweep was generated at 628 rad/sec and maintained at a constant bias voltage (diameter of the circle was 10.3 centimeters and sweep width was 0.56 millimeter). Twelve brightness measurements were made at each of six locations around the circumference of the sweep, as shown in figure 1.

The average values for the 12 readings in footlamberts for each measurement point are also shown. As you can see from these data, at one point on the circle the brightness of the scope face varied more than 50 percent of the smallest value obtained, and generally proved to be quite variable over the circumference. This is a significant difference under any circumstance, but it is especially critical when we attempt to obtain threshold measurements. If we intend to measure or predict visual performance using a display such as this, we should not assume that the output over the display surface is homogeneous. However, once the output characteristics of a particular medium or technique can be defined and its reliability can be specified, the ability of a display to transmit information to the man in a useful form depends on the capability, limitations, and idiosyncrasies of his visual mechanism.

THE VISUAL MECHANISM

Man's ability to detect visual signals is sufficiently understood in terms of the optical, photochemical, and neural events that occur within the eye itself. At this level we can explain his ability to detect electromagnetic energy that is within the visible spectrum and at sufficient absolute or differential intensities. Furthermore, the principles that apply to the basic visual discriminatory processes for relatively simple signals (that is, varying in a single dimension) are well documented and are available to the display designer (refs. 4 to 7). However, when we attempt to predict reliably an observer's response to the complex signals resulting from the majority of the media or techniques described in the preceding papers, we find that we quite frequently do not have the data to provide the answer. What can be even worse is that uncritical use of published data, especially that found in some of the more popular handbooks, can be seriously misleading when directly applied.

An example of this latter problem can be found in the faulty interpretation of some work conducted by Bartlett and Williams (ref. 8), and summarized in a later report by Williams (ref. 9). They were interested in determining the effect of image (target) size on visibility using a PPI presentation. In their study they

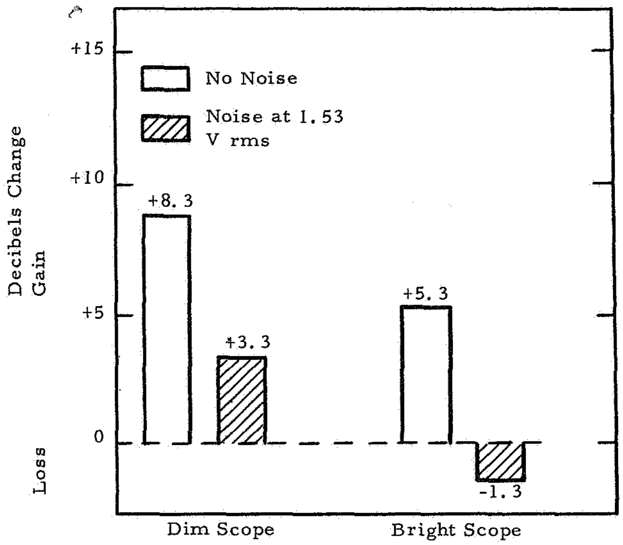

FIGURE 2.—Change in visibility for 6-inch viewing distance vs. 24 inches as standard for two brightness levels.

used a 7BP7 CRT with a viewing surface of approximately 6 inches in diameter. One of their methods for varying target size was to change the operator's viewing distance. They used two distances: 6 inches and 24 inches, and two variations in conditions of presentation; that is, low-high background brightness, and noise-no noise. A graphic summary of the results they obtained is shown in figure 2. The zero point on the graph represents the visibility obtained at 24 inches as a standard for two brightness levels.

These data show an increase in visibility of targets for the 6-inch viewing distance, regardless of background brightness when there is no noise. There is still an advantage to the 6-inch viewing distance under dim conditions, but there appears to be a small, though insignificant, difference in favor of the larger distance, under bright conditions, with noise. The authors themselves used considerable caution in interpreting their data, and carefully pointed out the limiting conditions used in this experiment. However, for several years a 6-inch viewing distance for CRT's was preferred, regardless of size, utilization, or viewing conditions, based on the results of this research. This interpretation of the data completely neglects the purpose of the study and the severe limitations in the data resulting from the conditions that were used. Because a fixed scope and target size were used, it is not surprising that viewing distance is directly related to visibility. With only two viewing distances so widely separated, we might, rather, be surprised that the superiority of the shorter distance was not more clearly demonstrated. However, there is little basis for applying the results obtained from these experiments to situations involving larger scopes, especially when target detection, not simple visibility, is the criterion. Where search is required, we can be sure that there is an optimum viewing distance related to scope size, partially unrelated to the size or brightness of targets. Furthermore, even though small, the reversal obtained with noise present on the scope should have been a warning to anyone attempting to apply such laboratory data to the operational situation, where noise is most frequently an integral part of the display output.

A brief review of several of the less well understood visual phenomena and some recent research studies will provide examples of areas where definitive data are still scarce and where the direct application of basic visual performance might result in faulty design decisions.

Flicker

When we consider that much of the basic data concerned with the human response to intermittent light or flicker derive directly from experiments begun by Helmholtz over 100 years ago, we must use care in the extrapolations we make to present and future display media. He used a rotating disk with black and white sectors as a stimulus, and bright sunlight versus bright moonlight as his experimental conditions, and determined that the critical frequency of intermittent stimulation for the human eye reached approximately 50 to 55 Hz at higher brightnesses, and might be as low as 5 cycles at very low light levels.

Of course, research into the problem of man's response to flickering light did not cease with the investigations of Helmholtz and the early psychophysicists. Much has been done experimentally to quantify the relationships between man's CFF and the intensity, spatial, and temporal parameters of the signal (ref. 10). However, Turnage (ref. 3) points out the shortcomings of the published data for CRT application and emphasizes the importance of phosphor persistence in establishing CFF for cathode-ray tubes. In order to relate the effects of phosphor persistence on CFF, he ranked six phosphors for the amount of residual light output after 33 milliseconds, as listed below, rather than by the conventional methods of specifying time of decay to 10 percent of peak output.

Phosphor:	*Residual, %*
P–28	[a] 85
P–12	70
P–7 (yellow component)	45
P–1	4.0
P–4 (silicate)	1.3
P–20	.1

[a] Not used in experiments.

The subjects varied the refresh rate of a spot 5/32 inch in diameter at six comparable brightness levels for each of the phosphors in question. They were asked to report when flicker appeared while reducing the rate, and when it disappeared in the other direction, using a modified method-of-limits technique. Except for the P–28 phosphor, which proved not to have the decay characteristics specified for it, Turnage obtained a direct and consistent relationship between persistence and the flicker fusion frequency (as determined by accurate published persistence curves) with the lowest frequency required for fusion with the longest persistence phosphor. The CFF for all phosphors was found to be lower than the values obtained by standard techniques (averaging 33 Hz at 10 ft-L and about 44 Hz at 100 ft-L).[1]

Anyone familiar with the functions of the visual mechanism should not be surprised by the results obtained by Turnage. Because the eye is known to integrate light energy over time, it is logical that with a given refresh rate, longer persistent signals will tend to lose their individual identities at lower frequencies. However, in addition to other quantitative relationships pointed out by Turnage, this study is important because it illustrates the difference between the data that are directly applicable by the display designer and the data he might find available in basic source material dealing with visual flicker phenomena.

[1] Author's estimates from published data.

Apparent Movement

The appearance of motion in the visual field, without objective movement of the visual stimulus, occurs under selected circumstances.

Autokinesis

Autokinetic movement describes that apparent motion of a stationary bright object viewed in darkness. It can be explained by tremor and drift of the eye which occurs normally, but is more prevalent when one looks at a totally unstructured field. Without fixation points to direct his orientation, these involuntary eye movements will occur unrecognized by the observer. As a result of its changing location on the retina, the image of a fixed object under these conditions appears to move. Eye tremor is a small, rapid oscillation of the eye; its frequency is approximately 140 Hz, and its amplitude about 3 minutes of visual arc. Drift normally reaches approximately 15 minutes of arc displacement from a fixation point before it is corrected. This occurs every 1 or 2 seconds in time and requires approximately 50 microseconds to complete. Knowledge of this phenomenon is of little positive use in the design of displays because of its transitory and unpredictable nature. It may, however, pose *problems* for display under darkened conditions. Large-area displays viewed under very low ambient conditions should contain some light distributed over the total display surface to minimize the chances for such autokinetic movement to occur. This will permit the eye to minimize drift or oscillations and thus stabilize any fixated image.

Phi-Phenomenon

The phi-phenomenon is probably the most familiar example of apparent motion to anyone concerned with displays—first, because one cannot work long in the field without encountering its effects and, second, because its causes and effects are sufficiently understood to be employed as a controllable information element in display. With the proper temporal intensity and spatial relationships between two sequentially activated stationary points, movement difficult to discern from real motion can be induced in the visual field. Variations on this phenomenon have been used to create a variety of displays presenting apparently moving signals, yet containing no moving elements on the display surface (or in the signal source). Electroluminescents have proved to be exceptionally satisfactory for this type of display, and are currently being employed using this principle for prototype aircraft instrumentation.

Jump Phenomenon

Stationary signals may appear to move on CRT's that use short-persistence phosphors. In a study for NASA reported by Bowen and Guinness (ref. 11), it was found that alphanumerics displayed on a CRT using a P–31 phosphor appeared to "jump" or "dance" under certain conditions. The user agency and its contractor had noted the phenomenon and had initially attributed it to instabilities in the equipment. Investigation showed, however, that the major cause of the phenomenon was due to the interaction of the observer's eye with certain temporal aspects of the display.

Briefly, their findings were as follows: using varied alphanumeric formats displayed on a P–31 phosphor (which decays to 10 percent of peak brightness in approximately 40 microseconds) at refresh rates varying from 20 to 50 Hz, two basically different types of induced motion were observed: the "jump" and the "shift" phenomena.

The shift phenomenon is of lesser consequence because of its instability and relatively small amplitude, and thus has not been investigated sufficiently to draw any firm conclusions concerning its influence on operator performance. It was found that some observers detected small shifts between the upper and lower halves of the display format, which appeared random in nature, although they most frequently appeared to be lateral displacements. The authors attributed them to the same kinds of involuntary eye movements noted above; that is, tremor and drift. Because of their small amplitude and their transitory nature, they would probably produce only minor annoyances for some operators, under special conditions.

The jump phenomenon, on the other hand, was found to be both consistent and compelling under the conditions investigated in this study. Symbol formats displayed on the CRT appeared to jump rapidly in the direction opposite to the motion of the eye when a viewer shifted his fixation across the screen. Where only one symbol, or a small group of symbols, was displayed, movement appeared to be a simple jump from the objective position and return after a very short time interval. At times, a symbol was seen repetitively during the jump, giving the appearance of a stream of replications.

When the display contained a more complicated format of lines of symbols, the phenomenon became perceptually more complex. The format structure appeared to interfere with the apparent jump, resulting in movement of parts of the display while other parts seemed to remain fixed.

The explanation for the apparent movement described as the "jump" phenomenon derives from two characteristics of the visual mechanism. The first, noted previously in discussing the phi-phenomenon, is the ability of the observer to perceive motion when two separate areas of the retina are stimulated in the proper space-time-intensity relationship. The second is the nature of the movement imparted to the eyeball when an observer is scanning the visual field without tracking a moving fixation point.

When the eye follows a moving target, its motion is smooth and continuous, closely following the actual motion of the target. However, in free scanning, even in a structured field, the eye moves in a series of abrupt (saccadic) steps from one fixation point to another. For instance, eye-movement studies have shown the kind of motion illustrated in figure 3 for an observer attempting to trace the circumference of a circle.

The nature of the jump phenomenon suggests that it results from the stroboscopic effect of successive light pulses striking different locations on the retina of the rotating eyeball. Furthermore, because the jump appears to be

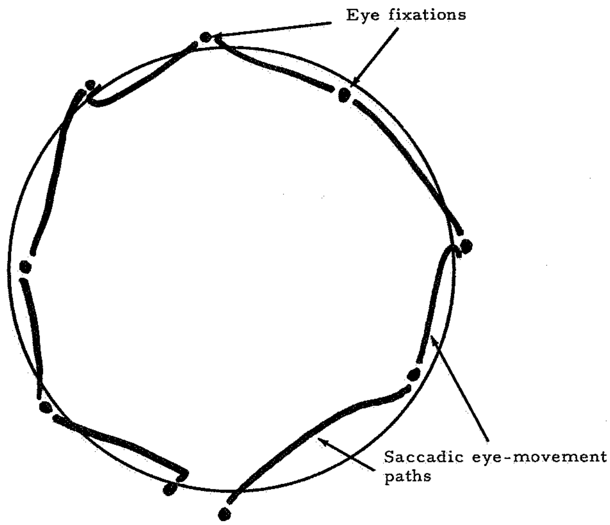

FIGURE 3.—Recorded eye movements of an observer "sweeping" the eyes around the circumference of a circle.

opposite in direction to the true motion of the eye, it would appear to be directly associated with its saccadic movements. In confirmation of this interpretation, it was found that the jump is more pronounced when saccadic movements are made over larger angles. It has been shown that peak angular velocities vary with the size of the angle of saccadic movement (from 320° per second for 15° shifts to 720° per second for 90° shifts). Therefore, it is to be expected that the amplitude of perceived jump should correlate with the angular translation of the eye between successive pulses.

To accept the interpretation of the jump phenomenon submitted by the authors, it is necessary for the eye to see while in motion. An uncritical review of the literature suggests that the eye is blind during saccadic movements. This has been used to explain why we do not experience blur between fixations while scanning. However, the fact is that the eye does see during a saccadic shift, but not very efficiently. This means that the brightness threshold would be significantly higher during a saccadic shift than when the eye is at rest and, therefore, suggests that the jump phenomenon would be more compelling with a bright signal and when the signal-to-background contrast is high.

A second study by the same researchers (ref. 12) was conducted in an attempt to verify the conclusions of the original study and to obtain quantitative data defining the threshold conditions necessary to produce movement in response to pulsed light.

Different experimental conditions were used to permit variations in pulse characteristics and PRF, but image size, ambient brightness, and the visual tasks requiring eye movements were analogous to what might be expected in an operational situation.

A light pulse produced by a flash tube had essentially the same shape and duration as the P–31 phosphor used in the first experiment, but its amplitude was controllable to change brightness. Refresh rates of 30, 40, 60, 80, and 100 Hz were used. Although limitations in the experiment preclude precise plotting of the relationship of the jump phenomenon to brightness and/or PRF, it was clearly demonstrated that the likelihood of "jump" did increase with the brightness of a single pulse to a 50-percent threshold at about 15 000 ft-L, as shown in figure 4. Although a slight tendency for the threshold to decrease with increasing PRF can be seen in the curve, there is no statistical significance in the differences obtained with the frequencies studied.

These data clearly demonstrate that the phenomenon occurs considerably above the CFF for these conditions (less than 50 Hz), and therefore suggests the independence of "jump" from

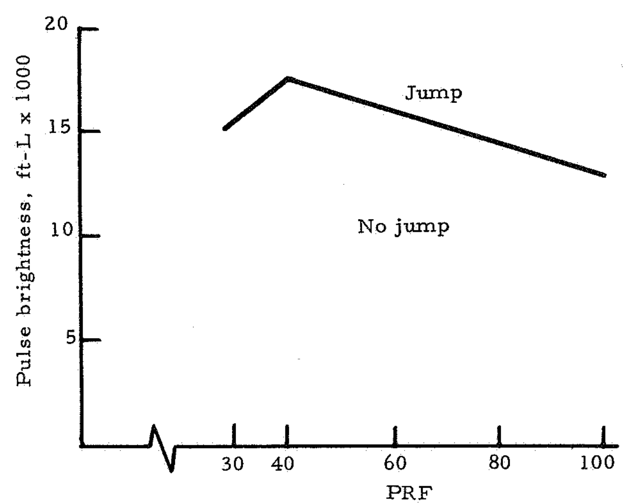

FIGURE 4.—Threshold for the presence of jump as a function of brightness of pulse.

any possible dependence on flicker in the traditional sense, as a determinant.

Whatever the theoretical implications of the jump phenomenon are for the field of visual perception, its presence under the proper mix of brightness, persistence and refresh rate has been established. Although we currently have insufficient data to define its influence on the accuracy, speed, or duration of operator performance, we can, at the very least, expect it to create annoyance and distraction when it occurs. Therefore, its likelihood should be minimized by proper design precautions.

These two studies described, together with the brief review of some of the less frequently discussed visual phenomena, are not intended as a survey of the unique or unsolved problems facing the engineering psychologist or the display designer. The goal has been to point out some rather clear-cut areas where familiarity with, or access to, some of the traditional sources of information dealing with human vision may fail to provide the display designer with the information he needs in a usable form. There are still many problems in the field of information display requiring the serious consideration of a qualified engineering phychologist.

Where the data are available, but do not explicitly answer the question in mind, background and experience may be the key to a proper interpretation. Where the data are unavailable, background and training provide the display designer with the proper tools to design definitive experiments for the development of the required data. However, he cannot perform either of these functions effectively without the close cooperation of the user, who should be responsible for an objective statement of the requirements, and the engineer-scientist, who has the responsibility of translating these requirements into the capabilities of selected media and techniques reflecting the state of the art.

Conclusions

The comments and discussion presented in this paper have not attempted to provide a definitive survey of the human-factors considerations that must be made in the use of advanced display media. Rather, they are intended to reaffirm the need for serious and continued concern on the part of engineer-scientists responsible for the development and implementation of these media for the ultimate human user.

The researcher should be sensitive to the fact that the human eye is a complex mechanism employing optical, photochemical, and neuroelectrical media in its functioning. It, therefore, can be expected to behave in reasonable approximation to the laws that apply to these media in its basic sensory response to simple signals. However, when the signal is complex (or its dimensions are not clearly understood), he must be cautious in his assumptions concerning the critical parameters for seeing and, therefore, the data that apply to his problem. He should also be aware that, when more complex tasks, involving recognition, interpretation, etc., are required by the operator, he is no longer dealing with simple sensory processes and, therefore, must concern himself with phenomena for which the data may not be available in sources dealing solely with "vision."

In conclusion, there are several considerations that might assist those responsible for display research and development in organizing their efforts toward the ultimate goal of providing useful information to man in a form compatible with his sensory perceptual capabilities. The first group is most appropriate to the researcher in the field of media development.

(1) Conceptualize the ultimate use, or uses, of the medium being investigated. This analysis should provide a spectrum of human visual requirements that can be used as guidelines in development, thus preventing "shotgun" investigations based on a fascination with the properties of the medium rather than its ability to meet the needs for display.

(2) Define the properties of the medium carefully, including the range(s) of output characteristics, and their stability for the conditions under which they might be used.

(3) Relate these properties to the known characteristics of the visual process to see how they can best provide a useful signal, and in what directions medium improvement will better satisfy the visual requirements.

(4) Remember that the visual mechanism is itself a very complex medium, over which we

have very limited control, and the degree to which its capabilities and limitations are satisfied is the ultimate criterion of effectiveness for the medium that represents the display interface.

In addition to these considerations pertinent to the media researcher, the following guidelines should assist the display designer in meeting the specific information and human-factors requirements for a specific display concept.

(1) Establish objective user information requirements for the display, keeping in mind the tendency to overstate the needs to match or exceed the state of the art.

(2) Convert these requirements into signal parameters necessary to carry the information.

(3) Determine the procedural and ambient conditions under which the display will be used, considering time constraints, lighting, etc.

(4) Relate the information needs (signal parameters) to the characteristics of the media available for implementation, and assess their relative effectiveness for display under the conditions specified.

(5) Implement design decisions for breadboard or prototype testing to determine any unanticipated idiosyncrasies arising from the mix of media, technique, and format, using human operators to perform realistic tasks with the display as the criterion.

References

1. Poole, H. H.: Fundamentals of Display Systems. Macmillan & Co., 1966.
2. Silver, C.; and Cruikshank, Lt. Col. R.: The Faster the Better? 6th National Symposium on Information Display, New York, N.Y., Sept. 1965, pp. 81–84.
3. Turnage, R. E., Jr.: The Perception of Flicker in Cathode Ray Tube Displays. Information Display, May/June 1966, pp. 38–52.
4. Baker, C. A., ed.: Visual Capabilities in the Space Environment. Pergamon Press, 1965.
5. Baker, C. H.: Man and Radar Displays. Pergamon Press, 1962.
6. Morgan, C. T., et al., eds.: Human Engineering Guide to Equipment Design. McGraw-Hill Book Co., Inc., New York, N.Y., 1963.
7. Wulfeck, J. W., et al.: Vision in Military Aviation. WADC Technical Report 58–399. Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, Nov. 1958.
8. Bartlett, N. R.; and Williams, S. B.: Signal Mark Size and Visibility of Radar Signals on a Plan Position Indicator. Rept. No. 166I–30, Systems Research Laboratory, The Johns Hopkins University, Sept. 20, 1947.
9. Williams, S. B.: Visibility on Radar Scopes. Rept. no. 166–1–82, Inst. for Cooperative Research, The Johns Hopkins University, Feb. 15, 1949.
10. Landis, C.: An Annotated Bibliography of Flicker Fusion Phenomena. Publ. by Armed Forces-National Research Council Vision Committee Secretariat, June 1953.
11. Bowen, H. M.; and Guinness, G. V.: Apparent Movement Phenomena on CRT Displays. Rept. no. 510–TM–2 (DRD–64–131), NASA Contract no. NASW–954, Oct. 1964.
12. Bowen, H. M., et al.: Apparent Movement Phenomena on CRT Displays: Threshold Determinations of Apparent Movements of Pulsed Light Sources. NASA CR–342, Dec. 1965.

N68-22317

DISPLAY MEDIA: SUMMARY, INTERPOLATIONS, AND EXTRAPOLATIONS

GEORGE KOVATCH AND EDWIN H. HILBORN
Electronics Research Center, NASA, Cambridge, Mass.

INTRODUCTION

Modern aerospace vehicles are placing new requirements on visual display media for the presentation of information to the pilot or astronaut. Conventional multi-instrument displays with their numerous moving dials, tapes, and indicators no longer have the potential for advanced applications either in terms of reliability, available panel space, or the selective presentation of the information required for a given mission phase. The advent of the modern aerospace computer has opened up avenues for information display never before thought possible. In turn, computers have placed a requirement for new display media and have required software programing specialists to develop effective means for transferring information from the computer to the display. Another major component of the system is, of course, man. His role is expected to become that of decision-maker, or manager of the system, rather than as an active component in the control loop. To design the best display system, an interdisciplinary group of engineers and psychologists, who can understand man and his capabilities, is required. Hence, for complete display system design, attention to all three major components — man, computer, and display — is essential.

This paper emphasizes the display portion of the system, and mainly visual displays, although other types such as auditory displays may be important to the final design. It attempts to summarize the requirements placed on display media by future high-performance vehicles and also reviews briefly some current developments. Finally, it extrapolates from these advanced developments and attempts to predict how they will fit into future display systems. The comments are the opinions of the authors and do not reflect any official NASA position.

DISPLAY REQUIREMENTS

Advanced Space Vehicles

A comparison of the state of the art in spacecraft display for the Apollo vehicle and the desired parameters for advanced manned operation shows that weight, volume, power, and lifetime requirements must be improved. In the Apollo command module the display system accounts for approximately 311 pounds. It is desired to reduce this to approximately 100 pounds, or a reduction of 66 percent. The same reduction factor is desired for the lunar module display system which currently weighs 151 pounds. Concurrent with the reduction in weight is a need for reduction in volume by a factor of about one-third. A power reduction from 128 watts to approximately 50 watts is also sought.

There are approximately 400 different display functions in the Apollo vehicle, 350 of which are on display simultaneously. For future spacecraft display systems, the same number of functions will be required, but it will be desirable to display simultaneously no more than 150 of these. Herein lies the interest in developing multiformat displays; that is, to use the same basic display for many functions. This can be achieved by calling the information out of the computer when required for a particular phase of the mission. In Apollo, certain instruments are used for only 10 minutes during an entire mission. If this same information could be displayed by calling it out of the computer during this 10-minute period, a separate display instrument would not be required. This would result in a reduction in power, space, and weight. Certain future displays will also require longer lifetimes. Typically, of the media suitable for use with a computer, those which have reached a stage of development such that they are essentially off-the-shelf items have useful lifetimes of 1000 to, at most, several thousand hours. For advanced space missions, displays with a reliability extending to several years will be required, if the desired mission objectives are to be accomplished.

Advanced Aeronautical Vehicles

Future high-performance aeronautical vehicles, such as the supersonic transport and vertical takeoff and landing aircraft, impose new contraints on display media. The supersonic transport will place new demands on the pilot in his role as systems manager. In this regard, he must be presented a variety of information in a form that he can readily interpret in making decisions which affect vehicle performance. In order to perform these functions, he must depend on a large-capacity, digital computing and storage system which can process data into appropriate forms. Interfacing hardware and software that will convert the basic data into a desired form for display will be needed. Finally, advanced display media will be required to supply an interface between the human operator on one side and the driving hardware on the other.

Computer-driven CRT's may offer the greatest range of flexibility for handling this display function. In commercial aircraft, power and weight constraints are not as stringent as they are in spacecraft; therefore, the physical bulk inherent in a CRT is tolerable. Ruggedized CRT's for airborne use can already be built. A redundant gun can be utilized in the design to offset possible catastrophic filament failure. Phosphor lifetime of the CRT is predictable, and preventive maintenance thus can be scheduled easily. The CRT offers other reliability advantages because of the small number of connections needed to drive it, and it is flexible in that either analog or digital deflection is possible. Improvements and simplifications in color CRT's will offer even wider potential for aircraft system applications. Finally, the moderate cost of a CRT is compatible with the large numbers required for multiple vehicles.

It is expected that multiformat displays will be necessary in supersonic aircraft. These will reduce considerably the numerous dials and indicators currently needed in large aircraft. The computer will be called upon to process information in pictorial or symbolic form and to transfer it to the display. This will include the presentation of status information and the presentation of flight profiles and trajectories. The latter will be required for both nominal planned mission profiles and for presentation of alternative flight profiles to handle unexpected occurrences while en route. Much of the flight will be handled automatically by the computer. The pilot will monitor the system and will take over the operation only when required. Hence, he must be presented information about the status of the flight continuously so that he can make prudent decisions about automatic, manual, or split-axis control.

The vertical or short takeoff and landing vehicle (V/STOL) offers additional new challenges to display system designers because of the peculiar flight regime in which it flies. This includes near-vertical landing, hopefully under all-weather conditions. The pilot, therefore, must essentially control with two different types of controls: aerodynamic control for horizontal flight, and direct, or deflected, lift for vertical flight. Therefore, the pilot needs appropriate information for these phases to make the proper decisions and to insert the appropriate commands into the system. There is a need for him to assimilate information both from instruments and from the outside world in order to orient himself and the vehicle and to keep informed about the status of the various control subsystems. For this we see again a strong role for an onboard computer to process data and to present them to the pilot in symbolic or pictorial form. In addition, we expect to see a need for "head-up" displays. Through these he can combine outside world information with electronically generated information. The overall display system might include a computer-driven CRT and an optical projection system, which transfers information to the windshield.

Our third area of interest is that of displays for general aviation. There is a need here for instrumentation which will allow operation by visual flight rules (VFR) trained pilots. Improved sensors and displays for general aviation vehicles should enjoy high reliability and, of course, low cost, commensurate with the cost of the vehicle itself. The general utilization of high-capacity computer equipment on board these vehicles immediately puts it beyond the apparent cost limitations of general aviation vehicles. We could not afford a computer and display system for an aircraft that costs as much as the aircraft itself. Therefore, we must emphasize display media that can be readily interpreted by the pilot, but with low-cost potential, and display media that require a minimal amount of onboard computer capability. One might expect to see a large ground-based, time-shared computer to satisfy computational requirements, the results of which are then transmitted to low-cost terminals in the cockpit of the small aircraft. Fluidic or magnetic displays may prove of value here because the amount of information requiring display in these relatively simple vehicles is much less than that which is required in large commercial aircraft and, hence, an economical system might be devised.

These improvements are required not only for the safety of the small plane operator, but also to improve commercial aircraft safety in our increasingly crowded airways.

Future Trends

To generalize now about future trends in advanced display media, the following potential paths are expected. First, a greater emphasis on computer-driven displays will be seen onboard advanced aerospace vehicles. For aircraft, it appears that the CRT will find increasing levels of acceptance because it offers a general multiformat display that can be computer driven and changed through software programing as a function of changing flight conditions. We expect to see developments in CRT's, making them more rugged and more reliable, and needing less maintenance. It is expected that future improvements will provide high-contrast ratio tube faces, allowing the CRT to be read without difficulty under a wide range of ambient conditions. Developments such as the black-face CRT are expected to find wide application, and we expect an increasing use of color CRT's, in order to distinguish among various types of information through color differentiation on the display tube. Simplification in means of generating color in a CRT will be required and improvements in the state of the art are expected. This may involve improved single-gun, colored CRT's or simplifications in other means of synthesizing color information.

For spacecraft, we expect to see multiformat, computer-driven display systems. However, because of weight and power limitations, it is expected that solid-state matrix displays will be used. Field-effect electroluminescent displays may find less application due to lifetime restrictions. Carrier injection EL offers a real potential should its rate of development be fast enough and successful enough to meet the demand. The plasma display may offer a realistic alternative. It can help in handling the lifetime limitations but it still will require high voltage drive.

Is is expected that for future aerospace display applications, a variety of display media will be needed to satisfy the variety of requirements. No single medium would appear to best satisfy the differing requirements of these several vehicles.

To develop the display system fully, consideration must also be given to displays other than visual displays. Audio displays may well be integrated into the overall design. We expect human factors or engineering psychologists to help establish the role of man and the wide variety of display media in future man-display-control systems. We also expect computer engineers and mathematical programers to play a significant role in display system design brought about by the demand for efficient programs for multiformat computer-driven systems.

LIST OF EXHIBITS AND DEMONSTRATIONS

Company	*Representative*	*Device Exhibited*
Canadian Westinghouse	R. E. W. Lake	Electroluminescence
CBS Laboratories	Richard Roule	TV aperture correction
Conductron Corp	Paul Rosengaard	Holography
G.D. (Stromberg-Carlson)	J. H. Redman	Computer I/O
General Electric Corp. (Schenectady)	Hans Stern	Fluidic display
General Electric Corp. (Syracuse)	Daniel Osborn	Electroluminescence
Farrand Optical	R. Draudin	Pancake window
Hartman Systems	K. Lally	Electroluminescence
	D. Amburger	High-contrast CRT
Honeywell	Charles Baker	Electroluminescence
IBM	Robert Rund	Electroluminescence
ITT (Fort Wayne)	T. Sisneros	Single-gun color CRT
	E. Smierciak	Narrowband TV
Lear-Siegler	Donald DeMyer	Electroluminescence
McDonnell-Douglas	Warren Merboth	Avian retina model
Martin (Denver)	John Polhemus	Multiparameter display
Martin (Orlando)	J. Van Der Heyden	Fluidic display
Monsanto	David Hillman	Electroluminescence
NASA-Ames	J. Christiensen	Electroluminescence
NASA-ERC	David Sawyer	Flying spot scanner
National Cash Register	Richard Petticrew	Liquid crystals
RCA (Somerville, N.J.)	Lawrence Murray	Electroluminescence
RCA (Van Nuys, Calif.)	T. V. Curran	Computer I/O
Sigmatron	Martin Reder	Electroluminescence
Simmonds Precision Products	William Dunn	Magnetic display
Sinnott Co	R. Sinnott	Magnetic display
Sperry-Phoenix	Lowell Schuck	Plasma display
Sylvania (Seneca Falls, N.Y.)	Elmer Stone	Single-gun color CRT
Thomas Electronics	Peter Seats	CRT
TRW Systems	Bowman Evans	Holography
United Aircraft (Norden)	Hal Wooding	Vertical situation display

U.S. GOVERNMENT PRINTING OFFICE: 1968 O—291-351