VOLUME 3 □ NUMBER 8 AUGUST 1966

IEEE spectrum

features

Departments: *Please turn to the next page*

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

departments

the cover

Mosaic of narrow-angle pictures of moon taken by Surveyor I's television camera forms a panoramic view of lunar terrain stretching about 115° across the horizon. Elongated shadow of the spacecraft was created by the low sun sinking on the horizon behind it. Each photo chip is two inches square and represents a 6° field of view as seen by the camera. The article beginning on page 54 gives details of the television camera system.

VOLUME 3 ☐ NUMBER 8 IEEE spectrum AUGUST 1966

Spectral lines

Another new journal. In June, the IEEE announced that a new periodical, the IEEE Journal of Solid-State Circuits, will be published on a quarterly basis beginning in September 1966. This publication will cover a field of growing importance to the IEEE membership—a field in which the passive and active circuit components can no longer be considered separately but must be considered as an integrated assemblage. This field is clearly one that fits squarely into the IEEE field of interest but cuts across the interests of several existing IEEE Groups.

The subject is logically a part of circuit theory and the Circuit Theory Group was largely responsible for the initiative that resulted in the development of the new Journal. The Journal's coverage also falls into the area of electron devices and so it is not surprising to find the Electron Devices Group as a sponsor. Other Groups with active interest in this area are the Computer Group and the Microwave Theory and Techniques Group. To administer the new Journal, an Advisory Council, which includes representatives from each Group, has been formed with the chairman and vice-chairman appointed by the Vice President for Technical Activities, Hendley Blackmon. He has named Dr. John G. Linvill of Stanford University and Dr. G. E. Moore of Fairchild Semiconductor Division to these posts. In addition, Dr. James Meindl, U.S. Army Electronics Command, is editor, and Arthur Lo of Princeton University, Dr. H. C. Lin of Westinghouse Electric Corporation, R. Engelbrecht of Bell Telephone Laboratories, and R. Webster of Texas Instruments Inc. are associate editors.

The new Journal and the IEEE Journal of Quantum Electronics, which was started in April 1965, represent two new publishing ventures of a different type for the IEEE. Undoubtedly many members wonder where this new process will lead and what its significance is to the overall Institute publications program, but it is not yet possible to give an adequate answer.

In both instances, the reasons for starting a new periodical were compelling as seen by the Institute's governing Boards.

In the quantum electronics case, rapid developments in solid-state physics were leading to a host of new devices and concepts of tremendous importance to IEEE members and to the profession they represent. It was also clear that neither the Proceedings of the IEEE nor any of the existing Transactions was attracting the many significant papers necessary to become a central repository for the field. This was evident even though the Proceedings had devoted a special issue to quantum electronics and the Electron Devices Group had attempted to include quantum electronics papers in its Transactions. The Institute felt that this field, which had such obvious technological significance, could be better served if its literature were not scattered over a large number of periodicals but rather was given an appropriate focus in a single journal. It is a pleasure to note that the Journal of Quantum Electronics has now been selected as the periodical for the publication of papers for the 1966 International Quantum Electronics Conference, which was held in Phoenix, Ariz., last April. These conferences, held on an every-second-year basis and sponsored by a number of professional societies including the IEEE, are the most important meetings for the quantum electronics field.

Similarly, in the case of solid-state circuits, where a design philosophy and technology have already developed into a host of important applications, as was so clearly shown in the December 1964 Proceedings, a special issue devoted to integrated electronics, it was evident that none of the existing IEEE publications was serving as a focus for the area. Again the literature was being scattered over a number of periodicals. It is expected that the new Journal will give those active in this field an appropriate central publication that will fulfill their needs. Active on-going conferences, such as the International Solid-State Circuits Conference, the Annual Meeting on Electron Devices, and the Solid-State Research Conferences, already service this field, and it is anticipated that papers from such conferences will find an outlet in the new Journal.

Although there are similarities, there are also differences between these two cases. With quantum electronics, many of the active contributors are physicists and chemists whose primary professional affiliations are in other societies, and it was felt that periodicals which did not include subjects not of interest to them would best attract these papers and subscriptions. In contrast, the solid-state circuits field includes mostly people who have a close association with the IEEE; yet no existing IEEE periodical served the field adequately. An alternate approach to starting a new Journal would have been to try to redirect an existing Transactions to service the new field. To date this procedure has not been successful in the Institute although some attempts have been made.

The Institute must be responsive to the changing publication needs of its members and of the profession if it is to fulfill its objectives. The initiation of the new Journal of Solid-State Circuits is another indication of the Institute's desire to fill needs as soon as they are recognized. The Institute owes appreciation to many individuals for the hard work that has resulted in the launching of the new Journal. Its Editors and the Advisory Council have our best wishes for success.

F. Karl Willenbrock

(*For subscription price information, see July, p. 36.*)

Authors

The Surveyor lunar landing television system (page 54)

Donald R. Montgomery (M) received the B.S.E.E. degree from California State Polytechnic College in 1952 and joined the U.S. Navy Electronics Laboratory in San Diego, Calif., where he participated in development activities related to underwater transducers and submarine sonar systems. The following year he became associated with the Pacific Division of the Bendix Corporation and continued to work on sonar system design. Later work at Bendix included design and development of the Bendix commercial marine radar and a K_u-band terrain avoidance radar (which led to a patent application for a radar display index generator), development of X-band radar beacons for use during Operation Hardtack, rendezvous-refueling beacon development for B-58 weapons system, and digital instrumentation encompassing PCM telemetry synchronization and decommutation. He joined the Jet Propulsion Laboratory of the California Institute of Technology, Pasadena, in 1963 and worked on the development of the Surveyor television system. He is now engineering group supervisor in the lunar and planetary instruments section, Space Science Division.

Frank J. Wolf received the B.S. degree in mechanical engineering from the California Institute of Technology in 1948, the M.B.A. degree from the University of Southern California in 1961, and the M.S. degree in engineering from the University of California at Los Angeles. He has been associated with the southern California aerospace industry since 1953, initially as production manager for Honeycomb Structures and subsequently as project engineer of the Bomarc Data Link Receiver with Lear, Inc. He has been with Hughes Aircraft Company since 1958 and associated with the scientific payload of the Surveyor since the inception of the program in 1961. During this time he has been experiment engineer on advanced payload items such as a gas chromatograph, an X-ray spectrometer, an X-ray diffractometer, and various geophysical instruments designed to determine fundamental properties of the lunar surface and subsurface material. As television manager and system engineer for the Surveyor television system, he has been responsible for the engineering definition of the performance and test of Surveyor's "eye" and the implementation and integration of the hardware as part of the spacecraft.

New means of communication (page 62)

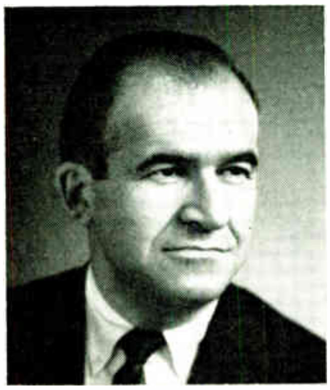
John R. Pierce (F) joined the Bell Telephone Laboratories in 1936 after receiving the Ph.D. degree from California Institute of Technology. From then until 1955 he worked on high-frequency electron tubes and, particularly, on traveling-wave tubes. He became director of electronics research in 1952 and at present is executive director, research, of the Communications Sciences Division at Bell Telephone Laboratories in Murray Hill, N.J., with responsibilities in such fields of research as radio, electronics, acoustics and vision, mathematics, and psychology. Dr. Pierce is the author of nine technical books and numerous articles and the recipient of the 1942 Eta Kappa Nu Award, 1947 Morris Liebmann Memorial Prize, 1960 Stuart Ballantine Medal, 1962 Golden Plate Award of the Academy of Achievement, 1963 Arnold Air Society General Hoyt S. Vandenberg Trophy, 1963 Edison Medal, 1963 Valdemar Poulsen Medal, 1963 National Medal of Science, and 1964 H. T. Cedergren Medal. He is a member of the National Academy of Sciences, National Academy of Engineers, and Air Force Association, and a Fellow of the American Academy of Arts and Sciences and the American Physical Society.

The peculiarities of high-voltage dc power transmission (page 76)

A. Uno Lamm (F) was graduated from the Royal Institute of Technology, Stockholm, Sweden, in 1927 and received the Doctorate of Technology in 1942. He joined Allmänna Svenska Elektriska aktiebolaget, Ludvika, in 1928 and worked on the development of mercury-arc rectifiers, becoming head of the rectifier department in 1929. He was named chief engineer in 1947 and director in 1955. Transferred to ASEA headquarters in Västerås, he became electrotechnical director in 1959. For his work in design and application of transductors and his development of the fundamentals of a new theory, he received awards from the Swedish Royal Academy of Engineering Sciences, Association of Engineers and Architects, and Royal Academy of Science. He received his first patent in the field of high-voltage dc transmission in 1929 and led a team working on its practical applications, such as the 20-MW HV dc cable from the Swedish mainland to the island of Gotland and 160-MW Cross Channel project. This work was recognized by the American Society of Swedish Engineers. In 1965, he received the IEEE Lamme Medal.

International standardization—interface with the future (page 91)

Alexander C. Grove (SM) received the B.S.M.E. degree in 1949 from Cooper Union and the M.S. degree in management engineering from Columbia University in 1966. He joined the General Precision Laboratory in 1952, working in the development of airborne navigation and air traffic control systems and, later, the design of ground checkout equipment for the mobile medium-range ballistic missile guidance system. In 1964 he organized the Defense Communications Agency Industry Review Program of the American Standards association. He is presently managing this program and, under his direction, 80 of the largest American communications equipment manufacturers are aiding the Agency to develop realistic state-of-the-art standards for the defense communications system. His experience in international standardization began in 1959; since then he has served on various International Electrotechnical Commission technical committees and helped to organize meetings of the International Organization for Standardization's Committee on Mechanical Vibration and Shock. This year ASA nominated him for the office of ISO Secretary General.



Quantitative lightning spectroscopy (page 102)

Martin A. Uman is presently a member of the technical staff of the high-voltage breakdown department of the Research and Development Center, Westinghouse Electric Corporation, Pittsburgh, Pa. He received the B.S.E. degree in electrical engineering in 1957 and the Ph.D. degree in 1961, both from Princeton University, Princeton, N.J. Following this, he accepted a position on the faculty of the University of Arizona, Tucson, Ariz., as associate professor of electrical engineering. His work at the University of Arizona also consisted of conducting theoretical as well as experimental studies of the behavior of low-energy electrons in gas mixtures and, in addition, he performed theoretical research relating to lightning discharge.

In January 1956 he became a member of the technical staff of Westinghouse Research Laboratories in Pittsburgh, Pa., and, at the present time, he is continuing his research into lightning phenomena. Dr. Uman is the author of approximately 20 technical articles, as well as the textbook *Introduction to Plasma Physics*, which was published in 1964 by the McGraw-Hill Book Company, Inc.

High-strength conductors for supermagnets (page 111)

D. Bruce Montgomery of the M.I.T. Magnet Laboratory, Cambridge, Mass., received the A.B. degree in physics from Williams College in 1955 and the B.S. and M.S. degrees in electrical engineering from the Massachusetts Institute of Technology in 1957; he graduated Phi Beta Kappa and was a member of Sigma Xi. From 1954 to 1957, he was a part-time engineer with Raytheon Manufacturing Company, and in 1957 he became a staff member of Arthur D. Little, Inc. He joined the M.I.T. Lincoln Laboratory in 1959 and, in 1961, he became associated with the M.I.T. National Magnet Laboratory. Working on the planning and research stages of this laboratory, he became leader and built up the present magnet research and development group. He is presently working in Europe for 14 months at the Clarendon Laboratory, Oxford University, and at the Laboratoire de Physique, University of Lausanne, Switzerland. His various publications include work on the topics of air core solenoids, iron magnets, pulse magnets, superconducting magnets, and field analysis. He received the 1965 Young Engineer of the Year Award of the Massachusetts Society of Professional Engineers.

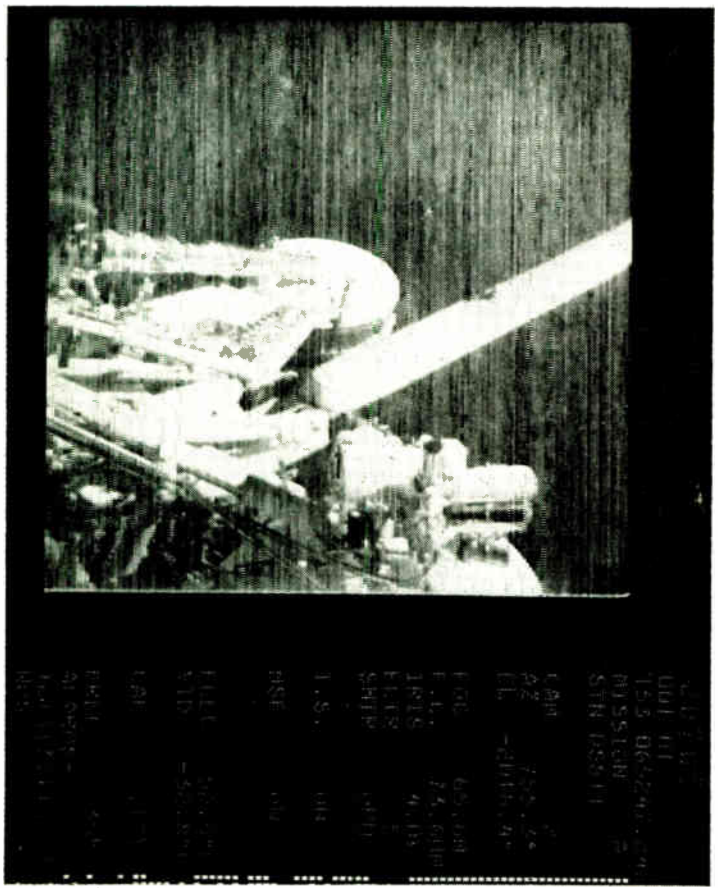


Fig. 1. The first television photograph from the Surveyor I spacecraft. Transmitted in 200-line mode, it shows a view of a spacecraft leg and footpad.

The Surveyor lunar landing television system

From the period of June 1, 1966, through June 14, and from July 6 through July 13, the Surveyor spacecraft television camera provided the United States a close-up view of the lunar surface at the millimeter scale. This article describes the functional engineering aspects of the camera and its performance capabilities

Donald R. Montgomery

Jet Propulsion Laboratory,
California Institute of Technology

Frank J. Wolf

Hughes Aircraft Company

The slow-scan television camera used in Surveyor I operates in a slow-scan mode that reduces the RF power requirements for the lunar distances involved. One 600-television-line frame is provided every 3.6 seconds in normal operation; one 200-line frame every 60.8 seconds in a second mode of operation. Upon command from earth, the camera's shutter opens, allowing light energy to reach a vidicon image sensor. The vidicon transduces the light energy to electric signals for subsequent transmission to earth. Detailed calibration information obtained prior to launch permits the correction of images received for geometric nonlinearities and distortions, frequency or aperture response, photometric nonuniformities, and coherent noise.

At exactly 07:41 hours (Pacific Daylight Time) on May 30, 1966, the first Surveyor spacecraft* was launched from pad 36A at Cape Kennedy, Fla., on a direct ascent trajectory to the moon. Following a near-perfect boost and injection by the Atlas–Centauer launch vehicle, the spacecraft accomplished a successful midcourse maneuver at a range of 155 000 kilometers from the earth, which provided the precise correction to achieve a landing accuracy within 15 km of the desired location on the lunar surface. A "textbook" touchdown was accomplished at 23:17 hours (PDT) on June 1, 1966, at a lunar coordinate position of 43.32° west longitude and 2.49° south latitude in the area called Oceanus Procellarum (Sea of Storms) in the crater Flamsteed.

The first television picture, Fig. 1, was transmitted in the 200-line mode 35 minutes after landing prior to positioning the spacecraft high-gain antenna toward the earth. Subsequent to the lunar landing, during the 14 days prior to the passing of the evening terminator, the television camera on the spacecraft provided engineers and scientists at the Jet Propulsion Laboratory in Pasadena, Calif., with 10 338 individual television frames of the lunar surface, including photographs of the surface upon which the spacecraft landed and subsequently came to rest and the area roughly 2.3 km to the local lunar horizon.

Camera optical system

The slow-scan television camera, shown in Figs. 2 and 3, provided images of the lunar surface over a 360° panorama. Each picture, or frame, is imaged through an optical system onto a vidicon image sensor whose electron beam scans a photoconductive surface, thus producing an electrical output proportional to the conductivity changes resulting from the varying receipt of photons from the object space. The camera is designed to accommodate scene luminance levels from approximately 0.008 to 2600 footlamberts, employing both electromechanical mode changes and iris control. Frame-by-frame coverage of the lunar surface provides viewing of 360° in azimuth and from +40° above the plane normal to the camera Z axis to −60° below this same plane.

Camera operation is totally dependent upon receipt of the proper command structure from earth. Commandable operation allows each frame to be generated by causing

---

* The Surveyor spacecraft was built by the Hughes Aircraft Company for the National Aeronautics and Space Administration and the Jet Propulsion Laboratory.

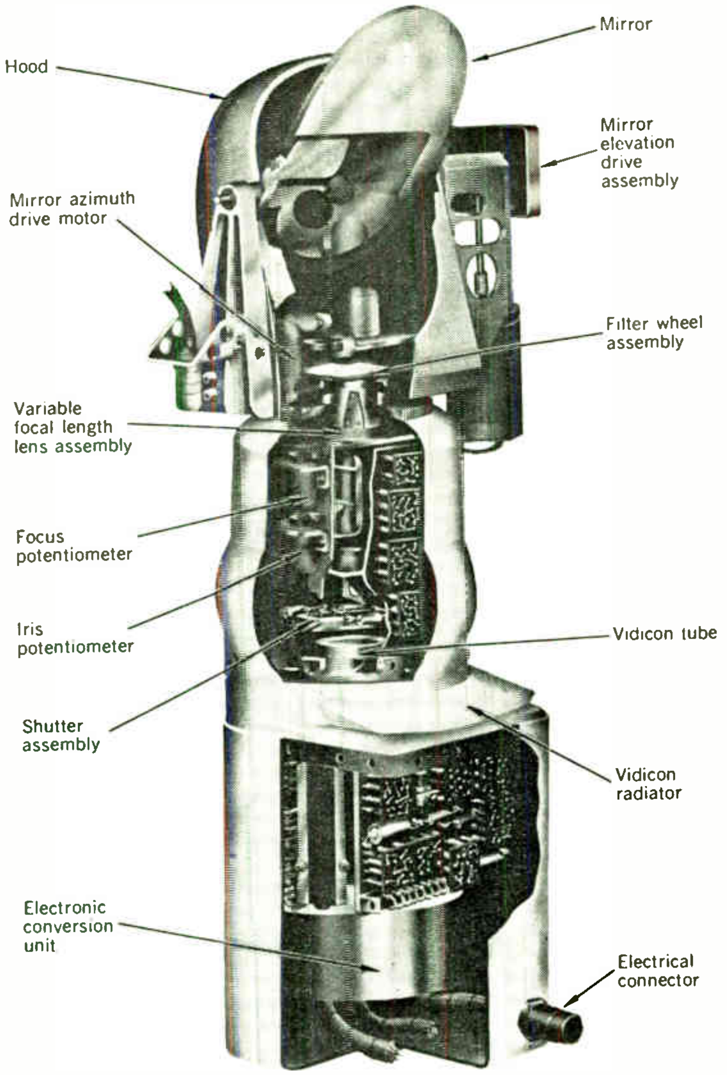


Fig. 2 (above). Cutaway drawing of the Surveyor camera, depicting major component location.

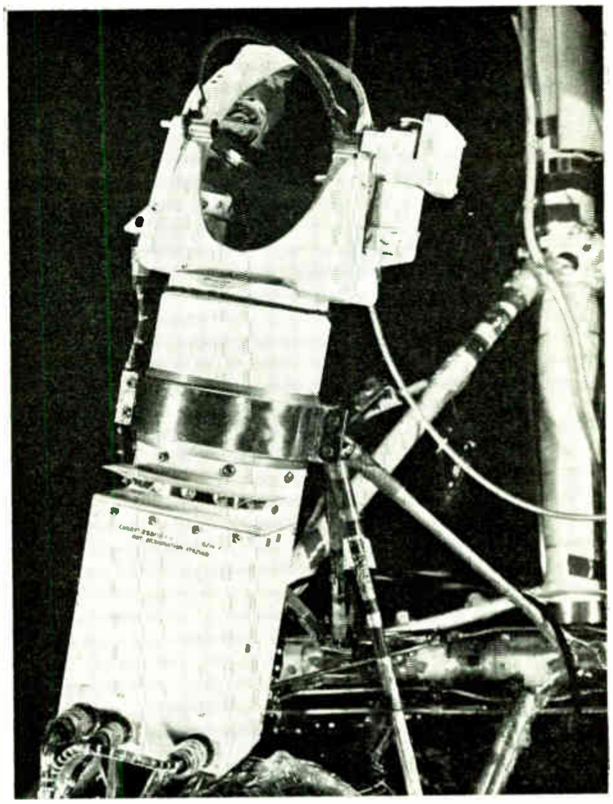


Fig. 3 (top right). Surveyor camera mounted on the spacecraft.

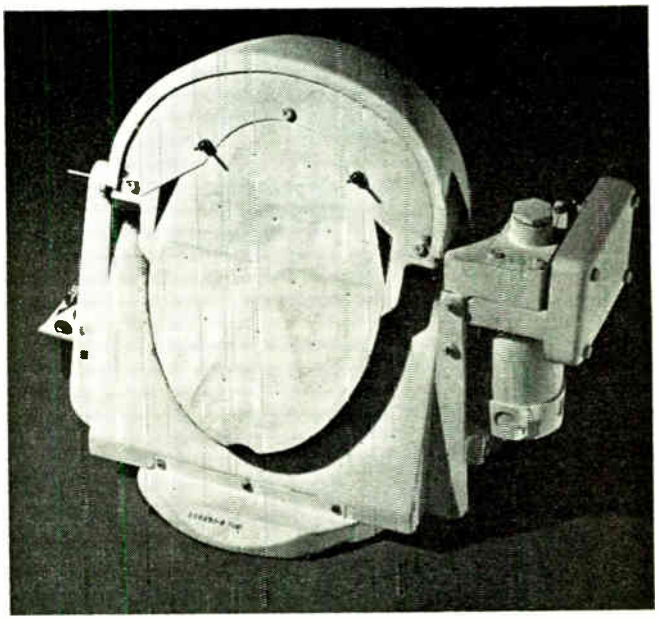


Fig. 4. Camera mirror assembly shown in the closed position.

sequencing of the shutter preceded by appropriate lens settings and mirror azimuth and elevation positioning to obtain adjacent views of the object space. Functionally, the camera provides a resolution capability of approximately 1 mm at 4 meters and can focus from 1.23 meters to infinity. The camera, weighing 16.1 pounds, consists essentially of six major items or subassemblies: the mirror, lens, shutter, filter wheel, vidicon, and the attendant electronic circuitry.

The mirror assembly shown in Fig. 4 consists of a 10.5- by 15-cm elliptical mirror supported at its minor axis by trunions. This mirror is formed by vacuum depositing an aluminum surface on the beryllium blank, followed by a deposition of Kanogen with an overcoat of silicon monoxide. The mirrored surface is flat over the entire surface to less than one-quarter wavelength at $\lambda = 550$ millimicrons (nanometers) and exhibits an average specular reflectivity in excess of 86 percent. The mirror is positioned by means of two drive mechanisms, one for azimuth and the other for elevation. The drive mechanism consisting of stepper motors with appropriate gear reduction, provides a mirror step size of 2.48° ± 0.1° in elevation and 3.0° ± 0.1° in azimuth. Angular step positions of both axes are sensed by position potentiometers, the outputs of which are digitized and transmitted to earth in PCM form.

The rotation of the mirror in the azimuth direction, while providing the azimuth coverage capability to the camera, creates as a result an image rotation proportional to the angular azimuth position of the mirror. This rota-

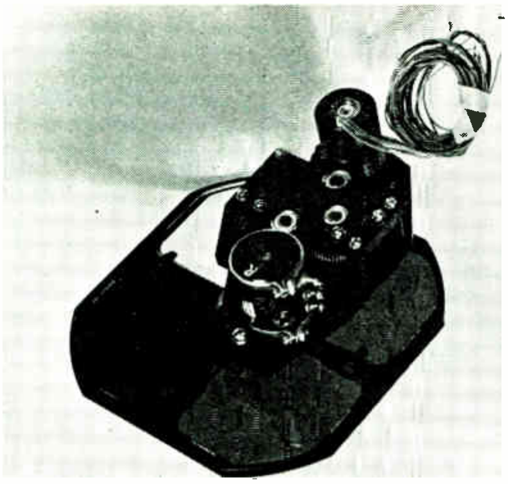
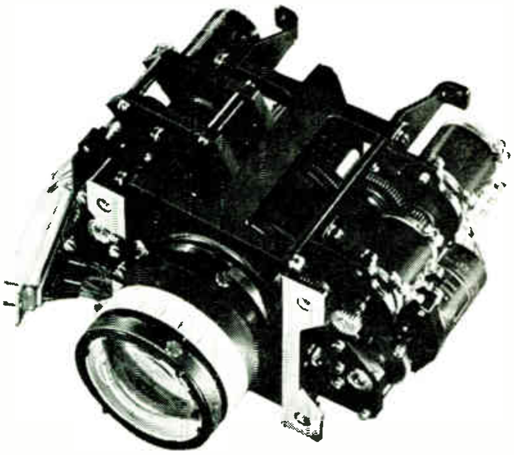


Fig. 5. Filter wheel mechanism.

Fig. 6. Variable-focal-length lens assembly.



tion is created by the simple fact that the image plane and the scanning raster of the image sensor (vidicon) are held stationary with respect to the mirror azimuth axis. These rotated images may be "corrected" in the ground video processing equipment.

In addition to the mirror itself, the mirror assembly contains a commandable filter wheel mechanism, shown in Fig. 5, which is capable of accommodating four separate sections of optical-quality glass filters. The Surveyor I spacecraft contained three such filters—red, green, and blue—in addition to a fourth section containing a clear element for nonmonochromatic observations. The filter characteristics are computer derived and tailored such that the camera responses, including the spectral response of the image sensor, the lens, and the mirror, match as nearly as possible the standard tristimulus value curves of the International Commission on Illumination (CIE). The filter wheel is sequentially placed in the field of view of the camera following the receipt of the proper earth-originated command. Color photographs of any given scene are then reproduced on earth after three video transmissions, each with a different filter element in the field of view.

The optical formation of the image is performed by means of a variable-focal-length lens assembly placed between the vidicon image sensor and the mirror assembly. Each lens, Fig. 6, is capable of providing a focal length of either 100 or 25 mm, which results in an optical field of view of approximately 6.43° and 25.3° respectively. Additionally, the lens assembly may vary its focus by means of a rotating focus cell from near 1.23 meters to infinity while an adjustable iris provides effective aperture change of from *f*/4 to *f*/22 in increments that result in an aperture area change of 0.5. While the most effective iris control is accomplished by means of command operation, a servo-type automatic iris is available to control the aperture area in proportion to the average scene luminance. As in the mirror assembly, potentiometers are geared to the iris, focal length, and focus elements to allow ground determination of these functions. A beam splitter integral to the lens assembly provides the necessary light sample for operation of the automatic iris component of the camera.

Two modes of operation are afforded the camera by means of a mechanical focal plane shutter located between the lens assembly and the vidicon image sensor and shown in Fig. 7. Upon receipt of an appropriate earth command, the shutter blades are sequentially driven by rotary solenoids across an aperture in the shutter base plate, thereby allowing light energy to reach the image sensor. The time interval between the initiation of each blade determines the exposure interval, nominally 150 milliseconds. An additional shutter mode allows the blades to be positioned to leave the aperture open, thereby providing continuous light energy to the image sensor. This mode of operation is useful in the imaging of scenes exhibiting extremely low luminance levels including star patterns.

The transducing process of converting light energy from the object space to an equivalent electrical signal in the image plane is accomplished by the vidicon tube. The vidicon employed is a hybrid device utilizing electrostatic focus and electromagnetic deflection. The principle by which the video signal is produced from the photoconductive surface is illustrated in Fig. 8. A low-velocity scanning beam strikes one side of the surface, the other side of which receives illumination through a signal plate from which the video signal is taken. When the photoconductive surface is scanned in darkness, electrons deposited from the scanning beam reduce the potential to zero. The conductivity becomes so low under these conditions that very little current flows across the surface. If on the other hand, the surface is illuminated, the conductivity increases and charge flows across the surface, and the scanned surface becomes more and more positive in the interval between successive scans. The beam then deposits sufficient numbers of electrons to neutralize the accumulated charge, thereby generating the video signal. The photoconductor incorporated in the vidicon sensor consists of a selenium derivative. Integral to the photoconductor surface is a 5-by-5 matrix of dots comprising a reseau, which is utilized in correcting the image information for nonlinearities and distortions. Additionally, a reference mark is included in each corner of the scanned format, which provides, in the video signal, an electronic level representing optical black for photometric reference.

The electronic circuits necessary to provide the timing, power, and amplification functions required for proper camera operation are solid state and are packaged in module form as depicted in Fig. 9. This circuitry comprises five functional groups consisting of the drive circuits for lens and mirror mechanical positioning, the video amplifier, the horizontal and vertical sweep circuits that create the scanning raster, the synchronization circuitry for ground recording and reproduction purposes, and an electronic conversion unit to provide the necessary voltages and regulation from the spacecraft central power source for camera operation. Thermal control devices are located within the camera; surrounding the vidicon faceplate, on selected electronic modules, and within the mirror assembly to provide and maintain operational temperatures when the camera experiences low transit and lunar temperature conditions.

Camera modes of operation

Functionally, the camera operates in what is termed a slow scan mode in contrast to a "standard" scan such as is utilized in commercial television. Such a reduced scan rate requires less information bandwidth from the spacecraft communications system for a given picture quality, thereby reducing the RF power requirements for the lunar distances involved. In the normal mode of operation, the camera provides one 600-television-line frame every 3.6 seconds. Each frame requires 1 second to be read from the vidicon. A period of 200 milliseconds is utilized to transmit the lens and mirror position information plus several temperature measurements. The remaining 2.4 seconds are utilized in erasing the image from the vidicon in preparation for the next exposure interval. A second mode of operation is available in the camera which provides one 200-line frame every 60.8 seconds. Each frame requires 20 seconds to complete the video transmission and utilizes a bandwidth of 1.2 kc/s in contrast to the 220 kc/s used for the 600-line mode. This 200-line mode is used in instances of omnidirectional antenna transmission from the spacecraft.

A third operational mode, used in instances of stellar observations and lunar surface observation under earthshine illumination conditions, is referred to as an integrate mode. This mode may be applied, by earth command, to either the 200- or 600-line scan mode and creates a condition whereby the scanning beam of the vidicon is cut off while the shutter is allowed to remain open. Such a configuration allows a continued charge buildup on the vidicon proportional to the received photon energy. Readout of the vidicon is commanded from earth after a given, predetermined "integrate" period with the resulting video output being proportional to the photons received and the integrate time. Scene luminances on the order of 0.008 footlambert are reproduced in this mode of operation, thereby permitting photographs under "earthshine" conditions. An example of such earthshine photography is shown in Fig. 10, which shows the spacecraft leg and footpad as well as the lunar surface illuminated by earth at a luminance level on the order of 0.05 footlambert.

Integral to the spacecraft and within the viewing capability of the camera are two photometric/colorimetric reference charts. These charts—one on an omnidirectional antenna and the other on a spacecraft leg adjacent to the footpad—are located such that the line of sight of the camera when viewing the chart is normal (±3°) to the plane of the chart. Each chart is identical and contains a

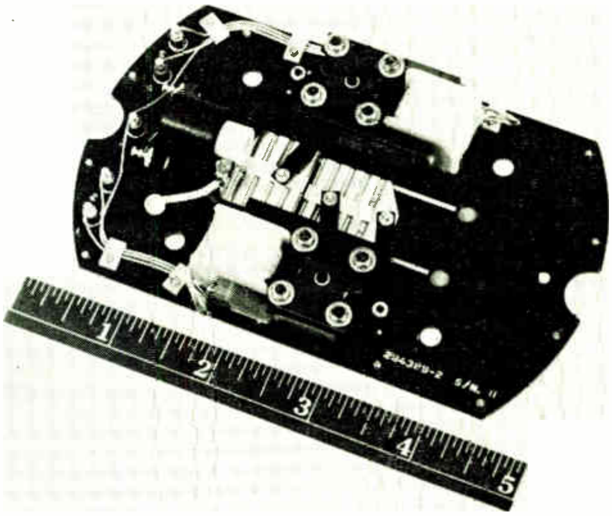


Fig. 7. Focal plane shutter assembly.

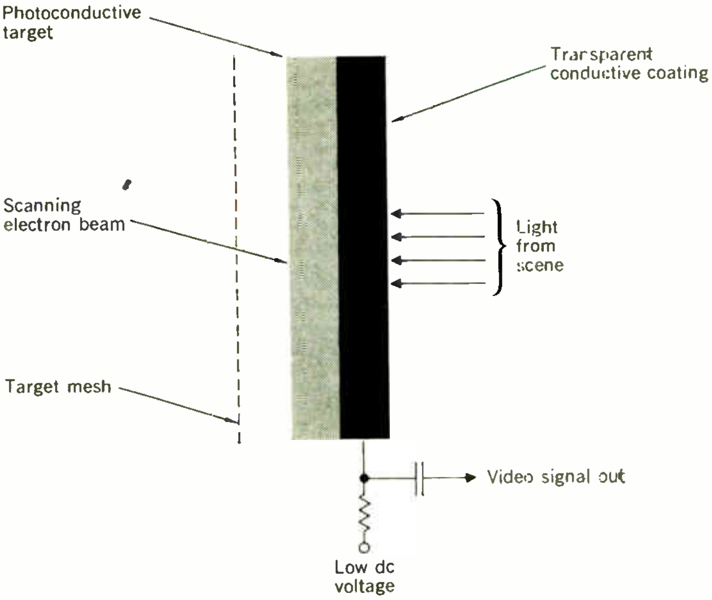
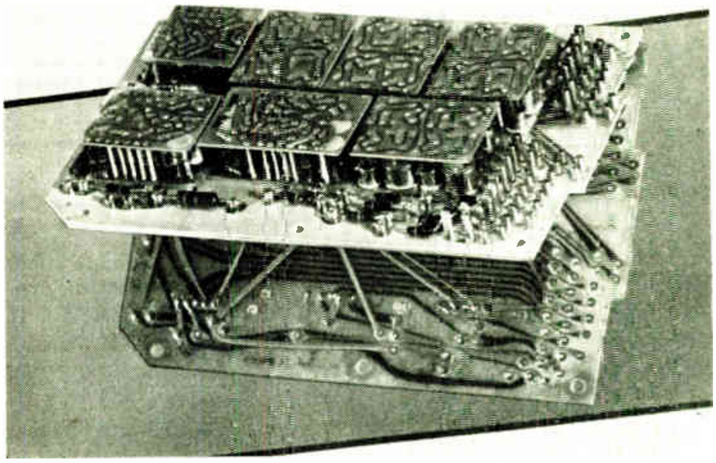


Fig. 8. Functional diagram of vidicon transducing process.

Fig. 9. Electronics chassis showing modular construction used in the camera assembly.



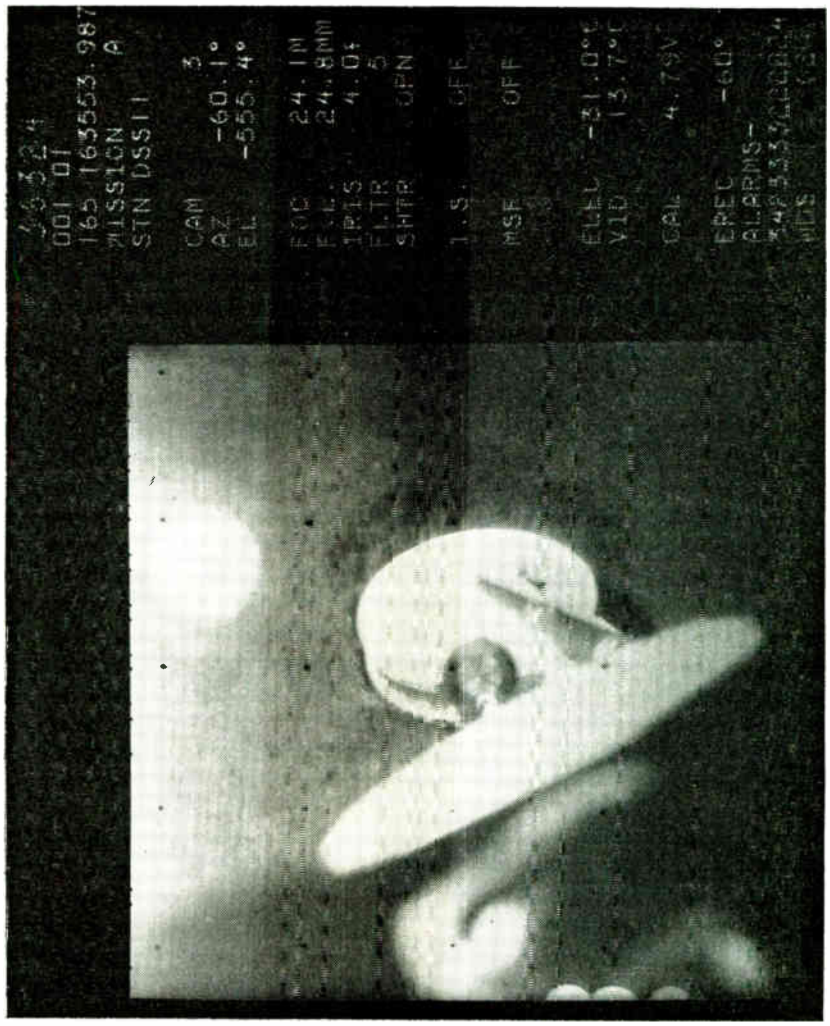


Fig. 10. Earthshine photograph of the spacecraft landing pad and lunar surface taken with a luminance level of 0.05 footlambert in the camera integrate mode. This was the last photograph following the lunar sunset.

series of 13 gray wedges arranged circumferentially around the chart. In addition, three color wedges, whose CIE chromiticity coordinates are known, are located radially from the chart center. A series of radial lines are incorporated to provide a gross estimate of camera resolution. Finally, the chart contains a center post that aids in determining the solar angles after lunar landing by means of the shadow information. Each chart, prior to launch, is calibrated goniophotometrically to allow an estimation of postlanding camera dynamic range.

Camera calibration

To derive fully the maximum scientific information available from a photograph, it is necessary to have precise quantitative information on the camera that obtained the photograph in terms of those parameters which serve to describe the quality of the image. A calibration was performed on the Surveyor I spacecraft with the camera mounted on the spacecraft. Each calibration utilizes the entire telecommunication system of the spacecraft, thereby including those factors of the modulator, transmitter, etc., that influence the overall image transfer characteristics. This calibration was performed at the launch complex, as close to the launch day as was practical consistent with the overall launch operation schedule of the spacecraft.

The calibration information is utilized both prior to the mission and during the postmission data-analysis period. Prior to launch, the entire television ground data-handling system was adjusted and calibrated utilizing the prerecorded spacecraft/camera video signal derived during the calibration of the camera. This allowed the ground equipment to be optimized for the particular spacecraft in terms of the real time receipt and processing of the image information. With respect to the postmission analysis, the camera calibration information is used to correct the images for geometric nonlinearities and distortions, frequency or aperture response, photometric nonuniformities, and coherent noise. Additionally, the photometric information is used to convert the video levels to absolute lunar luminance units[1] and subsequently to convert the pictures to elevation maps.[2,3]

Digital computer techniques[4,5] developed and utilized in conjunction with the Ranger and Mariner photographic experiments and applied to the Surveyor imagery, allow the correction factors to be applied to any selected frame of video in a preprogrammed manner. An example of the application of aperture correction by the use of digital techniques is shown in Figs. 11 and 12. Figure 11 depicts the original film data, whereas Fig. 12 illustrates the same photograph after the application of the sine wave correction. The correction shown in this particular photograph represents a "flat" response out to the 20 percent relative response point on the sine-wave response curve.

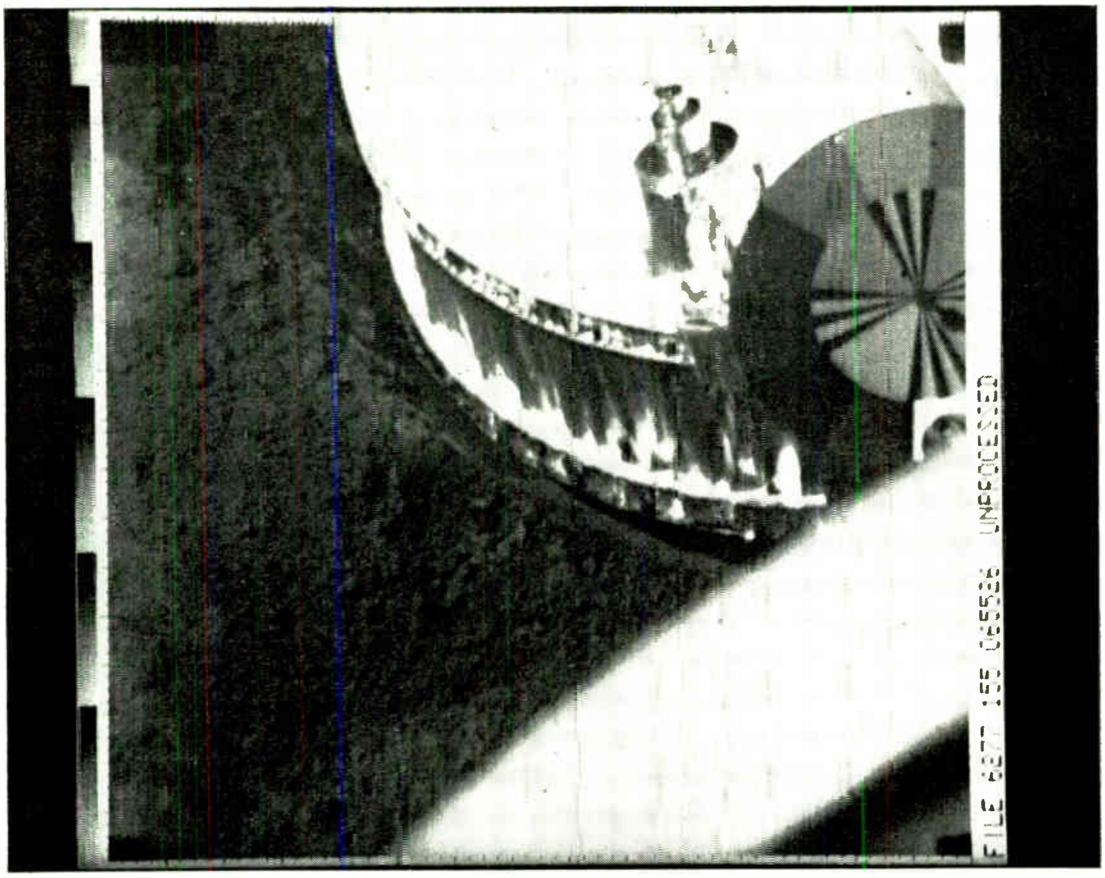
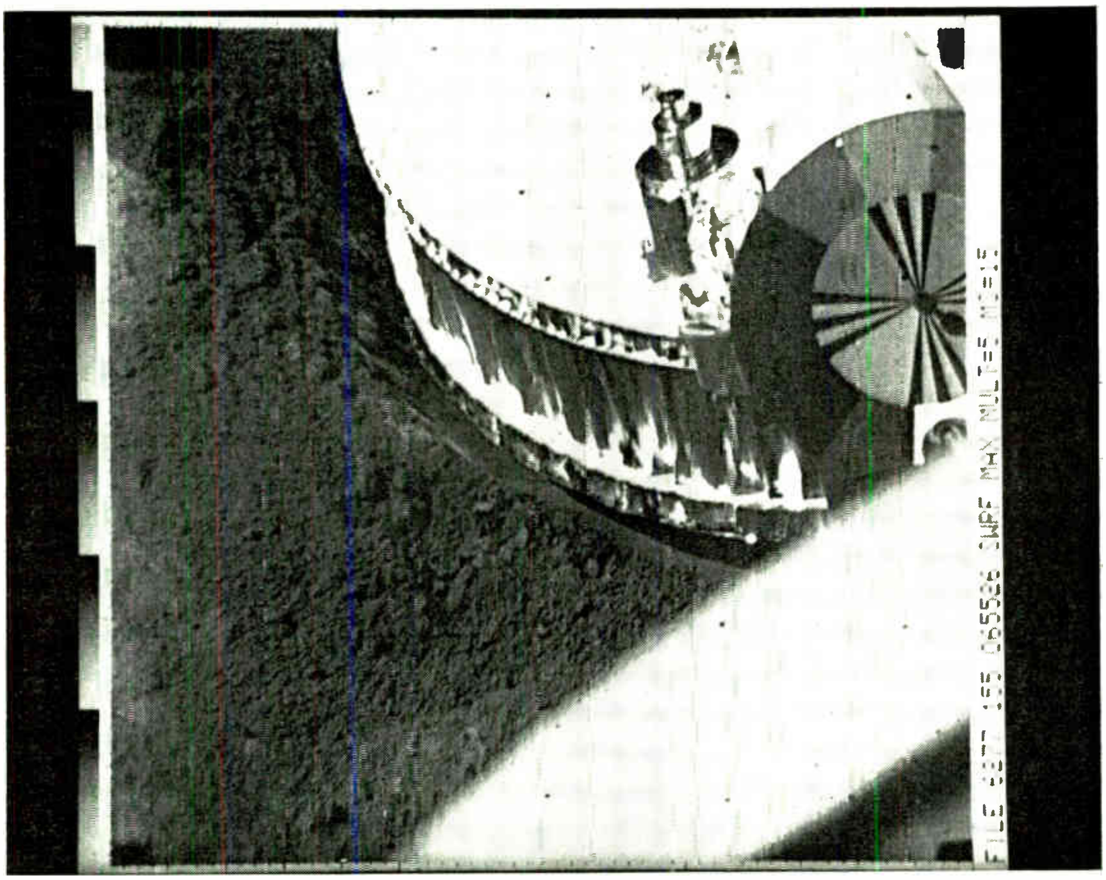


Fig. 11. Surveyor I unprocessed photograph of the lunar surface showing the spacecraft footpad, photometric chart, and the lunar surface disturbed by the spacecraft during landing.

Fig. 12. Same photograph as in Fig. 11 with sine-wave correction applied.



Those factors or parameters of the camera that control the first-order effects in the resulting images are the dynamic range or light transfer characteristic, the modulation transfer or sine-wave response, the geometric distortion, the signal-to-noise ratio and the shading and vignetting of the lens/vidicon combination. It is, therefore, primarily these parameters that are calibrated extensively on the Surveyor camera.

The calibration stimuli for the camera system consist of test slides accurately calibrated and configured to be placed in a special light source. Several sine-wave slides are used to determine modulation transfer or sine-wave system response. The true sine wave is used in contrast to the more often used square wave, thus enabling a determination of the true Fourier representation of the camera. Another slide consists of a series of gray scale wedges used to determine the erasure characteristics of the vidicon, thereby enabling a correction to be applied as a function of latent image level resulting from previous exposures. Finally, a grid pattern is used which, by means of either manual or computer techniques, allows the nonlinearities and distortions to be removed from each image. The light transfer characteristics and shading measurements are obtained by exposing the camera to a series of uniformly illuminated light fields, each progressively brighter, until a saturation point is experienced.

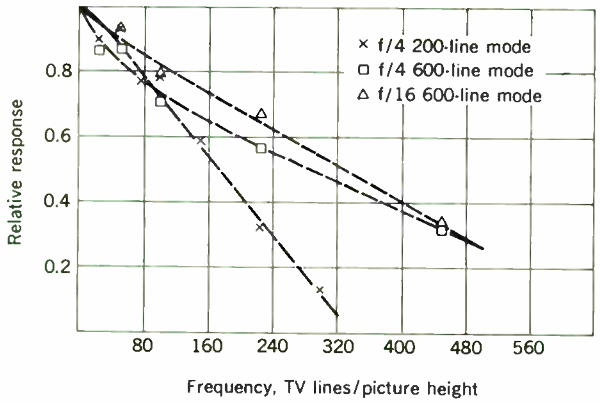


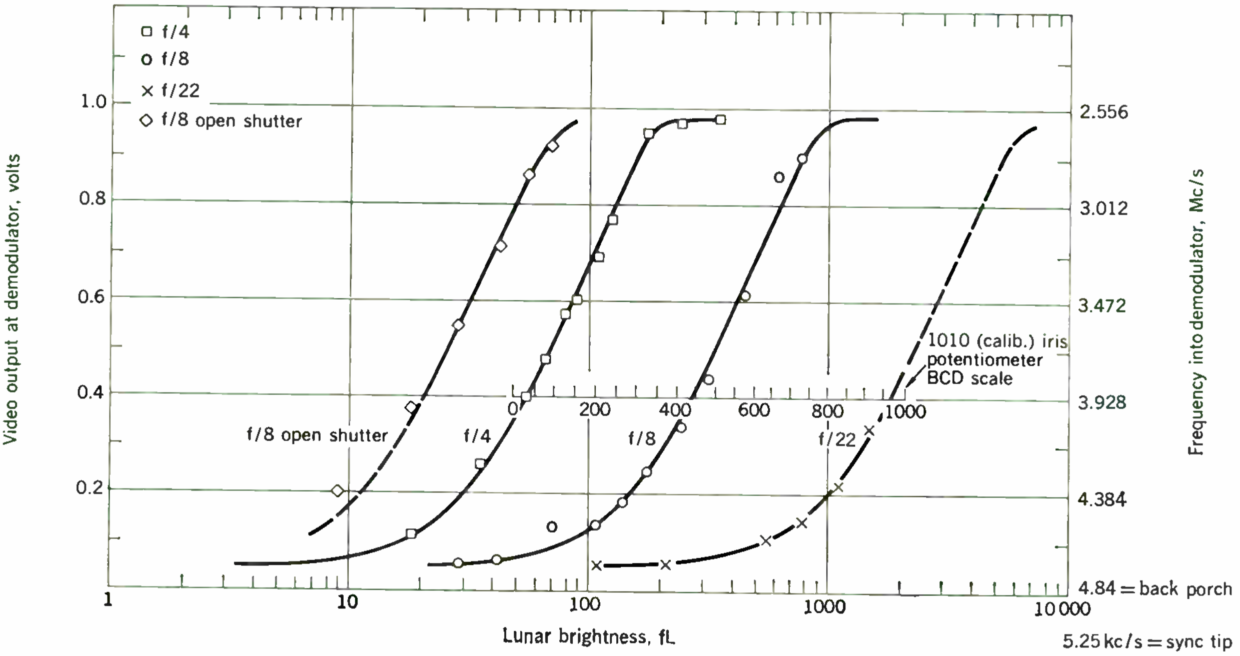
Fig. 13. Camera 600-line light transfer characteristics as a function of lunar brightness and plotted in terms of FM frequency deviation.

Typical data representative of the type obtained during the camera calibration include that shown in Figs. 13 and 14. Figure 13 indicates the light transfer characteristic of the camera in one mode of operation. It is based on the actual lunar scene brightness as determined through appropriate correction factor calculations. These correction factor calculations involve the spectra of the camera, the standard eye, the measuring photometer, the light source, lunar light, and a separate National Bureau of Standards calibration light source. Figure 14 illustrates the sine-wave response characteristic in terms of a relative response (normalized to the dc component) with respect to spatial frequency in television lines per picture height.

The calibration equipment was mounted adjacent to the spacecraft camera at the launch complex. The light source, which utilized mercury-zenon lamps in conjunction with an integrating hemisphere, with its attendant power supply was elevated on a test stand to bring it into position with the camera. A photometer was used for continuous monitoring of the light source luminance.

The recording playback configuration is depicted in the block diagram shown in Fig. 15. A predetection method of recording is used, thus providing calibration information free from unknown factors and nonlinearities associated with uncalibrated video test equipment. The tape recorder is an Ampex VR 1560 rotating head, helical scan machine appropriately modified to accept slow-scan-type video. Video information was obtained from the intermediate-frequency amplifier in the ground receiver, which provided a 50-Mc/s carrier. This 50-Mc/s signal is amplified and translated to 4 Mc/s. In the 200-

Fig. 14. Camera sine-wave response characteristic.



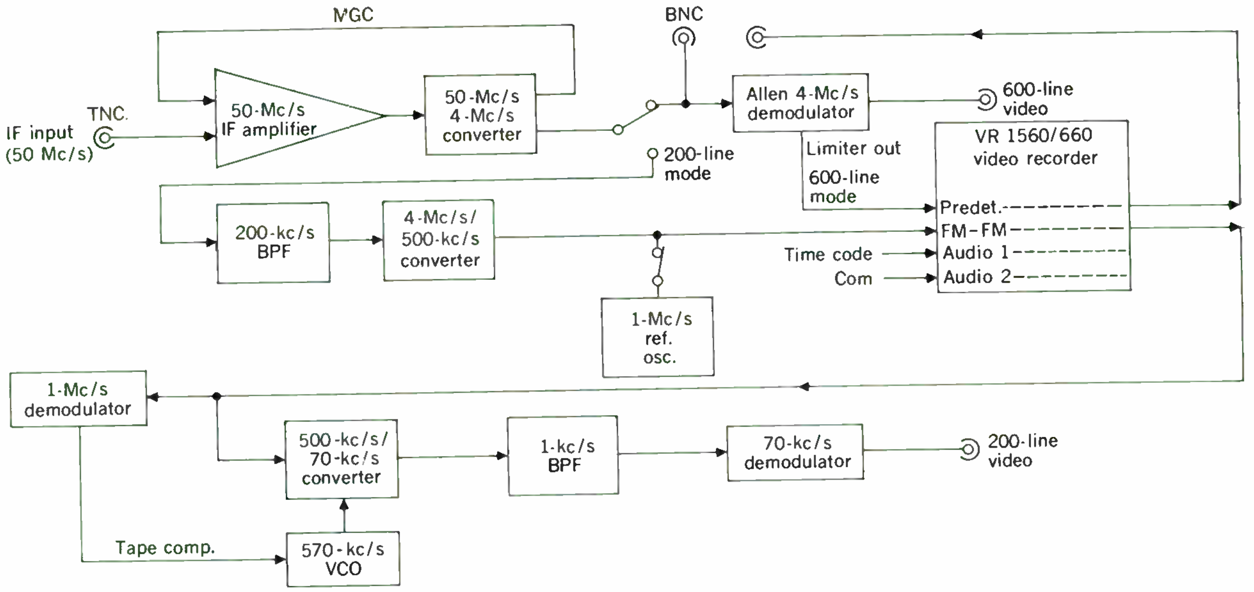


Fig. 15. Block diagram of the video recording configuration used for camera prelaunch calibration.

line mode, the 4-Mc/s signal is further translated to 500-kc/s and then applied FM-FM to the recording machine. A 1-Mc/s pilot signal is mixed with the 500-kc/s information for tape speed compensation during 200-line mode playback.

The 600-line mode playback is accomplished by demodulating the 4-Mc/s carrier through a carefully calibrated pulse-averaging-type demodulator to obtain the baseband video signal. The 200-line mode playback requires another frequency conversion to 70 kc/s for demodulation. An additional 1-Mc/s demodulator separates the tape speed compensation signal from the video and appropriately modifies the 500- to 70-kc/s conversion device.

The baseband video signals are provided to a precision photorecorder and processed on both 35- and 70-mm film for analysis. Additional analysis results from examination of the video signal by electronically selecting and oscilloscope photographing any given scan line within any frame. Additionally, the video signals from the tape recordings are digitized for computer processing as previously described.

The successful landing of the Surveyor I spacecraft provides the United States with a tangible demonstration of a suitable landing site for Apollo astronauts. Additional landing sites will be evaluated on subsequent flights. The television pictures, in conjunction with other data received at the time of touchdown and throughout the spacecraft life, provide significant data pertaining to lunar surface properties vital to the engineering aspects of the Apollo program. In addition, these photographs provide the scientific community with an abundance of data from which conclusions can be inferred relative to the nature and origin of this, our nearest celestial neighbor.

Postscript

Following the lunar night during which the camera temperature reached −280°F, attempts were made to interrogate the spacecraft as the solar angle increased The spacecraft failed to respond until the lunar noon (July 6) when, at 11:29:10 GMT, the Deep Space Station at Canberra, Australia successfully commanded the spacecraft transmitter on.

After completing an assessment period to determine the engineering status of the spacecraft, a series of 24 pictures were received from the camera. An analysis of these pictures indicated that the camera had adequately survived the lunar night without any observable deterioration in image quality.

Subsequent tracking periods from Goldstone, Calif., and other overseas tracking stations produced approximately 634 additional pictures. On July 13, contact was lost when the spacecraft suddenly stopped transmitting a picture to a South African tracking station.

This paper presents the results of one phase of research carried out by the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NASA-7100, sponsored by the National Aeronautics and Space Administration.

The authors wish to express their gratitude to the members of the Hughes Aircraft Company whose devoted efforts made this system and its attendant historical photography possible. In addition, they wish to extend special recognition to Dr. A. C. Dunk, D. E. Willingham, G. M. Smith, and H. E. Wagner of the Jet Propulsion Laboratory for their contributions in the areas of calibration and photographic data processing.

REFERENCES

1. Smith, G. M. and Willingham, D. E., "Ranger photometric calibration," JPL Technical Report 32-665, Aug. 15, 1965.

2. Rindfleisch, T., "Photometric method for lunar topography," *Photogrammetric Eng.*, Mar. 1966.

3. Rindfleisch, T., "A photometric method for deriving lunar topographic information," JPL Technical Report 32-786, Sept. 15, 1965.

4. Nathan, R., "Digital video data handling," JPL Technical Report 32-877, Jan. 5, 1966.

5. Billingsley, F. C., "Digital video processing at JPL," Paper No. 15, Seminar Proc. of Electronic Imaging Techniques for Engineering, Laboratory, Astronomical and Other Scientific Measurements, Society of Photo-optical Instrumentation Engineers, April 26–27, 1965.

New means of communication

Universal communication systems that have the flexibility and modalities of the human nervous system—with which men could, at will, talk, hear, see, gesticulate, write, examine documents, and even manipulate machines—are coming within reach. All we need to do is foresee what is really wanted or needed, a major problem in itself

John R. Pierce *Bell Telephone Laboratories, Inc.*

In the past, men viewed different modes of communication, such as telegraphy and telephony, as being fundamentally different in kind, a view that has been dispelled by technological advances and especially by the communications theory of Shannon. The way is now open to achieve a universality in electrical communications systems analogous to that of the human nervous system in which different sensory functions—vision, hearing, taste, etc.—are served by uniform nervous transmission lines that differ only in the different types of sensory transducers in which they terminate. There are technical problems and challenges in achieving a future universal communication system that will serve all man's needs, but there are also other obstacles rooted in obsolete social and legal concepts.

I want to begin by observing an interesting contrast between man's internal communication system and the electrical communication systems that serve him. Over the years, neurophysiologists have learned a good deal about the nervous system. They have found that nerves are the same, whatever sense they serve. Electric impulses travel along nerves, and are chemically regenerated as they travel. Sometimes there are pulses in the absence of a stimulus, sometimes only during the stimulus. But the pulses themselves are the same, whether we are listening, smelling, or hearing. Indeed, the same sort of nerve impulses are used in controlling our muscles.

The difference between the various modalities of communication or sense within the human being are merely differences in the types of transducers at one end of the nerve. Certain nerves are sensitive to touch, others to heat, others to light or to vibration, but all generate the same sort of nerve impulses. We do not understand why we have different sensations—of sight, of sound, of smell—but clearly this must be related to the fact that the nerves associated with different senses are connected to different receiving mechanisms, different sensory transducers.

Thus, in all the varied forms of communication that we use—speaking, seeing, gesturing, and so on—our internal communication signals are carried in the same way, and we see a difference in purpose, if you wish, only in transducers and receptors.

Finding unity amidst diversity

We can find a unity among different methods of electrical communication by recognizing that the different methods all use electric currents. However, in the early days of electrical communication we did not see this unifying principle. Instead, we were acutely conscious of special inventions of a rather primitive kind.

Samuel F. B. Morse invented a form of the telegraph that made use of off–on signals and the very ingenious Morse code. The signal from the key was effective over medium distances; to send the signal further, you provided relays at intermediate points. These relays would respond only to off–on signals. At the far end of the telegraph circuit, a sounder made a sound.

This was not, however, the whole of the communication system that was then known as telegraphy. There was also the messenger boy. At the sending end we had to write out the message, and at the receiving end the messenger boy had to take it to someone's home.

At one time, telegraph service of this type was an almost universal form of communication; it could reach from a telegraph office to almost any home in the country. Indeed, business offices were provided with electric signaling devices for summoning telegraph messenger boys so that businessmen need not go to a telegraph office to send a message.

When Alexander Graham Bell invented the telephone, it seemed very different from the telegraph. The messenger boy, who had been an integral part of telegraphy, was eliminated. Telegraphy would have meant nothing without the human link at the ends, for the message would not have reached anybody. However, the telephone was on the premises at both ends, and the telephone circuit was not just a circuit extended between two offices. Through switching, the telephone enabled one to communicate directly with any one of a number of people. Another difference between the two systems was that the telephone signal did not travel over a telegraph circuit equipped with relays.

You will notice that I had to retrogress many years to make telegraphy and telephony look different. As electrical communication developed, it was found possible to send telegraph signals over telephone lines. In fact, today telegraph and teletypewriter signals are carried over the same communication facilities that send voice. A small part of the bandwidth is allocated below the voice channel, or else the voice channel is filled with "carrier signals" of different frequencies, which enables many telegraph signals to be sent over one circuit. The contrary is also true. In pulse code modulation, which is now in commercial service, one reads the successive amplitudes of a continuously varying signal, and represents these by numbers, which in turn are represented by strings of binary, off–on impulses. In pulse code modulation the signal is not a continuous signal, but is discrete, limited to certain amplitudes sent periodically at predetermined times. This is certainly reminiscent of telegraphy.

The communication plant that has evolved in the telephone system, and which carries telegraph messages (or data, or teletypewriter), has not achieved the degree of uniformity that we find in the nervous system, where everything is sent as pulses. It has, however, achieved the uniformity of function. If you pick up your telephone or use a teletypewriter, you have no way of knowing how the signal is being transmitted. It may be transmitted by the on–off pulses of pulse code modulation, or by a baseband signal translated in frequency and sent via coaxial cable or via a microwave radio relay.

The electronic art has become so flexible that we can achieve the same overall effect from terminal to terminal with any of the circuits in existence. Furthermore, we can predict that future circuits will accommodate any signals we may wish to send.

Telegraph service as we once knew it is a thing of the past. The messenger boy is gone (probably working as a communications engineer by this time). What we now think of as telegraphy involves telephoning a message to the telegraph office.

Communication theory

We have at last achieved in electrical communication something that in effect is very much like the universality of communication within the human being. Different aspects of communication are served by communication circuits, and the purpose being served depends on the terminal equipment at the ends of the circuits, not on the circuits themselves.

This has all been formalized mathematically in Shannon's communication theory. Shannon defines a common unit of information that is equally applicable to voice signals, facsimile signals, television signals, or data signals. That measure or *bit* information is a binary choice, a yes or no. Shannon defines a *capacity* of communication circuits, likewise universal, and not dependent upon the modality of communication. This measure is the channel capacity of the circuit, measured in bits per second.

Shannon points out that we have in communication a source generating information at a certain rate that can in principle be deduced by examining the signal which the source generates. We can describe the capacity of circuits to transmit such signals in bits per second. This makes it much easier to talk about the present state of communication and the new possibilities of communication than it would be were we living in the day of primitive inventions when each system was viewed as a distinct and different type.

Today, from every place of business, from almost every home, there extend electric circuits that can be set up through the telephone plant. Various transducers can be connected at each end. It is the character of these transducers that determines the particular aspect of the human need for communication the circuits will serve. This, then, is closely analogous to human internal communication, in which we have nerve endings sensitive to sight, sound, touch, taste, and so forth, but a common transmission medium that serves all messages. Today, to the same communication network you can connect a keyboard, a typewriter, a telephone instrument, or a television device.

Let us consider such communication circuits. How much channel capacity must they provide for various sorts of communication? Telegraphy, for example, takes little channel capacity, voice takes a little more, fast data take a lot more, and television takes a great deal more. In meeting communication needs there are two problems. The first is that of expanding the capability of the already universal network that makes use of coaxial cables, pairs of wires, microwave radio relays, submarine cables, and satellites. The second has to do with providing transducers that will enhance the human process of communication through sight, sound, and, perhaps ultimately, touch and smell.

New communication systems

What can we see in the future? This broad universal commodity—electrical communication—is useful chiefly because it *is* universal; because it will connect anything with anything else. A new communication circuit would be of little use if only a few people could use it or if it could be used only in a restricted area. Not everyone need have a new form of communication at first, but it is essential that it be universally available, that it extend, say, from New York to Los Angeles.

Here we must confront the problem of channel capacity. How can we expand the channel capacity to provide more communication—faster and cheaper? We must disregard particular human uses of communication channels, because that is not relevant. All sorts of communication must go over the same circuits.

There are many promising new approaches to more economical communication. First, there is the extension and improvement of things we already have, through advances such as the transistor and, more recently, through the revolution of microelectronics. Components too costly yesterday are not costly now and will be cheaper tomorrow. Obvious examples are the pocket radio, alerting devices, and improved radio receivers in automobiles. Another example is a communication system called the T-1, that transmits some 1½ million plus or minus pulses per second over a pair of wires that might otherwise carry one voice conversation. Transistors have doubled the communication capacity of a coaxial cable with respect to what could be attained with vacuum-tube amplifiers.

The transistor is also bringing us microwave systems that are simpler and cheaper than those of the past. The first microwave systems were extremely expensive. They had huge standby power plants, and repeaters were located on expensive hilltop sites about 30 miles apart. Now you can make a microwave repeater that is self-contained. The standby power is a couple of storage batteries. You do not need expensive equipment or sites. You can use such systems on all kinds of routes, all over the country.

Further developments will make communications by microwaves even cheaper and more flexible. Beyond this, however, there is a problem of transmitting very broadband signals, such as those associated with television, across the continent. We know that the art is available for sending millimeter waves through pipes or waveguides between amplifiers perhaps 15 miles apart. One pipe could be used to transmit 100 000 two-way telephone channels or 200 channels of television bandwidth.

Such a waveguide system could come into being when there is an economic need for it. Technically it is already feasible, but it is not clear whether so much communication capacity is needed. Beyond millimeter waves we have coherent light, a truly new thing under the sun. Coherent light from lasers has inconceivable bandwidth. We know how to guide it through pipes with gaseous lenses. The art is not enough advanced to build a useful system, but a practical resource can be created through further work. Thus, by utilizing future developments, we can find ways to expand our electrical communication to meet future needs.

The unforeseen needs

I am sure that in the future there will be needs that we cannot foresee—but there will also be ways that we cannot foresee for providing more communication. Progress is a matter of working hard, and then recognizing the right time when economic conditions call for new things.

Some years ago it seemed to us that it would be very expensive to bridge the oceans with broadband communication so that we could communicate as readily across the world as we can across the continent. Today we have broadband submarine cables; and transistorized cables will have a capacity of an order of magnitude greater than the first vacuum-tube cables. We also have communication satellites. These are another piece of a complicated network that can send any signal to a far point over many media—partly over a pair of wires, a coaxial cable, microwave radio, and communication satellite. At the ends of the circuit you cannot tell what sort of circuit is being used.

The challenge of broader band and more economical communication circuits is a great challenge, but it is one I think will be met. There is another great challenge. That is to adapt this marvelous communication network to the needs of man—needs that can be foreseen, and needs that cannot. Needs are not necessarily something that exists independently of means. They come into being and grow. There was no need for the telephone before the invention of a telephone. People got along very well without it. If the telephone had been destroyed a year after it came into being, no one would have missed it. If it were destroyed now, society would be in chaos. The same can be said about the automobile. Thus, needs come into being and grow as new means come into being. We cannot always foresee needs. Nevertheless, we can see many of them.

Civilization today does not exist merely in a few big cities. It extends over the face of the country and over the face of the globe. In doing the world's work, people continually fly from the West Coast to the East Coast just to converse with one another for a short time. This is wasteful. The happy future would be one in which we could use electrical communication with something approaching face-to-face flexibility—a world in which we would not choose among modes of communication but would use them all. The communication system we need would enable us to talk, to hear, to see, to gesticulate, to write, to examine documents, and even to manipulate machines.

Why should you have to go to a bank or call a bank to tell somebody to do something to your account? Why should you have to call somebody in a department store to tell him your selection, whereupon he must enter it into the books and charge it to you? This human link seems as unnecessary as the telegraph messenger boy. Obviating this link, too, is a great challenge.

The challenge is partly technological, but it is also partly that we do not yet understand human beings as well as we need to. What things should come first? What things are most important? In a conference, for instance, you do not want everyone holding a telephone in his hand—or is this true? You would probably like distant talking devices. But this desire raises technological problems because of feedback from the loudspeaker to the microphone. This problem can be tackled in a number of ways, all of which have limitations. Furthermore, in a conference between two groups of people, how do you know who is talking? Experience indicates that if people know one another very well they can keep the conferees straight because they recognize the voices. If you are in a conference with unfamiliar people you are nevertheless able to see them, and you learn to identify them even before you hear them speak. That factor is missing in conferences via telephone today. Perhaps you need television, but what kind of television do you want? Must it provide sufficient detail so that you can read maps? Or do you want a keyboard along with your conference? Or do you want a telewriting machine? These things must be tried out. I do not think you can guess what will be needed.

Another thing that is missing from today's communications systems is the element of secrecy or security. For classified Washington business, a device is needed at the terminals that will frustrate wiretappers. This can be done, but it may be costly.

It is more difficult to evaluate such options than you might believe. You try a combination and find that some people get along with it. But how much can you generalize? How much is necessary? You provide people with a lot of things; how much can you take away? It is very difficult to judge. In the end, all these services will have to be offered to the public, but it is difficult to know when they should be offered, or in what order. It is certainly possible to have television from home to home, or from office to office, but at present there are obstacles. One is the cost and the reliability of the terminals, although transistors help greatly in reducing terminal cost and increasing reliability. Another obstacle is the cost of the circuits, although advances will make transmission less costly.

I predict a great many advances in improving both circuits and terminals. Probably we will have telewriting and Picturephones, such as were demonstrated at the 1964-65 New York World's Fair. Advances will also make it possible for us to communicate with machines—for instance to give problems to a computer over a telephone. In the future when you need information or service you will be able to call a computer. You will be able to query libraries as well as people.

Nontechnical obstacles

There are obstacles other than the technological obstacles in going toward the future in communications. There are legal obstacles to face. Traditionally, the Federal Communications Commission tends to be interested not in communication but in modalities of communication. Conforming to this pattern, Congress has set up satellite communication as a particular technique that is operated independently of common carriers although it ties into their circuits in providing a service to users.

Such thinking leads to queer questions such as: "Would you like some improved form of telegraphy, or would you like a telephone?" This is reminiscent of the question: "Would you rather be blind, or would you rather be deaf and dumb?" It seems to me to be an unnatural question. I hope that, in the long run, laws and regulations of communication will leave room for a universality and improvement of communication so that we can all use in our daily affairs all the modes of communication with which technology can economically supply us.

This article is an edited version of a talk given at the Columbia Engineering Centennial Symposium, Columbia University, New York, N.Y., May 5, 1965.

Earthquake or explosion?

The science of nuclear test

Nilo Lindgren *Staff Writer*

This article, which follows up on the December 1965 Proceedings of the IEEE special issue on nuclear test detection, outlines the problems of underground detection and identification, offers some background in geophysics that the nonspecialist needs to get a perspective on the niceties of the engineering problems, cites some of the criteria whereby underground events are discriminated, and assesses the present state of the art. It concludes with a description of LASA, the largest seismic array ever built, and speculates on the possible political effects of the recent scientific and technical developments.

Man has always had a lot of pests on his hands that have made his existence on this globe somewhat less than paradise. One of the most fearsome pests is the earthquake, which strikes without warning, often in unpredictable places, at irregular intervals, and with scourging effect. Its "attendant phenomena" have been calculated by the gods to bemuse mankind with terror and curiosity: mountains have been moved, islands have sunk or risen, massive tsunamis from the sea drive up onto the shores, waterspouts shoot out of the naked land, strange lights are seen on the horizon, and horrible grindings and roarings issue from the underworld. All very pleasant. One needn't go into the number of lives lost, buildings wrecked, or fires set, those gory statistics of which journalists never tire. In recent times, man himself has managed to create a pest of equally vivid dimensions, the nuclear bomb.

Since the partial test-ban treaty of 1963, which bans nuclear tests in the atmosphere, under water, and in space, the most significant connection between nuclear bombs and earthquakes, of course, is that the bomb, exploded under ground, is very much like an earthquake. Or is it? Just how different is an underground explosion from an earthquake, and how can you tell the difference? The effort to answer those questions has engaged some of the best scientific and political talent available, including leaders of the most powerful political aggregations on earth. A lot of work and money have been ploughed into the effort—upwards of 340 million dollars (in the total nuclear detection program) in the last six years in the United States alone. In this short time, under the guidance of the Advanced Research Projects Agency (ARPA) in the Department of Defense, in the so-called VELA UNIFORM program, the science of seismology, which had been simmering along for decades, living on a shoestring, has been brought abruptly into the 20th century of Big Science. A measure of the concomitant engineering achievements alone can be found in the December 1965 PROCEEDINGS OF THE IEEE, a special issue on nuclear test detection, constituting a comprehensive coverage of advances in detection and identification in all four environments in which nuclear tests can be made—in the atmosphere, in space, under water, and under ground.

Of these four regimes, it is now clear that nuclear explosions can be detected and identified with relatively great confidence in the first three.

The problems relating to the seismic detection and identification of underground explosions, however, have not proved as tractable in either the technical or political spheres. Politically, as everyone knows, the words that indicate the focus on which opposing views have been hung are "on-site inspection." From the scientific and engineering points of view, what these words signify is that the state of the seismic art of detection and identification thus far has not provided the confidence that certain underground events could be detected and positively identified through means of national systems alone.

The most recent re-enunication of the question came when the 18-nation disarmament conference reconvened in Geneva on June 14. Although the United States delegate, William C. Foster, and the Soviet Union delegate, Aleksei A. Roshchin, both gave priority to extending the present ban to underground explosions, they were once again in dispute on old grounds. Foster "acknowledged that science had made 'substantial progress' in developing ways to detect such explosions from a distance." He added, however, that "hard evidence still points to the need for on-site inspection to verify a comprehensive test ban."[1] Roshchin asserted "that the United States advanced groundless demands for international inspection in order, by preventing an accord, to have a free hand for carrying on dangerous experiments with nuclear weapons."[1] But both men called for a treaty soon to prevent the spread of nuclear weapons, reflecting a concern with what is sometimes called the *N*th-country problem.

Many of the proponents of a test-ban treaty view it as a potential constraint against the spread of nuclear weapons to still further countries, serving as a first step toward further efforts to build even stronger safeguards against proliferation. Only thus can the danger be reduced of accidental or *N*th-country malicious provocation of what everyone in his right mind wishes desperately to avoid.

The treaty environments

Figure 1 schematizes some of the methods whereby atmospheric and space explosions are detected and identified. In the lower atmosphere, the methods used include

detection

Despite substantial progress in the science of explosion seismology brought about under the aegis of ARPA's VELA UNIFORM *program, it appears now that there will always be some earthquakes that look like underground explosions. Potentially how important are such unidentifiable events?*

the detection of the primary bomb radiations and outputs (prompt X rays, gamma rays, neutrons, etc.). These outputs are particularly characteristic of nuclear events as contrasted to natural events largely because of their rapid onset and unique time dependences. Those techniques that detect the secondary and long-range effects must cope with more difficult problems of natural background effects. Atmosphere and space techniques include the following: acoustic; debris sampling by aircraft and rocket; radio flash; satellite detection of X rays, prompt gamma rays, and neutrons; atmospheric fluorescence; detection of VLF, LF, and HF; radio sounders; cosmic noise; magnetic–telluric methods; debris resonance. Natural background effects include meteors, earthquakes, volcanoes, chemical explosions, aurore, tornadoes, winds, lightning, solar radiation, cosmic ray showers, ionospheric disturbances, and so on.

The icosahedron at the top of Fig. 1 depicts a VELA satellite. These satellites, of which six have been launched into far earth orbit since October 1963, represent experimental efforts at developing a system that can detect nuclear detonations at distances up to two astronomical units. Although these satellites are considered experimental, they are reported to be operating successfully and serving as deep-space explosion monitors.

Not indicated in the Fig. 1 cartoon are underwater detection devices. Particular attention has been given to the development of unattended ocean-bottom observatories that can gather many different types of data for long periods. Although these observatories are being evolved experimentally (also under the aegis of the ARPA VELA program), and although ocean-bottom seismic data are still being gathered and evaluated, the operation of such

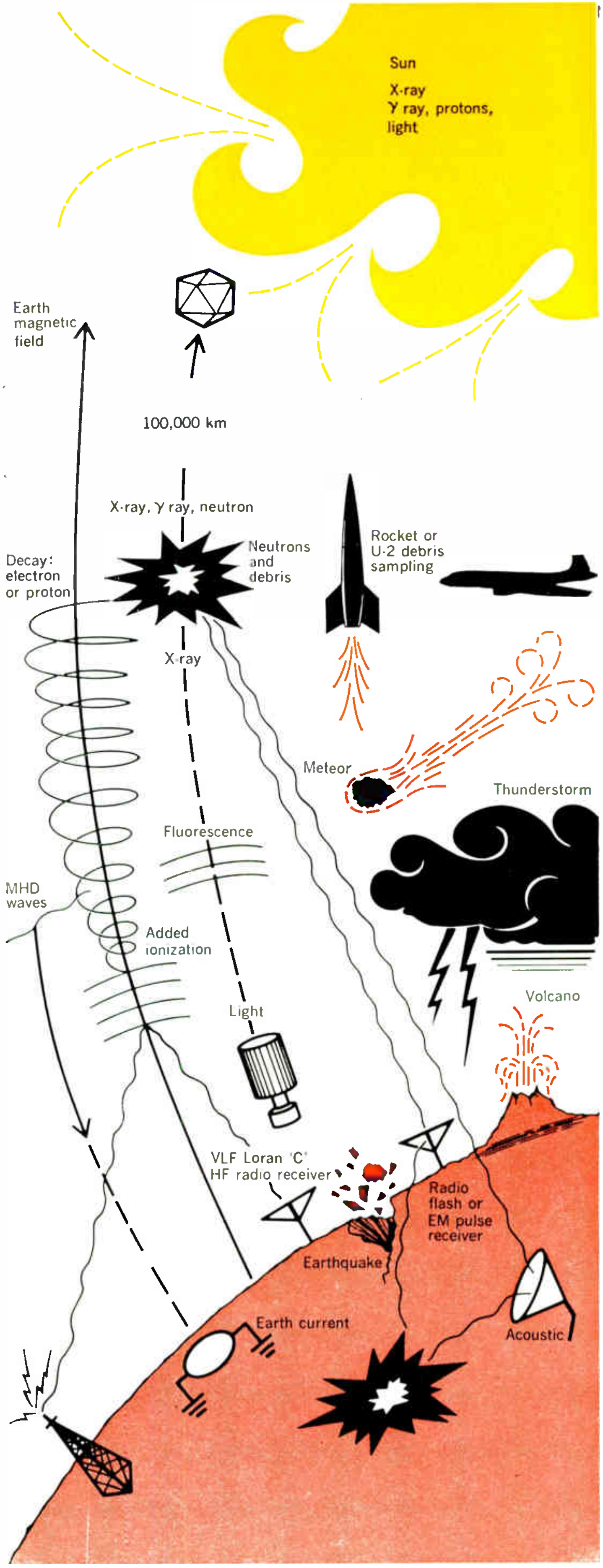


Fig. 1. Cartoon shows some methods of detecting and identifying nuclear explosions in the atmosphere and in outer space along with some natural events that can make detection difficult. Current research aims at improving U.S. capability of monitoring 1963 test-ban treaty.

permanent or semipermanent seismic observatories on the ocean floor is considered to be entirely feasible. These observatories can be equipped with various types of instrumentation. Hydrophones, for instance, can pick up acoustic effects from underwater explosions; seismographs can pick up seismic signals from underground earthquakes or explosions from the continental masses or from underwater blasts whose acoustic energy becomes transformed into seismic waves at the solid earth–water interface. In addition, there is a certain layer or underwater channel (the sofar channel) that acts as a waveguide for water-sound waves over enormous distances. It has been speculated that if a worldwide seismic network is eventually developed, ocean-bottom observatories could perform unique functions in such a network.

The new explosion seismology

The new science of detection and identification of underground nuclear tests involves a challenging and exciting problem in pattern recognition. Out of a mess of seismic hash (called microseisms), picked up by any seismometer (see Fig. 2), which to the nonspecialist may look little different from the recordings of brain waves, the explosion seismologist must detect that a significant event took place somewhere on the globe, and with records from other seismometers, he must determine as accurately as he can the location of the event. Then, by whatever means he has at his disposal, he must try to arrive at a decision as to whether or not the event was an explosion or an earthquake. Moreover, he must try to assess the yield or the energy released by the event.

Although the science of seismology had accomplished some pretty remarkable things in the sixty-odd years since it began making measurements of earth vibrations, it was hardly prepared to answer the questions that the new technological and political events thrust upon it. Explosion seismology was a kind of "gray area," falling between the knowledge and interests of earthquake seismologists and the exploration geophysicists.[2] The seismic energy released in underground explosions was two or three orders of magnitude less than that of the events earthquake specialists had largely studied. Likewise, both the explosive yields of atomic weapons and the distances from which they were to be measured were three or four orders of magnitude greater than anything employed by oil prospectors.

In 1958, when the first international negotiations began on the question of an effective ban on underground nuclear testing, the groups of specialists who were assembled together in the so-called Conference of Experts at Geneva made various recommendations and shrewd guesses, but they returned to their countries with the sharp realization that very much more needed to be known about the nature of underground events. Accordingly, by 1960 there was begun in the United States a vast program of research and development aimed at the advancement of the art of detection and identification techniques. The program was called VELA UNIFORM. (Research on atmospheric detection techniques went on under VELA SIERRA, and space detection techniques under VELA HOTEL.) Out of this program attacking the unknowns in the "gray area" there has grown this new science of explosion seismology. At first, a wide variety of research projects, both theoretical and experimental, were supported—studies of earthquake mechanisms and explosion characteristics, of the earth's crust and mantle and their effects on signal transmission (especially over long distances), of seismic noise; other studies aimed at the development of better detectors and signal processing techniques. Gradually, the program has become more selective, the community of experts has grown larger, and the hardware has become vastly more impressive. But despite the steady acquisition of knowledge and insight, there has been, as yet, nothing that might be called a breakthrough in explosion seismology. Each stage of an adequate nuclear test detection program—the detection in the first instance, its location to a small area such that verification efforts will be effective, its identification as an explosion and not an earthquake, and even the techniques of verification on site—presents unique complexities and difficulties. To appreciate these problems, one must know something about the structure of the earth, the nature of earthquakes, and the methodology of earthquake seismology.

Earthquakes vs. explosions

The story of how earth scientists have pieced together their picture of the interior of the earth through the study of earthquake phenomena is a fascinating one (see, for instance, John H. Hodgson's most readable book[3] or Richter's encyclopedic works[4]), although we shall concentrate here only on certain elements that suggest some relationship to the problem of explosion identification.

Theoretically, one might assume that earthquakes could occur anywhere. However, the accumulation of earthquake records over the past 60 years (instrumental studies began early in this century) has revealed that there are definite earthquake or seismic zones. These are regions where earthquakes occur with relatively great frequency, and which seem to be linked to some kind of principal failure pattern of our planet. About 90 percent of the world's largest earthquakes occur around the boundary of the Pacific Ocean, the so-called circum-Pacific zone, the same circuit that volcanologists call the ring of fire. A second major seismic zone extends from the Azores, through the Mediterranean, along the northern border of India, finally connecting with the circum-Pacific zone. Analogously, we may say that underground explosions are likely to occur only in certain places.

Earthquakes occur at different depths under the surface of the earth, some occurring as much as 600 kilometers below the surface. However, the most numerous occur in what is defined as the normal range, from the surface down to 70 kilometers. This is also the range having the most bearing on nuclear explosion seismology since underground explosions will be relatively shallow, usually no deeper than several kilometers. The point of onset of an earthquake, the place where the earth faults, is called the focus (or hypocenter), and the place on the earth's surface vertically above the focal point, where the earthquake damage is usually most severe, is called the epicenter. Very large earthquakes (over magnitude 8.0) are relatively few (only ten of magnitude 8.6 occurred between 1918 to 1955), but as earthquake magnitudes decrease, the number of earthquakes increases dramatically.

Without going into details of how the equivalence has been derived, a major problem in itself, we can begin to get a first rough idea of the problems of nuclear explosion detection through an examination of Fig. 3, which relates the number of shallow earthquakes that occur all over the world to the yield of an underground blast. Although the relationship shown in Fig. 3 would be modified somewhat in light of recent more accurate data, the import of this chart remains unchanged. It tells us that, theoretically, an underground explosion of one kiloton would have to be separated (in the detection and identification process) from nearly 10 000 events (viewed from a worldwide basis). That figure does not represent the magnitude of the "realistic" problem involved in detection, but it does give us a first rough cut at the shape of the problem. Smaller earthquakes are more numerous still, to be counted in the millions. Looking at the multitude of events at this end of the scale, one can then seriously wonder whether a detection system could ever be perfected to the point that all underground events can be detected. In answer to that point, there was evinced from Dr. Carl Romney of the Vela Seismological Center, Air Force Technical Applications Center, several years ago, this most provocative image: "One can picture the case where you can reduce the threshold (of the detection system) further and further until finally you hear a continuous roar from the millions of earthquakes which are going on around the world."[5] This continuous roaring is the background "noise," the irregular hash, which varies from site to site but inevitably sets a lower threshold below which specific underground events cannot be observed, much less identified as to source mechanism. This continuous underground noise (microseisms) comes from many sources—storms at sea, surf, small earthquakes, man-made noise, winds shaking trees and their roots, movement of cattle. This recitation of microseism sources may give the layman a qualitative notion of the relative sensitivity of seismographs; it may also make him take an uneasy look at the old terra firma.

It should be mentioned that there has been a lack of standardization in seismometer designs, which created problems in the calibration and correlation of data obtainable from stations all over the earth. The installation of 125 standard seismic stations around the world carried out under the VELA program, along with the recordings of actual underground nuclear tests (often of known energy, location, depth, and time), has brought a new level of confidence into the calibration and statistical findings of seismology.

As Dr. Romney pointed out in the 1963 Congressional Hearings on these problems, data from mobile stations constructed under the VELA programs, and data from the worldwide network of standard fixed stations, began to be available in quantity in the late spring and summer of 1962. He says: "The turning point in our understanding of the yield-magnitude-earthquake statistics problem followed the detonation of Aardvark, a 37 ± 7 kiloton underground explosion in Nevada on May 12, 1962. This was the largest underground explosion of known yield detonated up to that time. For the first time, high-quality, calibrated data from an explosion were available in quantity at distances comparable to distances at which earthquakes in the U.S.S.R. were, of necessity, recorded. A second large explosion, Haymaker (56 ± 8 kilotons), followed on June 27, 1962, and was of comparable importance."[5] By that summer, the explosion seismologists had established a new relationship between the estimated numbers of seismic events in the U.S.S.R. and their equivalence in terms of underground explosions. Whereas earlier it had been estimated that the equivalent of 2 kilotons in tuff (a form of volcanic material) was between 225 and 900 earthquakes in the U.S.S.R. (i.e., the number of events that would have to be screened above that threshold), the new estimate reduced the number of earthquakes larger than this threshold to something between 125 and 250. Those who would jump to the conclusion that the detection and identification problem had been simplified by a factor of two at that point should be warned of other complications. Somewhat earlier, it had been discovered that explosions could be effectively decoupled by setting them off in dry porous materials or in big holes; in alluvium, for instance, there appeared to be a decoupling of the nuclear blast yield by a factor of ten or so that a 20-kiloton explosion might look to a distant seismometer like a 2-kiloton event. Thus, one might say that the art of clandestine underground explosions, if it were going to practiced at all, was also making progress. It is clear from the Congressional Hearings on this problem that such progress aroused no little uneasiness.

There are many controversial views on the question of earthquake mechanisms.[6] Suffice to say that the mechanics involved are complicated. What is important to the explosion seismologist is that earthquakes constitute, in general, rich sources of complex signals. Whereas an ex-

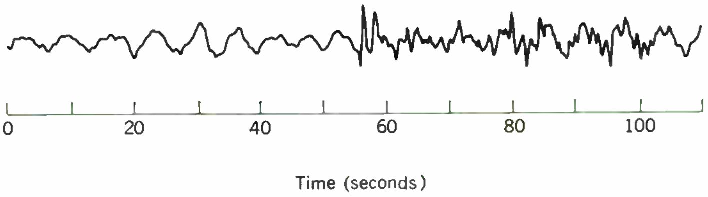
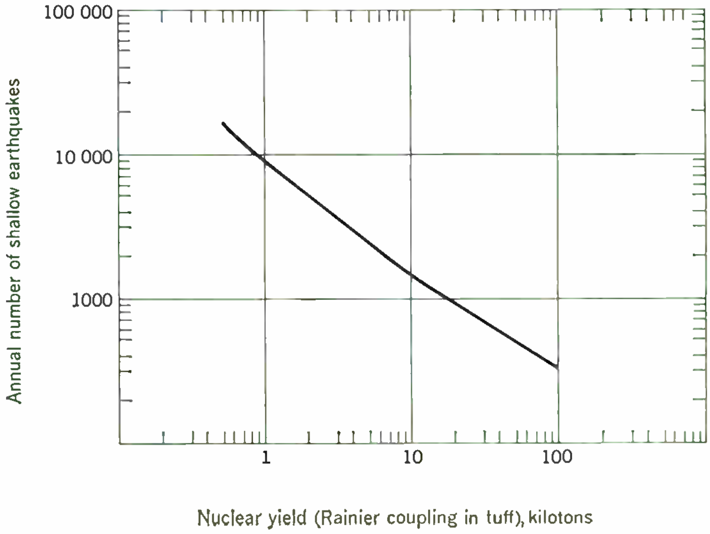


Fig. 2. Typical seismic trace from a single seismometer. The large peaked oscillations in center record an event. Prior to the event, the trace represents noise. After the onset of the event, the trace records signals that have arrived over different, longer paths and locally generated reverberation superimposed on the noise.

Fig. 3. Estimate of the annual number of shallow earthquakes occurring over the world plotted against the equivalent energy yields of underground nuclear explosions.



plosion goes off at a single point, producing in the first instance an outward compression in all directions, a large earthquake may work along a regional stress line of extreme length, and in the process produce complex rarefactional and compressional waves, shear or surface waves, and trains of aftershocks extending over relatively long periods of time.

The famous San Francisco earthquake of 1906, for instance, moved along the San Andreas fault for a length of 250 miles. However, small earthquakes (against which bombs are competing for recognition) may often look like point action, and even large earthquakes appear to start at a definite point (the focus) where the stresses first exceed the strength of the material. Moreover, the "simple" outward compressional wave produced by an explosion may be much modified by the complexities of the geological structure surrounding the explosion site and by the anomalies of the transmission path between the event and the particular receiving station. The point is that, as many people involved in this field have pointed out, and so far as present knowledge is concerned, it may be a fact of nature that there will always be some earthquakes that will look like explosions, and possibly vice versa—so that there may always be some underground events at lower magnitudes that cannot be confidently identified by *seismic means alone*.

Wave propagation. Figure 4* shows how some of the more common earthquake waves propagate (two events depicted). What should be noted is that the interior is composed of definite zones of increasing density. The crustal layers (of sediment, volcanic rock, granite, and basalt), which are the most inhomogeneous and therefore the cause of the greatest distortions in the source waves, vary in thickness anywhere from 10 to 50 miles. Generally, the velocity of the earthquake waves increases with depth, perhaps in discrete steps across the layers. Below the basalt layer is the well-known Mohorovičić discontinuity, and below it is the mantle, evidently a more homogeneous layer. With the exception of a low-velocity layer within it, earthquake-wave velocity generally increases with increasing depth until the outer core boundary is reached (at a depth of about 1800 miles), where the velocity drops abruptly, and increases only slightly thereafter. This outer metallic liquid core acts like a lens that refracts the waves striking it, creating what is called a shadow zone on the surface beyond, as shown in Fig. 4, in which a seismic station would not detect any direct signals from the event depicted. The wave paths throughout the mantle, as well as the core, are curved; this is not due to artistic rendering, but represents in a much-simplified fashion how the waves are gradually refracted upward by the velocity gradients. The situation is very much like that in the ionosphere except turned upside down, and the problems of identifying the discontinuities are also similar.

Seismic signals of many frequencies, consisting of both compressional and shear waves, are radiated in all directions from an earthquake source, along the surface, and down through the deeper layers of the earth. Although all the waves may start out at the same time, the compressional or longitudinal waves move nearly twice as fast as the shear waves (the earth moves sideways relative to the direction of propagation), so that over long distances the faster compressional wave arrives first at a seismic receiver; the greater the distance traveled, the greater will be the time difference between these two waves at the receiver. [For this reason, the compressional wave is called the P wave (for primus), and the later shear wave is called the S wave (for secundus).] The nomenclature for other common waves appears in Fig. 4 (e.g., PP and PPP are first and second reflections). At the interfaces, waves may undergo transformations, P waves giving rise to S and vice versa. Shear waves cannot penetrate the liquid core.

The interesting fact of these and more complex seismic waves is what seismologists have made of them. Over the years, not only have they constructed this picture of the earth's interior, but they have also painstakingly accumulated from earthquake data tables of wave travel times as a function of distance. As a result, given clear seismograms of the phases propagated from distant events (especially the time of onset of the first arriving phase), they have been able to calculate the distance of the event. Then, by combining the distance information from a number of stations of known location, they have been roughly able to box in the area of the event by a triangulation method.

This brief description does not mean to minimize the difficulties, it means merely to indicate the method. In actual practice, seismic records may be obscure and may have to be interpreted by a human trained seismologist; in many cases the location of the source could not be verified (i.e., the earthquake occurred in the wilds or under the ocean), and seismic stations were widely scattered and few in number, often employing no standard equipment. The marvel is that the picture of the earth's interior was assembled at all. Yet as Hodgson affirms: "The fact that all these secondary phases exist, are recorded in a routine way, and arrive at the times forecast by their travel-time curves is a final proof that the picture of the earth's interior is a sound and self-consistent one. There will be changes to the picture in the future as seismological research progresses, but they will be minor changes, refinements rather than additions."[3] Hodgson speaks as an earthquake specialist. To the explosion seismologist, the "refinements" encompass the whole range of his efforts, and represent the really big problems.

Criteria for identifying explosions

At this time, as has been mentioned, there is no unequivocal criterion whereby explosions can be discriminated from earthquakes. However, there has been assembled a number of criteria or diagnostic aids that can be applied against detected underground events to eliminate those events most likely to be earthquakes, thus reducing the number of unidentified events. These criteria relate to the determination of the geographic location of an event, its depth, its first motion, its complexity, and its aftershock activity. Their effective application depends on the obtaining of the clearest possible seismograms (high signal/noise), and preferably at many stations.

The accurate determination of the location of an event may eliminate it almost immediately from the suspicious category. For example, the epicenter may appear to be under the sea or within the boundaries of a country unlikely to be testing. If the epicenter is placed within the U.S.S.R. in an area where earthquakes only rarely occur, the event undoubtedly would warrant further examina-

---

* Figure 4 from John H. Hodgson, "Earthquakes and Earth Structure," 1964. Redrawn by permission of Prentice-Hall, Inc., Englewood Cliffs, N.J.

tion. Events in strong seismic areas, such as the Soviet Kuriles and the Kamchatka Peninsula, might very well be earthquakes, but these are also areas that could be exploited for screening explosions. Accurate location of the epicenter depends on an accurate assessment of the moment the first phase (P wave) arrives (a minimum of three stations is required to triangulate to the epicenter).

Potentially one of the most powerful of the criteria is the determination of depth. If the focus of the underground event is determined to be, say, deeper than five kilometers from the surface, it is undoubtedly an earthquake. Figure 5 suggests the principle of one method of determining the depth of an event. At long distances, the P wave arrives first at the receiver, followed shortly thereafter by the first wave to reflect from the surface somewhere immediately above the event (the pP wave). The time difference in the arrival of these two waves can be used in the calculation of the extra distance traveled by the pP. There are, of course, variations to this principle; the difficulties to its implementation are implicit in a statement made not long ago by Carpenter: "No details of the effectiveness of different methods of determining depth of focus have been published."[7]

In the classical picture, a nuclear explosion sends out compressive waves (in the first half cycle) in all directions so that all stations recording the event, whatever their direction from the source, should record the first arrival (P) as being compressive. For earthquakes, however, owing to the slipping or dipole mechanism of earthquakes, there will be some stations that record first arrivals as rarefactions (it would seem that the earth is being pulled away from the observer, not being pushed toward him). The classical picture is complicated, however, by what is called the cone of vision. Because of the way in which waves are propagated from the source, there will be a bounded zone—in the shape of a vertical cone below the source—defining all takeoff angles leading to distant stations. Thus, even a large number of stations will not get a view of the entire radiation pattern of the source—they see only that part emanating from the cone. For many earthquakes, this cone lies in the compressional part of the radiation pattern so that only compressions are seen the world over, and look like explosions.

The relative complexity of the first several tens of seconds of wave motion at a distant receiver may indicate that the source is an earthquake. As was mentioned earlier, earthquakes will be rich in shear energy, a portion of which will be transformed into compressional energy during reflections near the source into the wake of the P, giving it a longer and more complex tail (the coda).

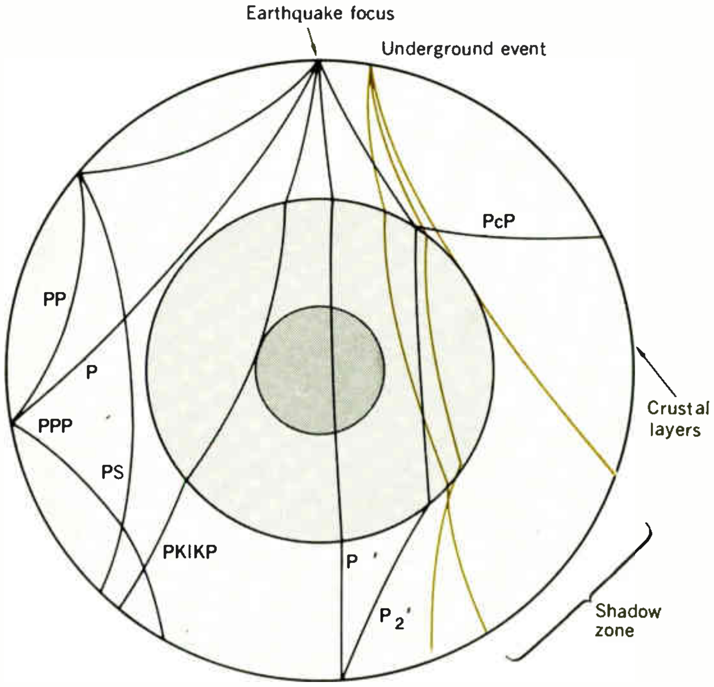
Aftershock activity is another useful determinant. Earthquakes are often followed by weaker quakes many hours and even days later, whereas with explosions aftershocks generally subside after a few hours.[5]

Other criteria are being sought. One particularly interesting one that has been discovered, and still being actively pursued, is that earthquakes and explosions generating equal amounts of compressional waves often produce different amounts of shear and surface-wave energy. Also, interesting experiments have been conducted with transforming seismic records into audio signals on the assumption that the unique pattern-recognition properties of the human ear would distinguish the sounds peculiar to explosions and earthquakes, but these studies have not proved remarkably fruitful. Learning-type pattern-recognition devices have been "trained" on earthquake and explosion seismograms, but when tested did no better than probability and considerably worse than trained humans. Nonetheless, the science of explosion seismology is still young, and the search for seismic criteria still in its early stages. Dr. Paul E. Green, Jr., head of the M.I.T. Lincoln Laboratory LASA project, remarks that "this is quite an active field; new criteria are constantly being proposed and tested, and old ones reexamined and made to work more efficiently."[8] Green feels sure, as do other experts in this field, that there are other criteria yet to be discovered.

Toward the big array

In terms of present knowledge of earthquake and explosion characteristics, the ideal arrangement for a system that could detect and discriminate between such events, and with a high degree of confidence, should consist of many seismic receivers located fairly uniformly around the globe. Thus, presumably for any event, a considerable number of stations would obtain unique records, which, through correlation, would determine the location and character of the event with a relatively high precision. One estimate, for instance, concluded that "with seven teleseismic stations well distributed around an event it should be possible to define areas of about 250 and 500 square kilometers for which the probabilities that one can find an epicenter located within these areas are 50 and 75 percent respectively."[7] With fewer stations, the accuracy would become progressively degraded (recordings must be made by a minimum of three or four stations), while with more stations, the event might be specified as having occurred in an area of 100 square kilometers. This kind of estimate is based on the premise

Fig. 4. How earthquake waves propagate. Event at right shows how outer core refracts waves, creating a "shadow" zone where direct seismic signals cannot penetrate.



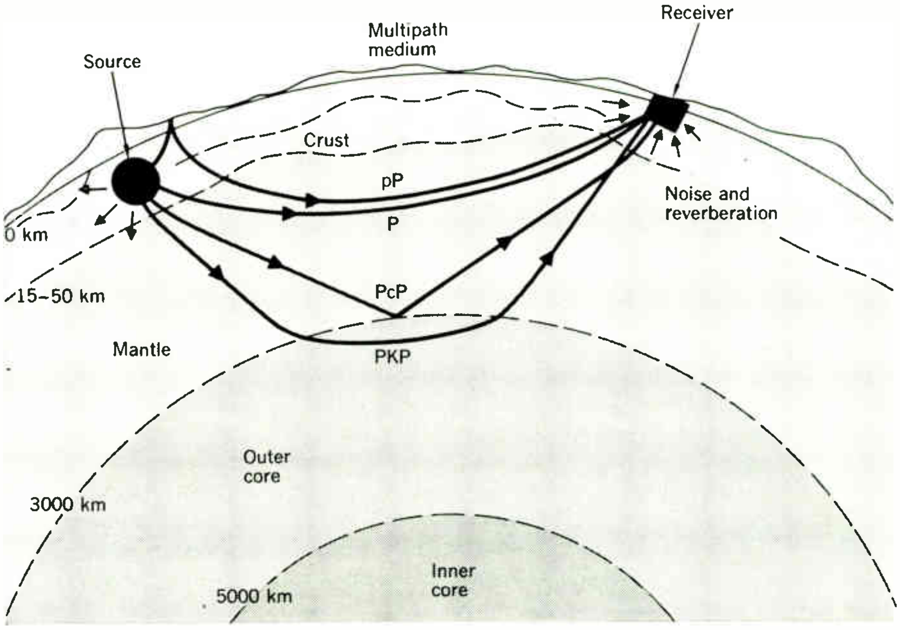


Fig. 5. Schematic of how some common earthquake waves propagate through the earth. Not shown here are surface waves, which are the most modified and diminished as they pass through the inhomogeneous crustal layers. The earth exhibits peculiar waveguide effects. For instance, some seismic stations will pick up an underground event "loud and clear," while another station in another direction from the event, although equally distant, will barely detect that an event has occurred. A potential clandestine nuclear tester can never be sure that a seismic station somewhere is not on a direct "pipeline" to his test site.

that the event be fairly strongly received relative to the noise; a high SNR is especially important in the case of the small events of interest to the explosion seismologist. With improved instrumentation, and calibrations of stations by means of explosions of known magnitudes, times, and locations, these accuracies could be (and are being) refined to still greater precision. Presumably, in such a system, once an event were accurately located, and its character determined to be sufficiently suspicious to warrant investigation, some form of verification, such as on-site inspection, would be in order.

The 1958 Conference of Experts did, in fact, propose a worldwide network of 180 stations distributed rather uniformly, including more than 20 within the borders of the Soviet Union; but the proposal was never implemented. Subsequently, there have been proposals and engineering designs for unmanned seismic stations—those notorious "black boxes"—on the territories of the parties involved, whose information would be integrated with that obtained from stations external to the state involved, but these have neither been accepted or implemented, as yet, except "in principle."

Meanwhile, in 1960, subsequent to the discussions of the Berkner Panel in 1959,[2] the United States embarked on its urgent program of research under the VELA program to determine, among other matters, how much could be achieved with national systems alone, that is, those systems that had to rely on data obtained from outside the territory of the state involved. Such national systems, which, by necessity, are located at much greater distances from the underground events of interest, in some respects imposed further difficulties on problems that were already difficult enough. Signal strength is reduced, so that signal-to-noise ratios are reduced, shear and surface waves, which can provide useful information, are much attenuated so that they go largely undetected, and so on. In effect, the amount of information propagated to the receiver has been reduced, and trickier things must be done at the receiver to compensate for such losses.

Mobile seismic stations have been developed to search for quiet sites and to provide new flexibility in the accumulation of seismic data, seismometers have been planted in deep boreholes where noise from local sources is reduced, all kinds of wave filtering have been devised; moreover, computers are opening the way to much fancier tricks in seismic data processing and filtering, are performing some seismogram reading operations formerly done entirely by humans, and are taking the hard work out of dealing with reams of seismic data and in calculating locations of seismic events. But the fanciest work of all has been done in the development of large arrays of seismometers incorporating sophisticated data-processing techniques; and among the arrays is the superarray, LASA, laced across a large chunk of Montana (see Fig. 6), installed at an urgent pace, with money drawn from ARPA emergency funds, during the winter of 1964–1965, the worst winter in 43 years.

What was all the rush about? It was to develop an instrument which was designed by M.I.T. Lincoln Laboratory engineers and scientists (jointly with ARPA and the Air Force Technical Applications Center) to provide "a substantial improvement in seismic discrimination capabilities and also to serve as a powerful new tool for advanced seismological research." At its dedication on October 12, 1965, Dr. Herbert Scoville, Jr., Assistant Director of the U.S. Arms Control and Disarmament Agency, said: "We have come a long way in the last seven years in our research program; with the development of LASA we can at last see clearly means of avoiding the placement of permanent stations within the borders of such countries as the Soviet Union. We estimate that ten to twelve stations similar to [LASA] would be satisfactory to monitor seismic events throughout the world. Furthermore, the quality of the data which would be obtained from such a system far surpasses that which was estimated to be attainable from the 1958 system." However, it was a scientific tool whose potentialities were not completely known.

The notion of using arrays of seismometers to enhance seismic signals is not a new one. Geophysical prospectors have used arrays since the 1920s, and the stations proposed by the Geneva Conference of Experts called for 10-element arrays. Between 1959 and 1963, five array stations, each containing between 10 and 31 individual seismometers, were installed in the United States by AFTAC under the ARPA VELA UNIFORM program. Relatively small though they were, in contrast with the LASA to come, these arrays demonstrated a significant improvement in signal-to-noise ratio, which is important in determining the character of weak signals from distant sources. These earlier arrays, all still functioning and supplying data, are shown in Fig. 6; they are the Wichita Mountains Seismological Observatory (WMO), which started operating in 1960, the Blue Mountains (BMO), the Uinta Basin (UBO), and Cumberland Plateau (CPO) Observatories, all operating by 1962, and the Tonto Forest Seismological Observatory (TFO), which is the largest of these early 60s arrays. Out of the experiments with these arrays (multichannel filtering, etc.), there gradually crystallized the acceptance of a few extremely large and powerful arrays replacing the network of many less sophisticated stations. Some individuals were strong proponents of developing large arrays from the beginning—Dr. Jack Ruina, now President of the Institute for Defense Analyses, and formerly the Director of ARPA, when the VELA program was started, is credited with being instrumental in persuading the M.I.T. Lincoln Laboratory to become interested in the problem in 1962. He evidently believed that the Lincoln investigators, with their pioneering background in the development and application of digital computers to problems of processing signals received through noisy and dispersive channels, could bring new ideas and insights to the seismic signal detection problem, and perhaps challenge existing concepts. Dr. Paul E. Green, who headed the small study group established at the Laboratory, says that after a few months of study of the problem they saw that the array techniques had not been sufficiently exploited, and that previous investigators had been too modest in the size of the arrays, which were roughly only one wavelength across.[5] They thought the arrays should be much bigger. They concluded that the techniques of directional (phased array) antennas and digital data processing, which had worked so well in radio communications, radar, and radio astronomy, should be applied to the seismic signal detection problem. By March 1964, Dr. R. A. Frosch of ARPA proposed that a large experimental array be built, and by October, the Lincoln Laboratory was authorized to proceed, in collaboration with ARPA and AFTAC, with its detailed design and implementation.

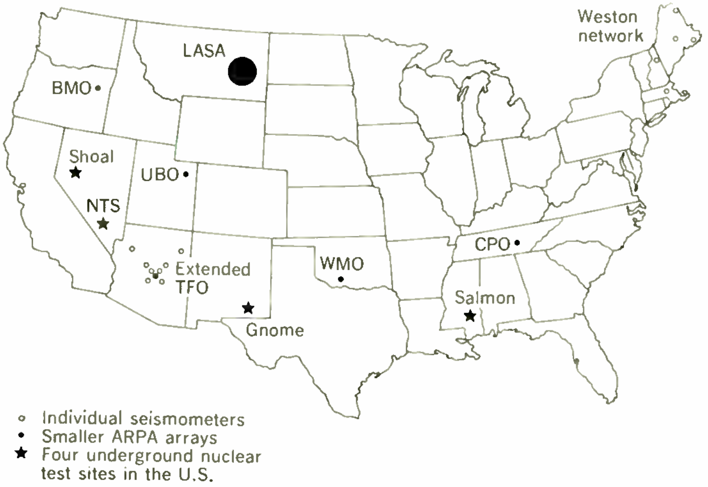


Fig. 6. Scaled drawing showing the relatively huge size of LASA, a giant array of 525 seismometers installed across the "steppes" of eastern Montana. Also shown are the sites of individual seismometers, five earlier, smaller arrays, and four underground nuclear test sites.

Large Aperture Seismic Array

Figure 6, which shows LASA (Large Aperture Seismic Array), is drawn to scale. One can immediately see that the "real estate" problems alone were not insignificant, and in fact were a major factor in the location of the site. With a diameter of 125 miles, consisting of 525 seismometers arranged in 21 subarrays, covering an area of 10 000 square miles, this array is like a giant eye (the large aperture) looking into the earth.

The general arrangement of the LASA components is shown in Fig. 7. Each of the black dots in the large array is in itself a subarray, roughly equivalent in size and capability to each of the earlier arrays mentioned—which is why LASA has come to be called an array of arrays. Each of the 525 seismometers composing the total array system is buried in a 200-foot-deep hole to reduce local noise effects. The output of each seismometer is transmitted to the Data Center, recorded on digital magnetic tape, and saved until an automatic event detection and source location computer program designates whether or not a specific time period contained an event of interest, in which case the relevant sections of tape may be submitted to further exhaustive analysis or erased.

If the reader will turn back to Fig. 5, he will note that at the receiver there is depicted both noise and reverberation. The small arrows indicate that both of these disturbances of the signal have different directional features; they also have different spectral features. The principal aim in the design of the array system and its processing techniques is to suppress this noise and reverberation and thus enhance the signals of interest. With LASA, the noise is suppressed in part by frequency filtering and in part by the array's directivity. (Directivity is in part a function of the spacing of the seismometers; the larger the array spread, the better the directivity.) Reverberation components, generated by the arrival of source signals, are generally of the same frequency spectra as the signal, and so are suppressed by array directivity. Earlier arrays were found to enhance signal over noise three to five times more than single seismometers, while LASA, each of whose 21 subarrays is equivalent to the earlier arrays, offers an *additional* factor of three to five improvement in signal-to-noise ratio.

An important aspect of LASA's search capacity is its directivity. Through the use of the on-site signal processing equipment, which includes two general-purpose digital computers and one small special-purpose computer, the array can be steered electronically to look in any direction, and thus it can be made to reject unwanted teleseisms. At the present time, it can look in five directions

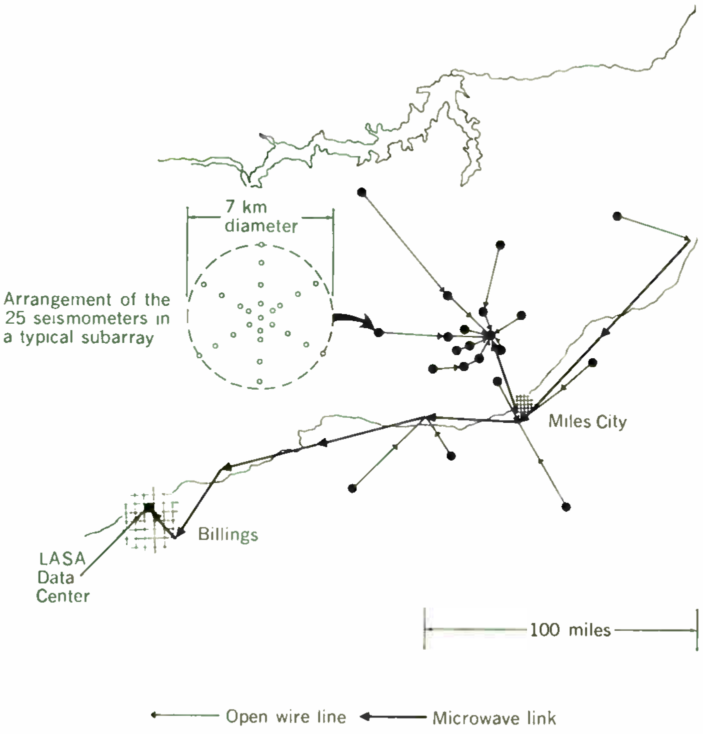
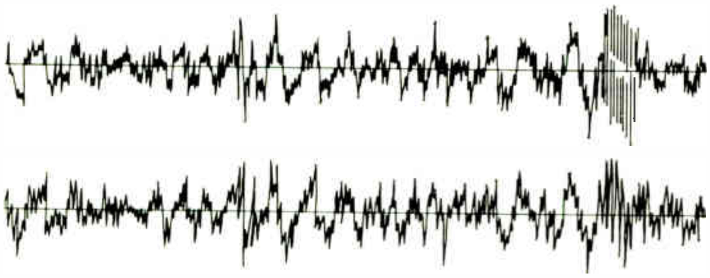
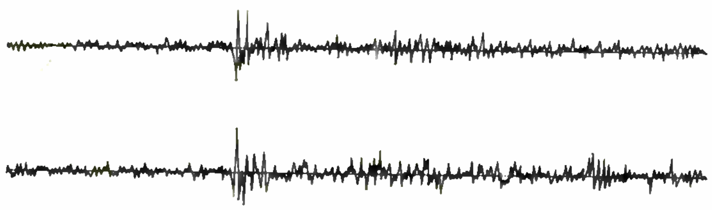


Fig. 7. Geographic arrangement of the experimental Large Aperture Seismic Array. Detail shows arrangement in a typical subarray. Each black dot represents subarray.

Fig. 8. These typical traces show how dramatically the LASA system can pull a signal out of noise and reverberation components. In the top two traces, the existence of an underground event would probably not even be suspected by a nonspecialist, but successive processing by the subarrays and the LASA Data Center produces the output at bottom. The event was a small earthquake in the Soviet Kurile Islands.



2 of 525 seismometer outputs



2 of 21 subarray outputs

LASA output

simultaneously (other studies are being made on how 150 or more beams might be processed simultaneously). Beam steering in an array is achieved not by moving physical elements but by the insertion of delays into the traces made by individual seismometers and adding the waveforms from many different seismometers. When the delays are added in such a way that the beam is looking in the direction of a signal, the signal will be maximized, while the noise, which is different in magnitude and direction at each seismometer in the array, will be suppressed. Figure 8 shows how LASA processing can pull signals out of noise.

By all accounts, it is still too early to judge whether or not the LASA system will live up to its expectations. Certainly, inasmuch as it is an experimental research tool, it holds great possibilities for the advance of seismology, and possibly also for new discoveries in the discrimination of explosions and earthquakes. Since it began operation, LASA has "seen" about 20 to 30 events a day; roughly 200 of these have been analyzed, and of these, about 10 have been underground explosions.[8,10] An important result of the system's directivity is that it can be programmed to look only in directions where events of real interest are likely to occur; furthermore, the gradual accumulation of data relating to travel times and travel-time anomalies from events in such regions certainly adds to the precision with which suspicious events can be located.

By this time, enough samples of events have been accumulated with LASA to provide an idea of its capabilities.[9,10] The detection capability of just the one LASA is estimated to extend down to approximately magnitude 3.5 at distances of the order of 3000 to 10 000 surface miles.

Not mentioned explicitly thus far is another problem that the implementation of LASA has set out to answer. How big should the biggest array be? Should there be a worldwide integrated network, a kind of super-LASA? The information-processing problem would be stupendous. Would the effort be worth it? Or would a few well-placed installations like LASA be sufficient? All those concerned with the problem appear agreed that the most effective array size (technically speaking) is still an open question. In any case—whether it's to be many small arrays or a super world array—who will pay for it?

Points on the political spectrum

An interesting question remains to be asked: What actual effect will these technical advances, including LASA-type arrays, have on the test-ban negotiations? Although technical advances make it increasingly unlikely that a potential violator can escape detection, the scientific people in the United States and the United Kingdom (where a program equivalent to VELA UNIFORM has also been under way) state fairly unequivocally that there will always be a threshold (say, magnitude 3.5 or so, which, under the right conditions, might be equivalent to a one- or two-kiloton nuclear explosion) below which there probably will be no detection capability. Below this threshold, there will occur every year thousands of earthquakes around the world amongst which there could be concealed (to be sure, at considerable cost and effort) some small underground nuclear explosions.

Such a probability, then, spawns a number of other questions. How potentially important to weapons developments could underground tests be in this lower kiloton range (viewing the entire testing scale as running from a fraction of a kiloton all the way up to 100 megatons)? Is it likely that advances could be made that would upset the current standoff between the U.S. and the U.S.S.R.? Even with a residuum of uncertainty, is not a possible clandestine treaty violation a lesser risk to the world's safety than the continuing rounds of weapons tests and the continual temptation for other nations to enter the nuclear club? In the world in which we live, in which nations exhibit such paranoid tendencies, is that not the inevitable direction?

The views on these questions might be seen as composing a political spectrum. At one end are those who are most worried about the danger of continued testing and about obtaining a comprehensive test ban as soon as possible. They argue that even if weapons tests are conducted in the low-kiloton range they would probably not bring important rewards, and that in any case the signatory nations would probably abide by a treaty and not attempt secret tests that would be costly and difficult to conceal (apart from seismic means, there are intelligence efforts, political "antennae," and so on). They argue that a treaty would discourage additional nations from undertaking costly nuclear development programs, especially if it could lead to further treaties banning assistance to nations that might wish to undertake such developments.

At the other end are those who are most worried about the technical advances that might be made clandestinely. They voice deep suspicions of the Soviet intentions, argue that weapons tests in the low magnitudes could be unsettling to the nuclear balance, and argue against the acceptance of a treaty that does not provide the hardest assurances against possible clandestine testing.

As to the relevance of the U.S. explosion seismology program to the issues involved in a test ban, Dr. Jack Ruina has this to say: "The real issues in this matter are not quantitative technical questions as much as they are matters of judgment. We must weigh the risk to our military security if the Soviet Union were indeed to carry out a series of undetected small underground nuclear explosions—which they could do if we must rely on seismic means of detection alone—against the very considerable benefits to us and the world of a comprehensive treaty. The risk depends in turn on our assessments of our nonseismic as well as our seismic capability for detection and identification, the advances in nuclear technology made possible by below-threshold testing, the improvements in weapons systems made possible by by such advances, and finally the increased military threat to the U.S. of such improved weapons systems.

"In this chain, seismic capability is the one most susceptible to quantitative analysis; nevertheless, improvements in our seismic capability that can be reasonably expected are least likely to affect the net result. Combining this with the fact that there will always be some magnitude below which testing can take place with impunity if we must rely on seismic means alone for detection, it is difficult to see why there seems to be such tremendous dependence on our seismic capability in the test ban considerations. The need for on-site inspections has been based on its deterring effect on a potential treaty violator and on its comforting and reassuring effect on us that a treaty is not being violated, and this need is not affected by the specific improvements in our seismic capability. Even the number of inspections desired is not strongly dependent on our seismic capability considering the large number of other factors involved here. This is not to say that we shouldn't have the best seismic capability possible, or that seismic improvements shouldn't be pursued with vigor, but only that the tight coupling of our technical capabilities in seismic detection and identification to the provisions of a test ban treaty is not warranted."[11]

What is the recent thinking of our scientific people on the improvements in weapons systems that might be possible through clandestine testing? One man long associated with disarmament efforts, the former Special Assistant for Science and Technology to President Kennedy, Dr. Jerome B. Wiesner, holds the position that no significant development work could be done with any underground testing now. That is, Wiesner says, nothing could be done that would seriously upset the weapons balance that now exists.[12] He adds that there is development that can be done in the lower yield range, but that this is work that can be done in other ways, without testing. (Dr. Wiesner's views on two treaty proposals that have been advanced very recently are reported elsewhere in this issue, page 156.)

The actual treaty negotiations are genuinely complex, and involve many interlocking issues, proposals, and changes of position.[13] Perhaps the new moves being reported (see page 156) will at last lead to a break in the long-standing East–West treaty impasse. Time will tell.

The writer expresses his appreciation to the following persons who generously gave their time and assistance during the preparation of this article: P. E. Green, Jr., J. Kessler, M.I.T. Lincoln Laboratory; S. J. Lukasik, H. Test, ARPA; J. Ruina, IDA; B. L. Basore, U.S. ACDA; C. F. Romney, J. Pantall, F. Pilotte, AFTAC; G. H. Sutton, D. D. Prentiss, Lamont Geological Observatory, Columbia Univ.; and J. B. Wiesner, M.I.T. The writer owes special thanks to Drs. Green, Ruina, Basore, and Lukasik, who read the manuscript and offered much helpful criticism.

REFERENCES

1. "U.S.-Soviet clash mars arms talk," *The New York Times*, June 15, 1966.

2. Bates, C. C., "Foreword to Special Issue I on VELA UNIFORM," *Geophysics*, vol. 29, p. 151, Apr. 1964.

3. Hodgson, J. H., *Earthquakes and Earth Structure*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

4. Richter, C. F., *Elementary Seismology*. San Francisco: W. H. Freeman, 1958.

5. "Developments in Technical Capabilities for Detecting and Identifying Nuclear Weapons Tests," Transcript of Hearings before the Joint Committee on Atomic Energy, Mar. 5–12, 1963, p. 108.

6. Carpenter, E. W., "Teleseismic methods for the detection, identification, and location of underground explosions," VESIAC State-of-the-Art Rept., 4410-67-X, Institute of Science and Technology, The University of Michigan, Apr. 1964.

7. Carpenter, E. W., "Explosion seismology," *Science*, vol. 147, pp. 363–373, Jan. 22, 1965.

8. Green, P. E., Jr., M.I.T. Lincoln Lab., Private communication.

9. Green, P. E., Jr., *et al.*, "Seismic discrimination," Semiannual Technical Summary, 1 July–31 Dec. 1965, Lincoln Lab., issued Mar. 15, 1966.

10. Briscoe, H. W., *et al.*, "Interim report on capabilities of the experimental large aperture seismic array," Tech. Note 1966–16, Lincoln Lab., Feb. 24, 1966.

11. Ruina, J. Institute for Defense Analyses, Washington, D.C., Private communication.

12. Wiesner, J. B., M.I.T., Private communication.

13. "Review of International Negotiations on the Cessation of Nuclear Weapon Tests," Sept. 1962–Sept. 1965, U.S. Arms Control and Disarmament Agency, Pub 32, released May 1966.

The peculiarities of high-voltage dc power transmission

A modern high-voltage dc link interconnects two systems that both carry alternating current. It is this combination of two different kinds of current that presents interesting problems and solutions

A. Uno Lamm ASEA

The cost of a basic dc line is always less than that of an ac line. For submarine installations the difference is particularly pronounced; for underground cables, the difference is somewhat smaller. For overhead transmission on land, many studies show that a dc line generally costs about one third less than an ac line of the same capacity. It may be entirely justifiable, however, to compare a two-pole dc line with a double-circuit ac line that has two 3-phase circuits carried on the same towers. The dc line cost, on this basis, would be about 55 percent of the cost of the ac line.

Direct-current transmission, in itself, should be the simplest of all transportation techniques for electric energy. Essentially, it is merely a straightforward application of Ohm's law. What lends peculiarity to a modern high-voltage dc link, however, is the fact that it does not connect a dc source to a dc load, but, instead, it interconnects two systems that both carry alternating current. It is this combination of two different kinds of current that presents the interesting problems. And these problems demand practical solutions to justify the practical application of high-voltage dc power transmission.

High-voltage convertors

Some of the properties of dc transmission are typical of all static power convertors that use grid control for varying the ratio between ac and dc voltages, and for operation as inverters. In principle, these convertors are switches that commutate the current in the three ac phases to collect it into a unidirectional current in the two dc poles (Fig. 1). In so doing, the convertor eliminates one third of the ac cycle from each phase-to-phase voltage and transposes it onto the dc terminals, and six such voltage sections follow immediately upon each other on the dc side during one ac cycle.

Without grid control, the crest part of the ac half wave is cut out and transposed. The dc voltage in such a rectifier operation obtains its maximum, which, in zero-load condition, is only 5 percent lower than the peak of the phase-to-phase ac voltage wave. By delaying the commutation through grid control, the transposed one-third part of the ac voltage wave will be displaced to comprise a section with a lower mean value; thus, the dc voltage will be lower. At 90 (electrical phase) degrees delay, the dc voltage will be zero, and, at greater delays, the dc voltage polarity will reverse. Since the current must retain the same direction, the power flow also reverses and the convertor functions as an inverter.

This is the simple principle of inverter operation—conversion of direct current into alternating current. The physical equipment is identical to that which is used for grid-controlled rectification. Rectifier–inverters of this kind have been used for the past 30 years for low-voltage industrial drives. The earliest types used mercury-arc tubes, and, more recently, controlled semiconductor cells, or "thyristors," have been employed.

The mercury-arc convertors were rarely connected in double-way arrangement as shown in Fig. 1. For reasons of efficiency, they were generally used in single-way connection in which the direct current had to pass through but one tube in series and, therefore, was exposed only to losses that corresponded to one arc–voltage drop.

The high-voltage convertors used for transmission purposes, which utilize the tubes or valves to their full-voltage capacity, however, take complete advantage of the double-way connection since it provides double dc voltage for the same ac voltage. A further important advantage of the double-way connection is the simplicity and good utilization of the convertor transformer, which, in principle, is a straightforward three-phase transformer of about the same size as one with the same insulation levels used directly for ac transmission.

The high voltage to which the valves are exposed during the periods when the anode is negative with respect to the cathode, and when, through the action of a negative grid, the anode carries high positive voltage and zero current, are peculiarities of convertors used for high-voltage dc transmission. Thus the development of such high-voltage mercury-arc valves has been the key to the practical application of the dc transmission system.

High-voltage valves

The early efforts to push the development of mercury-arc valves toward higher voltage with simultaneous current rating, sufficient for transmission purposes, always hit a voltage ceiling that was too low. Subsequent efforts, therefore, were directed to mechanical valves in which contacts were moved by synchronous motors. The timing of such contacts had to be very accurate since the opening had to take place while the current passed through zero value.

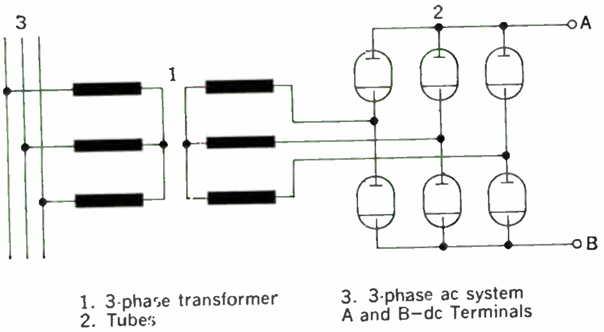
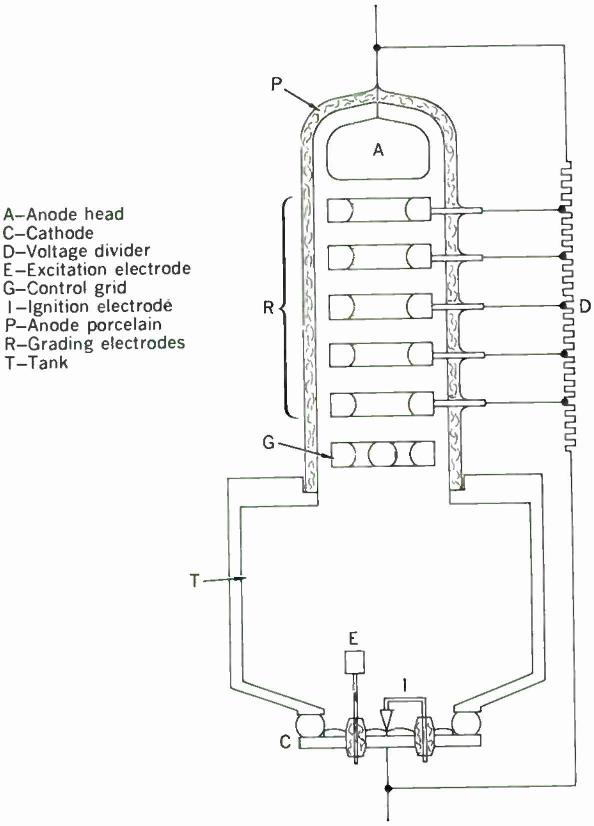


Fig. 1. Block diagram of a convertor showing a six-pulse two-way connection.

Fig. 2. Diagram showing the principle of grading electrodes in a high-voltage valve.



Even if creditable results were gained in steady-state operation of such mechanical convertors, it was obvious that the synchronous mechanical drive would never respond quickly enough to the transient conditions that would take place during disturbances and rapid load changes in the ac system.

Another device, the airblast valve, which, like the mercury-arc type, was fundamentally a "genuine" valve (opening and closing the circuit when necessary), was tried by Prof. Erwin Marx in Germany 20 to 30 years ago. The arc was artificially ignited by a high-pressure high-velocity air stream passed between the electrodes. The air blast produced the necessary deionization after the current dropped to zero. Although impressive results were achieved in large-scale tests, the methods seem to have had the main disadvantages of premature current cutoff with consequent overvoltage. Also, the huge consumption of compressed air represented excessive losses in the blower drive.

In Sweden, our early efforts were directed toward finding improvements in mercury-arc valves that would permit a rise in the apparent low voltage ceiling of conventional mercury-arc devices, especially when combined with currents of several hundred amperes.

The physical picture of the mercury-arc valve was—and to a large extent remains—not as clear as that of the hard electronic valves. Yet, it was obvious that the reason for the limited voltage capacity was a positive ion space charge that formed adjacent to the "anode" surface during the inverse period. Practically all of the anode-to-cathode voltage was concentrated in this thin space-charge sheath. The major part of the arc path had a very low field strength, and thus no improvement in voltage capability could be obtained by increasing the length of the path.

The means for eliminating the concentration of the voltage in the immediate vicinity of the "anode" surface was found by using a succession of intermediate or "grading" electrodes interposed in the path. These electrodes (Fig. 2) were connected to a voltage divider between anode and cathode, from which grading, intermediate voltages were impressed on the electrodes. Simple experiments in the early 1930s showed promising results. In the following decade, we had the opportunity of including in the experimental valves a technique, novel for its time, which ensured vacuum tightness and purity of materials, consistent with the strict requirements of any mercury-arc device.

The experimentation, started in the laboratory of ASEA's Ludvika works, could be continued on a larger scale in special test facilities made available by the Swedish State Power Board in one of their major hydro power plants, since the empirical nature of the valve development requires continuous full-scale testing. The trial-and-error procedure, which entailed the testing of more than 100 modifications of the interior valve design, was accompanied by physical research that gradually improved our knowledge of the phenomena involved in the operation of the valves. Such a physical concept, of course, is of great value as a guide toward improved experimental designs, choice of materials, and processing methods.

In 1950, the development of the valve itself, and its combination with other components, to complete high-voltage convertor stations, had advanced to a point where decision could be taken on a first practical installation. A submarine transmission cable from the Swedish mainland to Gotland Island offered a unique case of a moderate-sized commercial installation. See Table I for the main arrangement and data for this and succeeding high-voltage dc plants.

The writer would digress too far in describing herein all the various requirements of high-voltage ionic valves for dc transmission. It should only be mentioned that, like in all other ionic devices, arc-backs (failure of valve action) are not entirely avoidable. Arc-backs, however, do not produce any lasting effects on the valve, and their consequences on the transmission operation are eliminated by suitable measures that are described later in this article.

The losses in high-voltage valves are defined essentially by the arc drop of 40–50 volts, to which are added a smaller loss portion from auxiliaries such as the voltage divider for the grading electrodes. Thus the relative valve losses in a 150-kV convertor are of the order of 0.1 percent of the load.

The complete convertor

The transformer of a high-voltage convertor, which in principle has but one primary and one secondary winding per phase, poses a few special problems. The high-voltage winding of most ordinary EHV transformers carry near-ground potential at one end, thereby making it possible to use a graded insulation that builds up voltage from the yokes toward the center of the leg. A high-voltage dc convertor-transformer will also have a valve side winding, both ends of which carry high and varying potential to ground. This circumstance, and the special character of the dc voltage stress on the insulation, has required some important innovation work on high-voltage dc transformers.

Another feature, characteristic of the high-voltage con-

I. Data on commercial high-voltage dc schemes based on the ASEA technique

	Gotland	English Channel	Sardinia	New Zealand	Japan 50–60 Hz	Konti-Skan
Commissioning year	1954	1961	1965	1965	1965	1965
Power transmitted, MW	20	160	200	600	300	250
Direct voltage, kV	100	±100	200	±250	2×125	250
Valve groups per station	2	2	2	4	2+2	2
Direct voltage per valve group, kV	50	100	100	125	125	125
Direct current, amperes	200	800	1000	1200	1200	1000
Parallel anodes per valve	2	4	4	4	4	4
RMS value of current per valve, amperes	115	460	575	690	690	575
Reactive power supply	Synchronous condensers Capacitors	Capacitors	Synchronous condensers Capacitors	Synchronous condensers Capacitors	Capacitors AC grids	Synchronous condensers Capacitors
Convertor station location	Västervik Visby	Lydd, England Echinghen, France	Codrongianos San Dalmazio	Benmore Haywards	Sakuma	Göteborg Ålborg
AC grid voltage	Västervik 130 kV Visby 30 kV	Lydd 275 kV Echinghen 225 kV	Codr. 230 kV S. Dalmazio 220 kV	Benmore 16 kV Haywards 110 kV	275 kV, 50 Hz 275 kV, 60 Hz	Göteborg 130 kV Ålborg 150 kV
Length of overhead dc line	—	—	290 km (180 miles)	575 km (354 miles)	—	86 km (53 miles)
Cable arrangement	1 cable, earth return	1 cable per pole	2 parallel cables earth return	1 cable per pole	—	1 cable earth return
Length of cable	96 km (60 miles)	64 km (40 miles)	116 km (72 miles)	42 km (25 miles)	—	87 km (54 miles)
Earthing of the dc circuit	For full current in two sea electrode stations	Mid-point earthed in one station	For full current in two sea electrode stations	For full current in one earth and one sea electrode station	One point earthed direct	For full current in two sea electrode stations
Control	Constant frequency on Gotland	Constant power in either direction	Constant power or constant frequency on Sardinia, or a mixture of both	Constant power from Benmore to Haywards	Constant power in either direction	Constant power in either direction
Reversal of power flow	Effected manually	Controlled by power-setting device	Controlled by power-setting device and frequency regulator equipment, respectively	Possible, not normal	Controlled by power-setting device	Controlled by power-setting device
Emergency change of power flow	—	On manual or automatic order to preset value	—	—	On manual or automatic order to preset value	On manual or automatic order to preset value
Reason for choosing system	Long sea crossing, frequency control	Sea crossing, asynchronous link	Long sea crossing earth return	Long distance, including sea crossing	Rapid control, low losses, asynchronous link	Sea crossing, building in stages
Power company	Statens Vattenfalls-verk, Vällingby, Sweden	Central Electricity Generating Board, Guildford, England Électricité de France, Paris, France	Ente Nazionale per l'Energia Elettrica, Rome, Italy	New Zealand Electricity Dept., Wellington, New Zealand	Electric Power Development Co., Tokyo, Japan	Statens Vattenfalls-verk, Vällingby, Sweden Elsam, Pr. Fredericia, Denmark
Main supplier of convertor equipment	ASEA	ASEA	English Electric. ASEA subcontractor for valves and control equipment	ASEA	ASEA	ASEA

vertor, is the bigger role played by capacitance in transformers, bushings, and conductors. The steep voltage changes, occurring during commutation of current from one valve to the next, produce, in the *L-C* circuit formed, oscillations of frequencies from the audio to the megacycle range. When the full voltage capacity of the valves is to be utilized, these oscillations have to be damped by special *R-C* or *R-L* circuits (Fig. 3). The *R-C* circuits, in particular, which damp the oscillations produced by an outgoing valve, tend to produce losses that are an appreciable percentage of the total losses of the convertor. Therefore, an objective for further development is to attain improvements in the valves that will permit the omission of these damping circuits. The oscillations just mentioned do not propagate outside the convertor group.

Another specific characteristic of high-voltage convertors is that some components, especially the valves, have high and varying potential to ground. Hence, the power supply to the valve auxiliaries requires high-voltage insulation from ground. Although other methods have been suggested, such as generators driven by a long insulating shaft, or a belt from a motor on ground potential, or by hydraulic transmission using an insulating liquid, the simplest way to supply this power is to feed the auxiliaries over insulating transformers. The grid-control impulses also have to be fed through insulating devices or techniques such as transformers, light or radio links, or by inductive or capacitive transmission.

Combination of convertors into high-voltage dc terminals

A striking difference between a high-voltage dc installation and conventional power plants and substations is that the several convertor units, required to deal with the total capacity involved, are connected in series rather than in parallel. There are three independent reasons for this:

1. The protection against arc-back and commutation failures is more effectively arranged with series connections (Fig. 4). By having a bypass valve across the dc terminals of each convertor, an individual convertor, affected by such a fault, can be transiently eliminated from the circuit, thereby enabling the other convertors to continue their unaffected operation. And no current can feed into the fault from parallel units.
2. High dc line voltage, so important for the economy of the transmission line itself, can be achieved with a convertor voltage that is only a fraction of the line voltage.
3. The insulation to ground of the series-connected convertors can be diminished in steps, so that only the convertor next to the line has to be insulated for the highest dc voltage.

For each pair of convertors, a delta connection is used for one and Y connection for the other to give 12-pulse characteristics to the group and to the station. In this way, harmonics on the ac side and ripple on the dc side are reduced.

In industrial low-voltage rectifier plants, the several rectifiers are sometimes combined to give a higher pulse number than 12. Because of the unavoidable asymmetries, however, the useful advantage of the combination is relatively small. For high-voltage dc terminals, the small reduction in harmonics is not worth the considerable complication that would be required of the transformers.

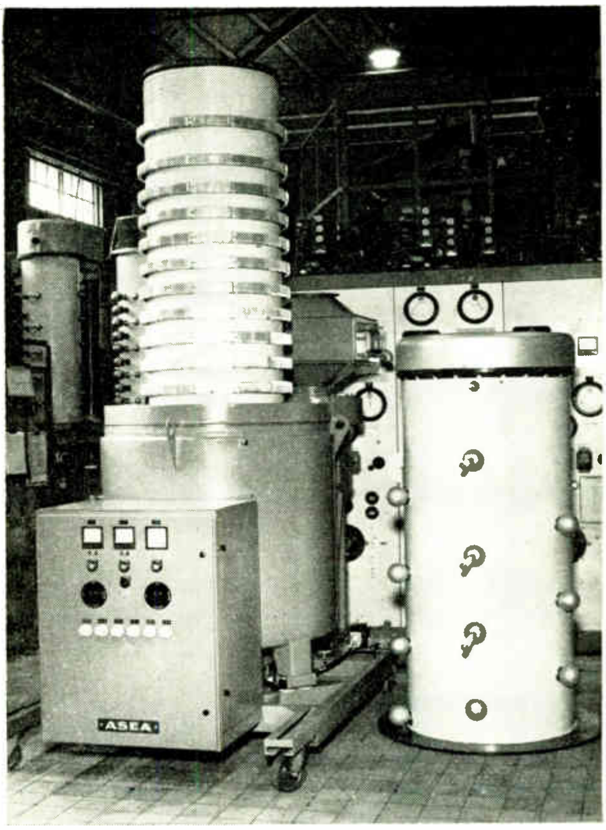
Because of the series connection of the different units, the method of cutting one unit in or out of operation differs from the conventional technique. A bypass isolating link must be connected across the dc terminals of each convertor, and other isolating links are interposed in both terminal conductors. A single convertor can be shut down by the grid-blocking of its ordinary valves and the simultaneous unblocking of its bypass valve. These operations are activated by a regular master controller. The two dc terminals of this convertor are thus short-circuited through the bypass valve, thereby enabling the current from the other convertors to pass freely. Subsequently, the bypass isolator is closed and takes over the current from the bypass valve. The isolators in the two poles of the convertor can now be opened. The ac-side connections can also be opened, either on the line side or the valve side of the transformer. The complete valve assembly may be grounded and made available for servicing.

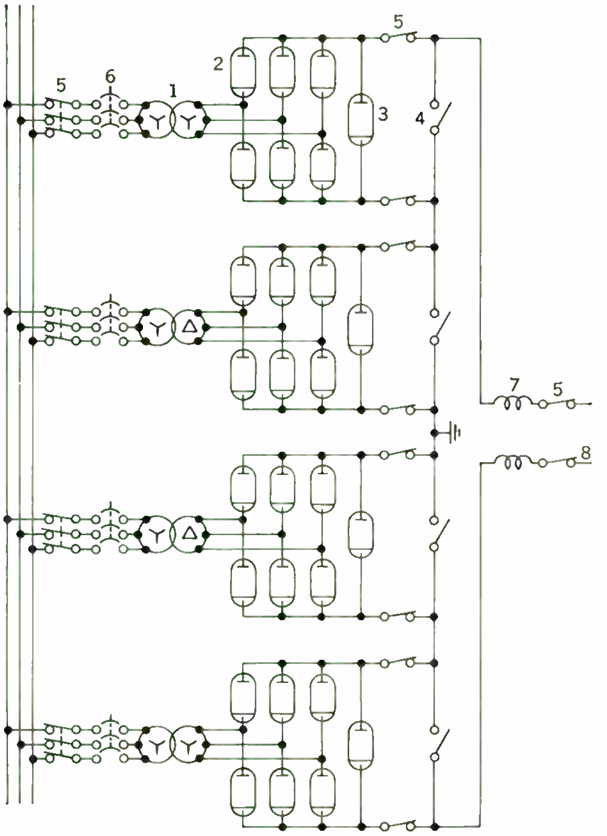
The start-up of a single convertor unit onto the energized line occurs in the opposite sequence, and the final step is the grid-blocking of the bypass valve and the unblocking of the six working valves.

Ground as a dc conductor

The ground is an excellent return conductor for a dc line. No inductive effect forces the ground current to follow the line route near the surface as is the case with alternating current. By contrast, the dc ground current,

Fig. 3. View of experimental mercury-arc anode for 200-kV convertors. The anode assembly is shown mounted on a single-anode vessel for testing purposes.





1. Transformer
2. Operating valve
3. Bypass valve
4. Bypass switch
5. Isolator
6. Circuit breaker
7. Smoothing reactor
8. Outgoing dc line

Fig. 4. Block diagram of convertor station, with four convertors connected in series on their dc side.

following the path of least resistance, penetrates into the good-conducting interior of the earth, leaving noticeable effects on the ground surface only in the vicinity of the earth electrodes.

The fundamental, straightforward dc line thus has only one insulated conductor with plus or minus polarity while the other pole of each terminal is connected to an electrode in the ground. Between line and ground, there are, in each terminal, a number of series-connected convertors just described. Several high-voltage dc installations, particularly those with submarine cable lines, have been built in this way. Other installations, especially those with overhead high-voltage dc lines for carrying very high power, have two insulated conductors, each of which is connected to its own series of convertors. The other end of the series is connected to the ground. If opposite polarities are chosen for the line conductors, the ground will carry zero current as long as the currents of the two poles are equal.

It should be noted, however, that each of the two poles can be generally regarded as an independent transmission in which the ground is arranged to take care of the current difference. To safeguard optimum reliability, the terminal equipment—including to a large extent the auxiliaries, controls, and protection gears—are made separately and individually for each pole (and generally for each convertor unit). But the remarks in the following section can be interpreted as referring either to a single pole or to a combination of two poles, depending upon the preferred arrangement in each case.

Operational characteristics and power control

The application of a new technique essentially requires some new thinking. This is particularly true where power control is concerned in the case of high-voltage direct current for power transmission.

We shall try to outline the main differences between the ac and dc methods, since the dc method is applied today in several installations. It should be noted that this mode of application and this arrangement of control is not the only possible arrangement with dc transmission, but it is the one combination that has been chosen so far in a careful consideration of the special new possibilities that are offered.

In many early articles devoted to the voltage and current characteristics of high-voltage dc transmission, the authors studied the "natural" characteristics of inverters and rectifiers combined with each other over a dc line (which, in itself, merely obeyed Ohm's law). The determined characteristics, however, did not seem to offer a practical solution, since the power flowing over the line would vary in an arbitrary way with the ac voltages of the two interconnected sytems. And the power flow, even if it were stable, would not correspond to any practical load-dispatch requirements.

This was the reason why our early choice was directed toward a transmission that "floated" on the automatic grid control of the convertors. The inverter, whose operation depends upon the termination of each commutation before a certain point on the three-phase voltage wave, has its grid control so arranged that the inverter has a margin of commutation sufficiently large for safe operation, but still not large enough to produce an unnecessary waste of reactive power. This arrangement is achieved by the *consecutive grid-control method*. The inverter so controlled works with an essentially constant ac-to-dc voltage ratio, and sets up an essentially constant EMF to the dc line.

The rectifier station is made to carry out the actual power control—and by grid control, its ratio of dc voltage to ac voltage can be varied at will. By such action, the dc output voltage is automatically adjusted so that the current pushed through the line against the counter EMF of the inverter station attains the desired value. This value, in turn, is prescribed by a small input electrical quantity, fed into the automatic regulator, which can be varied according to any desirable program.

Alternating-current transmission characteristics

In an ac line, it is the angular difference between the voltages of the two interconnected systems rather than the magnitude of the voltages that determines the amount of power transmitted. When several generators and load networks are incorporated in the system, the power fed into it from each generator is controlled only by the generator's turbine governor. At least one of the generating stations must be controlled in this manner in order to maintain constant frequency of the total interconnected system; and, in doing so, the station also delivers the balance between total generation and total load.

When two or more circuits are paralleled, the division of power between them is determined by their relative impedances, and, generally speaking, no operational measures can be taken in the lines themselves or in their terminals to influence this sharing of the active load.

The voltage of the system is controlled by generator excitation and tap changers on the transformers, and it has little influence on the load division. Should a line, generator, or load group trip out, the load sharing will readjust itself without the necessity of immediate intervention by an automatic regulator, although turbine and excitation regulators will go into action to restore correct frequency and voltage as required. Thus an interconnected ac system offers a very "comfortable" method of control as long as the interconnecting links are "rigid" enough to safeguard synchronism between all machines at all times.

In some cases, however, there is the problem of maintaining synchronism or stability. Such a problem can occur when a long line, having a high series impedance, forms a link between two separate systems so that the synchronizing force is weak. Also, if one particular line or its terminal equipment becomes overloaded, no corrective measures can be taken with respect to the lines or terminals themselves, and the line either will be tripped out automatically, or action must be taken with regard to the load or generation to eliminate the overloading condition.

Direct-current transmission characteristics

Direct-current transmission provides no immediate synchronizing force between the two ac systems at its ends, but strives, as previously described, to maintain the current set on its regulator. When the dc line forms the only connection between two separate ac systems, the power import and export at each time can be preset on a handwheel to a desired value that will be maintained regardless of what happens in the two systems.

This value also can be reset automatically and quickly by external orders if a sudden change of power is desired. The reason may be a loss of generation, noticeable as a drop of frequency in the affected system. This change of power flow can be made to include an eventual power flow *reversal*, which also can be actuated by simple grid-control action on the convertors. Should such a change of power flow, required by one of the two ac systems, have to be limited because of considerations of the other ac system, a limitation can be easily arranged and incorporated into the program regulator as an overriding control. Further, a dc transmission will not permit itself to be overloaded, since it is provided with an overriding control—rather than protective relays for line tripout—which limits the current to the capacity of the lines and the terminals.

When other ac connections exist between the systems at each end of the dc line, the aim of the dc control often may be to help the ac connection to maintain synchronism. For this function, the dc line can act in a more positive way than an ac line of the same capacity. The regulator, acting on the grid controls, can achieve a larger change of power flow and can act quickly before the phase lag approaches a critical value. Parenthetically, it is obvious that the flexibility of the convertor control would make it possible for a dc transmission to be controlled independent of the phase lag so that it would act in the same way as an ac line. It should be clear, however, from the above, that an even better stabilizing effect can be achieved more easily in other ways.

The principle of overriding controls already has been used in the first modern commercial installation, the 20-MW submarine high-voltage dc link between the Swedish mainland and Gotland Island. Most of the time, the dc transmission is the only power supply to the island. Thus the power flow must be controlled to maintain constant frequency of the island system. At times, however, the inverter station may work in parallel with a steam plant on the island which may—or may not—take over the frequency control. The control desk is therefore provided with two handwheels, one for setting the power to be transmitted, and the other for setting the frequency. The combined controls act in such a way that, if the frequency tends to drop below the set value, the power is increased automatically as required, but only up to the limit of the power setting. If power control is desired, the frequency handwheel is reset to a value a little above the frequency setting of the steam plant, while the power handwheel is set on the desired power level. In its efforts to reach the set frequency, the regulator is limited by the power control and maintains the prescribed power value. If a large part of the load is lost, however, the frequency control will intervene and prevent the frequency from rising above the preset level, and, in doing so, it will cause the transmitted power to drop.

Thus the transfer from power control to frequency control is made, without any switching operation, merely by resetting the two handwheels so that the frequency control makes contact with the actual frequency, while the power setting is made higher than the actual power transfer. In the Gotland Island case, the dc transmission supplies all power not provided by the steam plant, and it is constantly prepared to increase its power up to the preset value should any portion of the generation be lost.

This, of course, is only one example of the possible arrangements of the power control of dc transmission. It is possible to provide a program incorporating several more overriding controls, and also to have each control modified to compromise between several simultaneously active control inputs.

Interaction between rectifier and inverter station controls

As previously mentioned, the inverter station is normally controlled to meet its own requirements and to safeguard the inverter commutation, while the actual power control is performed on the grid control of the rectifier station. The power control is built primarily as a control of the direct current delivered from the rectifier station (Fig. 5). The dc voltage of a rectifier, however, always has an upper limit that corresponds to zero-phase retardation in its grid control. If, at some instant, this maximum voltage of the rectifier station should be insufficient in relation to the counter EMF of the inverter station, the current and power would drop to zero. Because of this, the inverter station is also provided with current control.

The setting of the inverter current control is always made somewhat smaller than that of the rectifier station

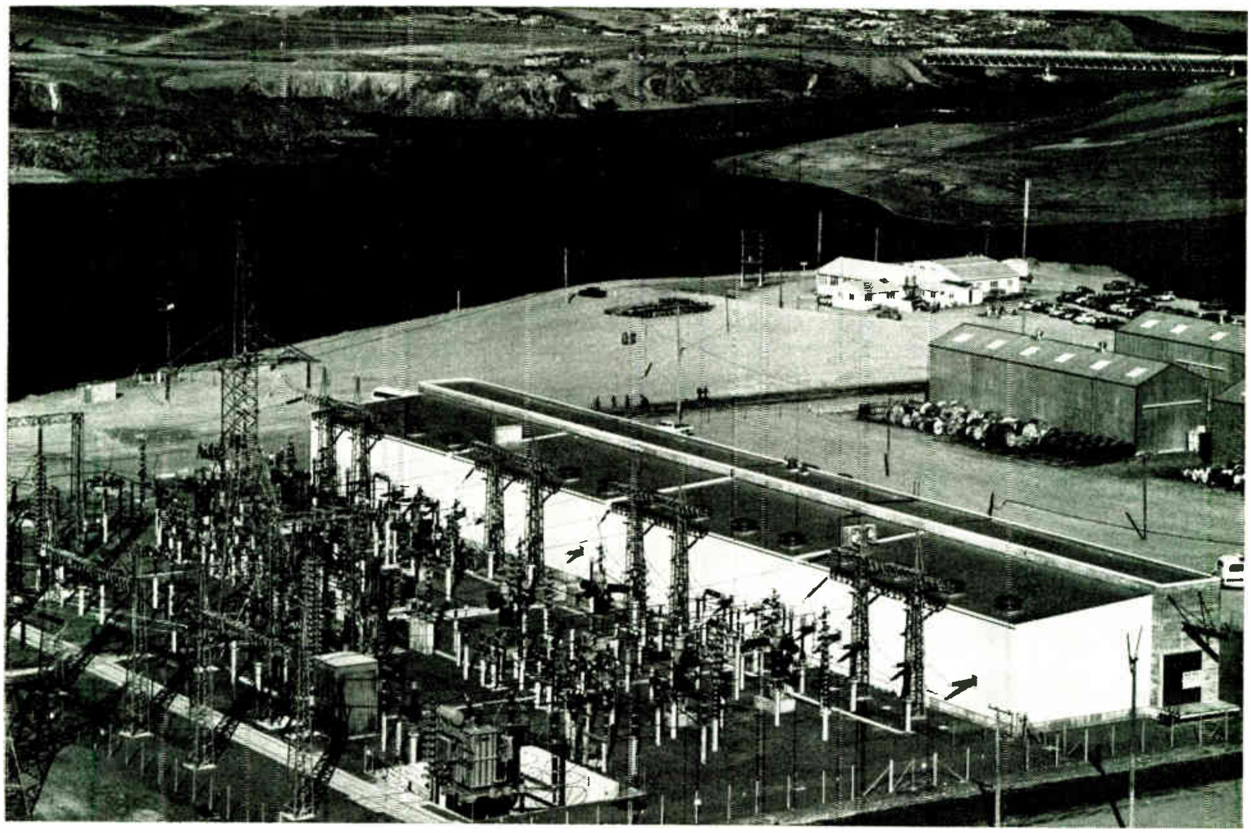


Fig. 5. View of outdoor switchyard and building containing valve and rectifier equipment at the high-voltage dc hydro power plant, Benmore, New Zealand.

control, the difference being called the *current margin*. Normally, the rectifier current control will govern when the inverter control, in its effort to reduce the current margin, will tend to raise its voltage until checked by the limiting consecutive grid control.

If the voltage supplied by the rectifier station drops transiently, the current control of the inverter station will intervene and lower the inverter voltage to such an extent that the preset current will be removed from the line. Consequently, the power will drop only by an amount that corresponds to the current margin and to the initial dc voltage drop. As will be seen, tap changers on the convertor transformers will eliminate any lasting deficiency in output voltage from the rectifier station, and such interference by the inverter station, therefore, will be only temporary.

Essentially, what we have been discussing here is the *primary function* of the control—the maintenance of the direct current at an approximately constant value, irrespective of outside influence on the voltages. But, as previously stated, the automatic regulator, for instance, can be readily supplemented to keep the *power amount* constant by having the current setting determined by a power regulator. Should the voltage fall, either by influences from the ac side or by the loss of one convertor in the series, the power regulator will increase the current sufficiently to maintain an unchanged power flow. (Naturally, this will only be accomplished within the limits of the load capability of the equipment, which would always be an overriding control quantity.)

The coordination of the current settings in the rectifier and inverter stations is generally achieved by remote control. This is particularly essential when the current setting (current order and reference current) is, in turn, the output quantity of a somewhat sophisticated program regulator.

The behavior of a dc link under faults

The principal enemy of any power transmission system is lightning. Although lightning arresters are available today which prevent permanent damage to the terminal equipment, and overhead ground wires take care of the majority of lightning strokes in the vicinity of the line without affecting the active conductors, there is still a certain statistical occurrence of strikes that lead to flashover.

In an ac line, the follow current in the flashover arc amounts to several times the normal current, depending upon the short-circuit capacity of the system. The follow current has to be eliminated by circuit breakers in both ends of the line. The breakers can be reclosed after a time interval that allows for the deionization of the arc path.

A dc line is similarly protected; however, because of the automatic current control applied to the rectifier and inverter stations, the current of a flashover arc will be only a fraction of the normal current, that is, equal to the current margin. Therefore, the risk of having insulators damaged by the arc is greatly reduced and the deionization time is short. As a result of the smaller current and more exact control, the outage can be considerably shorter than in the ac case. No breakers need to be opened, but the line is momentarily de-energized by resetting the grid control of the rectifier station into inverter operation. From the short period of practical experience gained thus far on this kind of ground-to-fault protection for a dc overhead line, there is some indication that the dc arc may be self-extinguishing—thanks to the low-current level.

As in the ac installation, the dc terminals must be protected against interior faults in the station equipment. In this respect, it is characteristic of the convertors that their valves are exposed to occasional, unavoidable interior short circuits. These transient faults, and arc-backs in the valves, form a short circuit on the ac side which has to be cleared by the grid blocking of all six valves of the affected convertors. Blocking occurs so quickly that the short circuit will last for only one-half cycle, and, after 0.5–0.75 second, normal operation is resumed. During the blocking period, the current is transferred to the bypass valve so that the other convertors of the pole can operate unaffected. The voltage, however, will dip in correspondence to the transient loss of one convertor of the pole, but the current, as described earlier, will be maintained at approximately the same value.

The design of the valves is so coordinated with the parameters of the circuits that they can stand the arc-back current without damage or any other lasting effect.

Commutation failure in an inverter represents a transient short circuit of the inverter's dc terminals. The inductance of the dc circuit, provided mainly by the dc reactors in the terminals, prevents the current from rising during the first moment until the current control of the rectifier station has time to intervene. Generally, the commutation failure—whether caused by a disturbance in the voltage wave of the ac system or by an arc-through in a valve—will be cleared within a fraction of a cycle. In fact, intricate measuring equipment is required to detect this kind of an occurrence.

Reactive power supply and harmonics

The power factor on the ac side of a static convertor will never reach unity value. Aside from the usual reactive consumption in the transformers, the rectifiers themselves have an inductive phase lag of their ac-side current because of the overlap in the commutation produced by the inductance of the transformer and the ac system. Moreover, the rectifiers normally operate with some phase retardation, produced by the grid control, in order to have a margin for a sudden increase in the voltage ratio that might be demanded by the automatic regulators.

The transformers, in at least one of the terminal stations, are provided with tap changers, which are automatically controlled to maintain a suitable margin of grid control without unnecessary waste of reactive power. The tap-changer control is a slower action since the primary control is always provided by the grid control. The power factor of a rectifier station will be about 0.9.

The inverters must work with a larger phase lag to safeguard their commutation. Their reactive power consumption is therefore larger; in practical cases it is 50–60 percent of the active power.

When the receiving system has a low short-circuit rating, it may be necessary to supply at least part of the reactive demand by a synchronous condenser. In other cases it can be furnished by static capacitors that are combined with filters for absorbing the current harmonics. In the latter cases, the dc link will cause no increase in short-circuit capacity in the two interconnected ac systems.

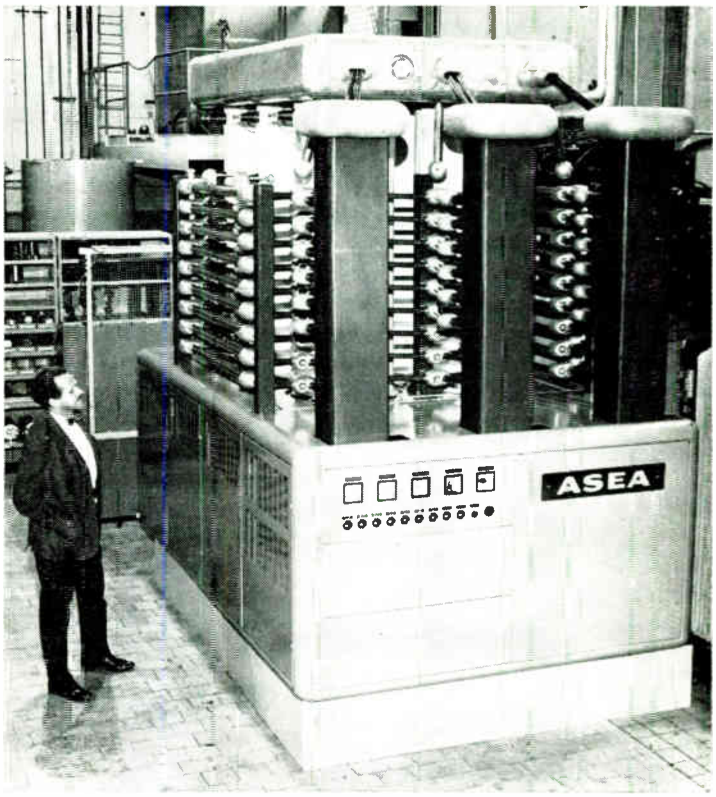


Fig. 6. View of the prototype six-anode valve for the 133-kV, 1800-ampere convertors for the Bonneville Power Administration's Pacific Intertie.

Actually, although the energy flow in a three-phase line is fundamentally even, just as in a dc line, it is an unfortunate circumstance that static convertors, in the course of their commutation, will break up this uniform flow and produce a ripple in the dc voltage and harmonics in the ac-side current. [Since the convertor itself does not store any energy, the *power overtones* are the same on the ac and dc sides of a six-pulse convertor (Fig. 6.) On the dc side, they appear as the product of the constant direct current multiplied by ac voltage components of frequencies that are 6 times, 12 times, 18 times, etc., the fundamental frequency. On the ac side, the identical power swings appear as the product of the fundamental voltage wave and the current harmonics of order numbers: 5 and 7, 11 and 13, 17 and 19, etc.]

It has been found practical, so far, in most high-voltage dc installations to provide filters on the ac side to absorb the harmonic currents. These filters generally contained tuned *L-C-R* links for each of the lower harmonics, and a high-pass assembly for the higher harmonics. The design of these filters—a complicated procedure—is important because of the considerable cost of the components. Resonances in the ac system must be avoided at any expected network configuration and at any normal frequency variation. The dc-side ripple is generally easier to handle since a large dc reactor is always required to facilitate fault elimination.

Economy of dc transmission vs. ac transmission

All dc lines are basically less expensive than ac lines insofar as the line itself is concerned. For submarine cables, the difference is particularly pronounced. The absence of a capacitive charging current and dielectric wear makes the dc cable so much cheaper that the savings will pay for the additional costs for the terminal stations that are spaced at distances of 20–25 miles.

Underground ac cables, on the other hand, can be provided with compensating reactors at regular intervals. The savings with direct current, therefore, is somewhat smaller, and the break-even distance is presently estimated at 30–50 miles.

It is significant that the first few practical applications of high-voltage dc systems have been with submarine cable transmission.

For overhead transmission on land, many efforts have been devoted to calculating the distances at which direct current becomes the more economical alternative to alternating current. It is not the purpose of this article to detail these rather difficult general comparisons, but it should be mentioned that many independent studies of line designs for direct current and alternating current, in diverse parts of the world, indicate that a dc line generally costs about 33 percent less than an ac line of the same capacity.

It should be noted that the ac line is a single one with three phase conductors, whereas the dc line has two poles with the neutrals grounded in each terminal. The ground will serve as a spare conductor if one pole is out of order. A dc line, therefore, should be regarded as equivalent to a double-circuit ac line from the viewpoint of reliability. Today, it is also possible to arrange the terminal stations in such a way that, upon failure of one pole, all the convertors are transferred to the other pole. This sound pole, by itself, can then transmit the total rated power of the link (except that transmission losses will be doubled). Hence, it may be entirely justifiable to compare a two-pole dc line with a double-circuit ac line that has two 3-phase circuits carried on the same towers. The dc line cost, on this basis, would be about 55 percent of the cost of the ac line.

In all comparisons, of course, each of the two alternative lines has to be optimized with regard to the operating voltages. It is interesting to note then that the main part of the saving in the dc alternative is achieved through the higher voltage to ground, which will result from the individual line optimizations. The higher voltage, in turn, will mean a smaller investment in conductors. This is to be expected since overvoltage and corona conditions are more favorable with direct current than with alternating current.

The relative costs just mentioned also include the capitalized value of losses in the line, which will be reduced in roughly the same proportion as the investment costs.

From the straightforward economical viewpoint, the high-voltage dc method becomes interesting when the savings on the line, including the cost of all losses, are greater than the additional cost for the terminal stations. The "additional cost" is then understood to be the difference between the cost of the dc terminal stations and all station costs for the ac alternative—including possible series-capacitor and shunt-reactor installations. Usually, however, the picture is more complicated since the configuration of existing and future systems must be considered, as well as differences in operational characteristics, expected reliability, etc.

It seems reasonable to expect that direct current will be chosen, in some cases, because of combined advantages such as savings in line and right-of-way costs, desirability of partial underground placement of the line, absence of synchronism problems, and possibility of automatic power control. Another factor that may weigh in favor of the dc alternative is that a dc interconnection interposed between two ac systems does not necessarily add to the short-circuit capacity of either system. In view of this, the expense entailed by the replacement of circuit breakers could be avoided or postponed.

Multiterminal dc systems

An obvious advantage of the ac transmission system is that an indefinite number of terminals can be connected to the same line at a moderate cost for each terminal. Some of these terminals, or branch-offs, may be added at a later date. Technically, a dc line also can have several terminals, but, if the distance between them is short, the cost will be high.

Actually, there is a quantitative difference—not a difference in principle—between alternating current and direct current in this respect. With higher voltage for an ac transmission, the cost of the terminals also will be greater. Naturally, the dc system has to compete economically in the area of the highest voltage links. As the power systems are extended and interconnected over ever-increasing distances, higher voltage lines, with greater distances between the terminals, are superimposed. In the future, the dc system may be found suitable for such superimposition wherever the longest transmission distances are involved.

The technical implications of multiterminal dc transmissions should be recapitulated here in brief. Fundamentally, the control principles are the same as for a two-terminal link. In the multiterminal case, one of the terminals (preferably one that is presently operating as an inverter station) is controlled to maintain constant voltage. All of the other terminals are controlled to maintain preset currents. Thus the first terminal will take care of the balance current; that is, the difference between the sum of all rectifier station currents and the sum of the remaining inverter station currents (comparable to the frequency controlling station of an interconnected ac system). Also, the inverter station that takes care of the balance current must have a current setting that is the function of all the current settings and is smaller than the algebraic sum of these settings.

Various quantities, required to achieve a desired overall power dispatch program and to provide potential changes to be carried out upon the occurrence of certain anticipated inner or outer faults, can be fed into the power-control computer. For this purpose, high-quality telecommunication links between stations are mandatory. Terminals with relatively low power rating may be suitably attached to only one of the poles of a two-pole line.

Selective fault clearance in such a multiterminal dc system is achieved by a combination of quick-acting isolators and grid-control action. In principle, a dc power breaker, capable of interrupting a heavy current, would be required only when two solidly paralleled dc circuits are to be separated.

International standardization—interface with the future

With the continued growth of new products—and of new nations—world trade has reached the point where the importance of international standardization must be recognized if order is to come out of the inevitable chaos. It is time we acknowledged standards as one of the prime catalysts of prosperity

Alexander C. Grove New York, N. Y.

International standardization has always been important to the smooth flow of world trade but today, with the breakup of empire and the rapid birth of new nations, the more affluent countries must re-examine their standardization policies. The first step toward international standards came in 1906 with the creation of the IEC, and in 1946 the ISO was founded. However, U.S. business has been notoriously apathetic in its attitude, with the result that many sales have been lost because of the incompatibility of American and foreign standards. With the predicted merger of IEC and ISO into what is hoped will be an imaginative, dynamic organization, the U.S. should find an incentive for more vigorous participation.

The subject of this article is so seemingly anticontroversial in the United States that it has collected its own envelope of mythology, and, therefore, has been put on a shelf to be worshipped one day a week. Like motherhood and the flag, international trade and standardization are concepts that no one speaks against. But in some parts of the world, even motherhood, or perhaps babyhood, is now being frowned upon. This noncontroversiality has had its price— apathy. Apathy has been the attitude of industry, which keeps standardization operating at a minimum level, and of the engineering profession, which views it as a tertiary bureaucratic function. The presence of some very articulate and dedicated men in standardization has not been able to overcome one of the basic illnesses of the profession, i.e., the failure to comprehend, exploit, and publicize the relationship of standardization to a growing and healthy economy. If this is generally true, it is doubly true with regard to international standardization. In the United States, which boasts the most viable and affluent economy in history, and in which the flow of gold and trade are looked upon as a sign of economic health, international standardization is not yet recognized as one of the prime catalysts of prosperity.

Perhaps the guideposts we use to judge our position in foreign trade are not quite as valid as in the past. Perhaps these guideposts, which were assumed stable, are shifting; and what is more important, were never very stable. Even the Rock of Gibralter, insurance advertisements not withstanding, has been revealed to be tubercular—porous to the core.

Motivations for international standardization

The primary motivation behind international standardization today, as it has always been, is money. It is the money that can be made via an increase trade flow between countries that is the carrot to most of the Western nations, and to some of the Communist ones as well.

The breakup of old empires and the realignment of nations into trade blocs have forced some of the affluent countries to re-examine their policies on standardization. In former years European nations could ram their products down the throats of their colonies but today they must do an excellent selling job in order to compete with other industrial powers. The same is true for the United States. New nations have been born with alarming rapidity during the last two decades—alarming because not all of these nations may develop viable economies. However, they will develop needs for industrial products, and hopefully the means by which to pay for the fulfillment of these needs. To whom will these countries turn for standards by which to measure the products and services they require? One would expect each industrial country to attempt to influence the young nations in their selection of standards, and that is exactly what is happening. The salesmen of the Western World descend upon their potential customers with all sorts of arguments in favor of competing specifications and standards. Confusion reigns.

As far back as 1904, there were thoughts about bringing some order into international electrical trade. It was then, during the International Electrical Congress in St. Louis, that the idea for the formation of an international group of technical societies was introduced. In 1906 the International Electrotechnical Commission was born. The initial mandate of the IEC was to "consider the question of the Nomenclature and Ratings of Electrical Apparatus and Machinery."

In 1946 the International Organization for Standardization was founded, with standardization responsibility for all nonelectrical items. In 1947, the IEC became affiliated with the ISO as its electrical arm; both have consultative status with the United Nations. Although these two groups have reduced the chaos somewhat in international standardization, many other international bodies are also active in this area. The World Health Organization, the International Labor Office, and the International Telecommunication Union are a few of the organizations doing outstanding but sometimes redundant work.

I. World export trade in electrical goods

	1955 $ Million	Percent of Total	1963 $ Million	Percent of Total
U.S.	773	30.0	1449	23.5
EEC	919	35.4	2472	40.0
EFTA	741	28.8	1401	22.6
Japan	31	1.2	519	8.5
Other	115	4.6	344	5.4
Total	2579	100.0	6185	100.0

Since the end of World War II the pace of international trade has increased so tremendously that the strain has been almost unbearable for the IEC and the ISO. Their staffs have grown, but hardly enough to cope with the demands of the member nations for increased efforts. It must be remembered that thousands of products and materials are in use today that were not even dreamed of prior to 1940. These are part of the new abundance of goods and standards are required for them.

Some indication of the growth in trade can be seen from Table I.[1] According to the table, in 1963 the world trade in electrical goods was $2.59 billion. It may also be noted that the United States share *dropped* from 30.0 to 23.4 percent between 1955 and 1963. Competition was setting in. All but one of the other blocs of nations increased their exports and Japan achieved a huge 500 percent export increase during this period, which is a phenomenal growth in any economic climate.

The Electronic Industries Association has compiled some interesting and dramatic statistics as part of its "1965–1970 Forecast for the Electronic Industries." Table II traces the change in the U.S. Gross National Product from 1961 to 1970.[2] It also combines history and prophecy for the same period to show the changes in dollar volume for consumer electronics and components. It is immediately evident that all categories of electronics have increases greater than that for the GNP during 1961–64 and that these categories will still increase faster than the GNP between 1964 and 1970. The rate of change for industrial electronics is estimated to be 60 percent higher than that for consumer electronics during the 1964–70 period. The question arises, "Where is all

II. Growth of the American electronics industry, 1961–1970

Year	1961	1964	1970*	1961–64		1964–70*	
Category	$\$10^6$	$\$10^6$	$\$10^6$	$\$10^6$	Percent	$\$10^6$	Percent
Gross National Product	520 100	620 700	800 000	108 600	20.8	171 300	27.3
Consumer electronics							
Total U.S. production	2 018	2 955	4 375	937	45.4	1 420	48.0
Exports	55.0	68.9	129.0	13.9	25.3	60.1	87.2
Imports	134.7	218.7	433.0	84.0	62.8	214.3	97.9
Percent exported	2.7	2.3	2.9				
Percent imported	6.7	7.4	9.9				
Industrial electronics							
Total U.S. production	2 380	3 565	6 360	1 183	49.9	2 795	78.1
Exports	436.0	723.9	760.0	287.9	66.1	36.1	5.0
Imports	22.6	47.5	82.0	24.9	110.1	34.5	72.8
Percent exported	17.6	20.3	12.0				
Percent imported	5.0	6.6	10.4				
Components							
Total U.S. production	3 600	4 034	5 488	434	12.0	1 454	35.9
Exports	113.9	148.3	325.0	34.4	30.5	176.7	119.0
Imports	41.5	81.7	219.0	40.2	96.9	137.3	170.0
Percent exported	3.2	3.7	5.9				
Percent imported	1.1	2.0	4.0				
All categories							
Total U.S. production	7 998	10 554	16 223	2 556	39.1	5 669	53.8
Exports	604.9	941.1	1214.0	336.2	55.8	272.9	29.0
Imports	198.6	347.9	734.0	149.3	75.0	387.1	111.5
Percent exported	7.5	8.9	7.5				
Percent imported	2.5	3.3	4.5				
Total electronics as percent of GNP	1.54	1.50	2.03				

* Estimated

this electronics business coming from?" The probable answer is that much of it will originate overseas from those countries trying to enter the cybernetic age. Even now each small country attempts to present the image of an economy that is on the verge of automation.

In 1961 our two-way foreign trade in nonmilitary electronics was $804 million; by 1970 it will be almost $2 billion. Although not usually presented in this way, exports and imports are combined in these figures for two reasons. First, many foreign companies are at least partially owned by American interests and the manufacturing profits come home to roost in American bank accounts. Second, and perhaps more important, any imported item provides income for Americans in the distribution and retailing trades. There are profits in any case. It is only now that the predictions of the authors of the Marshall Plan and the Point Four Program are coming true. The nations of Western Europe now give foreign aid whereas once they only received it. The expansion of Soviet Communism, once so widely dreaded, has not come to pass; no dent has been made in Africa or Western Europe. In South America there is still political churning but it is here that we have never established a wholehearted foreign aid effort.

Now let three of the most eminent men in American electrical standardization testify on motivation for standardization. Each, distinguished and extremely active in his own field, has felt compelled to promote support for international standardization.

Hendley Blackmon, President of the USNC, speaking to the 1966 IEEE International Convention, stated that American business operated for years according to this gospel:

1. The United States market is greater than that of the rest of the world.
2. The rest of the world needs U.S. goods more than we need theirs.
3. Mass production techniques and low unit costs are U.S. monopolies.
4. A favorable balance of payment situation will always exist for the United States.

These, he said, are now antiquated and invalid notions. He proposed these substitute guidelines:

1. The *rate* of economic growth is greater abroad than at home, and will continue to be so for some time to come.
2. The world is heading toward freer trade in all goods and services.
3. Future business will become increasingly competitive throughout the world.
4. In all free world countries, government intervention is increasing and must be faced as a fact of business life.
5. If the United States is not selling to the future expanding world market, other countries will take over.
6. The IEC Recommendations will be the language of international electrical commerce and the U.S. had better have a voice in how that language is written.

William McAdams, Vice President of the USNC, speaking to the same Convention said that "the U.S. electrical manufacturing industry is competing in a world market for electrical goods that is expanding at a very rapid pace. This growth has been at a much faster rate since 1959 and if it continues at this rate, we can expect a total world trade in electrical goods of something like $15 billion by 1970."[1]

Leon Podolsky, technical assistant to the president of Sprague Electric Company, warned that the U.S. electrical industry is under constant threat of being "engineered out of the market" by standards that are partial to products of other countries. This is a very common technique used within and outside of the United States. It has been successful with connectors, television sets, and household appliances. There is no accurate count of the millions of dollars lost to our economy.

Also beginning to be heard in the forums of international standardization is William E. Andrus, Jr., Group Director of Standards for International Business Machines Corporation. In May 1966 he spoke to the Business Equipment Manufacturers Association Data Processing Group Committee and pointed out that "standards based on United States needs could be a hidden international trade barrier because of the inherent nature of narrower United States acceptance parameters. Computers are basic to all information interchange. Thus, if they are to provide a total information interchange capability, they must be compatible on an international scale. This includes not only hardware and software but also languages and disciplines."

Unseen trade barriers

Recently the world has seen a reduction in the number and intensity of the tariff barriers between nations. While the millenium has not yet been reached, many tariff obstructions have come down within regional blocs such as the European Economic Community and the European Free Trade Association. There may be a day in the near future when even U.S. products may enter most markets with tariff restrictions no greater than for the products of any other country. While most eyes are on the tariff problems, however, few people even sense the insidious nontariff restrictions that affect the flow of goods between nations, restrictions that usually take the form of licensing special products to enter a country or of setting up standards and specifications for certain products such that only a single nation can supply the item. Sometimes the technique of standards writing is used to keep a *single nation out of the market*. Examples of this practice are being collected by the Department of Commerce. It is estimated that a $100 000 increase in domestic vacuum cleaner sales to Norway and Sweden would result if their electrical standards were reconciled with those of the United States. A $200 000 increase in other household appliance sales to Norway could be achieved in the same manner; and this is for a small country with a well-developed industry of its own. There should be no general inference that all other countries deliberately exclude U.S. products or write standards which, even though applied to imports, are inferior to U.S. standards. Often our standards are not applicable.

It is also true that the United States is sometimes out on an isolated limb while the rest of the world has agreed on an international standard, as is the case with television standards. In the Western Hemisphere, television standards are promulgated by the Electronic Industries Association, whereas most of the rest of the world conforms to the standards set up by the International Radio Consultative Committee of the International Telecommunications Union (CCIR). There are enough technical differences between these two sets of standards that the cost of modifying U.S. equipment to CCIR standards often results in noncompetitive pricing. One major U.S. manufacturer of television equipment estimates that this practice has cost him the following sales during the past few years: Ghana, $5–6 million; Belgium, $2–3 million; Jordan, $0.3 million; Kuwait, $0.3 million; Israel, $0.3 million; and Malaya, $1 million. The company states that "... the difference in standards played a large part in our decision not to make an offer." How often is this story repeated with other American companies?

IEC payoff to U.S. industries

It must be remembered that in our semifree enterprise system, the owners and stockholders of each enterprise always question motives: "What is in it for us?" In the heyday of the Marshall Plan and Point Four Aid Programs this question was also asked quite often by the American taxpayer. The only answer that could be given at that time was that in the future, which is with us now, increased foreign aid would provide a return of trade to the United States. Table II indicates that the predictions were quite justified for the electronics industry. Many of our citizens are direct beneficiaries of the foresight of those planners in the late 1940s.

We also have observed secondary payoffs of significant magnitude. With West European countries, an improved trading position has removed them from the rolls of U.S. aid recipients. Germany, France, England, and Italy, for example, are now themselves sponsors of aid programs. As the new developing countries achieve their industrial potential, they too will disappear as receivers and join the ranks of the donors. Even the Communist nations react similarly. Czechoslovakia, Poland, Yugoslavia, and Romania are beginning to feel semiaffluent and, tasting the benefits, are extremely concerned that their sister Communist countries might rock the boat.

Is there any correlation between the increase in international trade and the increased activity of the IEC? Specifically, has the United States benefited from its expanding participation in IEC work? From the viewpoint of engineering, which is accustomed to dealing in measurable causes and effects, one is doomed to disappointment. There is as yet no equation, no matrix, no nomograph, which will reveal any correlation. Such a study is being proposed by the American Standards Association but no support has been offered.

The National Bureau of Standards is making minor attempts to seek the truth but is also dealing with fragmentary information. At a recent meeting with the Secretary-General of the IEC in Geneva, the author initiated discussion on this problem. The only conclusion reached was that it is still a mystery. One of the main stumbling blocks is the retrieval of pertinent data, which lie in the files of embassies, commerce ministries, and private businesses all over the world. At this point it is worth quoting from a memorandum of the Technical Analysis Division of the National Bureau of Standards. Its subject is the problem of data collection to support the thesis of trade loss to the United States as a result of the barrier of hostile international standards.[3]

"This data represents the best representative responses from the 2000 U.S. firms surveyed by the Department of Commerce in preparation for the Kennedy round of tariff negotiations. The balance of our acqusitions is hardly satisfactory and merely serves to re-enforce what is already known—that there is a real scarcity of good examples of nontariff barriers and their effects, backed by hard data.

"Many of the developing nations, with their developing markets, have a European orientation due to their colonial history. They also increasingly have a sense (as their technological level rises) for their need to acquire standards. In addition, technology is creating new industries and new markets. All these facts point to our need to involve ourselves in the movement toward standardization that is developing elsewhere and must be coordinated.

"If we do not, and as the tide of standardization advances, we will find our domestic manufactures increasingly isolated and disadvantaged in export trade."

W. A. McAdams, again speaking to the 1966 IEEE International Convention, stated the problem in a slightly different manner. He said: "It is not always easy to measure in a tangible way the benefits which acrue to industry as a result of the IEC work. Sometimes the published results of long discussions on a standard seem quite small or trivial, but they usually involve fundamental engineering and design concepts that may affect the future course of a whole segment of the electrical industry. Take the standardization of extra high voltages for example. Some years ago the IEC adopted 400–420 kV as the next voltage step above 230 kV. The U.S. was not very concerned about the IEC standardization at the time and began to develop the 345–362-kV system instead. Eventually the U.S. standard became accepted by the IEC but it was in addition to the 400–420-kV system already in use in Europe. When time came for higher levels, the U.S. was anxious to avoid the previous experience which resulted in two acceptable systems, and became active in IEC work. As a result, agreement was reached on 500–525 kV in 1962 and more recently on 700–765 kV for the next higher step. Both manufacturers and users will benefit from this standardization since it will permit concentration of their research and development effort."

The following positive results of U.S. participation in IEC work can be stated:

1. Continued recognition of inch-based sizes for electric motors.
2. Acceptance of U.S. (ASTM) tests for insulating oils.
3. Recognition of U.S. rating practice for reactors.
4. Acceptance of U.S. test methods for volume and surface resistivity.
5. Acceptance of 90 percent of the U.S. recommendations for switchgear and control gear.
6. Compatibility of the new requirements for shipboard electrical installations with IEEE no. 45.
7. Acceptance of the tougher tests and ratings in use in the United States for static converters.
8. Acceptance of over 60 percent of the U.S. recommendations on fuses.
9. Acceptance of the tougher tests and ratings in use in the United States for lightning arresters.
10. Acceptance of the U.S. basic grid (0.1 inch) for printed circuit boards.
11. Compromise between U.S. and Japanese connector industries to prevent engineering of U.S. batteries out of the market.
12. Acceptance of most U.S. methods of environmental testing.

Actually, there are so many of these individual examples that it is akin to the cigarette-smoking vs. lung-cancer situation. It is known that the American participation in the work of the IEC has positive connotations relative to our foreign trade but we cannot get a correlation at this time. Thus it may be observed that the payoffs operate on both sides of the oceans. They may cause the international standard to be so written that our manufactures will be accepted overseas without drastic redesign or they may require that foreign manufactures headed for the United States be in accord with an international standard that closely resembles the domestic standard here. In this way, our own industry is protected against competition from inferior merchandise. But, as the Bureau of Standards and Mr. McAdams imply, the true test of the worth of international standardization can only be measured on a long-range basis.

U.S. participation in IEC

As stated earlier, U.S. participation in the IEC* is accomplished through the American Standards Association in New York, which has organized and staffed the United States National Committee for the IEC. Two trade organizations are the principal financial and technical support for IEC activities—the Electronic Industries Association (EIA) and the National Electrical Manufacturers' Association (NEMA). They provide most of the money and act as liaison for the technical manpower that comprises the committees of technical experts.

An index of U.S. interest in the work of the IEC is the number of Secretariats held. The Secretariat plans and organizes the work of the committee, drafts the initial documents for discussion, and consolidates the committee opinions into final draft Recommendations. At the present time, the U.S. National Committee holds the following Secretariats: TC 4—Hydraulic Turbines; TC 19—Internal Combustion Engines; TC 25—Letter Symbols and Signs; TC 37—Lightning Arresters; TC 53—Computers and Information Processing; TC 56—Reliability of Electronic Components; TC 1 (65)—Terminology (Radiology); TC 1 (70)—Terminology (Electrobiology); TC 15B—Endurance Tests for Insulating Materials; TC 31H—Equipment for Atmospheres Containing Explosive Dusts; TC 46B—Waveguides and Their Accessories; TC 53B—Digital Data Transmission; TC 59A—Kitchen Appliances.

Figure 1 shows the distribution of the Secretariats among the members of the IEC; of 129 Secretariats, England holds 25, Holland 23, France 19, Germany 14, and the United States 14. Figure 2 indicates that of 112 chairmanships, England holds 22, France 18, Switzerland 11, and the United States 11. This is hardly a record to brag about even if we did acquire eight of the 14 Secretariats since 1960.

Growth and cost of IEC

The IEC is now sponsoring more than 60 technical committees and their subcommittees. As the world's trade increases and new products arrive each year at an incredible rate, increased demands are put upon the organization. Approximately five tons of documents are processed every six weeks. A recent visit to the headquarters in Geneva revealed that such document handling is being done in corridors. How much does this effort cost and where are the cost centers located? The technical work is increasing at the rate of about 20 percent per year but because the sale of publications has increased, the amount of money from national contributions has only had to increase at the rate of 16 percent annually.

Figure 3 gives the breakdown of National Committee contributions while Table III shows the exports of electrical goods of six countries for January–June 1964.[4] If we examine the contributions for the top four members, we see that there is apparently no correlation between their contributions and the country's electrical and electronic exports.

The expenses of the IEC are not arbitrarily determined. They depend upon the amount of work that has been authorized for the Technical Committees to undertake. About 60 Recommendations were published in 1965 and another 70 are expected in 1966, which is only one Recommendation from each of the 129 Technical and Subcommittees in two years. It certainly would not be in the interest of IEC members to slow down the completion of work already under way because of a shortage of funds. The present average yearly increase of 16 percent in total national contributions only maintains the rate of progress of past years without providing for further acceleration and growth.

Cost of American participation. It is difficult to assess the total cost of American participation in the work of the IEC. We can only attach costs to the items that can be and have been measured. Costs applicable to the time that technical experts spend reviewing draft documents are not known because the time itself is unknown. However, the following cost data are available.

1. ASA will pay $24 500 in dues to the IEC in 1966.
2. The U.S. National Committee should be annually funded at $120 000 (it is now funded at about $75 000).
3. The costs for some 150 delegates to attend meetings each year is about $150 000.
4. Fourteen Secretariats are held at an approximate cost of $168 000.
5. Eleven International Chairmanships are held at an approximate cost of $11 000.
6. About three Technical Committee meetings should be held in the United States each year at a cost of $22 500.
7. The annual cost of holding a General Meeting in the United States every 15 years is about $16 000.

Thus the total direct annual cost is $512 000 or only $\frac{1}{20}$ of one percent of our electronic export business alone! This minor cost element must be considered as a major investment in future business.

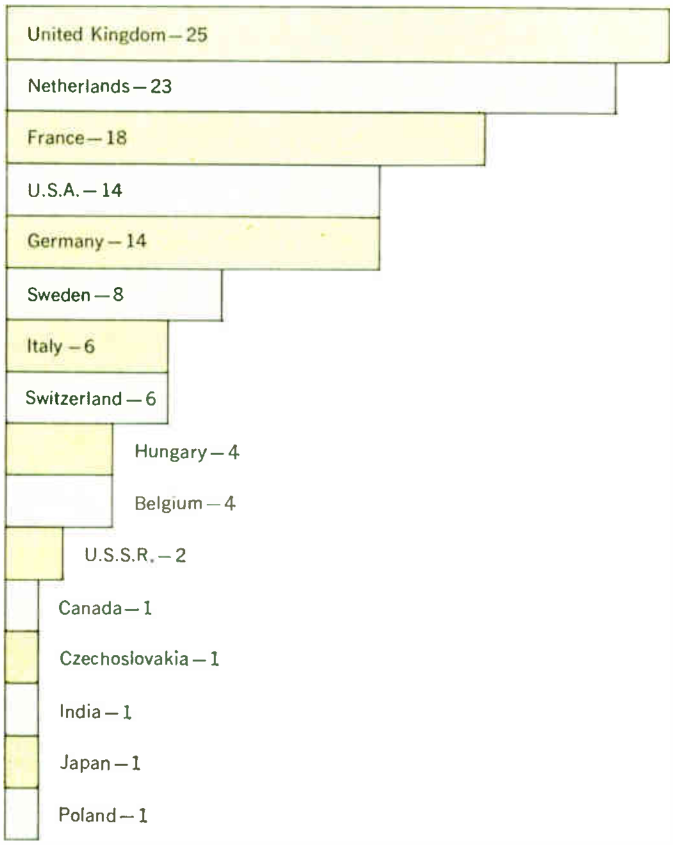
The future—innovation or suffocation?

Future predictions are always a hazardous adventure. With the IEC, however, there are such obvious problems that one can predict either their solution or the physical collapse of the Central Office staff. Figure 4 shows the structure of the IEC and the ISO with their separate staffs, heirarchy, and committees.

It is generally assumed that IEC and ISO will merge within the next decade and several committees are presenting plans to this end. For this to occur, concessions will have to come from both organizations concerning organization structure and procedures. Possibly the

---

* It is of interest to note that the AIEE was the group instrumental in obtaining American approval of the IEC.



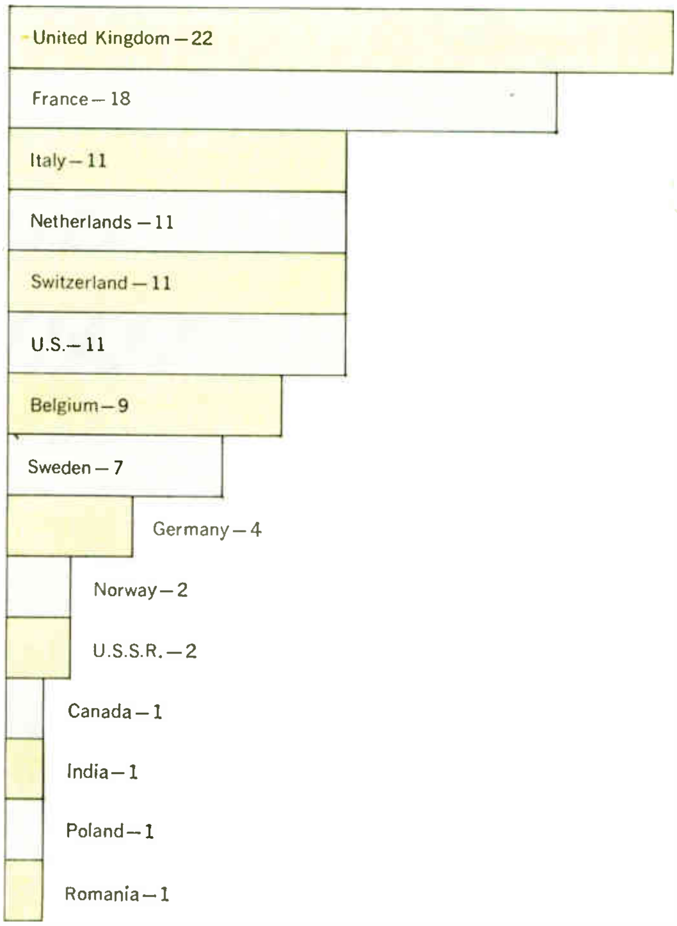
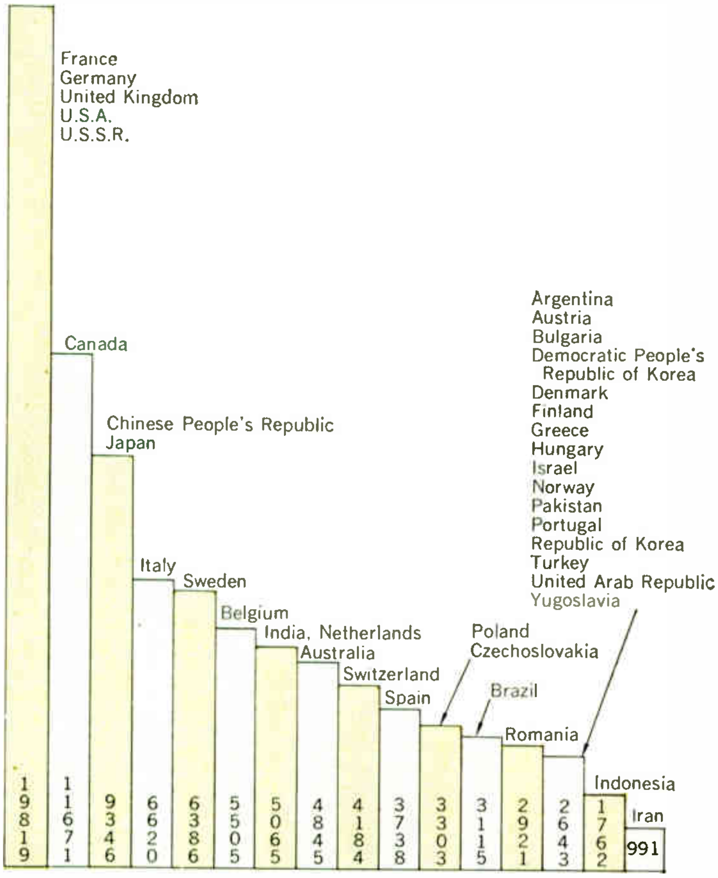


Fig. 1 (above). Distribution of IEC Technical Committee Secretariats.

Fig. 2 (right). Distribution of IEC Technical Committee chairmanships.

Fig. 3 (below). Contributions (in dollars) from National Committees—1965.



framers of the new ISO will act boldly enough to encompass all areas of standardization and set up divisions to handle the work in each area. Figure 5 shows a proposed configuration for the new organization. Under this plan, each Division would have its own Council and Secretary-General. Its operating procedures would be general and, with ISO approval, additional procedures could be adopted for specific work areas. The Electrotechnical Division (IEC) might necessarily be required to accomplish its work at a more rapid pace than the Agricultural Division and thus require a certain streamlining of procedures.

Before any case can be made for detailed new procedures, the aims and goals of the new organization must be defined. At present, these goals are rather few and are limited to production of Recommendations that will serve as performance documents in international trade.

III. Exports of electrical goods for January–June 1964

Country	Exports, thousands of dollars
United States	1 567 139
West Germany	1 195 771
United Kingdom	840 669
Japan	521 964
France	432 669
Italy	284 345

One suspects that the world will demand much more from the new ISO. Minimum goals would appear to be

1. To generate international Recommendations.
2. To establish and maintain a world standards information and data center.
3. To supply technical aid to newly developing countries.
4. To catalyze the exchange of standards engineers between member nations.
5. To stimulate the formal teaching of standardization techniques in universities throughout the world.

The new ISO must provide a much larger spectrum of services to the trading nations. It can no longer sit on a hill in Geneva far from the channels of trade. At least one field office is required to service areas on each continent. This means an immediate creation of five offices.

Generation of international Recommendations. The rate of document development is scandalously slow. If only 130 Recommendations are produced every two years, then why the hue and cry about the urgency of standardization, since this is but one document per committee? The document-per-committee index may not be the most valid one, but it is not totally unfair. Some committees have long since completed most of the standardization that they are prepared to under-

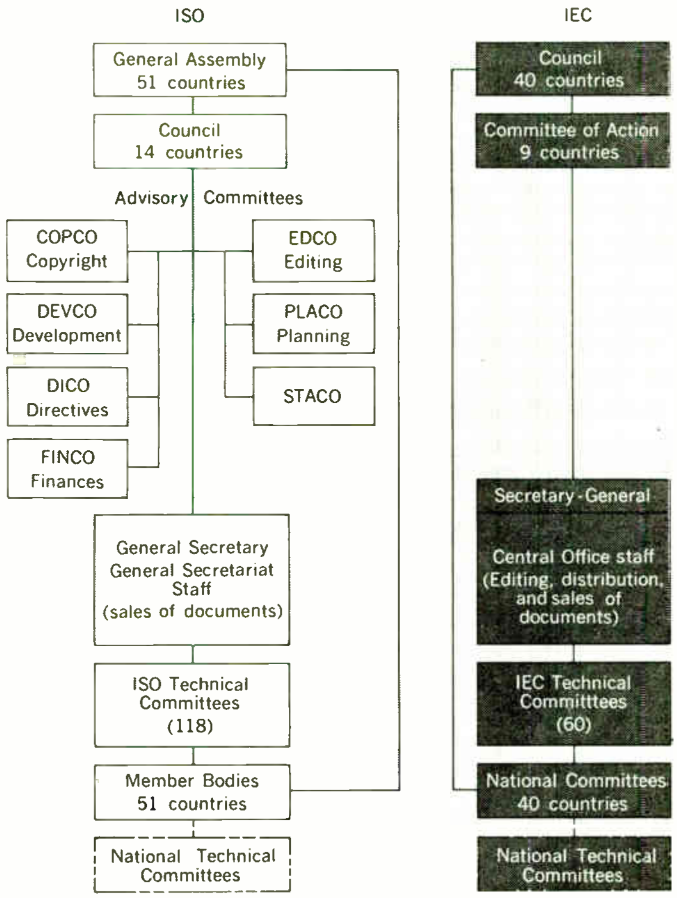
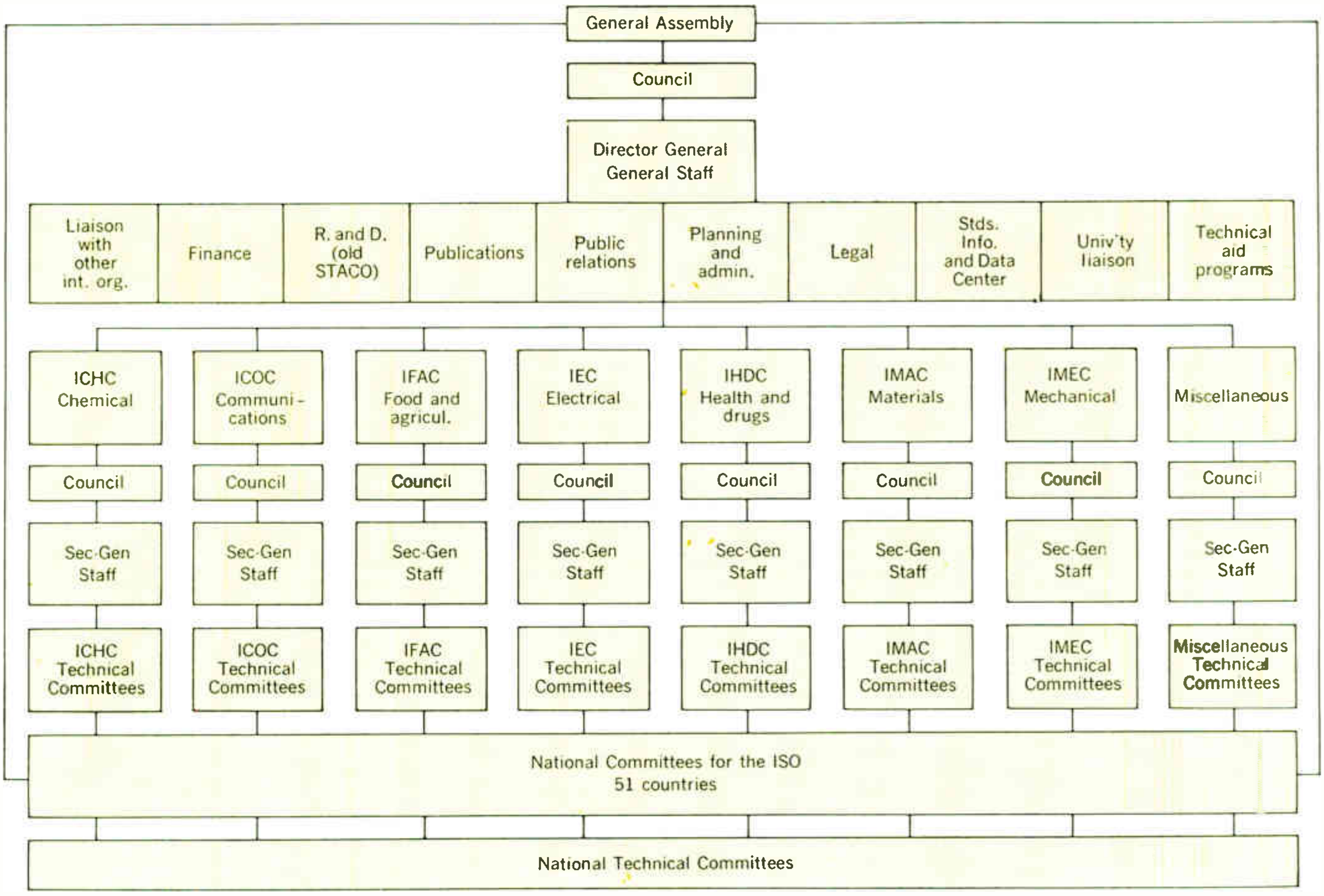


Fig. 4 (right). Present structures of the ISO and the IEC.

Fig. 5. Proposed structure of the ISO.



take. Others still have crowded agendas. A few committees work on documents for an overly long span of time and so several special committees of ISO and IEC have concerned themselves with the length of time required for document development. They investigated the details of committee procedure and office staff operations in Geneva and then came forth with recommendations that attacked the symptoms without hitting the root causes of delay. In examining the activities of technical committees, there is no mention of the obstructionism that sometimes crops up during the work. Occasionally one is reminded that participation in standardization has both positive and negative motivations. Some parties are present to abort the development of an international Recommendation, and situations exist where this actually may be a correct attitude. For example this may be true if many variations of a new product, test, or process are in use and enough data on each are not available. Again, one or two nations may be trying to get a standard drawn up that reflects specific products, and this type of standardization should be aborted. On the other hand, a majority of the nations may desire standardization on a specific item whereas one of two industrial nations may not. It is possible that a standard could be delayed from one to five years by manipulation of the procedures.

While it is not in the interest of good standardization to push documents through the mill at breakneck speed, technical committees must establish a timetable for each document and should be expected to meet their schedules. The penalty for delayed standardization is a multiplicity of higher costs to the consumer.

Thus far, there has been no mention here of the functions of the Central Office or General Secretariat staffs or of how these functions contribute to the acceleration or deceleration of the standardization process. Present studies, as previously noted, have not taken the systems approach to the problem; they have treated it as a complete entity instead of as a rather small segment of the process. Relatively minor improvements in equipment have been hailed as major breakthroughs, e.g., a tying machine for document wrapping. One cannot blame the organization because it has become used to living on handouts and the teeth of a gift horse are not to be counted. However, the world has lived through at least 15 years of modern data processing while international standardization is still in the age of the goose quill as far as its systems designs are concerned. Just as a bare start, and with very little investment, the following steps could be taken.

1. Place all documents on tape. Instead of sending tons of paper all over the world, send the tapes for the French and English versions of the documents. If there is an insistence on hard copy, then multilith-type masters can be sent via air mail. Again, if the masters differ from country to country, the receiving countries can supply their own types in bulk. With easily correctible tapes and lightweight masters, a significant saving of time and money could be achieved.

2. Establish formats for groups of standards that will permit a reduction in particular phraseology. It appears that almost all of the documents in the resistor–capacitor categories lend themselves to a standard format. Savings in time and money would accrue from: (a) Less editorial time per document. (b) Less committee time per document. (c) Less translation time per document.

Perhaps the most important step that can be taken prior to the merger of ISO and IEC is a systems analysis in depth. If this were undertaken then, for the first time, all of the sequential components and interfaces of international standardization could be studied; i.e., from the invention of a new item, test, or process, through the struggles for national acceptance to the hazards of international trade and finally to global consensus. The driving and counter forces could be examined, the intra- and intercountry information flow rates monitored, and the impedances identified as to type and value.

Establishment of an information center. An experienced delegate to meetings of almost any IEC and/or ISO technical committee must sometimes reflect upon the paucity of information regarding the standards on any particular item. It is not that this information does not exist but that the information is, at the moment of inquiry, inaccessible to the searcher. If all delegations were to bring all data relative to the items under discussion to the meetings, the freight charges would make an airline executive chuckle with joy. Yet it is not always the quantity but the quality and relevance of information that is critical.

The proposed ISO Information and Data Center (see Fig. 6) is expected to be the mother lode for all data about all national standards and product information relating to those standards. It should function somewhat like this:

1. Provide an updated index of all national, regional, and international standards; this index to be available in memory or hard copy.
2. Supply national standards documents in print, on tape, or on film in their original language. When requested to do so, it will provide translations of these documents into any language at a reasonable fee.
3. Answer inquiries about the products that conform to national and international standards.
4. Aid each technical committee by providing all available data pertinent to the items under discussion by each technical committee. By furnishing this information on film to each National Committee so that the delegates may arrive at their meetings with increased preparedness, considerable technical sparring can be avoided and the work can progress faster.
5. Maintain a library relative to the history and techniques of standardization so that persons in all technical disciplines may be trained to serve both national and international standardizing programs.

Technical aid for newly developing countries. Technical aid has come to mean almost anything that might help a recipient nation to achieve a viable and stable economy or agriculture. Often it starts at the preliterate level and so must first establish a basis of communications before any productive rapport results. Fortunately, in standardization the point of approach is at the postliterate level and there is usually a prostandards attitude. However, when this is not the case, the problems are compounded by the necessity for establishing an initial educational program that has nothing to do with standards.

Types of technical aid might vary from setting up a complete standards institution to training key people in Geneva. Essentially it means that the ISO must be able to provide personnel who will go out into the field and lead programs in almost any country on earth. Where will these trained engineers and technicians come from? First, it will be necessary to recruit, as permanent employees, a core group of men and women who have had experience in national standardization. After being trained by the ISO veteran staff, they will take an active role in the field, training others for national standardization in their own countries. This core group can be augmented by the groups who participate in the exchange programs.

Exchange of standards engineers. Though formal standardization has existed in the United States for almost 50 years, and in some countries for even longer, no cross-pollination of ideas and personnel has been attempted to any degree. Certainly the ISO should encourage such a trend, as do the professional teacher and scientific organizations. The short-range payoff would be a merging of methods, a two-way learning process. But the far-reaching effect comes from the bipartisan depth of understanding engendered by such a program, and understanding begets cooperation with subsequently improved standardization. There are situations in which the results of a good program can be regenerative, and this is one of them. It goes hand in glove with technical aid efforts and provides a burgeoning source of the manpower necessary to make a dent in the morass of present inaction.

Standardization in a worldwide curricula. In long-range corporate planning the question is often asked: "Who will be our customers 20 years from now?" Sometimes a second question is added: "Who will be our employees 20 years from now?" Successful, globally minded companies have been thinking this way for a long time and are prepared for future generations of customers and workers. Standardization—national and international—is not. Very little has been done to educate students or even the general public on the objectives and values of standards. If one were to survey the college curricula in the United States, it is doubtful that more than three could be found that include courses in standardization. The universities are not at fault; they usually respond to demand. It would be interesting to know why no demand exists when standardization is supposed to play such an important role in commerce. The standards engineering profession cannot achieve status if it exists only as fallout from the other engineering disciplines. Just how does one become a standards engineer, and is there really such an animal?

The ISO can offer a genuine service by cooperating in the creation of standards study programs on a global basis, with both engineering and economic orientations. It can offer source material from its Information Center and Publication Inventory. It can sponsor fellowships and a *Standards Journal* so that scholarly articles related to its work may receive wide readership. The opportunities in this area are limitless and so are the returns.

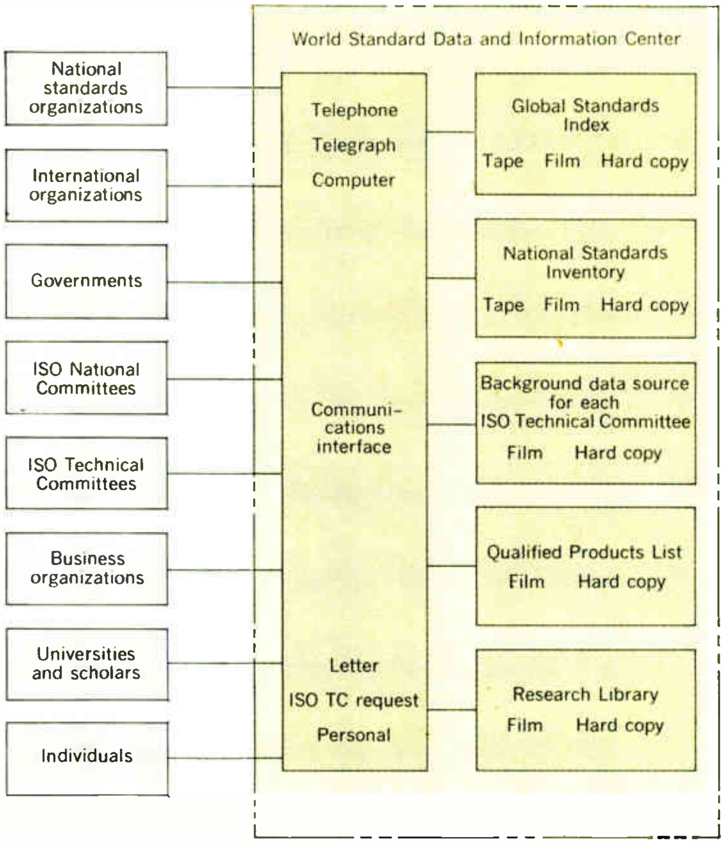


Fig. 6. Functional block diagram of the World Standards Data and Information Center.

What the United States can do

Today's international standardization movement is the product of an evolutionary process that started in antiquity. Although operating on a much more compressed time scale than that of animal evolution, its product is at about the same stage in its development that the pre-man apes are relative to man. It is time that the practitioners of standardization began to "walk erect," secure in the knowledge of their art and sensitive to the opportunities for future growth. Coincidently, the United States is now in a unique position to lead what could be an exponential increase in the scope of modern standardization. The American Standards Association will soon undergo a metamorphosis—in name, in structure, and in function. Due to the pioneering efforts of the La Que Committee, the U.S. Department of Commerce, and the industrial organizations that constitute ASA, the association is about to become the United States of America Standards Institute (USASI). A draft constitution has been written, as have new bylaws, and revised procedures are now being studied and should be submitted to the membership before the year is out. Next year USASI may be chartered by Congress, thereby conferring on it a recognition that it has long needed and which is the hallmark of almost every other national standardizing body in the world. What does all this mean to the management of international standardization in the United States? It can mean nothing, or everything. All of the documents will add up to zero if they are not implemented with imagination and talent—imagination that is freed from the gravitational field of conventionality and talent that does not fear experimentation.

Some present and future minimum goals were stated earlier without questioning whether there existed in the United States standards movement the capability to motivate these goals. Three new questions must be asked before any answers can be given. Is there any standardization philosophy? Do the practitioners of standardization in our country have the professional security to walk erect and articulate their philosophy? If articulate professionalism does indeed exist, will the business community recognize it as a force that can be extremely profitable and that should be enthusiastically supported? A superficial inquiry might lead to abject pessimism when one sees the status of standardization within individual companies. A look at the usual organization chart, showing the standards department or section as part of "engineering services" along with drafting and document reproduction, hardly encourages visions of innovation in international standardization. The stubborn fact is that the standards effort in this country has never (except for a very few hopeful exceptions) been considered more than a low-priced overhead operation remotely located from any profit-generating centers. Despite the herculean efforts of Messrs. Blackmon, McAdams, Podolsky, and Andrus, and of ASA and the Standards Engineers Society, the message of the future has hardly penetrated the fog of corporate apathy. However, it is just possible that until now a well-integrated plan for international standardization has not been presented for consideration by industry and government. Are the goals suggested earlier valid and does the brain of the primitive U.S. standards complex comprehend its real environment? Can an inspiring plan come from those who should father it? Any plan should assume that:

1. USASI must be able to place one of its own staff as Secretary of each technical committee for which the U.S. holds the Secretariat, which means that the USNC needs at least two more engineers immediately.

2. The USNC must be able to operate on a businesslike basis with regard to long-range planning, which means broadening the base of financial and technical support in industry. Every company that benefits from exports and imports must be convinced of the value of funding its own future prosperity.

3. Each technical committee must be staffed with qualified U.S. delegates so that continuity of participation can be maintained over at least five-year periods since it sometimes takes three years to develop a good delegate.

4. The United States must offer to take over more Chairmanships and Secretariats as they become available. The U.S. National Committee often loses these opportunities because it cannot be sure that industry will support its efforts.

The final plan will be a bold one that assumes the probable, embraces the possible, and reaches for the impossible. It will be predicated on solid data furnished by all those who will benefit from its operation. It will invite the participation of those professionals who can still experience a sense of excitement when they realize its potential for economic good. Perhaps best of all, such a plan will, for the first time, appear lucid enough and logical enough to motivate the support of the business community. After all, history is on its side. The industry of the United States has never rejected an opportunity for economic growth. American management will support innovative international standardization because it can be proved that this, like research and development, is a profit-generating center and that the payoff is not only evident on the corporate financial report but in more and better paying jobs at home, in more effective foreign aid to those nations that can use it, and in a more peaceful and prosperous world.[5]

Appendix. Basic phrase list of standardization

1. Standardization*

Standardization is the process of formulating and applying rules for an orderly approach to a specific activity for the benefit and with the cooperation of all concerned, and in particular for the promotion of optimum overall economy, taking due account of functional conditions and safety requirements.

It is based on the consolidated results of science, technique, and experience. It determines not only the basis for the present, but also for future development, and it should keep pace with advances.

Some particular applications are

(1) Units of measurement.

(2) Terminology and symbolic representation.

(3) Products and processes (definition and selection of characteristics of products, testing and measuring methods, specification of products for defining their quality, regulation of variety, interchangeability, etc.).

(4) Safety of persons and goods.

2. International standardization

Use the foregoing definition but extrapolate it beyond national boundaries to regional or global blocs of nations.

3. International Electrotechnical Commission (IEC)

The International Electrotechnical Commission came into being in 1906 as a result of the Resolution passed by the Chamber of Government delegates to the Saint Louis International Electrical Congress in 1904, on a motion by Col. R. E. B. Crompton (United Kingdom), "that steps should be taken to secure the cooperation of the technical Societies of the world by the appointment of a representative Commission to consider the question of the standardization of the Nomenclature and Ratings of Electrical Apparatus and Machinery."

Fourteen National Committees having been officially formed, the Council met for the first time in London in 1908 and approved the first statutes, which remained almost unchanged until 1949. The first President was Lord Kelvin and the first General Secretary was Charles le Maistre, who held this post until his death in 1953.

In 1947, the IEC became affiliated with the International Organization for Standardization (ISO) as its electrical division, while preserving its technical and financial autonomy. In this capacity, the Commission has at present consultative status (Category B) with the Economic and Social Council of the United Nations.

General meetings have been held yearly since 1947.[7]

Today the IEC has 60 Technical Committees covering the standardization of items from turbines to transistors. It is the electrical arm of the ISO.

4. International Organization for Standardization (ISO)

Origin

(a). The International Federation of the National Standardizing Associations (ISA), set up in 1926, comprised the national standardizing associations of about 20 countries.

(b). In 1944, the United Nations Standards Co-ordinat-

---

* Defined by the Standing Committee for the Study of Scientific Principles of Standardization (STACO), 1961.[6]

ing Committee (UNSCC), comprising the national organizations of 18 allied countries, succeeded the former ISA with a view to co-ordinating the activities of its members' national industries. It was above all of value as a wartime organization.

Creation

On October 14, 1946, the representatives of the members of UNSCC met in London, together with representatives of the standardization bodies of certain countries which were not members of UNSCC, in order to:

1. "Discuss and approve the constitution of a new international organization whose object shall be to facilitate the international co-ordination and unification of industrial standards."

2. "Draft recommendations concerning the technical work to be undertaken by the new Organization."

The Constitution and Rules of Procedure were subsequently ratified by other national standardization committees that had participated in the ISO. Other national standardization committees were thereafter admitted to membership in ISO.

Membership

The ISO Members are the national bodies most representative of standardization (one for each country), who have agreed to abide by the Organization's Constitution and Rules of Procedure.

Structure of the organization

The ISO consists of a General Assembly, a Council, a President, a Vice-President, a Treasurer, a General Secretary and a General Secretariat, Technical Committees, and Technical Divisions.

The General Assembly is constituted by a meeting of Delegates nominated by member bodies; it is convened at least once every three years.

Within five to ten years, ISO should merge completely with IEC.

5. Technical Committee (TC)

A Technical Committee is a group composed of delegates from all countries interested in the standardization of an area of equipment, materials, or testing methods. It is chartered by the IEC to deal only with those subjects within its scope. The Chairman and the Secretariat are elected and are the only permanent members. All other delegates may change from meeting to meeting but usually they maintain continuity over five-year periods. The primary output of a technical committee is draft documents, which become IEC Recommendations. It may form subcommittees and working groups to handle specific subareas of work.

6. TC Secretariat

The Secretariat is the body that interfaces between the national standardizing forces and the rest of the IEC. When a new technical committee is being formed, many nations often vie for the Secretariat. The Committee of Action of the IEC (its "executive committee") selects the nation for the job. For example, with TC56—Reliability, France and the United States both wanted the Secretariat. After serious discussions, France withdrew and the U.S. now holds the position.

The Secretariat is an influential position because it drafts all documents and governs the timing and direction of the work. While nominally impartial, it is in a position to judge the merits and deficiencies of the comments and proposals of other national committees. Running a Secretariat in the United States costs approximately $12 000 to $15 000.

7. National Committee

A National Committee is formed by any country desiring to participate in the work of the IEC.

The National Committees are required to be as representative as possible of all electrical interests in the country concerned: manufacturers, users, governmental authorities, and teaching and professional bodies. They are composed of representatives of the various organizations that deal with questions of electrical standardization at the national level. Most of them are recognized and supported by their respective governments.

There is only one National Committee for each country.[7]

8. U.S. National Committee (USNC)

The U.S. National Committee was founded in 1907. Since 1931, it has been affiliated both from an administrative and technical viewpoint with the American Standards Association (ASA), which is the United States member body of the ISO.

The USNC has in its membership all of the interested members of the Electrical Standards Board of the American Standards Association, and, in addition, representatives from acoustical and mechanical engineering groups and distinguished Members-at-Large. Thus, administratively, the participation of the United States in the work of the IEC goes hand in hand with the national standardization program.

From a technical viewpoint, the USNC generally utilizes sectional committees of the ASA as its Advisory Groups on each specific subject. These sectional committees, which are the national standardization committees in the United States for their particular subject, are responsible for formulating the draft opinions of the United States, which are transmitted to the IEC through the USNC. They are also responsible for nominating delegates to represent the USNC at meetings of particular IEC Technical Committees.

The National Electrical Manufacturers' Association (NEMA) and the Electronic Industries Association (EIA) are major supporters of the USNC.

This article is based on remarks made at the IEEE International Convention, New York, N.Y., March 21–25, 1966.

REFERENCES

1. McAdams, W. A., "Worldwide electrical standardization—A status report," presented at 1966 IEEE Internat'l Convention, p. 1.

2. "Forecast—1965–1970—for the electronic industries," Electronic Industries Assoc., Washington, D. C., Nov. 1965.

3. "The problem of data collection to support the thesis of trade loss from the barrier of international standards," U.S. Dept. of Commerce memorandum, Jan. 28, 1966.

4. "Trade by commodities statistical bulletins," Organization for Economic Cooperation and Development, Paris, France, series C, Jan.–June 1964, pp. 507–531.

5. Grove, A. C., "Management payoffs from international standardization," presented at 1966 IEEE Internat'l. Convention.

6. Hoffman, S. D., " 'Standardization' defined by ISO committee," *Mag. Std.*, vol. 32, p. 301, Oct. 1961.

7. "The IEC, What it is—What it does—How it works," IEC Publication, 1964, pp. 5–6.

Quantitative lightning spectroscopy

Even though lightning has always been a common, spectacular phenomenon, man has known little about its physical properties until fairly recently. Spectroscopy is now providing some of the solutions to this complex mystery

Martin A. Uman

Westinghouse Research Laboratories

The physical properties of the lightning return stroke have been investigated recently by a number of scientists, who have made detailed analyses of the optical spectrum emitted by the stroke. Several different time-integrated and time-resolved spectroscopy techniques were used. Through evaluation and comparison of the resulting data, this article presents important information on such lightning stroke properties as temperature and particle densities.

Lightning is a transient, high-current electrical discharge, having a path length of such a magnitude that it is generally expressed in kilometers. Each cloud-to-ground lightning discharge is made up of one or more intermittent partial discharges. A total lightning discharge (whose time duration is of the order of a half second) is called a *flash*; each component discharge (whose time duration is measured in hundreds of microseconds) is called a *return stroke* or *stroke*. There are usually three or four strokes per flash, the strokes typically being separated by about 40 to 50 ms. Each lightning stroke is preceded by a barely luminous predischarge, the so-called leader process, which produces a negatively charged and ionized path between cloud and ground for the return stroke to follow.

One method for determining information about the lightning stroke is by an analysis of the optical spectrum emitted by the stroke. Lightning return-stroke spectra suitable for detailed analysis have recently been obtained by Leon E. Salanave and co-workers,[1-5] by E. Philip Krider,[6,7] and by Richard E. Orville,[8,9] all of the University of Arizona. Salanave's slitless spectra were integrated on film over the time of an individual stroke, the various strokes in a flash being recorded separately. Krider's spectral data were obtained using filters to isolate particular spectral lines and photocells to measure intensity vs. time with a time resolution of about 5 μs. Orville's time-resolved spectra were obtained on the film of a high-speed streak camera, the time resolution of the system being about 5 μs.

In order to be able to determine the physical properties of the lightning return stroke from these spectral data,

it is helpful to form a rough physical model of the return stroke. When the lightning leader has deposited a negatively charged and ionized path from cloud to near ground, the high electric field between leader head and ground causes an intense breakdown there, and the return stroke begins. The ionizing return-stroke wavefront[10,11] propagates at a velocity of typically 1/3 to 1/10 the speed of light[12] up the path forged previously by the leader. The leader channel acts like a nonlinear, lossy transmission line supporting the luminous return stroke. We will concentrate our attention on a short length (meters) of the leader channel. That length is traversed by the return stroke in about 0.1 μs, a time very short compared with the resolution used in lightning spectroscopy to date. We will therefore consider that the physical events taking place along the length of this short section of lightning channel occur simultaneously. The development of a short section of the return-stroke channel is thought to take place as follows[13-21]: The initial increase of current in the leader channel caused by the return stroke quickly transforms the leader into a high-temperature,* comparatively narrow, current-carrying channel. The initial gas density in the return-stroke channel is probably near that for standard temperature and pressure (or somewhat less, depending on the state of the leader), but the temperature has risen sharply. The channel pressure (roughly nkT) exceeds the pressure of the surrounding air and the channel will expand. This expansion takes place with supersonic speed and hence produces a roughly cylindrical shock wave, which eventually becomes a part of the thunder we hear. The shock-wave phase of the channel expansion lasts about 5 to 10 μs. As the shock wave expands, the gas density in the current-carrying channel behind it decreases, and late in the shock-wave phase the temperature in the channel is probably near 30 000°K. After this phase of the channel expansion is completed, the high-temperature low-density channel approaches, in microseconds or a few tens of microseconds, a state of approximate pressure equilibrium with the surrounding air, and the current density in the channel stabilizes at about 10^3 A/cm^2, which is about two orders of magnitude greater than the current density in a free-burning arc in air carrying a similar current. The channel now slowly expands and the current density slowly decreases to the value characteristic of a stable arc. During this time thermal conduction and convection become important in determining the physical conditions in the spark (arc).

We can obtain a rough idea of the variation of the channel radius with time from the experiments of Norinder and Karsten[20] and of Flowers[15,16] and from the theory of Braginskii.[14] A typical lightning current measured at the ground rises to a peak of 20 000 amperes in about 1 μs or so, or less for strokes after the first in a multistroke flash,[22,23] and falls to half value in about 40 μs. It is not clear that the return-stroke current waveform above the ground is similar to that measured at the ground, but we shall assume that it is. Norinder and Karsten have measured the diameter of a long oscillatory discharge in air whose current rose to a peak of 10 000 amperes in 4 μs (1/4 period). Initially the luminous diameter was very small. At current maximum it had expanded to about 1 cm, at the first current zero (8 μs) to about 1.5 cm. Flowers found that a nonoscillatory spark, which reached a peak current of 22 000 amperes in 3.3 μs, exhibited a diameter of 1.3 cm at that time. The theory of Braginskii, which relates diameter to current, follows the experimental data quite closely during the shock-wave phase if it is assumed that the conductivity of the channel is constant at 150 mhos/cm[17,18,24] and that Braginskii's ξ factor is 4.5.† According to the theory of Braginskii, the typical lightning current just described would be characterized by a diameter of about 2 cm at 10 μs. The channel should reach a luminous diameter of about 4 cm when the current is half of peak value. (The 4-cm value comes from extrapolation of data of Norinder and Karsten and of Flowers.)

The particle energies in the return-stroke channel should be maximum at a time early in the shock-wave expansion phase, probably before current maximum.

---

* In this article the word "temperature" is often used in a very loose sense to indicate a measure of the particle average energies. Local thermodynamic equilibrium may not exist under all conditions within the return stroke.

† These calculations were performed by J. Hanlon, University of Arizona.

The maximum in particle energies will be reflected in the spectrum in that lines of high excitation potential will be visible. These will be lines of singly ionized nitrogen (N II), singly ionized oxygen (O II), and, possibly, doubly ionized nitrogen (N III). (The last mentioned is not yet identified in lightning.) During the early shock-wave phase, the high particle densities and energies may cause the channel to be opaque (optically thick) to its own radiation at certain wavelengths, at which the channel will radiate like a blackbody. As time increases and current decreases, the channel temperature will also decrease. A channel near atmospheric pressure whose temperature is between 15 000°K and 25 000°K will be characterized by radiation from both neutral and singly ionized nitrogen and oxygen as well as by recombination radiation. When the channel temperature falls below 15 000°K, primarily neutral radiation and recombination radiation should be evident. For a given pressure, each spectral line from a given species of particle will reach a maximum at a fixed temperature. Above that temperature the species becomes ionized and ceases to radiate that spectral line; below it the excitation is insufficient to populate the upper energy level characteristic of the spectral line. Thus, during the time that the channel is at approximately atmospheric pressure, after the shock-wave phase, the temperature decrease in the return-stroke channel will cause lines of lower excitation potential to reach maximum with increasing time.

It is not practical at present to attempt to measure radiation as a function of position within the lightning channel. Hence all lightning spectra thus far obtained yield at best the total radiation at a given wavelength emanating from a short length of channel regardless of where within the channel that radiation originated.

The lightning spectra can be analyzed only if we assume that, at a given time, physical conditions are constant throughout a cross section of the channel. The temperature profile of the channel is determined primarily by the means available for transporting heat out of the channel. If, for example, thermal conduction is the dominant mechanism, there must be a temperature gradient throughout the channel; i.e., the temperature decreases with increasing radius. If radiation that escapes the channel is the dominant energy-loss mechanism, the channel's temperature profile will, as assumed, be relatively flat; it also probably changes considerably with time.

Theory

In this section we will examine the techniques available for the determination of the particle densities and average energies within the return-stroke channel from an analysis of the return-stroke spectrum. The first step in analyzing a lightning spectrum must be a demonstration that the lightning channel is either optically thin or optically thick, or intermediate, to the particular wavelengths of interest. This determination can be made for the N II spectral lines by comparing the measured relative intensities of multiplet members with the predictions of theory and with previous laboratory measurements. Orville and the writer[25] have shown that for the stroke-integrated data of Salanave the lightning channels are optically thin to the N II lines during the time that the greater part of those line intensities are integrated on film. It has not yet been possible to make an opacity determination on the time-resolved lightning spectra because the wavelength resolution on these spectra is insufficient to allow the various multiplet members to be distinguished. Zhivlyuk and Mandel'shtam[26] report time-integrated lightning spectra from lightning channels that are optically thick at the centers of certain O II and N II lines.

If the radiation emitted by either optically thin or optically thick lightning channels is to be analyzed, it is usually necessary, from a practical point of view, to assume that local thermodynamic equilibrium (LTE) exists within those channels as a function of position and time. If this assumption is not made, one must resort to detailed calculations involving excitation and ionization cross sections, recombination rates, transition probabilities, etc., many of which are now known, in order to reduce the measured radiation to particle densities and energies. Unfortunately, proof of the existence of LTE requires a detailed calculation involving the same unknown atomic parameters. On the brighter side, whether LTE exists or does not exist, electron densities can be determined from a measurement of the Stark broadening of certain spectral lines.

In LTE the atomic energy levels within an ionization state are populated according to Boltzmann statistics:

$$N_n = \frac{Ng_n}{B(T)} \exp\left[-\epsilon_n/kT\right] \qquad (1)$$

where N_n is the number density of atoms in energy level n, N is the total number density of atoms, ϵ_n is the excitation potential of the nth level, k is the Boltzmann constant, T is the absolute temperature, g_n is the statistical weight of the nth level, and $B(T)$ is the partition function, given by

$$B(T) = \sum_j g_j \exp\left[-\epsilon_j/kT\right] \qquad (2)$$

The following relation, the Saha equation, is valid for a system of particles in LTE:

$$n_e = \frac{N^i}{N^{i+1}} \frac{2}{h^3} (2\pi mkT)^{3/2} \frac{B^{i+1}}{B^i} \exp\left[-\chi/kT\right] \qquad (3)$$

where n_e is the electron density, the superscripts give the state of ionization (N^0 is the number density of neutral atoms of a given type, N^1 is the number density of singly ionized atoms formed by ionization of those neutrals, etc.), χ is the ionization potential from the ith to the $(i+1)$th ionization state, h is Planck's constant, and m is the electron mass.

The intensity of an emission line from an *optically thin* gas per unit volume of gas at uniform temperature and density due to transitions from level n to level r is

$$I_{nr} = CN_nA_{nr}h\nu_{nr} \qquad (4)$$

where A_{nr} is the Einstein transition probability, ν_{nr} is the frequency of the emitted photon, and C is a geometrical factor. If we use Eq. (1), that is, assume that LTE exists, the intensity may be written

$$I_{nr} = \frac{CNg_nA_{nr}h\nu_{nr}}{B(T)} \exp\left[-\epsilon_n/kT\right] \qquad (5)$$

If the number densities in the gas are known and the absolute intensity is measured, temperature can be determined from (5). More often, (5) is used with a known temperature and measured intensity to determine number density. The ratio of the intensity of an emission line due to transitions from level n to level r to the intensity of an emission line due to transitions from level m to level p is given by

$$\frac{I_{nr}}{I_{mp}} = \frac{g_n A_{nr} \nu_{nr}}{g_m A_{mp} \nu_{mp}} \exp\left[-(\epsilon_n - \epsilon_m)/kT\right] \tag{6}$$

which can be solved for the temperature, as follows:

$$T = \frac{\epsilon_m - \epsilon_n}{k \ln\left(I_{nr} g_m A_{mp} \nu_{mp} / I_{mp} g_n A_{nr} \nu_{nr}\right)} \tag{7}$$

Thus the measured intensity ratio of two spectral lines emitted by the same type of atom from an optically thin gas along with tabulated atomic parameters is sufficient to determine the temperature of that gas. In practice, to obtain an accurate temperature determination using (7), $(\epsilon_m - \epsilon_n)$ should be chosen to be greater than kT. For (6) and (7) to be valid it is unnecessary to require "complete" LTE to exist, but only that the energy levels involved in the pertinent transitions be occupied according to Boltzmann statistics. Should only the latter be true, the temperature under consideration will be the electron temperature, since it is electron excitation and de-excitation collisions that are responsible for the maintenance of the Boltzmann distribution. It is, as previously mentioned, not possible to determine whether complete LTE exists in the return-stroke channel because of the uncertainty in many pertinent atomic parameters. It has been shown by the writer, however, that the particular energy levels of the N II ion that have been used to measure lightning temperatures are populated according to Boltzmann statistics for the conditions existing in the return-stroke channel. In addition, (7) remains valid on a time scale of 0.1 μs, except possibly during the initial few tenths of a microsecond of the return-stroke current increase.[27]

In the case of a gas that is *optically thick* at several wavelengths and in LTE, a blackbody temperature can be determined for that gas by comparing the measured intensities at wavelengths for which the gas is optically thick with the Planck radiation law

$$I_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \tag{8}$$

Consider now atoms of type C denoted by subscript C and atoms of type D denoted by subscript D. The intensity ratio of a spectral line due to atoms D to a line due to atoms C from an *optically thin* gas is, using (5),

$$\frac{I_D}{I_C} = \frac{N_D g_{Dj} B_C A_D \nu_D}{N_C g_{Ci} B_D A_C \nu_C} \exp\left[(\epsilon_{Ci} - \epsilon_{Dj})/kT\right] \tag{9}$$

If, for instance, atoms D are neutral atoms and atoms C are singly ionized ions formed by ionization of D, the elimination of N_D/N_C in combining (9) with the Saha equation, (3), yields an expression for electron density as a function of temperature and line intensity ratio. Once the temperature has been determined using (7) and the intensity ratio of a neutral line (e.g., N I) to a singly ionized line (e.g., N II) has been measured, the electron density can be calculated. As an alternative to this approach, one can use the tables of thermodynamic properties of air computed by Gilmore[28] and others[29] and various measured intensity ratios—for instance, neutral oxygen (O I) intensity to singly ionized nitrogen (N II) intensity—to determine electron density and additional properties of the return-stroke channel. Gilmore's tables, which represent the solution to a number of coupled Saha equations, equation of charge conservation, and equation of percentage composition, take the place of (3) in that they provide the theoretical ratio of the density of any two constituents of the air as a function of the temperature and mass density of the air.

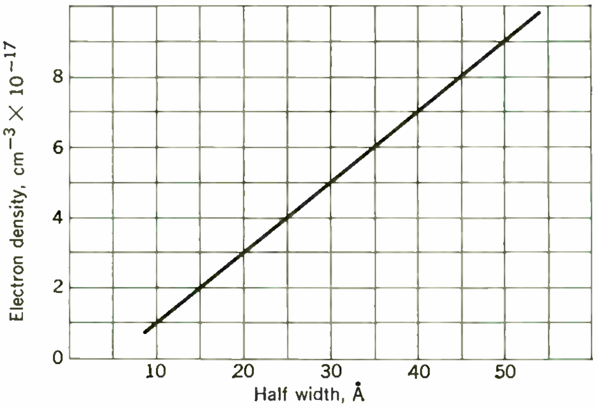
The use of the Saha equation to calculate electron density requires that complete LTE be maintained. There is some question whether complete LTE exists within the return-stroke channel as a function of position and time.

Fortunately, electron density can be determined without resort to the assumption of LTE. The width of a spectral line, which is broadened predominantly by the Stark effect, is primarily dependent upon charged particle number densities and only slightly dependent on particle energies. Stark shifts, widths, and profiles of many spectral lines are given by Griem.[30] Hydrogen is present in the lightning stroke by virtue of the decomposition of water vapor, and the H_α line is considerably Stark-broadened so that an accurate electron density may be determined from a measurement of its width. Figure 1 shows the full width of H_α at half intensity as a function of electron density. The theoretical Stark profiles for the Balmer series have been well substantiated by experiment.[31] The resolution of the available lightning spectra is inadequate to determine electron density from the Stark widths or shifts of the N II lines, the most prominent features of the lightning spectrum.

Results of time-integrated spectroscopy

Two slitless stroke-integrated lightning spectra obtained by Salanave and co-workers are shown in Fig. 2. In slitless spectroscopy the light source itself acts as the slit, and hence one obtains a spatial picture of the lightning channel at each wavelength where there is spectral emission. The wavelength resolution on the spectra shown in Fig. 2 is about 2 to 3 Å. We will consider now the analysis of a spectrum of this type. As previously mentioned, it was shown[25] that the N II lines of interest in the stroke-integrated spectra are emitted by a channel that is optically thin to those lines during the time that the greater part of those line intensities are integrated on film. The opacity determination is made as follows:

Fig. 1. A theoretical plot of half width vs. electron density for H_α at 20 000°K.[30]



According to Eq. (6), if two spectral lines originate in an optically thin gas from the same upper atomic energy level, their intensity ratio will be independent of temperature. Thus the intensity ratio will also be independent of spatial and temporal integration on film. Therefore, with the exponential term set equal to unity, (6) specifies the expected intensity ratio of two spectral lines with the same upper energy level emitted from an optically thin gas. Figure 3 shows measured line profiles for an N II multiplet whose component members originate from approximately the same upper energy level and have approximately the same wavelength. The intensity ratios should depend only on the ratios of the gf products (where the oscillator strength f is a measure of the transition probability A) if the lightning channel is optically thin to the multiplet. In Table I a comparison is given between measured intensity ratios within two N II lightning multiplets and the results of theory and laboratory experiment. The lightning channels would appear to be optically thin to the two N II multiplets.

Since the lightning channel is optically thin to the N II lines of interest, and since the energy levels which yield those lines are populated according to Boltzmann statistics on a 0.1-μs time scale,[27] a lightning "temperature" can be calculated from (7). To do so using (7) involves the implicit assumption that the return-stroke temperature is constant in time and is constant across the channel cross section. Since this is not the case, the calculated temperature will represent some average value. Prueitt[32] has calculated this average temperature, which he terms the excitation temperature, for five different strokes using a variation of (7) applied to five N II multiplets. His measured average temperatures range

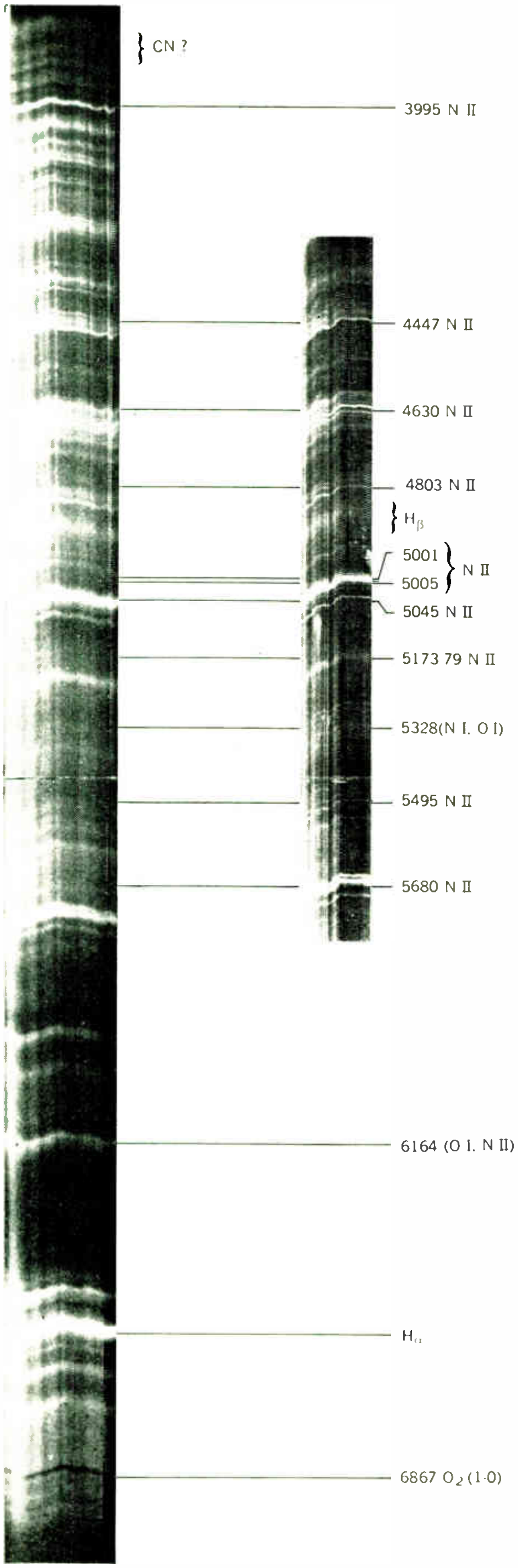
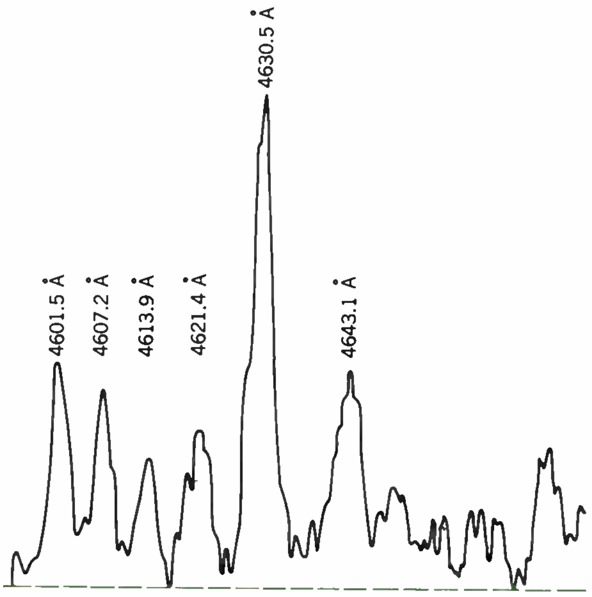


Fig. 2. Time-integrated slitless spectra of two lightning strokes.[2] (Courtesy Institute of Atmospheric Physics, University of Arizona, 1962)

Fig. 3. Relative intensity vs. wavelength within the N II (5) multiplet for lightning stroke A-3, no. 2. Dashed line indicates the line base chosen.[25]



from 24 200°K to 28 400°K; see Table I. These average temperatures are average electron temperatures if "complete" LTE does not exist in the channel.

Although it is not possible to determine the error incurred by assuming the channel properties to be uniform as a function of channel radius, it is possible to make an estimate of the error incurred by assuming the return-stroke temperature to be constant as a function of time. The writer integrated (5), by computer, for two N II multiplets for various assumed temperature vs. time characteristics.[34] The ratio of integrated multiplet intensities was matched with the measured ratio by adjusting the peak temperature for a given variational form of temperature vs. time. Using this technique it was found that if the average temperature was near 24 000°K, the peak temperature was within 10 percent of the average. This result is reasonable since in the temperature range of interest the N II line intensities increase rapidly with temperature, thus strongly weighting the average temperature toward the peak temperature. It should be noted, however, that the lightning temperature may rise to a very high value for a very short time without leaving a measurable line intensity on the time-integrated spectrum; thus, peak temperature really means the highest temperature that the return-stroke channel attains for about a microsecond or longer.

Once the return-stroke temperature has been determined, the techniques involving use of the Saha equation can be applied to determine electron density, channel pressure, mass density, electrons per air atom in the channel, and other channel properties. Three return strokes, found by Prueitt to have average temperatures near 24 000°K, have been analyzed.[35,36] It was assumed that the spectral lines of O I, N I, and N II considered in the analysis were primarily emitted at temperatures near 24 000°K. The existence of complete LTE, optical thinness, and uniformity of cross-sectional properties was also assumed. Partial results of the analysis are shown in Table II. The electron density calculated is of the order of 3 × 10^{18} electrons/cm^3. A fundamental uncertainty in the analysis is due to the assumption that the spectral lines of interest are emitted primarily at temperatures near 24 000°K. Although it is not unreasonable to expect the O I, N I, and N II line intensities to be near maximum at 24 000°K, the O I and N I lines will contribute more to the time-integrated spectrum at temperatures below 24 000°K than will the N II lines. Since we wish to know the return-stroke properties near 24 000°K, which is the temperature determined from an analysis of N II lines, the measured intensities radiated by the O I and N I lines at temperatures below 24 000°K should be subtracted from the total O I and N I radiation. The result of overestimating the O I and N I contribution at 24 000°K is an overestimation of the mass density and electron density in the channel at that temperature. The calculated values of mass and electron density in Table I should therefore be considered upper limits to the actual values at 24 000°K. The number of electrons per air atom is a relatively insensitive function of the mass density. For example, if $\rho/\rho_0 = 0.01$, where ρ_0 is the mass density at standard temperature and pressure, the number of electrons per air atom is 0.98. A further error may occur if the N II lines are emitted primarily from a hot central portion of the channel while the N I and O I lines originate from a cooler part of the channel.

Electron densities have been determined on time-integrated spectra by a comparison of the Stark profiles of H$_\alpha$ (Balmer series) with theory.[37] Figure 4 shows an example of that comparison. Electron densities of between 1 × 10^{17} cm^{-3} and 5 × 10^{17} cm^{-3} have been found for three return strokes. Note that both Fig. 4 and Tables I and II give data for a stroke labeled A-1 (by Prueitt). According to Drellishak,[38] by calculation and assuming LTE, the electron density in a nitrogen plasma at atmospheric pressure is between 1 × 10^{17} cm^{-3} and 2 × 10^{17} cm^{-3} for temperatures between 14 000°K and 35 000°K. Since the Stark profiles are strong functions of electron density and only weak functions of temperature, the H$_\alpha$ profiles emitted from the lightning channel would not be expected to change much from the end of the shock-wave phase when the channel has attained a near-atmospheric pressure until the time

I. Intensity ratios within N II (3) multiplet and N II (5) multiplet for five lightning strokes[25]

Stroke Destination	Average Excitation Temperature,* °K	N II (3)		N II (5)				
		A	B	C	D	E	F	G
a†		2.6	8.0	3.0	3.8	5.0	3.8	3.0
b‡		2.8	9.9	3.2	4.1	7.2	4.8	2.9
A-1	24 400 ± 800	2.8		2.6	2.7	4.1	3.1	2.3
A-2	27 800 ± 400	2.4		2.1	3.1	4.4	3.8	1.9
A-3, no. 1	24 900 ± 400	3.0		2.9	3.1	3.5	3.1	2.3
A-3, no. 2	26 600 ± 600	3.1		2.7	3.6	4.6	3.1	2.1
B-1, no. 1	26 900 ± 600	2.4	6.8	2.9	2.5	4.2	3.4	2.0
B-1, no. 2	28 400 ± 1000	2.7	7.6	2.4	2.8	4.3	3.5	2.2
B-1, no. 3	27 500 ± 900	2.8	6.5	3.0	2.8	4.0	3.5	2.2
C-1, no. 1	24 200 ± 900	2.5	6.7	2.8	3.8	4.6	3.5	2.4
C-1, no. 2	24 700 ± 500	2.8	7.2	3.1	3.2	6.1	3.3	2.3

Letters designate the following measured line intensity ratios:
A = (5679.6 + 5686.2 + 5676.0)/5666.6, B = (5679.6 + 5686.2 + 5676.0)/5710.8, C = 4630.5/4601.5,
D = 4630.5/4607.2, E = 4630.5/4613.9, F = 4630.5/4621.4, G = 4630.5/4643.1

* From Prueitt[32]
† Theoretical calculation of gf ratio by Griem[30]
‡ Measurement of gf ratio by Mastrup and Wiese[33]

II. Several properties of three lightning strokes at 24000° K[36]

Parameters	Stroke A-1	 A-3, no. 1	 C-1, no. 1
$(\rho/\rho_0)_{avg}$	0.18	0.076	0.12
P, atmospheres	56	24	38
Electrons per air atom	0.78	0.84	0.80
$n \times 10^{-18}$ in cm^{-3} (1)	3.9	2.3	2.7
$n \times 10^{-18}$ in cm^{-3} (2)	5.8	2.6	2.6
$n \times 10^{-18}$ in cm^{-3} (3)	7.5	3.4	5.1

Electron densities calculated by (1) comparison of intensities of nitrogen lines at 4935 Å and 3995 Å and use of the Saha equation, (2) comparison of intensities of nitrogen lines at 4935 Å and 4447 Å and use of the Saha equation, (3) use of Gilmore's tables in conjunction with two O I/N II line intensity ratios and two N I/N II line intensity ratios. The quantity ρ_0 is the mass density of air at standard temperature and pressure, ρ the mass density under conditions prevailing within the channel, P the channel pressure, and n the electron density within the channel. The stroke designation is that used by Prueitt.[32]

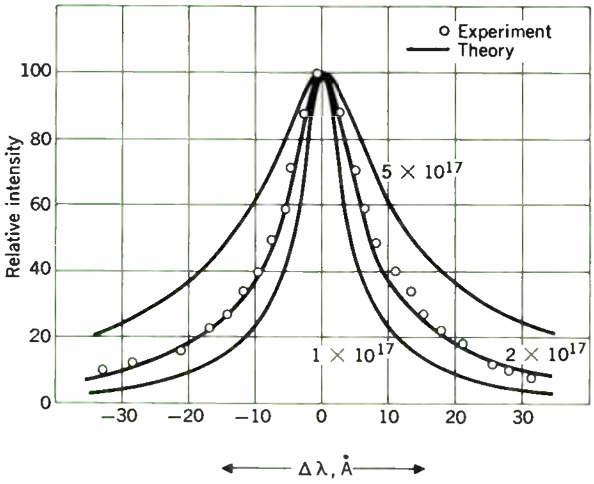
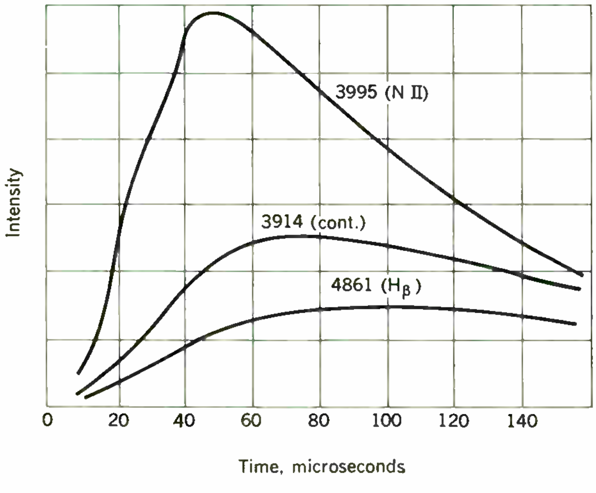


Fig. 4. Comparison of the measured and calculated line profiles for H_α in lightning stroke A-1.[37] Theoretical profiles are given for three electron densities at 20 000°K. The agreement between experiment and theory must be considered fortuitous in light of the various sources of error.

Fig. 5. Curves of intensity vs. time, obtained from approximately 80 meters of a return-stroke channel.[6] The intensity scale is different for each filter.



at which the temperature falls below 14 000°K. Thus the value of electron density determined from Stark-broadening of H_α would appear to be characteristic of a return-stroke channel at atmospheric pressure with temperature above 14 000°K. The electron density determined using Stark-broadening of H_α is consideraby more reliable than the electron density determined from the Saha equation, although the latter may be indicative of an earlier stage of the discharge than the former.

In addition to return-stroke properties determined from the slitless spectra of Salanave and co-workers,[1-5] measurements on time-integrated slit spectra have been made by Wallace[39] and by Zhivlyuk and Mandel'shtam.[26] The slit spectra are inferior to the slitless spectra in that the former represent the integrated radiation from a number of strokes and from all heights on the channels whereas Salanave's slitless spectra represent the integrated radiation as a function of channel height from single strokes. Wallace has determined a lightning temperature of 6000°K to 30 000°K from a study of the optically thin intensities within the N_2^+ bands near 39000 Å. In addition, Zhivlyuk and Mandel'shtam measured the relative intensities at the centers of several spectral lines and, assuming the channel to be optically thick to those line centers, used (8) to calculate an average blackbody temperature of 21 000°K. The return-stroke spectra obtained by Zhivlyuk and Mandel'shtam would appear to be significantly different from those obtained by Salanave and co-workers.

Results of time-resolved spectroscopy

A graph of lightning spectral intensity vs. time, measured with filters and photocells by E. P. Krider,[6] is shown in Fig. 5. Krider found that the N II lines reach intensity maxima first (in a variable time, typically 40 μs), followed in 25 to 35 μs by the continuum maximum, which in turn is followed by the peaking of the neutral hydrogen lines in 30 to 40 μs. Lines of lower excitation potential therefore reach a maximum later in time, consistent with a channel temperature that decreases with time. The indication is that the effective excitation potential of the continuum lies between that of the ions and the neutrals, consistent with a continuum due to radiative recombination or radiative attachment. Orville and the writer[40] have shown that the time-integrated continuum on Salanave's spectra is not due to blackbody radiation or electron–ion bremsstrahlung emitted at constant temperature. The time required for the N II lines measured by Krider to reach maximum intensity varied from stroke to stroke, and Krider suggests that this time may depend on the time for the wavefront of the return stroke to traverse the particular length of the channel section under observation. Further, Krider[41] has suggested that the length of channel under observation may have been considerably more than the 80 meters calculated for a channel about 10 kilometers away, since a return stroke traveling at 1/10 the speed of light[12] would traverse 80 meters in about 2.5 μs.

A time-resolved return-stroke spectrum is shown in Fig. 6. This spectrum, one of several obtained by Orville[8] during the summer of 1965, represents the first detailed spectral data ever obtained for lightning with both good spatial resolution and good temporal resolution. These spectra represent the radiation from about a 10-meter length of channel. The time resolution is approximately

5 µs and the wavelength resolution approximately 15 to 20 Å. Orville finds that the time for the N II line intensities to rise to peak is 10 µs or less and that these lines remain intense for as long as 30 to 60 µs. A preliminary temperature vs. time curve as determined from N II relative line intensities using (7) is shown in Fig. 7. The peak temperature is near 30 000°K, and the temperature decays to 12 000°K in about 40 µs. The temperature measurement made in the time interval 0 to 5 µs is subject to considerable error, and hence the apparent delineation of a rising channel temperature during the first 10 µs must be considered suspect. The wavelength resolution in these spectra is inadequate to permit determination of whether the channel is optically thin or is optically thick, or intermediate, to the N II radiation.

In his spectra Orville finds faint lines, due to neutral nitrogen and oxygen, which persist for as long as 200 µs. Unfortunately, these lines are too weak to allow a measurement of their intensity. The neutral hydrogen line H_α is prominent and reaches maximum intensity at about 40 µs. The wavelength resolution is insufficient to allow electron density to be determined from the width of H_α by means of the Stark effect.

Summary and conclusions

The lightning temperatures obtained by Prueitt[32] from Salanave's time-integrated spectra[1-5] are in reasonably good agreement with the temperatures determined by Orville[9] from time-resolved spectra. Prueitt found *average* lightning temperatures ranging from 24 200°K to 28 400°K. Orville integrated, with pencil and paper, his time-resolved line intensities and from the integrated intensities has determined an average lightning temperature that can be compared with Prueitt's. This average is about 25 000°K for a stroke whose peak temperature is 30 000°K. Orville found that the rise to peak temperature takes place in 10 µs or less. It is important to note that this peak temperature determination is in itself an average over about 5 µs. It is therefore possible that a considerably higher temperature was reached for a short time. The writer's analysis,[34] relating average to peak temperature, becomes invalid at temperatures much above 30 000°K, so a reliable estimate of the possible maximum temperature within the 5-µs interval cannot be made.

Our knowledge of the various particle densities present in the return-stroke channel is at present less reliable than our knowledge of the lightning temperature. The most reliable measurement of particle density is that made from the observation of the Stark-broadening of H_α in the time-integrated spectra. The electron density measured, about 2×10^{17} cm^{-3}, is probably indicative of a channel whose pressure is near atmospheric and whose temperature is above 15 000°K. Such conditions can be expected to exist within the channel after the shock-wave phase. If we are willing to assume that after the shock-wave phase the lightning channel exhibits a pressure of one atmosphere, is in LTE, and has essentially a flat temperature profile, the temperatures measured by Orville can be used in conjunction with the tables of thermo-

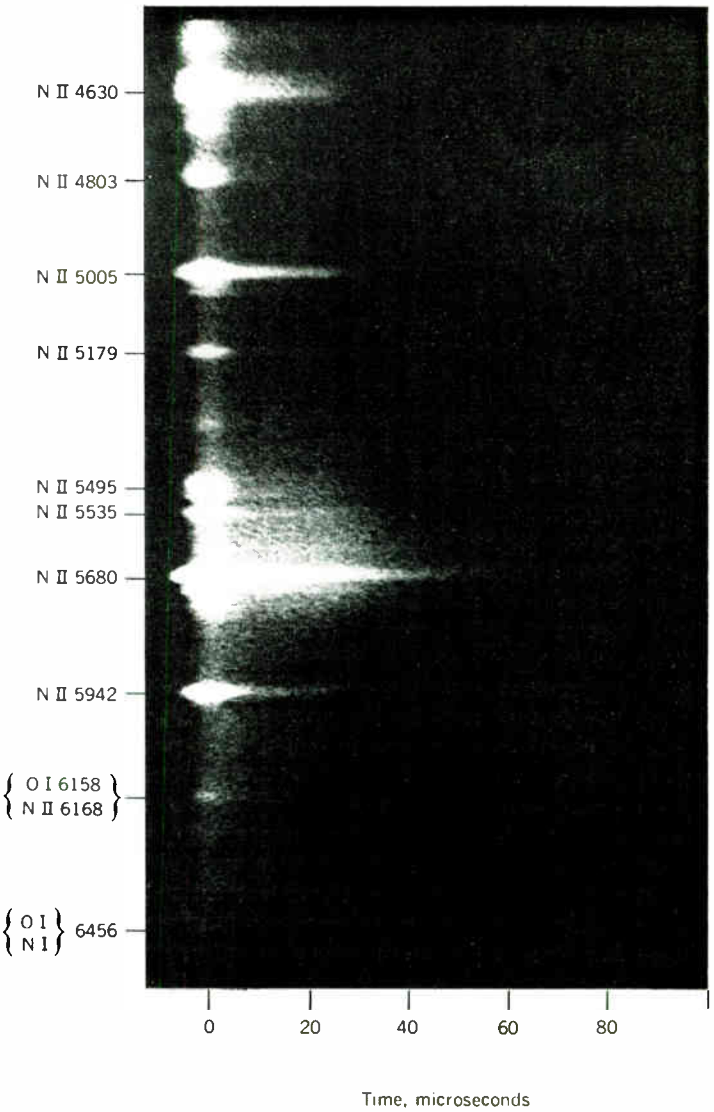
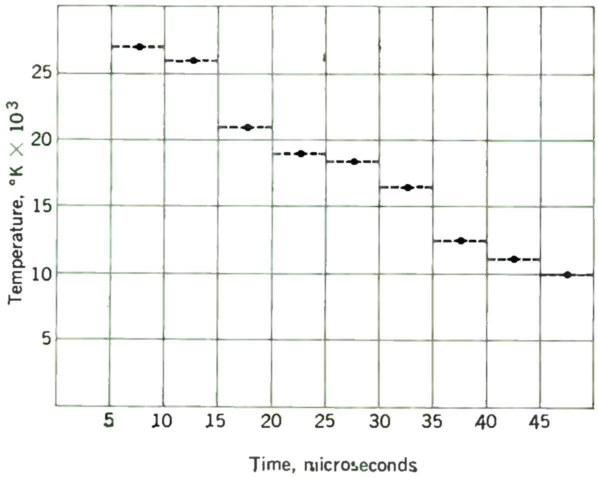


Fig. 6 (right). Time-resolved spectrum of about 10 meters of a lightning return-stroke channel.[8] (Courtesy Institute of Atmospheric Physics, University of Arizona, 1965.)

Fig. 7. Lightning return-stroke temperature vs. time.[9] The dashed lines indicate the effective time over which radiation was integrated on one position on the film—that is, the time resolution of the spectrometer.



dynamic properties of air[28,29] to determine the number densities of all the constituents of the channel as a function of time.

Channel properties computed from time-integrated spectra using the Saha equation must be considered unreliable. Those properties computed from time-resolved spectra using the Saha equation would be most valuable, but unfortunately the weak intensity of the neutral nitrogen and oxygen lines on the existing time-resolved spectra makes this calculation impossible. The weak intensity of the neutral lines in the time-resolved spectra also precludes the possibility of making a quantitative estimate of the errors incurred in the analysis of the time-integrated spectra. It is, however, apparent that the neutrals radiate for a far longer time than do the ions.

It is to be expected that, in the near future, lightning spectra will be obtained with a time resolution of about 1 μs and a wavelength resolution of about 5 Å, enabling both a more detailed temperature determination and the simultaneous measurement of electron density via the Stark effect. Hopefully, spectra will be obtained from close strokes wth relatively strong neutral nitrogen and neutral oxygen line intensities, so that the Saha equation approach may be employed to calculate channel properties on a microsecond time scale. If the electron densities calculated by use of the Stark effect and by use of the Saha equation are in good agreement, the lightning properties determined from lightning spectroscopy may be considered very reliable.

Revised text of a paper presented at the IEEE Region 6 Annual Conference, Tucson, Ariz., April 26–28, 1966. The work described was sponsored in part by the Geophysics Branch of the Office of Naval Research and by the National Science Foundation.

The author expresses his thanks to L. E. Salanave, E. P. Krider, R. E. Orville, and L. B. Loeb for their helpful comments regarding this article.

REFERENCES

1. Salanave, L. E., "The optical spectrum of lightning," *Science*, vol. 34, pp. 1395–1399, Nov. 1961.

2. Salanave, L. E., Orville, R. E., and Richards, C. N., "Slitless spectra of lightning in the region from 3850 to 6900 angstroms," *J. Geophys. Res.*, vol. 67, pp. 1877–1884, May 1962.

3. Salanave, L. E., "The optical spectrum of lightning," *Advan. Geophys.*, vol. 10, pp. 83–97, 1964.

4. Meinel, A. B., and Salanave, L. E., "N_2^+ emission in lightning," *J. Atmospheric Sci.*, vol. 21, pp. 157–160, Mar. 1964.

5. Salanave, L. E., "The photographic spectrum of lightning; determination of channel temperature from slitless spectra," in *Problems of Atmospheric and Space Electricity*, S. C. Coroniti, ed. New York: Elsevier, 1965, pp. 371–383.

6. Krider, E. P., "Time-resolved spectral emissions from individual return strokes in lightning discharges," *J. Geophys. Res.*, vol. 70, pp. 2459–2460, May 1965.

7. Krider, E. P., "The design and testing of a photoelectric photometer for selected lines in the spectrum of lightning," M.S. thesis, University of Arizona, Tucson, 1965.

8. Orville, R. E., "Time-resolved slitless spectrum of a lightning stroke," *Science*, vol. 151, pp. 451–452, Jan. 1966.

9. Orville, R. E., "High-speed time-resolved slitless spectra of lightning strokes," presented at the 46th Ann. Meeting, American Meteorological Society, Denver, Colo., Jan. 24–27, 1966.

10. Loeb, L. B., "The positive streamer spark in air in relation to the lightning stroke," in *Atmospheric Explorations*, H. G. Houghton, ed. New York: Wiley, 1958, pp. 46–57.

11. Loeb, L. B., "Ionizing waves of potential gradient," *Science*, vol. 148, pp. 1417–1426, July 1965.

12. Malan, D. J., *Physics of Lightning*. London: English Universities Press Ltd., 1963, pp. 24–26.

13. Drabkina, D. I., "The theory of the development of the spark channel," *J. Exptl. Theoret. Phys. (USSR)*, vol. 21, pp. 473–483, Apr. 1951; English transl., AERE Lib/Trans. 621, Harwell, Berkshire, England.

14. Braginskii, S. I., "Theory of the development of a spark channel," *Sov. Phys. JETP (English Transl.)*, vol. 34, pp. 1068–1074, 1958.

15. Flowers, J. W., "The channel of the spark discharge," *Phys. Rev.*, vol. 64, pp. 225–235, Oct. 1943.

16. Flowers, J. W., "Lightning—measuring lightning current photographically, equipment procedure and examples of results," *Gen. Elec. Rev.*, vol. 47, pp. 9–15, Apr. 1944.

17. Higham, J. B., and Meek, J. M., "The expansion of gaseous spark channels," *Proc. Phys. Soc. (London)*, Sect. B, vol. 63, pp. 649–661, Sept. 1950.

18. Higham, J. B., and Meek, J. M., "Voltage gradients in long gaseous spark channels," *Proc. Phys. Soc. (London)*, Sect. B, vol. 63, pp. 633–648, Sept. 1950.

19. Andreev, S. I., Vanykov, M. P., and Kotolov, A. B., "Growth of the spark discharge canal for a discharge circuit with a rapidly increasing current," *Soviet Phys.—Tech. Phys. (English Transl.)*, vol. 7, pp. 37–40, July 1962.

20. Norinder, H., and Karsten, O., "Experimental investigations of resistance and power within artificial lightning current paths," *Arkiv. Mat.*, vol. 36, pp. 1–48, Jan. 1948.

21. Somerville, J. M., *The Electric Arc*. New York: Wiley, 1959, pp. 101–119.

22. Berger, K., "Front duration and current steepness of lightning strokes to the earth," in *Gas Discharges and the Electricity Supply Industry*, J. S. Forrest, ed. London: Butterworth, 1962, pp. 63–73.

23. Berger, K., "Messungen und Resultate der Blitzforschung der Jahre 1955–1963 auf dem Monte San Salvatore," *Bull. SEV*, vol. 56, no. 1, pp. 2–22, 1965.

24. Uman, M. A., "The conductivity of lightning," *J. Atmospheric Terrest. Phys.*, vol. 26, pp. 1215–1219, Dec. 1964.

25. Uman, M. A., and Orville, R. E., "The opacity of lightning," *J. Geophys. Res.*, vol. 70, pp. 5491–5497, Nov. 1965.

26. Zhivlyuk, Yu N., and Mandel'shtam, S. L., "On the temperature of lightning and the force of thunder," *Soviet Phys. JETP (English Transl.)*, vol. 13, pp. 338–340, Aug. 1961.

27. Uman, M. A., paper to be published.

28. Gilmore, F. R., "Equilibrium composition and thermodynamic properties of air to 24 000°K," Rand Corp. Res. Memo. RM-1543 (Defense Documentation Center, AD-84502, unclassified), Aug. 1955.

29. See, for example, Hilsenrath, J., and Klein, M., "Tables of thermodynamic properties of air in chemical equilibrium including second virial corrections from 1500°K to 15 000°K," AEDC-TR-65-58 (Defense Documentation Center, AD-612301, unclassified), Mar. 1965. This report contains references to all earlier significant work on the thermodynamic properties of air.

30. Griem, H. R., *Plasma Spectroscopy*. New York: McGraw-Hill, 1964, pp. 445–529.

31. Wiese, W. L., "Line broadening," in *Plasma Diagnostic Techniques*, R. H. Huddlestone and S. L. Leonard, eds. New York: Academic Press, 1965, pp. 265–317.

32. Prueitt, M. L., "The excitation temperature of lightning," *J. Geophys. Res.*, vol. 68, pp. 803–811, Feb. 1963.

33. Mastrup, F., and Wiese, W. L., "Experimentelle Bestimmung der Oszillatorenstarken eineger N II und O II-Linien," *Z. Astrophys.*, vol. 44, pp. 259–279, 1958.

34. Uman, M. A., "The peak temperature of lightning," *J. Atmospheric Terrest. Phys.*, vol. 26, pp. 123–128, Jan. 1964.

35. Uman, M. A., Orville, R. E., and Salanave, L. E., "The density, pressure, and particle distribution in a lightning stroke near peak temperature," *J. Atmospheric Sci.*, vol. 21, pp. 306–310, May 1964.

36. Uman, M. A., Orville, R. E., and Salanave, L. E., "The mass density, pressure, and electron density in three lightning strokes near peak temperature," *J. Geophys. Res.*, vol. 69, pp. 5423–5424, Dec. 1964.

37. Uman, M. A., and Orville, R. E., "Electron density measurement in lightning from Stark-broadening of H_α," *J. Geophys. Res.*, vol. 69, pp. 5151–5154, Dec. 1964.

38. Drellishak, K. S., "Partition functions and thermodynamic properties of high temperature gases," AEDC-TDR-64-22 (Defense Documentation Center, AD-428210, unclassified), Jan. 1964.

39. Wallace, L., "Note on the spectrum of lightning in the region 3670 to 4280 Å," *J. Geophys. Res.*, vol. 65, pp. 1211–1214, Apr. 1960.

40. Orville, R. E., and Uman, M. A., "The optical continuum of lightning," *J. Geophys. Res.*, vol. 70, pp. 279–282, Jan. 1965.

41. Krider, E. P., "Some photoelectric observations of lightning," *J. Geophys. Res*, vol. 71, 1966, to be published.

High-strength conductors for supermagnets

Magnets have been developed that produce 225 kG in continuous operation, and are designed to reach 300 kG. However, these high fields also entail high stresses and temperatures, which means that the magnet structure must be porous as well as strong

D. Bruce Montgomery *M.I.T. National Magnet Laboratory*

With the introduction of powerful magnets such as the National Magnet Laboratory's 225-kG device that recently went into continuous operation, it has become necessary to consider the structural problems that arise because of the high stresses—and high temperatures—involved in generating these high fields. Primarily, a magnet material must be provided that possesses strength as well as porosity. Various combinations of ETP copper, steel, beryllium-copper, and zirconium-copper have proved effective. Also, it is pointed out that stresses can be reduced by proper arrangement of the coils, and that there is an advantage to using high-strength materials in stress-limited magnets despite their increased resistivity.

Magnets now in operation at the Massachusetts Institute of Technology's National Magnet Laboratory, which is sponsored by the Air Force Office of Scientific Research, range in field strength from 50 kG (kilogauss) to 225 kG (Table I). When the 225-kG device began functioning recently, it became the world's most powerful continuously operating magnet, surpassing an NML magnet that produces 205 kG in its basic configuration. The new magnet, which is constructed of steel-supported zirconium-copper, will soon be pushed to still higher fields. In its ultimate form it will reach 300 kG and withstand shear stresses of 93 000 psi in its innermost coils by current densities that average 50 000 A/cm².

I. Magnets in operation or under design at the National Magnet Laboratory

Field Strength, kG	Number of Magnets	Bore, inches	Transverse Access, inches
	In Operation		
100	1	4	0
80	1	4	1⅛
100	3	2⅛	variable
110	4	2⅛	0
50	1	14	0
65 (high homogeneity)	1	4	0
150	2	2⅛	0
175	2	1¼	0
205*	1	2⅛	0
225	1	1¼	0
	In Design		
85	1	6	0
120	1	6	0
150	1	4	0
200	1	1¼	0
250	1	2⅛	0
300	1	1¼	0

* This magnet has produced 255 kG in a reduced volume by the inclusion of iron pole tips. Its inner coil will be replaced by one allowing a field of 225 kG.

To generate really high fields, high current must be forced through a conductor of small cross section, thus creating typical current densities of about 50 000 A/cm². This action produces heat and large mechanical forces, brought about by the Lorentz-force interaction between the very high current densities involved and the resultant magnetic field; see Fig. 1.

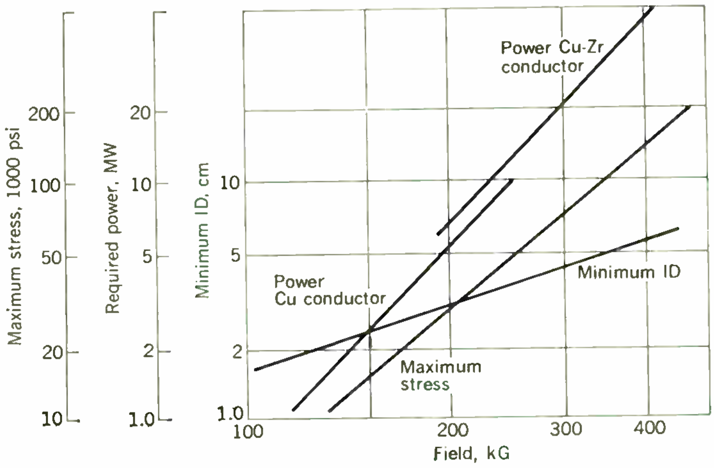


Fig. 1. Shear stress present in a continuous-duty, water-cooled magnet as related to the intensity of the generated field. Zr-Cu is the most practical conductor for magnets of about 250-kG field strengths and higher. It increases strength substantially but adds little to resistivity.

The need to withstand these stresses, and at the same time to remove significant amounts of heat, results in a design conflict—to provide a structure that is porous enough to allow room for cooling channels and electrical insulation and yet is strong enough to resist very severe stresses. The conflict is resolved by using structures with relatively high strength for their porosity, with much of the strength derived from the materials used, and by reducing stresses at the expense of efficiency through controlled reduction of current density in the innermost windings.

The basic structure used in the very-high-field magnets is shown in Fig. 2. The windings are built as a stack of conducting disks with insulators arranged to form a flat helix. The conducting disks are usually of full-hard electrolytic-tough-pitch (ETP) copper into which axial holes or radial slots are etched to serve as cooling passages. Interturn insulation is provided by insulated steel disks because the extreme forces would crush ordinary impregnated fiberglass insulators. As described later, the steel also furnishes additional support for the conductors if sufficient preload axial clamping is applied to the stack. When more strength is needed than is afforded by ETP copper conductors and steel, Zr-Cu is used.

The exact stresses in the structure are difficult to analyze. The Lorentz force varies greatly in the plate because of the variation in field and the nonuniform current distribution. The stress concentrations that occur because of the cooling holes and slits are also difficult to determine accurately. An effective way of evaluating the relative merits of various composite materials is to compress axially small composite stacks of disks scaled from the designs used in the magnets, and then to measure the axial and radial deflection of the stacks.

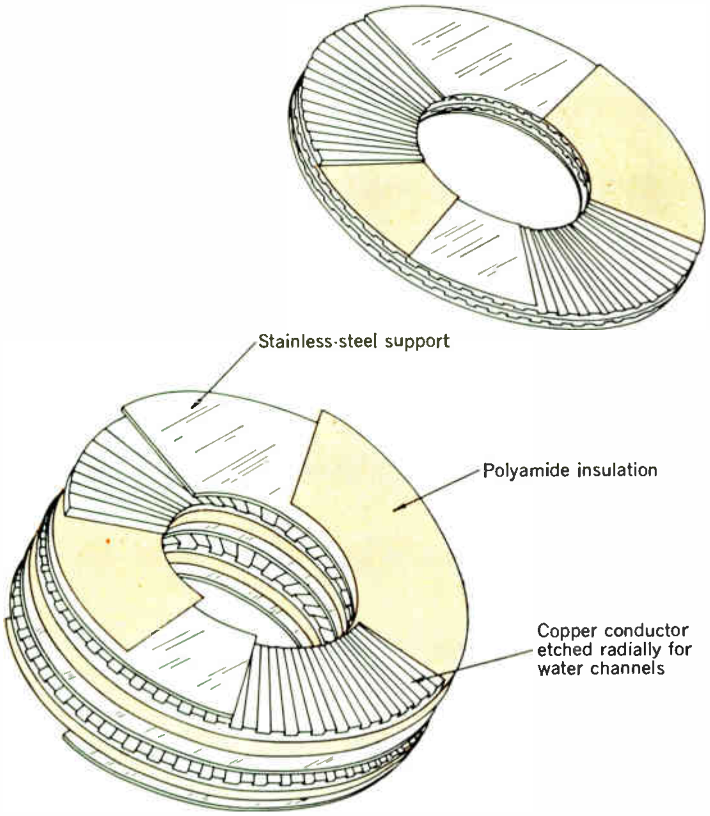


Fig. 2. Basic structure of highest-field continuously operating magnets at NML. Disks, slotted or drilled to pass cooling water, are stacked, with alternating disks of Amzirc copper and insulated steel. They are then compressed to form a conducting and an insulating interleaving helice.

The local maximum stress at the first sign of major yielding for various materials and composites is given in Table II. The table clearly shows that the addition of glass-cloth insulators reduces the strength of the stack, whereas steel insulating disks improve it. The last column in the table indicates the effective conductivity of the composite materials considered. If the stack is all ETP copper it is 100 percent conductive; if it were half ETP copper and half steel it would be only 50 percent conductive. An all-beryllium-copper stack with no steel would also be 50 percent conductive because the conductivity of Be-Cu is 50 percent of that of ETP copper.

Many variations are possible. Table II shows that the failure point of full-hard copper can be increased from 50 000 psi to 82 000 psi by using 180 000-psi steel plates for one quarter of the stack height. The failure point of a Zr-Cu stack can be increased to 120 000 psi by making one fourth of its length of a still-higher-strength 275 000-psi steel. This compares favorably with a 120 000-psi all-Be-Cu stack but requires only 75 percent as much power, because the higher effective conductivity more than compensates for the volume lost to the steel.

A 205-kG magnet can be used to illustrate some of the principles described. This magnet is built in three concentric sections: two inner sections patterned after Fig. 2 and an outer section of more conventional hollow-conductor design. The three solenoids have 2.625-inch, 6.5-inch, and 14-inch inner diameters, and generate, respectively, 73 kG, 82 kG, and 50 kG, with a total power input of 10.5 MW. The inner coil takes the largest stress. The present inner stack consists of 81.5 percent full-hard ETP copper, 12.5 percent 180 000-psi-yield stainless steel, and 6 percent polyanamid insulation

II. Failure points for various materials

Material	Stress at Failure, psi	Effective Conductivity, percent*
Soft copper	10 000	100
One-half-hard copper	27 000	100
Full-hard ETP copper	50 000	100
Zr-Cu	80 000	89.5
Be-Cu, 0.5 percent	120 000	50
Be-Cu, 2 percent	200 000	22.5
Composites:		
FHC and 14 percent glass cloth and epoxy	30–40 000	87
FHC and 19 percent 180 000-psi steel	73 000	81
FHC and 24 percent 180 000-psi steel	82 000	76
Zr-Cu and 20 percent 180 000-psi steel	100 000	71.5
Zr-Cu and 22 percent 275 000-psi steel	120 000	69.5

* The effective conductivity is an expression of the conductivity of the material diminished by the volume lost to noncurrent-carrying members; i.e., support steel or insulation.

(Dupont Pyre-ML). The maximum tensile stress at 205 kG is estimated at 50 000 psi, and the strength, based on tests that produced the data in Table II, is 67 000 psi.

If the present magnet were to be operated at 250 kG, the yield stress would reach 70 000 psi and there would be catastrophic yielding. At the 250-kG level the magnet would consume 14.5 MW of power, would have a current density of 50 000 A/cm^2 and a local power dissipation of 5 kW/cm^3, and would require 2000 gallons per minute of cooling water. To enable it to generate more than 205 kG, the ETP magnet was modified by the use of a new inner coil of Zr-Cu plates and by the addition of very-high-strength steel. This stack, which uses Amzirc* copper alloy, has a calculated yield strength of 93 000 psi.

Zirconium-copper, although necessary for strength, is not as conductive as ETP copper, and additional power is required. However, the use of a material of higher strength but also of higher resistivity does not always result in a need for more power, as we shall see when we consider reducing stress by redistributing currents.

How stresses can be reduced

Calculation of exact stress distribution is complicated, and often is not justified in view of uncertainties as to the actual yield point of the conductors. An approximate formula assumes that the distributed Lorentz forces are actually all operating on the inner bore of the winding in the manner of a high-pressure fluid. This equivalent "magnetic pressure," P_m, can be calculated from energy considerations as

$$P_m = \frac{B^2}{2\mu} = H^2\, 0.577 \text{ psi}$$

where H is field strength in kilogauss and μ is permeability. We now can consider the winding as a thick-walled cylinder pressurized with fluid inside.

The foregoing formula for magnetic pressure is based on an infinitely long solenoid. For solenoids of finite length, the current density must be higher for a given field by a factor γ; for a magnet with the type and shape of inner section that we have been discussing, this is typically about 1.4.

When a coil is to be divided into several concentric coils, it is essential to consider the differential magnetic pressure rather than the entire magnetic pressure. The pressure acting on a given element generating a field ΔH but exposed to a central field H is

$$\Delta P = \frac{H^2}{2\mu} - \frac{(H - \Delta H)^2}{2\mu}$$

$$\Delta P = \frac{2H\Delta H - \Delta H^2}{2\mu}$$

Having found the differential pressure, the maximum average shear stress T_m, which occurs at the inner lip, can be found from formulas for a thick-walled pressure vessel:

$$T_m = \Delta P \left(\frac{\alpha^2}{\alpha^2 - 1}\right) \gamma$$

therefore,

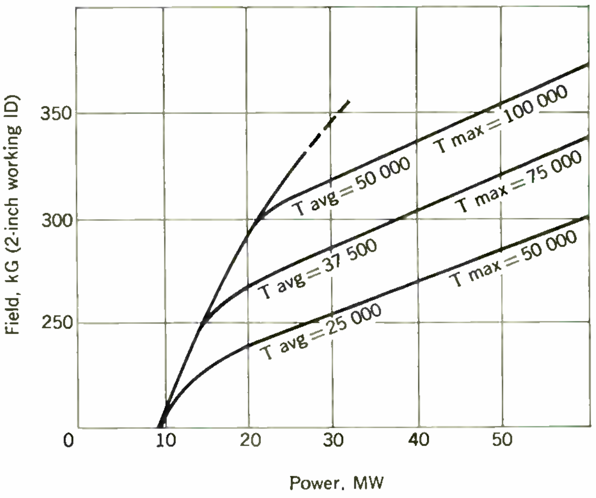
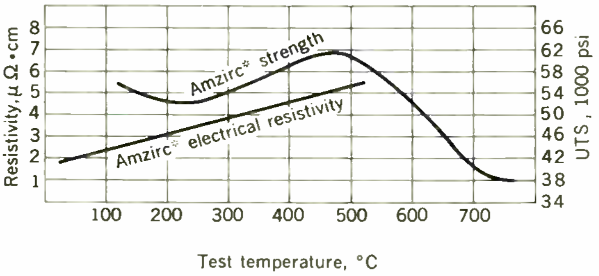


Fig. 3. Stresses associated with various magnetic fields in composite magnets as plotted at the National Magnet Laboratory. The graphs indicate that only a copper alloy such as a zirconium alloy will be able to survive the shear stresses imposed above 250 kG while keeping power consumption to reasonable levels. The maximum stresses, occurring at the innermost edges of the magnet, determine the failure point. Plot is shown to 400 kG.

Fig. 4. Zirconium-copper retains its strength well at elevated temperatures whereas electrical resistivity rises linearly with temperature, retaining 90 percent of the conductivity of pure copper throughout the range shown. Resistivity curve is for 0.016-inch Amzirc wire previously solution-annealed at 900°C for one hour and quenched, cold-drawn 75 percent, and aged 400°C one hour. Tensile strength curve is for 0.081-inch Amzirc wire previously solution-annealed at 925°C for 30 minutes, quenched and cold-worked 37 percent. (Data from American Metal Climax, Inc., Carteret, N.J.)



$$T_m = \gamma \left(\frac{\alpha^2}{\alpha^2 - 1}\right) [2H\Delta H - \Delta H^2]\, 0.577 \qquad (1)$$

where α is the ratio of the cylinder's outer to inner diameter and H is in kilogauss. Equation (1) indicates that for any given H, T_m can be reduced arbitrarily by limiting the ΔH across the element. However, any field not generated by the element in question must be generated by elements of greater radius, which will require greater power consumption.

To solve (1) for the allowable ΔH at a given H (assuming the coil to be stress limited), (1) must be inverted and the following quadratic equation solved:

$$\Delta H = H - \sqrt{H^2 - C}$$

---

*Registered trademark of American Metal Climax, Inc., for its Zr-Cu alloy.

where

$$C = \frac{T_m}{\gamma\left(\frac{\alpha^2}{\alpha^2 - 1}\right) 0.575} \qquad (2)$$

The power required to generate ΔH in any element is

$$W_n = \frac{\Delta H_n{}^2 a_n{}^2}{G_n{}^2} \frac{\rho_n}{\lambda_n} \qquad (3)$$

where a_n is the inner radius of the element, ρ_n is the resistivity, and G_n is 0.175 for our design.

In the 205-kG magnet, the inner section is exposed to a total field of 205 kG while it itself is generating 73 kG. For this coil the average stress at the inner diameter is 25 000 psi and the maximum stress is 50 000 psi. If this were the maximum permissible level, C in Eq. (2) would be 2.4×10^4, and increasing the field to 250 kG, for example, would require reducing ΔH from 73 kG to 59 kG to avoid failure. The next sections then would have to generate not only the additional 45 kG to reach 250 kG from 205 kG but also the 14 kG no longer produced by the inner section. The field interior to the second section is now 50 kG less than the central field, or 191 kG. Assuming that the second element is geometrically similar to the inner one, all constants are the same and ΔH can be calculated from Eq. (2). The value is 83 kG, which means that 108 kG still must be generated from one or more elements beyond the second. If the third element is geometrically similar to the other two, it will be self-supporting because the roots of (2) are imaginary; i.e., $H_3{}^2 > C$ at $H = 108$ kG. Having found the allowed ΔH for each element, the power for each can be found from (3) and the total power determined. The problem can then be solved again for a higher or lower allowable stress by changing C in (2).

This analysis has been carried out for average stress levels of 25 000, 37 000, and 50 000 psi (1, 1.5, and 2 times the present operating stress level). Since local stress, under slits and near cooling passages, is about twice the average stress because of the smaller current-carrying cross section at those points, it does not seem feasible to support average stresses beyond 50 000 psi, or double the present average stress level. The results of the analysis, up to 400 kG, are shown in Fig. 3. It is clear that enormous amounts of power would be required to reach the highest fields shown.

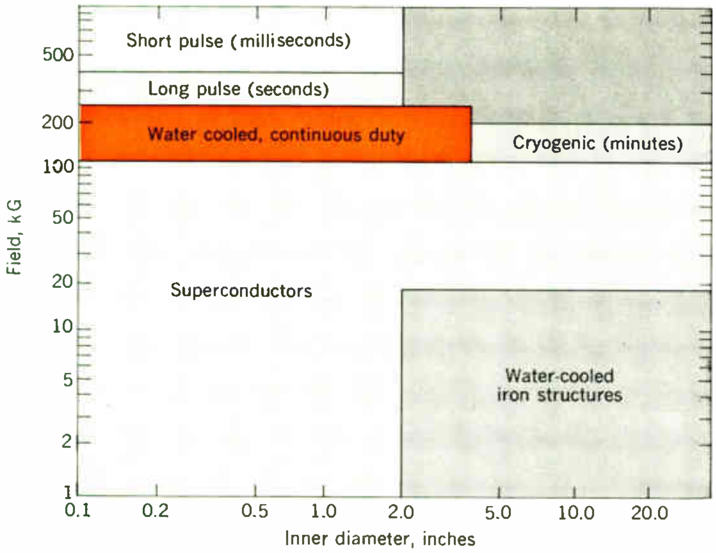


Fig. 5. Continuous-duty, water-cooled magnets discussed in this article as related to other types of high-field magnets. Ranges of fields and working areas that are now practical with various types of magnets are shown.

Importance of strong conductors

We can further use Fig. 3 to illustrate the effect of trading strength for conductivity. If higher-strength materials are used for the inner coils, the higher resistivity of those materials may be more than compensated for by the fact that we need not limit their field and thus have to generate so much field with the outer sections.

For example, if copper limited operations to a stress level of 25 000 psi and a field of 300 kG were to be generated, 60 MW of power would be required. If zirconium-copper such as Amzirc (Fig. 4) were used, which has an allowed stress level 50 percent higher than ETP copper but only a 10 percent higher resistivity, 300 kG could be generated with 41 MW for a saving in power of one third. (The 50 percent 37 000-psi allowable average stress contour calls for 37.5 MW at 300 kG for copper, and this must be increased by 10 percent to make up for the increased resistivity of the Zr-Cu; thus the net power is 41 MW.) Because Zr-Cu need not be used in all elements as the outer section does not require higher strength, an even greater saving is possible. If beryllium-copper, which has twice the allowable stress of ETP copper but double the resistivity, were used, the power required would be 21 × 2 MW or 42 MW, even if Be-Cu were used in all sections. Thus there is a real advantage in using high-strength materials in stress-limited magnets despite their increased resistivity.

This analysis, however, neglects the important fact that if Be-Cu were used in the inner element, the power required in that element to generate the allowable field would undoubtedly be more than could be dissipated at a reasonable temperature. Nevertheless, it does illustrate the trend to higher efficiency for higher-strength materials in stress-limited magnets.

Future targets

Our future magnets will utilize all the techniques discussed: use of zirconium-copper, of stronger steels, of larger ratios of steel to conductor, and of concepts of stress reduction. It is hoped that fields beyond 300 kG can be produced (Fig. 5), although there is not enough power at the laboratory at present to generate such fields continuously. However, by using the overload capacity of the power-supply system, the advanced magnets can be tested and some experiments performed during pulses that are several seconds in duration. Present dc power for the laboratory's magnets is supplied by two motor–generator sets; each consists of two 2.5-MW generators, a 6000-hp synchronous motor, and an 84-ton flywheel. The total capacity is 10 MW continuously, 16 MW for 1 minute, or 32 MW for 5 seconds, with the 32 MW supplied by energy stored in the flywheels.

The author acknowledges the contributions of Mat Leupold and Carl Weggel to many aspects of the high-field work at M.I.T.

Special Conference Report

Report on the A. S. Popov Society meeting

Norman Abramson *University of Hawaii*

The A. S. Popov Society plays a role in the U.S.S.R. somewhat analogous to that of the IEEE in the United States. Certain differences however were apparent at the 22nd annual meeting of the Popov Society, held from May 11 to May 14 in Moscow. Perhaps the most notable of these differences is in the coverage of the two organizations. The IEEE includes in its technical coverage a wide variety of material—ranging from the information sciences through computers, electron devices, and power to industry and general applications, engineering management, writing, and speech. The Popov Society restricts its attention to radio electronics and certain directly related fields such as signal theory and quantum electronics.

Another difference in these two organizations is that the Popov Society is a technical and scientific society of the U.S.S.R. while the IEEE is non-national in character. In response to a special invitation from Dr. V. I. Siforov, President of the Popov Society, a nine-member United States delegation of the non-national IEEE, headed by Dr. William G. Shepherd, President of the IEEE, attended this meeting as official representatives of the IEEE. Other members of the delegation were Prof. Norman Abramson (University of Hawaii), Dr. Louis R. Bloom (General Telephone and Electronics Laboratories), Prof. Robert J. Collins (University of Minnesota), Dr. Walter E. Morrow, Jr. (M.I.T. Lincoln Laboratory), Prof. Calvin F. Quate (Stanford University), Dr. Herbert Sherman (M.I.T. Lincoln Laboratory), Prof. Samuel Silver (University of California, Berkeley), and Prof. M. E. Van Valkenburg (University of Illinois). IEEE members from other countries also attended.

The conference opened in the main auditorium of the U.S.S.R. Army Museum with a plenary session at 10:00 A.M. on May 11. After a welcoming address by Dr. Siforov, three papers were presented at this session; several translators were provided for the American delegation. The first paper was "Quantum Radio Waves and Their Applications" by M. E. Zhabotinskii, A. A. Malenkov, and V. B. Steinshleiger. This consisted of a review of Soviet work in quantum amplifiers, primarily with applications in recent radar astronomy measurements. The second paper, "Prospects for the Application of Optical Generators in Television and Holography," was given by A. L. Mikaelyan. In this paper some possibilities for the application of holographic techniques were discussed. An experimental system for employing coherent light to enhance television brightness was mentioned as was the possibility of three-dimensional television by the use of a system of transparent and reflecting screens. The final paper of the plenary session was "Distant TV Transmission Using the Satellite, Molniya-I" by A. D. Fortushenko. After a brief description of some of the characteristics of the communication system employed (a solid-state system except for a 40-watt-output traveling-wave tube), an extensive description of tests on American, French, and U.S.S.R. color television systems was given. Dr. Fortushenko stressed the economic advantages of using satellite transmission for relaying television programs to distant portions of the Soviet Union.

Following this opening session, approximately 70 more-specialized sessions were held during the afternoon of May 11, the morning and afternoon of May 12, and the morning and afternoon of May 13. These parallel sessions, which featured some 300 papers, were divided into 19 separate technical areas indicating the interests of the Popov Society: information theory; propagation of radio waves; general radio technology; waveguide instruments; radio measurements; radiobroadcasting, electroacoustics, and sound recording; semiconductor devices; telemechanics; television; cybernetics; computer technology; antenna devices; radio transmitting devices; microelectronics; electronics; radio receiving devices and amplifiers; theory and technique of transmitting discrete signals; wire communications; and quantum electronics.

In addition to the papers listed on the program a number of papers were invited from the United States. Translators—usually well acquainted with the area under discussion—were provided for English-speaking authors. Considerable information was exchanged during the question period following each paper.

Because all sessions were conducted in Russian so that, in spite of excellent translation services, prolonged attendance by foreign visitors at these sessions was difficult, several visits to laboratories and technical institutes in the Moscow area were arranged for members of foreign delegations. Among the most interesting of the visits were those to the Moscow Computing Center of the Academy of Sciences of the U.S.S.R., the Moscow State University, and the Institute of Radio Engineering and Electronics of the Academy of Sciences of the U.S.S.R.

At the Moscow Computing Center visitors were shown the BESM 3M, a transistorized machine that performs 20 000 operations per second. It utilizes a 4000-word (12 bits per word) 10-microsecond core memory together with four drums of 16 000-word capacity each.

At Moscow State University the graduate program of the physics faculty, comprising 32 chairs and approximately 3000 students, was outlined. The organization of the radio physics portion of this faculty, comprising six chairs and perhaps 1000 students, was described in some detail. The number of students working in these fields is remarkable; when the inevitable comparisons are made, it seems clear that no single academic institution in the United States can match the quantity and scope of work in progress at Moscow State University. Particularly impressive to members of the American delegation was the high quality of the equipment available to students.

The Director of the Institute of Radio Engineering and Electronics is Academician V. A. Kotelnikov, a Fellow of the IEEE who is well known in the United States for his work in information theory. At the institute, the group of foreign visitors was addressed by Professor Zornov, assistant director of the institute and a corresponding member of the Academy of Sciences of the U.S.S.R. Work in progress includes radio astronomy, observations of planets, tropospheric propagation, quasi-optical systems, information transmission, channel capacity, randomly varying channels, coding, radio properties of plasmas, quantum electronics, and thin films. During our tours of the various individual laboratories, we were shown equipment used for measuring dielectric properties of materials at optical frequencies as well as some experimental millimeter and submillimeter detectors made from indium antimonide.

On the morning of May 14 a plenary session was held at which two papers were presented: "Application of Semiconductor Devices in the Physiological Measurement Apparatus of Space Vehicles" by I. T. Akulinchev and I. I. Popov, and "Some Problems of Microelectronics" by V. I. Stafeyev. Following this session a reception was given for foreign delegates at Moscow's House of Friendship. It featured addresses by Professor Siforov and the heads of each delegation, including Dr. Shepherd.

Although the Popov Society meeting (in common with the IEEE International Convention) has some of the disadvantages of a large technical meeting, the American delegates feel that attendance at this event is a valuable experience. It is probably true that English-speaking engineers and scientists are not as familiar with work published in other languages as they should be, whereas discussions at the Popov Society showed that Soviet engineers and scientists are generally aware of important work in their fields in the United States.

An important impression carried away from the Popov Society meeting by the IEEE U.S. delegation is of their hosts' hospitality and willingness to exchange information. This opportunity to hear of work and to establish contacts in the Soviet Union will undoubtedly result in increased communication and improved understanding in both technical and nontechnical matters.

Address by Dr. Shepherd

President Siforov, honored guests, and members of the Popov Society, let me begin by expressing appreciation on behalf of our delegation for the opportunity to join with your distinguished society on the occasion of your Twenty-Second Annual Meeting. It is my privilege and honor to bring you greetings of The Institute of Electrical and Electronics Engineers. Although the Institute is a non-national society with members all over the world, the delegation representing it consists of nine Americans. In visiting you, we are carrying forward an exchange of delegations that was first established between the Popov Society and the Institute of Radio Engineers in 1958. Your delegation visited the Annual Convention of IEEE in New York City in March of this year and it was my pleasure to introduce President Siforov at the Awards Banquet. It is a privilege for me personally to be able to return that visit and to be able to address this plenary session of your society.

These exchanges constitute an important aspect of the scientific and technological communications between us. The publications of our societies inform our memberships of new developments, but opportunities to become personally acquainted, to exchange views at first hand, to visit laboratories, are important not only within nations but across national boundaries. It is our hope that the relationship so well begun between IEEE and the Popov Society will not only continue and grow but will set an example in other areas of common endeavor.

It is a well-known truth that science and technology know no national boundaries. History has many examples of this and indeed our two societies are living examples through the work based on the researches of Popov, Marconi, and Hertz. Because of the scientific tradition of the sharing of the results of research in the literature, the basis for new advances is available to all. It is not surprising to scientific workers that new discoveries and inventions may be made virtually simultaneously by investigators in widely separated parts of the world. A recent example of this, which is a source of pleasure to Soviet and American citizens, was the work of Prokhorov and Basov in your country and of Townes in the United States that led to the award to them jointly of the Nobel Prize for Physics.

The work of Popov, Marconi, and Hertz changed the nature of the world by opening avenues of instant communication. It is too early for us to foresee all the implications of the work of Basov, Prokhorov, and Townes. But one thing we can say with assurance is that the time span for the development of practical applications of this recent work will be much less than that which was required for the development of radio communication. It was many decades after the initial work on radio before its impact was felt by society at large. The obvious reason was that the necessary technology and the manufacturing facilities needed for its implementation had to be created. By contrast it was approximately five years between the first ammonia maser and its initial use in a developmental communications system. It was less than one year between the first successful demonstration of the laser and its application as a surgical tool.

The shortening time span between fundamental discoveries and their applications is a consequence of the sophisticated technological base which the world now takes almost for granted. I have chosen only two examples familiar to the membership of our two societies but many others will come readily to your minds. The point is that the coupling between basic science and its application to the needs of society has become tight and immediate. This coupling has some important implications for our educational institutions and for the role that professional societies play in serving their memberships.

I should like to speak of the educational problems as I see them from the viewpoint of an educator associated with an American Land Grant University. The Land Grant Universities of the United States played an important role in the development of the economy of the country. These universities were established over a century ago by an act of the Congress of the United States. At that time the United States was expanding westward into an undeveloped wilderness. Its government was seeking means to aid in the development of its economy. As the title "Land Grant University" implies, these institutions were given grants of land by the government. The revenues from these lands were to be used to support institutions of higher learning in which curricula concerned with agriculture and the mechanic arts were to receive special emphasis. These universities represented a major philosophical departure from the traditions of earlier universities in the United States. The new universities were given the specific mission of carrying on research and education in agriculture and engineering to aid in the development of the economy of the regions in which they were located. The agricultural abundance of the United States is largely attributable to the efforts of the scientists and technicians of these Land Grant institutions.

It is worthy of comment that these universities did not confine their activities to the specific missions assigned to them. They have developed as large and complex institutions concerned with the broad spectrum of disciplines characteristic of the great universities of the world. Their rise to prominence in American education influenced the patterns of development of the earlier and more traditionally oriented American universities as well as universities in other parts of the world. The Land Grant Universities demonstrated that universities could serve as an important force for the economic development of a country in addition to fulfilling their traditional role as the custodians of man's cultural inheritance.

These thoughts have relevance to the problem of the rapid translation of scientific discoveries into devices and systems that serve society's needs. The techniques developed in Land Grant Universities proved successful in transferring advances in agricultural science to the farmers. Now we ask whether they can be applied to the transfer from laboratory to practice in other areas of university research. There are distinct differences now existing in the situations of academic agriculture relative to the other academic sciences and technology.

The traditional universities of the past held themselves apart from society and literally lived within walls. The great philosophical departure of the Land Grant Universities from tradition was that the agricultural scientists of these universities moved out into the communities and literally worked with the farmers in the fields in demonstrating research findings. This close working relationship between the farmer and the agricultural scientist stimulated an awareness by the scientist of the problems of the farmer and an appreciation by the farmer that he could benefit from science.

In other areas of science the interaction between the academician and the practitioner has been less close so that the translation of basic science into society's service has been less direct. Those responsible for the education and research training of engineers have too little contact with the technological problems arising in practice. In communication terms there is too little feedback in the loop. This situation became more acute following World War II when engineering education moved toward an increased emphasis on the basic sciences underlying engineering and away from the prewar emphasis on practice. Increasingly, engineering faculty were recruited from the science faculties or from those trained by science faculties. The result was an improved emphasis on fundamentals in the engineering curricula. But a faculty trained in the sciences lacks the essential experience of engineering that is the synthesis of the science specialties required in the development of devices or systems. Their research and their teaching have tended to emphasize disciplinary specialization. This kind of experience trains a student to ask why and how things happen but *not* to ask the important question: How can I use this understanding?

Unless we bring to our university faculties a greater number of individuals whose experience encourages them to present this challenge to students, our effort to translate basic findings into practical uses will be less efficient than it should be and unnecessarily delayed. We are attacking this problem by improving the communication between engineering faculties and industrial engineers and scientists who are responsible for development and production for community needs. Examples of these techniques are the exchange of personnel between university and industrial laboratories, the encouragement of a part-time involvement of university faculty in the activities of industrial organizations, the presentation to students by industrial personnel of real problem situations in colloquia and seminars, etc. We have no complete answer to this problem and our delegation has been interested in how your teaching institutions are meeting this challenge.

Professional societies such as the IEEE and the Popov Society can significantly aid in the solution of this problem by providing a common meeting ground where there is an opportunity for the exchange of ideas between engineers in universities and in industry. It is necessary in your society and in ours that we use this opportunity to maximum advantage. It is important that our organizations serve as communication links between the academic faculties and those who bring technology to the service of the people. It is equally important that our societies serve as channels of communication across national boundaries. This task can be accomplished through our publications and through the exchanges of scientific delegation.

My colleagues on the IEEE delegation and I have been impressed by the progress in science and engineering evident in the papers presented at this congress and as we have observed it in the laboratories which we have had the privilege to visit. We have also enjoyed the opportunity to experience some of your fine artistic presentations such as the Bolshoi Ballet at your magnificant Kremlin Palace of Congresses.

The relationship between the Popov Society and The Institute of Electrical and Electronics Engineers has been most felicitous. It is our sincere hope that this relationship will continue and grow in effectiveness.

On behalf of the IEEE, of myself, and of my colleagues, I express thanks for the opportunity to join with you on the occasion of this meeting.

William G. Shepherd, President IEEE

IEEE publications

scanning the issues
advance abstracts
translated journals
special publications

Scanning the issues

Remembrance of Things Past. Living as we do under a Niagara of technical publications, in which the onrushing works seem soon subsumed in a cloud of anonymity, it is both surprising and curiously soothing to come across a crystallized piece of the history of the field. Past events take on a novel subjective coloration and significance owing to the subsequent events across which they are seen. This happy effect is achieved in two historical sketches appearing in the July IEEE Transactions on Aerospace and Electronic Systems commemorating the 1966 Pioneer Awards to Robert J. Dippy and Otto Scheller. The sketches were prepared by Robert I. Colin, the Secretary-Historian of the G-AES Pioneer Awards Committee.

The 1966 Pioneer Award to Robert J. Dippy, an Englishman and now a resident of South Australia, is in recognition of his contributions to the conception and realization of the earliest hyperbolic radio navigation system. As the G-AES Awards Committee points out, not only did Mr. Dippy's "Gee" system render yeoman service in the defense of England during World War II but it was essentially the progenitor of pulsed hyperbolic radio navigation systems developed subsequently and now in widespread use. Notable among these is Loran A, to which Mr. Dippy also contributed. A fascinating account of the proposal of the Gee system and of its subsequent effectiveness in the war comes from Sir Robert Watson-Watt's book, *Three Steps to Victory*. He wrote: "We were still heavily loaded by our top priority work on chain and airborne systems when, in Oct. '37, R. J. Dippy brought forward, in a conference in my somewhat grandiose office at Bawdsey, a proposal for a radar-like system to aid approach and landing in conditions of bad visibility. It was to use two pulse transmitters a few (10 or so) miles apart, with a fixed relation between the times of sending out the evenly spaced pulses from each transmitter. A novel receiver in the aircraft, with a cathode-ray display, would measure the re'ative delay between the two sets of pulses, and thence, not the distance to one transmitter or to the other, but the difference between the distances from the aircraft to the two. If the difference was zero, the craft was clearly in the vertical plane containing the centre line of the landing runway. Every other difference defined a particular curved (in fact hyperbolic) surface on which the aircraft at the moment lay, these curves crowded together as they came near the line joining the transmitters, and a chart with these unchanging curves plotted on it would show the aircrew where they were relatively to the extension of the centre-line of the runway, enabling them to fly along that line or any chosen one near it. The curves gathered the homing flock into a precisely defined approach funnel, so to say." Although Sir Robert Watson-Watt thought the idea "of great novelty and ingenuity," and although it could meet "a pressing need," he shelved it. However, the strong-minded Dippy "brought it forward again, at a time of critical need, to become the Gee (G for Grid) system of navigational guidance which made the thousand-bomber raid practicable, and which made 'Gee' equipment the most widely installed of all radar and radar-like airborne equipment, with the one exception of that most widely fitted of all, I.F.F." Gee was being used operationally by 12 R.A.F. aircraft by August 1941, and was subsequently developed to such an extent that on May 31, 1942, the first thousand-bomber attack was launched against Cologne, "in which 80 percent of the force reached the target area with a concentration in time which saturated the defences." Before the end of the war, 80 percent of the U.S.A.A.F. was flying with Gee.

What is Mr. Dippy doing now? He says in a letter, quoted in Colin's account: "I have found it rather a depressing exercise to discover how far I have moved from direct technical work. The modern scientific world is so vastly more complex than the dawn of electronics with which I was concerned. So now I am, in effect, a Manager, making the obstructions for other people to struggle against." Rather than deprive the reader any further pleasure of a fresh reading of, among other things, Mr. Dippy's autobiography in the Transactions, we have done a bit of hasty research among some of Dippy's old American colleagues (from M.I.T. Radiation Lab days), and turned up the following memorabilia which will not be found in the Transactions: Dippy was known to be a demon puzzle expert and a man of irreverent mien. Like Pooh-Bah in *The Mikado*, it is said, he was born sneering; for instance, after a controversy with the military, he is reported to have uttered this weighty reflection on the military mind: "*A little pants for little behinds, Little thoughts for little minds.*"

Dippy was also remembered both as a punster and as a collector and dispenser of "Little Willie" stories. For