X-320-64-341

NASA TM X-55168

FINAL REPORT OF THE GODDARD SUMMER WORKSHOP PROGRAM IN MEASUREMENT AND SIMULATION OF THE SPACE ENVIRONMENTS

COMPILED BY

ELIAS KLEIN

JUNE 15 TO SEPTEMBER 15, 1964

GODDARD SPACE FLIGHT CENTER

GREENBELT, MARYLAND

N66-14101 – N66-14131

X-320-64-341

FINAL REPORT OF

THE GODDARD SUMMER WORKSHOP PROGRAM

IN MEASUREMENT AND SIMULATION

OF THE SPACE ENVIRONMENTS

JUNE 15 TO SEPTEMBER 15, 1964

". . . For no venture is more ultimately profitable than adding just a little bit to man's knowledge of the universe in which he lives."

Lee A. DuBridge

GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland

FOREWORD

Started as an experiment in the summer of 1962, the Summer Workshop has emerged in 1964 as a helpful adjunct to Goddard's research program. Its catalytic effect has produced many useful results from university and government scientists working together in teams at Goddard.

Several important problems were solved in the brief summer period, such as those relating to equipment and techniques for high-precision calibrations of the sensitive optical experiments carried by our orbiting observatories. The solutions of still more problems undoubtedly will emanate from the groundwork initially laid by the Workshop.

Of even greater significance are the intangible benefits that accrue from this program. We have been able to attract high-calibre professors and graduate students as principal investigators and team participants. Their enthusiasm and ideas have had an invigorating effect on our in-house research. In turn, the academic people have expressed their appreciation for the opportunities afforded by the Summer Workshop to become aware of Goddard's problems. This awareness in some instances has influenced university teaching programs toward more effective training of future aerospace scientists.

The key to success of the Summer Workshop has been the excellent cooperation between the university people and Goddard staff. This is amply demonstrated by the authorships of the papers contained in this Final Report of the 1964 Summer Workshop.

HARRY J. GOETT
Director

PREFACE

The manuscripts assembled in this Final Report represent the experimental and theoretical studies conducted by the three 1964 Summer Workshop teams during their 12-week stay at Goddard. Each team investigated a number of subjects in a broad technical area to provide answers to specific questions which were required by various GSFC engineers and scientists. The results achieved in the 1964 Workshop, which was handicapped by a late start, attest to the marked effectiveness in method and procedures of this program for conducting in-house joint Goddard-university studies. Advance planning and preparation, together with the established modus operandi of this effort, made it possible to obtain solutions of essential problems.

As in previous Summer Workshops, the problems to be studied in 1964 by the different teams were chosen from suggestions offered by the various scientific groups at Goddard. This selection was made in January, about the time when the entire program had to be curtailed because of manpower ceilings which limited summer employment. Notwithstanding these sudden changes in plans, a number of university scientists was mobilized pending a management decision on personnel. In April, Dr. Goett made possible the implementation of the third successive Summer Workshop at Goddard; but the list of project-topics was reduced, and there was a shift in emphasis of the subjects to be investigated. At this time, top priority was placed on preparing the Vacuum Optical Bench Facility for test and evaluation of the optical experiments on the several orbiting observatories.

Despite the late start, Goddard was able to attract topflight specialists in the fields of optics, photometry, and spectroscopy, whose contributions to the team effort made possible the fruitful results that were obtained. It should be noted, however, that the academic experts were available on short notice only because most of them were tentatively lined up early in January. This situation indicates the importance of advance planning and preparation in all phases of the Summer Workshop Program, especially in establishing suitable relationships with university professors long before they are needed on the job at Goddard. For effective results, much depends upon the attitudes and interests of the professors who participate in this collaborative activity. The team individuals tend to catalyze and stimulate each other so as to foster professional awareness among the group.

While this Report emphasizes the real domain of problem-solving and outlines the useful results obtained in the Summer Workshop, it also includes the long-range benefits (not readily apparent) and the intangible values which accrue to Goddard as well as to those who contribute to this effort. By working on these problems in groups or as teams composed of academic scientists and Goddard practical engineers, there is developed a lively interchange of ideas relating to basic knowledge and applied design. Perhaps the Workshop participants should speak for themselves. Here is what a faculty specialist said: "The end result of the Summer Workshop is not only an item of new knowledge or novel equipment, but also a learning and searching process which begets a more practical viewpoint. Under this light, the Workshop program has all the advantages of a conventional university contract with NASA and several other benefits besides. In addition to what the Summer Workshop might contribute to Goddard, there is another long-term benefit no less important: By working closely with Goddard people, we learn about the

NASA program more deeply than from books and journals, more accurately than from newspapers. And what we learn will naturally color our teaching."

In a letter from a graduate student: "I would like to say that I am proud I was able to be a part of the team effort at Goddard in the Summer Workshop. I will always remember the first exposure to research work; this exposure will influence my future plans concerning the field of engineering I choose. I feel that I have learned much more about space exploration and its related problems, about the role of the federal government in this exploration, and about how these problems are attacked. The Summer Workshop seems to be an ideal method of attack and one that works well. In general, I was impressed by the high degree of organization at Goddard."

Another facet of this year's Goddard Summer Workshop was the participation in each team of a National Science Foundation high-ability secondary school student. As in the 1963 program, these students were actively involved in the different problem areas associated with their respective teams. Each student was included in the planning sessions of his or her team. Being under the guidance of experienced academic researchers and teachers as well as of the Goddard permanent staff, the high-school students received exceptional supervision and training. Here is the response of one of these students after eight weeks of scientific work with one of the teams:

"My summer experience gave me an opportunity to work with creative and dedicated scientists who helped me understand what scientific research is like. I learned of the various areas of training necessary to carry out an investigation, and how each scientist in a team depends upon the results and the work of the others. This program made me realize how important mathematics, chemistry, and physics are to a scientist."

SENIOR FACULTY MEMBERS

BEYER, Gerhard H., Dr.,—Professor and Chairman of Chemical Engineering Department in the University of Missouri. Member of numerous scientific societies. Author of many publications in chemical engineering, thermodynamics and energy balance. During 1962-63 visiting professor in chemical engineering in University of Wisconsin. Participated in Goddard 1963 Summer Workshop.

CHEN, Yu, Dr.,—Associate Professor of Engineering Mechanics, Rutgers, The State University, New Nersey. Member of numerous honor and scientific societies. Author of various publications on vibration and structural dynamics.

MOHR, Eugene I., Dr.,—Professor and Head of Physics Department, Columbia Union College. Member of numerous honor and scientific societies. Author of some publications in Spectroscopy. Participated in Goddard 1963 Summer Workshop.

MORRIS, Joseph B., Dr.,—Associate Professor of Chemistry, Howard University. Member of numerous honor and scientific societies. Author of various publications in analytical chemistry.

THEKAEKARA, Matthew P., Dr.,—Associate Professor of Physics at Georgetown University specializing in spectroscopy and space physics. Belongs to numerous honor and scientific societies. Author of many publications in physics and astrophysics. Participated in Goddard 1962 and 1963 Summer Workshops.

WAIDELICH, Donald, Dr.,—Professor of Electrical Engineering, University of Missouri. Formerly visiting Professor at University of New South Wales, Kensington, NSW, Australia, on Fulbright grant. Member and fellow of numerous professional engineering and scientific societies. Many awards and listings in honor societies. Author of numerous scientific publications and co-author of book entitled "Transients in Electrical Circuits." Participated in Goddard 1962 and 1963 Summer Workshops.

JUNIOR FACULTY MEMBERS

JOHN, James E. A., Dr.,—Assistant Professor Mechanical Engineering, University of Maryland. BSE Princeton University 1955, MSE Princeton University 1956, PhD University of Maryland, 1963. Participated in Goddard 1962 and 1963 Summer Workshops.

VEGUILLA-BERDECIA, Luis A., Dr.,—Teaching Fellow and Research Associate in chemistry and solid state physics, Howard University. BSc Brooklyn College 1954; MA Brandeis University 1957; PhD Howard University 1964.

GRADUATE STUDENTS

CROOKS, Jonathan B.—Johns Hopkins University—Chemical Engineering. BSE Johns Hopkins University 1962. Working toward PhD on University Fellowship. Participated in Goddard 1962 and 1963 Summer Workshops.

DRALLE, Anthony V.—Carnegie Institute of Technology—Physics. SB Massachusetts Institute of Technology 1963. Working toward PhD in Physics at Carnegie Institute of Technology.

HARDESTY, Donald Roy—University of Maryland—Mechanical Engineering. BS University of Maryland 1964. Expects to work for PhD in Mechanical Engineering (Jet Propulsion) at Princeton University.

HEINE, John C.—Massachusetts Institute of Technology—Mechanical Engineering. SB Massachusetts Institute of Technology 1962 and SM 1964. Working toward PhD at MIT.

MACKENZIE, Donald W.—University of Californis (Berkeley)—Engineering. BS Civil Engineering 1965. Working toward MS in Civil Engineering, University of California.

MORRISON, John C.—Johns Hopkins University—Physics. BS University of Santa Clara 1963. Working toward PhD in Physics, Johns Hopkins University.

OWENS, Frank J.,—University of Connecticut—Physics. BS Manhattan College 1962. MS University of Connecticut 1964. Working toward PhD in Physics, University of Connecticut.

PEARCE, S. Beacher—Michigan State University—Physics. BA Hiram College 1962. Working toward PhD in Physics, Michigan State University.

PIZIALI, Robert L.—University of California (Berkeley)—Engineering. BS Mechanical Engineering 1965. Working toward MS in Mechanical Engineering, University of California.

SAPERSTONE, Stephen H.—University of Maryland—Mathematics. BEP Cornell University 1962; MA University of Maryland 1964. Working toward PhD in Mathematics, University of Maryland.

SHAPIRO, Stephen M.—Johns Hopkins University—Physics. BS Union College 1963. Working toward PhD in Physics, Johns Hopkins University.

SHARP, Robert L.—University of Maryland—Mathematics. BEE Ohio State University 1959. Working toward PhD in Mathematics, University of Maryland. Participated in Goddard 1963 Summer Workshop.

YAFFE, Barry M.—Johns Hopkins University—Fluid Dynamics. BES Johns Hopkins University 1962. Working toward PhD in Fluid Mechanics, Johns Hopkins University.

PART-TIME CONSULTANTS

ALPERT, Daniel, Dr.,—Professor, Coordinated Science Laboratory at the University of Illinois. Has made many outstanding contributions to the work in gaseous and physical electronics and in their applications to high vacuum measurements.

GARDNER, Irvine C., Dr.,—For over 40 years, Chief Division Optics and Metrology, National Bureau of Standards. Eminent authority on testing and use of optical instruments. Author of many publications in optics and spectroscopy.

SIMONS, John D., Dr.,—Independent consultant in vacuum science and technology. Has made important contributions to the development of high vacuum systems and gage calibration. Chairman of ASTM Committee on Space Simulation.

PROJECT A: SPACE OPTICS, RADIATION, AND SOLAR SIMULATION

CONTENTS

ILLUSTRATIONS

TABLES

SUMMER WORKSHOP 1963

Program Outline and Team Participants

PROJECT A: Space Optics, Radiation, and Solar Simulation

Study Topics

A-1 Measurement of small angles with an accuracy of 0.01 second of arc

A-2 Calibration of an angle generator using a sine bar and a laser beam

A-3 Gravity-reference system for precise alignment of instruments

A-4 Degradation of energy sensors of TIROS satellites

A-5 Evaluation of 6-foot integrating sphere as a wide-angle source

A-6 Interferometric spectrophotometry with finite range of travel of the mirror

TEAM A

Academic Personnel	Goddard Personnel	Code
Dr. Matthew P. Thekaekara, principal investigator	Dr. William Nordberg, staff advisor	651
Dr. Irvine C. Gardner	John W. Larmer	333
Stephen M. Shapiro	James L. Diggins	325
Donald W. Mackenzie	Norman Wade	325
Dr. Eugene I. Mohr	Robert Appler	325
Dr. Joseph B. Morris	Andrew W. McCulloch	651
Robert L. Sharp	Nathan J. Miller	651
John C. Morrison	Ernest Hilsenrath	651

NSF Summer Science Student
Kathleen Thompson

OPTICS AND SPECTROSCOPY IN SPACE ENVIRONMENT SIMULATION AND CALIBRATION

M. P. Thekaekara

INTRODUCTION

The titles of the papers in this section of the 1964 Summer Workshop Report might give the impression that there is little in common between them, that they represent endeavors in widely scattered areas by men working in isolation. But there is a strong thread of unity; none of the work was done in isolation without some cooperation at least from others of Team A or of the Goddard permanent staff. And some of the work was fully a team effort, the joint achievement of many participants. The purpose of this introduction is to show how each of these projects came to be started, and how these separate tasks fit into the long-range objectives of Goddard Space Flight Center and the National Aeronautics and Space Administration.

All these are problems in optics and spectroscopy, the field of physical sciences which deals with electromagnetic radiation and instruments for measuring it. As they are part of the NASA venture, these problems necessarily include optical procedures and equipment for testing experimental packages before launching and for ensuring that they will perform satisfactorily throughout the life of the satellite. The instruments, both spaceborne and ground-based, are delicate yet rugged. They cover a wide range of electromagnetic spectrum – from 900 A in the vacuum ultraviolet to 50 microns in the far infrared – a gamut of ten octaves. The energy sensors are of many different kinds: photomultiplier tubes, thermistor bolometers, thermopiles, photographic plates and films. The monochromators and spectrographs are prism instruments and grating instruments, with concave gratings and plane gratings, and prisms of glass, quartz, and rocksalt. Spectrophotometers, integrating spheres, autocollimators, multilayer interference filters, interferometers, and instrumentation at every level of sophistication are the ordinary tools of the trade.

SOLAR SIMULATION RADIOMETRY

During the two previous Summer Workshops, the solar-simulation studies occupied the major share of our effort. This summer limited project in the same area was attempted, to further improve certain techniques initiated last summer and developed during the winter.

Trouble-free, reliable performance of space experimental packages requires that the temperatures of the instruments remain within certain safe limits. The equilibrium temperature and thermal balance of spacecraft are governed by the heating due to energy absorbed mostly from the sun, also from the earth and other sources, and the cooling due to energy radiated to the space environment of about 4°K. Prelaunch tests of thermal balance expose the satellite to radiation which, for total energy, spectral distribution of energy, uniformity, and collimation, closely approximate solar radiation outside the earth's atmosphere. The Space Environment Simulator in Building 10; with its battery of 127 Hg-Xe lamps, each of 2.5 kw, each in a collimation module of high complexity, has been designed to provide adequate approximation to radiation that the satellite in orbit will receive from the sun. As a support and complement to the extensive time-consuming

measurements made in earlier Summer Workshops on the high-resolution spectral irradiance of solar-simulation sources, techniques are also necessary for relatively rapid measurements on the short-term variations of spectral irradiance. The technique of using a series of 12 bandpass multilayer interference filters designed for the Hg-Xe spectrum by Eppley Laboratories had been investigated during the last year, and a computer program for automatic reduction of data had been developed.

What we have now attempted is a further advance in the filter method, to give better spectral resolution. In addition to a few Eppley filters of the range below 4000 A, we have used thirty monopass filters manufactured by the Optics Technology Lab. These filters have certain advantages: namely, narrower bandpass, a symmetrical transmittance profile, and more complete spectral coverage. The suitability of these filters has been established experimentally. Measurements of sufficient internal consistency also have been made using a tungsten-coil lamp of known spectral irradiance and a Hg-Xe lamp. Computer programs have been set up for obtaining filter factor from known spectral-radiant flux and filter signals, and for graphic display of spectral radiant flux of an unknown source from measured values of filter signals.

TIROS SATELLITES RADIOMETRY

Two problems concerning the TIROS satellites engaged the attention of some of the members of our group. The TIROS satellites, in addition to their well-known activities of photographing cloud cover and predicting hurricanes, are also engaged in detailed measurements over widebands of the energy distribution of earth-emitted radiation. The energy sensors are thermistor bolometers, each with five channels, each channel having the appropriate filters to view specified ranges of the spectrum.

The first problem was prelaunch calibration; that is, relating the radiant-energy input to the electrical-voltage output. Conventional methods of calibration with NBS tungsten-lamp standards, which are quasi-point sources, cannot yield meaningful results for a radiometer with a rather wide acceptance angle. An earlier modification of the method was to use a high-intensity projection lamp and a curved-paper diffusing screen. Although the results were satisfactory, the new generation of TIROS radiometers seemed to call for improved calibration techniques. An integrating sphere, which has long been used in measurements of illuminating engineering, was chosen by the TIROS group scientists as the best solution. But, when a sphere has a 6-foot diameter, the source of radiation is twelve standard lamps distributed on the inside, and the spectral range is extended beyond the visible of 0.4 to 0.6 micron to the UV-visible-IR range of 0.2 to 3.0 micron, many unforeseen problems arise.

A related problem in radiometry was the need for a horizon sensor to trigger the television cameras of a new series of TIROS satellites which will be known as the TIROS Cartwheels, since they are spin-stabilized to turn as a cartwheel, and not space-stabilized as are many other satellites. The relative orientations of the television camera and the horizon sensor will permit the former to look directly down at the earth when the latter sees the horizon. A comparative study has been made of the energy received by the horizon sensor when it sees the clear horizon or a cold cloud above the earth, so that its response characteristics can be adjusted to the level necessary for trigger action only when the camera is pointing at the earth.

The second problem concerning the TIROS radiometers comes from an analysis of their behavior in flight. A radiometer which, in the laboratory, seems to have an almost unlimited life of reliable performance, of constant output for a given input, seems to degrade in a few days or weeks when exposed to the rigors of space. Since the experimental

package cannot be recovered, the degradation must be diagnosed on similar packages under simulated space environment. A prime suspect, according to the staff of the Physics Branch, was the black paint which might undergo vacuum evaporation; this would affect the optical components by lessening the transmittance of the lenses and filters and the reflectance of the mirrors. As the extremely low pressures of space environment cannot readily be duplicated in the laboratory, a satisfactory compromise is the medium-low pressure of a vacuum chamber. In the chamber is placed a metal surface coated with the same paint as that used on the radiometer; mirror and lens surfaces are mounted close to the paint; and the painted surface is maintained at about 100°C. The experiments showed some highly meaningful results.

Another research area presented to the Summer Workshop concerned the use of the Michelson interferometer as a spectrophotometer. The Michelson interferometer, an instrument familiar to students of college physics, has to its credit certain spectacular achievements: for example, standardizing the meter in terms of the wavelength of light, and giving the experimental basis for the Einsteinian theory of relativity through the famous Michelson-Morley experiment. The conventional use of the Michelson interferometer is the precise measurement of wavelength of strictly monochromatic light. The instrument has in recent years been put to an entirely new application, the spectral-radiance analysis of composite light. An experimental package developed for interferometer measurements of the spectral radiance of earth albedo must be modified significantly when the same instrument is to be used on a fly-by mission to Mars or Venus, when the object is at a greater distance and subtends a smaller angle. Preliminary work has been done on a telescopic system which gives the required change in the acceptance angle.

Several problems of a theoretical nature arise in the transformation of the interferometer data to a spectral-radiance curve. The data telemetered from the satellite give a time-dependent Fourier integral. The end result of the computational analysis should be a wavelength-dependent spectral-energy curve. A high-speed electronic computer can reliably perform the Fourier analysis, provided, however, that the telemetered data gives a true Fourier integral. Unfortunately, experimental errors enter at different stages of data collection; also, most importantly, the telemetry does not give all of the Fourier integral, but only a truncated central portion of the integral. A mathematical analysis has been attempted. It is shown that the incompleteness of the experimental data can be corrected for by adding an approximating function to the data and an iteration procedure to the data analysis.

VACUUM OPTICAL BENCH

Other problems in space optics attempted by the Summer Workshop concern the highly precise instrumentation of the vacuum optical bench.

In the design of the vacuum optical bench, a Cassegrainian telescopic system receives radiation from the exit aperture of a modified Paschen-mounting monochromator, and illumines the test plane of the VOB with a high degree of uniformity, monochromaticity, and collimation at any desired wavelength between 900 and 5000A. Early performance tests showed that the rather exacting specifications were not being met as far as uniformity of illumination was concerned. Part of the failure could be attributed to the tripartite concave grating which has been blazed for a wavelength near Lyman-alpha. A plane grating in one of the many possible mountings (as, for example, the Czerny and Turner mounting) should in principle be free of the annular pattern of brightness. The design specifications of such a monochromator were examined.

Because testing for alignment of optical instruments is a major concern of the permanent staff of the vacuum optical bench, several high-precision autocollimators are available for measuring small-angle deflection. One of these which embodies many advanced features and gives promise of a high level of performance is the Midarm Model 14 autocollimator, supplied by the Razdow laboratories. An instrument of such high precision, however, requires several fine adjustments and elaborate calibration procedures. Two methods of absolute calibration of small angles were also developed: one of these measures the angular deflection of a mirror by measuring the linear displacement of a laser beam reflected from the mirror and focused at a distance of about 100 feet. The other produces small angular displacements in increments of 0.05 sec. by placing graded gauge blocks below one of the legs of a massive sine bar.

In all small-angle measurements, a fixed direction independent of all laboratory frames of reference is highly desirable. Such a direction is provided by the force of gravity. The direction of the normal to a mercury surface, while coinciding with the gravity reference, has the added advantage of being readily detectable by all autocollimators. Theoretical studies have been made on the reliability of the gravity reference. The direction of the local vertical (that is, the direction in which a plumb line comes to rest) is affected by the perturbations of the moon and the sun. According to the relative positions of these perturbing masses with respect to the earth, the local vertical swings around the direction of the earth center. The maximum possible change is of the order of 0.01 second of arc per hour, that is, of the same order of magnitude as the accuracy of the angular measurements which we are aiming at.

Several techniques have been tried for obtaining a mercury surface stable, ripple-free, and of constant high reflectance. With a shallow dish of sloping edges as container for the mercury and a layer of highly viscous and transparent oil floating on the mercury, a satisfactory gravity-reference system has been obtained.

USE OF INTERFERENCE FILTERS FOR SPECTROPHOTOMETRY OF SOLAR SIMULATORS

Matthew P. Thekaekara and Donald MacKenzie

INTRODUCTION

The problem of determining the short-period fluctuations in the spectral-radiant flux of solar simulators has suggested the desirability of a filter technique for spectroradiometry. Although the conventional method[1] using a monochromator and a tungsten standard source provides a high degree of accuracy and a high wavelength resolution, these advantages are to some extent sacrificed for the sake of rapidity in making the measurements and of facility in analyzing the data. For this purpose we choose multilayer interference filters which transmit the energy over a narrow wavelength band and absorb all energy outside that band. The energy transmitted by the filter is measured by a suitable detector, a thermopile for example. The signals of the thermopile are entered directly into a computer program which prints out normalized values of the spectral radiant flux for a large number of narrow wavelength bands, and which also displays the energy distribution as a graph with wavelength on the x-axis and spectral-radiant flux per unit wavelength on the y-axis. Approximations are made at different stages of the data analysis in order to facilitate the solution.

In the course of the research, different types of filters and techniques of data reduction were tried. Most of the earlier work[2] was done with sets containing 12 filters each, manufactured by Eppley Laboratories, Newport, Rhode Island, for spectral ranges which had been chosen with special reference to the mercury-xenon arc on the carbon arc. The spectral resolution in the infrared range beyond one micron seemed inadequate, since only two of the Eppley filters are in this range. A set of thirty filters manufactured by Optics Technology, Inc., known as Spectracoat monopass filters, was found to give more satisfactory results.[2] The main disadvantage of these filters is that the lowest wavelength range reached is about 4000 A, whereas the energy of solar-simulator sources in the range below 4000 A is relatively high and specially significant. In the final set of measurements, the only one reported here, a combination of four Eppley filters in the UV range and thirty Monopass filters was used.

EXPERIMENTAL PROCEDURE

The experimental arrangement is quite simple. A thermopile (water-cooled Eppley thermopile No. 5888) is mounted on an optic bench at a suitable distance from the source of radiation. Filters are interposed in succession in front of the thermopile, and the voltage output is read as accurately as possible. A digital voltmeter or a strip chart recorder gives a quick reading, but is not sufficiently accurate; a Leeds and Northrup K-3 potentiometer was found to be most suitable for our purpose. Great care must be taken in mounting the filters so that all filters are in the same position relative to the thermopile and are normal to the radiant flux. Shutters were placed in front of the source, to shield the thermopile from radiation from the lamp envelope or lamp housing. Two types of sources were used: tungsten-coil standard lamps (known as quartz-iodine lamps) standardized at the National Bureau of Standards,[4] and 2.5-kw mercury-xenon lamps manufactured by Hanovia Co.

The quartz-iodine lamps served for purposes of calibration and for testing the methods of data analysis. Although these lamps are calibrated for a distance of 43 cm, the signal is too small for accurate measurement at such a long distance, especially in the visible range where both the spectral-radiant flux of the lamp and the transmittance of the filters are small. Three sets of readings, taken at distances of 15, 25, and 30.3 cm, were reduced to what they would have been at 43 cm, assuming the inverse-square law to hold true. The average deviation of the reduced readings was 2.2 percent. Measurements made at different times on the signal without any filter showed an average deviation of 1.25 percent.

The mercury-xenon lamp was housed in a large metal box and was viewed through a quartz lens. The lamp was operated at different power levels in order to study the dependence of spectral-radiant flux on power input.

DATA ANALYSIS

The filter technique would give accurate values of the energy in discrete spectral ranges, if the following conditions were satisfied: (1) Filters are available which have a constant spectral transmittance over a small wavelength range and which are totally opaque outside this range; (2) thermopiles are available with output dependent solely on the incident energy and independent of the wavelength. Though neither of these conditions is fully satisfied, our experimental conditions are assumed to give a close enough approximation.

The filters are multilayer dielectric film filters. Their principle of operation has been discussed in detail in standard textbooks.[5,6,7] The spectral transmittance is approximately bell-shaped, as shown in Figure A-1. If P°_{λ} is spectral radiant flux incident on the filter, τ_{λ} the spectral transmittance, the energy transmitted is

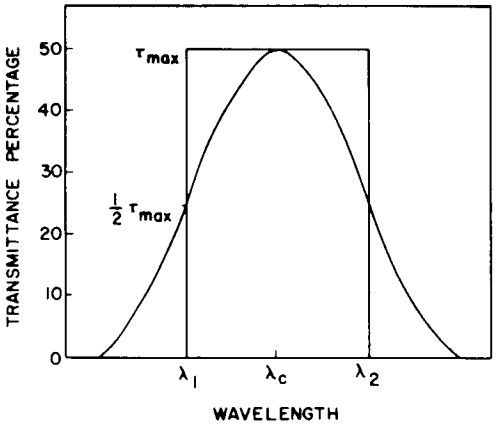


Figure A-1-Typical Filter Transmittance Curve

$$P = \int_{\lambda_1}^{\lambda_2} P^{\circ}_{\lambda}\, \tau_{\lambda} d\lambda$$

λ_1, λ_2 are the wavelengths on the short and long wavelength sides respectively, where τ_{λ} drops down to zero. The energy incident in the same range is

$$P^{\circ} = \int_{\lambda_1}^{\lambda_2} P^{\circ}_{\lambda}\, d\lambda .$$

We may define a filter factor F by the equation $F = P^{\circ}/P$. If a constant F can be uniquely defined for a filter, the value of P° can be calculated from the observed value of P.

But the filter factor F can be uniquely defined only if the transmittance τ_{λ} is constant for the range λ_1 to λ_2. If τ_{λ} is dependent on λ, as is always the case, F for a given filter depends on both τ_{λ} and P°_{λ}. However, F can be either empirically determined or theoretically calculated from a spectrum of known P°_{λ}. For the empirical determination, P is measured in an actual experiment and P° is found by numerical integration. For the theoretical calculation, both P and P° are found by numerical integration, τ_{λ} being known from independent measurement. F thus determined for a known

spectrum will be approximately valid for another spectrum which is very close to the first. This is one of the basic assumptions of the filter technique.

On the other hand, τ_λ falls to zero more or less asymptotically, so that λ_1 and λ_2 cannot be determined precisely. The wider the range, the greater will be the values of F. The dependence of the computed value of F on the width of the range and the spectral distribution is shown in Table A-1, which gives the filter factors for the 12 filters of Eppley Filter Set A-2. For each filter, 4 values of F are given, the first three based on our calculations and the fourth from the manufacturers. For our calculations three different wavelength ranges were used: the range of more than 5-percent transmission, the entire range of non-zero transmission, and the range assigned by the manufacturers. The function P_λ° is that of a mercury-xenon arc, and the values are those given in the Final Report of the Summer Workshop, 1963.[1] The function τ_λ was averaged from curves supplied by the manufacturers and those we obtained by using a Beckman DK-2 spectrophotometer. The integration for P° and P was performed by series summation of the values of $P_\lambda^\circ \tau_\lambda d\lambda$ and $P_\lambda^\circ d\lambda$ over the short wavelength ranges between flexion points of the spectrum. In some cases, as many as 100 terms were computed for the series summation. All calculations were made on a Burroughs E101 Computer. The manufacturer's values of F are based on a different function for P_λ° and the summation was made by dividing the transmittance range into 12 subranges. As is evident from the table, Eppley's values of F are lower than ours, calculated for the same range, for all filters except the third. The average deviation between our values and Eppley's is 35 percent. Changing the range to 5-percent transmission or non-zero transmission greatly changes the filter factor; but the relative increase is not the same for all filters, being as small as 6 percent for the fifth filter and as high as 85 percent for the first filter. The strong dependence of F on the arbitrarily assigned limits λ_1 and λ_2 shows that a computed value of F cannot be used to find the energy P° in a range from the transmitted energy P. It would also seem that, by including in the computation of F subranges where transmittance is 10 or 5 percent of maximum transmittance, we would be giving an unduly high weight to the energy P_λ° in these subranges.

For the technique of data analysis which we have adopted here, we define the effective range of a filter as follows: Let λ_c (Figure A-1) be the wavelength at which the transmittance of the filter is maximum; let τ_{max} be the maximum transmittance; let λ_1, λ_2 be two wavelengths on either side of and equidistant from λ_c such that $(\lambda_2 - \lambda_1)\tau_{max}$ is equal to area $\int \tau_\lambda d\lambda$ under the transmittance curve. We consider the range λ_1 to λ_2 as the effective range of the filter. Let $\lambda_2 - \lambda_1 = \lambda f$. We also assume, as a first approximation, that the total energy transmitted by the given filter is the same as that of an ideal filter which has $\tau = \tau_{max}$ for $\lambda_1 < \lambda < \lambda_2$ and $\tau = 0$ for $\lambda_1 > \lambda > \lambda_2$. This is a heuristic assumption, and its validity will have to be established by the final results.

The filter factor F for the range λ_1 to λ_2 is determined empirically as follows: Assume that we have a source (for example, a quartz-iodine lamp) of which the spectral radiant flux is known to a sufficiently high degree of accuracy. Let P° be the radiant flux of the lamp in the range λ_1 to λ_2. Let P be the energy transmitted by the given filter. P° is obtained by numerical integration of the known energy-distribution function. P is determined experimentally from the measured voltage output and the calibration constant of the thermopile. The empirical value of F is defined to be P°/P.

Using the quartz-iodine lamp QL 74 and the thermopile No. 5888, we determined the filter factors of the thirty monopass filters and compared them with the calculated values of the filter factors. The calculated values of F are the ratio of P°/P, where P is also obtained by numerical integration, as given by the expression $P = \Sigma P_\lambda \tau_\lambda \Delta\lambda$. τ_λ is known from the data supplied by the manufacturers, which agreed rather closely with our own measurements with a Beckman DK-2A.

Table A-1
Filter Factors of Eppley Filters A-2

(Calculated for different wavelength ranges (1, 2 and 3) compared with manufacturer's values (4)). (1) for wavelength range of more than 5-percent transmission; (2) for entire wavelength range of non-zero transmission; (3) for wavelength range assigned by manufacturers (4) filter factors supplied by the manufacturer.

Filter	Range		Energy		Filter Factor	Filter	Range		Energy		Filter Factor
	λ_1	λ_2	Incident	Transmitted			λ_1	λ_2	Incident	Transmitted	
1	1-2500	2967	1099	112	9.81	7	1-4195	4525	3640	1517	2.40
	2-2500	3150	3172	130	24.4		2-4128	4580	4205	1551	2.71
	3-2500	3090	1540	122	12.6		3-4150	4550	4030	1550	2.60
	4-2500	3090			11.4		4-4150	4550			2.02
2	1-2795	3075	829	122	6.77	8	1-5171	5775	4623	2136	2.16
	2-2700	3126	1642	134	12.2		2-5146	5859	7540	2230	3.38
	3-2750	3120	1530	133	11.5		3-5150	5750	4116	2099	1.96
	4-2750	3120			4.03		4-5150	5750			1.48
3	1-3038	3305	2266	506	4.48	9	1-5641	6002	4964	2238	2.22
	2-2900	3400	3180	520	6.12		2-5604	6038	5720	2325	2.46
	3-2950	3380	2891	517	5.59		3-5650	6000	4647	2283	2.04
	4-2950	3380			7.14		4-5650	6000			1.88
4	1-3280	3510	799	112	7.17		1-6044	7960	11257	5083	2.21
	2-3220	3607	1863	120	15.6	10	2-5966	9010	22205	5374	4.13
	3-3250	3550	1084	117	9.29		3-6000	8500	87243	4939	3.41
	4-3250	3550			4.07		4-6000	8500			3.32
5	1-3550	3846	3919	714	5.49	11	1-7960	11625	36148	17923	2.02
	2-3500	3900	8304	1436	5.78		2-7620	13520	52796	15195	2.90
	3-3550	3860	7843	1433	5.47		3-8000	12500	46374	18052	2.57
	4-3550	3860			5.32		4-8000	12500			2.30
6	1-3919	4195	2140	768	2.79	12	1-11841	18450	33771	16352	2.06
	2-3842	4280	2944	797	3.69		2-11100	21000	41121	16536	2.79
	3-3910	4150	2076	721	2.88		3-11500	19500	37465	16462	2.28
	4-3910	4150			2.19		4-11500	19500			2.03

The results of the comparison are presented in Figure A-2, where the ordinates are the percentage differences between the observed and calculated values of the energy transmitted through the thirty filters, and the absissae are the wavelengths λ_c of maximum transmittance. Only for two of the filters in the wavelength range beyond 2500 nm (one nanometer = 10^{-9} meter = 10 A) is the observed value of the energy transmitted by the filter less than the calculated value. In the range 500 to 2500 nm, the observed value is about 20 percent higher; and, in the range between 400 and 500 nm, the percentage difference is considerably higher. These large differences show that empirical values of F should be used rather than calculated values; they also show the dependence of the filter factors on the spectral distribution of the source.

The reason for the lack of agreement between calculated and observed values can be readily understood. The central transmittance peak of the multilayer interference filters is necessarily accompanied by secondary transmittance curves at shorter and longer wavelengths. (See Figure 0-3 in Reference 5.) The latter are removed, almost but not entirely, by colored glass filters. The transmittance curves of monopass filters supplied by the manufacturers show a very low transmittance tail on either side of the peak for a wide wavelength range. No account was taken of this in our calculation. The tungsten-coil lamp has only 1.5 percent of its energy in the wavelength region below 500 nm; filters in this range transmit relatively more energy of longer wavelength.

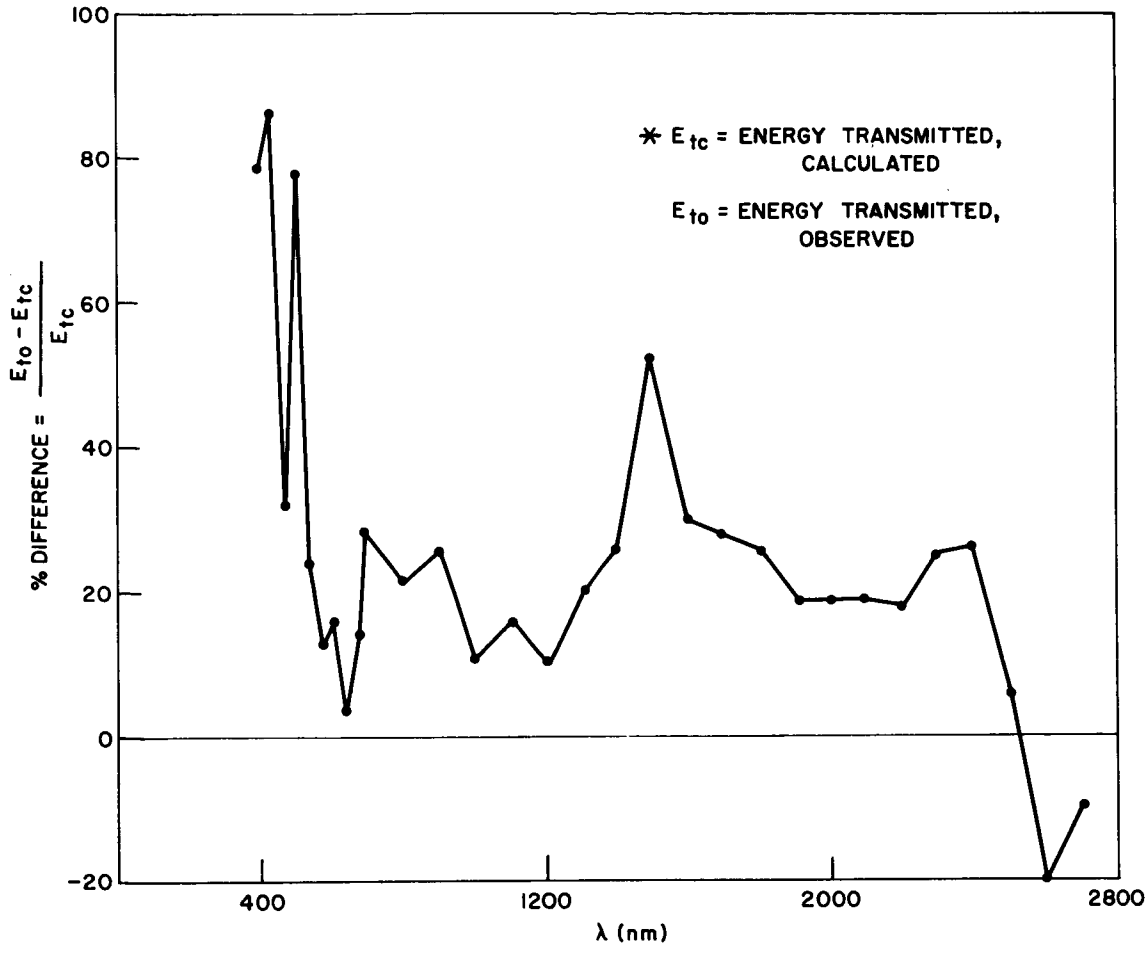


Figure A-2-Differences in Energy Transmitted by the Filters

The low values of the observed transmitted energy in the range beyond 2500 nm are probably due to a different cause, which also is highly significant for the reliability of the filter technique. It is usually assumed that the response of the Eppley thermopiles is independent of wavelength, and manufacturer's bulletins have frequently stated that the absorptance of the thermopile surface is perfectly flat. Careful studies which we have made on different thermopile surfaces seem to indicate that this is not strictly true. This is confirmed by independent measurements made by Stair and Schneider[8] and by Eisenman, Bates, and Merriam.[9] Stair and Schneider have recorded the ratio of the response of two Eppley thermopiles to that of a cavity detector; the ratio, which is a maximum at 700 nm, falls by 5 percent as the wavelength increases to 3000 nm. The absorptance of cavity detectors also decreases steadily as the wavelength increases from 1000 nm to the distant infrared, as shown by Eisenman, Bates, and Merriam.

Figure A-3 shows the spectral irradiance of the tungsten lamp QL 74 and the calculated values of the energy transmitted by the monopass filters. Each filter is represented by a thin rectangular strip located at the respective λ_c . The width of each rectangle is the effective wavelength range λ_f , and the area of the rectangle is the transmitted energy.

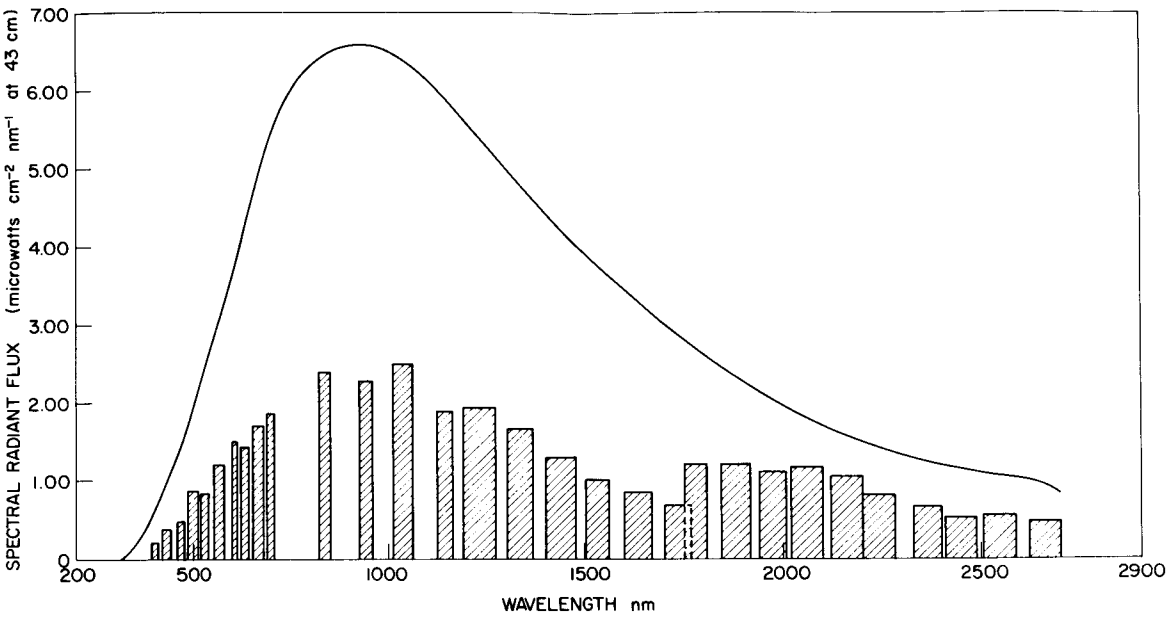


Figure A-3-Spectral Radiant Flux of Standard Lamp vs. Energy Transmitted by Monopass Filters

The procedure for evaluating the spectral-radiant flux of an unknown source is as follows: The wavelength ranges λ_1, λ_2 for each of the filters are determined from the transmittance curves; as stated earlier, they are the wavelengths at which transmittance is half the maximum transmittance. The filter factors F_i are determined from measurements made on a source which is spectrally similar to the unknown source and of which the spectral-radiant flux is known with sufficient accuracy. The energy p_i transmitted by the ith filter ($i = 1$ to N, $N = 34$, 4 Eppley filters and 30 monopass filters) is measured. If P_i° is the incident energy and P_i the transmitted energy in the range of the ith filter from the known source, and if p_i° is the incident energy in the same range from the unknown source, then, since the two sources are similar,

$$F_i = \frac{P_i^\circ}{P_i} = \frac{p_i^\circ}{p_i} \quad ,$$

so that

$$p_i^\circ = F_i \, p_i \, .$$

Thus is obtained the energy in 34 narrow-wavelength bands, each of width λ_{fi}, centered around λ_{ci}.

The filters do not yield information directly about the ranges intermediate between those covered by the filters. Since the sources are similar and the ranges are small, we assume the following relation:

$$\frac{P_{ij}^\circ}{P_i^\circ + P_j^\circ} = \frac{p_{ij}^\circ}{p_i^\circ + p_j^\circ}$$

where the subscript ij refers to the range between the ith and jth filters. Hence, we define a filter factor F_{ij} for the intermediate range:

$$F_{ij} = \frac{P^\circ_{ij}}{P^\circ_i + P^\circ_j}, \text{ so that } p^\circ_{ij} = F_{ij}\,(p^\circ_i + p^\circ_j)\,.$$

For the wavelength range shorter than λ_1 of the first filter (subscript 01), and longer than λ_2 of the Nth filter (subscript $N\infty$),

$$p^\circ_{01} = F_{01}\, p^\circ_1, \text{ where } F_{01} = \frac{P^\circ_{01}}{P^\circ_1}, \text{ and}$$

$$p^\circ_{N\infty} = F_{N\infty}\, p^\circ_N, \text{ where } F_{N\infty} = \frac{P^\circ_{N\infty}}{P^\circ_N}$$

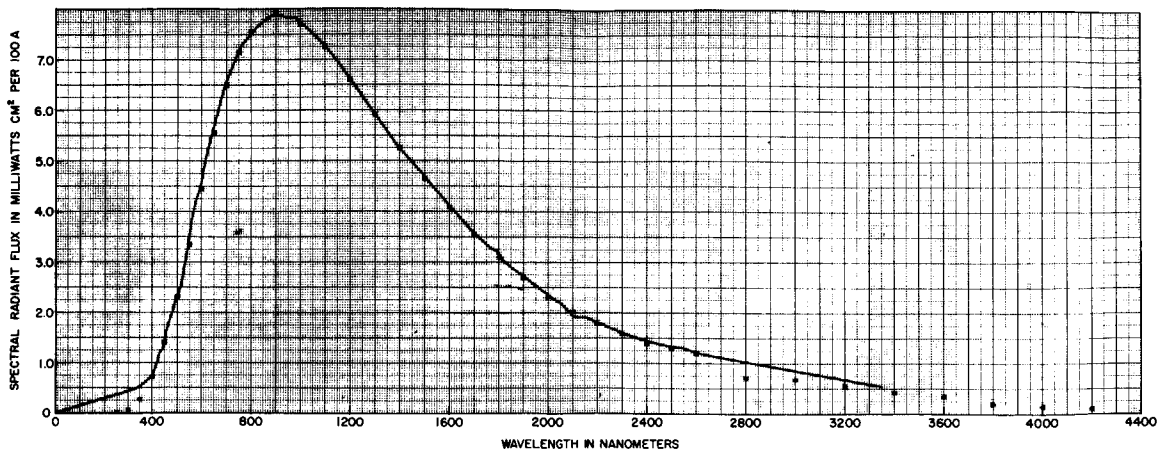
Two computer programs have been developed, one for calculating the filter factors and the other for giving the spectral-radiant flux of an unknown source. In the actual computation, we assume that the energy is zero below 200 nm and above 4000 nm, so that p°_{01} is for the range 200 to 282.8 nm, and $p^\circ_{N\infty}$ is for the range 2701 to 4000 nm.

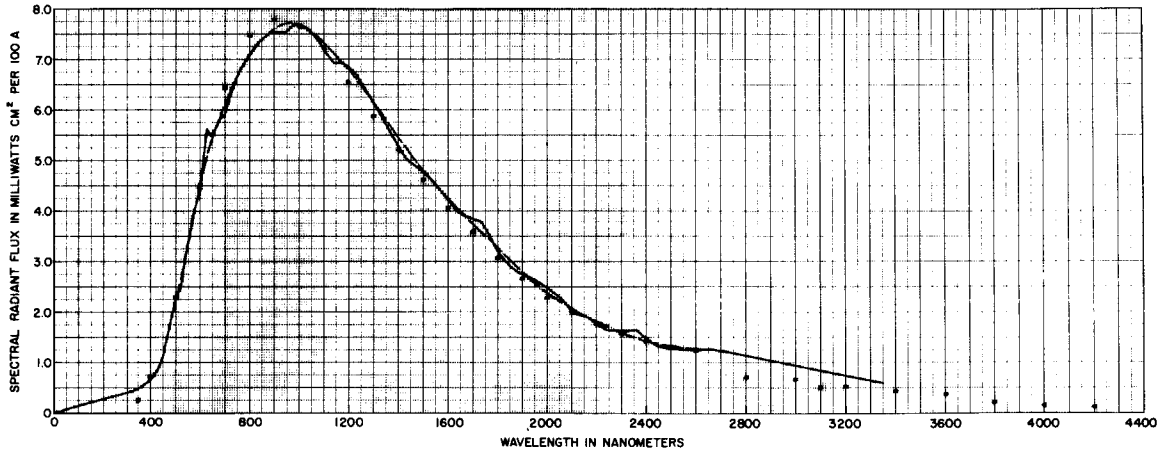
For normalization, each p°_i and p°_{ij} may be divided by the sum of all the values. An alternate procedure for normalization which we followed in the present computer program is to divide each p°_i by $(\Sigma P^\circ_i + \Sigma P^\circ_{ij})\,\Sigma p^\circ_i / \Sigma P^\circ_i$. Thus p°_i's are first normalized and p°_{ij}'s are calculated from these. This gives the fractional energy in each range. Dividing the fractional energy by the width of the wavelength band, we obtain the energy per unit wavelength or the ordinate for a histogram representing the spectral-radiant flux. The spectral-radiant flux may also be represented as a graph drawn by the x-y plotter of the computer, x being the wavelength at the midpoint of each range and y the normalized spectral-radiant flux in units of watts cm^{-2} nm^{-1}. The x-y points are joined together by straight lines. The area under the graph is unity by virtue of the normalization procedure.

If the assumption is valid that the energy incident on the thermopile is proportional to the voltage output, the input data for the computer may be the voltage signals instead of the energy P_i and p_i. The output of the computer program for the filter factors gives for each of the $(2N + 1)$ ranges the filter factors, the calculated value of the signal, and the percentage difference between the calculated and observed signals. The output of the computer program for the spectral-radiant flux gives for each of the $(2N + 1)$ ranges the fractional energy in the range, the ordinate for the histogram, and the coordinates for the x-y plotter. The program also plots the graph of spectral radiant flux.

The graphs for six typical cases are presented in Figures A-4 and A-5. The continuous lines in all graphs are those drawn automatically by the x-y plotter. The numbers along the x and y axes, the dots in the three graphs of Figure A-4, and the dashed curve in the second and third graphs of Figure A-4 were added later.

The first graph in Figure A-4 is an accuracy check on the two computer programs. Measurements were made on the quartz-iodine lamp, QL-74, operated at 6.5 amp. The filter factors were computed. These filter factors and the original values of the thermopile voltage were supplied to the second computer program, which resulted in the graph shown here. The dots indicate the NBS values (multiplied by a normalization factor) of





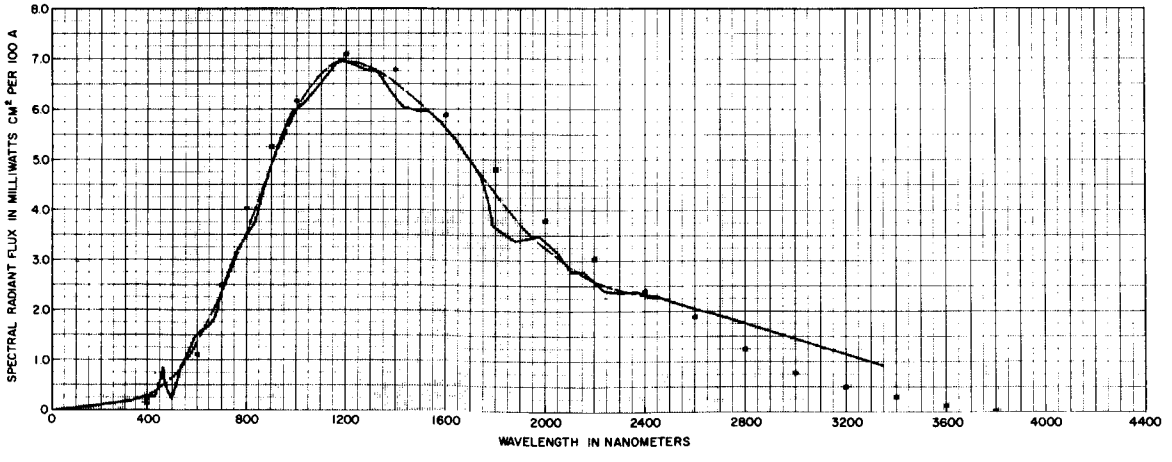


Figure A-4–Spectral Radiant Flux of Tungsten Lamps by Filter Method

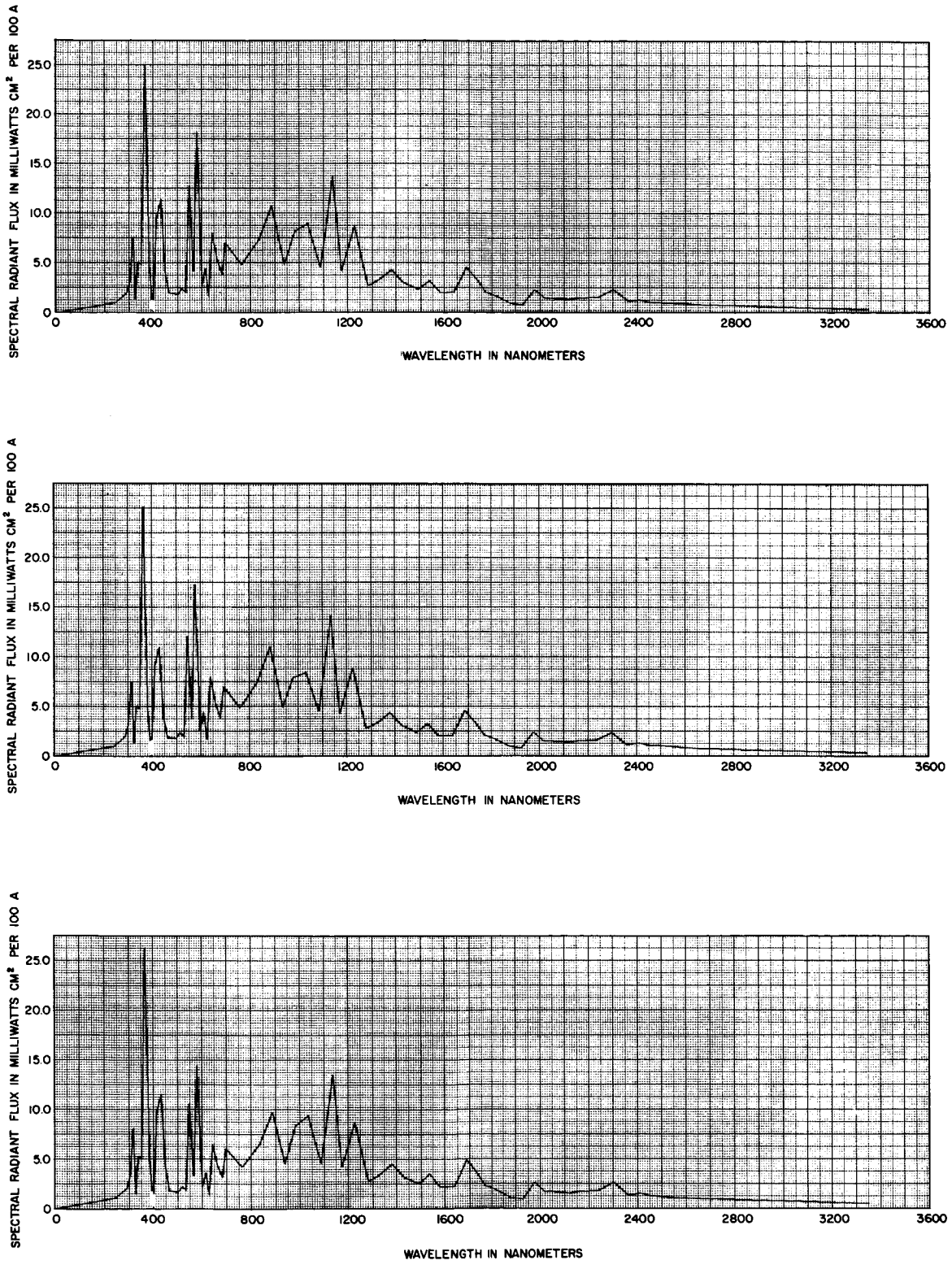


Figure A-5–Spectral Radiant Flux of Mercury-Xenon Lamp by Filter Method

the spectral-radiant flux, on which was based the computation of the filter factors. The close agreement between the graph and the original data points shows the accuracy of the program.

Another set of measurements was made at a later data on the same lamp, operated at the same current. The output signals were slightly different. The continuous curve of graph 2 in Figure A-4 was obtained when these signals were fed to the computer program along with filter factors determined earlier. The dashed curve has been drawn by hand to smooth out the minor discontinuities in the computer graph. The dots are the NBS values, the same as those in graph 1. The average difference between the computer results and the original data is about 4 percent.

Graph 3 shows the results of measurements made on a standard tungsten-ribbon lamp (lamp 185, operated at 25 amp) viewed through the pyrex envelope. The continuous curve was drawn by the x-y plotter; the dash curve was drawn by hand. The dots show the values of spectral-radiant flux given by the NBS calibration table of the same lamp. The computer curve shows clearly the shift in wavelength of maximum output; lamp 185 was operated at a lower temperature than lamp QL-74, and the maximum occurs around 1200 nm.

In the wavelength range up to 3983 and beyond 27010, the agreement between the computed results and the original data is poor, because the computer gives an averaged value for the range. These graphs are based on results from the thirty monopass filters only.

Figure A-5 shows the results of measurements made with thirty monopass filters and four Eppley filters on the 2500-kw mercury-xenon arc. Values of P°_{λ}, the incident spectral-radiant flux, for computation of the filter factors were taken from the 1963 Summer Workshop Report.[1] Graph 1 of Figure A-5 was obtained by using as input data the same values of the thermopile signals as were used for calculation of the filter factors. Graph 2 is the spectral-radiant flux of the same source, the lamp operated at 2500 kw, based on a different set of measurements. Graph 3 is the spectral-radiant flux of the 2500-kw lamp when operated at 1500 kw, calculated using the same filter factor as for the two other graphs. All the major groups of lines of the Hg and Xe spectrum are shown clearly in the graphs. Although the graphs lack the high degree of wavelength resolution obtainable from a monochromator scan, the wavelength resolution of these graphs is adequate for a study of the energy absorbed by test surfaces. The considerable saving in time fully justifies this simpler procedure.

Table A-2 presents the numerical values of wavelengths, filter factors, signal strengths, and computed results in one particular case, Hg-Xe lamp operated at 1500 watts. The first five columns are the input data, and the last six are the output results. The first two columns give the range λ_1 to λ_2 for each filter. The first and last rows are terminal ranges 2000-2828 A and 27010-40000 A. The filter factors F_i, F_{ij} determined from measurements on the Hg-Xe arc at 2500 kw are given in the next two columns. These data are common to all Hg-Xe spectra. The signal strengths s_i for the 1500-kw source are given in column 5.

Columns 6 and 7 show p°_i, fractional energy in the wavelength range of a filter, and p°_{ij}, fractional energy in the range intermediate between a filter and the next. The last four columns give the results λ and p°_{λ} required for the x-y plotter. As explained earlier

$$p^{\circ}_i = F_i \, s_i \, N \text{ , where } N \text{ is a normalization factor;}$$

Table A-2
Spectral Radiant Flux of Mercury-Xenon Arc Lamp (1500 watts)
(computed by the filter technique – input and output data of the computer)

Range		Filter Factor		Signal Energy × 10^4			Wavelength Flux × 10^5		Wavelength Flux × 10^5	
from λ_1	to λ_2	range F_i	interm. F_{ij}	$\mu\omega$ s_i	range p_i°	interm. p_{ij}°	λ	p_λ°	λ	p_λ°
2000			2.22			86.0	2414.0	1.04		3.31
2828	3012	.225	.112	212	38.7	21.2	2920.0	2.11	3044.0	1.41
3076	3264	.1457	.023	1275	150.8	5.4	3170.0	8.03	3283.0	4.95
3302	3458	1.615	.113	62	81.3	66.8	3380.0	5.21	3525.5	4.72
3583	3787	.801	.139	783	509.5	71.2	3690.0	26.26	3862.5	
3938	4081	.0036	.805	325	1.45	158.5	4009.5	.102	4163.0	9.66
4245	4415	.331	.298	728	195.5	66.5	4330.0	11.50	4496.5	4.08
4578	4742	.264	.256	129	27.7	18.0	4660.0	1.69	4797.0	1.63
4852	5128	.185	.331	282	42.4	24.4	4990.0	1.54	5185.5	2.12
5243	5418	.167	1.602	232	31.4	151.9	5330.5	1.79	5490.0	10.55
5562	5758	.158	1.826	493	63.4	202.3	5660.0	3.23	5828.5	14.35
5899	6121	.420	.353	139	47.4	26.8	6010.0	2.14	6158.5	3.57
6196	6404	.312	.283	112	28.4	48.2	6300.0	1.37	6442.0	6.34
6480	6815	.980	.074	178	141.7	19.6	6647.5	4.23	6846.5	3.11
6878	7082	1.018	1.485	148	122.4	466.4	6980.0	6.00	7641.5	4.17
8201	8499	.858	2.215	275	191.7	723.7	8350.0	6.43	8874.0	9.65
9259	9553	.573	.814	290	135.0	469.4	9401.0	4.44	9834.0	8.35
10115	10586	1.009	.268	539	441.7	273.6	10350.5	9.38	10885.5	4.57
11185	11615	2.003	.092	357	580.8	117.3	11400.0	13.51	11756.5	4.14
11898	12703	1.184	.072	727	699.2	67.7	12300.5	8.69	12826.0	2.75
12949	13652	.443	.336	676	243.0	153.7	13300.5	3.46	13823.0	4.50
13994	14702	.552	.217	481	215.7	89.3	14348.0	3.05	14885.0	2.44
15068	15631	.531	.235	454	195.9	85.5	15349.5	3.48	15829.0	2.16
16027	16774	.525	.256	396	168.8	105.7	16400.5	2.26	16883.0	4.85
16992	17707	.758	.000658	397	244.5	0.21	17349.5	3.42	17707.5	2.15
17708	18115	.248	.376	410	82.6	57.7	17911.5	2.03	18283.5	1.71
18452	19147	.221	.086	397	71.1	21.3	18799.5	1.02	19273.5	.84
19400	20100	.503	.086	436	178.0	24.8	19750.0	2.54	20175.5	1.65
20251	20949	.315	.173	430	110.0	41.2	20600.0	1.58	21086.5	1.50
21224	21977	.478	.0238	330	128.0	6.34	21600.5	1.70	21995.0	1.76
22013	22786	.874	.481	195	138.5	113.6	22399.5	1.79	23005.0	2.59
23224	23976	.389	.084	310	97.8	16.6	23600.0	1.30	24032.0	1.48
24088	24913	.601	.0911	205	100.0	17.0	24500.5	1.21	24990.0	1.10
25067	25932	1.086	.154	98	86.4	24.6	25499.5	1.00	26061.0	.95
26190	27010	2.588	6.70	35	73.6		26000.0	.89		
	40000					492.9			33505.0	.38

$$p^{\circ}_{ij} = F_{ij} \left(p^{\circ}_{i} + p^{\circ}_{j} \right) .$$

$$p^{\circ}_{\lambda} = \frac{p^{\circ}_{i}}{\lambda_2 - \lambda_1} \text{ or } \frac{p^{\circ}_{ij}}{\lambda_{j1} - \lambda_{i2}}$$

It should be emphasized that the filter technique is not an independent method, but is complementary to the more complete and accurate monochromator method. The filter factors are known from the measurements made on a known source of similar spectral-energy distribution. The filter technique gives a rapid evaluation of short-time fluctuations in the spectral-radiant flux.

REFERENCES

1. Thekaekara, M.P., Ojalvo, M.S., and Galli, A.J., "Solar Radiation and Simulation" in "Final Report of the Goddard Summer Workshop Program," X-320-63-264, edited by E. Klein, Goddard Space Flight Center, Greenbelt, Md. (1963) pp. A-1 to A-63

2. Final Report on Contract NAS-5-5490, June 8, 1964, to Goddard Space Flight Center (unpublished)

3. MacKenzie, D., "Use of Interference Filters for Spectrophotometry of Solar Simulators" in Midsession Progress Report, July 28, 1964, Goddard Space Flight Center (unpublished) p. A-2

4. Stair, Ralph, Johnston, Russel G., Halback, E.W., "Standard of Spectral Radiance for the Region of 0.25 to 2.6 Microns," J. Res. Natl. Bur. Std. (U.S.) 64A, 291 (1960)

5. Greenler, Robert G., "Optical Filters," Appendix O in Strong, J., Concepts of Classical Optics, W. H. Freeman and Co., San Francisco (1958) p. 580

6. Born, M. and Wolfe, E., Principles of Optics, Pergamon Press, New York (1959) p. 346

7. Jenkins, F.A., and White, H.E., Fundamentals of Optics, McGraw-Hill Book Co., Inc., New York (1957) p. 284

8. Stair, R. and Schneider, W.E., "Standards, Sources, and Detectors in Radiation Measurements," paper presented at the Symposium on Thermal Radiation of Solids, Section III, Measurement Techniques, San Francisco, Calif. (March 1964)

9. Eisenman, W.L., Bates, R.L., and Merriam, J.D., "Black Radiation Detector," J. Opt. Soc. Am. 53, 729 (1963)

EVALUATION OF A 6-FOOT SPHERICAL INTEGRATOR AS A WIDE-ANGLE SOURCE

Eugene I. Mohr and Nathan J. Miller

INTRODUCTION

The TIROS meteorological satellites contain several radiometers, one of which is a five-channel medium-resolution scanning radiometer designed to collect radiation in five wavelength bands by optically filtering the radiation. Channel 3 is sensitive to the visible part of the spectrum. The filter in this channel transmits radiation in the range 0.2-4.5 microns.

The Physics Branch of the Aeronomy and Meteorology Division originally calibrated channel 3 by using light from a 2740-candlepower tungsten-filament lamp, diffusely reflected from a sheet of white Kodak paper.

Team A of the 1963 Summer Workshop, in collaboration with members of the A&M Division Staff, made an independent calibration by use of standard tungsten-ribbon lamps. The resulting calibration agreed reasonably well with the previous calibration.[1]

When in flight, the radiometer scans a wide-angle surface (about 5 by 5 degrees). It is therefore desirable to calibrate the radiometer by use of an extended source. For this reason, the Physics Branch of the A&M Division purchased a 6-foot integrating sphere (Figure A-6) for possible use as a wide-angle source.

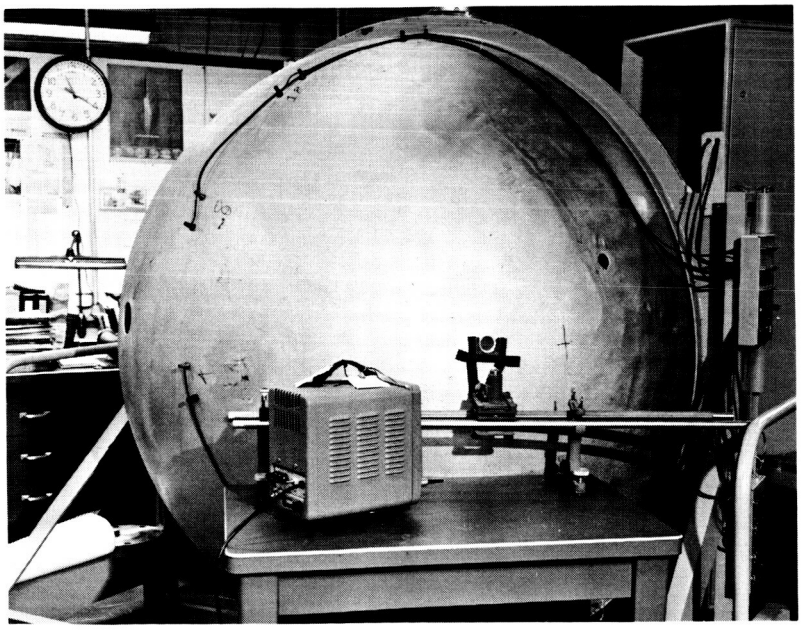


Figure A-6-A & M Division's 6-Foot Integrating Sphere

THEORY OF THE INTEGRATING SHPERE

The principle of spherical integration was proposed in 1892 by Sumpner[2] who showed that, if a source of light is placed inside a hollow sphere coated internally with a perfectly diffusing coat, the luminance of any portion of the surface due to light reflected from the rest of the sphere is everywhere the same and is directly proportional to the total flux emitted by the source.

It is easy to show[3] that the theoretical expression for this relation, in terms of the single reflectance ρ of the perfectly diffusing coat of the sphere wall due to an infinite number of reflections, is given by

$$\Phi = \frac{F}{4\pi r^2} \frac{\rho}{1-\rho} \tag{1}$$

where Φ is the total radiant flux reaching unit area of the spherical surface by reflection, F the total flux emitted by the source, and r the radius of the sphere.

THEORETICAL SPECTRAL IRRADIANCE OF THE INTEGRATING SPHERE

Assume that the A&M spherical integrator is one for which the flux falling on unit spherical surface, due to the multiple reflections, may be represented by equation (1). The spectral irradiance J_λ^c, of the diffusing coat due to these reflections, may be expressed in terms of the spectral irradiance of the quartz iodine lamp, J_λ^s, and of the spectral reflectance of the paint, ρ_λ. The resulting expression may be written as

$$J_\lambda^c \propto \frac{J_\lambda^s}{W_s} \frac{\rho_\lambda}{1-\rho_\lambda} \tag{2}$$

where W_s is the total flux emitted by the quartz-iodine lamp. The spectral irradiance of this surface as observed through a windowless port in the opposite wall will be given by

$$J_\lambda^P \propto \rho_\lambda J_\lambda^c \propto \left(\frac{J_\lambda^s}{W_s} \frac{\rho_\lambda}{1-\rho_\lambda}\right) \rho_\lambda \tag{3}$$

This may be normalized to the total energy, W_t, which a thermopile detects at the windowless port by choosing a normalizing factor, N_p, so that

$$W_t = N_p \int_{\lambda_1}^{\lambda_2} \frac{J_\lambda^s}{W_s} \frac{\rho_\lambda^2}{1-\rho_\lambda} d\lambda \tag{4}$$

Equations (2) to (4) refer only to the spectral irradiance of the paint on the surface opposite the port due to the multiple reflections from the spherical surface. To this should be added the contribution made by the source directly. This will modify equation (4) to the form

$$W_t = \int_{\lambda_1}^{\lambda_2} \left[N_p \frac{J_\lambda^s}{W_s} \frac{\rho_\lambda^2}{1-\rho_\lambda} + N_p' \frac{J_\lambda^s}{W_s} \rho_\lambda \right] d\lambda \tag{5}$$

The normalizing factors involve the inverse-square correction, effect of finite size of lamps, shields, holes, etc. In the case of a dimensionless source placed at the center of the sphere, N_p and N'_p are the same and equation (5) will simply reduce to the expression

$$W_t = N_p \int_{\lambda_1}^{\lambda_2} \frac{J_\lambda^s}{W_s} \frac{\rho_\lambda}{1 - \rho_\lambda} \, d\lambda \qquad (6)$$

Equation (6) is the expression for the total radiant flux which would fall on unit area of the surface at the port, had it not been cut out to make the windowless port. Hence this is the energy which should be received by a detector located at the port. However, in the case of a source which is not located at the center of the sphere, the normalizing factors may be different, so that equation (5) will not reduce to equation (6).

Figure A-7 shows four spectral curves which represent different spectral distributions normalized to the same total energy W_t. Curve 1 shows the Johnson curve of solar energy, and curve 4 shows the spectral-flux distribution of a quartz-iodine lamp. Curve 2 shows the spectral curve of the spherical integrator which may be expected on the basis of equation (4), while curve 3 shows the corresponding distribution which may be expected on the basis of equation (5) by assuming completely arbitrary normalizing factors. Curves 2 and 3 show the modification in the spectral distribution of a source, since the integrating action of the sphere amplifies any nonuniformity in the spectral reflectance. This causes the maximum of the curve to shift toward shorter wavelengths. In other words, the maximum of the spectral-radiant flux of the spherical integrator may be expected to shift away from the maximum of the spectral-radiant flux of the source (quartz-iodine lamp), and to shift toward the maximum value of the Johnson curve. Moreover, the maximum is somewhat sharper than that of the source because the integrating action amplifies the intensity in the range of high reflectance and reduces the intensity in the ultraviolet and infrared where the reflectance is low.

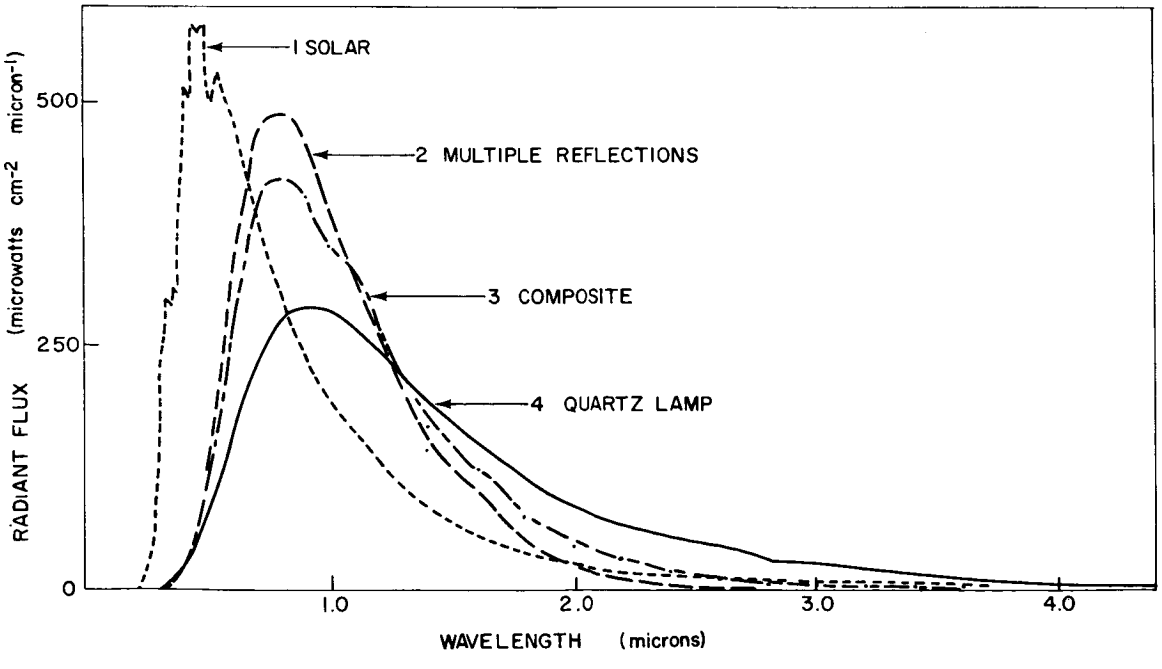


Figure A-7-Theoretical Spectral Curve of Sphere

In this theoretical treatment, it is assumed that the sphere is perfectly empty so that the flux is not disturbed by any objects, nor by any openings in the wall, and that internally the wall is perfectly diffusing. In practice, it is necessary to have some holes for making measurements. In addition, a source must either be placed inside the sphere, or be able to send its radiant flux into the sphere through some hole. Moreover, there is no paint which follows Lamberts' cosine law exactly. Because of these uncertainties, one cannot make absolute measurements of flux.

THE A&M SPHERICAL INTEGRATOR

The spherical integrator shown in Figure A-6 has an inside diameter of 182.5 cm. It consists of two hollow fiber-glass hemispheres, one mounted on a rigid but movable framework. The second hemisphere is hinged to the first and held shut by means of a latch, which permits easy access to the interior. A small opening at the top, with a fan mounted over it, provides ventilation.

Several coats of a white paint with a magnesium carbonate base, manufactured by the Burch Co., were applied to the inside surface to provide the diffusing surface. The spectral reflectance of this paint was determined by using a Beckman DK-2A spectrophotometer and a Beckman IR7 spectrophotometer. Samples were prepared by spraying small squares of sheet aluminum with the paint. This gave the reflectance relative to M_gO. Absolute values of spectral reflectance were obtained by multiplying the relative values by the known spectral reflectance of M_gO. Figure A-8 shows the spectral reflectance of the Burch paint.

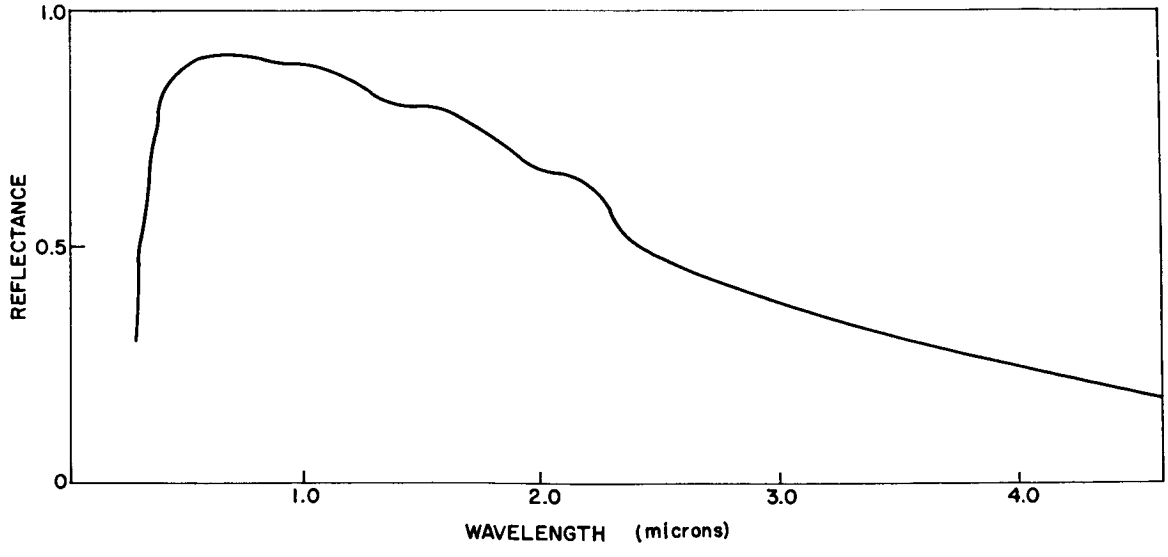


Figure A-8–Reflectance of White Paint

Preliminary trials with external sources gave negative results, making it necessary to turn to some internal source. In order to hold to a minimum the size of objects introduced into the sphere, and also to have some means of varying the intensity of the source, twelve quartz-iodine tungsten-filament lamps, rated at 200 watts each, were mounted inside the sphere. These lamps[4,5] have several characteristics which make them superior to conventional tungsten lamps: They do not blacken. The bulb remains clear during its entire life, since an "iodine cycle" deposits evaporated tungsten back on the filament. The filament can be operated at a much higher temperature than the conventional lamp for the same life, which means greater efficiency. The use of the quartz envelope makes it possible to make the lamps more compact for a given power rating.

The output of a quartz-iodine lamp has been compared wavelength by wavelength with a standard tungsten lamp operated at 3000°K.[6] Its spectrum through the iodine vapor is relatively lower in the green-yellow region and higher in the violet region below 0.43 microns than that of the conventional lamp. This effect is small if there is only enough iodine to maintain the iodine cycle. Because of its size and characteristics, the quartz-iodine lamp has been found suitable as a standard of spectral irradiance.[7]

The quartz-iodine lamps were arranged in a circle (Figure A-9) around two small ports, 3.8 cm in diameter, provided to make intensity measurements. Small shields were used so that the lamps could not be seen directly from either port. The lamps, operated at 29.8 volts, were controlled by separate switches so that the total flux available in the sphere could be varied in steps. Diffusing quartz windows, 6 mm thick and 3.45 cm in diameter, were made to fit exactly into the observation ports.

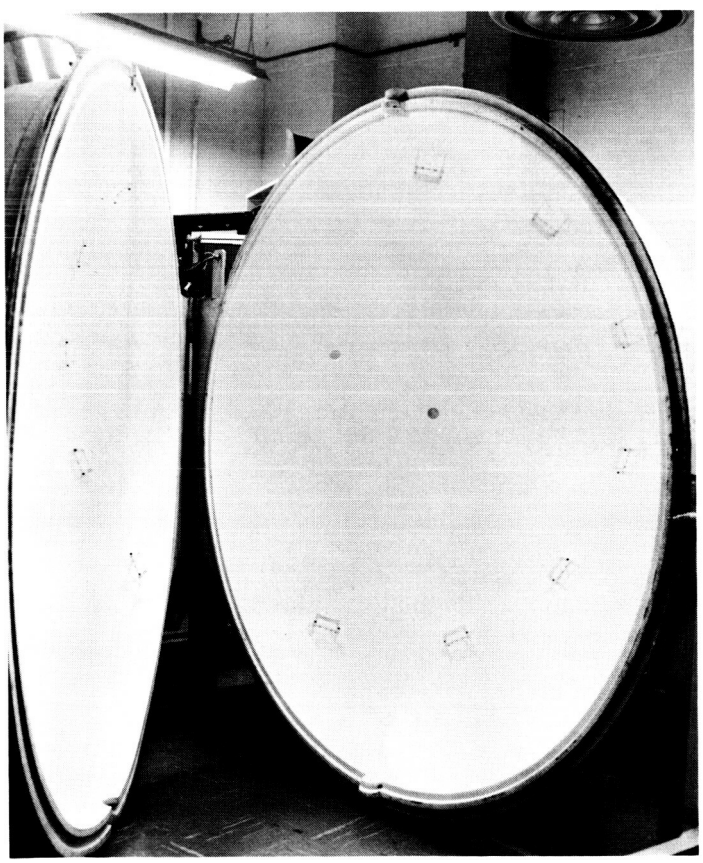


Figure A-9—Arrangement of Quartz-Iodine Lamps in Spherical Integrator

CALIBRATION OF THE SPHERICAL INTEGRATOR

The calibrated Eppley water-cooled thermopile No. 4928A was used to measure the radiant flux at the windowless port, and then to determine the flux through the diffusing quartz window. The absorber of the thermopile was placed 6 cm from the port. The exact distance is immaterial as long as its field of view is completely filled. Measurements were made when the twelve quartz-iodine lamps were operating at 29.8 volts. Additional measurements were made after turning off one of the lamps at a time until only one lamp acted as a source. In addition, flux from the sphere was measured for each of the twelve lamps by using only one lamp at a time. This was repeated for various combinations of two and three lamps. The average value for each combination was used for the calibration.

The calibration of the thermopile is known in terms of the energy falling on the absorber expressed in watts/cm^2. In order to know the power received by the thermopile, the above must be multiplied by the area of the detector.

The power received by the thermopile must be given in terms of the power per square meter radiated by the diffusing surface. Let x watts/m^2 be diffusely reflected into a solid angle 2π. Then the diffuse radiation in the normal direction is x/π. The energy dP_t falling on the effective aperture of the thermopile, dA' from an annulus of area dA of the diffusely reflecting surface (Figure A-10) is equal to

$$dP_t = \frac{x}{\pi} \frac{dA \cos\theta \cdot dA' \cos\theta}{r^2}$$

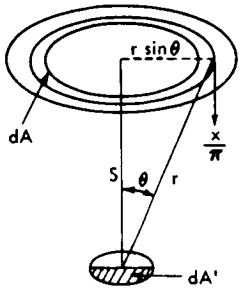


Figure A-10–Energy Incident on Thermopile

where

$$dA = 2\pi r \sin\theta \cdot r d\theta$$

$$\therefore dP_t = 2x\, dA' \cos^2\theta \sin\theta\, d\theta$$

$$P_t = -\int_0^{\theta} 2x\, dA' \cos^2\theta\, d(\cos\theta)$$

$$\therefore P_t = -2x\, dA' \frac{\cos^3\theta}{3}\Big]_0^{\theta} \qquad (7)$$

The water-cooled thermopile 4928A has values of dA' and θ respectively equal to 0.83 cm^2 and 10°35'. This gives the result

$$P_t = -2x\,(8.3 \times 10^{-5}) \frac{\cos^3\theta}{3}\Big]_0^{10°35'} = 2.78 \times 10^{-6}\, x \text{ watts.}$$

The value of x (the power diffusely reflected by unit surface of the sphere into the solid angle 2π) will be

$$x = 3.603 \times 10^5\, P_t \text{ watts/m}^2 \qquad (8)$$

Equation (8) was used to calculate the power per square meter at the port in terms of the power (P_t) received by the thermopile. The resulting data is given in Table A-3, and plotted in Figure A-11. The columns designated by V_0 and x_0 give the data for the windowless port; V_q and x_q give the data with the quartz window in place.

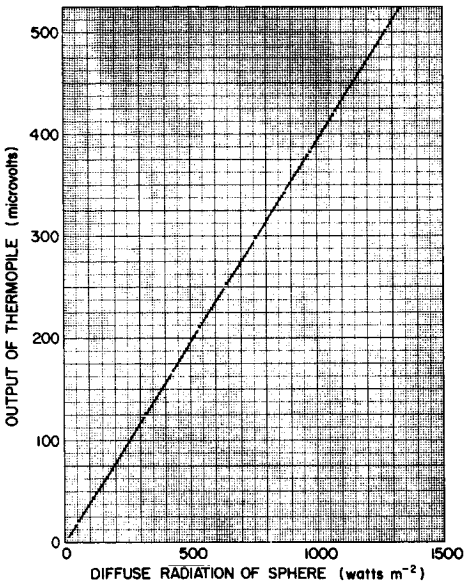


Figure A-11–Calibration of Sphere

Table A-3
Calibration for Energy per Unit Area of Spherical Integrator
(v in microvolts; x in watts/m²)

Number of Lamps	V_0	x_0	V_q	x_q
12	501	1273	413	1050
11	464	1180	382	972.8
10	423	1076	348	886.4
9	382	972.8	315	803.2
8	342	871.7	282	719.2
7	298	759.6	247	629.8
6	253	645.3	209	533.8
5	211	538.9	176	449.8
4	168	429.1	139	355.3
3	126	322.4	105	262.0
2	86	214.0	71	175.0
1	44	106.0	37	87.0

SPECTRAL CURVE OF THE SPHERICAL INTEGRATOR

A Perkin-Elmer double-pass monochromator, Model 99, was used to measure and plot the spectral-radiant flux of a quartz-iodine lamp placed about 43 cm from the slit. A similar chart was made of the flux at the windowless port of the spherical integrator when twelve quartz-iodine lamps were turned on. The ratio of the deflections in the latter case to the former was obtained over the wavelength range measured. The product

of the spectral radiance of the quartz-iodine lamp (as supplied by NBS) by the above ratio, wavelength by wavelength, gave a relative spectral radiance of the spherical integrator. This relative value was normalized to the total radiant flux, W_t, which was measured at the port by means of the thermopile. The spectral distribution was thus normalized to the same total energy used to normalize the theoretical curves in Figure A-7. The above measurements were repeated when the quartz window was in the port of the sphere.

The spectral distribution of the integrator, viewed at the windowless port, was determined in terms of microwatts per square centimeter per micron as seen by the thermopile and plotted in Figure A-12. Similar calculations were made for the distribution when viewed at the quartz window. The resulting data for both cases were changed to watts per square meter of diffuse radiation originating at the extended diffuse source (sphere wall or quartz window) by use of equation (8). Table A-4 gives the resulting data for the spectral distribution when ten quartz-iodine lamps were used. The columns designated by x_0 give the data for the distribution when the sphere was viewed through the open port; the columns designated by x_q show the corresponding data when the sphere is viewed through the quartz window.

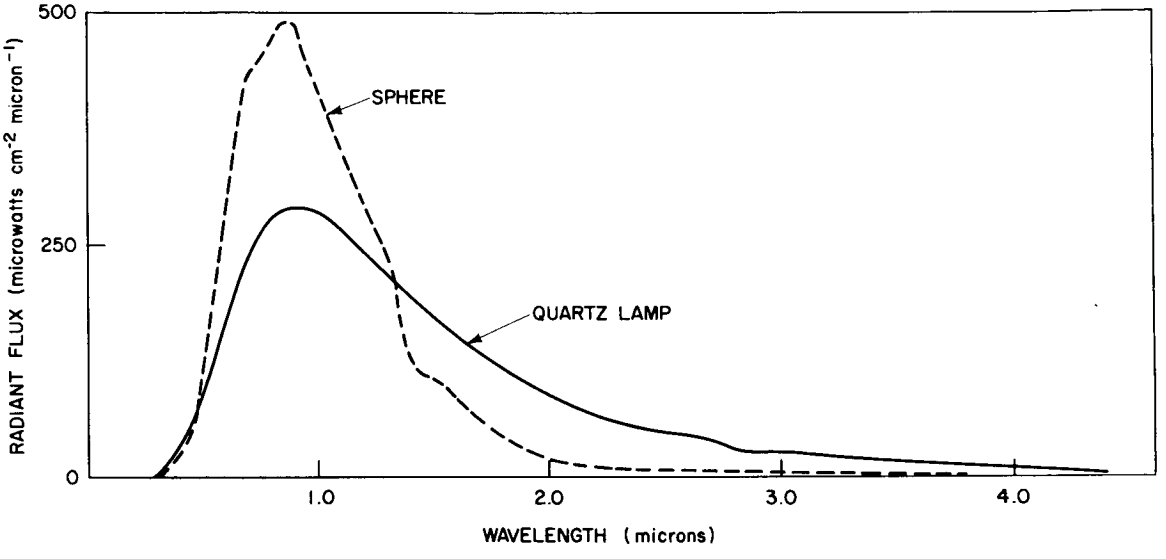


Figure A-12–Experimental Spectral Curve of Sphere

To obtain the spectral distribution for any number of lamps from 1 to 12, the data in Table A-4 may be multiplied by the ratio of the intensity for n lamps to the intensity for ten lamps. These ratios are given in Table A-5. R_0 gives the ratio in the case of the windowless port, and R_q in the case of the quartz window.

FLUX MEASUREMENTS WITH THE RADIOMETER

Each channel of the radiometer has a built-in amplifier. The gain of the amplifier is adjusted so that the output of the radiometer has a range of 0.0 to 5.25 volts when viewing an extended source having the most likely range of intensities. A similar range of intensities must therefore be used in the integrating sphere in order to avoid saturation of channel 3 of the radiometer. For this reason, from one to seven quartz-iodine lamps

Table A-4
Spectral Distribution of Spherical Integrator

(λ in microns; x in watts $m^{-2}\ \mu^{-1}$)

λ	x_{λ_0}	x_{λ_q}	λ	x_{λ_0}	x_{λ_q}
0.3	.380	.314			
0.4	6.872	5.733			
0.5	32.543	22.335			
0.6	83.183	63.229	2.6	2.171	1.482
0.7	126.602	97.213	2.7	1.842	1.190
0.8	135.960	110.677	2.8	1.462	.887
0.9	142.392	112.022	2.9	1.346	1.020
1.0	122.072	101.967	3.0	1.256	1.003
1.1	103.214	86.451	3.1	1.169	.938
1.2	85.962	72.551	3.2	1.053	.865
1.3	69.003	58.774	3.3	.936	.789
1.4	38.302	44.490	3.4	.789	.606
1.5	31.343	29.576	3.5	.703	.511
1.6	25.670	22.100	3.6	.586	.426
1.7	17.076	14.214	3.7	.526	.356
1.8	11.989	10.701	3.8	.440	.210
1.9	7.632	5.708	3.9	.350	.162
2.0	5.380	4.565	4.0	.323	.122
2.1	4.357	3.894	4.1	.233	.060
2.2	3.038	2.590	4.2	.203	.020
2.3	2.865	2.468	4.3	.147	.008
2.4	2.515	2.166	4.4	.117	.004
2.5	2.282	1.961	4.5	.087	.003

Table A-5
Ratio of Intensities of Spherical Integrator
to Number of Lamps

n	R_0	R_q
12	1.18282	1.18522
11	1.09613	1.09750
10	1.00000	1.00000
9	.90386	.90620
8	.80994	.81140
7	.70575	.71052
6	.59961	.60222
5	.50069	.50742
4	.39872	.40080
3	.29952	.29558
2	.19888	.19743
1	.09851	.09815

were used at a time when the windowless port was viewed by the radiometer. When the diffusing quartz window was in the port, from one to nine of the lamps were used at a time. This gave sixteen pairs of values for calibration purposes.

THE INTEGRATING SPHERE AS A WIDE-ANGLE SOURCE

Each channel of the radiometer, when in orbit on TIROS, looks at an extended or wide-angle source having a field-of-view of about 5 by 5 degrees. A similar wide-angle source ought to be used when calibrating channel 3. When the radiometer looks through the windowless port at the opposite inside wall of the sphere, it looks at such a source. This was checked by moving the radiometer across the port horizontally as well as vertically. It was found that the axis could be displaced in either direction from the center of the port by more than one centimeter without any change in detector response when the leading edge of the radiometer was 4 cm from the port. Moreover, with the axis of channel 3 in line with the normal to the port at its center, a rotation of the channel axis through 5 degrees toward either side produced no change in detector response. One may assume that, when channel 3 is centered at the port, with the leading edge not more than 4 cm away, the field-of-view is completely filled.

CALIBRATION OF THE RADIOMETER

The output, in volts, of channel 3 of the radiometer was measured for various values of the radiant flux emerging from the windowless port and also through the diffusing quartz window.

To obtain the energy detected by the radiometer which gives rise to a given voltage output, it is necessary to know the total responsivity, φ_λ, of channel 3 as a function of wavelength. This responsivity is composed of the product of the following factors: the reflectance of the prism, the reflectance of the chopper, the transmittance of the lens-filter system, and the responsivity of the detector flake. The values of φ_λ for the radiometer #306 are shown in Table A-6.

Table A-6
Total Spectral Responsivity of Channel 3
(Radiometer #306)

λ	φ_λ	λ	φ_λ	λ	φ_λ
0.3	.181	1.8	.542	3.3	.465
0.4	.294	1.9	.544	3.4	.445
0.5	.318	2.0	.546	3.5	.432
0.6	.348	2.1	.532	3.6	.413
0.7	.400	2.2	.515	3.7	.379
0.8	.391	2.3	.510	3.8	.352
0.9	.464	2.4	.508	3.9	.313
1.0	.549	2.5	.495	4.0	.280
1.1	.565	2.6	.492	4.1	.240
1.2	.564	2.7	.490	4.2	.199
1.3	.552	2.8	.491	4.3	.154
1.4	.548	2.9	.495	4.4	.116
1.5	.551	3.0	.495	4.5	.084
1.6	.552	3.1	.491	4.6	.038
1.7	.547	3.2	.478	4.7	.018

The total energy originating at the diffuse surface which produces the actual radiometer response may be found by use of the relation,

$$x = R \int_{\lambda_1}^{\lambda_2} x_\lambda \varphi_\lambda \, d\lambda \simeq R \sum_{\lambda_2}^{\lambda_2} x_\lambda \varphi_\lambda \, \Delta\lambda \qquad (9)$$

where x_λ is the radiant flux for ten lamps at wavelength λ (Table A-4), R is the ratio of intensity for n lamps compared to the intensity for ten lamps (Table A-5), and φ_λ is the responsivity of channel 3 of the radiometer (Table A-6). Equation (9) was applied to each intensity viewed by the radiometer at the windowless port and also at the quartz window. The resulting data are given in Table A-7 and plotted in Figure A-13. For purposes of comparison, the data obtained by use of the Kodak paper is included in Table A-7 (labelled x_p), and in Figure A-13 (by a dotted curve).

Table A-7
Calibration of Channel 3 (Radiometer #306)

(V in volts; x in watts/m²)

V	x_s	x_p
5.25	380	
5.00	360	282
4.75	340	264
4.50	321	249
4.25	302	234
4.00	283	219
3.75	265	206
3.50	248	193
3.25	231	179
3.00	214	166
2.75	198	153
2.50	182	139
2.25	166	126
2.00	150	113
1.75	134	100
1.50	118	86
1.25	101	73
1.00	85	59
0.75	65	45
0.50	45	

DISCUSSION OF THE RESULTS

The calibration curve for radiometer #306, obtained by using the spherical integrator as a wide-angle source, shows that the output of the radiometer departs slightly from a completely linear relation with the intensity. From theoretical considerations, one would expect this relation to be completely linear; the slight departure may be inherent in the geometry of the radiometer.

In order to check the accuracy of the calibration, an independent calibration was made by the Physics Branch of the A&M Division. This was done by means of the usual

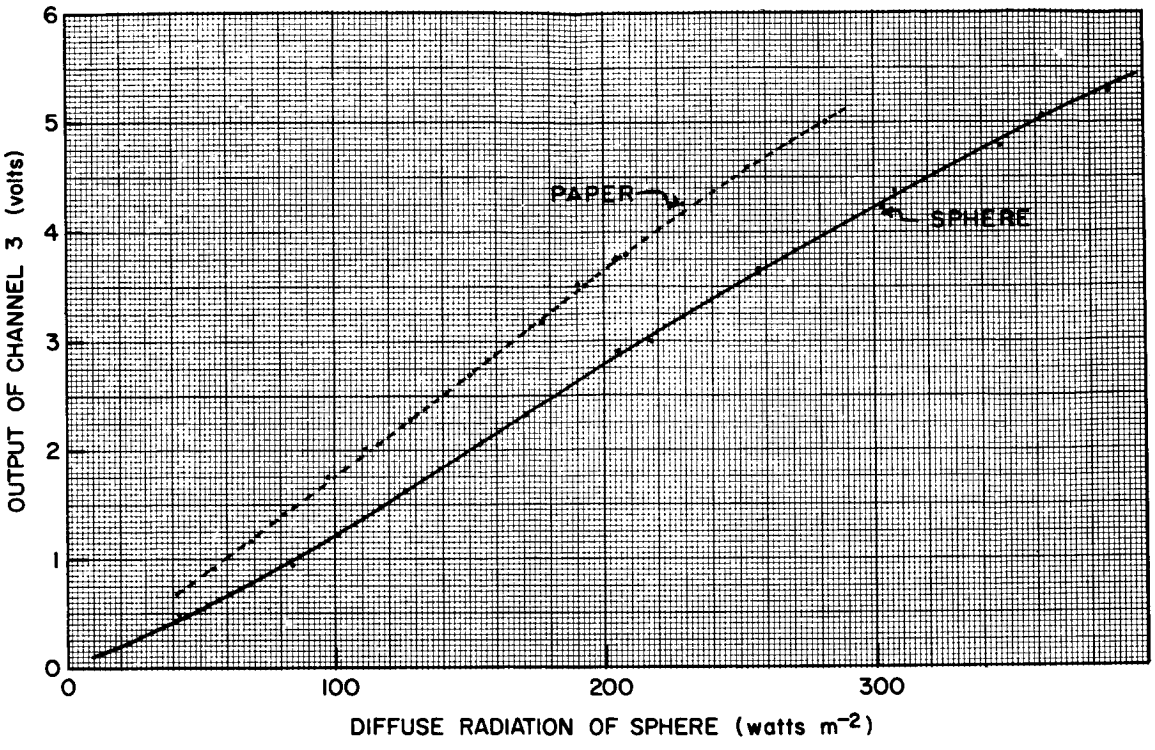


Figure A-13-Calibration of Radiometer #306

A&M procedure of using a sheet of diffusing white Kodak paper illuminated by a tungsten-filament lamp. The resulting calibration has a slightly different slope. Intensities for the corresponding output voltages of the radiometer are from 28 to 35 percent less than the values obtained by means of the spherical integrator. Both sets of data have been checked for internal consistency. From this standpoint, both would seem to be correct within experimental error. Obviously, one or the other set, or possibly both sets, of calibration data must be incorrect.

In spite of the internal consistency of the calibration by means of the Kodak paper, its accuracy has been questioned. The reason for this is to be found in the information which has been received from the orbiting TIROS weather satellites. In applying the calibration of the radiometer to this information, one obtains a value for the earth's albedo which is considerably below the value it ought to have, as determined from the heat balance of the earth. Some of this discrepancy can apparently be explained in terms of degradation of the radiometer for several days after the satellite has been launched.

It was precisely this problem which led to a restudy of the calibration by means of the integrating sphere. For this reason, we would like to believe that the data obtained by means of the sphere are more nearly correct than those obtained by means of the Kodak paper. If this is true, the problem in regard to the albedo would at least be partially solved.

CONCLUSIONS

It is apparent that further study must be given to the two methods of calibration. It is necessary to reevaluate the experimental techniques. Furthermore, the theoretical implications must be reexamined.

REFERENCES

1. GSFC Final Report of the Summer Workshop Program, 1963
2. Sumpner, W.E., Phys. Soc. Proc. 12, 10 (1892)
3. Walsh, J.W.T., "Photometry," London, Constable & Co. (1958), p. 258
4. Zubler, E.G., and Mosby, F.A., Illum, Engr. 54, 734 (1959)
5. Moore, J.A., and Jolly, C.M., General Electric Co. J. 29, 99 (1962)
6. Studer, F.J., and Van Beers, R.F., JOSA 54, 945 (1964)
7. Stair, R., Schneider, W.E., and Jackson, J.K., Applied Optics 2, 1151 (1963)

Page intentionally left blank

HORIZON SENSOR: RATIO OF SIGNAL TO NOISE

Eugene I. Mohr

INTRODUCTION

When a spacecraft is equipped with television cameras, the latter must be triggered when their line-of-sight is normal to the earth's surface. For this reason, a horizon sensor, properly adjusted, is mounted on the satellite to trigger the camera at the instant that the sensor sees the horizon.

The spacecraft spins at some predetermined rate. Because of this spin, the horizon sensor looks at empty space during a portion of its revolution and consequently gives a zero signal (actually a signal equal to the amplifier noise) during this time. As the sensor's field-of-view sweeps from space across the horizon, it starts to scan what may be thought of as essentially a blackbody at a temperature of more or less 300°K. As a result, the signal put out by the sensor rises rapidly above amplifier-noise level to a fairly high value, determined by the energy received from the earth. The output of the sensor remains at a relatively high value until the field-of-view sweeps off the earth at the opposite horizon.

The amplifier electronics in the sensor is designed to cut off any pulse which rises above a predetermined voltage level. This has the effect of cutting off the long high-output plateau, and of leaving only a short pulse produced at the instant when the energy being detected rises quite rapidly above noise level as the field-of-view sweeps from space across the horizon. It is this short pulse, produced as the sensor looks at the horizon, which triggers the camera.

Suppose that, as the field-of-view of the sensor sweeps across the ground, it suddenly sees a much colder cloud. This will cause a sharp decrease in the energy seen by the horizon sensor, followed by a rapid rise in energy as the field-of-view again looks at the ground. This rise in energy would produce another pulse, seen by the sensor as a false horizon, which could also trigger the camera at a moment when it might be looking at space. To prevent this, the circuit is adjusted so that the latter pulse will usually be above cutoff. It becomes evident that the cutoff point must be adjusted so that the sensor will not signal a false horizon as it looks at a cold cloud, and also so that the signal given by the true horizon is below cutoff, but still well above the amplifier-noise voltage.

THE ELECTRICAL NOISE IN THE THERMISTOR BOLOMETER

Any resistive element exhibits an electrical noise due to random motions of the charge carriers. This is known as thermal noise or Johnson noise. This noise is independent of the frequency and may be shown[1] to be equal to

$$V_j = (4kTR_d\Delta f)^{1/2} \tag{1}$$

At ambient temperature (300°K), equation (1) has the value of

$$V_j = 1.287 \times 10^{-10} (R_d \Delta f)^{1/2} \text{ volts.} \tag{2}$$

The resistance of the bolometer in the horizon sensor Model 13-205Sc is a function of temperature, and has a value of 2.2×10^5 ohms at 300°K. It is customary to analyze noise in terms of a frequency bandwidth of one cycle per second ($f = 1$ cps). This gives a value of the Johnson noise voltage of

$$V_j = 1.287 \times 10^{-10} (2.2 \times 10^5 \times 1)^{1/2} = 6.04 \times 10^{-8} \text{ volts.}$$

AMPLIFIER NOISE OF THE HORIZON SENSOR

The noise figure of the amplifier was measured by Barnes Engineering at 300°K and "at source impedance of 125K ohms." By means of these measurements they calculated the noise figure for a bandwidth of one cps at two frequencies. These values are listed below.

Frequency (f)	Noise Figure (F_N)
30 cps	19.3 db
200 cps	12.35 db

Keeping in mind that the bridge circuit cuts the amplified signal in half, and making use of the definition of a decibel, the noise voltage V_n for a bandwidth of one cps may be calculated by use of equation (3).

$$V_{ni} = (\text{antilog } .05\, F_N)\, V_j \text{ volts.} \tag{3}$$

Using equation (3), the noise voltages at 30 cps and 200 cps respectively are

$$V_{ni} = (\text{antilog } 0.05 \times 19.3) \times 6.04 \times 10^{-8} = 5.57 \times 10^{-7} \text{ volts}$$

$$V_{ni} = (\text{antilog } 0.05 \times 12.35) \times 6.04 \times 10^{-8} = 2.51 \times 10^{-7} \text{ volts.}$$

Assuming that the amplifier noise voltage has the form, $V_{ni} = k f^{-1/2}$, the value of the constant k is given by $k = V_{ni}\, f^{1/2}$. This gives values of the constant k at frequencies of 30 cps and 200 cps respectively of

$$k = 5.57 \times 10^{-7} (30)^{1/2} = 3.05 \times 10^{-6}$$

and

$$k = 2.51 \times 10^{-7} (200)^{1/2} = 3.55 \times 10^{-6}.$$

By assuming an intermediate value for the constant k, equation (4) will be

$$V_{ni} = 3.25 \times 10^{-6} f^{-1/2}. \tag{4}$$

A number of values of the noise voltage were calculated by means of equation (4). These values are listed in Table A-8 and are plotted in Figure A-14.

Table A-8
Noise Voltage Versus Frequency

f(cps)	V_{ni} (volts)	f(cps)	V_{ni} (volts)
5	14.50×10^{-7}	50	4.60×10^{-7}
10	10.30×10^{-7}	75	3.75×10^{-7}
15	8.40×10^{-7}	100	3.25×10^{-7}
20	7.29×10^{-7}	125	2.91×10^{-7}
25	6.51×10^{-7}	150	2.66×10^{-7}
30	5.94×10^{-7}	175	2.46×10^{-7}
40	5.14×10^{-7}	200	2.30×10^{-7}

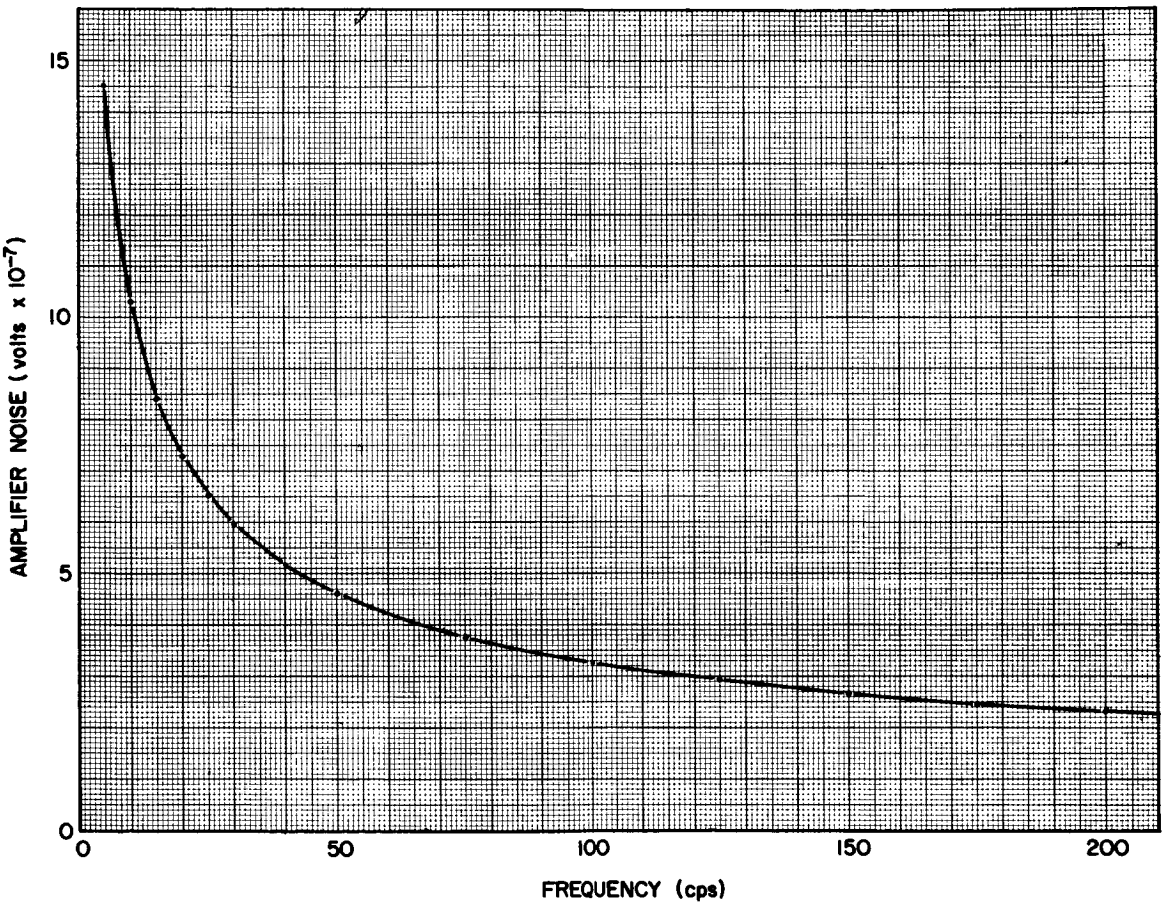


Figure A-14–Spectrum of Amplifier Noise

The average amplifier noise voltage for a bandwidth Δf may be found by use of equation (5).

$$V_n = (\Sigma V_{ni}^2)^{1/2} \tag{5}$$

Values of V_{ni} were read from the curve in Figure A-14, at intervals of one cps, from 4 cps to 200 cps. Since the nominal bandwidth of the amplifier is 4-30 cps, the noise voltage for this bandwidth, by equation (5), is

$$V_n = (2254 \times 10^{-14})^{1/2} = 4.75 \times 10^{-6} \text{ volts.}$$

The bandwidth of the amplifier can be as high as 4-200 cps, in which case the corresponding voltage is

$$V_n = (4239 \times 10^{-14})^{1/2} = 6.51 \times 10^{-6} \text{ volts.}$$

The corresponding value of the Johnson noise voltage for the 4-30 cps band is 3.08×10^{-7} volts, and for the 4-200 cps band is 8.47×10^{-7} volts. The amplifier-noise voltage is approximately ten times as large as the Johnson noise voltage, and is therefore the significant factor in determining the signal-to-noise ratio S/N.

EVALUATING THE TOTAL ENERGY SEEN BY THE HORIZON SENSOR

Making use of Simpson's rule, the total energy received by the bolometer may be found by equation (6):

$$W = \frac{\Omega}{\pi} \int_{\lambda_1}^{\lambda_2} P_\lambda(T_i)\, \varphi_\lambda \, d\lambda \simeq \frac{\Omega}{\pi} \sum_{\lambda_1}^{\lambda_2} P_\lambda(T_i)\, \varphi_\lambda \, \Delta\lambda \qquad (6)$$

where $P_\lambda(T_i)$ is the spectral-radiant power of a blackbody at temperature T_i, φ_λ the relative spectral responsivity of the horizon sensor, Ω the solid-angle field-of-view. The value of Ω is 4.92×10^{-4} steradian for the horizon sensor, which gives a value of 1.57×10^{-4} for Ω/π. The values of the spectral responsivity are shown in Figure A-15.

An energy slide rule was used to evaluate $\Sigma P_\lambda(T_i)\, \varphi_\lambda \Delta\lambda$ for temperatures of 200°K 240°K and 300°K. These values, in microwatts/cm^2,were respectively 2,004, 5,746, and 18,208. Values of the total energy received by the sensor at these temperatures were found by use of equation (6).

SIGNAL VOLTAGES VERSUS CLOUD TEMPERATURES

The signal voltage of the horizon sensor is a function of the area of the detector optics, A_d, of the energy received from the source, W, and of the responsivity of the bolometer, $\Re$. This is shown in (7).

$$V_s = \Re \cdot A_d W \text{ volts} \qquad (7)$$

The value of A_d is 2.24 cm^2. Barnes Engineering supplied two values for the bolometer responsivity: 240 volts/watt and 2000 volts/watt. This would give low and high values in a bridge circuit of 120 volts/watt and 1000 volts/watt. The first value would seem too low and the second rather high. The responsivity of a bolometer is proportional to the square root of the ratio of the time constant to the area. Comparing the bolometer of the horizon sensor with an immersed and an unimmersed bolometer of known responsivity, values of 1300 to 1400 volts/watt were obtained. This gives 650 to 700 volts/watt in a bridge circuit.

To be on the conservative side, calculations have been made using a bolometer responsivity of 600 volts/watt. Substituting the given values in equation (7), signal voltages were found for the three temperatures.

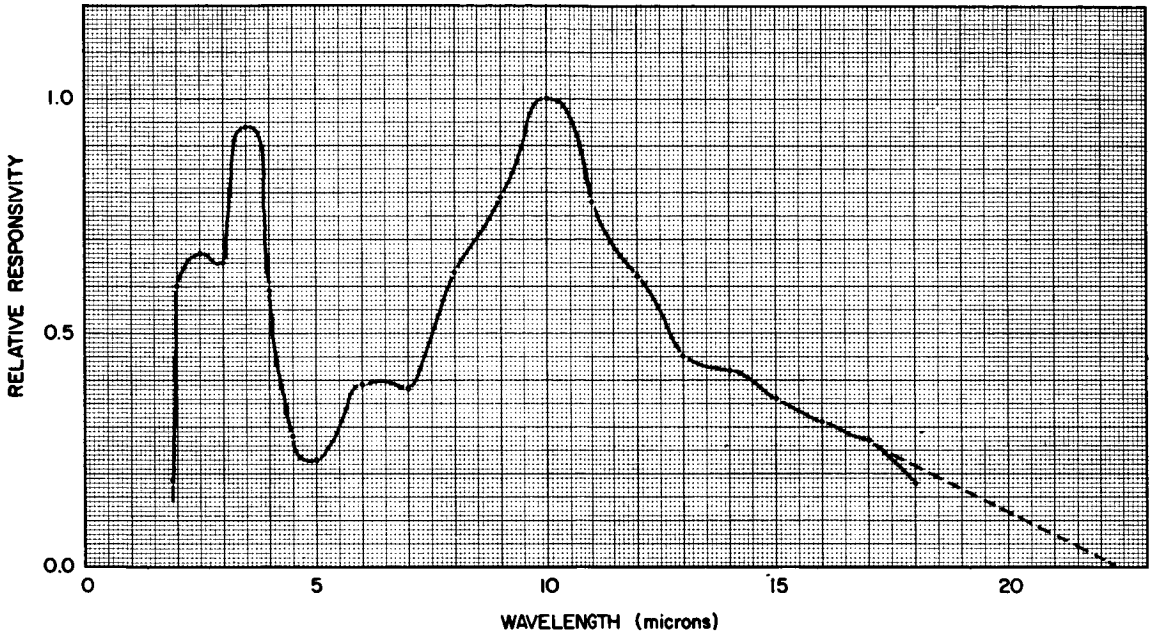


Figure A-15–Spectral Responsivity of the Sensor

Table A-9
Total Energy Seen by Sensor

T (°K)	W (microwatts/cm²)
200	0.313
240	0.898
300	2.85

Table A-10
Signal Voltage Versus Source Temperature

T (°K)	V_s (microvolts)
200	421
240	1,206
300	3,830

SIGNAL-TO-NOISE RATIO

The signal-to-noise ratios for the three temperatures and in the two bandwidths are given in Table A-11.

Table A-11
Signal-to-Noise Ratio

T (°K)	S/N (4-30 cps)	S/N (4-200 cps)
200	90.7	64.7
240	254	185
300	805	588

The spectral responsivity of the horizon sensor as supplied by Barnes Engineering are only relative values normalized to unity. These values may be too high by a factor of two. If this assumption is made, the ratio of signal-to-noise in the horizon sensor listed in Table A-11 must be divided by two. The resulting ratios are given in Table A-12.

Table A-12
Probable Signal-to-Noise Ratios

T (°K)	S/N (4-30 cps)	S/N (4-200 cps)
200	45.4	32.4
240	127	92.5
300	402	294

CONCLUSIONS

Since the signal-to-noise ratio is more than 30:1 for a cloud as cold as 200°K, which may on occasion be seen by the horizon sensor, it would seem safe to assume that the cut-off value of the amplifier may be set at 200°K. The signal-to-noise ratio for the pulse given by the true horizon would be well above a satisfactory level; the pulses of false horizons given by cold clouds would usually be above cutoff, and would thus be eliminated.

ACKNOWLEDGMENTS

The helpful discussions and suggestions of Andrew McCulloch are gratefully acknowledged.

REFERENCE

1. Kruse, R.W., McGlauchlin, L.D., and McQuistan, R.B., "Elements of Infrared Technology," N. Y., Wiley (1962) pp. 236-41

STUDIES ON THE DEGRADATION OF THE FIVE-CHANNEL RADIOMETER

Joseph B. Morris

INTRODUCTION

Housed on NASA'S TIROS meteorological satellites is a five-channel medium-resolution scanning radiometer.[1] Two channels, numbers 3 and 5, are sensitive to reflected solar radiation (UV-Vis-NIR); the other three channels, numbers 1, 2, and 4, respond to radiation emitted from the earth and its atmosphere. The effective spectral response (ESR) of each channel is different and is a function of the spectral region. This is accomplished by the various filters and lenses for each channel. Common to all channels is the aluminum-silicon monoxide-coated prismatic mirror. Each channel has its own half-aluminized, half-black chopper disk. The radiometer is mounted in the satellite with its optical axis inclined 45 degrees to the satellite spin axis. Thus, when one side of the prism views the earth, the other views outer space, the latter being used as a reference.

After launch, all channels appear to degrade, although their degradation patterns are definitely different[1,2] (see Figures 47, 48, 50, and 51 in the TIROS IV Radiation Data Catalog and Users' Manual). However, prelaunch degradation has been observed in the solar channels. This, alone, appears to be evidence of chemical activity or environmental factors which affect the optical system, if all other parameters are controlled.

The purpose of this study is to determine whether any chemical factors are directly or indirectly related to the observed degradation of the five-channel radiometer. Should any of the materials degrade, the optical system could be affected; consequently, changes should be observed in its response. It is assumed that particle bombardment does not have a great effect on the radiometer, and that there is no appreciable degradation of the electronic circuits. The aluminized mirrors on the prism and chopper are coated with SiO (silicon monoxide). The properties of SiO as a protective coating on optical materials has been thoroughly studied by Haas.[3] These protective properties, a function of the rate of deposition, pressure of the vacuum system, and thickness of the film, govern the rate at which SiO is oxidized to SiO_2. Oxidation appears to have some disadvantages: Transmission decreases, and reflectance increases between 9-11 microns. When SiO undergoes disproportionation ($2SiO \rightarrow Si + SiO_2$), higher absorbances are observed (absorbance = $\log_{10} 1/T$, where T is transmittance), and reflectance and transmission change only slightly in the IR.

One primary material which may affect optics is the type of paint used in the radiometer. On TIROS III, "Sicon" flat black was used on the shields; "Murphy's Black" on the choppers, and "Zapon" flat black on the thermistor-bolometers. On TIROS IV and VII, the same paints were used on the choppers and detectors.

Any outgassing of these paints with subsequent deposition on any part of the optical system would probably change the optical properties of the system. Because of the geometry of the radiometer, outgassing by Murphy's Black could affect the radiometer in three ways: (1) change of reflectance of the prism, (2) change of reflectance of the aluminized half of the chopper, and (3) decrease in the transmission properties of the nearest filter (or, as in case of channel 3, the BaF_2 lens). Any outgassing of "Zapon" flat black could

affect the transmittance of the last lens and probably the characteristics of the thermistor-bolometer. "Sicon" outgassing from the shields would probably have the smallest effect.

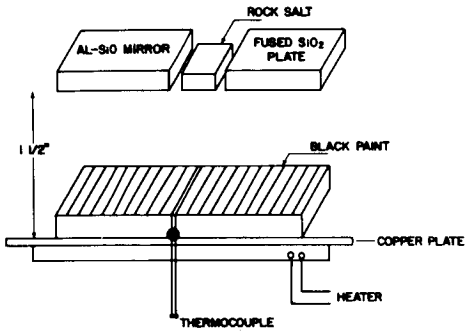


Figure A-16–Experimental Arrangement to Study Outgassing of Paint

EXPERIMENTAL

The experimental arrangement (Figure A-16) to determine the effect of outgassing on the optical properties of the radiometer was simple, but highly effective. As the effective spectral response of all channels falls within the 0.2- to 15-micron region, optical materials were chosen which could be used to make reflectance and transmission measurements for this region. The mirrors used for reflectance measurements were 2 by 2-inch aluminized glass plates covered with SiO. Fused silica plates (2 by 2 inches) were used to cover the 0.2-2.5-micron region for transmission studies, and rock-salt plates were used to cover the 2.5 - 15-micron region. The paints were coated on Al-SiO mirrors and allowed to dry.

The entire arrangement was placed in a vacuum chamber. Before heat was applied, the chamber was evacuated to the lowest pressure obtainable. For the preliminary studies, the pressure was about 1.5×10^{-5} mm; however, for most of the studies, the lowest pressure obtainable was 5×10^{-4} mm. The temperature of the plates was $(100 \pm 5)^{\circ}$C. Short runs for one hour were begun with new mirrors and plates. For 3-hour studies, the same plates and mirrors were used.

The Beckman IR-7 was used to make all IR-transmission and IR-spectral reflectance measurements (only a few reflectance measurements were made in the IR region). The Beckman DK-2A spectroreflectometer was used to make transmission studies on the fused-silica plates and reflectance measurements on the Al-SiO mirrors. Runs were made on all optical materials both before and after exposure to outgassing in the vacuum chamber.

RESULTS

Visual Observations

In every experiment, outgassing was confirmed. The film in most cases was quite uniform, with semi-solid characteristics; it appeared to be polymeric but of low molecular weight. When blank optics were used in another part of the chamber, their characteristics remained unchanged except in one of the 24-hour exposures. Some oil got back in the chamber. Short-time runs eliminated the oil problem.

Murphy's Black

From a chemical viewpoint, outgassing of Murphy's black would have the greatest effect on the optical system. Therefore, it was studied more thoroughly and systematically. Outgassing was studied as a function of time in 1-, 3-, and 24-hour exposures. Figures A-17 and A-18 show the results of transmission studies for the 1- and 3-hour exposures. Transmission studies are not shown for the region 1.2-2.5 microns; however, they remained fairly constant over these regions.

"Page missing from available version"

It appeared that outgassing is rapid at first, and then slows down. The 1-hour exposure had an overall lower transmittance in all spectral ranges. The 3-hour exposure and the 24-hour exposure show a general decrease in absorption, which indicates that further outgassing of the originally deposited outgassed product may have occurred. Furthermore, the background in the IR spectra showed lower absorption. Experiment showed that, after 3 hours, thermal equilibrium was reached between the heated plates and the collector, which was 1-1/2 inches above the heated plate. This temperature was about 85°C. At this temperature, further outgassing could occur on the collecting plates. It was also observed that the deposited film at the end of the first hour was non-uniform, whereas, after the 24-hour exposure, the film was quite smooth. In absorption spectroscopy, particle size and particle orientation have a large effect on the scattering of radiation. Smoother films show less scattering effects than non-uniform films. This probably accounts for the lower background in the IR spectra.

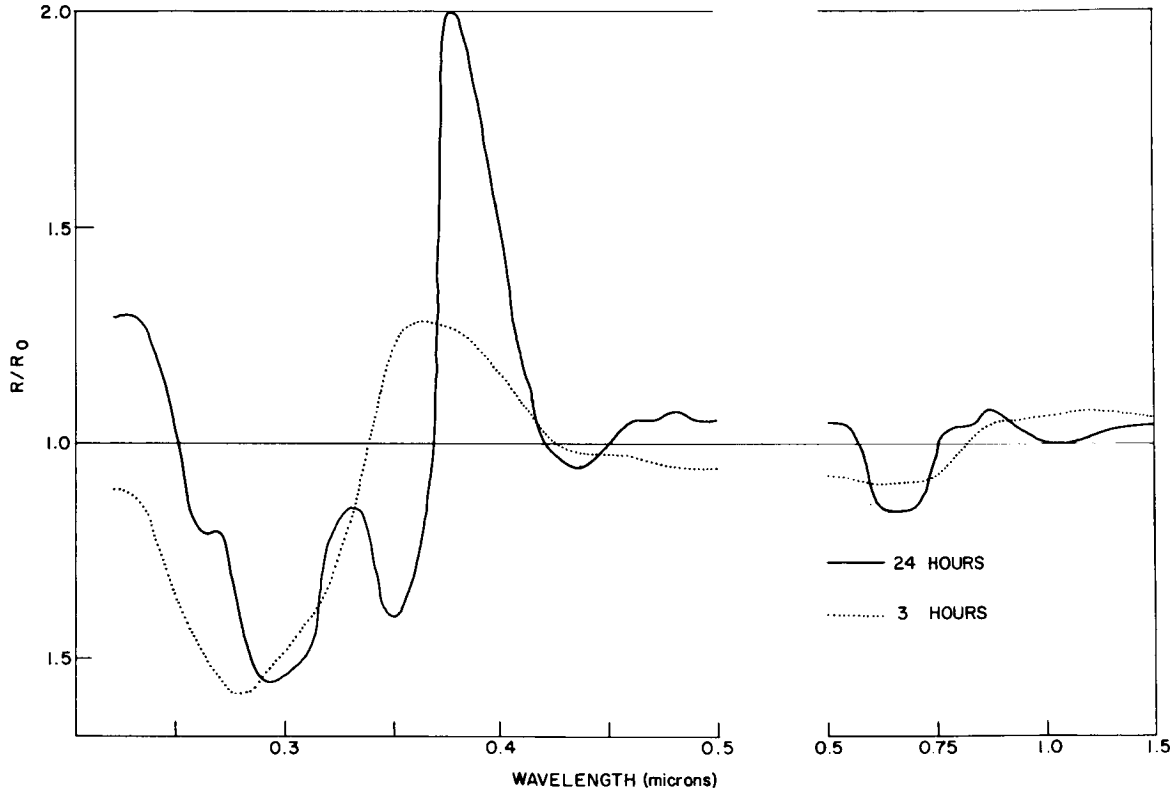


Figure A-19–UV-Vis-NIR Reflectance Ratio of Outgassed Film (Murphy's Black)

Figure A-19 shows the UV-Vis-NIR reflectance ratio of the outgassed film. R_0 is the reflectance of the mirror before exposure. The IR reflectance measurements were made on the 3-hour exposure, but not on the 24-hour exposed film. After 1.5 microns, the R/R_0 decreased steadily to about 0.85 at 2.5 microns, remaining at 0.85 ± 0.05 to 15 microns. The extreme transition between 0.35 and 0.40 microns in R/R_0 values cannot be explained definitively. Since this is the UV cutoff region for Al-SiO mirrors, it may account for the phenomenon. Further experimentation indicated two important facts: Exposure of a blank mirror under the same conditions, without Murphy's black, gave

reflectance measurements, before and after exposure, which were practically identical; reflectance measurements made on the silica plate indicated that the reflectance for pure silica and silica plus film are practically identical. These two experiments rule out the extreme transition in reflectance in the UV as being due to changes in the SiO film and the outgassed film as individual entities. It does not, however, rule out possible interaction of the two components.

A predicted total transmission effect has been calculated on the basis of the changes in the reflectances of the mirrors and the transmission of the two plates after film formation. When radiant energy enters the radiometer, it is first reflected from the Al-SiO prismatic mirror; it is then reflected from the aluminized surface of the chopper and transmitted through the first filter. If the intensity of the radiation is P_0 before it enters the radiometer and P after the first filter, then the actual transmission at this point would be P/P_0 and is designated as T_0. It would take the following form:

$$P/P_0 = T_0 = R_{om} \times R_{oc} \times T_{of}$$

where RO_m and RO_c are the reflectances of the mirror and chopper, respectively. TO_f is the transmission of the filter.

After exposure to outgassing, the film would cause changes in the reflectances of the mirror and chopper and the transmission of the first filter. The transmission of the optical system would then be as follows:

$$T'_0 = R_m \times R_c \times T_f$$

where R_m and R_c are the reflectances of the film and mirror, and film and chopper, respectively; T_f would be the transmission of the film and filter. The overall effective transmission is, therefore, T'_0/T_0, and is given as the calculated T value, T_{calc}.

Values of T_{calc} at small-wavelength intervals for the entire spectral range 0.22 to 14 microns are given in Table A-13. They are also presented graphically in Figures A-20 and A-21. The observed values of R_m/R_{om} and R_c/R_{oc} are the same; they are given in the third column of Table A-13 as R/RO. These results are based on measurements made on samples exposed for 3 hours to Murphy's black. Table A-14, which gives the values averaged over the wavelength ranges of the channels of TIROS IV, shows the correlation between effective spectral response and effective transmission.

In all regions except between 0.35 and 0.42 microns, a decrease in the predicted transmission of energy is noted. From calculations made for the 24-hour exposure, the effective transmittance would be about the same in the UV, but somewhat higher values in the Vis-NIR region. IR reflectance measurements were not made on the 24-hour exposed mirrors.

Zapon Flat Black

Zapon flat black paint outgassed profusely. This is the paint on the detector. The data for the 3-hour exposure are given in Figures A-22 and A-23. Transmission and reflectance data are presented for the entire spectral range. The IR spectrum indicates a series of organics which have outgassed which highly affect the reflectance spectrum. Otherwise, the general effect is much like Murphy's black.

Table A-13
Data on Outgassing from Murphy's Black Paint

λ_μ	T/T_0	R/R_0	$T_{calc.}$	λ_μ	T/T_0	R/R_0	$T_{calc.}$	λ_μ	T/T_0	R/R_0	$T_{calc.}$
.22	.38	.90	.31	1.05	.95	1.07	1.08	5.4	.98	.87	.74
.23	.37	.87	.29	1.10	.94	1.06	1.06	5.6	.98	.87	.74
.24	.51	.78	.31	1.15	.94	1.05	1.04	5.7	.97	.78	.59
.25	.68	.63	.27	1.20	.94	1.05	1.04	5.8	.93	.81	.61
.26	.65	.53	.18	1.25	.94	1.04	1.02	6.0	.98	.84	.70
.27	.58	.46	.12	1.30	.95	1.03	1.02	6.2	.95	.84	.67
.28	.59	.43	.11	1.35	.95	1.02	.99	6.4	.97	.84	.69
.29	.75	.46	.16	1.40	.95	1.01	.97	6.6	.86	.84	.61
.30	.81	.52	.22	1.45	.96	1.01	.98	6.8	.95	.84	.67
.31	.83	.58	.28	1.50	.96	1.00	.96	7.0	.96	.86	.71
.32	.84	.67	.38	2.4	.97	.87	.73	7.2	.95	.85	.68
.33	.85	.79	.53	2.5	.98	.85	.71	7.4	.96	.86	.71
.34	.86	1.0	.86	2.6	.98	.84	.70	7.6	.96	.85	.69
.35	.87	1.23	1.34	2.7	.98	.85	.71	7.7	.91	.85	.66
.375	.90	1.21	1.32	2.8	.98	.84	.70	7.8	.92	.83	.63
.40	.92	1.15	1.22	2.9	.98	.83	.68	8.0	.87	.81	.57
.425	.94	1.0	.94	3.0	.98	.82	.66	8.2	.94	.86	.70
.45	.94	.98	.90	3.1	.98	.82	.66	8.4	.91	.85	.66
.475	.95	.95	.86	3.2	.98	.82	.66	8.6	.95	.86	.70
.50	.94	.94	.84	3.3	.97	.79	.60	8.8	.93	.85	.67
.525	.95	.94	.84	3.4	.92	.77	.54	9.0	.94	.88	.72
.55	.95	.94	.84	3.5	.95	.82	.64	9.3	.95	.87	.72
.575	.95	.94	.84	3.6	.97	.83	.67	9.5	.97	.87	.74
.60	.96	.95	.87	3.7	.98	.84	.70	9.6	.92	.85	.66
.625	.96	.95	.87	3.8	.98	.84	.70	10.0	.97	.84	.69
.65	.96	.95	.87	3.9	.98	.85	.71	10.5	.97	.90	.79
.675	.95	.95	.86	4.0	.98	.85	.71	11.0	.96	.89	.76
.70	.95	.95	.86	4.1	.98	.85	.71	11.5	.97	.87	.74
.725	.94	.95	.84	4.2	.98	.85	.71	12.0	.92	.87	.70
.75	.93	.95	.83	4.3	.98	.86	.73	12.5	.97	.91	.81
.80	.95	1.0	.95	4.4	.98	.87	.74	13.0	.97	.90	.79
.85	.95	1.04	1.03	4.5	.98	.87	.74	13.5	.97	.91	.81
.90	.95	1.06	1.06	4.8	.98	.87	.74	14.0	.97	.92	.82
.95	.95	1.07	1.08	5.0	.98	.87	.74				
1.00	.95	1.06	1.06	5.2	.98	.87	.74				

Table A-14
Average of Calculated Transmittance for Spectral Ranges
of TIROS IV Channels

Channel	Wavelength for max. θ microns	$T_{calc.}$ (average)
1	6.0-6.5	0.65 ± 0.04
2	8.0-12.0	0.85 ± 0.05
3	0.2-6.0	0.28 ± 0.15 (0.22-0.33 microns)
		1.10 ± 0.10 (0.34-0.45 microns)
		0.85 ± 0.02 (0.45-0.75 microns)
		0.95 ± 0.05 (0.75-2.5 microns)
		0.75 ± 0.05 (2.5-15 microns)
4		time reference channel in TIROS IV
5	0.55-0.75	0.85 ± 0.02

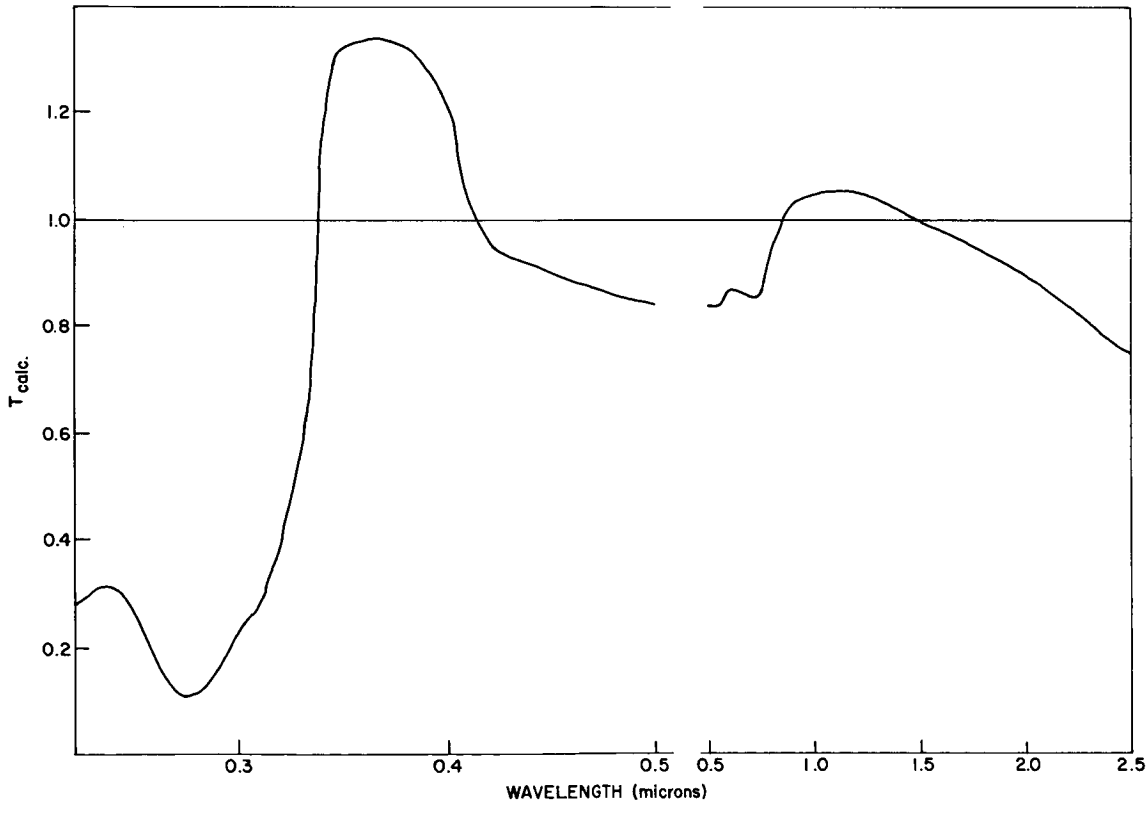


Figure A-20–Calculated UV-Vis-NIR Spectral Distribution of Radiant-Energy Transmission

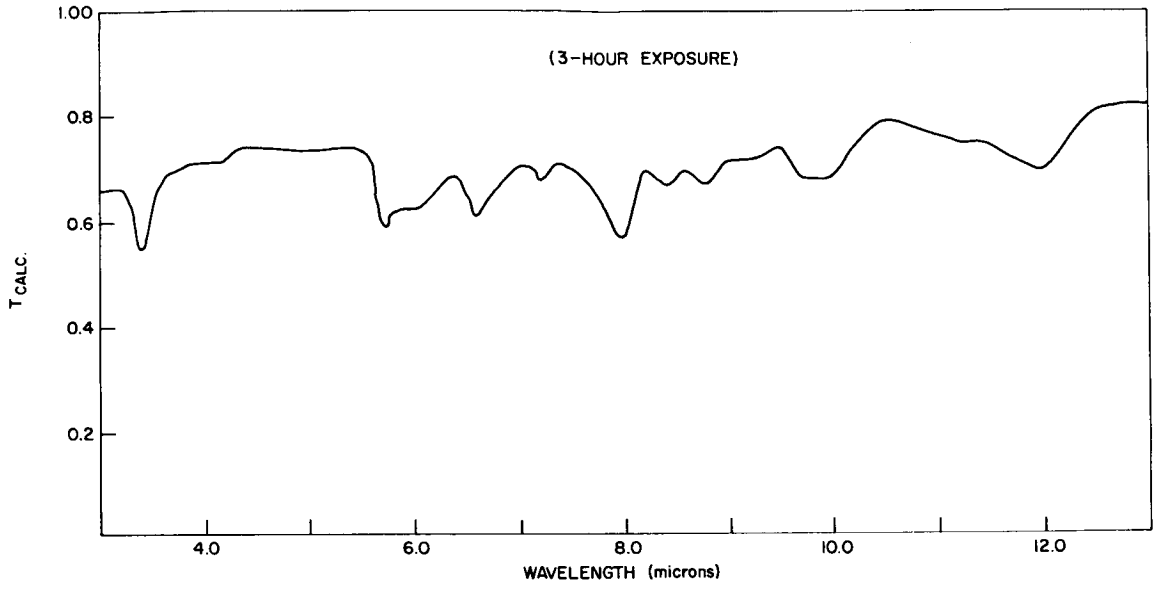


Figure A-21–Calculated IR Spectral Distribution of Radiant-Energy Transmission

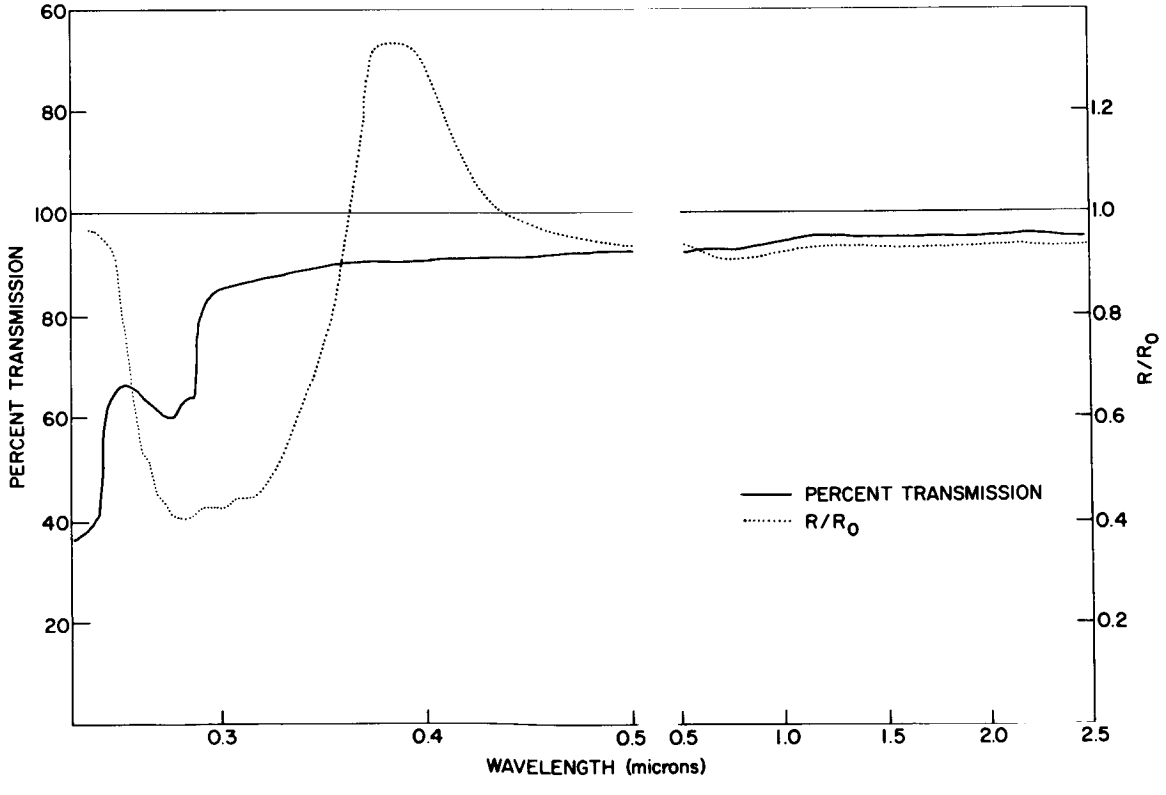


Figure A-22–UV-Vis-NIR Transmission Spectra and Reflectance Ratio of Outgassed Film (Zapon Flat Black)

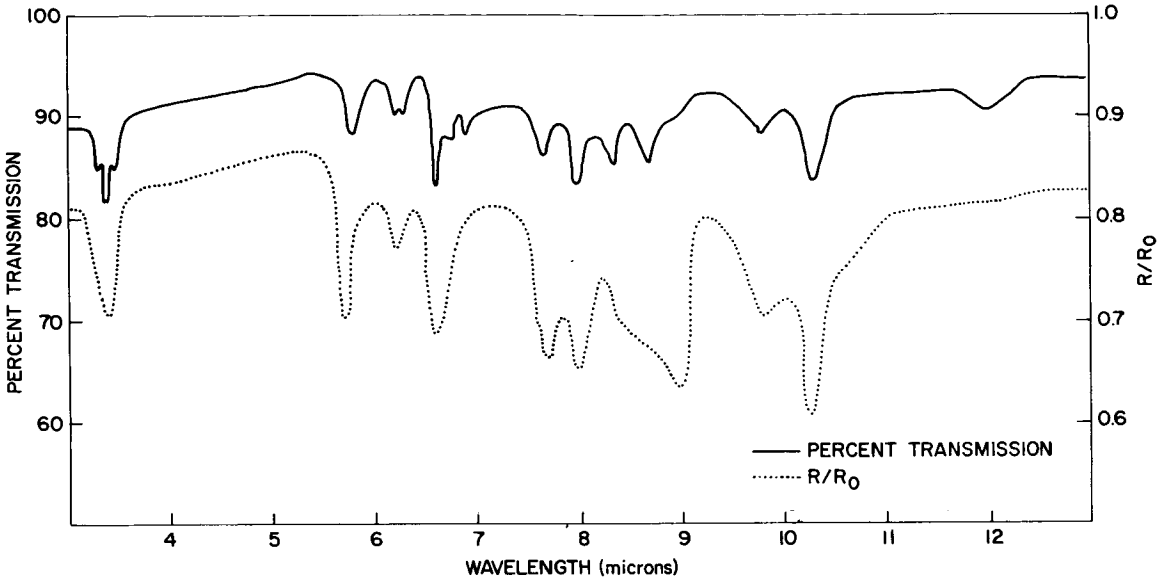


Figure A-23–IR Transmission Spectra and Reflectance Ratio of Outgassed Film (Zapon Flat Black)

Sicon Flat Black

Sicon flat black also outgasses, and the transmission spectrum for a 24-hour exposure is given in Figure A-24. Reflectance measurements were made only in the UV-Vis-NIR. These are given in Figure A-25. Again, the general effect is much like Murphy's black.

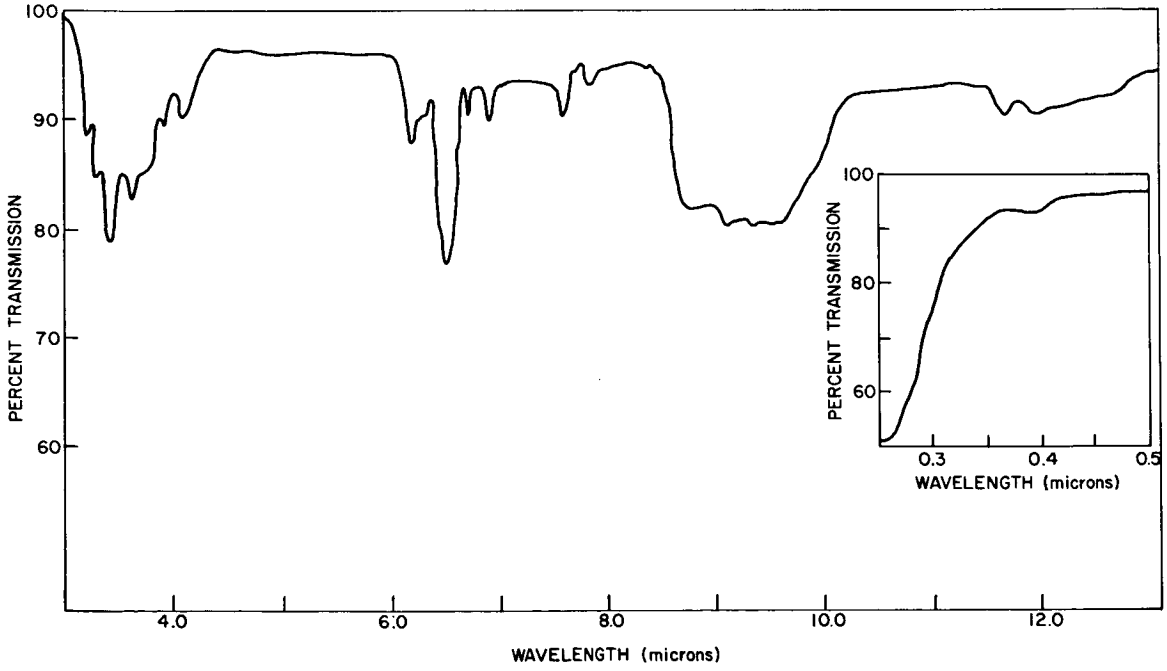


Figure A-24–UV-Vis-IR Spectra of Outgassed Film (Sicon Flat Black)

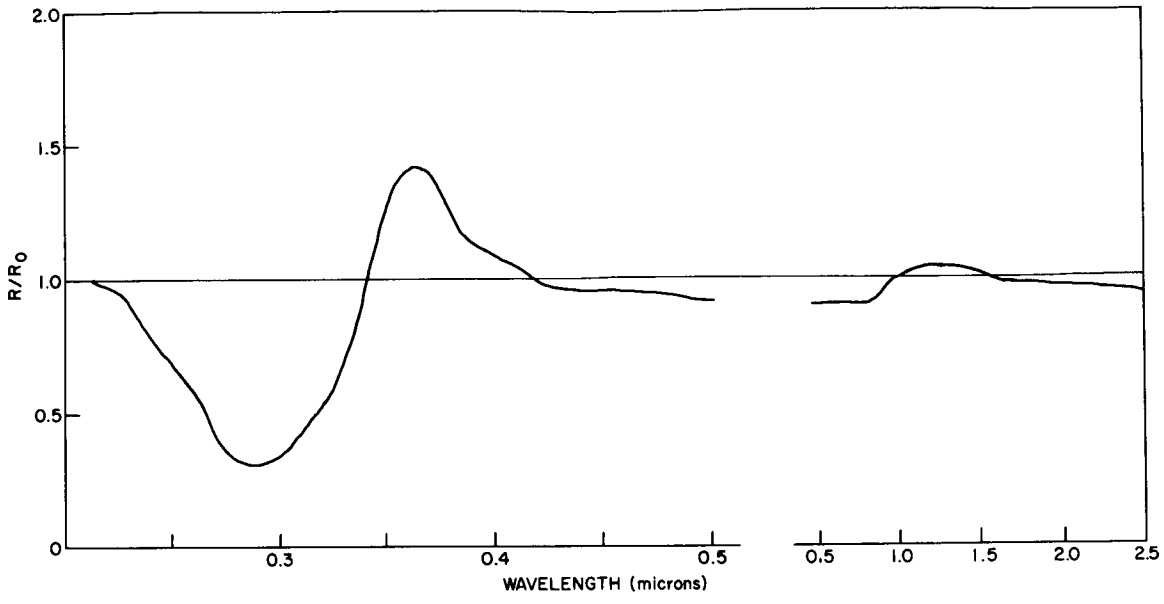


Figure A-25–UV-Vis-NIR Reflectance Ratio of Outgassed Film (Sicon Flat Black)

CONCLUSIONS

All paints used on the radiometer in the TIROS series of satellites outgas under reduced pressure at 100°C. Based on reflectance and transmission measurements, a predicted decrease in the transmission of radiant energy throughout the entire spectral region was verified, and this experimental decrease agrees closely with what was actually observed in all channels of the TIROS radiometers and in all spectral ranges except a narrow band centered at 0.375 microns. Because outgassing appears to be rate-controlled, it appears that the effect does have some correlation with the general phenomenological degradation of signals from the radiometer. Further investigations will require better simulated conditions. It appears further that profuse outgassing of "Zapon flat black" may affect the detecting properties of the thermistor-bolometer. An investigation of the latter seems advisable in the future.

ACKNOWLEDGMENTS

The investigator wishes to acknowledge the help of R. Carter in all experimental vacuum work. The advice and cooperation of Andrew McCulloch and James McLean of the A&M Division are also acknowledged.

REFERENCES

1. "TIROS III Radiation Data Catalog and User's Manual," A&M Division, GSFC, NASA (August 1962)
2. "TIROS IV Radiation Data Catalog and Users' Manual," A&M Division, GSFC, NASA (December 1963)
3. Haas, G., and Salzberg, C.D., J. Opt. Soc. Am., 44, 181-7 (1954)

TELESCOPIC DESIGN FOR AN INTERFEROMETRIC SPECTROPHOTOMETER

Ernest Hilsenrath

As part of Goddard's program for the exploration of planetary atmospheres, the Aeronomy and Meteorology Division is developing an infrared interferometer spectrometer. The interferometer spectrometer will have two missions; exploration of the earth's atmosphere, and exploration of the other terrestrial-type planets, Mars and Venus. The spectrometer is the same for both missions except that the planetary probe instrument will include a telescope.

The instrument without the telescope has an acceptance of 8 degrees. Therefore, for a satellite such as Nimbus, the instrument would cover an area about 100 km across at the surface of the earth. However, an 8-degree acceptance angle is too large for a planetary fly-by. The Mars probe is expected to pass about 5000 km. from the planet; to cover 250 cm on the planet would require that the acceptance angle of the instrument be reduced to 2.8 degrees and this is exactly the purpose of the telescope.

Before discussing the configuration of the instrument with the telescope, it would be instructive to give a brief description of the interferometer spectrometer. Figure A-26 is a schematic diagram of the Michelson interferometer. Light enters from the left and is divided by the beamsplitter. A portion of the beam is reflected to mirror M_1; the other portion is transmitted through the beamsplitter to mirror M_2. Upon reflection from the mirrors, the beams are recombined, giving the well-known light and dark concentric fringes when focused at the plane of the detector. This instrument can be used to make spectrometric measurements by oscillating mirror M_1 back and forth in a line perpendicular to the surface. The relative path length of the two beams is thus changed continuously. If the light source is monochromatic, the detector output is sinusoidal. However, if the source is other than monochromatic, a superposition of sine waves from each wavelength in the source would be detected. A plot showing the variation of detector output as a function of mirror movement is called an interferogram. The spectrum of the source (that is, the function of intensity vs. wavelength) can be obtained by performing a Fourier analysis on the interferogram.

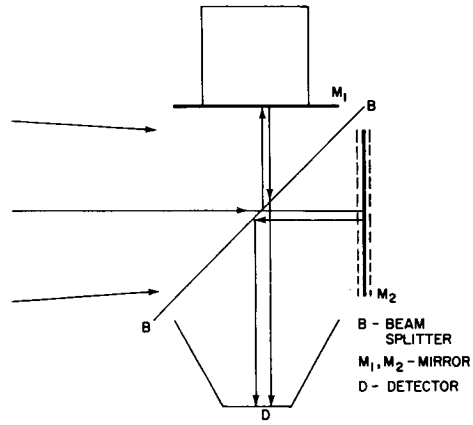


Figure A-26–Michelson Interferometer

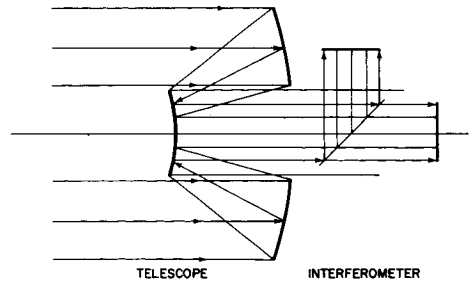


Figure A-27–Telescope and Interferometer

Figure A-27 shows the interferometer with telescope. First-order theory showed that a Cassegrain-type telescope would be adequate if

the diameter of the secondary mirror could be kept to a minimum. For this telescope, the obscuration is about 12 percent. From what was stated earlier, it can be seen that this telescope is really a collector and does not require the precision of an imaging instrument. This type of telescope can be considered the reflective analog of the Galilean telescope, which is effective only at small angles; as in the Galilean telescope, the focal points of the primary and secondary are coincident. The magnification in this case is the ratio of the angles subtended by the image and object, which is approximately 3.

Dr. Irvine C. Gardner did the actual third-order calculations and computed the exact curvature of the mirrors. It turns out that they are both parabolas. The f-number is about 1.3 and since the telescope is really acting as a collector, aberration is not a serious problem.

To conserve weight, and since the surfaces are parabolas which facilitates fabrication, the mirrors will be made of metal. The first step is to grind glass masters conforming to the computed shapes and then to electroform the mirrors from nickel. The final reflecting surface will be gold evaporated on the nickel. The telescope housing and mirror mounts are being designed at Goddard, and the final instrument will be ready for testing before the end of the year.

THE FOURIER TRANSFORM IN INTERFEROMETRIC SPECTROPHOTOMETRY

Robert L. Sharp and Matthew P. Thekaekara

INTRODUCTION

A Michelson-type interferometer is now being designed for use as a spectrophotometer aboard Nimbus and Mariner spacecraft.

Figure A-28 is a sketch of the Michelson interferometer. Light from the spectral source S, whose spectrum is to be analyzed, is collimated by a lens. The beam then is split at G_1, taking two paths, being reflected from mirrors M_1 and M_2 and then being recombined at G_1. Mirror M_2 is fixed, so that the path of that portion of the light reflected by M_2 is constant. Mirror M_1 is movable and therefore the light reflected by it has a variable path length. The two portions of the initial beam have a definite phase relationship determined by the optical paths they traverse in the interferometer.[1]

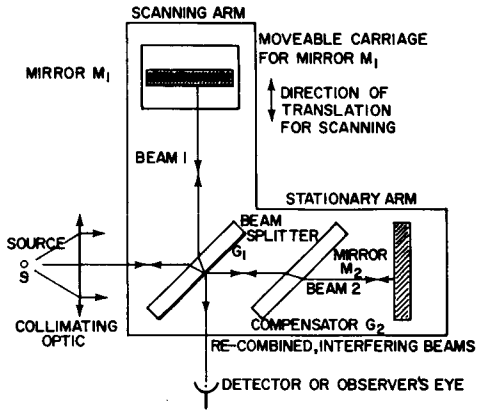


Figure A-28-Michelson Interferometer, Schematic Diagram

Assuming that S is a monochromatic point source, and that x the difference in path length is 0, the two portions of the beam will recombine in phase and strong intensity will be detected. If $x = \lambda/2$, λ the wavelength of the source, then the two portions will be exactly out of phase and the detector will read a minimum. The minimum is zero if the two beam portions are of equal intensities. It can readily be shown that, for a monochromatic point source, the detector output will vary sinusoidally with path-length difference.

If S is not monochromatic but emits many different wavelengths, the detector will produce a signal representing the superposition of sinusoidal signals due to each of the component wavelengths. A plot showing the variation of intensity of signal with path difference is called an interferogram.

In practice, the source of illumination is not a point source but is an extended source. If the two mirrors M_1 and M_2 are adjusted to be exactly perpendicular to each other, a system of circular fringes is observed in the plane where the two beams are brought to a focus. When the Michelson interferometer is used as a spectrophotometer, a detector is placed at the principal focus of the condensing lens so as to observe a small area at the center of the fringe system. Fourier analysis will produce from the interferogram an intensity analysis of the spectrum; namely, the relative intensity as a continuous function of the wavenumber.[2]

Assume it is possible to let the path length difference x vary from $-\infty$ to $+\infty$, and let the function $I(x)$ be the unterminated interferogram of the spectral source under examination. Then the true spectrum $G(w)$ is the Fourier transform of $I(x)$, which is given by

$$G(w) = 2\int_0^{\infty} I(x) \cos wx \, dx,$$

where $w = 2\pi\sigma$, σ being wavenumber.

This interferogram will embody several sources of distortion, one of which is examined in this paper. The distortion with which we are concerned is that caused by the finiteness of the path-length difference. The physical limitations of any interferometer prevent a path-length difference varying between $-\infty$ and $+\infty$.

To analyse this termination of the interferogram, we introduce the terminating function,

$$T(x) = \begin{cases} 1 & \text{if } |x| \leq L \\ 0 & \text{if } |x| > L. \end{cases}$$

If $I'(x)$ denotes the terminated interferogram, then $I'(x) = I(x)\,T(x)$. We call the Fourier transform of $I'(x)$ the measured spectrum, and denote it by $G'(w)$; i.e.,

$$G'(w) = 2\int_0^{\infty} I'(x) \cos wx \, dx.$$

The measured spectrum $G'(w)$ which we obtain is definitely different from the true spectrum $G(w)$, but we would like to measure quantitatively this difference and examine possible methods of improving $G'(w)$.

A commonly used method of changing the observed interferogram $I'(x)$ into another $I''(x)$ is to make $I''(x) = I'(x)V(x)$, where $V(x)$ is an arbitrary weighting function which makes the effect of the sharp discontinuity at $|x| = L$ less significant. This process is called apodization, a word which means etymologically "removing or trimming the feet." A triangular pulse function, $V(x) = 1 - |x|/L$ is a simple form of such a weighting function.

Several authors have written in favor of apodization. A.S. Filler[3] states that the unweighted interferogram may distort the spectrum too much, so that one or more additional spectra with different apodizations may be needed. Athanasios Papoulis[4] states in his book "The Fourier Integral and Its Applications" that the triangular pulse used as a weighting function gives a more satisfactory approximation to the true spectrum in the vicinity of a discontinuity. However, adequate criteria for measuring satisfaction or distortion of these approximations are lacking.

In the next section we shall set up a criterion in terms of minimizing the least-square error and discuss an alternate approach to the problem.

THEORY

We shall examine the effect on $G'(w)$, the Fourier transform spectrum, of changing the observed interferogram $I'(x)$ to any other arbitrary interferogram, $I''(x)$.

There are two possible cases which can be handled analytically,

(i) $$I''(x) = I'(x)V(x) \tag{1}$$

where $V(x)$ is an arbitrary even function. Since $I'(x) = 0$ for $|x| > L$, $I''(x) = 0$ for this same range; and for $|x| \leq L$, $I''(x) \neq I'(x)$, and $I(x) = I'(x)$

(ii) $$I''(x) = I'(x) + V(x), \tag{2}$$

where $V(x)$ is an arbitrary even function, equal to zero for $|x| \leq L$, and equal to an assumed value of $I(x)$, for $|x| > L$.

If the true interferogram, $I(x)$, were known for the whole range $-\infty$ to $+\infty$, its Fourier transform would be the true spectrum $G(w)$.

$$G(w) = 2\int_0^{\infty} I(x) \cos wx \, dx.$$

What is obtainable experimentally is the spectrum corresponding to the weighted interferogram; namely,

$$G''(w) = 2\int_0^{\infty} I''(x) \cos wx \, dx.$$

The effect due to the weighting may to a certain extent be determined by evaluating the mean square error of $G''(w)$, which we shall denote by M.

$$M = \int_{-\alpha}^{\alpha} |G(w) - G''(w)|^2 \, dw. \tag{3}$$

For evaluating M, it is necessary to recall some of the basic theorems in Fourier analysis.

Let $I(x)$ and $J(x)$ be two functions of (x) which are absolutely integrable in the range $-\infty$ to $+\infty$. The convolution of $I(x)$ and $J(x)$ is denoted by $C(x) = I(x) * J(x)$ and defined as

$$C(x) = \int_{-\infty}^{\infty} I(a) \, J(x - a) \, da. \tag{4}$$

The convolution theorem[5] concerns the Fourier transform,

$$\int_{-\infty}^{\infty} C(x) \, e^{iwx} \, dx \quad \text{of} \quad C(x).$$

$$\int_{-\infty}^{\infty} C(x) \, e^{iwx} \, dx = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} I(a) \, J(x - a) \, da \right) e^{iwx} \, dx, \tag{5}$$

which by transposing the variables gives

$$\int_{-\infty}^{\infty} da \; I(a) \int_{-\infty}^{\infty} dx \; J(x - a) \, e^{iwx}. \tag{6}$$

Let $b = x - a$, so that $x = a + b$.

Substitution in (6) gives

$$\int_{-\infty}^{\infty} da\, I(a) \int_{-\infty}^{\infty} db\, J(b)\, e^{iw(a+b)}$$

$$= \int_{-\infty}^{\infty} da\, I(a)\, e^{iaw} \int_{-\infty}^{\infty} db\, J(b)\, e^{ibw}. \tag{7}$$

Combining (4), (5) and (7)

$$\int_{-\infty}^{\infty} I(x) * J(x)\, e^{iwx}\, dx$$

$$= \int_{-\infty}^{\infty} I(x)\, e^{iwx}\, dx \int_{-\infty}^{\infty} J(x)\, e^{iwx}\, dx. \tag{8}$$

This reduces to a simpler form if $I(x)$ and $J(x)$ are even functions. Writing

$$e^{iwx} = \cos wx + i \sin wx,$$

and equating the real parts on both sides,

$$\int_{-\infty}^{\infty} I(x) * J(x) \cos wx\, dx$$

$$= \int_{-\infty}^{\infty} I(x) \cos wx\, dx \int_{-\infty}^{\infty} J(x) \cos wx\, dx - \int_{-\infty}^{\infty} I(x) \sin wx\, dx \int_{-\infty}^{\infty} J(x) \sin wx\, dx.$$

Now,

$$\cos wx = \cos w(-x) \quad \text{and} \quad \sin wx = -\sin w(-x),$$

so that, the integrals involving $\sin wx$ on the right hand side vanish, and those involving $\cos wx$ are twice the same integrals in the limit 0 to ∞. Thus we have the convolution theorem for even functions,

$$2 \int_{0}^{\infty} I(x) * J(x) \cos wx\, dx$$

$$= \left(2 \int_{0}^{\infty} I(x) \cos wx\, dx \right) \left(2 \int_{0}^{\infty} J(x) \cos wx\, dx \right) \tag{9}$$

If $G(w)$ is the Fourier integral of $I(x)$,

$$G(w) = \int_{-\infty}^{\infty} I(x)\, e^{iwx}\, dx,$$

and since $I(x)$ is even,

$$G(w) = 2\int_{0}^{\infty} I(x) \cos wx\, dx,$$

which shows that $G(w)$ is even.

The Fourier transform

$$I(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(w)\, e^{-iwx}\, dw.$$

And since $G(w)$ is even,

$$I(x) = \frac{1}{\pi}\int_{0}^{\infty} G(w) \cos wx\, dw.$$

Similarly, if

$$H(w) = 2\int_{0}^{\infty} J(x) \cos wx\, dx,$$

$$J(x) = \frac{1}{\pi}\int_{0}^{\infty} H(w) \cos wx\, dw.$$

Substituting the values $G(w)$ and $H(w)$ for the two integrals on the right-hand side of (9),

$$2\int_{0}^{\infty} I(x) * J(x) \cos wx\, dx = G(w)\, H(w).$$

Thus $G(w)\ H(w)$ is the Fourier integral of the convolution $I(x) * J(x)$. The Fourier transform is

$$I(x) * J(x) = \frac{1}{\pi}\int_{0}^{\infty} G(w)\, H(w) \cos wx\, dw. \qquad (10)$$

The left-hand side is, from (4)

$$\int_{-\infty}^{\infty} I(a)\, J(x - a)\, da,$$

which, since I and J are even, equals

$$2\int_0^\infty I(a)\, J(x-a)\, da,$$

and since

$$J(x-a) = J(-x+a),$$

from (10) we have

$$2\int_0^\infty I(a)\, J(-x+a)\, da = \frac{1}{\pi}\int_0^\infty G(w)\, H(w)\cos wx\, dw.$$

Now, setting $x = 0$,

$$\int_0^\infty I(a)\, J(a)\, da = \frac{1}{2\pi}\int_0^\infty G(w)\, H(w)\, dw.$$

Now we change the variable on the left-hand side from a to x, and substitute for $J(x)$ the weighted interferogram $I''(x)$, and for $H(w)$ the spectrum $G''(w)$.

$$\int_0^\infty I(x)\, I''(x)\, dx = \frac{1}{2\pi}\int_0^\infty G(w)\, G''(w)\, dw. \tag{11}$$

And for the case $I(x) = I''(x)$

$$\int_0^\infty |I(x)|^2\, dx = \frac{1}{2\pi}\int_0^\infty |G(w)|^2\, dw. \tag{12}$$

Equation (12) is a form of the completeness theorem for even functions. Its physical significance is that the sum of the energy received by the interferometer for all possible positions x of the mirror is the sum of the energy in all possible values of the wave number

$$\sigma = \frac{w}{2\pi}.$$

Equations (11) and (12) will now be applied to evaluate the mean square error M of equation (3);

$$M = \int_{-\infty}^{\infty} |G(w) - G''(w)|^2\, dw.$$

Since $G(w)$ and $G''(w)$ are even,

$$\frac{M}{2} = \int_0^\infty |G(w) - G''(w)|^2 \, dw$$

$$= \int_0^\infty |G(w)|^2 \, dw + \int_0^\infty |G''(w)|^2 \, dw - 2\int_0^\infty G(w)\, G''(w) \, dw.$$

From (11) and (12),

$$\frac{M}{4\pi} = \int_0^\infty |I(x)|^2 \, dx + \int_0^\infty |I''(x)|^2 \, dx - 2\int_0^\infty I(x)\, I''(x) \, dx. \tag{13}$$

(i) If $I''(x) = I'(x)\, V(x)$ as given in (1), for $|x| < L$, $I(x)$ is the observed interferogram $I'(x)$; and for $|x| > L$, $I''(x) = 0$. Hence (13) gives

$$\frac{M}{4\pi} = \int_0^L |I(x)|^2 \, dx + \int_L^\infty |I''(x)|^2 \, dx - 2\int_0^L I(x)\, I''(x) \, dx.$$

$$- 2\int_0^L I'(x)\, I''(x) \, dx$$

$$= \int_L^\infty |I(x)|^2 \, dx + \int_0^L |I'(x) - I''(x)|^2 \, dx. \tag{14}$$

The first term on the right-hand side is the integral from L to ∞ of the square of the unobserved part of the interferogram; it is a positive quantity. The second term is the integral of the square of the difference between the observed interferogram and the weighted interferogram. This term vanishes when $I''(x) = I'(x)$. It thus appears that any attempt to change the interferogram by multiplying it with a weighting function only increases the mean square error.

(ii) If $I''(x) = I'(x) + V(x)$ as given in (2), since $I'(x) = I(x)$ for $|x| < L$, and $V(x) = 0$ for $|x| < L$, equation (13) gives

$$\frac{M}{4\pi} = \int_0^L |I(x)|^2 \, dx + \int_L^\infty |I(x)|^2 \, dx + \int_0^L |I(x)|^2 \, dx$$

$$+ \int_L^\infty |V(x)|^2 \, dx - 2\int_0^L |I(x)|^2 \, dx - 2\int_L^\infty I(x)\, V(x) \, dx$$

$$= \int_L^\infty |I(x) - V(x)|^2 \, dx. \tag{15}$$

The first case gives a minimum least-square error of

$$\int_L^\infty |I(x)|^2 \, dx,$$

and the second case gives a least-square error

$$\int_L^\infty |I(x) - V(x)|^2 \, dx,$$

which tends to zero as the assumed interferogram $V(x)$ tends to $I(x)$ over the unobserved part of the interferogram.

CONCLUSION

The spectrum which the interferometric spectrophotometer tries to measure is known to a certain degree of approximation. It is better known in the case of the earth-emitted radiation than in the case of the radiation from Mars or Venus. Even in this case, certain assumptions concerning the reflectance coefficient of the planet for the wavelength ranges of the solar spectrum, the temperature of the planet, and the absorptance of its atmosphere, permit a sufficiently close approximation. From an assumed value of $G''(w)$, the interferogram $I''(x)$ can be obtained by numerical integration. In the range 0 to L, $I''(x)$ is replaced by the observed interferogram $I'(x)$ to give a new value of $I''(x) = I'(x) + V(x)$. A Fourier transform is now made of $I''(x)$ to give a closer approximation to $G(w)$. An iteration process will further reduce the mean-square error.

The justification for weighting by adding $V(x)$ for $|x| > L$ to $I'(x)$ may also be seen from a qualitative point of view as follows: If $G(w)$ is a narrow spectral line, $I(x)$ is a simple cosine function, and the successive maxima occur at increments of $2\pi/w$ in the path difference of the two beams. An actual spectrum can be considered equivalent to the superposition of a large number of narrow spectral lines. The interferogram has large amplitude for values of $|x|$ close to zero. As $|x|$ increases, for any given large path difference several wavelengths are superposed in multiple orders of interference, so that the variations of amplitude with changes in x are small. The contribution to $G(w)$ from this tail-end of the interferogram, $|x| > L$, is relatively small compared to that from the central section $|x| < L$. How much smaller it is depends on the magnitude of L and on the complexity of the spectrum $G(w)$. The two cases we have considered, (i) $I''(x) = I'(x)\,V(x)$ and (ii) $I''(x) = I'(x) + V(x)$, are completely different in the way they alter the observed interferogram $I'(x)$. The first case changes the interferogram in a range where the spectrum is highly sensitive to such changes. The method applies further apodization to what is already an apod. The second case might be termed anapodization; it retains all of the valid experimental data and supplies approximate data in a range of $|x|$ where experimental data are unavailable.

The iteration process of computation improves the accuracy of the results obtainable from the anapodized interferogram $I''(x) = I'(x) + V(x)$. Let $G_1(w)$ be the assumed spectrum in first-order approximation, and $I_1(x)$ the corresponding interferogram. $I_1(x)$ is determined by performing the integration

$$I_1(x) = 1/\pi \int_0^\infty G_1(w) \cos wx \, dw. \tag{16}$$

We introduce a step function

$$T_0(x) \begin{cases} = 0 \text{ for } |x| < L \\ = 1 \text{ for } |x| \geq L . \end{cases} \tag{17}$$

For first-order approximation,

$$V_1(x) = I_1(x)\ T_0(x), \text{ and } I_1''(x) = I'(x) + V_1(x).$$

From $I_1''(x)$ is obtained the second-order approximation for $G(w)$, namely

$$G_2(w) = 2 \int_0^\infty I_1''(x) \cos wx\, dx . \tag{18}$$

By performing the integration on $G_2(w)$ as in equation (16), we obtain $I_2(x)$, and therefrom also

$$V_2(x) = I_2(x)\ T_0(x), \text{ and } I_2''(x) = I'(x) + V_2(x).$$

The Fourier transform $G_3(w)$ of $I_2''(x)$ obtained as in equation (18) will be a closer approximation to $G(w)$. The iteration process may be continued through as many steps as the accuracy of the experimental data $I'(x)$ will permit.

It is suggested that the validity of the method be tested by solving typical cases. An assumed spectrum of earth-emitted radiation or a calibrated spectrum of a standard infrared source may well serve as a test case.

REFERENCES

1. For more details, refer to Jenkins, F.A., and White, H.E., Fundamentals of Optics, McGraw-Hill Book Co., New York (1950) 239-52

2. Strong, J., Concepts of Classical Optics, (W.H. Freeman and Co., San Francisco, 1958) p. 428

3. Filler, A.S., "Apodization and Interpolation in Fourier-Transform Spectroscopy," J. Opt. Soc. Am., 54(1964), p. 763

4. Papoulis, A., "The Fourier Integral and Its Applications," (McGraw-Hill Book Co., New York, 1962) p. 33

5. Apostol, T.M., Mathematical Analysis, (Addison-Wesley, Inc., Reading, Mass. 1958) p. 488

Page intentionally left blank

THE VACUUM OPTICAL BENCH

AND

PRECISION MEASUREMENT OF ANGLES

Matthew P. Thekaekara, Stephen M. Shapiro,
John C. Morrison and Robert Appler

INTRODUCTION

This paper is a report of the research done by several members of the Summer Workship Team A, in collaboration with additional members of the Goddard permanent staff, on uniformity of spectral-radiant flux in the vacuum optical bench and on associated problems of space optics. A major share of the effort was spent on developing techniques for highly precise measurement of small angles. Alignment of the optic axes of experimental packages along precisely defined directions is of the utmost importance in any optical testing facility. The next section deals with the problem of the monochromator of the VOB, and the following sections describe the performance tests on autocollimators and the absolute calibration of a subsecond positioner.

THE VACUUM OPTICAL BENCH

The vacuum optical bench is essentially a system of reflecting optics in a vacuum chanber. It is a testing facility for the optical experimental packages of the Orbiting Astronomical Observatory program. The experimental packages may be a single large-aperture telescope, or several small optical systems mounted side by side in a satellite. The VOB can test at one time an experimental package with an aperture as large as 38 inches in diameter for the incoming light.

Figure A-29 shows a cutaway drawing of the VOB, which was manufactured according to NASA specifications by J. W. Fecker Division of the American Optical Company. The tall cylindrical chamber can be evacuated to a pressure of 10^{-5} mm of mercury. The two vacuum pumps are shown on either side. Inside the chamber is a smaller cylinder which holds the optical components. It can be taken out of the vacuum chamber and can be turned upside down, rotating through 180 degrees. In the normal position, shown in the figure, at the lower end is a source of light illumining the entrance slit of a monochromator. From the exit aperture of the monochromator, the light enters an inverted Cassegrainian telescope and emerges as a collimated beam. The collimated beam of monochromatic light is sent upwards to illumine the test package of the satellite. An artist's conception of a sample test package is shown in the upper half of Figure A-29.

The monochromator and the collimating system are shown in more complete detail in Figure A-30. The vacuum ultraviolet monochromator now in use is a grating instrument in a modified Paschen mounting; it was manufactured by Tropel, Inc., Fairport, N.Y. The concave tripartite grating of the monochromator was ruled by Bausch and Lomb, Inc., Rochester, N.Y., according to specifications by Tropel, Inc. The light from the hydrogen discharge tube or some other suitable source falls on the entrance slit and then on a plane mirror, and is reflected towards the concave grating. The diffracted beam is focussed on the exit aperture located within the large paraboloidal mirror. The exit aperture, the

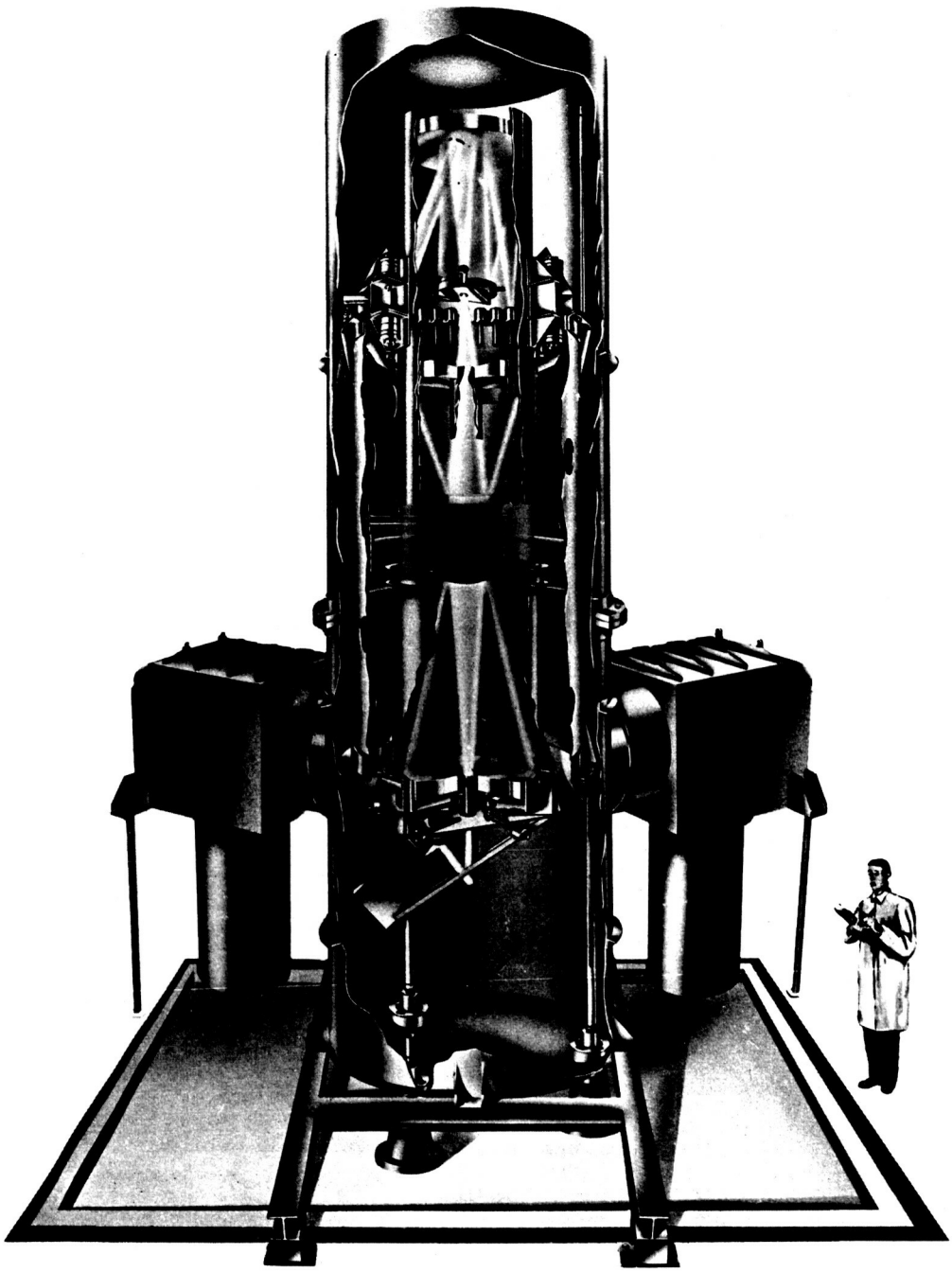


Figure A-29–Vacuum Optical Bench, Artist's Conception

grating, and the entrance slit are on the Rowland circle. To focus different portions of the spectrum on the exit aperture, the only adjustment necessary is a rotation of the plane mirror. The plane mirror is pivoted about an axis which coincides with the center of the Rowland circle, so that, when the mirror rotates, the virtual image of the entrance slit moves along the Rowland circle.

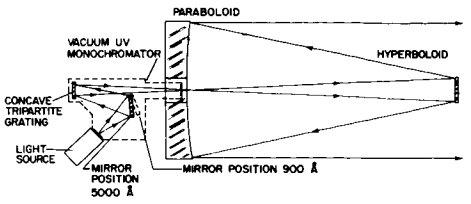


Figure A-30–VOB Monochromator and Collimating System

The difference between the optical system of this monochromator and that of the conventional Paschen mounting spectrograph is that, in the Paschen mounting, the entrance slit is fixed and the photographic plate is mounted to coincide with the circumference of the Rowland circle; here, the exit aperture is fixed and the source virtually moves along the circumference of the Rowland circle. As shown in Figure A-30, the light emerging from the exit aperture falls on the small hyperboloidal mirror and then on the large paraboloidal mirror.

A monochromator of this type has several advantages: There are only two reflecting surfaces, the plane mirror and the concave grating. In the UV, below 2000A, the reflectance of most surfaces goes down rapidly as wavelength decreases. The reflectances of the primary and secondary of the Cassegrainian system with the present surface coatings are given in Table A-15 (data supplied by Fecker Division, American Optical Co.)

Table A-15
Reflectance Percentage of the Mirrors
of the Collimating System

Wavelength (A)	Paraboloid (%)	Hyperboloid (%)
1606	71	74
1400	70	73
1216	66	72
1113	29	38
1025	17	16

The radiant flux available in the test area is proportional to the incident intensity multiplied by the series product of the reflectance of each surface. In the UV, near 1000A or lower, where sources are relatively weak and reflectance of surfaces very low, it is of the utmost importance to have as few reflecting surfaces as possible. Other advantages of the present system are that the spectrum can be scanned by remote control; a digital readout gives the wavelength; the gaseous discharge of the UV source is directed away from the grating, and a high degree of spectral resolution is possible.

On the other hand, one of the essential requirements of the VOB is that the radiant flux of the collimated beam should be uniform over the whole test area, departing by not more than 5 percent from the average. Figures A-31 and A-32, prepared by G. Comeyne, R. Crist, and H. Hoehn, show the results of studies on uniformity of illumination made on the present system.

Figure A-31 shows the variation of radiant flux in the collimated beam after reflection from the primary. The source is a small mercury-discharge tube, the pen-ray lamp (Model 11 SC-1, Spectroline Ultraviolet Lamps, Spectronics Corporation, Westbury, L.I.,

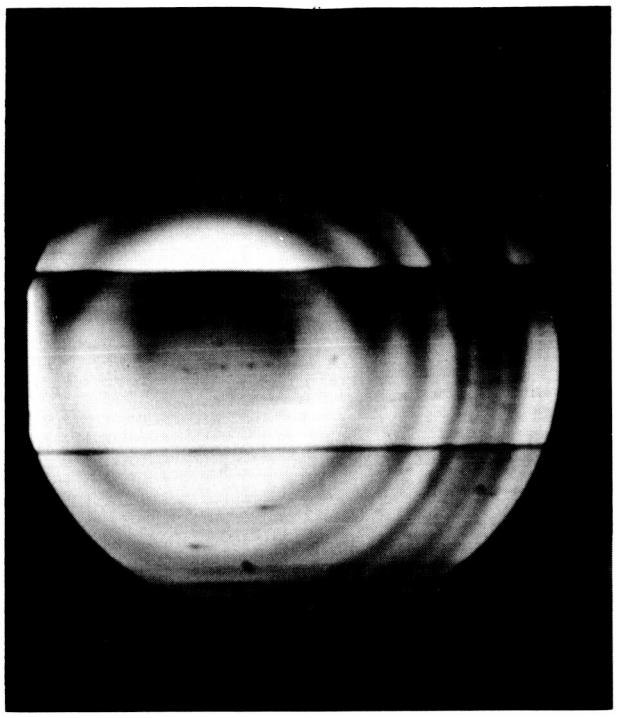


Figure A-31–Intensity Distribution in VOB Beam

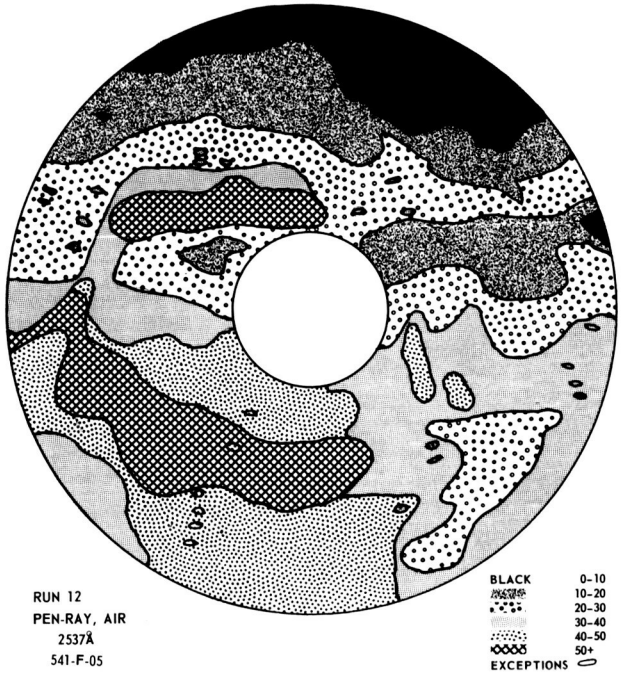


Figure A-32–Intensity Distribution of VOB Monochromator Output

N.Y.). The monochromator is set at 2537 A. The units for the legend at lower right are 10^3 photons cm^{-2} sec^{-1}. Thus the black area at top of picture receives less than 10^4 photons cm^{-2} sec^{-1}. The energy sensor is a photomultiplier tube 541F-05 (Serial No. 343, Electro-Mechanical Research, Inc., Princeton, N.J.), carried by a radial arm which pivots about an axis on one side of the VOB. The distance of the phototube along the radial arm and the rotation of the arm can be remotely controlled. Conversion of the phototube current to absolute units of photons per cm^2 per sec. is based on manufacturer's calibration. The results presented in Figure A-31 are typical of several runs made with different sources and at varying wavelengths. The radiant-flux density, far from being uniform is relatively weak at the top and extremely strong in certain small hot spots. The white circle in the center is the area occluded by the secondary hyperboloid.

The lack of uniformity shown by Figure A-31 is due to three causes, the collimator system, the monochromator and the light source. The photograph in Figure A-32 is typical of the effect due to the monochromator alone. To produce this photograph, the monochromator was taken out of the VOB and placed in a dark room. The source and the wavelength setting are the same as for the scan in the VOB. A photographic plate, Kodak emulsion type 103 a-0, is mounted at a certain distance away from the exit aperture to receive the emerging cone of light. The plate is developed after an exposure of 20 minutes.

Figure A-32 shows that the radiant flux has an annular pattern. The center of the ring system is displaced a short distance from the center of the grating. Two well-defined lines indicate the three sections of the tripartite grating. The intensity pattern changes considerably from one section to another. Another replica grating made from the same master grating was available for visual examination; that grating showed an identical ring pattern. If the wavelength of the light is changed, say from 2537 A in the UV to 4358 A in the blue, the annular pattern remains essentially the same, but the relative intensity in the three sections changes. This is what might be expected from a blazed concave grating.

Replacement of the small circular aperture by a tall slit at the exit causes a significant change in the ring pattern; the portions of the dark rings across the length of the slit become relatively brighter, and the arcs of the circle parallel to the slit seem more elongated. This again is what might be expected from the pinhole-camera effect; the tall slit is a superposition of many pinholes.

One method of securing greater uniformity of radiant flux is to have relatively wide entrance and exit apertures; however, this would entail a loss of both wavelength resolution and collimation beyond tolerable limits. The test specifications require that the VOB should simulate a star over a narrow spectral band of 1A and with a collimation of ±1 arcsecond. The diameter of the exit aperture required for eliminating the strongest of the rings is 3.55 mm. As the grating is of 75-cm radius of curvature, 5900 lines per cm, a 3.55-mm aperture would give a wavelength resolution of about 77 A and a collimation of ±16 minutes of arc.

The variation in reflectance across the surface of the grating seems to be a defect common to all concave gratings. Several explanations have been suggested: for example, defects in the blank on which the master grating is ruled, or defects in the application of the replica film to the concave surface. A bipartite concave grating of a McPherson monochromator Model 225 (McPherson Instrument Corporation, Acton, Mass.) was also examined; the ring pattern was also observed in this case, with the difference, however, that the changes between bright and dark areas occurred along more sharply defined circles, and that the rings were not continuous in the two halves of the grating. An explanation which seems plausible is that the irregularities arise while the master grating is being ruled. While the diamond point moves across the convex surface, a small percentage of

the thin metal film is gathered up in front of the cutter. Little work has apparently been done to determine the cause of these so-called "cosmetic defects" of the grating, since they are of no consequence in ordinary applications of a concave-grating instrument. If the light from the whole area of the grating is focused to a single point, the non-uniformity of reflectance does not affect the quality of the image.

This question will be treated in further detail in the paper by Dr. Irvine C. Gardner on "The Illuminating System for the Vacuum Optical Bench." A plane grating in the Czerney-Turner[1] type of mounting or one of its later modifications[2] might be more effective in securing uniformity of illumination over a large area. The plane grating is free of surface irregularities in reflectance. The wavelength of the light focussed on the exit aperture is given by the equation $\lambda = A \sin \theta$, where A is a constant of the instrument and θ is the angle which the grating normal subtends to the bisector of the angle between the directions of the incident and diffracted rays. This simple relation permits remote control and digital readout of the spectral line focused on the exit aperture. The main disadvantage is the introduction of one more reflecting surface into the optical system, with a corresponding loss in radiant flux.

The VOB is a testing facility that tries to meet some very exacting specifications: uniformity of illumination over a wide area; high degree of vacuum; capability of operating the optical bench upside down; wavelength resolution of one Angstrom unit, covering the entire wavelength range from visible near 5000 A to UV near 900 A; and collimation of the radiant flux to within an arc-second. What will actually be realized will at best be a compromise; the plane-grating monochromator seems to be a compromise more readily acceptable than the present concave-grating monochromator.

THE MIDARM MODEL 14 AUTOCOLLIMATOR

Collimation of the radiant flux to within one arc-second was listed among the requirements of the VOB. Such highly precise tolerance also requires instrumentation capable of detecting and measuring such small angles. Alignment problems arise at various stages in the testing of experimental packages of satellites like OAO and AOSO. Autocollimators now in use in the Test and Evaluation Division, Engineering and Optical Facilities Section, can measure to within one arc-second. Some of the AOSO experiments require a capability of measuring with an accuracy of 1/20 arc-second or better.

In view of these requirements, we made several calibration tests and measurements with a Midarm model 14 autocollimator, optical unit 1404-W 110 P-5. (Midarm[3] is an acronym for Micro-Dynamic Angle and Rate Monitor.) The Midarm model 14 was manufactured by Razdow Laboratories, Inc., Newark, N.J., according to specifications of NASA. It was intended to measure very small angles with an accuracy of the order of 0.05 arc-second. The objective in our calibration tests and preliminary measurements was to determine how to operate the instrument so as to yield maximum accuracy, what precautions are to be taken, and how to convert the output signals and counter readings of the Midarm to angles in arc-seconds. One arc-second corresponds approximately to five millionths of a radian, or more precisely to $4.848\ 136\ 8 \times 10^{-6}$ radian; it is the angle subtended by one foot at a distance of about 38 miles. One hundredth of an arc-second is approximately the angle subtended at the center of the earth by one foot on the earth's surface.

In measuring angles of such small magnitude, several problems arise which are not encountered in ordinary measurements of angles and directions.

The Midarm model 14 is a combination of an optical unit and an electronic unit. The optical unit is shown schematically in Figure A-33. Light from a monochromatic point

source passes through a grid (1), which is essentially a plane-transmission grating having about 6000 lines per cm. After transmission through the grid, the beam passes through a beamsplitter (7) and through a collimating lens to emerge from the instrument. At a short distance from the instrument is placed a plane mirror (4) rigidly attached to the object whose angular deflections are to be measured. The mirror surface is normal to the axis of the instrument to within a degree. The beam of light on reflection from the plane mirror returns to the instrument, passes through the collimating lens, is deflected through 90 degrees by the beamsplitter, and falls on two grids (2) and (3) which are identical to but smaller than the first grid. The rulings of all three grids are exactly parallel to each other. The grids (2) and (3) receive two adjacent portions of the reflected wave front. After transmission through the grids the light falls on photomultiplier tubes (5) and (6). The current from the photomultiplier tubes is led to the electronic unit where it is amplified and made to actuate a servomechanism which registers directly on a counter the angular deflections of the mirror (4). The amplified current from either of the phototubes may also be led to a resistor; also, a digital voltmeter may be used to display the voltage difference between the terminals of the resistor.

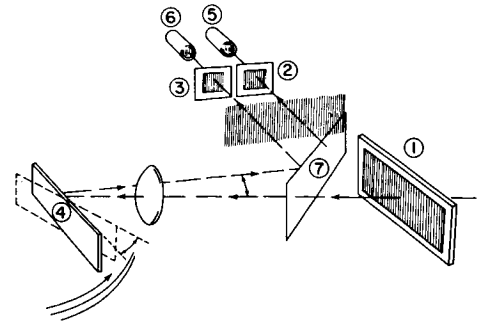


Figure A-33-Autocollimator, Midarm Model 114

The effect of the angular deflections of the mirror (4) on the output voltage of the phototube and the readings of the Midarm servo may be readily understood by considering the virtual image of grid (1) formed by the beamsplitter (7) in front of the two grids (2) and (3). A deflection of the external mirror causes a lateral displacement of the virtual image. Opaque and transparent portions of the image pass alternately across the lines of the grids (2) and (3), and the meshing effect causes fluctuations in the intensity of the light received by the phototube. The two smaller grids are positioned so that the light output from one has a 90-degree phase lag behind that of the other. The corresponding voltage output of the two phototubes may be considered as two vectors oscillating in directions at right angles to each other. The resultant of the two is a vector which rotates through an angle 2π during each cycle of the component vectors. The problem is analogous to that of producing circular polarization of light from two plane-polarized beams. The servomechanism is actuated by the resultant of the output from the two photomultipliers, and therefore can display on a counter the angle of rotation of the external object.

The operating range of the instrument is one degree of arc. The least division of the servo counter is 0.2 arc-second. The rated value of the period (that is, the angle corresponding to one complete cycle of the output voltage) is 10.355 arc-seconds. The optimum distance of the reflecting mirror from the collimating lens of the Midarm model 14 is between 11 and 13 inches.

For calibrating and making performance tests on the Midarm, a necessary piece of apparatus is an instrument which can generate known small angles. The Two-Axis Subsecond Positioner, Model 397, Serial No. 2024, manufactured by Optomechanisms, Inc., Plainview, L. I., New York, was found to be most convenient for this purpose. Figure A-34 is a photograph of the angle-generator. The instrument carries a plane mirror, shown left, at the end of a pivot tube. The mirror can be rotated about the vertical axis or the horizontal axis by means of two micrometers. The horizontal micrometer closest to the

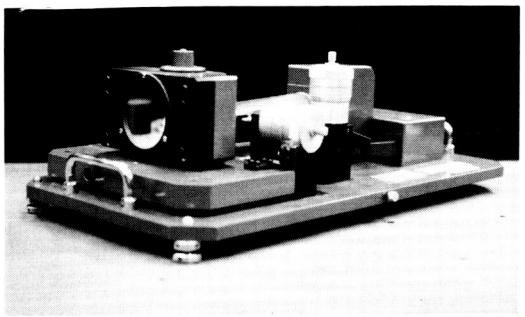


Figure A-34—Angle-Generator, Subsecond Positioner

mirror is for rotation about the vertical axis, and the other micrometer which is mounted vertically is for rotation about the horizontal axis. The total travel in each coordinate is 1000 arc-seconds, 25 seconds for each rotation of the micrometer. The smallest division on the micrometer is 0.1 of an arc-second and each major division one arc-second. The total distance by which the micrometer advances for 1000 arc-second rotation of the mirror is one inch. Thus, the effective lever arm of the angle generator is 17.2 ft. According to the manufacturers, the angle-generator has an accuracy of 0.25 arc-second settability. Our calibration of the angle generator is discussed in the next section.

Our early tests on generating and measuring small angles with the Midarm model 14 showed that, to ensure accuracy and repeatability, all mountings should be very rigid and stable and that temperature fluctuations in the room and turbulence in the air should be kept to a minimum. A massive optical table weighing 2500 lbs., having a flat 8- by 4-foot steel-plate surface, was used for mounting the instrument. A heavy aluminum angle had its horizontal part bolted to the optical table and the vertical part bolted to the baseplate of the Midarm model 14, the Midarm being rotated about its own axis 90 degrees. This provided a highly stable mount for the instrument, and it was possible to measure rotation about the horizontal axis of the angle generator. The mirror of the angle generator was placed 12 inches from the collimating lens of the Midarm. Turbulence was reduced considerably by attaching a light tube to the front of the lens and by putting a further heavy cover of black velvet cloth over the Midarm and the angle generator. When suitably mounted, the Midarm is a highly sensitive instrument which is capable of showing significant voltage fluctuations due to the body heat of an observer changing his relative position with respect to the instrument, or exerting a slight pressure with a finger on the massive table.

For most of the performance tests of the Midarm, the voltage output of one of the phototubes, Channel B (junction J 107-109 on the electronic unit), was monitored by a digital voltmeter. This gives the output before it enters the servomechanism and its associated amplifiers. Balancing of the two amplifiers according to the manufacturer's instructions had no effect on this output voltage, which is dependent only on the light entering the phototube. A lack of balance does, however, affect the servo reading.

Several tests were made on the stability and repeatability of the voltage output. In a typical test where all precautions were taken to eliminate external disturbances, with the Midarm rigidly clamped to the optical table, and a magnetic-back mirror in place of the angle generator, also mounted rigidly on a bracket clamped to the table, the drift in output was 1 volt in 10 min. This corresponds to a change in angle at the rate of 0.006 arc-second per minute. This change is presumably due to springing of the steel brackets and screws, or to small temperature gradients. The change is small enough to be neglected. Performance tests on the Midarm are described in the next section.

CALIBRATION OF THE ANGLE-GENERATOR USING A LASER BEAM

Because most of the performance tests on the Midarm model 14 were dependent on the angle generator, an absolute calibration of the angle generator was considered most important. Of several methods which were attempted, the one finally adopted as most practicable and capable of giving the required accuracy was a laser beam.

We use a laser beam instead of an ordinary source of light because it gives a large output of energy along a given direction with very small angular spread. A helium-neon gas laser manufactured by Spectra-Physics, model 130, was used for our measurements. It gives a strong red light due to the transition $3p'_{12} - 5s'_{01}$ of neon, wavelength 6328.1646 A. At a distance of 100 feet the beam spreads out to a distance of only one inch. This can be focused almost to a point by placing a telescope of high optical quality in front of the exit window of the laser. A Wilde telescope, Heerbrugg, Switzerland, (NH 3-96774) of the type used in geodetic survey was found adequate. The light enters through the eyepiece of the telescope and emerges from the field lens as a beam of 50-mm aperture. By adjusting the focusing screw of the telescope, a well-defined image can be obtained at any desired distance.

The experimental arrangement is shown schematically in Figure A-35. The light from the laser, after passing through a neutral filter, a shutter, and the telescope, falls on the mirror of the angle generator. It is reflected to a mirror mounted on an optic bench and then falls on a photographic plate. A Kodak spectroscopic plate with type 103°F emulsion was selected as most suitable for giving well-defined images under the neon red light. A camera shutter was used to give accurately repeatable exposure times of 1/5 sec. Even with such short exposure times, the intensity is too high, so that a neutral-density filter was used to cut down the intensity by a factor of 100.

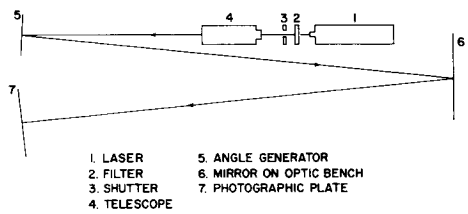


Figure A-35-Laser for Angle Measurement

The theory underlying the experiment is simple: If L is the distance between the mirror of the angle generator and the photographic plate, and θ the angle through which the mirror is rotated, the displacement s of the image on the plate is given by $s/L = 2\theta$. Considerably greater displacement can be obtained if the light is reflected several times between the angle generator and the mirror on the optic bench. But this was not feasible, because the beam is about 2 inches wide when it first falls on the angle generator and multiple reflection would require a much larger size for both mirrors. The error in measuring the path-length L between two distant points in a room is considerably greater than that in measuring the distance between images of a photographic plate. Hence, the mirror which returns the image to the photographic plate is mounted on a long optical bench. A 243-cm-long Beck Precision Research Optical Bench, Model 23-000, was available in the darkroom. A magnetic-back mirror was rigidly fixed on a clamp first at one end and then at the other end of the optical bench. If L_1, L_2 be the path length of the laser beam in the two positions of the mirror, and s_1, s_2 are the corresponding displacements of the image for the same deflection θ of the angle generator,

$$2\theta = \frac{s_1}{L_1} = \frac{s_2}{L_2} = \frac{s_2 - s_1}{L_2 - L_1} . \qquad (1)$$

The distance $(L_2 - L_1)$ can be read with a scale and vernier correct to 1/10 mm.

The other major sources of error are in setting the drum of the angle-generator to the divisions marked on its scale, and in measuring the distances between the images on the plate. The smallest division on the drum of the angle-generator is 0.1 sec, as was stated earlier. An experimental procedure of obtaining a large number of exposures, and a least-squares method of reducing the data, were adopted to reduce these errors to a minimum.

A reading lens was mounted in front of the vertical drum of the angle-generator to improve accuracy of setting without parallax. The photographic plate was mounted on a rigid stand with C clamps. The angle-generator was set successively at different angles between 0 and 1000 arc-seconds with equal increments of 50 arc-seconds, and in each position an exposure of 1/5 second was made. This procedure was repeated using the same photographic plate and the other position of the mirror on the optical bench. In all, there are two series of 21 images on a plate. A second photographic plate was taken with the same experimental procedure. A test exposure had shown that the primary image is too strong, but that each image is accompanied by a lighter, well-defined image caused by reflection from the back glass surface of the angle-generator mirror.

The positions of the images were measured on a precision comparator, series 292 AY, manufactured by Gaertner Scientific Company. Four sets of measurements were made on each row of images and the values were averaged. Measurements were made on the first plate by two observers and on the second plate by three observers. A preliminary data analysis was made, using the conventional methods of reducing wavelengths of spectrum plates with a desk calculator. The results showed that the data were sufficiently accurate to permit a more elaborate least-squares method with an electronic computer.

Table A-16 gives the results of the least-squares computation on one of the sets of data. These are from one of three measurements on the second plate for the more distant position, distance L_2, of the mirror on the optical bench. The first column gives the setting ϑ of the angle generator in arc-seconds; the second column gives the position s of the laser image averaged from four sets of readings. There are three major sources of error: The graduations of the angle generator are not exact; the settings of the angle generator were not exactly on the markings; the centers of the images were not measured precisely on the comparator. The two latter are experimental errors, and the first is instrumental. The aim of the calibration experiment is to determine the error, if any, in the graduations. If all three sources of error were absent, a graph of s versus θ would be a straight line. The third source of error is lessened by averaging the data sets of different observers on the same series of images. Similarly, the second source of error is lessened by comparing the results of the four series of images. The method of least squares applied to the 21 data points further reduces both these sources of error.

Below Table A-16 are given the constants a_0, a_1 of a linear equation $s = a_0 + a_1 \theta$, as well as the constants A_0, A_1, A_2 of a quadratic equation $s = A_0 + A_1 \theta + A_2 \theta^2$, which give the best fit to the data points. The difference between the calculated and observed values of s are given in column 3 (Δs) for the linear equation, and in column 5 ($\Delta_2 s$) for the quadratic equation. An error analysis of Δs and $\Delta_2 s$ for all sets of data showed that the linear equation is preferable.

Column 4 gives the values of $\Delta\vartheta = \vartheta_{cal.} - \vartheta_{obs.}$. $\vartheta_{cal.}$ is the calculated value of θ given by the equation

$$\vartheta_{cal.} = \frac{s_{obs.} - a_0}{a_1}$$

Table A-16
Calibration of the Angle Generator
(Data on Laser Images, plate no. 2; distance L_2
least-squares solution - linear and quadratic equations)

θ_{obs} sec.	s_{obs} inch	$\Delta s = s_{cal} - s_{obs} \times 10^4$	$\Delta\theta = \theta_{cal} - \theta_{obs}$	$\Delta_2 s = s_{cal} - s_{obs} \times 10^4$
0	1.79 8300	- 9.424	+0.2146	+14.666
50	2.01 6775	+ 1.237	-0.0282	+18.100
100	2.23 7925	-14.852	+0.3382	- 4.455
150	2.45 5575	+ 4.060	-0.0925	+ 8.751
200	2.67 7125	-16.029	+0.3650	-16.283
250	2.89 6300	-12.367	+0.2817	-16.806
300	3.11 5900	-12.956	+0.2951	-20.818
350	3.33 4675	- 5.295	+0.1206	-15.819
400	3.55 2750	+ 9.367	-0.2133	- 3.060
450	3.77 3025	+ 2.028	-0.0462	-11.539
500	3.99 1375	+13.940	-0.3175	- 0.008
550	4.21 0800	+15.101	-0.3439	+ 1.533
600	4.42 9225	+26.262	-0.5981	+13.835
650	4.64 9325	+20.674	-0.4709	+10.149
700	4.86 8175	+27.585	-0.6283	+19.723
750	5.08 6275	+41.997	-0.9565	+37.558
800	5.31 0325	- 3.092	+0.0704	- 3.347
850	5.53 0500	- 9.431	+0.2148	- 4.740
900	5.75 0175	-10.769	+0.2453	- 0.373
950	5.97 0850	-22.108	+0.5035	- 5.245
1000	6.19 2775	-45.946	+1.0464	-21.857

$s = a_o + a_1\theta$ $\quad a_1 = 0.0043908228$ $\quad A_0 = 1.7997666$ $\quad A_2 = .000000015215$
$a_o = 1.7973576$ $\quad s_2 = A_0 + A_1\theta + A_2\theta^2$ $\quad A_1 = .0043756078$

If it were possible to eliminate the two experimental sources of error, and if, further, the average of the second intervals of the angle-generator is exactly one arc-second, the value of $\Delta\theta$ would give exactly the error in the graduations at each of the 50-second markings.

The average value of the second intervals in the graduations can be calculated from the values of a_1 for the distances L_1 and L_2. Let θ_c be the average value in radians for a rotation of the mirror through one arc-second of the angle-generator; let s_1 be the corresponding displacement of the laser image for the shorter path-length L_1, and s_2 the displacement for the longer path-length L_2. Let the values for the constant a_1 of the linear equation, determined by the least-squares method and averaged for all measurements on the same plate, be a_{11}, a_{12} respectively for the distances L_1, L_2.

Since $s = a_0 + a_1\theta$, $ds/d\theta = a_1$, θ is measured in units of θ_c, the angle corresponding to one second of the angle-generator. Hence a_{11}, a_{12} are the increments s_1, s_2 for deflection θ_c.

By equation (1), $\theta_c = \frac{1}{2}\frac{s_2 - s_1}{L_2 - L_1}$.

i.e. $\theta_c = \frac{1}{2}\frac{a_{12} - a_{11}}{L_2 - L_1}$.

The mirror on the optical bench was moved through a distance of 230 cm, and a correction factor of 0.2 cm must be added because the reflected beam is inclined at an angle 8.7×10^{-4} radian to the optical bench. Thus $L_2 - L_1 = 460.2$ cm.

The averages a_{12}, a_{11} are substituted in the equation for θ_c, after conversion from inches per sec. to cm per sec. The weighted average of θ_c is calculated from the results of the two plates.

The final result shows that one second interval of the angle-generator is on an average

4.837 140 5×10^{-6} radian, or

0.997 731 8 arc-second.

The probable error in this value is ±.000 018 arc-second, and hence it is preferable to retain only six significant figures. This error of 0.23 percent is not surprising, in view of the delicate balancing of the pivots in the angle-generator. This result was confirmed by further measurements on the Midarm model 14 and the Hilger-Watts autocollimator.

From the calibration of the average interval, we may now proceed to calibrate each 50-second interval. Such a calibration is possible only if the experimental errors in the angle-generator settings and in the comparator measurements are small compared to the instrumental errors in the angle-generator. That this is the case is shown by the data analysis and the final results which are presented in Table A-17 and in Figures A-36 and A-37. In Table A-17, the first column gives angle as measured on the angle-generator, and the next six columns give the values of $\Delta\theta$, the difference between θ_{cal} and θ_{obs}. Only a few sample sets of values have been chosen. The third column in Table A-17 is the same as the fourth column in Table A-16. The degree of accuracy in the comparator measurements is shown by the four columns of values for the second plate at distance L_2. The results from three observers and the average are given for each of the angles. The average deviation of the results of any observer from the average of all three is 0.020 arc-second, which corresponds to an accuracy of $\pm 8.8 \times 10^{-5}$ inch in reading the position of the image. This is slightly better than the accuracy obtainable in spectrographic plates. In columns 5, 6, and 7 are given three sets of values: namely, the average values of $\Delta\theta$ for the second plate at distance L_2, the same at distance L_1, and the weighted average for all measurements. A comparison of these data shows the repeatability of the angle-generator settings. The average deviation is 0.16 arc-second. These data are also presented in Figure A-36. We observe that though there is a certain degree of scatter of the points about the average, all three curves have the same general shape.

Figure A-37 shows the calibration curve of the angle-generator. Data points of the graph are given in the last column of Table A-17. The correction factor $d\theta$ which should be added to a given reading of the angle-generator is the sum of two correction factors: the average correction factor, which is -0.002268 times the reading on the drum, and a variable correction factor depending on the random irregularities. This factor is obtained by multiplying the values in the 7th column of Table A-17 by 0.997732.

The calibration curve shows that, for small angles of the order of 50 or 100 seconds, the error is within the rated settability of 0.25 arc-second claimed by the manufacturers. But, in generating an angle of 500 seconds or more, the correction factor should be applied. Because of improvised arrangements for taking the photographs, and for developing and measuring them, the calibration procedure proved to be rather laborious. The computational analysis is quite simple, and most of the data reduction can be performed automatically. The main advantage is that we have here a method of absolute calibration, capable of yielding a high degree of accuracy.

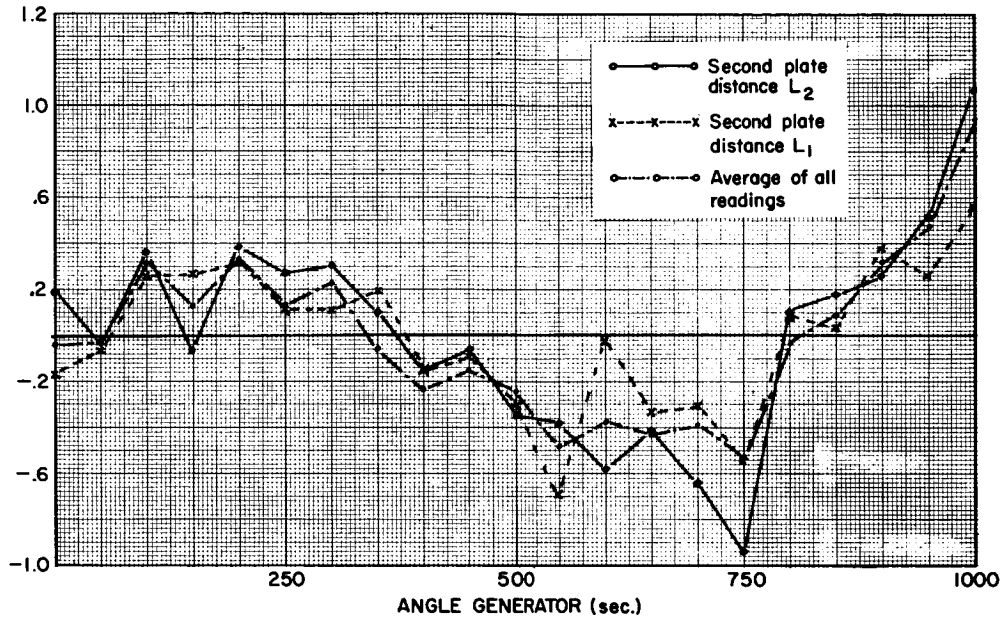


Figure A-36–Curve Showing Repeatability of Angle-Generator Readings

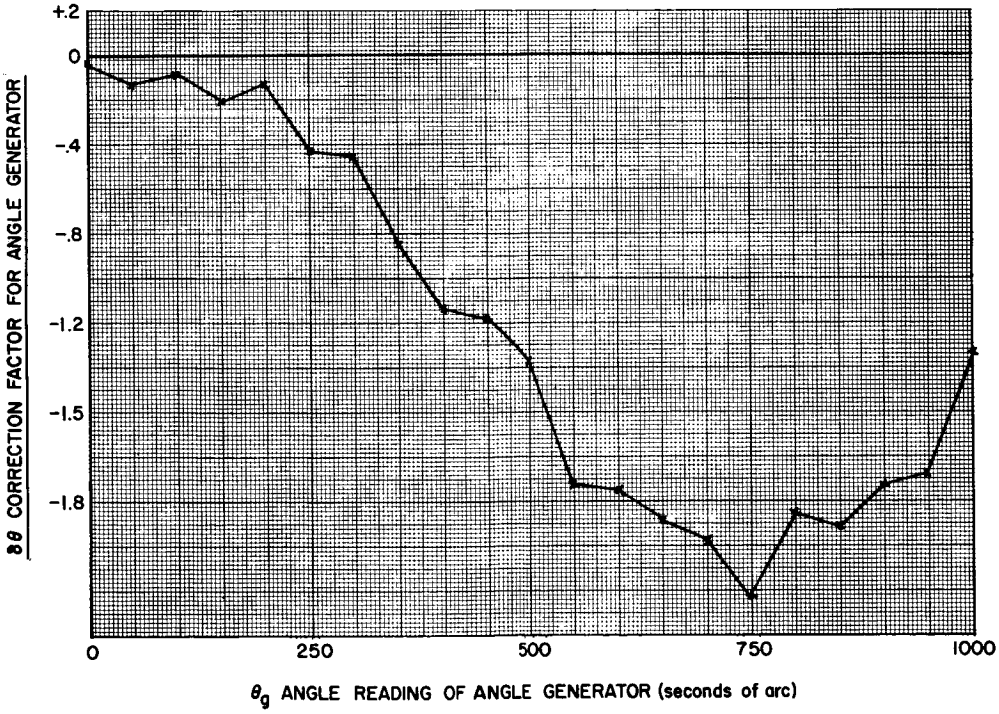


Figure A-37–Calibration Curve for Angle-Generator

Table A-17
Calibration of Angle-Generator
(data analysis and results)

θ_g	$\Delta\theta = \theta_{cal} - \theta_{obs}$						$d\theta$ Correction Factor
	Plate 2, Distance L_2				Plate 2 Distance L_1 Average	Average of all Values	
	1	2	3	Average			
0	0.183	0.215	0.172	0.190	-0.176	-0.040	-0.040
50	-0.053	-0.028	-0.007	-0.029	-0.073	-0.023	-0.136
100	0.394	0.338	0.337	0.356	0.260	0.316	-0.088
150	-0.036	-0.092	-0.061	-0.063	0.264	0.132	-0.209
200	0.331	0.365	0.451	0.382	0.320	0.321	-0.133
250	0.255	0.282	0.292	0.276	0.115	0.134	-0.433
300	0.291	0.295	0.320	0.302	0.116	0.221	-0.460
350	0.135	0.121	0.047	0.101	0.196	-0.059	-0.853
400	-0.112	-0.213	-0.175	-0.167	-0.154	-0.235	-1.142
450	-0.086	-0.046	-0.056	-0.063	-0.092	-0.154	-1.184
500	-0.357	-0.317	-0.380	-0.351	-0.305	-0.241	-1.376
550	-0.388	-0.344	-0.409	-0.380	-0.789	-0.481	-1.727
600	-0.567	-0.598	-0.591	-0.585	-0.025	-0.386	-1.746
650	-0.507	-0.471	-0.546	-0.508	-0.330	-0.424	-1.897
700	-0.641	-0.628	-0.642	-0.637	-0.304	-0.394	-1.981
750	-0.923	-0.956	-0.984	-0.954	-0.538	-0.539	-2.239
800	0.094	0.074	0.137	0.102	0.087	-0.030	-1.845
850	0.154	0.215	0.183	0.184	0.039	0.090	-1.908
900	0.236	0.245	0.291	0.257	0.381	0.317	-1.725
950	0.513	0.503	0.530	0.514	0.266	0.478	-1.678
1000	1.085	1.046	1.077	1.069	0.560	0.940	-1.330

ABSOLUTE CALIBRATION USING A SINE BAR

Angles of the order of a fraction of an arc-second cannot be measured with sufficient accuracy by the laser-beam method described in the previous section. The Midarm model 14 has a period of about 10 arc-seconds. To determine the variation of voltage output as a function of the angle, we need to generate and measure angles of the order of 1/10 second or less. The sine bar with precision gauge blocks is a suitable piece of apparatus for this purpose.

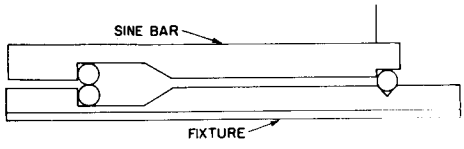


Figure A-38–Sine Bar and Support Fixture

Figure A-38 is a line drawing of the sine bar mounted over its support fixture. The sine bar is a massive bar of steel supported by two cylindrical rods whose axes are 20.025 inches apart. By placing precision gauge blocks below one of the rods, the bar may be made to rotate about the other rod. A magnetic-back mirror is rigidly clamped to the front end of the bar to deflect a beam of light through an angle twice that of the sine bar. The autocollimator (Midarm model 14) which is to be calibrated is placed in front of the mirror. Angles of the required magnitude and accuracy can be generated by a set of twenty Weber AA precision gauge blocks which are calibrated in increments of 5×10^{-6} inch, and are guaranteed in flatness

and thickness to $\pm 10^{-6}$ inch. By replacing a gauge block by another next in the series, the normal to the mirror rotates in increments of nearly 0.05 ± 0.01 arc-second; the twenty blocks provide 20 calibration points over a range of one second.

One major source of error in this procedure is the variable air layer between the gauge block and the metal surfaces above and below it. Such an air layer might well be of the order of magnitude of the increments in the gauge blocks. To avoid this, we have designed a support fixture shown in the lower half of Figure A-37. The fixture appears as a complement or "shadow" of the sine bar, with the exception that the front support is replaced by a V-shaped groove. This design, which was chosen after trying several others, is patterned after the suggestion made by J. C. Evans and C. O. Tylerson[4] of the National Physical Laboratory of England. When the gauge block is in position, there is a line contact with the upper and lower support rods and the air layer is thereby eliminated.

A problem exists, however, in that the hind support may not rest vertically above its support fixture, because of a backward or forward displacement of the sine bar. But it can be readily seen that, provided these displacements can be kept within 10^{-3} inch, the corresponding error in the vertical displacement of the hind support is smaller than 10^{-6} inches. To minimize these displacements, an added precaution is taken in making the support fixture. The V-groove surface is lapped to the cylinder support, so that the two surfaces make an 80-percent line contact through a minimum of three degrees of rotation. The field-of-view of Midarm model 14 is one degree, so that this would be adequate. A thumbscrew will be provided on the support fixture close to the hind rod, to raise or lower the sine bar gently and thus to minimize the chances of lateral displacement.

This arrangement permits an accurate calibration of the autocollimator over an interval of one arc-second. However, the period of the Midarm model 14 is 10 seconds. To calibrate over this wide range, we need a means of generating angles longer than one second. This can be achieved by providing the mirror attached to the sine bar with a tangential screw which permits the mirror to be rotated through very small angles. An interval of one second is first scanned by successively placing thicker blocks under the hind rod. The final voltage output of the Midarm is noted. The last gauge block is replaced by the first, and the mirror is adjusted to restore the former voltage. In this way, it is possible to obtain calibration points over a relatively wide range.

To avoid random voltage fluctuations due to the vibrational motion of the mirror relative to the Midarm, the sine bar support fixture and the Midarm will be rigidly fastened to a massive table; to avoid fluctuations due to turbulence in the air, a plexiglass case will be placed over the entire apparatus.

PERFORMANCE TESTS ON MIDARM MODEL 14

As was stated earlier, our objective in making the performance tests on Midarm model 14 was to determine how best to operate the instrument so as to yield maximum accuracy. The variables in the experimental setup are: V the voltage output of the phototube, θ_s the angle as recorded by the Midarm servomechanism, and θ the angle which the axis of an external mirror makes to fixed direction of reference. For most of our tests the external mirror was the mirror of the angle-generator, and we denote θ as θ_g. The variation of V with θ_g is periodic. We need to determine accurately the period, and to know whether the period is constant over the whole range of the Midarm. The rate of change $dV/d\theta_g$ is maximum when V is in the middle of its range; at this setting, the instrument has maximum sensitivity. This rate of change and its dependence on the amplitude of V have to be determined.

The period of the Midarm according to the manufacturers is 10.355 sec. This was determined by measuring the angle through which an external mirror has to be rotated so that the V output completes 100 cycles. The error is 1/100 of the error in the readings of the initial and final settings of the mirror, and the value is an average for the whole range of about 1000 arc-seconds.

For a more precise determination of the period, the following procedure was adopted: The Midarm was rigidly fixed on the massive table; the angle generator was placed at a distance of one inch from the light tube. The usual adjustments for minimum air turbulence and maximum sensitivity were made. The output voltage fluctuated between 62.75 volts and 17.92 volts, so that the midpoint of the range is about 40v. The angle generator was slowly rotated and readings of the angle were recorded each time the increasing voltage showed a value just above 40v or the decreasing voltage showed a value just below 40v. If the value of V at which the angle reading was taken differed from 40 by more than 0.1 volt, a correction based on an independent measurement of $d\theta/dv$ was applied to find what θ would have been for V equal to 40 volts. A total of 192 readings of θ_g was taken, 96 for decreasing voltage and 96 for increasing voltage. The difference between the nth reading and the (n + 96)th reading gives the average period for 48 cycles; such values are available in all 96 readings, and from these a highly precise value of the average period is obtained. These data also permit an analysis of the variation of the period over the range of 1000 seconds.

Let t be the period per cycle and T the difference of readings of θ_g for the nth and (n + 96)th setting. T = 48t. The average of all the values of T is 497.1483 sec. Hence, the value of t for the range of 1000 sec = 10.3573 sec, with an average deviation of 0.009 sec. The value of T is near 495.8 in the first half of the range, and increases to 497.4 after the first 20 cycles. Table A-18 gives the values of $(T_{av} - T)$, and the deviation from the average of each of the observed values of T. The deviation is high at the beginning of the range, but later becomes almost constant at about 0.4 sec. A few entries are missing because readings for the corresponding cycles were not recorded. The period measured over a more limited range gives a more steady value. Thus, if T is calculated for the cycles N + 48 to N (N varying from 22 to 31), all values lie between 497.42 and 497.64. The average is 497.525. This yields a value of the period t = 10.365. The average deviation is considerably less, being equal to 0.001 second. All values are recorded in terms of the seconds of the angle generator. Because one second on the angle-generator is on an average 0.2268 percent less than the correct value, the average period of the Midarm model 14 for the 1000-sec. range is 10.334 ± .009 sec. From these studies we conclude that, for all precise measurement, the period should be determined for the range over which the Midarm is going to be used.

The variation of V with θ was determined also using the angle-generator. The angle was changed in steps of half a second and the voltage was recorded. Because the reading of the digital voltmeter fluctuates by a few hundredths of a volt, four or five readings were taken for each setting of θ_g. The shape of the V - θ_g curve depends necessarily on the maximum and minimum values of V; these values in turn depend on the distance of the mirror, the reflectance of the mirror, and certain variable factors of the electronic unit of the Midarm. Measurements of V versus θ were made for three different settings of the voltage range. The voltage range was changed by moving the angle generator with respect to the Midarm. The results are shown graphically in Figure A-39. The data are presented partially in Table A-19, where voltages V_1, V_2, V_3 are not shown for the half-integral values of θ_g; all the data points are, however, used in drawing the graphs of Figure A-39.

A preliminary examination shows that the shape of the curve is sinusoidal. A more precise set of measurements over a range of 3 cycles was made, again at half-second

Table A-18
Change of Period of Midarm Model 14
(over a range of 1000 seconds)

Range of Cycles		$\Delta T = T_{av} - T$ sec.		Range of Cycles		$\Delta T = T_{av} - T$ sec.	
from	to			from	to		
49	1	-1.398	-1.368	73	25	+0.412	+0.402
50	2	-1.248	-1.138	74	26	+0.422	+0.492
51	3	-1.118	-1.028	75	27	+0.412	+0.492
52	4	-1.078	-0.988	76	28	+0.302	+0.302
53	5	-0.888	-0.718	77	29	+0.382	+0.432
54	6	-0.768	-0.808	78	30	+0.402	+0.402
55	7	-0.878	-0.808	79	31	+0.412	+0.382
56	8	-0.608	-0.618	80	32	+0.412	+0.402
57	9			81	33	+0.322	+0.352
58	10	-0.468	-0.398	82	34	+0.292	+0.402
59	11	-0.498	-0.298	83	35	+0.352	+0.352
60	12	-0.228	-0.148	84	36	+0.352	+0.412
61	13	-0.138	-0.138	85	37	+0.392	+0.422
62	14	-0.182	-0.132	86	38	+0.312	+0.382
63	15	-0.048	-0.078	87	39	+0.312	+0.712
64	16	-0.018	+0.022	88	40	+0.762	+0.392
65	17	+0.082	+0.052	89	41	+0.202	+0.502
66	18	+0.152	+0.102	90	42	+0.422	+0.522
67	19	+0.022	+0.192	91	43	+0.292	
68	20	+0.192	+0.212	92	44	+0.352	+0.392
69	21	+0.192	+0.322	93	45	+0.262	+0.342
70	22	+0.452	+0.272	94	46	+0.172	+0.262
71	23	+0.312	+0.312	95	47	+0.192	+0.262
72	24	+0.302	+0.352	96	48	+0.122	+0.072

intervals. The observed values of V were compared with values calculated from a sine curve given by the equation

$$V - v_0 = v \sin 2\pi \frac{\theta - \theta_0}{t} \qquad (2)$$

where v is the amplitude = $\frac{V_{max} - V_{min}}{2}$,

v_0 is the mean value = $V_{max} + V_{min}/2$, t is the period, and θ_0 is a phase angle which gives the best fit between the calculated and observed curves. The results are presented in Table A-20 and Figure A-40. The values of V for half-integral values of V are not listed in Table A-20, but all data points have been used in drawing the two graphs. The crosses in Figure A-40 indicate the observed values of V and the circles indicate the calculated values. By changing θ_0 or v_0, it is possible to shift the sine curve (V_{cal} versus θ) parallel to the θ axis or the V axis; by changing t or v, it is possible to contract or expand the sine curve along these axes. The constants were calculated with approximation methods using a desk calculator. Table A-20 shows that most of the values of ($V_{cal} - V_{obs}$) are positive. A more precise determination of the constants, using a

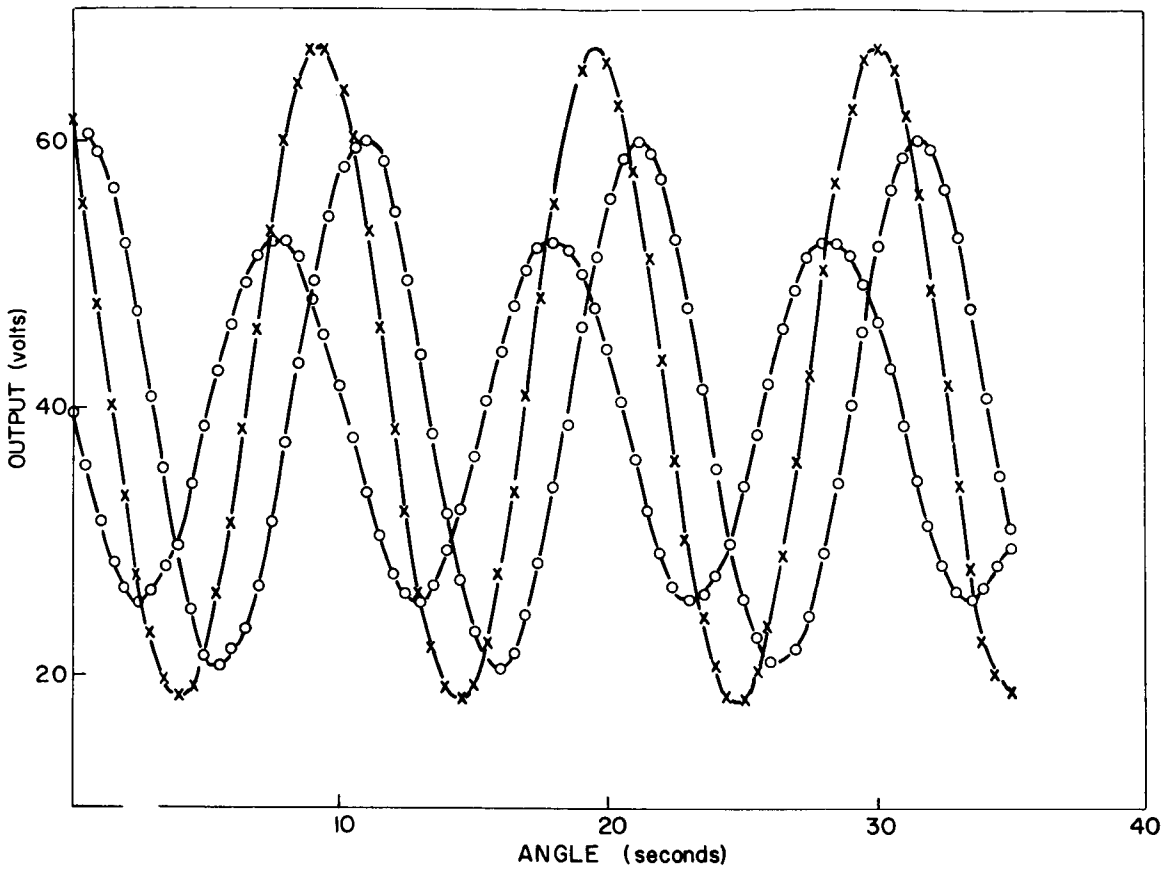


Figure A-39–Voltage Variation, Output vs Angle

least-squares method and an electronic computer, would give an even closer agreement between the observed curve and a true sine curve. The maximum differences occur where the slope $dV/d\theta$ is the greatest, and they are of the order of one volt, which corresponds to 0.07 arc-second. We would therefore conclude that, for measuring small angles with an accuracy of 0.01 arc-second, the constants of the sine curve should be determined by a separate experiment in each case, and there is no guarantee that the amplitude v and the phase θ_0 will remain constant. Variations in the period t are less significant.

A test run was made to study the variation of V_{max} and V_{min} over a range of 1000 seconds. The angle-generator was kept at the same fixed distance and every precaution was taken to secure repeatable readings. The average of several sets of readings is presented in Table A-21. While V_{max} goes steadily down, V_{min} goes to a minimum and then goes up. The value of v_0 changes by 1.45 volts.

Tests were made on the repeatability of the slope of the curve under identical conditions of mirror distance and electronic unit adjustment. To increase the accuracy of the measurements, instead of changing increments, an alternate procedure was adopted. The drum of the angle-generator was rotated until the voltage was very close to 10 volts

Table A-19
Change in Slope with Change in Amplitude

θ_g	V_1	V_2	V_3	θ_g	V_1	V_2	V_3
.0	61.37	59.22	39.57	20.0	66.40	55.92	44.27
1.0	48.19	59.57	31.49	21.0	58.00	60.07	36.20
2.0	33.84	52.44	26.52	22.0	44.00	57.02	29.03
3.0	23.09	40.80	26.30	23.0	30.10	47.70	25.76
4.0	18.55	29.09	31.11	24.0	21.00	35.60	27.64
5.0	21.85	21.76	38.88	25.2	18.66	25.32	34.16
6.0	31.87	21.08	46.50	26.0	23.99	20.99	41.84
7.0	46.30	26.88	51.59	27.0	36.09	21.84	48.95
8.0	60.17	37.20	52.69	28.0	50.83	29.21	52.48
9.0	67.09	49.49	49.19	29.0	62.89	40.69	51.60
10.0	64.41	58.01	41.90	30.0	67.15	52.08	46.75
11.0	53.40	60.05	33.83	31.0	62.08	58.87	38.90
12.0	39.00	54.87	27.53	32.0	49.44	59.48	31.11
13.0	26.30	44.06	25.79	33.0	34.57	52.95	26.36
14.0	19.32	31.97	29.24	34.0	22.90	40.84	26.46
15.0	19.62	23.25	36.34	35.0	18.51	29.26	31.87
16.0	27.83	20.57	44.39				
17.0	41.19	24.57	50.50				
18.0	55.64	34.19	52.71				
19.0	65.47	46.24	50.44				

Table A-20
Voltage Calculated from Sine Function, vs. Observed Voltage

θ	V_{cal}	V_{obs}	$V_{cal} - V_{obs}$	θ	V_{cal}	V_{obs}	$V_{cal} - V_{obs}$
4.0	56.61	55.60	1.01	25.0	52.10	51.70	.40
5.0	42.46	40.93	1.53	26.0	38.16	36.37	1.79
6.0	28.49	27.83	.66	27.0	25.22	24.70	.52
7.0	19.69	19.99	-.30	28.0	18.62	18.80	-.18
8.0	19.19	19.24	-.05	29.0	20.70	20.43	.27
9.0	27.17	26.22	.95	30.0	30.73	29.26	.4
10.0	40.78	38.80	1.98	31.0	45.12	42.97	2.15
11.0	55.17	53.80	1.37	32.0	58.73	57.57	.16
12.0	65.20	64.67	.53	33.0	66.71	66.21	.50
13.0	67.28	67.02	.26	34.0	66.21	65.59	.62
14.0	60.68	60.32	.36	35.0	57.40	56.27	1.13
15.0	47.74	46.74	1.00	36.0	43.43	42.19	1.24
16.0	33.10	32.17	.93	37.0	29.29	28.33	.96
17.0	21.97	22.27	-.30	38.0	20.03	20.00	.03
18.0	18.34	18.51	-.17				
19.0	23.49	22.71	.78				
20.0	35.59	33.97	1.62				
21.0	50.32	48.59	1.73				
22.0	59.49	-	-				
23.0	67.57	67.17	.40				
24.0	63.93	63.26	.67				

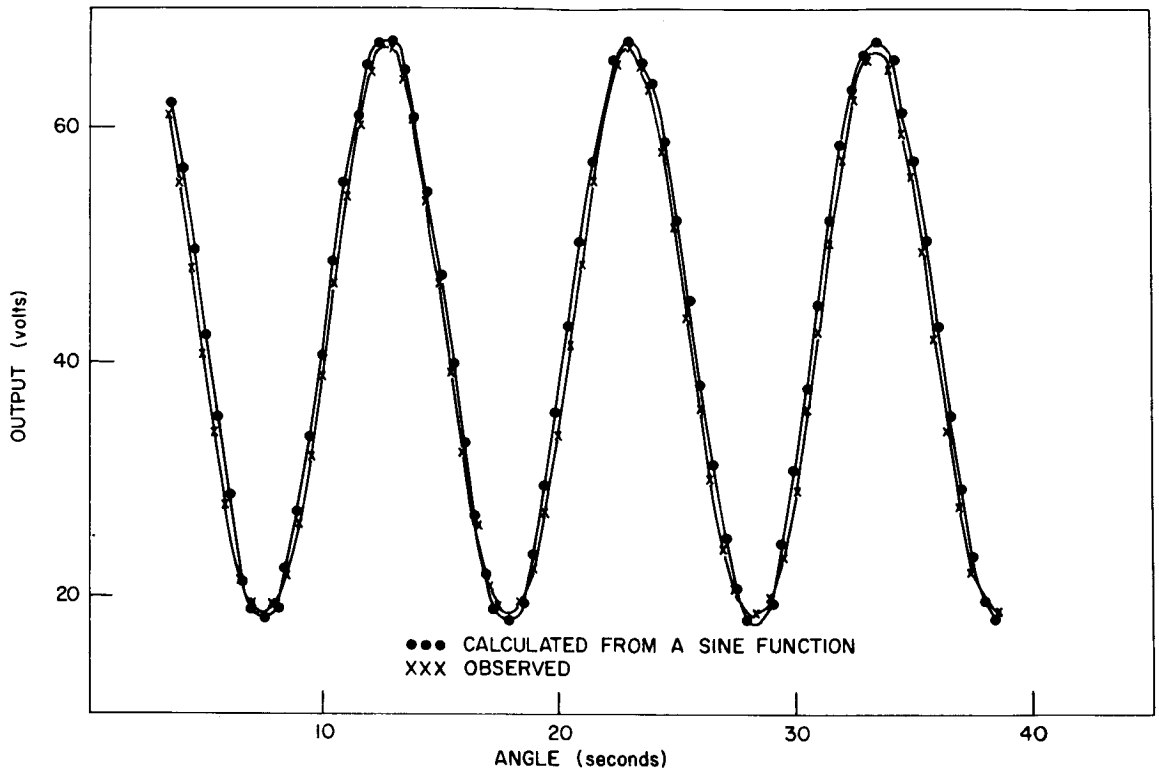


Figure A-40–Observed and Calculated Voltages, Output vs Angle

Table A-21
Variation of V_{max} and V_{min}
(over a scanning range of 1000 seconds)

θ_g	V_{max}	V_{min}	v_0
0	62.90	18.45	40.68
100	62.86	18.27	40.57
200	62.89	18.03	40.44
300	62.80	17.85	40.33
400	62.62	17.74	40.18
500	62.43	17.70	40.07
600	62.24	17.70	39.97
700	61.92	17.73	39.83
800	61.52	17.71	39.62
900	61.02	17.74	39.36
1000	60.63	17.80	39.23

above v_0 for V decreasing, or 10 volts below v_0 for V increasing. The reading of the drum was viewed through a low-power microscope, so that the angle could be estimated to a hundredth of a second. The drum was further rotated until a change of very nearly 20 volts was produced in the output, and the reading of the angle was again taken. The readings were repeated for five successive cycles and on two different days. The results are presented in Table A-22. We observe that the reciprocal slope is practically the same on both days, and that no value differs from the average by more than 0.003 arc-second per volt. This indicates a higher degree of repeatability of measurements than is claimed by the manufacturers of the Midarm and the angle generator.

A precise knowledge of the slope of the V-θ curve permits us to measure large angles using a method of successive increments. The method can best be illustrated by an actual test run made over the range from 200 to 800 seconds of the angle generator. The first ten steps in the addition of the increments are shown in Table A-23 as being sufficient to illustrate the method. Complete results for the 600-second range are shown in Figure A-41.

Table A-22
Reciprocal Slope of V-θ Curve, Measured on Two Different Days
(θ_g in arc-seconds)

V_{max} 63.07 V; V_{min} 17.46 V					V_{max} 63.49 V; V_{min} 17.96 V				
First Series					Second Series				
V	θ_g	ΔV	$\Delta\theta$	$\Delta\theta/\Delta$V	V	θ_g	ΔV	$\Delta\theta$	$\Delta\theta/\Delta$V
29.99	446.79				30.00	469.91			
49.99	448.35	20.00	1.56	.0780	50.00	471.43	20.00	1.52	.076
50.10	451.89				50.00	475.09			
30.10	453.37	20.00	1.48	.074	30.00	476.59	20.00	1.50	.075
29.99	457.26				30.15	480.34			
50.10	458.80	20.11	1.54	.0766	50.16	481.88	20.01	1.54	.077
49.95	462.30				50.08	485.51			
30.05	463.81	19.90	1.51	.0759	30.00	487.00	20.08	1.49	.0742
29.96	467.62				29.99	490.70			
50.15	469.25	20.19	1.63	.0807	50.10	492.22	20.11	1.52	.0756
49.96	472.71				49.99	495.86			
30.05	474.18	20.09	1.47	.0732	29.99	497.39	20.00	1.48	.074
30.15	478.10				29.96	501.12			
49.95	479.61	19.80	1.51	.0762	49.90	502.66	19.94	1.54	.0772
50.05	483.11				30.15	511.52			
30.00	484.61	20.05	1.50	.0748	50.15	513.08	20.00	1.56	.078
30.12	488.50				49.99	516.70			
50.08	489.01	19.96	1.51	.0757	29.99	518.15	20.00	1.45	.0725
49.93	493.56				29.99	521.89			
29.93	494.01	20.00	1.45	.0725	50.10	523.45	20.11	1.56	.0766
Average $\Delta\theta/\Delta$V				.0758	Average				.0757

Table A-23
Sample Data: Comparison of Midarm Model 14
with Angle-Generator

(V_{max} 63.70 v; V_{min} 17.46 v; $d\theta/dV = 0.0758$ sec. v^{-1}.
Period t = 10.345 sec.)

Angle generator reading θ_g	Output Voltage v	Change in output voltage $V_{i+1} - V_i = \Delta V$	Corresponding change in angle $\Delta V \times \frac{d\theta}{dV} = \Delta\theta$	Increment in Midarm angle $t - \Delta\theta = \delta\theta$	Midarm angle $\theta_i + \delta\theta = \theta_m$
200	55.79				200.
210	50.55	5.24	.390	9.955	209.955
220	46.30	4.25	.323	10.022	219.977
230	41.90	4.40	.314	10.031	230.008
240	36.93	3.87	.370	9.975	239.983
250	31.95	4.88	.371	9.974	249.957
260	27.53	4.22	.321	10.024	259.981
262	53.40	25.77	1.953		261.934
272	48.47	5.03	.391	9.954	271.888
282	42.90	5.57	.423	9.922	281.810

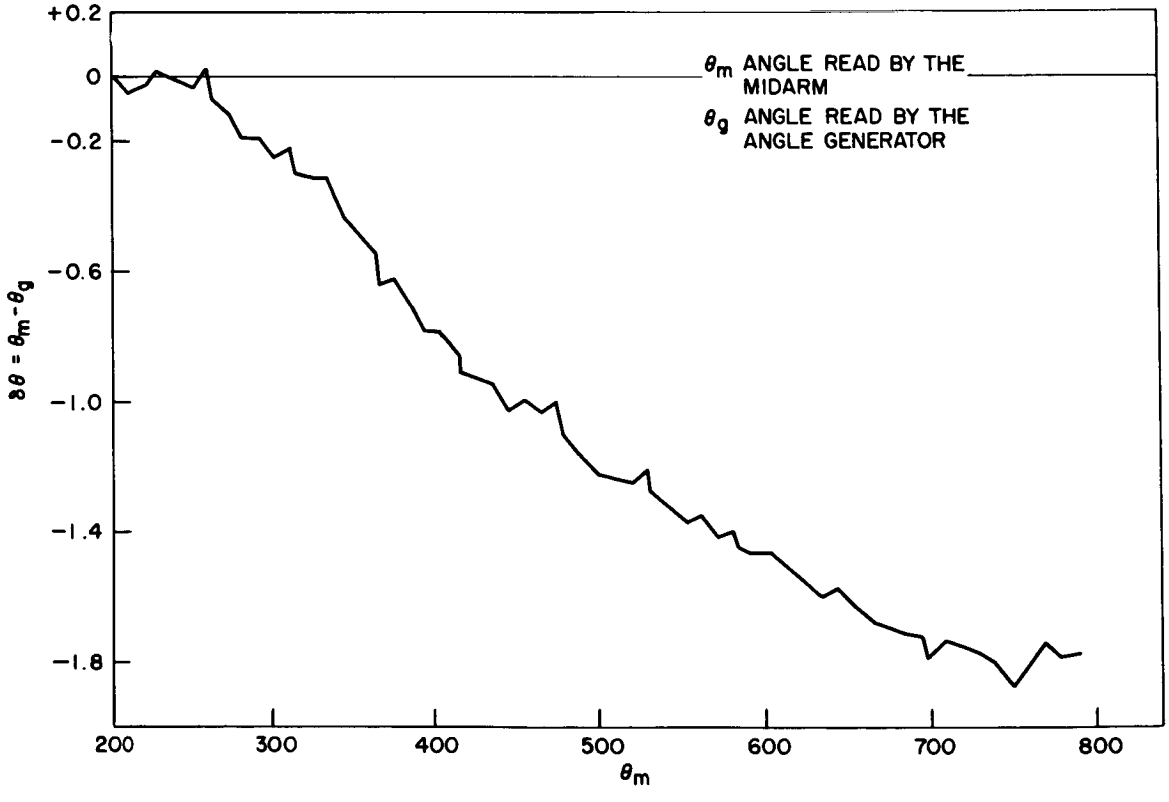


Figure A-41–Comparison of Angle Readings, Midarm Model 14 vs. Angle-Generator

The angle-generator is set initially at θ_g = 200 seconds. The Midarm output is 55.79v. The corresponding angle of the Midarm θ_m is arbitrarily assumed to be also 200 sec., so that $\theta_m = \theta_g$ at the beginning of the scan. On the V-θ curve of the Midarm, this corresponds to a point on the ascending part of the sine curve. The observed values of V_{max}, V_{min}, $d\theta/dV$ and period t are given in Table A-23 above the column entries. If the angle-generator is rotated through 10.345 sec., a complete Midarm period, the final voltage output should be the same as the initial voltage, 55.79v. Actually the rotation is only 10 seconds, and the observed value of the output is 50.55v, that is, 5.24v less. Since $d\theta/dV = 0.0758$ sec V^{-1}, we conclude that the rotation as recorded by the Midarm is through 0.390 (= 5.24 × 0.0758) sec. less than a complete period, that is, through 9.955 sec. This gives a value 209.955 sec. for θ_m. Five more increments of 10 sec. each in θ_g brings θ_g to 260 sec., and the output voltage down to 27.53v. This is sufficiently close to the mean value v_o so that the linear approximation for the change of θ_m with V is still valid. Addition of successive increments of $\delta\theta$ brings θ_m to 259.981 sec. At this point the angle generator is turned through 2 sec., which gives an increase in voltage of 25.77, corresponding to a change 1.953 sec. in θ_m. The new value of θ_m is 261.934 sec. We are once again at a high point in the ascending part of the V-θ curve, so that another series of increments nearly equal to 10 sec. can be calculated and added to θ_m.

The calculations for the range of θ_g 282 sec. to 800 sec. are not shown in Table A-23, but the final results for the whole range are presented in Figure A-41, which is a graph of the difference $\delta\theta = (\theta_m - \theta_g)$ versus θ_m. The difference goes on increasing until it reaches a value about 1.8 sec., at θ_m equal to 800 sec. The similarity is striking between this curve and the calibration curve of the angle-generator by the laser-beam method, shown in Figure A-37. Over the range 200 to 750 sec., $\delta\theta$ in Figure A-37 changes by about 1.8 sec. There is also some indication in Figure A-41 of a decreasing difference beyond 750 sec. The minor flexions in Figure A-41 are probably not as meaningful as the overall downward trend. These flexions may be caused by errors in observation or by the linear approximation's not being valid over a range of 26 volts. Since the V-θ curve is approximately sinusoidal, the minimum value of $d\theta/dV$ occurs at $V = v_o = 45.6$v. At the two extremes the value is about 10 percent lower. From equation (2),

$$\frac{d\theta}{dV} = \left[\frac{2\pi v}{t} \cos \frac{2\pi(\theta - \theta_0)}{t}\right]^{-1} \qquad (3)$$

The value of $\cos 2\pi(\theta - \theta_0)/t$ is 0.894 at 57.8v and 0.918 at 27.5v. These errors affect the values of the correction factors $\Delta\theta$, and since over a wide range the positive and negative values of $\Delta\theta$ are nearly equal, the errors are not cumulative. Thus the general downward trend of Figure A-41 is truly meaningful, and confirms the validity of both the laser-beam calibration method for the angle-generator and the method of successive increments for the Midarm.

We also made a comparison between a Hilger-Watts autocollimator (TA 3-1 155876) and the angle-generator. The Hilger-Watts was mounted very close to the angle-generator so that it could measure rotation about a horizontal axis; it was rigidly fixed on the massive optical table, and the whole assumbly was enclosed in a box covered with aluminized mylar to keep thermal gradients at a minimum. The zero point of the Hilger-Watts is at the center of the screw, at 5 minutes of arc; near this position, readings cannot be taken because of the superposition of the returning image and the cross hairs of the instrument. The Hilger-Watts was first set at θ_h = 322.2 sec. and the angle-generator at θ_g = 650 sec. θ_g was decreased in equal steps of 10 sec. down to 380 sec., and θ_h was measured for each setting. A similar series of readings was taken with starting point of θ_h slightly below the midpoint, at 285.95 sec., θ_g being 350 sec. θ_g was increased 10 sec. at a time to 600 sec. and the corresponding values of θ_h were recorded.

Table A-24
Comparison between the Hilger-Watts
Autocollimator and the Angle-Generator

θ_g	θ_h	$\Delta\theta$	θ_g	θ_h	$\Delta\theta$
650	322.2	0	350	285.95	0
640	332.4	0.20	360	276.10	-.15
630	342.40	.20	370	265.95	0
620	352.10	-.10	380	256.05	-.10
610	362.3	.10	390	246.25	-.30
600	372.25	.05	400	236.10	-.20
590	382.05	-.15	410	226.20	-.25
580	392.20	0	420	216.40	-.45
570	402.10	-.10	430	206.20	-.25
560	412.00	-.20	440	196.30	-.35
550	422.15	-.05	450	186.50	-.55
540	432.05	-.15	460	176.35	-.40
530	441.90	-.30	470	166.40	-.45
520	452.05	-.15	480	156.65	-.70
510	462.05	-.15	490	146.45	-.50
500	471.80	-.40	500	136.50	-.55
490	481.90	-.30	510	126.70	-.75
480	491.95	-.25	520	116.50	-.55
470	501.65	-.55	530	106.60	-.65
460	511.80	-.40	540	96.85	-.90
450	521.70	-.50	550	86.60	-.65
440	531.50	-.70	560	76.65	-.70
430	541.65	-.55	570	66.85	-.90
420	551.50	-.70	580	56.65	-.70
410	561.50	-.70	590	46.70	-.75
400	571.50	-.70	600	36.80	-.85
390	581.40	-.80			
380	591.35	-.85			

The results are recorded in Table A-24, where to the left are given the values for the second half of the range of the Hilger-Watts, and to the right are given the values for the first half of the range. For both ranges the Hilger-Watts was compared with the same central range between 350 and 650 sec. of the angle-generator. As θ_g increases, θ_h decreases, and vice versa. As a first approximation, we assume a linear relationship between θ_h and θ_g, given by the equation $\theta_h = (a_0 + a_1 \theta_g) + \Delta\theta$ where a_0 is a constant, $a_1 = -1$, and $\Delta\theta$ is a small correction factor. It is convenient to take a_0 equal to the $\theta_{go} + \theta_{ho}$, the sum of the values of θ_g and θ_h for the initial setting in each series of measurements. The values of $\Delta\theta$ $(= \theta_g + \theta_h - a_0)$ are given in the third and sixth columns. These values are also plotted as a function of θ_h in Figure A-42. The $\theta_h - \Delta\theta$ points are joined together by the continuous lines.

There are no data for the range of the Hilger-Watts between 285 and 322 seconds; this is the center of the range where readings cannot be taken. The relation between $\Delta\theta$ and θ_g is not linear. There is a cumulative error shown by the upward slope in the first half of the graph and by the downward slope in the second half of the graph. Superposed on this cumulative error is a smaller periodic error of ±0.2 second amplitude, with a period of 30 seconds. One complete turn of the main screw of the Hilger-Watts corresponds to 30 seconds. This periodic error is obviously due to the Hilger-Watts. The dashed

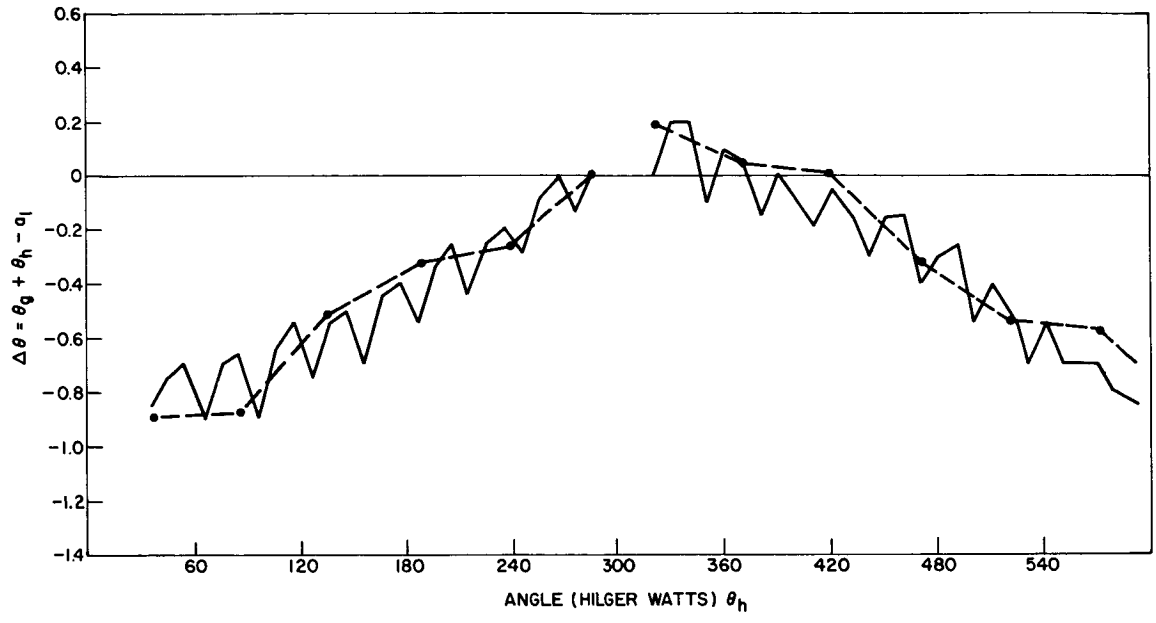


Figure A-42–Comparison of Angle Readings, Hilger-Watts Autocollimator vs. Angle-Generator

curve in Figure A-42 is the calibration curve of the angle-generator, reproduced from Figure A-37 with the necessary shifting of the axes. The y-axis for the dashed curve is $\delta\theta = \theta - \theta_g$, with $\delta\theta$ being assumed to be zero for $\theta_g = 350$ sec. for the first half, and for $\theta_g = 550$ sec. for the second half. We observe that the dashed curve follows closely the general slope of the $\Delta\theta$ curve. Here we have a further confirmation of the laser-beam calibration of the angle-generator. We also conclude that the Midarm, even with the approximation of a linear relation between v and θ, is a more precise method for the measurement of small angles.

Several studies were made on the performance of the Midarm servomechanism. It is actuated by the signals from both the phototubes and is intended to display on a counter the total angle through which the external mirror rotates. The main conclusions of performance tests will be summarized here, without giving the extensive data and the graphs. The variation of the output voltage v with the reading of the Midarm servo, θ_s, is strictly periodic, as is the variation of v with θ_g. In the absence of a precision instrument for measuring angles, the period t can therefore be determined from the servo, provided the angular rotation of the external mirror is sufficiently slow. On the other hand, the servo records the angle only with an accuracy of ± 1 second. Although the accuracy can be improved by balancing the output of the two photomultipliers, this proved to be very time-consuming and not always reliable. A curve of the difference between the readings of the angle-generator and the servomechanism versus θ_g is shown in Figure A-43 for a 20-second range of θ_g. As θ_g increases from 0 to 20 sec., θ_s decreases from 1386.7 to 1366.2 sec. Hence $\Delta\theta$ is defined as $\theta_s + \theta_g - C$, where C is a constant equal to 1386. We observe that, within each cycle of the Midarm, $\Delta\theta$ has two oscillations of unequal amplitude. This form of the curve repeats itself for cycles, provided the adjustments in the electronic circuits remain the same. But, as the amplitude of the output voltage in the two channels changes, $\Delta\theta$ also changes.

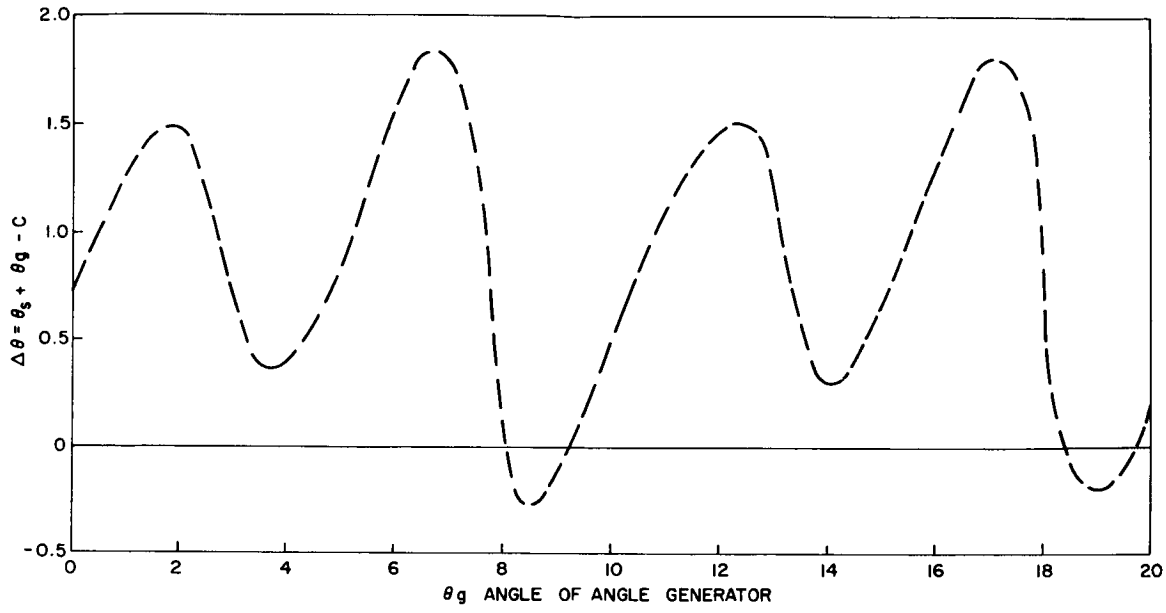


Figure A-43–Comparison of Angle Readings, Midarm Servomechanism vs. Angle-Generator

In conclusion, for measuring angles with the Midarm autocollimator, several alternate procedures are available. The more painstaking and time-consuming the method, the higher is the order of accuracy attainable. For measurements with a relatively poor accuracy of ±1 sec., the angle may be read directly from the Midarm servomechanism. The correction curve of the servo may be used to diminish the margin of error. For greater accuracy, the method of successive increments discussed earlier may be applied. This requires an arrangement whereby the angle is changed in small increments, so that a reading of the voltage output can be taken once or twice in each cycle. A more elaborate procedure to ensure accuracy of the order of 0.01 arc/sec is to assume the sinusoidal variation of v with θ and to apply the equation (2). Since

$$V - v_0 = v \sin \frac{2\pi}{t} (\theta - \theta_0),$$

$$\theta = \theta_0 + \frac{t}{2\pi} \sin^{-1} \left(\frac{V - v_0}{v} \right).$$

If V_1 and V_2 are the initial and final output voltage, the corresponding difference in angle is

$$\Delta\theta = \theta_2 - \theta_1 = \frac{t}{2\pi} \left[\sin^{-1} \frac{V_2 - v_0}{v} - \sin^{-1} \frac{V_1 - v_0}{v} + n\pi \right].$$

The term $n\pi$ within the brackets is added since the angles are to be evaluated from the inverse sine functions. n is an integer whose value can be determined unambiguously

from the total number of periods between θ_1 and θ_2 and from the relative phase of V_1 and V_2. The accuracy is relatively poor if V_1 or V_2 are close to the maximum or minimum values of v; but, since there are two phototubes whose outputs are 90 degrees out of phase, one or the other phototube will give a sufficiently accurate value of the two angles.

A GRAVITY-REFERENCE SYSTEM FOR PRECISE ALIGNMENT OF INSTRUMENTS

A fixed direction to which angles measured by autocollimators can be referred is as necessary in the alignment of the spacecraft experiments as is the instrumentation for precise measurement of angles. Assume, for example, the simple case of three or four optical instruments in a satellite, all of which have to view a distant star. The optic axes of all the instruments should be exactly parallel to each other. The optical testing facility should be able to determine whether the instruments are truly aligned and whether any misalignment has been introduced during the severe prelaunch tests of solar simulation, vibration and shock, exposure to high and low temperatures, and the like.

One obvious method of testing the parallelism of the optic axes of the instruments is to sight a reference mirror on each instrument through an autocollimator. The autocollimators themselves should be precisely aligned; for this, we would require a fixed reference direction which can be sighted by all autocollimators in their respective positions. A large optical flat, aluminized for maximum reflectance, would provide such a reference direction; normals to the mirror surface at different points are all parallel. But, given the large area over which the experiments will be distributed and the accuracy with which the angles are to be measured, the optical flat of such a large size will raise serious problems of structural rigidity and stability of support; such an optical flat may be impossible to build and test, even at a prohibitive cost.

An alternate approach to the problem is the gravity-reference system. The surface of a pool of mercury is a good reflector, and the normal to the surface is along the local vertical or the direction of the force of gravity. The autocollimators fixed at suitable locations to view the optical instruments will first be aligned along the vertical by sighting them on the mercury surface. The mercury pools will then be replaced by the instrumental package and the misalignment, if any, of each instrument can be read by the autocollimators.

The principle of the mercury-pool gravity reference has long been used by astronomers for measuring positions of stars. The so-called reflex zenith tube was designed by Sir George Airy[5] in 1854. The photographic zenith tube based on the same principle was developed as an instrument of high accuracy by Frank E. Ross.[6] Several modifications of the photographic zenith tube have since been added, notably by W. Markowitz[7] and his co-workers at the U. S. Naval Observatory, Washington, D. C. A slightly modified application of reflection from a mercury surface to measure directions has been described by André Danjon.[8] It is the so-called Impersonal Astrolabe which, under favorable atmospheric conditions can record the transit of a star with an accuracy of 0.09 arc-sec. (standard deviation).

When the mercury surface is used to reflect light to an autocollimator, stability of the surface is an important consideration. The small vibrations of the building and of the support set up standing waves on the mercury surface. Only after a very long time does the surface become sufficiently flat to give a strong reflected image. We were interested in finding out how best the mercury surface might be used with autocollimators.

A theoretical problem closely connected with this is the stability of the gravity reference itself. The normal to the mercury surface is along the direction of the force of gravity. The gravity vector has three components which are essentially different; $\vec{g}$ is a vector sum, $\vec{g}_1 + \vec{g}_2 + \vec{g}_3$. The first component $\vec{g}_1$, the most important one, is due to the attraction of the earth; in the absence of surface anomalies, it is directed towards the center of mass of the earth. The second component is due to the centrifugal force. The third is the most significant for our purpose, because it varies with the hour of the day; it is the tide component due to the attraction of the moon and the sun. Hence the local vertical, the direction of $\vec{g}$, rotates about the direction of the earth center as a conical pendulum of very small and variable angle ϕ. When the vector $\vec{g}_2 + \vec{g}_3$ is along the horizontal, the deflection ϕ of the vertical is a maximum. The magnitude and the direction of ϕ vary with the zenith angle z; $\phi = k \sin 2z$. k is a constant equal to 8.5×10^{-8} radian for the moon and 3.9×10^{-8} radian for the sun.[9] For $z = 45°$, $\phi = k$. If the forces due to the sun and the moon are in the same direction, $\phi = 12.4 \times 10^{-8}$ radian which is equal to 2.55×10^{-2} arc-second. Thus the maximum possible deflection of the gravity vector about the direction of the earth center is 2.55×10^{-2} arc-sec. As in the case of tidal movements, the angle ϕ oscillates about the earth center direction; there are two complete oscillations per day, and the amplitudes of the two oscillations are unequal. The maximum rate of change of the gravity reference line is of the order of 4 thousandths of an arc-second per hour.

This effect, though very small, is not altogether negligible when we are attempting measurements with an accuracy of 1/100 arc-second. It should also be remembered that, as stated earlier, the directions of the earth center at two points one foot apart on the surface of the earth differ by 1/100 arc-second. There are no measurable changes in the direction of the gravitational force due to local movements of tidal waters or movements within the crust of the earth.[10]

Having established that the local direction of gravity is sufficiently constant to serve as a reference system, we look for a mercury surface having high reflectance, free of ripples, and of negligible oxidation upon exposure to the atmosphere. The major problems with the mercury surface are ripples which are caused too easily and dampen out too slowly, and oxidation which reduces the reflectance unless the mercury is extremely pure and its container very clean.

R. Bunnagel,[11] in a paper presented at the Colloquium on "Optics in Metrology" in Brussels in 1956, suggested that a layer of mercury about 0.2 mm thick can be used as a standard flat, because it is insensitive to vibrations. According to the author, by silvering the inner surface of the vessel, the mercury surface is prevented from oxidizing and maintains its high reflectance for a long time. Bunnagel tested the flatness of the mercury surface by two interferometric methods. He also found that a very rigid support is essential, lest slight deviations from the horizontal make the mercury flow and the film disappear. Our own experience in building such a thin-layer optical flat was inconclusive. For controlling the thickness of the mercury film, a vessel having a movable piston as its bottom surface was tried. We found that a film thickness of about 0.2 mm is difficult to maintain, that rigidity and horizontality of the system could not readily be achieved, and that the long exposures to air required in our alignment tests caused oxidation of the mercury surface and loss of reflectance.

The two major problems were ripples and oxidation. We felt that both problems could be solved by pouring over the mercury a liquid which was sufficiently high in viscosity and transparency, and which would not react chemically with mercury. An oil known under the trade name Molvac A, supplied by F. J. Stokes Corporation, Philadelphia, Pa., was found to satsify the requirements. It has a kinematic viscosity of about 500 centistokes,

which is very high compared to that of most liquids; mercury has a viscosity of 0.12 centistoke, and most industrial oils and lubricants have viscosities ranging from 4 to 20 centistokes.[12]

A quantity of triple-distilled mercury is poured into a flat dish made of glass to form a pool of about one cm depth. Over the mercury is poured the Molvac A oil to about 1/2-cm thickness. The amount of oil and mercury is not critical, except that there should be enough of both so that the mercury surface is flat over a sufficiently large area, and the oil is sufficiently heavy to dampen the ripples in the mercury. Figure A-44 shows the shape of the flat dish, which is an ash tray, and the relative amounts of mercury and oil. The dish has 15.5-cm internal diameter at top of the rim, 13.5-cm diameter for the flat portion of the base, and 3-cm depth. The vertical sides curve gently to meet the base, with a radius of curvature of about 1 cm. This curvature of the sides is an additional factor for effectively dampening out the ripples. The edges of the mercury pool have a convex meniscus. The contact between the edges and the container is along the concave portion of the sides of the container. If ripples are started on the mercury surface, they reach the glass walls at such an angle that the energy of the waves is almost completely absorbed. It is relatively more difficult for standing waves to be formed in a vessel of this shape than in one with vertical sides. The use of an ash tray was accidental; it happened to be ready at hand. Later tests with vessels having vertical sides, a large-diameter paper cup, and a glass beaker, were less satisfactory.

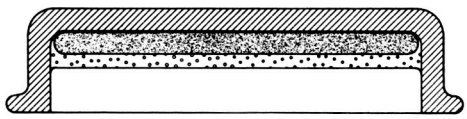


Figure A-44–Mercury Pool with Viscous Oil

Figure A-45 is a photograph of the experimental setup for measuring the stability of the mercury-pool gravity reference. A Hilger-Watts autocollimator is mounted on a tripod of the kind used in geodetic survey. Below the autocollimator is the ash tray with mercury and oil. The measurements showed a high degree of repeatability, with a maximum deviation of 0.3 arc-second, which is probably all due to the relative instability of the tripod. The return image was almost as clear as when the mercury pool was replaced by an aluminized optical flat. An independent experiment with a photometer showed that the reflectance of the mercury surface, as viewed through the oil film, is 80 percent. Exposure to the air for several days produces no appreciable change in the reflectance; we therefore conclude that there is no oxidation at all of the mercury surface. When the oil is first poured over the mercury, a large number of small air bubbles are seen in the oil. Fifteen or twenty minutes are required before all air bubbles disappear and the oil becomes totally transparent.

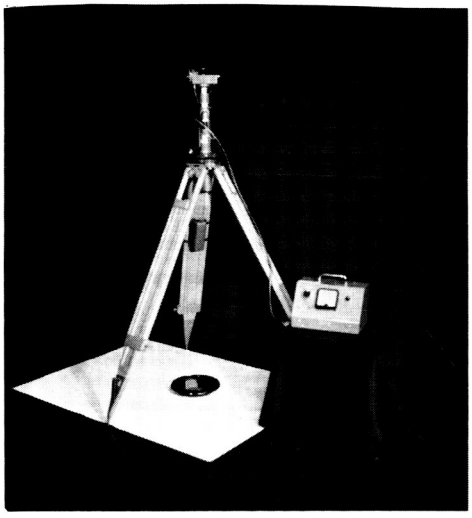


Figure A-45–Apparatus for Testing Mercury Pool

If the container is tilted or otherwise disturbed, and then replaced below the autocollimator, the return image is at first indistinct; in a few seconds, a faint image appears which oscillates about the central cross hairs. This is due to the vibrating mercury surface. In about a minute the vibrations are dampened out, and the image becomes quite bright and distinct. While the mercury surface is becoming steady, a rather faint and less

unsteady image appears on the side. This is due to reflection from the top surface of the oil. Gradually this faint image approaches the brighter image and coincides with it, as the surface of the highly viscous oil becomes horizontal and parallel to the mercury surface. Within about 2 or 3 minutes the system is sufficiently steady for another measurement. Tests were made on the mercury surface without any oil film and with a different fluid considerably more viscous, a 200 000 centistoke fluid manufactured by the Dow Chemical Corporation. In neither case could a ripple-free surface of mercury be obtained.

Thus we have, in the mercury surface covered with a layer of Molvac A oil, a gravity-reference system which combines the advantages of constancy within a few hundredths of an arc-second, reflectance almost the same as that of an aluminized surface, and easy reproducibility in any desired location.

ACKNOWLEDGMENTS

Our thanks are due to several members of the Engineering and Optical Facilities Section and in particular to the section head, James L. Diggins, for their help in procuring as expeditiously the items of auxiliary equipment which were needed at different stages, and for helpful discussion throughout the course of this study. Figures A-29, A-30, A-31, and A-32 are taken from a paper presented by J. L. Diggins at the midsession meeting of the Summer Workshop. The comparison between the angle-generator and the Midarm, and between the angle-generator and the Hilger-Watts, was done by R. Crist in cooperation with the second author. Computer programming for the laser-beam calibration by the least-squares method was done by John Nielsen. Kathleen Thompson, NSF summer science student, cooperated closely with the authors for data handling and analysis throughout the course of the research. To all of them the authors wish to express their appreciation.

REFERENCES

1. Czerny, M., and Turner, A.F., "Uber den Astigmatismus bei Spiegelspektrometern," Z. Physik, 61, 792 (1930); J. Strong, Concepts of Classical Optics, W.H. Freeman and Co., San Francisco (1958) p. 371

2. Fastie, W.G., "A Small Plane Grating Monochromator," J. Opt. Soc. Am. 42, 641 (1952); W.G. Fastie, "Experimental Performance of Curved Slits," J. Opt. Soc. Am. 43, 1174 (1953)

3. Hamilton, A.S., "Improving Rate Tables for Gyro Testing," Electronic Industries, 22, No. 9, 72 (September 1963)

4. Evans, J.C., and Tylerson, C.O., Measurement of Angle in Engineering, Notes on Applied Science, No. 26, National Physical Laboratory, Teddington, Middlesex, England (1961)

5. Airy, G.B., Astr. Mag. and Meteor. Obs. of the Greenwich Observatory, Appendix I (1854)

6. Rose, F.E., Latitude Observations with the Photographic Zenith Tube at Gaithersburg, Md. Spec. Pub. No. 27, U. S. Coast and Geodetic Survey (1915)

7. Markowitz, W., "The Photographic Zenith Tube and the Dual-Rate Moon Position Camera," Chapter VII of Telescopes, edited by G. P. Kuiper and B.M. Middlehurst, University of Chicago Press, Chicago, Ill. (1960) p. 88

8. Danjon, André, "The Impersonal Astrolabe," Chapter VIII of Telescopes, edited by G.P. Kuiper and G.M. Middlehurst, University of Chicago Press, Chicago, Ill. (1960) p. 115

9. Jensen, H., "Formulas for the Astronomical Correction to the Precise Levelling," Bulletin Geodésique No. 17, Afdelingsgeodaet ved Geodaetisk Institut, Copenhagen (1950) pp. 267-277

10. Jeffreys, H., The Earth, Cambridge University Press, England (1959) p. 228, to which reference may be made for a more complete discussion of the topic

11. Bunnagel, R., "Herstellung und Verwendung eines gegen erschütterüngen unempfindlichen Quecksilberspiegels als Ebenheitsnormal" in Optics in Metrology, edited by Pol Mollet, Pergamon Press, New York (1960) pp. 172-178

12. Gray, D.E., editor, American Institute of Physics Handbook, McGraw Hill Book Company, Inc., New York (1957) p. 2-168

Page intentionally left blank

THE ILLUMINATING SYSTEM FOR THE VACUUM OPTICAL BENCH

Dr. Irvine C. Gardner

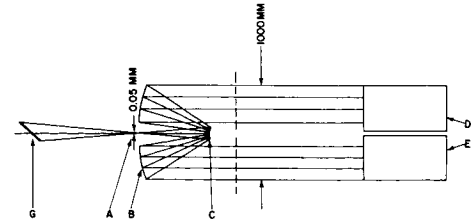
The Vacuum Optical Bench (VOB) is essentially a very large collimator system mounted in a vacuum chamber. The beam of parallel light of large diameter that is produced is intended to simulate a beam of the character that an optical instrument carried by a satellite will receive from a star when in outer space. The performance requirements that are established make demands for which there is no body of scientific engineering data; many of the requirements go beyond the frontier of known optical performance.

The spectral range to be provided extends from 5000 to 900 angstroms. The earth's atmosphere limits the radiation we receive from celestial bodies to wavelengths greater than 3000 angstroms. As the wavelength becomes shorter than this value, the experimental difficulties increase. For wavelengths shorter than 1800 angstroms, the air absorption, even for the short distances encountered in laboratory instruments, becomes important. For wavelengths shorter than 1200 angstroms, there is no known solid that is transparent. Hence, all control of the course of the rays must be by reflecting elements, and there can be no solid material to separate the interior of the electric-discharge tube serving as a source of light from the optical system.

A beam of parallel light large in diameter is required because some of the optical systems to be tested have apertures of very large diameter: in other cases, several systems of lesser aperture are mounted in parallel so that, collectively, a very large aperture is presented. To give an idea of the difficulty presented by this demand for a large aperture, it should be noted that in some applications the aperture of the monochromator by which the light enters the optical system is only 0.050 mm (0.002 inch) in diameter, and the final beam of parallel light is approximately 1000 mm (39 inches) in diameter. The ratio of the two corresponding areas is $1{:}4 \times 10^8$. Rather an imposing ratio! Over this very large beam it is further required that the value of $d\varphi/dA$, the flux per unit area, shall not vary by more than ±5 percent. This uniformity is required because many of the optical systems to be placed in orbit are to make photometric measurements, and this uniformity is necessary in order to calibrate these instruments.

A third requirement is that it be possible to isolate from any part of the spectral range, radiation having a bandwidth not greater than 1 angstrom. The isolation of radiation having so small a spectral bandwidth presents no great difficulty because it makes no severe demand on the resolving power of the monochromator system; but is truly difficult to provide a source of such intensity that it will provide a measurable amount of power with this small bandwidth over a beam of so large a cross section.

For the photometric calibration to which reference has been made, it is desirable to be able to vary the illumination over the cross



G– DIFFRACTION GRATING (OTHER PARTS OF MONOCHROMATOR NOT SHOWN
A– APERTURE OF MONOCHROMATOR
B–C MIRRORS OF CASSEGRAINIAN COLLIMATING SYSTEM
D–E OPTICAL SYSTEMS IN POSITION FOR TESTING

Figure A-46–Schematic Sketch of Vacuum Optical Bench

section of the beam from 50 to 10^5 photons per (cm^2 sec) in known steps. This is a requirement that will be extraordinarily difficult to meet.

Furthermore, consideration must be given to the possibility, at a future date, that a VOB will be required which operates not only in a vacuum but at a temperature of -60° Centigrade, in order to more completely simulate the environmental conditions that will be encountered in outer space.

It is well to know that the justification for these requirements derives directly and logically from the tests that are to be made. Otherwise, one might well think that the conditions were formulated with the purpose of presenting a difficult problem. The monochromator and source of light can advantageously be considered as a single unit. Each of these requirements that has been listed exercises control upon the design, construction, and operation of the source-monochromator unit.

First, the requirements upon the light source will be considered. It has been mentioned that there is no solid material known that is transparent to the shorter wavelength portion of the spectral range that is to be available. At present, an electric discharge in hydrogen appears to be the most promising source, at least for the shorter wavelengths. Indeed, such a source has been constructed and is now in use on the VOB with the present monochromator. The pressure in the vacuum chamber enclosing the VOB is 10^{-5} Torr, which is very much too low to suitably conduct the electric discharge in the lamp chamber. The discharge tube is all enclosed except for the aperture, a few thousandths of an inch in diameter, by which the light enters the monochromator. Experience has shown that, by the controlled admission of hydrogen to the discharge tube, and by reason of the restricted outflow of hydrogen from the discharge tube through the small aperture into the larger vacuum chamber, it is possible to have the pressure in the discharge tube adequate for the electric discharge, and at the same time to maintain the pressure in the outer vacuum chamber as low as 10^{-5} Torr. This is not the only problem to be encountered in the design of the lamp; as can be readily seen, a very large luminous flux must be provided for adequate illumination over the entire cross-section of the beam. In the lamp presently in use, the capillary is of heavy ceramic material, a large electric current is passed through the lamp, and water cooling is required. There is evidence of heating of parts of the monochromator in the neighborhood of the aperture for the passage of the light, which suggests that the glow of the lamp has extended beyond the limits of the lamp into the monochromator space.

The luminous portion of the discharge cannot be right up against the slit, and it is necessary that it be sufficiently large to permit an even illumination over all the surface of the grating. If the luminous flux does not fall equally upon all parts of the grating, the illumination over the large cross-section of the collimated beam will certainly not be uniform; this is a condition that must be avoided. The current density in the capillary of the lamp is so great that the spectral lines appear upon a strong background of continuous radiation, which is a great advantage. It is quite possible that the hydrogen lamp with the differential pressure that has been described will not be satisfactory over the entire spectral range; if this should be true, a supplementary source must be provided. This offers less difficulty than does the first lamp, because transparent solid materials for the longer wavelengths are available, and the lamp can be sealed off from the vacuum chamber in the conventional manner.

The variation of the illumination from 50 to 10^5 photons per (cm^2 sec) is particularly difficult to obtain in the shorter wavelength portion of the spectral range. The most obvious method of varying the intensity is to vary the electric current passing through the lamp, but this probably will not be a satisfactory method of varying the intensity over the large range that is required. As the current is diminished, a stage will probably arrive

when the source becomes so small that the entire grating area will not be illuminated, and the illumination for different parts of the large cross section of the beam of parallel light will no longer be uniform. For the portion of the spectral range for which transparent solid substances are available, filters of different densities may be mounted on a disc to permit different ones in sucession to be brought into the beam to obtain a variation of intensity. For the portion of the spectral range lying between 1850 and 1200 angstroms, it will be possible to have a filter cell with parallel windows of lithium fluoride, and to vary the transmission by varying the pressure of some gas (say oxygen) in the cell. The satisfactory operation of such a cell is relatively certain.

An alternate method is control of the area of the aperture by which the light enters the monochromator by means of a variable diaphragm. After the slit reaches a certain width, further opening will increase the size of the illuminated area without increasing the value of $d\varphi/dA$. The low resolving power of the television circuit by which the observations are transmitted to the earth somewhat restricts this difficulty. A fourth method for modulating the illumination is the use of a rotating sector. However, it is very difficult to vary the transmittance in the space available over the range from 50 to 10^5, and it is not a simple matter to install a high-speed electric motor to operate in a vacuum of 10^{-5} Torr. The rotating sector has one attractive feature: the attenuation for any setting will be the same for all wavelengths, and no solid material need be introduced into the optical path.

In the present monochromator, the grating is of the so-called tri-bar type. The grating is concave and, as the light is incident on different parts of the grating at different angles, the angle of blazing should vary continuously from one edge to the other. As this is impracticable, the grating is divided into three sections and the blazing angle is changed in three steps, each blazed at a different angle for the same wavelength. It is difficult, if not impossible, to do this precisely; even if this were accomplished, it would be only an approximate approach to the ideal blazing with a continuously varying angle. The result is that the illumination over the large section of the beam of parallel light is very non-uniform. A concave grating was used because it not only produces a spectrum, but it also focusses the light on the exit slit, so that it is necessary only to use two reflections in the monochromator with a consequent moderately high transmittance. This is particularly advantageous in the shorter wavelength portion of the spectrum because the reflection coefficient for all materials for these wavelengths is very low, and each additional mirror results in a large additional loss of light. Despite this, it now seems advisable to construct the monochromator with a plane grating and two concave mirrors to focus the light. With a plane grating, for any wavelength, all the rays are incident on the grating at the same angle, the blazing should be the same over the entire surface, and one should therefore obtain greater uniformity of illumination over the entire cross section of the large beam.

It is obvious that, if one is to deal with wavelengths shorter than 1200 angstroms, a reflecting optical system must be employed. Even if these shorter wavelengths were not to be utilized, it is still probable that a reflecting system would be employed. There are many reasons for this: When making an optical system of large aperture, discs of glass sufficiently homogeneous to transmit the light uniformly are extremely difficult to make and extremely expensive, and not less than four surfaces must be ground and polished for the production of one achromatic component. A refracting telescope 40 inches in diameter – the largest refractor in the world – was constructed about 50 years ago, and is now installed at Yerkes observatory. No large refractors have been constructed since that time, and all large telescopes have been reflectors. For a given relative aperture, the curves are much shallower on a reflector than on a refracting telescope; this simplifies the correction of the aberrations and the grinding and polishing of the surface. Furthermore, only one reflecting surface is required to replace the four surfaces of the achromatic component.

A reflecting surface is perfectly achromatic, whereas the best refracting system approaches – but does not attain – complete achromatism. Also, the focus of a reflecting system does not change with a change of index of refraction of the ambient medium. In other words, a reflecting system is perfectly achromatic and remains in perfect focus as the pressure of the surrounding air is reduced from normal atmospheric to 10^{-5} Torr. For a system which is to be placed in orbit, this is a great advantage. The instrument can be focussed and adjusted at normal atmospheric pressure for a single convenient wavelength, and it will then be in perfect focus when in outer space. This assumes that the reduction of air pressure will not cause relative displacements of the different parts of the optical system.

It has been mentioned that a future VOB may be built which will simulate both the pressure and temperature of outer space. This will greatly increase the problems of design. Lowering the pressure without changing the temperature may be expected to produce little or no change in the dimensions and curvatures of the optical components. Changing the temperature from our ambient temperature to -60°C will cause a significant change in the curvatures of the reflecting elements, and will bring about an even more important change in the relative positions of the parts. If temperature equilibrium is not attained before testing begins (that is, if the temperature of a reflecting element is not uniform), an even greater change in the optical performance is to be anticipated. After the temperature is reduced, it will be necessary to wait a considerable period of time in order that all parts may reach a uniform temperature. It will certainly be advisable and probably necessary to incorporate in the VOB an auxiliary optical system by which the correction of focus can be tested, as well as an externally operated mechanism for adjusting the focus. For testing the focus, use will probably be made of the principles utilized in the military rangefinder.

PROJECT B: HEAT TRANSFER AND CRYOGENIC PUMPING IN VACUUM TECHNOLOGY

CONTENTS

ILLUSTRATIONS

TABLES

SUMMER WORKSHOP 1964

Program Outline and Team Participants

PROJECT B: Heat Transfer and Cryogenic Pumping in Vacuum Technology

Study Topics

B-1 Theoretical analysis and experimental investigation of cryosorption

B-2 Performance evaluation of space thrusters

B-3 Establishment of criteria for pressure level and duration of tests on spacecraft systems; defining role of vacuum in space simulation

B-4 Interpretation of flight data from Explorer XVII

TEAM B

Academic Personnel	Goddard Personnel	Code
Dr. G. H. Beyer, principal investigator	Dwight C. Kennard, staff advisor	320
Luis A. Veguilla-Berdecia	Dr. J. E. A. John	322
J. B. Crooks	William Hardgrove	322
F. Owens	Ron Sheffield	322
Donald R. Hardesty	Kenneth L. Rosette	322
S. B. Pearce	George Newton	651
J. C. Heine	Carl Reber	651
Barry Yaffe	Archimede Panagopolous	623
Dr. Yu Chen	Carl A. Wagner	623
	James M. Bridger	623

NSF Summer Science Student
Erik Schultz

Part-time consultants: Dr. Daniel Alpert; Dr. John Simons, Jr.; Dr. F. M. Propst

ACKNOWLEDGMENTS

Team B of the 1964 Summer Workshop appreciates the splendid cooperation of Drs. Daniel Alpert, Norman Peacock, and Franklin Propst of the University of Illinois Coordinated Science Laboratory. The wide experience of each in the field of vacuum measurements has been most helpful. The papers by Dr. Propst, "Anomalous Effects in Vacuum Gauges" and "Application of Ceramic Fiberfrax Paper in Vacuum Pumping," bear directly on the work of Team B, and the Work-Shop is grateful for their inclusion in the Final Report. The suggestions of Dr. Peacock regarding instrumental techniques for measuring the backstreaming of diffusion-pump oil are also gratefully acknowledged.

The investigation on "Criteria for Vacuum Testing" undertaken by Dr. John C. Simons, Jr., involves, in part, the study on "The Measurement of Outgassing" by Frank Owens and Dr. Yu Chen, as well as "The Analysis of Spacecraft Failure" by Kenneth Rosette. From the results of these studies, together with other available data, Dr. Simons will analyze and evaluate the thermal-vacuum test criteria currently in use. This information will be summarized in a report by Dr. Simons which will be issued in December 1964 as a supplement to the present report. The cooperation and interest of Dr. Simons in this complex task are much appreciated.

PREVIEW

G. H. Beyer

Vacuum technology is so broad a subject that it is not surprising that Project B of the Summer Workshop was concerned with a very wide range of problems. Increasing sophistication is becoming evident in spacecraft design - particularly the inclusion of optical systems whose performance can be markedly degraded by minute amounts of contamination. To test such systems effectively calls for the careful control of environmental conditions and the incorporation of new methods into conventional testing procedures.

One new method of producing the high vacuums needed for testing spacecraft, with a minimum of contamination by diffusion-pump oil, is cryosorption. At temperatures of the order of 20°K, substantial quantities of hydrogen can be tightly held on the surface of charcoal and molecular sieves. Because the absorbed material exhibits only a small fraction of its normal vapor pressure, cryosorption panels can be used to immobilize the hydrogen commonly outgassed by spacecraft in a high-vacuum space chamber. The first three reports that follow deal with the role of cryosorption in vacuum systems, a better understanding of the theoretical aspects of cryosorption, and a literature survey on recent articles in that field.

Another problem studied by Group B was the performance evaluation of space thrusters. Satellites such as Syncom III are accurately positioned by short bursts of gas expanded through small nozzles. Better methods are being sought for testing such thrusters on a sensitive stand in a vacuum environment.

A third area of interest to Group B involved means for improving the vacuum testing of spacecraft. Two approaches were used: First, to apply mass spectrometer measurements on a more routine basis to identify the gases evolved during testing; second, to examine the experience accumulated with various satellites and to evaluate the effectiveness of the testing procedures currently used.

Interpretation of flight data from Explorer XVII was the fourth problem studied by Group B. During its 100-day lifetime, this satellite sent back a phenomenal number of pressure readings at altitudes varying from 260 to 900 kilometers above the earth. Such data were complicated by the fact that the satellite was revolving rapidly while traveling at a speed of 7.5 km/sec through an atmosphere of changing composition.

The last section of this report on the activities of Group B includes some comments and suggestions which may be helpful in continuing the efforts outlined above.

Page intentionally left blank

ROLE OF CRYOSORPTION IN ULTRAHIGH-VACUUM SYSTEMS

Dr. J. E. A. John

It is generally accepted that the pressure in interplanetary space is in the order of 10^{-12} to 10^{-14} torr. Duplication of the environment that a satellite encounters in space would require a chamber capable of this level of vacuum. However, it is really not necessary to reproduce the spatial environment in a simulator, only the effects which occur as a result of this environment. As an example, for thermal balance studies, a simulator capable of 10^{-5} torr is adequate, this level of vacuum being sufficient to eliminate convective heat transfer. For surface effects studies, much lower pressures would be required. What are needed are test criteria, specifying the level of vacuum necessary for testing a given satellite with a given mission. To establish these criteria, a chamber is needed that is capable of space vacuum. Such a chamber, with a cylindrical test volume 20 by 20 inches, has been designed by the Vacuum Research Section, Thermodynamics Branch, and is now under construction.

In a vacuum chamber capable of maintaining pressures of 10^{-14} torr in the presence of a test object, high pumping speeds are required at high vacuum, to handle both gases initially present in the chamber and those outgassed during the test. Provisions also must be made for reducing the gas leak rate to a low level.

Cryogenic pumping seems best for this application, because it provides high pumping speeds and the volumetric pumping speed is independent of pressure. Furthermore, large pumping surfaces can be provided, the cryogenically cooled walls acting as pumping area.

Disadvantages of cryogenics are the expense involved, the cost of operating a liquid helium system at 4.2K being almost prohibitive; also, only condensable gases can be pumped. If it is assumed that, from cost considerations, a gaseous helium refrigerator is to be used, and if the system is capable only of 10K to 20K, difficulty would be experienced in pumping hydrogen, with vapor pressure at 20K of 760 mm of mercury. In order to pump hydrogen and still not have to resort to a 4.2K system, the system design has incorporated a cryosorption panel.

Cryosorption refers to the adsorption of certain gases on cryogenically cooled surfaces. Several adsorbents have a great affinity for hydrogen when cooled to 20°K and below; however, two problems have prevented its use: first, containment of an adsorbent bed in a vacuum chamber: second, insurance that the entire bed is maintained at 20K or below, as adsorbents inherently possess a low thermal conductivity. Both these problems have been solved by the bonding of a thin layer of a suitable adsorbent, Molecular Sieve 5A, to a metal plate. The bond has been produced by the Linde Co. The resultant panel produces high pumping speeds for hydrogen, demonstrating a large capacity for the hydrogen molecule and providing a minimum of contamination for a chamber. The Thermodynamics Branch has followed the work of the Linde Co. and has negotiated for the purchase of cryosorption panels, both for testing and for inclusion in the vacuum chamber mentioned previously.

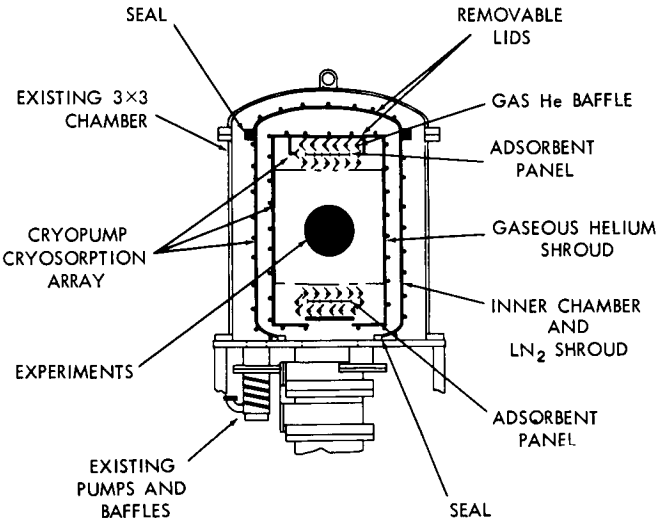


Figure B-1-Conceptual Design for Cryosorption Chamber

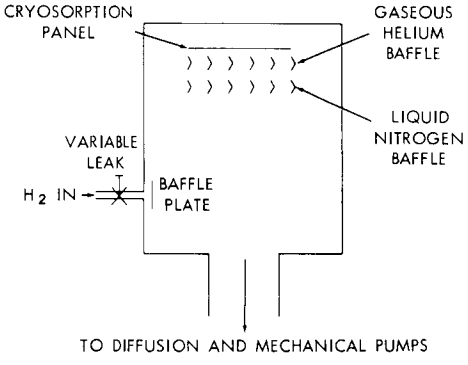


Figure B-2-Experimental Cryosorption Array

Figure B-1 shows the design of this ultrahigh-vacuum facility using cryosorption. The chamber consists of a series of concentric chambers, each a vacuum vessel, thus minimizing the real gas-leak rate into the test volume. The outer chamber is cooled with liquid nitrogen to act as a heat shield for the inner gaseous-helium-cooled shroud. The cryosorption panel is baffled to prevent contact with condensable gases and reduction of its capacity for hydrogen. The system provides high pumping speeds at ultra-high vacuum, both for condensable gases (by gaseous-helium cryogenics) and for non-condensable gases (by cryosorption). It is estimated that the chamber will be capable of 10^{-15} torr with a reasonable gas load from a test object.

While the system described above is being built, the Thermodynamics Branch will conduct a series of experiments with a cryosorption panel, both to gain familiarity with its operation and to provide additional data.

Essentially, two types of experiments are planned: a dynamic experiment, to determine pumping speeds of the panel for hydrogen; and a static experiment to determine the panel's capacity for hydrogen. Figure B-2 shows the apparatus to be used for the dynamic experiment. Hydrogen will be allowed to enter the chamber at a controlled rate through a variable leak. The hydrogen will strike a baffle plate, to prevent beaming of the hydrogen flow to the panel. For a series of leak rates, the ultimate pressure in the chamber will be measured with an ionization gauge. Pumping speed will be calculated from:

$$S = \frac{Q}{P}$$

where Q = rate of inbleed and P = ultimate pressure. These experiments must be run first without cryogenics, to establish the pumping speed of the diffusion pumps for hydrogen. Next, similar experiments must be run to measure the pumping speeds of the cryogenic baffles. Finally, with the cryosorption panel cooled to 10°K to 20°K, the pumping speed of the panel can be found from the difference between the total speed and that of the diffusion pump and baffles. The unbaffled pumping speed of the cryosorption panel can then be found by correcting the baffled speed for the transmission coefficients of the chevron baffles.

To obtain data on the volumetric capacity of the panel, a fixed volume of hydrogen will be admitted to the chamber, with the diffusion pump blanked off. The equilibrium pressure in the chamber will be measured, thus obtaining one point on the adsorption isotherm. These experiments will be run for a series of volumes and at several panel temperatures.

The Thermodynamics Branch is concerned mainly with the engineering problems of designing an ultrahigh-vacuum facility to utilize cryosorption. The Summer Workshop was able to approach cryosorption from a more theoretical viewpoint; i.e., what is the nature of the forces between the adsorbent and the hydrogen molecule that enable cryosorption to prove useful? Further, isotherm data are available only down to 10^{-10} torr. It is hoped that theoretical justification can be provided for extrapolating these data down to 10^{-15} torr, the success of cryosorption for this application depending on its providing adequate capacity for hydrogen at ultrahigh-vacuum.

Page intentionally left blank

SOME THEORETICAL ASPECTS OF CRYOSORPTION

L. A. Veguilla-Berdecia

INTRODUCTION

When a gas or vapor is allowed to come to equilibrium with a solid surface, the concentration of gas molecules is always found to be greater in the immediate vicinity of the surface than in the free gas phase, regardless of the nature of the gas or of the surface. The process by which this surface excess is formed is termed adsorption. In any solid, the atoms at the surface are subject to unbalanced forces of attraction normal to the surface plane; the balance of forces is partially restored by the adsorption of gas molecules. If the forces responsible for adsorption are of van der Waals type, the process is called physical adsorption. Furthermore, if the solid (and its surface) are maintained at relatively low temperatures (say 50°K or below), the term cryosorption is used.

The formation of a physically adsorbed layer may be likened to the condensation of a vapor to form a liquid (or a solid, in some instances). The heat of physical adsorption is roughly of the same order of magnitude as that of liquefaction (or sublimation); physically adsorbed layers, especially those many molecular diameters thick, behave in many respects like two-dimensional liquids. Physical adsorption, being related to the process of liquefaction, occurs to an appreciable extent at pressures and temperatures close to those required for liquefaction. Thus, if P is the equilibrium pressure of the adsorbed film, and P_0 is the vapor pressure of the bulk liquid at the temperature of the experiment, it is generally found that below $P/P_0 = 0.01$ no significant adsorption takes place. There are, however, some notable exceptions, e.g., molecular sieves and adsorbents having very fine pores.

The forces causing adsorption are no different from those involved in any other interatomic or intermolecular interaction phenomenon. The singular fact that the adsorbent-surface atoms partake of two types of interactions (with other adsorbent atoms, and with absorbate molecules) makes physical adsorption particularly difficult to treat mathematically. The interactions between an atom or molecule and a solid surface are electromagnetic in origin, involving the electrons and nuclei of the system, whose states are determined from quantum mechanical considerations. When the equilibrium charge distribution is such that there is no transfer or sharing of electrons among the participating atoms and the individuality of the interacting species is maintained, the forces are of van der Waals type and can be subdivided into two main categories: (1) dispersion, and (2) dipole. Each of these categories can be handled, in principle, by the proper solution of the Schrödinger equation for the system in question. However, theoreticians have yet to arrive at this computational Valhalla, and gross approximations must be attempted.

The physical interaction between an atom or molecule and a solid surface is caused by the attractive van der Waals forces which, according to Margenau[1], may be defined simply as the forces which give rise to the constant a in the van der Waals' equation of state,

$$\left(P + \frac{an^2}{V^2}\right)(V - b) = nRT, \qquad (1)$$

and the repulsive forces which arise when atoms come close enough together to allow interpenetration of the electron clouds.

When the adsorbed atom or molecule possesses no permanent dipole or multipole moment, the attractive interaction with the solid surface is due to non-polar dispersion forces only unless the solid itself has an external electric field, as, for example, ionic solids. In order to understand the dispersion forces, consider an inert gas atom in the ground state: The charge distribution is spherically symmetrical and the atom possesses no permanent dipole or multipole moment, and thus no external field. The kinetic energy of this state is not zero, however, because of the existence of the zero-point energy, a purely quantum mechanical effect. This zero-point energy implies non-vanishing instantaneous dipole and multipole moments. The instantaneous moments induce in any neighboring atom resonant moments which are in phase with those of the first atom, and there is thus a force of attraction between the atoms. If the individual atoms possess no resultant angular momentum, and a perturbation calculation is carried out, the first-order perturbation energy is zero, whereas the second-order perturbation is given by the London formula

$$E = -\frac{C}{r^6}, \tag{2}$$

which is the dispersion energy. In equation (2), r is the distance between atoms and C is constant,

$$C = \frac{3}{2}\alpha_1 \alpha_2 \frac{h \nu_1 \nu_2}{\nu_1 + \nu_2} \tag{3}$$

where α is the polarizability, ν is the characteristic frequency, and h is Planck's constant.

Equation (2) arises from dipole-dipole interactions only. If dipole-quadrupole and quadrupole-quadrupole interactions are included, then one has

$$E = -\frac{C}{r^6} - \frac{C'}{r^8} - \frac{C''}{r^{10}} \tag{4}$$

where C' and C'' are constants, calculable, in principle, from the wave functions of the species in question. Using equation (4), Kiselev et al[2] have calculated the heats of adsorption of several gases on graphite, and have concluded that the dipole-quadrupole term contributed about 10 percent and the quadrupole-quadrupole term about 1 percent of the total value of the dispersion interaction. It is generally true that this type of contribution holds for other systems.

A problem of some importance is to find the value of C_{AB} between unlike atoms A and B, if values of C_{AA} and C_{BB} are known. A useful approximation is

$$C_{AB} = k(C_{AA}\, C_{BB})^{1/2}. \tag{5}$$

When three or more species are to be considered, say A, B, and C, the dipole-dipole constant may be evaulated by the interactions between pairs, in each instance neglecting the presence of the third species and adding the resulting values.

This principle of additivity allows the atoms of a solid to be considered as force centers. Therefore, the total interaction energy of an adsorbed atom with a solid surface can be written as

$$E_{TOTAL} = \sum_j E_j = - \sum_j \frac{C_j}{r_j^6}, \tag{6}$$

where E_j is the interaction energy between the adsorbed species and the jth atom of the solid. London and Polanyi[3] have pointed out that the dipole and multipole forces involved in calculating E are precisely the forces assumed to exist in the potential theory of adsorption. In pursuing this idea further, London[4] has changed equation (6) to an integral

$$E_{TOTAL} = N \int C r^{-6} \, dv, \tag{7}$$

where C is the same constant as in equation (2), N is the number of adsorbent atoms per unit volume, and dv is a volume element at a distance r from the adsorbed molecule. It is clear that approximating a sum (equation 6) by an integral (equation 7) can be valid only at large interatomic separations; this is clearly violated when the species considered is adsorbed (i.e., in its equilibrium position) on the surface of the solid.

It seems clear from the foregoing discussion that detailed calculations of the factors influencing physical adsorption are doomed to failure at present. Steps to remedy this situation are under way. A particularly promising approach is that of Halsey[5], who has used a theoretical approach to calculate surface areas of various adsorbents from the virial expression and potential functions of several gases. How well his theory will hold at relatively low temperatures is unknown.

Whenever detailed mechanistic processes cannot be treated mathematically to yield relationships between observables, gross or macroscopic effects can usually be handled either from a thermodynamic point of view or from a more down to earth approach, i.e., empirically. This is indeed the situation here. It has been shown that detailed atomistic considerations yield untractable equations. In view of this, it behooves us to seek thermodynamic relationships between laboratory observables in order to arrive at mathematical relationships linking the various parameters into economical (i.e., succinct) formulae. These desiderata are met by Polanyi's Potential Theory.[6]

The forces holding a molecule adsorbed on a surface have been shown to decay with distance from the surface. Therefore, a multilayer film can be thought of as lying in a potential gradient, analogous to the situation of the earth and its atmosphere. This implies the existence of equipotential surfaces and adsorption volumes. The force of attraction at any given point in the adsorbed film is conveniently measured by the adsorption potential ϵ, defined as the work done by the adsorption forces in bringing a molecule from the gas phase to that point. This is the idea conveyed by Figure B-3, a cross section of an adsorbed layer, the solid, and surrounding gas. The adsorption volumes, φ_i, are defined as the volume between the equipotential ϵ_i and the adsorbent surface. Clearly an inverse relation exists, at equilibrium, between φ_i and ϵ_i. As φ increases from zero to φ_{max}, ϵ decreases from its maximum value ϵ_0 at the surface to zero at the outermost layer. The process of building up the adsorbed film may therefore be represented by the innocuous-looking function

$$\epsilon = f(\varphi) \tag{8}$$

which is in reality a distribution function.

Table B-1
Distinct Temperature Ranges for
Potential-Theory Calculations

Case	Temperature	State of adsorbed film
I	$T \ll T_c$	liquid
II	$T \sim T_c$	liquid & compressed gas
III	$T > T_c$	compressed gas

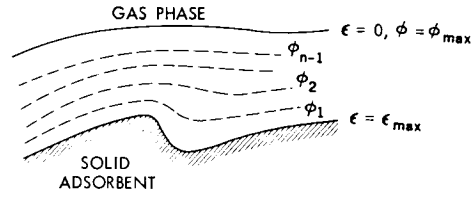


Figure B-3–Cross Section of Gas – Solid Interface

The adsorption potential is postulated to be independent of temperature, so that $\epsilon = f(\varphi)$ should be the same for a given gas at all temperatures; hence, its name "characteristic curve." If the characteristic curve is found for one temperature, then it can be used for all other temperatures, and a very nice extrapolating procedure is obtained without having to perform a whole series of experiments at each value of the temperature. We see then that the function $\epsilon = f(\varphi)$ replaces the usual isotherm or P vs V plot. In practice, however, a P vs V plot is obtained, and from it a $\epsilon = f(\varphi)$ "master curve" is calculated. The procedure by which a P vs V plot is changed into a characteristic curve depends on the range of temperatures to be covered by the experimenter. Polanyi[7] recognized three different cases, given below.

The easiest case to treat is Case I. As a first approximation, it is assumed that the vapor behaves ideally, that the liquid film is incompressible, and that no work is expended in forming a liquid surface. Then the work done in compressing one mole of vapor from the equilibrium pressure P to the vapor pressure P_0 is simply

$$\epsilon_i = RT \ln\left(\frac{P_0}{P}\right). \tag{9}$$

Therefore, for each point on the P vs V plot, one obtains a value of ϵ_i through equation (9) and a value of φ through

$$\varphi = kV \tag{10}$$

where k is a constant to be determined. It has been found by Dubinin et al[8], and by Hobson[9], that for many gases the characteristic curve has a Gaussian form, i.e.,

$$\varphi = \varphi_0 e^{-k\epsilon^2}, \; k, \; \varphi_0 = \text{constants} \tag{11}$$

so that a plot of

$$\ln \varphi \text{ vs } \left[\ln\left(\frac{P_0}{P}\right)\right]^2$$

yields a straight line. There is no theoretical justification yet for this interesting dependence. The fact remains, however, that this Dubinin-Raduskevitch plot, as equation (11) is called, fits a wealth of experimental data.

The method of calculating $\epsilon = f(\varphi)$ for Cases II and III has been published by Brunauer.[10]

APPLICATION OF POTENTIAL THEORY TO CRYOSORPTION

An interesting application of the potential theory to the problem of cryosorption at very low pressures (below 10^{-6} torr) has been put forward by Manes and Grant.[15] These workers content themselves with the calculation of characteristic curves for $T < T_c$ for N_2 on PCB and BPL carbons. They claim that the characteristic curves of H_2 and He and other gases can be obtained from the nitrogen curves by multiplication of the adsorption energies by a proper "affinity ratio." This idea dates back to the work of Polanyi,[7] and Berenyi[11] who tried to use a master curve for all gases (on the same adsorbent) by inclusion of the ratio of the respective van der Waals constants. This simply means using

$$\epsilon = \beta f(\varphi) \tag{12}$$

instead of (8), and β is the so-called affinity factor. It is not surprising that the results of Manes and Grant are not very good for $T > T_c$, since no corrections were made for this.

If the Polanyi potential theory is to be used in determining the capacities of regenerative cryosorption pumping systems, then the method of calculation must show in detail how each of the ranges in temperature is handled. In particular, consider the case of hydrogen cryosorption on molecular sieves. If the raw data are obtained at 50°K, then a different type of calculation must be used in order to evaluate $\epsilon = f(\varphi)$, from that used if the raw data is obtained at 20°K. This is evident because $T_c = 33$°K for H_2. The decision as to whether a particular system is able to show sufficient capacity for cryosorption (i.e., to extend or maintain present limits of attainable vacuum) must be made on the basis of the most critical evaluation of the data.

In this report we wish to do precisely that. The system to be considered is the cryosorption of H_2 on Linde molecular-sieve 5A, an aluminosilicate. Details of this system have been published in several AEDC reports.[12] We begin by establishing some of the properties of H_2 gas. Tables B-2 and B-3 give the most recent vapor pressure of H_2 as a function of temperature.

Table B-2
Vapor Pressure of H_2 from the Triple Point to the Critical Temperature

Temperature Range (in degrees Kelvin)	Pressure (in atmospheres)
13.8°K (triple point) to 20.27°K	$\log_{10} P = 1.772454 - \frac{44.36888}{T} + .02055468\ T$
20.27°K to 29°K	$\log_{10} P = \log_{10} P_x = 2.00062 - \frac{50.09708}{T + 1.0044} + .01748495\ T$
29.0°K to 32.976°K (critical point)	$P = P_x + 1.317 \times 10^{-3}\ (T-29)^3 - 5.926 \times 10^{-5}\ (T-29)^5 + 3.913 \times 10^{-6}\ (T-29)^7$

Table B-3
Vapor Pressure of H_2, Calculated from Equations of Table B-2

T(°K)	P(atm)
13.8	.06946
14.0	.07781
15.0	0.1327
16.0	0.2103
17.0	0.3251
18.0	0.4759
19.0	0.6726
20.0	0.9228
20.27	1.000
21.0	1.233
22.0	1.613
23.0	2.069
24.0	2.610
25.0	3.245
26.0	3.982
27.0	4.829
28.0	5.794
29.0	6.886
30.0	8.117
31.0	9.501
32.0	11.044
32.976	12.700

The truncated virial expansion for H_2 (para) from 2 to 100 atm, and from the triple point to 100°K, can be written as

$$PV = RT\left\{1 + \frac{B(T)}{V} + \frac{C(T)}{V^2}\right\} \qquad (13)$$

where

$$B(T) \equiv \sum_{i=1}^{4} B_i \left(\frac{T_0}{T}\right)^{\frac{2i-1}{4}},$$

$$C(T) \equiv C_0 x^{-\frac{1}{2}}(1 + cx^{-3})\left[1 - e^{(1-x^3)}\right],$$

$T_0 = 109.7806$ °K,

$B_1 = 42.46385$ cm³/gm-mole,

$B_2 = -37.11724$ cm³/gm-mole,

$B_3 = -2.298187$ cm³/gm-mole,

$B_4 = -3.048415$ cm³/gm-mole,

$C_0 = 1310.5$ (cm³/g-mole)²

$c = 2.1486$, and

$x = T/20.615$°K.

This expression has been found to be valid for p-H_2 vapor and gas at densities below 0.007 g-mole/cm³.[13]

Table B-4 contains adsorption data for H_2 on molecular-sieve 5A at 20°K.

Table B-4
20°K Adsorption Isotherm of H_2 on Molecular Sieve 5A*

P(torr)	V (cm³ STP/gram adsorbent)
10^{-9}	2.6
10^{-8}	9.8
10^{-7}	28.
10^{-6}	63.
10^{-5}	100.
10^{-4}	138.

*(See also Figure B-4.)

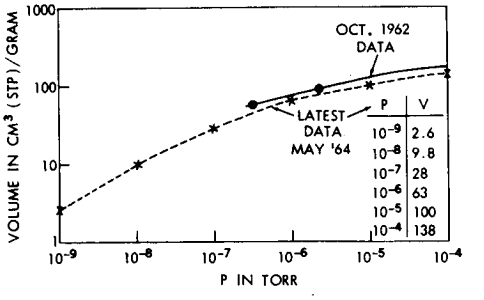


Figure B-4-Adsorption of H_2 on Molecular Sieve 5A (20.2°K)

Since 20°K is well below the critical point of 33°K, we can use a straightforward identification

$$\epsilon \Rightarrow -RT \ln\left(\frac{P}{P_0}\right) , \tag{14}$$

and

$$\varphi \Rightarrow \frac{2.016}{1.568} \times 10^{-3} \text{ V.} \tag{15}$$

Table B-5 shows the calculated values of ϵ, ϵ^2, and ϵ^4.

Table B-5
Calculated Values of Adsorption Energy
as a Function of Pressure

P (torr)	ϵ	ϵ^2	ϵ^4
10^{-9}	-470.49	2.2136×10^5	4.88×10^{10}
10^{-8}	-430.89	1.8566×10^5	3.45×10^{10}
10^{-7}	-391.29	1.5311×10^5	2.34×10^{10}
10^{-6}	-351.69	1.2369×10^5	1.51×10^{10}
10^{-5}	-312.09	9.7400×10^4	$.95 \times 10^{10}$
10^{-4}	-272.49	7.4251×10^4	$.55 \times 10^{10}$

The Linde data is best represented by the empirical formula

$$\log_{10} V = 2.35 - 3.98 \times 10^{-11} \epsilon^4 \tag{16}$$

so that the expected capacity of the molecular sieve 5A at $P = 10^{-15}$ torr is

$$V = 2.2 \times 10^{-8} \text{ cm}^3 \text{ STP/gram.}$$

Figure B-5 shows the plotted data of $\log_{10} V$ vs ϵ^4. The characteristic function would then be

$$\varphi = \varphi_0 e^{-k\epsilon^4}, \tag{17}$$

a rather severe type of dependence. Partly because of disbelief in such strong dependence on ϵ, and because of the unusually low capacity predicted at $P = 10^{-15}$ torr, the experimental setup used by Linde (Figure B-6) was re-examined in great detail. One can see that the chamber holding the adsorbent panel has a set of baffles held at some undetermined temperature (but certainly above 77°K), that the cylindrical walls are held at 77°K by LN_2 (liquid nitrogen), and that only the adsorbent is maintained at (or near) 20°K.

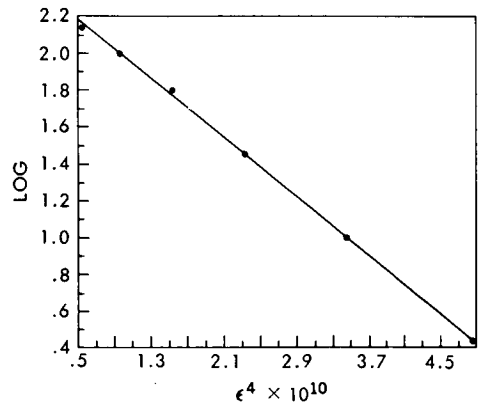


Figure B-5–Linde Data, Adsorption of H_2 on Sieve 5A (log V vs ϵ^4)

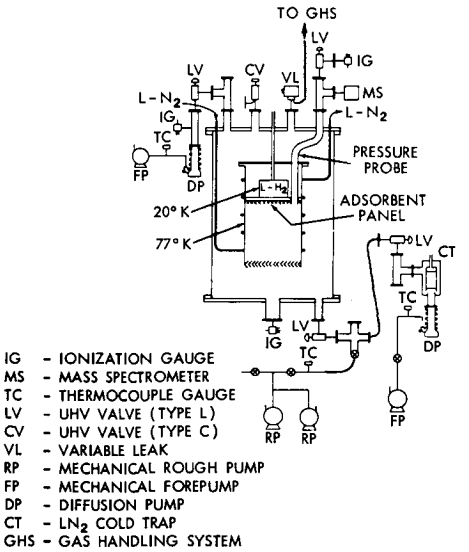


Figure B-6–Diagram of Apparatus for Testing Cryosorption Panels

This unusual experimental setup for obtaining cryosorption data at an alleged 20°K will certainly lead to low values of adsorbed volumes of H_2. This means that the molecular sieve-5A will be an even more efficient material for the physical adsorption of H_2 than is indicated by Table B-4.

Since the H_2 gas above the adsorbent panel is certainly at a temperature higher than 20°K, it might be instructive to try to correct the data in Table B-4 to cover this contingency. In the following analysis, an arbitrary temperature of 40°K was used for the H_2 gas above the panel. It is recognized, however, that the most likely situation for this setup is one where a temperature gradient exists in the chamber. Since 40°K is greater than the critical temperature of H_2 (33°K), the adsorbed hydrogen can be considered to be in a "compressed gas" state, i.e., Case III. This is not necessarily true, however, if coverage is below the monomolecular-layer capacity, unless the adsorption is sufficiently heterogeneous for "patch" adsorption or capillary condensation to have taken place. In these latter cases, it would be expected that the volumes of H_2 (STP) adsorbed at relatively high pressures (10^{-4} and 10^{-5} torr) would experience the biggest corrections. In trying to apply these corrections, we employ a modified Berenyi method of calculation which we outline in detail below. Assume a P-V isotherm for hydrogen at $T > T_c$. Berenyi suggests using

$$\varphi = \frac{22400 \times b}{M}, \tag{18}$$

and

$$\epsilon = RT \ln\left(\frac{0.14T}{Pb}\right) \tag{19}$$

where

$x \equiv$ grams of gas adsorbed at a pressure P/gm adsorbent,

$M \equiv$ molecular weight of gas, and

$b \equiv$ van der Waals constant for the gas. Now we wish to calculate a function

$$\epsilon = f(\delta) \tag{20}$$

where δ is the density of the adsorbed gas. In order to do this, we must perform the integration

$$\epsilon = \int_{V_1}^{V_2} V dP, \quad V_1 \equiv \frac{RT}{P} \tag{21}$$

and $V_2 \equiv M/\delta$. This gives a function $\epsilon = f(\delta)$ for each value of P. If there are 10 data points for the P-V isotherm, there are 10 distinct $\epsilon = f(\delta)$ curves. The analytical form of $\epsilon = f(\delta)$ will depend on the $V = V(P)$ dependence used in the integrand. Berenyi suggests using either a numerical integration or the van der Waals equation of state. Instead, we are going to use the truncated virial, equation (13). Therefore

$$P = RT\left\{\frac{1}{V} + \frac{B(T)}{V^2} + \frac{C(T)}{V^3}\right\} \tag{22}$$

$$VdP = -RT\left\{\frac{1}{V} + \frac{2B(T)}{V^2} + \frac{3C(T)}{V^3}\right\}dV \tag{23}$$

$$\epsilon = -RT\left\{\ln\left(\frac{M}{\delta}\right) - \frac{2B(T)}{M}\,\delta\right\} + \text{const.} \tag{24}$$

In equation (24) the term $C(T)$ has been neglected because it can be shown that, if $T \geq 40°K$ $C(T)$ does not contribute significantly, the constant in equation (24) is

$$\text{const.} = RT\,\ln\left(\frac{RT}{P}\right) - 2B(T)P. \tag{25}$$

Once these curves are obtained, we choose an arbitrary energy value, say ϵ_i, and obtain a value of φ_i corresponding to it from the graph of $\epsilon = f(\varphi)$; we also obtain a value of δ_i, corresponding to that ϵ_i, from equation (24).

We thus get a $\varphi = f(\delta)$ curve by repeating the process of choosing a range of values of ϵ_i. Once $\varphi = f(\delta)$ is obtained for each $\epsilon = f(\delta)$, one calculates

$$x_{new} = \int_0^{\varphi_{max}} \delta\, d\varphi. \tag{26}$$

The corrected φ now becomes

$$\varphi_{corrected} = \varphi_{uncorrected}\,\frac{x_{new}}{x_{old}} \tag{27}$$

and the entire procedure is repeated until $\varphi_{corr} = \varphi_{uncorrected}$. This last set of φ's is plotted vs ϵ to obtain the characteristic curve. We will now apply this procedure to the Linde data (Table B-4). Using

$$R = 82.057\ \frac{\text{atm-cm}^3}{\text{gm-mole-}^\circ\text{K}}, \qquad M = 2.01572\ \frac{\text{gm}}{\text{mole}},$$

$T = 40°K$, and δ in gm-mole/liter, we obtain Table B-6. Upon calculation of $\epsilon = f(\delta)$, we noticed that ϵ_{max} for this function was lower than the values of ϵ on Table B-6. This

Table B-6
φ vs ϵ, Utilizing Linde Data

$\varphi(cm^3)$	P_{atm}	$-\epsilon$
69.1	1.31579×10^{-12}	8.558×10^4
260.4	1.31579×10^{-11}	7.802×10^4
744.1	1.31579×10^{-10}	7.046×10^4
1674.2	1.31579×10^{-9}	6.290×10^4
2657.5	1.31579×10^{-8}	5.535×10^4
3667.3	1.31579×10^{-7}	4.779×10^4

meant that no $\delta = f(\varphi)$ could be calculated. In order to circumvent this difficulty, consider the function $\epsilon = f(\delta)$ given by (24):

$$\epsilon = -RT\left\{\ln\left(\frac{M}{\delta}\right) - \frac{2B(T)}{M}\,\delta\right\} + \text{const.} \tag{24}$$

$$\frac{d\epsilon}{d\delta} = -RT\left\{-\frac{1}{\delta} - \frac{2B(T)}{M}\right\} = 0,$$

giving

$$\delta = -\frac{M}{2B(T)}. \tag{28}$$

Note that $B(T) < 0$. Therefore, ϵ_{max} for $\epsilon = f(\delta)$ becomes

$$\epsilon(\text{max}) = RT \ln\left\{\frac{RT}{-2PB(T)\,e}\right\} - 2B(T)P \tag{29}$$

$$e = 2.718\cdots .$$

Let us change Berenyi's

$$\epsilon = RT \ln\left(\frac{0.14T}{PD}\right)$$

to

$$\epsilon = RT \ln \frac{aT}{P} \tag{30}$$

and find the value of $\underline{a}$ in (30) for which

$$RT \ln \frac{aT}{P_2} \leq RT \ln \left\{ \frac{RT}{-2B(T)\, e} \right\} - 2B(T)P_1 \tag{31}$$

where $P_2 = 10^{-9}$ torr

$P_1 = 10^{-4}$ torr.

Therefore

$$a \leq \left(\frac{P_2}{P_1} \right) \frac{R}{-2B(T)\, e} = .299 \times 10^{-5} \tag{32}$$

Using the modified Berenyi expression

$$\epsilon = RT \ln \frac{.299 \times 10^{-5}\, T}{P} \tag{33}$$

one can now obtain $\varphi = f(\delta)$ curves for the Linde data. Tables B-7 to B-12 give the results of this calculation.

Table B-7
δ vs ϵ at $P = 10^{-4}$,
Using Modified Berenyi Expression

δ	ϵ	
10^{-7}	1.600×10^{3}	$P = 10^{-4}$
10^{-6}	9.158×10^{3}	
10^{-5}	1.671×10^{4}	
10^{-4}	2.426×10^{4}	
10^{-3}	3.167×10^{4}	
10^{-2}	3.775×10^{4}	

Table B-8
δ vs ϵ at $P = 10^{-5}$,
Using Modified Berenyi Expression

δ	ϵ	
10^{-8}	1.600×10^{3}	$P = 10^{-5}$
10^{-7}	9.158×10^{3}	
10^{-6}	1.672×10^{4}	
10^{-5}	2.427×10^{4}	
10^{-4}	3.181×10^{4}	
10^{-3}	3.922×10^{4}	

Table B-9
δ vs ϵ at $P = 10^{-6}$,
Using Modified Berenyi Expression

δ	ϵ	
10^{-9}	1.600×10^{3}	$P = 10^{-6}$
10^{-8}	9.158×10^{3}	
10^{-7}	1.672×10^{4}	
10^{-6}	2.427×10^{4}	
10^{-5}	3.183×10^{4}	
10^{-4}	3.937×10^{4}	

Table B-10
δ vs ϵ at $P = 10^{-7}$,
Using Modified Berenyi Expression

δ	ϵ	
10^{-10}	1.600×10^{3}	$P = 10^{-7}$
10^{-9}	9.158×10^{3}	
10^{-8}	1.672×10^{4}	
10^{-7}	2.427×10^{4}	
10^{-6}	3.183×10^{4}	
10^{-5}	3.939×10^{4}	

Table B-11
δ vs ϵ at $P = 10^{-8}$,
Using Modified Berenyi Expression

δ	ϵ	
10^{-11}	1.600×10^3	$P = 10^{-8}$
10^{-10}	9.158×10^3	
10^{-9}	1.672×10^4	
10^{-8}	2.427×10^4	
10^{-7}	3.183×10^4	
10^{-6}	3.939×10^4	

Table B-12
δ vs ϵ at $P = 10^{-9}$,
Using Modified Berenyi Expression

δ	ϵ	
10^{-12}	1.600×10^3	$P = 10^{-9}$
10^{-11}	9.158×10^3	
10^{-10}	1.672×10^4	
10^{-9}	2.427×10^4	
10^{-8}	3.183×10^4	
10^{-7}	3.939×10^4	

The dependence of ϵ on φ can now be exhibited for comparison. We recall that we are using the modified Berenyi expression

$$\epsilon = RT \ln \frac{.299 \times 10^{-5} T}{P}.$$

Table B-13
Energy of Adsorption vs Adsorption Volume
for Modified Berenyi Expression

$\varphi (cm^3)$	ϵ
69.1	3.838×10^4
260.4	3.082×10^4
744.1	2.326×10^4
1674.2	1.570×10^4
2657.5	0.8145×10^4
3667.3	0.0587×10^4

Figure B-7 is a graph of $\epsilon = f(\varphi)$, and Figure B-8 is the $\epsilon = f(\delta)$ graph for $P = 10^{-9}$. Now we pick ϵ_i and select one value of δ_i and another of φ_i. Using some foresight, we pick the ϵ_i's so that $\Delta\varphi_i = 600$; the reason for this is that we want to use Simpson's rule for the numerical integration of equation 26,

$$x_{new} = \int_0^{\varphi_{max}} \delta \, d\varphi. \qquad (26)$$

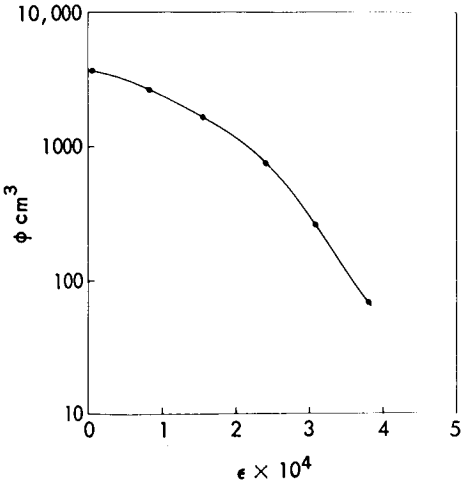


Figure B-7–Adsorption of H_2 on Sieve 5A (40°K)

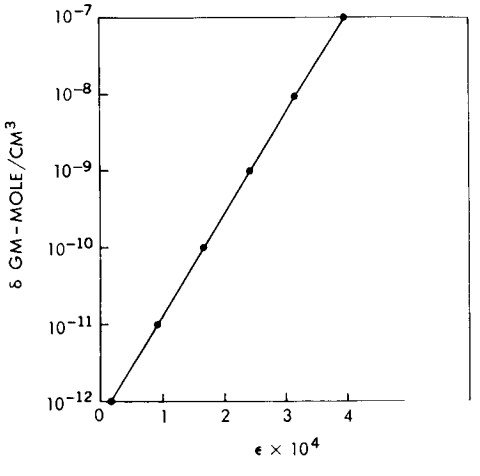


Figure B-8–Theoretical $\epsilon = f(\delta)$ Curve for H_2 Adsorption (40°K)

Table B-14
$\epsilon_i - \varphi_i - \delta_i$ Values for $P = 10^{-9}$ and $\Delta\varphi = 600$

φ_i	ϵ_i	δ_i
69.1	3.838×10^4	10^{-7}
669.1	2.50×10^4	1.3×10^{-9}
1269.1	1.92×10^4	1.9×10^{-10}
1869.1	1.40×10^4	4.3×10^{-11}
2469.1	1.00×10^4	1.3×10^{-12}
3069.1	0.62×10^4	4×10^{-12}
3669.1	0.059×10^4	7×10^{-13}

Simpson's rule gives us

$$x = \frac{600}{3}\left\{10^{-7} + 4\left[1.3 \times 10^{-9} + 4.3 \times 10^{-11} + 4 \times 10^{-12}\right]\right.$$

$$\left. + 2\left[1.3 \times 10^{-11} + 1.9 \times 10^{-10}\right] + 7 \times 10^{-13}\right\} = 2.02 \times 10^{-5}.$$

We notice that the only contributing terms in the Simpson's rule expansion are y_0 and y_1,[14] the first two terms. From a glance at Tables B-7 to B-12, we notice further that this will be the case for all pressures ($P = 10^{-4}$ to $P = 10^{-9}$). The new value of φ becomes

$$\varphi = 69.1 \times \frac{2.02}{2.34} \times 10^{-1} = 5.91 \text{ cm}^3 \quad (P = 10^{-9}).$$

For $P = 10^{-8}$ we obtain a table of values of $\epsilon_i - \delta_i - \varphi_i$ almost identical to Table B-14, except that the powers of 10 of δ_i are changed by one and the energies ϵ change by a small amount; i.e.,

$$\varphi_i = 69.1 \quad \epsilon_i = 3.838 \times 10^4 \quad \delta_i = 10^{-7}, \quad P = 10^{-9}$$

becomes

$$\varphi_i = 69.1 \quad \epsilon_i = 3.939 \times 10^4 \quad \delta_i = 10^{-6}, \quad P = 10^{-8}$$

so that a single integration suffices to calculate all the new φ's. This means that

$$x_{new} \sim 2 \times 10^{-4} \quad P = 10^{-8}$$

and the new φ

$$\varphi = 260.4 \times \frac{2.00}{2.34} = 222.$$

The next value of φ for $P = 10^{-7}$ will be greater than the old value because the ratio x_{new}/x_{old} is larger than 1. This effect is precisely opposite to that expected. Therefore, the new dependence of $\varphi_{corrected}$ vs ϵ will be more drastic than that exhibited by equation (17). This will mean even lower adsorption at $P = 10^{-15}$ than the value of 2.2×10^{-8}, using straightforward extrapolation of the uncorrected Linde data (Table B-5).

We conclude, therefore, that the corrections implied in the Berenyi calculations are not applicable to this system. This might have been expected on the basis of the low adsorption volumes (and hence low coverages) observed at low temperatures. The choice of $T = 40°K$ for the effective gas temperature is not the culprit here in vitiating the results. It can be shown that, unless $T < 20°K$, essentially the same results are obtained in the corrections. We conclude, therefore, that a value of 2.2×10^{-8} cm^3 STP/gram of H_2 be used for the expected adsorption of hydrogen at $P = 10^{-15}$ torr. Furthermore, we suggest obtaining more creditable tables on H_2 adsorption on molecular sieves; for example, the system used by Linde should be at 20°K throughout. There are indications that the adsorption volumes for such a system will be significantly greater than the values quoted in Table B-5.

REFERENCES

1. Margenau, H., Rev. Mod. Phys. 11, 1 (1939)
2. Kiselev, A.V., et al., Izvest. Ahad. Nauk. S.S.S.R., Otdel. Khim. Nauk. 1314 (1957)
3. London, F., and Polanyi, M., Naturwiss. 18, 1099 (1930)
4. London, F., Z. Phys. Chem. 11, 222 (1930)
5. Halsey, G.D., Jr., Amer. Vac. Soc. Vol. I, 119, 1961 Transactions, of Eighth Vac. Symposium, Pergamon Press, N. Y. (1961)
6. Polanyi, M., Verh. Deutsch. Phys. Geo. 16, 1012 (1914)
7. Polanyi, M., Verh. Deutsch. Phys. Geo. 18, 55 (1916)
8. Dubinin, M.M., et al., C. R. Acad. Sci. U.S.S.R. 54, 701 (1946); 55, 137, 327 (1947)
9. Hobson, J.P., J. Phys. Chem. 67, 2000 (1963)
10. Brunauer, S., "The Adsorption of Gases and Vapors," Clarendon Press, Oxford, and Princeton University Press, Princeton (1945)
11. Berenyi, L., Z. Phys. Chem. 105, 55 (1923)
12. Hemstreet, R.A., et al., "The Cryosorption Pumping of Hydrogen at 20°K," AEDC — TDR-64-100 (May 1964) Linde Division of Union Carbide Corporation, Tonawanda, New York
13. "Advances in Cryogenic Engineering," Vol. 9, Plenum Press, New York (1964) page 234
14. Margenau, M., and Murphy, G.M., "The Mathematics of Physics and Chemistry," D. Van Nostrand Co., Inc., New York (1956) pages 477-478
15. Manes, M., and Grant, R.J., "Calculation Methods for the Design of Regenerative Cryosorption Pumping Systems," paper presented at 1963 Symposium of the American Vacuum Society (October 17, 1963)

Reject

LITERATURE SURVEY

J. B. Crooks

A search was conducted of recent literature pertaining to ultrahigh-vacuum measurement, cryopumping, and cryosorption. The appended bibliography with abstracts represents the results of this effort.

Some indication of the scope of the search is given in the following outline:

I. Vacuum Measurement

- A. General reviews
- B. Contributions to the theory behind the operation of a vacuum measuring device – why it works, and what it measures
- C. Calibration
  1. Satellite and spacecraft measurement of neutral gas density (for comparison with satellite drag calculations, or comparison of simultaneous readings of gauges aboard spacecraft)
  2. Laboratory comparison of, and interaction between, vacuum gauges or mass spectrometers
  3. Novel methods of vacuum-gauge calibration
  4. Reviews on subject

II. Cryosorption – Cryopumping

- A. General theory
- B. Adsorption at low temperature and pressure; both theory and experimental measurements (cryopumping or cryosorption not necessarily the immediate concern of the article)
- C. Investigation of phenomena occurring in cryosorption, cryopumping, or cryotrapping
- D. Works of developmental nature – improvements for cryogenic-vacuum pumping systems (new adsorbents, improved configuration of condensing surfaces, etc.)

III. Useful bibliographies – e.g. heat transfer, cryogenics, satellite measurement of upper atmosphere, etc.

The abstracts that have been checked so far are:

Technical Abstract Bulletin (Report generated by U. S. Defense Dept. and some foreign military reports)	June 15, 1964 to November 15, 1963 June 15, 1964 to January 1, 1964 (white section only)
International Aerospace Abstract (covers open literature of interest to field)	July 1, 1964 to January 1963
Sci. and Technical. Aerospace Reports (U. S. Govt. and some foreign government reports abstracted)	June 8, 1964 to January 8, 1963
Vacuum (abstracts)	May 1964 to June 1963

Although the period June 1964 – January 1963 has not been completely searched, approximately 250 articles and reports were found to have some pertinency to the fields previously outlined. Only those references judged to be most significant are included in the bibliography. The broad scope of the search itself, together with the numerous articles encompassed, made progress fairly slow.

The bibliography has been divided into three sections: A (Cryosorption), B (Vacuum Instrumentation), and C (Miscellaneous). Under each section, items are arranged first by year, then by author. Reports are referenced either by their STAR "N numbers" or their ASTIA accession (AD) numbers, as this seemed to be the most convenient uniform method of identification, and is sufficient to obtain the report from either NASA or the Defense Documentation Center.

The abstract is either the author's original, or is taken from one of the abstract journals previously mentioned.

A. CRYOSORPTION

A621 Narcisi, R. S. et al., Zeolite adsorption pump for rocket-borne mass spectrometer. Trans. 9th National Vacuum Symposium (1962) 232-236

A631 Anon., Sorption pumps (molecular sieve). Electronic Equip. News, 5, p24 (November 1963)

A632 Anon., Basic mechanisms of noncondensable gas cryotrapping. STAR, N63-18734 (1963). 66 p.

Systematic measurements are reported on the rate at which argon and nitrogen are trapped by water as it condenses under vacuum onto a 77°K surface. For condensation rates between 10^{15} and 10^{18} molecules per cm^2-sec and noncondensable partial pressures between 10^{-6} and 10^{-1} torr, trapping is proportional to condensation rate and concentration of adsorbed gas. The latter are described by Langmuir isotherms with characteristic pressures for saturation of 2.5×10^{-3} torr and $3. \times 10^{-3}$ torr for nitrogen and argon respectively. Independent adsorption measurements show that the quantity of gas adsorbed is proportional to mass of condensate (~ 600 $meters^2$ per gram) and depends on pressure.

"Apparent trapping" due to adsorption on surface area generated during condensation is only about 7 percent of actual trapping. All the results are consistent with a mixed phase for the condensate, of which approximately 10 percent is the 12 A gas hydrate. Rates for the trapping of nitrogen and helium by water and nitrogen by 1 nitropropane were below the limits of detection for the equipment in use.

A633 Bliven, C. M. and Florio, J. V., Study of adsorption of gases in the high vacuum range. STAR, N63-22199 (1963)

Experimental data on the adsorption of nitrogen on tungsten have been obtained. These data have been analyzed, and the sticking probability has been obtained from consideration of the mass balance relationship in the reaction chamber. The chemical pumping speeds of the omegatron filament and oxide cathode assembly, due to reaction with the incoming nitrogen, were estimated. The sticking probability and number of adsorbed molecules of oxygen on the molybdenum anode were calculated from the pressure - time characteristics.

A634 Degras, D. A., Adsorption and desorption phenomena and their role in particle accelerators and space simulators. 2nd European Vac. Symposium 5-7 June, 1963, 54-67, Rudolph A. Lang Verlag, Esch (1963)

A635 Degras, D. A., Adsorption of binary mixtures of gases at very low temperatures. 2nd European Vac. Symposium, 5-7 June, 1963, 95-107 (see A634)

It has been observed that the physical adsorption of condensable gases on cold surfaces is accompanied by the adsorption of other gases, which under the same circumstances can be adsorbed by themselves only with great difficulty. Results of earlier experiments, carried out at 78°K, are presented; the gas desorption of a Hopkin's condenser cooled down to 78°K, was measured, during its natural warming up by means of an omegatron. By this method the amount of nitrogen adsorbed at the low temperature can be evaluated. The phenomenon has been investigated recently by Schmidlin, Heflinger and Garwin. Other experiments were made at very low temperature with another binary gas mixture in a dynamic system. The results so far lead to desorption characteristics of great interest.

A636 Degras, D. A., The adsorption of a binary gaseous system at low temperature. 2nd Intern. Symp. residual gases in electron tubes, Milan (March 1963) Nuovo Cimento (in press)

A637 Grant, H. L., and Davey, J. E., Multistage sorption pump. Rev. Sci. Instrum., 34, No. 5, 587-588 (1963)

A638 Hemstreet, R. A., et al., Research study of the cryotrapping of helium and hydrogen during 20°K condensation of gases. Phase I and Phase II. STAR, N63-16547, (1963) 109 p.

A study of the variables affecting cryotrapping of helium and hydrogen during 20°K condensation of oxygen and nitrogen is presented. The mechanism of helium trapping appears to involve burying helium atoms in solid oxygen or nitrogen. The process is most efficient at high condensation rates, but is not likely to find application for the removal of helium in a large space chamber. Experimental results indicate that hydrogen trapping occurs by adsorption of hydrogen molecules on the solid oxygen or nitrogen, and is most efficient at low condensation rates and low heat flux to the solid surface. These conditions indicate that hydrogen trapping may be applicable for hydrogen removal in a large space chamber. The trapping of hydrogen by nitrogen condensed on a 20°K surface has been

studied under conditions where the surface was shielded from ambient temperature radiation. The process is shown to be an inefficient means for hydrogen pumping, and is likely to have value only as a bonus in the operation of a large space chamber. Hydrogen trapping by methane, some simple fluorocarbons, and carbon dioxide condensed on a 20°K surface not shielded from ambient-temperature radiation has been studied. $C Cl_2 F_2$ and carbon dioxide show rather large capacities for hydrogen, and may be useful for the pumping of hydrogen in a space chamber and in special applications.

A639 Hobson, J. P., and Armstrong, R. A., J. Phys. Chem. 67, 2000-2008 (1963)

Ultrahigh-vacuum techniques have been used in the measurement of the adsorption isotherms of nitrogen and argon on Pyrex glass in the range of pressures 10^{-13} to 10^{-6} torr. Temperatures between 63.3°K and 90°K were used. The Dubinin-Radushkevich isotherm equation was found empirically to describe all the data. This equation was found empirically to describe the isotherm data of others at low relative coverage on a variety of heterogeneous adsorbents. The energy distribution function giving rise to the Dubinin-Radushkevich isotherm is discussed. A simple analytical hypothesis is suggested for the extrapolation of isotherms on heterogeneous surfaces to the region of Henry's law.

A6310 Hoch, H., Studies of sorption phenomena in ultrahigh-vacuum systems. 2nd European Vacuum Symp., 5-7 June, 1963 (see A634)

A6311 Kindal, S. M., A study of vacuum cryosorption by activated charcoal. STAR, N63-20590 (1963) 41 p.

An experimental program was carried out in the 10^{-8} to 10^{-2} torr region to supply basic information on the adsorption of noncondensable gases at 77°K and to establish the functional dependence of adsorption rate on pressure and amount of gas adsorbed. From the experimental results, analytical expressions were developed which describe the behavior of adsorption rate in terms of adsorption parameters for high-vacuum, low-temperature conditions. These expressions are shown to agree with the observed pressure-time variations resulting from charcoal adsorption, provided the charcoal remains well below saturation.

A6312 Read, P. L., Sorption pumping at high and ultrahigh-vacua. Vacuum, 13, No. 7, 271-275 (1963)

It is demonstrated that the operating range of a liquid-nitrogen chilled synthetic zeolite pump can be extended into the high and ultrahigh-vacuum regions, without sacrificing system cleanliness. Pressures as low as 5×10^{-11} torr were produced in a small system by sorption pumping with liquid nitrogen and chilled Linde 13-X molecular sieve (a synthetic zeolite) after pre-evacuation of the system by a small oil-diffusion pump. The design of this new type of sorption pump, its operating characteristics, and its application to, ion-getter-pumped vacuum systems are briefly discussed.

A6313 Windsor, E. E., Sorption pumping using zeolites. 2nd European Vacuum Symp. 5-7 June, 1963, 278-283 (see A634)

Sorption pumps are usually used as fore pumps to ion pumps. An examination has been made using zeolite under various conditions of activation and refrigeration, and it is shown that pressures of 10^{-9} torr can be obtained using zeolite alone. An examination made by a simple radio-frequency mass spectrometer indicates that the residual gases are largely the rare gases. Problems associated with the design and use of zeolite pumps are considered and a radical system described.

A641 Bliven, C. M., Florio, J. V., and Polyanyi, T. G., Study of adsorption of gases on solid surfaces in the high vacuum range. STAR, N64-18187 (1964) 21 p.

A642 Budnikov, P. P., and Petrovykh, I. M., Zeolites – molecular sieves. ASTIA, AD-426-902 (1964)

B. VACUUM INSTRUMENTATION (Gauges and Mass Spectrometers)

B621 Damoth, D. C., and Burgess, R. G,, Residual vacuum analysis with time-of-flight mass spectrometer. Trans. 9th Nat. Vac. Symp., 1962, p. 418-420

B622 Farrah, H. R., et al., Study of instrumentation for a space simulation chamber. STAR, N62-14912 (1962)

The purpose of this report is to document the salient information realized through a study of the instrumentation and data collection of pertinent parameters in a large ultrahigh-vacuum aerospace chamber. Existing techniques for the measurement of cryogenic temperature, residual gas pressure, solar and planetary radiation, composition and density were evaluated. Methods of modifying, designing, or developing sensing devices to make them suitable for use in an aerospace chamber were studied. Although the report emphasizes the evaluation of instruments under the particular environmental constraints imposed by the proposed aerospace chamber, the basic information reported is also pertinent to other aerospace or vacuum chambers.

B631 Anon., Calibration of ultrahigh-vacuum gauges. New Scientiest 19, No. 350, p. 248 (August 1963)

A pump is used to draw air through a series of very accurately machined holes, and the gauge to be calibrated is coupled between the last hole and the pump. A McLeod vacuum gauge is coupled further up the line of holes when the pressure is higher. The height of the mercury column in this reference gauge is read with a cathetometer. It is claimed by the National Research Corporation of Cambridge, Massachusetts, who offer this service, that the system can calibrate gauges on an absolute basis down to 10^{-9} torr.

B632 Bailey, J. R., Residual gas spectra in high and very high vacuum systems. 2nd Intern. Symp. Residual gases in electron tubes. Milan, March 1963. Nuovo Cimento (in press)

The mass spectrum of gas in vacuum systems usually indicates the presence of several contributing sources. To help in identifying these, a number of gases, solvents and pump oils have been individually examined. For this purpose a 2-inch-radius commercial mass spectrometer having a resolution of about 50, mass range 2-100, and ionizing electron-beam energy of 20 v was used. The results obtained help interpret results obtained with other mass spectrometers of the same type. They are also found to be qualitatively applicable to omegatron results when the same electron-beam energy is used.

B633 Günther, K. G., 2nd European Vac. Symp., 5-7 June, 1963 p. 21-41 Rudolph A. Lang Verlag, Esch (1963)

The possible methods used in vacuum technology for partial pressure analysis are described. These methods either utilize certain sorption and desorption processes, or concentrate on procedures already well-known in mass spectrometry.

The potential applications of the first group are rather limited and have gained some importance only in special cases, mainly in conjunction with work in ultrahigh-vacuum systems. A variety of methods, however, has been developed for qualitative and quantitative partial pressure analysis in close connection to the methods of mass spectrometry. Besides the omegatron, which historically is the oldest gauge of this kind, we have to mention here the time-of-flight mass spectrometer, further spectrometer gauges based on the principle of resonant ions, the so-called "mass filter," and ion-discriminating systems using a cycloidal ion path. The special features of the various gauges, their particular advantages, and their particular sphere of applications are discussed and compared.

B634 Heimburg, R. L., Huang, A. B., et al., Measurement of pressure in ultrahigh vacuum. STAR, N63-13796 (1963)

B635 Huber, W. K., Partial pressure measurement in high and ultrahigh vacuum systems. Part I. Vacuum, 13, No. 10, 399-412 (1963)

B636 Huber, W. K., Partial pressure measurement in high and ultrahigh vacuum systems. Part II. Vacuum, 13, No. 11, 469-73 (1963)

B637 Huber, W. K., Comparative measurement between Bayard-Alpert gauge and Lafferty's magnetron gauge. 2nd European Vacuum Symp. 5-7 June, 1963, 240-244, Rudolph A. Lang Verlag (1963)

The x-ray limits of various Bayard-Alpert gauges have been measured on a glass ultrahigh-vacuum pumping unit. A magnetron gauge with a hot cathode as described by Lafferty was used as a reference gauge in order to prevent any influence on the gauges during measuring. The time- and emission-dependent variations of the residual current in nude Bayard-Alpert gauges have been investigated on an ultrahigh-vacuum pumping unit made of stainless steel. The attempt was made to locate and partially explain the errors in measuring. (author)

B638 Kreisman, Wallace, Calibrating vacuum gauges below 10^{-9} torr. Research and Development, 14, 58-60 (December 1963)

Description is provided of a system of calibrating vacuum gauges at pressures of 1×10^{-12} torr. The general arrangement of the vacuum-system components is shown in a block diagram. The experimental procedure used in making flow and comparison calibration is described. It is stated that absolute gas-flow measurements have already yielded believable calibrations in the 10^{-11} and 10^{-12} torr regions, and that the method can be extended to lower pressures as better vacuum systems are constructed.

B639 Lafferty, J. M., Hot-cathode magnetron ionization gauge with an electron-multiplier ion detector for the measurement of extreme ultrahigh vacua (10^{-14} and beyond). Rev. Sci. Instrum., 34, No. 5, 467-476 (May 1963)

A hot-cathode magnetron gauge capable of measuring pressures down to 10^{-14} torr has been described in previous publications. The low-pressure limit of this gauge has been extended by the addition of an electron-multiplier ion detector. Measurements of sensitivity and x-ray photocurrent indicate that this device should have a linear response down to pressures of 2×10^{-17} torr. Ability to measure low pressures with the gauge appears to be limited by the inherent dark current of the multiplier.

B6310 Lafferty, J. M., and Vanderslice, T. A., Instruments and Control Systems, 36, 90-96 (March 1963)

B6311 Mulkey, M. R., Klautsch, R. E., et al., Space simulation chamber instrumentation. STAR, N63-17068 (1963) 70 p. 22 ref.

B6312 Schueman, W. C., A photocurrent suppressor gauge for the measurement of very low pressures. STAR, N63-10551 (1963) 13 p.

An ionization gauge has been designed which removes the low-pressure limitation caused by x-rays found in conventional ionization gauges, while maintaining a comparable sensitivity. In this gauge, an additional electrode establishes an electric field which repels photoelectrons ejected from the ion collector. Preliminary experimental results have verified the suppression of photocurrents; pressures as low as 2×10^{-12} torr have been measured.

B6314 Trendelenburg, E. A., Progress in the ultrahigh-vacuum field. I. Principles of measurement and measuring instruments. Phys. Blatt, 14, No. 10, 463-470 (October 1963)

B6315 Vanderslice, T. A., Ultrahigh-vacuum instrumentation. Science, 142, 178-184 (1963) 16 ref.

General review of measurement techniques associated with ultrahigh vacuum, including instruments that measure the total pressure of all components in a system, and devices that measure the individual components which contribute to the total pressure. Of the former instrument, hot-cathode devices and cold-cathode ionization gauges are described. Partial-pressure analyzers described are time-of-flight mass spectrometer and magnetic-deflection mass spectrometers.

B641 Alpert, D., Theoretical and experimental studies of the underlying processes and techniques of low-pressure measurement. STAR, N64-17582 (1964)

The program in low-pressure measurement centered about two major activities: first, a program of laboratory research in the general area of problems and techniques of low-pressure measurement; second, a survey of current activities in aeronomy, particularly those involved with experimental studies of the neutral density and composition of the upper atmosphere. Among the topics discussed are studies of anomalous effects in ionization gauges, studies in new methods of pressure measurement, thin-film electron sources, and a survey of measurement of neutral and ion composition of the upper atmosphere.

B642 Apgar, Edward, Differences between nude and enclosed ion gauges in an oil-free system. STAR, N64-12063 (1964)

High-conductance "nude" gauges and low-conductance glass-enclosed tubulation ion gauges were compared in an oil-free, all-metal, bakable ultrahigh-vacuum system over the pressure range of 10^{-10} torr to 10^{-5} torr. Large differences in pressure indications were observed with CO, the tubulation gauge reading consistently less than the nude gauge. Smaller differences in the same direction occurred with H_2 and N_2 and none with He. Heating the tubulation or maintaining the pressure above 10^{-7} torr reduce the difference. An explanation is offered in terms of adsorption in the tubulation and a low effective conductance.

B643 Brunnee, C., et al., Quadrupole high-frequency mass spectrometer for residual gas analysis in high-vacuum systems. Vakuum Technik 13, No. 2, p. 35-42 (1964)

The quadrupole tube is 20 cm long, of all metal construction and bakable up to 450°C. No deflecting magnet is used, separation being performed by means of a high-frequency electric field. The ion source is in the form of a nude system, and the normal ion collector can be exchanged against a secondary electron multiplier.

The mass range is from 1 to 120, with a resolution of approximately 100. For 100-percent transmission, the sensitivity is of the order of 2×10^{-4} A/torr. The lowest detectable partial pressure is in the region of 10^{-13} torr.

B644 Carr, P. H., Problems in establishing standards for vacuum measurement and in calibrating vacuum gauges. Vacuum, 14, No. 2, 37-45 (1964)

An evaluation of the basic standards of vacuum measurement and a survey of recent developments in the technique of calibrating vacuum gauges are given. Methods for minimizing and correcting for systematic errors in the McLeod gauge are reviewed; a new precision manometer is described. Methods for checking the linear response of ionization gauges and extrapolating the calibration against a McLeod down to 10^{-10} torr are reviewed. Recent developments in the use of the orifice in calibration systems are evaluated.

B645 Cavaleru, A., et al., Adsorption measurement in high vacuum Brit. J. Applied Physics, 15, No. 2, 161-168 (1964)

B646 Shuemann, W. C., de Segovia, J. L., and Alpert, D., ASTIA, AD 427-830 (1964)

A systematic study of an anomalous nonlinearity in ionization gauges was made, indicating the magnitude of possible errors in ionization-gauge readings. The errors are particularly enhanced after O_2 has been introduced into the system. Evidence was provided for interpreting the effect as due to dissociative ionization by electron impact at the grid surface at low values of grid correct. There is an enhanced rate of production of such ions accompanied by a reduced rate of collection in a Bayard-Alpert gauge. The anomalous readings in ionization gauges can be reduced greatly by thorough outgassing of the grid by electron bombardment. Thereafter, operation at electron currents of 5 ma or greater gives reliable readings. In the presence of O_2, a number of other effects may be present to give erroneous ionization-gauge reading; composition of the gas may be altered both by chemical reactions at the hot filament and by electron bombardment of adsorbed gases.

C. MISCELLANEOUS

C621 Bachler, W., Cryogenic pump systems operating down to 2.5°K. Transactions 9th Natl. Vacuum Symposium, 1962, 216-219

C631 Adam, H., The significance of cryogenic techniques for the production of low pressure. Revue de la Société Royale Belge des Ingénieurs et des Industriels, 11, 457-463 (1963)

The performance of cryogenic devices in the production of low pressures depends on a number of physical parameters, such as condensation and accommodation coefficients, adsorption isotherms and saturation pressures of the gases in question. Starting from first principles, the author applies the method of Langmuir and de Boer to the problem of static and dynamic equilibrium, and deduces formulae for pumping speeds and ultimate pressure. Drastic final pressure reductions are obtained by producing fractional monolayers absorbed at liquid-helium temperature. Simultaneous pumping of various gases and the effect of cryogenic trapping are among the typical cryogenic problems being studied today.

C632 Alpert, D., Ultrahigh Vacuum; a survey. STAR, N63-16869 (1963), 34 p., 23 ref.

Low-pressure measurement devices and instruments for producing high vacuum are reviewed. It is found that these two types of instrument systems are approximately equivalent to pressures down to 10^{-10} torr. Two new classes of ionization instruments, sensitive to 10^{-11} torr for the measurement of total pressures and for the mass analysis of partial constituents as low as 10^{-16}, are explained. The study of the physical and chemical processes occurring at 10^{-11} torr is summarized and includes: The kinetics of gas surface interaction at the interphase, the interaction of atomic particles, electrons and photons with surfaces and the nature of the electronic bonds between adsorbed molecules and surfaces.

C633 Ansley, S. P., Jr., Cryogenics as applied to the design and fabrication of a space simulation test cell. N63-11606 (1963) 41 p., 10 ref.

C634 Collins, I. A., Jr., and Dawson, J. P., Cryopumping of 77°K nitrogen and argon on 10-25°K surfaces. STAR, N63-16033 (1963)

Results are presented on the first phase of an experimental apparatus designed to study the effects of gas and cryosurface temperatures on the capture coefficient in cryopumping. Experimental pumping speeds and capture coefficients were obtained for 77°K nitrogen and argon on a spherical cryosurface with temperatures ranging from 10°K to 25°K. The capture coefficients for the two gases were found to follow experimental functions of the cryosurface temperature. These functions were found to be similar to the Boltzman distribution.

C635 Farkass, I., and Horn, G. W., Cryogenic pumping in space simulators. A.I.Ch.E. Annual Meeting, Symposium on Molecular Properties at Cryogenic Temperatures (Houston), Dec. 1-5, 1963, Paper 456, 12 p.

Experimental investigation of cryogenic pumping techniques for space simulators. The results show that large space simulators utilizing vapor-diffusion pumps and liquid-nitrogen-cooled trap alone in the pumping duct can achieve an end-vacuum pressure of about 10^{-11} torr without the use of any type of cryogenic surface in the chamber. The addition of a small helium cryopump can lower the pressure in the chamber into the 10^{-12} torr region.

C636 Hengevoss, J., and Trendelenburg, E. A., Cryogenic pump tests with hydrogen under ultrahigh vacuum conditions with range 4.2°K and 2.5°K. Z. Naturforsch. 18a, 481-489 (1963)

C637 Hood, C. B., Cryopumping speeds in large simulation chambers. Inst. Environ. Sciences, Proc. Annual Technical Meeting (1963), p. 81-87 (IAA A63-18269)

C638 Klipping, G., and Macher, W., Production of vacuum by condensation of gases on low temperature surfaces — cryogenic pumps. STAR, N63-23371 (1963)

C639 Mark, H., and Sommers, R. D., The combined use of liquid and gaseous helium to provide near actual space environment. Advances in Cryogenic Engineering, 8, 93-99 (1963), 2 ref. (also STAR, N63-19618)

A facility has been designed, built, and successfully operated which has as its purpose the reproduction of the environment of interplanetary space. During operation at space conditions, liquid helium fills a jacket which completely surrounds the space volume and thus maintains the wall at liquid-helium temperatures. To economize on the use of

liquid helium, gaseous helium is circulated through the system to reduce material temperatures below 35°K, before pouring the liquid. Liquid helium boil-off rate during space-simulation operation has been maintained at about twice the solar-plus-conduction heat load, and runs of 8 hours have been made with solar radiation, 4°K background, and pressure less than 10^{-13} torr. Longer runs are possible and require only a more accurate liquid-helium-level indicator to be incorporated, since most of the boil-off occurs near the top of the chamber. The major cost of providing the refrigeration is that of the helium gas, and means for reclamation of the boil-off must be incorporated.

C6310 Mullen, L. O., and Hiza, M. J., AICHE Annual Technical Meeting, Symposium on Molecular Properties at Cryogenic Temperatures. (Houston, Texas), Dec. 1-5, 1963, paper 45a, 16 p. 33 ref. (also IAA, A63-26054)

Review of cryopumping methods for producing high and ultrahigh vacua required in numerous applications, ranging from vacuum deposition to space simulation. Application to unbaked elastomer-flanged systems are stressed. Discussed are the types of gas loads encountered in typical systems; some of the limitations of conventional diffusion pumps; and advantages and limitations of cryogenic pumping methods.

C6311 Reikhrudel, E. M., and Smirnitskaia, G. V., Modern methods for obtaining ultrahigh vacuum. Zhurnal Tekhnicheskoi Fiziki, 33, 1405-29 (1963) (Russian)

Presentation of some of the basic methods used for obtaining ultrahigh vacuum both for research purposes and for industrial applications. The principles of operation and characteristics of several ultrahigh-vacuum pumps are described, including sorption-cryosorption pump, sorption-ion-evaporation pumps, sorption-ion sputter pumps, and turbomolecular pumps. Particular attention is paid to sorption-ion pumps with cold cathodes and turbomolecular pumps, which possess considerable advantages over the other types of pumps.

C6312 Simons, J. C., Vacuum Technology in Environmental Simulation. STAR, N63-20884. (1963), 25 p., 44 ref.

Vacuum technology requirements for high-altitude and space-environment simulation are reviewed. The characteristics of the total space environment are outlined. Since total simulation of the space environment is rarely attempted, the extent of vacuum required for simulation of thermal radiation, shortwave electromagnetic radiation, energetic-particle bombardment, and surface phenomena (fatigue, bearing wear, cohesion, etc.) is indicated. The relationships of the total gas load and the volumetric speed of the pumping system in conventional high-vacuum systems are discussed; techniques for vacuum pumping are classified according to the principles involved. Current ultrahigh-vacuum techniques are reviewed, and pumping techniques for the 10^{-10} torr range are evaluated. Molecular drag pumps, ion-gettering pumps, pumping by surfaces cooled to liquid-nitrogen or liquid-hydrogen temperatures and cryosorption are considered, and their appropriate applications indicated; but the oil-diffusion pump, properly trapped to eliminate backstreaming oil vapor, is considered to be the most suitable general-purpose pumping means. Achievement of still lower pressures (10^{-13} to 10^{-15} torr) will depend upon reducing the nonessential gas load, perhaps by using double-walled chambers, and reducing the backward flow common to all pumps. Measurements of low pressures and nonisotropic situations are also considered.

C6313 Wallace, D. A., and Rogers, K. W., Molecular flow trap. Research and Development, 14, No. 9, 64-67 (1963)

A gas in which the mean free path of the molecules is large compared with some significant surface characteristic of the enclosure is termed a free molecular gas; in this regime, molecule-surface collisions are more significant than encounters between molecules. It has been determined experimentally that a cryopump does not condense each molecule as it strikes, but that only a fraction will stick. Thus, sticking coefficient can be increased by shaping the condenser surface to ensure multiple collision and reduce diffuse reflection. Such a cryopump is termed a molecular trap. The authors illustrate an example designed to deal with the directed nozzle flow of a rocket exhaust. This model achieved a sticking coefficient of 99-1/2 percent; a cryopump three orders of magnitude larger would have been required if the gas had been allowed to diffuse normally.

Page intentionally left blank

SPACECRAFT AUXILIARY THRUST SYSTEM: PERFORMANCE-MEASURING TECHNIQUES

J. M. Bridger

The advent of the stabilized spacecraft has brought with it totally new requirements for thrusting devices. These new requirements include very low thrust levels; periodic operational modes involving very short "on-times;" long lifetimes; rigorous repeatability; "no" leakage; and operation in an environment including hard vacuum, extremely low temperatures, and high radiation and particle fluxes. Although these requirements are individually difficult of attainment and verification, they are so interdependent as to render even the definition of levels by analytical means of questionable accuracy. The very large expenditures involved in engineering, and the almost prohibitive cost of flight testing, in addition to the rigorous performance requirements, have made imperative the evolution of flight-simulation techniques which will serve to improve both flight-performance capability verification, and system and component analysis and selection criteria.

Techniques and equipment developed in the past to assess the performance of thrust devices have been designed for continuous operation at sea-level ambient pressures, and are inadequate for pulse-mode operation at simulated altitude measurements. Even empirical thrust coefficients which have been experimentally established to relate sea-level to vacuum performance have made assumptions which do not hold for pulse-mode operation. The addition of environmental factors (vibration, principally) generated by test facilities make the task of evaluating thrusters operating in the millipound range a very formidable one.

To illustrate the magnitude of the task, consider some typical ranges of performance required for attitude stabilization and station-keeping of a synchronous satellite. Actual levels depend upon the control accuracies specified and the control method employed.

- Thrust: 10^{-3} to 10 lbs.*
- Response time: 5 - 15 milliseconds
- Normal impulse bit: 3×10^{-4} lb. sec.
- Duty cycle: 1 - 5%
- Lifetime: 3 to 5 yrs.
- Total cycles: 250,000 or more
- Pulse reproducibility: <5% deviation
- Ultimate vacuum: $<1 \times 10^{-11}$ torr
- Temperature range: 20 to 3000°R**

*For some applications, a maximum allowable thrust of 9×10^{-6} lbs. has been calculated.

**This is nominal temperature range from space ambient to equilibrium for thruster operating in continuous mode.

The choice of the measurement technique and the basic design of the measuring instrument must take into account the environment, from the points of view both of the degree of simulation required and of the undesirable ground-generated excitation functions present. The equipment also must be designed to produce data whose accuracy and frequency response are consistent with the way these data are to be employed.

Measurements of thrust and impulse bit are those most commonly desired. As an initial target design for a thrust stand for vacuum operation, the following performance characteristics were defined:

- Thrust range: 5 to 200 millipounds
- Frequency response: 250 cps
- Accuracy: ±5%
- Operating pressure range: ambient to 10^{-8} torr

Other very desirable information indicative of the performance of a thrusting device includes exhaust velocities, exhaust (stagnation) temperatures, and the shape of the exhaust plume. The performance objectives assumed here were:

- Stagnation temperature range: 1000 - 3000°R
- Absolute temperature range: 1000 - 3000°R
- Accuracy: ±5%
- Response adequate to measure temperature for single 15-millisecond pulses

A vibration survey of the 4- by 5-foot thermal-vacuum space simulator was made. Measurements were taken on an overhead T section and on the shroud rails where experiment installations are mounted. Measurements in three axes were made at the door end and at the middle of the chamber, with various combinations of exciting equipment in operation. The maximum accelerations observed (0.05 g zero-peak) were in the frequency range of 75-80 cps on the door end of the T section along the chamber's horizontal axis. The maximum acceleration observed on the shroud rails (0.04 g zero-peak) was at 60 cps, also along a horizontal axis at the door end of the chamber. Vibrations were observed also at lower amplitudes in other frequency ranges: 25-30 cps, 50-70 cps, 75-80 cps, 200 cps, and 240-260 cps. A survey of the data indicated some dependence of the accelerations on the phases of the different exciting equipments.

A number of questions need answers in order to assess the suitability of the target performance objectives: Are thrust levels, temperatures, response times, and accuracies adequate to the evaluations planned, and are they consistent with the current state of the measurement art? What are the optimum measurement approaches? Are the target requirements defined in unequivocal terms, so that there can be no misunderstanding of their meaning?

Thrust and impulse measurements are discussed by J. C. Heine; and jet exhaust measurements are discussed by Barry Yaffee, in subsequent sections of this report.

THRUST-MEASUREMENT STAND DESIGN CONSIDERATION

J. C. Heine

Several different types of satellite attitude-correction system are available and have been incorporated into satellite designs. One of these employs small rockets (with very low peak thrusts) located on the perimeter of the satellite. Short bursts from these rockets apply specific torsional impulses to the satellite at regular intervals in time to maintain, as closely as possible, a desired attitude in the face of perturbing torques (such as gravity or solar torques).

It has been suggested that the original analysis of problems of satellite dynamics involving thrusters be performed on an analog computer, the impulse used in the computer solution being obtained from the actual firing of a rocket. This would be accomplished by mounting the thruster on a suitable thrust-measuring device in a vacuum environment, and determining the impulse from the measured thrust.

This report will consider in detail one such thrust-measuring device, and will discuss the practicality and accuracy of several others.

MODEL

Assume that the thrust of the rocket to be tested acts in one direction. Little generality is lost if the rocket is fastened to the same apparatus which is constrained to move in the direction of the thrust. Let the total mass of the rocket, fasteners, and constraining apparatus directly attached to the rocket be M.

The rocket may now be fired under various levels of constraint in the thrust direction. Rocket constraint may be of three types: total (no motion), partial (some small motion), or unrestrained motion. The type of constraint depends on the type of measurement to be made. This section will be concerned mainly with a particular form of partial restraint, other types and forms being briefly considered under a separate heading.

Systems which partially constrain the thruster fall into two categories: active or closed-loop systems, and passive or open-loop systems. The two forms are equivalent, in that gross performance characteristics of the former may be explained in terms of system parameters of the latter. Because we are primarily interested in the effect of system parameters on system performance, we will consider only the passive, i.e., the "spring-mass-dashpot," systems.

Because the attitude-orientation process is essentially cyclic, it is desirable that the rocket-motor portion of an analog simulation be a system which operates with a minimum of recalibration of instruments or other time-consuming operations. If, for instance, the rocket-mass system always returned to the same position after a thrust pulse, instruments to detect position would have an immediate reliable zero reading. One system which has this property is a single-degree-of-freedom spring-mass-dashpot system such as that shown in Figure B-9.

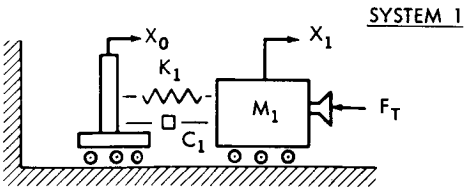


Figure B-9–Single-Degree-of-Freedom Thrust Stand

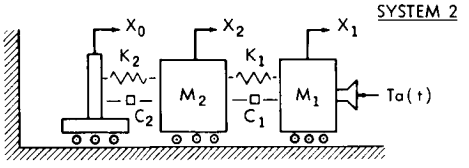


Figure B-10–Isolated Thrust Stand

EQUATIONS OF MOTION

If: F_t = the thrust of the rocket as a function of time,

X_2 = the displacement of the rocket-mass relative to some stationary coordinate system,

X_0 = the motion of the ground relative to the same coordinate system,

K_2 = the spring constant,

C_2 = the dashpot constant,

M_2 = the total mass of the rocket-mass system, and

$Y = X_2 - X_0$,

then the equation of motion of the system is

$$M_2\ddot{Y} + C_2\dot{Y} + K_2Y = -M_2\ddot{X}_0 + F_T \qquad (1)$$

The first term on the right side of equation (1) may be interpreted as the background noise in g's, and may or may not be small. If we assume it is large, we must isolate the thrust stand from the noise by placing in series with it, and between it and the "ground," a large mass with a weak spring (low natural frequency). This system will filter the noise out; that is, not transmit it to the thrust stand. Such a system is pictured in Figure B-10. If

K_1 = spring constant of isolator spring,

C_1 = dashpot constant of isolator dashpots,

M_1 = mass of isolator, and

$Y_1 = (X_1 - X_0)$,

when X_1 = absolute displacement of isolator mass and X_0 is as defined above, and $Y_2 = (X_2 - X_1)$, then the equations of motion for this system are

$$M_2\ddot{Y}_2 + C_2\dot{Y}_2 + K_2Y_2 = -M\ddot{X}_1 + F_T \qquad (2)$$

$$M_1\ddot{Y}_1 + C_1\dot{Y}_1 + K_1Y_1 = -M_1\ddot{X}_0 + C_2\dot{Y}_2 + K_2Y_2. \qquad (3)$$

These are the two models which will be considered in the following sections. It will be shown, however, that the second model may be reduced to the first after certain parameters have been fixed, and that in fact only the first need be considered when determining test-stand performance.

Finally we will discuss the thrust function F_T.

Two thrust-time functions to be used in equations (1), (2), and (3) were selected. They were:

$$F_T = F(1 - e^{-rt}) \qquad 0 \le t \le \tau \tag{4a}$$

$$= 0.99\, Fe^{-r(t-\tau)} \qquad \tau \le t \le 2\tau$$

$$F_T = 0 \qquad 2\tau \le t$$

$$F_T = \frac{1}{2} F \qquad 0 < t \le 2\tau \tag{4b}$$

$$= 0 \qquad 2\tau \le t$$

where F is the peak force exerted by the thrustor and r and t are chosen so that $F - F_T \le .01$ when $t = \tau_0$.

The reason for selecting these particular functions is that practical experience has shown that the thrust of a small rocket tends to increase and decrease approximately exponentially. Function (4a) has such properties. It is quite probable, however, that the actual thrust may differ quite markedly from (4a); it is therefore desirable to obtain some upper bound (which will be discussed in more detail in a later paragraph).

Equations (1), (2), and (3) were solved on a Pace 231 analog computer, with (4a) and (4b) as forcing functions.

SYSTEM RESPONSE CHARACTERISTICS

In this section, we will first determine the response of system 1 to the thrust functions (4a) and (4b). Secondly, we will examine the possibility of obtaining the value of the impulse applied by the thrustor to the system. Finally, we will determine the complex frequency response of the two systems.

If $I(t)$ is the response of any system to a unit impulse, the response of the system to an arbitrary excitation F_T is[1]

$$Y(t) = \int_0^t F_T\, I(t - \tau)\, d\tau. \tag{5}$$

One can easily show (cf. [1]) that the impulse of system 1 is

$$I(t) = -\frac{1}{\rho_2} e^{\zeta_2 \omega_2 t} \sin \rho_2 t \tag{6}$$

Where $\rho_2 = \omega_2 (1 - \zeta_2^2)^{1/2}$

$\omega_2^2 = K_2/M_2$

$\zeta_2 = C_2/2\,(K_2 M_2)^{1/2}$.

[1] Crandall, S.H. and Mark, W.D., "Random Vibrations," Academic Press, New York, 1963.

With (5), (6), (4a), and (4b) the solutions may be obtained and are

$$Y_A(t) = \frac{F}{K}\left[1 - e^{-\zeta_2\omega_2 t}\left\{\frac{\zeta_2\omega_2}{\rho_2}\sin\rho_2 t + \cos\rho_2 t + \right.\right. \tag{7a}$$

$$\left. + \frac{\omega_2^2}{\rho_2}\,\frac{(r - \zeta_2\omega_2)\sin\rho_2 t - \rho_2\cos\rho_2 t}{r^2 - 2r\,\zeta_2\omega_2 + \omega_2^2}\right\} -$$

$$\left. - \frac{\omega_2^2\, e^{-rt}}{r^2 - 2r\,\zeta_2\omega_2 + \omega_2^2}\right]$$

$$0 \le t \le \tau$$

Similarly, expressions may be written for the regions

$$y_B(t) = \frac{F}{K}\left(1 - e^{-\zeta_2\omega_2 t}\cos\rho_2 t - \frac{\zeta_2\omega_2}{\rho_2}\, e^{-\zeta_2\omega_2 t}\sin\rho_2 t\right) \tag{7b}$$

$$= A(t) \quad \text{for} \quad 0 \le t \le 2\tau$$

and $$y_B(t) = A(t) - A(t - 2\tau) \quad \text{for} \quad 2\tau \le t.$$

It is not possible to write similar closed-form solutions for a general two-degree-of-freedom system such as for system 2.

We will now establish a general property of linear spring-mass systems that are constrained to have motion in one direction only. This property may be studied as follows:

If a force acts on a mass M_i for a time 2τ, and if (at $t = 0$) the total kinetic and potential energy of the system of which the mass is a part is zero, then as long as some form of damping exists in the system

$$\int_0^t F_t\,dt = \int_0^T \left(\sum_j F_{s_{ij}}\right)dt \tag{8}$$

where $F(t)$ is the force acting on the mass

T is some time $T > 2\tau$

$F_{s_{ij}}$ is the jth spring force acting on the ith mass; i.e., $F_{s_{ij}} = K_{ij}\,Y_i$

and Y_i is the displacement of the jth mass relative to the ith mass.

To prove this, we use the following reasoning: If F_T is finite and lasts for a finite time, a finite amount of energy is added to the system. If a form of damping exists, that energy will be dissipated in a time T. At T, therefore, the kinetic and potential energy must again be zero. These initial and final conditions may be written

$$y_i(0) = \dot{y}_i(0) = 0 \tag{9}$$

$$y_i(T) = \dot{y}_i(T) = 0 \tag{10}$$

where $Y_i(t)$ is the relative displacement of the ith mass attached to M_i by a spring and/or dashpot from M_i.

The differential equation of motion of $\underline{M}$ may be written

$$M_i \ddot{X}_i + \sum_j C_{ij} \dot{Y}_j + \sum_j K_{ij} Y_j = F_T \tag{11}$$

where X is the absolute displacement of M_i. Integrating (11) with respect to time from 0 to t, one gets

$$M_i \int_0^T \ddot{X}\,dr + \int_0^T \left(\sum_j C_{ij} \dot{Y}_j \right) dr + \int_0^T \left(\sum_j F_{s_{ij}} \right) dr = \int_0^T F_T(r)\,dr$$

and from (9) this may be integrated to the form

$$M_i \dot{X}(t) + \sum_j C_{ij} Y(t) + \int_0^T \left(\sum_j F_{s_{ij}} \right) dr = \int_0^T F_T(r)\,dr. \tag{12}$$

If we now let t approach T, we see from (10) that the first two terms on the left of (12) are zero, and we are left with (8). Since $F_T = 0$ when $t \geq 2\tau$, we may write

$$\int_0^{2\tau} F_T\,dt = \int_0^T \left(\sum_j F_{s_{ij}} \right) dt. \tag{13}$$

It can be seen that, for systems 1 and 2, (11) becomes:

$$\int_0^{2\tau} F_T\,dt = \int_0^T K_2 Y_2(t)\,dt. \tag{14}$$

Finally, for systems 1 and 2 respectively, the complex frequency responses may readily be found to be cf [1]).

$$H_{21}(\omega) = \frac{\omega_2^2 + 2i\,\zeta_2 \omega_2 \omega}{\omega_1^2 - \omega^2 + 2i\,\zeta_2 \omega_2 \omega} \tag{15}$$

and

$$H_{12}(\omega) = \frac{\left[\begin{array}{l} -i\,\omega^3\,2\zeta_1\omega_1 - \omega^2\,(\omega_1^2 + 4\zeta_1\zeta_2\,\omega_1\,\omega_2) \\ +\, i\omega\,(2\,\zeta_1\,\omega_1\,\omega_2^2 + 2\,\zeta_2\,\omega_2\omega_1^2) + \omega_1^2\omega_2^2. \end{array}\right]}{\left[\begin{array}{l} \omega^4 - i\,\omega^3\,[2\zeta_1\omega_1 + 2(1+r)\,\zeta_2\omega_2] - \omega^2\,[\omega_1^2 + \\ +\,(1+r)\,\omega_2^2 + 4\zeta_1\zeta_2\,\omega_1\omega_2] + i\omega\,[2\,\zeta_1\,\omega_1\omega_2^2 + \\ +\,2\,\zeta_2\omega_2\omega_1^2] + \omega_1^2\omega_2^2. \end{array}\right]} \qquad (16)$$

where $H_{21}(\omega)$ is the complex frequency response of M_2 in system 1, and H_{12} is the complex frequency response of M_1 in system 2, and where

$$\zeta_i = C_i/2\,(K_i M_i)^{1/2}$$

$$\omega_i = (K_i/M_i)^{1/2}, \quad r = M_2/M_1$$

and ω = input acceleration circular frequency.

ERROR

The primary objective of the thrust stand is to produce as exact a representation of the thrust impulse as is possible by the time the thrust pulse is completed. A second objective is to produce the exact thrust-time characteristic of the rocket. Obviously, if the latter is obtained, the former is easily determined by an integration. We note, however, that if F_T is to be produced, each of the four other terms in equation (1) must be determined.

We can reduce this number to three by eliminating the first term on the left side of (1). We do this by showing that, if we use system 2 instead of system 1, the acceleration of the isolator mass X in (2) (see Figure B-10) may be neglected if the parameters K_1 and M_1, the isolation spring constant and mass, are chosen properly.

Consider the complex frequency response for mass M_1 of the second system as given by equation (16). The absolute value of $H_{12}(\omega)$ may be considered to be the ratio of the peak mass acceleration to the peak input acceleration of a sinusoidal excitation. The peak mass acceleration is therefore simply

$$\ddot{X}_1 = |H_{12}(\omega)|\ (\ddot{X}_0)_\rho$$

where $(\ddot{X}_0)_\rho$ is the peak input acceleration. To show the effect of isolation, let us consider an example:

The "ground" acceleration $\ddot{X}_0$ has been measured and found to be very close to sinusoidal, with a peak of approximately 1/32 g at 30 cps. Assuming the isolator parameters to be $\zeta_1 = 0.01$ (low damping) and $\omega_1 = 3.14\ \text{sec}^{-1}$ (1/2 cps); the thrust stand parameters to be $\zeta_2 = 0.6$ (high damping) and $\omega = 1570\ \text{sec}^{-1}$ (250 cps), and the mass ratio r to be 200, we find

$$|H_{12}(\omega)| \cong 2.3 \times 10^{-6}.$$

This indicates that the ground acceleration that acts directly upon the thrust stand (e.g., $\ddot{X}_1$ in system 2) can be easily reduced by a factor of approximately ten thousand by isolation. We conclude that it is entirely reasonable to neglect the $\ddot{X}_1$ term in equation (2), if: (a) the natural frequency $(K^2/M_2)^{1/2}$ of the isolator is low, and if (b) the mass ratio r is high. This last condition will be discussed further.

IMPULSE AND THRUST TIME CHARACTERISTICS

The displacement $Y_2(t)$ may be determined very accurately with equipment that is readily available. If $Y_2(t)$ and $\dot{Y}(t)$ were continuous and differentiable in the region of interest, one could obtain all the terms in equation (2) (neglecting $\ddot{X}_1$) and determine F_T to within the accuracy of the $Y(t)$ measurement. Because of background noise the rocket exhaust, however, the existence of the desired continuity in $Y_2(t)$ is highly unlikely. Electronic differentiation of discontinuous signals leads to errors much larger than are acceptable for this application. Let us examine the parameters which determine the error in measured impulse and thrust-time characteristics if we make the approximation:

$$F_T = F_{s_2}(t) \tag{17}$$

where $F_{s_2}(t) = K_2 Y_2(t)$, the force in the spring K_2.

If we examine the responses (7a) and (7b), we see from the latter that the transient terms (i.e., the terms which determine the deviation from a step) decay exponentially with $\zeta_2\omega_2$. Note that this product is independent of K_2; i.e., $\zeta_2\omega_2 = 2C_2/M_2$. Thus, if the transient terms are to decay in a short time, either C_2 must be large or M_2 must be small. This latter restriction is the obvious one of increasing the ability of an instrument to follow a signal by decreasing the total inertia of the system. Similarly, from (7a) we see that transients die out exponentially with $\zeta_2\omega_2$; but here we also see that $\zeta_2\omega_2$ must be as large as possible relative to r, the inverse of the time necessary for the thrust to be within 63 percent of its final value. If $\zeta_2\omega_2 >> r$, as can be seen from Figure B-11, the transients last an insignificant length of time, and the force in the spring lags behind the thrust by a time (γ) given by

$$\gamma = \frac{1}{r} \ln \left(\frac{1}{1 + \left(\frac{r}{\omega_n}\right)^2 - 2\zeta \frac{r}{\omega_n}} \right)$$

It can be seen that the time lag approaches zero as ω_n becomes large compared to r. Thus, the restriction on K_2 is that $K_2 >> M r^2$, yet K must not be so large that the deflections caused by the small thrust are undetectable.

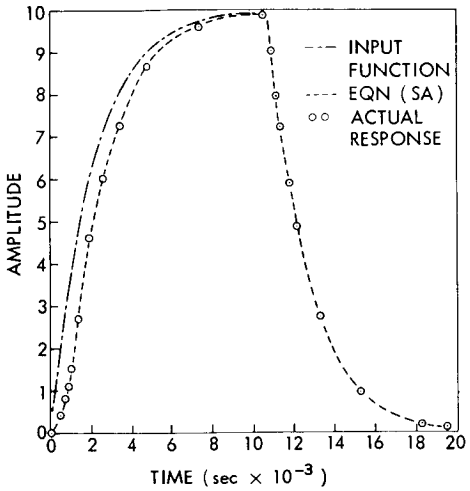


Figure B-11–Decay of Transient Terms

To determine the error in measured impulse that results from the approximation (15), we set T in equation (12) equal to 2τ. Then

$$I \lessgtr \int_0^{2\tau} K_2 Y_2 \, dt$$

If we let

$$J = \int_0^{T} K_2 Y_2 \, dt$$

where

$$I = \int_0^{2\tau} F_T \, dt.$$

Then the percentage of error between I and J is

$$E = \frac{I - J}{I}$$

By solving equation (2) on an analog computer with (4a) and (4b) as forcing functions, E was continuously computed for a series of thrust stands with natural frequencies between 25 and 300 cps and damping ratios between 0.1 and 0.7. In all cases, the total input impulse was 10 volt-seconds and the value of E at $t = 2\tau$ was recorded.

It can be shown that it is not possible to determine meaningful upper or lower bounds for the percent difference E if F_T is completely arbitrary. The engineering assumption was made that the actual thrust-versus-time characteristic of the rocket would be a linear combination of a smooth increase and decrease, which would be at least as severe as (4a) but not as severe as the step input (4b) plus some noise.* It was believed that some approximation to the bounds of the error measured in impulse (and to the manner in which the bounds converge) after time 2τ could be obtained.

Typical results are shown in Figure B-12. Dotted lines represent the maximum error of either (4a) or (4b), and the shaded area is the area bounded by the actual error-versus-frequency curves. It should be noted that the curves converge more rapidly for systems with low damping. In all cases with damping factor 0.3 or lower, if the natural frequency is greater than 150 cps, the error in measured impulse is less than 3 percent.

* The noise input was Gaussian white noise with 1.5-volt RMS and a bandwidth of 60 cps about 0. No significant change in results was noted when this noise was added to (4a) or (4b).

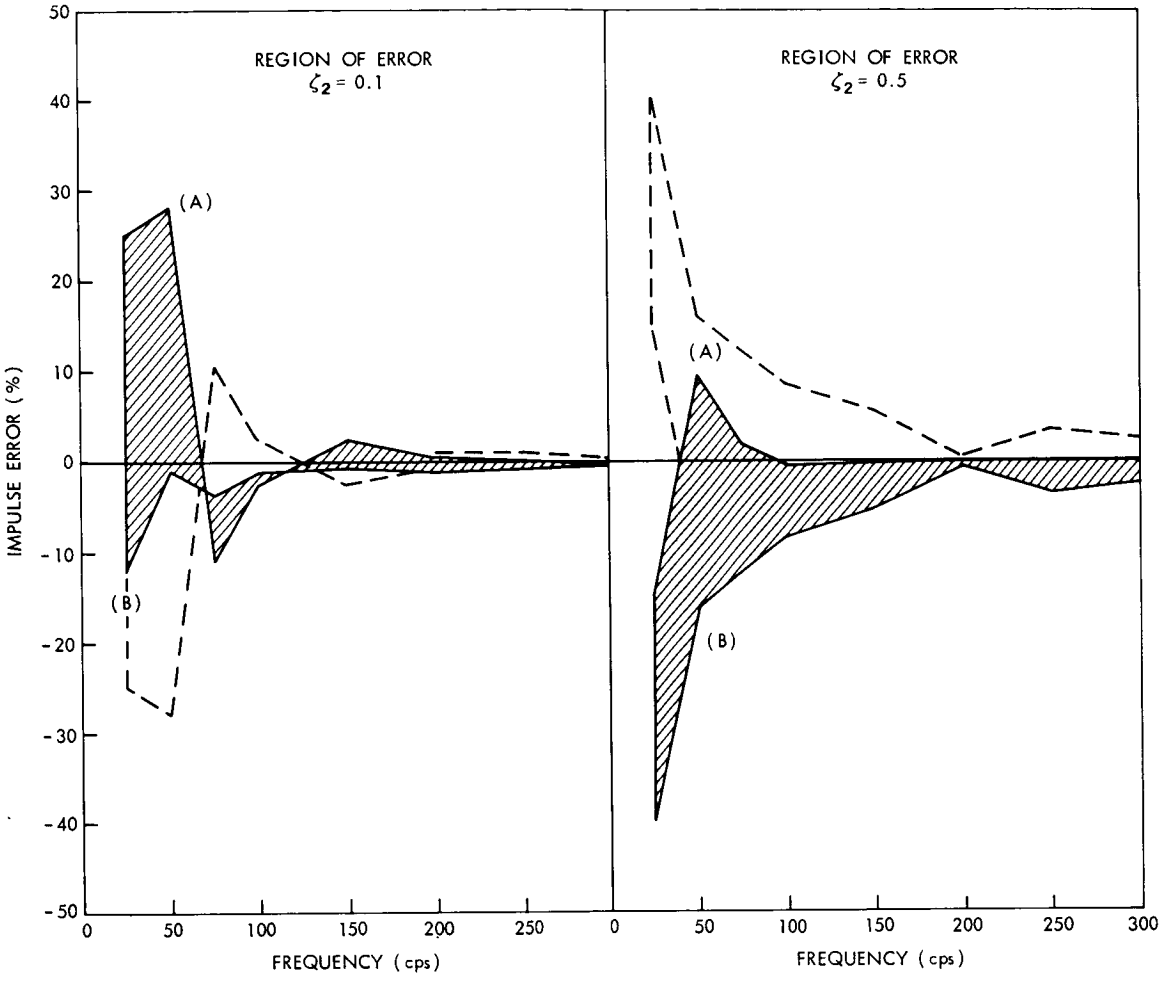


Figure B-12–Input-Response Errors Used as Upper and Lower Bounds

OTHER SOURCES OF ERROR

Possible nonlinearities in system components such as friction or nonlinear spring dashpot will not be considered. It is necessary that the spring constant K_2 be relatively high, which necessarily limits the motion of the system so severely that the assumption of linear components seems justified. There is no possibility of air drag, as all testing will take place in a vacuum. The possibility of interaction between the thrust stand and isolator seems possible; analog computer solutions of equations (2) and (3) have shown that this is indeed the case for large or moderately small values of the ratio M_2/M_1. If this parameter is below 5×10^{-3}, however, the isolator mass is so massive that the motion of the thrust stand causes a negligible acceleration ($\ddot{X}_0$ in equation (2)).

SENSORS

General instrumentation problems (such as amplifier noise), will not be considered here; however, since we have specified that motion must be detected, a section on motion sensors is included.

It is obvious from the above discussion that the spring force will be determined by finding the spring constant (or equivalent spring constant) of the system, and then computing the product of this number and a displacement-versus-time function.

In all probability, this displacement function will be produced by a capacitance-type displacement detector, most simply a capacitor made up of the moving body and a plate separated by some distance. As the distance varies, the capacitance is changed, and this is detected by a bridge circuit. This method, which has been used effectively in the past, shows promise of eventually being able to detect displacements down to 10^{-10} inches. Thus, having decided to measure spring force as a function of time, we find that, with suitable system parameters and instrumentation, a good approximation to applied impulse and a reasonable approximation to thrust-versus-time function are produced.

Now let up briefly consider some of the other approaches mentioned in an earlier section. In fact, any method of producing a thrust-versus-time function must involve motion, as sensors can detect only change. All sensors, in principle, detect displacement; it is only what is done with the displacement signal that determines whether the sensor is called an accelerometer, an integrating accelerometor, or a displacement sensor. (Thrust transducers are essentially capacitive displacement sensors). We therefore conclude that it is not possible to measure thrust if the thruster is totally constrained.

What about unrestrained motion? By the law of conservation of momentum, the final velocity of the thruster would be a measure of the impulse. It is not possible, however, to measure velocity of a freely moving object in as short a time as is required for the computer applications; therefore, this possibility is also eliminated. Finally, let us consider another possible method of detecting the motion of the partially restrained system.

As we have already noted, an accelerometer is a displacement-measuring device. An accelerometer will be driven by the acceleration of the thruster, and as can be seen from (4a) and (7a) an exponentially increasing thrust produces a very high initial acceleration of the thruster. Now the question is: Can the accelerometer (a mechanical one degree of freedom system) follow this acceleration, and produce a better measure of the thrust-versus-time function, by integrating the signal to produce velocity and displacement, and finally multiplying these signals by the appropriate (experimentally produced) constants to determine all the terms in equation (2)?

The answer is definitely not. Just as the displacement of the thruster cannot exactly reproduce the thrust-versus-time function (disregarding lag), the displacement of the mass in the accelerometer cannot follow the acceleration input, although it might be a closer approximation because the natural frequency of the acceleratometer will generally be higher than that of the thrust stand.

Thus, a record produced by an accelerometer would have its maximum error in the same regions as the maximum error produced by displacement measuring; but, because of a higher natural frequency, the duration of time during which the signals would differ greatly (again neglecting phase) would be reduced.

CONCLUSIONS

The study of an isolated single-degree-of-freedom thrust stand with linear spring and dashpot has indicated that a good approximation to the total impulse of a rocket may be obtained by integrating the spring force with respect to time, if:

(a) the undamped natural frequency of the thrust stand is greater than 150 cps.

(b) the damping coefficient of the thrust stand is moderate (0.1 to 0.3).

(c) the ratio of the thrust-stand mass to isolator mass is low (5×10^{-3}).

(d) the natural frequency of the isolator is low compared to the lowest frequency of essentially sinusoidal ground motion.

An exact measurement of a short-thrust pulse with a single-degree-of-freedom thrust stand is impossible, unless the displacement signal can be differentiated.

In general, the higher the magnitude of the product of the damping coefficient and the undamped natural frequency, the more exact will be the approximation of setting spring-force equal to thrust.

There seems to be no reason to assume that a form of sensor other than a displacement sensor would produce measurably better approximations to the thrust-time and impulse characteristics of the rocket.

Page intentionally left blank

JET EXHAUST MEASUREMENTS

Barry Yaffe

It is suggested that shielded thermocouples operated in conjunction with compensating circuitry offer an acceptable method for determining the total energy contained in the exhaust plume of a small resistojet thruster operating in a vacuum environment. Interest is focused on the transient mode of the system.

OUTLINE OF PROBLEM

One measure of the efficiency of devices producing thrust by exhausting matter to the environment is to determine the total energy of the exhausted matter as a function of any input energies required to initiate and sustain the flow of such matter. The thruster of interest herein is a thermal-resistance jet shown in Figure B-13. Gas (usually hydrogen, helium, or ammonia) is stored in a tank before operation. Upon energizing a solenoid valve in the inlet system, the fluid is metered through a critical orifice. The valve requires about 10 milliseconds to go from closed to open to closed; therefore, since the valve is the lower limit on pulse duration, pulses of 10-millisecond duration may be instigated. This pulsed mode of operation is what is required in attitude control and station-keeping, for which a short pulse of small thrust is all that is needed for proper positioning. Such pulses may be applied intermittently to maintain position over extended periods of time.

As the pulse advances through the thruster, it is heated by means of a tungsten-rhenium wire embedded in a core of boron nitride, a substance which is a poor electrical conductor and at the same time a good heat conductor. The heated gas is then exhausted through a converging-diverging nozzle into the vacuum environment. Temperatures of the order of 2000°R and thrusts of the order of millipounds are attained.

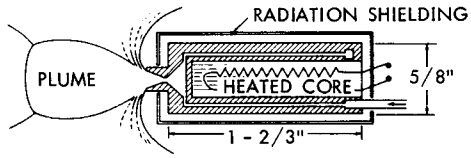


Figure B-13–Thermal-Resistance Jet Thruster

The measure of efficiency we seek involves measurement of the total energy in the exhausted gas as a function of the known electrical input to the resistance heater. The specification of a method for such measurement involves knowledge of the shape and structure of the exhaust plume (i.e., how much spread is there; does the plume curl back to contaminate satellite surfaces; what are the importance of shock phenomena?), since we would expect to use a finite number of sensing devices placed in regions of the plume which contain significant contributions to the total energy.

In the exhaust plumes of jets operating at vacuum conditions and at low mass-flow rates, pressures can be very low. It is expected that continuum fluid mechanics as a tool for analysis must be supplemented by rarefied gas-flow theory.

The problem posed is: Specify a method to obtain the total energy of the exhaust plume of a thermal-resistance jet operating under conditions of low mass-flow rates, unsteady state, and vacuum environment.

ENERGY CONSIDERATIONS

Consider a particle p of ejected gas which leaves the thruster with an instantaneous temperature $T_p(t)$ and instantaneous speed $V_p(t)$. If the mass of this particle is m_p, and its heat capacity at constant pressure is $C_p(T)$, then:

$$E_p(t) = m_p C_p(T)\, T_p(t) + \frac{1}{2}\, m_p V_p(t)^2 \tag{1}$$

Other energies, such as gravitational potential energy, are not considered, because they are either invariant or negligible. Since we will be considering a volume distribution of such particles, we write the differential energy per unit volume at a fixed point (X, Y, Z) as:

$$\underset{dxdydz}{dE(x,\ y,\ z;\ t)} = \rho(x,\ y,\ z;\ t)\ \{C_p(T)\, T(x,\ y,\ z;\ t) + \frac{1}{2}\, V(x,\ y,\ z;\ t)^2\} \tag{2}$$

Now, if we perform a process in which the p^{th} particle is brought to rest without a change in its energy level, we can write:

$$m_p C_p(T)\, T_p(t) + \frac{1}{2}\, m_p V_p(t)^2 = m_p C_p(T)\, T_{p,\ stag}(t) = E_p(t) \tag{3}$$

where $T_{p,\ stag}(t)$ — actually a constant with respect with time, since the energy is conserved — is the temperature rise associated with this energy conversion. For the case of the volume distribution:

$$dE(x,\ y,\ z,\ t) = \rho(x,\ y,\ z;\ t)\, C_p(T)\, T_{stag}(x,\ y,\ z;\ t)\, dxdydz \tag{4}$$

Integration over the entire volume occupied by the aggregation of particles yields the total instantaneous energy of the exhaust gas:

$$E(t) = \underset{vol}{\iiint} \rho(x,\ y,\ z:\ t)\, C_p(T)\, T_{stag}(x,\ y,\ z;\ t)\, dxdydz \tag{5}$$

To obtain the total energy imparted to the exhaust gases, we would bring the moving fluid to rest losslessly at every point and there determine the temperature; integration over all these points would yield the total energy. Of course, the quantity $\rho(x,\ y,\ z;\ t)\, C_p(T)$ is also needed in the determination of total energy.

STAGNATION REGION

A blunt obstacle placed in a supersonic stream (the exhaust gases of the thruster are exited supersonically) will set up a flow field similar to that depicted in Figure B-14. A detached shock upstream of the obstacle appears, through which the flow goes from supersonic to subsonic. At the nose of the blunt obstacle, a stagnation region exists in which the fluid is practically at rest. In any method for the measurement of stagnation temperature, we form a stagnation zone by introducing an obstacle into the exhaust flow, and then place a temperature sensor in this stagnation region. The stagnation temperature is then monitored through the sensor. The course of action is this: Choose a temperature sensor which fits the requirements of this problem. Upon determination of an

acceptable sensor, and after an investigation into errors to be expected, we would decide where and how many of these sensors could be placed in the exit plume in order to obtain an accurate measure of the time-dependent total energy.

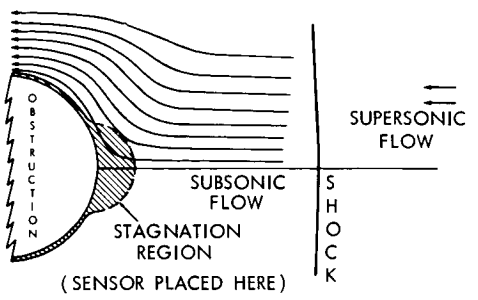


Figure B-14–Flow Field, Blunt Obstacle in Supersonic Stream

TEMPERATURE SENSORS

Many techniques have been employed in temperature measurement, and each was investigated as a possible method.

Acoustic Pulse Thermometer: This sensor utilizes the temperature dependence of the speed of sound in a gas. To determine temperature, the speed of propagation of sound waves (usually of high frequency) is determined by measuring the time of travel of the wave over some known path. This method holds promise as an alternative to the method finally selected, but seems unnecessarily complex as a first choice.

Radiation Techniques: These methods utilize the temperature dependence of intensity and wavelength of emitted radiation. Usually optical methods focus the radiation on a detector, although this is optional. Radiation techniques are classified according to the type of detector they employ: some detectors are resistance thermometers, thermocouples or thermopiles, and photosensitive detectors. The radiation detectors are of the type which need not be inserted into the region of measurement (although the stagnation obstacle is inserted), but they have the disadvantage of requiring complex measurement to compensate for the averaging process inherent in their operation. Radiation sensors are only as good as the radiation impinging upon them; gases which are poor emitters will be difficult to analyze in this way. Hydrogen and helium are poor emitters.

Magnetohydrodynamic Thermometer: The electrical conductivity of a gas depends upon the degree of ionization, which, in turn, depends upon the temperature of the gas. From this, a measurement of ion current is a means of obtaining the temperature of the gas. This method has been used successfully at elevated temperatures at which ions exist in quantity, but is not as useful at temperatures to be encountered in the resistojet.

Resistance Thermometer: The resistance of a length of electrical conductor (resistor) is a function of the temperature. A method of temperature measurement is indicated, and it is found that such a method has been used with great accuracy. As will be discussed subsequently, all heat-transfer sensors exhibit a time lag in following a transient temperature environment. Reducing the size of the sensor is one way of reducing this lag time, and such reductions have been utilized in resistance thermometers; however, these reductions have resulted in sensors which are not overly rugged and which must be replaced frequently. Modifications to be treated later indicate that this problem may have a solution; however, preference is given to a sensor which is more rugged than the resistance thermometer and which is of comparable calibration complexity.

Thermocouple: This method utilizes the contact potential which is set up between two dissimilar metals and the temperature dependence of this EMF. Such probes have been used successfully in a wide variety of temperature measurement situations, among which is supersonic gas-flow total-temperature measurement. Their characteristics in such situations are well documented. It is suggested that the thermocouple be used as the basic component in a system of total-temperature measurement.

ERRORS

The temperature reported by a thermocouple immersed in the stream of supersonic fluid will be the result of a heat balance among the sensor, the fluid, the thruster, and the environment, as well as a balance among the various components of the thermocouple support and lead wires. These modes of heat transfer are radiation, conduction, and convection. Figure B-15 shows an example of a thermocouple probe in which steps have been taken to reduce heat losses which impair accuracy. A shield is incorporated which reduces the heat lost through radiation from the hot junction and gas to the cooler environment. The geometry of the probe minimizes conductive errors within the probe itself.

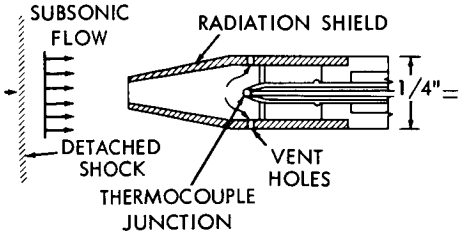


Figure B-15–Winkler Shielded Thermocouple Probe

Convective heat transfer is the process by which the hot, practically stagnant fluid transfers heat to the thermocouple junction; this process is governed by fluid momentum and energy considerations, and has been investigated by others for many geometrical configurations. The results are usually described by referring to the "recovery factor" of the probe in question. The recovery factor, r, is given by:

$$r = \frac{T_{probe} - T_{static}}{T_{stag} - T_{static}} \tag{6}$$

where r is a measure of the degree of energy conversion in bringing the fluid to rest from a velocity v and a temperature T static. Actually, for the probe in the diagram, $r = 0.994$ over a large range of flow conditions. In practice, since r is empirically determined, it may include other losses because it is difficult to determine precisely the partitioning of energy transfer through the three mechanisms. As long as r is known and relatively constant for a particular probe, we may correct the probe reading to obtain stagnation temperature with reasonable accuracy.

The foregoing discussion of thermocouple error refers to cases of steady flow. Under such conditions, it is assumed that sufficient time has elapsed so that equilibrium among the energy-exchange processes has been attained and temperature readings are constant in time. However, we must also be concerned with errors occurring as a result of the transient nature of the thruster pulse. If we assume that the only relevant mode of heat transfer from gas to sensor is convection, we may form a heat balance between the time rate of change of the heat stored in the sensor and the heat flux across its boundaries. Let $Q(t)$ be the heat in the sensor. Then the heat flux across its boundaries is usually written as $q(t) = hA[T_i(t) - T_0(t)]$ where A is the wetted surface area of the sensor, h is the so-called heat transfer coefficient, T_i is the temperature of the environment, and T_0 is the temperature of the sensor (i.e., T_0 is the temperature reported by the sensor). If v is the volume of the sensor, C its constant heat capacity, and ρ its density, then:

$$Q(t) = \rho V C T_0(t) \tag{7}$$

The time rate of change of heat stored in the sensor is equal to the flux of heat across its boundaries:

$$\frac{dQ(t)}{dt} = q(t) \tag{8}$$

Substitution yields:

$$T_0(t) + \rho \frac{VC}{hA} \frac{dT_0(t)}{dt} = T_i(t) \tag{9}$$

This equation is based on the assumption that the only relevant mode of heat transfer is convection. Because of the rarefied nature of the plume, it is possible that radiation is more important. The thermal-lag equation for the radiation case is non-linear and will be a subject of future investigation. For the present, we note that there exist electrical circuits which are governed by the same differential equation as (9). Two simple examples are shown in Figure B-16. Such circuits can obviously be used to solve the differential equation for arbitrary $T_i(t)$ if the solution is desired. More useful, however, is the fact that these circuits or more sophisticated operational amplifiers may be used in reverse as a method of determining the forcing temperature $T_i(t)$: Circuits which operate on a signal which lags an original input signal, so as to yield the original signal unaltered in time, are called compensating circuits. Such a circuit (Figure B-17) is suggested as a means of reducing time-lag errors incurred in transient measurements.

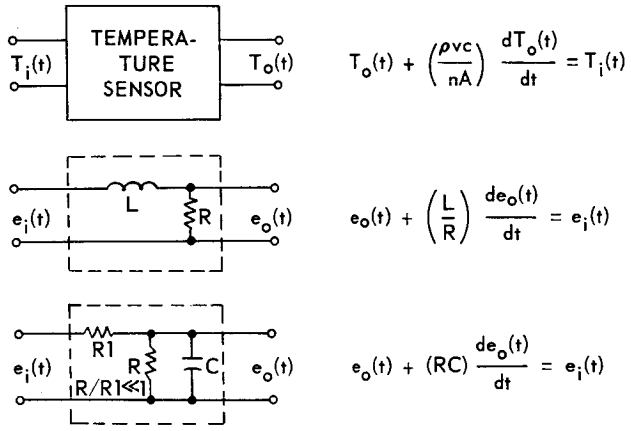
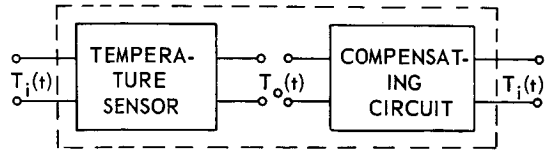


Figure B-16–Electrical Circuits

Figure B-17–Compensating Circuits



CONCLUSIONS AND RECOMMENDATIONS

The shielded compensated thermocouple has been suggested as a method of measuring the stagnation temperature in the gas plume of a low-level thruster. In arriving at this recommendation, many properties of the plume have been implicitly assumed: for example, the use of any convective heat-transfer sensor such as the thermocouple assumes that this mode of heat transfer is both dominant and subject to measurement and analysis. It is known, however, that a gas of very low density exhibits convective heat transfer to a

much smaller degree than does an ordinary gas. The degree and mode of heat transfer in low-density supersonic gas plumes will be met often in problems of this sort, and an analysis would be recommended.

Another problem requiring further analysis – determination of the shape of the plume and its structure in the presence of obstacles – also changes complexion if the gas is of very low density.

All these phenomena are properly studied under the heading of rarefied gas dynamics, and such an investigation was initiated. However, it was discovered that the research literature on the subject presupposes (and rightly so!) a sophisticated grasp of the basic fundamentals, which are, unfortunately, widely dispersed in various texts and references on fluid dynamics and kinetic theory. As many workers in the field of vacuum technology may not be conversant with these basic concepts, it was suggested that a primer on rarefied gases was needed as a prerequisite to consideration of these often-met problems. The primer which follows is an attempt to fill this need.

PRIMER OF RAREFIED GAS FLOW

Barry Yaffe

This primer attempts to present a sketch of the behavior of rarefied gases and a synopsis of the analytical tools employed in their study. Because the subject of rarefied gas flow draws heavily upon classical fluid mechanics and the kinetic theory of gases, we must take time at the outset to establish a common fund of knowledge composed of pertinent concepts from these two areas of study. Then we will be in a position to discuss rarefied gas phenomena usefully.

author

INTRODUCTION

As a tool in the discussion of the mechanics of rarefied gases, consider first a frequently encountered flow situation under both normal conditions and rarefied conditions: the laminar flow of fluid through a tube of circular cross section of radius r and length L induced by a pressure difference $P_A - P_B$ (Figure B-18). Ignoring end effects, and assuming that the flow has achieved steady-state, we seek an expression for the mass flow rate G (mass per unit time) in terms of the fluid properties (coefficient of viscosity, density, and so on), geometry (L, r), and imposed conditions (P_A , P_B).

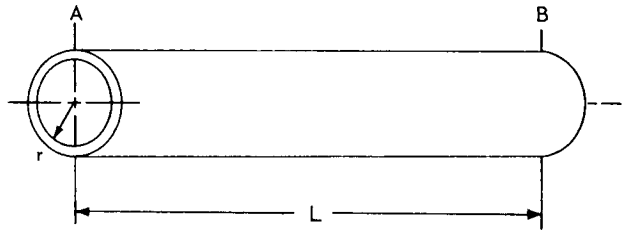


Figure B-18–Laminar Flow Through a Tube

CONTINUUM MECHANICS

The classic approach to the problem generally means use of the proper form of the Navier-Stokes equations, subject to the "no-slip" boundary condition at the wall, or Euler's equations of inviscid flow if they may be assumed valid. Either set of equations is based on several fundamental assumptions about the flow which define the limits of application. These assumptions are:

1. The physical variables (i.e., pressure, density, temperature, velocity) may be represented as continuous mathematical functions of the physical coordinates. In other words, the fluid is regarded as a continuum.

2. The stress acting on any small element of volume of fluid at a point is proportional to the rate of strain of the element. This says that the fluid is Newtonian.

3. Latent in these assumptions are the requirements that the fluid be homogeneous and isotropic, and depart only slightly from equilibrium.

The Navier-Stokes equations are statements of the transport of momentum principle applied to a small volume element of the fluid under the above assumptions. Another equation, the continuity equation, is obtained by applying the conservation of mass principle to the element. In vector notation, these may be written:

Navier-Stokes equations

$$\frac{d\vec{u}}{dt} = \vec{X} - \frac{1}{\rho}\nabla P + \frac{\mu}{\rho} \quad \nabla^2 \vec{u} + \frac{\mu}{3\rho}\nabla(\nabla\cdot\vec{u}) \tag{1}$$

Continuity equation

$$\frac{d\rho}{dt} + \nabla\cdot\vec{u} = 0 \tag{2}$$

where ρ is the density of a parcel of fluid at a point, $\vec{u}$ is the vector velocity of the parcel, P is the thermodynamic pressure, $\vec{X}$ is the body force per unit mass, μ is the coefficient of viscosity (here assumed to be constant) and d/dt is the rate of change with respect to time following the motion. In regions for which viscous effects may be neglected, the relevant set of equations is:

Euler's equations

$$\frac{d\vec{u}}{dt} = \vec{X} - \frac{1}{\rho}\nabla P \tag{3}$$

Continuity equation

$$\frac{d\rho}{dt} + \nabla\cdot\vec{u} = 0 \tag{4}$$

In order to have as many equations as there are unknowns, an equation of state must also be specified. Classical fluid mechanics has concerned itself with obtaining solutions of these sets of equations, either exactly or approximately, under transformations and boundary conditions pertinent to the geometry and requirements of the problem posed.

The "no-slip" boundary condition requires that the velocity of fluid masses directly adjacent to any solid surface be the same as that of the solid surface. This fact (that the fluid right at a wall does not move relative to the wall) is derived from a preponderant number of observations of fluid phenomena under innumerable conditions; its use as a boundary condition for the differential equations of viscous flow has been justified in practically every instance. However, flow situations have been observed in which the no-slip condition becomes inappropriate, and it was these "exceptions to the rule" which prompted investigations into the basic structure of fluids and surfaces.

TUBE FLOW

Returning to the problem originally posed, we assume that the fluid meets all the requirements of the Navier-Stokes equations, in addition to which it is assumed to be incompressible (a valid assumption for liquids and slow-moving gases, but not for fast-moving gases). Application of this mathematical model of the flow yields, for the mass flow rate G;

$$G = \frac{\pi r^4 \rho (P_A - P_B)}{8\mu L} \tag{5}$$

If the fluid is a gas, we replace the density in the expression by $\tilde{\rho}$, which is the mean density. The mean density is then related to the mean pressure P through the ideal gas law:

$$P = \frac{\tilde{\rho} RT}{m} \tag{6}$$

The quantity P will be called the pressure level; the mass flow rate for a gas can then be written in the form:

$$G = \frac{\pi r^4 m (P_A - P_B)}{8\mu LRT} \; P \tag{7}$$

These expressions for the mass flow rate, called the Poiseuille formulae, are known to be valid for many conditions of flow.

Experiment with very thin capillaries has shown, however, that the mass flow rate in such tubes is often greater than that predicted by the Poiseuille relations. It is as if the fluid does not obey the no-slip condition and actually slides along the wall; this is the exception to the rule to which we previously alluded. In 1860 it was shown[1] that a slip of this kind actually does occur in liquids; and in 1875 Kundt and Warburg[2] showed that gases flowing in very thin capillaries exhibit the same behavior. If a theory is to be in accord with experiment, the assumptions upon which it is based must be valid. In this instance, the no-slip boundary condition becomes invalid and must be relaxed; in other instances, the entire set of equations may require modification or replacement, based on a new set of assumptions.

A successful theory for this "slip flow" was first given in 1879 by James Maxwell.[3] He assumed a finite slip velocity V_0 at the wall, such that the total drag force D on the fluid is proportional to the product of V_0 and the contact area between fluid and walls, as given by:

$$D = \epsilon (2\pi rL \, V_0) \tag{8}$$

where ϵ is a constant. He was then able to deduce that,

$$V_0 = \frac{(P_A - P_B) \, r}{2\epsilon L} \tag{9}$$

The mass flow rate G′ for this slip-flow is then given by:

$$G' = \frac{\pi\, r^4\, m\,(P_A - P_B)}{8\,\mu LRT}\, P \left(1 + \frac{4\mu}{\epsilon\, r}\right) \tag{10}$$

Introducing the coefficient of slip $\xi = \mu/\epsilon$, G′ may be written:

$$G' = \frac{\pi\, r^4\, m\,(P_A - P_B)}{8\,\mu LRT}\, P \left(1 + \frac{4\xi}{r}\right) \tag{11}$$

The slip at the walls thus increases the mass flow rate by the fraction $4\xi/r$. This fraction, known as the Kundt and Warburg correction, becomes important in cases of very thin capillaries for which r becomes small. In such cases, the fraction may significantly exceed unity, and the mass flow rate becomes proportional to r^3 rather than to r^4. Note that the analysis retains the Navier-Stokes equations in the interior of the fluid, but modifies conditions at the wall.

The Kundt and Warburg formula is found to agree well with experiment, as long as the pressure level of the fluid is sufficiently high; but it fails entirely for gases at very low pressures (rarefied gases?) for which it is observed that the mass flow rate becomes independent of pressure level and viscosity. We are faced again with the job of checking the assumptions to see what must be done to arrive at a theory which will be in accord with observation. In this case, a major overhaul in our approach to fluid phenomena is required.

Recall that the continuum theory of matter treats the physical parameters as continuous variables of the physical coordinates. Each system is divided into small volume elements; it is assumed that parameters like density, pressure, or temperature have constant values throughout the element, but may vary from one element to the next, according to a specific mathematical law. No assumption is made regarding the size of these volume elements except that they are small. To obtain truly continuous functions for these parameters, the volume of these elements must be permitted to approach zero without limitation. On the other hand, the discrete, coarse molecular structure of matter limits the minimum size of these elements to a volume containing a number of particles which is large enough to allow an interpretation of the macroscopic concepts (pressure, temperature, and so on). At low pressures, the size of elements containing enough molecules for such a macroscopic of phenomenological description becomes a significant fraction of characteristic volumes of the flow, and the molecules themselves become important. Realizing that we may no longer ignore the basic molecular makeup of matter, we therefore turn to matters molecular (the kinetic theory of gases).

MOLECULAR MECHANICS

The molecular theory of matter considers a solid, liquid, or gas as being composed of a large number of distinct particles called molecules. A molecule is taken to mean a regular arrangment of atoms, electrons, or ions, existing either singly or in combination. No attempt is made to explain the internal structure of molecules; rather, the molecule is usually treated as the fundamental particle of a gas, having no substructure of its own. This assumption of "no substructure" is expected to break down eventually, just as it does in continuum mechanics, and we will then have to consider the "inside" of the molecule. For now, however, consider the molecule as the basic building block.

In the gaseous state these molecules are separated by distances which are large compared to their characteristic dimensions, and they are in a continual state of motion. Someone able to see a gas on the molecular level would observe vast, almost empty space in which "dots" are whizzing around in all directions at all speeds, bumping into each other and into the walls of the container. This molecular activity is the mechanism from which all observable macroscopic phenomena derive.

Because we are not concerned with any substructure of the molecule, we ignore quantum effects and assume that each molecule obeys Newton's classical equations of motion. According to these principles, a particle will move in a straight line at constant speed unless acted upon by external forces. For convenience, these forces may be divided into:

1. Action at a distance forces. The forces of gravitational attraction and intermolecular attraction and repulsion are of this type.

2. Contact forces. The forces generated during collisions are of this type.

In those cases for which intermolecular forces are relatively weak, the molecule may be conceptualized as a rigid, elastic sphere of definite size whose mass is concentrated at its center and which is acted upon by contact forces only. For those gases in which intermolecular forces are important, the molecule is pictured as a point center of a spherically symmetric force field, possessing no definite radius. In either case, the molecule is essentially a mass point with three degrees of freedom (i.e., kinetic energy may be associated with translatory motion measured along three independent coordinates). In those cases in which the "inside" of the molecule is not ignored, one postulates some representative geometrical configuration for which more than three degrees of freedom are possible (e.g., a "dumbbell" molecule has two extra degrees of freedom, because it may exhibit rotary and vibratory motions in addition to translatory motion). It should be clearly understood that a molecule possesses only kinetic energy and (if subject to action at a distance forces) potential energy. Such terms as "heat energy" or "internal energy" as commonly used find no place in the molecular model. They are, in fact, macroscopic terms, based on macroscopic manifestations of invisible molecular energies.

The Ideal Gas

Consider now a cubical container of volume L^3 enclosing a gas as represented by a large number of rigid elastic spheres of mass m and diameter σ acted upon by contact forces only. The walls of the container are assumed to be perfectly smooth* and the conservations of momentum and energy are assumed to be valid for all processes. We seek a correlation between macroscopic and microscopic properties.

At this point, we realize that we cannot hope to catalog the history of each and every molecule as it bounces around inside the container. We turn therefore to statistics to obtain a measure of the average behavior of these molecules: Consider a small volume element $\Delta\tau$ inside the total volume of the cube. The number of molecules which find themselves inside $\Delta\tau$ at any time will fluctuate rapidly because molecules are constantly "whizzing" in and out. If we make the size of the element large enough to contain a great many

*In actuality, the coarse molecular structure of the walls is quite important. The phenomena of adsorption and desorption familiar to ultrahigh-vacuum technologists, as well as "slip flow", depend upon the molecule-surface interaction ignored here.

molecules at any time, but still small when compared to the total volume,* the fluctuating part of the molecular head-count will become a small part of the total head count at any time. Thus we compute the average number of molecules contained in this volume over a time interval Δt, which is small compared to the characteristic times of macroscopic behavior, but large compared to the time it takes a molecule to pass through the volume. Dividing this average by $\Delta\tau$, we obtain $\bar{n}$, the number density of molecules (molecules per unit volume). Multiplying $\bar{n}$ by m, the mass of a molecule, we may obtain the density ρ (mass per unit volume):

$$\rho = m\bar{n} \tag{12}$$

In this discussion we assume that there is no macroscopic motion of the gas: that is, there is no visible mass motion or flow; all motion is on the invisible molecular level. Because of this, we invoke the principle of equipartition of energy, which means generally that energy in a system is divided equally among all possible carriers. In this application, we mean that the total energy of the gas is equally divided among the three degrees of freedom. This follows from the assumption of no visible motion: since the gas is not moving on the macroscopic level, there should be no preferred directions on the molecular level. This condition is alternately called perfect chaos or perfect randomness. Note that, if there is no visible motion, the sum of all vector velocities must be zero, but that the sum of all speeds is not zero. We thus define an average speed $\bar{c}$ as the average over all of the speeds possessed by molecules passing through the volume element $\Delta\tau$ in the time interval Δt. The molecular motion is found to be completely characterized by m, σ, $\bar{n}$ and $\bar{c}$.

Going back to the cubical container, consider molecular impacts with the wall under the assumption that $\bar{n}$ and $\bar{c}$ are uniform through the cube; i.e., perfect chaos exists in the entire volume – in some problems this may be true only locally. Under these conditions, we may say that the random distribution of velocities is equivalent to a situation in which one-sixth of all molecules are travelling toward each of the six faces with a speed $\bar{c}$. In a time Δt, $\bar{n}\bar{c}L^2\Delta t/6$ molecules will impact on each face whose area is L^2. During the impact these molecules suffer a reversal of velocity, so that their momentum changes by

$$\frac{\bar{n}\bar{c}L^2\Delta t\,(-m\bar{c})}{6} - \frac{\bar{n}\bar{c}L^2\Delta t\,(m\bar{c})}{6}$$

or

$$-\frac{m\bar{n}\bar{c}^2L^2\Delta t}{3}$$

This change of momentum results from the force exerted on the molecules by the wall during the time of impact. If F_w denotes the magnitude of this force, we may define $P = F_w/L^2$ as the force per unit area of pressure exerted by the molecules on the wall. Since the rate of change of momentum with time is equal to the force acting, we obtain:

$$-\frac{m\bar{n}\bar{c}L^2\Delta t}{3} = -F_w\Delta t = PL^2\Delta t \tag{13}$$

*We should be able to do this since, for example, there are 3×10^{10} molecules in a cube whose edge is 0.01 mm (STP).

or

$$P = \frac{1}{3}\rho \bar{c}^2 \tag{14}$$

The total kinetic energy contained in the gas is

$$\frac{1}{2}\rho L^3 \bar{c}^2$$

so

$$P = \frac{\frac{2}{3}\left(\frac{1}{2}\rho L^3 \bar{c}^2\right)}{L^3} = \frac{(2/3)\text{kinetic energy}}{\text{volume}} \tag{15}$$

Pressure may thus be regarded as two-thirds of the kinetic energy of the molecules per unit volume. If we increase the volume of a vessel and maintain the same molecular speeds, the pressure decreases, according to Boyle's law. Since kinetic energies are additive, the addition of other gases of different kinetic energies increases the pressure, according to Dalton's law of partial pressures. We may obtain a measure of the average molecular speed by inverting the relation for the pressure:

$$\bar{c} = \left(\frac{3P}{\rho}\right)^{1/2} \tag{16}$$

Air at ordinary temperatures has a density of 0.00123 grams/liter at one atmosphere, and calculation yields $\bar{c}$ = 1625 ft/sec. Hydrogen at ordinary temperatures, but at a pressure of 10^{-4} atm, has a $\bar{c}$ of about 6100 ft/sec.

The distance that a molecule is able to travel between successive collisions with other molecules is called the free path ℓ. The mean free path $\bar{\ell}$ is defined as the average distance between successive collisions. With the aid of a formula presented later, mean free paths may be computed from macroscopic properties. In ordinary air, $\bar{\ell}$ is about 3×10^{-7} ft; at pressures of about $\frac{1}{2}$ mm of mercury, $\bar{\ell}$ is about 10^{-3} ft. In interstellar space, densities may be of the order of 10^{-24} gram/cc (one molecule per cubic centimeter) and $\bar{\ell}$ will be of the order of 10^{15} ft (100,000 million miles): in internebular space, where densities may be as low as 10^{-29} gram/cc (one molecule per hundred thousand cubic centimeters), the mean free path will be of the order of 10^{16} miles (1000 light years!).

If two unlike gases of the same temperature are mixed, it is observed that no temperature change occurs. Under the assumption of equipartition of energy in the mixed state, it follows that the average kinetic energy per molecule must be the same for all gases at the same temperature; i.e., in a mixture of unlike gases at the same temperature, the heavier molecules move slower than the lighter molecules, on the average. It also follows that the rate of increase of kinetic energy with temperature must likewise be the same for all gases. This leads to the conclusion that temperature is proportional to the average kinetic energy of a molecule:

$$T \propto \frac{1}{2} m \bar{c}^2 \tag{17}$$

This proportionality may be expressed by:

$$\frac{3}{2} kT = \frac{1}{2} m \bar{c}^2 \tag{18}$$

where k is the Boltzmann constant. Substitution of this into the relation for pressure yields:

$$P = \bar{n} kT \tag{19}$$

If the gas occupies a volume V, and molecular chaos prevails throughout the entire volume, we may say:

$$P = \frac{N}{V} kT \tag{20}$$

where N is the total number of molecules in V. Letting N_0 represent Avogadro's number (6×10^{23} molecules/gmol), we define the universal gas constant R by the relation:

$$R = k N_0 \tag{21}$$

The molecular weight M (grams/gmol) can be expressed as:

$$M = m N_0 \tag{22}$$

Letting n be the number of moles of gas ($n = N/N_0$), we may write, finally:

$$pV = n RT \tag{23}$$

which, of course, is the Ideal Gas Law. The important things to be gleaned from the derivation of the ideal gas law is the number of assumptions that go into it, and the corresponding number of reasons why it should not hold. The amazing thing about it is that it holds as well as it does!

Maxwell's Approach

In the model used to obtain the ideal gas law, we tacitly assumed that we could find some average speed $\bar{c}$, but we offered no way of determining it from the actual molecular velocities themselves. We now look into the more fundamental approach of Maxwell to obtain a statistical spectrum of molecular speeds: that is, we will obtain a relation which tells us what percent of the molecules will probably be moving at a specified speed.

Consider the physical space (X, Y, Z) occupied by the molecules of the gas at a specific instant t, and focus attention on the molecules contained in an arbitrary volume element dV as shown in Figure B-19. At time t there will be $\bar{n} dV$ molecules in dV, where $\bar{n}$ is the number density of molecules now considered as a function of space and time. These $\bar{n} dV$ molecules will all be moving at different velocities at time t. If we construct a coordinate system (Figure B-20) whose axes represent velocity components (u, v, w), we may plot the velocity of every molecule in dV at time t as a point in the (u, v, w) space (called the phase space). The number of points in an arbitrary volume

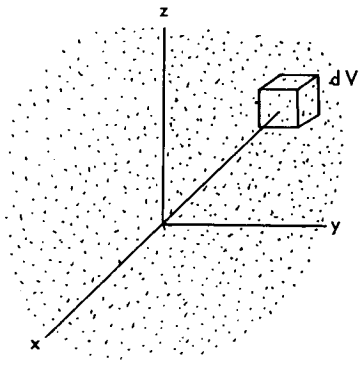


Figure B-19–Physical Space Occupied by Gas Molecules at Instant t in Volume Element dV

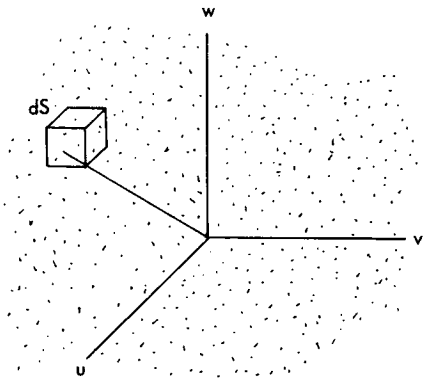


Figure B-20–Phase Space

element dS of phase space will be proportional to the number of molecules in dV, and, in general, will also depend on the position of dV in the physical space X, Y, Z, the position of dS in the phase space (u, v, w), and the time t. We may express this by saying that the number of points in a unit volume of phase space is equal to $f\bar{n}dV$ where f, called the velocity distribution function, is a function of the position of dV, the position of dS and time. The number of points in the arbitrary elements dS is then equal to $f\bar{n}dVdS$; that is, it is equal to $f\bar{n}dXdYdZd(udvdw)$.

Usually we seek the average value per molecule of some parameter F which is a function of the molecule's velocity only (e.g., average speed, kinetic energy, or momentum per molecule.) Every molecule in dV has a value of F associated with it at any time, so that there is a value of F associated with every point in phase space. Because there are $f\bar{n}dVdS$ points in dS, there will be an amount $f\bar{n}FdVdS$ contained in dS. Every point in phase space represents a molecule in dV, so the total amount of F contained in dV is the integral of $F\bar{n}fdVdS$ over every volume element of phase space. That is:

$$\int_S (F\bar{n}fdVdS) = \text{Total amount of F contained in dV.} \tag{24}$$

Since we seek the average value of F per molecule, we divide the total amount of F contained in dV by the total number of molecules in dV ($\bar{n}dV$). If F represents the average value of F per molecule, we write:

$$\bar{F} = \frac{\int_S (F\bar{n}fdV)\, dS}{\bar{n}dV} \tag{25}$$

But $\bar{n}dV$ has been defined, so that it is independent of velocity; it may therefore be taken out of the integral and divided by the denominator to yield:

$$\bar{F} = \int_S FfdS \tag{26}$$

Once the distribution of velocities f is known, it will be possible to compute all of the desired average properties of the flow. Note that the averaging process defined here is not an average over space or time but is an average over all velocities.

Removing the restriction of "no flow," we may consider the velocity of a molecule to be composed of a molecular agitation (which must average out to zero since the molecule suffers no net displacement from it), and that due to its motion as part of a macroscopic transport of molecules (mass flow). The velocity of a molecule therefore may be written:

$$\mathbf{c} = \bar{\mathbf{c}} + \mathbf{C} \tag{27}$$

where $\mathbf{c}$ is the total velocity (vector) of the molecule, $\bar{\mathbf{c}}$ is the velocity (vector) of mass motion (not to be confused with $\bar{c}$, the average speed) and $\mathbf{C}$, the molecular agitation velocity (vector). Written in component form:

$$\begin{aligned} u &= \bar{u} + U \\ v &= \bar{v} + V \\ w &= \bar{w} + W \end{aligned} \tag{28}$$

It is understood that:

$$\bar{u} = \int_S u f \, dS, \quad \bar{U} \equiv 0 \tag{29}$$

and so on for v and w. The kinetic energy E of a molecule of mass m can be written:

$$E = \frac{1}{2} m c^2 = \frac{1}{2} m (u^2 + v^2 + w^2) \tag{30}$$

The average kinetic energy per molecule $\bar{E}$ can be written:

$$\bar{E} = \frac{1}{2} m \overline{(u^2 + v^2 + w^2)} = \frac{1}{2} m (\overline{u^2} + \overline{v^2} + \overline{w^2}) \tag{31}$$

Now,

$$u = \bar{u} + U \tag{28}$$

so

$$u^2 = \bar{u}^2 + 2\bar{u} U + U^2 \tag{32}$$

and

$$\overline{u^2} = \bar{u}^2 + \overline{U^2} \tag{33}$$

and so on for v and w. Therefore the average kinetic energy per molecule may be written as:

$$\bar{E} = \frac{1}{2} m \left\{ (\bar{u}^2 + \bar{v}^2 + \bar{w}^2) + (\overline{U^2} + \overline{V^2} + \overline{W^2}) \right\} \tag{34}$$

which is the same as:

$$\overline{E} = \frac{1}{2} m \overline{c}^2 + \frac{1}{2} m \overline{C^2} \tag{35}$$

kinetic energy due to mass motion	kinetic energy due to molecular motion

Note the difference between $\overline{c}^2$ and $\overline{C^2}$: $\overline{c}^2$ is the square of the mean velocity, and $\overline{C^2}$ is the mean of a squared velocity. The latter term is sometimes called "thermal energy of a molecule" but, as has been said, it really is a kinetic energy of the molecule responsible for the macroscopic quantity. In order to determine these mean squared quantities, we must determine f, the velocity distribution function. Through a fairly involved derivation,[7] it is possible to obtain a condition on f called the Boltzmann equation, and to form equations for mass, momentum and energy transport dependent upon f. The problem of predicting macroscopic behavior from a statistical picture of molecular behavior becomes the problem of determining f as a solution of the Boltzman equation. This forms a substantial portion of the subject matter of rarefied gas flow analysis.[12, 13, 14]

Under conditions of steady state, a solution of the Boltzmann equation may be obtained in the form:

$$f(c) = \left(\frac{m}{2\pi RT}\right)^{3/2} e^{-\frac{c^2 m}{2RT}} \tag{36}$$

Given an arbitrary speed c, the fraction of molecules (f) moving at this speed can be found. This velocity distribution, called Maxwellian, can be shown to be a necessary and sufficient condition for equilibrium in a macroscopically motionless gas. For gases exhibiting mass motion, the Maxwellian distribution yields Euler's equations (no viscous effects), the isentropic energy equation (no heat conduction), and the ideal gas law. The Maxwellian distribution may be shown to yield the following relations as well:

$$\text{mean free path: } \overline{\ell} = \frac{1}{\pi \sqrt{2}\, \overline{n} \sigma^2} \tag{37}$$

$$\text{mean speed: } \overline{c} = \sqrt{\frac{8RT}{m\pi}} \tag{38}$$

$$\text{most probable speed: } \overline{c}_m = \sqrt{\frac{2RT}{m}} \tag{39}$$

$$\text{root mean square speed: } (\overline{c^2})^{1/2} = \sqrt{\frac{3RT}{m}} \tag{40}$$

We have not yet discussed the nature of the molecule-surface interaction, and it remains unspecified. The assumption of specular reflection* of molecules is not inconsistent with a Maxwellian distribution but, in fact, such reflection permits no adjustment of mass velocity or temperature at the wall. More complicated surface interactions allow an accommodation of fluid temperature and velocity to those of the wall, still within the context of Maxwellian flow.

When the symmetric point center of force molecular model is assumed, the velocity distribution deviates from Maxwellian. The transfer equations take the form of the Navier-Stokes equations and Fourier's linear law of heat conduction; the equation of state is altered by the presence of intermolecular forces, but reverts to the ideal gas law for rarefied gases. Since these results are obtained by assuming that any deviations from Maxwellian (equilibrium) flow are small, this flow is sometimes called "almost-Maxwellian". In such a flow, the following conditions are found to be valid:

coefficient of viscosity: $$\mu = \frac{1}{2}\rho\bar{\ell}\bar{c} \tag{41}$$

speed of sound: $$a = \sqrt{\frac{\pi\gamma}{8}}\,\bar{c} \tag{42}$$

where γ is the ratio of specific heats C_p/C_v.

coefficient of heat conduction: $$\lambda \propto \mu \propto T^N \tag{43}$$

and $\bar{c}$, $\bar{c}_m$ and $(\overline{c^2})^{1/2}$ have the same definitions as for Maxwellian flow.

In regions of flow in which departures from equilibrium are not small (in shock waves, for example), higher order approximations to f and hence to the transport equations have been suggested. Notable among these are the equations of Burnett (1935) and Grad (1949) documented in the references.

With this brief outline of the bases of fluid mechanics and kinetic theory, we are by no means adequately prepared to approach the subject of rarefied gas flow in depth; we proceed to it anyway, realizing that the gaps left open will serve well as guideposts of further study.

Rarefied Gas Flow

Rarefied gas flow is flow in which the length of the mean free path $\bar{\ell}$ in a gas is comparable to some significant dimension L of the flow field. The gas then does not behave entirely as a continuous fluid, but exhibits some characteristics of its coarse molecular structure. If we define the Knudsen number K as $\bar{\ell}/L$, we may say that a rarefied gas flow exists when K is not negligibly small. For some considerations, L may be a characteristic physical dimension (radius of curvature of an obstacle or the radius of an internal fluid conduit); for other considerations, L may be taken as the boundary layer thickness δ (if such a layer exists; it won't if the Reynolds number $Re = \rho UL/\mu$ is small) or the thickness of a shock transition zone. In particular, one may expect to encounter rarefied gas effects in those regions of the flow possessing very sharp gradients: i.e.,

*The angle of reflection equals the angle of incidence, as in geometrical optics.

regions in which the velocity, pressure, or temperature change appreciably in the space of a few mean free paths, regardless of whether the absolute density of the gas flow is especially low.

From the relations obtained through kinetic theory, the Knudsen number may be related to the Reynolds number and Mach number $M = U/a$ (U is a characteristic flow velocity) of the flow field. Recall that:

$$\mu = \frac{1}{2}\rho\bar{\ell}\bar{c} \tag{41}$$

and

$$a = \sqrt{\frac{\pi\gamma}{8}}\bar{c} \tag{42}$$

Now,

$$K = \bar{\ell}/L \tag{44}$$

$$= \gamma\sqrt{\frac{\pi}{2}}\frac{\mu}{\rho a L} \tag{45}$$

$$= \gamma\sqrt{\frac{\pi}{2}}\frac{\mu}{\rho U L}\cdot\frac{U}{a} \tag{46}$$

so that

$$K = \gamma\sqrt{\frac{\pi}{2}}\frac{M}{Re} \tag{47}$$

where K and Re are based on the same characteristic length L.

Regimes of Flow

The division of gas dynamics into various regimes based on characteristic ranges of values of the appropriate Knudsen number is convenient in analyzing flow phenomena. Continuum Flow, Slip Flow, Transition Flow and Free Moelcule Flow are defined so that they correspond to flow in which, roughly speaking, the density levels are respectively ordinary, slightly rarefied, moderately rarefied, and highly rarefied.

For flows of high Reynolds number ($Re >> 1$), a boundary layer is formed and the characteristic length is δ. Since it is known that:

$$\frac{\delta}{L} \sim \frac{1}{\sqrt{Re}} \tag{48}$$

we may obtain:

$$K \sim \frac{M}{\sqrt{Re}} \tag{49}$$

Hence, for $Re >> 1$, ordinary gas dynamics (continuum theory) prevails for

$$K \sim \frac{M}{\sqrt{Re}} << 1 \tag{50}$$

For flows of low Reynolds number $(Re << 1)$, we get Stokes or "creeping flow". Here no boundary layer is formed, and L is the only characteristic length. Then,

$$K \sim \frac{M}{Re} \tag{51}$$

and for

$$K \sim \frac{M}{Re} << 1 \tag{52}$$

ordinary gas dynamics (low-speed continuum mechanics) prevails.

For flows in which the value of the appropriate K is small but not negligible, a departure from classical continuum mechanics can be expected. Such an effect is the phenomenon of "slip" which has been noted. The change from ordinary continuum gas dynamics to this region is of course gradual, but we may use experimental evidence to define the limits of slip flow as:

$$0.01 < \frac{M}{\sqrt{Re}} < 0.1 \qquad \text{when } Re > 1 \tag{53}$$

$$0.01 < \frac{M}{Re} < 0.1 \qquad \text{when } Re < 1 \tag{54}$$

In the slip flow regime so defined, the mean free path is of the order of one to ten percent of δ or L and slip flow effects may be expected also to be of this order. It should be noted that in the slip flow regime either Re must be very small (strong viscous effects) or M very large (strong compressibility effects). Thus true rarefied gas flow effects such as slip occur only in coincidence with either strong viscous or compressibility effects or both. In the slip flow regime as defined above these latter phenomena very often dominate rarefaction effects associated with the coarse molecular structure of a gas and really large scale deviations from continuum behavior are not apparent until the "transition" regime, defined below, is reached. It might also be noted that in the hypersonic range, δ is not given by $L/\sqrt{R_e}$, so modifications in K must be made.

For extremely rarefied flow, $\bar{\ell} >> L$. Under these conditions no boundary layer is formed. Molecules returning to the gas after a surface impact do not collide with free stream molecules until they are far away from the body. One can then neglect the disturbance of the free stream velocity distribution due to the presence of the body. Here the flow phenomena are mostly governed by molecule-surface interaction and the regime is called "free-molecule" flow. (A molecule travels about virtually unhindered by other molecules and so is "free".) It may be defined on the basis of experimental evidence by the requirement:

$$\frac{M}{R_e} > 3 \tag{55}$$

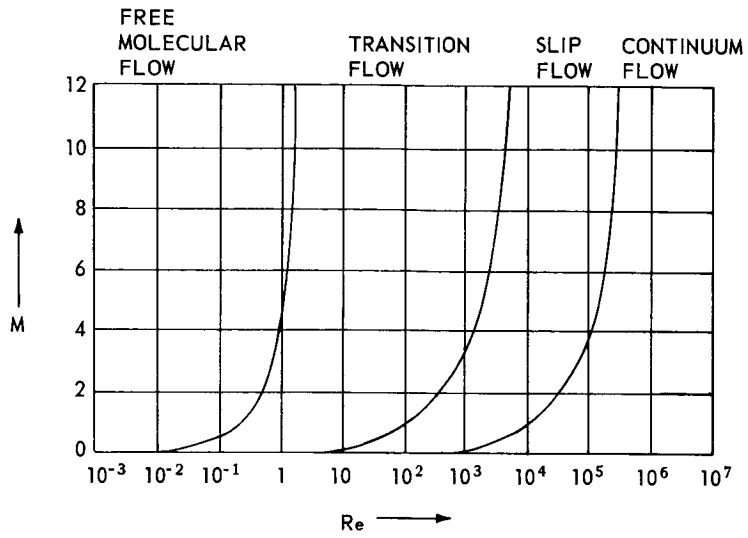


Figure B-21–Regimes of Rarefied Gas Flow

In the transition regime between slip flow and free-molecule flow, $\bar{\ell}$ is of the same order as a typical body dimension. Surface collisions and free stream intermolecular collisions are of more or less equal importance and the analysis becomes very complicated. A chart (Figure B-21) from Reference 6 may be of use in recognizing the regimes of rarefied flow.

The main usefulness of Reynolds and Knudsen numbers does not lie in their absolute values but in their use as scaling factors. For example, if we consider the flow of gas through an orifice of diameter d at a certain pressure level p, the flow regime prevailing will be unchanged if we change d and p while maintaining the Knudsen number constant; that is, an increase in p accompanied by a proportionate decrease in d retains the original Knudsen number. Therefore low density behavior may exist in a gas at very low pressure with respect to an orifice of large diameter and also in a gas at atmospheric pressure with respect to extremely fine pores or holes.[5]

Free Molecule Flow

In this regime of extremely low density, the mean free path is much larger than any characteristic length of the flow region, and intermolecular forces of either type are insignificant. Molecules moving about interact with the boundary walls long before they have a chance to collide with each other. A molecule interacting with the wall may be reemitted (i.e., it returns to the gas) or it may be adsorbed (i.e., it doesn't return to the gas). Reemitted molecules interact very little with incident molecules (those heading for the wall) and, as a result, the incident stream of molecules has no way of knowing that the wall lies in its path. In fact, the basic assumption of free molecule theory is that the incident stream of molecules is entirely undisturbed by the presence of the solid wall; hence, no shock waves or boundary layers are expected to form in the vicinity of an object. The molecular motion remains Maxwellian (undisturbed from equilibrium); viscous phenomena inside the fluid disappear (velocity gradients are not present); and heat conduction inside the fluid vanishes (temperature gradients are not present). The theories developed for free molecule flow are concerned with the all-important mechanism of the molecule-surface interaction, which emerges as the governing factor of the flow.

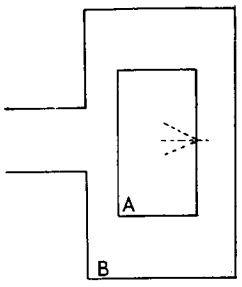


Figure B-22–Effusion

Before returning to the problem of gas flow in tubes under extremely rarefied conditions, consider the effusion of free molecules through an orifice. In Figure B-22, vessel A contains gas at pressure p_A surrounded by a vacuum envelope B whose very low pressure ($p_B \simeq 0$) is maintained by pumps. One of the walls of A is equipped with a circular orifice of diameter d, and the wall thickness is very small compared to d. If the gas pressure p_A or the orifice diameter d is so small that $K = \bar{\ell}/d$ is large, a molecule from the inside of vessel A moving toward the orifice will normally pass through without collisions. The rate of flow through the orifice can be found by computing the number of molecules which impinge upon the area $S = \pi d^2/4$ per unit time, under the assumption that the molecules obey the Maxwellian velocity distribution corresponding to the temperature T_A. It can be shown[5] that $\frac{1}{4}\bar{n}\bar{c}$ molecules impinge on a unit area of the wall in unit time, so that the mass flow rate G through the orifice is given by:

$$G = \frac{1}{4} m \bar{n} \bar{c} S \tag{56}$$

Now, since

$$p = \bar{n} k T, \quad k = R/N_0, \quad \bar{c} = \sqrt{\frac{8RT}{\pi m}}$$

we may write:

$$G = \frac{p_A m N_0 S}{\sqrt{2\pi RT/m}} \tag{57}$$

Now reconsider the gas flow in the tube (Figure B-18) under the conditions that p_A and p_B are both very small (i.e., the pressure level p is low). A molecule entering from the A side will strike the inner wall of the tube, except in the rare case when its velocity is parallel to the tube and it is able to pass through without collisions. The progress of an arbitrary molecule thus depends on its interaction with the wall. If the impact between the molecule and the wall is perfectly elastic, every molecule will proceed along a zigzag path through the tube, maintaining always its component of velocity parallel to the axis of the tube. Since for large Knudsen numbers ($\bar{\ell} >> r$) the number of molecules leaving A per unit time is the same as if the tube were replaced by an orifice, and since all of these molecules would pass through the tube, the rate of flow would be equal to that obtained for effusion through an orifice as obtained above, diminished by the flow of molecules from B to A (which, under free molecule conditions, is independent of the flow from A to B). Experiments have shown, however, that the rate of flow is much smaller, and that it depends inversely on the length of the tube. The assumption of specular reflection (elastic collision) is found to be an oversimplification; a more detailed explanation is therefore required.

In 1909, Martin Knudsen assumed that a molecule impinging on a surface may not be reflected like an elastic body, but may be scattered from the surface according to a cosine law (i.e., the probability of its leaving the surface in a direction having an angle θ with the normal is proportional to $\cos\theta$ and independent of the angle of incidence). As a result, the probability of a molecule's continuing in the forward direction after an impact is equal to its probability of continuing in the backward direction, velocities being relative to the wall. A fraction of molecules which enter the tube will therefore return to A, and the number of molecules which succeed in passing through the whole length of the tube

becomes considerably smaller than the number entering and decreases with the length of the tube. The same process will occur for molecules going from B to A. Knudsen determined that the mass flow rate would then be given by:

$$G' = \frac{8\pi r^3 (p_A - p_B)}{3L \sqrt{2\pi RT/m}} \tag{58}$$

where $L >> r$.

It is seen that for extremely rarefied gas flow (free molecule flow) the mass flow rate through a tube is independent of the pressure level; viscosity, of course, plays no part in this regime.

An interesting phenomenon of free molecule flow is that of thermal transpiration. Assume that the tube we have been discussing links vessels A and B containing gas A and gas B respectively. Under ordinary conditions in which classical continuum theory is valid, a condition for equilibrium is that the total pressures in A and B be equal. If the gases in A and B are different and have the initial pressures p_A and p_B, gas will flow from A to B (assuming p_A exceeds p_B) until the pressures are equal, regardless of the composition of the resulting gas mixtures or the temperatures in A or B. It is true that the composition and temperatures in A and B will ultimately equalize by diffusion and heat conduction, but these processes are so slow compared to equilibration of the total pressures that they generally can be disregarded. If, however, the Knudsen number is large, the flows from A to B and from B to A will be independent, and the net mass flow rate will be given by Eq. (58). If the temperatures in A and B are different, however, this will no longer be the case. The flow from A to B will be given by:

$$G''_{A\to B} = \frac{r^3}{3L}\sqrt{\frac{32\pi m}{RT_A}}\, p_A \tag{59}$$

and the flow from B to A will be given by:

$$G''_{B\to A} = \frac{r^3}{3L}\sqrt{\frac{32\pi m}{RT_B}}\, p_B \tag{60}$$

Equilibrium will be reached when these flow rates are equal; this requires that:

$$p_A = p_B\sqrt{\frac{T_A}{T_B}} \tag{61}$$

and not $p_A = p_B$. This interesting result means that, even if initially the pressures were equal but the temperatures unequal, a flow would ensue until the pressure satisfied Eq. (61). Similarly, if the system is originally in equilibrium (temperatures and pressures the same) and the temperature in one of the parts is changed, gas flow will result. First pointed out by Knudsen in 1910, this phenomenon, called thermal transpiration or thermal molecular flow, is of importance in high-vacuum pressure technology. A similar phenomenon, observed if the gases are of unlike composition, is used in isotope separation.[5]

Transport of energy and momentum in free molecule flow can take place only during surface interactions, because the mechanisms of viscosity and heat conduction are not present inside the gas. If one assumes that molecular interaction with the wall is purely

specular, one removes even the possibility of viscous stress and heat transfer at the wall, because molecules reflected in this way neither gain nor lose energy (hence no heat transfer) and exchange momentum only in the normal direction (pressure forces are present but not viscous forces). However, it is known that heat transfer and viscous stresses are present at the wall in a free molecule flow, and one must therefore assume that molecular reflection from walls is of a diffuse or scattered nature. (Recall the cosine law, for example.) The model of free molecule flow in conjunction with diffuse reflection has been used as a first approximation to high-altitude aerodynamic problems.[12, 13, 14]

In heat transfer studies of rarefied gases for which the wall is at a temperature T_w, the incident gas molecules at a temperature T_i (with corresponding energy level E_i), and the reemitted molecules at a temperature T_r (with corresponding energy level E_i), it is convenient to define the coefficient of accommodation α as a measure of energy transfer:

$$\alpha = \frac{E_i - E_r}{E_i - E_w} \tag{62}$$

E_w is the energy a molecule would have if it accommodated completely to the wall temperature. In the case of pure specular reflection, $\alpha = 0$ (i.e., there is no accommodation). Generally α has a value less than unity but greater than zero; there is then a temperature jump at the wall, as well as a velocity jump (slip). Both of these depend ultimately upon the mechanism of interaction which is now only poorly understood but hopefully studied.

In the discussion of free molecule flow, we completely neglect the effects of intermolecular influence due either to intermolecular forces or to intermolecular collisions. With increasing density, these effects begin to assume importance, and the low density part of the transition regime is entered. In this regime, surface interaction and intermolecular interaction occur simultaneously in a most complex way. We therefore defer discussion of the transition regime and consider next the slip flow regime about which much more is known.

Slip Flow

The flow regime referred to as slip includes those flows for which the gas density is just slightly less than that associated with a continuum flow. We encountered the phenomenon of slip when we first surveyed the flow of gases through tubes, and we outlined Maxwell's treatment of the problem at that time. We now consider the molecular behavior of a gas at a wall, and, through kinetic theory, arrive at a better understanding of slip, after which we apply it to tube flow.

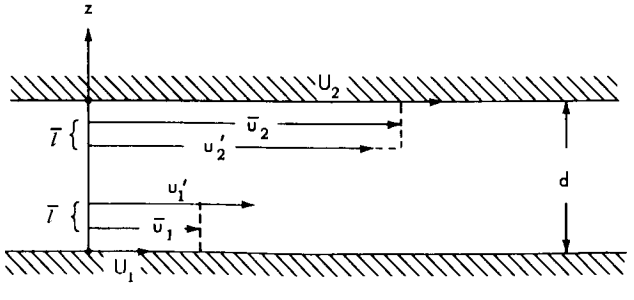


Figure B-23–Slip Flow Between Parallel Plates

Consider the steady flow between two infinite flat plates separated by a distance d (Figure B-23) and moving parallel to each other with velocities U_1 and U_2. We confine attention to the molecules in the layer of thickness $\bar{\ell}$ immediately adjacent to the wall. The average parallel component of molecular velocity at the distance $\bar{\ell}$ from the lower wall is u_1', so that all molecules approaching the wall in this layer have, on the average, a parallel component of velocity u_1'. (Note that the last collision suffered by an incident molecule was not within this layer.) We call to mind the cosine law of diffuse reflection, and assume that the fraction f of incident molecules interact with the wall according to this law. The remaining fraction (1-f) is assumed to reflect specularly. After N molecules having interacted with the wall, Nf of them will have a parallel component of velocity U_1 and $N(1-f)$ of them will have a parallel component of velocity u_1'. (Figure B-24) The layer contains an equal number of incident and reflected molecules because of the steady-state hypothesis, so the average velocity component parallel to the wall $\bar{u}_1$ in the layer will be given by:

$$\bar{u}_1 = \frac{1}{2} \{u_1' + fU_1 + (1 - f)\, u_1'\} \tag{63}$$

or

$$\bar{u}_1 = \left(1 - \frac{f}{2}\right) u_1' + \frac{f}{2} U_1 \tag{64}$$

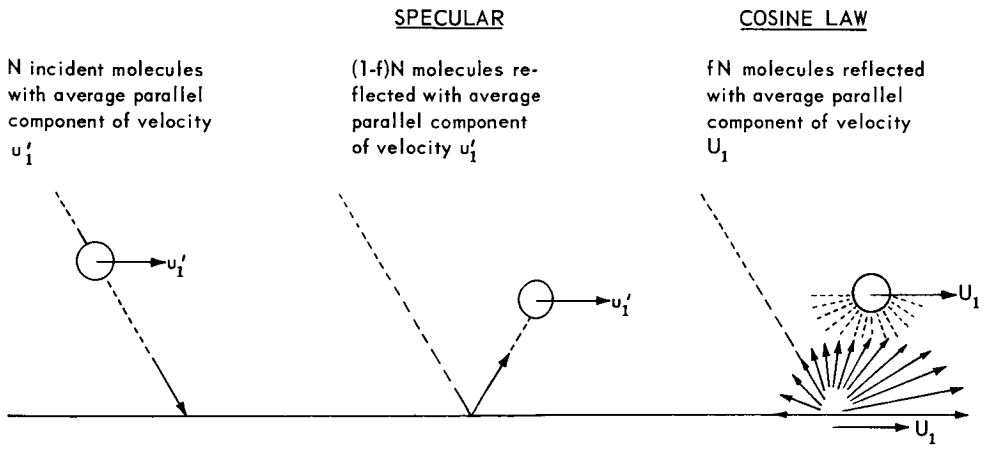


Figure B-24–Specular and Diffuse Molecular Reflection at a Surface

Since we're interested in the drag of the wall on the fluid, we must know the momentum of an average molecule before collision and its momentum after collision, the difference of which we will equate to the impulse of the drag force at the wall. Before collison a molecule has a parallel component of momentum equal to $m\, u_1'$ and after collision its momentum is given by:

$$m\, \{(1 - f)\, u_1' + fU_1\}$$

The difference of these momenta is $mf\,(u_1' - U_1)$. Since there are $\frac{1}{4} \bar{n}\, \bar{c}$ impacts per unit area per unit time, the total momentum transfer at the wall is:

$$\frac{1}{4} m \bar{n} \bar{c} f (u_1' - U_1)$$

Maxwell postulated an external frictional force per unit area, F_F, equal to $\epsilon(\bar{u}_1 - U_1)$; where ϵ has been already introduced and is called the coefficient of external friction. Equating the external frictional force per unit area to the transfer of momentum per unit area, we may obtain:

$$F_F = \frac{1}{4} m \bar{n} \bar{c} f (u_1' - U_1) \tag{65}$$

and recalling the definition of $\bar{u}_1$, we may also write:

$$F_F = \epsilon \left(1 - \frac{f}{2}\right) (u_1' - U_1) \tag{66}$$

Substitution then yields:

$$\epsilon = \frac{m \bar{n} \bar{c}}{2} \left(\frac{f}{2 - f}\right) \tag{67}$$

To put this into more useful form, recall that

$$\mu = \frac{1}{2} m \bar{n} \bar{c} \bar{\ell} \tag{41}$$

so that

$$\epsilon = \frac{\mu}{\bar{\ell}} \left(\frac{f}{2 - f}\right) \tag{68}$$

The coefficient of slip defined for tube flow may then be written:

$$\xi = \bar{\ell} \left(\frac{2 - f}{f}\right) \tag{69}$$

and the expression for the mass flow rate (Eq. 11) takes the form:

$$G' = \frac{\pi r^4 m (p_A - p_B) p}{8 \mu L R T} \left\{1 + 4 \left(\frac{\bar{\ell}}{r}\right)\left(\frac{2 - f}{f}\right)\right\} \tag{70}$$

We easily see the direct dependence of slip flow phenomena upon the magnitude of the mean free path. Also note the emergence of the Knudsen number $\bar{\ell}/r$ as the significant nondimensional parameter of the flow, and its value in predicting the contribution of slip to the overall flow.

If we now make the simplifying assumption that all molecules interact according to the cosine law, $f = 1$ and the expression for the mass flow rate reduces to:

$$G' = \frac{\pi r^4 m (p_A - p_B) p}{8\mu LRT} \left\{1 + \frac{4\bar{\ell}}{r}\right\} \tag{71}$$

Substitution of

$$\bar{\ell} = \frac{\mu}{p\sqrt{\frac{2m}{\pi RT}}} \tag{72}$$

from the relations:

$$\bar{\ell} = \frac{2\mu}{\rho\bar{c}}, \quad \bar{c} = \sqrt{\frac{8RT}{\pi m}}, \quad p = \rho\frac{RT}{m}$$

yields an expression for G' in terms of macroscopic parameters:

$$G' = \frac{\pi r^4 m (p_A - p_B) p}{8\mu LRT} \left\{1 + \frac{4\mu\sqrt{\frac{RT}{2m}}}{rp}\right\} \tag{73}$$

in which the dependence of the slip effect upon temperature and pressure is clearly seen. As the pressure level gets lower and lower, the correction term assumes dominance and, in the limit of vanishing pressure, the mass flow rate will approach:

$$\lim_{p\to 0} G' = \frac{\pi r^3 (p_A - p_B)}{L\sqrt{\frac{8RT}{\pi m}}} \tag{74}$$

This formula has all the essential features of flow encountered in the free molecule ranges, but is not in agreement with the expression for the mass flow rate G'' as obtained from free molecule theory (Eq. 58):

$$G'' = \frac{8\pi r^3 (p_A - p_B)}{3L\sqrt{2\pi RT/m}}$$

being smaller by the factor $3\pi/16$. Although there is no rigorous reason to expect that the limiting value of G' is G'', it would have been satisfying. Actually the trouble lies in the transition region intermediate to slip flow and free molecule flow. If we draw a graph of $G_p = G/(p_A - p_B)$ versus p (mass flow per unit pressure drop versus pressure level) as predicted by the theories we have developed, it would resemble Figure B-25 in which the ambiguity concerning the transition region is plainly seen. As the behavior of the curve cannot be obtained from the foregoing analysis, we must call upon experiment to give us a clue. Knudsen did this about fifty years ago and discovered that the curve goes through a minimum in this region, after which it links smoothly with the constant value predicted by free molecule theory. This led to a theoretical investigation from which an expression for the entire curve could be written in the form:

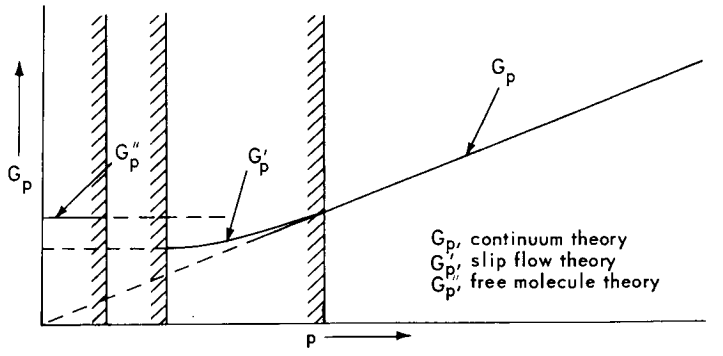


Figure B-25–Theories for Mass Flow Rate of Gas Through Tubes

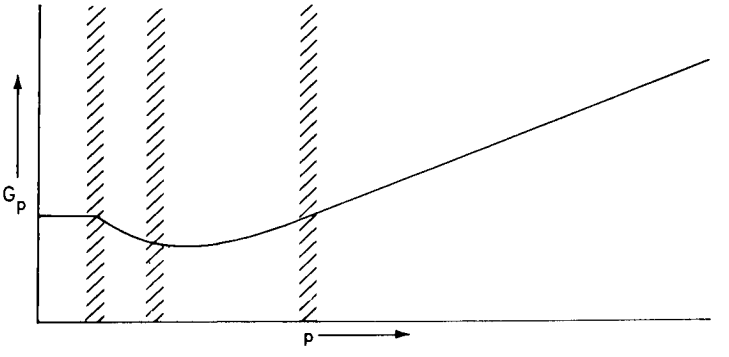


Figure B-26–Graph of Mass Flow Rate of Gas Through Tubes

$$G_p = \frac{\pi r^4 mp}{8\mu LRT} + \frac{r^3}{3L}\sqrt{\frac{32\pi m}{RT}} \left\{ \frac{1 + \frac{2r}{\mu}\sqrt{\frac{m}{RT}}\, p}{1 + \frac{2.47\, r}{\mu}\sqrt{\frac{m}{RT}}\, p} \right\} \qquad (75)$$

and whose graph resembles Figure B-26.

Because of present interest in high-altitude and space flight, much experimental and theoretical work is being done on rarefied gas flow in general and slip flow in particular. In aerodynamic slip flow, strong compressibility and viscosity effects are always present, and they often serve to mask slip flow effects. In those cases in which slip flow effects are not hidden, they are substantially influenced by interactions between compressibility and viscous phenomena (such as the interaction between a viscous boundary layer and a supersonic inviscid free stream). Because of these complex interactions, few, if any, solutions are available which properly handle viscosity, compressibility, and rarefaction effects simultaneously. In fact, the formulation of the basic differential equations and boundary conditions governing the slip flow regime has not been agreed upon. Recent experimental and theoretical work has suggested that the Navier-Stokes equations,

together with the velocity and temperature-jump boundary conditions, are superior to higher order equations (such as the Burnett equations or Grad's "thirteen-moment" equations) in the slip flow regime; their utility in the transition regime is still in doubt.

APPLICATIONS

Current studies of rarefied gas behavior fall into categories corresponding to the regimes of flow outlined above. Problems in ultrahigh-vacuum technology, for example, are generally in the free molecule regime, while high-altitude aerodynamics is largely concerned with the regimes of slip and transition, free molecule flow becoming important at altitudes of seventy-five miles or more.

A method capable of producing ultrahigh vacuum is that of cryosorption, which uses a low-temperature surface to condense the gases in a vacuum chamber and thereby reduce the pressure. The surfaces are cooled by liquefied gases such as helium. At levels of very high vacuum, such factors as the rate at which molecules are adsorbed and desorbed from the surface, how many molecules remain at the surface, and for how long, are crucial in improving the efficiency of the system when considered as a vacuum pump. One way of studying these important surface interactions utilizes molecular beams impinging on surfaces under controlled conditions. A molecular beam is a stream of molecules constrained to move in almost parallel trajectories in a highly evacuated space.[5]

In studying the effects of the upper atmosphere on the characteristics of aircraft, the high-altitude environment is simulated by low-density supersonic wind tunnels. The usual methods of measurement and flow visualization used in ordinary wind tunnels must be abandoned, however; Schlieren and interferometric flow visualization techniques, so useful at normal densities, are no longer practicable because of the low optical density of rarefied gases. These methods have been superseded by new techniques, among which are electron density probes and afterglow techniques detailed in Reference 12.

REFERENCES

1. Helmholtz & Piotrowski, Wien. Sitzunger. 40 (1860) p. 607

2. Kundt & Warburg, Pogg. Annalen, 155 (1875) p. 337

3. Maxwell, J.C., Collected Works, 2, p. 708

4. Jeans, Sir James, An Introduction to the Kinetic Theory of Gases, Cambridge University Press, London (1940) (A superior conceptual introduction to the subject)

5. Estermann, I., In Thermodynamics and Physics of Matter (F.D. Rossini, ed.), Vol. I, Sec. I, Princeton Series on High Speed Aerodynamics and Jet Propulsion, Princeton University Press, Princeton, New Jersey (1955) (An excellent section on rarefied gas flow problems of a hydraulic rather than aerodynamic nature)

6. Schaaf, S., & Chambre, P., In Fundamentals of Gas Dynamics, (H.W. Emmons, ed.), Vol. III, Sec. H, Princeton Series on High Speed Aerodynamics and Jet Propulsion, Princeton University Press, Princeton, New Jersey (1958) (Rarefied gas flow problems from an aerodynamic point of view)

7. Patterson, G. Molecular Flow of Gases, John Wiley & Sons, New York (1956) (A concise, rigorous mathematical treatment of its subject)

8. Dushman, S., Scientific Foundations of Vacuum Technique, John Wiley & Sons, New York (1962) (An engineering approach to rarefied gas problems in vacuum systems)

9. Loeb, L.B., The Kinetic Theory of Gases, 2d ed., McGraw-Hill Book Co., New York (1934) (One of several comprehensive treatments of the subject)

10. Chapman, S., and Cowling, T.G., The Mathematical Theory of Non-Uniform Gases, 2d ed., Cambridge University Press, London (1952) (This book and the following one are bibles of transport phenomena.)

11. Hirschfelder, J.O., Curtiss, C.F., and Bird, R.B., Molecular Theory of Gases and Liquids, John Wiley & Sons (1954)

12. Devienne, F.M., Rarefied Gas Dynamics, Proceedings of the First International Symposium, Pergamon Press, New York (1960)

13. Talbot, L., Rarefied Gas Dynamics, Proceedings of the Second International Symposium, Academic Press, New York (1961)

14. Laurmann, J.A., Rarefied Gas Dynamics, 2 vol., Proceedings of the Third International Symposium, Academic Press, New York (1963) (These Proceedings contain most of the important papers presented in the last few years on the subject of rarefied gas flow.)

INTERPRETATION OF VACUUM-GAUGE DATA FROM EXPLORER XVII

Donald R. Hardesty, S. B. Pearce and G. P. Newton

INTRODUCTION

The Explorer XVII, an aeronomy satellite, was designed to obtain measurements of the earth's local atmospheric structure parameters: neutral particle density, composition, and electron density and temperature. The satellite orbit had the following characteristic parameters:

orbital inclination	58°
perigee altitude	256 km.
initial apogee altitude	920 km.
orbital period	90 min.

The satellite was spin stabilized and had a rotational period of 0.67 sec.

During its 100 day active lifetime, the satellite was activated for four-minute intervals by command as it passed over the Minitrack ground stations. During each active cycle measurements were made by the satellite instruments and the resulting data were transmitted by a PCM telemetry system to the Minitrack stations. Inactive periods of the satellite varied from several minutes to one day.

The satellite carried four ionization vacuum gauges: two modified Bayard-Alpert hot filament gauges and two modified Redhead cold-cathode magnetron gauges. Figure B-27 is a photograph of both types of gauges. A more complete description of the

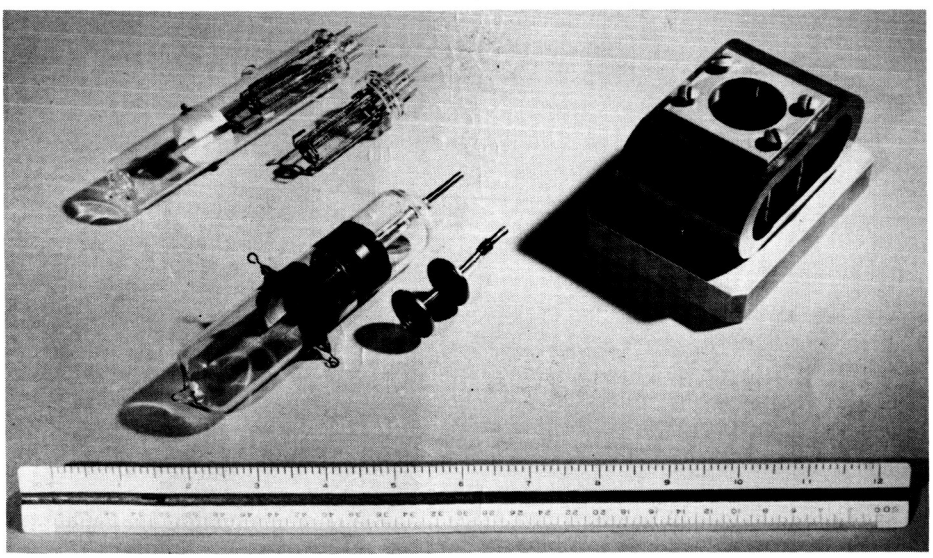


Figure B-27–Flight Gauges: Bayard-Alpert (top), Redhead (bottom)

experimental systems may be found in references 3 and 5. One gauge of each type was located on the satellite's spin equator and the remaining two gauges were symmetrically placed off the equator. The analysis below is restricted to the data from the two equatorial gauges, in order to compare the two types of gauges under similar environmental conditions.

GAUGE PRESSURE MODULATION

A gauge mounted on the equator of a spinning satellite experiences a pressure modulation because the satellite velocity ($\approx$ 8 km/sec) is greater than the most probable thermal velocity of the atmospheric gas molecules ($\approx$ 1 km/sec). When the gauge-opening faces in the direction of the satellite velocity, the pressure in the gauge is greater than the ambient because the gauge is overtaking atmospheric particles and hence essentially experiences an 8 km/sec neutral beam. Correspondingly, when the gauge faces opposite to the direction of motion, the gauge pressure is less than the ambient because very few atmospheric particles can overtake and enter the gauge.

Amplitudes of the spin modulation of pressure representing variations of a factor of 30 over the 340 msec. (1/2 spin) period have been measured for the Explorer XVII. Theory, however, predicts a 21 decade variation. The discrepency is attributed primarily to surface phenomena. The atmospheric density is proportional to the peak to peak pressure over a satellite spin cycle. Thus when there is a large (i.e. a factor of 30) maximum-to-minimum pressure ratio the measured density is only slightly (3%) affected by the discrepency between the measured and theoretical pressure ratio. However, an understanding of the causes for the observed ratio is necessary in order to improve the atmospheric density measurement technique. The purpose of this paper is to describe and define the phenomena causing the measured gauge response, particularly in the rarefaction region of the spin cycle.

ATMOSPHERIC COMPOSITION

Due to dissociation of molecular oxygen primarily by solar radiation, atomic oxygen is the predominent atmospheric constituent over a large percentage of the altitudes covered by the satellite. Because of the high affinity for molecular association exhibited by the atomic species, it is difficult to retain atomic oxygen in controlled laboratory experiments. Thus, an accurate laboratory calibration of the gauge's response to atomic oxygen is not possible. The flight data were examined for departures from expected behavior based upon molecular oxygen laboratory data. No deviation was observed that could be attributed to atomic oxygen.

PRELIMINARY ANALYSIS

The data, as explained before, consists of pressure time curves. The gauge pressure is modulated at the spin rate of the satellite. The deviation of the pressure from the theoretical value is largest in the rarefaction region of the spin cycle. In this region an apparent exponential structure can be recognized in the pressure curves. The preliminary analysis was designed to assign approximate time constants to the pressure curves and to establish the time constant repeatability.

Since approximately 15 million data points were obtained during the active lifetime of the satellite, it was necessary to establish criteria for the selection of data. The criteria used were:

1. The data should exhibit a large ratio of ram pressure to rarefaction pressure
2. The data must have been analyzed previously to obtain atmospheric densities. This insures an evaluation of the S1N, the satellite attitude, the data quality.
3. Certain passes were selected on the basis of orbit number, pressure regime and satellite attitude. This allowed an examination of the possible effect on the time constants of time in orbit, pressure, and composition.

Initially sixteen passes were selected: six at high altitudes (pressure range 10^{-9} to 10^{-10} torr), and ten at low altitudes (pressure range 10^{-6} to 10^{-7} torr).

Figure B-28 shows a typical pressure-time characteristic for the Bayard-Alpert gauge. Exponential curves were matched to the raw data (without subtracting the residual pressure) in the region 90° past the maximum ram position near the beginning of rarefaction. Time constants (the time required for the pressure to decay to 1/e of the initial value chosen) for the Bayard-Alpert curves ranged from 80 to 110 milliseconds, considerably longer than predicted if only the volume and conductance of the gauge are considered in computing the time constant. The existence of a minimum residual pressure observed in the Bayard-Alpert gauge response and the nearly exponential pressure decay are indicative of sorption phenomena. The repeatability of the Bayard-Alpert spin cycle time constants is within ±20%.

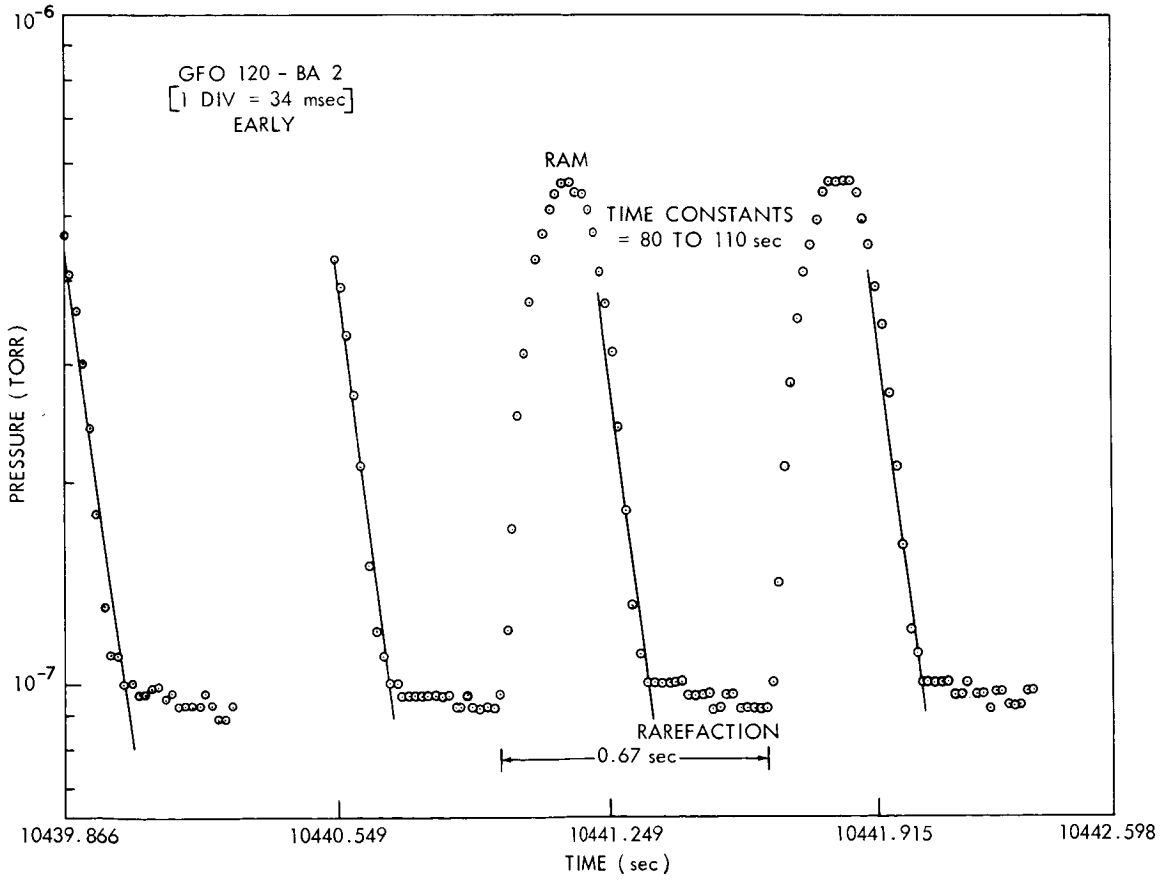


Figure B-28–Plot of Bayard-Alpert Gauge, Pressure vs. Time

Figure B-29 shows a pressure-time plot for the Redhead gauge. Analysis of the Redhead gauge data was carried out in a manner similar to that described above. Time constants in this case exhibited a repeatability within ±30% throughout each pass and over the satellite lifetime.

The time constants for both gauges increase with altitude while the low altitude time constants show more repeatability over the lifetime of the satellite.

In the Bayard-Alpert gauge, there was observed a limiting residual pressure which appeared to decay slowly over each active period.

ADDITIONAL CORRELATIONS AND THEORETICAL INVESTIGATIONS

An investigation was made of possible correlations of the observed gauge responses with the characteristic behavior of the following:

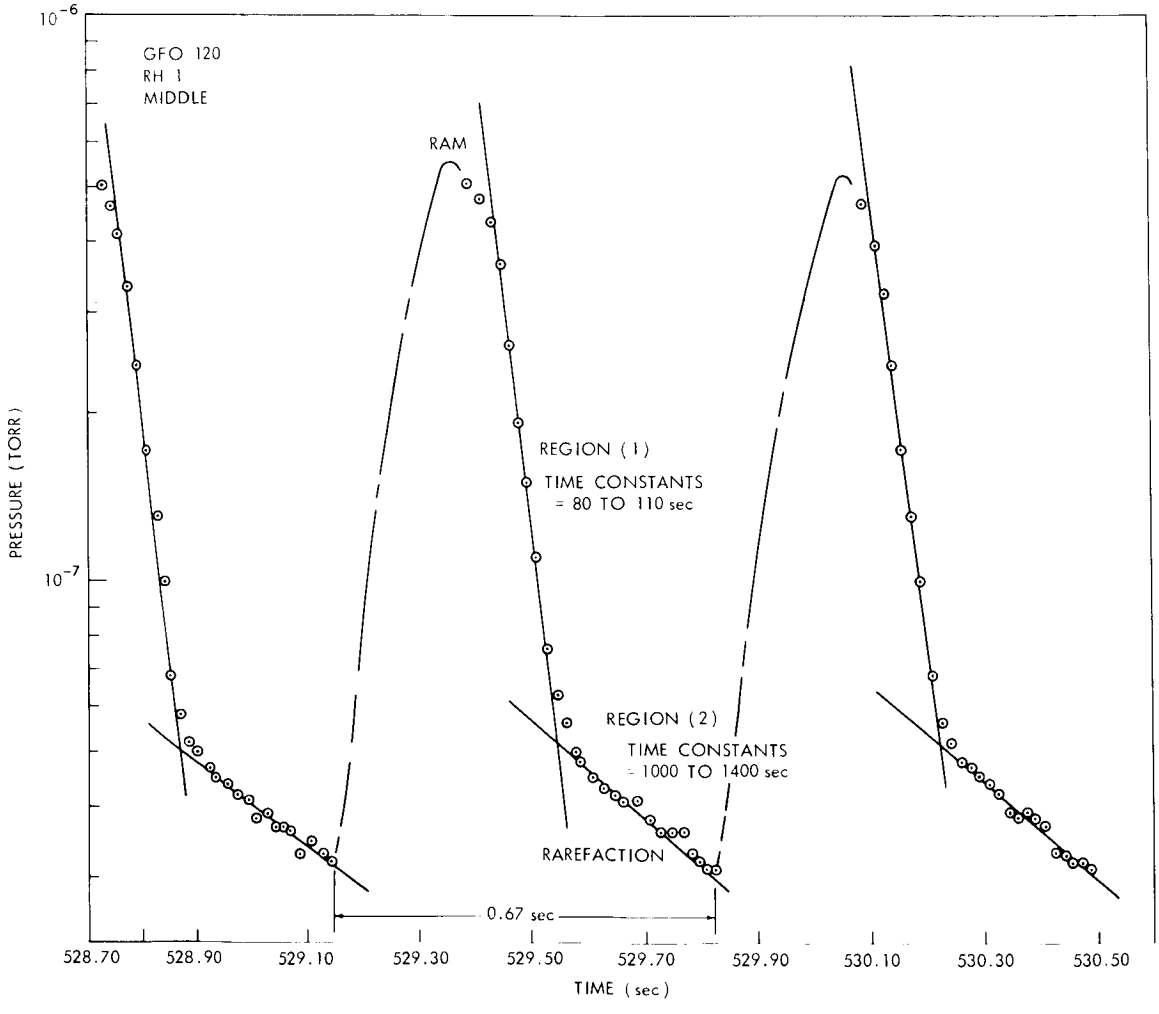


Figure B-29–Plot of Redhead Gauge, Pressure vs. Time

Redhead gauge –

1. Dependence of sorption phenomena on temperature fluctuations of the gauge envelope as reflected in the outgassing rates.

2. Time response of the gauge as limited by the structural configuration of the gauge.

Bayard-Alpert gauge –

1. Correlation of time constants with additional environmental parameters such as satellite orientation, length of preceding inactive period, altitudes during inactive periods.

2. Surface-ionization-determined residual of Bayard-Alpert gauges.

REDHEAD GAUGE TEMPERATURE EFFECTS

A correlation of observed gauge time constants with gauge temperature was attempted since physical desorption of gas from a surface is strongly temperature dependent. This was restricted to the Redhead gauge since the Bayard-Alpert gauge hot filament introduces analytical difficulty.

An examination of the satellite gas temperature over the satellite lifetime revealed a variation in temperature of only 15°C. While the satellite skin temperature variation was greater, the temperature of the gas contained in the satellite is believed to dominate the gauge temperature. Since the observed 15°C temperature variation was too small for a meaningful correlation, this study was abandoned.

THEORETICAL TIME RESPONSE OF THE REDHEAD GAUGE

The theoretical time response to pressure variations was calculated for the Redhead gauge. A diagram and pertinent dimensions of the Redhead gauge used in Explorer XVII are shown in Figure B-30(a). The figure (b) also presents a simplified schematic diagram of the gauge with volumes and conductances labelled.

Since standard vacuum system equations were applied, it was necessary to calculate the pertinent volumes and conductances.

Volumes (V_1, V_2, and V_3) were computed for short cylindrical tubes. The conductance to atmosphere, C_1, was approximated as a series combination of a short-tube conductance, C_t, and the conductance of an orifice, C_o. All conductances were calculated for molecular nitrogen at room temperature (300°K).

Thus,
$$\frac{1}{C_1} = \frac{1}{C_0} + \frac{1}{C_t}.$$

Using equations 2.29a and 2.37a (cf. Reference 1, pages 88 and 91), the conductance C_1 was calculated as C_1 = 6.5 liters per second.

The conductances linking V_1 and V_2, and V_2 and V_3, respectively, were approximated by considering $C_2 = C_3$ and using equation 2.48 (cf. Reference 1, page 101) for the conductance of a short annulus. This gives,

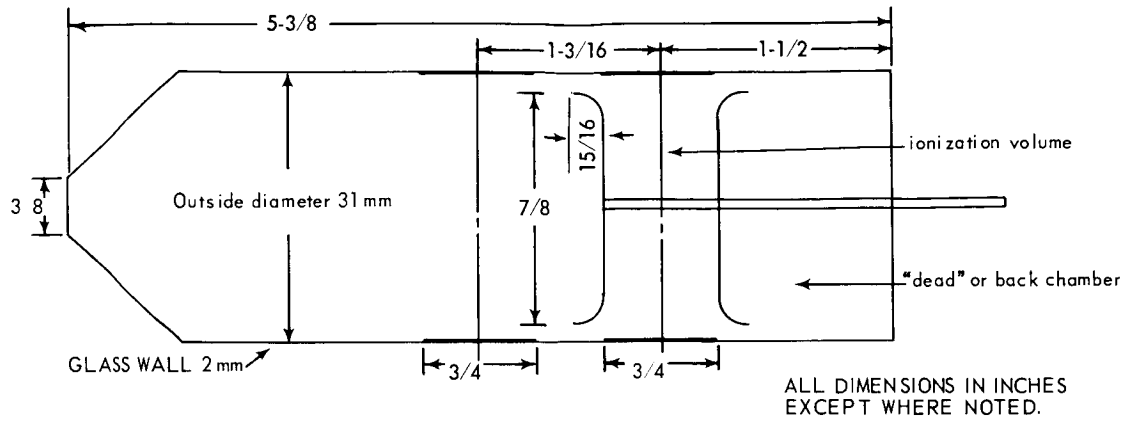


Figure B-30(a)–Pertinent Dimensions of Flight Redhead Gauge

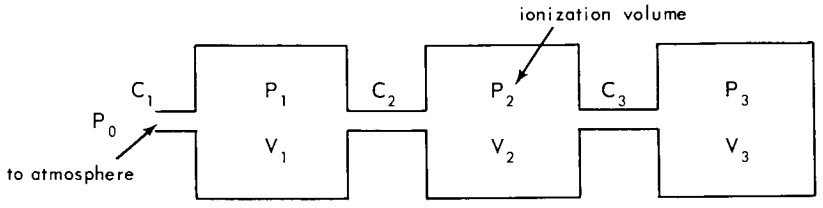


Figure B-30(b)–Schematic of Flight Redhead Gauge

$$\frac{1}{C_2} = \frac{1}{C_3} = \frac{1}{C_a} + \frac{1}{C_0}$$

where C_a is the annulus conductance based on a long-tube derivation, and C_o is the conductance of an orifice included to correct for end effects.

Two calculations for the desired conductance were made; the first considered one orifice to be present, and the second accounted for a series combination of two orifices. Or,

$$\frac{1}{C_{0(total)}} = \frac{1}{C_{01}} + \frac{1}{C_{02}} = \frac{2}{C_0}.$$

Using the referenced equations, C_a = 24.7 liters per second and C_o = 22.1 liters per second. Thus, for one orifice, $C_2 = C_3$ = 11.7 liters per second; for two orifices, $C_2 = C_3$ = 7.7 liters per second.

To determine the gauge time response for the "worst case" conductance, the smaller value of $C_2 = C_3$ = 7.7 liters per second was used. Based on the schematic diagram of Figure B-30 and the established vacuum-system equation,

$$\left[V \frac{dP}{dt} = PC\right]$$

the three differential equations governing the Redhead gauge system can be written as:

$$\begin{cases} V_1 \dfrac{dP_1}{dt} = -C_1(P_1 - P_0) + C_2(P_2 - P_1) \\ V_2 \dfrac{dP_2}{dt} = -C_2(P_2 - P_1) + C_3(P_3 - P_2) \\ V_3 \dfrac{dP_3}{dt} = -C_3(P_3 - P_2) \end{cases} \tag{1}$$

where, P_1 = pressure in V_1
P_2 = pressure in V_2
P_3 = pressure in V_3
P_0 = external varying pressure

These equations can be rewritten in standard form as,

$$\begin{cases} \dot{P}_1 = a_{11}P_1 + a_{12}P_2 + KP_0 \\ \dot{P}_2 = a_{21}P_1 + a_{22}P_2 + a_{23}P_3 \\ \dot{P}_3 = a_{32}P_2 + a_{33}P_3 \end{cases} \tag{2}$$

where $a_{11}, a_{12}, \ldots\ldots K$ represent constants composed of various combinations of V_1, V_2, V_3, C_1, C_2, and C_3. Evaluation of these constants yields:

$$\begin{cases} \dot{P}_1 = -278P_1 + 151P_2 + 127P_0 \\ \dot{P}_2 = 705P_1 - 1410P_2 + 705P_3 \\ \dot{P}_3 = 468P_2 - 468P_3 \end{cases} \tag{3}$$

These equations can be solved for the ionization volume pressure P_2, which the data represents, using familiar operator techniques. Writing the equations in the form

$$\begin{cases} (D + 278)P_1 - 151P_2 = 127P_0 \\ -705P_1 + (D + 1410)P_2 - 705P_3 = 0 \\ -468P_2 + (D + 468)P_3 = 0 \end{cases} \tag{4}$$

the systematic elimination of P_1 and P_3 yields

$$[D^3 + 2156D^2 + 7.45(10^5)D + 4.1(10^7)]P_2 = 705(D + 468)127P_0. \quad (5)$$

The complementary solution for P_2 is

$$P_{2(complem.)} = C_1 e^{-64t} + C_2 e^{-1738t} + C_3 e^{-353t} \quad (6)$$

Hence, the transients in the solution for P_2 vanish rapidly relative to the spin period of the satellite.

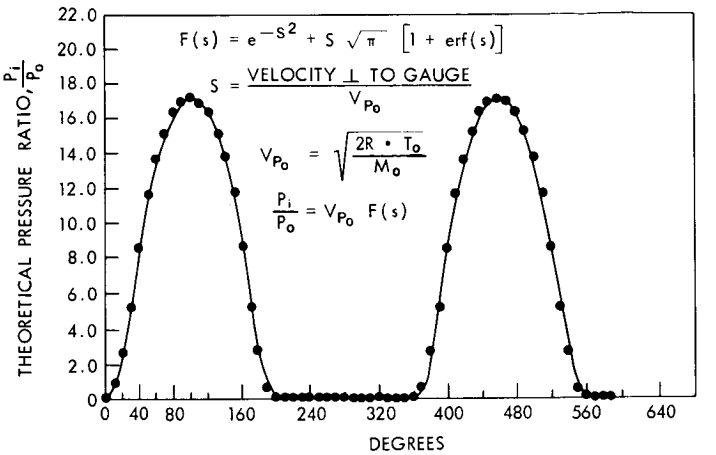
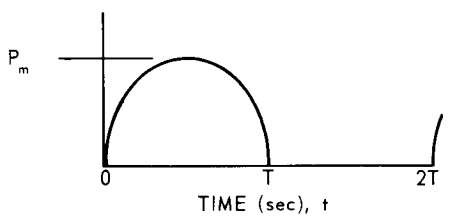


Figure B-31–Plot of Calculated Theoretical Equatorial Gauge, Pressure vs. Satellite Roll Angle

To obtain a particular integral for P_2, the external pressure P_0 was specified. Figure B-31 shows a simplified theoretical forcing function for the pressure P_0 experienced by an equatorial gauge. To account for the spin modulation of the gauge pressure, P_0 is assumed to be a half-wave rectified sine function as indicated below.



$$P_0 = P_m \sin \omega t \qquad 0 < t < T$$

$$P_0 = 0 \qquad T < t < 2T$$

P_0 may then be expanded in Fourier series:

$$P_0 = \frac{P_m}{\pi}\,[1 + 1.57 \sin 9.37t - 0.67 \cos 18.74t - 0.13 \cos 37.48t - .06 \cos 56.22t]$$

where

$$w = \frac{2\pi}{T_{satellite}} = \frac{2\pi}{0.67} = 9.37°/\text{sec}. \quad (7)$$

Evaluation of P_0 and dP_0/dt on the right side of the P_2 equation [5] and a projected solution for P_2 on the left side of the form of P_0 (i.e.,

$$P_2 = A + B_1 \sin 9.37t + B_2 \cos 9.37t + B_3 \sin 18.74t + \cdots, \quad (8)$$

gives the desired particular solution:

$$P_{2(part.)} = \frac{P_m}{\pi} - 0.729\,P_m \sin(9.37t - 1.3\pi) - 0.225\,P_m \sin(18.74t - 1.4\pi)$$

$$- 0.20\,P_m \sin(37.48t - 0.75\pi). \qquad (9)$$

Figure B-32 is a plot of P_0 and P_2 showing the fundamental and harmonics. It is seen that P_2 exhibits 31.5 millisecond maximum phase lag relative to P_0. No appreciable phase shift was observed in the telemetered data, however, but this is expected, since the data resolution for peak pressure is approximately 2 data points, corresponding to 18° or 34 milliseconds.

In the above analysis, the pumping speed of the gauge (i.e., the rate of ionic pumping of molecules by the surfaces in P_2) has been neglected. The effect on P_2 of this pump speed, S, would be included in the dP_2/dt equation as $-P_2S$. This pump speed has been measured as 1 liter per second for N_2 for Redhead gauges, while the conductances to and from V_2 are much larger (8 liters per second). Thus the effect of this pumping upon the calculated results is small.

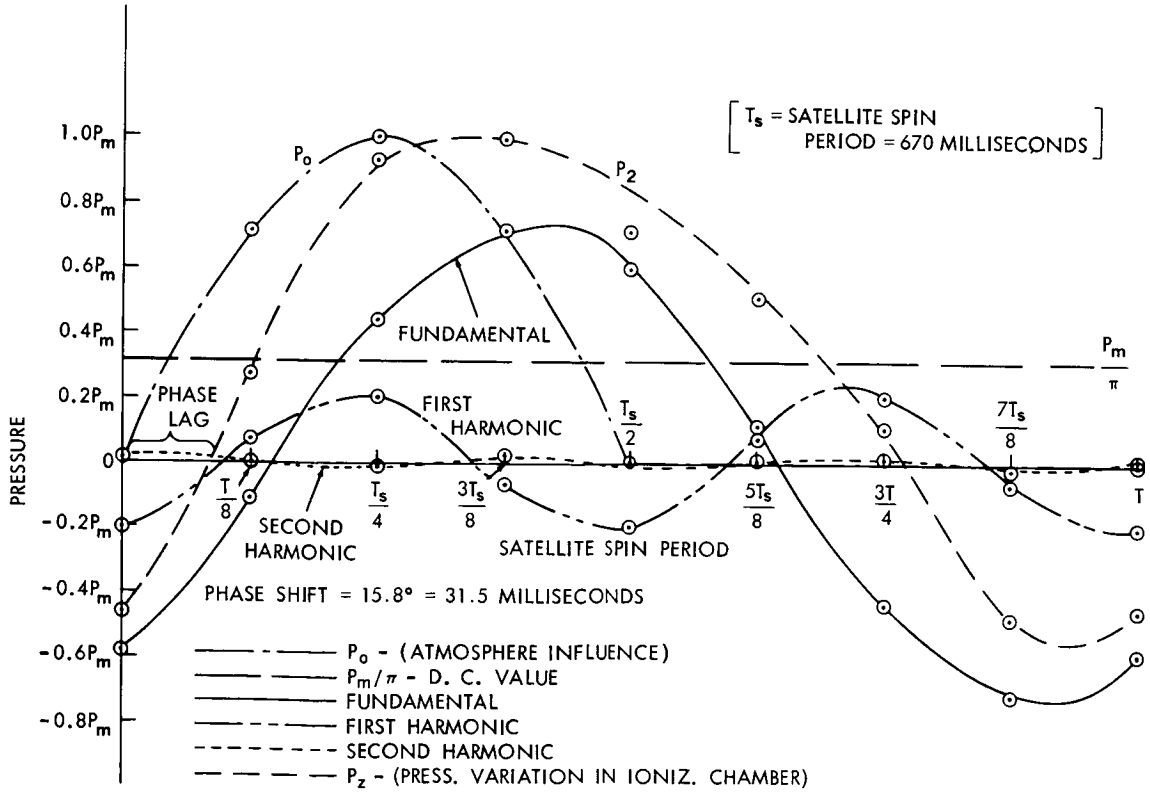


Figure B-32–Plot of Redhead Gauge, Ionization Chamber Pressure Response vs. Satellite Spin Period

THE BAYARD-ALPERT GAUGE ADDITIONAL CORRELATIONS

The examination of the Bayard-Alpert gauge response was continued by undertaking an extensive survey of high-altitude passes. At altitudes above 350 kilometers the Bayard-Alpert gauge residual pressure is determined by effects within the gauge and not by the atmosphere. At these altitudes the Bayard-Alpert gauge shows no spin modulation of the pressure, although the Redhead gauge continues to do so. Also, the Bayard-Alpert residual pressure is observed to decay slowly over each individual active period. This decay seems related to adsorption and desorption phenomena in the gauge. In the analysis below, the data was correlated to theoretical calculations and laboratory-determined desorption constants for the Bayard-Alpert gauge.

First, thirty-eight passes were selected with altitudes ranging from 370 to 920 kilometers. These 38 passes and pertinent parameters are catalogued in Table B-15. The Bayard-Alpert gauge pressures were plotted against time at 5-second intervals on semi-log paper. Figure B-33 is a typical plot of such a high-altitude pass with a

Table B-15
Explorer XVII Passes and Parameters

Pass	No.	Altitude		Aspect	Greenwich Mean Time		Prior Off Time	Prior Alt	Time Constant	Pressure Range (torr)
		Start	Finish		Start	Finish				
1 WINK	1	373	465	51.8° → 66.5°	0356	0400	54 min	D930	254 sec	$3.5\text{-}12(10^{-9})$
2 JOB	2	812	880	135° → 148°	0555	0559	115 min	A373	226 sec	$5.5\text{-}15(10^{-9})$
3 JOB	3	850	905	145° → 160°	0735	0739	96 min	A812	174 sec	$7\text{-}20(10^{-9})$
WOOMERA	3	719	632	111° → 96°	0804	0808	25 min	A850	Flat	$3.5\text{-}4(10^{-9})$
4 WOOMERA	13	896	919	149° → 162°	2342	2345	64 min	D668	225 sec	$1\text{-}2(10^{-8})$
5 WINK	16	389	464	45° → 63°	0403	0407	109 min	D266	308 sec	$1.1\text{-}2.4(10^{-8})$
6 NFL	18	393	467	46° → 63°	0716	0720	69 min	A842	246 sec	$1\text{-}2.5(10^{-8})$
JOB	18	897	925	151° → 164°	0744	0748	24 min	A393	Flat	$5.5\text{-}6(10^{-9})$
7 BPO	19	420	502	51° → 67°	0854	0858	38 min	D676	224 sec	$9\text{-}20(10^{-9})$
JOB	33	920	933	153° → 168°	0753	0757	26 min	A410	Flat	$7\text{-}9(10^{-9})$
8 SANTIAGO	37	917	932	150° → 165°	1418	1422	104 min	A768	263 sec	$1.5\text{-}2.5(10^{-8})$
JOB	303	778	698	93° → 107°	0951	0955	18 min	A905	Flat	$3\text{-}4(10^{-8})$
9 WOOMERA	411	732	807	79.5° → 64.5°	1440	1444	165 min	D909	918 sec	$4\text{-}5.5(10^{-8})$
10 WOOMERA	445	914	888	16.6° → 1.3°	2137	2141	157 min	D358	774 sec	$5\text{-}7(10^{-8})$
11 WOOMERA	588	646	731	179° → 163°	1038	1042	35 min	D361	625 sec	$8\text{-}13(10^{-8})$
12 WOOMERA	622	923	907	115° → 102°	1733	1737	163 min	D556	580 sec	$7\text{-}12(10^{-8})$
13 WOOMERA	623	914	884	109° → 92°	1912	1916	48 min	A261	585 sec	$6\text{-}8(10^{-8})$
14 QUITO	632	807	737	77° → 61°	0948	0952	72 min	D411	685 sec	$7\text{-}10(10^{-8})$
15 WOOMERA	647	569	657	135° → 151°	0913	0917	34 min	D421	606 sec	$7\text{-}11(10^{-8})$
16 FT MYER	649	612	529	48° → 33°	1138	1142	101 min	D782	420 sec	$5\text{-}8(10^{-8})$
17 WOOMERA	667	919	901	138° → 124°	1744	1748	121 min	A515	586 sec	$4.7\text{-}6.5(10^{-8})$
18 WOOMERA	706	538	627	100° → 116°	0749	0753	34 min	D395	496 sec	$7.5\text{-}15(10^{-8})$
19 WINK	765	498	422	82° → 66°	0545	0549	538 min	A511	458 sec	$4.5\text{-}6.0(10^{-8})$
20 WOOMERA	765	507	573	69° → 82°	0624	0627	35 min	D498	437 sec	$4.5\text{-}5.5(10^{-8})$
21 BPO	768	550	476	92° → 78°	1031	1035	99 min	D689	543 sec	$3\text{-}4.5(10^{-8})$
22 WOOMERA	770	912	911	160° → 170°	1451	1455	241 min	D298	868 sec	$4.5\text{-}6.5(10^{-8})$
		Apogee								
23 SANTIAGO	794	913	905	149° → 162°	0521	0525	507 min	A534	473 sec	$5\text{-}9(10^{-8})$
		Apogee								
24 WOOMERA	815	909	911	133° → 149°	1459	1503	47 min	D290	754 sec	$4\text{-}6\ (10^{-8})$
25 QUITO	825	828	768	170° → 161°	0715	0719	72 min	D454	638 sec	$4.5\text{-}6.5(10^{-8})$
26 QUITO	839	869	823	158° → 169°	0538	0542	72 min	D525	655 sec	$5\text{-}7(10^{-8})$
27 WOOMERA	874	905	906	100° → 115°	1334	1338	49 min	D308	805 sec	$5\text{-}7(10^{-8})$
28 WOOMERA	913	424	507	54° → 48°	0335	0339	35 min	D584	845 sec 585 sec	$6.5\text{-}10(10^{-8})$ $9\text{-}18(10^{-8})$
WOOMERA	18	676	586	111° → 95°	0812	0816	24 min	A897	Flat	$3.5\text{-}5(10^{-9})$
29 WOOMERA	293	767	820	No Aspect	1716	1719	34 min	D271	1180 sec	$5\text{-}7(10^{-8})$
WOOMERA	474	922	893	39° → 24°	2008	2012	20 min	A673	Flat (sl incr)	$3\text{-}4(10^{-8})$
QUITO	508	503	569	166° → 153°	0210	0213	0 min	A291	Flat	$9\text{-}11(10^{-8})$
WOOMERA	548	922	896	72° → 58°	1852	1856	22 min	A614	Flat (sl incr)	$3\text{-}4.5(10^{-8})$
QUITO	618	749	691	57° → 42°	1124	1128	73 min	D370	Flat (sl incr)	$3.5\text{-}4.5(10^{-8})$

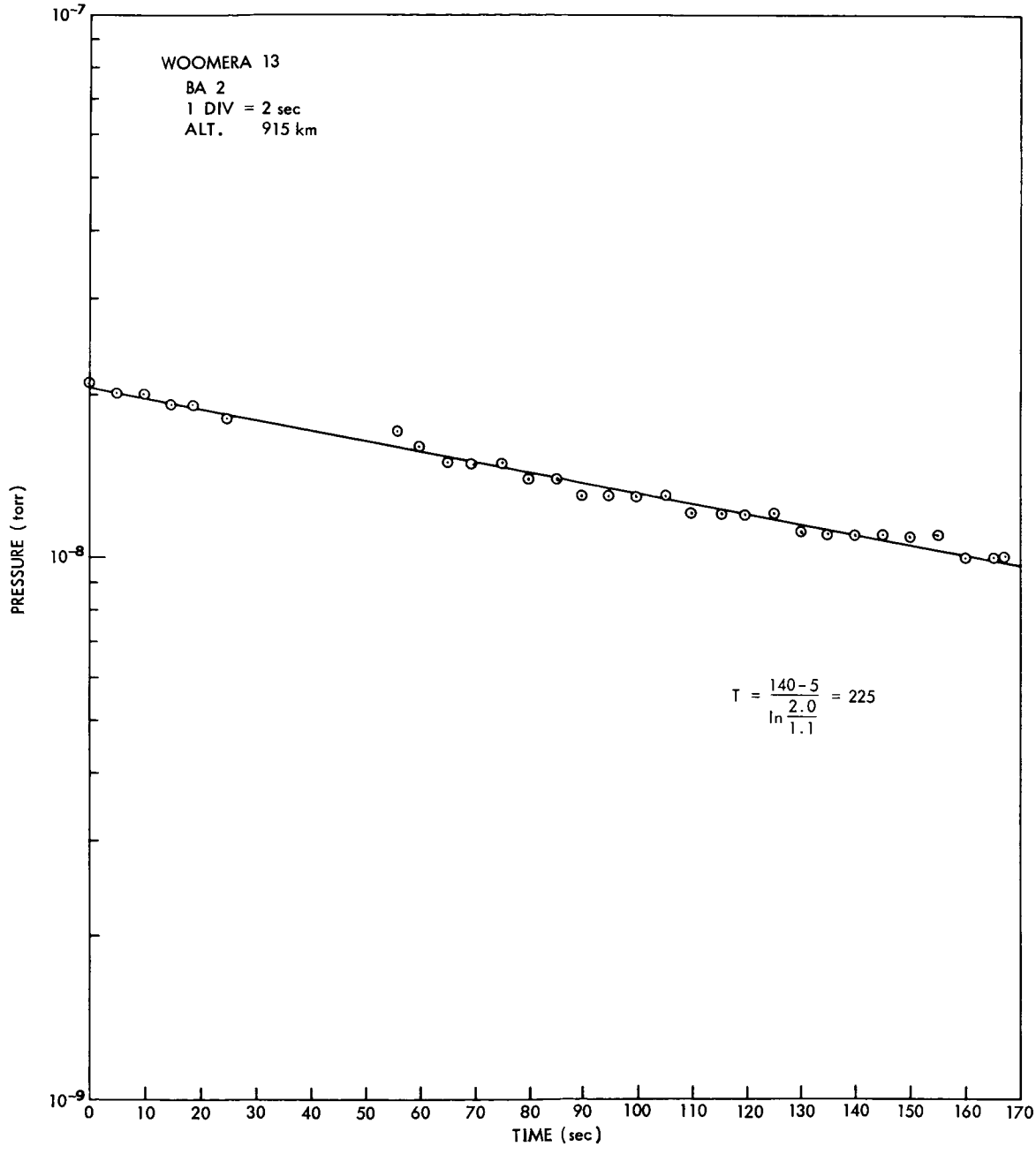


Figure B-33–Plot of High-Altitude Bayard-Alpert Gauge Response

straight-line fit to the curve. For each plot, a time constant was computed, these constants were then used to characterize the decay of each indicated residual pressure. In order to determine the factors governing the magnitudes of these residuals, the derived time constants were correlated to environmental parameters. The parameters examined include: off-time of the gauge since its last operation; off-time preceding the data-producing pass below 350 kilometers (where the atmospheric pressure is large enough to significantly affect monolayer formation times); the altitude and pressure regime during the active period; the altitude during the last active period; the orientation of the gauge normal relative to the satellite velocity vector during the data-producing pass; and the orbit number corresponding to the pass under consideration. Figures B-34 through B-38 indicate the relationship between the residual decay time constants and the several parameters. From these plots, the following conclusions were drawn:

1. There is no discernible correlation of the residual time constants with gauge orientation, with total off-time between passes, or with the altitude of the pass.

2. The residual pressure decays more rapidly early in the life of the gauge than during the later period. Passes through orbit number 40 have time constants on the order of 240 ± 25% sec; those during the later life are on the order of 640 ± 30% sec. Thus the gauge experienced a more rapid clean-up early in its operational life.

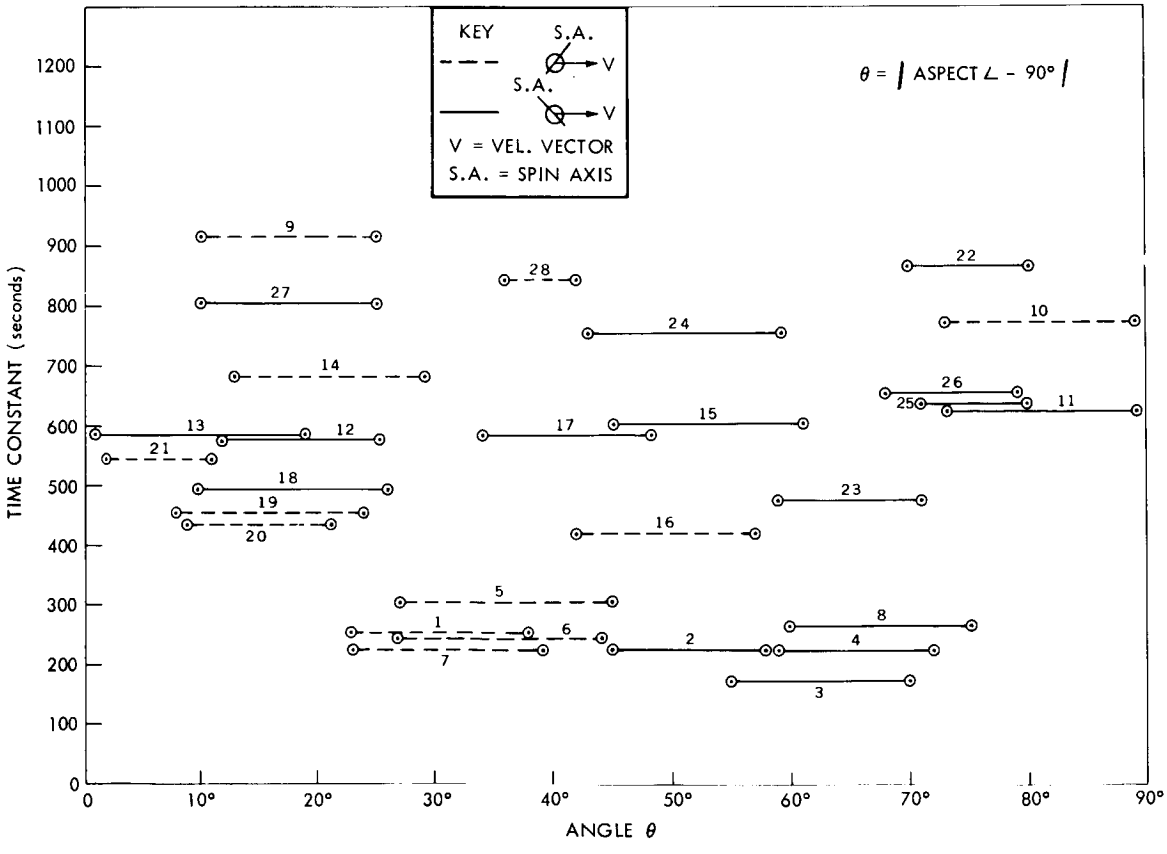


Figure B-34–Time Constants vs. Equatorial-Plane/Velocity-Vector Angle

In connection with this comparison, the maximum value of the Redhead gauge pressure during each pass was used to estimate the effect of the atmosphere upon the indicated Bayard-Alpert gauge pressure during the pass. For the early passes, the ratio of the Redhead maximum pressure to the Bayard-Alpert indicated that residual pressure ranged from 0.12 to 0.70; for all the later passes, the ratio was less than 0.10. This corroborates the "clean-gauge" interpretation stated previously for the early passes. A clean gauge would give consistently lower pressures and exhibit a lower residual (thus producing pressure readings in agreement with those of the Redhead gauge, which recorded lower residual pressures than the Bayard-Alpert gauge consistently throughout the satellite lifetime.

3. Comparison of the decay time constants with the off-time preceding the pass below 350 kilometers yielded an interesting statistic: Of the 38 passes examined, 8 had flat or constant residuals; but these were the only passes with no off-time between turn-ons below 350 kilometers. The flat residual pressures were interpreted in light of the calculated theoretical time constant, T_m, for monolayer adsorption, although it is recognized that the gauge probably operated with gas-covered surfaces, possibly with several adsorbed monolayers. According to the theory, T_m varies from 33 minutes to 330 minutes (for sticking probabilities, from 0.1 to 0.01) for a pressure of 10^{-8} torr. Since all the preceding off-times between passes for the flat residuals were less than 33 minutes, and all in a pressure regime less than 10^{-8} torr, it appears that there was insufficient time between turn-ons for the readsorption of a monolayer. This behavior supports the "clean-gauge" theory.

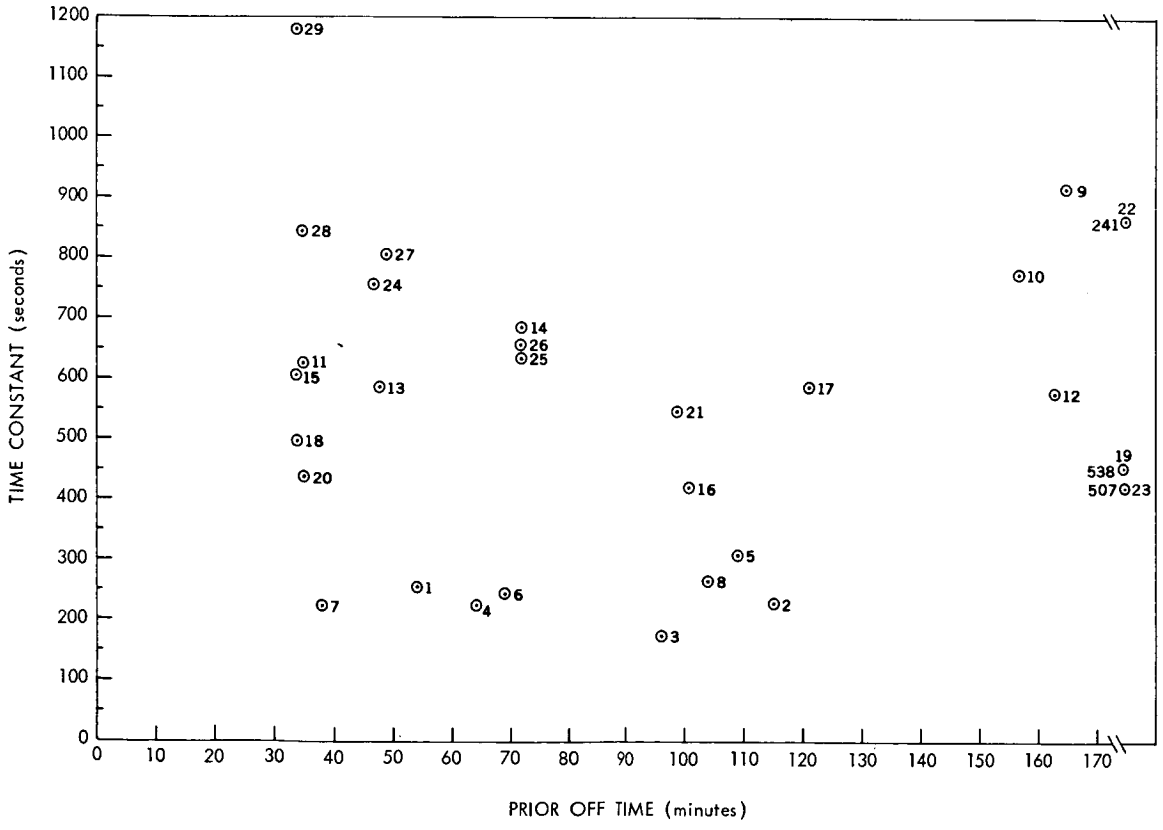


Figure B-35–Plot of Time Constants vs. Preceding Off-Time

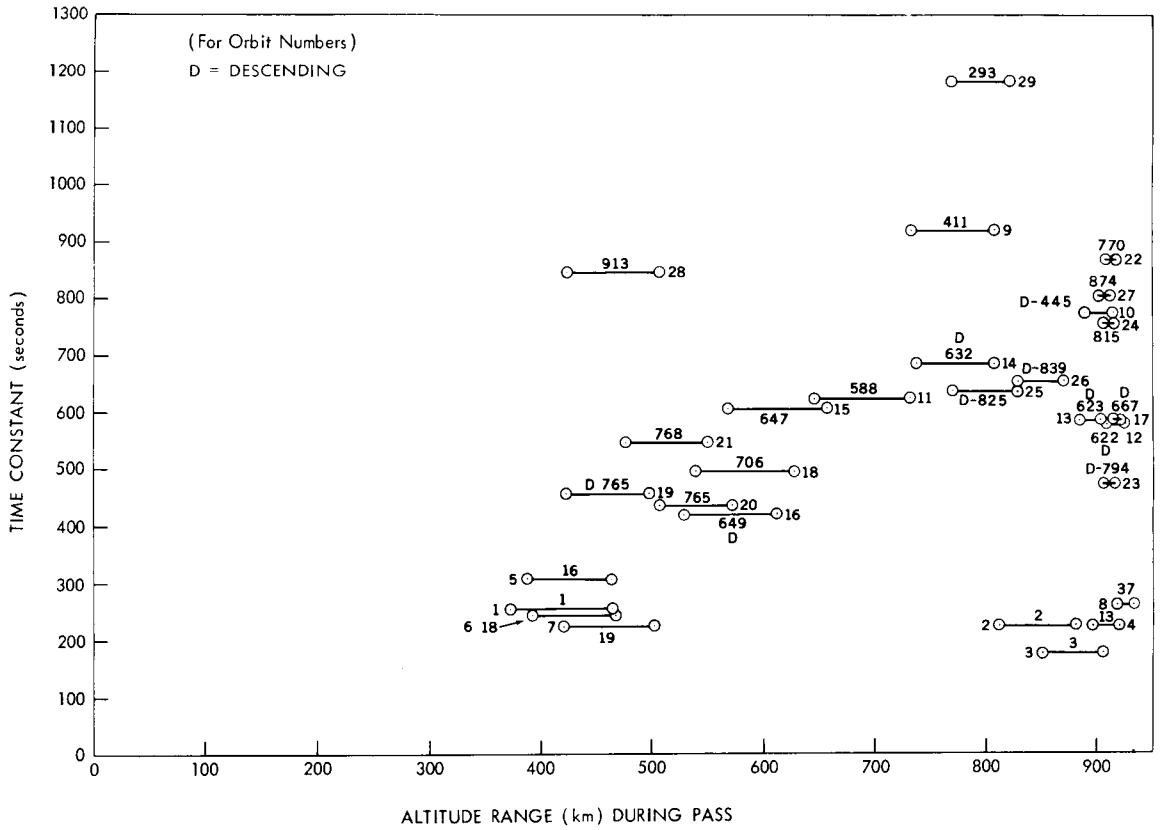


Figure B-36–Plot of Time Constants vs. Pass-Altitude Range

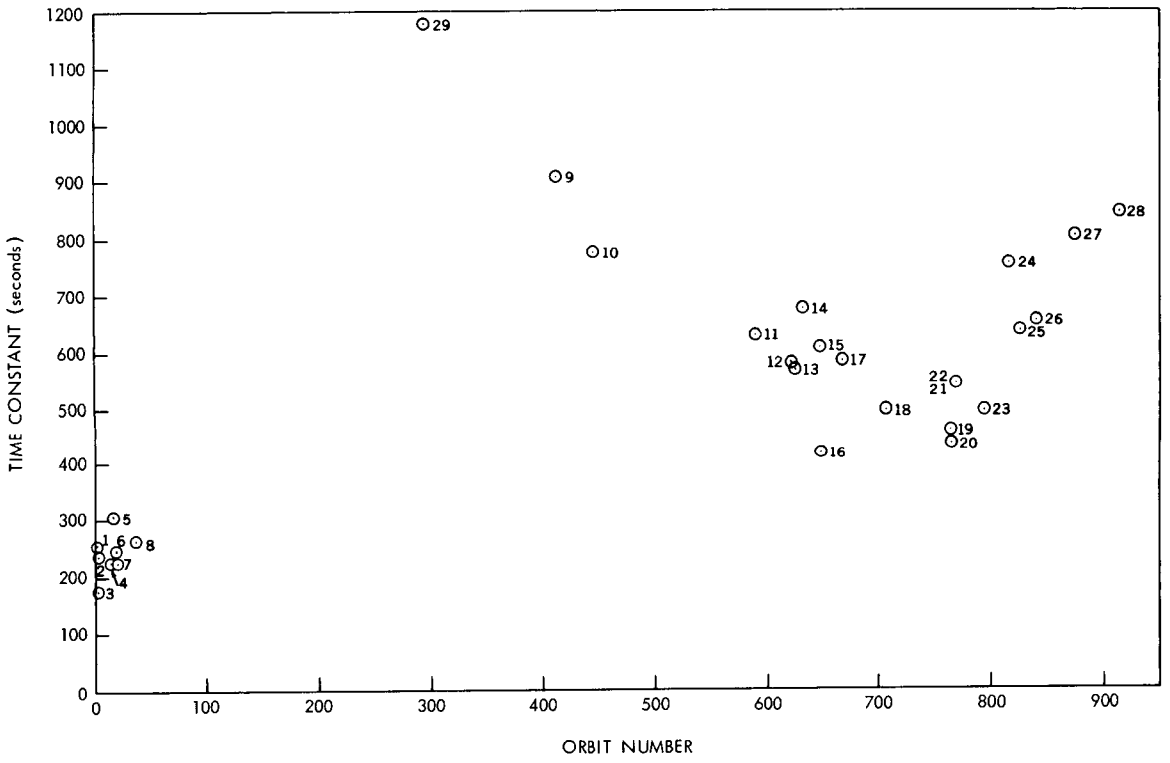


Figure B-37–Plot of Time Constants vs. Orbit Number

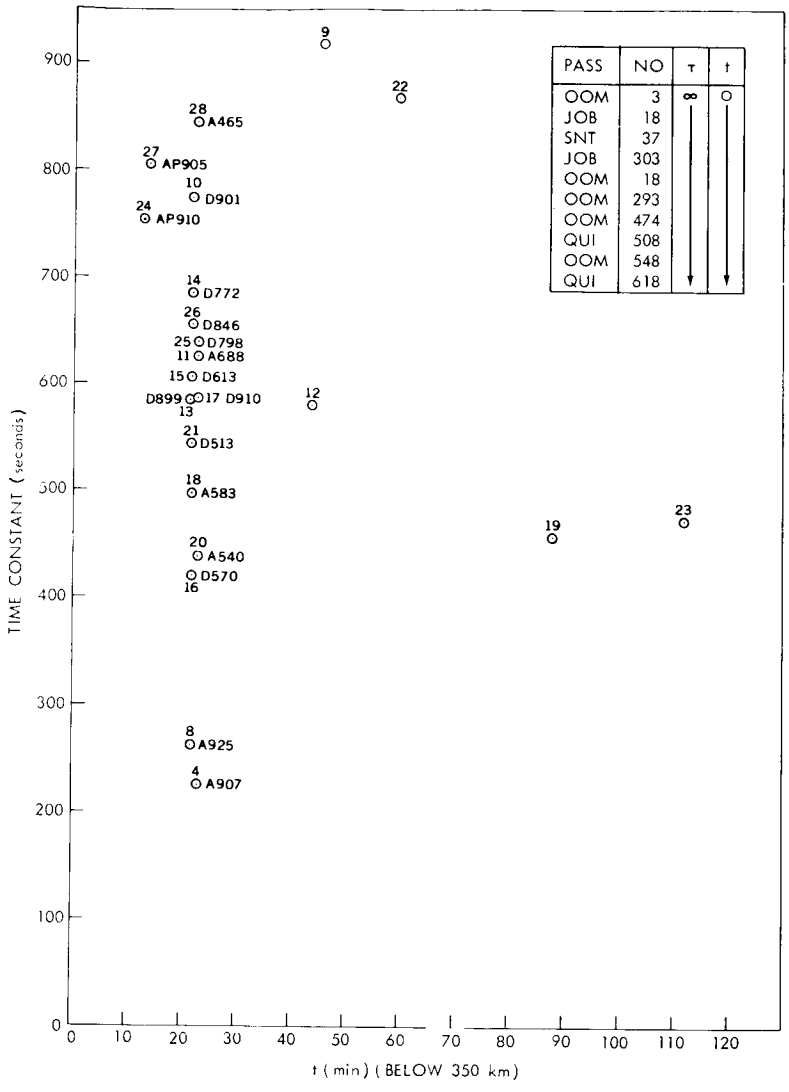


Figure B-38–Plot of Time Constants vs. Time Below 350 Km

NOTE ON BAYARD-ALPERT ELECTRONIC SYSTEM WARM-UP

It is important to point out the effect which the gauge electronics has on the pressure readings. The electronics experiences a warm-up during the entire 4-minute active period. The effect of the warm-up is to increase the magnitudes of the sorption time constants. The increase is non-linear, but the maximum error that could be introduced in the time constants is thought to be 30 percent.

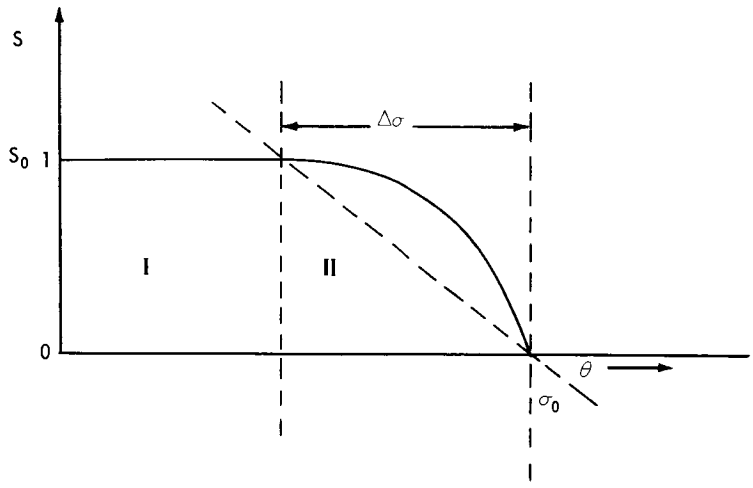
THEORETICAL APPROACH TO ADSORPTION AND DESORPTION PHENOMENA

The approach presented here is a simplified analysis: the coverage of the gauge surfaces not being known, monolayer adsorption is assumed. Because of the magnitude of the residual current it was decided to examine the Bayard-Alpert high-altitude residual data in the light of adsorption and desorption theory due to electron bombardment.

For pure adsorption we have:

$$\frac{d\sigma}{dt} = PVS(\theta) \tag{10}$$

where σ is the number of adsorbed atoms per square centimeter, P is the pressure, $V = 1.98 \times 10^{21}/M^{V2}$, $S(\theta)$ is the sticking probability, θ is σ/σ_0 the fractional coverage, and σ_0 is a monolayer coverage. In general, $S(\theta)$ has the form:



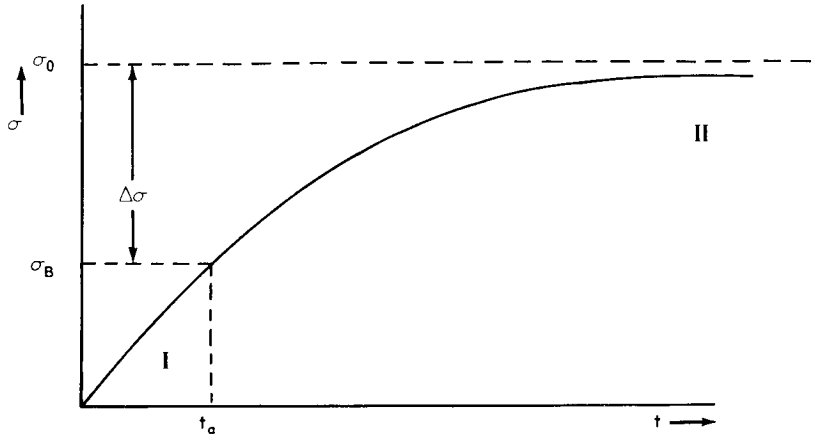
If $S(\theta)$ is approximated by two straight lines, we have two differential equations for two regions:

$$\frac{d\sigma}{dt} = PVS_0 \qquad \text{(region I)} \tag{11}$$

and

$$\frac{d\sigma}{dt} = PV[a + b\theta] \qquad \text{(region II)} \tag{12}$$

The total solution will be of the form:



where

$$\sigma = PVS_0\, t \qquad \text{(region I)} \qquad (13)$$

$$\sigma = \sigma_0 \left[1 - \frac{\Delta\sigma}{\sigma_0} e^{-\Delta t \Big/ \left(\frac{\Delta\sigma}{PVS_0}\right)}\right] \qquad \text{(region II)} \qquad (14)$$

The time required to reach a monolayer coverage is approximately:

$$\frac{\sigma_B}{PVS_0} + 3\tau = \frac{\sigma_B}{PVS_0} + \frac{3\Delta\sigma}{PVS_0} = \frac{1}{PVS_0}\left[\sigma_B + 3(\sigma_0 - \sigma_B)\right]$$

$$= \frac{3\sigma_0 - 2\sigma_B}{PVS_0} \qquad (15)$$

or from $\dfrac{3\sigma_0}{PVS_0}$ to $\dfrac{\sigma_0}{PVS_0}$

for $\sigma_B = 0$ to σ_0

generally $\sigma_B \cong \dfrac{\sigma_0}{2}$

Therefore, the time for a monolayer is:

$$T_m \cong \frac{2\sigma_0}{PVS_0} \qquad (16)$$

for atomic oxygen $$T_m = \frac{20 \times 10^{-6}}{P(\text{torr})} \text{ sec} \quad \text{for} \quad S_0 = 0.1$$

$$T_m = \frac{200 \times 10^{-6}}{P(\text{torr})} \text{ sec} \quad \text{for} \quad S_0 = 0.01.$$

The S_0's (sticking probabilities) used here are order-of-magnitude approximations; they depend on many parameters and there is little concise information on them in the literature.

Now electron bombardment desorption of gas from the grid will be considered. For ion emission, we have:

$$n_\epsilon = \frac{I^-}{\epsilon A} = \frac{\text{incident electrons}}{\text{cm}^2 \text{ sec}}$$

where ϵ is the electronic charge, I^- is the electron current, and A is the area. If the cross section is φ and the coverage is σ, we have:

$$n_\epsilon \varphi\sigma = \frac{\text{excitations}}{\text{cm}^2 \text{ sec}}$$

and for P = escape probability

$$n_\epsilon \varphi\sigma P = \frac{\text{ions escaping}}{\text{cm}^2\text{-sec}}$$

or

$$I^+ = \epsilon n_\epsilon \varphi\sigma PA = \frac{\epsilon I^-}{\epsilon A} \varphi P\sigma A$$

$$= (\varphi P)_I \sigma I^- = [(\varphi P)_I \sigma_0] \theta I^-$$

$$I^+ = g^+ \theta I^- \tag{17}$$

Similarly, for neutrals,

$$I^\circ = (\varphi P)_N \sigma I^- \tag{18}$$

$$\frac{dN}{dt} = \frac{I^\circ}{\epsilon} = (\varphi P)_N \sigma \frac{I^-}{\epsilon}$$

or

$$\left.\frac{d\sigma}{dt}\right|_N = \frac{1}{A} \frac{dN}{dt} = (\varphi P)_N \sigma \frac{J^-}{\epsilon}$$

$$= g_N \theta \frac{J^-}{\epsilon} \tag{19}$$

$$\left.\frac{d\sigma}{dt}\right|_I = \frac{1}{A}\frac{I^+}{\epsilon} = g^+\theta\,\frac{J^-}{\epsilon} \tag{20}$$

thus

$$\left.\frac{d\sigma}{dt}\right|_{total} = g\theta\,\frac{J^-}{\epsilon} \tag{21}$$

where J^- = electron-current density.

$$g = [(\varphi P)_N + (\varphi P)_I]\,\sigma_0 \tag{22}$$

For the case of concurrent adsorption and electron desorption, we have:

$$\frac{d\sigma}{dt} = -g\theta\,\frac{J^-}{\epsilon} + 2PVS(\theta) \tag{23}$$

At equilibrium

$$\frac{d\sigma}{dt} = 0$$

Therefore

$$\frac{g\theta J^-}{\epsilon} = PVS(\theta)$$

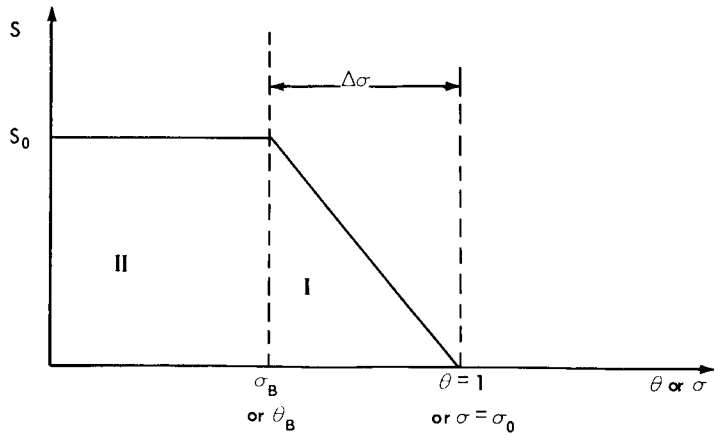
but

$$I^+ = g^+\theta\, I^- = g^+\theta\, J^- A$$

hence

$$I^+_{eq} = \frac{\epsilon\, g^+}{g}\,[2PVS(\theta)] \tag{24}$$

The total equation (15) can be solved assuming P = constant, and as before, $S(\theta)$ has the form



For region I ($\sigma_B < \sigma < \sigma_0$), we have:

$$S = \frac{(\sigma_0 - \sigma)}{\Delta\sigma} S_0 \tag{25}$$

thus

$$\frac{d\sigma}{dt} = - \frac{g\sigma}{\sigma_0} \frac{J^-}{\epsilon} + \frac{2PVS_0}{\Delta\sigma} (\sigma_0 - \sigma) \tag{26}$$

or

$$\frac{d\sigma}{dt} = - \frac{\sigma}{\tau_D} + \frac{(\sigma_0 - \sigma)}{\tau_A} \tag{27}$$

where τ_D = desorption time constant = $\frac{\sigma_0 \epsilon}{gJ^-}$ (28)

and τ_A = adsorption time constant = $\frac{\Delta\sigma}{2PVS_0}$ (29)

$$\frac{1}{\tau} = \frac{1}{\tau_A} + \frac{1}{\tau_D} \tag{30}$$

thus

$$\frac{d\sigma}{dt} = - \frac{\sigma}{\tau} + \frac{\sigma_0}{\tau_A} \tag{31}$$

We have the solution (with $\sigma = \sigma_0$ at $t = 0$),

$$\sigma = \sigma_0 e^{-\frac{t}{\tau}} + \frac{\sigma_0 \tau}{\tau_A} \left(1 - e^{-\frac{t}{\tau}}\right)$$

or

$$\frac{\sigma}{\sigma_0} = \left(1 - \frac{\tau}{\tau_A}\right) e^{-\frac{t}{\tau}} + \frac{\tau}{\tau_A} = \frac{\tau}{\tau_D} e^{-\frac{t}{\tau}} + \frac{\tau}{\tau_A} \tag{32}$$

For region II ($0 < \sigma < \sigma_B$) we have:

$$\frac{d\sigma}{dt} = - \frac{gJ^-}{\epsilon\sigma_0} \sigma + 2PVS_0 \tag{33}$$

$$= - \frac{\sigma}{\tau_D} + \frac{\Delta\sigma}{\tau_A} \tag{34}$$

taking $\sigma = \sigma_B$ at $t = t_B$

$$\sigma = (\sigma_B - \alpha\Delta\sigma)\, e^{-\frac{(t-t_B)}{\tau_D}} + \alpha\Delta\sigma \tag{35}$$

where $\alpha = \frac{\tau_D}{\tau_A}$

The collector efficiency can be calculated from an equation from a paper by P. A. Redhead.[1] Of all ions released at the grid surface, the fraction reaching the ion collector is

$$F = \frac{2 r_c}{\pi r_g} \left(\frac{V}{V_0}\right)^{1/2} \tag{36}$$

where r_c is the collector radius, r_g is the grid, V is the potential between the grid and collector, and V_0 is the energy of the ions ejected.

Final Analysis of Bayard-Alpert High Altitude Residual Data

Calculations were made of the reaction probability, g, using the above theory and the experimental data. These calculations were then compared to modified values from several papers on desorption of oxygen and dissociation of carbon monoxide on metal surfaces due to electron bombardment. From considerations of the upper atmospheric composition and the flight Bayard-Alpert gauge, these two processes were thought to be the most probable. As was mentioned earlier, this simplified analysis assumes a monolayer coverage. In determining time constants for monolayer coverage, an average pressure below 350 km was used, this was calculated by taking three-fourths of the maximum pressure for a near-perigee pass. The average pressure was 5.6×10^{-7} torr; thus, for $S_0 = 0.1$, $T_m = 35.7$ seconds, and for $S_0 = 0.01$, $T_m = 357$ seconds (Figure B-39). The total time the satellite was below 350 km is 22 minutes or 1320 seconds. Assuming the above sticking probabilities, the gauge surfaces had ample time to achieve at least a monolayer coverage.

Using the dimensions of the flight Bayard-Alpert gauge ($r_c = .006$ in, $r_g = .264$ in) and $V = 127$, $V_0 = 6v^1$ and using equation (35), we obtain $F = 6.70$ percent for the percentage of all ions released at the grid surface that reach the ion collector.

Now if the adsorption time constant is large at high altitudes, then $\tau = \tau_D$ from equation (29). Two average τ_D's were determined: one for passes early in the lifetime, group I; and one for passes later in the lifetime, group II, (i.e., orbit number >40). From these two values of τ_D values for g, the reaction probability, were calculated using equation (27). These experimental values were then compared to a value of g, modified for these comparisons from a paper by P. A. Redhead.[1] This g was modified by the factor F (Table B-16) which reflects the ion-collection efficiency. The value of g so obtained is for desorption of oxygen from clean molybdenum. In Redhead's experiment the molybdenum surface was treated very carefully to be clean and to be free from carbon. The grid in the flight Bayard-Alpert gauge was molybdenum, but its state of cleanliness is not known.

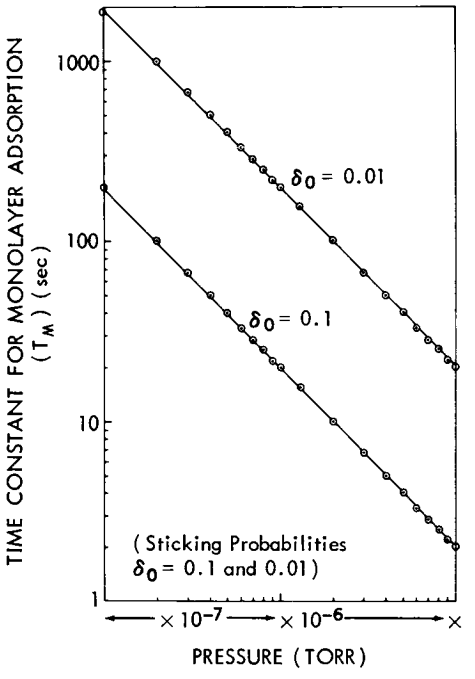


Figure B-39–Plot of Theoretical Times for Oxygen-Monolayer Adsorption (Room Temperature) at Various Pressures (Sticking Probabilities S_0 = 0.1 and 0.01)

Table B-16
Comparison, Experimental Reaction
Probability vs. Modified Reaction Probability

Grid area (cm^2)	V_e (volts)	I^- (amp)	Ave. τ_D	g P. A. Redhead	Collection efficiency F	F × g	Exp. g	% Diff.
2.40	100	5×10^{-3}	240 (I)	1.33×10^{-3}	.067	8.90×10^{-5}	16.0×10^{-5}	80.0
↓	↓	↓	638 (II)	↓	↓	↓	6.02×10^{-5}	32.4

As seen from Table B-16, the average reaction probabilities for Group I does not agree closely with the theory. Considering the average τ_D for group II, the value of g is within the error of the data which is ±50 percent. The results are consistent with the theory that oxygen is adsorbed on the grid and then desorbed during electron bombardment.

From the Bayard-Alpert high-altitude residuals, which were flat, the ion currents were determined. From these ion currents, g^+'s were calculated from equation (8). (See Table B-17). These experimental values were compared to a modified g^+ from P. A. Redhead's[1] paper and to a modified g^+ from a paper by Moore.[2] The g^+'s from the two papers were also modified by a multiplying factor of F where F = .067.

Table B-17
Comparison, Experimental Ion Reaction Probability vs.
Modified Ion Reaction Probabilities

I^- (amp)	I^+ (amp)	Exp. g^+	g^+ (CO) Moore	% Diff.	g^+ (0) Redhead	% Diff.
5×10^{-3}	1.78×10^{-10}	3.56×10^{-8}	2.34×10^{-6}	98.7	7.44×10^{-7}	95.2
	1.59×10^{-10}	3.18×10^{-8}		98.6		95.7
	2.42×10^{-10}	4.84×10^{-8}		97.6		93.5
	3.34×10^{-10}	6.68×10^{-8}		97.0		91.0
	1.46×10^{-9}	2.92×10^{-7}		87.5		60.8
	1.46×10^{-9}	2.92×10^{-7}		87.5		60.8
	4.18×10^{-9}	8.36×10^{-7}		64.1		11.0
	1.57×10^{-9}	3.14×10^{-7}		86.7		57.8
↓	1.67×10^{-9}	3.34×10^{-7}	↓	85.7	↓	55.1

The experimental g^+'s from Table B-17 are listed in order of increasing orbit number. Also, as a point of interest, the g^+ with the 11-percent difference, and the 64.1-percent difference is from a pass with no time off between two successive passes, again indicating clean-up of the gauge.

Comparison of the percentage differences in Table B-17 shows that there is better agreement of the experimental g^+'s with the g^+ for the desorption of oxygen than the g^+ for the dissociation of carbon monoxide. This result is also consistent with the result from Table B-16.

Summary Conclusions and Suggestions for Future Work

A method for selection of data has been determined, and the repeatability of the data verified over both individual passes and the satellite lifetime. The temperature effect on the Redhead gauge response has been seen to be small. Theoretical calculations have supported the observed phenomena of no phase shift in the Redhead gauge response to the spin-modulation of the pressure. An extensive study has been made of the Bayard-Alpert gauge response at high altitudes where the expected spin-modulation of pressure is absent. Correlations have been made of time constants obtained from straight-line fits to the indicated Bayard-Alpert residual pressure decays. From these correlations, strong indications have been obtained of more rapid gauge clean-up early in the life of the gauge, and of cleaner gauge operation when the satellite has remained above 350 km with an off-time less than the theoretical time for monolayer adsorption. However, an enhanced residual current was consistently obtained. Equations have been used from the presented adsorption and desorption theories to calculate ion currents, g's, and g^+'s from the experimental data. Comparisons of these values with adjusted published laboratory data seem to indicate the desorption of oxygen from the grid rather than the dissociation of carbon monoxide on the grid.

Since there seems to be a good indication that oxygen is desorbing from the grid, further experimental work should be done in the laboratory to verify this. An experiment should be constructed to measure decay rates, residual currents, electron currents, and fractional coverage for a flight Bayard-Alpert gauge in an oxygen atmosphere in a pressure range of 10^{-7} to 10^{-11} torr, using times comparable to the length of an Explorer

XVII turn-on. Further studies should be made of the Bayard-Alpert gauge high-altitude residual pressures, the lack of correlation not being understood. The combined effect of several parameters should also be examined.

ACKNOWLEDGEMENTS

The authors thank Dr. F. M. Propst, University of Illinois, for many helpful discussions and valuable suggestions, particularly in the areas of adsorption and desorption of gases. Thanks also to Mr. Brian Kritt, GSFC, for his helpful suggestions in preparing the paper.

REFERENCES

1. Dushman, S., Scientific Foundations of Vacuum Technique, Nitey and Sons, New York (1962)

2. Moore, G.E., "Dissociation of Adsorbed CO by Slow Electrons," Journal of Applied Physics 32 (1961)

3. Newton, G.P., Pelz, D.J., Miller, G.E., Horowitz, R., "Response of Modified Redhead Magnetron and Bayard-Alpert Vacuum Gauges Aboard Explorer XVII," NASA Technical Note, TN D-2146 (February 1964)

4. Redhead, P.A., "The Effects of Adsorbed Oxygen on Measurements with Ionization Gauges," Vacuum 13 (1963)

5. Spencer, N.W., Newton, G.P., Reber, C.A., Brace, L.H., Horowitz, R., "New Knowledge of the Earth's Atmosphere from the Aeronomy Satellite," GSFC, X-651-64-114 (May 1964)

6. Horowitz, R., and Kleitman, D., Upper Atmosphere Research Report No. XVIII, NRL Report 4246, U.S. Naval Research Laboratory, Washington, D.C. (October 1953)

Page intentionally left blank

PROCESSES IN HOT-CATHODE IONIZATION GAUGES
(In Reference to Explorer XVII)

Dr. F. M. Propst

Coordinated Science Laboratory

Because of the large number of "anomalous" processes that can take place in a hot-cathode ionization gauge, it is perhaps worthwhile to look in a general way at the various types of such processes that can occur. We must first decide what is meant by "anomalous." We shall adopt the definition that "any current collected at or leaving the collector of an ion gauge that is not proportional to the volume density of gas constitutes an 'anomalous effect.' " We see immediately that all anomalous effects are secondary to cathode processes (with the exception of capacitive current induced in the collector circuit by work function changes of the grid and collector upon gas adsorption), because no currents flow when the cathode is cold.

When the cathode is heated, we find the emission of

- electrons
- ions (positive and negative)
- neutral particles
- photons

All phenomena, both normal and anomalous, can be predicted by following the interaction of these particles with the surfaces and volume gas in the gauge.

Electrons

As an example, let us trace the history of the electrons emitted from the cathode. For simplicity, we shall assume that the walls of the gauge do not participate in the gauge processes. The effects due to the walls can be treated in exactly the same manner outlined below.

The electrons emitted from the cathode pass through the volume of the gauge and are collected at the grid. During transit, the electrons can collide with and ionize the volume gas. They can also partake in collective motion resulting in space charge oscillation. Upon striking the grid, the electron can cause the emission of photons, electrons, ions (positive and negative), and neutrals. We have, then, the processes secondary to the electron emission from the cathode:

1. Volume ionization
   a. positive
   b. negative
2. Oscillations

3. Particles emitted from the grid

   a. electrons

   b. photons

   c. ions

      i. positive

      ii. negative

   d. neutrals.

The question is, then, to determine which of these possibilities give rise to observable effects.

1.a–The positive ions formed within the grid volume during the transit of the electron are drawn to the collector. The current due to these ions is proportional to the pressure; this is the pressure-measuring feature of the gauge. The positive ions formed outside of the grid are lost (assuming no walls) and are dropped from consideration. When the ions strike the collector, they cause the emission of photons, ions (positive and negative), neutrals, and electrons. We can then trace the processes involving these particles also to look for anomalous effects; however, they are all secondary to the ion current and for constant collector conditions proportional to the ion current (and hence proportional to the pressure), so we shall not consider them here. When the positive ions are formed in the volume of the gauge, equal numbers of electrons are formed. At higher pressures, the density of these electrons becomes of the order of the density of electrons originating at the cathode, and non-linearity can occur; however, below $\sim 10^{-5}$ T, any effects due to these electrons are small compared to the effects due to the primary electrons.

1.b–The negative ions formed in the flight of the electrons are attracted to the grid. The impact of these ions can cause the emission of photons, ions (positive and negative), electrons, and neutrals. Again, these particles can cause anomalous readings, but will not be considered here for the same reasons given in the case of particles secondary to the positive ions striking the collector.

2.–Collective oscillations of the electrons can cause large anomalous readings which can in general be recognized and corrected. We shall not attempt a discussion of these effects here.

3.a–Electrons ejected from the grid by the primary electrons will be attracted back to the grid. Any effects due to these secondaries will be equivalent to the effects due to the primaries. They can be accounted for by realizing that the effective electron current to the grid can be larger than the measured current.

3.b–Photons emitted from the grid by the bombarding ions can strike other parts of the grid, the collector, and the filament. These photons can cause the emission of electrons, ions (positive and negative), photons, and neutrals from these surfaces (again we are neglecting the walls). Since these particles are secondary to the photons (tertiary to the primary electrons) and not related to the gas pressure, the effects due to each should be investigated for possible "anomalies."

The electrons ejected from the collector by the impinging photons are drawn to the grid and give rise to the familiar "x-ray" limit. This effect has been commonly recognized and will not be discussed further.

The positive ions ejected from the collector will be drawn back to the collector. These can cause effects similar to those discussed above due to positive ion bombardment from the gas phase. These effects would be tertiary to the photon flux to the collector and will not be considered further.

Any negative ions emitted from the collector would give an effect similar to the "x-ray" effect, and probably substantially smaller.

Neutrals ejected from the collector would give an addition to the gas flow into the system. This could be important in the case of a heavily contaminated collector.

In a similar way, the particles ejected by photons striking the grid and filaments can be traced and the resulting effects determined.

Processes 3.c. (i and ii) and 3.d remain to be considered. 3.c.i. is of particular interest because it has been realized only recently that, under certain circumstances, this effect can give orders-of-magnitude error in the pressure as read by an ionization gauge. This process is of particular interest in the case of the Explorer XVII satellite, and is discussed in greater length in another section of this report.

The processes discussed above are indicated on Figure B-40. This discussion and the diagram illustrate the complexity of the processes in an ionization gauge; they also illustrate that, although the situation is complicated, all the possible processes can be predicted. Finally, as the probability of volume ionization continues to decrease as the pressure decreases, we see that some of the third or even fourth order processes can become important at low pressures.

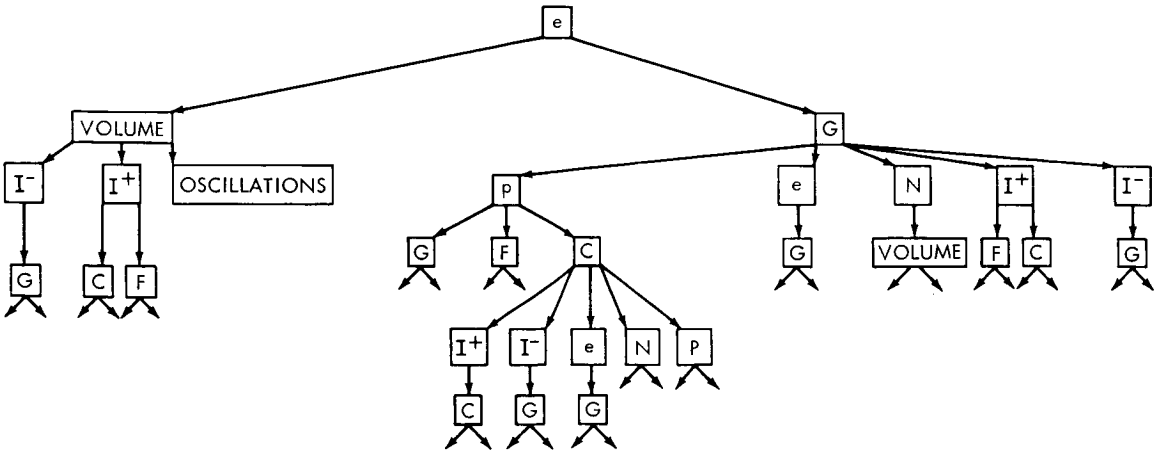


Figure B-40–Processes Caused by Electrons Emitted by the Filament

We shall not pursue the general discussion of anomalous effects further. The procedure outlined above, in following only the processes caused by the electrons emitted from the cathode, can be applied to the photons, ions, and neutrals also emitted from the cathode, and the importance of a given effect can be qualitatively determined. In particular, we note that no attention has been given the photons from the hot filament. A quick calculation shows that these may indeed contribute substantial errors under certain circumstances; for example, photon desorption of gases in cryogenically pumped systems at very low pressure may well be the primary source of gas, making hot-cathode gauges impractical to use in such systems.

Page intentionally left blank

MEASUREMENT OF OUTGASSING FROM SPACECRAFT

F. Owens and Y. Chen

OBJECTIVE

Before a spacecraft is launched, it is subjected to a series of environmental tests. The purpose of these tests is generally to observe the performance of the equipment and experiments aboard under the conditions the spacecraft will see in orbit. The object of our undertaking was to see if the present testing procedures could be expanded to include a determination of the gas-evolution pattern from a spacecraft, commonly referred to as outgassing. We were also interested in evaluating necessary instrumentation with emphasis on operating procedures and on methods of data acquisition and data analysis.

Emission of gases from a spacecraft in vacuum is due to sublimation of materials or desorption of gases from the body of the spacecraft in high vacuum. The value of an outgassing analysis in a test procedure lies mainly in the fact that a knowledge of gases coming from spacecraft can provide some idea of what is happening inside the spacecraft; in particular, it may indicate degradation in components in spacecraft, or some other abnormal process such as a leak in the battery. Also, the evolving gases are a factor in the corona-discharge phenomenon. Furthermore, the evolving gases could contaminate the spacecraft; for example, the gases could return and form layers on lenses changing their transmitting properties. These considerations make an outgassing analysis of interest.

DESCRIPTION OF EQUIPMENT

A mass spectrometer was used in conjunction with a large vacuum chamber. The spectrometer is set up so as to be able to detect the residual gases in the chamber.

A Stokes 8 by 8-foot thermal vacuum chamber was used to simulate the environmental conditions imposed on a vehicle during a typical launch and flight. The stainless steel thermal vacuum chamber is initially pumped down by a mechanical forepump which also backs a 32-inch and 4-inch vapor pump during high-vacuum operation. Between the chamber and vapor pumps are located the disc-type water-cooled baffle and the chevron-type liquid nitrogen baffle that aid pumpdown. A small mechanical pump acts as a hold pump to keep the vapor pumps evacuated at all times. The chamber has a temperature-control system which circulates trichlorethylene through tracing circuits on the chamber shell. Twelve individually adjustable tracing circuits allow balancing of temperature under non-uniform heat loads. The system is capable of maintaining chamber-wall temperatures between -65°C and 100°C.

The total pressure measurement in the chamber is made by an ion gauge which is atop the chamber. The operation of the ionization gauge is based on the fact that electrons from a filament ionize the entering gas molecules, and these ions are accelerated through a potential to a collector. When the electron current and electrode potentials are maintained at a constant value, the relationship between pressure and ion current can be sufficiently linear for practical measurements of pressure. When making measurements with an ion gauge, it is necessary to keep the electrode potentials constant, and more particularly to keep the electron current constant by controlling the filament current. As the

filament current required to give a certain electron current depends on the existing pressure and on the condition of filament, an electron current control is employed.

The spectrometer was a 21-613 Consolidated Electro-dynamics gas analyzer. The gas entering the spectrometer is ionized by electrons from a filament. These ions are then accelerated through a potential difference into a magnetic field which is perpendicular to their path. The magnetic field deflects the ions to a collector. There is also an electric field perpendicular to the magnetic field which causes the ions to move in a cycloidal path. This enables the ions to be deflected 360 degrees rather than 180 degrees, giving a better resolution. Each particular $m/_e$ requires a different voltage V to be deflected through a fixed radius R. If $m/_e$ is the mass-to-charge ratio, and B the flux density of the magnetic field, then

$$m/_e = \frac{2R^2B^2}{V}$$

The product $m/_e \cdot V$ is then a constant. A plot of $m/_e$ versus accelerating voltage necessary to have a beam deflect in a fixed radius is shown in Figure B-41 for our spectrometer. The partial pressure is thus measured by the number of ions arriving at the collector. This current reading is read on a chart recorder. The amount of deflection or peak height is proportional to the partial pressure of a particular gas. The proportionality constant then reflects the sensitivity of the instrument for a specific gas. The sensitivity S_K for a particular gas is given by

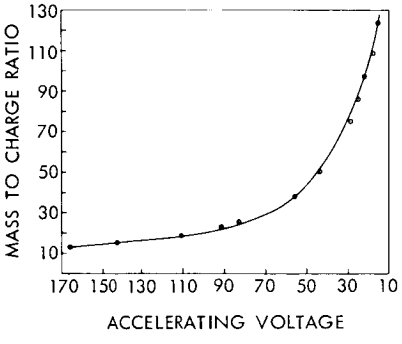


Figure B-41—Mass-to-Charge Ratio vs. Voltage

$$S_K = \frac{\Gamma_K L I_e Q_K}{kT}$$

where Γ_K is the percent of ions collected which are formed by the electron beam current I_e, Q_K is the ionization cross section, L is the length of the electron beam parallel to mass spectrometer slit system, T is the temperature in the ionization chamber, and k is Boltzmann's constant. This shows the factors on which the sensitivity of a spectrometer depends. The dependence on electron beam current and temperature should be noted. The raw data obtained from a scan are peak height and voltage.

EXPERIMENTAL WORK

The experimental aspect of the undertaking consisted of two parts: The first was the actual measurement of a gas-evolution pattern from a spacecraft; the second involved the development and performance of an in-test calibration of the spectrometer.

Outgassing Measurement

In the measurement of a gas-evolution pattern, the experimental procedure involved scanning an empty vacuum chamber with the spectrometer at pressures in the order of 10^{-6} torr, at three temperatures -15°C, +25°C and +35°C. The high and the low temperatures correspond to the extremes of temperature the spacecraft sees in orbit. Then the

chamber was scanned with an actual spacecraft in the chamber (IMP, Interplanetary Monitoring Platform). We also observed the total amount of outgassing by noting ion-gauge reading with and without spacecraft, as well as the behavior of some specific $m/_e$ peaks over a period of time.

Our results show that peak heights with spacecraft in the chamber increase considerably over empty chamber scans. Figure B-42 shows a few typical peak heights with and without spacecraft. Approximately 75 $m/_e$ values were detected at +35°C with spacecraft in chamber; at 35°C in the empty chamber, 45 peaks were present. Thus, there were 30 peaks due entirely to the spacecraft material.

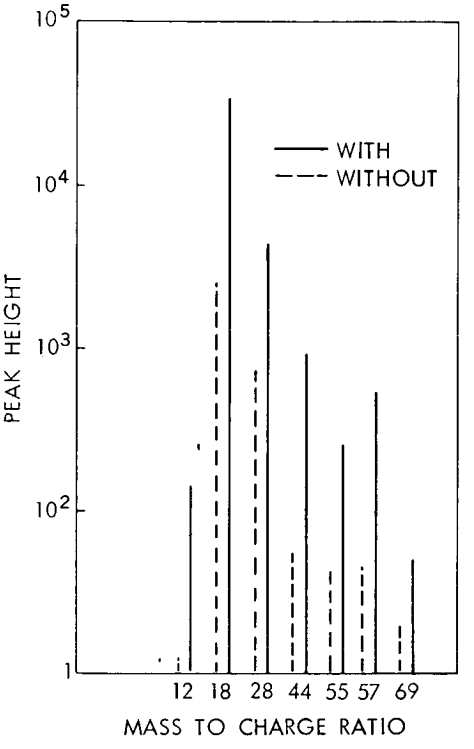


Figure B-42–Comparison of Peak Heights With and Without Spacecraft at 38°C

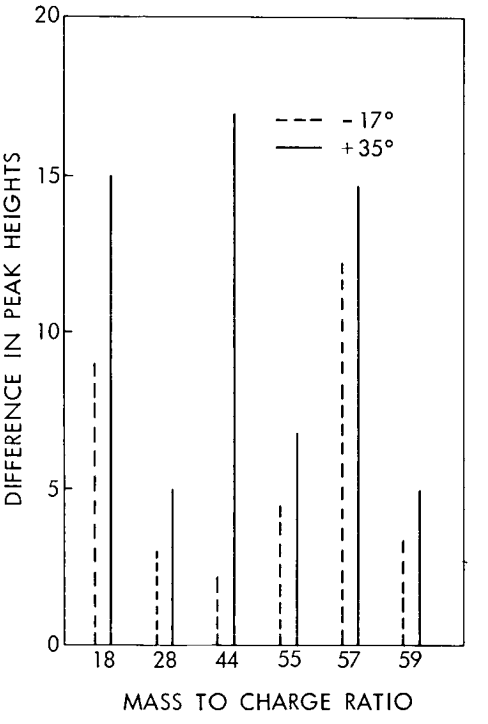


Figure B-43–Comparison of Difference in Peak Heights With and Without Spacecraft at 35° and -17°C

Figure B-43 shows the difference between peak heights with and without spacecraft at 35°C and -17°C. It is clear that the amount of gas coming from spacecraft is greater at higher temperatures. This can be seen from another point of view when we look at the steady-state total pressure as read on the ion gauge with and without spacecraft at the different temperatures. Figure B-44 shows a comparison of segments of the pumpdown curves at steady state with and without spacecraft at +35°C and -17°C. We can calculate the quantity Q of gas coming from spacecraft by

$$Q = (\Delta P)\, S.$$

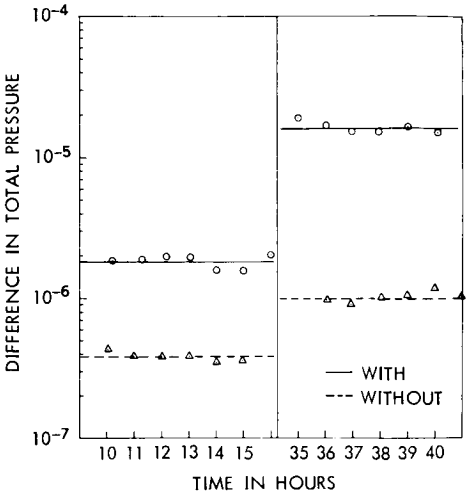


Figure B-44—Comparison of Difference in Total Pressure With and Without Spacecraft at 35° and -17°C

ΔP is the difference in pressure of chamber with and without spacecraft, and S is the pumping speed of the vacuum system (5.0×10^3 liters/second).

At -17°C we have

$$Q = (1.8 \times 10^{-6} - 3.8 \times 10^{-7}) \text{ torr} \times 5.0 \times 10^3 \text{ liters/sec}$$

$$(1.42 \times 10^{-6}) \ (5.0 \times 10^3) \text{ torr - liters/sec}$$

$$= 7.1 \times 10^{-3} \text{ torr - liters/sec}$$

At +35°C

$$Q = (1.6 \times 10^{-5} - 1 \times 10^{-6}) \text{ torr } (5.0 \times 10^3) \text{ liters/sec}$$

$$= 7.5 \times 10^{-2} \text{ torr - liters/sec}$$

Thus, about 10 times as much outgassing takes place at the higher temperature.

The principal conclusions drawn from this analysis are that the use of a spectrometer is feasible in determining the outgassing pattern from a spacecraft, and that the instrument clearly shows up the various characteristics of the outgassing.

Calibration

The second part of the experimental work involved a determination of a calibration of the instrument: i.e., a determination of the relationship between partial pressure of a specific gas and the number of divisions of deflection on the recorder. This calibration was designed so that it could actually be incorporated into a complete spacecraft-testing procedure without appreciably disturbing the spacecraft-testing program. This was done to enable the response of the instrument to be checked from time to time during the period of spacecraft testing. The calibration procedure used to answer these purposes is the following:

1. Pump down the vacuum chamber to a steady-state pressure of about 10^{-6} torr.
2. Set the wall temperature at -15°C.
3. Obtain a complete evolution pattern with spectrometer.
4. Read the nitrogen peak on the spectrometer.
5. With a calibrated leak, introduce enough nitrogen to raise pressure to 1.5×10^{-5} torr.
6. Read the nitrogen peak on the spectrometer.
7. Introduce more nitrogen to raise pressure to 4.0×10^{-5} torr.
8. Read the nitrogen peak on the spectrometer.

The calibrated leak was a Granville-Phillips variable-conductance valve; such valves are variable from conductances of 1 liter/sec to 10^{-14} liters/sec. The choice of numerical values is somewhat arbitrary within three limitations: first, that not too much nitrogen be introduced, so that the nitrogen peak will not be off-scale on the recorder; second, that the pressure in the chamber not exceed 10^{-4} torr, as this is the limiting operating pressure for the spectrometer filament; third, that ΔP not be too small. The -15°C wall temperature was chosen because less background gas would be present. Step 3 is incorporated to provide a check on the makeup of this background. The sensitivity is obtained by noting the change in the nitrogen peak heights corresponding to change in pressure on the ion gauge:

$$C = \frac{P_F - P_i}{H_F - H_i}$$

where P_i and H_i are pressure and peak height before leak. The nitrogen sensitivity obtained was 4.3×10^{-10} torr/div. Once the absolute sensitivity of nitrogen is determined, the absolute sensitivities of the other gases can be obtained by using relative sensitivities of the other gases to nitrogen. These are available.[2,3] Note that the determined sensitivity is valid only for a given filament current, i.e., the electron current necessary to ionize the molecules. It is important to realize that a different filament current will yield a different sensitivity. We were operating at 50 ya. Also, this calibration is for a definite ion-chamber temperature. Once the leak of nitrogen was stopped, it took only about a minute for the pressure of the chamber to return from 4×10^{-5} to a steady state at 1×10^{-6}. The pumping speed for nitrogen is apparently very high. This point is significant because an in-test calibration check can be performed with very little loss of time to the spacecraft-testing program. A spectrometer calibration check would take at most about 30 minutes.

DATA ANALYSIS

Our purpose in discussing methods of data analysis is a part of our object to put the spectrometer to use in analyzing the outgassing of spacecraft. Our specific aim in this section is twofold: first, to demonstrate that it is possible to make a simple direct determination of partial pressures for the predominant gases evolving from the spacecraft; second, to establish the basis for a more complete analysis by means of a computer program. Here no attempt at completeness is made, but rather a starting point or foundation is set for further study in this area.

A Simple Analysis

A quantitative analysis is of importance because it can give us some idea as to the order of magnitude of the evolving gases. An exact quantitative analysis is made difficult by the fact that, at any one peak, more than one gas may be present, each with a different sensitivity. For example, mass number 28 can result from nitrogen, carbon monoxide, and ethylene. Difficulty is further encountered because of the existence of "cracking." Certain compounds are dissociated at the hot filament of the spectrometer, the amount of dissociation depending on the particular instrument. The distribution of the cracked components of a compound with respect to the parent (largest peak of that compound) is called the spectral pattern of the material. Each compound has a unique spectral pattern. It distributes its components in a particular way over a number of $m/_e$ values. Table B-18 is a typical chart of spectral-pattern data. The spectral pattern coefficient, the numbers in this table, is defined by

$$A_{KN} = \frac{\text{height of peak at the Nth spectral line for the Kth gas}}{\text{height of the parent peak for the Kth gas}}$$

Thus, any peak may be a mixture of contributions from different gases, each gas having a different sensitivity. The height of any peak is the sum of the heights of the contributing gases at that peak.

If H_N is the height of the Nth peak in divisions, P_K is the total partial pressure of the Kth gas in torrs: A_{KN} is the pattern coefficient of the Kth gas at the Nth peak; and CS_K is the absolute sensitivity of the Kth gas in torrs/division (C is the absolute sensitivity of nitrogen, S_K is sensitivity of Kth gas relative to nitrogen). Thus we have

$$H_N = \sum_{K=1} \frac{P_K A_{KN}}{C S_K}$$

This provides a set of simultaneous equations enabling solution for P_K . Note that, if is greater than N(i.e., more gases than peaks), the equations cannot be solved. In general this would not be the case. An approximate analysis using the predominant peak heights was performed. The gases used were CO , CO_2 , N_2 , H_2O and O_2. All other gases present were neglected. The S_K and A_{KN} are obtained from Table B-18. C , the absolute sensitivity of nitrogen, is obtained by a calibration procedure such as previously outlined. We define

$$a_{KN} \equiv \frac{A_{KN}}{C S_K}$$

These will be the coefficients of the partial pressures of the gases in the equations and are obtained simply by dividing the absolute sensitivity of the gas into the spectral pattern coefficient. To solve a particular problem for a limited number of gases, let us reproduce the spectral pattern coefficient table just for the gases we have.

Table B-18
Spectral Pattern Data

A_{KN}

N	K	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Gas / Mass	Hydrogen	Helium	Neon	Carbon Monoxide	Nitrogen	Argon	Carbon Dioxide	Ammonia	Water	Sulfur Dioxide	Air	Xenon	Krypton	Oxygen
1	2	100.00							.25	.10		.05			
2	4		100.00												
3	12				3.21			5.80							.11
4	13				.03i		.02	.08i							
5	14				1.23d	7.90			2.47			8.12			
6	15								7.26						
7	16				1.02			8.04	92.02	3.07	3.60	1.88			9.23
8	17								100.00	27.01		.11			
9	18								2.25i	100.00		.53			
10	19									.19i					
11	20			100.00			17.08d					.20			
12	21			.33i											
13	22			9.90i				2.10d							
14	24										1.90d				
15	27					.94									
16	28				100.00	100.00		8.85				100.00			2.02
17	29				1.10i	.88i		.08				1.01i			
18	30				.25i										
19	32										8.80	21.02			100.00
20	33										.10i				.09i
21	34										.41i				.41i
22	36						.36								
23	38						.07								
24	39													.22d	
25	40						100.00					1.45		1.16d	
26	41													5.71d	
27	41.5													5.64d	
28	42													27.89d	
29	43												.58t		
30	44							100.00				.17	.53t		
31	45							1.14i							
32	46							.41i							
33	48										51.35				
34	49										.46				
35	50										2.42i				
36	64										100.00		1.27d		
37	64.5												18.28d		
38	65										.85i		2.83d		
39	65.5												14.58d		
40	66										4.79i		18.42d		
41	67												7.12d		
42	68												6.02d		
43	78													.65	
44	80													4.14	
45	82													20.70	
46	83													20.19	
47	84													100.00	
48	86													29.71	
49	128												7.45		
50	129												100.00		
51	130												15.50		
52	131												77.76		
53	132												97.00		
54	134												36.06		
55	136												29.45		
S_K	Relative Sens.*	1.7	5.3	3.5	.99	1.0	.86	.89	1.8	1.3	1.6	1.4	2.8	1.6	1.4

SYMBOLS

i = isotope peak

m = metastable ion (diffuse peak)

d = doubly-charged ion

t = tripley-charged ion

*To obtain true sensitivity, multiply relative sensitivity by value obtained for nitrogen.

Table B-19
Spectral Pattern Coefficients for Predominant Gases Used, A_{KN} (X100)

N	K	1	2	3	4	5
	gas / m/e	CO	N_2	CO_2	H_2O	O_2
1	12	3.21	*	5.80		
2	14	1.23	7.90			
3	16	1.02	8.04	8.04	3.07	9.23
4	17				2.70	
5	18				100.00	
6	28	100.00	100.00	8.85		2.02
7	32					100.00
8	44			100.00		

*Note that the entries in the empty boxes are zeros.

Table B-20
Coefficients of Partial Pressure for Predominant Gases, A_{KN}

N	K	1	2	3	4	5
	gas / m/e	CO	N_2	CO_2	H_2O	O_2
1	12	0.65×10^8		1.3×10^8		
2	14	0.25×10^8	1.6×10^8			
3	16	0.21×10^8		1.8×10^8	0.47×10^8	1.30×10^8
4	17				4.2×10^8	
5	18				15.4×10^8	
6	28	$22.\times10^8$	20.0×10^8	1.97×10^8		0.29×10^8
7	32					14.3×10^8
8	44			22.4×10^8		

Now, if we divide each of these elements in Table B-19 by their respective absolute sensitivity given at the bottom of Table B-18, we then have the coefficients in the simultaneous equations. For example,

$$a_{11} = \frac{.0321}{(5 \times 10^{-10})\ (.99)}$$

We are using 5×10^{-10} torr/div as the nitrogen sensitivity. The result of this calculation for our limited number of gases is shown in Table B-20.

Once the sensitivity has been determined, assuming it does not change too much in the course of operation, the values of a_{KN} in Table B-18 are fixed so that the only change from scan to scan is the peak height values H_N. The simultaneous equations resulting from Table B-18 are immediately written.

$$H_1 = 0.65\times10^8 P_1 + 1.30\times10^8 P_{CO_2}$$

$$H_2 = 0.25\times10^8 P_1 + 1.60\times10^8 P_2$$

$$H_3 = 0.21\times10^8 P_1 + 1.8.10^8 P_3 + 0.47\times10^8 P_4 + 1.30\times10^8 P_5$$

$$H_4 = 4.20\times10^8 P_4$$

$$H_5 = 15.4\times10^8 P_4$$

$$H_6 = 22.0\times10^8 P_1 + 20.0\times10^8 P_2 + 1.97\times10^8 P_3 + 0.29\times10^8 P_5$$

$$H_7 = 14.3\times10^8 P_5$$

$$H_8 = 22.4\times10^8 P_3$$

A sample calculation is exhibited by using one set of data of H_N (taken with the presence of IMP):

$$H_1 = 122,\quad H_2 = 93,\quad H_3 = 180\quad H_4 = 1{,}350,$$

$$H_5 = 5{,}000,\ H_6 = 1{,}265,\quad H_7 = 250,\quad H_8 = 180.$$

Since there are five unknowns $P_1, P_2, \cdots P_5$, we have to choose from the eight values of H_N only five values to form a system of five equations. If we take H_1, H_2, H_4, H_7, H_8, for example, we then have the following system:

$$122 = .65 \times 10^8 P_1 + 1.30 \times 10^8 P_3 \quad (1),$$

$$93 = .25 \times 10^8 P_1 + 1.60 \times 10^8 P_2 \quad (2),$$

$$1{,}350 = 4.20 \times 10^8 P_4 \quad (3),$$

$$250 = 14.3 \times 10^8 P_5 \quad (4),$$

$$180 = 22.4 \times 10^8 P_3 \quad (5).$$

We observe that in the above system the value of P_3, P_4, and P_5 can be obtained immediately. After that, P_1 and P_2 can be solved from (1) and (2) respectively. The answers are:

$$P_1(CO) = 60.4 \times 10^{-8} \text{ torr.},$$
$$P_2(N_2) = 47.5 \times 10^{-8} \text{ torr.},$$
$$P_3(CO_2) = 8.0 \times 10^{-8} \text{ torr.},$$
$$P_4(H_2O) = 321.4 \times 10^{-8} \text{ torr.},$$
$$P_5(O_2) = 17.4 \times 10^{-8} \text{ torr.},$$

By summing up all the values of P, we obtain a total pressure of 4.54×10^{-6} torr as compared with the measured value of 1.80×10^{-6} torr. This does not seem to be a bad check. However, one must be guarded against drawing strong conclusions from such a calculation, because of the fact that many unknown gases, whose existence in the chamber is evidenced by the spectral diagram, have been ignored in the calculation. On the other hand, the result of calculation does give information about the order of magnitude of the outgassing. As our ability to identify other existing gases grows, more complete analysis can be done by using the digital computer program. To summarize, a procedure for a simple calculation is:

1. Reproduce the spectral Table B-16, for limited number of gases that are being considered.

2. Divide each element of the table by the absolute sensitivity of the gas. The absolute sensitivity $S_K C$ is the product of the relative sensitivity of nitrogen (bottom of Table B-16) times nitrogen sensitivity (5×10^{-10} torr/div). These new elements are then the coefficients of the partial pressures of the respective gases.

3. Write the simultaneous equations and solve.

The Digital Computer Program:

In calculating the partial pressures of the gases in a gas mixture from the data resulting from spectrometer measurements, the following formula is used:

$$H_N = \sum_{K=1} \frac{P_K A_{KN}}{C S_K}$$

The notations have been explained previously. Such calculations can best be done by a digital computer. To this end, a digital computer program is written for the purpose of quantitative analysis. Since information is incomplete as to the precise description of the content of a test chamber, especially when the spacecraft is present, a tentative dimension of 30 is assigned to the P_K - matrix. This number can be changed easily to suit the need. A tentative dimension for H_N is taken to be 60.

It must be pointed out that a quantitative analysis of the partial pressures will give exact answers only when the nature of the individual gases in the mixture is completely specified. Otherwise, the analysis is approximate, and it gives order of magnitude only. For diagnostic purposes this order-of-magnitude analysis may be just what is necessary.

Only further studies can elucidate this point. It becomes clear in the discussion that the outgassing pattern of the components of a spacecraft is a prerequisite to any quantitative analysis of the outgassing of the complete spacecraft.

To illustrate what the digital computer program does, we begin by assuming CO, N_2, CO_2, H_2O, O_2, Ar, methane, and acetone to be the total content of a certain mixture of gases in a test chamber. There are altogether 22 H_n's which cover all the contributions from these 8 gases. After assigning a K - index to each gas and a N - index to each ($m/_e$) - value, we make up a tabulation of A_{KN}'s for the digital program. Since the number of H_N's is usually larger than that of the unknown P_K's, in order to form K equations the computer actually selects K values of H_N from the complete list. Theoretically speaking, if we know the exact description of a mixture, then taking any set of K values of H_N will result in a unique answer. This is not the case if the precise break-down of the constituents of the mixture is unavailable. Criteria must be set to make the selection. A reasonable criterion is to take the first K values of an ordered list of H_N's. The program actually orders the H_N values and then selects according to this criterion.

If the selected system of K equations turns out to possess a vanishing determinant, the program provides an alternate route to form other systems with non-vanishing determinants. Normally, it should be a rare occurrence to have vanishing determinants when the A_{KN} table contains many organic compounds as entries.

Summarizing, it is clear that quantitative analysis of data from spectrometer is definitely feasible with the aid of a digital computer; however, its limitation must also be recognized. The analysis gives exact partial pressures only when the mixture is precisely specified as to its constituents. Without this break-down the analysis at best can be only approximate. Clearly, one of the many problems for further study is to determine the outgassing pattern of typical components of spacecrafts. Foam products are an example in this area. When component information is made available by further experimentation, the quantitative analysis of the synthesized pattern of the complete spacecraft will be made more meaningful.

DISCUSSION

Having established the spectrometer as a diagnostic tool in a spacecraft outgassing analysis, we seek now to offer some precautions and recommendations regarding its use, as well as some comments on outgassing analysis in general.

In using the spectrometer, certain precautions must be considered. The effect of scanning rate on peak height should be noted. The scanning rate (rate at which different $m/_e$ values pass the collector) is controlled by the rate of change of the accelerating voltage. This voltage selects the $m/_e$ that will be deflected through a fixed radius to a collector slit. If our rate of scan is too fast, the beam corresponding to a particular $m/_e$ will sweep over the collector slit too fast, so that a representative amount of ions will not be collected and the height of the peak will not be at its optimum value. On the 21-613 CEC residual gas analyzer, the scanning can be operated either manually or automatically. The automatic scan peak heights should be checked against a slow manual-scan result to see that the automatic scan is not too fast. This would be indicated by the peak heights being larger on the manual scan. This dynamic response should be checked from time to time. Floor vibration can be a source of difficulty. This can cause a vibration of the beam, so that fewer ions will be collected.

Care must be taken when spectral-pattern data are used. The cracking pattern of a pure gas varies from one instrument to another and, in a given instrument, depends on

precise operating conditions. Strictly speaking, to use spectral-pattern data, the instrument used should be operating under the same conditions under which the spectral-pattern data have been obtained. The instrument should have the same ionization current and filament temperature. It has been reported[4] that, over long periods of operation (weeks), cracking patterns show variations. Changing a filament usually changes the electron-beam position and often the cracking pattern.

The sensitivity of the instrument is also a function of operating conditions such as heater current and ionization current. Variations of 5 percent from week to week have been noted in some instruments.[5] It is for this reason that an in-test calibration has been discussed, so that the calibration can be checked from time to time throughout a spacecraft test. It should be emphasized that a particular value of the nitrogen sensitivity holds only for a given ionization current.

The overall purpose of an outgassing test is to provide information as to what is happening inside the spacecraft. Such information can be gained from a quantitative determination of the amount of gases coming off the spacecraft and an identification of their types. In the quantitative approach, the sum of the partial pressures of the gases in the chamber is compared to the total pressure on the ion gauge. There are some inherent difficulties in such a comparison: the sensitivity of any given design of gauge depends on the nature of the gas used, mainly because each gas has a unique ionization potential.[6] In an outgassing analysis in a large chamber, the total gauge reading is the response to a mixture of gases. As the sensitivity of a gauge depends on the specific gas used, the sensitivity of a gauge to a mixture of gases depends on the kinds of gases present, and the number of kinds present is not constant (some gases are pumped out quicker than others). Thus, the significance of a total gauge reading of a mixture is not clear. The gauge is usually calibrated only for a limited number of gases; the validity of extending this calibration to read a complex mixture of gases is questionable. When comparing total gauge pressure to sum of partial pressures obtained on spectrometer, the question should be raised as to whether the gauge has the ability to ionize all the gases the spectrometer can.

The quantity of gas evolving from the spacecraft can be calculated by noting steady-state pressure before and after introduction of spacecraft.

$$Q = (\Delta P)\, S,$$

where (ΔP) is the pressure difference between empty chamber and chamber with spacecraft. Q is the quantity of gas then evolving from spacecraft. Such a calculation, however, has an inherent problem in it.[7] Since the mean free path of molecules at pressures of 10^{-6} torr is much larger than the dimensions of the chamber, the ion gauge is measuring not just a representative number of molecules coming from spacecraft but is measuring also an excess due to molecules that have rebounded back and forth between wall of chamber and spacecraft. A calculation of the quantity of gases coming out by the difference in total pressure reading is an overcount.

In the present effort, H_2 and He were neglected in the measurement. However, since we are fairly certain about their presence in the chamber, measurements should be taken in further studies.

GENERAL CONCLUSIONS

1. The results of measurements performed show that there is considerable outgassing from a spacecraft in a vacuum chamber. This fact establishes the basis of using outgassing measurements as diagnostic tools.

2. The mass spectrometer is an instrument capable of giving an order-of-magnitude determination of the residual gases inside a test chamber. The nature and the type of the evolving gases can be ascertained to some degree.

3. More meaningful quantitative information can be obtained with further studies of component outgassing patterns. A computer program has been prepared for a quantitative analysis.

4. With qualitative and quantitative information made available by outgassing measurement, the phenomenon of corona discharge can be studied to determine the relationship between type of gases and probability of occurrence of corona discharge.

5. The spectrometer can detect certain equipment failure aboard the spacecraft. For example, a sudden increase in the peak of a certain gas might indicate a leak such as a battery or pressurized recorder.

6. The nature of component degradation can be studied by an outgassing analysis.

7. Studies of self-contamination can be undertaken. In space simulation, the evolving gases very often return to the spacecraft because of the rebound effect. Here the identity of the gases can provide useful information.

FURTHER STUDY

Our study is in no sense complete and at the most sets the basis for further work in this area.

More data are needed on the evolution pattern of spacecraft with the aim of getting a typical spectral pattern if such exists. The effect of temperature cycling on the outgassing should be looked at. Questions such as how long it has been exposed to atmospheric pressure before testing are important.

The best possible placement of outgassing detectors with respect to spacecraft should be considered. The response of the spectrometer should be considered as a function of its distance from the outgassing source.

Thought should be given to an operational procedure for making maximum use of instrumentation; for example, the peak heights and voltages could be read directly into a data central computing system, and a partial pressure of the various gases be instantaneously calculated and read out. Methods of monitoring various peaks should be considered. The operational procedures should incorporate checks on the various precautions mentioned, such as an in-test calibration.

The program of actual quantitative analysis for which the basis has been established here should be further developed. For this to be done best, the identity of all outgassing materials should be known and a catalogue of their spectral patterns made.

The authors wish to thank, for their help in this undertaking, Ronald Sheffield of the Thermodynamics Branch and Erik Schultz, National Science Foundation summer student.

REFERENCES

1. Caswell, H. J., "Analysis of Residual Gases in Several Types of High-Vacuum Evaporators," IBM Journal, (April 1960)

2. "Operation and Maintenance Manual, Type 21-613 Residual Gas Analyzer," Consolidated Electrodynamics Corporation, Pasadena, California

3. "Catalog of Mass Spectra Data," American Petroleum Institute Research Project 44

4. Robertson, A. J., "Mass Spectrometry," p. 89, Wiley and Sons

5. Nicholson, D. L., "Some Factors Affecting Mass Spectrometer Analysis," Institute of Petroleum, London (1952) p. 99

6. Dushman, S., "Scientific Foundation of Vacuum Technique "p. 321, John Wiley and Sons

7. Santeler, D. J., "Pressure Simulation of Outer Space," Vacuum Symposium Transaction, American Vacuum Society (1959)

8. Caruso, S. V., and Looney, W. C., "A Study of the Outgassing and Evaporation Products of Some Materials upon Exposure to Reduced Pressure," George C. Marshall Space Flight Center, MTP-p & VE-M-62-7

ANALYSIS OF SPACECRAFT FAILURES DURING THERMAL-VACUUM TESTING

Kenneth L. Rosette

INTRODUCTION

The primary objective of the scientific satellite is to make fundamental measurements of various types of phenomena that cannot be adequately measured from the earth; for example: energetic particles, cosmic rays, micrometeoroid flux, and magnetic-field strength.[1]

High system reliability is a prime requirement for an effective space program. This requirement is underscored by the high cost of launch vehicles and spacecraft. Also, launch opportunity for certain space studies is dependent on interplanetary relations; therefore, success on the first launch attempt is necessary for timely acquisition of needed space data. One of the most effective techniques of increasing the reliability of our space systems is the application of laboratory tests to simulate, as nearly as possible, the many environmental conditions which will be encountered by the spacecraft.[2]

This report contains the results of a study of the laboratory tests of three space programs: Ariel-1 (S-51), Ariel-2 (S-52), and the Interplanetary Monitoring Platform (IMP).

TEST PHILOSOPHY

The test philosophy[1] employed has been the use of environmental tests to gain information from which the suitability of a spacecraft for flight can be assessed. To this end, the systems test program for a scientific satellite has six goals:

1. Verification that novel or unproven designs meet performance requirements and have a satisfactory life expectancy
2. Verification that particular samples of previously employed hardware are suitable in a new application
3. Elimination of defects in design, material, or workmanship
4. Discovery of unexpected interactions between subassemblies when the system is exposed to environmental stress
5. Training of personnel to be responsible for the satellite at the launch site and for data reduction and analysis
6. Generation of information that will serve as a guide in making new designs and in assessing their reliability

THE SCIENTIFIC SPACECRAFT

The following is a brief description of the types of subsystems which a typical scientific satellite comprises. Referring to the electronics only, there are two main groups: (1) experiments and (2) housekeeping or support electronics.

1. Experiments

As previously mentioned, the prime purpose of the scientific satellite is to place in orbit the instrumentation necessary to measure several different types of phenomena which cannot adequately be measured from earth. The experiments may vary in number from three to as many as seven or eight per spacecraft.

2. Housekeeping or Support Electronics

To operate the experiments, and transmit the data, a complex system of support electronics is necessary. These may be divided into the following main groups:

Data Link	Power System
1. Encoders	1. Solar array
2. Programmers	2. Batteries
3. Tape recorders	3. Dumping circuit
4. Transmitters	4. Converters
5. Command receivers	5. Regulators
6. Decoders	6. Undervoltage detectors
7. Data storage	7. Recycle timers

TYPICAL SYSTEMS TEST

A typical systems thermal-vacuum test consists of the following:

Setup Phase

a. Attachment of the spacecraft to a mounting fixture

b. Installation of the necessary electrical leads required to operate the spacecraft and monitor its performance

c. Attachment of temperature sensors in predetermined positions

d. Installation of various sources (ultraviolet lamps, radioactive material, etc.) to exercise the various experiments

Ambient Checkout

After completion of the test setup, a complete systems checkout is conducted to ensure proper operation of the spacecraft before starting the test.

Corona Check

This check is conducted to ensure that the spacecraft, while operating, will survive the effect of pressure reduction during the launch phase. The spacecraft is operated in

the same manner that it will be during launch, while the test facility pressure is reduced. This test is normally conducted between the range of 760 torr to 1×10^{-4} torr.

Cold Soak

With the spacecraft not operating, its temperature is reduced to the predicted minimum mean orbital temperature (0°C to -20°C). When stabilization occurs, the spacecraft is turned on and operated in the same manner in which it will function in orbit. The duration of this test is normally 3 to 4 days.

Hot Soak

Upon completion of the cold soak, the temperature of the spacecraft is increased to the predicted maximum mean orbital temperature (45°C to 60°C) and the same mode of operation as was used during the cold soak is continued. The duration of this test is usually 4 to 6 days.

Temperature Cycling

In this test the temperature of the spacecraft is cycled, while operating, between the maximum and minimum temperature levels. This type of test provides the experimenter with information on the operating characteristics of his experiment during periods of transient temperatures.

THE FAILURE STUDY

The foregoing has been a brief description of the test item, the type of environmental test, and the test philosophy. The results presented in this report are based on a study of three space programs: the Ariel 1 (S-51), the Ariel 2 (S-52) and the Interplanetary Monitoring Platform (IMP).

For the S-51 and S-52 programs, a prototype and two flight spacecraft were available. In each of these programs the units were tested in succession. The S-74 program contributed a prototype and a flight system. In this case, the prototype was launched into orbit and the flight system was tested about 6 months later.

A breakdown of the number of failures encountered in the various subsystems is given in the following table:

Table B-21
Subsystem Failures in Three Programs

System	Experiments	Tape Recorders	High Voltage Supply	Dumping Circuit	Transmitter
S-51	7	7	4	5	0
S-52	3	1	4	0	0
S-74A	10	-	0	0	1
S-74B	8	-	0	0	3
Totals	28	8	8	5	4

Other subsystems in which failures were encountered were:

	Combined (for 8 spacecraft)
1. Harness	1
2. Performance parameters	1
3. Optical aspect sensors	1
4. Thermal coating	2
5. Undervoltage detectors	3
6. Recycle timers	2
7. Converters	2
8. Batteries	1
Total	13

From the above tables it appears that the experiments are the most prominent sources of failure in the thermal-vacuum environment. Of a total of 66 failures for eight systems, 28 or 43 percent were encountered in the experiments.

Referring to the experiment subsystems, the following is a breakdown of the types of failures encountered during test of eight spacecraft:

Component failure	10
Design problem	6
Mechanical	3
Unknown (not repeated on bench check)	6
No action	3
Total	28

From this table, it is apparent that component failures (resistors, capacitors, etc.) account for nearly 36 percent of the experiment problems.

To go one step further, the following table is a breakdown of the number of failures vs. temperature for the experiment subsystems.

Table B-22
Experiment Subsystem Failures vs. Temperature

	Temperature		
	Room	Hot	Cold
Component	0	2	8
Design	1	2	3
Mechanical	0	3	0
Unknown	0	3	3
No action	1	1	1
Totals	2	11	15

From this, it appears that the cold environment is more severe in uncovering the weak links. However, this may not be necessarily true; the bit factor may be time. In most spacecraft testing, the cold soak is conducted first, because it is easier to obtain lower test pressures and, at the end of the high-temperature test, no warm-up period is required before the test facility can be opened.

In most cases, the failures which were encountered appeared within the first 2-1/2 to 3 days. In a few cases, failures appeared after 4 or 5 days. An example of this was the battery-charging and protection circuitry of the S-51 prototype, in which a temperature problem appeared about the fifth day. This, however, was because it was the first time the item had been operated at its maximum level.

CONCLUSIONS

Present subsystem test specifications call for 24 hours at each temperature extreme. On the basis of the results of these spacecraft tests, it appears that a series of comprehensive subsystem tests of at least 3 days at each level would reduce the number of failures encountered during the system testing.

A preferred parts list recently published by the Quality Assurance Branch lists the most suitable, or reliable, components for spacecraft use. With the incorporation of this list, and tighter manufacturing and design procedures, more reliable spacecraft systems will be built.

SOLAR SIMULATION

The foregoing study was based on the results of the thermal-vacuum tests conducted on eight spacecraft systems. Recently, solar simulation testing has become an integral part of the spacecraft evaluation program. This type of test is conducted to evaluate the thermal design of the spacecraft; that is, to verify that the computer program established to predict thermal conditions in orbit is correct.

The spacecraft covered in this report, which have been subjected to this type of testing are the Ariel-2 prototype and the IMP flight system.

In this type of test, the spacecraft is attached to a gimbal system which is within a vacuum facility. A liquid nitrogen shroud provides the "black space." Twin carbon arcs directed through a quartz window simulate the energy of the sun.

The gimbal system has the capability of varying the spin-axis/sun-line angle by ±60 degrees. This system also provides the capability of rotating the spacecraft about its spin axis at its orbital rate.

During the solar test, the spacecraft is operated in its orbital mode. Slip rings provide a method of data acquisition and spacecraft control.

The duration of this test depends on the spacecraft thermal time constant; the test is continued until equilibrium is established for a given sun/spin-axis angle. The solar simulation test of Ariel 2 was conducted at three different solar aspect angles (45, 90, and 135 degrees); the IMP-B spacecraft was limited to the 45-degree aspect because of the physical dimensions of the test facility.

Thus far, only one failure has occurred in the solar environment. During the 90-degree solar aspect test of Ariel 2, an adherence problem with the passive thermal-control paint was encountered. It is noteworthy that the solar test is conducted after completion of the thermal-vacuum isotropic test. Therefore, all the problems arising from design or workmanship should have been eliminated at this point. The types of failures which may still be encountered are the random component failure (resistor, capacitor, etc.) and variations in spacecraft performance caused by temperature gradients.

COMMENTS

For the purpose of statistical analysis, eight spacecraft are not sufficient. It is recommended that this type of study be continued and made a part of every spacecraft evaluation program. Much more data are required to strengthen the results obtained in this study.

The Quality Assurance Branch is now compiling failure data in the form of a malfunction report. This report is completed by an investigating team composed of the personnel responsible for the particular item. Although this type of data is an excellent source of information, it must be completely and accurately filed if maximum value is to be obtained.

An analysis of the pressure-time curves for the right spacecraft studied in this report will be made by Dr. John Simons. He will also comment on the failures experienced and further extend the failure study.

REFERENCES

1. Boeckel, John H., "The Purpose of Environmental Testing for Scientific Satellites," NASA Technical Note D-1900 (July 1963)

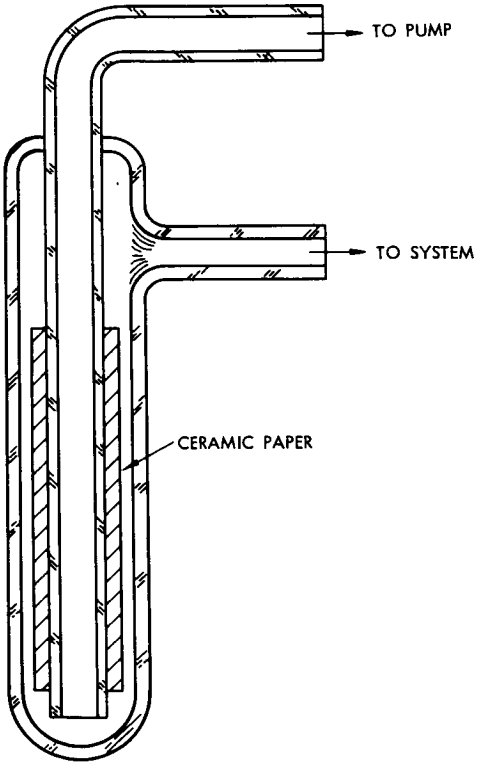
2. Timmins, A.R., and Rosette, K.L., "Experience in Thermal-Vacuum Testing Earth Satellites at Goddard Space Flight Center," NASA Technical Note D-1748 (August 1963)

APPLICATION OF CERAMIC-FIBER PAPER TO UHV TRAPS

Dr. F. M. Propst
Coordinated Science Laboratory

A ceramic paper manufactured by the Carborundum Corporation under the tradename of Fiberfrox Ceramic Fiber Paper has been applied to the construction of refrigerated and nonrefrigerated traps and baffles for high- and ultrahigh-vacuum systems. The paper is manufactured in grades 970-A, F, and J and 970-AH, FH, and JH, and contains approximately 51.2 percent Al_2O_3, 47.1 percent SiO_2, and small amounts of other materials. The grades 970-A, F, and J contain up to 5 percent organic binder. The paper has a felt-like texture and a large trapping area-to-volume ratio. The Al_2O_3 and SiO_2 are very stable and can be fired to high temperatures without decomposition. Complicated baffle and trap structures can be fabricated in a simple manner using the paper; also, the walls of pumping tubulation can be converted to effective trapping surfaces by lining them with the paper.

Figures B-45 and B-46 illustrate two of the baffles which can be constructed with this paper. The method of construction is obvious from the figures. The ceramic paper is prepared for the baffles by firing in air at 1,000 - 1,200°C for from 5 to 25 minutes.



← Figure B-45-Glass Reentry Traps

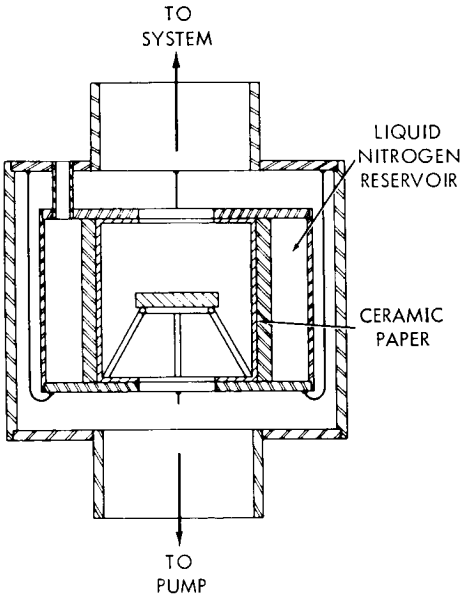


Figure B-46-Liquid-Nitrogen-Cooled Ceramic Trap

The performance of two vacuum systems using this material in the traps has been found to be exceptional. One system, using a trap similar to that illustrated in Figure B-45 and pumped by a glass two-stage oil-diffusion pump, has produced pressures of approximately 3 to 5×10^{-12} torr with the trap cooled to liquid nitrogen temperatures. This same system produced pressures of approximately 1 to 2×10^{-11} torr with the trap at room temperature. During the time of operation of this system, no increase in pressure was noted due to a saturation of the trapping element.

Preliminary work has been performed using this paper as a "filter" in the pumping lead; that is, the pumping lead was completely closed with a layer of this material, and the system evacuated through this layer. The results of this test were encouraging, in that the pumping speed through the paper was reasonably high and pressures in the 10^{-10} torr region were obtained.

This technique may be of importance in systems using cryogenic pumping in conjunction with oil-diffusion pumps. In such systems the noncondensable gases can diffuse through the paper (cooled to cryogenic temperatures) to the pump, whereas it would be extremely unlikely that oil or cracking products would pass in the opposite direction.

Reject

COMMENTS AND SUGGESTIONS

G. H. Beyer

Several comments are in order on a program for testing the effectiveness of cryosorption panels in removing noncondensable gases such as hydrogen.

Data gathered by the Linde Division[1,2] of Union Carbide are shown in Figure B-47. The capacity per gram of adsorbent is shown to be greater for charcoal than for molecular sieve 5A. At lower pressures, particularly, charcoal can hold appreciably larger quantities of hydrogen than can molecular sieve 5A. For example, at a pressure of 10^{-9} torr it is reported that charcoal can hold 100 cc (STP) of hydrogen, as contrasted to 2.8 cc (STP) for molecular sieve 5A. It is also apparent that the capacity of charcoal decreases much more slowly at lower pressures than does the capacity of molecular sieve 5A.

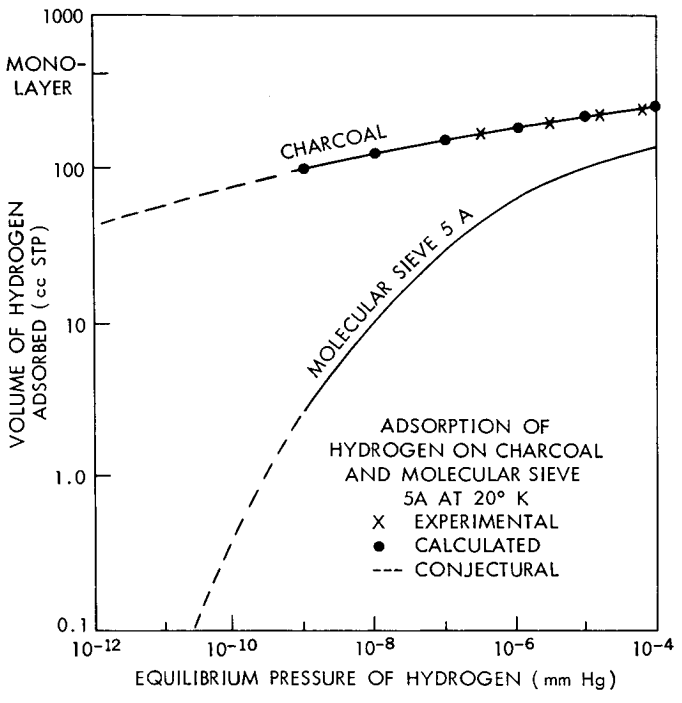


Figure B-47–Adsorption of Hydrogen at 20° K, Charcoal and Sieve 5A

A preliminary analysis of the charcoal data along the lines suggested by Hobson[3] resulted in the equation:

$$\log V = 2.60 - 1.03 \times 10^{-5} T^2 \left[\log \left(P_H/P_{sat}\right)\right]^2$$

where V is the volume of hydrogen (STP) adsorbed per gram of charcoal, P_{sat} is the vapor pressure of hydrogen at absolute temperature T in °K, and P_H is the equilibrium hydrogen pressure. This equation was correctly stated – but incorrectly plotted – in the report of the 1963 Summer Workshop. Points calculated using this equation are shown as circles, while experimental points are shown as crosses. The agreement is not surprising, since the charcoal data were used to determine the constants. Additional data would be desirable to establish confidence in extrapolation to lower pressures.

Such extrapolation of adsorption equilibrium data to lower pressures is of considerable practical interest. Cryosorption-panel capacity at very low pressures hopefully may be significant, even though numerically small. But careful conservation of potential pumping speed at low pressures will be necessary to avoid premature use of the forces at the surface of the adsorbent. Operational procedures will have to be rigorously controlled and considerable experimental finesse developed. The need for substantial and sustained experimental effort is evident, since experimentation at low temperatures and pressures is far from routine.

Pointing up the need for special precautions is the reported susceptibility of molecular sieve 5A to preadsorbed water and air. A minimum activation of 100 °C for at least 12 hours, and no more than 30 cc (STP) of preadsorbed air per gram, has been suggested.[2]

Thus, careful activation of the adsorbent and shielding from nitrogen contamination will be essential. Condensables frozen on the surface of the adsorbent are detrimental to hydrogen pumping speed.

It should also be remembered that the vapor pressure of nitrogen at 20°K limits the pressure attainable to 10^{-11} torr. Evaporation of nitrogen cryodeposits from chevron baffling will prevent attainment of lower pressures unless lower temperatures are used. Because air is a likely contaminant in most vacuum systems, it may well become necessary to enlist the help of adsorbent surface forces to immobilize nitrogen as well as hydrogen if lower pressures are desired.

Considering the cryosorption evaluation program as a whole, experimental emphasis could take one of two major directions: Emphasis on achieving ultrahigh vacuum (10^{-13} torr, for example) seems perhaps premature, when the techniques and eccentricities of cryosorption-panel operation have yet to be mastered. It may be better first to emphasize understanding cryosorption pumping under less stringent conditions (say, 10^{-8} torr), allowing a gradual buildup of operating experience at pressures where measurements and equipment limitations are less of a problem. The higher vacuums will probably be achieved only after substantial experience at more modest pressures.

Cryosorption-panel design is still so new that many interesting alternatives remain to be tried. Bonding techniques leave much to be desired until they can be shown to survive repeated cycling between bakeout and 20°K temperatures. Cracking, spalling, or dusting may become evident only after extended use. The occurrence of such contamination would focus attention on alternate bonding techniques and other methods of displaying extended surface where adsorption can take place. All-metal panels would offer distinct advantages in heat transfer and cleanliness. Honeycomb materials such as Hexcel[4] are now available in various shapes, made of various metals, which can be used to extend manyfold the surface of flat plates. Etching such honeycomb can markedly increase the surface presented to a space chamber, with good heat transfer ensuring temperatures close to those of the base plate. Then, too, combinations of honeycomb thermally bonded to a flat plate and surface-treated with platinum-black, Glidden-black, or molecular sieve may circumvent some of the problems inherent in current panel designs.

Increased concern about contamination during vacuum testing has focused attention on the backstreaming of oil from diffusion pumps. Most space chambers currently operating are equipped with such pumps. Opportunities for oil contamination are known to be greatest during startup and shutdown,[5] as indicated in Figure B-48. During pump warmup, the jet velocity is too low for controlled streaming to the condensing surface of the pump; during pump shutdown, a similar interruption of jet action occurs again, with random gas flow of oil molecules markedly increasing the chances for contamination of the system. With proper baffling between the pump and the chamber, such contamination can ordinarily be held to low levels by careful attention to approved operating procedures.

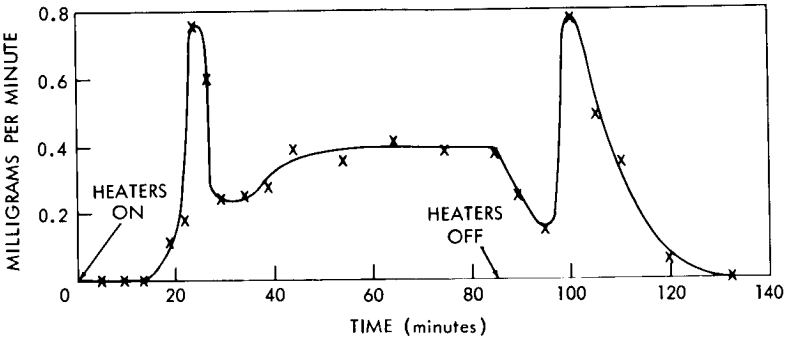


Figure B-48–Relative Rate of Backstreaming from Typical Diffusion Pump

Measurements of backstreaming have demonstrated that a cold cap directly above and surrounding the top jet can materially reduce oil migration.[6] It would be highly desirable to monitor any such oil migration so that corrective action can be taken before appreciable quantities of oil reach spacecraft under test. The quartz-crystal microbalance shows promise of making such measurements possible on a routine basis.

The basic principle of the quartz-crystal microbalance[7] is that the vibration frequency of the crystal is changed by the deposition of material on its surface. The frequency change is proportional to the mass deposited. In its most refined form, such an instrument employs two crystals, one covered and the other exposed. By arranging that both crystals experience identical temperature variations, temperature can be virtually eliminated as a variable. The difference (or "beat") frequency is then taken as a measure of the rate of oil migration. Dr. R. N. Peacock of the University of Illinois has suggested the possible application of this instrument to vacuum testing when oil contamination is of serious consequence. The electrical readout of such an instrument offers a distinct advantage over the use of mirrors or wettable surfaces, although it is unlikely that the microbalance will be able to detect as minute quantities of oil as optical methods can.

As indicated in Dr. Propst's second paper in this report, limited experimental work at the University of Illinois has also shown that placing Fiberfrax ceramic-fiber paper in the throat of diffusion pump may be helpful in controlling oil backstreaming. This inert alumina-silica paper is available in 20-, 40-, and 80-mil thicknesses. It is permeable to gases but quite opaque to oil molecules, at some sacrifice of pumping speed. The development of easily-replaceable cartridges would probably be required, since it is not practical to disconnect large diffusion pumps frequently.

There is a possibility of using cryosorption pumping to calibrate vacuum gauges in place, as suggested by Hobson.[8] The method uses a reference gauge connected to a reservoir volume at, say, 10^{-5} torr. (In practice, the reference gauge could be a McLeod gauge, but most likely it would be a conventional ion gauge periodically standardized by comparison with a McLeod gauge.) Gas in the reservoir volume is leaked into the system through a calibrated conductance. Before admitting gas through this conductance, the system pressure is reduced to very low values, using a properly-activated absorbent immersed in liquid helium. Cryosorption is proposed for this application because of its fixed pumping speed at pressures well below those at which gauges are now calibrated.

REFERENCES

1. Stern, S.A., et al, "The Cryosorption Pumping of Hydrogen and Helium at 20°K Interim Report," Arnold Engineering Development Center Report AEDC-TDR-62-200 (October 1962)

2. Hemstreet, R.A., et al, "The Cryosorption Pumping of Hydrogen at 20°K," Arnold Engineering Development Center Report AEDC-TDR-64-100 (May 1964)

3. Hobson, J.P., "Physical Adsorption of Nitrogen on Pyrex at Very Low Pressures," J. Chem. Phys., 34, 1850 (1961)

4. Hexcel Honeycomb, Hexcel Products, Inc., P.O. Box 700, Havre de Grace, Maryland

5. Giles, G., "Diffusion Pumped Systems," Test Engineering, p. 30 (August 1964)

6. Burnett, S.G., and Hablanian, M.H., "Diffusion Pumps," Journal of Environmental Sciences, p. 7 (May 1964)

7. Lins, S.J., and Kukuk, H.S., "Resonance Frequency Shift Thin-Film Thickness Monitor," AVS Trans (1960) 7th National Vacuum Symposium, 333-338, Pergamon Press, New York

8. Hobson, J.P., "Physical Adsorption in Ultra-High-Vacuum Systems," AVS Trans. (1961) Vol. I, 8th National Vacuum Symposium, 146-150, Pergamon Press, New York

SUMMER WORKSHOP 1964

Program Outline and Team Participants

PROJECT C: Magnetic and Electric Fields in Environment Simulation

Study Topics

C-1 Investigate large magnetic environments for testing meteoroid spacecraft

C-2 Study the measurement of small magnetic torques by means of electrostatics

C-3 Explore attitude-control test facility

TEAM C

Academic Personnel	Goddard Personnel	Code
Dr. Donald L. Waidelich, principal investigator	Robert Gebhardt, staff advisor	325
Anthony V. Dralle	C. Leland Parsons	325
Stephen H. Saperstone	Joseph Boyle	325
Robert L. Piziali	William D. Kenney	325

NSF Summer Science Student
Francine E. Wright

Page intentionally left blank

PROJECT C: MAGNETIC AND ELECTRIC FIELDS IN ENVIRONMENT SIMULATION

CONTENTS

ILLUSTRATIONS

TABLES

MAGNETIC AND ELECTRIC FIELDS

D. L. Waidelich

INTRODUCTION

To simulate magnetic fields in spacecraft testing, the Attitude Control Test Facility at Goddard Space Flight Center will produce a volume of uniform magnetic field large enough to test an entire satellite. The coils in the facility, approximately 40 feet in diameter, will be housed in a non-magnetic building. Since the shape of the coils may be either circular or octagonal, both types of coils have been studied.

The parameters of various circular coil systems have been calculated for rational turns-ratios, and contours of the axial flux density have been plotted. Similar calculations were made for the four-octagonal-coil systems. Details of the analysis for the octagonal coils are included in this report because this material had not been published previously.

The need for a very large volume of uniform field has suggested the use of approximation theory in obtaining the parameters of the coils systems. Up till now, practically every analysis has used the idea of making zero as many coefficients as possible of the power series for the magnetic field along the axis of the system. For a two-coil system, it is possible to show that a larger volume of uniformity is obtained when the coefficients are not zero. Similar conclusions may be drawn for other coil systems, and a study was made of the application of approximation theory to this problem. Some success was achieved for a few of the simpler coil systems, but an attempt to apply least-square theory to the problem did not turn out very well. It is believed that the method applied to these simpler coil systems can be modified and extended to the more complicated systems.

For the Pegasus and OAO satellites, a study was made to find coil configurations that would enable the satellites to be tested in nearly zero magnetic field. Several types of coil systems were tried, including two- and four-dipole, two square coils, two rectangular coils, dipole arrays, and one coil with two dipoles. The latter seems the best, and an example of its use with the OAO satellite is given. To adapt the one-coil and two-dipole system to the Pegasus satellite will require some additional analysis.

In the Attitude Control Test Facility it is proposed to float a satellite on water and measure the very small torque exerted on the craft by the magnetic field. An opposing torque must be supplied to balance the torque from the magnetic field so that the whole system is in stable equilibrium. The amount of torque can then be determined by a measured voltage (in the case of the piezoelectric method) or by the measured angular deflection (in the case of the other methods).

In working with electric fields, it would be convenient to have an instrument to measure the electric field intensity at any point in space. This field intensity should be measured in magnitude and direction or as three components. An investigation was made to determine what measuring instruments had been and are being used, and to suggest possibilities for a more compact instrument to determine the three components of the field intensity.

In calibrating electric field meters, a facility to provide a uniform electric field would be useful. One such facility consisting of four charged rings is suggested and a brief analysis was made to determine the possible parameters. Unfortunately, not enough time was available to determine the percentage of uniformity of the possible systems.

A question was raised as to the relationship of electric to magnetic fields, and the effect of one type of field on an instrument designed to measure the other type of field. This relationship depends upon the velocity of the instrument with respect to the field.

FOUR-CIRCULAR-COIL SYSTEMS

S. H. Saperstone

In order that the current in both coil pairs of the four-circular-coil system be the same, it is necessary that the turns-ratio N_2/N_1 (where N_1 and N_2 are the number of turns on coil 1 and 2 respectively) be a rational valued quantity. A computer program was written to determine the parameters of the four-coil system for specified rational values of N_2/N_1. The analysis of this system was done by Waidelich and Pittman.[1] By specifying N_2/N_1, the coil spacing and radius of each of the coils are automatically determined. Table C-1 contains these results. Notice that, as the turns-ratio decreases, the larger #1 coils are moving away from the center of the system, the smaller #2 coils are moving toward the center of the system, and the radius of the #2 coils is decreasing.

Another computer program was written to calculate the flux-density components of the coil system, based upon the parameters in Table C-1. A thorough analysis of this problem is given in Waidelich and Speiser.[2] Figure C-1 shows contours of the axial component of the flux density plotted on an X-Y plotter for the system with the one-to-one turns-ratio. The abscissa Z represents the axial direction and the ordinate V the radial direction. The distances along the axes are all in terms of the unity radius of the larger #1 coils. The value of the flux density at the origin was taken as unity, and the numbers on the curves are the values of the flux density on the curves referred to that at the origin. For example, the 1.001 curve means that the axial flux density at the curve is greater than that at the origin by a ratio of (1.001/1.000) or by 0.1 of one percent. Figure C-2 shows a similar result for the one-to-two turns-ratio. The curves of Figure C-2 indicate a slightly smaller volume of uniformity than those for Figure C-1; for example, the 0.99 contour at the Z axis in Figure C-2 has moved toward the origin, compared to the same curve in Figure C-1. In both figures, the curves very near to the origin become distorted. This distortion was caused by rounding-off error in the computer.

REFERENCES

1. Pittman, M.E., and Waidelich, D.L., "Three- and Four-Coil Systems for Homogeneous Magnetic Fields," NASA Technical Note D-2095 (January 1964)

2. Waidelich, D.L., and Speiser, M., "Four-Circular-Coil System," in "Final Report of the Goddard Summer Workshop Program," Goddard Space Flight Center Report X-320-63-264, pp. C-77 to C-90 (1963)

Table C-1
Parameters for Four Circular Coils

I	16/16	15/16	14/16	13/16	12/16	11/16	10/16	9/16	8/16
d_1	.27802801	.28123548	.28471355	.28849859	.29263615	.29716967	.30220061	.30777691	.31402395
d_2	.84566385	.83774669	.82923668	.82004694	.81007810	.79917237	.78724994	.77400986	.75920762
R	.76389938	.74748629	.73033752	.71237642	.69351438	.67364267	.65262085	.63028922	.60642682
$\cos \alpha_1$	.26786775	.27073265	.27383117	.27719353	.28085741	.28485788	.28927989	.29415973	.29959932
$\cos \alpha_2$	.74207038	.74615969	.75044003	.75492812	.75964490	.76460218	.76986395	.77542406	.78134013
b_1	1.0379304	1.0387941	1.0397412	1.0407840	1.0419385	1.0432208	1.0446651	1.0462918	1.0481464
b_2	1.1396006	1.1227445	1.1050006	1.0862583	1.0663905	1.0452133	1.0225832	.99817622	.97167366

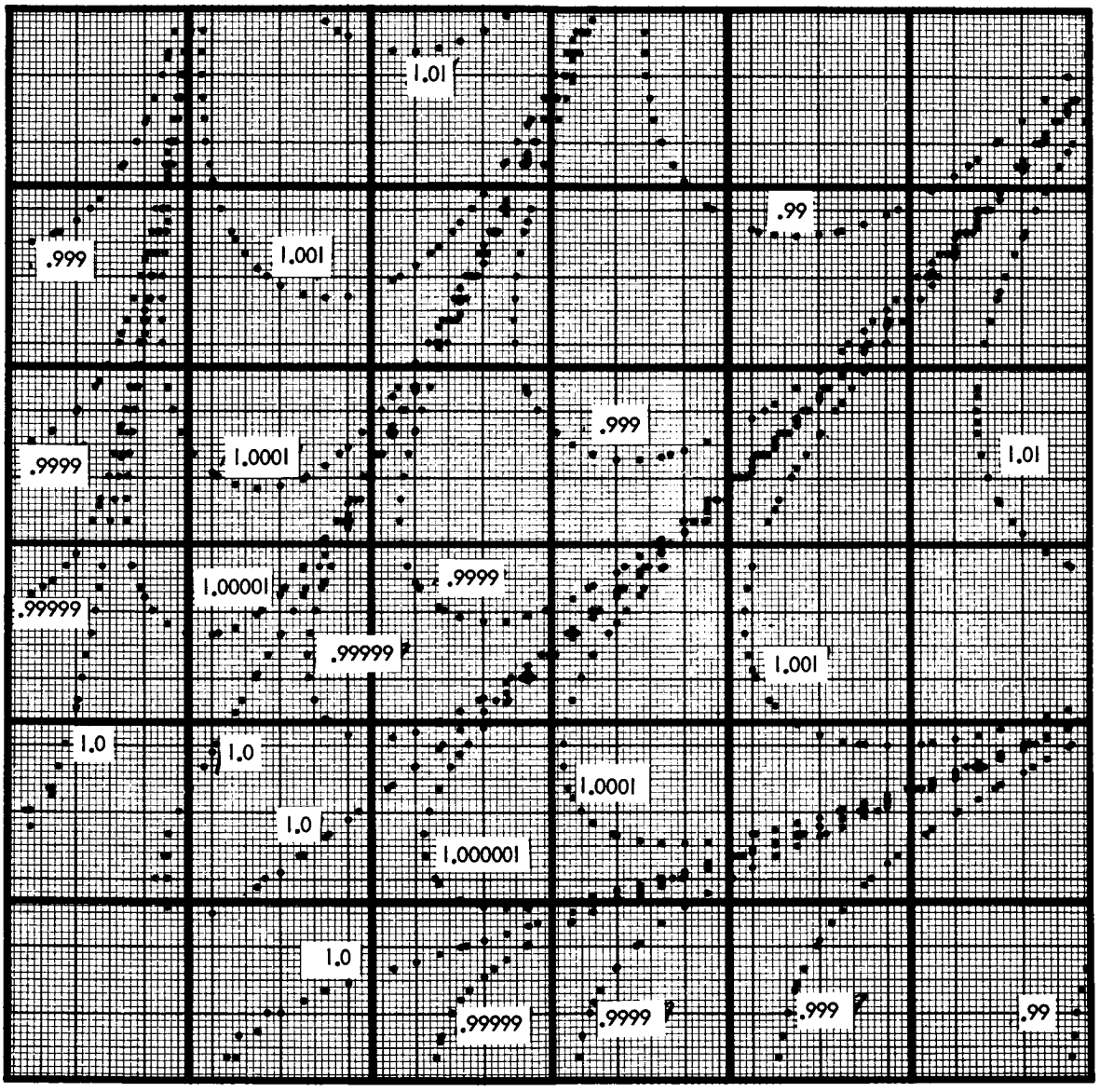


Figure C-1–Flux Density Contours for the One-to-One Turns-Ratio

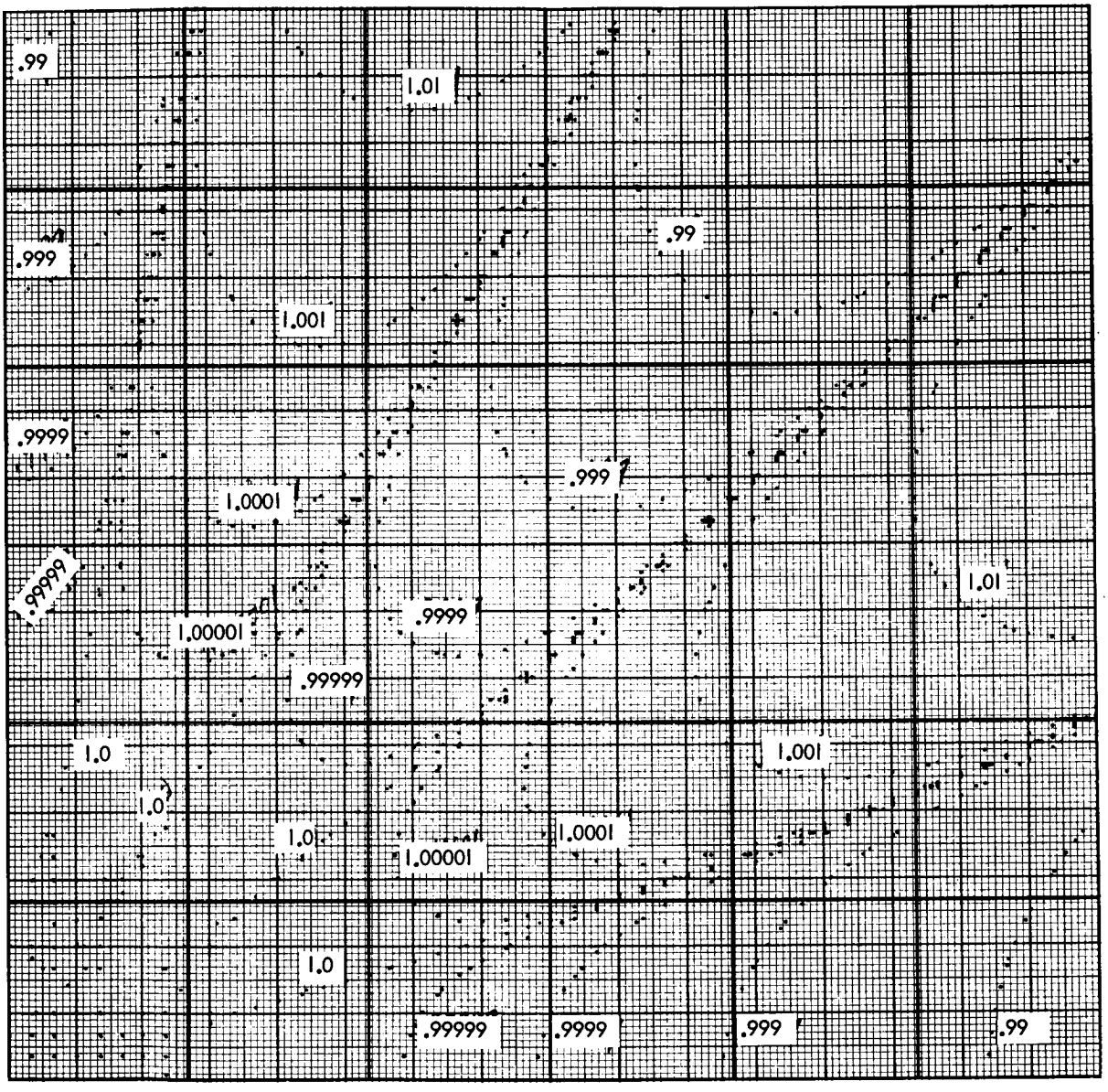


Figure C-2–Flux Density Contours for the One-to-Two Turns-Ratio

FOUR-OCTAGONAL-COIL SYSTEMS

S. H. Saperstone

INTRODUCTION

In the production of uniform magnetic fields, the regular octagonal-coil system is a compromise between the square-coil and circular-coil systems. Circular coils[1] and square coils[2,3] have been used to obtain uniform magnetic fields. Analysis of octagonal coils to determine the parameters of a four-coil system has already been carried out.[4]

The problem is to calculate the parameters of the coil system for specified rational valued turns-ratios, and to use these values for analysis of the field gradients.

ANALYSIS

The magnetic-field intensity along the axis for a single octagonal coil is given by

$$H(z) = \frac{NI}{RK} f(z) \tag{1}$$

where

$$f(z) = \frac{1}{(1 + z^2)\sqrt{B^2 + z^2}} \tag{2}$$

and NI is the ampere-turns of the coil, R is the "radius" (Figure C-3), and

$$B = \sqrt{1 + \tan^2 \frac{\pi}{8}} = 1.0823922.$$

For a four-coil system, shown in Figure C-4, the field intensity along the axis is

$$H(z) = \frac{N_1 I_1}{R_1 K}\left[f(d_1 - z) + f(d_1 + z)\right] + \frac{N_2 I_2}{R_1 K}\left[f(d_2 - z) + f(d_2 + z)\right] \tag{3}$$

Letting $x_1 = d_1/R_1$ and $x_2 = d_2/R_2$, we can expand $H(z)$ in a Taylor series about $z = 0$; hence,

$$H(z) = \sum_{n=0}^{\infty} a_{2n} z^{2n}, \quad a_n = \frac{1}{n!} \left.\frac{d^n H}{dz^n}\right|_{z=0} \tag{4}$$

For suitable values of x_1 and x_2, the terms a_2, a_4, and a_6 may be made zero.

Set $I = N_2 I_2 / N_1 I_1$ and $R = R_2/R_1$. If $R_1 = 1$, then Figure C-5 relates x_1, I, and R_2 vs. x_2 for the four-coil system. These results were obtained by use of a digital computer.

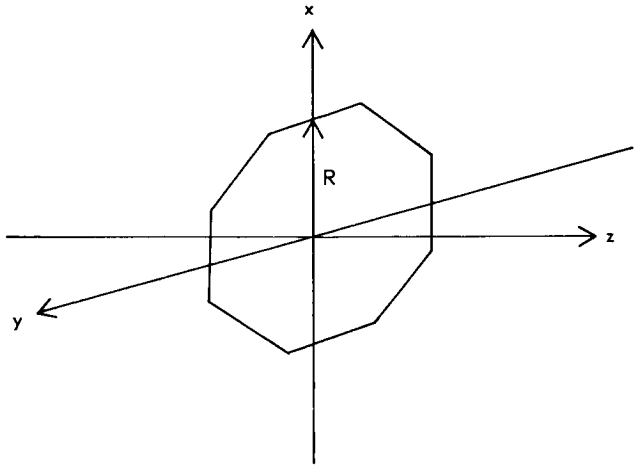


Figure C-3–Single Octagonal Coil

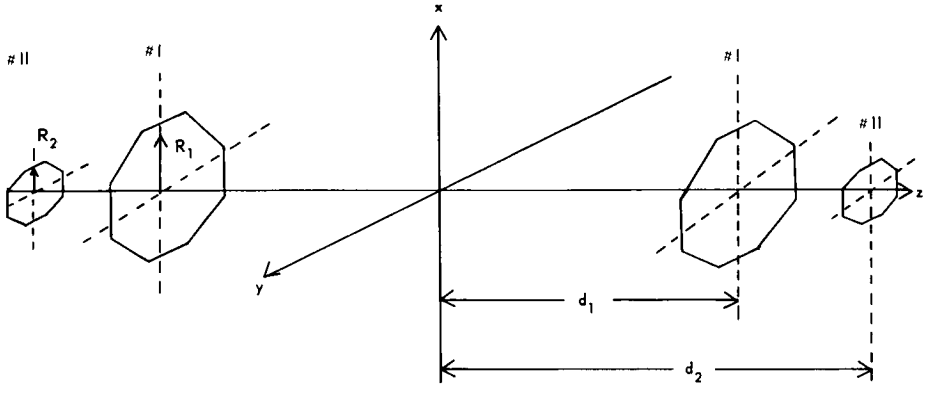


Figure C-4–Four Octagonal Coils

Equation (4) can be written as

$$H(z) = a_0 \left[1 + a_8/a_0 \, z^8 + a_{10}/a_0 \, z^{10} + \cdots \right] \tag{5}$$

Then the homogeneity of the system (Figure C-5), in the case when $R_1 = 1.0$ is given by

$$\text{HOMO} = \left[10^{-5} \, a_0/a_8 \right]^{1/8} \tag{6}$$

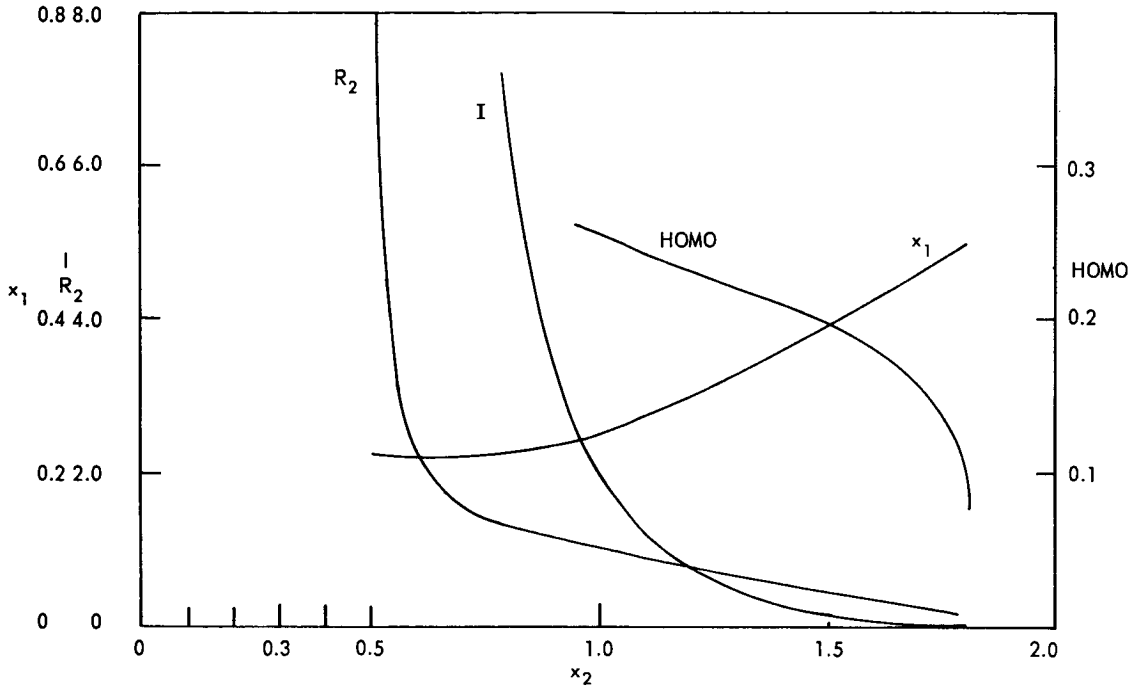


Figure C-5–Curves of Parameters of Four-Octagonal-Coil Systems

RESULTS

In order that $I_1 = I_2$, it is necessary for I to be a rational number. A computer program was written to calculate the parameters of the four-coil system in order to realize rational values of I. Table C-2 contains the results for some specific values of I.

Table C-2
Results for Octagonal Coil Systems
with Rational Turns-Ratios

$I = \frac{N_2 I_2}{N_1 I_1}$	16/16	15/16	14/16	13/16	12/16	11/16	10/16	9/16	8/16
d_1	.28435432	.28764020	.29120383	.29508335	.29932539	.30398098	.30913377	.31485706	.32126987
d_2	.86560663	.85751313	.84881337	.83942090	.82923067	.81809756	.80589842	.79237031	.77724546
R	.76317254	.74677648	.72964492	.71170181	.69285597	.67299845	.65200038	.62968647	.60584340

To learn how the octagonal coil shape affected the magnetic field in the circumferential direction, the investigators decided to obtain plots of the field components for points both on and off the axis of the system. A computer program was written to trace contour lines of constant field intensity, for the turns-ratios given in Table C-2. Details of the analysis of this problem are given in the appendix.

REFERENCES

1. Pittman, M.E., and Waidelich, D.L., "Three- and Four-Coil Systems for Homogeneous Magnetic Fields," NASA Tech. Note D-2095 (January 1964)

2. Yildirim, G., "Production of Uniform Magnetic Fields by Using Square Coils," Research Report, University of Missouri Engineering Library (June 1963)

3. Speiser, M., and Waidelich, D.L., "Four-Square Coil Systems," NASA GSFC Pub. X-320-63-264 (September 18, 1963) pp. C-91 to C-111

4. Chiang, B.A., and Waidelich, D.L., "Four-Octagonal-Coil Systems for Uniform Magnetic Fields," Report, University of Missouri (May 1964)

APPENDIX – Field Intensity Calculation

The magnetic field due to segments (1) and (5) in Figure C-6 are calculated. The field contributed by the other 6 components are determined in a similar manner. We assume that R represents the radius of the inscribed circle, and hence each segment has length $2L = 2(\sqrt{2} - 1)R$.

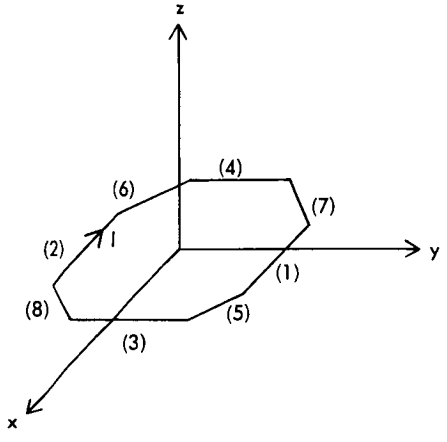


Figure C-6–Geometry of One Coil

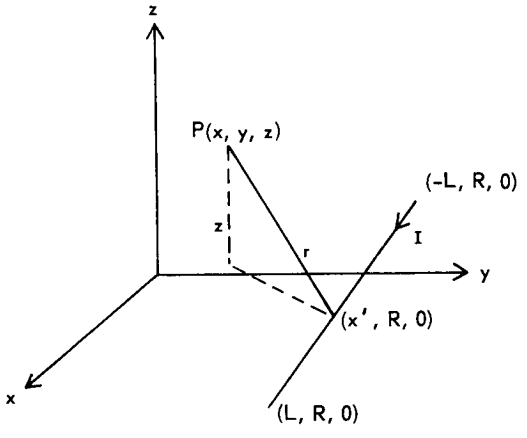


Figure C-7–Coordinates of Side 1

Then

$$H_y^{(1)} = \frac{I}{4\pi} \int \frac{(d\vec{\ell} \times \vec{r})_y}{r^3}$$

$$= -\frac{Iz}{4\pi} \int_{-L}^{L} \frac{dx'}{\left[(x - x')^2 + (y - R)^2 + z^2\right]^{3/2}}$$

Letting $u = x - x'$, $k^2 = (y - R)^2 + z^2$,

$$H_y^{(1)} = \frac{Iz}{4\pi} \int_{x+L}^{x-L} \frac{du}{[u^2 + k^2]^{3/2}}$$

$$= \frac{Iz}{4\pi} \frac{1}{(y - R)^2 + z^2} \left\{ \frac{x - L}{[(x - L)^2 + (y - R)^2 + z^2]^{1/2}} - \frac{x + L}{[(x + L)^2 + (y - R)^2 + z^2]^{1/2}} \right\}$$

Similarly,

$$H_z^{(1)} = \frac{I}{4\pi} \int \frac{(d\vec{\ell} \times \vec{r})_3}{r^3}$$

$$= \frac{I(y - R)}{4\pi} \int_{-L}^{L} \frac{dx'}{[(x - x')^2 + (y - R)^2 + z^2]^{3/2}}$$

$$H_z^{(1)} = \frac{I}{4\pi} \frac{y - R}{(y - R)^2 + z^2} \left\{ \frac{x + L}{[(x + L)^2 + (y - R)^2 + z^2]^{1/2}} - \frac{x - L}{[(x - L)^2 + (y - R)^2 + z^2]^{1/2}} \right\}$$

$$H_x^{(1)} = 0.$$

Now consider segment (5) as shown in Figure C-8:

Then

$$H_y^{(5)} = \frac{I}{4\pi} \int \frac{(d\vec{\ell} \times \vec{r})_y}{r^3}$$

$$= -\frac{Iz}{4\pi} \int_{L}^{R} \frac{dx'}{[(x - x')^2 + (y - \sqrt{2}R + x')^2 + z^2]^{3/2}}$$

Letting $u = x - x'$, $v = y - \sqrt{2}R + x'$, and $t = x + y - \sqrt{2}R$,

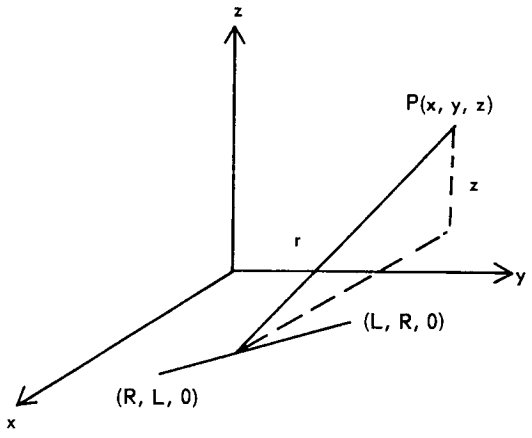


Figure C-8-Coordinates of Side 5

$$H_y^{(5)} = \frac{Iz}{4\pi} \int_{X-L}^{X-R} \frac{du}{[u^2 + (t - u)^2 + z^2]^{3/2}}$$

$$H_y^{(5)} = \frac{Iz}{8\sqrt{2}\,\pi} \int_{x-L}^{x-R} \frac{du}{\left[\left(u - \frac{t}{2}\right)^2 + p^2\right]^{3/2}} \quad \text{where} \quad p^2 = \frac{t^2}{4} + \frac{z^2}{2}$$

$$H_y^{(5)} = \frac{Iz}{2\sqrt{2}\,\pi} \frac{1}{t^2 + 2z^2} \left\{ \frac{x - R - \frac{1}{2}t}{\left[\left(x - R - \frac{t}{2}\right)^2 + \left(\frac{t}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{x - L - \frac{1}{2}t}{\left[\left(x - L - \frac{t}{2}\right)^2 + \left(\frac{t}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

Similarly,

$$H_z^{(5)} = \frac{I}{4\pi} \int \frac{(d\vec{\ell} \times \vec{r})_z}{r^3}$$

$$= \frac{I}{4\pi} \int_L^R \frac{(x + y - \sqrt{2}R)\, dx'}{\left[(x - x')^2 + (y - \sqrt{2}R + x')^2 + z^2\right]^{3/2}}$$

or $H_z^{(5)} = -\frac{t}{z} H_y^{(5)}$, where $t = x + y - \sqrt{2}R$

Also,

$$H_x^{(5)} = \frac{I}{4\pi} \int \frac{(d\vec{\ell} \times \vec{r})_x}{r^3}$$

$$= \frac{Iz}{4\pi} \int_L^R \frac{dy'}{\left[(x - \sqrt{2}R + y')^2 + (y - y')^2 + z^2\right]^{3/2}}$$

so,

$$H_x^{(5)} = \frac{Iz}{2\sqrt{2}\,\pi} \frac{1}{t^2 + 2z^2} \left\{ \frac{y - R - \frac{1}{2}t}{\left[\left(y - R - \frac{t}{2}\right)^2 + \left(\frac{t}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{y - L - \frac{1}{2}t}{\left[\left(y - L - \frac{1}{2}t\right)^2 + \left(\frac{t}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

If each of the coordinates is normalized with respect to the "radius" R of the coil, then the following set of equations represents the field components for each segment of the coil.

$$H_x^{(1)} = 0$$

$$H_x^{(2)} = 0$$

$$H_x^{(3)} = \frac{I}{4\pi} \cdot \frac{z}{(x-1)^2 + z^2} \left\{ \frac{y - \sqrt{2} + 1}{[(x-1)^2 + (y-t)^2 + z^2]^{1/2}} - \frac{y + \sqrt{2} - 1}{[(x-1)^2 + (y+t)^2 + z^2]^{1/2}} \right\}$$

$$H_x^{(4)} = \frac{I}{4\pi} \cdot \frac{z}{(x+1)^2 + z^2} \left\{ \frac{y + \sqrt{2} - 1}{[(x+1)^2 + (y+t)^2 + z^2]^{1/2}} - \frac{y - \sqrt{2} + 1}{[(x+1)^2 + (y-t)^2 + z^2]^{1/2}} \right\}$$

$$H_x^{(5)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(x + y - \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(y - x + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(y - x + \sqrt{2}) - 1\right)^2 + \left(\frac{x + y - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(y - x - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(y - x - \sqrt{2}) + 1\right)^2 + \left(\frac{x + y - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_x^{(6)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(x + y + \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(y - x - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(y - x - \sqrt{2}) + 1\right)^2 + \left(\frac{x + y + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(y - x + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(y - x + \sqrt{2}) - 1\right)^2 + \left(\frac{x + y + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_x^{(7)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(y - x - \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(y + x - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(y + x - \sqrt{2}) + 1\right)^2 + \left(\frac{y - x - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(y + x + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(y + x + \sqrt{2}) - 1\right)^2 + \left(\frac{y - x - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_x^{(8)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(y - x + \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(y + x + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(y + x + \sqrt{2}) - 1\right)^2 + \left(\frac{y - x + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(y + x - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(y + x - \sqrt{2}) + 1\right)^2 + \left(\frac{y - x + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_y^{(1)} = \frac{I}{4\pi} \cdot \frac{z}{(y - 1)^2 + z^2} \left\{ \frac{x - \sqrt{2} + 1}{[(x - t)^2 + (y - 1)^2 + z^2]^{1/2}} - \frac{x + \sqrt{2} - 1}{[(x + t)^2 + (y - 1)^2 + z^2]^{1/2}} \right\}$$

$$H_y^{(2)} = \frac{I}{4\pi} \cdot \frac{z}{(y + 1)^2 + z^2} \left\{ \frac{x + \sqrt{2} - 1}{[(x + t)^2 + (y + 1)^2 + z^2]^{1/2}} - \frac{x - \sqrt{2} + 1}{[(x - t)^2 + (y + 1)^2 + z^2]^{1/2}} \right\}$$

$$H_y^{(3)} = 0, \quad H_y^{(4)} = 0$$

$$H_y^{(5)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(x + y - \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(x - y + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(x - y + \sqrt{2}) - 1\right)^2 + \left(\frac{x + y - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(x - y - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(x - y - \sqrt{2}) + 1\right)^2 + \left(\frac{x + y - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_y^{(6)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(x + y + \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(x - y - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(x - y - \sqrt{2}) + 1\right)^2 + \left(\frac{x + y + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(x - y + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(x - y + \sqrt{2}) - 1\right)^2 + \left(\frac{x + y + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_y^{(7)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(x - y + \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(x + y + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(x + y + \sqrt{2}) - 1\right)^2 + \left(\frac{x - y + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(x + y - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(x + y - \sqrt{2}) + 1\right)^2 + \left(\frac{x - y + \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_y^{(8)} = \frac{I}{2\sqrt{2}\,\pi} \cdot \frac{z}{(x - y - \sqrt{2})^2 + 2z^2} \left\{ \frac{\frac{1}{2}(x + y - \sqrt{2}) + 1}{\left[\left(\frac{1}{2}(x + y - \sqrt{2}) + 1\right)^2 + \left(\frac{x - y - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right.$$

$$\left. - \frac{\frac{1}{2}(x + y + \sqrt{2}) - 1}{\left[\left(\frac{1}{2}(x + y + \sqrt{2}) - 1\right)^2 + \left(\frac{x - y - \sqrt{2}}{2}\right)^2 + \frac{z^2}{2}\right]^{1/2}} \right\}$$

$$H_z^{(1)} = \frac{I}{4\pi} \frac{y - 1}{(y - 1)^2 + z^2} \left\{ \frac{x + \sqrt{2} - 1}{[(x + \sqrt{2} - 1)^2 + (y - 1)^2 + z^2]^{1/2}} - \frac{x - \sqrt{2} + 1}{[(x - \sqrt{2} + 1)^2 + (y - 1)^2 + z^2]^{1/2}} \right\}$$

$$H_z^{(2)} = \frac{I}{4\pi} \frac{y + 1}{(y + 1)^2 + z^2} \left\{ \frac{x - \sqrt{2} + 1}{[(x - \sqrt{2} + 1)^2 + (y + 1)^2 + z^2]^{1/2}} - \frac{x + \sqrt{2} - 1}{[(x + \sqrt{2} - 1)^2 + (y + 1)^2 + z^2]^{1/2}} \right\}$$

$$H_z^{(3)} = \frac{I}{4\pi} \frac{x - 1}{(x - 1)^2 + z^2} \left\{ \frac{y + \sqrt{2} - 1}{[(x - 1)^2 + (y + \sqrt{2} - 1)^2 + z^2]^{1/2}} - \frac{y - \sqrt{2} + 1}{[(x - 1)^2 + (y - \sqrt{2} + 1)^2 + z^2]^{1/2}} \right\}$$

$$H_z^{(4)} = \frac{I}{4\pi} \frac{x + 1}{(x + 1)^2 + z^2} \left\{ \frac{y - \sqrt{2} + 1}{[(x + 1)^2 + (y - \sqrt{2} + 1)^2 + z^2]^{1/2}} - \frac{y + \sqrt{2} - 1}{[(x + 1)^2 + (y + \sqrt{2} - 1)^2 + z^2]^{1/2}} \right\}$$

$$H_z^{(5)} = -\left(\frac{x + y - \sqrt{2}}{z}\right) H_y^{(5)}$$

$$H_z^{(6)} = -\left(\frac{x + y + \sqrt{2}}{z}\right) H_y^{(6)}$$

$$H_z^{(7)} = -\left(\frac{y - x - \sqrt{2}}{z}\right) H_y^{(7)}$$

$$H_z^{(8)} = -\left(\frac{y - x + \sqrt{2}}{z}\right) H_y^{(8)}$$

For a four-coil system as illustrated in Figure C-4, the total magnetic field components are given below, where the coordinates of a point $P(x, y, z)$ are normalized by R_1, the radius of the inner coil, and where $z_1 = d_1/R_1$, $z_2 = d_2/R_2$.

$$H_x(x, y, z) = \sum_{i=1}^{8} H_x^{(i)}(x, y, z - z_1) + \sum_{i=1}^{8} H_x^{(i)}(x, y, z + z_1)$$

$$+ \sum_{i=1}^{8} H_x^{(i)}\left(\frac{x}{R}, \frac{y}{R}, \frac{z}{R} - z_2\right) + \sum_{i=1}^{8} H_x^{(i)}\left(\frac{x}{R}, \frac{y}{R}, \frac{z}{R} + z_2\right)$$

$$H_y(x, y, z) = \sum_{i=1}^{8} H_y^{(i)}(x, y, z - z_1) + \sum_{i=1}^{8} H_y^{(i)}(x, y, z + z_1)$$

$$+ \sum_{i=1}^{8} H_y^{(i)}\left(\frac{x}{R}, \frac{y}{R}, \frac{z}{R} - z_2\right) + \sum_{i=1}^{8} H_y^{(i)}\left(\frac{x}{R}, \frac{y}{R}, \frac{z}{R} + z_2\right)$$

$$H_z(x, y, z) = \sum_{i=1}^{8} H_z^{(i)}(x, y, z - z_1) + \sum_{i=1}^{8} H_z^{(i)}(x, y, z + z_1)$$

$$+ \sum_{i=1}^{8} H_z^{(i)}\left(\frac{x}{R}, \frac{y}{R}, \frac{z}{R} - z_2\right) + \sum_{i=1}^{8} H_z^{(i)}\left(\frac{x}{R}, \frac{y}{R}, \frac{z}{R} + z_2\right)$$

APPROXIMATION METHODS FOR OBTAINING UNIFORM MAGNETIC FIELDS

S. H. Saperstone

INTRODUCTION

It has become necessary to study methods of enhancing the volume of homogeneity of the magnetic field produced by circular coils. Present techniques demand very large coil systems for a corresponding increase in the volume of uniformity.

This paper attempts to show that some methods of approximation theory applied to the uniformity problem can yield results significantly better than those obtained with established techniques presented by Waidelich and Pittman.[1]

ANALYSIS (Magnetic Field for a Pair of Circular Coils)

At first, in order to simplify the problem and to acquire experience, approximation methods were applied to a pair of circular coils.

The magnetic field along the axis of the two coil system is given by

$$H(z) = \frac{NI}{2R}\left\{\frac{1}{[1+(z-s)^2]^{3/2}} + \frac{1}{[1+(z+s)^2]^{3/2}}\right\} \tag{1}$$

where R is the radius of each of the coils, NI the ampere – turns of each of the coils, $s = d/R$, where d is the distance from the origin of the system to each of the coils, and z is the point (relative to R) on the axis where the field is computed. Figure C-9 illustrates the geometry of the system.

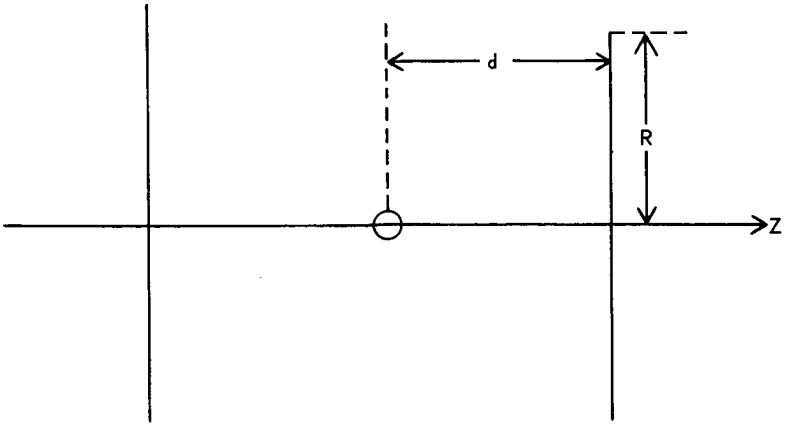


Figure C-9-Geometry of a Two-Coil System

A. Taylor Series Method for Two Circular Coils

If $H(z)$ is expanded in a Taylor series about $z = 0$, we get

$$H(z) = \sum_{n=0}^{\infty} A_{2n} z^{2n} \tag{2}$$

where

$$A_n = \frac{1}{n!} \left. \frac{d^n H}{dz^n} \right|_{z=0}$$

and the odd terms, a_{2n+1} for $n = 0, 1, 2, \ldots$ are zero by symmetry considerations.

We may set $a_2 = 0$ by requiring $s = 0.5$. This yields the familiar Helmholtz configuration. Figure C-10 shows $H(z)/H(0)$ as a function of z for $s = .5$.

If we require that the field vary no more than a prescribed amount, say δ, in a neighborhood of the origin, then we may tentatively define the homogeneity of the configuration as that largest value of z which makes the variation in $H(z)/H(0)$ less than or equal to δ.

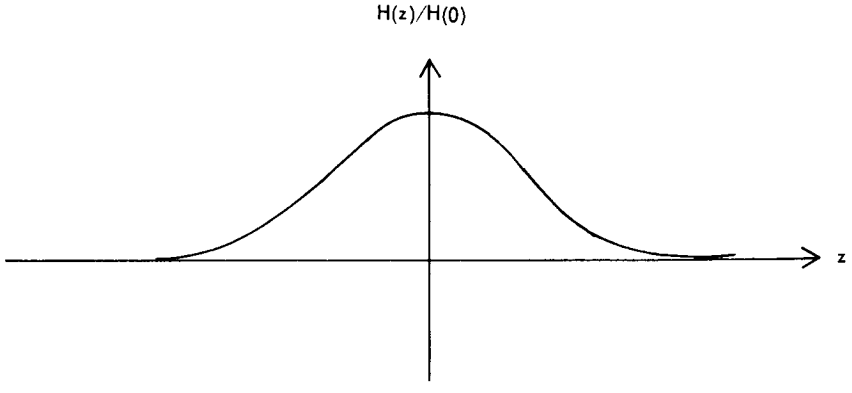


Figure C-10-Axial Field for s = 0.5

B. Optimal Spacing for Two Circular Coils

A paper by Crownfield[2] suggests that the homogeneity may be increased significantly by increasing the coil spacing. Heller[3] observes that, in the case $s = 0.5$, there is a maximum of the field at the origin but a relative minimum for larger spacings. Figure C-11 illustrates this condition.

By setting

$$\frac{d}{dz}\left[\frac{H(z)}{H(0)}\right] = 0 \tag{3}$$

we may determine the point where $H(z)/H(0)$ has a relative maxima. Let such a point be $z = z_0$. Then, for a given field variation, we must find a value of s that will make $H(z_0) - H(0) = \delta$. That this is possible follows from the fact that $H(z)$ is a continuous function of z on the interval $[-1, 1]$. This suggests the following definition of homogeneity (for a fixed value of $s > 0.5$):

$$\text{Homogeneity} = \max \{z \mid H(z) - H(0) = 0\} . \tag{4}$$

i.e., the homogeneity is the largest root of the equation, $H(z) - H(0) = 0$.

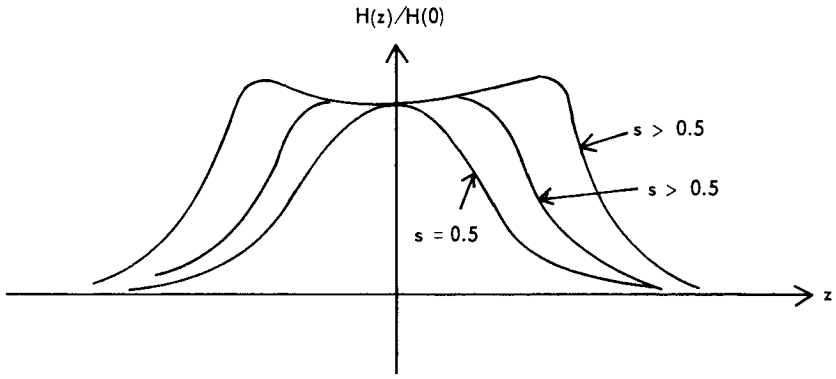


Figure C-11-Axial Field for s > 0.5

A comparison of the homogeneities based upon the two methods just discussed yields the results shown in Table C-3.

Table C-3
Results for Two Circular Coils

δ	s	HOMO	HOMO (for s = 0.5)
0.1%	.518	.247	.085
0.5%	.541	.376	.131
1.0%	.558	.451	.157
2.0%	.583	.546	.188
3.0%	.603	.613	.211
4.0%	.621	.669	.238
5.0%	.636	.714	.245
6.0%	.650	.754	.254

C. Applications of Some Methods of Approximation Theory

1. Least Squares

Obviously, the method used to obtain optimal spacing for two coils cannot be extended to 4, 6, 8 . . . . . coils, but a modification of this method might work. With such a generalization in mind, a technique must be found that can be applied to four or more coils.

The method of least squares was tried, though the results were unsuccessful. Noting that, in equation (1), H is a function of s as well as z, we wish to minimize the following function with respect to s:

$$f(s) = \int_{-a}^{a} [H(z, s) - H(0, s)]^2 \, dz$$

where the interval of integration $[-a, a]$ is suitably chosen in terms of the desired homogeneity. We thought that, by setting $df/ds = 0$, we could obtain the desired spacing of the coils with respect to the fact that $H(z, s)$ differs from $H(0, s)$ (the field at the origin) over the the interval $[-a, a]$ in the sense of least squares.

Unfortunately, the roots of $df/ds = 0$ that were calculated on a digital computer yielded points of relative maxima for the function $f(s)$. It was subsequently determined that $f(s)$ had no relative minima in the useful range for s (i.e., $.5 < s < .6$).

2. Chebyshev Approximation

The theory of synthesis of passive networks[4] employs Chebyshev polynomials to approximate an ideal "flat" response over a given frequency range. An attempt was made to apply these techniques to the two-circular-magnetic-coil problem. Though no conclusive results were obtained, a paper by Darlington[5] indicates how this problem might be attacked.

REFERENCES

1. Pittman, M.E., and Waidelich, D.L., "Three- and Four-Coil Systems for Homogeneous Magnetic Fields," NASA Tech. Note D-2095 (January 1964)

2. Crownfield, F.R., "Optimum Spacing of Coil Pairs," Rev. Sci. Inst., vol. 35, pages 240-241 (February 1964)

3. Heller, Carl, "Über die Erzeugung grossräumiger homogener Magnetfelder zum Studium des Verhaltens von Magnetkompassen und Kompensiermitteln auf verschiedenen magnetischen Breiten," Deutschen Hydrographischen Zeitschrift, Vol. 8, No. 4, pages 157 to 164, (1955)

4. Weinberg, L., "Network Analysis and Synthesis," McGraw-Hill, New York (1962)

5. "Chebyshev Polynomial Series," Bell System Tech. J., Volume 31, (1952) pp. 613-665

LARGE MAGNETIC ENVIRONMENTS

R. Piziali and D. L. Waidelich

INTRODUCTION

To determine the magnetic properties of a large satellite, it is necessary to set up a coil system which can cancel or nearly cancel the earth's magnetic field throughout the volume of the satellite. The earth's field is about 50000 gammas and makes an angle of approximately 70 degrees with the horizontal. One large satellite which was studied is the Pegasus,[1] a meteoroid detection satellite, whose approximate horizontal dimensions are shown in Figure C-12.

A number of possible coil systems were investigated with the object of obtaining a 5 percent variation of the field over the volume of the satellite. The results of these investigations are included here because they may be useful in other situations.

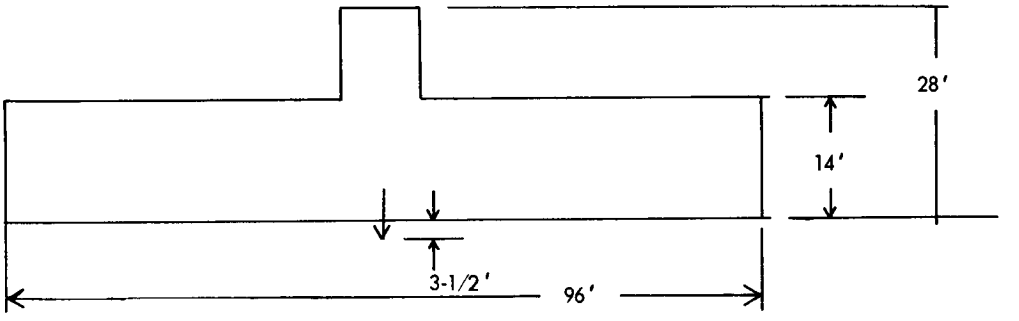


Figure C-12-Dimensions of the Pegasus

TWO- AND FOUR-DIPOLE SYSTEMS

The first work on the dipole system assumed that the coils were magnetic dipoles with the axial variation[2] of the magnetic flux density B as

$$B = \frac{\mu M}{2\pi Z^3} \tag{1}$$

where μ = permeability of the medium

M = magnetic moment of the dipole

= (ampere-turns) (area)

Z = distance measured along the axis from the dipole.

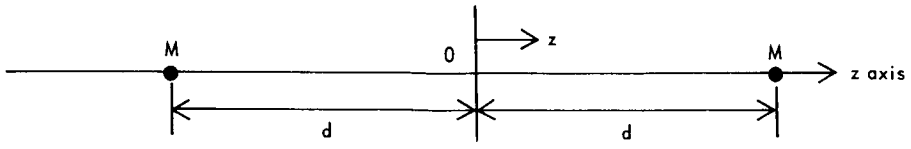


Figure C-13-Two-Dipole System

For two dipoles of equal strength located as shown in Figure C-13, the axial variation of the magnetic flux density B is

$$B = \frac{\mu M}{2\pi(d - Z)^3} + \frac{\mu M}{2\pi(d + Z)^3} = \frac{\mu M}{\pi d^3} f(s) \tag{2}$$

where $s = (Z/d)$

$$f(s) = \frac{1 + 3s^2}{(1 - s^2)^3}$$

$2d$ = axial spacing of two dipoles. At the center of the system, the flux density B_0 is

$$B_0 = \frac{\mu M}{\pi d^3} \tag{3}$$

The flux density of a system along the z axis may be written

$$B = a_0 + a_1 Z + a_2 Z^2 + a_3 Z^3 + a_4 Z^4 + \cdots \tag{4}$$

where the a's are constants. If the system is symmetrical on both sides of the origin, all the odd a's are zero. To improve the uniformity the first step would be to make a_2 zero. This is not possible for the two-dipole system of (2), but it is possible for a four-dipole system. Figure C-14 shows the four-dipole system in which the #1 dipoles have the strength M, and the #2 dipoles have the strength M_2. The axial flux density variation is, from (2)

$$B = \frac{\mu M_1}{\pi d_1^3} f(s_1) + \frac{\mu M_2}{\pi d_2^3} f(s_2) \tag{5}$$

where

$$s_1 = (Z/d_1) \quad \text{and} \quad s_2 = (Z/d_2)$$

The coefficient a_2 of (4) may be made zero for the four-dipole system. From (5), the result is

$$M_2 = -(d_2/d_1)^5 M_1 \tag{6}$$

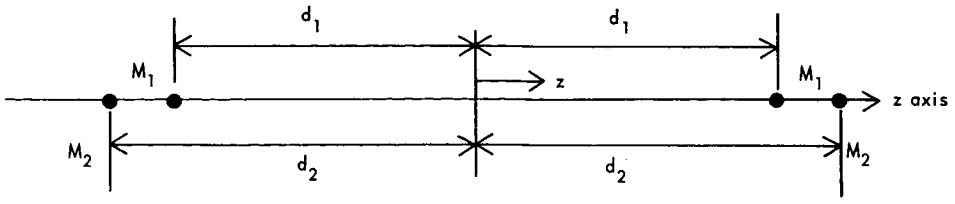


Figure C-14-Four-Dipole System

As an example, if $(d_2/d_1) = 1.1$, then $M_2 = -1.61051\ M_1$. The second dipole must therefore have a strength which is about 61 percent greater than that of the first dipole and which is aimed in the opposite direction. The flux density B at the center from (5) and (6) is

$$B_0 = \frac{\mu}{\pi}\left(\frac{M_1}{d_1^3} + \frac{M_2}{d_2^3}\right) = \frac{\mu M_1}{\pi d_1^3}\left(1 - \frac{d_2^2}{d_1^2}\right) \tag{7}$$

Thus, if $(d_2/d_1) = 1.1$ in (7), the flux density of the four-dipole system at the center is 21 percent of that for the two-dipole system as given by (3), but the volume of uniformity is considerably larger.

The flux density at various points both on and off the axis has been calculated for a four-dipole system, and the contours of the axial component of the field have been plotted. This plot indicates that the four-dipole system has a field near the center very similar to that of a two-coil (Helmholtz) system. The plot may be found in the article by Francine Wright. Calculation of the ampere-turns required on the dipoles indicated that this system would not be practical.

FINITE-SIZE FOUR-COIL SYSTEM

The system investigated was the four-circular-coil system shown in Figure C-15. The investigation was carried out for coils with finite radii, but with zero cross sections. The field on the axis of the coils is given by:

$$H = N_1 I_1 A_1^2 \left\{\frac{1}{\left[(d_1 - X)^2 + A_1^2\right]^{3/2}} + \frac{1}{\left[(d_1 + X)^2 + A_1^2\right]^{3/2}}\right\}$$

$$+ N_2 I_2 A_2^2 \left\{\frac{1}{\left[(d_2 - X)^2 + A_2^2\right]^{3/2}} + \frac{1}{\left[(d_2 + X)^2 + A_2^2\right]^{3/2}}\right\} \tag{8}$$

Equation (8) can be written in the form of a Taylor series as

$$H = \sum_0^\infty a_n X^n \tag{9}$$

where

$$a_n = \frac{1}{n!} \frac{d^n H}{dX^n} \tag{10}$$

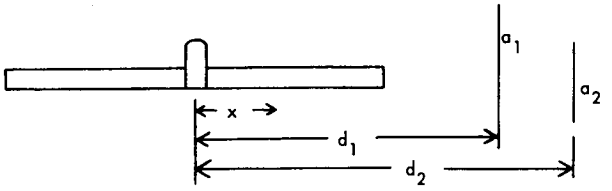


Figure C-15-Four-Circular-Coil System

From the symmetry of the system, the odd-numbered terms will cancel, and the resulting equation is:

$$H = a_0 + a_2 x^2 + a_4 x^4 + a_6 x^6 + \ldots . \tag{11}$$

The a_2 and a_4 terms were set equal to zero; the $a_6 x^6$ term was considered much larger than the remaining terms, and was used to determine the uniformity of the system.[3] The coils' parameters and the field properties were calculated over a wide range of values in dimensionless form. A system was selected with an inner coil separation of 200 feet. This result required that the coils have radii of R_1 = 14.95 feet and R_2 = 10.2 feet. The two coils are separated by 2.01 feet and have a uniform field over a length of 61.5 feet. The major drawback of this system is that the fields of the two coils on each end of the system must oppose each other; this necessitates a large number of ampere-turns to produce the necessary field at the center of the system. It was decided that the ampere-turns were too large for practical use.

TWO SQUARE COILS

This system would have the same general configuration as the four-coil system, but, by eliminating the opposing fields, the ampere-turns requirement would decrease substantially. Square coils were chosen over circular coils because the square coils are easier to construct and have a greater length of uniformity along their axis.

The expression giving the ratio of the field, at any point along the axis, to the field at the center of the system is:

$$H_r = \frac{1}{2}\left\{\left[(D - X)^2 + 1\right]^{-1} \left[(D - X)^2 + 2\right]^{-1/2} + \right.$$

$$\left.\left[(D + X)^2 + 1\right]^{-1} \left[(D + X) + 2\right]^{-1/2}\right\} \left\{(D^2 + 1)\,(D^2 + 2)^{1/2}\right\} \tag{12}$$

where $2d$ = coil separation

x = distance from the center of the system

21 = length of a coil side

$X = x/1$

$D = d/1$

This equation has a family of curves of the form[4] shown in Figure C-16.

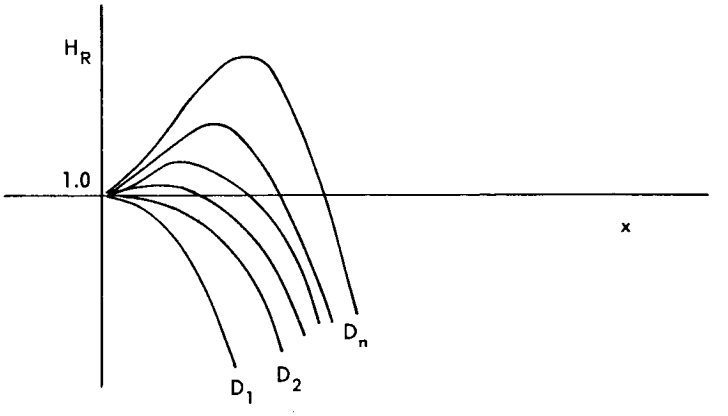


Figure C-16-Axial Field Intensity for Two Square Coils

To find the length of uniformity, the maximum of each curve was found first. Then, 0.05 was subtracted from this maximum value to give the lower limit of the field ratio. The value of x for this lower value of H_r gave the length of uniformity. The largest length of uniformity was $x = .782$ for $D = .695$.

To produce the required field strength and uniformity would require the coils to be 256 feet on a side and separated by 178 feet. The coils would have to carry about 100,000 ampere-turns, and the large size and separation of this system appeared to be impractical.

At this point it was decided to try to orient the spacecraft perpendicular to the earth's field, rather than parallel to it, which would greatly increase the ease of handling and testing the spacecraft.

TWO RECTANGULAR COILS

One system investigated for this orientation was two rectangular coils, as shown in Figure C-17.

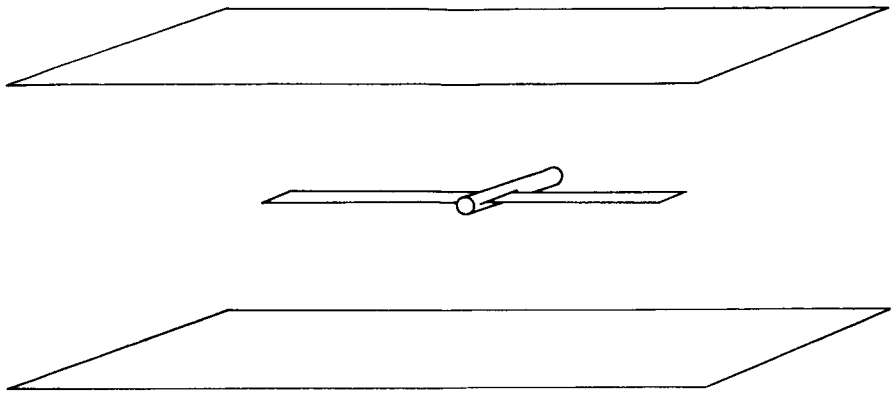


Figure C-17-Two Rectangular Coils

The equation for the axial field at any point between two rectangular coils[5] is given by

$$H_z = \frac{I}{4\pi}\left[\frac{x-a}{r_2(y+b+r_2)} + \frac{x+a}{r_4(y-b+r_4)} - \frac{x-a}{r_3(y-b+r_3)}\right.$$

$$\left. - \frac{x+a}{r_1(y+b+r_1)}\right] - \frac{I}{4\pi}\left[\frac{y+b}{r_1(x+a+r_1)} + \frac{y-b}{r_3(x-a+r_3)} -\right.$$

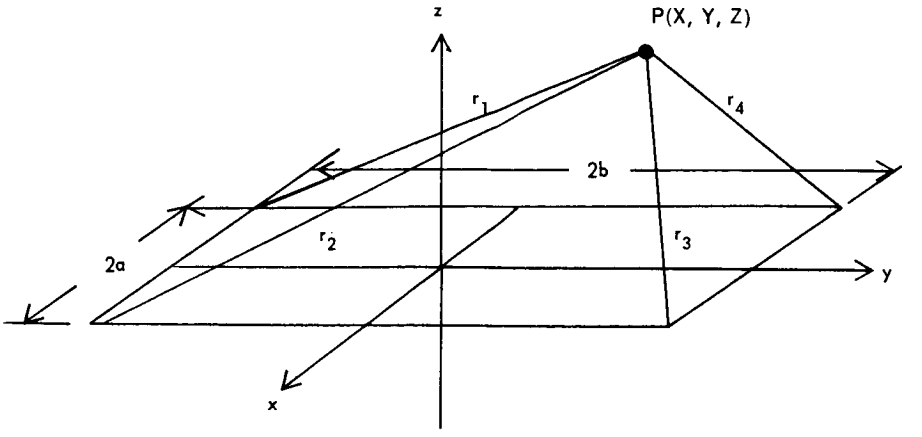
$$\left.\frac{y+b}{r_2(x-a+r_2)} - \frac{y-b}{r_4(x+a+r_4)}\right] \qquad (13)$$

The parameters are those shown in Figure C-18. The length of uniformity along each of the two axes, x and y, must be found and an optimum set of parameters obtained. The field along the x axis decreases monotonically while the field along the y axis goes up to a maximum and then decreases. Some promising preliminary results were obtained, but this work was stopped when another configuration appeared to have better possibilities.

DIPOLE ARRAYS

The idea behind the dipole-array method is to obtain a configuration of sets of dipoles which will produce a large area of uniformity in the plane perpendicular to the coils' axes. For sets of two dipoles, the basic configuration is shown in Figure C-19. For one set of two dipoles, the equation for the field at any point, perpendicular to the coil axis at the center point, is:

$$H_z = \frac{\mu M}{2\pi d^3}\left[\frac{2-x^2}{(1+x^2)^{5/2}}\right] \qquad (14)$$



$$r_1^2 = (x+a)^2 + (y+b)^2 + z^2$$

$$r_2^2 = (x-a)^2 + (y-b)^2 + z^2$$

$$r_3^2 = (x-a^2) + (y-b)^2 + z^2$$

$$r_4^2 = (x-a)^2 + (y-b)^2 + z^2$$

Figure C-18-Geometry of One Rectangular Coil

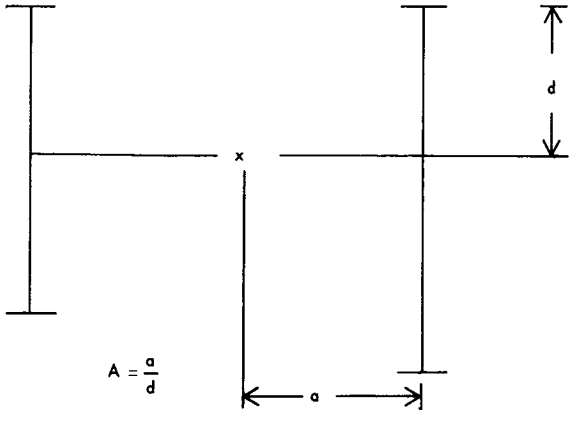


Figure C-19-Array: Two Sets of Two Dipoles

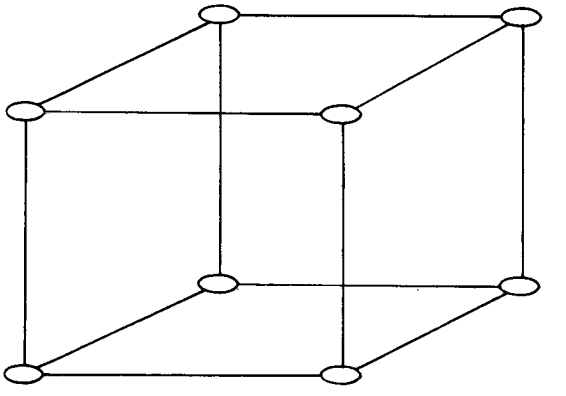


Figure C-20-Array: Four Sets of Two Dipoles

where x_1 = distance from the coil's axis

$x = (x_1/d)$

For two sets of two coils, as shown in Figure C-19, the field between the sets is uniform to 5 percent for A< about 0.12, and is also uniform for A slightly larger than 0.4. Calculations were then made on four sets of two dipoles in the form of a square, as shown in Figure C-20. For A = 0.45, the square has the largest percentage of uniformity with

only small areas near the coils exceeding 5-percent uniformity. Arranging three sets of two dipoles in the form of a triangle, the optimum value of A is 0.49; only small areas, again near the dipoles, exceed 5-percent uniformity.

The final system must also have 12 feet of uniformity along the coil axes; therefore, sets of four dipoles, as shown in Figure C-21, were investigated because they have a greater uniformity in this direction than do sets of two dipoles. The configurations investigated, with $D = (d_2/d_1) = 1.1$ and $M = -1.61$, were the same as those investigaged for sets of two dipoles. The equation used for sets of four dipoles is:

$$H_z = \frac{(2 - x^2)}{(1 + x^2)^{5/2}} + \frac{M(2D^2 - x^2)}{(1 + x^2)^{5/2}} \quad . \tag{15}$$

The optimum separation for two sets of four dipoles is $A = 0.80$; for three sets of four dipoles in triangular form, $A = 0.80$; for four sets of four dipoles in the form of a square, $A = 0.70$. The areas within the 5-percent uniformity range are shown in Figures C-22 and C-23.

Using both configurations (sets of two dipoles and sets of four dipoles), a rectangular system using six sets of dipoles was investigated. For both instances, the configuration failed to produce a substantial uniform field.

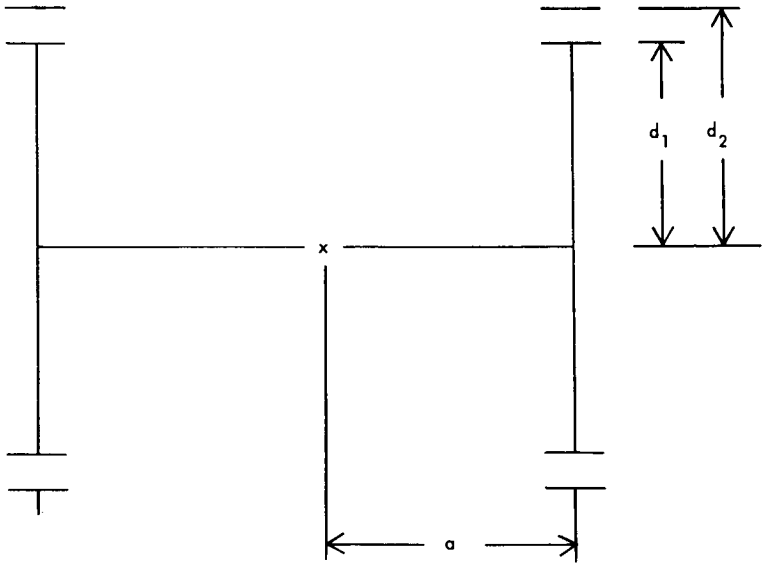


Figure C-21–Two Sets of Four Dipoles

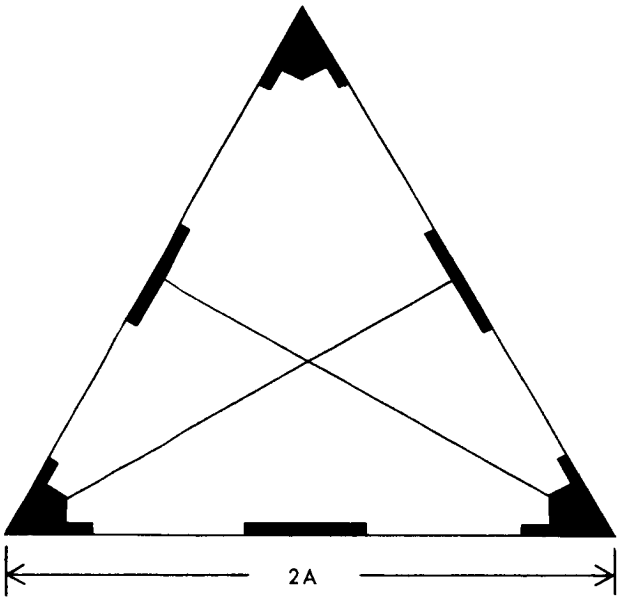


Figure C-22–Triangular Array

A = .80

Areas exceeding 5-percent uniformity

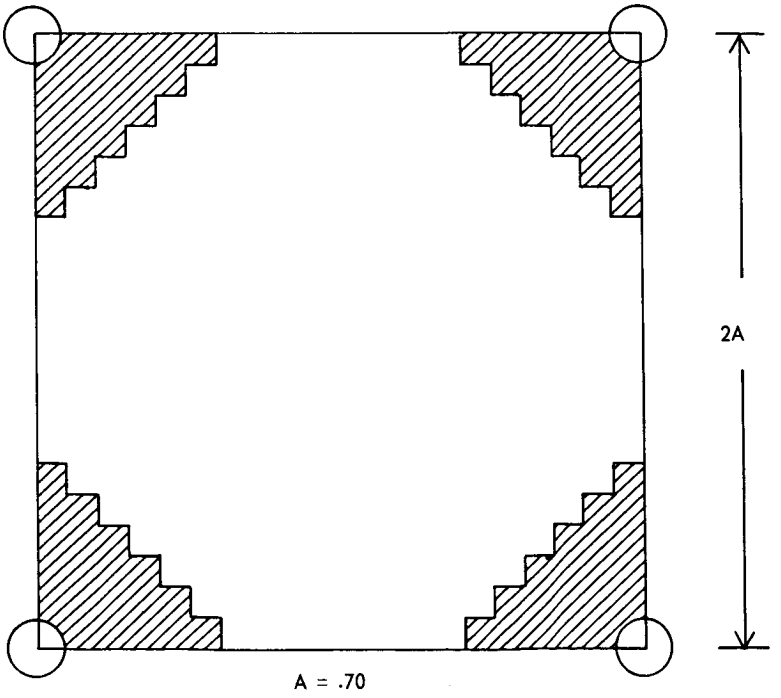
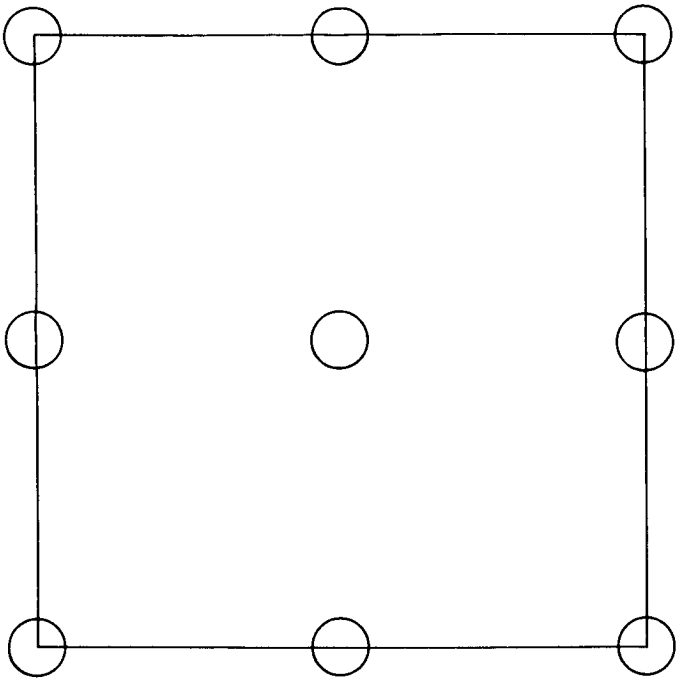


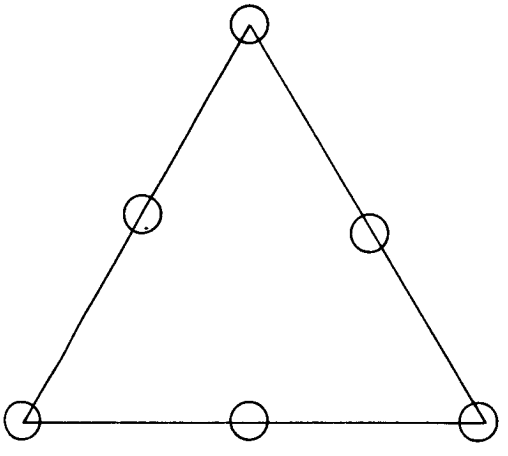
Figure C-23–Square Array

A = .70

Areas exceeding 5-percent uniformity



(a)



(b)

Figure C-24-Enlarged Arrays:
(a) Square, (b) Triangular

The next systems to be investigated should be enlargements on the square and triangular configurations. Figure C-24 shows top views of these configurations. The method of finding solutions for these systems is to set the field at the center equal to the field at various points in the area. By choosing enough points, A can be found along with the ampere-turns ratio of the dipoles. Field plots can then be made with these parameters in order to determine the area of uniformity. It was felt that the large number of dipoles required in these arrays would make them impractical for the problem at hand.

ONE COIL, TWO DIPOLES

Another configuration that appears promising is that of one large coil with two dipoles symmetrically placed along the axis of the coil, as shown in Figure C-25. Along the vertical axis the flux density is:

$$B = \frac{\mu M}{2\pi d^3}\left[\frac{1}{(1-Z)^3} + \frac{1}{(1+Z)^3} + \frac{kb^5}{(Z^2+b^2)^{3/2}}\right] \tag{16}$$

where M = magnetic moment of each dipole

$k = (\pi NI d^2/Mb^3)$

NI = ampere-turns of the coil

$b = (a/d)$

$Z = (Z_1/d)$

Now for the flux density B_0 at the origin:

$$B_0 = \frac{\mu M}{2\pi d^3}(2 + kb^2) \tag{17}$$

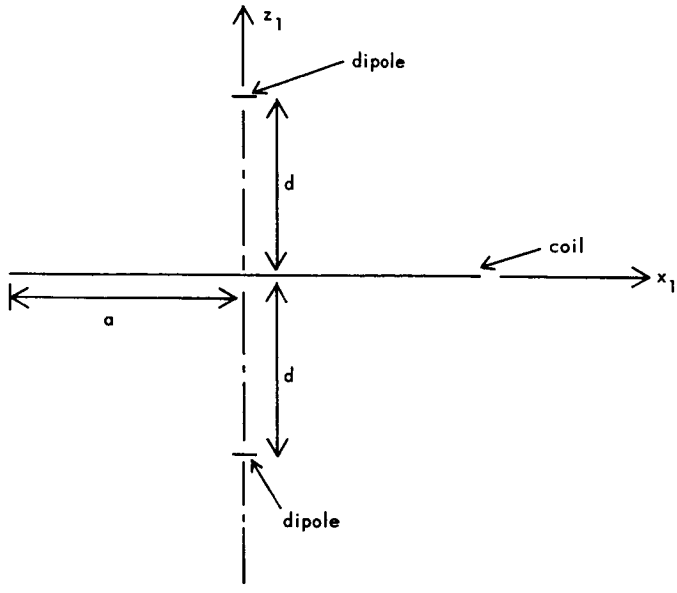


Figure C-25-Array: One Coil and Two Dipoles

A plot of (B/B_0) versus Z is shown in Figure C-26. The variation of (B/B_0) for $Z \leq Z_2$ is within 5 percent. Table C-4 gives corresponding values of b, k, Z_m and Z_t for the optimum systems having a maximum variation of 5 percent.

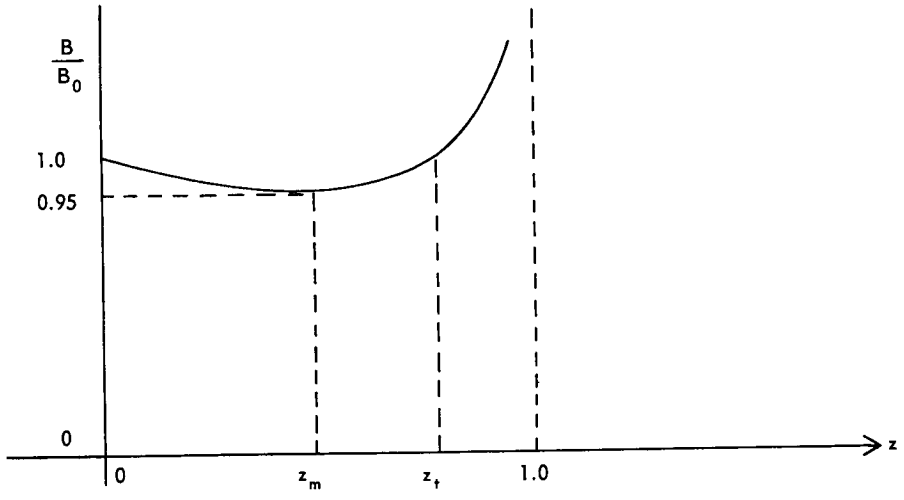


Figure C-26-Variation of Flux Density Along the Z-Axis

Table C-4
Parameters for Optimum Systems with
One Coil and Two Dipoles

k	b	Z_m	Z_t
147.27	0.033784	0.05	0.11509
31.008	0.12081	0.10	0.16761
19.322	0.25003	0.15	0.22968
16.793	0.40772	0.20	0.29376
16.655	0.58869	0.25	0.35722
17.752	0.79399	0.30	0.41905
19.952	1.0252	0.35	0.47880
23.511	1.2821	0.40	0.53624
29.022	1.5628	0.45	0.59119
37.599	1.8643	0.50	0.64354
51.346	2.1831	0.55	0.69321
74.470	2.5156	0.60	0.74015

As an example, this system might be used for testing the OAO satellite with the satellite placed at the center of the system of Figure C-25. The axis of the satellite would be placed along the axis of the system. From Table C-4, take K = 37.599, b = 1.8643, Z_m = 0.50 and Z_t = 0.64354. Assuming the axial length of the satellite to be 12 feet, then

$$Z_t d = \frac{12}{2}$$

or $d = \frac{6}{.64354} = 9.32$ feet or 2.84 meters

Now $b = (a/d) = 1.8643$

or $a = (1.8643)(9.32) = 17.32$ feet or 5.29 meters.

Take the flux density at the origin to be 0.6 gauss or 0.6×10^{-4} weber per meter^2. Then, from (17)

$$B_0 = 0.6 \times 10^{-4} = \frac{\mu M}{2\pi d^3} (2 + kb^2)$$

or $M = 51.8$ ampere-turns-meter^2. Assuming the dipole is a coil one meter in diameter, then the area of the coil is $(\pi/4)$ meter^2. The ampere-turns of the dipole then would be

$$N_d I_d = \frac{51.8}{(\pi/4)} = 66 \text{ ampere-turns.}$$

The dipole might then be a 66-turn coil having a diameter of 1 meter and carrying a current of 1 ampere. The ampere-turns of the large coil may be obtained from (16):

$$K = 37.599 = \frac{\pi N I d^2}{M b^3}$$

or $NI = 498.$ ampere turns.

Hence, the large coil might be a 10-turn coil 34.64 feet in diameter and carrying a current of 49.8 amperes.

To apply the system of one coil and two dipoles to the Pegasus satellite, a similar analysis must be carried out for the field in the χ_1 direction, as shown in Figure C-25. This analysis has been started but has not been finished. In addition, a plot of the contours of flux density would be helpful.

REFERENCES

1. La Fond, C.D., "Meteoroid Detection Satellite Mocks-up Shown," Missiles and Rockets, pages 32-33 (June 24, 1963)

2. Hallen, E., "Electromagnetic Theory," J. Wiley & Sons, New York (1962) page 154

3. Pittman, M.E., and Waidelich, D.L., "Three- and Four-Coil Systems for Homogeneous Magnetic Fields," NASA Tech. Note D-2095 (January 1964)

4. Heller, Carl, "Über die Erzeugung grossräumiger homogener Magnetfelder zum Studium des Verhaltens von Magnetkompassen und Kompensiermitteln auf verschiedenen magnetischen Breiten," Duetschen Hydrographischen Zeitschrift, Vol. 8, No. 4, pages 157 to 164 (1955)

5. Yildirim, Gungor, "Production of Uniform Magnetic Field by Using Square Coils," Research Report, University of Missouri Engineering Library (June 1963)

Page intentionally left blank

CONTOUR CURVES FOR THE AXIAL MAGNETIC FIELD OF A FOUR-DIPOLE SYSTEM

Francine Wright

The curves shown in Figure C-27 are the plotted results of the axially-directed H field. The horizontal axis of the graph represents the distance along the axis of the system, and the vertical axis of the graph is the radial distance. The graph is one quadrant (i.e., one-fourth) of the total picture.

The 1.000 contours, which are nearly straight lines radiating from the origin, have the same field intensity (HZ) as that of the center (origin) of the system. The contours at about 45°, as indicated by their values greater than unity, have greater field intensities than that at the center. The contours along the axes are of smaller field intensity than the contour at the center.

Note that the field decreases more rapidly in the axial direction than in the radial direction. This is because the intensity approaches $-\infty$ at the M_1 dipole; therefore, the field must decrease sufficiently rapidly to meet this restriction. This limitation is not present in the radial direction.

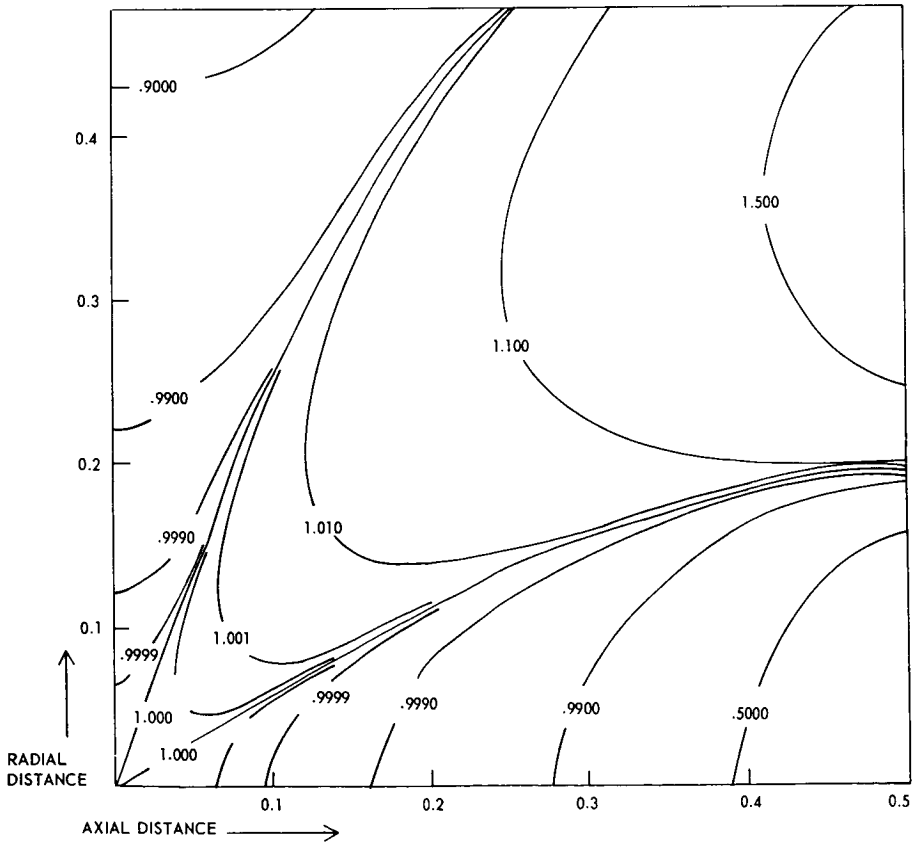


Figure C-27-Axial Magnetic Field of a Four-Dipole System

Page intentionally left blank

MEASUREMENT OF SMALL MAGNETIC TORQUES: PARALLEL PLATE METHODS

D. L. Waidelich

A project proposed for the Attitude Control Test Facility is to float a spacecraft on water and measure the torque exerted on the craft by the magnetic field. This torque is expected to be very small, of the order of 10, to 10000. dyne-centimeters; the problem is to provide a torque in the opposite direction that will stabilize the whole system so that the original torque may be measured.

One way to provide a torque of this magnitude would be to use parallel plates with a voltage between the plates, and the resulting force between the plates would supply the restoring torque. Figure C-28 shows one possible circuit in which C_1 is a fixed capacitor acting as a reservoir of charge, supplying the capacitor C whose upper plate may be moved a distance x. Both capacitors are charged to the same voltage V and the force F is exerted on the upper plate of C. An analysis of the circuit of Figure C-28 indicates that the force F is

$$F = \frac{1}{2} V^2 \frac{dC}{dx} \tag{1}$$

no matter how large or how small C_1 is compared to C. If w is the distance between the plates of capacitor C in Figure C-28, then (neglecting the fringing flux) the capacitance C is

$$C = \frac{\epsilon A}{w - x} \tag{2}$$

where ϵ = permittivity of the dielectric between the plates

A = area of one plate

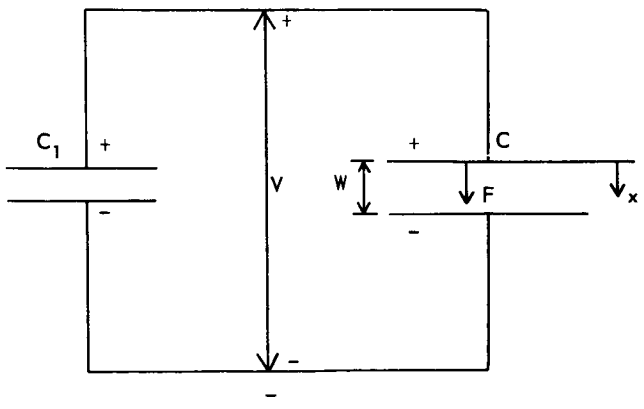


Figure C-28–Circuit for the Forces between Plates of a Capacitor

from (1) and (2), the force F at x = 0 is

$$F\Big|_{x=0} = \left[\frac{V^2 \epsilon A}{2(W-x)^2}\right]_{x=0} = \frac{V^2 \epsilon A}{2W^2}$$

which indicates that the force tends to move the plates together (x increases) and the force increases as the two plates approach each other. For plates in air separated by one centimeter and with an area of one square meter the force for a voltage of 1000. volts is 4420. dynes.

A more practical configuration is the three-plate arrangement shown in Figure C-29. The middle plate is assumed to be movable by reason of the forces exerted from each outer plate. Unfortunately, the system is in an unstable equilibrium at x = 0 because, as x increases, the force from the right-hand plate increases and that from the left decreases, tending to increase x further. If fringing is neglected, the capacitances C_A and C_B are

$$C_A = \frac{\epsilon A}{a + x} \tag{3}$$

$$C_B = \frac{\epsilon A}{a - x}$$

where A is the area of each plate and ϵ is the permittivity of the medium. Then, by the use of (1) and (3), the force F acting on the middle plate is

$$F = \frac{1}{2} V^2 \left(\frac{dC_A}{dx} + \frac{dC_B}{dx}\right)$$

$$= \frac{1}{2} V^2 \left[-\frac{\epsilon A}{(a+x)^2} + \frac{\epsilon A}{(a-x)^2}\right] \tag{4}$$

$$= \frac{2aV^2 \epsilon A x}{(a^2 - x^2)^2}$$

where V is the voltage from the middle plate to each of the outer plates. The force F from (4) increases nearly linearly with x for small values of x with the slope

$$\text{slope} = \frac{dF}{dx}\bigg|_{x=0} = \frac{2V^2 \epsilon A}{a^3} \tag{5}$$

Again, using a = 1 centimeter, A = 1 square meter, and V = 1000 volts, the slope is 1768 dynes per millimeter.

In the electrostatic gyroscopes[1,2] two methods have been used to transform the unstable system of Figure C-29 into stable systems. The first one (which might be called the dc servo method) consists in applying a much higher voltage V to C_A of Figure C-29 and a much lower voltage V_B to C_B as the distance x increases. The effect will be to produce a force F in the direction opposite to the increase in x, so that a stable system will result. Suppose that

$$V_A = V + Kx$$
$$V_B = V - Kx \tag{6}$$

where K is a constant indicating how many volts per meter the distance sensor and servo amplifier feed back to V_A and V_B. Then, from (1), (3), and (6),

$$F = \frac{1}{2} V_A^2 \frac{dC_A}{dx} + \frac{1}{2} V_B^2 \frac{dC_B}{dx}$$

$$= \frac{\epsilon A}{2} \left[-\left(\frac{V+Kx}{a+x}\right)^2 + \left(\frac{V-Kx}{a-x}\right)^2 \right] \tag{7}$$

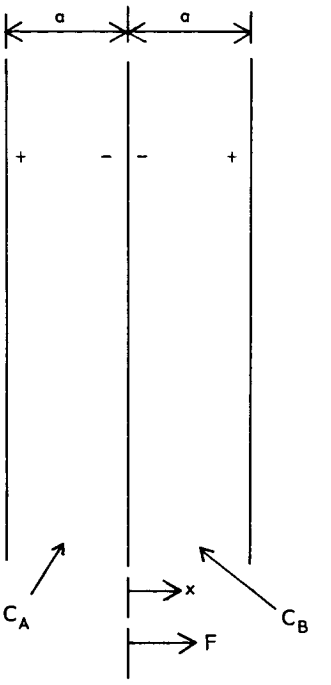


Figure C-29–Three-Plate Configuration

and the slope for x small is

$$\text{slope} = \left.\frac{dF}{dx}\right|_{x=0} = \frac{2 \epsilon A V (V - Ka)}{a^3} \tag{8}$$

The system will be stable if the slope of (8) is negative; this requires that

$$K > (V/a) \tag{9}$$

The servo amplifier must have enough gain to satisfy the inequality (9).

The second method of producing a stable system is to make C_A part of an ac-resonant circuit, as shown in Figure C-30. Then

$$V_A = \frac{V_T}{1 - \omega^2 L_A C_A} \tag{10}$$

where ω is the angular frequency of the alternating current. Put (3) in (10):

$$|V_A| = \frac{V_T (a + x)}{\omega^2 L_A A\epsilon - (a + x)} = \frac{V_T (a + x)}{K_T - x} \tag{11}$$

where

$$K_T = \omega^2 L_A A\epsilon - a > 0 \tag{12}$$

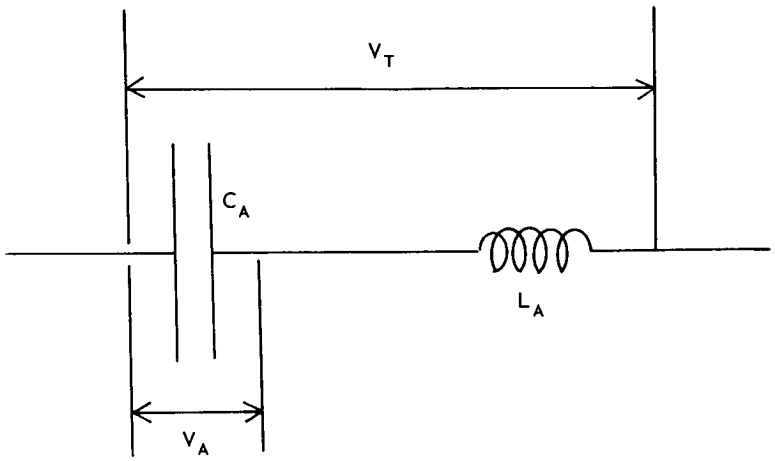


Figure C-30–AC-Resonant Circuit

and the value of L_A has been chosen large enough to satisfy the inequality of (12). If a similar circuit is used for C_B,

$$|V_B| = \frac{V_T (a - x)}{K_T + x} \tag{13}$$

Then, by use of (3), (4), (11), and (13),

$$F = -\frac{2 \epsilon A V_T^2 K_T x}{(K_T^2 - x^2)^2}, \tag{14}$$

and

$$\text{slope} = \left.\frac{dF}{dx}\right|_{x=0} = -\frac{2 V_T^2 \epsilon A}{K_T^3} \tag{15}$$

which is negative, indicating that the system is in stable equilibrium.

REFERENCES

1. Knoebel, H.W., "The Electric Vacuum Gyro: Pinpoint for Polaris Launching," Control Engineering, Volume 11, pages 70-73 (February 1964)

2. Riesgraf, M.H., "Electrically-Suspended Gyros," Aircraft Engineering, Volume 36, pages 18 to 20 (January 1964)

MEASUREMENT OF SMALL MAGNETIC TORQUES: DIELECTRIC-PLATE, PIEZOELECTRIC, AND ELECTRET METHODS

A. V. Dralle

The system shown in Figure C-31 may be employed in the measurement of a small torque on a suspended body. In the figure, S is the suspended body, free to rotate about axis 0. Instrument D is a device designed to provide an equilibrium position under action of the torque τ and a restoring torque τ_R, and to measure some quantity dependent on τ.

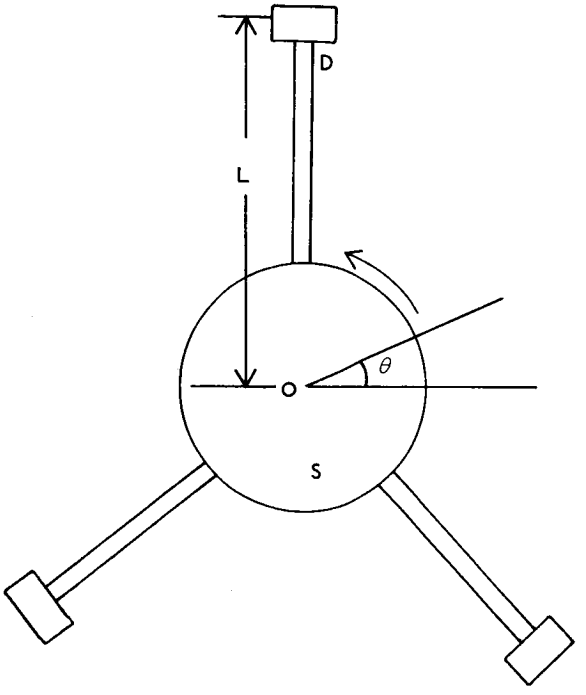


Figure C-31–Suspended Body Subject to Torque

The device D must be sensitive enough to detect the action of the very small torques anticipated, τ = 10 to 10^4 dyne-cm. Three such devices studied are the dielectric plate, piezoelectric crystal, and electret.

DIELECTRIC PLATE

Energy considerations show that a dielectric plate of a shape similar to that of the plates of a parallel-plate capacitor is in stable equilibrium[1,2] when centered within the capacitor. Thus, as shown in Figure C-32, the dielectric, when displaced from its centered position, is acted on by a restoring force F, tending to return it to the equilibrium position.

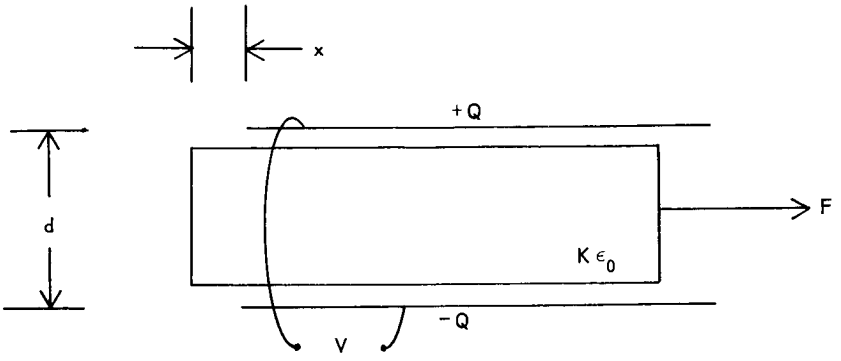


Figure C-32–Dielectric Plate Stabilization

If D in Figure C-31 consists of a dielectric between two capacitor plates, the deflecting torque τ will be opposed by the restoring torque τ_R arising from F. In the state of equilibrium, then, the system will have rotated through an angle θ, and $\tau = \tau_R$. If τ_R is a function of θ, a measurement of θ would determine τ. The angle θ could be measured to within 2 seconds of arc by use of an autocollimator.

The restoring torque τ_R was found in all cases to be a sum of two terms: $\tau_R = \tau_0 + \tau_f$, where τ_0 = torque calculated neglecting fringing fields of the capacitor, and τ_f = contribution of fringing fields to τ_R.

Two situations were considered in the calculation of τ_R: (a) constant voltage across the capacitor, and (b) constant charge Q on the capacitor plates.

(a) The restoring force on the dielectric is found to be

(neglecting fringing fields):

$$F_R = -\frac{\epsilon_0 (K - 1)}{2} S\epsilon^2$$

where ϵ_0 = permittivity of space, K = dielectric constant of a dielectric of cross section area S, and ϵ = electric field between capacitor plates. For plates of area 0.5 square meter; dielectric of thickness 1 cm. and K = 10, and voltage V = 1000 volts, a restoring torque $\tau_R = 2 \times 10^4$ dyne-cm. would be obtained, sufficient to balance the largest anticipated τ. In this case, however, τ_R is independent of θ, and there is no unique equilibrium position.

(b) The restoring force on the dielectric is found to be (neglecting fringing fields):

$$F_R = -\frac{Q^2 d}{2\epsilon_0 W} \frac{K - 1}{[x + K(\ell - x)]^2}$$

where d = separation of capacitor plates, W and 1 are dimensions of the dielectric, and x = displacement of the dielectric from its center position. For values of these quantities closely related to those in a), τ_R is again found to be about 2×10^4 dyne-cm., again

adequate to balance τ. In this case, F_R is a function of the displacement of the dielectric, in the form

$$F_R \propto \frac{1}{(a + y)^2}$$

However, the dependence on x is very weak, since $a >> x$, and thus F_R is essentially independent of the small values of x anticipated.

It thus appears that the forces calculated for this geometry, neglecting fringing fields, are practically independent of the position of the dielectric, and measurements of θ do not determine τ_R and τ. This problem may be overcome either by finding a new capacitor geometry in which F_R is more strongly related to θ, or by calculating the contribution of the fringing fields τ_f.

PIEZOELECTRIC CRYSTAL

A piezoelectric crystal[3] has the property of developing voltages across its volume upon the application of a mechanical stress, as shown in Figure C-33.

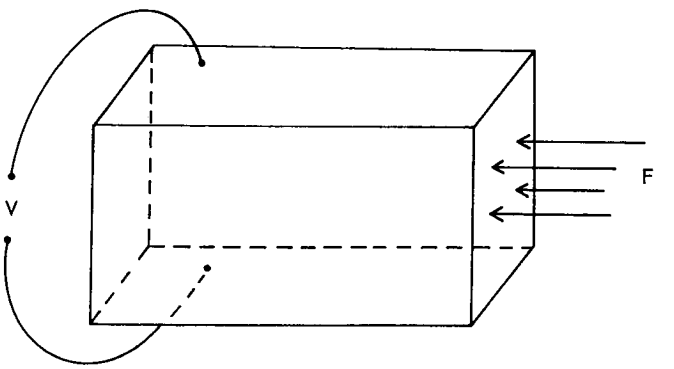


Figure C-33-Piezoelectric Crystal

In Figure C-31, D may consist of a rigid block clamped between piezoelectric crystals on either side which provide the balancing torque τ_R against τ. Torque τ_R arises from the reaction of the crystal against the pressure on it of the rigid block, and voltage between two of its faces is thus a measure of the torque τ.

The voltage V to be measured is found to be

$$V = \frac{dt}{K\epsilon_0} \frac{\tau}{AL}$$

where t, A, and K are the thickness, surface area and dielectric constant, respectively, of the crystal, d is a tabulated quantity for piezoelectric crystals, and L is the distance of the crystal from the system's rotation axis. For $t = 1$ cm., $A = 50$ sq. cm., $K = 30$, $d = 48 \times 10^{12}$ coul per square meter/newtons per square meter, and $L = 2$ meters, then $V = 10^{-4}$ through 10^{-7} volts, corresponding to the expected range of values of τ. These

voltages appear to be measurable, at least within a few percent, and indicate the apparent applicability of this technique to determination of τ.

ELECTRET

An electret[4] is a dielectric crystal with the property of being permanently electrically polarized along a single direction. It is essentially a dipole, and behaves as the electrical analog of a magnet. An important consequence of its polarization is the presence of a surface charge on the end faces normal to the polarization direction. This change may be as large as 3 e.s.u. per cm.^{2}.

In Figure C-31, D may consist of two such electrets, positioned with their charged faces in opposition, and interacting by simple Coulomb forces. If the faces carry charge of the same sign, they will tend to repel one another, giving rise to the desired restoring torque τ_R against the disturbing torque τ.

A first-order analysis of the repulsive Coulomb forces assumes that the charges are concentrated at the center of their surfaces, which then interact like point charges. Using surface charge density σ = 2 e.s.u. per cm.^{2}, surface area, S = 10 cm^2, and distance between electret faces, d = 20 cm, the force of repulsion becomes

$$F_R = \frac{(\sigma S)^2}{d^2} = 1 \text{ dyne},$$

or a restoring torque of about 100 dyne-cm. This is exactly in the desired range, and, most importantly, depends upon the distance between faces. Measurements of the disturbance of the system (from a previously established equilibrium position), under the action of torque τ, would uniquely determine τ.

It is of interest that electrets, particularly the recently developed ceramic electrets,[5,6] maintain their polarization and surface charge virtually unchanged over a period of several years. A single calibration of an electret measuring device would thus be sufficient for measurements over a long period of time.

Electrets may be prepared in a laboratory following the method outlined by Gubkin and Skanavi.[6]

REFERENCES

1. Frank, N.H., "Introduction to Electricity and Magnetism," New York, McGraw-Hill (1950)
2. Feynman, , "The Feynman Lectures on Physics," Volume II, Reading, Mass., Addison-Wesley (1963)
3. Mason, W.P., "Piezoelectric Crystals and their Application to Ultrasonics," New York, Van Nostrand (1950)
4. Gutman, F., "The Electret," Reviews of Modern Physics, 20(3):457-472 (1948)
5. Scleicker, E., "Investigation on the Practical Applicability of Ceramic Electrets," Experimentelle Technik der Physik, 10(3):224-230 (1962)
6. Gubkin, A.N., and Skanavi, G.I., "Some New Electrets from Inorganic Dielectrics," J. Exptl. Theoret. Physics (U.S.S.R.), 32, 140-142 (1957)

MEASUREMENT OF ELECTRIC-FIELD INTENSITY

D. L. Waidelich

Working with electrostatic fields requires a method of measuring electric-field intensity. Ideally this would mean that, at a given point, the three components or the magnitude and direction of the field intensity would be measured with no disturbance to the field itself.

Much previous work has been concerned with measurement of the field intensity at the surface of the earth or at heights attainable by airplanes and balloons.[1,2] As the direction of the electric field in this case is vertical, or nearly so, this led to the development of various instruments using an antenna or probe. If several wires were located above the ground and insulated from it, the wires would gradually become charged by the natural ionization of the air to a potential corresponding to the potential difference between the wires and the ground. The amount of charge (and thus the potential on the wires) could be determined by connecting the wires to the ground through an appropriate instrument to measure charge. The time involved for the wires to reach an equilibrium potential was rather long; to accelerate this process, various ionization and other methods were used, such as water drops, flames, and point discharges, as well as ultraviolet, X-ray and radioactive[3,4] sources.

A second instrument for measuring the electric-field intensity is the field mill[1,2,5,6,7,] This consists of an insulated stationary flat circular metal plate divided into four or more sectors with every other sector removed. A similar plate is rotated immediately above the fixed plate so that the fixed plate is alternately exposed to, then shielded from, the electric field. The fixed plate is connected to an amplifier which amplifies the resulting alternating current. To measure a potential gradient, two of these instruments are required[7] or one double field mill.[8] For three axes, three double field mills would be needed,[9] or possibly all three could be incorporated into one instrument.[10]

Several field mill experiments have been flown in rockets and satellites.[11,12,13,14] Difficulty was experienced in all these experiments because the plasma sheath about the instrument effectively shielded the instrument from the electric field outside. An attempt to separate the effects of the sheath and the electric field has been made[9] by the use of two frequencies in a field mill.

Probes[9,14,15,16] have also been studied as a means of measuring electric-field intensity. Two or more of the probes should be spaced far enough apart so that their sheaths will not meet. This means that the electric field between the two sheaths would affect the currents flowing in the probes.

Another method that appears promising is the use of a vibrating charged probe.[17,18] This method appears to be used in a recently announced "dynamic electrometer".[19] Other possible ways of measuring the electric-field intensity are by means of the fluxgate method used in magnetometers,[20] by the use of electrets,[21] and by use of the Stark effect.[22]

A problem closely connected to that of measuring electric-field intensity is that of determining the electric potential of a satellite.[23,24] A computer solution[25] and two model experiments[26,27] have been made on this problem. Some methods of finding the potential are reviewed in reference 9. Several experimental results have been obtained by probes on rockets[28,29] and on a satellite.[30]

REFERENCES

1. Chalmers, J.A., "Atmospheric Electricity," Macmillan Co., New York, pages 81 to 127 (1957)

2. Reiter, R., "Meteorobiologie und Elektrizität der Atmosphäre," Akademische Verlagsgesellschaft, Leipzig, pages 53 to 61 (1960)

3. Rein, G.C., Stergis, C.G., and Kangas, T., "An Airborne Electric Field Meter," IRE Transactions, volume I-6, pages 195 to 199 (September 1957)

4. Paltridge, G.W., "Measurement of the Electrostatic Field in the Stratosphere," Journal of Geophysical Research, volume 69, pages 1947 to 1954 (May 15, 1964)

5. Gunn, R., "Electric Field Meters," Review of Scientific Instruments, volume 25, pages 432 to 437 (May 1954)

6. Mapleson, W.W., and Whitlock, W.S., "Apparatus for the Accurate and Continuous Measurement of the Earth's Electric Field," Journal of Atmospheric and Terrestrial Physics, volume 7, pages 61 to 72 (1955)

7. Clark, J.F., "Airborne Measurement of Atmospheric Potential Gradient," Journal of Geophysical Research, volume 62, pages 617 to 628 (December 1957)

8. Smiddy, M., and Chalmers, J.A., "The Double Field-mill," Journal of Atmospheric and Terrestrial Physics, volume 12, pages 206 to 210 (1958)

9. Chalmers, J.A., Hutchinson, W.C.A., Wildman, P.J.L., and Edwards, M.G., "Measurement of Low Electric Fields under Upper Atmosphere Conditions," Final Technical Report (November 1962) Contract No. AF61(052) – 278, Defense Documentation Center AD-418211

10. Kasemir, von H. – W., "Die Feldkomponentenmühle," Tellus, volume 3, pages 240 to 247 (1951)

11. Gdalevich, G.L., "Measurement of the Electrostatic Field Strength at the Surface of a Rocket Flying through the Ionosphere," NASA TTF-8527

12. Imyanitov, I.M., and Shvarts, Ya. M. "Electrostatic Field Strength Measurement with Sputnik III," NASA TTF-8528

13. Imyanitov, I.M., Gdalevich, G.L., and Shvarts, Ya. M., "Measurement of the Electrostatic Field Intensity at the Surface of Geophysical Rockets Moving in the Upper Atmospheric Layers," NASA TTF-8529

14. Bourdeau, R.E., Donley, J.L., and Whipple, E.C., Jr., "Instrumentation of the Ionosphere Direct Measurements Satellite (Explorer VIII)," NASA TN D-414 (April 1962)

15. Boyd, R.L.F., "Some Techniques of Physical Measurement," Proceedings of the Royal Society, Series A, volume 253, pages 516 to 522 (1959)

16. Johnson, D.W., and Kavadas, A., "A Rocket-Borne Electric Field Probe," Canadian Journal of Physics, volume 41, pages 1980 to 1990 (1963)

17. Gershtein, G.M., "Application of a Vibrating Probe in the Simulation of Fields by the Induced-Current Technique," Soviet Physics – Technical Physics, volume 5, pages 688 to 690 (1960)

18. Gershtein, G.M., and Fedonin, G.K., "Simulation Device with Vibrating Charged Probe for Investigation of Two-Dimensional Fields," Soviet Physics – Technical Physics, volume 7, No. 1, pages 74 to 78 (July 1962)

19. Waller, L., "Device to Find Space Vehicle Electrostatic Buildup Set," Electronic News, page 37 (August 26, 1964)

20. Geyger, W.A., "Ring-Core Flux-Gate Magnetometers with Self-Saturating Circuits," AIEE Transactions, Communications and Electronics, No. 68, pages 508 to 514 (September 1963)

21. Gemant, A., "The use of Electrets in Electrical Instrumentation," Review of Scientific Instruments, volume 11, pages 65 to 71 (February 1940)

22. Condon, E.V., and Odishaw, H., "Handbook of Physics," McGraw Hill, New York, N.Y. (1958) pages 7-43 to 7-47

23. Brundin, D.L., "Effects of Charged Particles on the Motion of an Earth Satellite," AIAA Journal, volume 1, pages 2529 to 2538 (November 1963)

24. Kurt, P.G., and Moroz, V.I., "The Potential of a Metal Sphere in Interplanetary Space," Planetary and Space Sciences, volume 9, pages 259 to 268 (May 1962)

25. Davis, A.H., and Harris, I., "Interaction of a Charged Satellite with the Ionosphere," in "Rarefied Gas Dynamics" edited by L. Talbot, Academic Press, New York, (1961) pages 691 to 699

26. Hall, D.F., Kamp, R.F., and Sellen, J.M., Jr., "Plasma – Vehicle Interaction in a Plasma Stream," AIAA Journal, volume 2, pages 1032 to 1039 (June 1964)

27. Knechtel, E.D., and Pitts, W.C., "Experimental Investigation of Electric Drag on Satellites," AIAA Journal, volume 2, pages 1148 to 1151 (June 1964)

28. Sagalyn, R.C., Smiddy, M., and Wisnia, J., "Measurement and Interpretation of Ion Density Distributions in the Daytime F Region," Journal of Geophysical Research, volume 68, pages 199 to 211 (January 1, 1963)

29. Sagalyn, R.C., and Smiddy, M., "Electrical Processes in the Nighttime Exosphere," Journal of Geophysical Research, volume 69, pages 1809 to 1823 (May 1, 1964)

30. Bourdeau, R.E., and Donley, J.L., "Explorer VIII Satellite Measurements in the Upper Ionosphere," Goddard Space Flight Center Publication X-615-63-165, (August 1963)

Page intentionally left blank

ELECTRIC FIELD OF FOUR CHARGED RINGS

A. V. Dralle

INTRODUCTION

The need for controlled electric fields, highly homogeneous over a large volume, could arise from the calibration of electric field instruments or from future complete environmental studies of spacecraft. Investigations were made of a possible method of achieving these fields, that of four symmetrically spaced charged circular rings. Many solutions were found covering the entire range of parameters involved; this article reviews those of particular interest. The close analogy of this problem with the corresponding magnetic-field problem[1] enabled use of techniques already applied in the magnetic case.

THEORY

The electric field on the axis of two charged circular rings[2] is

$$E(z) = \sum_{\ell=0}^{\infty} a_{2\ell} z^{2\ell} \tag{1}$$

where z is the distance on the axis measured from the origin 0 shown in Figure C-34, with $z < B$. In this expression,

$$a_{2\ell} = \frac{Q}{2\pi\epsilon_0} \frac{(2\ell + 1)}{B^{2\ell + 2}} P_{2\ell+1}(x) \tag{2}$$

where Q is the magnitude of charge on the oppositely charged rings, B is the slant height from the origin to the rings, $P_{2\ell+1}(x)$ is the $(2\ell+1)$ th order Legendre polynomial, and $x = \cos\theta$. $(1/2\pi\epsilon_0)$ is the coefficient entering from the use of MKS units, and will henceforth be set equal to unity. Equations (1) and (2) are derived by extending the well-known one-ring solution[1] to include two oppositely charged rings symmetrically placed about the origin 0.

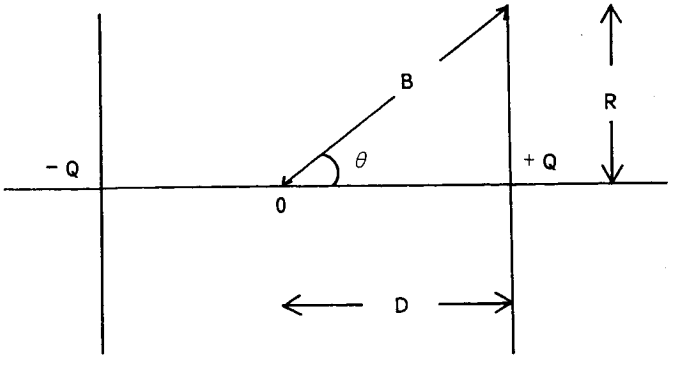


Figure C-34-Geometry for Two Charged Rings

For a symmetrical four-ring system, the axial field is

$$E(z) = \sum_{\ell=0}^{\infty} A_{2\ell} z^{2\ell} \tag{3}$$

where

$$A_{2\ell} = (2\ell + 1)\left[\frac{Q_1}{B_1^{2\ell+2}} P_{2\ell+1}(x_1) + \frac{Q_2}{B_2^{2\ell+2}} P_{2\ell+1}(x_2)\right] \tag{4}$$

in which the quantities Q, B, and x have the same meanings as in the two-ring system, and the subscripts 1 and 2 refer to the ring pairs 1 and 2 respectively, as shown in Figure C-35.

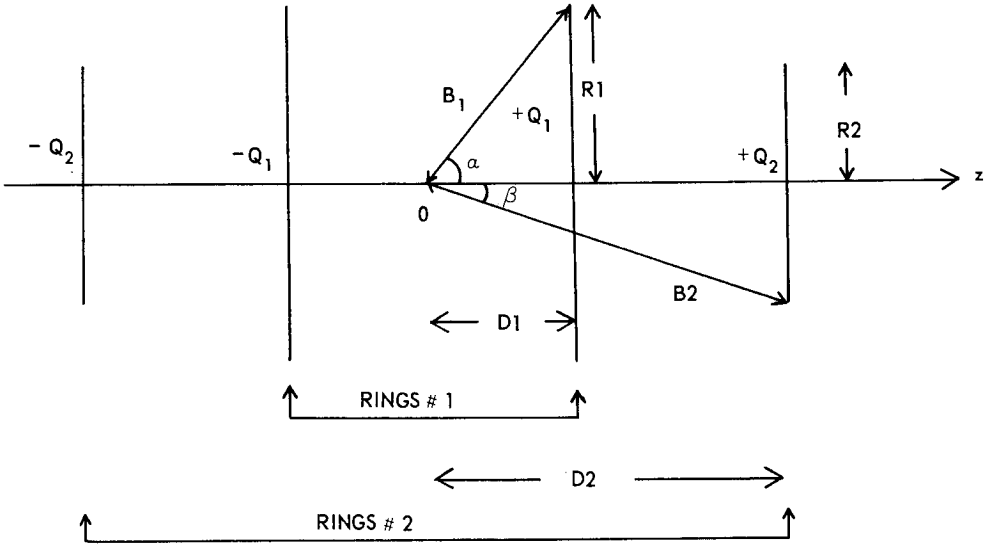


Figure C-35-Four-Ring System

Figure C-35 also indicates the quantities D1, D2, R1, and R2, which are the distances from the origin and the radii of rings 1 and 2 respectively. In the four-ring results which follow, Q1 and R1 are set equal to unity, enabling a more direct comparison with D1, D2, Q2, and R2.

FOUR-RING RESULTS

A high degree of field uniformity near the origin occurs in four-ring systems with parameters obtained by setting A2, A4, and A6 of Equation (3) to zero. It was found that two families of solutions exist, corresponding to $Q_2/Q_1 > 0$ and $Q_2/Q_1 < 0$. Table C-5 gives the results of solutions detailed in the Appendix.

Table C-5
Parameters for a Four-Ring System, $Q_2/Q_1 > 70$

x_1	x_2	R_2	Q_2	D1	D2	Endpoint
.40584	.77460	∞	0	.44	.30	Endpoint
.44	.8417	1.028	.0598	.49	1.60	Nearly equal R's
.53	.9025	.5248	1.004	.625	1.099	Nearly equal Q's
.64	.9370	.3100	9.883	.833	.832	Nearly equal D's
.758	.9494	.1169	4270.	1.162	.3534	Maximum x_2
.77460	.9491	0	∞	1.22	.0699	Endpoint

$Q_2/Q_1 > 0$: Solutions were found for x_1 ranging from 0.40584 to 0.77460, in which interval x_2 increased from 0.77460 to 0.94938, and then decreased to 0.94911.

Based on R1 = 1, D1 ranged from about 0.44 to about 1.22, while D2 ranged from about 3.00 to about 0.07, in the above interval for x_1 . The rings are equidistant from the origin at $x_1 = 0.64$, when D1 = D2. At this point, the two rings on one side of the origin lie in the same plane, with the #2 rings having a diameter about 0.31 that of the #1 rings. It also follows that the #2 rings lie closer to the origin than the #1 rings for $0.64 < x_1 < 0.77460$.

Based on Q1 = 1, Q_2 ranged from zero ($x_1 = 0.40584$) to infinity ($x_1 = 0.77460$). Near $x_1 = 0.53$, Q2 = 1, so that all rings would have the same magnitude of charge. Rings #2 have $Q_2 < Q_1$, or $Q_2 > Q_1$, depending on whether $x_1 < 0.53$ or $x_1 > 0.53$, respectively.

Based on R1 = 1, R2 ranges from infinity ($x_1 = 0.40584$) to zero ($x_1 = 0.77460$). Near $x_1 = 0.44$, R2 = 1, so that all rings have equal diameters at this point. For all $x_1 > 0.44$, the #2 rings are smaller than the #1 rings.

$Q_2/Q_1 < 0$. A second family of solutions exists in the short interval $0.742 < x_1 < 0.758$, in which the change on ring set #2 is opposite to that on set #1. x_2 for this set of solutions decreases from 0.774 to the maximum value of x_1 , 0.758.

With $Q_1 = 1$, Q_2 decreases from 0 at $x_1 = 0.742$ to -1 at $x_1 = 0.758$.

D1 remains within a few percent of the value D1 = 1.1 over this range of x_1 , while D2 decreases from its maximum value of 9.72 ($x_1 = 0.742$), becoming equal to D1 at $x_1 = x_2 = 0.758$.

Assuming R1 = 1, R2 decreases from 7.95 at $x_1 = 0.742$ to unity at $x_1 = 0.758$.

The limiting values of the parameters show that, when $x_1 = x_2 = .758 \ldots$, ring set #2 is superposed on set #1, and with opposite charge. This is the trivial situation in which the electric field everywhere vanishes.

APPENDIX – Details of Four-Coil Solution

The condition that $A_2 = A_4 = A_6 = 0$ in (3) leads to the relations:

$$\frac{Q_1}{B_1^4} P_3(x_1) + \frac{Q_2}{B_2^4} P_3(x_2) = 0 \tag{A1}$$

$$\frac{Q_1}{B_1^6} P_5(x_1) + \frac{Q_2}{B_2^6} P_5(x_2) = 0 \tag{A2}$$

$$\frac{Q_1}{B_1^8} P_7(x_1) + \frac{Q_2}{B_2^8} P_7(x_2) = 0 \tag{A3}$$

where

$$P_3(x) = \frac{1}{2}(3x^3 - 5x)$$

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

$$P_7(x) = \frac{1}{16}(429x^7 - 693x^5 + 315x^3 - 35x) \tag{A4}$$

Introducing the substitutions $Q = \frac{Q_1}{Q_2}$ and $B = \frac{B1}{B2}$, and combining (A1)-(A3):

$$Q = -B^4 \frac{P_3(x_2)}{P_3(x_1)} = -B^6 \frac{P_5(x_2)}{P_5(x_1)} = -B^8 \frac{P_7(x_2)}{P_7(x_1)} .$$

The simplest expressions for Q and B^2 are thus

$$Q = -B^4 \frac{P_3(x_2)}{P_3(x_1)} \tag{A5}$$

$$B^2 = \frac{P_3(x_2)\ P_5(x_1)}{P_3(x_1)\ P_5(x_2)} \tag{A6}$$

Also from Eq. (A4):

$$\frac{P_3(x_1)\,P_7(x_1)}{[P_5(x_1)]^2} = f(x_1) = \frac{P_3(x_2)\,P_7(x_2)}{[P_5(x_2)]^2} \tag{A7}$$

or $$P_3(x_2)\, P_7(x_2) - f(x_1)\, [P_5(x_2)]^2 = 0 \qquad \text{(A8)}$$

(A8) may be solved for as many as four values of x_2 corresponding to each value of x_1. From (A6), the requirement $B^2 > 0$ permits selection of solutions with physical significance. Many solutions were obtained in the interval $x_1 = .40 \ldots$, which is a root of $P_7(x) = 0$ to $x_1 = .77 \ldots$, which is a root of $P_3(x) = 0$. There was found to be one physically significant x_2 for each x_1, corresponding to $Q > 0$, except in the interval $x_1 = .74 \ldots$, another root of $P_7(x) = 0$, through $x_1 = .77 \ldots$, in which an additional solution corresponding to $Q < 0$ was found.

For each set of values (x_1, x_2), the ratio of slant heights B and of charges Q may be found from Eqs (A5) and (A6). From Figure 2 may then be found expressions for the radius R2 of Ring No. 2, and for D1 and D2, the distances of the rings from the origin. All quantities are expressed in terms of R1, which was set equal to unity throughout:

$$\frac{R2}{R1} = \sqrt{\frac{1 - x_2^2}{1 - x_1^2}} \qquad \text{(A9)}$$

$$\frac{D1}{R1} = \frac{x_1}{\sqrt{1 - x_1^2}} \qquad \text{(A10)}$$

$$\frac{D2}{R1} = \frac{x_2}{\sqrt{1 - x_1^2}} \qquad \text{(A11)}$$

Attempts at homogeneity determinations were inconclusive. The method employed was the first of those of Pittman and Waidelich, Ref. 3, and possibly a slightly different expression should be used.

REFERENCES

1. Glaser, W., "Über die zu einem vorgegebenen Magnetfeld gehörende Windungsdichte einer. Kreisspule " Zeitschrift für Physik, volume 118, pages 264 to 268, 1941

2. Smythe, W.R., "Static and Dynamic Electricity," New York: McGraw-Hill, 1950

3. Pittman, M.E., and Waidelich, D.L., "Three and Four Coil Systems for Homogeneous Magnetic Fields," NASA TN D-2095, Jan. 1964

Page intentionally left blank

INTERACTION OF MAGNETIC AND ELECTRIC FIELD MEASUREMENTS

A. V. Dralle

Another question that has arisen in the measurement of electric fields is the effect of an observer traveling through a magnetic field and measuring an electric field. Similarly an observer traveling through an electric field would measure a magnetic field.

The magnetic and electric fields $\vec{B}$ and $\vec{E}$ measured by an observer at rest at a given point in a frame of coordinates 0 differ from those measured at the same point by an observer moving with velocity V through 0. The fields B′ and E′ seen by the moving observer are related to $\vec{B}$ and $\vec{E}$ by the Lorentz transformation relations (MKS units):

$$E'_{||} = E_{||} \qquad E'_{\perp} = \gamma(\vec{E} + \vec{V} \times \vec{B})_{\perp}$$

$$B'_{||} = B_{||} \qquad B'_{\perp} = \gamma(\vec{B} - \frac{\vec{V}}{C^2} \times \vec{E})_{\perp}$$

The subscripts $||$ and $\perp$ denote directions parallel and perpendicular to velocity V, respectively, and

$$\gamma = \left(1 - \frac{V^2}{c^2}\right)^{-1/2}$$

(See references 1, 2, and 3.)

Magnetic (electric) fields will thus be measured by instruments moving through regions in which instruments at rest measure no magnetic (electric) field, only an electric (magnetic) field.

A magnetic field as large as 4 gammas may be induced by motion of an instrument at 40 km/sec through an electric field of 10^4 v/m. Similarly, an electric field of 0.3 v/m is induced by motions of an instrument at 8 km/sec in a magnetic field of 0.5 gauss.

To produce a measurement of 1 gamma by motion of a magnetometer with velocity V through a purely electric field E (as seen by an observer at rest), the condition

$$V \cdot E \cong 10^8$$

must be met. (V is measured in m/sec, E in v/m.) Some experiments of this kind have been done previously,[4] but they must be done quite carefully to eliminate other effects.

REFERENCES

1. Jackson, J.D., "Classical Electrodynamics," John Wiley and Sons, Inc., New York, N.Y. (1962)

2. Fano, R.M., Chu, L.J., and Adler, R.B., "Electromagnetic Fields, Energy, and Forces," John Wiley and Sons, Inc., New York, N.Y. (1960)

3. Rawer, K., "Plasma Effects on Satellites," in "Radio Astronomical and Satellite Studies of the Atmosphere," edited by J. Aarons. North Holland Publishing Company, Amsterdam (1963)

4. Joos, G., "Theoretical Physics," G.E. Stechert and Co., New York (1934) pages 446 and 447