Research Summary No. 36-14

for the period February 1, 1962 to April 1, 1962

jpl
JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

May 1, 1962

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
CONTRACT NO. NAS 7-100

Research Summary No. 36-14

for the period February 1, 1962 to April 1, 1962

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

May 1, 1962

Preface

This is the last issue of the *Research Summary*. Future reports of JPL/NASA supporting research activities will be incorporated in the *Space Programs Summary* series of bimonthly periodicals. Effective July 1, 1962, material that would have appeared in Vols. 1 and 2 of *Research Summary 36-15* will be found in Vols. 4 and 5 of *Space Programs Summary 37-15*. In addition to the supporting research that is normally published under JPL Division and Section titles, this new series will include various advanced development projects that were formerly reported in Vols. 1 and 2 of the *Space Programs Summary*.

W. H. Pickering, *Director*
Jet Propulsion Laboratory

Research Summary No. 36-14

Contents

SPACE SCIENCES DIVISION

SYSTEMS DIVISION

GUIDANCE AND CONTROL DIVISION

TELECOMMUNICATIONS DIVISION

PROPULSION DIVISION

SPACE SCIENCES DIVISION

I. Experimental Space Science

A. Effects of Ultrahigh Vacuum on *Bacillus Subtilis Variety Niger*

F. A. Morelli, F. P. Fehlner, and C. H. Stembridge**

1. Introduction

The effects created by exposing organisms to a vacuum have long been a subject of interest to biologists (Ref. 1). The use of vacuum in storing microorganisms is an essential part of a lyophilization preservation technique (Ref. 2). In this method, the organisms are frozen at −80°C and dehydrated by a vacuum of 10^{-2} to 10^{-3} mm Hg. The tubes containing the organisms are sealed off under vacuum and may be stored for several years.

A study was made by Bakanauskas (Ref 3) to determine the effect of a prolonged dynamic vacuum on the viability of selected spore-forming fungi and bacteria. The microorganisms, *Aspergillus niger, Aspergillus flavus, Bacillus globigii (Bacillus subtilis), Bacillus mycoides,* and *Bacillus cereus,* were exposed to pressures ranging from 1×10^{-5} to 5×10^{-7} mm Hg for periods of 2, 4, 8, 16, and 32 days. With the exception of *B. cereus,* all of the microorganisms remained viable after 32 days exposure.

Despite the above work, the ability of microorganisms to withstand ultrahigh vacuums of the order of 10^{-8} mm Hg or better has been questioned. Brueschke, Suess, and Willard (Ref 4) report that exposure of microorganism spores (*Bacillus subtilis, Aspergillus niger, Aspergillus terreus,* and *Penicillum citrium*) to pressures of 1.2×10^{-8} mm Hg and lower for a sufficient period of time ($10 < t < 31$ days) will cause their destruction. However, recent work by Portner, Spiner, Hoffman and Phillips (Ref 5) failed to show any effects on three types of microorganisms exposed to a vacuum of less than 10^{-9} mm Hg for 5 days. The present work seeks to verify one of these conclusions for the microorganism *B. subtilis.*

2. Test Procedure

The apparatus in Fig 1 was constructed from Pyrex glass and is designed to expose at room temperature eight samples of a microorganism to a vacuum of 10^{-8} mm Hg or better. The samples of *B. subtilis* were prepared, introduced into the vacuum, and recovered as described in the following paragraphs.

A 1-ml sample of a stock culture of an aerobic spore former, *B. subtilis var. niger,* was diluted by adding 99 ml of sterile distilled water. The suspension thus formed was placed in a 60°C water bath for 30 min to kill vegetative organisms. Average control counts of the remaining spores diluted 10^5 and 10^6 were each made in triplicate. Average values of 35×10^6 and 329×10^5

*Chemistry Section, Physical Sciences Division.

organisms/ml were found. Further dilutions were so arranged that inocula containing 300, 3000, 30,000, and 3×10^6 organisms/ml were obtained.

Each of the eight sample tubes attached to the apparatus diagrammed in Fig 1 was then inoculated with 0.1 ml of inoculum as delivered through the entry port by a tuberculin syringe and needle. After sealoff of the entry port, a slight vacuum (20 mm Hg) was pumped and the water in the cultures was allowed to evaporate. The sample tubes were immersed in a pan of warm water to prevent freezing of the microorganisms during this drying process. As soon as the water in the cultures had been removed, an ultrahigh vacuum was applied to the spore samples.

This vacuum was obtained with apparatus diagrammed in Fig 2. Standard vacuum components manufactured from Pyrex glass were utilized. Connections were 40-mm OD tubing. One grease joint and stopcock were included in the system, but outgassing from this source was easily handled by the high-capacity mercury diffusion pump. To insure that mercury was thoroughly trapped, a gold plating (Fig 1) was so placed that molecular backstreaming from the region of the pump and traps was

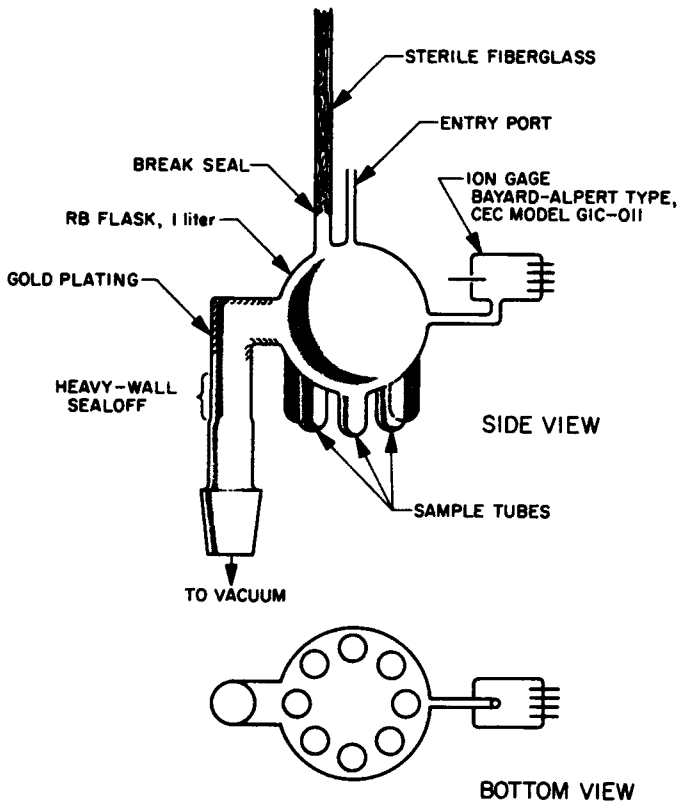


Fig. 1. Apparatus for exposing *B. subtilis* to ultrahigh vacuum

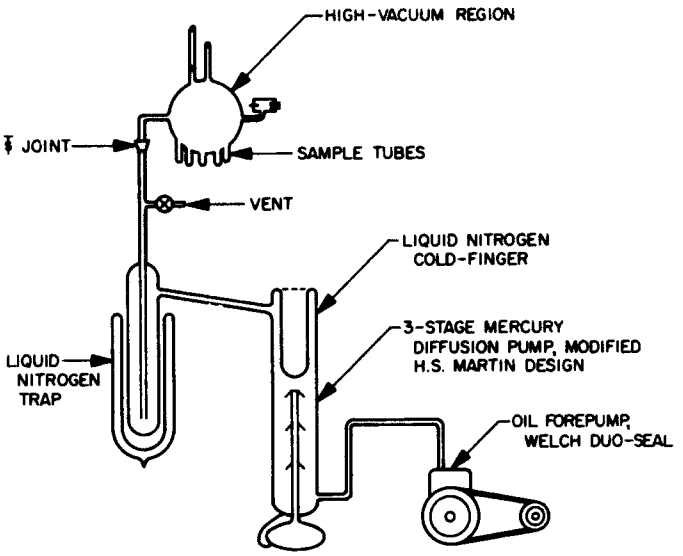


Fig. 2. Ultrahigh vacuum system

intercepted. The gold removed mercury vapor by amalgamation to prevent the vapor from reaching the microorganisms.

In order to minimize outgassing from the glassware in the high-vacuum region (Fig 1), the glass was cleaned in the following manner:

(1) Degrease with acetone

(2) Wash with Na_2CO_3-NaOH solution

(3) Rinse with boiling distilled water

(4) Wash with hot 1*N* nitric acid

(5) Rinse with boiling distilled water

(6) Wash with hot 1.5*N* ammonium hydroxide

(7) Rinse with boiling distilled water

When the apparatus was dry, the gold plating was applied by vacuum deposition and an ion gage was attached. The finished piece was stored in a sealed polyethylene bag until inoculated.

The pressure in the high-vacuum region was monitored by a Bayard-Alpert type hot filament ionization gage. It was located as far away from the pumping system as possible and remained in operation throughout the experiment. After the initial outgassing, accurate readings were obtained. These are recorded in Fig 3 for the 35 days during which the samples of *B. subtilis* were exposed to ultrahigh vacuum. The lowest pressure which can be measured directly with this type of gage is approximately 10^{-9} mm Hg.

The microorganisms were removed from the vacuum system at the end of the 35-day period by sealing off the heavy-wall section shown in Fig 1. The chamber was brought to atmospheric pressure by opening the break seal and allowing air to flow through the 4-in. plug of sterile fiber glass. To each sample tube, 2 ml of sterile distilled water were added through the entry port, and the whole apparatus was placed in a 25-kc ultrasonic transducer containing 3 in. of distilled water. Vibration for 5 min with the entry port closed sufficed to remove the microorganisms from the glass and disperse them into the distilled water. These suspensions were aseptically transferred into sterile Petri dishes. Warm trypticase soy agar was then added, and the mixtures were allowed to solidify into pour plates.

The samples originally containing 3×10^5 organisms were diluted 10^2 before pour plates were made.[a] Then 1 ml of the 10^2 dilution was further diluted to 10^3, and 1-ml aliquots in triplicate were made into pour plates. All pour plates were incubated at 37°C for 5 days before counting was carried out.

Control samples were prepared from 0.1 ml of each of the inocula containing 300, 3000, 30,000, and 3×10^6 organisms/ml. The suspensions were pipeted into test tubes, dried under vacuum in a bell jar, and placed in room-temperature desiccators, each of which contained one of the following atmospheres:

[a]A 1-ml sample of the recovered suspension was diluted with 99 ml of sterile distilled water. After mixing, 1-ml aliquots in triplicate were made into pour plates.

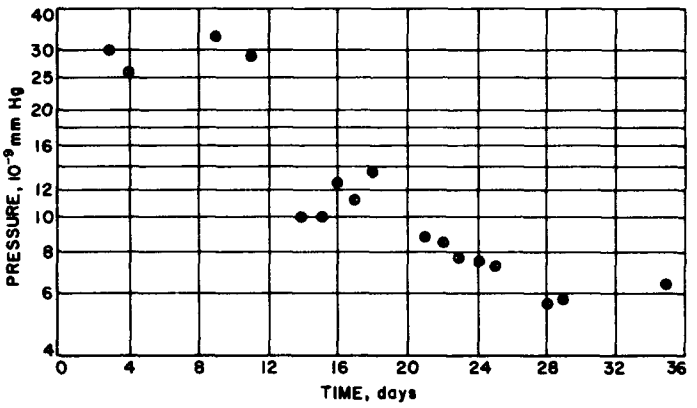


Fig. 3. Ion gage readings during ultrahigh vacuum exposure period

(1) Slight vacuum, 300 μ Hg of air

(2) Nitrogen at 1 atm

(3) Argon at 1 atm

(4) Carbon dioxide at 1 atm

(5) Air at 1 atm

The control samples were recovered and counted in the same manner as the specimens exposed to ultrahigh vacuum.

3. Results

Table 1 lists the bacteriological recoveries from controlled room-temperature atmospheres. In all cases, including ultrahigh vacuum as well as controls, approximately half of the microorganisms were recoverable.

Table 1. Recovery of *B. subtilis*

Inoculum microorganism count	Bacteriological recovery from controlled room-temperature atmospheres					
	Ultrahigh vacuum[a]	Slight vacuum[b]	Nitrogen[c]	Argon[d]	Carbon dioxide[e]	Air[f]
30	22	49	80	144	53	60
30	14	78	82	82	64	48
300	135	158	150	143	126	120
300	135	156	167	109	139	138
3000	150×10^1	500	1410	1500	1140	1080
3000	138×10^1	1700	1530	2500	1090	1910
3×10^{5g}	142×10^{2g}	107×10^{2g}	152×10^{2g}	73×10^{2g}	44×10^{2g}	143×10^{2g}
3×10^{5g}	142×10^{2g}	142×10^{2g}	190×10^{2g}	63×10^{2g}	41×10^{2g}	149×10^{2g}

[a]Ultrahigh vacuum for 35 days; 5 min of 25-kc ultrasonic agitation.
[b]Slight vacuum (300 μ Hg of air) for 36 days; 5 min of 25-kc ultrasonic agitation.
[c]Nitrogen at 1 atm for 38 days; 3 min of 25-kc ultrasonic agitation.
[d]Argon at 1 atm for 38 days; 3 min of 25-kc ultrasonic agitation.
[e]Carbon dioxide at 1 atm for 38 days; 4 min of 25-kc ultrasonic agitation.
[f]Air at 1 atm for 38 days; 4 min of 25-kc ultrasonic agitation.
[g]Average value of 3 counts.

The tabulated counts are within the range expected for a biological experiment. Spurious results were obtained with inoculum containing 30 organisms/ml, for more organisms were recovered than were thought to have been in the original sample. However, these counts are not considered significant, because they reflect the magnified errors[a] encountered in working with low concentrations of microorganisms. Samples containing statistically significant numbers of microorganisms produced consistent results.

4. Conclusions

The results in Table 1 demonstrate that *B. subtilis var. niger* will survive exposure to an ultrahigh vacuum (10^{-8} mm Hg) for a period of 35 days. The experiment also shows that ultrasonic vibrations and radiation encountered during sealoff of the apparatus do not kill the microorganisms.

Since the results of this work completely contradict the conclusions reported in Ref 4, further work on the problem is anticipated. An apparatus is under construction which will produce and directly measure pressures to 10^{-10} mm Hg. Even though a pressure decrease from 10^{-8} to 10^{-10} mm Hg (which lengthens the average mean free path of a gas molecule from 5×10^{5} to 5×10^{7} cm) is not significant in an apparatus having a radius of 12 cm, it will be interesting to expose *B. subtilis* to a vacuum approaching that encountered in outer space. A second experiment will involve exposing an anaerobic spore former and a vegetative organism to a vacuum between 10^{-8} and 10^{-9} mm Hg to determine whether the resistance of these forms to high-vacuum exposure is as great as that of aerobic spores.

[a]Errors are multiplied by manipulations such as transfers, dilutions, recovery by ultrasonic techniques (Ref. 6), and counting procedures. Coagulation, plating, and death of the micoorganisms also decrease the quantitative accuracy of work at low concentrations.

References

1. Chan, E. C. S., and Holz, F. M., *Report of Conference on Factors that Influence Viability and Function in Microorganisms*, DA 18-064-404-cml-344 and DA 18-064-cml-2737, National Research Council, Washington, D. C., November 1958.
2. Flosdorf, E. E., and Mudd, S., "Procedure and Apparatus for Preservation in 'Lyophile' Form of Serum and Other Biological Substances," *Journal of Immunology*, Vol. 29, p. 389, 1935.
3. Bakanauskas, S., *Resistance of Microorganisms to High Vacuum*, WADC TN 59142, Wright-Patterson Air Force Base, Ohio, September 1959.
4. Brueschke, E. E., Suess, R. H., and Willard, M., "Survival of Microorganisms in Ultrahigh Vacuum," *Planetary Space Science*, Vol. 8, p. 30, 1961.
5. Portner, D. M., Spiner, D. R., Hoffman, R. K., and Phillips, C. R., "Effect of Ultrahigh Vacuum on Viability of Microorganisms," *Science*, Vol. 134, No. 3495, p. 2047, Dec. 22, 1961.
6. Newton, N., Painter, R., et al, *Selected Applications of Ultrasonics to Biological Systems*, DA 18-064-cml-2375, Battelle Memorial Institute, Columbus, Ohio, November 1953.

II. Planetary Studies

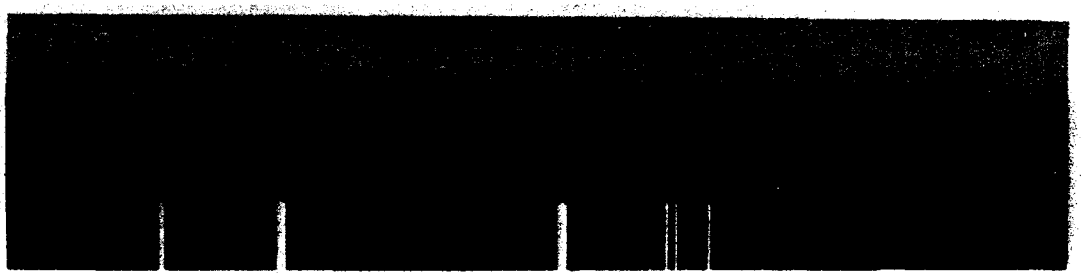
A. New CO_2 Band in Venus Spectrograms

Hyron Spinrad

During a detailed examination of Venus spectrograms in the Mount Wilson Observatory plate files, a band of CO_2 with a head at λ 7158 has been discovered. This band had been observed by Herzberg and Herzberg (Ref 1) as only a very weak head on laboratory spectra taken under conditions of high pressure and a very long path length.

The new Venus band was found on Ce 3028 taken on May 17, 1943. The dispersion was 5.6 A/mm on ammoniated IV- N emulsion. Figure 1 illustrates the λ 7158 CO_2 band on this spectrogram. On this date the λ 7820 CO_2 band was anomalously strong. The new band is marginally visible on a few other Venus spectrograms.

The equivalent width of the λ 7158 CO_2 band, corrected for three weak blends is about 130 mA. The λ 7820 band is approximately eight times stronger on Ce 3028. Since the depth of the head at λ 7158 is only some 4% on a microphotometer tracing, the optical depth in this band is probably quite large.



WAVELENGTH, λ

Fig. 1. Venus spectrogram taken at Mount Wilson on May 17, 1943

Reference

1. Herzberg, G., and Herzberg, L., *Journal of the Optical Society of America*, 43:1037, 1953.

SYSTEMS DIVISION

III. System Analysis

A. Trajectory Studies

W. E. Bollman

1. Low-Energy Solar Trajectories

In interplanetary space exploration, a type of trajectory may be needed wherein the spacecraft reaches a radial distance from the Sun which is as different from the Earth-Sun distance as it is possible to attain with available launch vehicles. If the object of the mission is to minimize the perihelion distance of the spacecraft (pass as close as possible to the Sun), the optimum direction in which the spacecraft would leave the Earth (direction of outgoing geocentric asymptote) is opposite to the direction of the Earth's heliocentric orbital velocity. However, if the object is to maximize aphelion distance, The optimum direction is along the direction of the Earth's heliocentric orbital velocity. Thus the trajectory of the probe with respect to the Earth and Sun would always be one which lies in the ecliptic plane.

a. Outgoing asymptote directed opposite to Earth's orbit motion. Figure 1a shows the probe's perihelion distance and approximate flight time to traverse from launch to perihelion vs C_3 (twice total geocentric energy per unit mass), where the outgoing geocentric asymptote is directed opposite to the Earth's heliocentric velocity for launchings near July 4 and January 3 (aphelion and perihelion times of Earth, respectively). Figure 1b shows the approximate Earth-probe communication distance at the time the probe attains perihelion distance. This distance corresponds to the maximum distance between Earth and probe up to the time of perihelion passage. Note from the two figures that for launching when the Earth is near aphelion (July 4) with a C_3 of 10 km^2/sec^2, the probe can attain a perihelion distance of 96,000,000 km at 140 days past launch at which time the Earth-probe distance will be 100,000,000 km. Launching when the Earth is near perihelion (Jan. 3) with a C_3 of 10 km^2/sec^2, the probe can attain a perihelion distance of approximately 101,000,000 km at 138 days past launch at which time the Earth-probe distance will be 105,000,000 km. As C_3 is increased, the probe's perihelion distance decreases, the flight time to perihelion decreases, but the Earth-probe distance at perihelion passage increases. Notice the variation in the trajectory parameters for January 3 and July 4 launchings (Earth's aphelion and perihelion passages). At first one might be surprised that these variations exist, but remember the Sun-Earth distance varies during a given year by 5,000,000 km, from about 147,000,000 km at Earth's perihelion to 152,000,000 at Earth's aphelion.

b. Outgoing asymptote directed along Earth's orbit motion. Figure 2a shows a plot of the probe's aphelion distance and flight time to aphelion for various C_3's, assuming the outgoing geocentric asymptote is along the Earth's heliocentric velocity, when again launchings occur at Earth's aphelion and perihelion. Figure 2b is a plot of Earth-probe distance at the time of aphelion

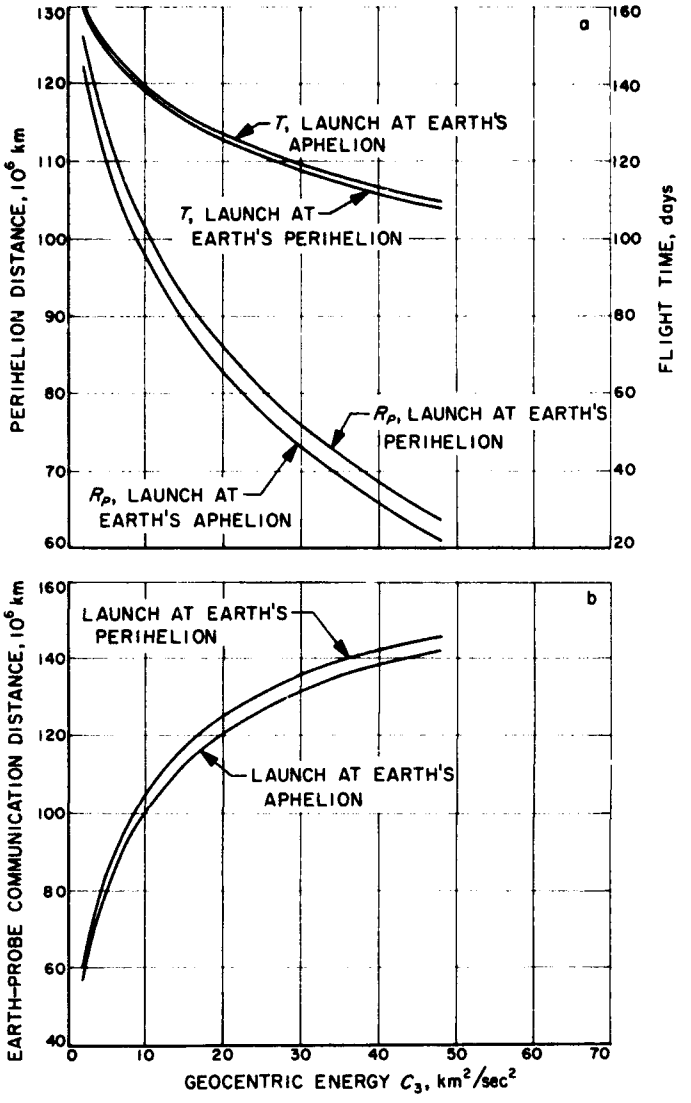


Fig. 1. Spacecraft perihelion distance, flight time, and communication distance vs C_3, where the outgoing asymptote is directed opposite to Earth's orbital velocity

passage of the probe. This distance is the maximum Earth-probe distance up to the time of aphelion passage for C_3's less than 30 km²/sec². Note that for a C_3 of 10 km²/sec², the probe can attain a maximum aphelion distance of 243,000,000 km at about 271 days past launch, at which time the Earth-probe distance will equal 276,000,000 km. As C_3 is decreased, the probe's aphelion distance is reduced, the flight time to aphelion is reduced, and the Earth-probe distance at aphelion passage is reduced for C_3's less than 30 km²/sec². All plots are based on two-body theory.

c. Launch on time. For the mentioned space missions, compensation for delay in launch time must be made

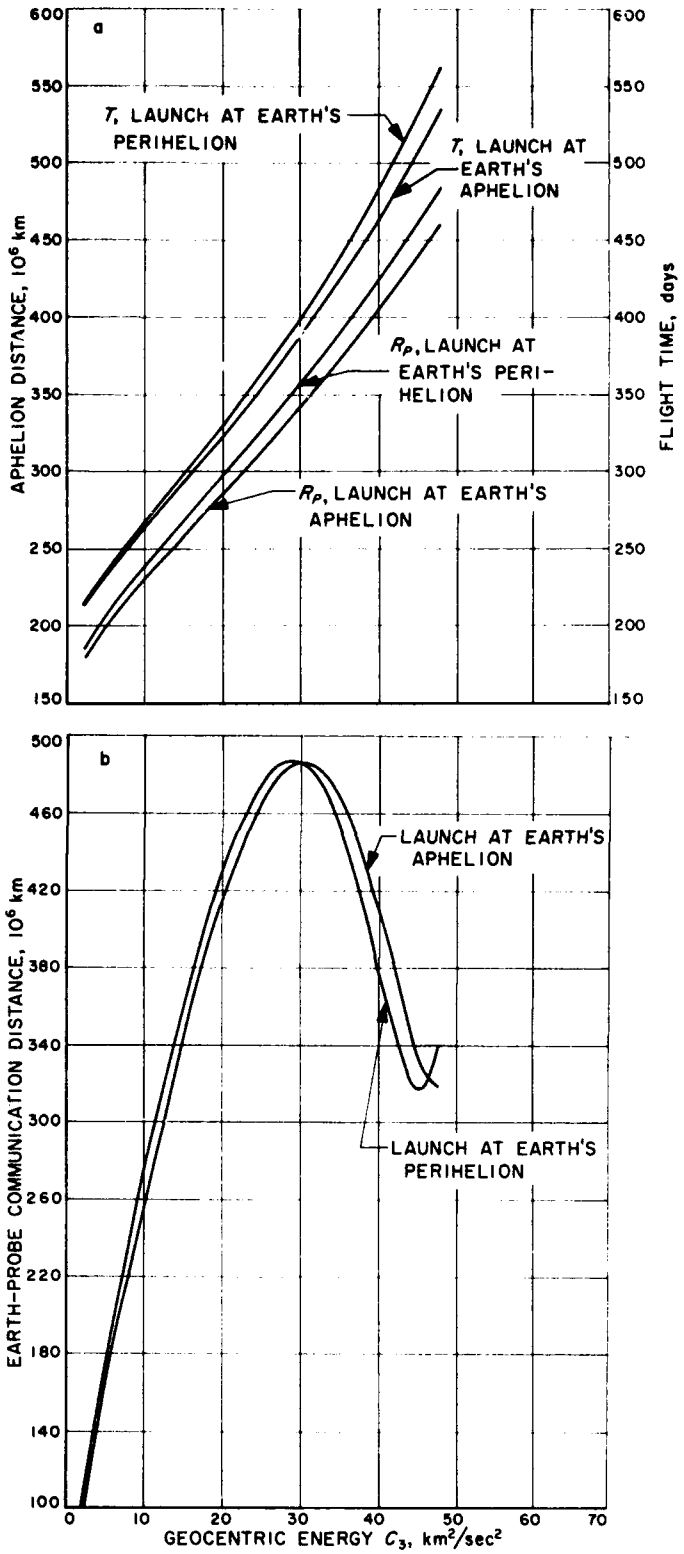


Fig. 2. Spacecraft aphelion distance, flight time, and communication distance vs C_3, where the outgoing asymptote is directed along Earth's orbital velocity

only if a precise trajectory is desired. In other words, if one is only interested in achieving a perihelion or aphelion distance of broad range, the launch azimuth and parking-orbit coast time may be fixed for a given launch day and perhaps throughout the firing period. If required, compensation could be accomplished by varying the launch azimuth and parking-orbit coast time with launch delay. The amount of time available in the firing windows for a given range of launch azimuths will depend on the declination of the outgoing geocentric asymptote. As the asymptotic declination increases algebraically, the firing window increases and the parking-orbit coast time for each launch azimuth increases.

Figure 3 shows a plot of the outgoing asymptotic declination vs launch date for firings (asymptotic direction) both opposite to and along the Earth's heliocentric velocity. For an asymptotic declination of −23.5 degrees, the firing window would be approximately 1 hr for the launch azimuth span of 93 to 111 degrees east of north from AMR, with injection loci in the North Atlantic Ocean. However, for the +23.5-degree asymptotic declination, a firing window of about 3 hr will exist for the 93- to 111-degree launch azimuth span, with the injection loci off the eastern coast of southern Africa. A smaller firing window would result, of course, in utilizing a smaller range of launch azimuths.

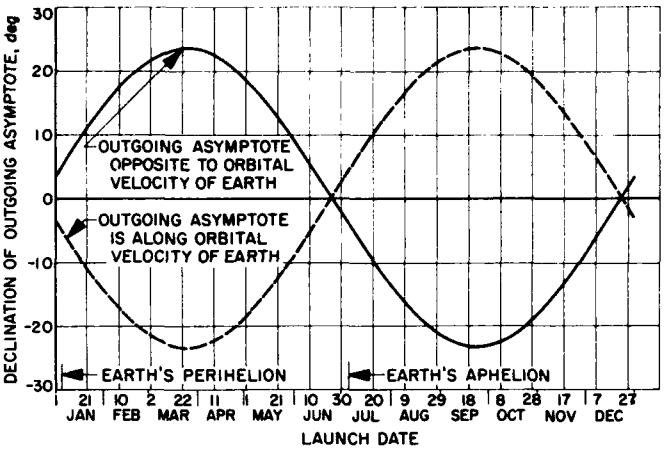


Fig. 3. Declination of outgoing asymptote vs launch date

B. Orbit Determination

C. B. Solloway and W. Kizner

1. Optimal Linear Estimates: Continuous Case,

C. B. Solloway

Any systematic investigation of the orbit determination problem and the concomitant search for a "best" estimate of the orbit parameters from observational data, almost invariably discrete, will naturally lead to a consideration of the analogous continuous problem in the hope of simplifying the study or, at the very least, in attaining absolute lower bounds on the best possible estimate. Much of the theory is well-known, but rarely have exact or explicit solutions been obtained for any given problem. It is worthwhile therefore to exhibit these when and if they can be achieved together with some of the attendant techniques, particularly when the given problem is of practical importance.

Consider the following simplest of this class of problems: let a be an unknown scalar parameter and let $x(t), 0 \leq t \leq T$, be a continuous observation of the form $x(t) = a + n(t)$ where $n(t)$ is the noise in the observation. It is assumed that the noise is a stationary random process over the interval with mean zero and autocorrelation function $R(\tau)$. That is,

$$\widetilde{n(t)} = \int n(t)\, p_1(n)\, dn = E[n(t)] = 0 \qquad 0 \leq t \leq T$$

$$\widetilde{n(t)\, n(t+\tau)} = \int n_1 n_2 p_2(n_1, n_2, \tau)\, dn_1 dn_2$$

$$= E[n(t)\, n(t+\tau)] = R(\tau)$$

where $\widetilde{\quad}$ denotes ensemble average, $p_1(n)$ and $p_2(n_1, n_2, \tau)$ are the first and second probability density functions of the random process and E is the expected value operator. We seek an estimate a^* of a which is linear, unbiased and of minimum variance. Thus we seek a function $h(t)\, 0 \leq t \leq T$ such that

$$a^* = \int_0^T h(\tau)\, x(\tau)\, d\tau \qquad \text{(linear)} \tag{1}$$

with

$$\int_0^T h(\tau)\, d\tau = 1 \qquad \text{(unbiased)} \tag{2}$$

such that

$$\sigma_{a^*}^2 = E[(a^* - a)^2] = \int_0^T \int_0^T h(u)\, h(v)\, R(u - v)\, du\, dv \tag{3}$$

is minimized.

Introducing the Lagrange multiplier -2λ with the constraint Eq (2), we minimize the function

$$\psi(h) = \int_0^T \int_0^T h(u)\, h(v)\, R(u-v)\, du\, dv - 2\lambda \int_0^T h(u)\, du \tag{4}$$

by the usual calculus of variations techniques to obtain the integral equation for $h(t)$

$$\int_0^T h(u)\, R(t-u)\, du = \lambda \qquad 0 \leq t \leq T \tag{5}$$

and we note that with this choice of $h(t)$

$$\sigma_a^2 = \lambda \tag{6}$$

The usual (theoretical) technique at this point is to solve the equations

$$\phi_n(t) = \nu_n \int_0^T R(t-u)\, \phi_n(u)\, du \qquad 0 \leq t \leq T \tag{7}$$

to obtain a complete set of orthonormal functions over the interval $[0, T]$ from which we can obtain the Fourier series for $h(t)$

$$h(t) = \sum_n C_n \phi_n(t)$$

with

$$C_n = \lambda \nu_n \int_0^T \phi_n(\tau)\, d\tau$$

$$\lambda = \sum_n \frac{C_n^2}{\nu_n} \tag{8}$$

or

$$\lambda^{-1} = \sum_n \nu_n \left[\int_0^T \phi_n(\tau)\, d\tau \right]^2$$

The problem is now theoretically solved since we have both $h(t)$ and λ. In practice, however, it is difficult, if at all possible, to obtain the orthonormal set $\{\phi_n(t)\}$, or to "recognize" $h(t)$ from its Fourier series expansion, or to obtain the C_n and ν_n to compute λ for any given autocorrelation function $R(\tau)$.

We consider below two cases of practical importance in which $h(t)$ and $\phi_n(t)$ may be obtained explicitly in closed form. In the interest of brevity we only mention the pertinent mathematical steps, leaving the details to the interested reader.

Case 1. Exponentially correlated noise:

$$R(\tau) = \sigma^2 e^{-\beta|\tau|}$$

The pertinent equations are

$$\int_0^T h(\tau)\, \sigma^2 e^{-\beta|t-\tau|}\, d\tau = \lambda \qquad 0 \leq t \leq T \tag{9}$$

$$\int_0^T h(\tau)\, d\tau = 1 \tag{10}$$

$$a^* = \int_0^T x(\tau)\, h(\tau)\, d\tau \qquad \sigma_{a^*}^2 = \lambda \tag{11}$$

If we extend Eq (9) over the rest of the positive t axis by periodicity and take Laplace transforms, we obtain the equation [equivalent to taking Laplace transforms over a finite range]

$$(-2\beta)\, \bar{h}(S) = \frac{(S^2-\beta^2)}{S} \frac{\lambda}{\sigma^2} (1 - e^{-ST}) - (S+\beta)\, \bar{h}(\beta) + (S-\beta)\, \bar{h}(-\beta)\, e^{-\beta T} e^{-ST} \tag{12}$$

where $\bar{h}(S)$ denotes the Laplace transform of $h(t)$. Now Eq (9) yields the results

$$\frac{\lambda}{\sigma^2} = \bar{h}(\beta) = \bar{h}(-\beta)\, e^{-\beta T} \tag{13}$$

by setting $t = 0, T$, respectively, and from Eqs (10), (12) and (13) we obtain

$$\bar{h}(0) = 1 = \left(\frac{\beta T}{2} + 1 \right) \frac{\lambda}{\sigma^2}$$

or

$$\lambda = \sigma^2 \frac{2}{2+\beta T} = \sigma_{a^*}^2 \tag{14}$$

and Eq (12) becomes

$$\bar{h}(S) = \frac{1}{2+\beta T} \left[\frac{\beta}{S} (1 - e^{-ST}) + (1 + e^{-ST}) \right] \tag{15}$$

from which

$$h(t) = \frac{1}{2+\beta T} \left[\beta + \delta(t) + \delta(T-t) \right] \qquad 0 \leq t \leq T \tag{16}$$

so that

$$a^* = \frac{1}{2+\beta T} \left[x(0) + x(T) + \beta \int_0^T x(\tau)\, d\tau \right] \tag{17}$$

The result is intuitively gratifying. If the data is perfectly correlated ($\beta = 0$), simply average the first and last readings $[x(o) + x(T)]/2$. If the data is perfectly uncorrelated ($\beta = \infty$), simply average all readings; i.e.,

$$\frac{1}{T} \int_0^T x(\tau)\, d\tau$$

For those who prefer to think in terms of correlation times, recall that the correlation time $= 1/\beta$, so that perfectly correlated data implies infinite correlation time while perfectly uncorrelated data implies zero correlation time.

The eigenfunctions $\phi_n(t)$ can be obtained in an analogous manner. Transforming the equation

$$\nu_n \int_0^T \phi_n(\tau)\, \sigma^2\, e^{-\beta|t-\tau|}\, d\tau = \phi_n(t) \tag{18}$$

we obtain

$$\bar{\phi}_n(S) = \sigma^2 \nu_n \frac{(S+\beta)\,\bar{\phi}_n(\beta) - (S-\beta)\,\bar{\phi}_n(-\beta)\, e^{-\beta T} e^{-ST}}{S^2 - \beta^2 + 2\beta\sigma^2 \nu_n} \tag{19}$$

$$= \frac{\phi_n(0)(S+\beta) - \phi_n(T)(S-\beta)\, e^{-ST}}{S^2 - \beta^2 + 2\beta\sigma^2 \nu_n}$$

It follows that in the range $[0, T]$

$$\phi_n(t) = \phi_n(0)\left[1 + \frac{\beta^2}{\omega_n^2}\right]^{1/2} \cos(\omega_n t - \psi_n) \tag{20}$$

where

$$\omega_n^2 = 2\beta\sigma^2 \nu_n - \beta^2$$
$$\psi_n = \tan^{-1}\frac{\beta}{\omega_n} \tag{21}$$

and imposing the constraint required in Eq (19) yields

$$\tan(\omega_n T - \psi_n) = \tan\psi_n \tag{22}$$

or

$$\tan\frac{\omega_n T}{2} = \frac{\beta}{\omega_n}; \quad \omega_n = \frac{2\psi_n}{T} \tag{23}$$

The solution to Eq (23) determines the (countable) set of eigenvalues ν_n and phase angles ψ_n (Eqs 21), hence the eigenfunctions $\phi_n(t)$. It is easily verified that they are orthogonal. The constants $\phi_n(0)$ are determined from the normality requirement, and there results

$$\phi_n(t) = \left(\frac{T}{2} + \frac{1}{2\sigma\nu_n}\right)^{-1/2} \cos(\omega_n t - \psi_n) \tag{24}$$

Of course, knowing the form of $\phi_n(t)$, it is probably as easy to determine the constants ω_n and ψ_n directly from Eq (18) without recourse to transforms. The main advantage of the transform was to establish this form.

Case 2. Exponential and white noise:

$$R(\tau) = \sigma_1^2 e^{-\beta|\tau|} + \sigma_2^2 \delta(\tau)$$

The case of noise consisting of an exponentially correlated noise source plus an additive white noise is of practical interest. The integral equation for $h(t)$ now becomes

$$\int_0^T h(\tau)\, \sigma_1^2\, e^{-\beta|t-\tau|}\, d\tau + \sigma_2^2 h(t) = \lambda \tag{25}$$

Taking transforms as before yields the equation

$$\bar{h}(S)\left[\sigma_2^2(S^2 - \beta^2) - 2\beta\sigma_1^2\right] = \frac{\lambda(S^2 - \beta^2)(1 - e^{-ST})}{S}$$
$$- \sigma_1^2 \bar{h}(\beta)(S+\beta) + \sigma_1^2 \bar{h}(-\beta)\, e^{-\beta T}(S-\beta)\, e^{-ST} \tag{26}$$

so that $h(t)$ must have the form

$$h(t) = A + Be^{\omega t} + Ce^{-\omega t} \tag{27}$$

with

$$\omega^2 = \beta^2 + 2\beta\frac{\sigma_1^2}{\sigma_2^2} \tag{28}$$

After suitable manipulations, we find that A, B and C will satisfy the equations

$$\frac{A}{\beta} + \frac{B}{\omega+\beta} + \frac{C}{\beta-\omega} = 0$$
$$\frac{A}{\beta} + \frac{e^{\omega T}}{\beta-\omega} B + \frac{e^{-\omega T}}{\omega+\beta} C = 0 \tag{29}$$
$$AT + \frac{e^{\omega T} - 1}{\omega} B + \frac{1 - e^{-\omega T}}{\omega} \smile \quad 1$$

from which we can deduce that $B = e^{-\omega T} C$, and $h(t)$ actually becomes a translated catenary in the range $[0,T]$

$$h(t) = A + B\left(e^{\omega t} + e^{\omega(T-t)}\right) \tag{30}$$

and further that

$$\lambda = \frac{A\sigma_2^2\omega^2}{\beta^2} = \frac{A}{\beta^2}(\sigma_2^2\beta^2 + 2\beta\sigma_1^2) \tag{31}$$

It is easily verified that as $\sigma_2 \to 0$ the solution tends to that of Case 1.

The eigenfunctions for this kernel can be obtained in a manner analogous to Case 1. They are the eigenfunctions of Case 1 if we replace $\sigma^2\nu_n$ by $(\sigma_1^2\nu_n)(1 - \sigma_2^2\nu_n)^{-1}$. The details are left to the interested reader.

It is interesting to note that the function $h(t)$ in both cases is symmetric about the point $t = T/2$, and indeed a moment's reflection will suffice to understand why this must be so.

2. Selection of Parameters To Be Estimated in Orbit Determination, W. Kizner

In orbit determination studies, as in the general problem of estimation of functions of unknown parameters, there is always the problem of choosing which parameters to estimate. For instance, should one try to estimate biases and physical constants as well as the standard dynamical parameters for the problem of orbit determination? A solution is given here which can be used to test whether or not the calculated changes in any given

variable or set of variables are statistically significant. This is accomplished by means of the so-called F test. The test assumes a particularly simple form in the case where many independent observations have been made and only one parameter is to be tested. Here the test involves the familiar normal distribution.

Suppose there are n parameters $x_1, \cdots x_n$, and x_n is the suspect parameter. Let λ represent the covariance matrix of x, which is assumed known if many independent observations have been made. Then a new set of independent variables $\hat{y}$ can be obtained and normalized by an application of the Schmidt process. These can be normalized to form the set y. Each y_i depends only on the preceding x up to x_i, or

$$y_i = y_i(x_1 \cdots x_i)$$

Since the y are orthogonal, y_n contains only the difference between x_n and its projection on the preceding variables. The given set of data is first fitted with the first $n-1$ parameters, and then with the full set of n parameters. The difference of the parameters is then tested, and y_n is evaluated. Since y_n is a unit normal variable of zero mean, the usual confidence intervals can be used to test the hypothesis that y_n is zero. For instance, if $|y_n| > 1.96$ and the true value of the variable were zero, only 5% of the cases would fall outside of this confidence interval, $-1.96 \leq y_n \leq 1.96$. It should be noted that this test does not attempt to say whether the change in the nth parameter is significantly large, as the usual method of significance test does. Even if the change in the nth parameter is significantly large it may still not be estimable if it is closely correlated with the other parameters. Therefore, the test chosen is made on an uncorrelated variable.

The equations for obtaining the y are given

$$y_1 = x_1/[\mathrm{Var}(x_1)]^{1/2}$$

$$y_2 = (x_2 - a_{21}\hat{y}_1)/[\mathrm{Var}(x_2 - a_{21}\hat{y}_1)]^{1/2}$$

$$y_3 = (x_3 - a_{32}\hat{y}_2 - a_{31}\hat{y}_1)/[\mathrm{Var}(\hat{y}_3)]^{1/2}$$

$$y_n = (x_n - a_{n,n-1}\hat{y}_{n-1} - \cdots - a_{n1}\hat{y}_1)/[\mathrm{Var}(\hat{y}_n)]^{1/2}$$

where Var stands for variance.

The a are obtained from the auxiliary quantities b:

$$b_{i1} = \lambda_{i1} \qquad a_{i1} = \frac{b_{i1}}{\mathrm{Var}(\hat{y}_1)}$$

$$b_{i2} = \lambda_{i2} - a_{21}b_{i1} \qquad a_{i2} = \frac{b_{i2}}{\mathrm{Var}(\hat{y}_2)}$$

$$\cdots\cdots\cdots\cdots$$

$$b_{ij} = \lambda_{ij} - \sum_{t=i-2}^{1} a_{jt}b_{it} \qquad a_{ij} = \frac{b_{ij}}{\mathrm{Var}(\hat{y}_j)}$$

where

$$j \leq i-1$$

and

$$\mathrm{Var}(\hat{y}_i) = \lambda_{11} - \sum_{j=1}^{i-1} a_{ij}b_{ij}$$

These relations have been known (Ref 6) but not used, partly because the new set of variables must be obtained recursively from the old set. To remedy this, the following transformation has been obtained which will directly transform the original set into the new set.

Let

$$y_i = C_i x_i + C_{i-1}x_{i-1} + \cdots C_1 x_1$$

then

$$C_i = \frac{1}{[\mathrm{Var}(\hat{y}_i)]^{1/2}}$$

$$C_{i-1} = -C_i a_{i,i-1}$$

$$C_{i-2} = -(C_i a_{i,i-2} + C_{i-1}a_{i-1,i-2})$$

$$\cdots\cdots\cdots$$

$$C_1 = -(C_i a_{i,1} + C_{i-1}a_{i-1,1} + \cdots + C_2 a_{2,1})$$

Thus the C can be calculated once and for all from the a, which are determined by the covariance matrix of the parameters.

Of course the significance level of the test must be determined by the nature of the resulting Type I and Type II errors. The penalty resulting from the Type I error stems from the fact that an additional parameter is estimated which is essentially zero. This leads to an increase in the variance of the estimate of the parameters. The penalty resulting from the Type II error is that a significant parameter is not estimated and the resulting estimate for the remaining parameters is biased.

The y_i used in the test may not have unit variance, since it is based on the difference of parameters of two approximations. However, in order to insure that y_i is of unit variance one would have to examine the method of convergence for the estimation of the parameters. Instead the test given here, which is based on an *a priori* probability, is a simple yardstick which can be used to indicate whether or not a variable should be included.

C. Space Guidance Theory

C. E. Kohlhase

1. Probability of Planetary Capture

a. Introduction. It is important that the exploration of other planets shall not result in their contamination by the transport of Earth organisms, at least until such time as it could be established that contamination of certain planets would be allowable. During the period of unmanned exploration, every effort must be made to prevent contamination of the planets under investigation so that (1) there would be no chance of upsetting any existing ecology, and (2) our own life experiments and related investigations could proceed without the danger of erroneously rediscovering Earth organism types.

We must either produce a sterile spacecraft or we must insure a low probability of planetary capture (impact) for an unsterile spacecraft. Currently, the most feasible method for achieving sterility requires heating the entire spacecraft (with all subsystems present) to 135°C for 24 hr. It appears, however, that component reliability may be severely jeopardized by the application of any rigid heat sterilization program; component type approval may require as high as three cycles through 145°C for 36 hr. Should some spacecraft components exhibit a significantly reduced reliability because of heat treatment, they would have to be subjected to other, perhaps less effective, sterilization procedures.

Since there would always be some reasonable doubt that the spacecraft was truly sterile, the trajectory aiming point would have to be selected at an adequate distance from the destination planet to effect a low probability of impact. It is currently hoped that capture probabilities of 0.001 to 0.0001 can be realized for Venus and Mars, respectively. The Venus requirement is less stringent as the present estimate of Venusian surface temperatures is in the neighborhood of 300°C, and it is believed that Earth organisms would not survive in such an environment.

The probability of planetary capture depends upon a large number of parameters, most of which can only be estimated. Therefore, the reader will appreciate the limitations of the numerical results of this study. If the injection and midcourse guidance systems should function as currently designed, then the results listed in the accompanying tables are representative of capture probabilities.

b. Analysis. Let us begin by assuming that geocentric injection has "occurred." Injection will be considered to have occurred if there was sufficient final acceleration to insure escape velocity. Injection may have occurred either "properly" or "improperly." If injection occurred properly, the injection guidance system has performed its task within the designed or specified bounds of expected error; whereas, if injected improperly, the vehicle has achieved escape velocity but, due to some component malfunction, has failed to maintain the injection errors within the design limits. Midcourse corrective maneuvers can also be performed properly or improperly, or they may fail altogether. A midcourse maneuver "failure" means that no velocity increment has been imparted to the spacecraft. It can be shown that the probability of capture P_c (see nomenclature Table 1) is given by the expression

$$\begin{aligned} P_c = {} & p_1 p_2 \{A [a_3 p_3 X_{\bar{3}} + (1 - a_3) p_3 X_3 + (1 - p_3) X_{\bar{2}}] \\ & + [1 - A] [a_3 p_3 X_3 + (1 - a_3) p_3 X_3 + (1 - p_3) X_2]\} \\ & + p_1 [1 - p_2] [a_1 X_1 + (1 - a_1) X_{\bar{1}}] \\ & + [1 - p_1] [a_0 X_0 + (1 - a_0) X_{\bar{0}}] \end{aligned} \quad (1)$$

where $A = a_1 + a_2 - 2a_1a_2$, and

P_i = the probability that the ith midcourse maneuver will be performed in some fashion (properly or improperly), given that the (i-1)st maneuver was performed in some fashion.

X_i = the probability of planetary capture (impact) after the proper execution of the ith maneuver, geocentric injection being the 0th maneuver.

$X_{\bar{i}}$ = the probability of planetary capture after the improper execution of the ith maneuver.

a_i = the probability that the ith maneuver will be performed either properly or improperly, given that the ith maneuver has been performed and, further, that the $(i - 1)$st maneuver was performed either properly or improperly, respectively.

The definition of a_i means that, if a given maneuver is performed at all, then it is just as likely that the given maneuver will be performed properly if the previous maneuver was performed properly, as it is that the given maneuver will be performed improperly if the previous maneuver was performed improperly. The following equations were used in computing the X_i:

$$X_0 = \alpha B_c^2 \exp [-2 (\alpha R_0)^2 (B_{11} \sin^2 \theta - B_{12} \sin 2\theta + B_{22} \cos^2 \theta)]$$

where

$$\alpha = [4 (B_{11}B_{22} - B_{12}^2)]^{-1/4}$$

$$B_c = [r_p (r_p + 2\mu/V_\infty^2)]^{1/2}$$

Table 1. Nomenclature

a_i	= probability that ith maneuver will be performed either properly or improperly, given that $(i-1)$st maneuver has been performed either properly or improperly, respectively
B_c	= capture radius of destination planet
p_i	= probability that ith midcourse maneuver will be performed, given that $(i-1)$st maneuver was performed
P_c	= overall mission probability of spacecraft capture, defined by Eq (1)
r_p	= actual radius of destination planet
R_i	= nominal aiming distance of ith maneuver from center of destination planet
V_∞	= hyperbolic excess velocity relative to target planet
X_i	= probability of planetary capture after proper execution of ith maneuver
X_i^-	= probability of planetary capture after improper execution of ith maneuver
μ	= product of universal gravitational constant and mass of target planet
θ	= angle measured counterclockwise from B_1 to radial direction in question
σ_i	= radius of 1σ dispersion circle; assumed equal to $\sqrt{2}$ times the 1σ dispersion of the semi-major axis of the actual uncertainty ellipse remaining after the ith maneuver

All of the above quantities are defined in Table 1 except for B_{11}, B_{12}, and B_{22}. These are statistical (Ref 7) parameters which describe the size and orientation of the dispersion ellipse at the target due to injection errors. Given that $\mathbf{B}_1$ and $\mathbf{B}_2$ define a plane normal to the asymptote of the approach hyperbola relative to the destination planet and that $\mathbf{B}_2$ is parallel to the ecliptic plane, then

$$B_{11} = E(B_1^2)$$

$$B_{12} = E(B_1 B_2)$$

$$B_{22} = E(B_2^2)$$

where E indicates the expected or average value. Figure 4 illustrates the dispersion ellipse for a typical 1964 Mars trajectory due to the "pessimistic" estimates of the *Centaur* injection guidance system component errors.

If it is assumed that the dispersion at the target planet that remains after the execution of the ith midcourse correction is circularly distributed (a conservative assumption) with an rms (1σ) value denoted by σ_i, then it can be shown that, if the aiming point is sufficiently removed from the planet such that the dispersion circle and capture area are reasonably separated, then the probability of planetary capture is given approximately by

$$X_i \simeq B_c (\pi R_i)^{-1} \left\{ \exp\left[-\left(\frac{R_i - B_c}{\sigma_i} \right)^2 \right] - \exp\left[-\left(\frac{R_i + B_c}{\sigma_i} \right)^2 \right] \right\} \tag{3}$$

where R_i is the nominal aiming distance of the ith maneuver from the center of the planet. If the aiming point is chosen directly at the planet's center, then the probability is given exactly by

$$X_i = 1 - \exp\left[-(B_c/\sigma_i)^2 \right] \tag{4}$$

c. Results. The numerical results given in Tables 2 and 3 are based upon the *Mariner B* mission to Mars in 1964. Table 2 presents the overall probability of capture P_c as a function of (1) the "realistic" and "pessimistic" *Centaur* injection errors, (2) $\sigma_1 = 12{,}000$ km and $\sigma_2 = 4{,}000$ km, (3) various aiming point selections with and without biasing, and (4) a range of midcourse

Table 2. Capture probabilities for 1964 Mars flyby spacecraft, $\sigma_1 = 12{,}000$ km and $\sigma_2 = 4{,}000$ km

Aiming scheme	Confidence Level I (very pessimistic)	Confidence Level II (pessimistic)	Confidence Level III (realistic)	Confidence Level IV (optimistic)
20-20-20	8.1 (−3)[a]	3.7 (−3)	1.9 (−3)	2.5 (−4)
$\theta = 120°$	6.6 (−3)[b]	3.0 (−3)	1.6 (−3)	1.7 (−4)
25-25-25	2.6 (−3)	1.3 (−3)	6.5 (−4)	9.4 (−5)
$\theta = 120°$	1.9 (−3)	8.8 (−4)	4.6 (−4)	5.7 (−5)
50-20-20	6.3 (−3)	2.8 (−3)	1.5 (−3)	1.6 (−4)
$\theta = 120°$	6.5 (−3)	2.9 (−3)	1.5 (−3)	1.7 (−4)
50-30-20	2.9 (−4)	1.3 (−4)	7.0 (−5)	7.8 (−6)
$\theta = 120°$	5.2 (−4)	2.5 (−4)	1.3 (−4)	1.9 (−5)
50-30-20	—	—	1.1 (−3)	2.1 (−4)
$\theta = 106°$	—	—	1.7 (−4)	2.8 (−5)
100-30-20	2.7 (−4)	1.2 (−4)	6.5 (−5)	6.8 (−6)
$\theta = 120°$	3.2 (−4)	1.5 (−4)	7.7 (−5)	9.3 (−6)
100-30-20	—	—	9.4 (−4)	—
$\theta = 106°$	—	—	1.7 (−4)	—

[a] Assuming "realistic" injection errors.
[b] Assuming "pessimistic" injection errors.

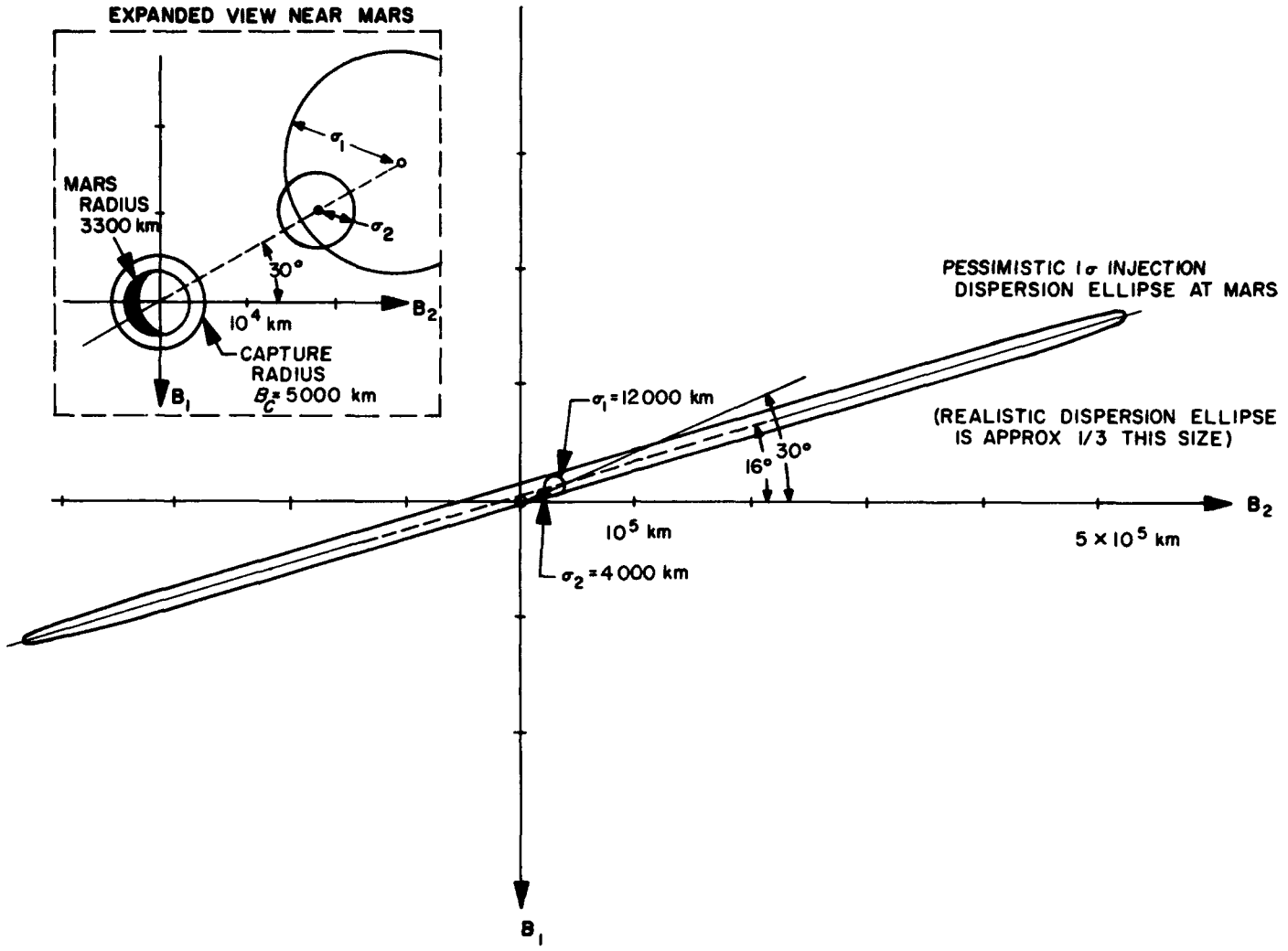


Fig. 4. Orientation and size of dispersion areas for 1964 Mars flyby mission, assuming 235-day trajectory

maneuver confidence levels (Table 4). Due to certain near-planet constraints for the 1964 Mars flyby trajectory, the radial direction of the nominal aiming point (s) should lie about 30 degrees above the B_2 direction. All of the aiming points were therefore selected along this line. This may be seen in Figure 4, which illustrates the particular biasing situation "50-30-20" and pessimistic 1σ midcourse dispersion circles (includes uncertainty in knowledge of the orbit and future solar disturbances). The 50-30-20 simply means that the injection guidance system aims 50,000 km from the center of Mars, the first midcourse maneuver corrects for injection dis-

Table 3. Effect of X_i^- upon P_c for 50-30-20 ($\theta = 120°$) situation

Injection errors	Confidence level	$X_i^- = 0$*	$X_T^- = 0.0001$	$X_T^- = 0.001$	$X_T^- = 0.01$
Realistic	Level III	6.3 (−5)	7.9 (−5)	2.3 (−4)	1.7 (−3)
	Level IV	7.7 (−6)	9.6 (−6)	2.7 (−5)	2.0 (−4)
Pessimistic	Level III	1.1 (−4)	1.3 (−4)	2.8 (−4)	1.8 (−3)
	Level IV	1.9 (−5)	2.1 (−5)	3.8 (−5)	2.1 (−4)

*For $i > 0$; X_3^- was always assumed to be 10^{-4}.

Table 4. Assumed midcourse maneuver confidence levels

Confidence level	α_0	α_1	α_2	α_3	p_1	p_2	p_3*
Level I, very pessimistic	0.5	0.5	0.5	—	0.8	0.5	—
Level II, pessimistic	0.8	0.8	0.8	—	0.9	0.8	—
Level III, *realistic*	0.9	0.9	0.9	—	0.95	0.9	—
Level IV, optimistic	0.95	0.99	0.99	—	0.99	0.99	—

*The analysis showed that P_c was essentially independent of p_3; this happened because the performance of a second midcourse maneuver so lowered the probability of capture (for the relatively large aiming distances involved) that a third maneuver contributed nothing further. It would be reasonable to assume $X_3 \simeq X_2$ and $X_3^- > X_3$, such that the addition of a third maneuver would actually increase P_c.

persions and also moves the aiming point closer to 30,000 km, and the second midcourse maneuver corrects for execution errors created by the first maneuver and also moves the aiming point in to 20,000 km.

The numerical results of Table 2 are based upon $X_{\bar{i}} = X_i$. There are possibly only two types of "improperly" executed maneuvers which might cause $X_{\bar{i}} > X_i$. These are (1) midcourse motor burning to fuel depletion when trying to maneuver closer to the planet, and (2) an explosion of the propulsion system which could produce a large number of fragment trajectories similar to a shotgun pattern. Possibility (1) could be handled by keeping track of the amount of unused propellant, noting the direction and magnitude of the planned maneuver. Possibility (2) has an extremely low chance of occurring, as no midcourse propulsion system has ever exploded either during testing or at any other time. Table 3 illustrates the effect of $X_{\bar{i}}$ upon P_c for the particular biasing situation 50-30-20.

d. Conclusions. Trajectory biasing will have to be used in order to (1) maintain the probability of spacecraft impact to one part in 10,000, (2) perform a capsule mission to Mars in 1964, and (3) pass close enough ($<$30,000 km) to the planet to obtain particularly good scientific measurements. The particular biasing scheme 50-30-20 currently appears the most attractive. The amount of additional midcourse fuel required to perform these aiming point displacements does not exceed 5 meters/sec.

References

1. Grenander, U., and Szegö, G., *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, 1958.
2. Petersen, E. L., *Statistical Analysis and Optimization of Systems*, John Wiley and Sons, New York, 1961.
3. Shinbrot, M., "On the Integral Equation Occurring in Optimization Theory with Nonstationary Inputs," *Journal of Math and Physics*, July 1957.
4. Thomas, J. B., and Wong, E., "On the Statistical Theory of Optimum Demodulation," *PGIT, I.R.E.*, Vol. IT-6, No. 4, September 1960.
5. Wiener, N., *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, John Wiley and Sons, New York, 1949.
6. Rao, C. R., *Advanced Statistical Methods in Biometric Research*, John Wiley and Sons, New York, 1952.
7. Noton, A. R. M., Cutting, E., and Barnes, F. L., *Analysis of Radio-Command Midcourse Guidance*, Technical Report 32-28, Jet Propulsion Laboratory, Pasadena, Sept. 8, 1960.

GUIDANCE AND CONTROL DIVISION

IV. Guidance and Control Research

A. Thin Permalloy Films

F. B. Humphrey

Very thin films of permalloy, evaporated onto a glass substrate, exhibit a uniaxial, magnetic anisotropy within the plane of the film (*RS 36-8*). The reason for the anisotropy is not known, although it is possible to control the direction of the anisotrophy and, to a lesser extent, its magnitude (*RS 36-12*). The anisotropy is important in that all thin-film devices that have either been proposed for or actually used in a digital computer depend upon the existence of the anisotropy for their operation. A detailed study of the anisotropy, using a torsion magnetometer (*RS 36-11*), has been initiated in the hope that such studies will extend our understanding into the basic mechanisms involved. Examples of rotational hysteresis loops (*RS 36-13*) have been discussed, as well as other magnetic characteristics that can be investigated by observing the torque perpendicular to the film plane created by a field rotated in this plane. Now the discussion will be extended by considering the torque parallel to the plane of the film that can be observed when the film is in the presence of an applied magnetic field at some angle to the film plane.

For this case, experimentally, the film is held by a quartz bracket which is then hung on the quartz fork as is shown schematically in Fig 1. With the torsion axis parallel to the film plane and perpendicular to the

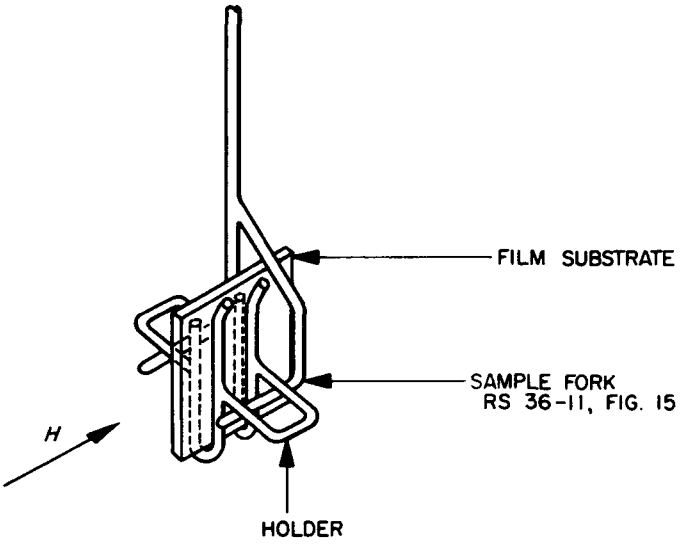


Fig. 1. Sample holder showing provision to hold sample perpendicular to field

easy axis, the energy of interaction with an external field is given by

$$E = -MH\cos(\alpha - \theta) + \tfrac{1}{2}NM^2\sin^2\theta \qquad (1)$$

where α is the angle between the plane of the substrate and the applied field H, θ is the angle the magnetization makes with the substrate, and N is the demagnetizing factor which for this case is 4π. The second term is the magnetostatic energy of demagnetization (Ref 1). At equilibrium,

$$\frac{dE}{d\theta} = 0 = -MH\sin(\alpha - \theta) + NM^2\sin\theta\cos\theta \qquad (2)$$

$$H\sin(\alpha - \theta) = NM\sin\theta\cos\theta \qquad (3)$$

The first concern should be as to the magnitude of θ, i.e., as to whether the magnetization remains in the plane of the film. The most extreme case will be when $\alpha = 90$ deg. At that time Eq (3) becomes

$$H\sin(90 - \theta) = H\cos\theta = NM\sin\theta\cos\theta \qquad (4)$$

Therefore,

$$\sin\theta = \frac{H}{NM} \qquad (5)$$

Since $4\pi M$ is typically 10^4 for permalloy films, it can be assumed that the magnetization remains in the plane of the film for fields of less than, say, 100 oe.

For the geometry presented schematically in Fig 1, the torque per unit volume is

$$\mathbf{L}/V = \mathbf{M} \times \mathbf{H} = -MH\sin\alpha \qquad (6)$$

where the negative sign indicates that the tendency is to decrease α. Clearly then

$$M = \frac{L_\alpha}{HV\sin\alpha} \qquad (7)$$

as long as $H << 4\pi M$. If H is large compared to any anisotropies, then $M = M_s$, the saturation magnetization. Generally two measurements are made: one with $\alpha = 0$ to indicated alignment, and one at some small angle, say 10 deg.

Clearly, from Eq (6), if the field is reversed so as to oppose M, then the torque will be positive. As the magnitude of such a field is increased, the torque will continue to be positive until the coercive force is reached. The magnetization will then reverse and the torque will be negative. The value of field H_c', at which the torque is zero, is related to the true coercive force of the material by

$$H_c = H_c'\cos\alpha \qquad (8)$$

By using a resistor in series with the field producing coils, the X-Y recorder can be connected to plot H vs L directly. An example of such a measurement is shown in Fig 2 for the same film as was used in *RS 36-13*, Fig 5. This film was made by vacuum evaporation from a melt with a composition of 80% nickel and 20% iron. It is a 1-cm-D film with a composition of 76.2% Ni, 23.8% Fe, and a thickness of 35 Å as measured by X-ray fluorescence techniques. As can be seen, $H_c' = 1.69$ so that $H_c = 1.58$. Accuracy to 1% is easily obtained by expanding the scale on the chart recorder. It should be noticed that this film is only 35 Å thick, yet large deflections are obtained. This method provides a convenient way of measuring the coercive force that is considerably more sensitive than conventional hysteresis loop tracers.

It was previously shown (*RS 36-13*, Eq 1) that

$$-L = KV\sin 2\phi \qquad (9)$$

where ϕ is the angle between the easy axis and the magnetization with the field in the plane of the film, and L is perpendicular to the plane. If the field is large enough, say greater than H_K, then the magnetization will follow around as the applied field is rotated. At some point, ϕ will be 45 deg and the torque will be maximum. Then:

$$-L_{max} = T = KV \qquad (10)$$

It is customary to plot T vs $1/H$ and extrapolate to $H = \infty$ (Ref 2). Such a procedure seems to be unnecessary for films observed by us since T_∞ can be determined at fairly low fields (Ref 3). In either event, Eq (7) and Eq (10) lead to an important quantity:

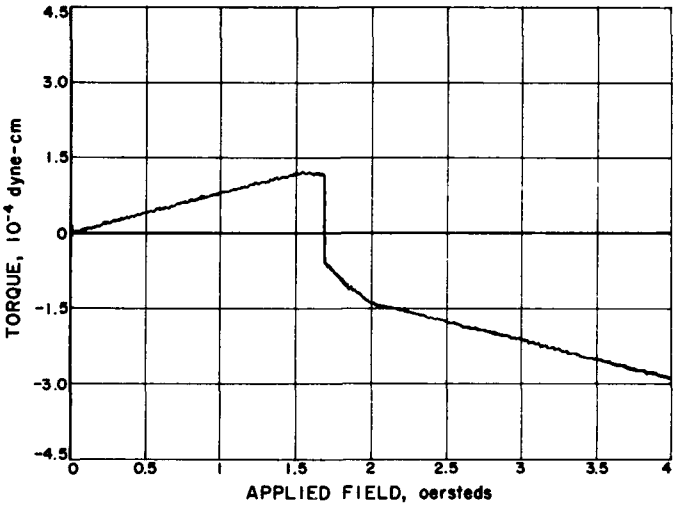


Fig. 2. Torque vs applied field parallel to easy axis and 20 deg to film surface

$$\frac{2K}{M_s} = \frac{2T_\infty}{L_\alpha} H \sin\alpha \tag{11}$$

It is convenient to call this quantity H_K as has been done by many (Refs 3, 4), especially when considering thin films. However, this method uses high fields to determine $2K/M_s$ so that it is not surprising that a different value of H_K is obtained by extrapolating the low-drive hard direction hysteresis loop (Ref 5). For this film, the hysteresis loop yields $H_K = 3.0$, whereas $2K/M_s = 2.65$ oe.

There is yet another way to measure H_K with the torque meter that more nearly corresponds to the loop tracer measurement. Consider the film hanging as before, only with the easy axis now perpendicular to the applied field. For $H\cos\alpha < H_K$ the torque along the easy (Z) axis is

$$\mathbf{L}/V = \mathbf{M} \times \mathbf{H}$$

$$-L_z = MV\frac{H^2}{H_k}\cos\alpha\sin\alpha \tag{12}$$

and for $(H\cos\alpha) > H_K$ the torque is given by Eq (6). An example is shown in Fig 3. It can be seen that the extrapolated parabola (for small H) intersects the extrapolated straight line (for large H) at a value of field very near that given for H_K by the hysteresis loop. The accuracy is clearly lacking for this method as is the case for the hysteresis loop.

A quantity that is of particular interest to computer engineers is the squareness, where

$$Sq = \frac{B_r}{B_s} \tag{13}$$

that is, the ratio of the remanent to saturation flux. It has been shown (*RS 36-13*, Eq 1) that

$$-L = HMV\sin(\alpha - \phi) \tag{14}$$

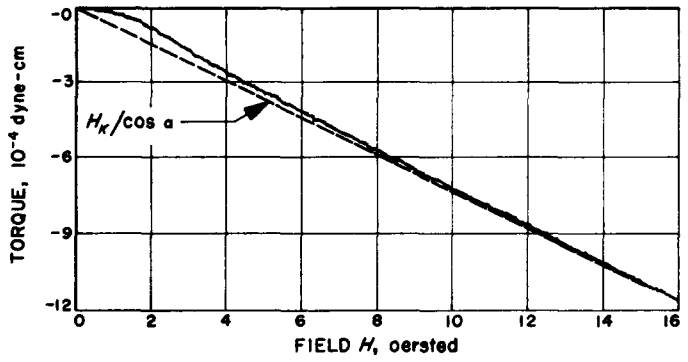


Fig. 3. Torque vs applied field perpendicular to easy axis and 20 deg to film surface

If $H << 0.5H_K$, the magnetization will be pulled away from the easy axis and then fall back again as α is increased. There will be no reversal, and the torque curve will be much like $\sin\alpha$ since ϕ will remain small. At some angle α of a torque curve, $(\alpha - \phi)$ will be equal to 90 deg; then

$$-L_{max} = T = HMV \tag{15}$$

The maximum value of the torque T_0 will be a measure of the remanent magnetization M_r or very nearly so if $H << H_c$. From Eq (7) which gives the high field or saturation magnetization, the squareness will be:

$$Sq = \frac{T_0}{H_0}\frac{H_\alpha \sin\alpha}{L_\alpha} \tag{16}$$

Squareness values for this film were calculated from Fig. 5i in *RS 36-13* (p. 21) and from Fig. 3 accompanying this article. It was found that Sq = 0.88.

It can be seen that a sensitive torque meter can be a powerful tool in the study of magnetism in thin films of magnetic materials, especially when the torque perpendicular to the plane of the film can be measured as well as the torque parallel to the plane. Not only can the anisotropy constant K be measured, but also the rotational hysteresis loss W_r, the remanent magnetization M_r, the saturation magnetization M_s, and the coercive force H_c. Accurate values of $2K/M_s$ and the squareness can be calculated.

B. Computer Study of Design Criteria for Control Systems

R. A. Bruns

In order to establish criteria whereby control system performance might be evaluated, a quantitative study of some of the more promising indices has been made. Although it is planned to apply these studies to more realistic models involving individual and collective spacecraft control systems, at the outset a simple model has been used to provide a quantitative standard for basic comparisons.

A comparison between *IES* (integral-error-squared) and *IAE* (integral-absolute-error) for a second-order linear attitude control system was presented in *RS 36-11* for disturbance step function inputs. Two additional criteria (Refs 7, 8, 9) which weight the time function are: *ITES* (integral-time-error-squared) and *ITAE* (integral-

time-absolute-error). These are expressed mathematically in terms of the nomenclature of *RS 36-11* as

$$ITES = \phi_T = \int_0^T tx^2(t)\,dt \qquad (1)$$

$$ITAE = \theta_T = \int_0^T t|x(t)|\,dt \qquad (2)$$

in which the disturbance is applied at $t = 0$, and the response is allowed to stabilize in a sufficiently extended period T.

In Fig 4 are shown the analog computer programs for quantitative comparisons of measurements of *ITES* and *ITAE* using the same model. As indicated previously in *RS 36-11*, iterative procedures were utilized to obtain the results listed in Table 1. The first column shows the setting of the rate constant K_r; the second column shows the voltage read on the output of integrator *A*, Fig 4, after transients had subsided; and the third column shows the set of numbers normalized to the minimum of the analytical value obtained from the *IES* criterion. These columns supply the data for the curves of Fig 5 in which the minima now necessarily appear at the same $f(K_r)$ for all four criteria. It will be noted that the *IES* is the flattest and the *ITAE* the sharpest. Occurrence of minima for each is shown in Table 2 together with the corresponding value for ζ, the damping constant as normally used in a second-order servomechanism.

Extensions of more complex weighting functions will not be attempted because of the accumulation of instru-

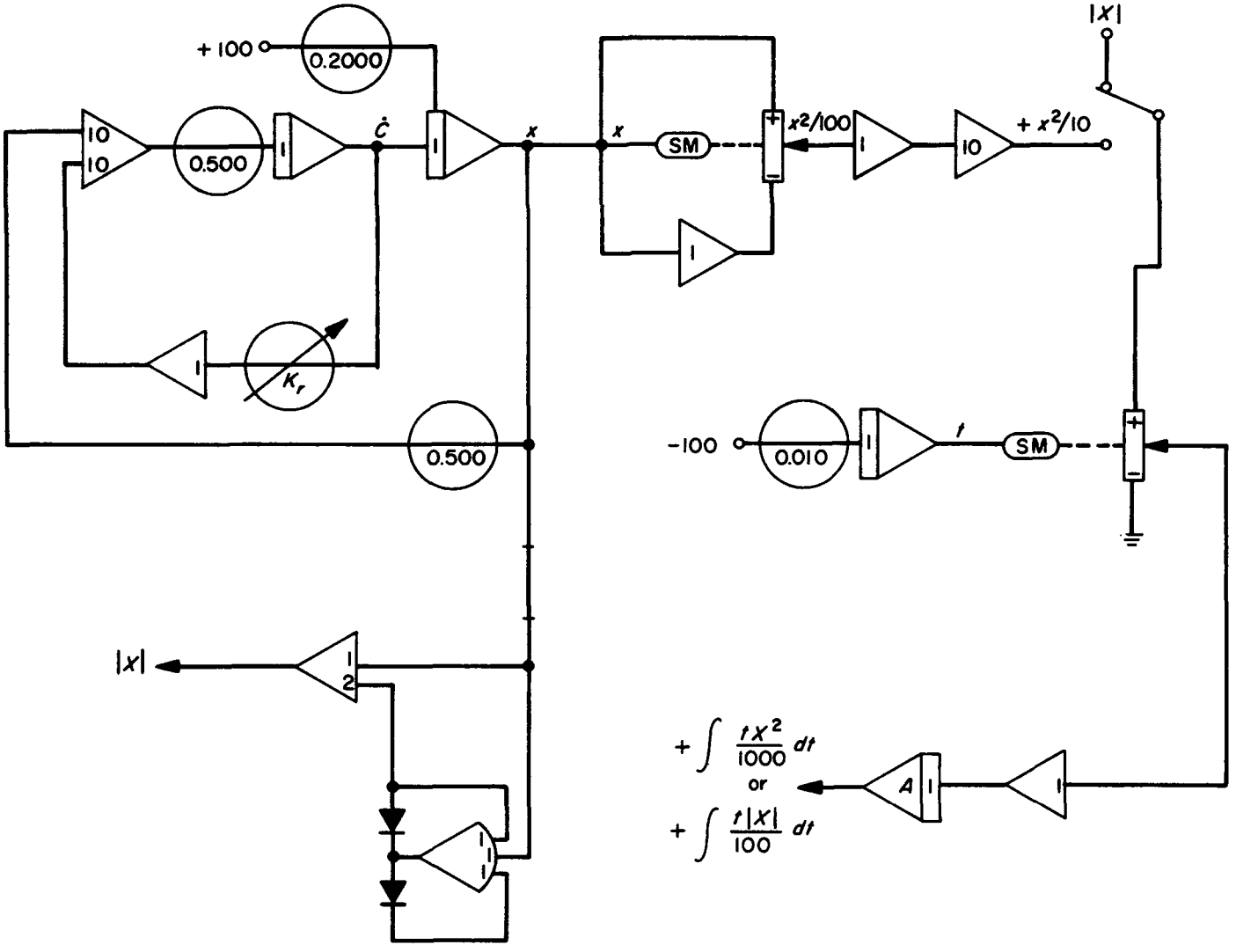


Fig. 4. Analog computer program for *ITES* and *ITAE* runs

Table 1. Analog computer data for *ITES* and *ITAE* runs

K_r	$ITAE = \frac{1}{10^2}\int_0^T t\lvert x(t)\rvert\,dt$		$ITES = \frac{1}{10^2}\int_0^T t\,x^2(t)\,dt$		$EPCE = \frac{1}{10}\int_0^T (x^2+\dot{c}^2)\,dt$	
	Volts	Norm. units	Volts	Norm. units	Volts	Norm. units
1.0	20.29	8.58	7.87	4.77	144.01	1.90
2.0	5.19	2.20	2.05	1.24	78.13	1.30
2.5	3.41	1.44	1.43	0.87	66.17	0.87
3.0	2.47	1.04	1.15	0.70	58.85	0.78
$\sqrt{10}$	2.26	0.96	1.11	0.67	57.14	0.76
3.5	1.93	0.82	1.05	0.64	54.20	0.72
4.0	1.64	0.69	1.04[a]	0.63	51.20	0.68
4.5	1.50	0.63	1.11	0.67	49.34	0.65
5.0	1.49[a]	0.63	1.23	0.75	48.25	0.64
6.0	2.00	0.85	1.56	0.95	47.64[a]	0.63
7.0	3.02	1.28	2.01	1.22	48.32	0.64
7.5	3.60	1.52	2.30	1.39	49.01	0.65
8.0	4.21	1.78	2.59	1.57	49.87	0.66
9.0	5.53	2.34	3.26	1.98	51.96	0.69
10.0	7.06	2.99	4.00	2.42	54.47	0.72

[a]Minimum reading for run.

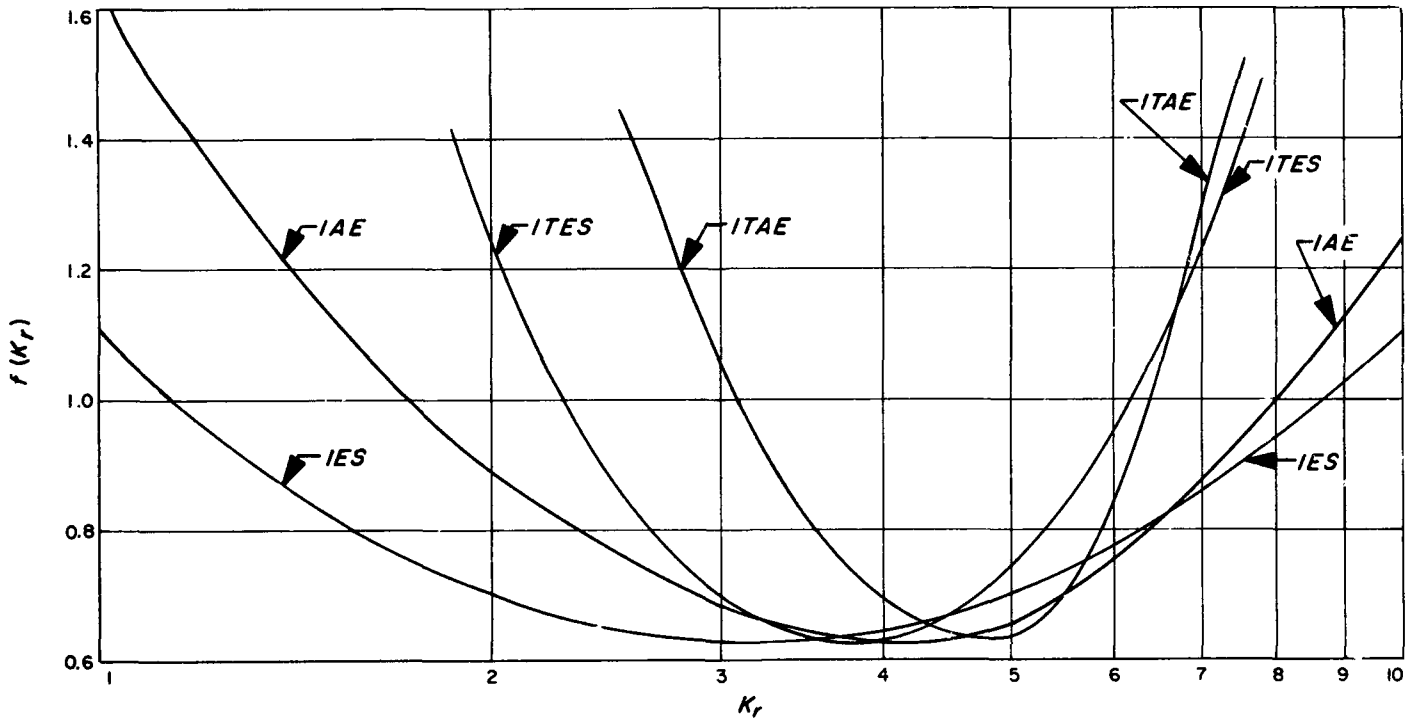


Fig. 5. Variations of four criteria for proportional attitude-control system

mentation errors which defeat the purpose of the criteria.

However, a criterion of the general form

$$J_1 + \lambda J_2 = \text{Min} \tag{3}$$

where

$$J_1 = \int_0^T x^2(t)\,dt \tag{4}$$

the usual *IES* criteria, and

$$J_2 = \int_0^T \dot{C}^2(t)\,dt \tag{5}$$

Table 2. Damping coefficients for minimum values of criteria for linear attitude-control system

Criteria	$(K_r)_{\min f(K_r)}$	$\zeta_{\min f(K_r)}$ [a]
IES	$\sqrt{10}$	0.5
IAE	4.25	0.67
ITES	3.85	0.61
ITAE	4.75	0.75
EPCE	6.00	0.95
[a] $\zeta = \frac{K_r K_1}{2(K_1 K_2 K_P)^{1/2}} = 0.158 K_r$		

which is a measure of the corrective effort (Ref 10), was applied to the control system. The case in which error and corrective effort are weighted equally ($\lambda = 1$) is shown in Table 1 and Fig 6. It will be noted from a comparison with the *IES* curve that the tendency is to flatten the curve and move the minimum toward larger damping. Once again, data is normalized to give equal minima for all criteria. This criterion is identified in Fig 6 and Table 1 as *EPCE* (error-plus-corrective-effort).

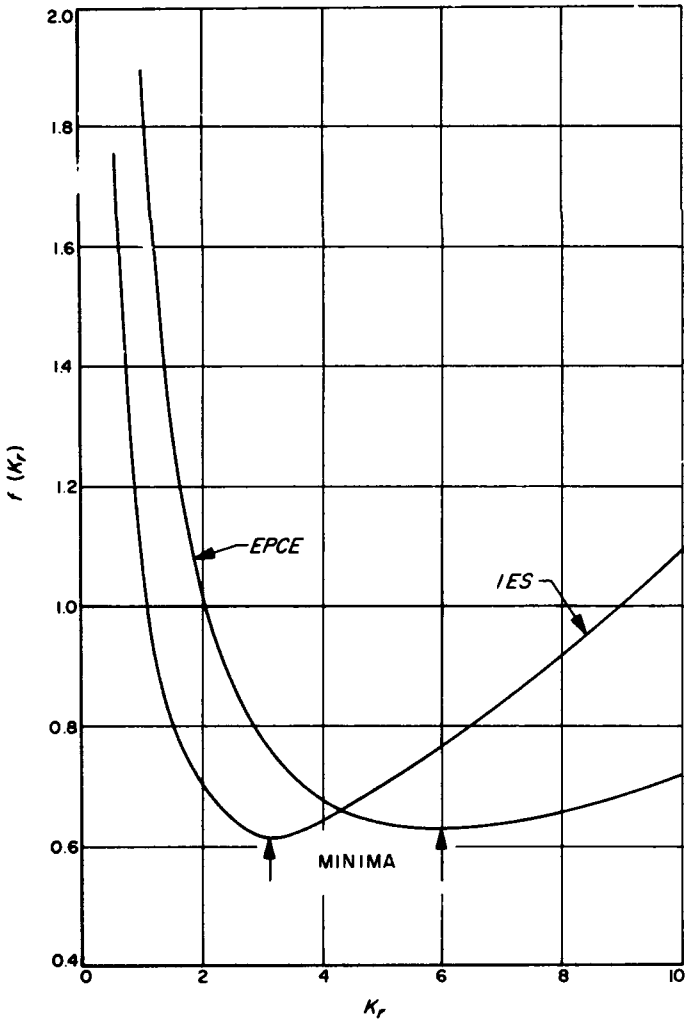


Fig. 6. Comparison of *J*- and *IES*-criteria for proportional attitude-control system

The *EPCE* criterion can, for example, include the effect of both spacecraft pointing error and fuel consumption, the importance of the latter being weighted by the value assigned to λ which in turn depends on overall mission objectives.

C. Measurements of AC Losses in Superconductors

J. T. Harding, R. H. Tuffias

Recent measurements of the deceleration rates of spinning superconducting spheres, levitated in a vacuum by a DC magnetic field, indicated an anomalously large dissipation of energy in the superconductor whenever the rotation caused the specimen to cut lines of flux. These losses were evidenced by rapid deceleration and ultimately by the loss of superconductivity caused by the rise in temperature.

A number of experiments have been conducted to determine the magnitude and nature of the AC losses. In these experiments, the specimen is stationary but is subjected to an alternating magnetic field. The length of time required to raise the thermally isolated superconductor to its transition temperature is the quantity measured. If the specific heat of the specimen in the superconducting state is known, this information leads to the rate of energy dissipation.

Tests were made of losses in both spherical and cylindrical specimens. In the former, the sphere was supported in an evacuated chamber by a nonuniform DC-levitation field. Superimposed on this was a uniform AC field. The ball was cooled to the temperature of the liquid helium bath surrounding the chamber, with the ambient magnetic field reduced to less than 1 milligauss. The transfer gas was pumped out, and the pressure (as read at room temperature) reached 10^{-5} mm Hg. The AC field was turned up to the desired level, and the length of time that the sphere remained levitated was recorded.

In Figs 7 and 8 are shown log-log plots of the time to fall vs rms AC field for several values of frequency; Fig 7 shows plots for Nb and Fig 8 for Nb_3Sn. A rough estimate of the quantity of heat required to heat the

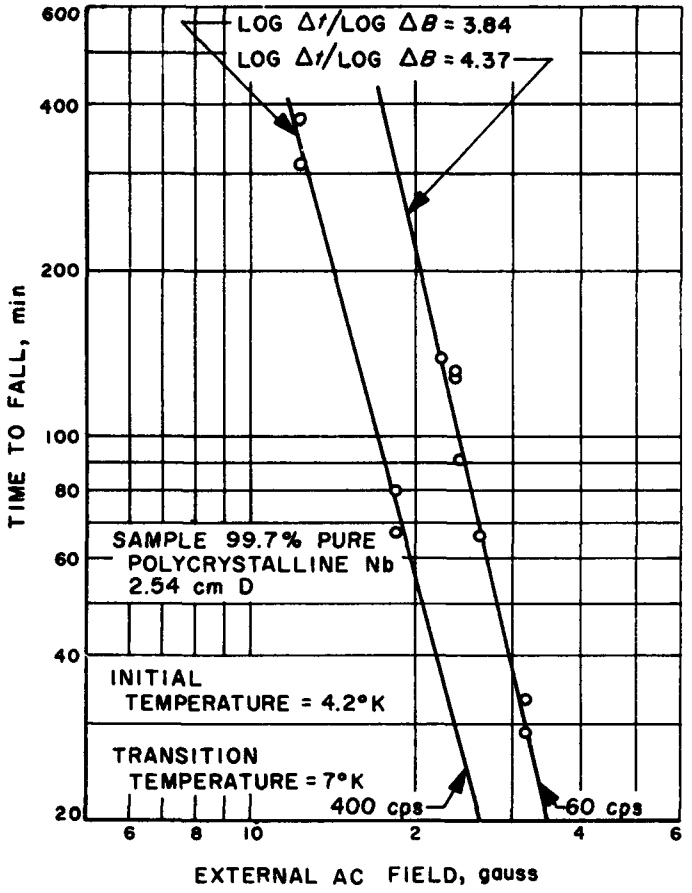


Fig. 7. AC losses in a solid Nb sphere

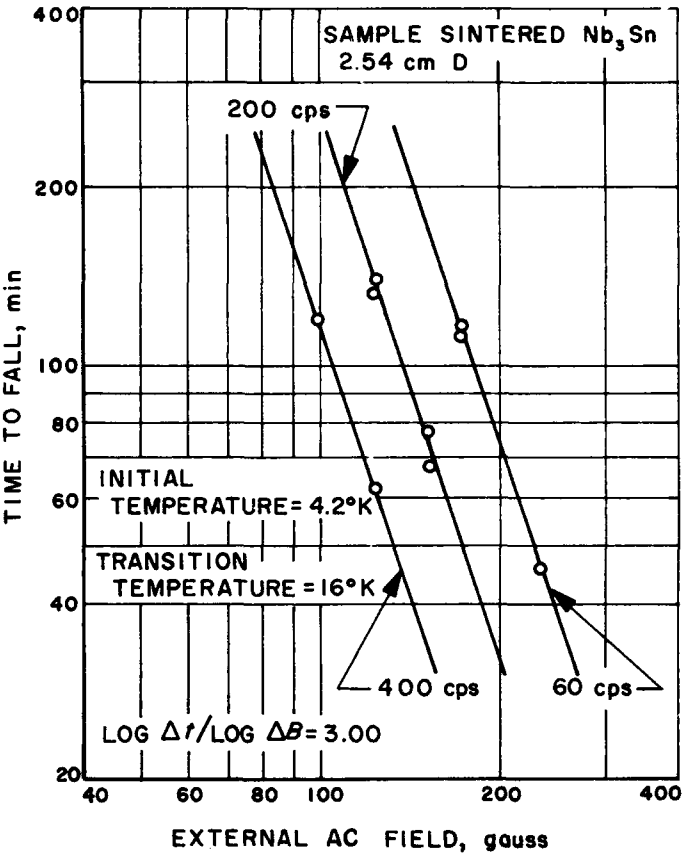


Fig. 8. AC losses in a solid Nb_3 Sn sphere

sample to the point where the critical field is exceeded and the ball falls is

for Nb:

$$Q = \frac{72\ \text{gm}}{92.9\ \text{gm/mole}} \int_{4.2°K}^{7°K} C_s\, dT = 0.18\ \text{joule}$$

The value of the specific heat in the superconducting state

$$C_s = 4.34 \times 10^{-4} T^3\ \text{joule/(°K)}^4\text{-mole}$$

was deduced from critical field data and the normal specific heat;

for Nb_3Sn:

$$Q = \frac{M}{m} \int_{4.2°K}^{16°K} C_s\, dT = 1.0\ \text{joule}$$

The value for Nb_3Sn was deduced from the time for fall when the Nb_3Sn was subjected to a known rate of radiative heating.

The average power dissipation is given by:

$$P = Q/t$$

for Nb (Fig 7) at 200 gauss AC field:

$$\overline{P}_{60\ \text{cps}} = 13.3\ \mu\text{w}$$

$$\overline{P}_{400\ \text{cps}} = 54.3\ \mu\text{w}$$

for Nb_3Sn (Fig 8) at 150 gauss AC field:

$$\overline{P}_{60\ \text{cps}} = 94.7\ \mu\text{w}$$

$$\overline{P}_{200\ \text{cps}} = 220\ \mu\text{w}$$

$$\overline{P}_{400\ \text{cps}} = 490\ \mu\text{w}$$

Several tests were run on the niobium sphere in which 40 to 60 gauss was trapped. The results indicated that the time required for the sphere to become normal was very dependent on the trapped field.

The necessity to DC-levitate the spheres severely restricts the range of AC field that can be used and also eliminates tests on materials with too low a critical field to permit support.

In the second series of tests, the specimen was in the shape of a long thin cylinder with rounded ends. This

cylinder was suspended by cotton thread within an evacuated tube and was subjected to a coaxial magnetic field. The normally coaxial components are shown side by side in Fig 9. Blackened welts on the sides of the glass tube just above the sample served as a radiation baffle. Relative susceptibility measurements made with a balanced set of oppositely wound coils revealed whether the specimen was normal or superconducting. Tests have been made with Pb and Sn rods (with the ambient magnetic field cancelled to less than 1 milligauss during transition), and the results are shown in Figs 10 and 11, respectively, where the time to become normal is plotted vs rms magnetic field for several frequencies.

The average power dissipation is given by

$$\overline{P} = Q/t$$

where Q, the heat required to reach the transition temperature, is

$$Q = \int_{T_0}^{T_c} C_s \, dT$$

for Pb (Fig. 10) at 225 gauss rms field ($B/B_{crit} \simeq 0.57$)

$$P_{60\,cps} = 14.3\ \mu w$$

$$P_{400\,cps} = 62.2\ \mu w$$

For Sn (Fig. 11) at 68 gauss rms field ($B/B_{crit} \simeq 0.73$)

$$P_{60\,cps} = 0.28\ \mu w$$

$$P_{400\,cps} = 1.59\ \mu w$$

Use has been made of the values for specific heat

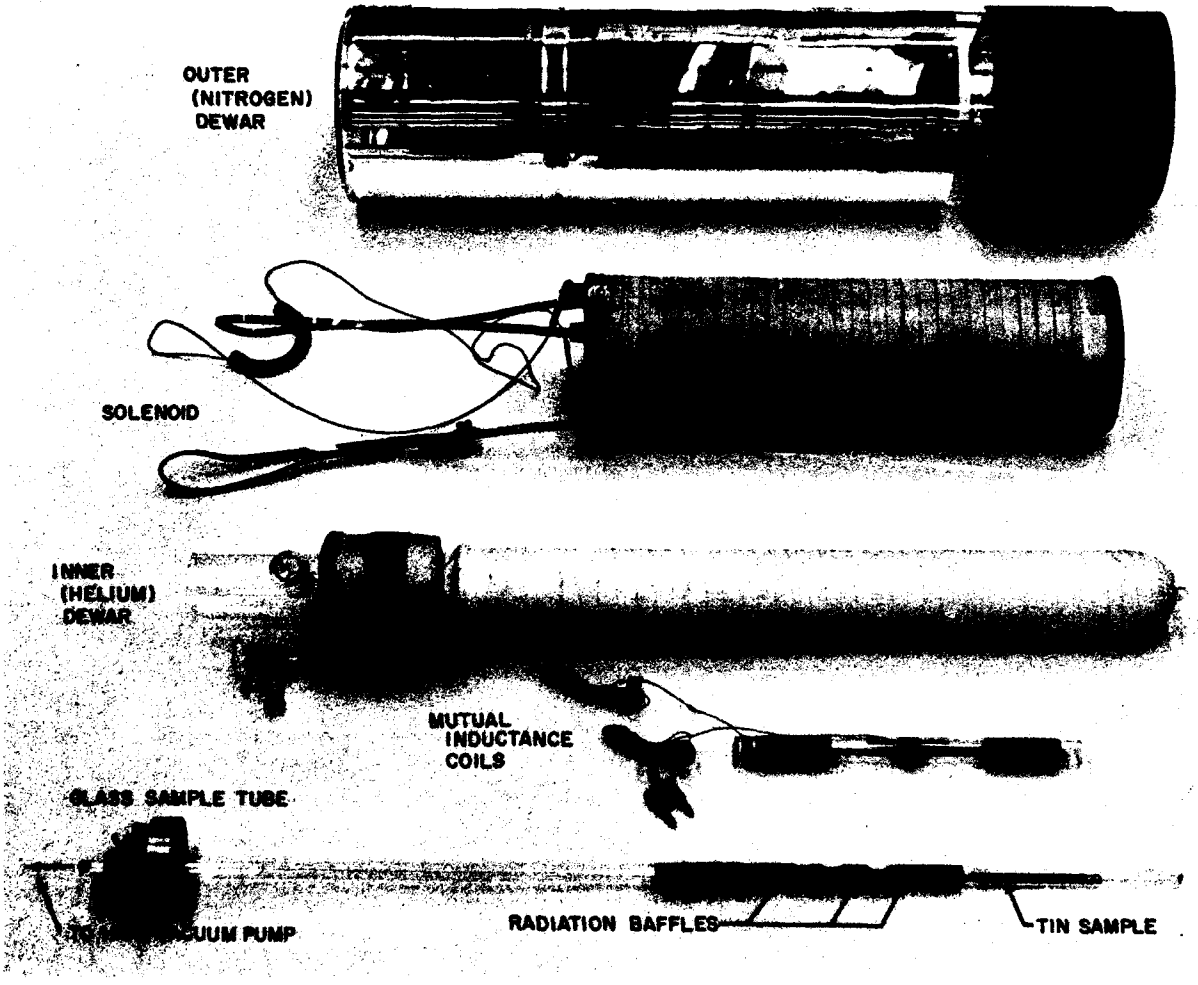


Fig. 9. AC loss test apparatus

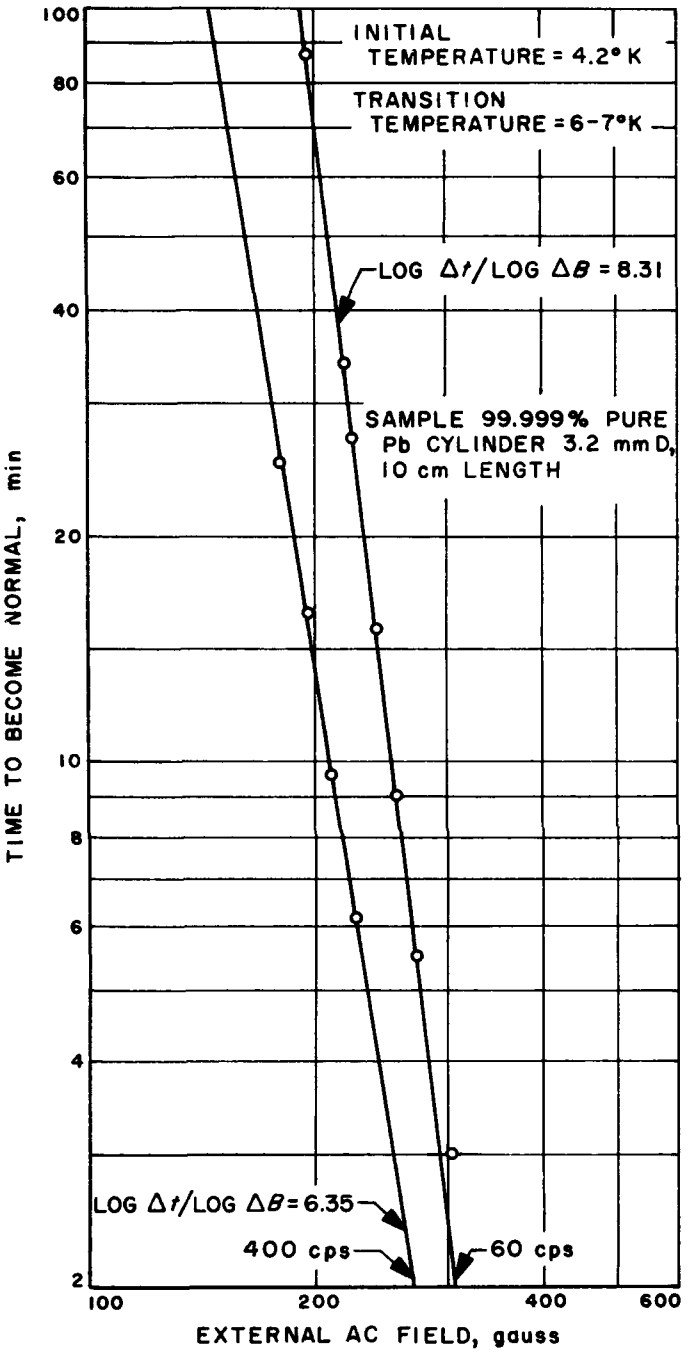


Fig. 10. AC losses in a solid Pb cylinder

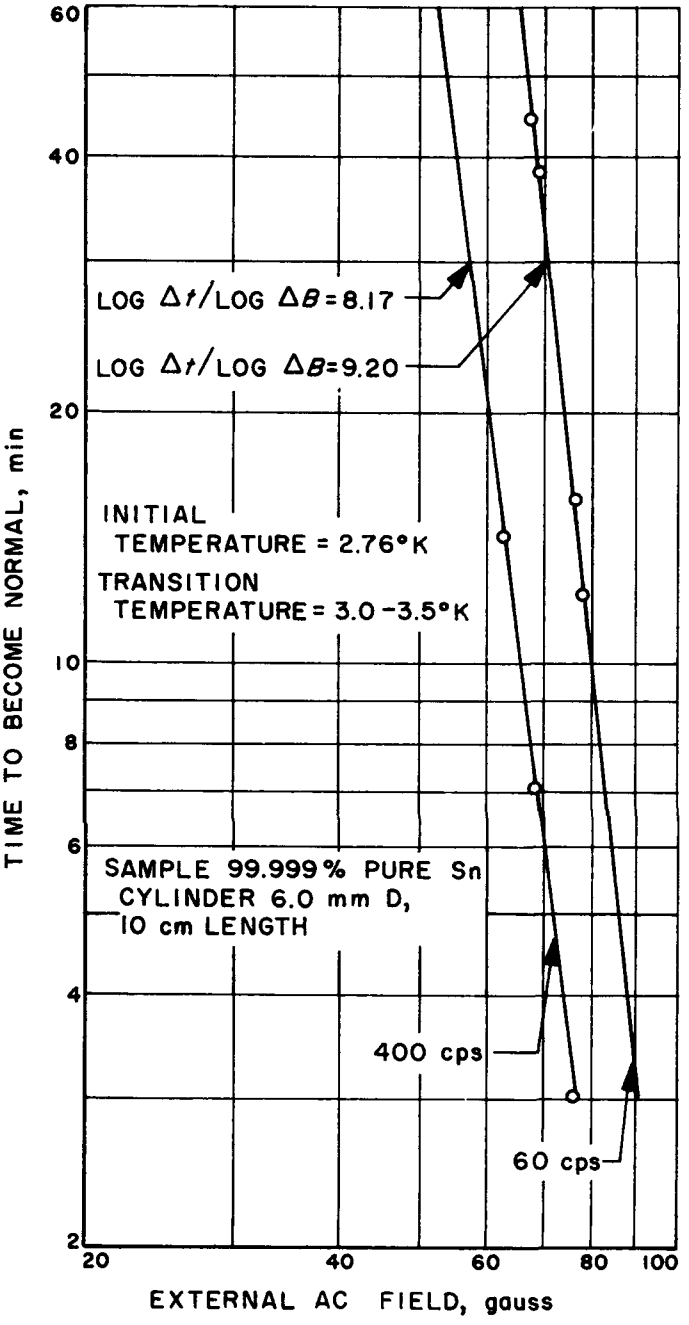


Fig. 11. AC losses in a solid Sn cylinder

Pb: $C_s = 2.47 \times 10^{-3} T^3$ j/(°K)⁴-mole

Sn: $C_s = 6.55 \times 10^{-4} T^3$ j/(°K)⁴-mole

There are several explanations proposed for the lossy behavior of these superconductors at such low frequencies.

Explanation 1. A normal layer on the surface will cause eddy-current losses. This layer might arise from compounds formed on the surface from exposure to atmospheric contaminants. On the other hand, at least for Nb, tunneling experiments indicate that machined surfaces are not superconducting. A damaged layer remains normal. Eddy current losses go as B^2 and are independent of frequency for thin shells, so this explanation does not seem to agree with the results.

Explanation 2. Because there is a non-zero penetration of magnetic field into an ideal superconductor, some

dissipation due to normal electrons is expected on theoretical grounds. A calculation shows that this expected value is less than 10^{-3} of the values obtained.

Explanation 3. Anomalous penetration of magnetic field into the superconducting specimens may cause additional dissipation. In particular, if the material (at least in the vicinity of the surface) is a filamentary structure of normal and superconducting regions, one expects both eddy current and hysteresis losses. Such a composition might be expected for niobium which is a hard superconductor; and perhaps the surfaces of the ultrapure Sn and Pb samples have contamination and lattice damage, which also produce characteristics of a hard superconductor. However, hysteresis losses depend linearly or frequency which is only approximately true for these losses; furthermore, the extreme dependence on B is unexplained.

References

1. Bozorth, R. M., *Ferromagnetism*, 6th Ed, p. 814, Van Nostrand Co., Princeton, New Jersey.
2. Ibid., p. 566.
3. Doyle, W. D., Rudisill, J. E., and Shtriknan, S., "Angular Dependence on Torque in Anisotropic Permalloy Films," *Journal of Applied Physics*, 32:1785, 1961.
4. Olson, C. D., and Pohm, A. V., *Journal of Applied Physics*, 29:274, 1958.
5. Smith, D. O., *Journal of Applied Physics*, 29:264, 1958.
6. Doyle, W. D., *Determination of the Anisotropy in Thin Permalloy Films*, Quarterly Progress Report Q-A2277-10, Franklin Institute, October 31, 1961.
7. Schultz, W., and Rideout, V., "Control System Performance Measures: Past, Present, and Future," *IRE Transactions*, PGAC, pp. 22–35, February 1961.
8. Graham, D., and Lathrop, R., "The Synthesis of 'Optimum' Transient Responses: Criteria and Standard Forms," *AIEE, Applications and Industry*, Vol. 72, Part II, pp. 273–286, 1953.
9. McDonnell, J., *Analytic Evaluation of the IT^2AE Performance Measure and Comparison of Analytic and Experimentally Determined Values of IT^2AE, IE^2, ITE^2, and IT^2E^2 for Zero Position-Error Second-Order Systems*, TM 97, Systems Technology, Inc., Inglewood, Calif., June 1961.
10. Chang, S.S.L., *Synthesis of Optimum Control Systems*, McGraw-Hill Book Co., 1961.

V. Flight Computers and Sequencers

A. Flexible Implementation of Digital Computer Arithmetic

A. Avizienis

The description of flexible implementation in *RS 36-12* presented a review of signed-digit number representations and the algorithms for addition, subtraction, multiplication, and division. These algorithms describe operations with operands of arbitrary multiple precision with respect to the length of the parallel adder in an arithmetic unit. The following sections deal with the generation of multiples and with significant digit operations in the proposed flexible implementation.

1. Adder With Multiple-Forming Circuits

a. Formation of multiples. Multiplication and division are implemented as sequences of additions and shifts. In both operations it is necessary to add a multiple $\pm cY$ of the addend Y to the augend Z. Multiples $\pm cy_i$ $(0 \leq c \leq a)$ of the addend digits y_i are to be added to the augend digits z_i. When the previously described flexible adder (Fig 53 in *RS 36-12*) is employed, $\pm Y$ is added c times to Z to complete one addition in the multiplication or division algorithm. However, it is also possible to devise a *multiple-forming circuit* which forms the multiples $\pm cy_i$ at once. When these multiples are available, the addition of $\pm cY$ is completed in one addition time. The addition of multiples is possible because more than two signed-digit numbers may be added simultaneously.

An adder for k digits with multiple-forming circuits for digits $\pm y_i$ is shown in Fig 1. The operation of the multiple-forming circuit M_i is described by the equation

$$c(\pm y_i) = v_i + rh_{i-1}$$

That is, two digits (v_i and h_{i-1}) are formed in M_i. Subsequently, the augend digit z_i is added to the two new digits in the adder circuits A_i^* and B_i:

$$z_i + v_i + h_i = rt_{i-1} + w_i \qquad \text{(circuit } A_i^*\text{)}$$

$$w_i + t_i = s_i \qquad \text{(circuit } B_i\text{)}$$

Minimal redundancy representations are preferred because the least number of multiples is required. For the radix r, digit magnitudes from 0 to $a_{\min}$, and $|t_i| \leq 1$, the following conditions must be satisfied by the values of v_i and h_i:

$$h_{\max} + v_{\max} \leq r - 1$$

$$rh_{\max} + v_{\max} \geq a_{\min} c_{\max}$$

$$v_{\max} \geq a_{\min} - 1$$

When the sets of allowed values of h_i: $\{-h_{\max} \cdots, -1, 0, 1, \cdots, h_{\max}\}$ and of v_i: $\{-v_{\max}, \cdots, -1, 0, 1, \cdots, v_{\max}\}$ are chosen so that these conditions are satisfied, addition of z_i, v_i and h_i will be totally-parallel and $s_i = f(z_i, \pm cy_i, z_{i+1}, \pm cy_{i+1}, \pm cy_{i+2})$ will hold. The derivation of the above conditions and a general discussion of multiple-forming circuits will be presented in a technical report.

For minimal redundancy representations, we have,

$$c_{\max} = a_{\min} = \tfrac{1}{2}(r_0 + 1)$$

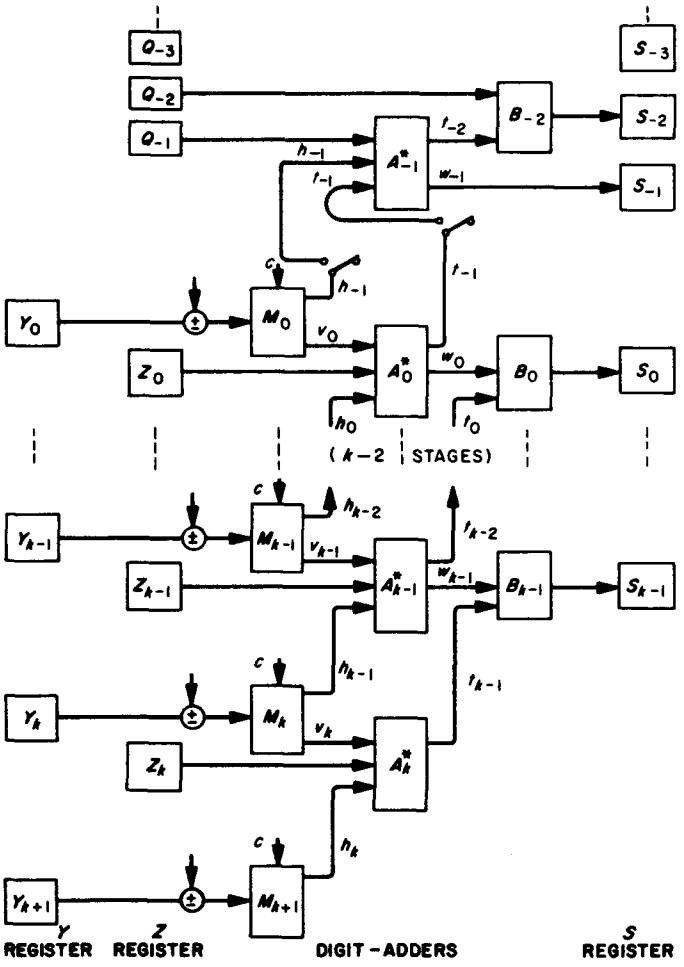


Fig. 1. An adder for K digits with multiple-forming circuits

for odd radices $r_e \geq 3$, and

$$c_{max} = a_{min} = \tfrac{1}{2}r_e + 1$$

for even radices $r_e \geq 4$. The choice of the least allowed value of $v_{max} = a_{min} - 1$ satisfies the conditions for all odd radices $r_e \geq 3$ and all even radices $r_e \geq 8$. For the even radices $r_e = 6$ and $r_e = 4$ it is necessary to establish $c_{max} = a_{min} - 1$, that is, to recode the multiplier digits serially into the canonical form. For greater values of r ($r_e \geq 10$ and $r_o \geq 7$), the choice of $v_{max} = a_{min}$ also satisfies the totally-parallel requirements. This choice offers the advantage that for $c = 1$ we have $h_{i-1} = 0$; that is, $\pm y_i$ is not recoded when $c = 1$. Two choices exist for the radix $r = 10$; one is $v_{max} = 5$ and $h_{max} = 4$ and the other is $v_{max} = 6$ and $h_{max} = 3$. Still more choices of the pairs (v_{max}, h_{max}) exist when r increases further.

b. Addition. The addition $Z \pm cY$ in the $k \geq 1$ digit length adder of Fig 1 is similar to addition $Z \pm Y$ which was described for Fig 53 of *RS 36-12*. The adder consists of k digit-adder circuits (M_0, A_0^* and B_0 to M_{k-1}, A_{k-1}^* and B_{k-1}); the storage registers Y, Z, and S; and the auxiliary circuits M_k, A_k^* and M_{k+1}. The circuits A_{-1}^* and B_{-2} and associated storage locations are used in multiplication and will be discussed later.

Given the n digit length operands Z and Y, the addition begins with the k most significant digits (y_0 to y_{k-1} and z_0 to z_{k-1}) held in storage register locations Y_0 to Y_{k-1} and Z_0 to Z_{k-1}. The $(k+1)$th digits of Y and Z (y_k, z_k) and the $(k+2)$th digit of Y (y_{k+1}) serve as inputs to the auxiliary circuits which supply the adder inputs h_{k-1} and t_{k-1}. During a multiplication, the inputs h_{-1} and t_{-1} are connected to the circuit A_{-1}^*. The inputs "$\pm$" and "c" to each circuit M_i are set for the desired multiple of Y. The k most significant digits of the sum are generated one addition time later; overflow is immediately detectable. The next k digits of Y and Z now may enter the adder, with the digits y_k, z_k moving to storage locations Y_0, Z_0, and y_{k+1} to Y_1. Inputs h_{-1} and t_{-1} are now disconnected from circuit A_{-1}^*. The addition is completed after α steps, where α is the least integer such that $\alpha \geq n/k$.

c. Multiplication. Multiplication is implemented as a sequence of additions and shifts, described by the algorithm:

$$P_j = r(P_{j-1} + x_j Y) \text{ for } j = 0, 1, 2, \cdots, m$$

where r is the radix, Y is the multiplicand, x_0 to x_m are the digits of the multiplier X, P_j ($j = 0, 1, \cdots, m-1$) are partial products, $P_{-1} = 0$, and $P_m = r^{m+1}$ XY is the product (with respect to the radix point location of Y).

When the adder of Fig 1 is used to perform multiplication, the Y register holds the multiplicand, while the partial products are accumulated in the Z register and its extension to the left, the Q register. The multiplier X is held in a shift register X and its digits are sensed sequentially, starting with the most significant digit x_0. One step of multiplication consists of an addition of $x_i Y$, followed by a left shift of the sum (all digits of the partial product move from storage locations S_i into Z_{i-1} or Q_{i-1}). One product digit moves into Q_{-1} during each shift. The circuits A_{-1}^* and B_{-2} add the incoming transfer digits h_{-1} and t_{-1} when the next partial product is formed. Circuits A_{-1}^* and B_{-2} are used and h_{-1} and t_{-1} are connected to A_{-1}^* only when Y_0 holds the most significant digit y_0 of the multiplicand Y. After the $(m+1)$th step, Q holds $m+1$ most significant digits of the product P_m, the last of which is in Q_{-1}, while the less significant digits of P_m are held in Z, that is, we have $P_m = r^{m+1}XY$ with respect to the radix point location of Y.

The totally parallel addition requirement for the circuit A^*_1 of Figure 1 is expressed as:

$$|Q_{-1}| + |b_{-1}| + |t_{-1}| \leq r + w_{max}$$

Upon the choice of the greatest possible values $w_{max} = a - 1, |Q_{-1}|_{max} = a$ and $|t_{-1}|_{max} = 1$, we have the requirement $|h_{-1}|_{max} \leq r - 2$, which can be satisfied for all $r \geq 3$.

d. Division. Given the radix r dividend Z and divisor Y, division is described by the algorithm:

$$R_j = r(R_{j-1} - Yq_j) \text{ for } j = 0, 1, \cdots, m$$

where $R_{-1} = Z$ is the dividend, R_j $(j = 0, 1, \cdots, m-1)$ are partial remainders, $R_m = r^{m+1}(Z - QY)$ is the remainder, q_0 to q_m are digits of the quotient Q, and $m+1$ is the required number of quotient digits (*RS 36-12*, pp. 46–49). When the adder of Fig 1 is employed, one step of division consists of the choice of q_j, followed by the addition of $-Yq_j$ and a left shift of the sum. During each step of division the value of the quotient digit q_j must be chosen to be such that the partial remainder R_j is within the same allowed range as the previous partial remainder R_{j-1}. This range is a function of the magnitude of divisor Y and of the allowed values of quotient digits q_j; that is, the requirement

$$|R_j| \leq rK \cdot |Y|$$

must be satisfied, where K may be any value in the range

$$\tfrac{1}{2} \leq K \leq \frac{|q_j|_{max}}{r-1}$$

A detailed analysis of this method of division is presented in the Appendix of Ref 1.

Since the quotient is in a redundant representation, the selection of quotient digit values may be based on the comparison of approximate (truncated) magnitudes Y'_n and R'_{j-1} of the normalized divisor Y_n and a partial remainder R_{j-1} (or dividend Z). To complete one step of division in one addition time, we employ a total of b $(b = |q_j|_{max})$ comparison circuits, in which the test $|R'_{j-1}| \leq (g + \tfrac{1}{2})|Y'_n|$ is performed for the values of $g = 0, 1, \cdots, b-1$. The least value of g which satisfies the test gives $|q_j| = g$; if no value of g satisfies the test, then $|q_j| = b$. The sign of q_j is chosen to be such that the signs of R_{j-1} and Yq_j agree.

For the ranges $-1 \leq Z \leq 1$ and $r^{-1} \leq |Y_n| \leq 1$, it is sufficient to employ the four most significant digits of R_{j-1} and Y_n to form $|R'_{j-1}|$ and $|Y'_n|$, respectively. The selection of the quotient digit is completed after the inspection of first four digits of the operands Y_n and R_{j-1}, regardless of their length. Since addition of $-Yq_j$ and a left shift now completes one step of division, the proposed implementation is especially effective for multiple precision operands.

2. Floating-Point Representation of Numbers

a. Introduction. In a positional number system with a constant radix r, each numerical value

$$Z = \sum_{i=b}^{a} z_i r^i$$

is represented by a finite number of significant digits z_i, where the significant (non-zero) digit z_b with the greatest weighting factor r^b is designated as the *leading digit*. In a *fixed point* number system, a unique weighting factor r^i is associated with each position i in a multidigit number (one *machine word*). The weighting factor of the leading digit of any number is established by its location with respect to a fixed reference point, the leftmost position of the word. (The reference is often the radix point, and the range of algebraic values is then $-1 < Z < 1$.) All positions between the reference point and the leading digit are filled by *leading zeros*. When the word length is fixed, it may be necessary to fill the positions to the right of the lowest-weighted significant digit z_a (designated as the *trailing digit*) with non-significant zero digits. These zeros usually are treated as significant during subsequent arithmetical operations. Conversely, the number of significant digits may exceed the word length; in this case either a round-off is performed or multiple-precision arithmetic is necessary. The use of leading zeros is a rather inefficient method to indicate the weighting of the leading digit; however, it keeps the digital positions aligned in addition of two numbers.

In a *floating point* number system a numerical value Z^* is represented as a fixed-point number Z (usually the *fractional part* with the range $-1 < Z < 1$) and an integral *exponent* E, such that the pair (Z, E) represents $Z^* = Zr^E$. The exponent E (together with the reference point of Z) now designates the weighting of the leading digit of Z^*. This makes leading zero digits unnecessary, and the fractional part Z is usually required to be without leading zero digits, that is, to be in the range $1 > |Z| \geq r^{-1}$. When $|Z|$ is in this range, the number is said to be in *normalized* form. In a fixed-length word, non-significant zero digits are used to fill the remaining positions to the right of the trailing digit. When leading zero digits develop in Z during a computation, Z is normalized by left shifts and more non-significant zeros enter at the right. As a result, error propagation is erratic,

and the number of significant digits may become uncertain after a long series of arithmetical operations.

To control the uncertainty in error propagation, a *significant digit arithmetic* has been proposed (Ref 2, 3). In this system, floating point operations are performed so that the number of the significant digits in the result is approximately equal to the number of significant digits in the operand having fewer significant digits. The number of significant digits is indicated by using leading zeros to align the trailing digit with the last position of a fixed-length word.

b. Signed-digit representations for floating-point arithmetic. For the discussion of a flexible implementation of floating-point arithmetic, it is convenient to select a preferred type of signed-digit representations. Minimal redundancy representations are preferred because they require least storage for the values of one digit and have the least magnitude of a ($a = a_{min}$ is the greatest allowed magnitude of a digit). Radix 10 representations are the most convenient choice for examples and practical applications. The radix 10 numbers with 13 digit values (-6 to 6) require only four binary storage elements per digit, (same as conventional radix 10 numbers) and have the advantage of the general usage of radix 10. The rules of implementation, however, may be extended for all choices of radix and redundancy.

The allowed range of the algebraic values

$$Z = \sum_{i=0}^{m} z_i r^{-i}$$

for fixed-point numbers is required to cover the range $1 \geqq Z \geqq -1$. An easily implemented method of overflow detection is also necessary. The two most significant digits z_0 and z_1 are inspected to detect overflow. *Positive overflow* occurs when: $z_0 > 1$; or $z_0 = 1$ and $z_1 > 0$; and *negative overflow* occurs when: $z_0 < -1$; or $z_0 = -1$ and $z_1 < 0$. For these overflow detection rules, no overflow will be indicated for the algebraic values of

$$|Z| < 1 + \frac{1}{r} - \frac{a}{r-1}(r^{-1} - r^{-m})$$

but overflow will be always indicated for

$$|Z| > 1 + \frac{a}{r-1}(r^{-1} - r^{-m})$$

where z_m is the least significant digit (the number is $m + 1$ digits long). In the range of values of $|Z|$ between these limits, overflow may or may not be indicated, dependent on the specific representation of Z. This *potential* overflow indication range exists because signed-digit representations are redundant and some algebraic values may be represented in more than one way. For example, given radix $r = 10$ and $a = 6$ (minimal redundancy), we shall have certain overflow indication for $|Z| > 32/30$ and no overflow indication for $|Z| < 31/30$.

In a floating-point (radix r) number system, a numerical value Z^* is represented as a *fractional part* Z and an integral *exponent* E such that the pair (Z, E) represents $Z^* = Zr^E$. We require the fractional part to be *normalized* since leading zeros are not significant. A normalized form of the fractional part Z is defined as the form in which the application of the overflow rules indicates no overflow when z_0 and z_1 are inspected; and it indicates overflow when z_1 is substituted for z_0, and z_2 is substituted for z_1. According to this definition, the range of the normalized fractional part Z is:

$$r^{-1} + r^{-2}\left(1 - \frac{a}{r-1}\right) + r^{-m}\frac{a}{r-1} \leqq |Z| \leqq 1 + r^{-1}\frac{a}{r-1} - r^{-m}\frac{a}{r-1}$$

Minimal redundancy representations with $a = a_{min} = \frac{1}{2}(r_0 + 1)$ for odd radices $r_0 \geqq 3$ and $a = a_{min} = \frac{1}{2}r_e + 1$ for even radices $r_e \geqq 4$ will be employed to represent the fractional part Z. Representation and arithmetic of the exponents E will be considered in a subsequent discussion.

3. Indication of Precision

In the proposed flexible implementation the addition of two numbers of arbitrary length commences with the most significant parts of the operands. Furthermore, the first step of a multiplication or division yields the most significant digit of the product or quotient, respectively. The time required by an operation will be minimized if a completion signal occurs as soon as the required number of significant digits of a sum, product, or quotient has been generated.

To determine the number of significant digits in the result, the number of significant digits in the input operands must be known. This information may be incorporated into the representations of numbers by a special digit ϕ, designated as the *space-zero*. The nonsignificant digit positions at the right (low significance) ends of the input operands and partial results are identified by this special digit ϕ. The relative locations of the ϕ digits in all operands supply the information concerning when to conclude an arithmetical operation.

There exist two methods to apply the digit ϕ. In the first method, the addition rules:

$$\text{(A)} \quad z_i \pm \phi = rt_{i-1} + \phi$$

$$\text{(B)} \quad t_i + \phi = \phi$$

apply to the new digit ϕ. In Rule (A), we have $w_i = \phi$ for all values of z_i, while $t_{i-1} = 1$ if $z_i > w_{max}$, $t_{i-1} = -1$ if $z_i < -w_{max}$, and $t_{i-1} = 0$ if $|z_i| \leq w_{max}$ or if $z_i = \phi$. In Rule (B), $s_i = \phi$ is the value of the sum digit when $w_i = \phi$, regardless of the value of t_i. When Rules (A) and (B) apply, the sum of two signed-digit numbers is rounded off by truncation to the length of the number with fewer significant digits. If every allowed value of the digit z_i occurs with the same probability, then the average error which is introduced by truncation is zero and the round-off is without bias. End of addition is signalled by the detection of ϕ as an input digit to the adder, or as a sum digit.

In the second method of addition, the digit ϕ is interpreted as an ordinary zero digit:

$$(C_1) \quad z_i \pm \phi = z_i \pm 0 = rt_{i-1} + w_i, \text{ for all } z_i \neq \phi$$

$$(C_2) \quad \phi + \phi = 0 + \phi$$

An exception occurs in Rule (C_2), where $t_{i-1} = 0$ and $w_i = \phi$ is generated. Rule (B) holds for $w_i = \phi$. When Rules (C_1) and (C_2) apply, round-off is avoided and the sum has the length of the number with more significant digits. End of addition is signalled by the detection of the digits ϕ as both inputs to (or the output from) one position of the adder. Rule (C) will be applied in the implementation of multiplication and division.

In order to define a normalized form of $Z = 0$, we add the rule that $z_1 = 0$ and $z_2 = \phi$ is an "overflow indication" in the test for normalization. Then $Z = 0$ is represented by $z_0 = 0$, $z_1 = 0$ and $z_2 = \phi$, and the value of the exponent E indicates that $Z = 0$ may be an approximation.

The application of the ϕ digit leads to an implementation of arithmetic in which the operations are performed with significant digits only. The propagation of error during a sequence of calculations may be more readily estimated in this case; furthermore, the least time will be taken by an arithmetical operation if a completion signal occurs as soon as the required number of significant digits of a sum, product, or quotient has been generated. A floating-point arithmetic will be described in the next *Summary* in which the numbers are normalized and the ϕ digits indicate the precision of fractional parts. The proposed implementation follows the rules of significant digit arithmetic which have been developed by Metropolis and Ashenhurst (Refs 2, 3), whose implementation employs leading zero digits in a fixed-length unnormalized fractional part to indicate the number of significant digits.

References

1. Avizienis, A., "Signed-Digit Number Representations for Fast Parallel Arithmetic." *IRE Transactions on Electronic Computers, EC-10*, pp. 389–400, 1961.
2. Metropolis, N. and Ashenhurst, R. L., "Significant Digit Computer Arithmetic," *IRE Transactions on Electronic Computers, EC-7*, pp. 265–267, 1958.
3. Ashenhurst, R. L. and Metropolis, N., "Unnormalized Floating Point Arithmetic," *Journal of the Association for Computing Machinery, 6*, pp. 415–428, 1959.

TELECOMMUNICATIONS DIVISION

VI. Communications Systems Research

A. Telemetry Coding

E. C. Posner, T. Kailath, A. L. Duquette, and G. Solomon

1. Coding for Source and Criterion, E. C. Posner

In considering the problem of when and how to code, attention is usually paid (for reasons of mathematical simplicity) to uniform random sources, i.e., to sources which choose symbols from an alphabet at random with the same probability for each symbol. More important is the "criterion of fidelity," first defined rigorously in Ref 1. This criterion is adopted to judge the efficacy of the communication system and is usually a word or bit error probability. However, there are many practical situations where the source is not random, and/or the criterion of fidelity should not be taken as an error probability.

Consider the example (*RS 36-12*, p 76) where a white Gaussian channel is used to transmit the results of independent coin tossings, the object being to estimate the probability of a head. The appropriate criterion of fidelity here is not an error probability, but rather the variance of the estimate for the probability of a head (with a given time of observation). It was shown (*RS 36-12*, p 76) that, relative to this criterion of fidelity, it is better not to code if the channel is to be made into a binary symmetric channel. Instead, the results should be sent as they happen even if the coding (ignoring a certain fraction of the outcomes) would be error free and at channel capacity.

However, considering the same channel but with the fidelity criterion of error probability, one knows that it does pay to code in many circumstances. As shown in *RS 36-12*, pp 66–67, a certain class of error-correcting codes coupled with bit-by-bit detection always gives improvement for high output signal-to-noise ratios. Limits for the effective increase in output signal-to-noise ratio which such codes furnish are given in *RS 36-13*, pp 35–38.

Coding to make the source random is sometimes called data compression. After compression, one can use ordinary coding techniques with confidence, provided one is convinced that the criterion of fidelity should be an error probability. The real difference occurs when error probability is not the right criterion. For example, if the data are to be used for parameter estimation, the variance of the estimate is a proper criterion.

The coin-tossing problem mentioned earlier is an extreme form of this; here, one is definitely not interested in the individual outcomes but only in the total number of heads. Many scientific space-borne experiments are of this general nature. Thus, in order to design a space communication system effectively, the communications

engineer should have knowledge of the experiments (as data sources) and also the purposes of the experiment (as fidelity criteria).

2. Error-Correcting Codes at Very Low Signal-to-Noise Ratios, E. C. Posner

If a space probe transmitting at a fixed power level is receding from the Earth, the signal-to-noise ratio at the output of the receiver can be less than one, and in fact approaches zero[a]. If a (linear) error-correcting code is used to encode the telemetry, one is interested in the asymptotic performance of the code as the signal-to-noise ratio approaches zero. This asymptotic result is needed in order to compare the thresholding properties of error-correcting encoding with the performance of other encoding systems. The following definition will be used (*RS 36-13*, pp 32–38): the *expected bit error probability* is defined as the bit error probability on an uncoded channel which results in the same expected number of information bits in error (out of k bits, where k is the number of information bits in the code) as one has when using the code. The *power gain* (or *loss*) resulting from use of a code is defined as follows: the code has n bits of which k are information bits. Suppose one chooses to increase the symbol time by a factor of n/k and then does not code the k information bits. The reciprocal of the factor by which the original signal-to-noise power ratio must be multiplied is the power gain or loss and is usually expressed in db. (Throughout, the white Gaussian channel is assumed.) This is done so that the expected bit error probability or word error probability using the code is equal to the bit or word error probability in the uncoded channel using the longer symbol time.

The criterion of word error probability shall be considered first. Let f_j, $0 \leq j \leq n$, denote the number of error patterns of weight j corrected by the code, so that

$$\sum_{j=0}^{n} f_j = 2^{n-k}$$

Consider a probability of symbol error $p = \frac{1}{2} - \delta$, where $\delta\,(>0)$ approaches zero as the signal-to-noise ratio approaches zero. Let

$$q = 1 - p = \tfrac{1}{2} + \delta$$

Then the probability α of correct decoding when using the code is expressible as

$$\alpha = \sum_{j=0}^{n} f_j\, p^j\, q^{n-j}$$

After several intermediate steps, one finds the following theorem:

Theorem 1.

$$\alpha = (\tfrac{1}{2})^k + \delta\left(\frac{2n}{2^k} - \frac{4\gamma}{2^k}\right) + 0\,(\delta^2)$$

where

$$\gamma = (\tfrac{1}{2})^{n-k} \sum_{j=0}^{n} j f_j$$

the average weight of a correctible error (γ is difficult to evaluate for most codes).

Now define α' as the probability of correctly receiving k bits using no coding on a channel in which the symbol time is n/k times as long as before. Let $p' = \frac{1}{2} - \delta'$, $q' = \frac{1}{2} + \delta'$ be the probability of symbol error and of correct symbol reception, respectively, where δ' corresponds to the same output signal-to-noise ratio to which δ corresponds. (As the output signal-to-noise ratio approaches zero, both δ and δ' approach 0.) One has

$$\alpha' = (q')^k = (\tfrac{1}{2} + \delta')^k = (\tfrac{1}{2})^k + k\,(\tfrac{1}{2})^{k-1}\,\delta' + 0\,(\delta')^2$$

Referring to *RS 36-13*, pp 37–38, for the output signal-to-noise (power) ratio s^2, one has

$$p = \tfrac{1}{2} - \int_{v=0}^{s} \frac{1}{(2\pi)^{1/2}}\, e^{-v^2/2}\, dv$$

$$p' = \tfrac{1}{2} \int_{v=0}^{(n/k)^{1/2}s} \frac{1}{(2\pi)^{1/2}}\, e^{-v^2/2}\, dv$$

or

$$\delta = \int_{v=0}^{s} \frac{1}{(2\pi)^{1/2}}\, e^{-v^2/2}\, dv$$

$$\delta' = \int_{v=0}^{(n/k)^{1/2}s} \frac{1}{(2\pi)^{1/2}}\, e^{-v^2/2}\, dv$$

As $s \to 0$,

$$\delta = s \cdot \frac{1}{(2\pi)^{1/2}} + o\,(s)$$

$$\delta' = \left(\frac{n}{k}\right)^{1/2} \frac{s}{(2\pi)^{1/2}} + o\,(s)$$

so that

$$\alpha = (\tfrac{1}{2})^k + \frac{s}{(2\pi)^{1/2}}\left(\frac{2n}{2^k} - \frac{4\gamma}{2^k}\right) + o\,(s)$$

[a]There will be a problem in maintaining phase reference in a phase-modulated system as the signal falls off; an envelope detector avoids this problem, but the assumption of Gaussian noise is then only approximately true.

$$\alpha' = (\tfrac{1}{2})^k + \left(\frac{n}{k}\right)^{1/2} \frac{s}{(2\pi)^{1/2}} k\,(\tfrac{1}{2})^{k-1} + o(s)$$

The factor by which s^2 must be multiplied in the formula for α' to make α' equal to α, i.e., the effective power gain G, is now given by the following theorem:

Theorem 2.

$$G = \frac{(n - 2\gamma)^2}{nk}$$

(This will be a gain if and only if $G > 1$.)

For the Hamming (7, 4) single error-correcting code, $m = 7$, $k = 4$, $\gamma = \frac{7}{8}$, one has

$$G = \frac{\left(7 - \frac{14}{8}\right)^2}{28} = \frac{63}{64} < 1$$

Thus there is a slight loss as $s \to 0$. For the higher Hamming codes, $n = 2^m - 1$, $k = 2^m - m - 1$,

$$\gamma = \frac{n}{n+1} = \frac{2^m - 1}{2^m} = 1 - (\tfrac{1}{2})^m$$

one has

$$G = \frac{\left[(2^m - 1) - 2\left(1 - \frac{1}{2^m}\right)\right]^2}{(2^m - 1)(2^m - m - 1)}$$

which approaches 1 as $m \to \infty$, but is greater than 1 for $m \geq 4$. A gain can therefore occur with a given code as $s \to 0$. For the Golay (23, 12) triple error-correcting code $n = 23$, $k = 12$,

$$\gamma = \frac{1 \cdot 0 + 23 \cdot 1 + 2\left(\frac{23 \cdot 22}{2}\right) + 3\,\frac{23 \cdot 22 \cdot 21}{3 \cdot 2 \cdot 1}}{2^{11}} = 2.858$$

and

$$G = \frac{(23 - 5.716)^2}{23 \cdot 12} = 1.037 > 1$$

a slight gain. Thus using the criterion of word error probability, there can be a power gain (but not much for these word lengths, evidently).

The picture is clearer when the criterion of expected bit error probability is adopted; this criterion is a reasonable one for many telemetry situations.

Only cyclic codes will be considered. In this case, it will be shown that there is always a power loss as $s \to 0$. The computations, however, will be more difficult.

The formula for expected number of bits in error, information or otherwise, will now be derived. If the zero code word is transmitted and is decoded as a code word a, there are $W(a)$ bits in error, where $W(a)$ denotes the weight of the word a. The expression for the expected number of bits in error where the code is used, β_1 say, is

$$\begin{aligned}
\beta_1 &= \sum_{\text{code word } a} W(a) \left(\sum_{\substack{\text{error vector } b \\ \text{decoding as } a}} q^{n - W(b)} \right) \\
&= \frac{1}{2^n} \sum_{\text{code word } a} W(a) \left(\sum_{b \text{ decoding to } a} 1 \right) \\
&\quad + \frac{\delta n}{2^{n-1}} \sum_{\text{code word } a} W(a) \left(\sum_{b \text{ decoding to } a} 1 \right) \\
&\quad - \frac{2\delta}{2^{n-1}} \sum_{\text{code word } a} W(a) \cdot \sum_{b \text{ decoding to } a} W(b) + 0(\delta^2)
\end{aligned}$$

Since there are 2^{n-k} correctible errors, 2^{n-k} words b decode to a, i.e.,

$$\sum_{b \text{ decoding to } a} 1 = 2^{n-k}$$

Furthermore,

$$\sum_{\text{code word } a} W(a) = n \cdot 2^{k-1}$$

by Ref 2, p 47, Prob. 3.4, or by fact that as $p \to \frac{1}{2}$, the expected number of received bits in error is $n/2$ and also is

$$\begin{aligned}
\sum_{\text{code word } a} W(a) \cdot (&\text{pr that error is made leading} \\
&\text{to } a \text{ being decoded}) = \sum_{\text{code word } a} W(a)\, \frac{1}{2^n}\, 2^{n-k} \\
&= \sum_{\text{code word } a} W(a) \cdot \frac{1}{2^k}
\end{aligned}$$

so that one now has

$$\begin{aligned}
\beta_1 &= \frac{n}{2} \\
&\quad + \delta n^2 - \frac{\delta}{2^{n-2}} \sum_{\text{code word } a} \sum_{\substack{\text{correctible} \\ \text{error } c}} W(a)\, W(a + c) + 0(\delta^2)
\end{aligned}$$

where one has used the observation that b decodes to a if and only if $b = a + c$, c being a correctible error.

The following result (Ref 3, p 19, Cor 2) shall be used: The sum of the squares of the weights of the code words in a cyclic code is $(n + 1) \cdot n \cdot 2^{k-2}$. Using the facts that the code is cyclic, shift each code word through its n shifts to the right, holding c fixed, considering the double sum as

$$\sum_{\text{code word } a} W(a) \left(\sum_{\text{correctible } c} W(a + c) \right)$$

division by n then yields the same value as the original sum. As each a is shifted, the total number of 1's in

$$\sum_c W(a+c)$$

holding c fixed, is $n[W(a)+W(c)]-2W(a)W(c)$. For two 1's add to zero every time a 1 in a meets a 1 in c. Since there are $W(c)$ 1's in c, and $W(a)$ 1's in a, twice $W(a)W(c)$ 1's are lost. Thus

$$\sum_{\substack{\text{correctible}\\ c}} W(a+c) = \frac{1}{n}\sum_{\substack{\text{correctible}\\ c}} \{n[W(a)+W(c)]-2W(a)W(c)\}$$

The double sum thus becomes

$$\sum_{\substack{\text{code word}\\ a}} W(a)\left\{\sum_{\substack{\text{correctible}\\ c}}\left[W(a)+W(c)-\frac{2}{n}W(a)W(c)\right]\right\}$$

and one derives

$$\beta_1 = \frac{n}{2} - \delta(n-2\gamma) + 0(\delta^2)$$

To obtain β, the expected number of information bits in error, one multiplies β_1 by k/n, as in *RS 36-13*, pp 32-35, to obtain the following theorem:

Theorem 3.

$$\beta = \frac{k}{2} - k\delta\left(1-\frac{2\gamma}{n}\right) + 0(\delta^2)$$

But β', the expected number of bits in error in the second case, is just $\beta' = kp' = k(\tfrac{1}{2}-\delta')$, which of course agrees with the formula of Theorem 3 for β when $n=k$, i.e., for no coding. Recalling

$$\delta = s\cdot\frac{1}{(2\pi)^{1/2}} + o(s)$$

$$\delta' = \left(\frac{n}{k}\right)^{1/2}\cdot\frac{s}{(2\pi)^{1/2}} + o(s)$$

one has

$$\beta = \frac{k}{2} - \frac{ks}{2}\left(1-\frac{2\gamma}{n}\right) + o(s)$$

$$\beta' = \frac{k}{(2\pi)^{1/2}} - k\left(\frac{n}{k}\right)^{1/2}\frac{s}{(2\pi)^{1/2}} + o(s)$$

Thus the zero-signal power gain G_1, with respect to the criterion of expected bit error probability, is given by the following theorem:

Theorem 4.

$$G_1 = \frac{k}{n}\left(1-\frac{2\gamma}{n}\right)^2$$

Since $\gamma < n/2$ and $k < n$, one has the final theorem

Theorem 5. $G_1 < 1$ for all cyclic codes.

For the Hamming codes,

$$\gamma = \frac{n}{n+1}$$

and

$$G_1 = \frac{k}{n}\left(1-\frac{2}{n+1}\right)^2 = \frac{k}{n}\left(\frac{n-1}{n+1}\right)^2$$

which agrees with a result of *RS 36-13*, pp 35–37, derived in another way.

Since the same results are likely to be approximately true even for non-cyclic codes (even ignoring the fact that most coding and decoding theory assumes a cyclic code), it appears that for communication at very low signal-to-noise ratios, say less than one, the use of error-correcting codes (coupled with bit-by-bit detection) is to be avoided.

3. Orthogonal Codes at Very Low Signal-to-Noise Ratios, E. C. Posner

Here the same problem as in the preceding section is considered, but for orthogonal codes using correlation detection instead of for error-correcting codes using bit-by-bit detection. Several results from Ref 4 shall be used, with the notation altered to conform to that of the preceding note.

An expression in s must be found for the word error probability p_W where a given orthogonal code is used on a given white Gaussian channel. Let the code have length 2^n with n information bits. First p_W shall be determined as a function of s, where s^2 is the output signal-to-noise power ratio per information bit. Then Eq (9) of Ref 4 becomes in this notation

$$p_W = 1 - \int_{v=-\infty}^{\infty}\frac{1}{(2\pi)^{1/2}}e^{-v^2/2}\left(\int_{z=-\infty}^{v+(n)^{1/2}s}\frac{1}{(2\pi)^{1/2}}e^{-z^2/2}\,dz\right)^{2^n-1}dv$$

As $s\to 0$, it can be seen from this formula, together with integration by parts, that $p_W \to 1-\tfrac{1}{2}n$, but this is also demonstrable without such computations: at zero signal, the channel transmits no information; there are 2^n words in the dictionary, and all errors are equally likely. What is really desired is the coefficient of s in the expression for p_W. Ignoring powers of s higher than the first, one has

$$p_W = 1 - \int_{v=-\infty}^{\infty}\frac{1}{(2\pi)^{1/2}}e^{-v^2/2} \times\left[\Phi(v)+(n)^{1/2}s\cdot\frac{1}{(2\pi)^{1/2}}e^{-v^2/2}\right]^{2^n-1}dv + o(s)$$

where $\Phi(v)$ denotes the cumulative normal distribution, and the integration of $o(s)$ from $-\infty$ to ∞ is justified by the integrability of

$$\frac{1}{(2\pi)^{1/2}} e^{-v^2/2}\, dv$$

from $-\infty$ to ∞. Thus, using the binomial expansion,

$$p_W = 1 - \int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2/2} [\Phi(v)]^{2^n-1}\, dv$$

$$- \int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2/2} [\Phi(v)]^{2^n-2} \cdot (2^n - 1)(n)^{1/2}$$

$$s \cdot \frac{1}{(2\pi)^{1/2}} e^{-v^2/2}\, dv + o(s)$$

The first two terms as mentioned add to $1 - 1/2^n$; letting $2^n - 1 = \nu$, one has

$$p_W = 1 - \frac{1}{2^n} - \frac{(n)^{1/2}\,\nu}{(2\pi)^{1/2}} s \int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2} [\Phi(v)]^{\nu-1}\, dv + o(s)$$

Define

$$A_\nu = \int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2} [\Phi(v)]^{\nu-1}\, dv \qquad \nu = 1, 2, \cdots$$

A_ν shall be determined for $\nu = 1, 2, 3, 4$; also, an expression asymptotic in ν as $\nu \to \infty$ shall be obtained. To this end, define

$$g_\nu(\alpha) = \int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2} [\Phi(\alpha v)]^{\nu-1}\, dv$$

which is shown to be an analytic function of α. One wishes to find $g_\nu(1) - A_\nu$. Note that $g_\nu(0) = 1/[\nu(2)^{1/2}]$ by integration by parts coupled with the result used in evaluating the constant term of $p_W(s)$; and

$$g_\nu(\infty) = \int_{v=0}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2}\, dv = \frac{1}{2(2)^{1/2}}$$

for $\nu > 1$, and $= 1/(2)^{1/2}$ for $\nu = 1$ (since g_1 is a constant function of α), since $\Phi(\alpha v)$ converges uniformly to the characteristic function of the right half line outside any fixed neighborhood of the origin.

A first order differential equation for g_ν shall now be obtained in terms of $g_{\nu-2}$ evaluated not at α but at a function of α; $g_\nu(0)$, $g_\nu(2^{1/2})$, or $g_\nu(\infty)$ can each be used to determine the constant of integration. It is easy to justify

$$\frac{d}{d\alpha} g_\nu(\alpha) = \frac{\nu - 1}{(2\pi)^{1/2}} \int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2(1+\alpha^2/2)} [\Phi(\alpha v)]^{\nu-2}\, v\, dv$$

For $\nu = 2$, one obtains

$$\frac{d}{d\alpha} g_2(\alpha) = \frac{1}{(2\pi)^{1/2}} \int_{v=-\infty}^{\infty} e^{-v^2(1+\alpha^2/2)}\, v\, dv = 0$$

Thus, $g_2(\alpha)$ is a constant equal to $1/[2(2)^{1/2}]$ since

$$g_2(0) = g_2(2^{1/2}) = g_2(\infty) = \frac{1}{2(2)^{1/2}}$$

Also, $g_1(\alpha)$ is the constant function $1/(2)^{1/2}$. Now let $\nu > 2$. In the integral for $d/d\alpha\,(g_\nu\,\alpha)$, integrate by parts letting

$$dU = \frac{1}{(2\pi)^{1/2}} e^{-v^2(1+\alpha^2/2)}\, v\, dv$$

and

$$V = [\Phi(\alpha v)]^{\nu-2}$$

Then after several intermediate steps

$$\frac{d}{d\alpha} g_\nu(\alpha) = \frac{(\nu-1)(\nu-2)}{2\pi} \frac{\alpha}{2+\alpha^2}$$

$$\int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2(1+\alpha^2)} [\Phi(\alpha v)]^{\nu-3}\, dv$$

Upon introducing $w = v(1+\alpha^2)^{1/2}$, the above expression will yield the following:

Theorem 1.

$$\frac{d}{d\alpha} g_\nu(\alpha) = \frac{(\nu-1)(\nu-2)}{2\pi} \frac{\alpha}{(2+\alpha^2)(1+\alpha^2)^{1/2}}$$

$$g_{\nu-2}\left(\frac{\alpha}{(1+\alpha^2)^{1/2}}\right) \text{ for } \nu \geqq 3$$

Now $g_3(\alpha)$ and $g_4(\alpha)$ will be determined (the higher ones are not expressible in closed form). One has

$$\frac{d}{d\alpha} g_3(\alpha) = \frac{2}{2\pi} \frac{\alpha}{(2+\alpha^2)(1+\alpha^2)^{1/2}} g_1(\alpha)$$

$$= \frac{1}{\pi} \frac{\alpha}{(2+\alpha^2)(1+\alpha^2)^{1/2}} \frac{1}{(2)^{1/2}}$$

This integrates exactly to

$$g_3(\alpha) = \frac{1}{\pi(2)^{1/2}} \arctan(1+\alpha^2)^{1/2} + C$$

since

$$g_3(0) = \frac{1}{2^{5/2}} = \frac{\pi/4}{\pi(2^{1/2})} C = 0$$

and

$$g_3(\alpha) = \frac{1}{\pi(2^{1/2})} \arctan(1+\alpha^2)^{1/2}$$

$$g_3(1) = \frac{1}{\pi(2^{1/2})} \arctan(2)^{1/2} = 0.215 \cdots$$

Similarly

$$\frac{d}{d\alpha} g_4(\alpha) = \frac{6}{2\pi} \frac{\alpha}{(2+\alpha^2)(1+\alpha^2)^{1/2}} \frac{1}{2(2^{1/2})}$$

the additive constant is, however, not zero this time, and

$$g_4(\alpha) = \frac{3}{2}(2)^{1/2} \frac{\arctan(1+\alpha^2)^{1/2}}{2\pi} - \frac{1}{12} \qquad A_4 = .144\cdots$$

Now an asymptotic expression for A_ν as $\nu \to \infty$ shall be found. One obtains by integrating by parts

$$A_\nu = \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \int_{v=-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} e^{-v^2/2} (\nu+1) [\Phi(v)]^\nu v\, dv$$

$$= \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \lim_{T\to\infty} \int_{v=-T}^{T} v\, d[\Phi(v)]^{\nu+1}$$

$$= \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \lim_{T\to\infty} \left\{ T - \int_{v=-T}^{T} [\Phi(v)]^{\nu+1} dv \right\}$$

since $\Phi(T) \to 1$ rapidly as $T \to \infty$, and $\Phi(-T)$ approaches 0

$$A_\nu = \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \lim_{T\to\infty} \left\{ \int_{v=0}^{T} [1-\Phi(v)]^{\nu+1} + [1-\Phi(v)]^{\nu+1} \right\} dv$$

using $\Phi(-v) = 1 - \Phi(v)$

$$A_\nu = \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \int_{v=0}^{\infty} \{1 - [\Phi(v)]^{\nu+1} + [1-\Phi(v)]^{\nu+1}\}\, dv$$

(The use of the intermediate limit expression was necessary due to the divergence of certain integrals.) The contribution to A_ν from

$$\int_{v=0}^{\infty} [1-\Phi(v)]^{\nu+1}\, dv$$

is readily seen to be arbitrarily small in comparison to the contribution from

$$\int_{v=0}^{\infty} \{1 - [\Phi(v)]^{\nu+1}\}\, dv$$

thus one has the asymptotic expression

$$A_\nu \sim \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \int_{v=0}^{\infty} \{1 - [\Phi(v)]^{\nu+1}\}\, dv$$

Let $0 < \delta < 1$ be arbitrary, and consider the set of $v \geqq 0$ with $1 - [\Phi(v)]^{\nu+1} > \delta$, an interval to the right of 0 with right hand endpoint v_0 given by $\Phi(v_0) = \delta^{1/\nu+1}$ (where ν is so large that $\delta < 1 - 1/2^{\nu+1}$). Let δ be close to 1 and ν be large. Then v_0 must be large; furthermore,

$$1 - \Phi(v_0) \sim \frac{e^{-v_0^2/2}}{(2\pi)^{1/2} v_0}$$

by Ref 5, p 106, Lemma 2, and

$$\delta^{1/\nu+1} = e^{\log\delta/\nu+1} = 1 + \frac{\log\delta}{\nu+1} + o\left(\frac{\delta}{\nu+1}\right)$$

as $\delta \to 1$ and $\nu \to \infty$. Thus an asymptotic formula for v_0 is given by

$$1 - \frac{1}{(2\pi)^{1/2} v_0} e^{-v_0^2/2} = 1 + \frac{\log\delta}{\nu+1}$$

$$\frac{1}{(2\pi)^{1/2} v_0} e^{-v_0^2/2} = \frac{\log\delta^{-1}}{\nu+1}$$

$$-\frac{v_0^2}{2} - \log[(2\pi)^{1/2} v_0] = \log\log\delta^{-1} - \log(\nu+1)$$

and asymptotically

$$\frac{v_0^2}{2} \sim \log(\nu+1), \ (\delta \text{ fixed}), \ v_0 \sim [2\log(\nu+1)]^{1/2}$$

Write

$$A_\nu \sim \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \int_{v=0}^{v_0} \{1 - [\Phi(v)]^{\nu+1}\}\, dv$$

$$+ \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \int_{v=v_0}^{\infty} \{1 - [\Phi(v)]^{\nu+1}\}\, dv$$

The second integral may be shown to be arbitrarily small with respect to the first as $\nu \to \infty$, δ fixed, so that

$$A_\nu \sim \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} \int_{v=0}^{v_0} \{1 - [\Phi(v)]^{\nu+1}\}\, dv$$

On $[0, v_0]$, however,

$$1 - \frac{1}{2^{\nu+1}} \geqq 1 - [\Phi(v)]^{\nu+1} \geqq \delta$$

so

$$\int_{v=0}^{v_0} \{1 - [\Phi(v)]^{\nu+1}\}\, dv$$

lies between $[1 - 1/(2^{\nu+1})]\, v_0$ and δv_0. That is, for each fixed δ with $0 < \delta < 1$ and for large ν with $\delta < 1 - 1/2^{\nu+1}$,

$$A_\nu \left[\frac{(2\pi)^{1/2} v_0}{(\nu+1)(\nu)} \right]^{-1}$$

lies between 1 and δ. Since δ is arbitrary,

$$A_\nu \sim \frac{(2\pi)^{1/2} v_0}{(\nu+1)(\nu)}$$

Thus Theorem 2 has been proved.

Theorem 2.

$$A_\nu \sim \frac{(2\pi)^{1/2}}{(\nu+1)(\nu)} [2\log(\nu+1)]^{1/2}$$

As a check, let $\nu = 63$; then $A_{63} = 0.0015$ by integration using Simpson's rule; the asymptotic formula gives 0.0012.

Getting back to p_W,

$$p_W = 1 - \frac{1}{2^n} - \frac{(n)^{1/2}\,\nu}{(2\pi)^{1/2}}\,s\,A_\nu$$

and

$$\left(1 - \frac{1}{2^n}\right) - p_W \sim \frac{(n)^{1/2}(2^n - 1)}{(2\pi)^{1/2}} \frac{(2\pi)^{1/2}}{(2^n)(2^n - 1)} (2\log 2^n)^{1/2}\,s$$

$$= \frac{n}{2^{n-1/2}} (\log 2)^{1/2}\,s$$

as $s \to 0$, if n is large. Thus expected bit error probability, equal to $[2^{n-1}/(2^n - 1)]\,p_W$ from Ref 4, becomes

$$\left(1 - \frac{1}{2^n}\right) - \frac{2^n - 1}{2^{n-1}}\,p_B \sim \frac{n}{2^{n-1/2}} (\log 2)^{1/2}\,s$$

so that

$$\tfrac{1}{2} - p_B \sim \frac{2^{n-1}}{2^n - 1} \frac{n}{2^{n-1/2}} (\log 2)^{1/2}\,s = \frac{1}{(2)^{1/2}} \frac{n}{2^n - 1} (\log 2)^{1/2}\,s$$

Without coding, the bit error probability is, as in the above summary, $\tfrac{1}{2} - p \sim s/(2\pi)^{1/2}$. Thus for large n there is a power loss as $s \to 0$ if the criterion is expected bit error probability.

Theorem 3.
With p_B as the criterion, the loss factor is asymptotic in n to $(\pi \log 2) \cdot (n/2^n - 1)^2$.

This is to be expected since orthogonal codes are especially bad from the standpoint of bit error probability; in fact, when an error is made, all error patterns are equally likely (Ref 1, Sec V).

If p_W is the criterion, in the uncoded case the word error probability is derived from

$$(1 - p)^n = [\tfrac{1}{2} + (\tfrac{1}{2} - p)]^n = (\tfrac{1}{2})^n + n\,(\tfrac{1}{2})^{n-1} \frac{s}{(2\pi)^{1/2}} + o(s)$$

$$1 - [\tfrac{1}{2} + (\tfrac{1}{2} - p)]^n = 1 - (\tfrac{1}{2})^n - n\,(\tfrac{1}{2})^{n-1} \frac{s}{(2\pi)^{1/2}} + o(s)$$

and in the coded case

$$1 - \frac{1}{2^n} - p_W \sim \frac{n}{2^{n-1/2}} (\log 2)^{1/2}\,s$$

for fixed large n, as $s \to 0$.

Theorem 4.
The power factor with p_W as the criterion is a gain approaching $\pi \log 2 = 2.18$ (or 3.4 db).

In Fig 1, using computed values of A_ν, this power factor is plotted with p_W as the criterion for $n = 2\,(1)\,6$, using the exact formula $\nu^2 A_\nu^2 4^{n-1}/n$. Note that there is a loss for $n = 1, 2$, corresponding to the fact that not coding when $n = 1, 2$ is exactly the same as using a biorthogonal (rather than just orthogonal) code.

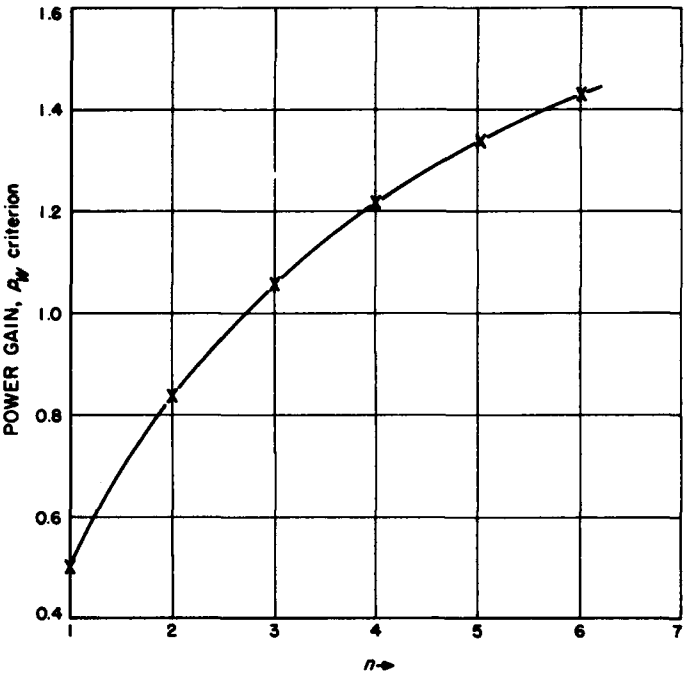


Figure 1. Power gain versus n

Thus, if bit error probability is the correct requirement for a given telemetry system, one concludes that a system using orthogonal codes should not be used for output signal-to-noise ratios less than one and approaching zero; a preprogrammed command or a command from the ground should be given to override the coding system and return to a system using no coding when the output signal-to-noise ratio drops below one.

4. Error Probabilities for Optimal Codes in a Gaussian Channel, T. Kailath

In a paper in 1959 Shannon (Ref 6) derived several lower and upper bounds on the error probability, $P_{e\,\mathrm{opt}}$, for an optimal code signalling over a channel perturbed by additive white noise. Although quite ingenious, the derivation was rather long and complex, but the final results were remarkably elegant. Much current work in Information Theory is concerned with obtaining results of this sort, and it seems desirable to obtain as much of an understanding of the results and the methods of obtaining them as possible. In 1956 Shannon (Ref 7) described a general and powerful technique for obtain-

ing such results. These techniques were delineated earlier in a paper by Cramer (Ref 8). In this report we shall give a brief discussion of the essentials of this method and show how it can be used to obtain, in a special case, the results of Shannon. However, the derivation is rather more transparent now.

The mathematical problem is to evaluate asymptotically, as $n \to \infty$, the distribution

$$Q(\alpha) = Pr\left[\sum_{i=1}^{n} \frac{n_i^2}{[m + (nP)^{1/2}]^2} < \tan^2 \alpha\right] \tag{1}$$

where m, $n_i\,(i = 1, \cdots n)$ are independent Gaussian random variables with mean zero and variance N. It is easily shown that the mean of the distribution occurs at $\alpha = \alpha_0 = \cot^{-1}(P/N)^{1/2}$.

Now for values of α *near the mean* α_0, the Central Limit Theorem can be used to obtain a Gaussian approximation to $Q(\alpha)$. However, this approximation breaks down for α not close to α_0, i.e., for values on the tail of the distribution. Shannon (following Cramer and Chernov) has described a technique for evaluating probabilities in this case. In the next section we outline the philosophy behind this method and then apply it to $Q(\alpha)$ in Sect 2.

a. The method of tilting. For convenience let us rewrite Eq (1) as

$$Q(\alpha) = Pr\left\{y = \sum_{i=1}^{n} y_i < n\rho\right\} \tag{2}$$

where $\beta = \tan^2 \alpha$ and the $y_i = n_i^2/[(P)^{1/2} + m/(n)^{1/2}]^2$ are independent, identically distributed, random variables. Near the asymptotic mean value, nN/ρ, of $y = \Sigma y_i$, we can use for $Q(\alpha)$ the Gaussian approximation previously discussed. For values away from the mean we "tilt" the distribution function of y by multiplying it by a "tilting" factor $e^{sy}, s > 0$. This has the effect (Fig 2) of moving the distribution over towards positive values of y. By this means we can move the mean of the distribution over to the point $n\beta$ at which $Q(\alpha)$ is to be evaluated, and now for the tilted variable we can apply the Gaussian approximation quite satisfactorily because we shall be working near the mean of this new (tilted) variable.

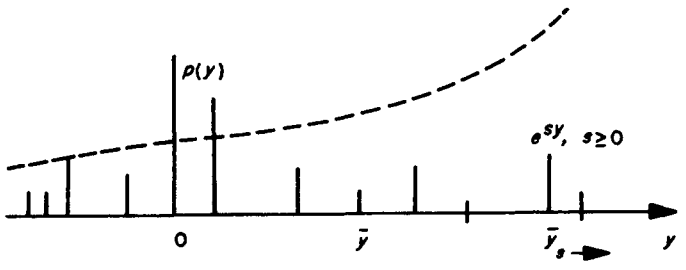


Figure 2. Tilting the distribution p (y)

To carry out the above procedure we first obtain the "tilted" distribution

$$dF_s(\beta) = \frac{e^{s\beta}\, dF(\beta)}{\int_{-\infty}^{\infty} e^{s\beta}\, dF(\beta)} \tag{3}$$

where $F(\beta) = Pr\{y \leq \beta\}$; $F_s(\beta) = Pr\{y_s \leq \beta\}$; and s is some at present unspecified real number. The moments of $dF(y_s)$ are easily obtained. The mean of the tilted variable, y_s, is given by

$$\bar{y}_s = \int \beta\, dF_s(\beta) = \frac{\int \beta e^{s\beta}\, dF(\beta)}{\int e^{s\beta}\, dF(\beta)} \tag{4}$$

We notice that the denominator of Eq (4) is the moment generating function of the original (untilted) variable y and may be denoted by $g(s)$. [$g(s)$ may only exist for certain values of s and this limits the range of s.] Then the mean of the tilted variable, y_s, is given by

$$\bar{y}_s = \frac{g'(s)}{g(s)} = \frac{d\mu(s)}{ds} = \mu'(s) \tag{5}$$

where the prime denotes differentiation with respect to s, $g'(s) = d\,g(s)/ds$, and where

$$\mu(s) = \ln g(s) = \ln\left[\int_{-\infty}^{\infty} e^{s\beta}\, dF(\beta)\right] \tag{5'}$$

$\mu(s)$ is usually known as the log-moment-generating function or cumulant-generating function. Similarly the variance of y_s is found to be

$$\bar{y}_s^2 - (\bar{y}_s)^2 = \frac{g''(s)}{g(s)} - \left[\frac{g'(s)}{g(s)}\right]^2 = \frac{d^2\mu(s)}{ds^2} = \mu''(s) \tag{6}$$

The moments of the untilted variable y are given by

$$y = g'(0) \text{ and } \bar{y}^2 - (\bar{y})^2 = g''(0) - [g'(0)]^2 \tag{7}$$

Therefore

$$\bar{y}_s = \mu'(s) \begin{cases} > \bar{y} & s > 0 \\ = \bar{y} & s = 0 \\ < \bar{y} & s < 0 \end{cases} \tag{8}$$

Therefore by varying s [within the range permitted, viz, for all s such that $g(s)$ exists] we can vary the location of the mean, $\bar{y}_s$, of the tilted variable (Fig 2). We also notice that, in terms of $\mu(s)$, we can rewrite Eq (3) as

$$dF_s(\beta) = e^{[s\beta - \mu(s)]}\, dF(\beta) \tag{9}$$

or

$$dF(\beta) = e^{-[s\beta - \mu(s)]}\, dF_s(\beta) \tag{10}$$

By choosing s suitably we can make $F_s(\beta)$ have its mean at the desired point on the tail of the distribution $F(\beta)$. We can now validly apply the central limit theorem to $F_s(\beta)$ and then use Eq (10) to obtain an approximation for the tails of the distribution $F(\beta)$.

b. Result. The detailed application of the above ideas to our problem is too lengthy for discussion here. We quote only the final result. Asymptotically

$$Q(\alpha) \sim \frac{1}{[2\pi n\mu''(0)]^{1/2}} e^{-n[s\mu'(s)-\mu(s)]} \tag{11}$$

where as defined above

$$\mu(s) = \text{log-moment generating function of } \frac{n_1^2}{[m + (nP)^{1/2}]^2}$$

and s is the solution of the equation

$$\mu'(s) = \tan^2\alpha \tag{12}$$

The coefficient of $-n$ in the exponent of Eq (11) is the quantity of most interest. It has been called the reliability and will be denoted by $E_L(\alpha)$.

Now in beginning to explicitly obtain $E_L(\alpha)$ in terms of α, P, N, we meet with an obstacle because $\tilde{\mu}(s)$ is quite difficult to compute. We therefore make the simplifying assumption that P/N is very large so that we need only to compute $\mu(s)$ for n_1^2/P, which is a chi-squared random variable. For this case we can show that

$$\mu(s) = -\tfrac{1}{2}\ln\left(1 - 2s\frac{N}{P}\right) \qquad s \leq \frac{P}{2N} \tag{13}$$

$$\mu'(s) = \frac{N/P}{1 - 2s(N/P)} \tag{14}$$

The value of s to be used in Eq (11) is given by Eq (12).

$$\mu'(s) = \frac{N/P}{1 - 2s(N/P)} = \tan^2\alpha \tag{15}$$

which yields

$$s = \frac{1 - (N/P)\cot^2\alpha}{2N/P} \leq \frac{P}{2N} \tag{16}$$

Therefore we can find a suitable value for s, viz, an s for which $\mu(s)$ exists. With this value for s, we have

$$E_L(\alpha) = \frac{P}{2N}\tan^2\alpha - \tfrac{1}{2}\ln\frac{P}{N} + \ln(\cot\alpha) - \tfrac{1}{2} \tag{17}$$

$$E_L'(\alpha) = \frac{P}{N}\tan\alpha\left[\sec^2\alpha - \frac{N}{P}\csc^2\alpha\right] \tag{18}$$

The general shape of $E_L(\alpha)$ is shown in Fig 3. It is easy to show that

$$E_L(\alpha_0) = 0 = E_L'(\alpha_0) \tag{19}$$

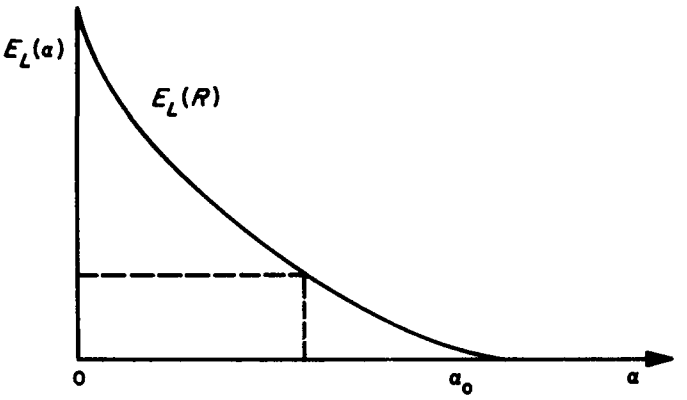


Figure 3. The reliability curve $E_L(\alpha)$

If we examine the behavior of $E_L(\alpha)$ near $\alpha = \alpha_0$, we get, by a Taylor series expansion about α_1,

$$E_L(\alpha) = \frac{P}{N}(\alpha - \alpha_0)^2 \qquad \alpha \doteq \alpha_0 \tag{20}$$

It can be shown that Eq (20) coincides with the expression that would be found by direct use of the Central Limit Theorem near the mean.

It can also be verified that, under the assumption of large signal-to-noise ratios (large P/N), the results we have obtained agree with those derived by Shannon. However, Shannon also has results valid for all P/N. It is possible that these general results may also be obtained by the tilting method. A basic connection between Shannon's original method and the tilting method may hinge on the fact that both methods are essentially related to saddle-point techniques for evaluating integrals.

5. An Error-Correction Procedure for a Class of Bose-Chaudhuri Codes, A. L. Duquette and G. Solomon

Bose-Chaudhuri codes are cyclic codes that are best defined in terms of the roots of the generator polynomial. The symbols are in general assumed to be elements of $GF(p^n)$; however, throughout this paper they will be assumed in $GF(2)$. Let m_0 be any non-negative integer and α be any non-zero element of $GF(2^m)$. Then the code consisting of all vectors $[f(X)]$ over $GF(2)$ for which $\alpha^{m_0}, \alpha^{m_0+1}, \alpha^{m_0+2}, \cdots, \alpha^{m_0+d-2}$ $(d_0 - 2 < e =$ order of $\alpha)$ are roots of $f(X)$ is a Bose-Chaudhuri code. The fundamental theorem (Ref 11, p. 162) states: The Bose-Chaudhuri code, for which $[f(X)]$ is a code vector if and only if $f(X)$ has $\alpha^{m_0}, \alpha^{m_0+1}, \cdots, \alpha^{m_0+d-2}$ as roots, has minimum distance at least d_0. There are cases, however, where the Bose-Chaudhuri lower bound d_0 does not coincide

with the minimum non-zero weight d of the code vectors (Ref 9); e.g., for the (17, 9), (23, 12), and (47, 24) Bose-Chaudhuri codes d_0 is equal to 3, 5, and 5, respectively, while d is equal to 5, 7, and 11. At the present time there appears to be only one manageable error-correction procedure known for Bose-Chaudhuri codes over $GF(2)$, that given by Peterson (Ref 10). This procedure is based on the roots of the generator polynomial; and, as a result, a Bose-Chaudhuri code is t-error correcting if, and only if, $\alpha, \alpha^3, \cdots, \alpha^{2t-1}$ are roots of a code polynomial; clearly, $d_0 = 2t + 1$. Hence, this procedure corrects only patterns of t or fewer errors where the minimum distance between code words is assumed to be the Bose-Chaudhuri lower bound, even though some codes are capable of correcting all patterns of s errors where $s > t$.

In this paper an error-correction procedure will be given that will apply to a large class of Bose-Chaudhuri codes and that will correct *all* error patterns that are correctible by a given code. This procedure uses the field properties of $GF(2^m)$ in a deep way; in essence, the geometry of the n-dimensional space $V_n[GF(2)]$ comes into play here. (The Peterson decoding procedure, however, is more general because it does not use the special algebraic properties of the individual codes.) It will be shown that the perfect Hamming (7, 4), the quasi-perfect (17, 9), the perfect Golay (23, 12), and the (47, 24) Bose-Chaudhuri codes are members of this class. The codes to which attention shall be restricted have previously been studied by Solomon and Mattson (Ref 9), and now a few results from that work will be utilized. Let n be an odd prime written as p, where $p = 2h + 1$, such that

$$x^p + 1 = (x + 1 \cdot f_0(x) f_1(x)$$

where $f_0(x)$ and $f_1(x)$ are irreducible over $GF(2)$. Then it was shown in Ref 9 that

$$f_0(x) = x^h + 0 \cdot x^{h-1} + \cdots + 1$$

and

$$f_1(x) = x^h + x^{h-1} + \cdots + 1$$

with suitable choice of the indexing 0, 1. Furthermore, one has necessarily (but not sufficiently) that $p \equiv \pm 1 \pmod 8$. In what follows $f_1(x)$ will be the generator polynomial of the Bose-Chaudhuri codes and the code polynomials will be "reversed." [Instead of Peterson's (Ref 11, p. 137) code polynomial $f(x)$, $x^p f(1/x)$ shall be used.] Choose β to be a root of $f_1(x)$ if $p \equiv 1 \pmod 8$, a root of $f_0(x)$ if $p \equiv -1 \pmod 8$. For each code vector $a = (a_0, a_1, \cdots, a_{p-1})$, there exists a unique polynomial $g_a(x)$ with coefficients in $GF(2_h)$ such that $a_i = g_a(\ ^i)$ for $i = 0, 1, \cdots, p - 1$. Moreover,

$$g_a(x) = \begin{cases} c_0 + c_1 \cdot x^{r_1} + c \cdot x^{r_2} + \cdots + c_h x^{r_h} & p \equiv -1 \pmod 8 \\ c_0 + c_1 \cdot x^{s_1} + c_2 x^{s_2} + \cdots + c_h x^{s_h} & p \equiv +1 \pmod 8 \end{cases}$$

where the r's form the set of least positive quadratic residues mod p and the s's the set of least positive quadratic non-residues mod p. The coefficients $c_0, c_1, \cdots, c_h$ of $g_a(x)$ are given by the formulas

$$c_0 = \sum_{i=0}^{p-1} a_i \quad [c_0 \in GF(2)]$$

$$c_j = \sum_{i=0}^{p-1} a_i \beta^{-i e_j} \quad [c_j \in GF(2^h)] \qquad j = 1, \cdots, h$$

where

$$e_j = r_j \text{ for all } j = 1, \cdots, h \qquad p \equiv -1 \pmod 8$$

$$e_j = s_j \text{ for all } j = 1, \cdots, h \qquad p \equiv +1 \pmod 8$$

In particular

$$c_1^2 = c_2, c_2^2 = c_3, \cdots, c_h^2 = c_1$$

Also, more generally (using n again instead of p), utilizing the results of Solomon (Ref 12) one can and does associate with each vector $a = (a_0, a_1, \cdots, a_{n-1})$ in $V_n[GF(2)]$ a unique polynomial $g_a(x)$ of degree $\leq n - 1$ such that $g_a(\beta^i) = a_i$ where β is a positive nth root of unity. Assuming that

$$g_a(x) = \sum_{i=0}^{n-1} c_i x^i$$

the c's are given by the formula

$$c_k = \sum_{i=0}^{n-1} a_i (\beta^i)^{-k}$$

Since $c_2 = c_1^2$ $c_4 = c_1^4$ etc., it is clear that the coefficients of $g_a(x)$ will be determined by a number of independent parameters equal to the number of irreducible factors in $h(x)$, where $h(x)$ is the polynomial which divides $x^n + 1$ and is such that the coordinates of the vector a form a linear recursive sequence determined by its coefficients. Henceforth the additive isomorphism which exists between a vector $a = (a_0, a_1, \cdots, a_{n-1})$ and its associated polynomial $g_a(x)$ will be written $a \leftrightarrow (c_0, c, d, e, \cdots)$. For the special class of Bose-Chaudhuri codes where $n = p$, a prime $\equiv \pm 1 \pmod 8$, one shall always have $a \leftrightarrow (c_0, c, d)$ for any vector a in $V_p[GF(2)]$; in particular, for a code vector one shall have

$$a \leftrightarrow (c_0, c, 0) \text{ if } p \equiv -1 \pmod 8$$

$$a \leftrightarrow (c_0, 0, d) \text{ if } p \equiv +1 \pmod 8$$

To correct an error in a given received vector is tantamount to finding the coset leader of the coset in

which the received vector lies, since the sum of these two vectors is the correct code vector. We shall now employ the Golay (23, 12) perfect, triple error-correcting code to demonstrate the error-correction procedure, although any code of the above class of Bose-Chaudhuri codes of the form $(p, p+1)/2$ could have been employed. Since $23 \equiv -1 \pmod 8$ we choose β to be a primitive 23rd root of unity which is a root of

$$f_0(x) = x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1$$

If two vectors are in the same coset, then the parameter d must be equal for both vectors since their difference is in the code. Given a received vector a, calculate d immediately by the formula

$$d = \sum_{i=0}^{22} a_i (\beta^i)^{-5} = \sum_{i=1}^{23} \bar{a}_i \beta^{5i} \pmod{23}$$

where $a_i = a_{17-i}$. The coset leader must have the same d as the received vector when d is expressed in powers (≤ 10) of β, for otherwise the minimal polynomial of β would have degree <11 which is contradictory. The coset leader will obviously be found if one can express d in the minimum possible number of powers (from 1 to 23) of β. [If d reduces to one, two, or three powers (≤ 10) of β, one immediately has the coset leader.] In order to achieve this, first assume that all powers are ≥ 11 when d is expressed in minimum form. If d is of this form it will be called "virtuous," otherwise "unvirtuous." Therefore

$$d = X_1\beta^{x_1} + X_2\beta^{x_2} + X_3\beta^{x_3} = b_{10}\beta^{10} + b_9\beta^9 + \cdots + b_1\beta + b_0$$

where the b_i are given and in $GF(2)$; the X_i are unknown and in $GF(2)$; the x_i are unknown and $11 \leq x_i \leq 22$. Clearly

$$\begin{aligned} &X_1\beta^{x_1} + X_2\beta^{x_2} + X_3\beta^{x_3} + b_{10}\beta^{10} + b_9\beta^9 + \cdots + b_1\beta + b_0 \\ &\quad = (a_{11}\beta^{11} + a_{10}\beta^{10} + \cdots + a_1\beta + a_0) \\ &\qquad \cdot (\beta^{11} + \beta^9 + \beta^7 + \beta^6 + \beta^5 + \beta + 1) \end{aligned}$$

since $x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1$ is the minimal polynomial of β. Expanding one has

$$\begin{aligned} (*)\ &X_1\beta^{x_1} + X_2\beta^{x_2} + X_3\beta^{x_3} + b_{10}\beta^{10} + b_9\beta^9 + \cdots + b_1\beta + b_0 \\ &= a_{11}\beta^{22} + a_{10}\beta^{21} + (a_{11} + a_9)\beta^{20} + (a_{10} + a_8)\beta^{19} \\ &\quad + (a_{11} + a_9 + a_7)\beta^{18} + (a_{11} + a_{10} + a_8 + a_6)\beta^{17} \\ &\quad + (a_{11} + a_{10} + a_9 + a_7 + a_5)\beta^{16} \\ &\quad + (a_{10} + a_9 + a_8 + a_6 + a_4)\beta^{15} \\ &\quad + (a_9 + a_8 + a_7 + a_5 + a_3)\beta^{14} \\ &\quad + (a_8 + a_7 + a_6 + a_4 + a_2)\beta^{13} \\ &\quad + (a_{11} + a_7 + a_6 + a_5 + a_3 + a_1)\beta^{12} \\ &\quad + (a_{11} + a_{10} + a_6 + a_5 + a_4 + a_2 + a_0)\beta^{11} \\ &\quad + (a_{10} + a_9 + a_5 + a_4 + a_3 + a_1)\beta^{10} \\ &\quad + (a_9 + a_8 + a_4 + a_3 + a_2 + a_0)\beta^9 \\ &\quad + (a_8 + a_7 + a_3 + a_2 + a_1)\beta^8 \\ &\quad + (a_7 + a_6 + a_2 + a_1 + a_0)\beta^7 \\ &\quad + (a_6 + a_5 + a_1 + a_0)\beta^6 + (a_5 + a_4 + a_0)\beta^5 \\ &\quad + (a_4 + a_3)\beta^4 + (a_3 + a_2)\beta^3 + (a_2 + a_1)\beta^2 \\ &\quad + (a_1 + a_0)\beta + a_0 \end{aligned}$$

It is now apparent that the a_i for $0 \leq i \leq 10$ can be determined in turn simply by equating coefficients. There are two possibilities for a_{11}: 0 and 1. Hence the coefficients of β^{11} through β^{22} can also be determined. Once these coefficients are determined for $a_{11} = 0$, the coefficients may easily be determined for $a_{11} = 1$ by changing the coefficients of β^{11}, β^{12}, β^{16}, β^{17}, β^{18}, β^{20}, β^{22} and leaving the other coefficients the same. If d is virtuous, one of the two possibilities will give a sum of one, two, or three powers (all ≥ 11) of β since the Golay code is triple error-correcting; otherwise, for an unvirtuous d, both possibilities will yield sums of four or more powers (all ≥ 11) of β. Note that all single errors are now corrected.

The unvirtuous d's consisting of two or three powers of β are now divided into two categories: (1) "degenerate" [consisting of two powers (>10) of β] and (2) "undegenerate" [consisting of one power (>10) of β]. One finds the undegenerate d by working from left to right on the right side of (*), assigning the value 1 to a single coefficient of a power (≥ 11) of β and assigning the value 0 to all the other coefficients of powers (≥ 11) of β. The coefficients of the powers (≤ 10) of β will always be uniquely determined, and to them one adds the given value of d which is already in powers (≤ 10) of β. Any undegenerate d will be found by at most 12 repetitions of this process since some power (≥ 11) of β plus the given value of d will yield a minimum number (one or two) of powers (≤ 10) of β. Observe that any single or double error has now been corrected as well as any undegenerate triple error.

The case of the degenerate triple error can easily be transformed into the case of the undegenerate triple error by multiplying the given d by β^{11}. If this does not yield the minimum, one realizes that the minimum d contained a β^{11} term and therefore multiplies the given d by β^{12}. Clearly, if the minimum d is still not attained, one knows that the minimum d is equal to $\beta^{11} + \beta^{22}$ plus some power (≤ 10) of β which can be immediately determined. One remarks in passing that, in multiplying the unvirtuous d by either β^{11} or β^{12}, one can always choose one of the two forms of d given by $a_{11} = 0$ and $a_{11} = 1$ in such a way that the transformed d consists only of powers (≤ 10) of β; hence, no reductions mod β^{11} are needed.

Since the Golay code does not correct any quadruple errors, one need not carry the procedure any further.

The same technique would be used for correcting all patterns of four or more errors. One should always work from left to right on the right side of (*) for any unvirtuous undegenerate error, where in general an undegenerate error is one having no greater number of powers ≥ 11 than the number of powers ≤ 10. Also, the same transformations as above would be employed to convert degenerate d's to undegenerate d's. Various simplifying modifications of the above procedure could undoubtedly be made depending upon the known algebraic peculiarities of the code under consideration. Since in the above work one actually calculates the powers (>10) of β for each unv[illegible]s d, one could avoid this by originally storing t[illegible]wers of β. If the code corrects quadruple errors, etc., [illegible]e could very well store the sums of powers (>10) of β added two at a time, and so forth.

Any cyclic code whose generator polynomial[a] $g(x)$ consists of one irreducible polynomial would have the same set of parameters associated with its vectors as the n-dimensional vector space containing it, with the restriction that one of the parameters associated with a given vector is zero if, and only if, a vector is in the code. Therefore the first part of the error-correction procedure described above for the special class of Bose-Chaudhuri codes would be applicable to this large class of cyclic codes. The sums of higher powers of β, for a suitable choice of β (Ref 12), would then be taken one at a time, two at a time, etc. Whether a simplification of the type utilized above for the special class of Bose-Chaudhuri codes considered in the paper exists for these cyclic codes is still an open question which is now being investigated.

B. Combinatorial Coding

M. Hall, J. J. Stiffler, S. W. Golomb, S. Rosen, J. Kohler, and H. Rumsey

1. The Nonexistence of Comma-Free Binary Orthogonal Sets of Order 8, *M. Hall, Jr., and J. J. Stiffler*

Consider an 8×8 array of *ones* and *zeros* with the property that the sum (term-by-term addition modulo 2) of any two rows of this array contain exactly four *ones*. Such an array is called an orthogonal set. If this set of rows has the additional property that no overlap formed from any two rows is a third row or the complement of a third row, it is called strongly comma-free of index *one*. This paper shows that no such set exists.

First observe that if a set of 8 rows, of 8 binary digits each, is to be comma-free, all cyclic permutations of all 8 rows and their complements must be distinct. To verify this, let $x_i(k)$ represent the ith row of the set permuted k positions to the left. Then if $x_i(k) = x_j(r)$, clearly $x_i(0) = x_j(r-k)$, in violation of the condition for comma freedom. Thus 2^7 distinct 8-tuples must be generated by cyclic permutations and complementation of the eight 8-tuples (or vectors) in the set. Secondly, all of these 8-tuples must be of the same parity (i.e., all must contain an even number of ones or all must contain an odd number of ones) since $x_i(0)$ and $x_j(0)$ must differ in an even number of positions for all $i \neq j$. Since there are only 2^7 even and 2^7 odd binary vectors of length 8, then all even or all odd vectors must be represented. But the even vectors cannot be used since, for example, the vector with components all of which are zero is not distinct from any of its cyclic permutations. Clearly then, if a set with the desired property is to exist, the 8-tuples containing an odd number of *ones* must divide into eight equivalence classes of 16 members each under the operation of cyclic permutation and vector complementation. This, in fact, does occur, and it may be verified that the following 8 vectors are representatives of different equivalence classes:

$$\begin{aligned} \{x_1\} &= \{00001110\} \\ \{x_2\} &= \{01100111\} \\ \{x_3\} &= \{00110010\} \\ \{x_4\} &= \{00111101\} \\ \{x_5\} &= \{00000001\} \\ \{x_6\} &= \{01101000\} \\ \{x_7\} &= \{01011011\} \\ \{x_8\} &= \{01010100\} \end{aligned} \tag{1}$$

Thus, a comma-free set must contain one, and only one, member of each class.

We now enumerate the number of orthogonal sets which can be constructed subject to this constraint. Since the orthogonality relationship between two vectors is invariant under complementation of either or both of the vectors, it is sufficient to consider only the sets that can be generated by excluding one of the two complements from consideration. It is noted that for every set

[a]Here the definition of generator polynomial found in Ref 11, which is distinct from that given in Ref 9, is used.

thus obtained, there are $2^8 - 1$ additional sets which result from complementing the rows in all possible ways.

We begin by observing that there is only one member of $\{x_3\}$ which is orthogonal to any given member of $\{x_2\}$. That is, the following two vectors represent the class of all orthogonal pairs from $\{x_2\}$ and $\{x_3\}$

$$\begin{matrix} 01100111 \\ 00110010 \end{matrix} \qquad (2)$$

Thus there is only one element for $\{x_3\}$ which can be used after one of the elements of $\{x_2\}$ has been chosen. Now choose 00001110 from $\{x_1\}$ and consider those cyclic permutations of the pair in Expression (2) which are orthogonal to it. It may be verified that the following three are the only triples with this property:

$$\begin{matrix} 00001110 \\ 01100111 \\ 00110010 \end{matrix} \qquad \begin{matrix} 00001110 \\ 01110110 \\ 00100011 \end{matrix} \qquad (3) \qquad \begin{matrix} 00001110 \\ 10011101 \\ 11001000 \end{matrix}$$

We now observe that no element of $\{x_4\}$ is orthogonal to all vectors of the last triple, only 00111101 is orthogonal to the first triple, and only 00111101 to the second triple.

It has been shown (*RS 36-9*) that if any one column of an 8×8 orthogonal set contains only *zeros* (or only *ones*), the set is a group coset. Consider a set containing four known vectors and four unknown vectors. Complement the necessary rows so that the first column contains only *zeros*. This is then a group coset. Label the known vectors $y_1 = x_1 + c$, $y_2 = x_2 + c$, $y_3 = x_3 + c$, $y_4 = x_4 + c$ where c is the vector added to every group element to make it a coset. Observe that the vectors $y_1 + y_2$, $y_2 + y_3$, and $y_3 + y_4$ are three linearly independent elements of the group if the y's are all distinct and if the sum of all four vectors is not the vector containing all *zeros*. Thus, if these conditions are satisfied (as they are by the two sets of four vectors discussed above) the group, if it exists, and hence the coset, is completely determined by these vectors. It may be verified that these two sets do indeed generate a group coset in the manner described, and one in which each of the above equivalence classes is represented, as required. Thus there are only two groups with the necessary properties. There are eight cyclic permutations and 2^8 row complementations of each group. However, only half the row complementations need be investigated since half the sets are simply complements of the other half. Similarly only half the permutations need be considered since, for example,

$$01110000$$

is just 00001110 read backwards, and it is apparent that if there is any comma-free set containing the first vector, its reflection (i.e., the set obtained by reading the given set from right to left) is also comma-free. Thus for every comma-free set containing 01110000 there is also one containing 00001110. This is obviously also true for the remaining pairs:

$$\begin{matrix} 11100000 \\ 00000111 \end{matrix} \qquad \begin{matrix} 11000001 \\ 10000011 \end{matrix} \qquad \begin{matrix} 00011100 \\ 00111000 \end{matrix}$$

There are then 8 cosets and consequently 8×2^7 sets which need to be investigated. This is a sizeable reduction from the 122,880 distinct orthogonal 8×8 sets.

The eight cosets may be represented by $G^{(0)}$, $G^{(1)}$, $G^{(-1)}$, $G^{(-2)}$, $H^{(0)}$, $H^{(1)}$, $H^{(-1)}$, $H^{(-2)}$

where

$$G^{(0)} = \begin{matrix} 00001110 \\ 01100111 \\ 00110010 \\ 00111101 \\ 01011011 \\ 01010100 \\ 00000001 \\ 01101000 \end{matrix}$$

$$H^{(0)} = \begin{matrix} 00001110 \\ 01110110 \\ 00100011 \\ 00111101 \\ 01011011 \\ 01000101 \\ 00010000 \\ 01101000 \end{matrix}$$

and where $G^{(i)}$, $H^{(i)}$ represents the cosets $G^{(0)}$ and $H^{(0)}$, respectively, with the columns permuted cyclically i positions to the left. Let g_j^i represent the jth row of $G^{(i)}$ (numbering consecutivelv from the top down) and $\bar{g}_j^i$ its complement, and similarly for h_j^i and $\bar{h}_j^i$, the elements of $H^{(i)}$ and their complements.

Now consider $G^{(0)}$ and observe that an overlap of g_2^0 followed by $\bar{g}_5^0$ is equivalent to g_4^0. Further, g_5^0 followed by g_4^0 contains the word g_2^0. Thus any set containing g_4^0 and g_2^0 can not be comma-free since both g_5^0 and $\bar{g}_5^0$ cannot be excluded from it. Similarly, any set containing $\bar{g}_2^0$ and g_4^0 cannot be comma-free due to the necessity of including either g_6^0 or $\bar{g}_6^0$. Clearly, identical arguments exclude from consideration the sets containing $\bar{g}_4^0$, $\bar{g}_2^0$ or $\bar{g}_4^0$, g_2^0. Since there are no other possibilities, no comma-free orthogonal set can be obtained from the coset $G^{(0)}$. Similar investigation eliminates the remaining cosets and hence establishes the non-existence of any comma-free binary orthogonal set of eight 8-tuples. Table 1 summarizes the source of the conflicts which make it impossible to select an orthogonal set from the union of $G^{(i)}$ and $\bar{G}^{(i)}$ or $H^{(i)}$ and $\bar{H}^{(i)}$ for all eight cases.

The superscripts in Table 1 have been excluded for convenience. Note that it was unnecessary to exclude all possibilities explicitly. That is, as mentioned earlier, one element can be selected at random, since the complement of any comma-free code must be comma-free. Thus, for example, in investigating G (see *RS 36-9*) no search need be made for a comma-free code which includes $\bar{g}_1$ since none exists containing g_1.

Table 1. Sources of conflict

Coset	Any set containing	Cannot contain
$G^{(0)}$	g_2 , g_4 $\bar{g}_2$, g_4	g_5 , $\bar{g}_5$ g_6 , $\bar{g}_6$
$G^{(1)}$	g_1 $\bar{g}_4$, g_8 g_1	g_3 , g_4 , $\bar{g}_8$ $\bar{g}_3$ g_3 , g_3
$G^{(-1)}$	g_7 g_7 , $\bar{g}_1$ g_2 , $\bar{g}_4$ g_2 , g_4 g_7	$\bar{g}_1$ $\bar{g}_2$, $\bar{g}_8$ g_6 , $\bar{g}_6$ $\bar{g}_8$ g_8 , g_8
$G^{(-2)}$	g_8 g_8 , g_1 $\bar{g}_4$, g_1 g^8	$\bar{g}_1$ g_4 , g_3 $\bar{g}_3$ g_3 , $\bar{g}_3$
$H^{(0)}$	h_4 h_4 , h_7	$\bar{h}_7$ h_1 , $\bar{h}_1$
$H^{(1)}$	h_8 , $\bar{h}_4$ h_8 , h_4	h_3 , $\bar{h}_3$ h_5 , $\bar{h}_5$
$H^{(-1)}$	h_1 h_1 , h_2 h_4 h_1	$\bar{h}_2$ $\bar{h}_4$, h_7 h_7 h_7 , $\bar{h}_7$
$H^{(-2)}$	h_4 h_4 , $\bar{h}_1$	h_1 h_2 , $\bar{h}_2$

2. The Sizes of Certain Code Dictionaries[a]

S. W. Golomb, S. Rosen, and J. Kohler

Because of its relevance to the geometric structure of error-correcting codes, the following problem was considered in *RS 36-13*, pp 23–24. The cells of a hyperchessboard of size n^k are to be colored in a maximum number of colors $w(k, n)$ such that a rook placed anywhere will cover (attack or occupy) at least one cell of each color. Several new results will be discussed here.

It is clear that $w(k, n) \leq 1 + k(n-1)$ as a cell does not "see" more than $1 + k(n-1)$ cells including itself.

Theorem 1. $w(k, n) < 1 + k(n-1)$, if $1 + k(n-1)$ does not divide n^k.

Proof. Suppose $w(k, n) = 1 + k(n-1)$. Consider all the cells containing the color A. No one of these may see any of the other A-cells, or it will not see $1 + k(n-1)$ different colors. And no two A-cells may see the same B-cell, say, where B is any other color, for then the B-cell will not see $1 + k(n-1)$ different colors. Thus since every cell must see an A-cell, $[1 + k(n-1)] \cdot$ number of A-cells $= n^k$, so that $1 + k(n-1)$ divides n^k.

Theorem 2. $w(3, 3) = 5$

Proof. By the figure below and Theorem 1, $5 \leq w(3, 3) < 7$. Consider the following arrangement:

A B D	E C E	B D A
C E A	D X B	A F C
B D C	E A E	C B D

Here the blank cell, X, may be left blank or may be filled with any one of the five colors and still preserve the required property. This shows $w(3, 3) \geq 5$.

It may be shown without too much difficulty that $w(3, 3) \neq 6$. Briefly, this is proved by contradiction, showing first that there may be no blank cells in the entire configuration; then that each color must appear at least once in each plane; and finally, that each of

[a]Continued from *RS 36-13*, pp 23–24.

the four possible (non-equivalent) ways of placing the three non-repeated colors that must appear in each plane leads to a contradiction.

Theorem 3. $w(3,4) = 8$

Proof. By the figure below (derived in *RS 36-13*) and Theorem 1, $8 \leq w(3,4) < 10$.

E	F	D	C
H	G	A	B
D	C	E	F
A	B	H	G

G	H	B	A
F	E	C	D
B	A	G	H
C	D	F	E

D	C	E	F
A	B	H	G
E	F	D	C
H	G	A	B

B	A	G	H
C	D	F	E
G	H	B	A
F	E	C	D

As in Theorem 2, it may be demonstrated without much difficulty that $w(3,4) \neq 9$. Again one must show that there may be at most one blank cell in each plane and that each color must appear at least once per plane. This will result in a linear expression, $9X + 8(9 - X)$, for the number of cells which must be occupied, where $X \geq 0$. But this number always exceeds 64. If there are no blank cells the expression is even larger, $10X + 8(9 - X)$, with $X \geq 0$.

Theorem 4. $w(4,3) = 9$

Proof. Since $1 + 4 \cdot 2 = 9$, this result is best possible. Consider the following representation of a pair of 3×3 orthogonal Latin squares:

	0	1	2
0	0	1	2
1	2	0	1
2	1	2	0

	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Consider the nine terms: 0000, 0111, 0222, 1021, 1102, 1201, 2012, 2120, 2201, where the first two members denote the position of the cell, the third is the occupant of the cell in the first Latin square, and the fourth is the occupant of the cell in the second Latin square. These nine terms form a group of 4-tuples *ABCD* under term-by-term addition modulo 3. The coset leaders of this group as a subgroup of ternary four-space are: 0000, 0001, 0010, 0100, 1000, 0002, 0020, 0200, 2000. One may then fill in nine 3×3 squares with nine colors, where the nine appearances of color *A* are obtained from the nine members of one of the nine cosets as follows: the terms *abcd* will give the cell in which an *A* is to be placed by using *c* to mark the "square-row," *d* the "square-column" (thus determining a particular 3×3 square at this stage); then *a* marks the row within the square and *b* the column; this determines a particular cell in the configuration. The final result is the figure below, which conveniently represents $3 \times 3 \times 3 \times 3$ "space."

A	F	G
H	B	E
I	D	C

B	E	H
D	C	I
F	G	A

C	I	D
G	A	F
E	H	B

D	C	I
F	G	A
B	E	H

G	A	F
E	H	B
C	I	D

H	B	E
I	D	C
A	F	G

E	H	B
C	I	D
G	A	F

I	D	C
A	F	G
H	B	E

F	G	A
B	E	H
D	C	I

It is clear that each cell sees each of the other colors *once*, and thus $w(4,3) = 9$. The configuration above is indeed a Latin square (of order 9) with interesting symmetry properties.

Theorem 4 may also be obtained as a special case of the following result:

Theorem 5. If n is a power of a prime, and if $k = n + 1$, then $w(k,n) = n^2$.

Proof. By the upper bound formula, $w(k,n) \leq 1 + (n+1)(n-1) = n^2$. To exhibit a coloring in n^2 colors, consider the vector space spanned by the following set of $n - 1$ vectors of length $n + 1$

$$
\begin{array}{l}
0\,0\,0 \cdots 0\,0\,1\,1\,1 \\
0\,0\,0 \cdots 0\,1\,0\,1\,2 \\
0\,0\,0 \cdots 1\,0\,0\,1\,3 \\
\cdots\cdots\cdots\cdots \\
0\,0\,1 \cdots 0\,0\,0\,1\ (n-3) \\
0\,1\,0 \cdots 0\,0\,0\,1\ (n-2) \\
1\,0\,0 \cdots 0\,0\,0\,1\ (n-1),
\end{array}
$$

over the field $GF(n)$, the finite field of n elements which exists since n is a prime power, where $0, 1, 2, \cdots, n-1$ denote the elements of this field. (These are *not* the integers modulo n, unless n is a prime.) It is clear that these vectors are linearly independent over $GF(n)$, and that any non-trivial linear combination of them yields a vector with at least three non-zero components. Hence they generate a single-error-correcting group code of n^{n-1} elements, which has n^2 cosets in the entire space of n^{n+1} elements. Each of the n^2 cosets may then be assigned a

distinct *color*. The current status of this problem is summarized in Table 2.

Table 2. Current results on w (k, n)

k \ n	1	2	3	4	5	6	7	8	9	10	
1	1	2	3	4	5	6	7	8	9	10	→
2	1	2	3	4	5	6	7	8	9	10	→
3	1	4	5	8							
4	1	4	9								
5	1			16							
6	1				25						
7	1					?					
8	1						49				
9	1							64			
10	1								81		
11	1									?	↘ n^2 for prime power values of n
	↓										

3. Signals by Flares, H. C. Rumsey

Consider a ship at sea wishing to signal another vessel with a sequence of flares of several colors. For simplicity sake assume that the ship has flares of three colors: say a red, b white, and c blue flares. It is clear that a signal should not consist of an arbitrary sequence of these flares, since two consecutive flares of the same color might be mistaken for a single flare. We agree to call a sequence of these flares of length $a + b + c$ a *signal* if no two consecutive flares in the sequence are of the same color.

A generating function for the number of possible signals shall be obtained, as well as a formula for computing this number. An asymptotic formula for the number of signals in the "cubic case" ($a = b = c$, large) shall be derived. The generating function and formula which are obtained have obvious generalizations to any number of differently colored flares, but at present there is no simple extension of the asymptotic result to the case of more than three colors.

Let $N(a, b, c)$ be the number of possible signals of length $a + b + c$ that can be made with a red, b white, and c blue flares. Let $S(a, b, c)$ be the number of arbitrary sequences of length $a + b + c$ that can be formed with these same flares. Let $P_n(m)$ be the number of ways of writing the non-negative integer m as the sum of n non-negative integers (counting permutations as distinct representations).

It is easy to see that each sequence counted by $S(a, b, c)$ can be uniquely associated with some signal of (possibly) shorter length by simply "collapsing" consecutive flares of the same color. Thus the *sequence rbbwrrr* of 4 red, 1 white and 2 blue flares is associated with the *signal rbwr* of 2 red, 1 white and 1 blue flares. Conversely, it is easily seen that each signal counted by $N(a', b', c')$ $(a' \leq a, b' \leq b, c' \leq c)$ is associated with exactly $P_{a'}(a - a')\ P_{b'}(b - b')\ P_{c'}(c - c')$ sequences counted by $S(a, b, c)$. Combining these facts, one obtains the relation

$$S(a, b, c) = \sum_{\substack{a' \leq a \\ b' \leq b \\ c' \leq c}} P_{a'}(a - a')\, P_{b'}(b - b')\, P_{c'}(c - c')\, N(a', b', c') \tag{1}$$

This relation will be used to find the generating function of N from those of S and P. Define these generating functions as follows:

$$\mathcal{N}(x, y, z) = \sum_{a, b, c = 0}^{\infty} N(a, b, c)\, x^a y^b z^c$$

$$\mathcal{S}(x, y, z) = \sum_{a, b, c = 0}^{\infty} S(a, b, c)\, x^a y^b z^c$$

$$\mathcal{P}_n(x) = \sum_{m = 0}^{\infty} P(m)\, x^m$$

Since these series and all others used are easily shown to be absolutely convergent in a sufficiently small neighborhood of the "orgin," the routine justifications of the manipulations performed on them shall be omitted.

To obtain a closed form for $\mathcal{S}(x, y, z)$, observe that

$$S(a, b, c) = \frac{(a + b + c)!}{a!\, b!\, c!}$$

Hence

$$\mathcal{S}(x, y, z) = \sum_{a, b, c = 0}^{\infty} \frac{(a + b + c)!}{a!\, b!\, c!}\, x^a y^b z^c$$

$$= \sum_{n = 0}^{\infty} \sum_{a + b + c = n} \frac{n!}{a!\, b!\, c!}\, x^a y^b z^c$$

$$= \sum_{n = 0}^{\infty} (x + y + z)^n$$

$$= \frac{1}{1 - (x + y + z)}$$

The corresponding formula for $\mathcal{P}_n(x)$ is obtained from the obvious recursion relation

$$P_n(m) = \sum_{j = 0}^{m} P_{n-1}(j)$$

$$P_0(m) = \begin{cases} 0 & m \neq 0 \\ 1 & m = 0 \end{cases}$$

One has

$$\mathcal{P}_n(x) = \sum_{m=0}^{\infty} P_n(m)\, x^m = \sum_{j,\, m=0}^{\infty} P_{n-1}(m-j)\, x^m$$

$$= \sum_{j=0}^{\infty} x^j \sum_{m=0}^{\infty} P_{n-1}(m-j)\, x^{m-j}$$

$$= \sum_{j=0}^{\infty} x^j\, \mathcal{P}_{n-1}(x) = \frac{\mathcal{P}_{n-1}(x)}{1-x}$$

Therefore

$$\mathcal{P}_n(x) = \frac{\mathcal{P}_0(x)}{(1-x)^n} = \frac{1}{(1-x)^n}$$

Theorem. The generating function of $N(a, b, c)$ is given by

$$\mathcal{N}(x, y, z) = \frac{1}{1 - \dfrac{x}{1+x} - \dfrac{y}{1+y} - \dfrac{z}{1+z}} \tag{2}$$

Proof. Multiplying by $x^a y^b z^c$ and summing, Eq (1) becomes

$$\mathcal{S}(x, y, z) = \sum_{a,b,c=0}^{\infty} \sum_{\substack{a' \leq a \\ b' \leq b \\ c' \leq c}}^{\infty} N(a', b', c')\, P_{a'}(a - a')$$

$$\times P_{b'}(b - b')\, P_{c'}(c - c')\, x^a y^b z^c$$

$$= \sum_{a', b', c' = 0}^{\infty} N(a', b', c')\, x^{a'} y^{b'} z^{c'}$$

$$\times \mathcal{P}_{a'}(x)\, \mathcal{P}_{b'}(y)\, \mathcal{P}_{c'}(z)$$

$$= \sum_{a,b,c=0}^{\infty} N(a, b, c) \left(\frac{x}{1-x}\right)^a \left(\frac{y}{1-y}\right)^b \left(\frac{z}{1-z}\right)^c$$

$$= \frac{1}{1 - x - y - z}$$

Substitute $x/1+x$, $y/1+y$ and $z/1+z$ for x, y, and z, respectively, to obtain Eq (2), which completes the proof of the theorem.

By a direct expansion of the right hand side of Eq (2) one obtains the formula

$$N(a, b, c) = \operatorname*{coef}_{x^a y^b z^c} \left\{ \frac{1}{1 - \dfrac{x}{1+x} - \dfrac{y}{1+y} - \dfrac{z}{1+z}} \right\}$$

$$= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} (-1)^{a+b+c-i-j-k}$$

$$\binom{a-1}{i-1}\binom{b-1}{j-1}\binom{c-1}{k-1} \frac{(i+j+k)!}{i!\, j!\, k!} \tag{3}$$

The formula is obviously not well suited for computing $N(a, b, c)$. In this paragraph another generating function for $N(a, b, c)$ shall be derived which yields a formula more amenable to numerical work. Observe that $N(a, b, c)$ can also be computed as follows:

$$N(a, b, c) = \operatorname*{coef}_{x^a y^b z^c} \left(\frac{(1-x)^{a-1}(1-y)^{b-1}(1-z)^{c-1}}{1 - x - y - z} \right) \tag{4}$$

This equation is proved by the fact that a direct computation of the prescribed coefficient immediately produces the expression appearing on the right of Eq (3). By separately expanding the numerator and denominator of the function on the right hand side of Eq (4) one obtains

$$N(a, b, c) = \operatorname*{coef}_{x^a y^b z^c} \left\{ \sum_{m, i, j, k = 0}^{\infty} (-1)^{i+j+k} \right.$$

$$\left. \binom{a-1}{i}\binom{b-1}{j}\binom{c-1}{k} x^i y^j z^k (x + y + z)^m \right\} \tag{5}$$

Since one need only consider terms of degree $a + b + c$ in x, y, z, Eq (5) may be written

$$N(a, b, c) = \operatorname*{coef}_{x^a y^b z^c} \left\{ \sum_{i, j, k = 0} (-1)^{i+j+k} \binom{a-1}{i}\binom{j}{b-1} \right.$$

$$\left. \binom{c-1}{k} x^i y^j z^k (x + y + z)^{a+b+c-i-j-k} \right\}$$

$$= \operatorname*{coef}_{x^a y^b z^c} \left\{ \left(1 - \frac{x}{x+y+z}\right)^{a-1} \left(1 - \frac{y}{x+y+z}\right)^{b-1} \right.$$

$$\left. \left(1 - \frac{z}{x+y+z}\right)^{c-1} \right\}$$

$$N(a, b, c) = \operatorname*{coef}_{x^a y^b z^c} \{(x+y+z)^3 (y+z)^{a-1} (z+x)^{b-1} (z+y)^{c-1}\} \tag{6}$$

(This was derived also in Ref 13, p 376, Eq 4.12.) In Eq (6) the generating function referred to at the beginning of this paragraph is displayed. It has an obvious extension to any number of colors. This extension needs no formal proof since none of these techniques actually depend on the assumption of three colors.

A close inspection of Eq (6) will show that $N(a, b, c)$ can be expressed as a sum on essentially only one variable. One way of writing this sum is displayed in the following theorem, along with the form it takes in the cubic case ($a = b = c$). The derivation of these formulas is quite straightforward from Eq (6) and is omitted. One finally has

Theorem. $N(a,b,c)$ is the coefficient of $x^a y^b z^c$ in the polynomial expansion of $(x+y+z)^3 (y+z)^{a-1} (x+z)^{b-1} (x+y)^{c-1}$, which gives

$$N(a,b,c) = \sum_{i+j+k=3} \frac{3}{i!\,j!\,k!} \sum_m \binom{a-1}{b-j-c+1+1} \binom{b-1}{b+i+a-1+m} \binom{c-1}{m}$$

In case $a = b = c$, this reduces to

$$N(a,a,a) = 6 \sum_{j=0}^{a-1} \binom{a-1}{j}^3 + 18 \sum_{j=0}^{a-2} \binom{a-1}{j}^2 \binom{a-1}{j+1} + 3 \sum_{j=0}^{a-3} \binom{a-1}{j} \binom{a-1}{j+1} \binom{a-1}{j+2}$$

Here is derived the asymptotic formula

$$N(a,a,a) \sim \frac{9(3)^{1/2}}{4\pi} \frac{8^a}{a}$$

valid for large a (this is not done in Ref 1). The approach is to write $N(a,a,a)$ as a contour integral, reduce this integral to what is essentially a real integral, and then estimate the rate of growth of the real integral by means of two lemmas.

To begin, recast Eq (6) in the form

$$N(a,a,a) = \underset{1/xyz}{\text{coef}} \left(1 + \frac{1}{x} + \frac{1}{y} + \frac{1}{z}\right)^3 \left(\frac{x}{y} + \frac{x}{z}\right)^{a-1} \left(\frac{y}{x} + \frac{y}{z}\right)^{a-1} \left(\frac{z}{x} + \frac{z}{y}\right)^{a-1} \tag{7}$$

where "coef" refers to the coefficient of $1/xyz$ in the Laurent expansion of the indicated function about the origin. Observe that the right side of Eq (7) is homogeneous in (x,y,z). It follows that

$$N(a,a,a) = \underset{1/xy}{\text{coef}} \left(1 + \frac{1}{x} + \frac{1}{y}\right)^3 \left(\frac{x}{y} + x\right)^{a-1} \left(\frac{y}{x} + y\right)^{a-1} \left(\frac{1}{x} + \frac{1}{y}\right)^{a-1}$$

One may write this last expression as a double contour integral

$$N(a,a,a) = \frac{1}{(2\pi i)^2} \int_{c_2} \int_{c_1} \left(1 + \frac{1}{x} + \frac{1}{y}\right)^3 \left(\frac{x}{y} + x\right)^{a-1} \left(\frac{y}{x} + y\right)^{a-1} \left(\frac{1}{x} + \frac{1}{y}\right)^{a-1} dx\,dy$$

where c_1 encircles the origin in the complex x-plane and c_2 does the same in the y-plane.

Substituting $x = e^{i\alpha}$, $y = e^{i\beta}$ and simplifying one obtains

$$N(a,a,a) = \frac{8^{a-1}}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} (1 + e^{-i\alpha} + e^{-i\beta})^3 \times e^{i\alpha} e^{i\beta} [F(\alpha,\beta)]^{a-1}\, d\alpha\, d\beta \tag{8}$$

where

$$F(\alpha,\beta) = \tfrac{1}{4}[1 + \cos\alpha + \cos\beta + \cos(\alpha - \beta)]$$

One needs the following two lemmas to determine the asymptotic behavior of this last integral. The proof of the first lemma is straightforward and is omitted.

Lemma. Let $F(\alpha,\beta)$ be a real-valued, continuous function of α and β over a region R in the (α,β) plane. Let $|F|$ have an absolute maximum at the point $(\alpha_0,\beta_0) \in R$. Let $G(\alpha,\beta)$ be a function continuous on R such that $G(\alpha_0,\beta_0) \neq 0$. Then the following asymptotic relation holds for large a:

$$\int_{R'} \int [F(\alpha,\beta)]^a G(\alpha,\beta)\, d\alpha\, d\beta \sim G(\alpha_0,\beta_0) \int_{R'} \int [F(\alpha,\beta)]^a\, d\alpha\, d\beta$$

where R' is any fixed neighborhood of (α_0,β_0) lying in R.

Lemma. Let R be any neighborhood of the origin $(\alpha = \beta = 0)$ in which

$$1 - \frac{\alpha^2 - \alpha\beta + \beta^2}{\rho^2} \geqq 0$$

Then as $a \to \infty$,

$$\int_R \int \left(1 - \frac{\alpha^2 - \alpha\beta + \beta^2}{\rho^2}\right)^{a-1} d\alpha\, d\beta \sim \frac{2\pi\rho^2}{a(3)^{1/2}}$$

Proof of second lemma: make the substitution

$$\alpha = \rho r \left(\frac{\cos\theta}{(3)^{1/2}} + \sin\theta\right)$$

$$\beta = \rho r \left(-\frac{\cos\theta}{(3)^{1/2}} + \sin\theta\right)$$

The Jacobian of this transformation is

$$J\left(\frac{\alpha,\beta}{r,\theta}\right) = \frac{2\rho^2}{(3)^{1/2}}\, r$$

The integral in the lemma therefore becomes

$$\frac{2\rho^2}{(3)^{1/2}} \int_{R(r,\theta)} \int (1 - r^2)^{a-1}\, r\, dr\, d\theta$$

[The region $R(r,\theta)$ is some neighborhood of the origin in the (r,θ) plane and lies inside the circle $1 - r^2 = 0$.] Thus, for some sufficiently small r_0 (depending on R) one has

$$\frac{2\rho^2}{(3)^{1/2}}\int_0^{2\pi}\int_0^{r_0}(1-r^2)^{a-1}\,r\,dr\,d\theta \leq \int_R\int\left(1-\frac{\alpha^2-\alpha\beta+\beta^2}{\rho^2}\right)^{a-1} d\alpha\,d\beta \leq \frac{2\rho^2}{(3)^{1/2}}\int_0^{2\pi}\int_0^{1}(1-r^2)^{a-1}\,r\,dr\,d\theta$$

The lemma now follows immediately.

Returning to Eq (8), one sees that the conditions of the first lemma are satisfied for the integral on the right hand side (F has an absolute maximum of 1 at $\alpha = \beta = 0$). One concludes that

$$N(a,a,a) \sim \frac{27}{(2\pi)^2}\,8^{a-1}\int_R\int (F(\alpha,\beta))^{a-1}\,d\alpha\,d\beta \qquad (9)$$

where R is any sufficiently small neighborhood of the origin. Now let ε be an arbitrary small positive constant. A simple application of Taylor's Theorem shows that the inequalities

$$1-\frac{\alpha^2-\alpha\beta+\beta^2}{4-\varepsilon} \leq F(\alpha,\beta) \leq 1-\frac{\alpha^2-\alpha\beta+\beta^2}{4+\varepsilon}$$

hold in a sufficiently small neighborhood R of the origin. Therefore one has

$$a\int_R\int\left(1-\frac{\alpha^2-\alpha\beta+\beta^2}{4-\varepsilon}\right)^{a-1} d\alpha\,d\beta \leq a\int_R\int F^{a-1}\,d\alpha\,d\beta \leq a\int_R\int\left(1-\frac{\alpha^2-\alpha\beta+\beta^2}{4+\varepsilon}\right)^{a-1} d\alpha\,d\beta \qquad (10)$$

The second lemma can be applied to the outside integrals to obtain

$$\frac{2\pi(4-\varepsilon)}{(3)^{1/2}} \leq \frac{\lim}{a\to\infty}\,a\int_R\int F^{a-1} \leq \overline{\lim_{a\to\infty}}\;a\int_R\int F^{a-1} \leq \frac{2\pi(4+\varepsilon)}{(3^{1/2})}$$

But since ε was arbitrary, it follows that

$$\int_R\int (F(\alpha,\beta))^{a-1}\,d\alpha\,d\beta \sim \frac{8\pi}{a(3^{1/2})}$$

Combining this with Eq. (10) the final theorem has been proved:

Theorem.

$$N(a,a,a) \sim \frac{9(3^{1/2})}{4\pi}\,\frac{8^a}{a}$$

C. Range Sequence Techniques

R. C. Titsworth and R. P. Loomba

1. Finding Sequences with Favorable Autocorrelation Functions, R. C. Titsworth

Binary cyclic sequences having certain periods of the form $4k$-1 possess autocorrelation functions which are ideal, in that the out-of-phase correlation values are all the same and negative. These are the so-called *pseudo-noise*, or *two-level autocorrelation*, sequences which have been described frequently in the literature and in Jet Propulsion Laboratory publications over the past few years (Refs 14, 15, 16, 17 and *RS 36-9*). These ideal sequences do *not* exist for *every* period of the specified form, the first counter-example being period 27. This article describes a search for the best (nearest-to-ideal) sequences of any specified length.

a. Autocorrelation function theorems. In this section, some theorems are stated concerning the autocorrelation function of a sequence. For brevity, the proofs will be omitted here.

Let $\alpha = \{\alpha_n\}$ be a cyclic sequence in $+1$ and -1. The correlation $R(m)$ of α is defined in the usual way:

$$R(m) = \sum_{n=1}^{p} \alpha_n\,\alpha_{n+m}$$

where p is the period of α. The difference d between the number of $+1$'s and -1's in α, per period, is called the *imbalance* of α.

$$d = \sum_{n=1}^{p} \alpha_n$$

Whenever $|d| \leq 1$, α is said to be *balanced*.

Next, for a given α let $R_{\alpha M}$ be the maximum out-of-phase correlation value

$$R_{\alpha M} = \max_{m \not\equiv 0 (p)} R(m)$$

and let R_M be the minimum such $R_{\alpha M}$ taken over all sequences α of period p.

$$R_M = \min_{\alpha} R_{\alpha M} = \min_{\alpha}\ \max_{m \not\equiv 0 (p)} R_\alpha(m)$$

Theorem 1. $R(m_1) - R(m_2)$ is divisible by 4.

***Theorem* 2.** R_M is bounded from below by

$$R_M \geq \begin{cases} 0; \text{ if } p \equiv 0\,(4) \\ 1; \text{ if } p \equiv 1\,(4) \\ 2; \text{ if } p \equiv 2\,(4) \\ -1; \text{ if } p \equiv 3\,(4) \text{ and a pseudonoise sequence exists having period } p \\ 3; \text{ if } p \equiv 3\,(4) \text{ and no pseudonoise sequence exists having period } p \end{cases}$$

Let α be a sequence whose maximum out-of-phase correlation is R_M. If the inequality above can be replaced by equality, α is called a *minimax* sequence. If α is a sequence such that R_M occurs fewer times in $R(m)$ than in the correlation function of any other sequence having the same period p, then α is *optimal*. Note that α may be minimax but non-optimal, and vice versa. However, we can connect these two ideas by the following theorem:

***Theorem* 3.** If α is a balanced sequence having a 3 (or less)-level autocorrelation function, whose out-of-phase levels are separated by no more than 4, then α is both minimax and optimal.

We use this theorem as a criterion in the search described in Sect b.

b. Computer search for sequences. One method of finding optimal sequences is an exhaustive search through all sequences of period p. This involves investigating approximately 2^{p-1} sequences. Even if we look at equivalence classes of sequences under an affine group (*RS 36-10*, pp. 17–21), we still need to look at roughly $k2^p$ sequences, where $k \approx 0.002$ for large p.

A method which has produced very good results is an iterative procedure which considers only about ½ p sequences using a technique which requires summing over p^2 terms, and hence requires computer time which increases as p^3 instead of 2^p.

The procedure used is as follows: Let α be a candidate for the optimal sequence, and let $R(m)$ be its autocorrelation. If we change α_k to $-\alpha_k$ in the sequence, the correlation $R_k(m)$ of the new sequence may be better or worse than $R(m)$. If it is better, we may make this change in α and repeat the procedure on the new sequence; if it is worse, or no better, we can try changing another α_k. If no such change produces a better correlation function, the process is terminated. The type of minimum, of course, depends on the "goodness" criterion we set.

In this method, we must compare $R(m)$ and $R_k(m)$ to see which is the better. The goodness criterion we use here is the $2s$th moment of $R(m)$ about an idealized correlation function value q. Let

$$E_k = \sum_{m=1}^{p} \{[R(m) - q]^{2s} - [R_k(m) - q]^{2s}\}$$

If $E_k > 0$, then $R_k(m)$ is a *better* correlation function than $R(m)$ as far as its $2s$th moment about q is concerned.

When the computer finally arrives at a sequence α such that $E_k \leq 0$ for all k, it can be checked manually for its number of levels, imbalance, minimaximality, optimality, etc. Such a method generally finds only a *relative* minimum of

$$\sum_{m=1}^{p} [R(m) - q]^{2s}$$

When $E_k > 0$ for *some* k, there are usually *many* such k; a decision is needed to indicate which α_k is changed to $-\alpha_k$. The following schemes were tried:

(1) least k such that $E_k > 0$

(2) least k such that $E_k > 0$ and $E_k \geq E_{k'}$ for all k'

(3) least k such that $E_k > 0$ and α_k preferably has same sign as previous changed element

The second of these methods seemed to give best results toward finding minimax and optimal sequences up to about $p = 20$ for $s = 1$, and up to about $p = 40$ for $s = 2$. The program has been modified for $s = 4$, and searches are now underway for sequences up to length 63. The value of q which seemed to work best was -2.

The results of the searches up to period 23 are contained in Table 3. Both the sequence α_m and its correlation function $R(m)$ are listed, starting with $m = 1$:

p	α_1	$R(1)$
	α_2	$R(2)$
	•	
	•	
	•	
	α_p	$R(p)$

Each term in the initial sequence was chosen by tossing a coin. Several starts with different initial sequences were required to find optimal sequences for some lengths, mainly those for which pseudonoise sequences exist. This seems to be due to the rarity of such sequences compared to other *relatively* good sequences of the same period.

Table 3. Optimal binary sequences[a]

p	α	R	s
3	+1	−1	1
	−1	−1	
	−1	3	
4	+1	00	1
	−1	−4	
	−1	00	
	+1	4	
5	+1	1	1[a]
	−1	−3	
	−1	−3	
	+1	1	
	+1	5	
6	+1	−2	1
	−1	−2	
	−1	2	
	+1	−2	
	−1	−2	
	+1	6	
7	+1	−1	1[b]
	−1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	+1	7	
8	−1	00	1
	−1	00	
	−1	−4	
	+1	00	
	−1	−4	
	+1	00	
	+1	00	
	+1	8	
9	−1	1	1
	−1	1	
	−1	−3	
	+1	−3	
	−1	−3	
	+1	−3	
	+1	1	
	+1	1	
	+1	9	
10	−1	2	1
	−1	−2	
	−1	−2	
	+1	−2	
	−1	−2	
	−1	−2	
	+1	−2	
	+1	−2	
	+1	2	
	+1	10	
11	−1	−1	1[c]
	−1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	+1	−1	
	+1	11	
12	−1	00	1
	−1	00	
	−1	00	
	+1	−4	
	−1	00	
	−1	−4	
	+1	00	
	−1	−4	
	+1	00	
	+1	00	
	+1	00	
	+1	12	
13	−1	1	1[d]
	−1	1	
	−1	1	
	+1	−3	
	−1	−3	
	−1	−3	
	+1	−3	
	−1	−3	
	+1	−3	
	+1	1	
	+1	1	
	+1	1	
	+1	13	
14	+1	−2	2[e]
	+1	2	
	−1	−2	
	+1	−2	
	+1	2	
	−1	−2	
	+1	−2	
	−1	−2	
	+1	2	
	−1	−2	
	−1	−2	
	−1	2	
	+1	−2	
	+1	14	
15	−1	−1	1[b]
	−1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	−1	−1	
	+1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	+1	−1	
	+1	15	
16	−1	00	1
	−1	00	
	−1	00	
	+1	−4	
	−1	00	
	−1	00	
	+1	−4	
	+1	00	
	−1	−4	
	+1	00	
	−1	00	
	+1	−4	
	+1	00	
	+1	00	
	+1	00	
	−1	16	
17	−1	−3	1[d]
	−1	1	
	−1	1	
	+1	1	
	−1	−3	
	−1	−3	
	+1	1	
	+1	−3	
	−1	1	
	+1	1	
	−1	−3	
	+1	−3	
	+1	1	
	+1	1	
	+1	1	
	−1	−3	
	+1	17	
18	−1	2	1
	−1	−2	
	+1	−2	
	+1	−2	
	−1	−2	
	−1	−2	
	+1	2	
	+1	−2	
	−1	−2	
	+1	−2	
	−1	2	
	+1	−2	
	+1	−2	
	+1	−2	
	−1	−2	
	−1	2	
	−1	18	
19	−1	−1	1[c]
	−1	−1	
	−1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	+1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	−1	−1	
	+1	−1	
	+1	19	
20	−1	00	2
	−1	00	
	−1	00	
	−1	00	
	−1	−4	
	+1	00	
	+1	00	
	−1	−4	
	+1	00	
	−1	−4	
	−1	00	
	+1	−4	
	+1	00	
	+1	00	
	−1	−4	
	+1	00	
	+1	00	
	−1	−4	
	−1	00	
	+1	−4	
	+1	00	
	+1	00	
	−1	−4	
	+1	00	
	+1	00	
	−1	−4	
	−1	00	
	+1	−4	
	+1	00	
	+1	00	
	−1	−4	
	+1	00	
21	−1	−3	2
	−1	1	
	−1	1	
	−1	−3	
	−1	−3	
	+1	−3	
	+1	1	
	−1	−3	
	+1	1	
	−1	1	
	−1	1	
	+1	1	
	−1	−3	
	+1	1	
	−1	−3	
	+1	−3	
	+1	−3	
	+1	1	
	−1	1	
	+1	−3	
	+1	21	
22	−1	−2	1
	−1	2	
	−1	−2	
	−1	−2	
	−1	−2	
	+1	−2	
	+1	−2	
	−1	2	
	+1	−2	
	−1	−2	
	−1	2	
	+1	−2	
	−1	−2	
	+1	2	
	−1	−2	
	+1	−2	
	+1	−2	
	+1	−2	
	−1	−2	
	+1	2	
	+1	−2	
	+1	22	
23	−1	−1	2
	−1	−1	
	+1	−1	
	+1	−1	
	−1	−1	
	−1	−1	
	+1	−1	
	+1	−1	
	−1	−1	
	+1	−1	
	+1	−1	
	+1	−1	
	−1	−1	
	−1	−1	
	−1	−1	
	−1	−1	
	+1	−1	
	−1	−1	
	+1	23	

[a] The symbol p is the period; α, the sequence; R, the correlation function; and $2s$ the moment of $R(m)$ about an idealized correlation function q.

[b] Linear shift register.

[c] Quadratic residue.

[d] Not quadratic residue.

[e] It is not known that this sequence is actually optimal since $d = 2$. However, many different initial conditions were tried, and each resulted in a 3-level, $d = 2$ sequence.

2. *A Concise Algorithm for the Sequence Generation*, R. P. Loomba

A sequence generator capable of generating all binary sequences of length up to 64 was implemented and reported in *RS 36-10*, pp. 27–31. The sequence generator consists of three parts: a shift register with a word detector, a function generator, and a reduction tree. A block diagram for the sequence generator is shown in Fig 4.

The first part contains a 6-stage shift register with linear feedback logic and a word detector, by means of which it is possible to generate one sequence of every possible length up to 64. This part of the sequence generator determines the desired length. For different length sequences, different word detector connections are required. Since there are 64 possible lengths, the word detector has 64 different connections, one for each length.

After the desired length has been achieved, the next step is to arrange the 1's and 0's properly to produce the desired sequence. This is done by means of the function generator and the reduction tree. The function generator generates all 16 functions of 2 variables, and the reduction tree combines selected functions of 2 variables into the desired function of 6 variables. The tree is called a reduction tree because it is derived (*RS 36-10*, pp. 27–31) by reducing the general function of 6 variables into a certain combination of functions of 2 variables. In order to obtain the desired sequence, suitable connections are made between the function generator and the reduction tree.

Previously, the word detector connections were given by a 2-page table which listed the appropriate connections for obtaining the length of the sequence. The connections between the reduction tree and the function generator were determined with the help of a 14-page table. A set of instructions and a special work sheet were used to obtain these connections.

The two tables, consisting of 16 pages, take a long time to prepare and are inconvenient to use; moreover, in the case of longer sequences, the two tables would contain hundreds of pages. Consequently, there has been a need for a concise table containing the necessary information.

A new table of connections. A single 1-page table, Table 4, has now been devised which replaces the two tables formerly comprising 16 pages. The first column in the table (*W.D.*) gives the length of the sequence, and in the same row the columns x_1, x_2, x_3, x_4, x_5 and x_6 give the word detector connections for that particular length. If there were no word detector, during each cycle the shift register would go through 63 different states represented by the columns x_1, x_2, x_3, x_4, x_5 and x_6 in Table 4. The first state with all 0's will not occur. The function of the word detector is to discard the unnecessary states and to allow only the appropriate states. The number of states that are allowed to appear during one cycle equals the length of the sequence. Therefore, for every length there is a first and a last state of the shift register. The first state of the shift register is given by column *B*, and the last state is the one below the word detector connections when Table 4 is read downwards. The reason for the last state being the one below the word detector connections is that there is one unit delay between the input and the output of the word detector. For example, in the case of a sequence of length 7, the word detector connections are 110111; the first state of the shift register is 101101 and the last state is 011011 (Table 4). Therefore, the word detector will allow the shift register to go through only 7 different states starting with 101101 and ending with 011011.

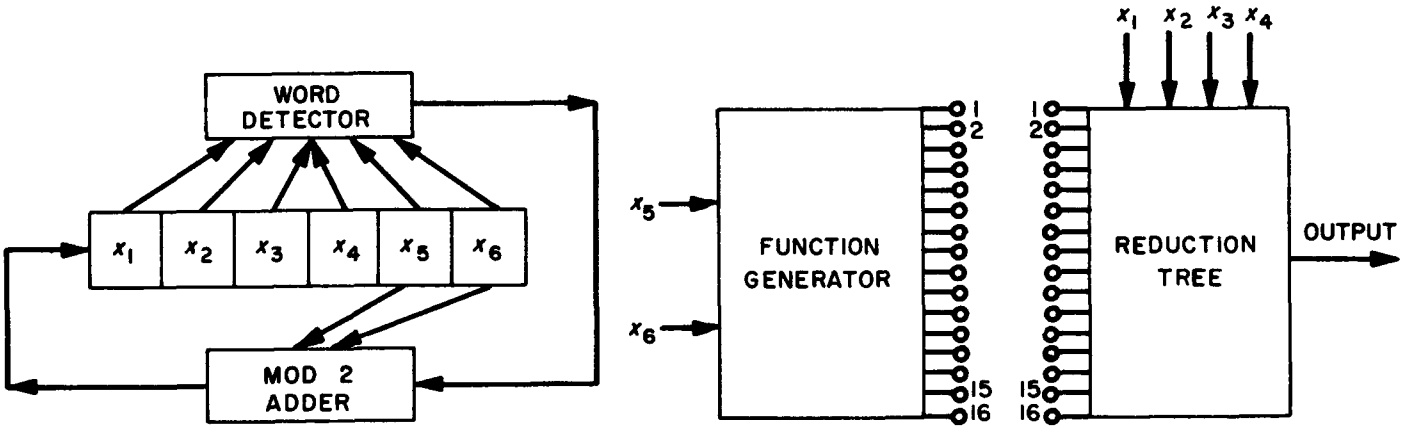


Fig. 4. Sequence generator

Table 4. Algorithm for the general purpose sequence generator

W.D.	B	X_1	X_2	X_3	X_4	X_5	X_6	L	M	b
63	64	0	0	0	0	0	0	0	1	
53	63	1	0	0	0	0	0	8	1	
34	5	0	1	0	0	0	0	4	1	
43	10	0	0	1	0	0	0	2	1	
6	29	0	0	0	1	0	0	1	1	
5	20	0	0	0	0	1	0	0	4	
44	57	1	0	0	0	0	1	8	2	
23	58	1	1	0	0	0	0	12	1	
27	19	0	1	1	0	0	0	6	1	
12	40	0	0	1	1	0	0	3	1	
49	36	0	0	0	1	1	0	1	4	
10	51	1	0	0	0	1	1	8	8	
41	14	0	1	0	0	0	1	4	2	
25	53	1	0	1	0	0	0	10	1	
55	22	0	1	0	1	0	0	5	1	
46	38	0	0	1	0	1	0	2	4	
33	8	1	0	0	1	0	1	9	2	
54	17	1	1	0	0	1	0	12	4	
26	30	1	1	1	0	0	1	14	2	
24	9	1	1	1	1	0	0	15	1	
42	37	0	1	1	1	1	0	7	4	
35	39	1	0	1	1	1	1	11	8	
8	21	0	1	0	1	1	1	5	8	
20	28	0	0	1	0	1	1	2	8	
14	55	0	0	0	1	0	1	1	2	
19	43	1	0	0	0	1	0	8	4	
9	49	1	1	0	0	0	1	12	2	
50	44	1	1	1	0	0	0	14	1	
32	54	0	1	1	1	0	0	7	1	
47	13	0	0	1	1	1	0	3	4	
60	31	1	0	0	1	1	1	9	8	
29	16	0	1	0	0	1	1	4	8	
17	3	0	0	1	0	0	1	2	2	
3	34	1	0	0	1	0	0	9	1	
22	46	0	1	0	0	1	0	4	1	
45	60	1	0	1	0	0	1	10	2	
56	41	1	1	0	1	0	0	13	1	
52	18	0	1	1	0	1	0	6	4	
59	7	1	0	1	1	0	1	11	2	1
48	11	1	1	0	1	1	0	13	4	1
13	4	1	1	1	0	1	1	14	8	1
21	15	0	1	1	1	0	1	7	2	0
4	50	1	0	1	1	1	0	11	4	1
7	42	1	1	0	1	1	1	13	8	0
36	59	0	1	1	0	1	1	6	8	0
16	56	0	0	1	1	0	1	3	2	
30	27	1	0	0	1	1	0	9	4	
40	47	1	1	0	0	1	1	12	8	
11	33	0	1	1	0	0	1	6	2	
28	23	1	0	1	1	0	0	11	1	
61	52	0	1	0	1	1	0	5	4	
38	35	1	0	1	0	1	1	10	8	
2	2	0	1	0	1	0	1	5	2	
18	25	1	0	1	0	1	0	10	4	
15	61	1	1	0	1	0	1	13	2	
37	45	1	1	1	0	1	0	14	4	
62	48	1	1	1	1	0	1	15	2	
1	26	1	1	1	1	1	0	15	4	
39	1	1	1	1	1	1	1	15	8	
31	62	0	1	1	1	1	1	7	8	
51	24	0	0	1	1	1	1	3	8	
57	32	0	0	0	1	1	1	1	8	
64	12	0	0	0	0	1	1	0	8	
58	6	0	0	0	0	0	1	0	2	

The desired sequence is written in column b starting with the row corresponding to the first state of the shift register and proceeding downwards. For example, in the case of a sequence of length 7, the starting point is the word 101101. Let us denote by $g_0, g_1, \cdots, g_{15}$, the 16 functions generated by the function generator. The 16 inputs to the reduction tree are denoted by $i_0, i_1, \cdots, i_{15}$. In order to obtain the desired sequence, the problem is to determine which g's should be connected to a particular i.

For a particular state of the shift register, column L gives the appropriate i and the product bM gives the corresponding g. The letter b represents the binary digit in the desired sequence correponding to a particular state of the shift register. Therefore, in order to obtain a particular b in the desired sequence, $i(L)$ should be connected to $g(bM)$. It is possible that the same number in the L column will occur more than once in the case of a particular sequence. In such situations the $i(L)$ should be connected to $g(b_1M_1 + b_2M_2 + b_3M_3 + b_4M_4)$. The reader should note that the same number cannot occur more than four times in the L column.

Suppose the sequence 1110100 of length 7 is desired. From Table 4 the word detector connections for a sequence of length 7 are 110111. The sequence is written in column b, starting with the first state of the shift register. From Table 4 the value of L for the first state of the shift register is 11. Also for the fifth state, L has the same value. The value of M for these two states is 2 and 4, respectively, and b is 1 in both cases. Therefore $i(11)$ should be connected to $g(1 \times 2 + 1 \times 4)$. In other words, $i(11)$ should be connected to $g(6)$. Similarly, L equal to 13 occurs twice in this sequence. The value of M for the two cases is 4 and 8. But b is 1 in the first case and 0 for the second. Hence $i(13)$ should be connected to $g(1 \times 4 + 0 \times 8)$, or $i(13)$ should be connected to $g(4)$. In the same way it can be shown that $i(14)$ should be connected to $g(8)$; $i(7)$ to $g(0)$; $i(11)$ to $g(4)$; and $i(6)$ to $g(0)$.

D. An Analysis of the Narrow-Band Spectra of Venus

R. L. Carpenter

Since the publication of *Radar Exploration of Venus* (Ref 18), efforts have been continued to reduce the large amount of data obtained with the narrow-band spectrum analyzer. The reduction consisted of computing the spectra of the radar return for each day's operation and then measuring their bandwidth and relative signal-to-noise ratio. The results of this reduction are presented and interpreted in terms of the rotation period of Venus.

If the returned signal was actually reflected from Venus' surface and if no hidden systematic error exists in the measured bandwidths, then the results would appear to suggest that Venus is rotating backwards!

1. Description of Experiment

Briefly, the experimental procedure, as far as the narrow-band spectrum analyzer was concerned, consisted of transmitting a 2388-mc cw signal for 15 min on each day of the experiment and recording the reflected energy. After each signal run, 10 min of noise were recorded with either the transmitter turned off or the receiving antenna pointed away from Venus. Of course these transmissions and recordings were phased to take account of the round-trip time of the radar signal. The returned signal was heterodyned down to the low-pass frequency range between dc and 125 cps. An ephemeris controlled oscillator was used to remove the doppler shift due to the relative velocity between Venus and the Earth. The accuracy of the ephemeris controlled oscillator and heterodyning system was on the average ± 1 cps.

The ephemeris controlled oscillator was purposely offset so that the center of Venus' spectrum fell in the middle of the low-pass frequency range at about 55 cps. The signal was sampled at 500 samples/sec with an analog-to-digital converter and was recorded on tape. These tapes were then processed on the IBM 7090 to produce the power spectrum of the recorded data. Due to the limitations of the storage capacity of the 7090, each 15-min run was broken into 10 parts of 90-sec duration and a separate spectrum was computed for each; consequently, each day's data was converted into 10 spectra whose relative signal level and bandwidth could be measured. The resolution of these spectra was approximately 2 cps. The procedure of computing 10 separate spectra for each day's run has the advantage of reducing the effect of frequency drifts that may have occurred in the ephemeris controlled oscillator or other parts of the receiving system. Also the probable error of the measurements made on each spectrum may be estimated.

2. Measurement of Bandwidth and Signal Level

Three different bandwidths and the relative signal level were measured for each of the Venus spectra. The bandwidth determinations treated the 3-db, 6-db, and effective bandwidths. The effective bandwidth is particularly useful since it tends to minimize the effect of the fluctuations on the spectrum. The 3-db and 6-db bandwidths were found by first drawing a horizontal line through the noise skirts on each side of the spectral peak. This line established a zero reference level and represents the best estimate of the base of Venus' spectrum. The bandwidths were measured at the half- and quarter-power levels relative to this base. The effective bandwidth was derived by dividing the area bounded by Venus' spectrum and its base by its height. The measureable part of the spectrum was taken to lie within ± 30 cps of the peak. Widening this range to ± 60 cps made no significant difference.

The area under Venus' spectrum is also a measure of the power of the received signal and hence establishes the relative signal level. Of course each spectrum must be calibrated. This is easily done by noting that the power in the noise spectrum upon which Venus' spectrum sits is directly proportional to the system's noise temperature. Since this temperature was measured each day of the experiment, any change in its value can be used to adjust the measured area to obtain the relative signal level.

3. Changes in the Relative Signal Level and Bandwidth

In Fig 5a, the changes in signal level are shown over the duration of the experiment for which narrow-band spectral data is available. The dashed curve shows the theoretical change in signal level due to Venus' changing distance from the Earth. These points have not been corrected for variations in system temperature or transmitter power. The mean transmitter power was 12.6 ± 0.5 kw (a variation of about ± 0.2 db) and the mean noise temperature was 64 ± 4 deg (a variation of about ± 0.27 db). The observed signal level matches the theoretical curve rather well considering the fluctuations that occurred from day to day. Since the distance to Venus is known from the ephemeris, the observed signal level may be normalized so that the fluctuations in the reflecting characteristics of the planet itself may be studied. In radar terminology the normalized signal level is equivalent to α, the radar cross section. It was determined

from the observed signal power by use of the radar equation:

$$\sigma = \frac{(4\pi)^3 d^4 P_r}{G_t G_r \lambda^2 P_t} \quad (1)$$

where

d = distance to Venus

P_r = received power

G_t = gain of the transmitting antenna

G_r = gain of the receiving antenna

λ = wavelength

P_t = transmitter power

In Fig 5b, the change in radar cross section during the course of the experiment is shown. This was obtained from Fig 5a by correcting for the variations in distance, transmitter power, and noise temperature. The average of the probable errors, based on the 10 spectra of each day's run for 10 days before and after conjunction, is approximately ±0.11 db. For the data taken before and after this period the average probable error increases to ±0.16 db because of the weaker signal. Including the probable errors of the transmitter power and noise temperature, the probable errors are, respectively, ±0.35 db and ±0.37 db. Both the ±0.16 and ±0.37 db probable errors are shown at the right of the plot.

The 3-db, 6-db, and effective bandwidths are shown in Fig 6. Mean probable errors of these measurements are depicted by the bar to the right of each graph. They are ±0.31, ±0.50, and ±0.34 cps respectively. The peak of the dotted curve shows the bandwidth on April 19. The dashed curve is a least-square fit of a 4th-order polynomial but with the 19th excluded because of its extreme value. The standard deviations of the least-square fits are indicated on each graph. They are 1.13, 2.14 and 1.85 cps for the 3-db, 6-db, and effective bandwidths, respectively.

The extreme bandwidth of April 19 is shown in Fig 7a which was taken at $20^h\ 48^m\ 30^s$ UT. It should be recalled that the bandwidths are based on the average of 10 such spectra for each day's run. The 3-db and effective bandwidths of this particular spectrum are 13.7 and 16.8, respectively. One might suspect that the ephemeris con-

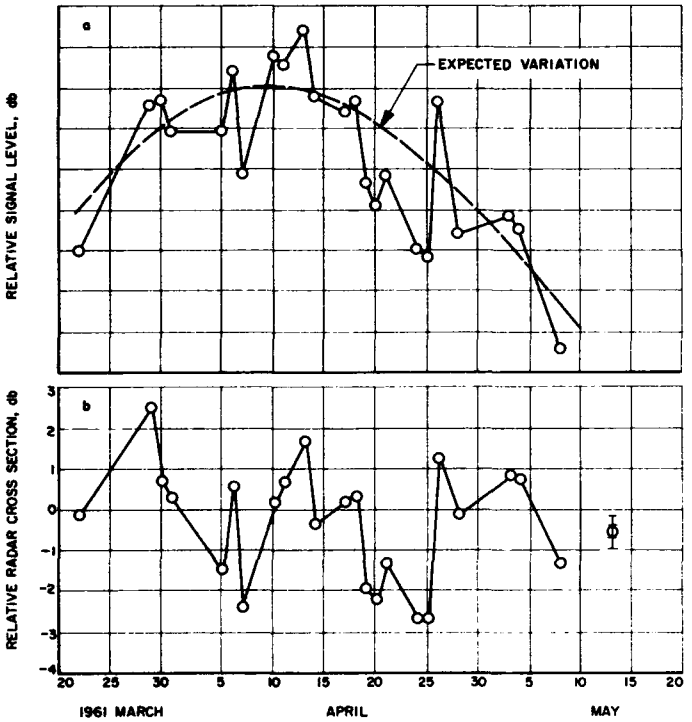


Fig. 5. Variations in relative signal level and relative radar cross section of Venus during experiment

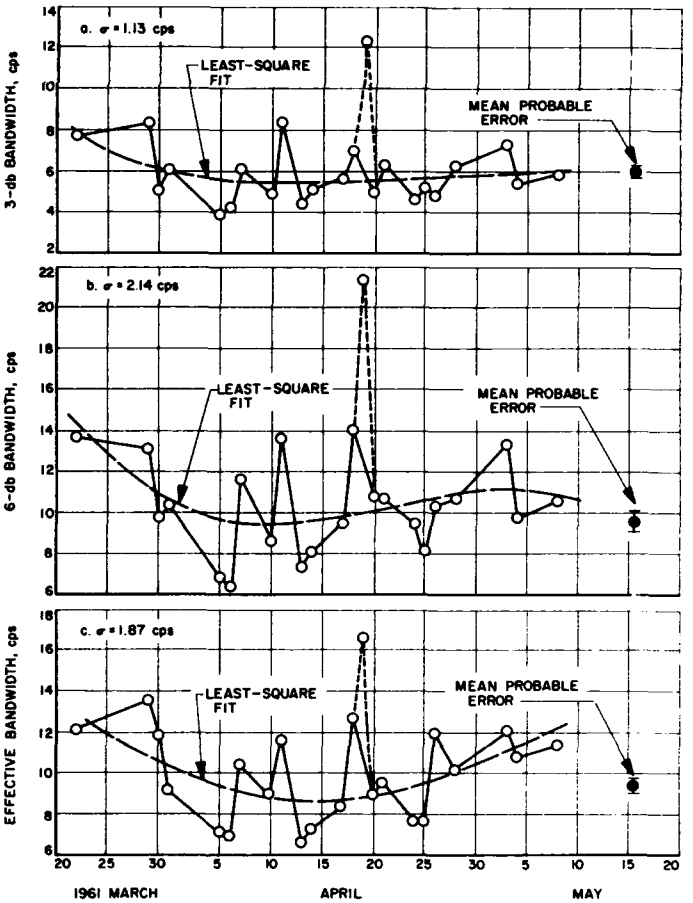


Fig 6. Variations in 3-db, 6-db, and effective bandwidths of Venus during experiment

trolled oscillator may have had excessive jitter and thus spread out the spectrum. However, no unusual jitter was observed in the recordings that monitored it.

For comparison, Fig 7b shows one of the narrowest observed spectra. It was taken April 13 at $21^h\ 21^m\ 30^s$ UT. Since the resolution of the spectral computations is 2 cps, it appeared that the spectrum may be narrower than indicated. Therefore, a spectrum was computed with a resolution of 0.25 cps. This high-resolution spectrum is shown in Fig 7c. The 3-db bandwidth is approximately 2.3 cps, a very narrow spectrum indeed, and changed by almost 6 to 1 between April 13 and 19. It is apparent that the 3-db bandwidth, and to some extent the 6-db and effective bandwidths also, are overestimated particularly for the period around conjunction.

4. Direction and Period of Rotation

From a study of the changes in bandwidth of the returned signal, the direction and period of Venus may be determined if the inhomogenieties in the scattering characteristics of its surface can be smoothed out over the 6.5 weeks of the experiment for which satisfactory narrow-band spectral data were obtained. This is due to the change in the apparent angular rotation of Venus as seen from the Earth as it passes from one side of conjunction to the other. The size of this change depends on the period and direction of the rotation of Venus and is not dependent on the scattering function of the surface. The computed variations in the bandwidth over the run of the experiment are shown in Fig 8 for several different periods of rotation. The bandwidths have been normalized so that the bandwidth at conjunction on April 11 was equal to 9.5 cps. This bandwidth was chosen since it corresponds to the observed 6-db and effective bandwidth around the time of conjunction. The maximum bandwidth, i.e., the bandwidth obtained if signals were received all the way to the limb of Venus, are shown in parenthesis beside each curve.

In cases III, IV, and V it was assumed that the axis of rotation was perpendicular to Venus' orbit plane. In case I the period equals 224.7 days and represents synchronous rotation. Several trials were computed with the axis tipped 30 deg relative to the normal of the orbit plane and oriented in a number of directions in longitude. All these trials lie within the shaded area. Case II corresponds to Kuiper's suggested period and axis orientation (Ref 19). He reported a period of 14 days with the axis oriented in the direction $\alpha = 3^h\ 34^m$ and $\delta = +80°7$. Cases III, IV and V are for retrograde rotations of 75 days, 150 days, and 225 days, respectively.

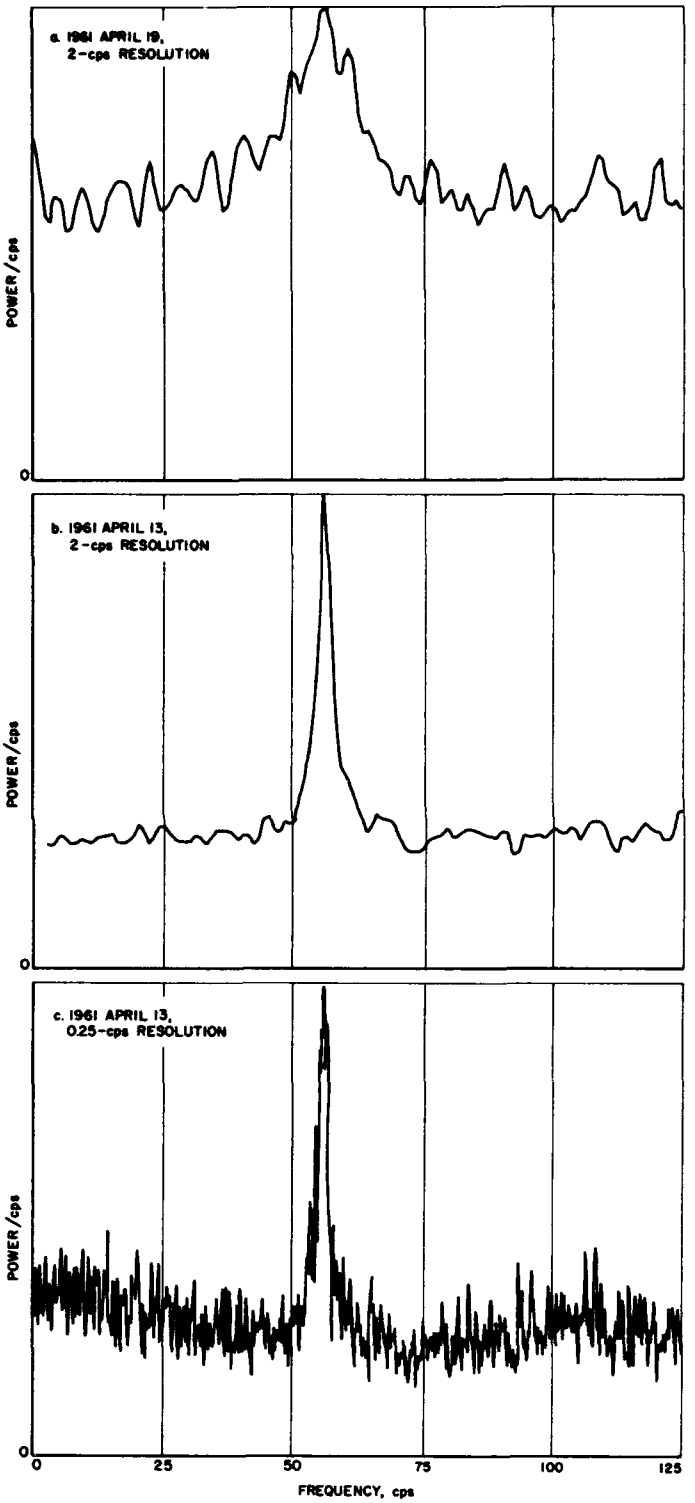


Fig. 7. Comparison of spectra for April 19 and April 13, 1961

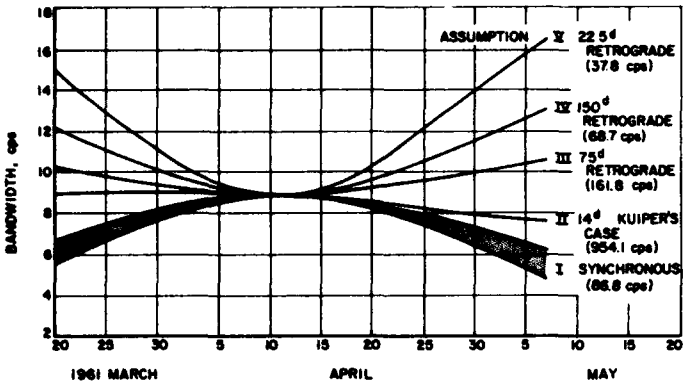


Fig. 8. Theoretical variation of bandwidth for different rotation periods and axis orientation of Venus during experiment

All three bandwidth measures show a trend to be narrower around conjunction, with the 3-db bandwidth showing the least trend. It is believed that this flattening in the 3-db case is due to the limited resolution of the spectral computations and the narrowness of Venus' spectrum. A comparison of these computed bandwidth curves with the observed 6-db and effective bandwidths data suggests that the period may be 150 ±75 days retrograde! The probable error is based on the standard deviation of the least square fits. It must be remembered that this assumes that the axis of rotation is perpendicular to the orbit plane. If the inclination of the axis approaches 90 deg, then the observed trend could arise even for forward rotation. This comes about because the bandwidth is proportional to the sine of the angle between the axis and the line of sight. As the axis appears to swing around toward the Earth and pass it, the bandwidth will decrease to a minimum and then increase. However, the chances that Venus' axis happened to be so oriented seem unlikely.

It is also possible that the bandwidth could be related to the changing phase of Venus and a possible consequent change in any turbulence in its atmosphere and/or possible redistribution of material on its surface. However, as far as the atmosphere is concerned, it is difficult to imagine a mechanism that would significantly broaden a signal at 2388 mc.

The trend of the observed bandwidth measurements could arise if the estimates were negatively correlated with the signal-to-noise ratio. In other words, the bandwidth could have been overestimated for small signal-to-noise ratios, and this could result in having an increase in bandwidths near the beginning and end of the experiment. To test this possibility, the correlation was measured between the effective bandwidth and the normalized signal level (radar cross section). This comparison should show any tendency for broad bandwidth to occur with poor signal level. The results are

$$R = +0.5, \quad \sigma_R = 0.2$$

σ_R is the one-sigma confidence level of the correlation. The correlation is significant, but with the *opposite* sign from that implied by the above reasoning.

Another possible source of systematic error should occur if the tracking of the doppler change was more erratic at the beginning and end of the experiment. A check was made on the accuracy of the doppler tracking by measuring the changes in the frequency of the peak of each day's spectra. The size of these changes, from day to day, was found to be uncorrelated with the bandwidth.

5. Conclusions

Unfortunately, a definitive answer cannot be given for the rotation period of Venus based on the present data. The trend in the change of the bandwidth during the experiment suggests the startling result that Venus rotates retrograde with a period of 150 ±75 days. When one considers the flattening of the bandwidth trend due to the two-cycle resolution of the spectra, a longer retrograde period is indicated. Also the polarization measurements described in Ref 18, which indicate that Venus scatters something like the Moon, strongly suggest a long rotation period. This is particularly true when the very narrow spectrum of April 13 is considered. However, it must be mentioned that opposing this conclusion is the rapid change in the bandwidth that occurred between the 19th and 20th of April. This change favors a period closer to ~50 days (forward or retrograde) than to 150 days.

E. Analysis of the Effect of a Limiter on a Square Wave Signal Plus Additive Gaussian Noise

M. Easterling

The system to be analyzed is shown in Fig 9. It is assumed that the added noise is white and that the filter limits the noise power to a finite amount, but is wide

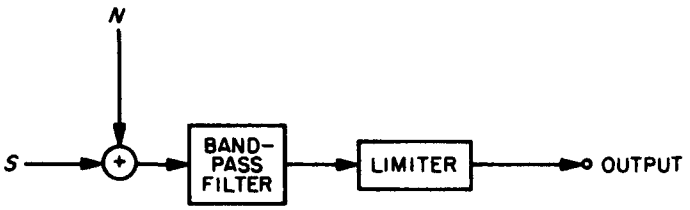


Fig. 9. Square-wave system being analyzed

enough that it does not significantly distort the square-wave signal. It is also assumed that the limiter is ideal; that is, the output is $+1$ if the input is positive and -1 if the input is negative. It is desired to obtain the SNR at the output of the limiter as a function of the SNR at the input to the limiter.

The relative output signal amplitude for any given signal-to-noise ratio at the input to the limiter can be computed as the cross correlation of the output with an uncontaminated version of the original signal. Since the two functions to be cross-correlated are restricted to having values of say, ± 1, the correlation C may be expressed as

$$C = \frac{\text{agreement} - \text{disagreement}}{\text{agreement} + \text{disagreement}} \tag{1}$$

where agreement means the fraction of the time that the two functions agree and disagreement means the fraction of the time the two functions disagree. Since we really desire the expected value of the correlation, it is appropriate to consider the expected value of the agreement, or more simply, the probability of agreement or disagreement. Therefore

$$C = \frac{P(A) - P(D)}{P(A) + P(D)} = P(A) - P(D) \tag{2}$$

By this definition, if there is no noise,

$$P(A) = 1$$

$$P(D) = 0$$

$$C = 1$$

and the relative amplitude of the output signal is unity. Conversely, if there is no signal,

$$P(A) = P(D) = \tfrac{1}{2}$$

$$C = 0$$

and the relative amplitude of the output signal is zero.

Since the output waveform is a square wave, it has constant unit power. The power in the signal component is equal to the amplitude squared. Therefore

$$P_S = C^2 \tag{3}$$

$$P_N = 1 - C^2 \tag{4}$$

$$SNR = \frac{C^2}{1 - C^2} \tag{5}$$

It remains to compute $P(A)$ and $P(D)$. The input noise power may be considered to be 1, in which case the σ of the Gaussian amplitude distribution of the noise component at the output of the filter is also 1. The amplitude of the signal may then be computed as $SNR^{1/2}$: call it A; consider the case when the signal has a value of $+A$. Then the probability that the output of the limiter is negative is the probability that the value of the noise is less than $-A$. Similarly, the probability of a positive output from the limiter when the signal has a value of $-A$ is the probability that the value of the noise is greater than $+A$. Because of the symmetry of the Gaussian distribution these can be combined as

$$P(D) = \int_A^\infty \text{Gaussian distribution with } \sigma = 1 \tag{6}$$

By a similar argument

$$P(A) = \int_{-\infty}^A \text{Gaussian distribution with } \sigma = 1 \tag{7}$$

These can be combined to give

$$C = \int_{-A}^A \text{Gaussian distribution with } \sigma = 1 \tag{8}$$

This last form is convenient since the Gaussian distribution is often tabulated this way.

Since A is the square root of the input signal-to-noise ratio, the output signal-to-noise ratio can be obtained as a function of the input signal-to-noise by combining Eqs (8) and (5) to give

$$SNR_0 = \frac{1}{\dfrac{1}{\left[\displaystyle\int_{-(SNR_I)^{1/2}}^{(SNR_I)^{1/2}} \text{Gaussian distribution with } \sigma = 1\right]^2} - 1} \tag{9}$$

This function for SNR_I is shown in Fig 10 over the range of -10 db to $+10$ db. The line of SNR_0 equal to the SNR_I is included for comparison.

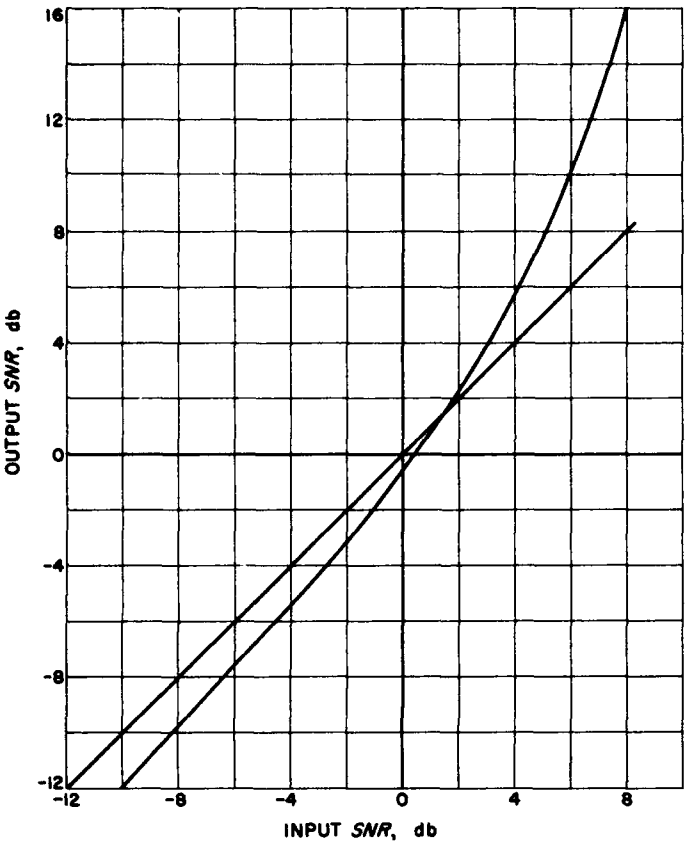


Fig. 10. Signal-to-noise ratio of a ranging code through a limiter

F. Control Theory

R. Goldstein, R. A. Winkelstein, and W. F. Gillmore, Jr.

1. The Minimization of Oscillator Noise, R. Goldstein

Phase-locked loops have found many applications at the Laboratory other than the detection of weak signals immersed in strong noise. Applications such as frequency synthesizers, single side-band modulators, and clean-up loops all operate under the conditions of strong signals and low noise. Yet the character and magnitude of this noise is very significant where high spectral purity of the output wave is important.

The primary frequency which is the input to the phase-locked loop always has some phase noise in it, as does the voltage-controlled oscillator (*VCO*) in the loop. Ordinarily, one would not expect that the loop filter which is designed to make the loop an optimum detector for weak signals would necessarily be the best for these other applications. This report is an investigation of what loop filter is best.

The circuit under consideration is shown in Fig 11, where the assumption of linear (Ref 20) operation has been made. The mathematical characterization of the phase noise of the input is represented by $n_1(t)$, and $n_2(t)$ is the phase noise associated with the *VCO*. The output phase noise is $c(t)$. The transfer function of the *VCO* is $1/jw$ and, assuming that the noise functions have Fourier transforms, it follows that

$$C(w) = \frac{F}{jw + F} N_1(w) + \frac{1}{jw + F} N_2(w)$$

The substitution

$$F = \frac{1 - jwG}{G} \tag{1}$$

will simplify the subsequent analysis. Thus

$$C = (1 - jwG) N_1 + GN_2$$

The frequency noise of the output, $d(t)$, is the derivative of $c(t)$. Expressed in terms of the Fourier transforms,

$$D(w) = (1 - jwG)\, jwN_1 + jwGN_2 \tag{2}$$

Let the total output noise energy, I, be the quantity to be minimized by a suitable choice of G.

$$I = \int_{-\infty}^{\infty} d^2(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} D(w)\, D^*(w)\, dw \tag{3}$$

where the * denotes complex conjugate. Substituting Eq (2) into Eq (3) and noticing that the cross product terms are zero (N_1 and N_2 are assumed independent), we have

$$I = \frac{1}{2\pi} \int_{-\infty}^{\infty} [(1 - jwG)\, jwN_1 (1 + jwG^*)(-jw) N_1^* + GjwN_2 G^* (-jw) N_2^*]\, dw \tag{4}$$

It is desired to find that function G, out of the class of realizable filters (no poles in the lower half w-plane),

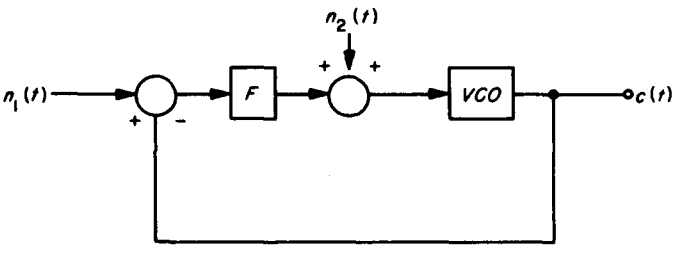


Fig. 11. Block diagram of linearized system

which minimizes I. To that end, we set the variation of I to zero.

$$\delta I = \frac{1}{2\pi}\int_{-\infty}^{\infty} \delta G^* \left[-(jw)^3 (1 - jwG) N_1 N_1^* - (jw)^2 G N_2 N_2^*\right] dw + \frac{1}{2\pi}\int_{-\infty}^{\infty} (\text{complex conjugate}) = 0$$

By combining the two integrals and rewriting, we have

$$\int_{-\infty}^{\infty} \delta G^* \left\{G\left[(jw)^4 N_1 N_1^* - (jw)^2 N_2 N_2^*\right] - (jw)^3 N_1 N_1^*\right\} dw = 0 \tag{5}$$

Choosing G so that the integrand is zero will not, in general, yield a realizable G function. However, for δI to be zero, it is sufficient to keep all of the poles of the integrand in the lower half of the w-plane.

Following the method of Rechtin (Ref 20), we factor the term

$$(jw)^4 N_1 N_1^* - (jw)^2 N_2 N_2^* = \psi^u \psi^l \tag{6}$$

where ψ^u has all of its poles and zeros in the upper half w-plane and ψ^l has its in the lower. Then

$$\int_{-\infty}^{\infty} \delta G^* \psi^l \left[G\psi^u - \frac{(jw)^3 N_1 N_1^*}{\psi^l}\right] dw = 0$$

Next, we expand into partial fractions the term

$$\frac{(jw)^3 N_1 N_1^*}{\psi^l} = \phi^l + \phi^u \tag{7}$$

where ϕ^u contains all of the poles in the upper half w-plane and ϕ^l contains the lower ones. Then

$$\int_{-\infty}^{\infty} \delta G^* \psi^l \left[G\psi^u - \phi^u - \phi^l\right] dw = 0 \tag{8}$$

Eq (8) is satisfied if

$$G = \frac{\phi^u}{\psi^u} \tag{9}$$

for then the integrand has poles only in the lower half w-plane.

As an example, consider the interesting case where N_1 is the phase noise from a *VCO* which has *R-C* filtered white noise for an input. Let N_2 also be white noise, similarly filtered. Then

$$N_1(w) = \frac{A_1}{jw\varepsilon + 1}\frac{1}{jw} \tag{10}$$

and

$$N_2 = \frac{A_2}{jw\varepsilon + 1} \tag{11}$$

where ε is the time constant of the *R-C* filter. Then, from Eq (6)

$$\psi^u\psi^l = \frac{(jw)^4 A_1^2}{-(jw)^2 (1 + jw\varepsilon)(1 - jw\varepsilon)} - \frac{(jw)^2 A_2^2}{(1 + jw\varepsilon)(1 - jw\varepsilon)}$$

so that

$$\psi^u = \frac{-(jw)^2 (A_1^2 + A_2^2)}{1 + jw}, \qquad \psi^l = \frac{1}{1 - jw\varepsilon} \tag{12}$$

From Eq (7)

$$\phi^u + \phi^l = \frac{(jw)^3 A_1^2 (1 - jw\varepsilon)}{-(1 + jw\varepsilon)(jw)^2 (1 - jw\varepsilon)}$$

so that

$$\phi^u = -\frac{jwA_1^2}{1 + jw\varepsilon}, \text{ and } \phi^l = 0 \tag{13}$$

Substituting Eqs (12) and (13) into (9) yields

$$G = \frac{A_1^2}{A_1^2 + A_2^2}\frac{1}{jw} \tag{14}$$

To find the desired filter F, we substitute Eq (14) into (1)

$$F(w) = jw\frac{A_2^2}{A_1^2} \tag{15}$$

It is surprising at first, but Eq (15) shows that the best filter is a differentiator. When $n_2 = 0$, it is obvious from Fig 11 that there will be no output noise if $F = 0$. This is verified by Eq (15). When $n_1 = 0$, it is clear that a fast loop is needed to reduce the output noise. Since the *VCO* is an integrator, the best compensation is a differentiator. Again this is verified by Eq (15).

The total error can be found by substituting Eq (15) into (3), which yields (for the case where $A_1 = A_2$)

$$I = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\left|\frac{1}{2}\frac{A}{1 + jw\varepsilon}\right|^2 + \left|\frac{1}{2}\frac{A}{1 + jw\varepsilon}\right|^2\right) dw = \frac{A^2}{4\varepsilon} \tag{16}$$

It is interesting to compare this number to the total error which would result if the filter F were simply an amplifier (F = constant). A calculation similar to that of Eq (16) shows that

$$I = \frac{A^2}{2\varepsilon} \tag{17}$$

For the case where F is an ideal tracking filter,

$$F = \frac{k(1 + jw\alpha)}{jw}$$

the corresponding error is

$$I \geq \frac{A^2}{2\varepsilon} \tag{18}$$

where equality is reached in the limit as $\alpha \to \infty$.

Comparison of Eqs (16), (17), and (18) shows that the optimum filter produces an improvement of at least 3 db in the integrated square of the frequency noise.

2. *Digital-RF Control System Analysis,* R. A. Winkelstein

An analysis has been made of the digitally controlled programmed local oscillator described in *SPS 37-14.* The purpose of the analysis is to determine the effect on system stability of two prime system variables. One variable is the amount of rate error correction which results from the system error signal. This variable is herein denoted by a and can be precisely controlled by the system design. The second variable is the magnitude of the voltage controlled oscillator (VCO) transfer characteristic B. Precise control of B throughout the entire frequency range of the VCO is difficult because of the inherent nonlinearity of the VCO control characteristic. In addition, the analysis provides a mathematical model of the local oscillator system which may be readily expanded to investigate second-order effects on system behavior resulting from frequency filtering elements and component deviation.

The analysis was carried out by first dividing the system into basic mathematically described functional blocks. For each block, the equivalent Laplace transform transfer characteristic was then determined. Using the z transformation method described by Jury (Ref 21), the z transform of the system closed-loop error signal was obtained for a ramp-type input reference signal. The real time error signal at the sampling instants was then obtained by means of the inverse z transformation. An examination of this real time error signal shows the stability of the system and the rate of convergence to steady-state values.

a. Physical system. A block diagram of the local oscillator physical system is shown in Fig 12. The VCO and RF system generates the local oscillator signal with the desired doppler offset. Also separately available is the low-frequency doppler component which is used to obtain the closed-loop error signal. It is desired to control this doppler frequency to an accuracy of 1 cps out of a continuous band of 50,000 cps. Band switching is used in the RF system to cover the complete doppler range of 1,600,000 cps. The counter digitizes the controlled doppler signal which is then subtracted by the comparator from the ephemeris doppler signal previously stored on paper tape. This subtraction takes place at the end of each 10-sec time interval and produces the system error signal.

The error signal is processed by the two error channels. Rapid correction is provided by the static error channel, and a corrected linear transition to the next expected doppler frequency is provided by the rate error channel. Signals from these channels are suitably summed and used to operate a stepping motor ganged to a precision 40-turn potentiometer. This digital-to-analog conversion technique permits the attainment of a voltage with an extremely fine coherent resolution to better than 1:100,000. Completion of the closed loop is obtained by using this analog voltage to control the VCO.

The manner of operation of the system in response to an arbitrary ephemeris doppler reference signal is shown in Fig 13. Time is shown as real time divided by the sample interval. The ephemeris doppler signal is

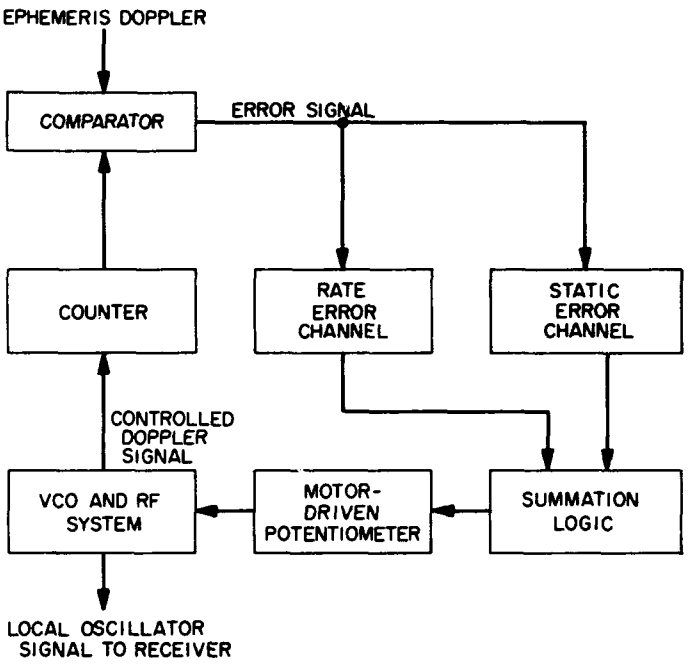


Fig. 12. Local oscillator physical system

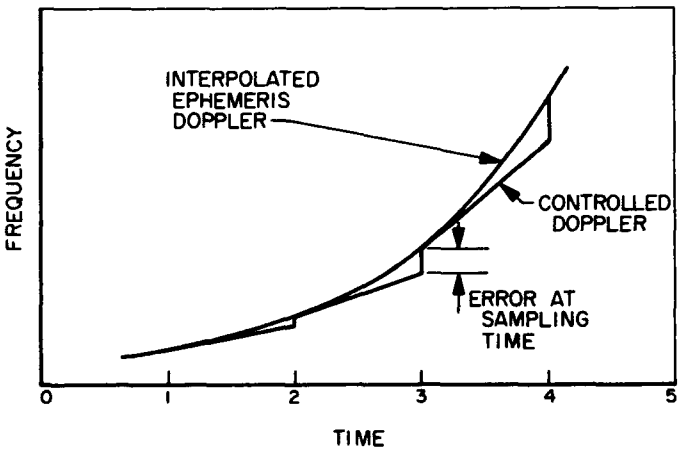


Fig. 13. Operational characteristics

shown as a smooth interpolated curve, although in the physical system it is only available at each sampling time. Between sample times, the controlled doppler frequency varies linearly as determined by the rate error channel. At the sampling time, an error signal is obtained, and the static error channel causes rapid correction of the controlled doppler frequency to the value of the ephemeris doppler frequency. Simultaneously, the slope of the controlled doppler signal is modified by the rate error channel to coincide with the slope of the ephemeris doppler signal. For a ramp ephemeris doppler, the slopes will become equal and the error at the sampling time will become zero.

b. Laplace transfer characteristic. A block diagram of the system Laplace transfer characteristic is presented in Fig 14. The ephemeris doppler, the controlled doppler, and the error signal are all assumed to be functions which are continuously available. Placing the sampling switch to sample the error signal is equivalent to having individual synchronized sampling switches sampling the controlled doppler and ephemeris doppler signals, as is the case in the physical system. $E^*(s)$ is the transform of the error signal at only the sampling time.

The system is divided into three main processing blocks consisting of the two error channels and the VCO transfer characteristic. The static error channel transfer characteristic is $e^{-\delta s}/s$ where s = Laplace transform complex frequency, e = Napierian logarithm base, δ = time delay much smaller than interval between sampling times. The term δ is included because the static error correction takes place a short time after the error signal is sampled. The response of this block to a unit step error signal is shown in Fig 15. At each sampling time the value of the output increases by the amount of the step.

The rate error channel transfer charactertistic is a/s^2 where a is equal to the magnitude of rate correction. For $a = 1$, the rate is exactly corrected to equal the ephemeris slope as determined by the last two samples. Undercorrection takes place for $a < 1$, and overcorrection takes place for $a > 1$. The response of this block to a unit step error signal is shown in Fig 16. At each sampling time the slope of the output increases by a times the amount of the step.

The VCO transfer characteristic is simply B and includes both the VCO control characteristic and the potentiometer characteristic. This assumes that for any finite time interval the transfer characteristic is linear and independent of the control signal rate of change. These assumptions are valid for normal system operation. When $B = 1$, the operation of the system will be as shown in Fig 13. For $B < 1$, the static correction will not quite cause the controlled doppler to reach the ephemeris doppler. For $B > 1$, the controlled doppler will over-

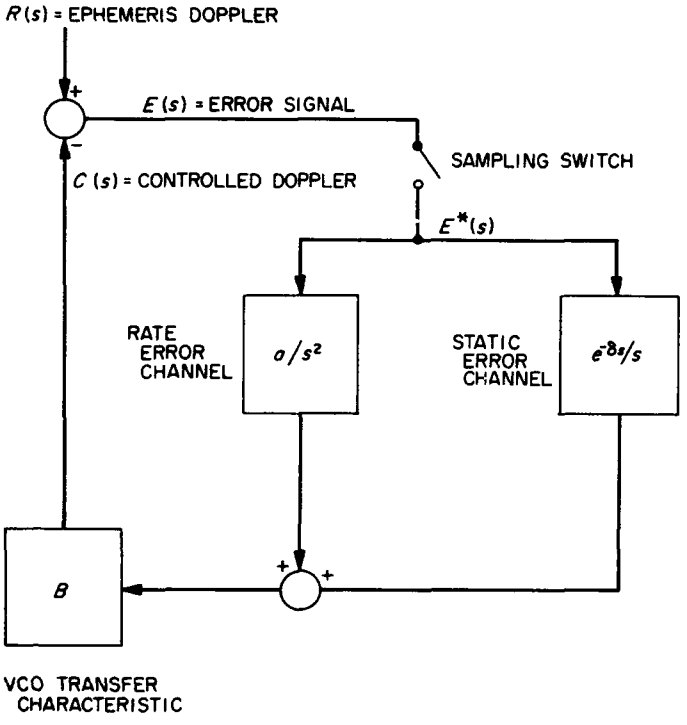


Fig. 14. Laplace transfer characteristic block diagram

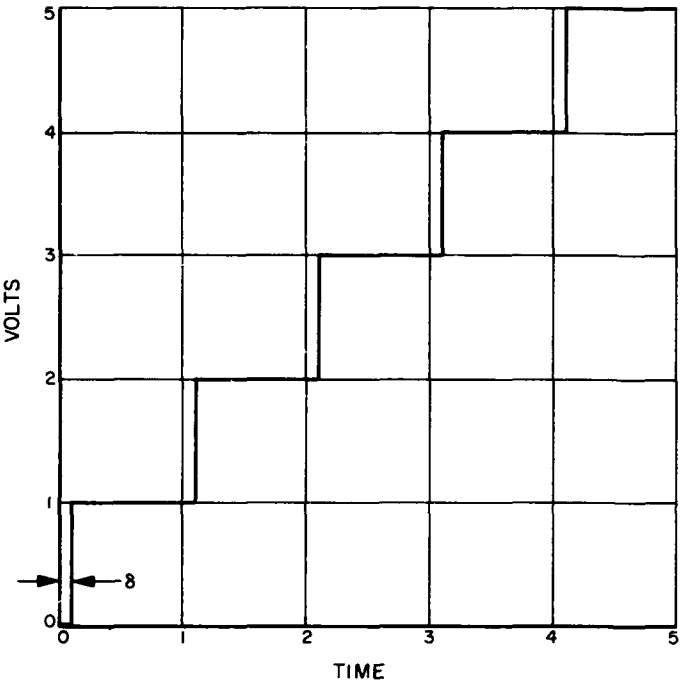


Fig. 15. Static error channel response to step input

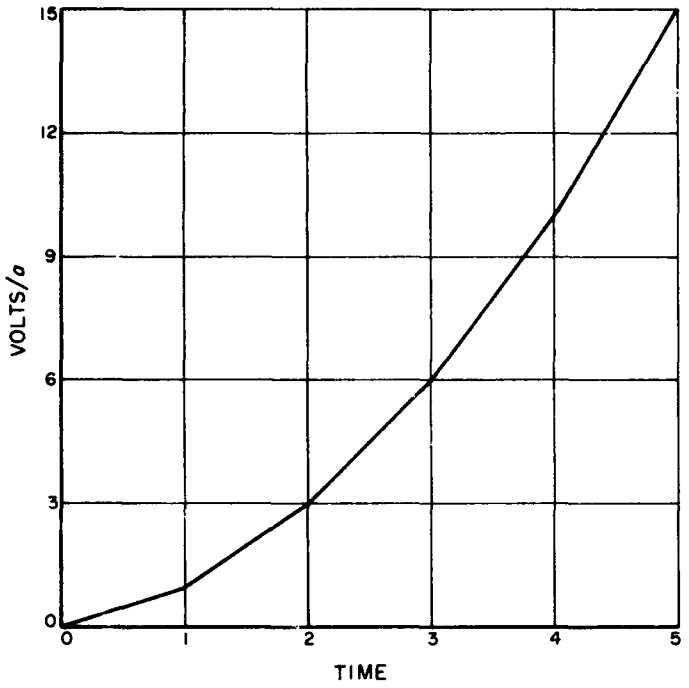


Fig. 16. Rate error channel response to step input

shoot the ephemeris doppler. Thus, although the system is designed for $B = 1$, B may become significantly different from 1 due to nonlinearity of the VCO control characteristic.

c. Closed-loop analysis. The investigation of optimum values for a and B was carried out independently for each variable. First, the closed-loop equation was solved with $B = 1$ and the optimum value was found for a. This value of a was put into the closed-loop equation which was then solved again for B. The criteria for optimization was taken to be those values of a and B which would cause the greatest rate of error decay, while still maintaining an adequate margin of safety from values which would cause the system to become unstable.

For the diagram shown in Fig 14, the z transform of $E^*(s)$ is $E(z)$ which is given by the equation

$$E(z) = \frac{R(z)}{1 + G(z)} \tag{1}$$

where

z = the transform variable e^{sT}, T is the sampling period and is normalized to unity for this analysis

$R(z)$ = the z transform of the reference ephemeris doppler $R(s)$

$G(z)$ = the z transform of the open-loop transfer function $G(s)$

The Laplace transform of an ephemeris doppler ramp function is

$$R(s) = \frac{A}{s^2}$$

where A is the slope of the ramp. The z transform of this function is

$$R(z) = \frac{Az}{(z-1)^2} \tag{2}$$

With $B = 1$, the open-loop transfer function is

$$G(s) = \frac{a}{s^2} + \frac{e^{-\delta s}}{s}$$

from which

$$G(z) = \frac{az}{(z-1)^2} + \frac{1}{(z-1)} \tag{3}$$

δ is made to equal unity since at the sampling times the value of the static error channel response as shown in Fig 15 is independent of δ for $0 < \delta \leqq 1$.

Substituting Eqs (2) and (3) into Eq (1) results in the z transform of the error signal

$$E(z) = \frac{A}{z - (1 - a)}$$

From the inverse z transformation

$$E(n) = A(1-a)^{n-1} \qquad n \geqq 2 \tag{4}$$
$$E(1) = A$$
$$E(0) = 0$$

where $E(n)$ is the value of the error signal at the nth sampling time.

Eq (4) is of the form

$$E(n) = c\alpha^n$$

where c is constant, and α^n is a decay term. The term α^n will decay to zero for all values of α with magnitude less than one. For $|\alpha| > 1$, α^n will increase with time and the system is said to be unstable. In Fig 17 the relationship of α to a is shown. Since $\alpha = 1 - a$, $|\alpha|$ is less than unity for $0 < a < 2$. For $a = 1$, $\alpha = 0$, causing $E(n)$ to be zero for all values of $n > 2$. Thus 1 is the optimum value of a.

The effect of the variable B on the system stability and rate of error decay can now be calculated. With $a = 1$

$$G(s) = \frac{B}{s^2} + \frac{Be^{-\delta s}}{s}$$

from which

$$G(z) = \frac{Bz}{(z-1)^2} + \frac{B}{(z-1)} \tag{5}$$

Substituting Eqs (2) and (5) in Eq (1) results in the z transform of the error signal

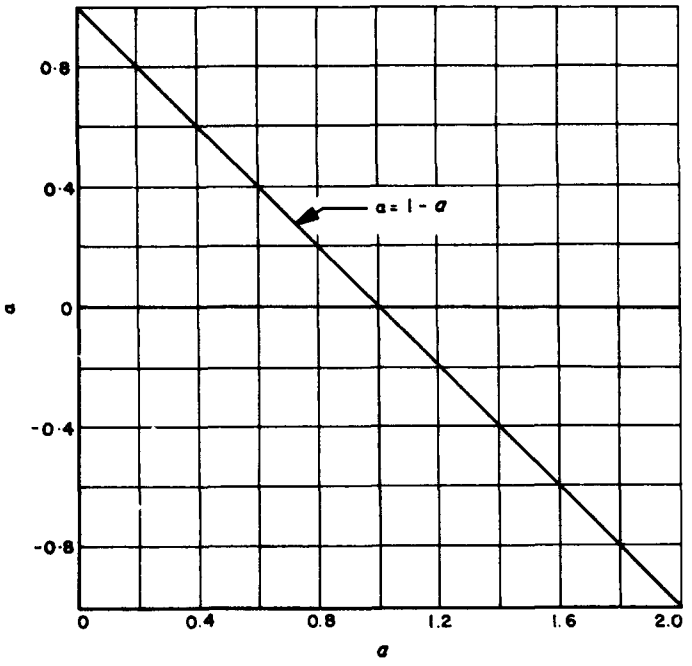


Fig 17. Exponential decay term $B = 1$

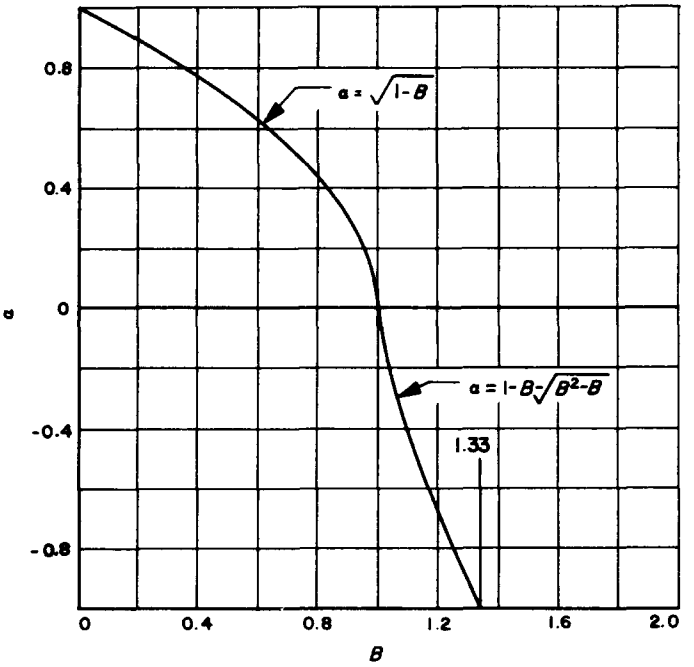


Fig. 18. Exponential decay term $a = 1$

$$E(z) = \frac{Az}{z^2 - 2z(1 - B) + (1 - B)}$$

For $B < 1$

$$E(n) = \frac{A}{(B - B^2)^{1/2}} (1 - B)^{n/2} \sin [n \cos^{-1} (1 - B)^{1/2}]$$

This may be written as $c\alpha^n$ where c oscillates between 1 and -1, and α^n is the decay term; $\alpha = (1 - B)^{1/2}$.

For $B > 1$

$$E(n) = \frac{A}{2(B^2 - B)^{1/2}} (\beta^n - \alpha^n)$$

where

$$\alpha = 1 - B - (B^2 - B)^{1/2}$$

$$\beta = 1 - B + (B^2 - B)^{1/2}$$

Both α^n and β^n are decay terms, but since $|\alpha| > |\beta|$, the stability of the system depends primarily on α^n. The dependence of α on the value of B is shown in Fig 18. The system is stable for all $0 < B < 1.33$. The fastest rate of error decay occurs when $B = 1$. However, slight deviation from 1 causes α to become appreciably > 0.

d. Application to system design. The analysis of the programmed local oscillator has shown that the optimum values of rate correction and VCO transfer characteristic are both 1. The value of rate correction may be precisely controlled by the system design, but the VCO transfer characteristic may deviate from unity because of nonlinearity of the VCO control characteristic. Since the transfer characteristic includes both the VCO control characteristic and the potentiometer characteristic, the latter will be used to compensate the former. Practical curve-shaping techniques will only result in perfect compensation at discrete values of VCO frequency. At other values, the analysis has shown that the transfer characteristic must be < 1.33 to insure stable system operation.

3. Design of a System for Tracking Linearly Polarized Signals, W. F. Gillmore, Jr.

A detection system capable of producing an error signal ε dependent upon the angle θ between the plane of polarization of an incoming signal and the plane of polarization of the receiving antenna has been described previously (*RS 36-12*). This error signal is linear for small values of θ, but for larger values of θ it becomes nonlinear and may even jump suddenly to a value of opposite sign. It will be shown here how a simple closed-loop system can be constructed using this error signal. The closed loop system has some interesting properties which are considerably different from typical linear systems.

a. The closed-loop system. The plane of polarization of the receiving antenna can be adjusted by mechanically rotating the antenna feed. If an electric motor is used to drive this system, it will have a certain mechanical time constant which can be combined with that of the load.

A suitable transfer function for representing the motor and load is given as Eq (1).

$$G(s) = \frac{K}{s(Ts+1)} \tag{1}$$

The gain K can include any amplification in the system, and T denotes the overall mechanical time constant.

A block diagram of the simplest closed-loop system is shown in Fig 19. The plane of polarization of the received signal is θ_1 in radians. Angular error θ is the difference $\theta_1 - \theta_2$. Error signal ε is given by Eq (2), in general

$$\varepsilon(\theta) = \frac{\cos 2\gamma \sin 2\theta}{(2 + 2\cos 2\gamma \cos 2\theta)^{1/2}} \tag{2}$$

where the parameter γ ranges from -45 to $+45$ deg and defines the ellipticity of the polarization of the incoming signal. This relation between ε and θ is illustrated by the curves in Fig 20. Clockwise and counterclockwise circular polarization are represented by letting γ be $+45$ and -45 deg, respectively. Perfectly linear polarization corresponds to a γ of zero.

A useful linear analysis of this system can be made for the special case where the magnitude of θ is small. Error $\varepsilon(\theta)$ is expanded in a Taylor series as shown in Eq (3)

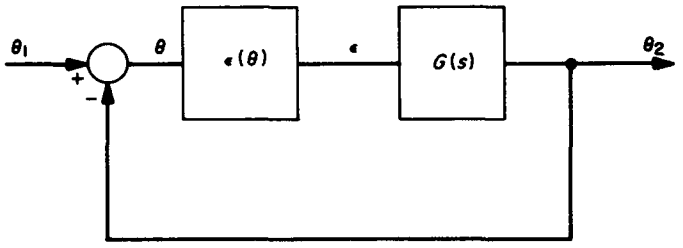


Fig. 19. Block diagram of the simplest closed-loop system

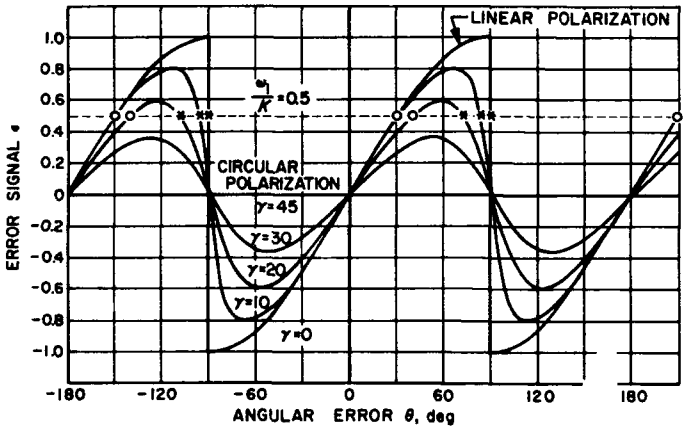


Fig. 20. A graphical method for locating singular points

$$\varepsilon(\theta) = \left[\frac{2\cos^2\gamma - 1}{\cos\gamma}\right]\theta + \left[\frac{(2\cos^2\gamma - 1)(2\cos^2\gamma - 3)}{\cos^3\gamma}\right]\frac{\theta^3}{3!} + \cdots \tag{3}$$

Using only the first term from this Taylor series makes the system linear. Error from this simple approximation is less than 0.25% when the magnitudes of θ and γ are less than 5 deg and less than about 0.5% when the magnitudes of θ and γ are less than 10 deg.

The response of the system in Fig 19 is described by a differential equation, Eq (4)

$$\frac{d^2\theta}{dt^2} + \frac{1}{T}\frac{d\theta}{dt} + \frac{K}{T}\varepsilon(\theta) = \frac{d^2\theta_1}{dt^2} + \frac{1}{T}\frac{d\theta_1}{dt} \tag{4}$$

for radian error θ in terms of input θ_1. It can be seen from this equation that when the input, θ_1, is constant the steady state error will be zero. For ramp inputs ($\theta_1 = \omega_1 t$), however, there is a steady state error given approximately by Eq (5)

$$\theta_{ss} = \frac{\omega_1}{K}\left[\frac{\cos\gamma}{2\cos^2\gamma - 1}\right] \tag{5}$$

This error may be made vanishingly small by choosing a sufficiently large gain K.

As the gain K increases, the system described by Eq (4) shows no tendency toward self-oscillation. The natural frequency of damped oscillations increases, however, and the effective damping decreases. A condition of critical damping occurs when K is given approximately by Eq (6)

$$K = \frac{1}{4T}\left[\frac{\cos\gamma}{2\cos^2\gamma - 1}\right] \tag{6}$$

Sometimes a practical limit on the gain is set by requiring the overshoot to be no more than that given by half the critical damping. This fixes the maximum gain at approximately 4 times the value given by Eq (6).

With the system gain adjusted for a damping of half the critical damping, the steady state error is given approximately by Eq (7)

$$\theta_{ss} = \omega_1 T \tag{7}$$

This equation helps to emphasize the importance of a small value of T. Although the T defined here is the fixed time constant associated with the motor and load, it is often possible to reduce the steady state error in Eq (7) by a factor of 10 or so. This can be done by adding a simple filter of the pole-zero cancellation type before the motor, thereby replacing the time constant T by a smaller one. Another method, theoretically capable of reducing the steady state error to zero, uses an additional integra-

tor in the forward path. Unfortunately, this type of system leads to very difficult stability problems and inevitable circuit complications.

It is useful to know the maximum value of $\dot{\theta}$ occurring at $\theta = 0$ which will just allow the system to come to rest at $\theta = \omega_1/K$. This is difficult to find in general, but simple bounds can be established by ignoring the damping and using the principle of conservation of energy. The result, for the case where $\gamma = 0$ and the linear approximation from Eq (3) is used, is given by Eq (8).

$$|\dot{\theta}_m| \leq \left(\frac{K}{T}\left[\frac{\pi^2}{4} - \frac{\pi\omega_1}{K}\right]\right)^{1/2} \tag{8}$$

Ignoring the damping tends to make this bound more conservative; therefore, the inequality in Eq (8) is shown with the possibility of equality.

b. Effects of nonlinearity. Equation (4) may be written in the form of Eq (9)

$$\frac{d\dot{\theta}}{d\theta} = \frac{\omega_1 - K\varepsilon(\theta) - \dot{\theta}}{T\dot{\theta}} \tag{9}$$

where it is assumed that $\theta_1 = \omega_1 t$. This equation relates θ to $\dot{\theta}$ and may be thought of as defining the slope of trajectories in a phase plane which is the $\dot{\theta} - \theta$ plane here. Points where both the numerator and denominator of Eq (9) vanish are called singular points. In this case singular points can only occur when $\dot{\theta} = 0$, but θ may have any of the values given by Eq (10)

$$\varepsilon(\theta) = \frac{\omega_1}{K} \tag{10}$$

Solutions of Eq (10) for θ may be found graphically. In Fig 20 the error signal ε is plotted as a function of θ for several values of γ. A dotted line is drawn with the ordinate ω_1/K. Intersections of this line with the appropriate ε curve give the desired solutions. Examples are given by the circles and crosses shown on the error curves. These intersections occur at points specified by Eq (11)

$$\cos 2\theta = \frac{-\omega_1^2 \pm (\omega_1^4 - 2\omega_1^2K^2 + K^4\cos^2 2\gamma)^{1/2}}{K^2\cos 2\gamma} \tag{11}$$

There is a theoretical maximum value of $|\omega_1|$ which this system can track. It can be seen graphically in Fig 20 as the highest or lowest possible position of the line $\varepsilon = \omega_1/K$ which will just intersect the particular error curve $\varepsilon(\theta)$. This is the maximum or minimum value of these curves which is given by Eq (12) and occurs at values of θ given by Eq (13).

$$\varepsilon_m = \pm(1 - |\sin 2\gamma|)^{1/2} \tag{12}$$

$$\theta_m = \pm\tan^{-1}(|\operatorname{ctn}\gamma|)^{1/2} \tag{13}$$

These expressions can be used to get a simple upper bound on the magnitude of ω_1 as shown in Eq (14)

$$\omega_1 < K\cos 2\gamma \tag{14}$$

It is also this condition which ensures that the quantity under the radical in Eq (11) will be positive.

The exact nature of the singular points can be investigated by linearly transforming them to the origin in the phase plane. When this is done the intersections marked with a circle in Fig 20 are found to be stable equilibrium positions, because at these points $\varepsilon'(\theta) > 0$. Conversely the intersections marked with crosses are unstable because at these points $\varepsilon'(\theta) < 0$. The unstable equilibrium positions are called saddle points. The stable equilibrium positions are called nodes or foci (spiral points) depending upon whether the system is overdamped or underdamped. A condition of critical damping for the special case where $\gamma = 0$ in the nonlinear system is given by Eq (15)

$$K = \frac{1}{4T}((4\omega_1 T)^2 + 1)^{1/2} \tag{15}$$

This expression is equivalent to Eq (6) when $\gamma = 0$ and $\omega_1 = 0$.

After understanding the nature of the singular points, it is easy to plot the trajectories in the phase plane determined by Eq (9). Typical examples are given by Figs 21 and 22 where the effects of different values of K and ω_1 are shown. The value of steady state error is given by Eq (16) and applies when $\gamma = 0$.

$$\theta_{ss} = \sin^{-1}\frac{\omega_1}{K} \tag{16}$$

Equation (16) still applies even when additional poles are used in the motor transfer function for increased accuracy, as shown in Eq (17).

$$G_2(s) = \frac{K}{s(T_1 s + 1)(T_2 s + 1)} \tag{17}$$

When this form is used, the new time constant T_2 comes from the electrical characteristics of the motor and is usually about one tenth the value of T_1.

The range of initial values of $\dot{\theta}$ at $\theta = 0$, for which θ will come to rest at the smallest positive value given by Eq (16), may be found from the appropriate conservation principle. The bound given by Eq (18) is conservative because damping is neglected.

$$\dot{\theta}_m \leq \left(\frac{2K}{T} - \frac{\pi\omega_1}{T}\right)^{1/2} \tag{18}$$

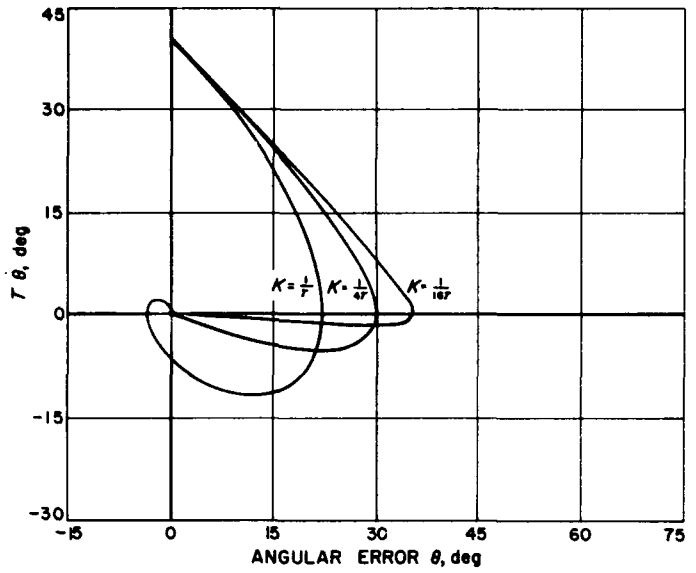


Fig. 21. Examples of phase trajectories for $\omega_1 = \gamma = 0$

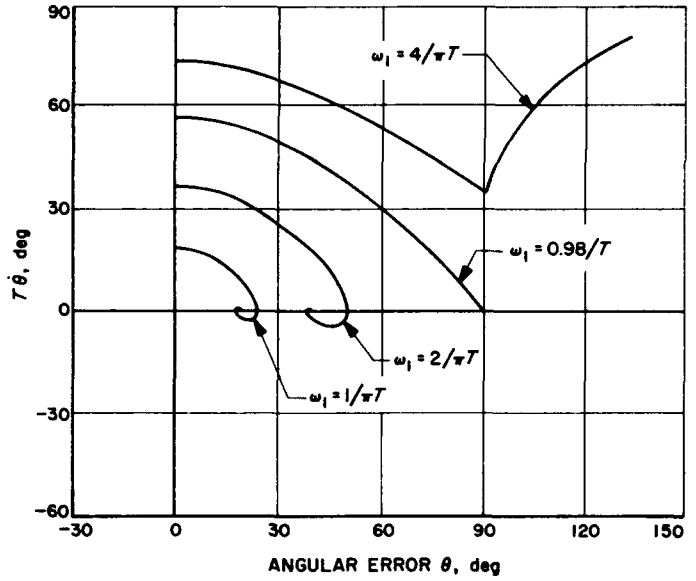


Fig. 22. Examples of phase trajectories for $K = 1/T$, $\gamma = 0$

It can be seen from Eq (18) that there is a maximum value of ω_1 given by Eq (19) beyond which a value of $\dot{\theta}_m$ cannot be found by this method.

$$\omega_1 \leq \frac{2K}{\pi} \tag{19}$$

Received signals having an ω_1 greater than this can sometimes still be tracked, however, if there is sufficient damping and if Eq (14) is satisfied. In this case it may be necessary for θ to go through several revolutions before coming to rest.

c. Determination of system stability. When a practical system such as this involving nonlinearities is actually built, it may be necessary to determine whether the system with certain types of tandem compensation is stable. Fortunately because of a theorem of Lyapunov, the linear analysis is useful here. Since the nonlinear characteristic ϵ is for all practical purposes an analytic function of θ, the theorem guarantees that the closed-loop nonlinear system will be stable for all values of gain K for which the previously described linearized system is stable. Sometimes a more exact knowledge of the degree of stability is necessary. For these cases the describing function method is useful.

Several describing functions have been computed for the non-linearity ϵ. Some of these which are useful for sinusoidal inputs are shown in Fig 23. Since the nonlinearity generates odd harmonics of a sinusoidal input, the third and fifth harmonics are also shown for comparison purposes. In general harmonics remain small until the amplitude of the input becomes excessive. Describing functions can be used with Nyquist, root locus, or gain phase diagrams to predict stability approximately. The gain phase diagrams are perhaps the quickest to use.

The two open-loop characteristics given by Eqs (1) and (17) have been plotted in Fig 24. The gain K in both cases is normalized to $1/T$. Superimposed on the -180 deg phase-shift line is the locus of the negative reciprocal of the describing function. Good stability is obtained when a gain K is chosen that allows an adequate margin of safety between these two curves. The gain phase diagram is also one of the quickest ways to check the design of compensating filters. For this purpose the filter characteristic can be included in the gain phase plot for the open-loop system.

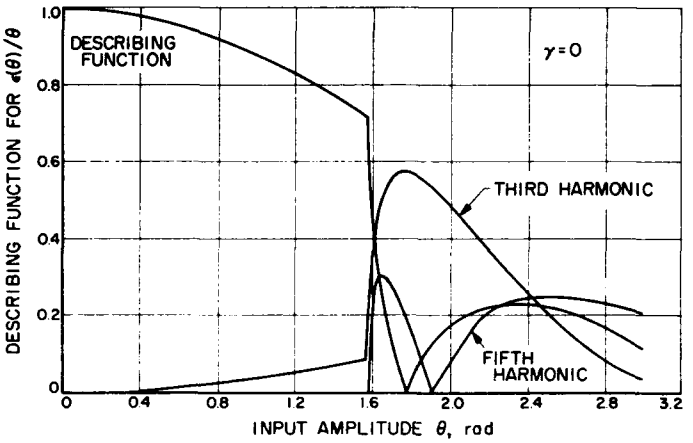


Fig. 23. Describing function for sinusoidal inputs

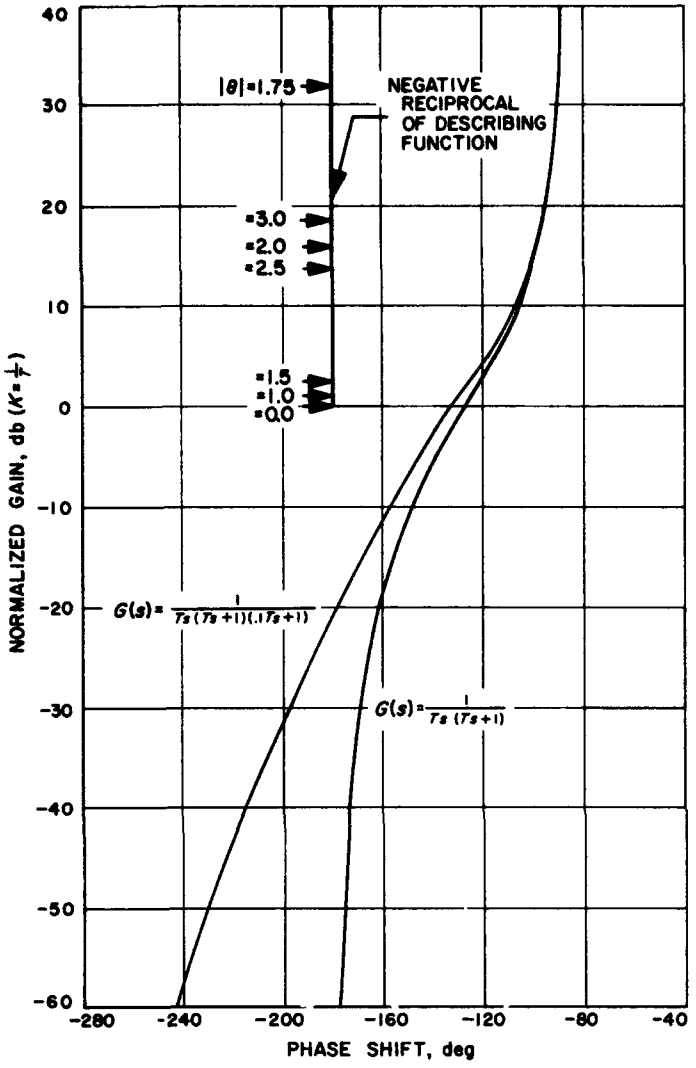


Fig. 24. Gain phase diagram

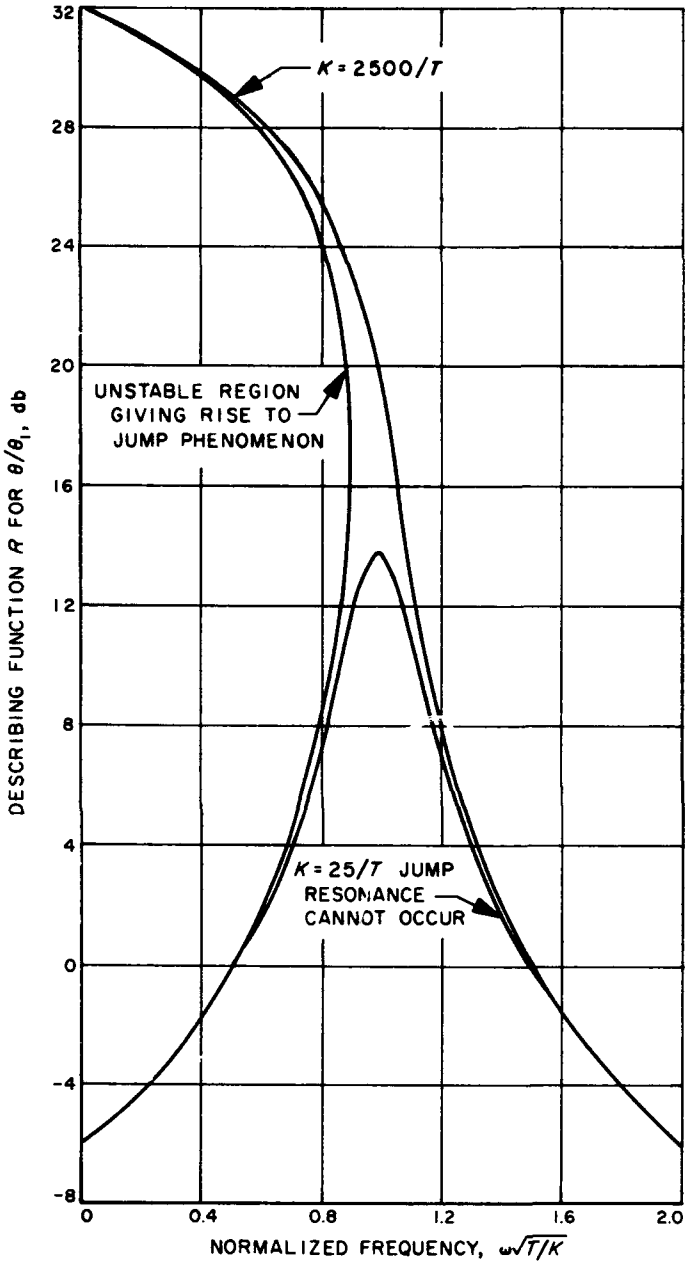


Fig. 25. Jump resonance phenomenon with input $\theta_1 = 0.1 \sin \omega t$

There is an interesting phenomenon called jump resonance which occurs in nonlinear systems. It has no counterpart in linear systems. In this system a jump resonance can occur when the damping is sufficiently small. Effects of this type of instability are shown in Fig. 25. These curves can be found in two ways. A perturbation method can be applied to the differential equation, Eq (4), giving an approximate expression such as Eq (20). Here R is used to represent the ratio of the fundamental amplitude of θ to the fundamental amplitude of θ_1. The input θ_1 is a pure sine wave of amplitude A and frequency ω.

$$1 - \frac{A^2R^2}{16} \pm \frac{1}{2}\left(\frac{1}{R^2} - \frac{1}{KT}\right)^{1/2} = \left(\frac{T}{K}\right)^{1/2} \omega \qquad (20)$$

It is necessary to take higher-order terms into account to get better approximations than Eq (20). Another alternate method is to use the describing functions shown in Fig 23 to obtain the system gain and predict the "describing function" R. This method can include the higher harmonics by a suitable iterative procedure.

At certain frequencies the value of the error signal is multiple-valued (Fig 25). When there are several values of the error, any slight disturbance will cause the system to seek the highest one. Jumps in output occur at different frequencies, depending upon whether the frequency of the input sine wave is being increased or decreased. Since inputs θ_1, which are useful, are either constants or

are of the form $\theta_1 = \omega_1 t$, purely sinusoidal inputs are really never encountered in practice. Although this type of instability can also be excited by noise, it cannot cause trouble. Appreciable jumps can only occur when the gain is so large that the system has negligible damping. Good performance can always be obtained by adjusting the gain to a value which gives about half critical damping.

G. Frequency Acquisition Techniques

A. J. Viterbi

The problem of determining the presence of a sinusoid in additive gaussian noise and of estimating its frequency is of considerable importance in the fields of communications and radar and has been treated by several authors (Refs 22, 23, 24). It may be stated simply as follows: It is desired to estimate the value of an unknown fixed frequency ω of the noisy received signal

$$y(t) = (2S)^{1/2} \sin(\omega t + \theta) + n(t) \tag{1}$$

where S is the received signal power; ω is the unknown fixed frequency, which is uniformly distributed over a given region; θ is the unknown fixed phase, uniformly distributed in the region $0 \leq \theta \leq 2\pi$; and $n(t)$ is white gaussian noise of zero mean and one-sided spectral density N_0. One approach is to assume an arbitrary initial frequency and phase, compare these with the received signal by a phase comparison device, and use the error estimate to correct the initial assumption. This concept is the basis of the phase-locked loop; its primary disadvantage is the excessive amount of time required to determine or lock onto the correct frequency and phase. In fact it can be shown (Ref 25) that, even in the absence of noise, the time τ required to achieve frequency lock is approximately

$$\tau \approx \frac{(\omega - \omega_I)^2}{\omega_n^3} \text{ sec} \tag{2}$$

where ω_I is the initial assumption (i.e., the initial frequency of the local oscillator) and ω_n is the natural frequency of the loop which is proportional to its noise bandwidth. Clearly then, if the noise level is high, ω_n must be made small and consequently the lock-in time becomes excessive. On the other hand, since a phase-locked loop is generally used in any case to track the received signal once its frequency is determined, it is clear that ω need not be determined precisely. Instead, it is sufficient to reduce the range of uncertainty to the point where the lock-in time of the phase-locked loop becomes reasonable. From Eq (2), it follows that reduction of the uncertainty by a factor of 10, say, reduces the maximum lock-in time by a factor of 100.

Let us first consider the related problem wherein one is given the received signal

$$y(t) = (2S)^{1/2} \sin(\omega_i t + \theta) + n(t)$$

where θ is uniformly distributed from 0 to 2π, but ω_i is equally likely to be any one of a set of N frequencies $(\omega_1, \omega_2, \cdots, \omega_N)$. Woodward (Ref 22) has shown that sufficient statistics for determining the *a posteriori* probabilities of the frequencies (without regard to phase) are obtained by forming the quantities

$$R_i^2 = \left[\frac{(2)^{1/2}}{T}\int_0^T y(t) \cos \omega_i t \, dt\right]^2 + \left[\frac{(2)^{1/2}}{T}\int_0^T y(t) \sin \omega_i t \, dt\right]^2 \quad (i = 1, 2, \cdots, n) \tag{3}$$

The most likely frequency is that for which this quantity is a maximum. The mechanization of a single element of the receiver bank is shown in Fig. 26a. A mechanization which is essentially equivalent provided $\omega_i T >> 1$ consists of a filter tuned to the frequency ω_i followed by an envelope detector (a linear rectifier and low-pass filter is generally sufficient). This is shown in Fig. 26b.

As Turin (Ref 26) has pointed out, the maximum likelihood estimation problem is really the continuous analog of maximum likelihood detection. That is, one allows the number of possible frequencies, N, to go to infinity and correspondingly the number of tuned filters to become infinite. The validity of this heuristic argument is substan-

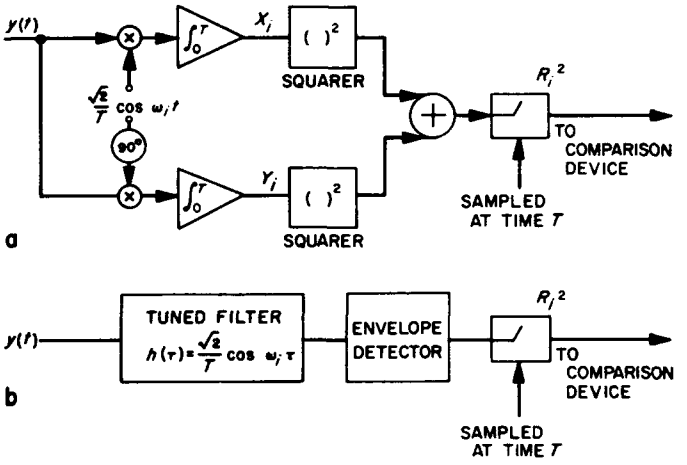


Fig. 26. Alternate mechanizations of an element of the receiver bank

tiated by rigorous developments (Ref 23). Actually if the observation time is T sec and the initial region of uncertainty is W cps, one can expect to reduce the uncertainty to not less than $1/T$ cps. This will require no more than WT tuned filters, as will be shown subsequently.

Let the region of uncertainty W be subdivided into N parts, and let there be $N + 1$ filters tuned at intervals of W/N cps from one band edge to the other (Fig. 27). Denote the frequencies $f_1 \ldots f_{N+1}$ and let $\omega_i = 2\pi f_i$.

Then

$$\omega_i - \omega_{i-1} = \frac{2\pi W}{N}$$

Let the received signal be that given by Eq (1). Referring to Fig. 26a, we see that the quantities X_i and Y_i are given by

$$X_i = \frac{(S)^{1/2}}{T}\int_0^T \sin\left[(\omega - \omega_i)\,t + \theta\right] dt + \frac{(2)^{1/2}}{T}\int_0^T n(t)\cos\omega_i t\, dt$$
$$Y_i = \frac{(S)^{1/2}}{T}\int_0^T \cos\left[(\omega - \omega_i)\,t + \theta\right] dt + \frac{(2)^{1/2}}{T}\int_0^T n(t)\sin\omega_i t\, dt \tag{4}$$

where the double frequency terms have been dropped under the assumption that the frequencies are very high. Then, taking the noise to have zero mean, we have for the means, variances, and co-variances of X_i and Y_i the following expressions:

$$\overline{X_i} = \frac{(S)^{1/2}\left\{\cos\theta - \cos\left[(\omega - \omega_i)\,T + \theta\right]\right\}}{T(\omega - \omega_i)}$$
$$\overline{Y_i} = \frac{(S)^{1/2}\left\{\sin\left[(\omega - \omega_i)\,T + \theta\right] - \sin\theta\right\}}{T(\omega - \omega_i)} \tag{5}$$
$$\sigma^2 = \overline{(X - \overline{X_i})^2} = \overline{(Y - \overline{Y_i})^2} = \frac{N_0}{2T}$$
$$\overline{(X - \overline{X_i})(Y - \overline{Y_i})} = 0$$

Of equal importance are the cross-moments between X_i and Y_i and X_j and Y_j, the variables of the jth tuned filter. These are

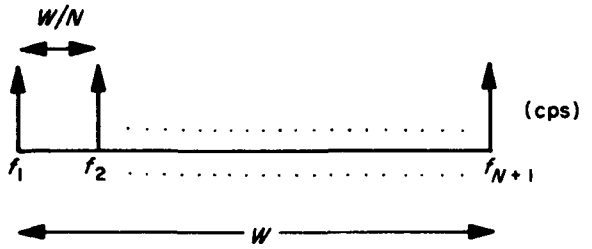


Fig. 27. Region of frequency uncertainty and location of tuned filter frequencies

$$\overline{(X_i - \overline{X_i})(X_j - \overline{X_j})} = \frac{N_0}{2T}\frac{\sin(\omega_j - \omega_i)\,T}{(\omega_j - \omega_i)\,T}$$
$$\overline{(X_i - \overline{X_i})(Y_j - Y_j)} = \frac{N_0}{2T}\frac{\left[1 - \cos(\omega_j - \omega_i)\,T\right]}{(\omega_j - \omega_i)\,T} \tag{6}$$

From this it is clear that if the frequencies, ω_i, are spaced at equal intervals of $2\pi/T$ radians, these quantities become zero so that the outputs of the various filters are independent. In other words, the output of one filter yields no information about the output of any other. In fact, one would expect by using arguments based on the sampling theorem (Ref 24) that it would be futile to attempt to estimate frequency to a tolerance of less than $1/T$ cps based on an observation time of T sec.

The variables X_i and Y_i have Gaussian distributions. Hence, using Eqs. (5) and (6) their joint probability density is given by

$$p(X_i, Y_i) = \frac{1}{2\pi\sigma^2}\exp - \frac{1}{2\pi\sigma^2}\left[(X_i - X_i)^2 + (Y_i - Y_i)^2\right]$$

Converting to polar coordinates by making the substitutions

$$X_i = R_i\cos\theta_i \quad \text{and} \quad Y_i = R_i\sin\theta_i$$

we obtain

$$p(R_i, \theta_i) = J\left(\frac{X_i, Y_i}{R_i, \theta_i}\right) p(X_i, Y_i) \tag{7}$$
$$= \frac{R_i}{2\pi\sigma^2}\exp - \frac{1}{2\sigma^2}\left\{R_i^2 + S\left[\frac{\sin(\omega - \omega_i)\,T/2}{(\omega - \omega_i)\,T/2}\right]^2 - 2(S)^{1/2}R_i\left[\cos\theta_i(\cos\theta - \cos\theta') + \sin\theta_i(\sin\theta' - \sin\theta)\right]\right\}$$

where

$$\theta' = (\omega - \omega_i)\,T + \theta$$
$$= \frac{R_i}{\pi N_0/T}\exp - \left\{\frac{R_i^2}{N_0/T} + \frac{ST}{N_0}\left[\frac{\sin(\omega - \omega_i)\,T/2}{(\omega - \omega_i)\,T/2}\right]^2\right\} \exp\frac{4R_i(S)^{1/2}\sin(\omega - \omega_i)\,T/2\sin\theta''}{N_0(\omega - \omega_i)}$$

where $$\theta'' = \theta_i + \theta + (\omega - \omega_i)\,T/2$$

It is seen in Figs. 26a and 26b that the sufficient statistic required to estimate the frequency is R_i. To determine its probability density we must integrate Eq. (7) with respect to θ_i. Recalling that the integral representation of the modified Bessel function of zeroth order is

$$I_0(x) = \frac{1}{2\pi}\int_0^{2\pi} e^{x\cos\theta}\, d\theta$$

we obtain

$$p(R_i) = \frac{R_i}{N_0/2T} \exp - \left\{ \frac{R_i^2}{N_0/T} + \frac{ST}{N_0} \left[\frac{\sin(\omega - \omega_i)T/2}{(\omega - \omega_i)T/2} \right]^2 \right\} I_0 \left[\frac{2R_i (ST)^{1/2} \sin(\omega - \omega_i)T/2}{N_0 (\omega - \omega_i)T/2} \right]$$

$$= \frac{v}{\sigma} \exp - \left(\frac{v^2 + a^2}{2} \right) I_0(av) \tag{8}$$

where

$$v = \frac{R_i}{(N_0/2T)^{1/2}} = \frac{R_i}{\sigma}, \quad a = \left(\frac{2ST}{N_0} \right)^{1/2} \frac{\sin(\omega - \omega_i)T/2}{(\omega - \omega_i)T/2}$$

We are now in a position to evaluate the performance of this frequency estimation technique at a given S/N_0 ratio with a given observation time T, and for an initial frequency uncertainty W. If we were merely interested in determining that frequency, ω_i, which is with maximum probability closest to the received signal frequency, ω, we would select the frequency which corresponds to the greatest R_i. However, there is another eventuality that must be avoided, namely, the possibility that no signal was present during the observation time. This may be effected by placing a threshold device after each filter so that if all the R_i's are less than the threshold, γ, the decision is made that no signal was present. The threshold device may then be used for the dual purpose of avoiding false alarms and of discriminating between the largest output and the other outputs when a signal is present. Also, this mechanization is generally simpler than that which determines the greatest output. The value of γ should depend on the probability that no signal is present, the cost of a false alarm, and the cost of an error. Unfortunately, these quantities are usually not known and the assignment of costs is often subjective. In the example to follow, we shall simply select a suitable threshold and evaluate the resulting performance.

The first step is to derive expressions for the probability of false alarm and error. If no signal is present, the distribution of R_i is given by

$$p(R_i) = \frac{R_i}{N_0/2T} \exp - \frac{R_i^2}{N_0/T} \tag{9}$$

This can be obtained by repeating the derivation leading to Eq. (8) with no signal present. However, it follows directly from Eq. (8) by setting S equal to 0. Then the probability of false alarm is simply

$$P_F = 1 - \text{Prob}(R_1, R_2, \cdots, R_{N+1} < \gamma)$$

$$= 1 - \left[\int_0^\gamma \frac{R_i}{\sigma^2} \exp\left(- \frac{R_i^2}{2\sigma^2} \right) dR_i \right]^{N+1}$$

$$= 1 - \left[1 - \exp\left(- \frac{\gamma^2}{2\sigma^2} \right) \right]^{N+1} \tag{10}$$

The error probability computation is somewhat more involved. First of all, it would be unreasonable to require that only the tuned filter nearest the received frequency, ω, produce an output greater than the threshold, γ. The reason is that ω might lie $(\pi/2T) - \epsilon$ from filter frequency ω_i and $(\pi/2T) + \epsilon$ from filter frequency ω_{i+1} (Fig. 28), and as ϵ approaches zero,

$$\text{Prob}(R_{i+1} > \gamma) \to \text{Prob}(R_i > \gamma)$$

Hence, we shall define the probability of correct decision as the probability that when $\omega_i < \omega < \omega_{i+1}$ either R_i or R_{i+1} or both exceed γ and that all other outputs are less than γ. This ensures that if the correct decision is made the received frequency will be no more than $1/T$ cps from the filter frequency selected, and thus the uncertainty will have been reduced by a factor of N. Since all the outputs are independent this may be written as

$$1 - P_E = P_c = \left[\text{Prob}(R_i > \gamma) + \text{Prob}(R_i < \gamma) \times \text{Prob}(R_{i+1} > \gamma) \right] \prod_{j \neq i, i+1}^{N+1} \text{Prob}(R_j < \gamma) \tag{11}$$

The probability densities for the R_i's are given by Eq (8). Hence,

$$\text{Prob}(R_i > \gamma) = \int_\gamma^\infty p(R_i)\, dR_i = \int_{\gamma/\sigma}^\infty p(v)\, dv$$

$$= \int_{\gamma/\sigma}^\infty v \exp - \frac{(v^2 + a^2)}{2} I_0(va)\, dv$$

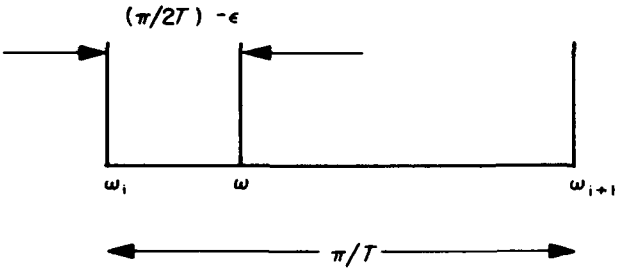


Fig. 28. Relative locations of received frequency and tuned filter frequencies

where

$$\nu = \frac{R_i}{\sigma} = \frac{R_i}{(N_0/2T)^{1/2}}$$

and

$$\alpha_i = (2ST/N_0)^{1/2} \frac{\sin(\omega - \omega_i)T/2}{(\omega - \omega_i)T/2} \qquad (12)$$

The integral

$$Q(\alpha, \beta) = \int_\beta^\infty \nu \exp - \frac{(\nu^2 + \alpha^2)}{2} I_0(\nu\alpha)\, d\alpha \qquad (13)$$

has been tabulated by J. I. Marcum (Ref 27). Using the terminology of Eq. (13) the error probability becomes

$$1 - P_E = \left\{Q(\alpha_i, \gamma/\sigma) + [1 - Q(\alpha_i, \gamma/\sigma)]\, Q(\alpha_{i+1}, \gamma/\sigma)\right\} \prod_{j \neq i, i+1}^{N+1} [1 - Q(\alpha_j, \gamma/\sigma)]$$

where

$$\alpha_j = \left(\frac{2ST}{N_0}\right)^{1/2} \frac{\sin(\omega - \omega_j)T/2}{(\omega - \omega_j)T/2} \qquad (14)$$

and

$$\sigma^2 = \frac{N_0}{2T}$$

It is clear that the error probability will depend on the received frequency, ω. In fact, because $Q(\alpha, \beta)$ is a monotonically increasing function of α, it is readily shown by examination of Eq. (14) that the worst case is that for which ω lies in the central interval; i.e.,

for N even,

$$\omega_{N/2} + \frac{\pi}{T} \leq \omega \leq \omega_{N/2+1} + \frac{\pi}{T}$$

for N odd,

$$\omega_{(N+1)/2} \leq \omega \leq \omega_{(N+3)/2}$$

However, within this interval it is not clear where P_E has a maximum. Equations (10) and (14) are sufficient to determine the performance as a function of S/N_0, observation time T, and the ratio of the threshold to the standard deviation, γ/σ.

We shall now consider an example. Let the initial uncertainty be W cps and the required final uncertainty be $W/4$ cps. This requires a bank of five tuned filters equally spaced from one band edge to the other and an observation time of $T = 4/W$ sec (the inverse of the final frequency uncertainty). It is desired that the probability of error be of the order of 10^{-2}. As was pointed out above, P_E varies with the position of the received frequency, ω, within the band of uncertainty, but the highest error probability occurs in the central interval,

$$\omega_3 - \frac{\pi}{T} \leq \omega \leq \omega_3 + \frac{\pi}{T}$$

Therefore, we evaluate Eq. (14) throughout this interval. It is found that, by using a threshold level of

$$\gamma = 5\sigma = 5(N_0/2T)^{1/2}$$

and a signal-to-noise ratio $S/N_0 = 69.5/T = 17.4\,W$, the error probability over the central interval is less than 1.8×10^{-2}, as shown in Fig. 29. At the same time the false alarm probability P_F (from Eq. 10) is 2×10^{-5}.

Further work is required to determine the optimum threshold level using this scheme. The primary drawback of the method is in the amount of equipment required when the final uncertainty must be a small fraction of the initial uncertainty (i.e., when N is large). However, the equipment problem can be avoided by using a single filter of the form of Fig. 26a and performing the operation serially (i.e., setting ω_i at ω_1 for T sec, at ω_2 for T sec, etc.). Of course, now the operation requires NT sec of observation instead of the original T *sec*. With this ex-

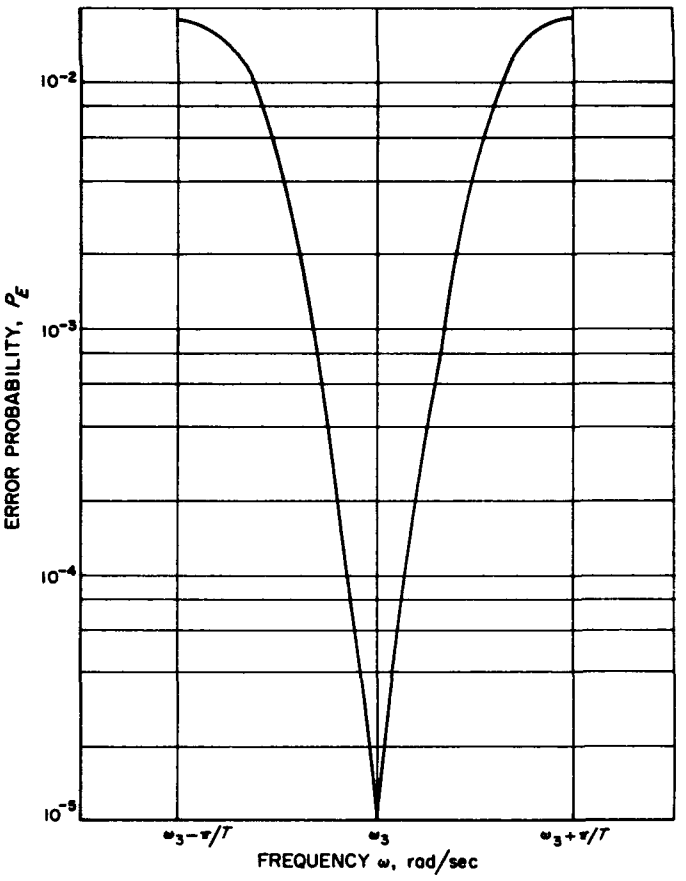


Fig. 29. Error probability as a function of received frequency in central region

panded observation time, it is quite possible that other techniques utilizing a single filter may be considerably more effective.

H. A Note on the Factorization Problem

T. Kailath

In a recent review article (Ref 28), Zadeh has summarized the few results known on the factorization problem and has pointed out some of the several unresolved questions on this topic.

The factorization problem is essentially the problem of solving the following non-linear integral equation

$$R(t,s) = \int_{-T_1}^{T_2} h(x,t)\, h(x,s)\, dx \qquad -T_1 \leq t, s \leq T_2 \tag{1}$$

$R(t,s)$ is a known symmetric positive definitive function, and $h(x,t)$ is to be determined. Either or both of T_1 and T_2 may be infinite or T_1 may be 0 and $T_2 = T$. $R(t,s)$ may be a difference kernel, $R(t,s) = \ddot{R}(t-s)$. There may be several constraints on the $h(x,t)$ functions; e.g.,

$$h(x,t) = f(x-t) \qquad f(t) = 0, t < 0 \tag{2}$$

or

$$h(x,t) = 0 \qquad x < 0 \tag{3}$$

The factorization problem arises in several fields, some of which are

a. Spectrum shaping techniques to "whiten" or otherwise "shape" certain non-stationary or stationary random processes.

b. Solution of a Fredholm integral equation of the first kind (especially the general Wiener-Hopf equation).

c. Simplification of the structure of certain ideal receiver configurations arising in multipath communications.

d. Certain problems of spectral estimation.

Finally, since so little is known about non-linear integral equations, any light shed on the problem might have a larger usefulness. In fact, as Zadeh points out, it still has not been shown that a solution to the problem of Eq. (1), with a kernel $R(t,s)$ of 2 variables, is at all possible in general.

In this note we shall discuss how this problem arises in multipath communication and show how a partial solution to the factorization problem may be obtained.

In a previous paper (Ref 29) it was shown that the receiver structure for the detection of Gaussian signals in Gaussian noise (of which multipath communication is a special case) was given by an expression of the form

$$\Lambda = \int_0^T \int_0^T y(t)\, y(s)\, R(t,s)\, dt\, ds \tag{4}$$

This formula can be implemented by passing $y(t)$ through a linear time-variant filter $R(t,s)$ and then cross correlating (i.e., multiplying and integrating) the filter output against $y(t)$ again. The multiplication operation required for the cross-correlation operation is generally the most difficult part of the implementation, and it would be very convenient if it could be avoided altogether. One method of achieving this is by factoring $R(t,s)$,

$$R(t,s) = \int_0^T h(x,t)\, h(x,s)\, dx \tag{5}$$

so that now

$$\Lambda = \int_0^T dx \int_0^T h(x,t)\, y(t)\, dt \int_0^T h(x,s)\, y(s)\, ds \tag{6}$$

If we impose the restriction

$$h(x,t) = 0 \qquad x < t \tag{7}$$

which insures that the filter $h(x,t)$ is physically realizable (no response before an input), we can write

$$\Lambda = \int_0^T \left[dx \int_0^x h(x,t)\, y(t)\, dt \right]^2 \tag{8}$$

This structure can be implemented as follows: Pass $y(t)$ through a time-varying filter $h(x,t)$, square the filter output, and integrate the result from time 0 to time T. In this implementation we have replaced the multiplication by a squaring operation, and the squaring can usually be done quite satisfactorily by a rectifier diode.

A similar simplification of structure can be obtained in certain planetary–surface mapping techniques (Ref 30).

Therefore the problem is to find a function $h(x,t)$ that satisfies Eq. (5). $R(t,s)$ does happen to be a symmetric positive definite kernel, so that Eq. (5) is really identical with the non-linear integral equation, Eq. (1), with $T_1 = 0, T_2 = T$. Now even *linear* integral equations of the first kind have only been solved in general for processes with rational spectra. Therefore, we shall first try to solve Eq. (5) for $R(t,s)$ corresponding to a first-order rational spectrum process.

$$R(t,s) = e^{-c|t-s|} \tag{9}$$

Since $R(t,s)$ is now effectively a function of a single variable, $u = t - s$, it might appear that the time-variant filter $h(x,t)$ could be replaced by a time-invariant filter, $f(x-t)$. Since time-invariant filters are usually much simpler in form than time-variant filters, it would be convenient if we could solve the equation

$$e^{-c|t-s|} = \int_0^T f(x-t)\,f(x-s)\,dx \qquad 0 \leq t, s \leq T \tag{10}$$

$$f(x-t) = 0 \qquad x < t \tag{11}$$

However, our efforts so far have been unsuccessful. On the other hand, when $T \to \infty$, a solution is quite easily obtained by Fourier transforms. This suggested, by analogy with the particular integral ("steady state") solution of differential equations, that we use the solution obtained for $T \to \infty$ as a particular integral type term and substitute this solution into the integral equation, Eq. (10). We thus obtain

$$e^{-c|t-s|} = 2c\int_0^T g(x-t)\,g(x-s)\,dx + 2e^{-c(T-t)}\,e^{-c(T-s)} \tag{12}$$

where

$$\begin{aligned} g(x-t) &= e^{-c(x-t)} \qquad x > t \\ &= 0 \qquad\qquad x > t \end{aligned} \tag{13}$$

is the solution to Eq. (10) when $T \to \infty$. The term $2e^{-c(T-t)}\,e^{-c(T-s)}$ is an end-effect ("transient") term that vanishes as $T \to \infty$. We can write this end-effect term in the form

$$2e^{-c(T-t)}\,e^{-c(T-s)} = \int_0^T g'(x,t)\,g'(x,s)\,dx \tag{14}$$

where

$$g'(x,t) = (2)^{1/2}\,e^{-c(T-t)}\,f(x) \tag{15}$$

and

$$\int_0^T f^2(x)\,dx = 1 \tag{16}$$

Now, however, we cannot have $g'(x,t) = 0, x < t$ unless we can set $f(x) = 0, x < t, 0 \leq t \leq T$. This is apparently impossible. If we abandon this restriction, but set

$$\int_0^T e^{-cx} f(x)\,dx = 0 \tag{17}$$

[note that this is compatible with Eq. (16)], we can write

$$e^{-c|t-s|} = \int_0^T h(x,t)\,h(x,s)\,dx \qquad 0 \leq t, s \leq T \tag{18}$$

where

$$h(x,t) = g(x-t) + g'(x,t) \tag{19}$$

with $g(x-t)$ and $g'(x,t)$ defined as above. However, now $h(x,t) \neq 0, x < t$. Thus we have been unable to factor $e^{-c|t-s|}$ into time-invariant and/or realizable factors. The fact that this cannot apparently be simply done for a simple exponential-type positive definite kernel leads us to conjecture the general impossibility of such a factoring. We make this last statement largely in the hope that it may spur someone to find a counterexample. We should point out that for certain (and not uninteresting) classes of kernels we have been able to find a solution. It is an interesting fact that in the discrete parameter case when we deal with (finite or infinite) matrices rather than integral equations a solution is always possible. This suggests that a limiting process might yield the continuous parameter solution. [In this connection it may be mentioned that Middleton's (Ref 31) so-called Bayes matched-filter-of-the-first-kind, which purports essentially to be the solution of Eq. (10), is incorrect.]

While we have not succeeded directly in our attempt at factorization, it turns out that (at least for rational spectrum processes) we can still find a solution to our desired multipath receiver transformation. Thus if we insert Eq. (18) in the receiver formula, Eq. (4), we obtain

$$\Lambda = 2c\int_0^T \left[\int_0^x e^{-c(x-t)}\,y(t)\,dt\right]^2 dx + 2\left[\int_0^T e^{-c(T-t)}\,y(t)\,dt\right]^2 \tag{20}$$

Notice now that the second term on the right side in Eq. (20) is the value of

$$\left[\int_0^x e^{-c(x-t)}\,y(t)\,dt\right]^2$$

at $x = T$. Therefore Λ can be obtained as shown in Fig. 30: the combination of a sampling switch [$\delta(t)$ is the Dirac delta function] and an amplifier provide the end-effect term, and we essentially have a filter-squarer receiver and have avoided any correlation operations.

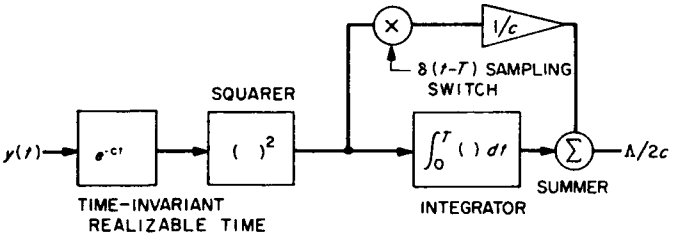


Fig. 30. Implementation for the ideal receiver that avoids a correlation operation

References

1. Shannon, Claude E., "A Mathematical Theory of Communication," *Bell System Technical Journal*, Vol 27, pp 379–423; 623–656, 1948.
2. Peterson, W. Wesley, *Error-Correcting Codes*, John Wiley & Sons, New York, 1961.
3. Assmuss, E. F. Jr., and Mattson, H. F., *Error-Correcting Codes: An Axiomatic Approach*, Sylvania Electronic Systems, ARM 269, Waltham, Mass., 1961.
4. Viterbi, A. J., *On Coded Phase-Coherent Communications*, Technical Report 32-25, Jet Propulsion Laboratory, Pasadena, California, 1960.
5. Feller, W., *An Introduction to Probability Theory and Its Applications*, 2nd ed, Wiley & Sons, New York, 1958.
6. Shannon, C. E., "Probability of Error for Optimum Codes in a Gaussian Channel," *Bell System Technical Journal*, Vol 3, pp 611–656, May 1959.
7. Shannon, C. E., "Certain Results in Coding Theory," *Information and Control*, Vol I, pp 1–25, September 1958.
8. Cramer, H., "Sur une Nouveau Theoreme-limite de la Theorie des Probabilities," *Actualities Scientifiques et Industrielles*, No. 736, Herman & Cie, Paris, 1938.
9. Mattson, H. F., and Solomon, G., "A New Treatment of Bose–Chaudhuri Codes." *S.I.A.M. Journal*, pp 654–669, December 1961.
10. Peterson,W.W., "Encoding and Error-Correction Procedures for the Bose-Chaudhuri Codes," *IRE Transactions, IT-6*, pp 459–470, 1960.
11. Peterson, W. W., *Error-Correcting Codes*, The M.I.T. Press and John Wiley & Sons, Inc., 1961.
12. Solomon, G., "A Note on A New Class of Codes," *Information and Control*, Vol 4, No. 4, pp 364–370, December 1961.
13. Mood, A. W., "The Distribution Theory of Runs," *Annals of Mathematical Statistics*, Vol 11, pp 367–392, 1940.
14. Golomb, S. W., *Sequences with Randomness Properties*, Report 639498, Glenn L. Martin Co., Baltimore, Md., June 14, 1955.
15. Baumert, L. D., *et al*, *Coding Theory and its Applications to Communications Systems*, Technical Report No. 32-27, Jet Propulsion Laboratory, Pasadena, Calif., March 31, 1961.
16. Zierler, N., "Linear Recurring Sequences," *Journal of the Society for Applied Mathematics*, Vol 7, No. 1, March 1959.
17. Hall, M., Jr., "A Survey of Difference Sets," *Proceedings of the American Mathematical Society*, Vol 7, pp 975–986, 1956.
18. Victor, W. K., Stevens, R., and Golomb, S. W., *Radar Exploration of Venus*, Technical Report 32-132, Jet Propulsion Laboratory, Pasadena, Calif., August 1, 1961.
19. Kuiper, G. P., "Determination of the Pole of Venus," *Astrophysical Journal*, Vol 120, p 603, 1954.
20. Jaffe, R. M., and Rechtin, E., *Design and Performance of Phase-Lock Circuits Capable of Near-Optimum Performance over a Wide Range of Input Signals and Noise Levels*, External Publication 263, Jet Propulsion Laboratory, Pasadena, Calif., December 1, 1954.

References (Cont'd)

21. Jury, E. I., *Sampled-Data Control Systems*, John Wiley and Sons, New York, 1958.

22. Woodward, P. M., *Probability and Information Theory, With Applications to Radar*, McGraw-Hill, N. Y., 1955.

23. Helstrom, C. W., *Statistical Theory of Signal Detection*, Pergamon Press, N.Y., 1960.

24. Turin, G. L., *A Review of Correlation, Matched-Filter, and Signal-Coding Techniques, with Emphasis on Radar Applications*, Vol I, Technical Memorandum 559, Hughes Systems Development Laboratories, April 1957.

25. Viterbi, A. J., "Acquisition and Tracking Behavior of Phase-Locked Loops," *B.I.P. Symposium Proceedings on Active Networks and Feedback Systems*, April 1960.

26. Turin, G. L., "An Introduction to Matched Filters," *IRE Transactions on Information Theory*, Vol IT-6, pp 311–329, June 1960.

27. Marcum, J. I., *Table of Q Functions*, Rand Corporation Research Memorandum RM-339, January 1, 1950.

28. Zadeh, L. A., "Time-Varying Networks—I," *Proceedings of the IRE*, Vol 49, pp 1488–1503, October 1961.

29. Kailath, T., "Correlation Detection of Signals Perturbed by a Random Channel," *IRE, PGIT-6;* pp 331–336, June 1960.

30. Kailath, T., paper to be presented at *Brussels Symposium on Information Theory*, September 1962.

31. Middleton, *Introduction to Statistical Communication Theory*, McGraw-Hill Book Co., New York, 1960.

VII. Communications Elements Research

A. Low Noise Amplifiers

C. T. Stelzried, C. J. Finnie, and W. H. Wells

1. Instrumentation for Noise Temperature Measurement, C. T. Stelzried

The precise measurement of absolute noise temperature has been investigated at JPL in parallel with the work on low noise antennas and amplifiers. The sensitivity of a "total power" radiometer is determined by the thermal fluctuations of the total system equivalent noise temperature and the amplifier gain instability as shown by

$$\Delta T = \frac{T_s}{(2)^{1/2}}\left(\frac{1}{(\tau B)^{1/2}} + \frac{\Delta G}{G}\right)$$

where

ΔT = rms fluctuations in the temperature measurement

T_s = total system equivalent noise temperature

τ = low-pass filter time constant

B = system bandwidth, cps

$\frac{\Delta G}{G}$ = rms system gain change

Schemes are under investigation to eliminate the effect of amplifier gain instability. This is not necessarily a cure-all for radiometer ills because a penalty must be paid in increased thermal jitter, increased circuit complexity, or both. A switching radiometer, as shown in Fig. 1 (assuming perfect switches and noise balance), has a sensitivity (Ref. 1) of

$$\Delta T = T_s\left(\frac{2}{\tau B}\right)^{1/2}$$

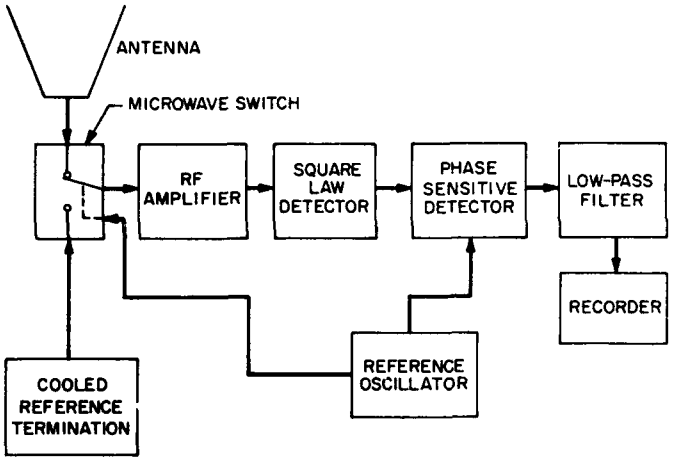


Fig. 1. Typical switching radiometer

This indicates that performance will not be improved over a total power radiometer when

$$\frac{\Delta G}{G} < \frac{1}{(\tau B)^{1/2}}$$

or rearranging

$$\tau < B\left(\frac{\Delta G}{G}\right)^2$$

The total power radiometers (with mixer front ends) constructed to date have achieved a value for $\Delta G/G$ of approximately 1/1000. With a receiver bandwidth of 2 mc, radiometer sensitivity will not be improved with a switching radiometer with a time constant less than about 2 sec [2 mc $(1/1000)^2$]. The advantage of each mode of radiometer operation is a question of the application. A radio-astronomy drift curve will typically require a radiometer time constant less than 1/10 (Ref. 2) the time of the drift curve. For a 0.1-deg antenna beam width, this is about 2.5 sec. The required radiometer gain stability must be maintained for several minutes (the total measurement time).

The switching radiometer of Fig. 2 was implemented at 960 mc using an American Electronics Laboratory (AEL) diode switch, a series resonant 960 mc degenerate parametric amplifier (*RS 36-10*), 416B triode amplifiers (*RS 36-11*), and a modified Model 14 Beckman DC amplifier. The equivalent noise temperature of the receiver was about 300°K (with 10 db paramp gain). When the input switch is closed (diode forward biased), the antenna is connected to the amplifier input. When open (diode reverse biased), signal from the reference termination

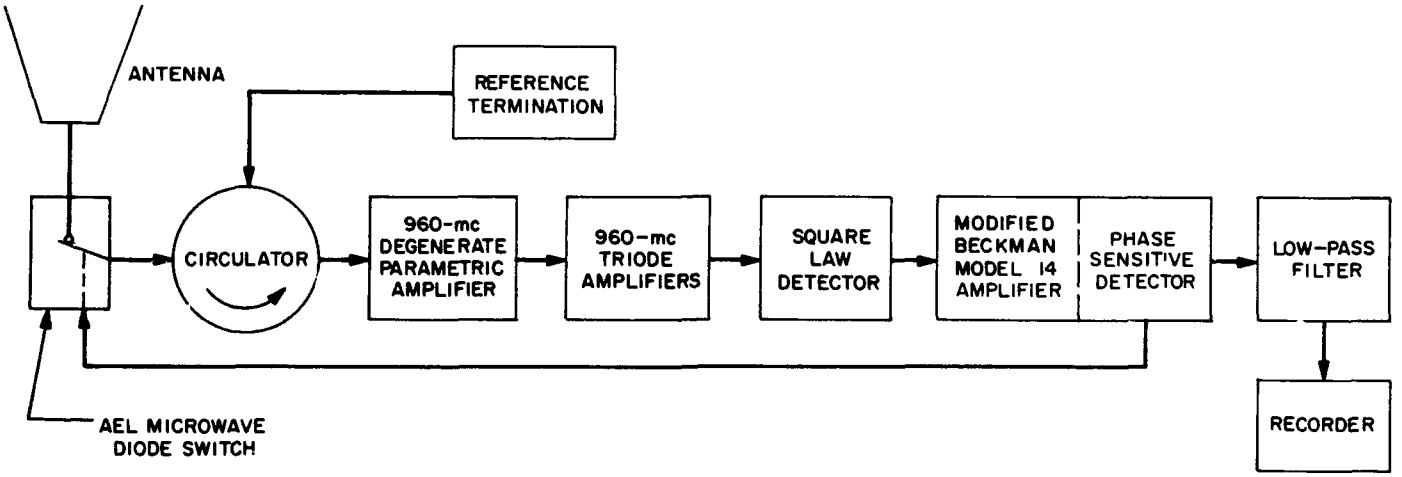


Fig. 2. 960 Mc switching radiometer

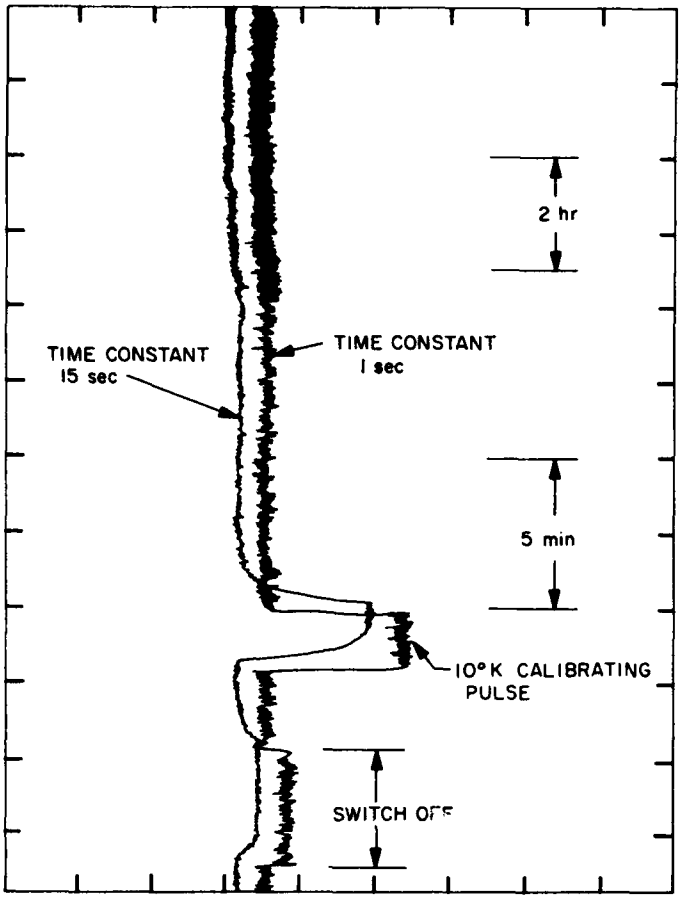


Fig. 3. 960 Mc radiometer stability recording switching between two ambient terminations

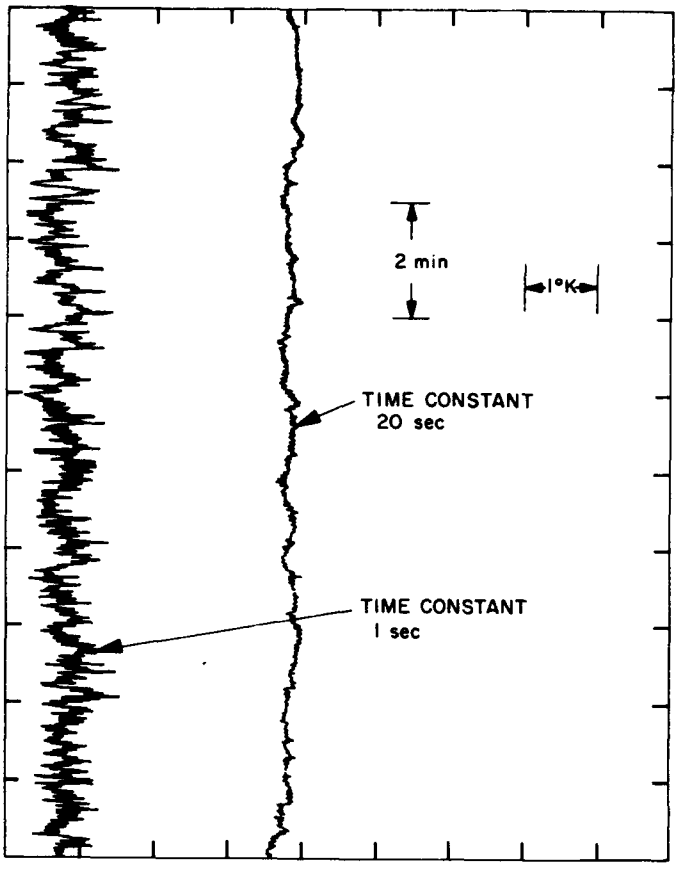


Fig. 4. 960 Mc radiometer stability recording switching between two liquid-nitrogen cooled terminations

connected to the circulator is reflected into the amplifier input. The AEL switch is controlled by the mechanical cam switching arrangement normally used at the input of the Beckman DC amplifier. The detected output of the radiometer is fed directly into the 8 cps AC amplifier (bypassing the input switch) of the Beckman and then into a phase sensitive detector using the synchronous cam-controlled switch at the output. A recording using a dual channel recorder with a 15 sec and 1 sec time constant is shown in Fig. 3. The two ambient loads used result in a 600°K total system temperature. A short run was made for comparison with no switching on the input. Since the jitter is substantially the same when the switch is turned on, the switch is not contributing to any instability (to the resolution of the recording). Both slow and fast recorder speeds were used to demonstrate short and long term gain stabilities on the order of 0.1 and 1°K, respectively.

In further experimentation, the Beckman cam-operated switch was replaced with a square-wave generator and a diode full-wave phase detector. Fig. 4 shows a recording with this arrangement and two liquid-nitrogen cooled terminations (resulting in a 400°K system noise temperature) on the radiometer input. The theoretical rms fluctuation with the 20-sec time constant is

$$\Delta T = \frac{400}{(2)^{1/2}} \frac{2}{[(20)(4\,mc)]^{1/2}} = 0.065°K$$

2. Lasers, *C. J. Finnie and W. H. Wells*

The ruby laser has been operated in an improved flash head shown in Fig. 5. The flash lamp is a helical FT 524 (General Electric). A gas flow system is included for optional cooling of the ruby with cold nitrogen gas boiled off of liquid nitrogen.

Various optical configurations are being tested in an attempt to observe laser action in rare earth chelates. The chelates are dissolved in organic glass, i.e., a frozen mixture of organic solvents. Mixtures tried thus far have not exhibited laser action.

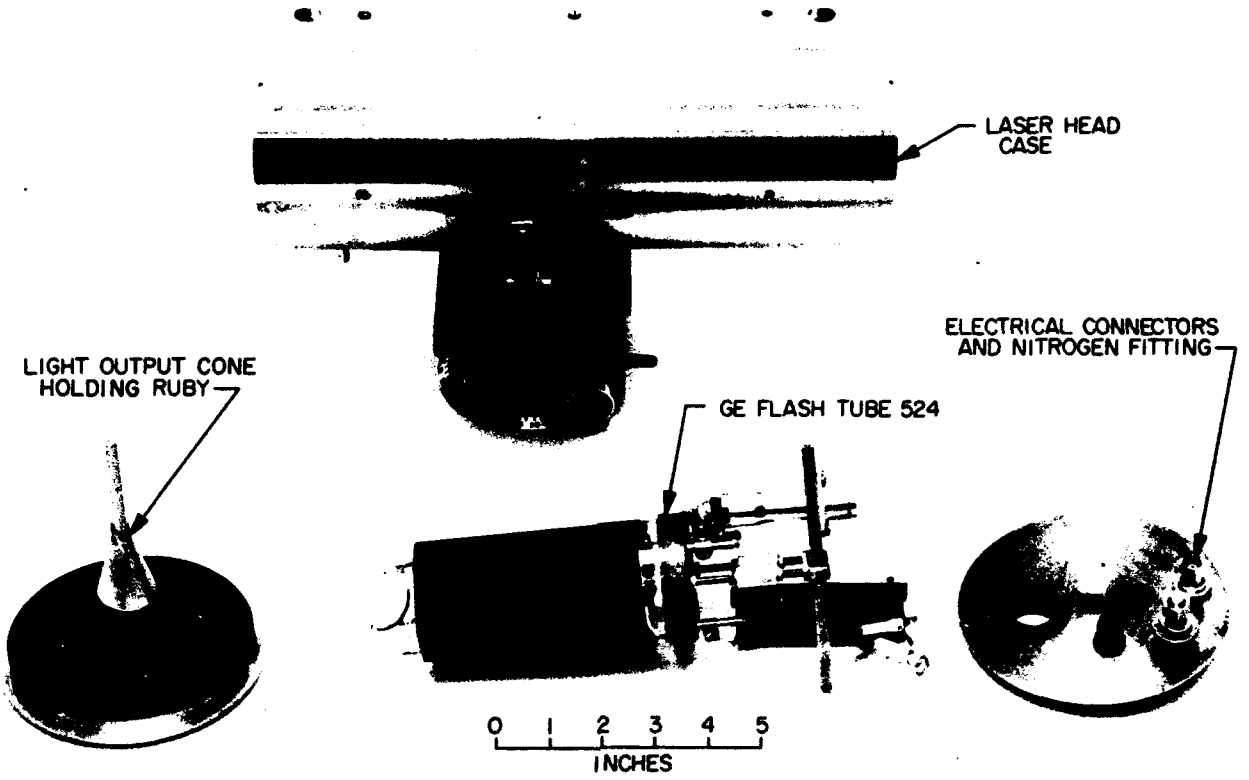


Fig. 5. Ruby laser and improved flash lamp

B. Antennas for Space Communications

P. D. Potter, T. Otoshi, and W. V. T. Rusch

1. *Antenna Feed Research*, P. D. Potter and T. Otoshi

a. Dual-channel X-Band radiometer. As an accurate means for far field antenna gain measurement, the application of radio astronomy techniques has proved very efficacious. These measurements basically involve determination of the ratio of the noise power received by a large antenna from a particular radio source to the noise power received by another small antenna whose gain is accurately known. For sources whose angular size is small compared to the beamwidths involved, a measurement of the above type determines the gain of the large antenna. The principal problem arising in this type of measurement is that of accurately determining the noise power received by the smaller antenna. At X-Band, for example, an 8.5-ft antenna will receive an excess noise power from Cassiopeia of 0.2 to 0.4°K. On the other hand, the Rayleigh far field distance for this antenna at 8450 mc is 620 ft, which is a rough upper range limit for straightforward accurate gain calibrations by conventional means. If, therefore, the small antenna is restricted to sizes of this order, the radio source measurement must be performed to an accuracy on the order of 0.01°K.

Very high resolution radiometers are almost always of the switching or Dicke type (Ref. 3). In order to obtain 0.01°K resolution with an ambient temperature switch, the differential loss between the switch ports must be stable to 0.00015 db during the measurement period. That this performance is practical in the field is not immediately obvious. As an alternate approach which contains no electrically or mechanically moving parts, a dual channel radiometer is being investigated.

In Fig. 6, the basic design of the dual channel radiometer is shown together with typical signal and noise powers at various points. The input signal is added to the noise power from a cryogenic load, and the total is then split between the two channels. The local oscillator

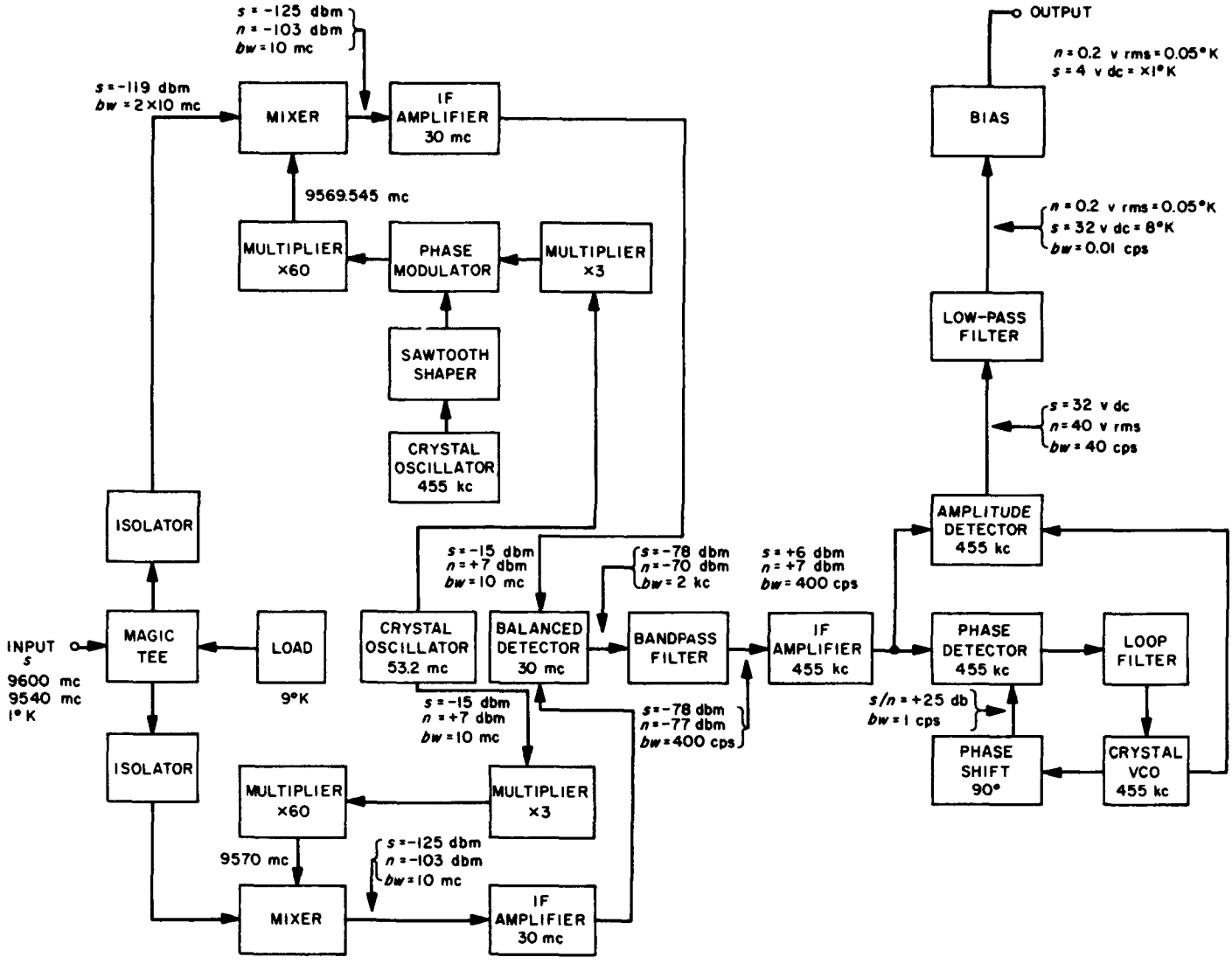


Fig. 6. X-Band dual channel radiometer

signals are separated by 455 kc which causes the two coherent signal spectra to be similarly separated in frequency. A CW 455 kc signal is thus constructed in the balanced detector and is tracked by the lock-loop circuitry. The lock-loop is utilized to eliminate the necessity for phase stability in the system. The use of 455 kc rather than DC also eliminates excessively stringent balance requirements on the 30 mc detector. Hard limiting is used in both inputs of the balanced detector to reduce the effect of gain instability in the local oscillators, mixers, and 30 mc IF amplifiers.

A simple analysis has been carried out on a CW basis and then extrapolated to the case of performance with noiselike signals. This analysis indicates comparable performance to a switching type of radiometer.

Most of the components in the block diagram exist from a previous program and are similar to those used in the TRAC(E) receiver system. The balanced detector has been modified, however, to provide both limiting and increased bandwidth. The circuit being used in this component is shown in Fig. 7. The two inputs are designated "signal" and "reference" by analogy to conventional balanced detector usage. The signal on the reference side acts to switch the diodes on and off once per cycle. Provided only that the reference signal level is much higher than the other side, the reference is thus automatically limited. The "signal" side of the detector is cutoff limited in the pentode by a combination of grid-leak bias and low screen voltage. The measured performance of the balanced detector is shown in Table 1.

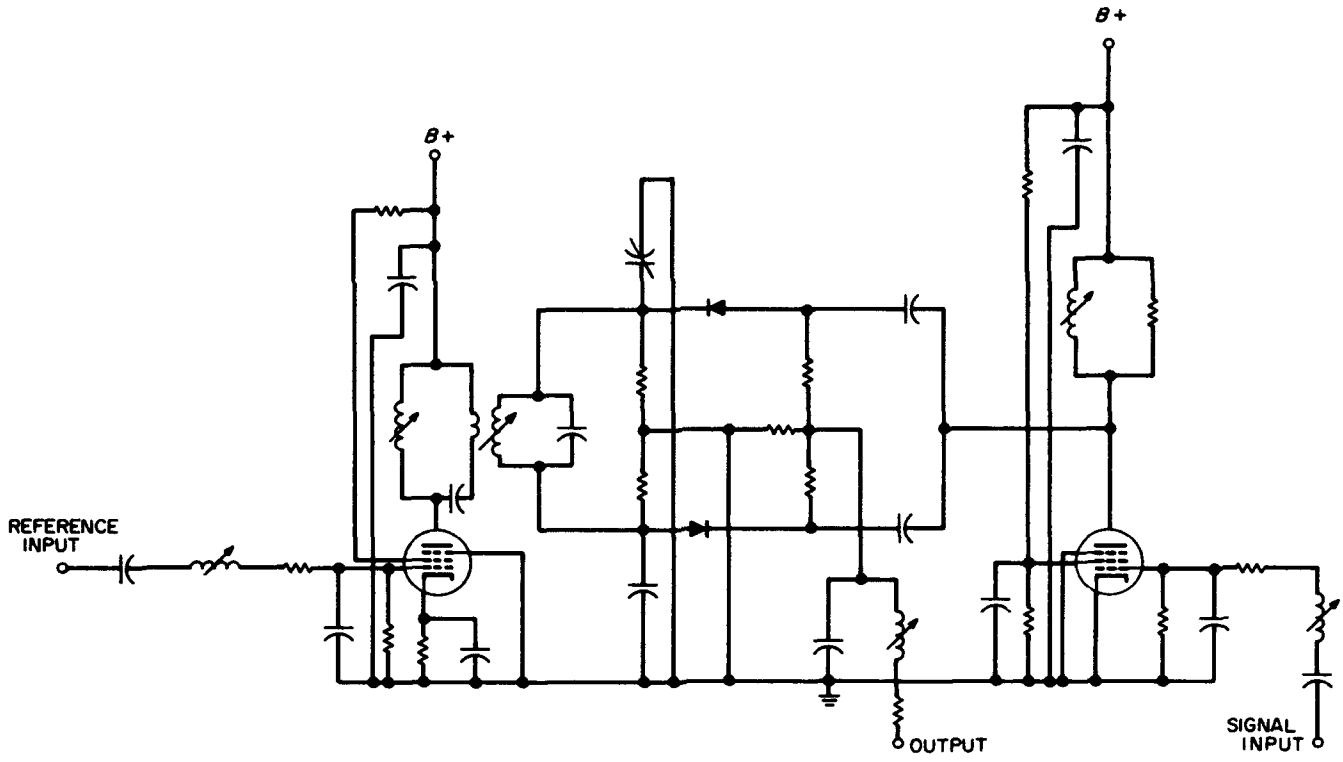


Fig. 7. Broadband 30 mc balanced detector-limiter

The 30 mc IF amplifiers are presently on order, and the 455 kc sawtooth shaper is under development. The remainder of the components are in existence and are presently being refurbished and adjusted.

b. Nitrogen load for X-band. To provide a reference standard for experimental X-band radiometer systems used in antenna feed research, the development of a liquid-nitrogen cooled termination for X-band waveguide has been initiated. A photograph of an experimental load which was tested at liquid nitrogen temperature (78°K) is shown in Fig. 8. The waveguide section was made from standard 1 × ½ in. brass waveguide approximately 22 in. in length. The terminating element (Hewlitt-Packard W 20F) is 4.4 in. in length, made from polyiron material, and is tapered along its length. The open end of the waveguide load was sealed with a waveguide window (Varian V1100-B), and the entire line was pressurized with helium gas through a valve connection. Pressure sealing the line in this manner prevented moisture condensation from forming inside the guide and also kept liquid nitrogen from leaking into the line. A compound (vacuum-pressure) gage was used to indicate any changes in pressure or vacuum in the line over extended periods of time.

Table 1. Balanced detector performance

Item	Reference side	Signal side	Output
Limiting input	+7 dbm	+4 dbm	—
Output level, 455 kc	—	—	−33 dbm
3 db bandwidth	10 mc	10 mc	—
Balance, 30 mc	76 db	52 db	—
35 mc	28 db	16 db	—
25 mc	32 db	18 db	—

Measurements of VSWR were first made at ambient temperature, and then the load was remeasured after immersing it into a 25-liter nitrogen-filled dewar for 0.5 hr. Results of measurements made between 9520 and 9620 mc are shown in Table 2. The poor VSWR's at ambient temperature are due to match of the waveguide window used. However, no changes in VSWR occurred after cooling the termination to liquid-nitrogen temperature.

Plans are presently being made to construct a line made from precision copper and stainless steel waveguides and then test it with a well-matched broadband window. A sketch of an improved design of a load to be constructed is shown in Fig. 9.

c. Noise temperature measurements using S-band radiometer. An experimental project has been proposed to

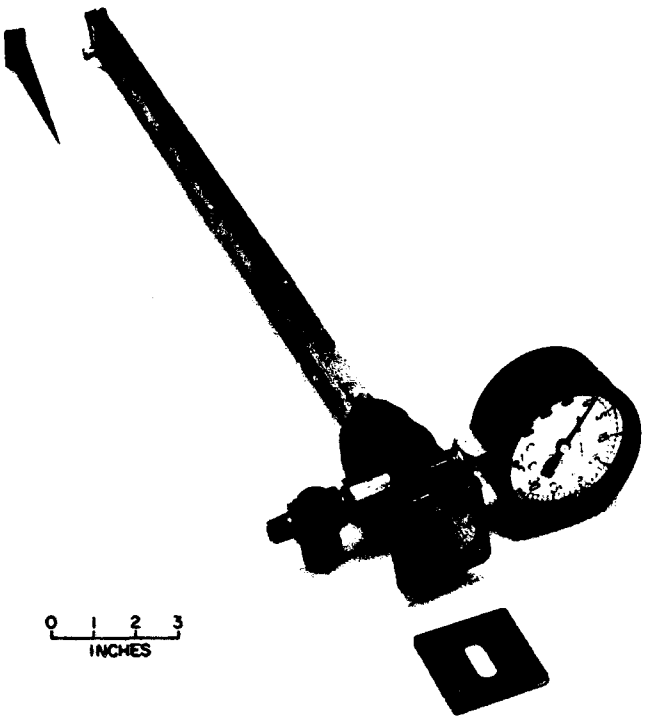


Fig. 8. Experimental waveguide nitrogen load for X-band

use a "total power" radiometer system to accurately measure thermal noise contributions from large reflector surfaces. At the present time, an improved S-band (2388 mc) radiometer system has been constructed and is being evaluated. Instrumentation and previous work on this radiometer system have been described in *RS 36-12* and *RS 36-13*.

A block diagram of the present radiometer system is shown in Fig. 10. The high-gain horn is pointed at the zenith position to measure sky background temperature. A liquid-nitrogen cooled termination, a liquid-helium cooled termination, and a gas tube noise source provide accurate reference temperatures for the calibration of temperatures received by the horn. A timer is used to remotely control the operation of two waveguide switches and a noise tube source. The switching sequence is automatically controlled by the timer, and temperatures are continuously recorded. The switching cycle repeats every 15 min. A sample recording of this switching sequence may be seen in Fig. 11.

Unexpected difficulties were encountered when the system was put into operation. The recorded temperature in the helium load channel was apparently high; but at

Table 2. VSWR measurements of experimental nitrogen load for X-band

Frequency, mc	VSWR at ambient temperature	VSWR at liquid-nitrogen temperature
9520	1.090	1.084
9540	1.085	1.084
9560	1.084	1.083
9580	1.085	1.084
9600	1.077	1.077
9620	1.080	1.081

the present time, the source of the difficulty has not been located. The system has been repeatedly rechecked for mechanical alignments, RF leakages through the flanges, and moisture in the line due to condensation and rain. Impedance matches at all the connections throughout the system were remeasured several times. Since unusually high insertion loss in the helium load channel is suspected, the insertion loss of the helium load assembly will be measured in a precision insertion loss measurement setup.

One of the future uses planned for this system, after successfully putting it into operation, is the measurement of surface loss of plane sheet reflectors and transmission loss through meshed or perforated surfaces. A setup for a possible experiment to measure loss of a solid surface plane sheet reflector is shown in Fig. 12. A high-gain horn is mounted on a pedestal which scans in the E or H plane ± 45 deg from zenith. In one position, the antenna is pointing at the reflector; and in the second position, the antenna points at the same part of the sky seen by the reflector in the first position. In the calculation of the thermal noise contributions from large reflector surfaces, let

T_{m1} = temperature measured when the antenna is pointing at the reflector

T_{m2} = temperature measured when the antenna is pointing at the sky in position 2

$T_{\text{sky}(\theta)}$ = temperature of that portion of the sky included in the angle θ

$T_{\text{background}}$ = temperature contributions from the remaining areas of the sky and from the ground

$T_{\text{reflector}}$ = temperature due to surface loss of the reflector

Assuming $T_{\text{background}}$ to be identical in both antenna positions (Fig. 12),

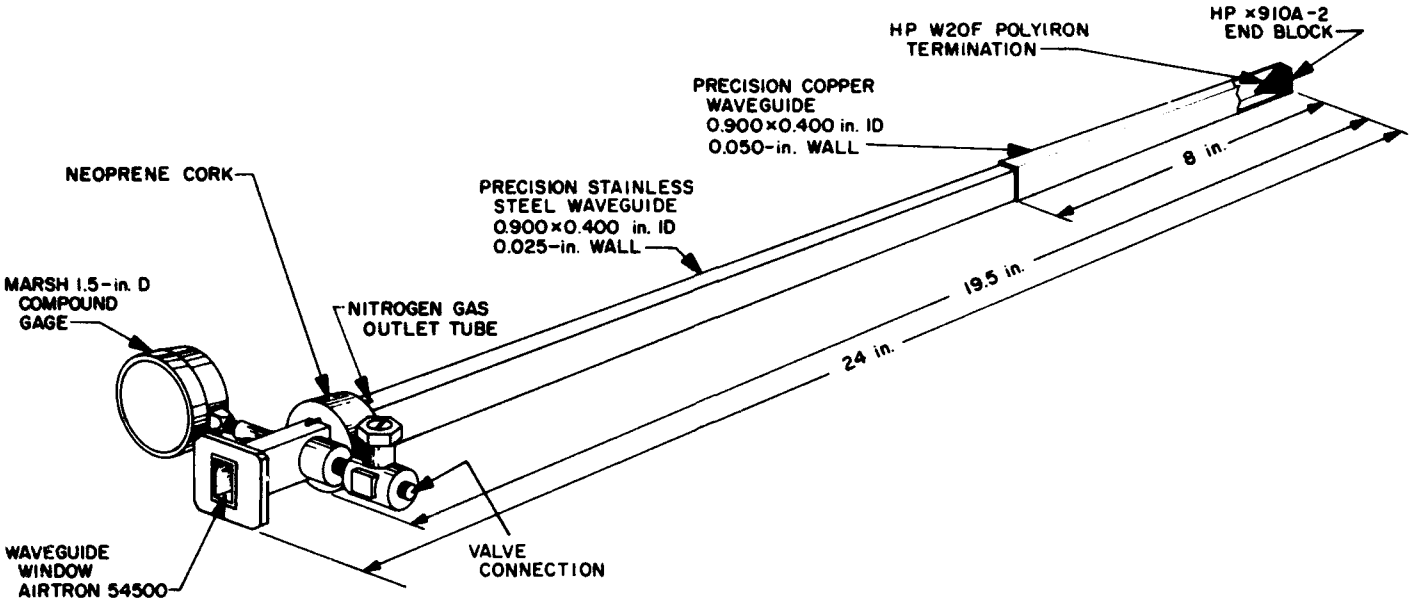


Fig. 9. Improved nitrogen load design for X-band

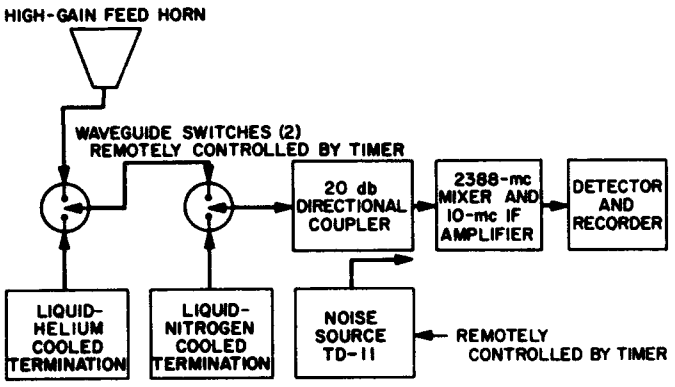


Fig. 10. Present 2388 mc "total power" radiometer system under test

$$T_{m1} = T_{\text{reflector}} + T_{\text{sky}(\theta)} + T_{\text{background}}$$

$$T_{m2} = T_{\text{sky}(\theta)} + T_{\text{background}}$$

Then

$$T_{\text{reflector}} = T_{m1} - T_{m2}$$

In practice, the reflector dimensions, antenna characteristics, and antenna-reflector spacing will be chosen so that 95% or more of the total thermal power received by the antenna will be contained in the solid angle subtended by the reflector. The greater portion of spillover thermal power will be radiating from the cooler parts of the sky (Fig. 12). An attempt will be made to select a test site where the ground terrain will contribute nearly identical amounts of thermal power in the two antenna positions. In preparation for this experiment, the design and construction of a pedestal has been completed. The photograph in Fig. 13 shows the high-gain horn in one of the 45-deg positions.

2. *Analysis of Low Noise Antennas*, W. V. T. Rusch

a. Scattering from an arbitrary surface of revolution. A ray-tracing analysis of a two-reflector antenna feed system shows that a subreflector of hyperboloidal shape is required to produce a spherical wave emerging from the secondary focus and incident on the principal reflector. However, at radio frequencies a truncated hyperboloid will neither produce a perfectly spherical reflected wave front nor eliminate spillover beyond the edge of the principal reflector. An empirical approach to this problem has led to the addition of a conical flange extending beyond the edge of the hyperboloid to improve cutoff of the principal reflector edge illumination (Ref. 4). A rigorous mathematical approach has been undertaken by calculating the field scattered from an axially symmetric subreflector of arbitrary shape. The results of this analysis are intended to yield an optimum subreflector shape for Cassegrainian feed systems.

In terms of the geometry of Fig. 14, the electric field of the wave emerging from the primary feed at O may be described by

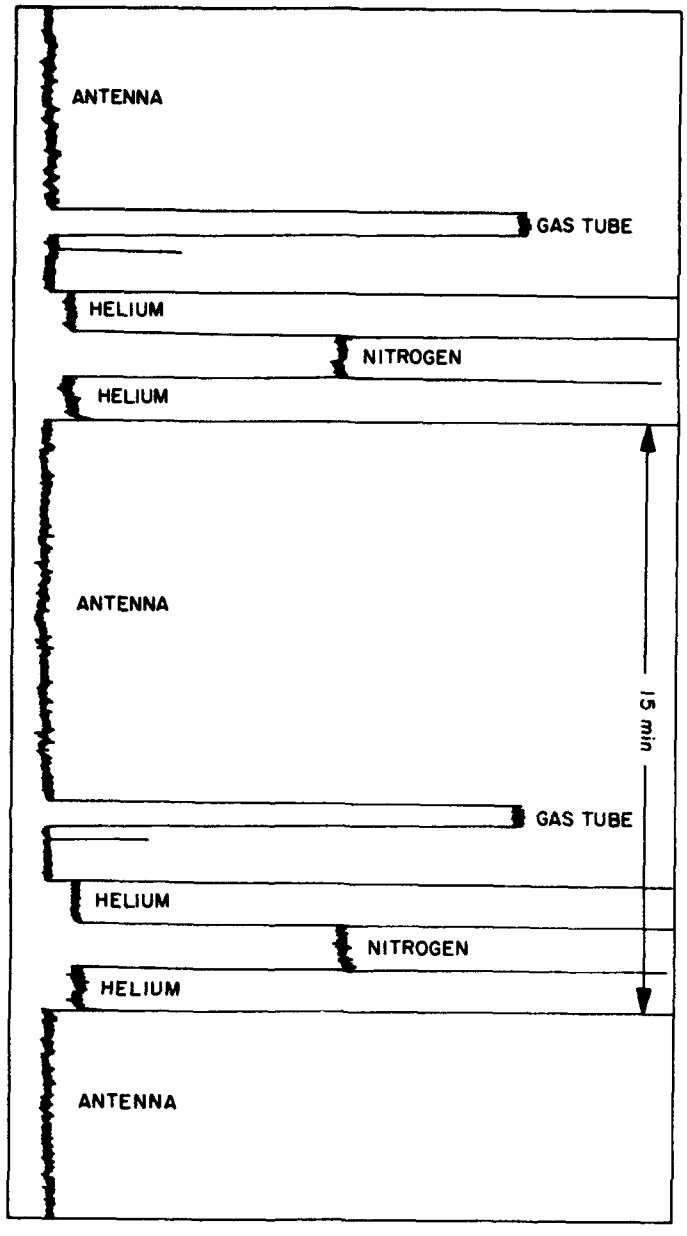


Fig. 11. Sample recording showing automatic switching sequence; recording is not calibrated in degrees due to presently unknown value of apparent temperature for the helium load line

$$\mathbf{E}_{pf}(\theta,\phi) = AP(\theta)\frac{e^{ikR}}{R}\,\mathbf{e}(\theta,\phi) \tag{1}$$

The unit vector $\mathbf{e}(\theta,\phi)$ describes the polarization of the primary field, and $P(\theta)$ is the axially symmetric pattern factor of the primary field. The axially symmetric subreflector is defined by the equation

$$k\rho g(\theta') = -1 \tag{2}$$

where for a paraboloid:

$$g(\theta') = \frac{1-\cos\theta'}{4\pi\dfrac{f}{\lambda}} \tag{3}$$

and for a hyperboloid:

$$g(\theta') = \frac{1+e\cos\theta'}{kep} \tag{4}$$

In general $g(\theta')$ is an unknown function which will be chosen to produce the desired amplitude and phase characteristics of the reflected wave. If the radius of curvature at every point on the subreflector is much larger than a wavelength, the scattered field may be calculated by integrating the surface currents over the front of the reflector. For a linearly polarized primary field, the H-plane component of the total field at the point P is

$$E_\phi\left(R,\theta,\frac{\pi}{2}\right) = \left(\frac{iA}{2}\right)\left(\frac{e^{ikR}}{R}\right)(R+iI) \tag{5}$$

where

$$R = \int_{\theta_0}^{\pi}\frac{P(\theta')\cos\alpha\sin\theta'}{[g(\theta')]^2}\{g(\theta')[J_0(\beta)-J_2(\beta)] + [g'(\theta')\sin\theta' - g(\theta')\cos\theta'][J_0(\beta)+J_2(\beta)]\}\,d\theta' \tag{6}$$

$$I = 2P(\theta) + \int_{\theta_0}^{\pi}\frac{P(\theta')\sin\alpha\sin\theta'}{[g(\theta')]^2}\{g(\theta')[J_0(\beta)-J_2(\beta)] + [g'(\theta')\sin\theta' - g(\theta')\cos\theta'][J_0(\beta)+J_2(\beta)]\}\,d\theta' \tag{7}$$

$$\alpha = \frac{\cos\theta'\cos\theta - 1}{g(\theta')} \tag{8}$$

$$\beta = -\frac{\sin\theta\sin\theta'}{g(\theta')} \tag{9}$$

The phase of the field with respect to a phase center at O' is

$$\phi = \tan^{-1}\left(\frac{I}{R}\right) - (2kc)\cos\theta \tag{10}$$

The amplitude and phase can then be determined by using Eqs. (6), (7), and (10).

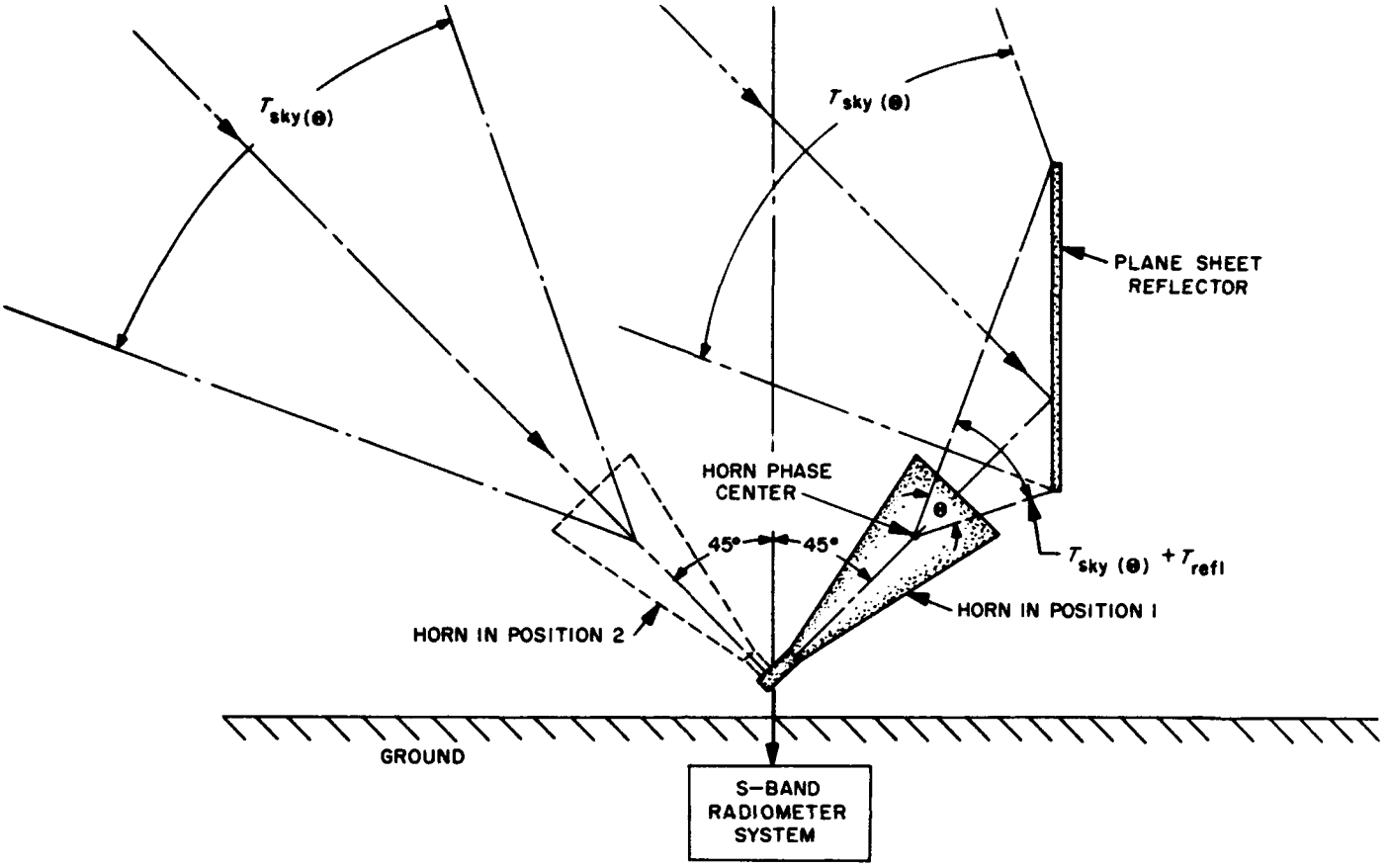


Fig. 12. Proposed experimental setup using a radiometer system to measure surface loss of plane sheet reflectors

Fig. 13. High-gain horn and pedestal for noise temperature experiments at S-band

C. Thin Film Techniques

R. E. Brantner, R. E. Frazer, J. Maserjian, and H. Erpenbach

1. Superconducting Components, *R. E. Brantner*

Superconductive specimens must be tested after fabrication to determine whether or not they actually become superconductors when cooled below the critical temperature of the material. Also, in the case of thin films, it is necessary to find whether or not pinholes or serious variations in film thickness are present. Finally, it is desirable to determine approximately the degree of purity of the specimens. To check these characteristics, three instruments are used: (1) a curve tracer, (2) a circulating current tester, and (3) a sensitive resistance bridge. The curve tracer being used is a Tektronix Model 575 transistor-curve tracer. This unit is capable of supplying up to 20 amp of current through the specimen. It is connected as for tests of transistor collector cutoff current and produces a display similar to Fig. 15, when the specimen

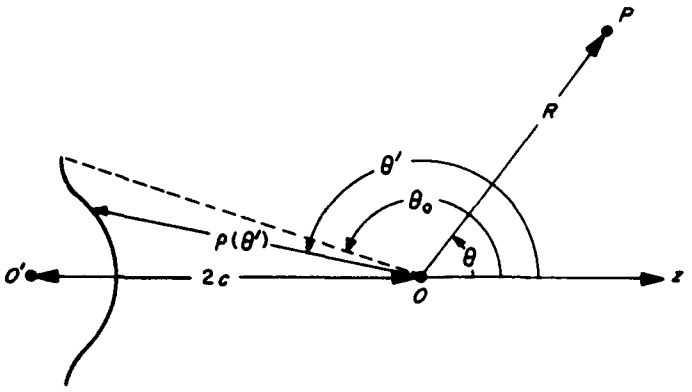


Fig. 14. Subreflector geometry

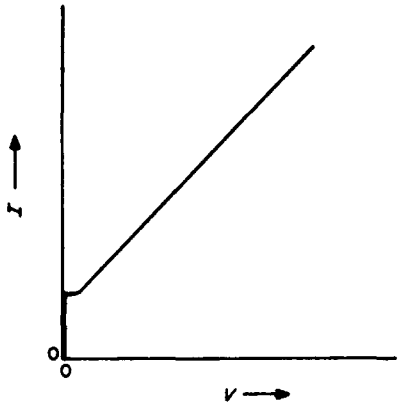


Fig. 15. Curve tracer display of superconductor

under test can be driven into the normal state by less than 20 amp of current. If the specimen is not superconductive, the display is that of an ordinary resistor.

The circulating current test system is shown in Fig. 16. The core of the sensing head was chosen to saturate at a fairly low level of DC current. A superconductive ring is placed in the gap of the sensing head (Fig. 16), and the external field is initially sufficient to drive the specimen into the normal state. The field is then decreased to zero; if the specimen becomes superconductive, part of the external field is trapped within the ring, and a circulating current is set up within the superconductive material. The trapped flux is confined, mostly within the core of the sensing head, and causes a distortion in the output signal when a sinusoidal input signal is applied. A DC current can be applied to a third winding, having a known number of turns. By adjusting and measuring the polarity and level of the DC current to minimize the distortion of the AC signal, the amplitude of the circulating current can be determined. With the specimen geometry known, its approximate critical field can be

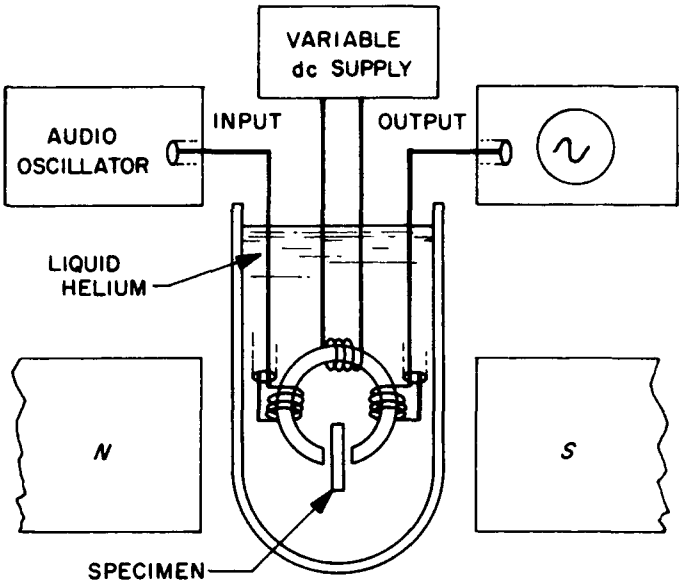


Fig. 16. Circulating current test system

calculated. When continuous films are to be tested, the same procedure is used. If a circulating current can be induced, either pinholes or serious variations in film thickness are present.

In characterizing superconductive specimens, it has been determined that the ratio of the normal resistance at room temperature to that at 4.2°K is dependent upon the amount of impurities present in the specimen. The standard nomenclature uses the relationship

$$\Gamma = \frac{R_{300°K} - R_{4.2°K}}{R_{4.2°K}}$$

to indicate this characteristic of the specimens. For hyperpure specimens, Γ may be in excess of 10^5, while for very impure specimens it may be as low as 3 or 4. To test the Γ ratio of the specimens, a sensitive resistance bridge is used. An external magnetic field is applied during the test to maintain the specimen in the normal state. Construction of permanent models of the last two test systems is now being completed.

2. Pyrolytic Thin Films, R. E. Frazer

A pyrolytic decomposition apparatus has been in operation for several months. This original apparatus was operated on a feasibility evaluation basis using the simplest gas handling manifolds. Heat was supplied by graphite and tantalum receptors in the reaction chamber. A surrounding RF field raised the receptors to the reaction temperature estimated to be 1000°C.

At these temperatures, films of several materials have been deposited. Tantalum and niobium halides have been decomposed to yield clean bright deposits of these metals on sapphire and quartz substrates. Silicon halides have been reduced to form thin amorphous films on quartz incapsulated receptors.

An improved apparatus is under construction. It will provide close control of source and reaction temperatures as well as flow rates and pressures of reacting gases.

3. *Single Crystal Films*, J. Maserjian

The investigation of microzone-growing of single crystal germanium films has continued (*RS 36-4* to *36-13* and Ref. 5). Recent efforts have been directed at controlling the area over which a continuous single crystal film is grown. This has been attempted with an electron beam scanning technique which limits the direction of growth so that the problem of seeding of new crystallites from the edge of the region is avoided (*RS 36-12*). A system designed for this purpose has been tried and has succeeded in demonstrating the feasibility of this approach, but its usefulness thus far has been limited by technical difficulties in the system.

The scanning pattern employed is shown schematically in Fig. 17, where for clarity only a fraction of the number of oscillations are indicated. The generator designed for this purpose involves a series of motor-driven pots which are geared to oscillate at 1 and 81 rpm. The outputs to the beam deflection coils are switched between the two frequencies at the proper phase to give this pattern. A 4 kc circular Lissajous figure is made to follow this pattern (*RS 36-12*). The specimen is also moved laterally during the scanning process.

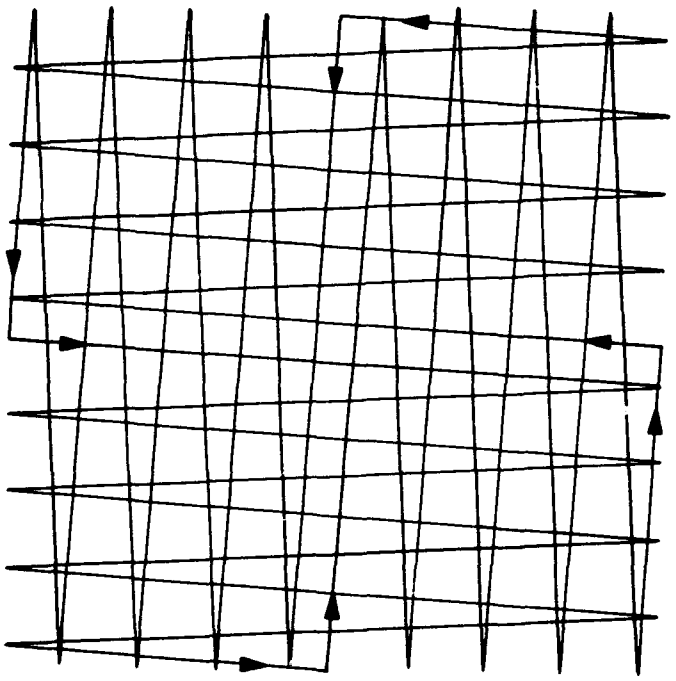


Fig. 17. Microzone-melting scanning pattern

Difficulty has been encountered in maintaining exact phase relationship during switching. This results in serious discontinuities in the pattern. These discontinuities in turn disrupt the growth process and prevent extended single crystalline growth. However, the pattern was maintained sufficiently long to demonstrate feasibility of this scanning technique.

A photomicrograph, shown in Fig. 18, was taken (under dark field, 30 × 43 mils) of a 4-μ thick germanium film that had been subjected to this technique. The circular forms in the figure are alloy spheres discussed below. In the dark field, the single orientations of crystal growth is revealed to extend over most of the width of the scanned region. The outer strips were expected to be polycrystalline because of the seeding from the edges (*RS 36-12*). These form when the zone makes its outermost trace, and growth occurs equally from the inner and outer edges of the zone. It is the intention to maintain continuous single crystal growth over the central part of the scanned region and along the traversed direction. The disruptions evidenced in the figure are believed directly related to the discontinuities which arose in the scan pattern as mentioned.

An effort was made to produce *p-n* junctions on these crystal films using a simple alloying technique. Gold-germanium eutectic alloy spheres doped with antimony or gallium were alloyed to the films, as can be seen in Fig. 18. High-power microscope observation revealed that a *p-n* junction did not form, however, as apparently all the germanium film was dissolved by the sphere and no regrowth occurred upon cooling. This technique which is so simple to apply to massive single crystals appears much too critical for these films. An interesting feature of this experiment is the absence of any cracking of the film in spite of rather severe thermal and mechanical stresses created by the alloying process. Although it still may be possible to obtain alloy junctions by proper control of alloy and temperature gradients, it does not seem worthwhile to spend more effort in this direction. This experiment was intended as a demonstrative one, and the device applicability is quite limited.

In work on the field effect transistor (*RS 36-13*), it will be necessary to apply planar techniques involving diffu-

Fig. 18. Microzone-melted germanium film

sion of impurities through masked regions in order to obtain the proper junction geometry. These techniques must still be developed in the laboratory before this application can be made.

4. Superconducting Thin Films, *H. Erpenbach*

Work has continued on developing the techniques for depositing superconducting thin films.

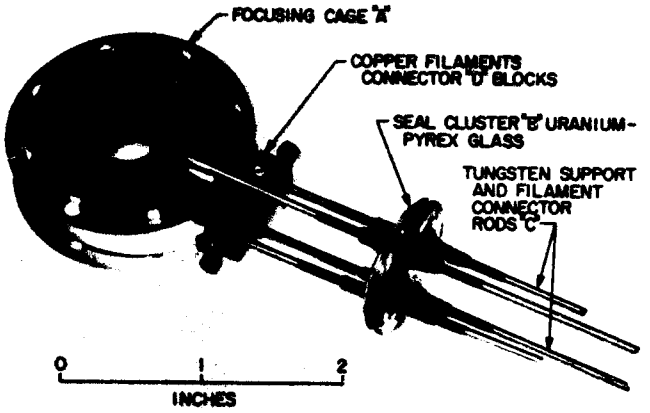


Fig. 19. Bakeable electron source

A bakeable electron source for the evaporation of refractory metals at ultrahigh vacuums is shown in Fig. 19. This electron source has been fabricated from materials with relatively low vapor pressures. The electron focusing cage "A" is molybdenum, the seal cluster "B" is uranium and pyrex glass, the support and filament connector rods are tungsten, and the filament connector blocks "D" are copper. This assembly is demountable and may be snapped apart for filament replacement.

After a 2-hr bakeout at a temperature of 500°C, tests made with a potential of 3000 v on the filament cathode and a negative bias on the cage "A" of 20 v, show that 600 w may be focused on the anode evaporant with negligible out-gassing. A schematic of how the assembly of Fig. 19 is used in a vacuum chamber is presented in Fig. 20.

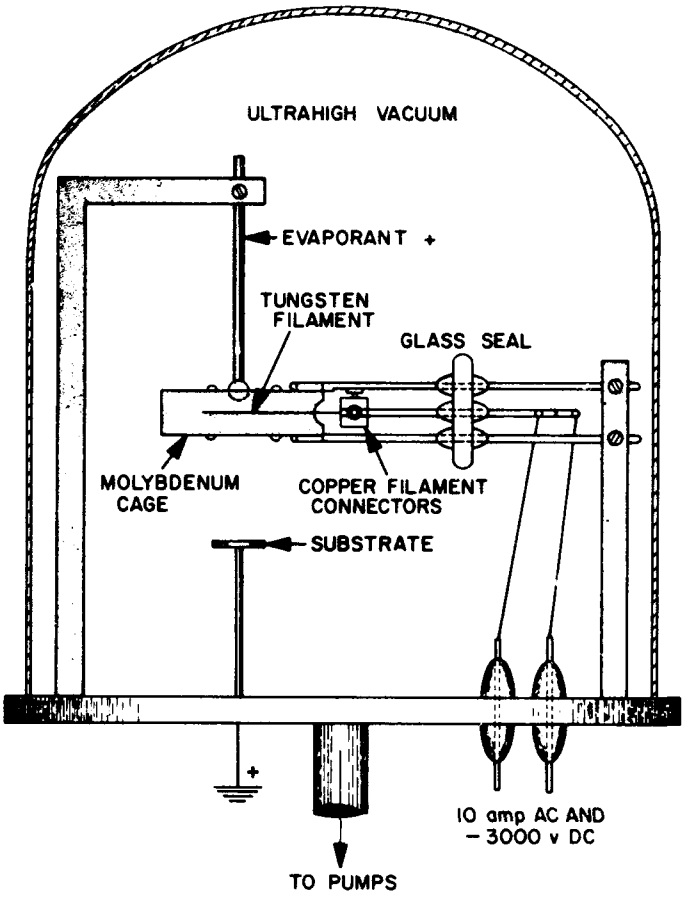


Fig. 20. Bakeable electron source for ultrahigh vacuum use

References

1. Strum, P. D., "Considerations in High Sensitivity Microwave Radiometry," *Proceedings of the IRE*, Vol 46, pp 44–53, January 1958.

2. Strandberg, M. W. P., *Microwave Spectroscopy*, John Wiley and Sons, New York, 1954.

3. Strum, Peter D., "Considerations in High-Sensitivity Microwave Radiometry," *Proceedings of the IRE*, Vol 46, No. 1, January 1958.

4. Potter, P. D., *A Simple Beamshaping Device for Cassegrainian Antennas*, Technical Report 32-214, Jet Propulsion Laboratory, Pasadena, Calif., January 31, 1962.

5. Maserjian, J., *A Semiconductor Approach to Microelectronics*, Technical Release 34-206, Jet Propulsion Laboratory, Pasadena, Calif., December 2, 1960.

PHYSICAL SCIENCES DIVISION

VIII. Chemistry Research

A. Thermodynamic Properties of Nitryl Fluoride

E. Tschuikow-Roux

In connection with our current interest in the kinetics of several reactions of nitryl fluoride, FNO_2, it was believed necessary to consider the thermochemistry of these reactions. Such thermochemical calculations were not possible until recently, since neither the heat of formation of nitryl fluoride nor its thermodynamic functions were available.

However, the infrared and Raman spectrum of nitryl fluoride has been studied by Dodd, Rolfe and Woodward (Ref. 1) and the microwave spectrum has been reported by Smith and Magnuson (Ref. 2). Using these spectroscopic data, the thermodynamic functions of nitryl fluoride were evaluated in the temperature range from 200 to 1200°K at 100°K intervals. Very recently, an approximate value for the standard heat of formation of FNO_2 ($\Delta H_{f,298} = -20 \pm 5$ kcal/mole) was reported by Anderson, MacLaren and co-workers (Ref. 3). This value was subsequently revised to -19 ± 2 kcal/mole (Ref. 4). Using this value for the standard heat of formation of nitryl fluoride and the thermodynamic data from the literature for the species FNO, NO_2, NO, O, F, F_2, CO, CO_2 it was possible to calculate the standard enthalpies, free energies, entropies, and equilibrium constants for the reactions of interest. The bond dissociation energies for the N-F and N-O bonds in FNO_2 have been calculated and are compared with the corresponding bond dissociation energies of the N-Cl and N-O bonds in $ClNO_2$.

We shall be interested in the following reactions:

$$NO_2 + F_2 = FNO_2 + F \quad (1)$$

$$FNO_2 + F = NO_2 + F_2 \quad (2)$$

$$FNO_2 = NO_2 + \tfrac{1}{2}F_2 \quad (3)$$

$$FNO_2 = NO_2 + F \quad (4)$$

$$FNO_2 = FNO + O \quad (5)$$

$$FNO_2 + NO = FNO + NO_2 \quad (6)$$

$$FNO_2 + CO = FNO + CO_2 \quad (7)$$

$$NO + F_2 = FNO + F \quad (8)$$

1. Calculations

The thermodynamic functions of nitryl fluoride were calculated according to the rigid-rotator harmonic-oscillator approximation. Nitryl fluoride was assumed to be a perfect gas and the pressure was taken as 1 atm. The molecular constants, determined from spectroscopic measurements, are summarized in Table 1. The fundamental frequencies were taken from the reported infrared spectrum (Ref. 1). The moments of inertia and the corresponding molecular dimensions are based on the microwave spectrum (Ref. 2). The symmetry number σ is 2. Nitryl fluoride is a planar molecule with a C_2 axis of

Table 1. Molecular constants for FNO_2

Fundamental frequencies,[a] cm^{-1}	Moments of inertia,[b] $g\text{-}cm^2$	Molecular dimensions[b]
ν_1 1312	$I_a = 63.562 \times 10^{-40}$	r (N—F) 1.35 Å
ν_2 822	$I_b = 73.313 \times 10^{-40}$	r (N—O) 1.23 Å
ν_3 460	$I_c = 137.116 \times 10^{-40}$	
ν_4 1793		∢ ONO 125°K
ν_5 570	$I_aI_bI_c = 6.3895 \times 10^{-113}\ g^3cm^6$	∢ ONF 117.5°K
ν_6 742		

[a]Ref. 1
[b]Ref. 2

symmetry (Refs. 1, 2). The ground-state electronic degeneracy corresponding to this point group is unity, and thus the electronic contribution to the thermodynamic functions has been neglected. The results of these calculations are presented in conventional form in Table 2.

In calculating the standard enthalpy, free energy, and entropy changes, as well as the equilibrium constants for Reactions (2) through (8), we have used the value of −19 kcal/mole for the standard heat of formation of nitryl fluoride at 298°K and the thermodynamic functions from Table 2. The results of these calculations are presented in Tables 3 and 4. Table 3 gives the standard enthalpy change at 0°K and the enthalpy, free energy, and entropy changes of Reactions (2) to (8) in their standard states at 298°K. Table 4 lists the logarithm of the equilibrium constant, $\log_{10} K_p$, for said reactions as a function of temperature from 200 to 1200°K.

Table 2. Thermodynamic functions of FNO_2

T, °K	Function, cal $mole^{-1}$ deg^{-1}			
	C_p^0	$(H^0 - H_0^0)/T$	$-(F^0 - H_0^0)/T$	S^0
200	9.857	8.421	49.427	57.848
273.16	11.428	9.019	52.137	61.156
298.16	11.919	9.241	52.937	62.178
300	11.954	9.258	52.993	62.251
400	13.626	10.150	55.780	65.930
500	14.903	10.979	58.135	69.114
600	15.877	11.718	60.204	71.922
700	16.621	12.368	62.068	74.436
800	17.193	12.936	63.749	76.685
900	17.638	13.435	65.302	78.737
1000	17.988	13.874	66.741	80.615
1100	18.265	14.261	68.081	82.342
1200	18.487	14.604	69.338	83.942

Table 3. Calculated thermodynamic data

Reaction	ΔH_0^0, kcal $mole^{-1}$	ΔH_{298}^0, kcal $mole^{-1}$	ΔF_{298}^0, kcal $mole^{-1}$	ΔS_{298}^0, cal deg^{-1} $mole^{-1}$
2	8.0	8.2	6.5	5.67
3	26.4	27.1	21.3	19.36
4	44.7	46.0	36.1	33.06
5	61.3	62.8	52.2	35.53
6	−10.4	−10.3	−11.5	4.04
7	−64.4	−64.4	−64.7	0.93
8	−18.4	−18.5	−18.0	−1.63

The standard heat of formation at 0°K and the free energy and equilibrium constant of formation at 298°K of FNO_2, all based on its reported (Ref 4) $\Delta H_{f,298}^0$, are presented in Table 5.

The enthalpy changes in Reactions (4) and (5) represent the heats of dissociation of the N-F and N-O bonds in FNO_2. It was thought instructive to compare these values with the corresponding bond energies in $ClNO_2$. Cordes and Johnston (Ref. 5) have reported the heat of the reaction

$$ClNO_2 = NO_2 + Cl \qquad H_0^0 = 29.5 \text{ kcal/mole}$$

Using this value together with the thermodynamic functions for nitryl chloride and nitrosyl chloride and the known heat of formation of nitrosyl chloride, the heats of the reaction

$$ClNO_2 = ClNO + O$$

have been calculated at 0 and 298°K. These data are also tabulated in Table 5.

2. Discussion

In making the thermodynamic calculations, our aim was to make a prediction of the feasibility of investigating the kinetics of some of the reactions given above. In particular, we were interested in the thermal decomposition of nitryl fluoride and its reaction with nitric oxide. As may be seen from an inspection of Table 4, the equilibrium of Reactions (3), (4), and (5) lies on the side of the reactants over the entire temperature range. The thermal decomposition of FNO_2 via a F-N bond rupture is favored over an O-N bond rupture; even so, we must conclude that the homogeneous thermal decomposition such as found in the case of nitryl chloride cannot be studied

Table 4. Equilibrium constants

T, °K	$\log_{10} K_p$ for Reaction No.						
	2	3	4	5	6	7	8
200	−7.72	−25.34	−42.96	—	—	—	—
298.16	−4.77	−15.63	−26.48	−38.23	8.44	47.55	13.21
300	−4.73	−15.51	−26.28	−37.95	8.39	47.12	13.13
400	−3.25	−10.57	−17.88	−26.49	6.51	35.38	9.77
500	−2.38	−7.61	−12.83	−19.60	5.38	28.34	7.76
600	−1.81	−5.63	−9.46	−14.99	4.62	23.65	6.43
700	−1.41	−4.23	−7.05	−11.70	4.07	20.29	5.48
800	−1.13	−3.19	−5.24	−9.22	3.66	17.77	4.79
900	−0.89	−2.36	−3.83	−7.30	3.34	15.81	4.23
1000	−0.72	−1.71	−2.71	−5.75	3.08	14.25	3.80
1100	−0.58	−1.18	−1.79	−4.50	2.87	12.97	3.44
1200	−0.46	−0.74	−1.02	−3.45	2.68	11.90	3.15

Table 5. Standard enthalpy, free energy and equilibrium constant of formation of FNO_2 and $ClNO_2$ and heats of dissociation of the N-X and N-O bonds at 0°K and 298°K

Function[a]	FNO_2	$ClNO_2$
$\Delta H^\circ_{f_0}$, kcal/mole	−17.6	7.7
ΔH°_{f}, kcal/mole	−19 (Ref. 4)	6.5
ΔF°_{f}, kcal/mole	−8.9	16.5
$\log_{10} K_{eq}$	6.51	−12.11
$\Delta H^\circ_{0(N-X)}$, kcal/mole	44.7	29.5 (Ref. 5)
$\Delta H^\circ_{(N-X)}$, kcal/mole	46.0	30.5
$\Delta H^\circ_{0(N-O)}$, kcal/mole	61.3	64.2
$\Delta H^\circ_{(N-O)}$, kcal/mole	62.8	65.6

[a]Except where otherwise indicated, all function values were calculated for this work.

in the case of FNO_2 at temperatures appreciably below 1200°K.

The equilibrium of the reaction between FNO_2 and CO (Eq. 7) lies overwhelmingly on the side of the reaction products. Here, the equilibrium shifts toward the reactants with increasing temperature; nevertheless, even at 1200°K the reaction is essentially 100% complete. One would expect the rate of this reaction to be extremely fast and probably of an explosive character.

The equilibrium of the reaction between nitryl fluoride and nitric oxide (Eq. 6) lies comfortably on the side of the reaction products. Here we note that the equilibrium shifts from right to left (toward the reactants) with increasing temperature. This reaction appears to be a promising one for kinetic rate studies, and we are presently engaged in this investigation.

Reaction (2) is of significance in the mechanism of the thermal decomposition of FNO_2. The evaluation of the equilibrium constant of this reaction provides a means of calculating its rate constant, since the rate constant for the reverse reaction (Eq. 1) has been determined experimentally (Ref. 6). Thus, if we write

$$k = A \exp(-E/RT)$$

then

$$K_p = k_2/k_1 = Z \exp(-E_2 + E_1)/RT$$

from which it follows that

$$k_2 = A_1 Z \exp(-E_2/RT)$$

and

$$E_2 = E_1 + RT(\ln Z - \ln K_p)$$

where Z is the ratio of partition functions, i.e.,

$$Z = Q_{NO_2} \cdot Q_{F_2}/Q_{NO_2F} \cdot Q_F$$

At 300°K, assuming ideal gas behavior and pressures of 1 atm for all species participating in the reaction, Z is found to be 11.612. Using this value, together with the corresponding K_p at 300°K ($\log K_p = -4.73$) and the reported value for k_1 [$k_1 = 1.59 \times 10^{12} \exp(-10470/RT)$ cc mole^{-1} sec^{-1}], we obtain the rate constant for Reaction (2):

$$k_2 = 1.85 \times 10^{13} \exp(-18430/RT) \text{ cc mole}^{-1} \text{ sec}^{-1}$$

B. Separation of Water from Hydrazine and Methylhydrazines by Gas Chromatography

D. M. Kuwada

In connection with the development of a mass spectrometric analytical method for the determination of hydrazine [NH_2-NH_2]; monomethylhydrazine [CH_3NH-NH_2]; 1,1-dimethylhydrazine [$(CH_3)_2N$-NH_2]; and their mixtures containing trace amounts of water, an independent method for the determination of water was required.

The Karl-Fischer method of water determination is not applicable because of side reactions. Several methods for separating water from hydrazine and methylhydrazines have been tried (Refs. 7, 8, 9) but have not been found to be satisfactory. We have found a gas chromatographic column suitable for the determination of water in hydrazine and water in mixtures of the methylhydrazines. However, we have not been able to separate water completely in a mixture of water, hydrazine, and the methylhydrazines with a single column.

The separation of water and hydrazine on a column of Fluoropak 80 loaded with 20% UCON Oil 550X is shown in Fig. 1. Allyl alcohol was used as an internal standard. The separation of water, CH_3NH-NH_2, and $(CH_3)_2N$-NH_2 on a column of Chromosorb P loaded with 15% Carbowax 300 and 15% tetra-hydroxyethylethylenediamine is shown in Fig. 2. Butyl alcohol was used as an internal standard. The internal standards are used to increase the precision of the analyses.

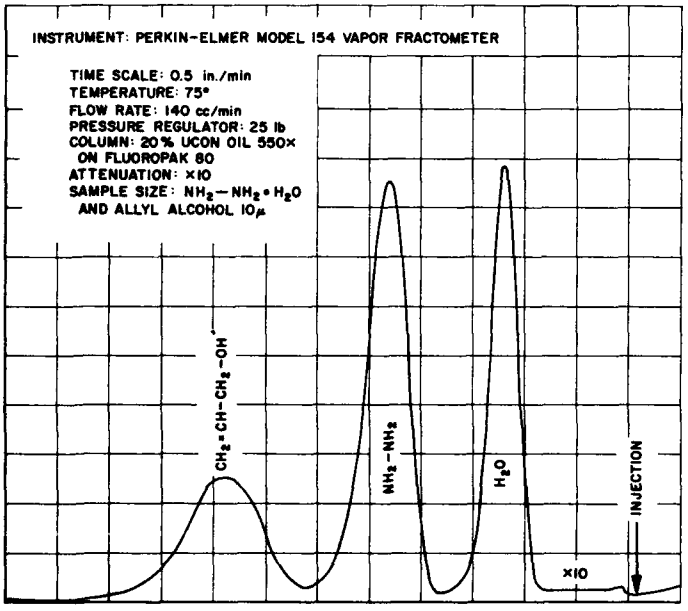


Fig. 1. Gas chromatographic separation of hydrazine, water, and allyl alcohol

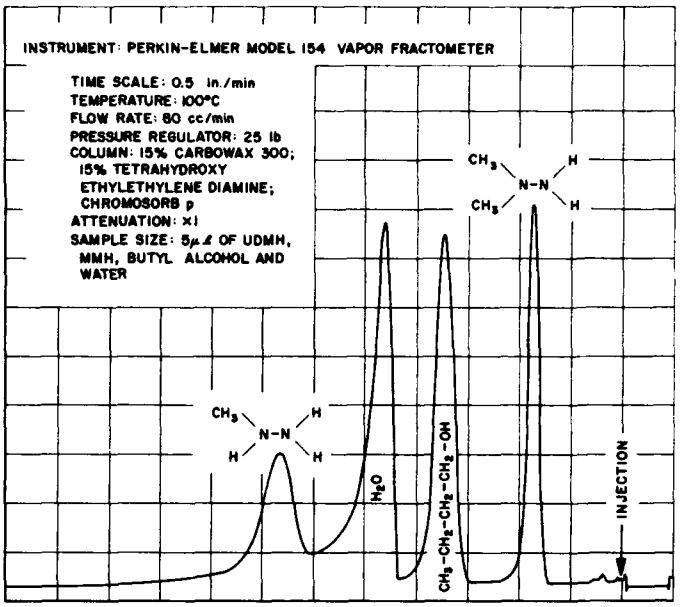


Fig. 2. Gas chromatographic separation of 1,1-dimethylhydrazine, monomethylhydrazine, water and butyl alcohol

Good agreement was obtained between the gas chromatographic results and the mass spectrometric results. The gas chromatographic technique is much more convenient than the mass spectrometric technique. An effort will be made to find a single column which will separate water, hydrazine and the methylhydrazines.

C. An NMR Study of Indene Using a Proton-Proton Decoupling Technique

D. D. Elleman and S. L. Manatt

The study of the high-resolution NMR spectrum of indene (Fig. 3a) was undertaken in the belief that knowledge of the coupling constants in this molecule would aid in the complete analysis of the NMR spectrum of cyclopentadiene (Fig. 3b) which is under study at this laboratory. The spectrum of cyclopentadiene is a complicated $A_2B_2X_2$ type (Ref. 10). We have simplified this spectrum to an A_2B_2 system by decoupling the methylene protons from the vinyl protons (*RS 36-8*) using the audio

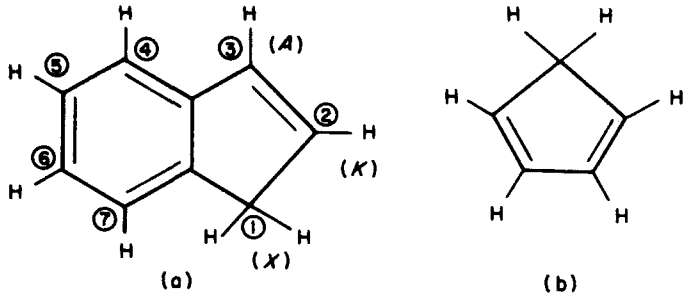


Fig. 3. Structure of indene (a) and cyclopentadiene (b)

side-band phase-detection spin-decoupling method (Refs. 11, 12). As the analysis of the A_2B_2 system does not appear to be trivial, we felt that information bearing on this problem could be gained by study of the NMR spectrum of the protons in the five-membered ring of indene (Fig. 3a). The latter protons have a rather simple AKX_2 type of spectrum. The magnitude of the three spin-coupling constants can be obtained directly from the measured spectra. However, the relative signs of the coupling constants can not be ascertained by means of a high-resolution analysis because the protons are rather weakly coupled. Evans (Ref. 13) and Freeman (Ref. 14) have shown that double irradiation experiments can be used to determine the relative signs of spin-spin couplings in many molecules. We have used an audio sideband phase detector (Refs. 11, 12) to perform double irradiation experiments on indene. From these experiments the relative signs of the coupling constants J_{AK}, J_{AX} and J_{KX} have been determined, simplification of the high-resolution spectrum was achieved, and information about certain long-range couplings of the protons in the aromatic ring with those of the five-membered ring was obtained.

1. Analysis of the Spectra

The complete 60 mc/sec spectrum of indene is shown in Fig. 4. A first-order perturbation treatment is adequate since in all cases the chemical shifts are considerably greater than the corresponding coupling constants. Thus line assignments are quite readily made. Resonance lines 1 through 6 of Fig. 4 are A transitions, lines 7 through 12 are K transitions, and lines 13 through 16 are X transitions. It is assumed resonance lines 14 and 15 are superposed and not resolved from one another because J_{AX} and J_{KX} are very nearly equal in magnitude.

A careful study of several traces of lines 1 through 6, the A resonance lines, reveals that these lines are very narrow doublets. This 0.52 $\pm$0.02 cps splitting of the A proton resonance lines must be due to one of the protons of the aromatic ring. The spectra of styrene and methyl substituted styrenes have been previously reported by other workers (Refs. 15, 16, 17, 18). The lowest-field vinyl proton signals in every case were found to be those due to the proton closest to the aromatic ring, i.e., the α-proton. Thus the A proton has been assigned as the 3-proton and the K proton as the 2-proton of indene on the basis of chemical shift alone. This assignment is then consistent with the observed long-range coupling of the A proton, and as will be shown below, with the rela-

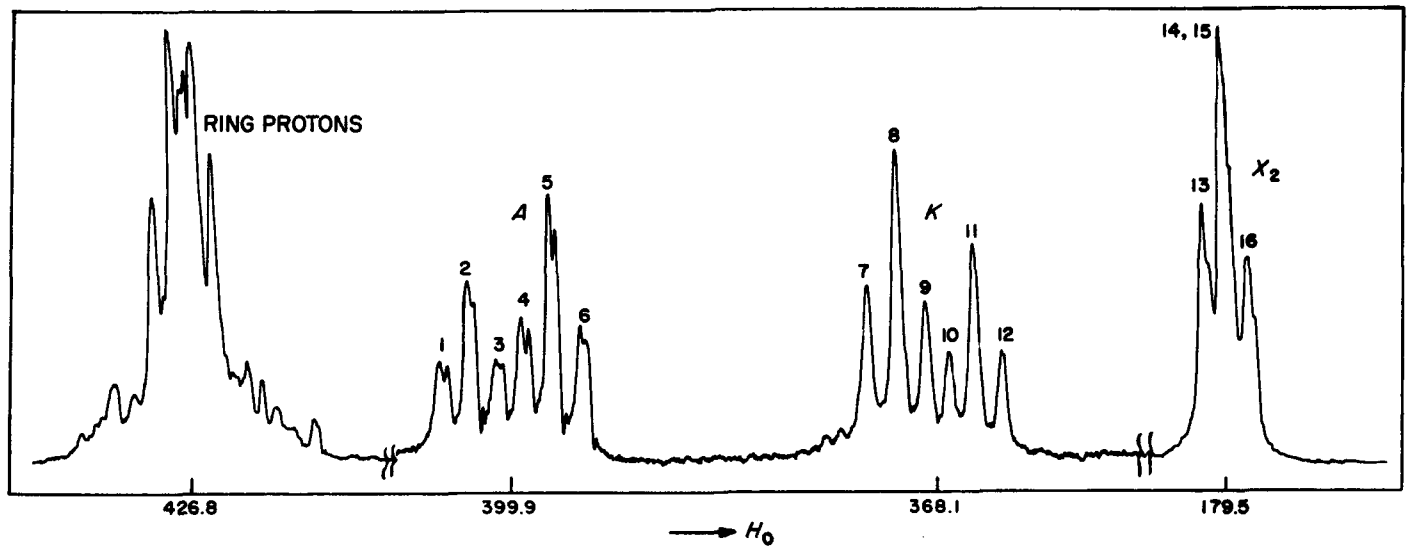


Fig. 4. The 60 mc NMR spectrum of indene; the *A and K* regions were run at twice the gain of the aromatic ring protons and the X_2 region

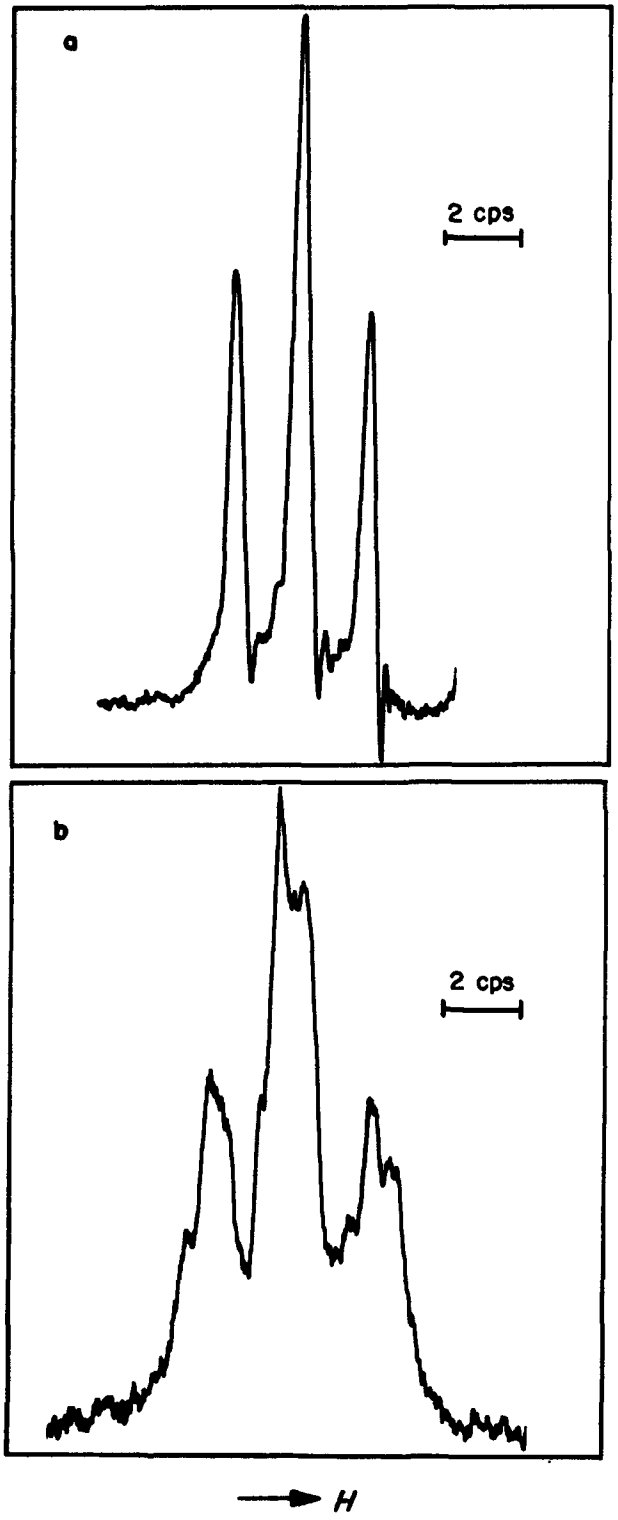


Fig. 5. The X_2 region of the spectrum: (a) double irradiation of the aromatic ring protons, $\lambda H_1/2\pi$ about 21 cps; (b) normal high-resolution spectrum of the X_2 region at a slightly slower sweep rate showing the structure due to long-range coupling with the aromatic protons

tive signs of the various coupling constants determined from double irradiation experiments. The proton at the 5-position of the aromatic ring would seem a logical choice for the origin of this long-range coupling (a coupling over four bonds).

It was noted that all of the other spectral lines show a rather large width (approximately 0.8 cps half-width) which is probably due to various long-range couplings with the protons of the aromatic ring. Double irradiation of the aromatic proton region while the *X* region is observed produces a much sharper triplet in the *X* region, as shown in Fig. 5. This verifies the hypothesis that long-range couplings are responsible for the large width of these lines. Under good resolution each of the components of the X_2 triplet appears itself to be a triplet, which suggests that two additional protons besides *A* and *K* are coupling with X_2 to the extent of about 0.5 cps (assuming that both long-range couplings have the same sign).

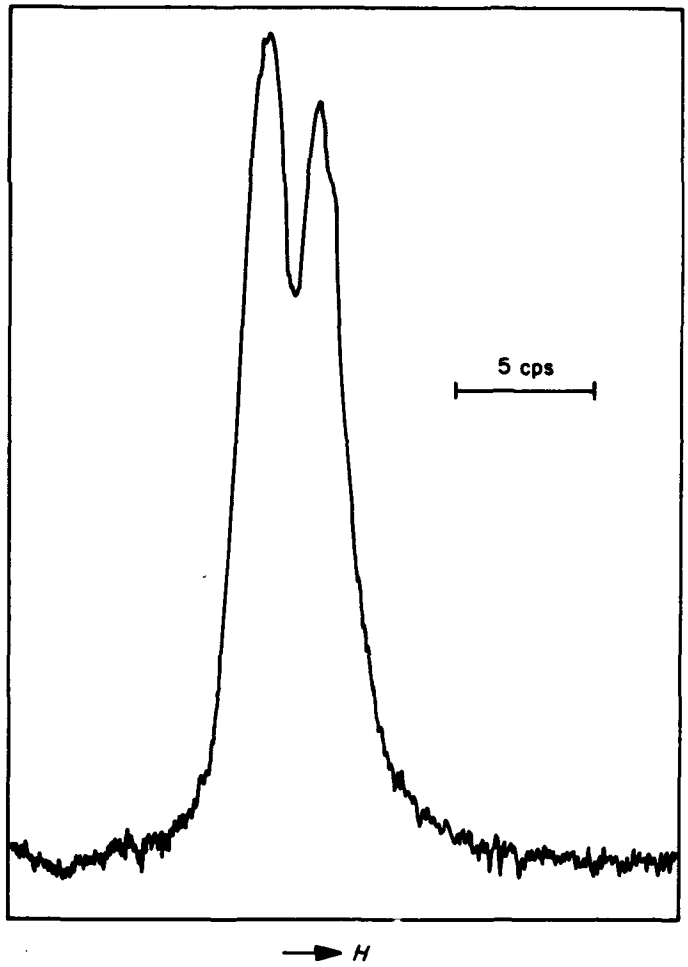


Fig. 6. The X_2 region of the spectrum of indene with double irradiation of the *A* proton, $\lambda H_1/2\pi$ about 21 cps

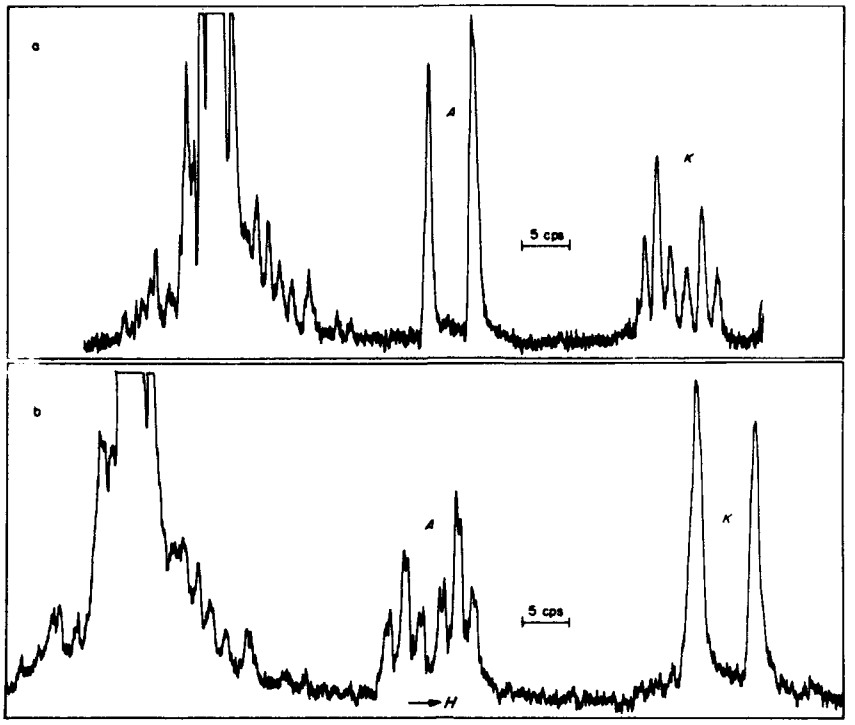


Fig. 7. Strong irradiation of the X_2 protons, $\lambda H_1/2\pi$ about 21 cps: (a) while observing the *A*-region; (b) while observing the *K*-region. Sweep rate of (b) slower than (a)

Double irradiation of the *A* region and observation of the X_2 region of the spectrum, as shown in Fig. 6, show that indeed the *X* triplet is the superposition of two doublets. Double irradiation of the *K* region while observing the X_2 region of the spectrum gave a similar result (not shown). In Fig. 7, the *A* and *K* regions are shown while the X_2 region was being strongly irradiated.

The values of the spin-coupling constants and the chemical shifts relative to tetramethylsilane (TMS) from a first-order treatment are as follows:

$J_{AK} = |5.58 \pm 0.02|$ cps $\quad \delta_A = 399.9 \pm 0.8$ cps

$J_{AX} = |1.93 \pm 0.02|$ $\quad \delta_K = 368.1 \pm 0.8$

$J_{KX} = |1.98 \pm 0.02|$ $\quad \delta_X = 179.5 \pm 0.2$

Complete solution of the ABC_2 problem on an electronic computer leads to several possible sets of parameters which depend on the relative signs of J_{AX} and J_{KX} and which are very similar.

	Set I	Set II	Set III	from TMS
J_{AK} =	$\|5.58\|$ cps	$\|5.58\|$ cps	$\|5.58\|$ cps	δ_A = 399.60 cps
J_{AX} =	± 1.93	$+1.97$	-1.98	δ_K = 368.37
J_{KX} =	± 1.98	-2.01	$+2.02$	δ_X = 179.51

These sets of parameters each gave a near-perfect fit to the experimental spectrum. The differences in the calculated intensities for the three sets of parameters were unquestionably beyond experimental detection.

2. Determination of Relative Signs of Spin Coupling Constants

We have carried out additional double irradiation experiments with a $\gamma H_1/2\pi \simeq J_{AX}$. Interpretation of the spectral changes by assigning neighboring nuclear spin states to each transition in a manner similar to that described by us in a previous summary (*RS 36-12*) enabled us to determine all the relative signs between the three coupling constants; i.e., (1) that the sign of J_{KX} is different from that of J_{AX} and (2) that the sign of J_{AK} is the same as that of J_{KX}. Thus J_{AX} has a different sign than J_{AK} and J_{KX}.

3. Discussion

This double irradiation work represents an experimental determination of the interrelationship of the signs of all the types of proton coupling constants of allylic systems. We have shown that the *cis*—H—C=C—H coupling and the C=CH—C—H coupling have a different sign than the long-range H—C—C=C—H coupling. Thus it is now possible to make relative sign assignments for all the allylic systems which have been previously reported (Refs. 16, 17, 18).

Our work is in complete agreement with Karplus' theoretical predictions for the relative signs of the sigma- (Ref. 19) and pi-electron (Ref. 20) contributions to proton-proton coupling constants. If one uses Karplus' theory, then the following absolute assignments of the signs of the coupling constants can be made for the five-membered ring protons of indene:

$$J_{AK} = +5.58 \pm 0.02 \text{ cps}$$
$$J_{AX} = -1.98 \pm 0.02$$
$$J_{KX} = +2.02 \pm 0.02$$

D. Molecular Structure and Synthesis

O. J. Klejnot

1. Some Alkoxysilanes Containing Silane Hydrogen

Alkoxylation of chlorosilanes containing silane hydrogen (Refs. 21 to 24) was reported to give low yields and side reactions (Ref. 24). Recent work on the hydrides of silicon[a] indicates that alkyl nitrites react quantitatively with such chlorosilanes at temperatures as low as −127°C (m.p. of trichlorosilane).

$$SiHCl_3 + 3RONO \rightarrow SiH(OR)_3 + 3NOCl \quad (1)$$

The silane hydrogen was found surprisingly stable toward nitrosyl chloride (Ref. 25) and all chlorine atoms of silanes were easily substituted. Cooling to a temperature between −20 and −60°C prevents gas evolution and thus prevents losses of trichlorosilane, commonly encountered with the alcohol method of alkoxylation and hitherto neglected (Ref. 24). The nitrite method was also applied to CH_3SiCl_3, CH_3SiHCl_2 and Si_2Cl_6; it has been reported (Ref. 25) for $SiCl_4$. Silicon tetrafluoride SiF_4 is inert. The application of vacuum technique to the silanes, after separation from nitrosyl chloride, was very useful. Table 6 shows constants obtained from some alkoxysilanes, alkoxymethanes and hexamethyldisiloxane. All data are new or improved over previously reported ones. Values in parentheses are extrapolated experimental values. It is observed that deviations of calculated boiling points from measured boiling points, as well as Trouton constants (*S*), increase with the degree of alkoxy substitution. Also, the alkoxysilanes are as volatile as, or more volatile than, the alkoxymethanes. This behavior also has been found with many alkoxides of transition metals such as Ti, Zr, and Hf (Ref. 26).

[a]Begun as postdoctoral research in 1954 at E. Wiberg's Laboratory, University of Munich, Germany.

2. Alkoxychlorosilanes Containing Silane Hydrogen

Little is reported (Ref. 27) on alkoxychlorosilanes containing silane hydrogen, a forgotten new class of highly reactive compounds. Their passing existence in the alcoholysis of trichlorosilane has been disregarded (Ref. 24), whereas actually the yield of trialkoxysilane in this reaction depends upon a disproportionation of alkoxychlorosilanes with exchange of ligands that is catalyzed by the present hydrogen chloride. The last chloro function of chlorosilanes reacts only slowly with alcohols. On the other hand the alkoxychlorosilanes tend to evolve hydrogen with alcoholic hydrogen chloride.

With the aid of the new alkyl nitrite method (Eq. 1), the alkoxychlorosilanes were easily obtained in high yields. Alkyl nitrite was passed as vapor of 10 mm pressure onto cold, stirred trichlorosilane at −60°C and was absorbed rapidly. The formation of trialkoxysilanes was prevented at a molar ratio of 1.0 silane to 1.1 nitrite. The mixture of alkoxychlorosilanes was separated in a vacuum column at tap-water temperature.

$$SiHCl_3 + RONO \rightarrow SiHCl_2OR + NOCl \quad (2)$$
$$SiHCl_2OR + RONO \rightarrow SiHCl(OR)_2 + NOCl \quad (3)$$

Table 7 shows constants of some chlorosilanes and alkoxychlorosilanes, interesting for comparison in regard to the effect of degree of substitution. The silanes $SiHCl_2OCH_3$ (some values estimated) and $CH_3SiHClOCH_3$ (obtained pure for a limited time with P_0 88 mm Hg) tend to disproportionate at room temperature even without catalysis by hydrogen chloride (exchange of OCH_3 and Cl). The chloro functions of alkoxychlorosilanes bearing silane hydrogen are hydrogenated with lithium borohydride. The results of studies on disproportionation and on silanes bearing several hydrogen atoms at the silicon atom will be reported shortly.

Table 6. Physical and thermodynamic constants[a]

Formula	log P (mm) = B − A/T		P_0, mm Hg	D_0, g/ml	MP, °C	BP, °C	BP, calcd	H_{vap}	S_{bp}	$n_D^{25°}$
	B	A								
$CH(OCH_3)_3$	8.3885	2039.0	8.5	0.974	−55.38	100.7	97.9	9330	24.96	1.3765
$SiH(OCH_3)_3$	8.3100	1929.0	17.7	0.932	−113.5	83.9	82.1	8827	24.73	—
$Si(OCH_3)_4$	8.2770	2128.2	2.5	1.040	+5.35	(123.0)	121.2	9738	24.58	1.3662
$CH_2CH(OCH_3)_3$	7.9287	1696.2	50.5	0.880	glassy	63.3	62.9	7762	23.06	1.3638
$CH_3SiH(OCH_3)_2$	8.0139	1712.5	55.7	0.896	glassy	61.1	60.5	7836	23.44	1.3574
$CH_3SiH(OC_2H_5)_2$	8.0628	1922.8	9.6	0.875	glassy	(99.4)	97.9	8799	23.62	1.3724
$(Me_2Si)_2O$	7.9875	1898.9	10.0	0.794	−69.8	(100.5)	98.7	8689	23.25	1.3745

[a] D_0 and P_0; at 0°C.

Table 7. Physical and thermodynamic constants[a]

Formula	log P (mm) = B − A/T		P_0, mm Hg	D_0, g/ml	MP, °C	BP, °C	BP, calcd	H_{vap}	S_{bp}
	B	A							
$(CH_3)_2SiHCl$	7.6161	1458.0	189.6	0.910	−103.0	34.9	34.8	6672	21.66
CH_3SiHCl_2	7.6784	1506.1	146.5	1.162	−92.5	41.1	40.8	6891	21.92
CH_3SiCl_3	7.6160	1602.7	55.8	1.287	−77.8	65.4	65.3	7333	21.66
$CH_3SiCl_2OC_2H_5$	7.8500	1807.3	11.6	1.132	−91.8	(100.1)	99.0	8462	22.66
$CH_3SiHClOC_2H_5$	7.9007	1737.8	33.4	1.016	−108.0	73.6	73.0	7951	22.93
$SiHCl_2OC_2H_5$	7.9004	1744.5	32.2	1.179	−114.0	75.7	74.4	7982	22.88
$SiHCl_2OCH_3$	(7.8400)	(1600.0)	95.3	1.204	−105.0	(50.5)	(49.5)	(7321)	(22.6)
$SiHCl(OCH_3)_2$	8.0616	1774.0	36.9	1.078	−119.0	70.5	69.3	8117	23.62

[a] D_0 and P_0; at 0°C.

References

1. Dodd, R. E., Rolfe, J. A., and Woodward, L. A., *Transactions of the Faraday Society*, 52:145, 1956.
2. Smith, D. F., and Magnuson, D. W., *Physical Review*, 87:226, 1952.
3. Anderson, R., and MacLaren, R. O., *2nd Quarterly Technical Summary Report*, Contract No. Nonr 3433(00); ARPA Order No. 184-61, United Technology Corporation, Sunnyvale, Calif., October 1961.
4. MacLaren, R. O., Private Communication, and Anderson, R., and MacLaren, R. O., *3rd Quarterly Technical Summary Report*, Contract No. Nonr 3433(00), ARPA Order No. 184-61, United Technology Corporation, Sunnyvale, Calif., February 1962.
5. Cordes, H. F., and Johnston, H. S., *Journal of the American Chemical Society*, 76:4264, 1954.
6. Perrine, R. C., and Johnston, H. S., *Journal of Chemical Physics*, 21:2205, 1952.
7. Cordes, H. F., et. al., *Analytical Chemistry*, 29:485–487, 1957.

References (Cont'd)

8. Weaver, R., et. al., *Analytical Chemistry*, 28:329–332, 1956.
9. Streim, et al., *Analytical Chemistry*, 33:85–88, 1961.
10. Nomenclature is that used in *High-Resolution Nuclear Magnetic Resonance* by Pople, J. A., Schneider, W. G., and Berstein, H. J., McGraw Hill Book Co., New York, 1959.
11. Kaiser, R., *Review of Scientific Instruments*, 31:963, 1960.
12. Freeman, R., *Molecular Physics*, 3:435, 1960.
13. Mahler, J. P., and Evans, D. F., *Proceedings of the Chemical Society*, 208, 1961.
14. Freeman, R., private communication; Freeman, R., and Whiffen, D. H., *Molecular Physics*, 4:321, 1961.
15. See Ref. 10, pp 238–241.
16. Fessenden, R. W., and Waugh, J. S., *Journal of Chemical Physics*, 30:944, 1959.
17. Bishop, E. O., and Richards, R. E., *Molecular Physics*, 3:114, 1960.
18. Banwell, C. N., and Sheppard, N., *Molecular Physics*, 3:352, 1960.
19. Karplus, M., *Journal of Chemical Physics*, 30:11, 1959.
20. Karplus, M., *Journal of Chemical Physics*, 33:1842, 1960.
21. Friedel, C., and Ladenburg, A., *Liebigs Annalen*, 143:118–128, 1867.
22. Taurke, F., *Berichte der Deutschen Chemischen Gesellschaft*, 38:1661–1670, 1905.
23. Helferich, B., and Hausen, J., *ibid* 57:795–799, 1924.
24. Havill, M. E., Jaffe, I., and Post, H. W., *Journal of Organic Chemistry*, 13, 280, 1948.
25. Beckham, L. J., Fessler, W. A., and Kise, M. A., "Nitrosyl Chloride," *Chemical Reviews*, 48:319–396, 1951.
26. Bradley, D. C., "Organo-Compounds of Transition Metals," *Record of Chemical Progress*, 21, No. 3:179–187, 1960.
27. Reilly, A. F., and Post, H. W., *Journal of Organic Chemistry*, 16:387, 1951.

IX. Gas Dynamics

A. The Laminar Boundary Layer on a Disk of Finite Radius in a Rotating Flow

L. M. Mack

The momentum-integral method developed by von Kármán for the turbulent boundary layer on a rotating disk was applied previously (*RS 36-9*, Vol I, p 68) to the calculation of the inward radial mass flow in the laminar boundary layer on a stationary disk of finite radius in a rotating outer flow. The calculation was performed for outer flows with zero radial velocity, and with the tangential velocity v_∞, given by the power law

$$v_\infty = \frac{1}{r^n} \tag{1}$$

In this equation, both v_∞ and r, the radius, are dimensionless (see Eq. 5 below). Both Schultz-Grunow (Ref. 1) and Taylor (Ref. 2) have used this method in the past for similar problems; Schultz-Grunow for the case where the outer flow is in solid-body rotation, $n = -1$; Taylor for the case of free-vortex outer flow, $n = 1$. Neither author computed the inward radial mass flow in the boundary layer. When it was computed (*RS 36-9*, Vol I, p 68), it was found, for $n = 1$, to increase monotonically from the edge of the disk, $r = 1$, to the center of the disk, $r = 0$. For $n < 0$, it was found to have a maximum at some radius, r_{max}, where $0 < r_{max} < 1$, and to be zero at $r = 0$.

On the contrary, if the inward radial mass flow is computed from the results of Cooke (Ref. 3), who investigated Taylor's problem ($n = 1$) with another version of the momentum-integral method, a different result is obtained. The mass flow, instead of increasing monotonically with decreasing r, is close to the result obtained for $n = 0.4$ with the method of *RS 36-9*; i.e., it has a maximum at some radius away from the center. It is not possible to resolve this contradiction within the framework of the momentum-integral method. A more exact solution is needed, such as that provided by the Blasius-type series of Stewartson (Ref. 4). Stewartson obtained and solved only the equations for the first terms of the series. This solution, which is a similarity solution, is independent of the radial pressure gradient and is valid only close to $r = 1$. The further development of the method will be described below, and the results obtained with it will be compared with those given by the two momentum-integral methods.

The dimensionless boundary-layer equations in cylindrical coordinates are

$$u\frac{\partial u}{\partial r} + w\frac{\partial u}{\partial z} - \frac{v^2}{r} = -\frac{v_\infty^2}{r} + \frac{\partial^2 u}{\partial z^2} \tag{2}$$

$$u\frac{\partial v}{\partial r} + w\frac{\partial v}{\partial z} + \frac{uv}{r} = \frac{\partial^2 v}{\partial z^2} \tag{3}$$

$$\frac{\partial}{\partial r}(ru) + \frac{\partial}{\partial z}(rw) = 0 \tag{4}$$

where z is the axial coordinate, and u, v, and w are the radial, tangential, and axial velocities. The notation here

differs somewhat from that used in RS 36-9. The definitions of the dimensionless quantities are

$$r = \frac{r^*}{r_1^*} \qquad z = \frac{z^*}{r_1^*}(Re_t)^{1/2}$$
$$u = \frac{u^*}{v_1^*} \qquad v = \frac{v^*}{v_1^*} \qquad w = \frac{w^*}{v_1^*}(Re_t)^{1/2} \tag{5}$$

where Re_t is the peripheral tangential Reynolds number $v_1^* r_1^*/\nu^*$. The asterisks refer to dimensional quantities. The reference quantities r_1^* and v_1^* are the radius of the disk and the tangential velocity of the outer flow at r_1^*. The quantity ν^* is the kinematic viscosity. The boundary conditions are

$$u = 0, v = 0, w = 0 \text{ at } z = 0$$
$$u \to 0, v \to v_\infty \text{ as } z \to \infty \tag{6}$$

One change is now made from Stewartson's formulation. The dependent variable v is replaced by $V = v/v_\infty$ in accordance with the ideas of Görtler (Ref. 5). The reason for this change is to simplify the boundary condition on v as $z \to \infty$. With V as the dependent variable, this boundary condition is satisfied by the Stewartson term, and all of the boundary conditions for the other terms in the series are homogeneous. The boundary layer equations become

$$u\frac{\partial u}{\partial r} + w\frac{\partial u}{\partial z} = -\frac{v_\infty^2}{r}(1 - V^2) + \frac{\partial^2 u}{\partial z^2} \tag{7}$$

$$u\frac{\partial V}{\partial r} + w\frac{\partial V}{\partial z} = -\frac{uV}{r}\left[1 + \frac{d(\log v_\infty)}{d(\log r)}\right] + \frac{\partial^2 V}{\partial z^2} \tag{8}$$

Several more changes of variable are now made. First a stream function, ψ, is introduced for the velocity components u and w.

$$u = -\frac{1}{r}\frac{\partial \psi}{\partial z} \qquad , w = \frac{1}{r}\frac{\partial \psi}{\partial r} \tag{9}$$

Next, the similarity variable of Stewartson, defined by

$$y = \frac{z}{(1 - r)^{1/4}} \tag{10}$$

and his similarity stream function, defined by

$$F(r, y) = \frac{\psi[r, z(r, y)]}{(1 - r)^{3/4}} \tag{11}$$

are used in place of z and ψ. Finally the radius is replaced by

$$x = 1 - r \tag{12}$$

and Eq. (1) is used for the outer-flow tangential velocity distribution. The transformed boundary layer equations are

$$(1 - x)^2\frac{\partial^3 F}{\partial y^3} + \frac{3}{4}(1 - x)F\frac{\partial^2 F}{\partial y^2} - \left[x + \frac{1}{2}(1 - x)\right]\left(\frac{\partial F}{\partial y}\right)^2$$
$$= x(1 - x)\left(\frac{\partial F}{\partial y}\frac{\partial^2 F}{\partial x\,\partial y} - \frac{\partial F}{\partial x}\frac{\partial^2 F}{\partial y^2}\right) - (1 - x)^{2(1-n)}(1 - V^2) \tag{13}$$

$$(1 - x)^2\frac{\partial^2 V}{\partial y^2} + \frac{3}{4}(1 - x)F\frac{\partial V}{\partial y} = x(1 - x)\left(\frac{\partial V}{\partial x}\frac{\partial F}{\partial y} - \frac{\partial V}{\partial y}\frac{\partial F}{\partial x}\right) - (1 - n)xV\frac{\partial F}{\partial y} \tag{14}$$

Near the edge of the disk, $x = 0$, these equations reduce to Stewartson's ordinary differential equations.

The stream function, F, and tangential-velocity ratio, V, are now expanded in power series in x.

$$F(x, y) = \sum_{p=0}^{\infty} x^p F_p(y) \tag{15}$$

$$V(x, y) = \sum_{p=0}^{\infty} x^p V_p(y) \tag{16}$$

When these series are substituted into Eqs. (13) and (14), and equal powers of x are equated, a recursive set of ordinary differential equations is obtained. The first terms, F_0 and V_0, satisfy the non-linear Stewartson equations; F_p and V_p, with $p > 0$, satisfy linear inhomogeneous equations whose coefficients contain $F_{p'}$ and $V_{p'}$ where $p' = 0, 1, \cdots, p - 1$. The boundary conditions are

$$y = 0: \qquad F_p = 0, \frac{dF_p}{dy} = 0, V_p = 0$$
$$y \to \infty: \qquad \frac{dF_p}{dy} \to 0, V_0 \to 1 \qquad V_p \to 0, p > 0 \tag{17}$$

For each value of p, there is a two-point boundary value problem to be solved for a fifth-order system of ordinary differential equations. It is necessary to find the unknown initial conditions $(d^2F_p/dy^2)_{y=0}$ and $(dV_p/dy)_{y=0}$ which satisfy the boundary conditions at infinity. These initial conditions were found on an IBM 7090 computer by a direct search procedure. Six terms of the series were computed, and the results for $n = 1$ are very good over most of the surface of the disk. As n decreased from unity, it became more difficult to obtain a good numerical solution. At $n = 0.5$, the solution is only good for $r > 0.35$.

When the F_p are determined, the dimensionless inward radial mass flow is found from

$$M(r) = (1 - r)^{3/4}\sum_{p=0}^{\infty}(1 - r)^p F_p(\infty) \tag{18}$$

The definition of M in terms of the dimensional mass flow, M^*, is

$$M = \frac{M^*(Re_i)^{1/2}}{2\pi\rho^* v_1^* r_1^{*2}} \tag{19}$$

where

$$M^*(r^*) = -2\pi r^* \rho^* \int_0^\infty u^* \, dz^* \tag{20}$$

The quantity ρ^* is the density.

The Stewartson similarity solution can also be used to improve the momentum-integral solution. Since the method used in *RS 36-9* represents each of the velocity components, u and v, by a single profile function which is independent of the radius, any reasonable function can be used for the velocity profiles. In particular, the Stewartson profile functions can be used. Near $r = 1$, the momentum-integral solution will then also be the exact boundary-layer solution. The details of this procedure and the results obtained with it are presented in Ref. 6.

The results for the inward radial mass flow obtained from the series solution and from the momentum-integral method when the Stewartson velocity profiles are used are shown in Fig. 1. Also shown is Cooke's result for $n = 1$. It is evident that for $n = 1$ the version of the momentum-integral method used in *RS 36-9*, and improved numerically in Ref. 6, is in the closest possible agreement with the series solution for $r > 0.1$, and that the result given by Cooke's method must be judged incorrect. For $n < 1$, the momentum-integral method and series method are in good agreement from $r = 1$ to a radius somewhat smaller than the radius at which the mass flow is a maximum. For smaller radii, the series solution is no longer adequate. However, it has established that, for at least $n \geq 0.5$, the momentum-integral method of *RS 36-9* and Ref. 6 can be used with confidence to predict the maximum inward radial mass flow carried in the boundary layer.

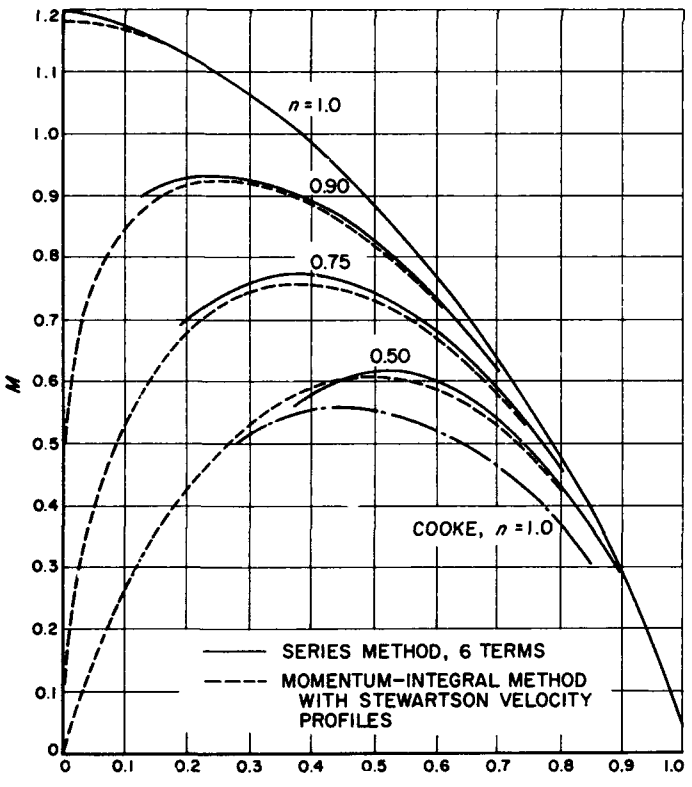


Fig. 1. Inward radial mass flow distribution in boundary layer as computed by series method and two momentum-integral methods

B. The Structure of a Strong Shock Wave in the Krook Collision Model

M. T. Chahine

The problem of the one-dimensional flow of a real gas through a shock wave is studied using a microscopic formulation similar to that proposed by Prof. Max Krook (Ref. 7) and is based on approximating the Boltzmann collision integral in terms of the first five moments of the distribution function.

The main purpose of the analysis has been to establish the iterative method of solution and then to apply it to find the velocity and temperature profiles for a steady-plane shock wave in a monatomic gas for Mach number $1 \leq M \leq 20$. This provides on one hand some insight into the range of validity of the other approximation methods, and on the other hand a check on the usefulness of Krook's collision model in this problem.

Theory

In standard notation, the Krook collision model for one-dimensional steady-state flow with no external forces is described by the equation (Ref. 8)

$$v_x \frac{\partial f}{\partial x} = An\,[F - f] \tag{1}$$

with

$$A = \frac{\bar{c}}{n\lambda} \tag{2}$$

where n is the number density of molecules, and $\bar{c}$ is their average speed, λ is the mean free path, and F corresponds to the local Maxwellian

$$F = \frac{n}{(2\pi RT)^{3/2}} \exp\left[- \frac{(\mathbf{v} - u)^2}{2RT} \right] \tag{3}$$

whose parameters n, u, T are given by the actual local conditions. The conservation equations are satisfied by Eq. (1), and A (Eq. 2) is chosen such that the number of collisions corresponds to a complete Maxwellian flow.

To determine the flow through a shock, we consider a plane stationary shock wave in a nonreacting gas; the boundary conditions are

$$f(\mathbf{v}, -\infty) = F_1 = \frac{n_1}{(2\pi RT_1)^{3/2}} \exp\left[- \frac{(\mathbf{v} - u_1)^2}{2RT_1} \right]$$

$$f(\mathbf{v}, +\infty) = F_2 = \frac{n_2}{(2\pi RT_2)^{3/2}} \exp\left[- \frac{(\mathbf{v} - u_2)^2}{2RT_2} \right] \tag{4}$$

Using different representation for f in the two regions of velocity-space defined by f_- in the half space $v_x < 0$ and f_+ in the half space $v_x > 0$, and applying the boundary conditions defined by Eq. (4), then Eq. (1) reduces to the form

$$f_-(\mathbf{v}, x) = \int_{+\infty}^{x} \frac{An}{v_x} F \exp\left[\int_x^y \frac{An}{v_x}\, dy' \right] dy$$

$$f_+(\mathbf{v}, x) = \int_{-\infty}^{x} \frac{An}{v_x} F \exp\left[\int_x^y \frac{An}{v_x}\, dy' \right] dy \tag{5}$$

The number density $n(x)$, the flow velocity $u(x)$, and the kinetic temperature $T(x)$ are then defined by the usual relations

$$\text{(a)} \qquad n = \int_{-\infty}^{+\infty} f\, d\mathbf{v}$$

$$\text{(b)} \qquad nu = \int_{-\infty}^{+\infty} v_x f\, d\mathbf{v}$$

$$\text{(c)} \qquad 3RnT = \int_{-\infty}^{+\infty} (\mathbf{v} - u)^2 f\, d\mathbf{v} \tag{6}$$

These are the basic equations used to determine the variations of flow parameters and the iteration procedure. The iterative solution is generated by the differential equation

$$v_x \frac{\partial f^{[m+1]}}{\partial x} = A^{[m]} n^{[m]} \left[F^{[m]} - f^{[m+1]} \right] \tag{7}$$

To start the iteration procedure with $F^{[0]}$, we derive from Eq. (1) the familiar macroscopic equations of motion. It is easy to verify that the definition of the parameters in F insures that

$$\text{mass} \qquad \frac{d}{dx}(nmu)^{[0]} = 0$$

$$\text{momentum} \qquad \frac{d}{dx}(nmu^2 + nmRT + P_{xx})^{[0]} = 0$$

$$\text{energy} \qquad \frac{d}{dx}\left[nmu \left(\frac{3RT}{2} + \frac{u^2}{2} \right) + u(nmRT + P_{xx}) + q_x \right]^{[0]} = 0 \tag{8}$$

A determinate set of equations is obtained by adding to these the relations

$$P_{xx} = -\frac{4}{3}\mu \frac{du}{dx} \qquad q_x = k \frac{dT}{dx} \tag{9}$$

Equations (8) and (9) are the familiar Navier-Stokes equations which are then solved (Ref. 9) to determine $n^{[0]}(x)$, $u^{[0]}(x)$, and $T^{[0]}(x)$ through the shock. These parameters define the function $F^{[0]}$ in Eq. (3) as

$$F^{[0]} = \frac{n^{[0]}}{(2\pi RT^{[0]})^{3/2}} \exp\left[- \frac{(\mathbf{v} - u^{[0]})^2}{2RT^{[0]}} \right] \tag{3'}$$

which define the functions $f_-^{[1]}$ and $f_+^{[1]}$ in Eq. (6) as

$$f_-^{[1]} = \int_{+\infty}^{x} \frac{A^{[0]} n^{[0]}}{v_x} F^{[0]} \exp\left[\int_x^y \frac{A^{[0]} n^{[0]}}{v_x}\, dy' \right] dy$$

$$f_+^{[1]} = \int_{-\infty}^{x} \frac{A^{[0]} n^{[0]}}{v_x} F^{[0]} \exp\left[\int_x^y \frac{A^{[0]} n^{[0]}}{v_x}\, dy' \right] dy \tag{6'}$$

and substitution of Eq. (6) into Eq. (7) determine the new values of the moments $u^{[1]}(x)$, $n^{[1]}(x)$ and $T^{[1]}(x)$, which generate the next step in the iteration procedure.

The difference between the Chapman-Enskog sequence and this iteration procedure should be noted here since, unlike the Chapman-Enskog method, this iteration is carried on the first five moments and not on the distribution function f as such.

Preliminary results for Mach number 3 tend to indicate that they are not significantly different from the Navier-Stokes solution on the higher density side of the shock. However, there is a clear disagreement on the lower density side, with the present theory converging slower to the upstream conditions.

A more detailed account of the work, together with the results of computations, will be published in the near future.

C. Heat Loss from Hot Wires in Transonic Flow

T. Vrebalovich

In *RS 36-1*, the wind tunnel technique which made use of a normal shock diffuser inlet to obtain both low Reynolds number and transonic Mach number flows was described. After the inlet was calibrated, hot-wire heat loss and recovery temperature measurements were made over the following approximate ranges of Mach number and Reynolds number: $0.5 \leq M \leq 1.6$ and $0.5 < Re_2 < 60$, where Re_2 is the wire Reynolds number based on wire diameter and local static temperature (T_2) for $M \leq 1$, or static temperature (T_2) behind the normal shock wave in front of the wire for $M > 1$. The hot wires were 90% platinum and 10% rhodium. Wires 0.00005, 0.00001, and 0.00015-in. in diameter and approximately 400 diameters long were used. The heat loss data for the 0.00005-in. wire is not included because of length errors and irregularities in the wire.

Kovásznay (Ref. 10) formulated a method for making end loss corrections on finite length hot-wire heat-loss measurements. He assumed that the needles to which the wire was attached were at the recovery temperature of the unheated wire. This assumption is permissible only when the wire Reynolds number (Re_2) is larger than 20. For Re_2 smaller than 20 the recovery temperature of the wire is greater than that for the needles, and the Kovásznay assumption leads to an error in both the recovery temperature and the corrected heat loss measurements (Ref. 11). The measured recovery temperature must be corrected and a slightly different end loss correction to the heat-loss measurements should be made. The results indicate that, in the range of Mach and Reynolds numbers used in this experiment, the recovery temperature but not the heat loss was affected by making this additional correction.

A proper end loss correction was made by the following method. For a small element of wire of length dx, the differential equation for the heat balance to the element may be written

$$\pi hD\,[T - T_a] - \frac{\pi D^2}{4} K_w \frac{d^2T}{dx^2} - i^2 r = 0$$

where the first term is convection to the airstream; the second is the conduction along the wire; the third is the heat generated in the element due to the current in the wire, and the radiation is neglected. The list of symbols follows

h = heat transfer coefficient

Nu = Nusselt number = hd/K, $(L = \infty)$

D = diameter of wire

L = length of wire

K = conductivity of air

K_w = conductivity of wire

r = local resistance per unit length

$r = r_0\,[1 + \alpha\,(T - T_{0°C})]$

i = current in the wire

α = temperature coefficient of resistivity, °C

T = local wire temperature

T_a = wire recovery temperature $(L = \infty)$

T_T = stagnation temperature of air

T_r = measured recovery temperature of wire $(i = 0)$

T_w = average temperature of heated wire

T_n = recovery temperature of needles

T_s = local static temperature in inlet

$\mathcal{R}$ = recovery factor for needles

$$\mathcal{R} = 0.85 = \frac{T_n - T_s}{T_T - T_s} \text{ (Ref. 12)}$$

$$\eta = \frac{T_a}{T_T}$$

By writing the equation in the following form

$$\frac{d^2T}{dx^2} - ST = -P$$

and using the boundary condition that the needle temperature is T_n, the solution is

$$T = \frac{T_n - P/S^2}{\cosh \frac{SL}{2}} \cosh sx + P/S^2$$

Since the average heated wire temperature is T_w,

$$T_w = \frac{2}{L} \int_0^{L/2} T dx$$

Therefore

$$T_w = (T_n - P/S^2) \frac{\tanh \frac{SL}{2}}{\frac{SL}{2}} + P/S^2$$

where

$$P = \frac{4}{D^2 K_w}\left\{KNuT_a + \frac{i^2 r_0}{\pi}(1 - \alpha T_{0°C})\right\}$$

$$S^2 = \frac{4}{D^2 K_w}\left\{KNu - \frac{\alpha i^2 r_0}{\pi}\right\}$$

For $T_r = T_a$, the Kovásznay solution is obtained. For $i = 0$, $T_w = T_r$. There are two unknowns, Nu and T_a, and it is clear that two equations may be obtained by letting $i = 0$ and $i \neq 0$. In fact, by operating at several currents, the second equation ($i \neq 0$) may be paired several times with the $i = 0$ equation. The result of this is the well-known fact that the Nusselt number is indeed slightly dependent on overheat, whereas the recovery temperature is constant, independent of which overheat equation is combined with the $i = 0$ equation.

The fact that the recovery temperature was not very dependent on Nusselt number made the solution difficult by ordinary IBM 7090 computer techniques for the simultaneous solution of the resulting pairs of equations. Fortunately each equation could be rewritten in the form $T_a = T_a(Nu)$. A first approximation to the Nusselt number was chosen by the Kovásznay method and Nu was varied in small steps near this value until the T_a's for both equations were equal in the fourth significant figure.

Even for the lowest wire Reynolds numbers, the difference between the end loss correction on Nu and that estimated by Kovásznay differed by less than 1%. This was not true for the recovery temperature. The data presented in Fig. 2 is a plot of the recovery temperature ratio, $\eta = T_a/T_T$ vs Re_2. This data is the result of fairing the corrected data and may be compared at $M = 1.6$ and $M = 0.6$ to the uncorrected data, shown by the dashed curves in Fig. 2. The correction for $M = 1.6$ was 4% at $Re_2 = 0.5$ and the curves match at $Re_2 > 12$, whereas

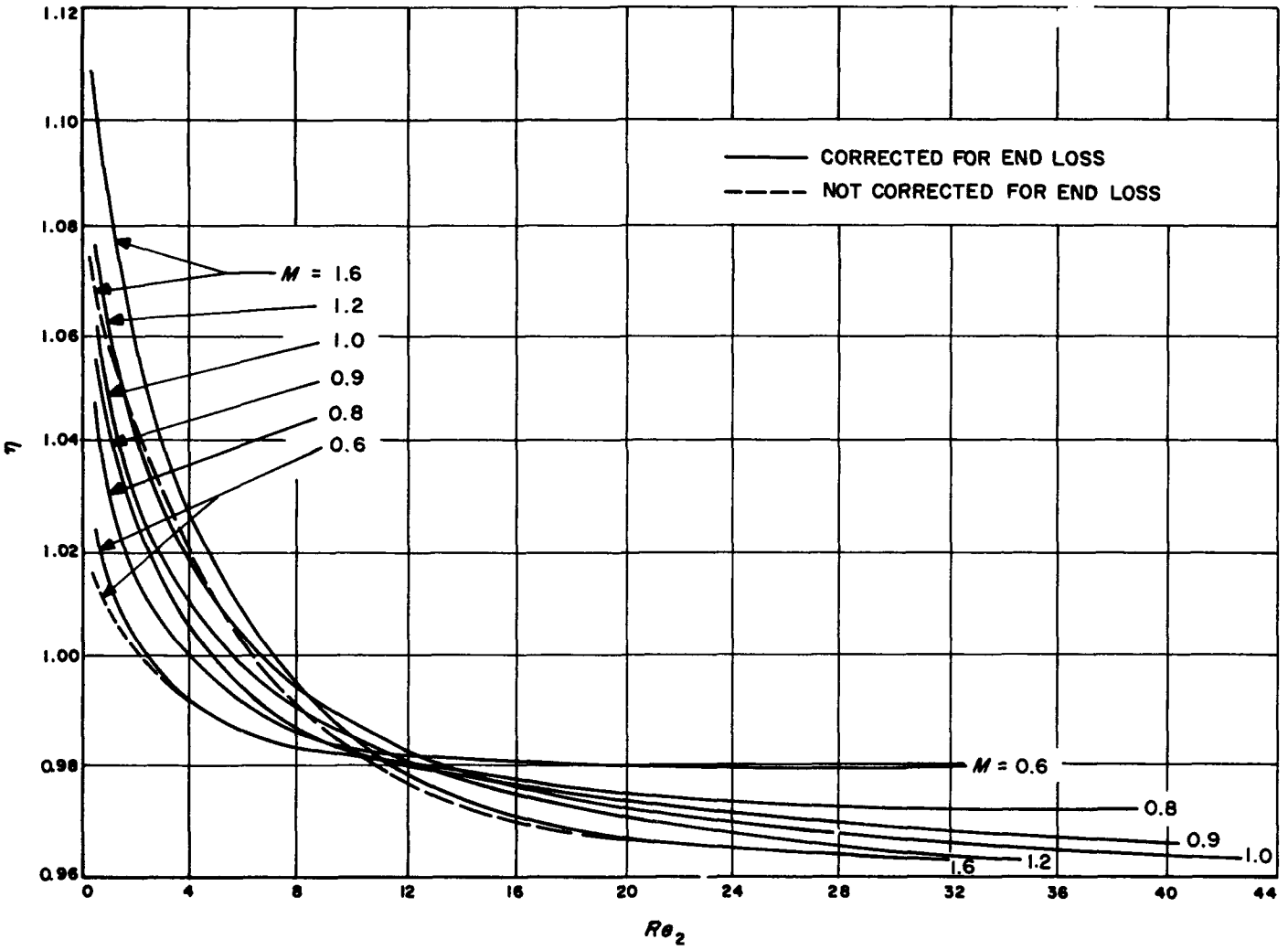


Fig. 2. Recovery temperature ratio in transonic flow

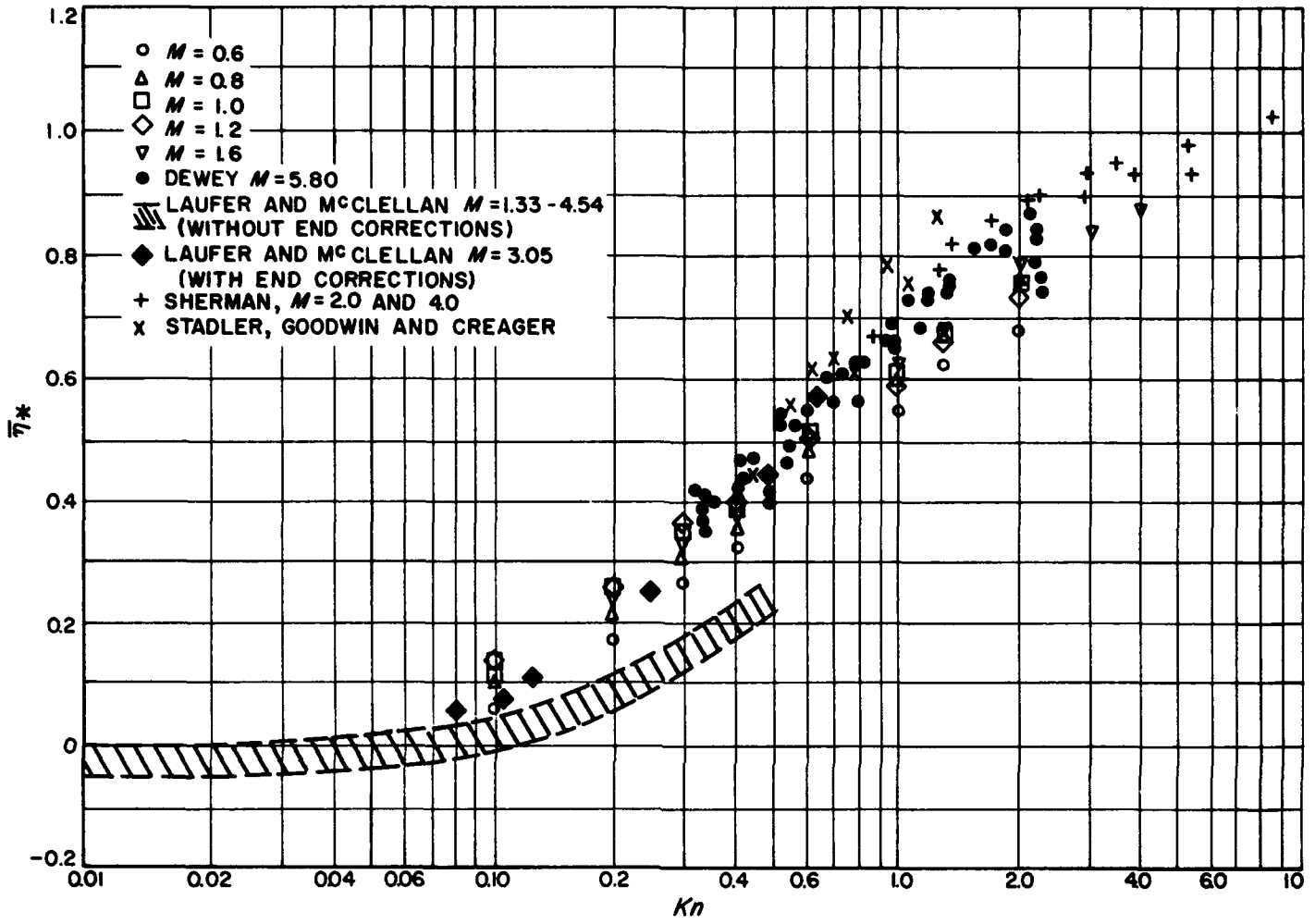


Fig. 3. Normalized recovery ratio vs free stream Knudsen number

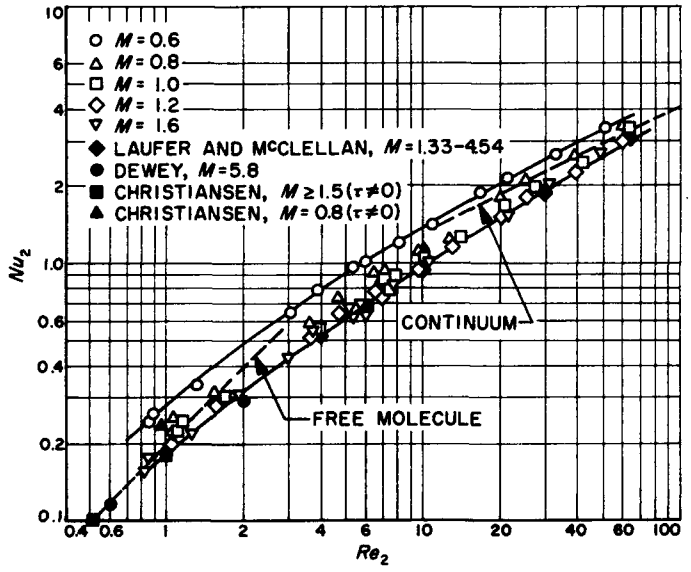


Fig. 4. Heat transfer from hot wires in transonic flow, $\tau = 0$

at $M = 0.6$ and $Re_2 = 0.5$ the error is less than 1% and corrections are negligible beyond $Re_2 > 2$.

The coordinate system used by Dewey (Ref. 11) in which the normalized recovery temperature ratio $\bar{\eta}_*$ is plotted against free-stream Knudsen number is a means of comparing the data of different experimenters, Fig 3. $\bar{\eta}_*$ is defined by

$$\bar{\eta}_* = \frac{\eta - \eta_c}{\eta_f - \eta_c}$$

where

$$\eta_c = \frac{T_c}{T_T}$$

for T_c = the continuum recovery temperature

and

$$\eta_f = \frac{T_f}{T_T}$$

for T_f = the free molecule recovery temperature

The data obtained by Laufer and McClellan at the Jet Propulsion Laboratory (Ref. 14) when corrected for end loss lies within the scatter of Dewey's data. The present data is self-consistent and the lower Mach number data lies below the higher Mach number data.

The heat loss data is presented in Fig. 4 and is a plot of Nu_2 vs Re_2 where Nu_2 and Re_2 are based on T_2. Nu_2 is also taken at zero overheat

$$\tau = \frac{T_w - T_a}{T_a} = 0$$

by fairing $\tau > 0$ data to the $\tau = 0$ intercept. The data checks reasonably well with that of Refs. 11, 13, and 14. The $M = 1.6$ data also seems to fair into the continuum and free molecule asymptotes.

The curves of Figs. 3 and 4 indicate that most of the data lies in what might be called the "slip flow" regime, which is usually defined as the region between continuum and free molecular flow. The transonic measurements furthermore provide hot-wire calibrations which may be used in determining fluctuations in supersonic turbulent boundary layers. The computational technique with some modification and with a radiation correction will be applicable in the hypersonic case.

References

1. Schultz-Grunow, F., "Der Reibungswidertand rotierender Scheiben in Gehäusen," *Zeitschrift für angewandte Mathematik und Mechanik*, Vol 15, p 191, 1935.
2. Taylor, G. I., "The Boundary Layer in the Converging Nozzle of a Swirl Atomizer," *Quarterly Journal of Mechanics and Applied Mathematics*, Vol 3, pp 129–139, 1950.
3. Cooke, J. C., "On Pohlhausen's Method with Application to a Swirl Problem of Taylor," *Journal of the Aeronautical Sciences*, Vol 19, pp 486–490, 1952.
4. Stewartson, K., "On Rotating Laminar Boundary Layers," *Symposium Freiburg*, pp 59–70, Springer Verlag, Berlin, 1958.
5. Görtler, H., "A New Series for the Calculation of Steady Laminar Boundary Layer Flows," *Journal of Mathematics and Mechanics*, Vol 6, p 1, 1957.
6. Mack, L. M., *The Laminar Boundary Layer on a Disk of Finite Radius in a Rotating Flow. Part I. Numerical Integration of the Momentum-Integral Equations and Application of the Results to the Flow in a Vortex Chamber*, Technical Report No. 32-224, Jet Propulsion Laboratory, Pasadena, Calif. (to be published)
7. Bhatnagar, P. L., Gross, E. P., and Krook, M., *Physical Review*, Vol 94, No. 3, May 1, 1954.
8. Narasimha, R., *Ph. D. Thesis*, California Institute of Technology, 1961.
9. Gilbarg, D., and Paolucci, D., *Journal of Rational Mechanics and Analyses*, 2, No. 4, pp 617–642, 1953.

References (Cont'd)

10. Kovásznay, L. S. G., "Turbulence Measurements," *Physical Measurements in Gas Dynamics and Combustion,* Vol IX, Princeton University Press, Princeton, New Jersey, 1954.

11. Dewey, C. F., Jr., "Hot Wire Measurements in Low Reynolds Number Hypersonic Flows," ARS *Journal,* pp 1709–1718, December, 1961.

12. Mack, L. M., "Summary of Existing Cone Temperature Recovery Factor Data Obtained in Continuous-Flow Wind Tunnels Plus Additional Data," *Journal of the Aeronautical Sciences, Readers' Forum,* Vol 22, No. 3, pp 212–214, March 1955.

13. Laufer, J., and McClellan, R., "Measurement of Heat Transfer from Fine Wires in Supersonic Flows," *Journal of Fluid Mechanics,* Vol 1, Part 3, p 276, September 1956.

14. Christiansen, W. H., *Development and Calibration of a Cold Wire Probe for Use in Shock Tubes,* GALCIT Hypersonic Research Project, Memorandum No. 62, July 1, 1961.

X. Physics Research

A. Superconducting Flux Pump. II

D. D. Elleman, A. F. Hildebrandt, R. Simpkins, and F. C. Whitmore

We have previously described a superconducting flux pump constructed from a pair of Nb_3Zr coils (Ref 1 and *RS 36-12*, p. 108). In this paper we discuss the results of experiments performed on a superconducting flux pump constructed from a block of 99.6% niobium machined into the configuration shown in Fig. 1. The block is 1.90 cm thick and 4.14 cm in diameter. There are two large cylindrical holes drilled through the block: one hole is 0.96 cm diameter and serves as the compression chamber; the other hole is 2.38 cm diameter and is the experimental chamber. In addition, very small holes are drilled in the wall which separates the two chambers and also in the outside wall of the compression chamber. Nichrome heating wires are placed in these small holes. When power is supplied to these heating wires, the temperature of the niobium is raised above the critical temperature and the niobium in the vicinity of the heating wires changes from the superconducting state to the normal state. Thus there are two superconducting "valves" which can be opened and closed to let magnetic flux into or out of a particular chamber. A niobium cylinder 2.12 cm long and 0.89 cm diameter is used as a piston to compress flux in the compression chamber. The sequence of operation of the valves and the piston is identical to that of the superconducting pump made of Nb_3Zr wire described in Ref. 1 and *RS 36-12*.

Fig. 1. Superconducting flux pump 99.6% niobium

In Fig. 2, the magnetic field measured in the experimental chamber is plotted vs the number of cycles the pump is operated. It is obvious that the experimental chamber may be made as large as desired, for the maximum magnetic field obtained depends only on the relative cross-sectional areas of the piston and the compression chamber. It is to be noted that, as the experi-

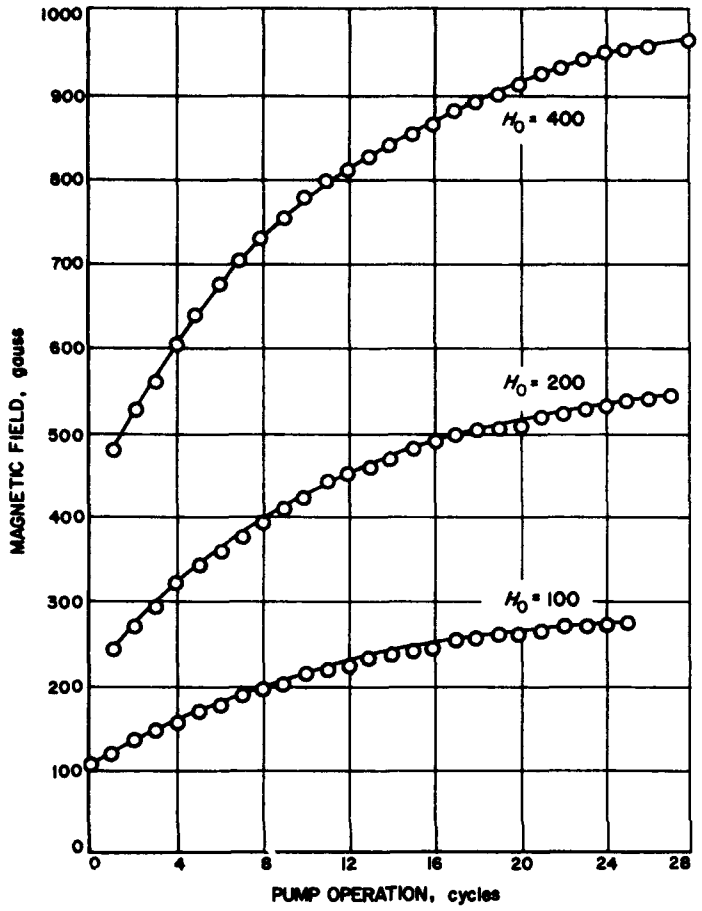


Fig. 2. Plot of magnetic field in experimental chamber vs number of cycles Nb flux pump is operated

mental chamber is made large compared to the compression chamber, the number of cycles required to reach the final field is increased. Also, the final magnetic field cannot exceed the critical field of the superconducting material.

Further experiments are being conducted with superconducting flux pumps constructed of Nb_3Sn material. In Fig. 3 a Nb_3Sn pump is shown which does not have any valves in it, so that all of the flux must be pumped into the small experimental chamber in one cycle. With this particular pump, we have obtained a magnetic field of 22.3 kilogauss.

The 99.6% niobium flux pump has been operated as a "magnetic vacuum pump"; i.e., magnetic flux has been pumped out of the experimental chamber. This was accomplished by changing the sequence of opening and closing the valves. The z component of the field was pumped down from approximately 600 milligauss (the Earth's field in the laboratory) to 0.005 milligauss. It was observed that as the magnetometer probe was moved

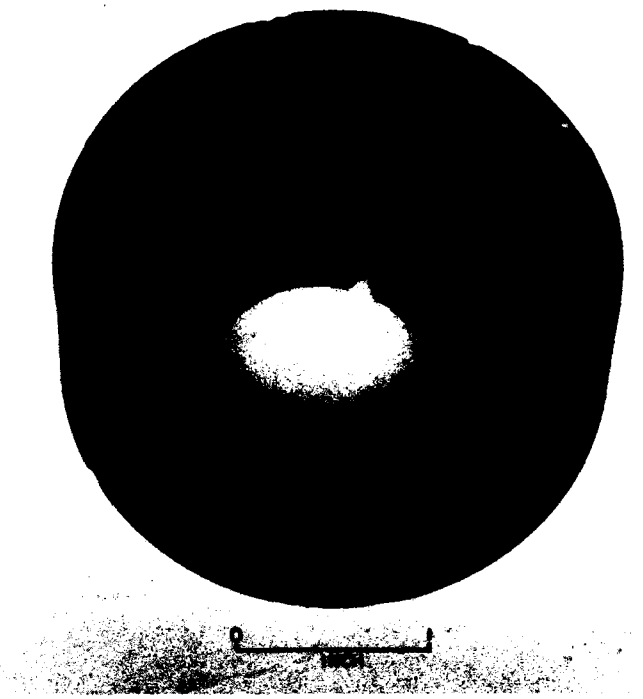


Fig. 3. Single stage Nb_3Sn flux pump

along the z axis, a negative value (opposite direction to that of the Earth's field) of the z component of the field was measured at certain positions of the probe. This difficulty is the result of trapping, in the body of the niobium, small magnetic fields produced by the currents in the heater wires.

It was also observed that the small field in the experimental chamber was not well shielded from fluctuations of the external field. This is the result of the small length-to-diameter ratio of the chamber.

A new pair of coils made of 20 mil Nb_3Zr has now been wound. These coils are 10.0 cm long and 1.50 cm diameter; the niobium piston is 10.0 cm long, 1.45 cm diameter. This system has been operated as a magnetic vacuum pump, and a plot of the magnetic field in the experimental chamber vs cycles of operation is shown in Fig. 4. The length-to-diameter ratio of the chamber is much larger in this case, and the resulting variation of the field at the center of the coil is correspondingly less than in the case of the niobium block. Preliminary tests indicate that this region of small magnetic flux has very good time stability and that the region is well shielded from external fluctuations of the magnetic field. A more detailed report of these experiments will be given at a later date.

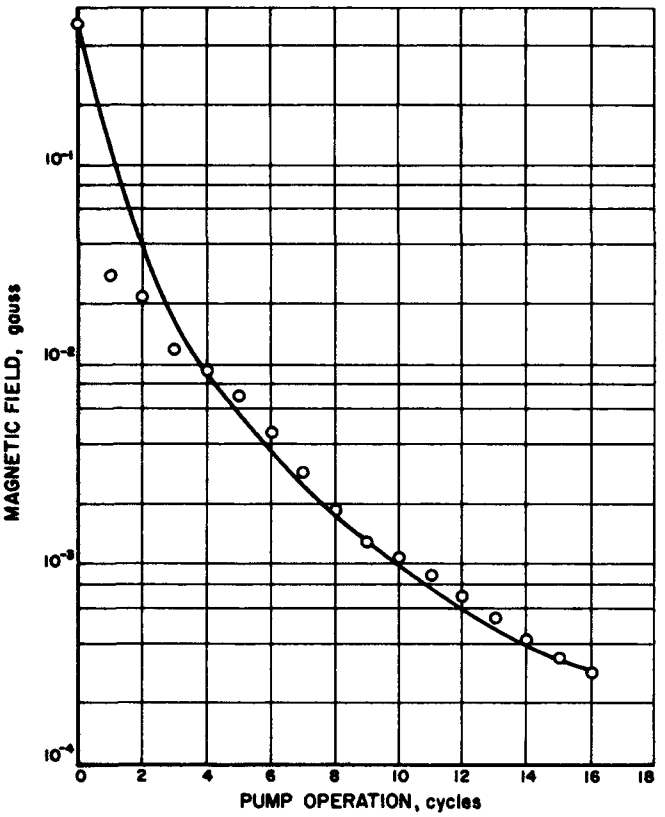


Fig. 4. Plot of magnetic field vs number of cycles that Nb_3Zr magnetic vacuum pump has been operated

B. Thermal Radiation Loss from Vortex Tubes

H. J. Stumpf

It has been shown (Ref. 2) that the ultimate performance potential of gaseous fission reactors is determined by the energy deposition rate in the solid members of the reactor. The physical processes which contribute are: direct nuclear radiation from the fission reactions, fission fragments which reach the tube wall with some residual kinetic energy, and thermal radiation from the hot gas mixture in the cavities. Estimates of the heat load due to nuclear radiation and fission fragments (Ref. 3) are available and should be adequate for preliminary analyses.

In earlier studies the thermal radiation from the gas to the tube walls was computed by assuming that the temperature of the gas is proportional to its enthalpy and that the gas mixture is transparent and radiates to the wall at its central or maximum temperature. These assumptions are quite conservative and handicap the system performance unnecessarily.

A recent treatment (Ref. 4) of this problem has shown that a more reasonable temperature-enthalpy relationship for the gas is

$$T(h) \sim h^{2/3} \tag{1}$$

In addition, the thermal radiation heat load was bracketed by considering the limiting cases of an opaque and a transparent gas.

It is apparent that a detailed treatment of the problem at the present time would be futile due to the absence of experimental data about the separation process and the effect of the fissionable species on the radiative characteristics of hydrogen. It is desirable, however, to consider the temperature dependence of the emissivity of hydrogen when estimating the thermal radiation heat load, since the emissivity has a minimum value at 4100°K (Ref. 5).

The thermal radiation heat load will be computed using the following simplifying assumptions:

Assumption 1. The gas in the vortex tube is hydrogen.

Assumption 2. The emissivity of hydrogen as a function of temperature is given by

$$\varepsilon_{\mathrm{I}} = 7.30 \times 10^{-2}\, e^{-1.17 \times 10^{-3} T} \qquad 2000^\circ\mathrm{K} \leq T \leq 4100^\circ\mathrm{K}$$

$$\varepsilon_{\mathrm{II}} = 7.98 \times 10^{-8}\, e^{2.18 \times 10^{-3} T} \qquad 4100^\circ\mathrm{K} \leq T \leq 7500^\circ\mathrm{K}$$

$$\varepsilon_{\mathrm{III}} = 1.00 \qquad T \geq 7500^\circ\mathrm{K}$$

(The data given in Fig. 18 of Ref. 5 has been approximated by the above expressions.)

Assumption 3. The radiation per unit volume from the gas is

$$k(T)\, e_g(T) = k(T)\, \sigma T^4$$

where T is the gas temperature, σ is the Stefan-Boltzmann constant, $e_g(T)$ is the emissive power of gas, and $k(T)$ is the absorption coefficient.

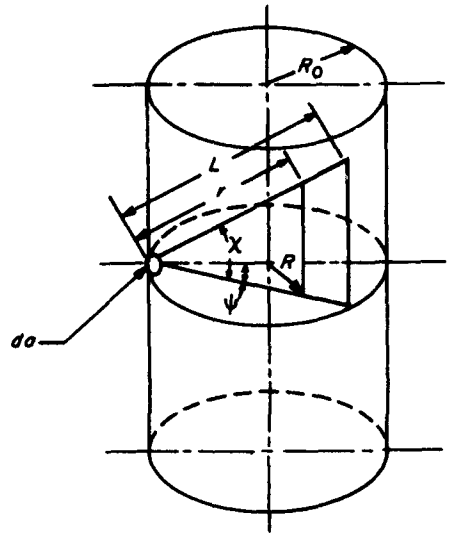


Fig. 5. Vortex tube coordinate system

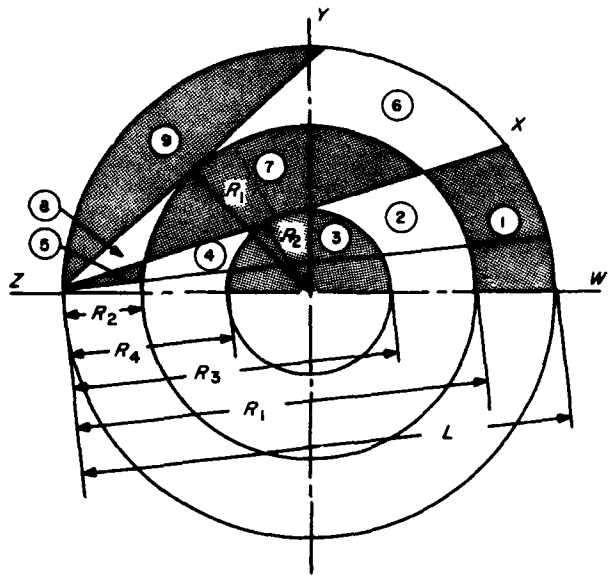


Fig. 6. Regions of integration

Assumption 4. The tube wall acts like a black body.

Assumption 5. The tube is long, and end effects are negligible (Fig. 5).

Assumption 6. The temperature distribution of the gas is

$$T(R) = T_s\left[I^2 - (I^2 - 1)\left(\frac{R}{R_0}\right)^2\right]$$

where T_s is the tube wall temperature, R_0 is the tube radius, $I^2 = T_c/T_s$, and T_c is the gas temperature at the center of the tube ($R = 0$).

Assumption 7. The radiation flux in any given direction is attenuated according to a simple exponential law

$$q(r) = q(0)\exp\left[-\int_0^r k(r')\,dr'\right] = q(0)\,e^{-\tau(r)}$$

where $q(0)$ is the thermal radiative flux at $r = 0$, $q(r)$ is the thermal radiative flux at r, and $\tau(r)$ is the optical thickness.

In general the net thermal radiative flux at dA is given by

$$Q = \frac{1}{\pi}\int_v k(T)\,e_g(T)\,\frac{\cos x\cos\psi}{r^2}\,e^{-\tau(r)}\,dv + \frac{e_w}{\pi}\int_s \frac{\cos x\cos\psi}{L^2}\,e^{-\tau(L)}\,ds - e_w \qquad (2)$$

where

$$e_w = \sigma T_s^4 \qquad L = \frac{2R_0\cos\psi}{\cos x}$$

Since the mathematical expression for the emissivity of the gas is different in each of the three temperature regimes specified in Assumption 2, the volume and surface integrals in Eq. (2) are composed of the contributions from the various regions shown in Fig. 6. In this figure, R_1 is the radius at which

$$T = 4100°\text{K} = R_0\left(\frac{I^2 - 2.05}{I^2 - 1}\right)^{1/2}$$

R_2 is the radius at which

$$T = 7500°\text{K} = R_0\left(\frac{I^2 - 3.75}{I^2 - 1}\right)^{1/2}$$

$$\mathcal{R}_1 = \frac{R_0\cos\psi + (R_1^2 - R_0^2\sin^2\psi)^{1/2}}{\cos x}$$

$$\mathcal{R}_2 = \frac{R_0\cos\psi - (R_1^2 - R_0^2\sin^2\psi)^{1/2}}{\cos x}$$

$$\mathcal{R}_3 = \frac{R_0\cos\psi + (R_2^2 - R_0^2\sin^2\psi)^{1/2}}{\cos x}$$

$$\mathcal{R}_4 = \frac{R_0\cos\psi - (R_2^2 - R_0^2\sin^2\psi)^{1/2}}{\cos x}$$

Since $R_1 = 0$ when $T_c \leq 4100°\text{K}$ and $R_2 = 0$ when $T \leq 7500°\text{K}$, the integrations are considerably simplified for the lower gas temperatures. The most general expression for Q is

$$Q = \frac{4}{\pi}\int_0^{\pi/2} dx \int_0^{\alpha_2} d\psi \cos^2 x \cos\psi \left\{ \exp\left(-\int_0^{\mathcal{R}_1} k_I(r)\,dr - \int_{\mathcal{R}_1}^{\mathcal{R}_2} k_{II}(r)\,dr - k_{III}(\mathcal{R}_3 - \mathcal{R}_4) - \int_{\mathcal{R}_4}^{\mathcal{R}_3} k_I(r)\,dr \right)\left[\int_{\mathcal{R}_3}^{L} k_I(r)\, e_g(r) \exp\left(-\int_{\mathcal{R}_3}^{r} k_I(r')\,dr' \right) dr \right.\right.$$

$$\left. + e_W \exp\left(-\int_{\mathcal{R}_3}^{L} k_I(r)\,dr \right)\right] + \exp\left(-\int_0^{\mathcal{R}_1} k_I(r)\,dr - \int_{\mathcal{R}_1}^{\mathcal{R}_2} k_{II}(r)\,dr \right)\left[\exp\left(-k_{III}(\mathcal{R}_3 - \mathcal{R}_4) \right) \int_{\mathcal{R}_4}^{\mathcal{R}_3} k_{II}(r)\, e_g(r) \exp\left(-\int_{\mathcal{R}_2}^{r} k_{II}(r')\,dr' \right) dr \right.$$

$$\left. + k_{III} \int_{\mathcal{R}_4}^{\mathcal{R}_3} e_g(r) \exp\left(-k_{III}(r - \mathcal{R}_4) \right) dr \right] + \exp\left(-\int_0^{\mathcal{R}_1} k_I(r)\,dr \right) \int_{\mathcal{R}_2}^{\mathcal{R}_1} k_{II}(r)\, e_g(r) \exp\left(-\int_{\mathcal{R}_2}^{r} k_{II}(r')\,dr' \right) dr + \int_0^{\mathcal{R}_2} k_I(r)\, e_g(r) \exp\left(-\int_0^{r} k_I(r')\,dr' \right) dr \Big\}$$

$$+ \frac{4}{\pi}\int_0^{\pi/2} dx \int_{\alpha_2}^{\alpha_1} d\psi \cos^2 x \cos\psi \left\{ \exp\left(-\int_0^{\mathcal{R}_1} k_I(r)\,dr - \int_{\mathcal{R}_1}^{\mathcal{R}_2} k_{II}(r)\,dr \right)\left[\int_{\mathcal{R}_3}^{L} k_I(r)\, e_g(r) \exp\left(-\int_{\mathcal{R}_3}^{r} k_I(r')\,dr' \right) dr + e_W \exp\left(-\int_{\mathcal{R}_3}^{L} k_I(r)\,dr \right)\right]\right.$$

$$\left. + \exp\left(-\int_0^{\mathcal{R}_2} k_I(r)\,dr \right) \int_{\mathcal{R}_2}^{\mathcal{R}_1} k_{II}(r)\, e_g(r) \exp\left(-\int_{\mathcal{R}_2}^{r} k_{II}(r')\,dr' \right) dr + \int_0^{\mathcal{R}_2} k_I(r)\, e_g(r) \exp\left(-\int_0^{r} k_I(r')\,dr' \right) dr \right\}$$

$$+ \frac{4}{\pi}\int_0^{\pi/2} dx \int_{\alpha_1}^{\pi/2} d\psi \cos^2 x \cos\psi \left\{ \int_0^{L} k_I(r)\, e_g(r) \exp\left(-\int_0^{r} k_I(r')\,dr' \right) dr + e_W \exp\left(-\int_0^{L} k_I(r)\,dr \right) \right\} - e_W$$

where $\alpha_1 = \sin^{-1}(R_1/R_0)$ and $\alpha_2 = \sin^{-1}(R_2/R_0)$.

The integrals are quite complex and will be evaluated by means of a computer.

C. A Simple Derivation of the Phonon Dispersion Relation in Metals

O. von Roos

In continuation of work done on correlation effects in an electron gas (*RS 36-13*), the electron-phonon interaction has been taken into account. Although in this report we will stay entirely within the framework of the random phase approximation and will therefore not obtain anything new, it is felt that an exceedingly simple derivation of the Bardeen-Pines result is of interest. Furthermore, the method to be employed lends itself to a straightforward extension which includes correlation effects. Attempts to include correlations have so far met with only limited success (Ref. 6). However, we will not dwell on this subject any further here.

In order to derive the phonon dispersion relation, we proceed as follows. If H is the full Hamiltonian in second quantization, including electron-electron as well as electron-phonon interaction, we define the Heisenberg operators

$$\rho(\mathbf{K}, \mathbf{K}', t) = \sum_s a_s^+(\mathbf{K}', t)\, a_s(\mathbf{K}, t) \tag{1}$$

where $a_s^+(\mathbf{K}, t)$ is an electron creation operator for spin S and momentum $\hbar\mathbf{K}$, and

$$b^+(\mathbf{Q}, t) \qquad b(\mathbf{Q}, t) \tag{2}$$

are the creation and destruction operators for a longitudinal phonon with wave vector $\mathbf{Q}$. The equations of motion of these three operators are easily obtained with the help of H, and read

$$i\hbar\frac{\partial}{\partial t}\rho(\mathbf{K}, \mathbf{K}', t) = \frac{\hbar^2}{2m}(K'^2 - K^2)\,\rho(\mathbf{K}, \mathbf{K}', t) + \frac{2\pi e^2}{V}\sum_{\substack{\mathbf{p}\neq 0 \\ \mathbf{k}}} p^{-2}$$
$$\times \{\rho(\mathbf{k}, \mathbf{k}+\mathbf{p}, t); [\rho(\mathbf{K}, \mathbf{K}'-\mathbf{p}, t) - \rho(\mathbf{K}+\mathbf{p}, \mathbf{K}', t)]\}$$
$$+ \frac{i}{V}\left(\frac{\hbar N}{2M}\right)^{1/2}\sum_{\mathbf{Q}} Q\,[\omega(Q)]^{-1/2}\,\phi(Q)\,[\rho(\mathbf{K}, \mathbf{K}'-\mathbf{Q}, t)$$
$$- \rho(\mathbf{K}+\mathbf{Q}, \mathbf{K}', t)]\,[b^+(\mathbf{Q}, t) - b(-\mathbf{Q}, t)] \tag{3}$$

$$i\hbar\frac{\partial}{\partial t}b(\mathbf{Q}, t) = -\hbar\omega(Q)\,b(\mathbf{Q}, t) - \frac{i}{V}\left(\frac{\hbar N}{2M}\right)^{1/2} Q\,[\omega(Q)]^{-1/2}$$
$$\times \phi(Q)\sum_{\mathbf{k}}\rho(\mathbf{Q}+\mathbf{k}, \mathbf{k}, t) \tag{4}$$

$$i\hbar\frac{\partial}{\partial t}b^+(\mathbf{Q}, t) = \hbar\omega(Q)\,b^+(Q, t) - \frac{i}{V}\left(\frac{\hbar N}{2M}\right)^{1/2} Q\,[\omega(Q)]^{-1/2}$$
$$\times \phi(Q)\sum_{\mathbf{k}}\rho(\mathbf{k}-\mathbf{Q}, \mathbf{k}, t) \tag{5}$$

In the above equations, V is the quantization volume, N the number of lattice sites, $\omega(Q)$ the unrenormalized phonon frequency, and $\phi(K)$ the Fourier transform of the electron-ion interaction energy, and M of course is the ionic mass. In order to solve the system of Eqs. (3–5) approximately we use the same method as employed in *RS 36-13*; i.e., we put the expectation values in the true ground state equal to

$$\langle 0|\rho(\mathbf{K}, \mathbf{K}', t)|0\rangle = 2S(k_F - K)\,\delta_{\mathbf{K}|\mathbf{K}'} + 2g_0(\mathbf{K}, \mathbf{K}', t)$$
$$\langle 0|b(\mathbf{Q}, t)|0\rangle = b_0(\mathbf{Q}, t)$$
$$\langle 0|b^+(\mathbf{Q}, t)|0\rangle = b_0(\mathbf{Q}, t) \tag{6}$$

where the unknown amplitudes g_0, b_0 and $\bar{b}_0$ are considered to be small in the sense that quadratic terms can be neglected. If we also replace the expectation value

of a product of operators by the product of the expectation values we arrive at a system of matrix equations from Eqs. (3–5) which is equivalent to the RPA. The last approximation is what was called the 0th order Tamm-Dancoff approximation in *RS 36-13*. From the resulting equations it is easy to eliminate the electronic degrees of freedom, i.e., the function g_0. In going to the continuum limit ($V \to \infty$) and making a Fourier transformation in time, one then obtains the following set of homogeneous equations for the amplitudes b_0 and $\bar{b}_0$:

$$[\omega(Q) - \omega]\, b_0(-\mathbf{Q}, \omega) = \frac{N}{16\pi e^2 NV} Q^4 [\phi(Q)]^2 A(\mathbf{Q}, \omega) \times [1 - A(\mathbf{Q}, \omega)]^{-1} [\bar{b}_0(\mathbf{Q}, \omega) - b_0(-\mathbf{Q}, \omega)] \quad (7)$$

$$[\omega(Q) + \omega]\, \bar{b}_0(\mathbf{Q}, \omega) = -\frac{N}{16\pi e^2 MV} Q^4 [\phi(Q)]^2 A(\mathbf{Q}, \omega) \times [1 - A(\mathbf{Q}, \omega)]^{-1} [\bar{b}_0(\mathbf{Q}, \omega) - b_0(-\mathbf{Q}, \omega)] \quad (8)$$

The function A has been defined in *RS 36-13*. Clearly Eqs. (7) and (8) have only a nontrivial solution if the determinant vanishes, which happens when ω satisfies the equation:

$$\omega^2 - [\omega(Q)]^2 = \frac{N}{8\pi e^2 MV} Q^4 [\phi(Q)]^2 \frac{A(\mathbf{Q}, \omega)}{1 - A(\mathbf{Q}, \omega)} \quad (9)$$

This result is just the dispersion relation of Ref. 7.

D. Relativistic Plasma Oscillations

P. B. Burt

The general dispersion relations for oscillations of an infinite, homogeneous plasma in a uniform magnetic field have been derived using the relativistic generalization of the Vlasov equation and Maxwell's equations (Ref. 8). A preliminary calculation of the dispersion relations for a beam of electrons moving parallel to the magnetic field has been performed in order to determine qualitatively the effects of relativistic particle velocities on unstable transverse oscillations.

The method of calculating the dispersion relations, similar to that used by Bernstein (Ref. 9), is well known and will not be given here. For the case considered, if the disturbance in the plasma propagates parallel to the magnetic field, the dispersion relation for transverse waves is

$$\omega^2 = c^2k^2 + \omega_{pe}^2 \frac{\omega + \alpha k \xi_z}{\omega + \alpha(k\xi_z \mp \omega_{ce})} + \frac{\omega_{pi}^2 \omega}{\omega \pm \omega_{ci}} \quad (1)$$

where ω is the frequency of the disturbance, k the wave number, c the velocity of light, and the uniform magnetic field B is in the z direction. In addition

$$\omega_{pi,e}^2 = \frac{4\pi N e^2}{m_{i,e}} \quad (2)$$

$$\omega_{ci,e} = \frac{eB}{cm_{i,e}} \quad (3)$$

$$\xi_z = \frac{v_z}{(1 - v^2/c^2)^{1/2}} \quad (4)$$

$$\alpha = \frac{1}{(1 + \xi^2/c^2)^{1/2}} \quad (5)$$

In the above, e and m are charge and mass, respectively; the subscripts i, e refer to ions or electrons; and N is the density of ions and electrons.

If $k\xi_z > \omega_{ce}$, Eq. (1) can have complex roots for ω. The condition that ω be complex is easily obtained if the upper sign is taken in the last two terms in Eq. (1) and is

$$\omega_{pe} > \frac{(\alpha^2 k^2 \xi_z^2 - c^2 k^2)}{(k\xi_z \omega_{ce})^{1/2}} \quad (6)$$

if, in addition,

$$k^2 \xi_z^2 > c^2 k^2 \quad (7)$$

If the conditions for instability are satisfied, one finds for the imaginary part of ω above the critical density

$$\mathcal{I}(\omega) \sim \frac{\omega_{ce}(k\xi_z - \omega_{ce})\, ck}{\omega_{pe}^2 \left(1 - \dfrac{[k\xi_z - \omega_{ce}]^2}{\omega_{pe}^2}\right)} \quad (8)$$

For magnetic fields $\sim 10^3$ to 10^4 gauss and velocities $0.7c$, one has

$$k > 10\ \text{cm}^{-1} \quad (9)$$

Taking $k = 100\ \text{cm}^{-1}$, Eq. (6) gives for the critical density

$$N_c \approx 3 \cdot 10^{18}\ \text{cm}^{-3} \quad (10)$$

Such a density is certainly unreasonably large for any relativistic plasma available in the laboratory. Consequently, this instability is not likely to occur in practical situations. In any case, the stream instability considered here is of little interest except as an heuristic example. Further work on instabilities of more intrinsic interest is in progress. In addition, within the formulation of the problem, it will be possible to examine Landau damping

of the oscillations and the coupling of transverse and longitudinal modes. These problems are under study and will be given in full in a subsequent report.

E. Covariant Expression of Lorentz and Rotation Tensors

F. B. Estabrook

In a thoroughgoing Minkowskian exposition of the theory of relativity, physics is placed in a $3+1$ geometrical continuum characterized by a metric tensor field $g_{\mu\nu}$, and, as the notation implies, coordinates are relegated to the role of descriptive convenience by insistence on covariant expression. One may then ask what has become of the Lorentz transformation, customarily formulated as relating "observer frames in relative motion"—by which is often meant certain very particular orthogonal rectangular coordinate grids (x, y, z, t systems).

The answer is found in analysis of orthogonal tensor fields in the given $3+1$ Riemmanian manifold. A tensor field $A_{\mu\nu}$ is said to be orthogonal if it satisfies

$$A_{\mu\sigma}\, g^{\sigma\tau}\, A_{\nu\tau} = g_{\mu\nu} \tag{1}$$

A sub-group is that of *proper* orthogonal tensor fields, where $\det|A_{\mu\nu}| = +\det|g_{\mu\nu}|$.

It can be shown that any proper orthogonal field in $3+1$ space can be generated in the following fashion:

$$A_{\mu\nu} = R^{*}_{\mu\sigma}\, R^{\sigma}_{\cdot\nu} \tag{2}$$

where the components of the tensor $R_{\mu\sigma}$ are complex, the asterisk denotes complex conjugation, and where $R_{\mu\sigma}$ is also proper orthogonal but of special form

$$R_{\mu\sigma} = D g_{\mu\sigma} + \underset{\vee}{R}_{\mu\sigma} \tag{3}$$

$$\underset{\vee}{R}_{\mu\sigma} = \Omega_{\mu\sigma} + \frac{i}{2(-g)^{1/2}}\, \epsilon^{\cdot\cdot\alpha\beta}_{\mu\sigma}\, \Omega_{\alpha\beta} \tag{4}$$

$\Omega_{\mu\beta}$ is a *real* antisymmetric tensor; $\epsilon^{\mu\nu\sigma\tau}$ is the permutation symbol, a numerically invariant tensor density. The proper orthogonal tensor $A_{\mu\nu}$ or the antisymmetric tensor $\Omega_{\mu\nu}$ are characterized by two invariants, E and H:

$$-\tfrac{1}{2}\,\Omega_{\mu\nu}\,\Omega^{\mu\nu} = E^2 - H^2 \qquad \frac{1}{2(-g)^{1/2}}\,\epsilon^{\mu\nu\sigma\tau}\,\Omega_{\mu\nu}\,\Omega_{\sigma\tau} = 4EH \tag{5}$$

and the complex scalar D is given by

$$1 - (H + iE)^2 = D^2 \tag{6}$$

If either the invariant E or the invariant H vanish, we arrive at a further subgroup—the group of proper orthogonal tensors having an unmoved 2-flat. That is, for such an $A_{\mu\nu}$ there exists a two-dimensional manifold of vectors λ^{ν} such that

$$'\lambda^{\mu} \equiv A^{\mu}_{\cdot\nu}\, \lambda^{\nu} = \lambda^{\mu} \tag{7}$$

When it is H that vanishes, the unmoved 2-flat is space-like, the 2-flat orthogonal to this is time-like, and the tensor $A^{\mu}_{\cdot\nu}$ operates on any vector λ^{ν} to produce a $'\lambda^{\mu}$ by what we might call "pure" Lorentz transformation; in fact, by suitably orienting orthogonal rectilinear x, y, z, t coordinates, we can see that the present vector transformation is at each point mathematically exactly equivalent to the usual $x, t \rightarrow x', t'$ (y and z unchanged) coordinate transformation equations of Lorentz and Poincare.

When it is E that vanishes, the unmoved 2-flat is time-like (e.g., a z-t plane), and the A^{μ}_{ν} tensor engenders a *true rotation* of vectors λ^{μ}.

We thus have an invariant characterization of both Lorentz tensors and rotation tensors.

A very useful expression for either of these can be given in any coordinate system where $g_{\mu\nu}$ is locally diagonal and equal to $+1, +1, +1, -1$: the most general proper orthogonal tensor having an unmoved 2-flat is

$$A^{\mu}_{\cdot\nu} = \begin{pmatrix} \gamma^2 - \varepsilon^2 - \zeta^2 + (\alpha^2 + \beta^2 - \delta^2)\cos\theta & (\gamma\beta + \delta\varepsilon)(1 - \cos\theta) + \alpha\sin\theta & (\alpha\gamma + \delta\zeta)(1 - \cos\theta) - \beta\sin\theta & (\alpha\varepsilon - \beta\zeta)(1 - \cos\theta) + \delta\sin\theta \\ (\gamma\beta + \delta\varepsilon)(1 - \cos\theta) - \alpha\sin\theta & \beta^2 - \delta^2 - \zeta^2 + (\alpha^2 + \gamma^2 - \varepsilon^2)\cos\theta & (\alpha\beta + \varepsilon\zeta)(1 - \cos\theta) + \gamma\sin\theta & (\gamma\zeta - \alpha\delta)(1 - \cos\theta) + \varepsilon\sin\theta \\ (\alpha\gamma + \delta\zeta)(1 - \cos\theta) + \beta\sin\theta & (\alpha\beta + \varepsilon\zeta)(1 - \cos\theta) - \gamma\sin\theta & (\alpha^2 - \delta^2 - \varepsilon^2) + (\beta^2 + \gamma^2 - \zeta^2)\cos\theta & (\beta\delta - \gamma\varepsilon)(1 - \cos\theta) + \zeta\sin\theta \\ (\beta\zeta - \alpha\varepsilon)(1 - \cos\theta) + \delta\sin\theta & (\alpha\delta - \gamma\zeta)(1 - \cos\theta) + \varepsilon\sin\theta & (\gamma\varepsilon - \beta\delta)(1 - \cos\theta) + \zeta\sin\theta & \alpha^2 + \beta^2 + \gamma^2 - (\delta^2 + \varepsilon^2 + \zeta^2)\cos\theta \end{pmatrix} \tag{8}$$

The parameters of the unmoved 2-flat satisfy the subsidiary conditions

$$\alpha^2 + \beta^2 + \gamma^2 - \delta^2 - \epsilon^2 - \zeta^2 = 1$$
$$\alpha\zeta + \beta\epsilon + \gamma\delta = 0 \qquad (9)$$

The unmoved 2-flat is swept out by points $x^\mu = (x, y, z, t)$ satisfying

$$\alpha y - \beta z + \delta t = 0$$
$$\alpha x - \gamma z - \epsilon t = 0$$
$$\beta x - \gamma y + \zeta t = 0$$
$$\delta x + \epsilon y + \zeta z = 0 \qquad (10)$$

When θ is real the unmoved 2-flat is time-like (pure rotation). When θ is imaginary the unmoved 2-flat is space-like (pure Lorentz transformation). The special case of the latter usually given in texts is $\alpha = \beta = \gamma = 0$, $\delta = iV_x/V, \epsilon = iV_y/V, \zeta = iV_z/V$, $\tan\theta = i(1 - V^2)^{-1/2}$, where V is interpretable as the relative velocity 3-vector of two parallel spatial reference triads—but this is a special case only in special coordinates, in which the unmoved 2-flat is normal to the t axis.

The above covariant characterizations of pure Lorentz or pure rotation tensors are being applied to several problems of relativistic kinematics, in particular to the Fermi-Walker transport expounded by Synge (Ref. 10) and to Thomas Precession.

References

1. Hildebrandt, A. F., Elleman, D. D., Simpkins, R., and Whitmore, F. C., *Journal of Applied Physics*, In print, 1962.
2. Meghreblian, R. V., *Gaseous Fission Reactors for Spacecraft Propulsion*, Technical Report 32-42, Jet Propulsion Laboratory, Pasadena, Calif., July 6, 1960.
3. Stumpf, H. J., *Fission-Fragment Energy Loss From Vortex Tubes*, Technical Report 32-188, Jet Propulsion Laboratory, Pasadena, Calif., March 12, 1962.
4. Meghreblian, R. V., *Thermal Radiation in Gaseous Propulsion Reactors*, Technical Report 32-139, Jet Propulsion Laboratory, Pasadena, Calif., July 24, 1961.
5. Olfe, D., *Equilibrium Emissivity Calculations for a Hydrogen Plasma at Temperatures up to 10,000°K*, California Institute of Technology, Guggenheim Jet Propulsion Center, TR 33, May 1960.
6. Hone, D., *Physical Review*, 120, 1600, 1960; von Roos, O., *Physical Review*, 124, 71, 1961.
7. Bardeen, J., Pines, D., *Physical Review*, 99, 1140, 1955.
8. Clemmow, P. C., and Willson, A. J., *Proceedings of the Royal Society*, London, A237, 117, 1956.
9. Bernstein, I., *Physical Review*, 109:10, 1958.
10. Synge, J. L., *Relativity: The General Theory*, pp 13, 123–132, North-Holland Publishing Co., Amsterdam, 1960.

ENGINEERING MECHANICS DIVISION

XI. Materials Research

A. Solid State Research

P. J. Shlichta

1. Growth and Perfection of Crystals

a. The contact rule in crystallography. A crystal structure may be treated as a space-group array of rigid spheres in contact. According to this approximation, the number of contacts in a given structure is to some extent an index of its binding energy, and therefore worth determining. Unfortunately, most structures have a large number of variable parameters—the (n-1) radius ratios of the n atomic species, the space-group position parameters, and the unit cell axial ratios and/or angles—so that exact calculation of all possible contact modes is impractical. It can be shown, however, that a simple topological relation exists which determines the maximum number of contact modes for any given structure.

A crystal structure, expanded so that none of its atoms are in contact, has a degree of freedom (i.e., an arbitrary variable) corresponding to each of its variable parameters, plus an additional one corresponding to the implicit variable of volume expansion (i.e., a_0/R_1). Let us assume that the structure then contracts until some of the atoms come into contact. Each mode of contact specifies a parameter (or a relationship between parameters) and therefore eliminates one of the degrees of freedom. Hence, for any crystal structure

$$F = P - C + 1$$

where F is the number of degrees of freedom, P the number of variable parameters, and C the number of contact modes. This equation is identical in form to the phase rule and has a comparable variety of applications.

The sodium chloride structure provides a simple illustration. It has only one variable, the anion-cation radius ratio, but there are three principal modes of contact: anion-anion, cation-cation, and anion-cation. According to the *contact rule*, stated above:

1. One mode of contact permits one degree of freedom; e.g., anion-cation contact is possible for any radius ratio between 0.414 and 2.413.
2. Two modes of contact permit no degrees of freedom, so that the structure is completely specified; e.g., simultaneous anion-anion and anion-cation contact fix the anion-cation radius ratio at 2.413.
3. It is impossible to achieve all three modes of contact simultaneously.

These results may be verified by geometric construction.

For a structure with two variable parameters, it is necessary to construct a *contact diagram*, similar in most respects to a phase diagram. The contact diagram shown in Fig. 1 is for the hexagonal AlB_2 structure in Fig. 2. Each area defines the "field" of a given contact mode, each boundary line defines the conditions for two modes

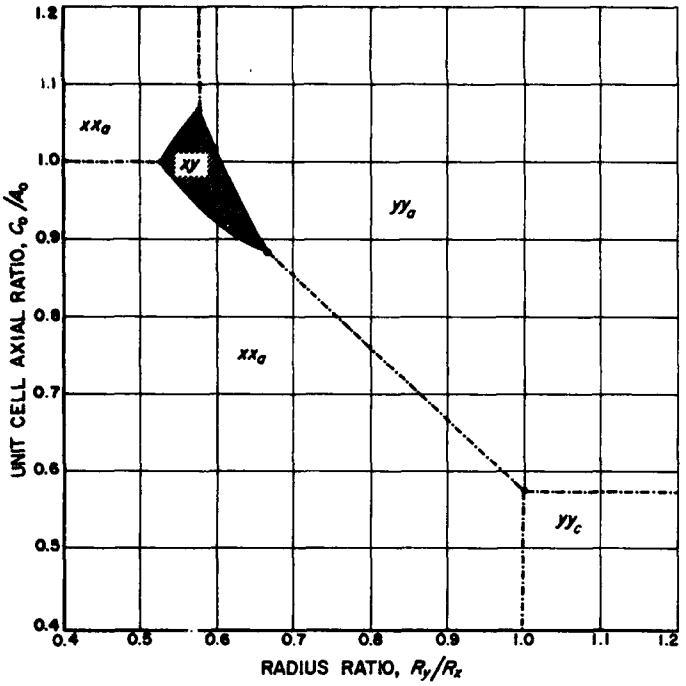


Fig. 1. Contact diagram of the AlB_2 structure

of simultaneous contact, and each point of intersection specifies the conditions for triple contact.

The concept of *connectivity* is useful in interpreting contact diagrams. This is achieved when all the atoms of a structure form a continuous contact-network with no unattached or "floating" moieties left over. Connectivity is usually a necessary requirement of a real structure, though there are apparent exceptions such as lithium iodide. Since connectivity depends qualitatively on the number and types of contact modes, the regions of connectivity in a contact diagram correspond to entire contact-mode fields and/or the boundaries between them; these may be designated by shaded areas and solid boundaries.

Contact diagrams may be used to predict the optimum parameters for a real structure. In the AlB_2 structure, for example, the requirement of connectivity would confine us to the *XY* field and its boundaries. For a metallic structure, for which it may plausibly be assumed that maximum contact is desirable, the triple points at the

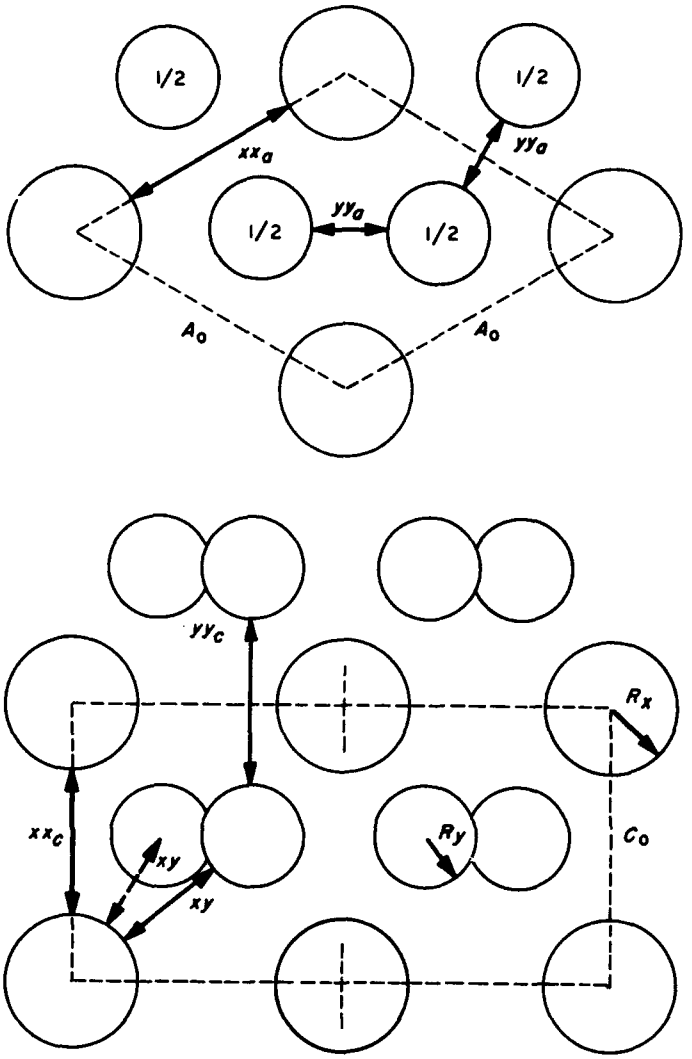


Fig. 2. The AlB_2 structure, showing parameters and contact modes

corners of the *XY* field would seem to be best. On the other hand, for an ionic structure one wishes to maximize anion-cation contacts while minimizing contact of like ions; this would be realized near the center of the *XY* field.

From the examples cited above, it would appear that the contact rule and contact diagrams may be of considerable value in predicting and interpreting crystal structures.

PROPULSION DIVISION

XII. Plasma Propulsion

A. Projects Summary

J. H. Rupe

Recent analyses have shown that the space exploration program proposed for the next decade would be strengthened considerably by the early development of an efficient, high-impulse propulsion system (Refs. 1, 2). These analyses have also shown that of the several schemes that have been suggested, the devices utilizing some form of electrical energy appear promising. In particular, it can be concluded that only the ion engine or one of the several forms of plasma accelerators are compatible with both the performance requirement and a reasonable development schedule.

A critical survey of the national effort in this field was conducted in 1960, and it was clear at that time that those investigations that had been initiated tended to emphasize the development approach. It was also clear that the probability of achieving the ultimate objectives could be improved considerably by supplementing these efforts with investigations of a more fundamental nature. Thus, it was concluded that the Laboratory's long-range supporting research program should include a substantial investigation into the phenomena associated with magneto-plasma accelerators.

A review of the status of knowledge of the phenomena associated with the device has revealed that the critical areas in which research is required are the heat transfer to the electrodes, the process of electrical discharge and plasma acceleration, and the initial ionized gas source. The attainment of high efficiency is critically dependent upon the extent of the heat transfer from the accelerating plasma to the electrode walls and the performance penalty that may be associated with affording the necessary thermal protection to these electrodes. The design and development of an optimum device will require specific quantitative knowledge relative to the interactions of the accelerating plasma with the applied electric and magnetic fields, so that the optimum shaping of these fields can be accomplished. Successful development of a reliable flight propulsion system based on this concept will also require enough knowledge of the initial ionization process to develop a plasma source capable of putting out a steady, uniform, clean plasma flow of predictable and controllable thermal and electrical properties over an extremely long period of operation.

A program of applied research which is intended to contribute knowledge to each of these areas was initiated during the first part of 1961. To date, this effort has been concerned primarily with the construction of a facility and some preliminary experiments with a plasma source

and plasma diagnostics. Some of the preliminary results are summarized in the following sections. It is expected that progress in this field will be reported in subsequent *Summaries*.

B. Plasma Sources

J. A. Gardner

At the beginning of the fiscal year 1962, the Jet Propulsion Laboratory initiated a series of investigations in the field of magnetohydrodynamics (MHD). The first phase of this long-range program was implemented by the procurement of a commercially available plasma jet which would serve, at least on an interim basis, as a plasma source for evaluating diagnostic techniques and materials.

In the following discussion, a brief description of the MHD laboratory is presented along with current preliminary investigations and experimental studies being conducted in the laboratory.

1. Present Status

The laboratory consists of 300 ft² of experimental working area, excluding the requirements of the vacuum and DC power systems. Both high- and low-pressure water distribution systems were installed in order to provide sufficient flow for cooling purposes. Argon and helium high-pressure manifolds and control systems were constructed to provide the working gas in sufficient quantity, and high-pressure nitrogen lines were installed for cooling purposes.

Various primary electrical power circuits 110 v, 220 v, and 440 v were installed for DC power generation, air conditioning system, vacuum system, high-pressure water pump, and supporting electronic equipment.

The basic plasma source is a Model M-4 arc plasma generator purchased in July, 1961 from Plasmadyne Corporation. This plasma unit (Fig. 1) is rated for 120 kw of DC power at efficiencies of energy transfer from the arc to the gas in excess of 70% of input power at atmospheric pressure with argon gas. With co-conical electrode design, the control of the arc is accomplished by gas dynamics which is a result of the vortex and nozzle geometry. This configuration offers a more stabilized arc

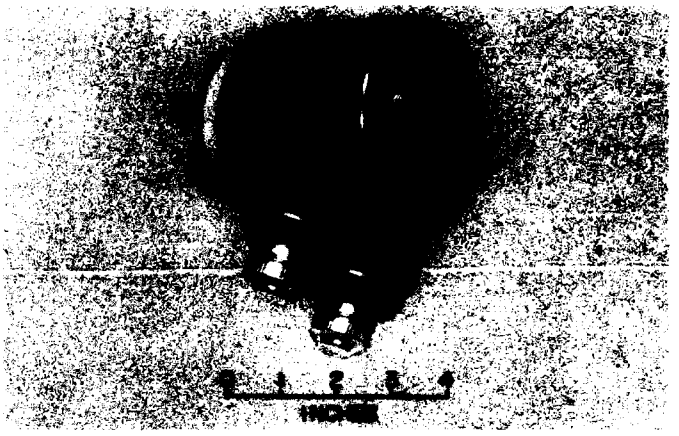


Fig. 1. M4 plasma generator

Fig. 2. DC power supplies

and enhances electrode life. In addition to the M-4 arc plasma generator, an AM-3 modular assembly was purchased in order to study argon plasma flow at supersonic velocities. The AM-3 assembly connects directly on the M-4 generator and is designed for Mach 2.5 in air.

The DC power requirements for the above mentioned systems is supplied by four portable Lincoln welders, Model S 7059 rated at 600 amps, which are capable of operating on 220 v or 440 v. These four welders (Fig. 2) are capable of supplying 3000 amp at a load voltage of 40 v. Each generator has a current selector switch which sets the upper and lower current values but indicates an approximate mid-range value. This current selector indicates from approximately 120 to 750 amp.

Since each generator has its own lower current value, the power system is arranged so that any number of

generators may be operated in a parallel manner, thereby allowing the operator to obtain low to high power settings. Unfortunately, the lowest and highest current setting may not be obtained with the same configuration. Power control is regulated by the variation of the generator excitation fields. The power control console (Fig. 3) is so designed to enable the operator to control field excitation and, therefore, generator output power of all four generators (or any combination such as one and three or two and two, depending on the positions of the generator selector switches) by two separate controls. This enables a distribution of the available power between two individual loads, the plasma source and the MHD accelerator section.

The control console (Fig. 3) purchased from Plasmadyne has facilities for regulation of gas flow and coolant water, with a visual indication of voltage, current, gas pressure, and gas flow, plus the automatic arc starting circuitry and a built-in safety interlock system.

This safety interlock system has been changed from the original Plasmadyne design in order to provide a fail-safe system when utilizing a motor-generator DC power supply. Besides the original water and gas pressure switches, which initiate automatic shutdown when either water or gas pressure falls below a preset minimum value, the safety interlock system was deleted from the original power control and extended to the power control console of the motor generator DC power system. This extension allows automatic shutdown of the 440-v, 3-phase motor and also interruption of the DC generator fields for faster power decay times. The system also requires, after shutdown, a recycle of the initial starting procedures before a restart can be initiated.

Low-pressure operation requirements are provided by the National Research Equipment Corporation, Model 200M, rotary gas ballast pump which has the design capability of maintaining a pressure of approximately 10 mm Hg with a pumping speed of 200 ft^3/min.

During initial phases of experimental study, the instrumentation requirements were satisfied by the recording center located in an adjacent room. This recording center (Fig. 4) is electrically connected by transmission cables terminating in two junction boxes located on opposite walls of the laboratory. Laboratory experimental parameters such as temperatures, pressures, voltage, and current are connected to the junction boxes and then through a patch panel, which can either feed directly to amplifiers, attenuators, and recorders at the MHD facility or can be sent by cable to the central recording facilities of the Jet Propulsion Laboratory.

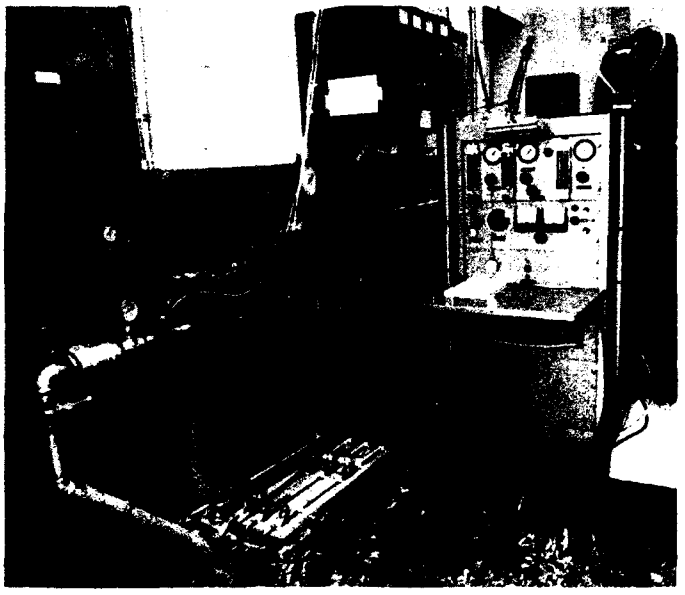


Fig. 3. Control center

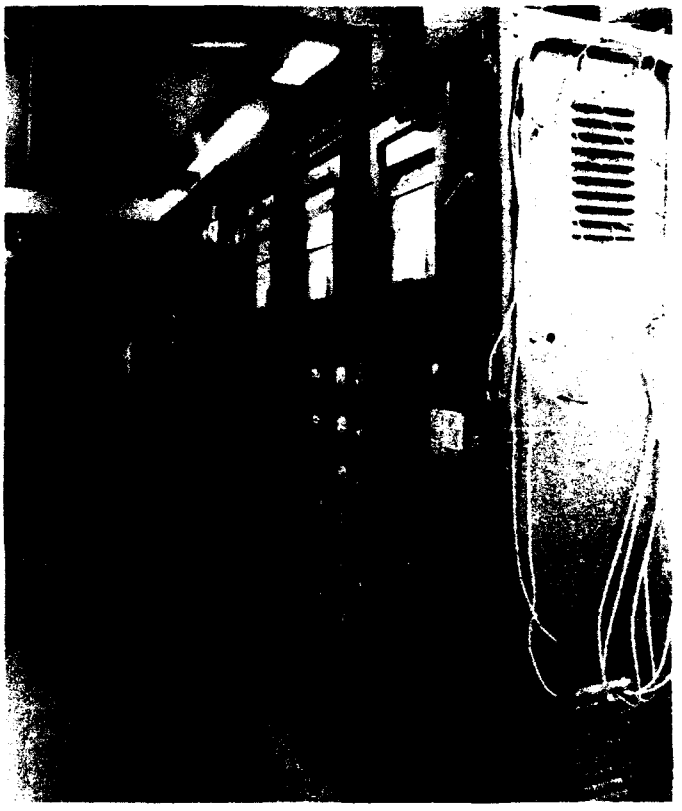


Fig. 4. Instrumentation and recording

2. Experimental Phase

Initial experiments were designed to determine operational characteristics of the M-4 and to provide sufficient information for the determination of enthalpy, argon flow, and input power ranges. This information then represents the basis for the calibration curves (Fig. 5) obtained by determining the energy gained per unit time by the argon gas. Quantitatively this is equal to the difference between the total electrical input power and the power loss to the cooling water. This information, together with a measurement of the argon gas flow rate and inlet conditions, provided the basis for calculating the gas enthalpy. These calibrations were performed for different anode exit bores over power ranges of approximately 10 to 100 kw. Initial anode exit bores were designed for subsonic flow, but later testing will be performed in the supersonic regions up to approximately Mach 3 by utilizing the AM-3 assembly or the JPL designed nozzle (Drawing A 911 2513).

In order to better understand the plasma jet arc characteristics, several different methods for observing the exhausting plasma, at both ambient and 10 to 30 mm Hg pressure regimes, were utilized. These methods included high-speed motion pictures taken at different angles and speeds, study of the visible light fluctuation as observed by a type 929 phototube recorded on an oscilloscope, and basic spectrographic and monochrometer investigations.

The high-speed color motion pictures, with speeds up to 7000 frames/sec, showed that the exhausting plasma flame moves in a whip-like motion when exhausting into ambient conditions but did not exhibit the blob-like characteristics as reported in Plasmadyne's report (PLR-16, Ref. 3). "Arc traces" were not distinguishable in the flow patterns of the M-4 when looking into the ¾-in. exit bore anode with a high-speed camera.

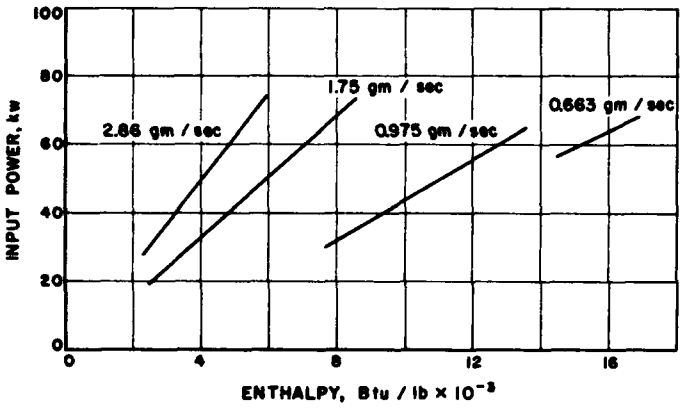


Fig. 5. Typical M4 plasma generator performance characteristics

When the output of an RCA 929 phototube was recorded on an oscilloscope, variations of visible light intensity of approximately 4% were observed at frequencies greater than 1 kc. The wave forms appeared random in nature for the condition at which the M-4 plasma jet was operating.

Preliminary spectroscopic measurements of the visible electromagnetic spectrum utilizing a Cenco grating spectrograph, Model PG3865, indicated that a high continuum level was observable at ambient conditions. However, in the lower-pressure range of 10 to 30 mm Hg, line structure could be resolved in the 4100 to 4400 A region. A Bausch and Lomb (Catalog No. 33-86-45) monochrometer verified the line structure and continuum information produced by the Cenco spectrograph.

By monitoring the intensity of the 4158 A argon line as a function of time using a photomultiplier tube and suitable amplifier circuit, one observes variation of light intensity of approximately 15% at frequencies greater than 5 kc when operating the M-4 at pressure of 10 to 30 mm Hg. Again these wave forms appear nonrepetitive.

Material problems at these temperatures are also being encountered; therefore, time has been utilized in solving the cooling and sealing problems. In the utilization of spectroscopy and microwave diagnostic techniques (*RS 36-14*, Sect. XIIC) it is imperative that the materials of the test section have certain properties. Spectroscopy requires material that exhibits good optical properties, such as transparency to required wavelengths, whereas microwave diagnostics require a material that is transparent to microwave radiation. Of course, both materials will have to withstand high temperatures. Initial experiments have utilized three materials, Lucite, Pyrex, and quartz. It was found that Lucite is unacceptable due to its low melting point and high absorption of ultraviolet radiation; Pyrex and quartz have withstood this required condition with cooling, but they present fabrication problems.

C. Diagnostic Techniques

A. J. Kelly

The interaction of an electromagnetic wave with the unbound electrons present in a body of ionized gas (e.g., a plasma) provides a means for monitoring some of the

plasma's characteristic features, viz., electron density (N) and electron heavy-component collision frequency (ν_c). Significant interaction will occur only when the frequency of the electromagnetic wave train (ω) is generally of the same order of magnitude as the so-called "plasma frequency" (ω_p), i.e., the frequency an electron in the plasma would have if perturbed and then allowed to oscillate freely. The plasma frequency is proportional to the square root of the unbound electron density (Ref. 4),

$$\omega_p = \left(\frac{e^2}{m\varepsilon_0}\right)^{1/2} (N)^{1/2}$$

where e is the charge of the electron, m the electron mass, and ε_0 the dielectric constant of vacuum.

Therefore, for the type of plasma being studied, where the electron density is anticipated to be in the range of $N = 10^{12}$ to $10^{14}/\text{cm}^3$, the probing electromagnetic wave should have a wavelength on the order of 1 cm or less.

A microwave circuit was developed to take advantage of this interaction, thereby permitting measurements to be made of the electron density and collision frequency of the plasma jet (used as a source of ionized gas for the crossed field accelerator program).

This endeavor is one phase of an effort to develop and utilize diagnostic techniques which would allow definitive measurements of the properties of a flowing plasma.

1. Microwave Circuit

An overall view of the circuit is shown in Fig. 6. This circuit has the following three primary functions:

(1) To provide a source of microwave energy at a fixed frequency in the range of 22 to 25 kmc stable to within ± 2 mc.

(2) To provide a means of detecting the microwave energy after it has propagated through the plasma.

(3) To provide for the detection of the reflected signal from the plasma.

By use of this apparatus it is possible to continuously monitor the microwave energy that is transmitted through the flowing plasma and that which is reflected from the plasma, at a given position in the flow stream. From these two measurements it is then possible to calculate the properties of interest (N, ν_c).

A Raytheon 2K-33 klystron operating in the K-band (at a nominal frequency of 24 kmc corresponding to

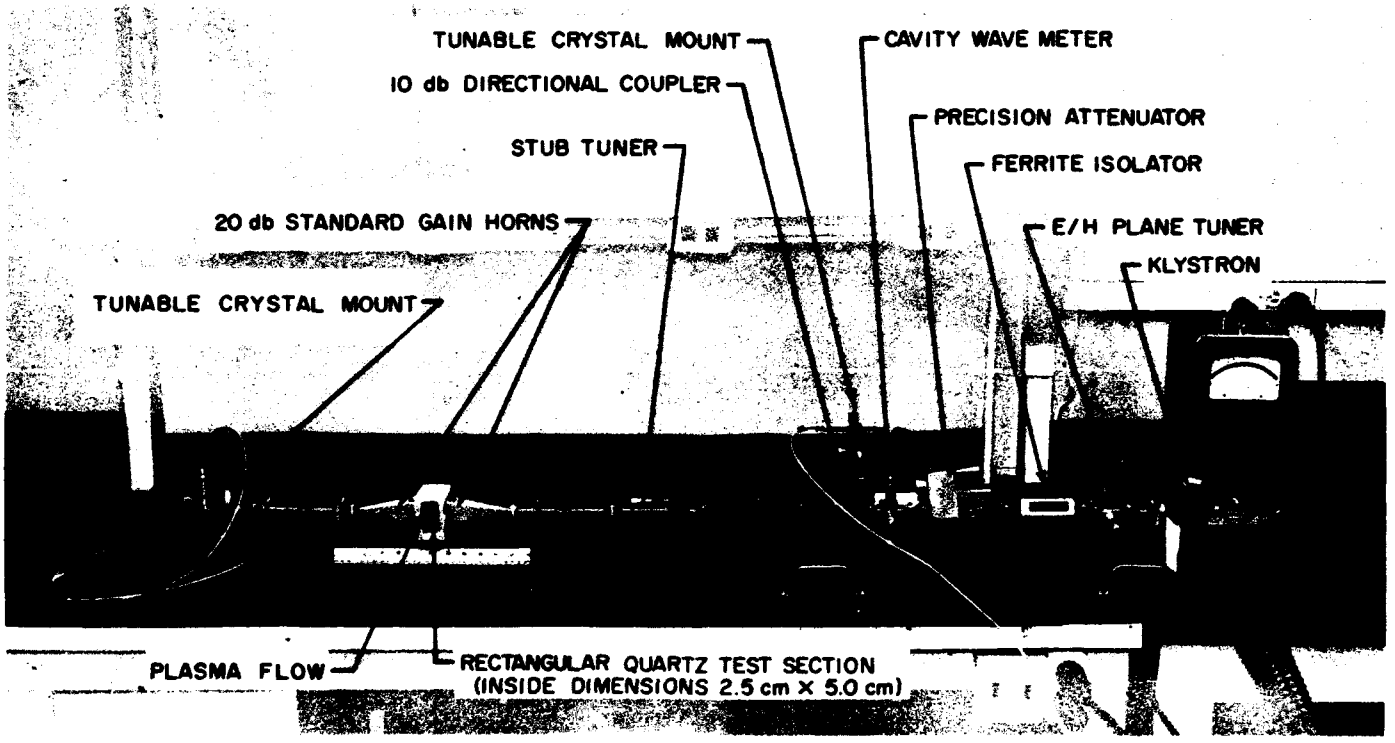


Fig. 6. Microwave probe circuit

~1.25 cm wavelength) supplies up to 40 mw of power to the circuit. The E/H plane tuner immediately adjacent to the klystron provides a matching element for the klystron's resonant cavity and is used to maximize the output signal. The next element in the transmission train, the ferrite isolator, prevents the reflected signal from adversely interfering with the operation of the klystron, which is load-sensitive. A precision attenuator allows the transmitted power to be adjusted to any level between essentially zero to the full output of the klystron. The frequency of the microwave signal is determined to within ~2 mc by the cavity wavemeter crosscoupled to the wave guide directly adjacent to the attenuator.

A 10-db directional coupler, which permits any reflected energy from the plasma to be transmitted up its vertical leg while effectively preventing the transmitted signal from propagating in this portion of waveguide, is contiguous to the crosscoupler. This reflected energy is detected by use of an IN 26 crystal detector in a tunable crystal mount. A stub tuner between the horn and the directional coupler is used to cancel out the small undesirable reflected signals arising from any uncorrected mismatch between the microwave horns and the quartz test section. DeMornay ~ Bonardi 20-db "standard gain" horns are used to transmit and receive the microwave energy to and from the test section. These horns have aperture dimensions of 3.5 × 4.5 cm, so that spatial resolution of approximately 7 cm along the axis of the quartz section is anticipated due to diffraction effects in the plasma.

The signal that is transmitted through the plasma is received by one of the horns and detected again by an IN 26 crystal in a tunable crystal mount.

The above circuitry is the simplest and most straightforward insofar as the number of components and mode of operation is concerned and yet it provides an adequate means for the determination of the properties of interest.

The calculated transmitted and reflected signal amplitudes $|T|$, $|R|$, as a function of the electron density, with the ratio of electron-atom collision frequency as a parameter, are shown in Figs. 7 and 8 for a channel 2.5 cm wide and free space wavelength $\lambda_0 = 1.25$ cm.

These calculations were based upon the theory developed in Ref. 4. This theory assumes that a plane monochromatic wave of electromagnetic energy interacts with a plane homogeneous slab of plasma. A simplified model of the electron heavy-component collision process was incorporated in this analysis, so that it is valid only for those plasmas where electron-neutral collisions are far more predominant than electron-ion collisions. This restricts the utility of this analytical development to plasmas with less than approximately 0.1% ionization.

In order to duplicate these conditions as closely as possible, it is necessary that the sides of the channel used to contain the plasma (to which the microwave horns are attached) be essentially lossless, halfwave plates and therefore not introduce spurious reflected signals. To determine whether quartz, because of its low loss tangent (at room temperature) and thermal expansivity, would

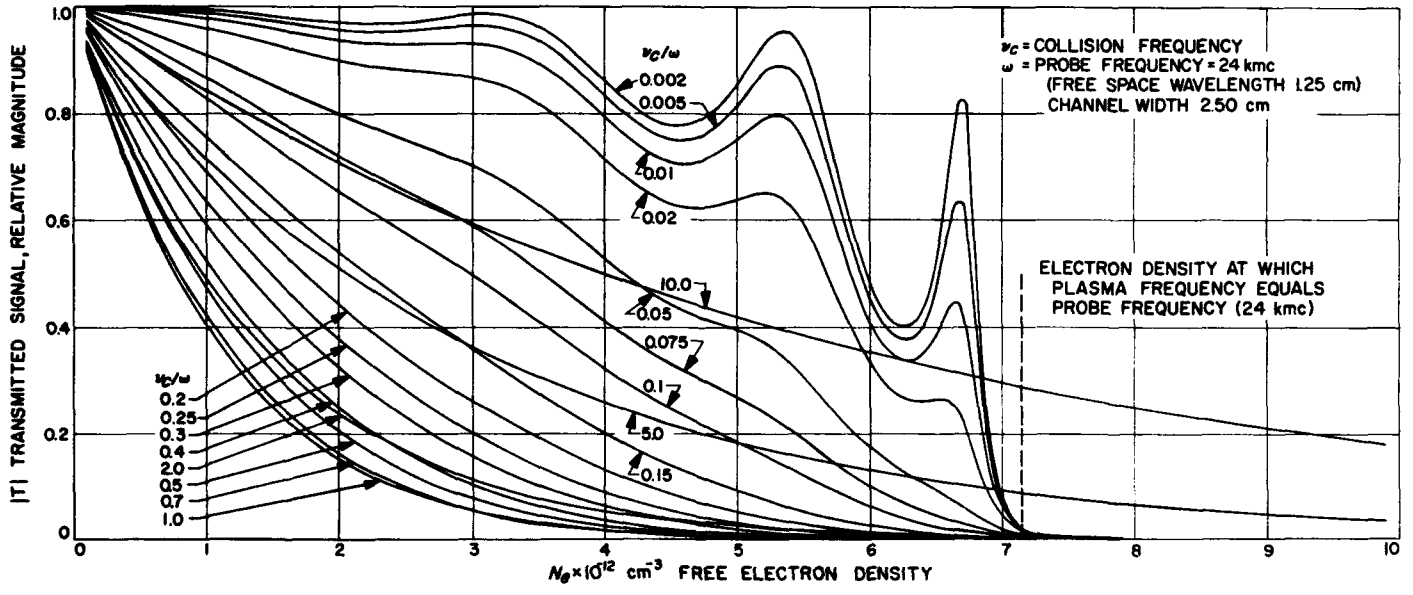


Fig. 7. Transmitted signal magnitude vs electron density for various collision frequencies

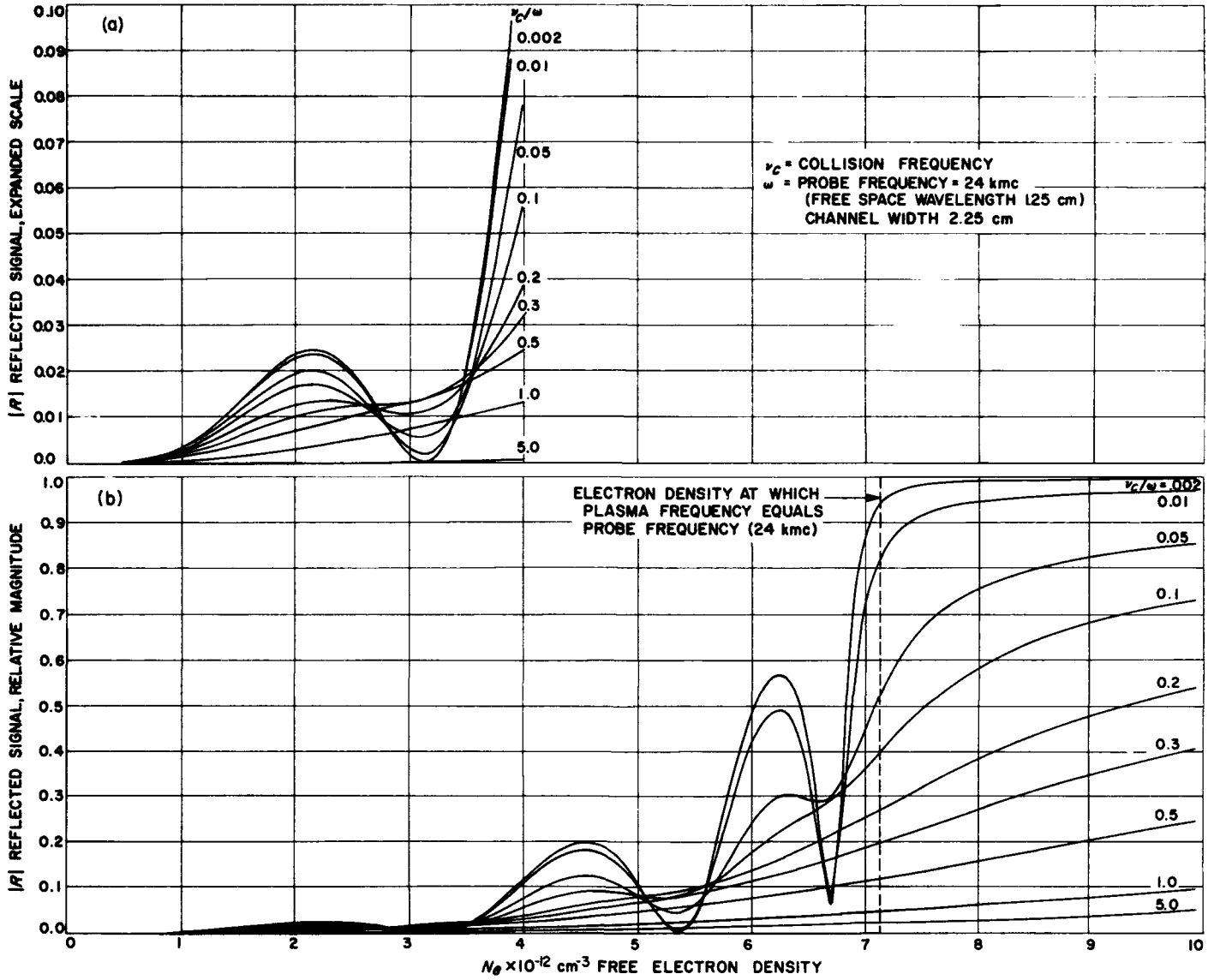


Fig. 8. Reflected signal magnitude vs electron density for various collision frequencies

be a suitable material for the construction of this channel, a rectangular slab about 1.25 cm thick was matched to the transmitting section with a nominal standing wave ratio of 1.04. Thermocouples were emplaced so that the temperature on both sides of the quartz plate could be monitored. One side of the slab was immediately against the aperture of a 20-db horn and the other was tangent to the geometrical extension of the ¾-in. exhaust nozzle of the plasma jet (*RS 36-14*, Sect. XIIB) with the centerline some 9 cm from the exit plane of the jet. The experimental setup is shown in Fig. 9. The crack that appears across one corner of the quartz slab occurred during the first run of the series and was attributed to thermal stresses developed in the plate because of the holes for the thermocouples (junction of thermocouples was within 2 mm of the plasma side) and the method of mounting. Temperatures up to 800°C as measured with the jet-side thermocouples (corresponding to a surface temperature of approximately 900°C) resulted in a negligible change of the reflected signal, i.e., from a nominal of 0.6 mv the reflected signal increased to 0.8 mv. By comparison the total reflected signal generated by covering the horn with an aluminum plate was 140 mv. The 0.2-mv change in signal level then corresponds to approximately 0.2% change in reflectivity due to temperature variation of the quartz plate. It was therefore concluded that insofar as microwave diagnostics were concerned, quartz would be a suitable material for the construction of the channel

Fig. 9. Test arrangement to determine microwave properties of heated quartz

because of its apparently negligible change in dielectric and dimensional properties as a function of temperature (in the range of interest). Currently the quartz test section shown in Fig. 6 is being matched to the microwave system so as to reduce the reflected signal to as low a level as possible.

This quartz section has a rectangular channel for the plasma flow 2.5 cm wide and 5.0 cm high with walls approximately 1 cm thick. Hopefully, a standing wave ratio of less than 1.04 can be achieved (without use of a stub tuner). This degree of matching is necessary to minimize error as any signal reflected from the quartz would add vectorially to that reflected from the body of the plasma, making the separation of these two signals difficult, if not impossible.

2. Preliminary Tests

Preliminary measurements were made with the microwave circuit shown and a 1-in. diameter quartz tube (2-mm wall thickness) through which the exhaust from the plasma jet was directed. Because of the circular geometry of the plasma and the lack of a consistent theoretical description for this configuration, it was not possible to deduce quantitative values for N and ν_c, but it was possible to provide assurance that there was no arc plasma generated or extraneous sources of microwave energy in the frequency region being used which would interfere with these measurements. Strong interaction between the plasma and the microwave signal was observed. This interaction was of a roughly periodic nature at a frequency of about 3000 cps. That is, the transmitted signal was observed to vary from 100% transmission to essentially zero transmission, while at the same time the reflected signal was greatly enhanced. This was attributed to pulsation in the free-electron level of the plasma produced by the arc jet.

As soon as the rectangular quartz section is matched and the plasma flow system completed, further checkout runs will be made.

References

1. Jaffe, L. D., Lucas, J. W., Merrill, O. S., Shafer, J. I., and Spencer, D. F., *Electric Spacecraft for Planetary and Interplanetary Missions,* Technical Memorandum 33-43, Jet Propulsion Laboratory, Pasadena, Calif., March 15, 1961 (Confidential).

2. Shafer, J. I. (Editor), *Utilization of Electric Propulsion in Spacecraft,* Technical Memorandum 33-21, Jet Propulsion Laboratory, Pasadena, Calif., October 1, 1960 (Confidential).

3. Von Jaskowsky, W. F., *Luminosity Fluctuations of the Plasma Jet,* Report No. PLR-16, Plasmadyne Corporation, Santa Ana, Calif., January 2, 1959.

4. Jahn, R. G., *Interaction of Electromagnetic Waves with Slightly Ionized Gases,* Part I, "Uniform Media," Technical Note No. 2, California Institute of Technology, August 1960.

XIII. Liquid Propellant Propulsion

A. Combustion and Injection

G. I. Jaivin

1. Atomization and Injection Hydraulics

As part of the continuing effort to categorize the hydrodynamic parameters of liquid sheets, four sets of photographs were taken to illustrate the visual characteristics of certain reference sheet configurations. The purpose of these photos (Figs. 1 to 4) is to illustrate the differences which exist in the several types of sheets currently under study. Because these sheets have markedly different fluid dynamic properties this work should serve to indicate, in a qualitative sense, the significant role played by such properties in the formation, stability, and ultimate breakup of such liquid sheets.

Four different flow conditions were evaluated. In three of the cases examined the liquid sheets were formed by the 180-deg impingement of two identical, opposed jets having known dynamic properties. The sheet formed by an axially symmetric, sharp-edged, annular slit was studied also. Table 1 indicates the pertinent characteristics of each test series. The photographs were made using an arc discharge light source which produced a flash duration of 2 or 3 μsec. Two representative photos from each test series are presented.

It should be noted that when two identical nonturbulent opposed jets having uniform velocity profiles impinge, each to the other appears as a flat frictionless wall much like an ideal boundary in a potential flow problem. Each half of the sheet can be thought of as a mirror image of its counterpart. The formation of the sheet is analogous, therefore, to the axially symmetric potential flow case of a circular jet perpendicularly striking a flat boundary (Ref. 1).

Viscid effects in the jets and resultant sheet must, of course, be considered in the actual case. The losses incurred in the turning of the flow near the impingement point should be quite small, however, since the flow does not experience marked velocity gradients except in the immediate vicinity of the sheet centerline. If it is assumed that such losses in the impingement zone are negligible, it is evident from conservation of mass and energy considerations that, to a first approximation, sheet thickness at a given radius is solely a function of jet diameter. If it is further assumed that the departure from velocity profile uniformity can be neglected in the case of fully developed turbulent velocity profile jets, a rough estimate of sheet thickness can be made for this case as well.

It is of interest to examine the data from Test Series TS 1 and TS 2 in some detail (Figs. 1, 2). Note that in TS 2 a sharp-edged orifice having a diameter of 0.1002 in. was used. Assuming a jet contraction ratio of 0.64 due to *vena contracta* effects, the jet diameter in TS 2 was approximately 0.080 in. as compared to a jet diameter of 0.0986 in. in TS 1. For these jet diameters, the sheet in TS 1 is approximately 1.52 times thicker than that in TS 2 at the same sheet radius. Based upon the observed radii at which the sheets disintegrated in these two cases, the sheet thickness at the breakup zone in TS 1 was approximately 7 times thicker than in TS 2. The appearance of the sheets produced in the two series is noticeably different as evidenced by Figs. 1 and 2. In Fig. 1 the entire surface of the sheet appears ruffled even at low

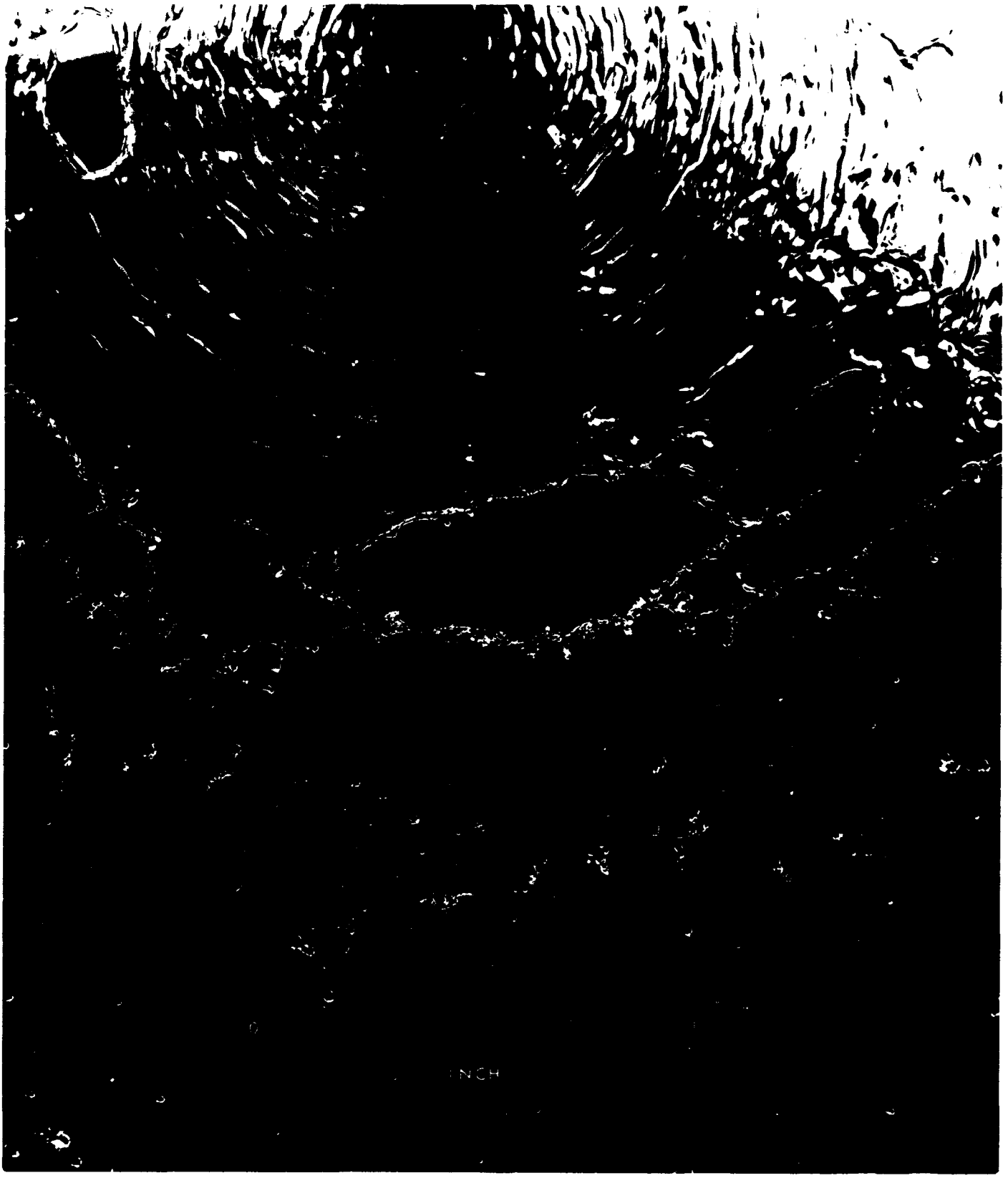


Fig. 1a. Liquid sheet formed by 180-deg impingement of two jets having fully developed turbulent velocity profiles. Sheet velocity $\simeq$ 16 ft/sec; flow rate = 0.106 lb/sec

Fig. 1b. Liquid sheet formed by 180-deg impingement of two jets having fully developed turbulent velocity profiles. Sheet velocity $\simeq$ 59 ft/sec; flow rate = 0.400 lb/sec

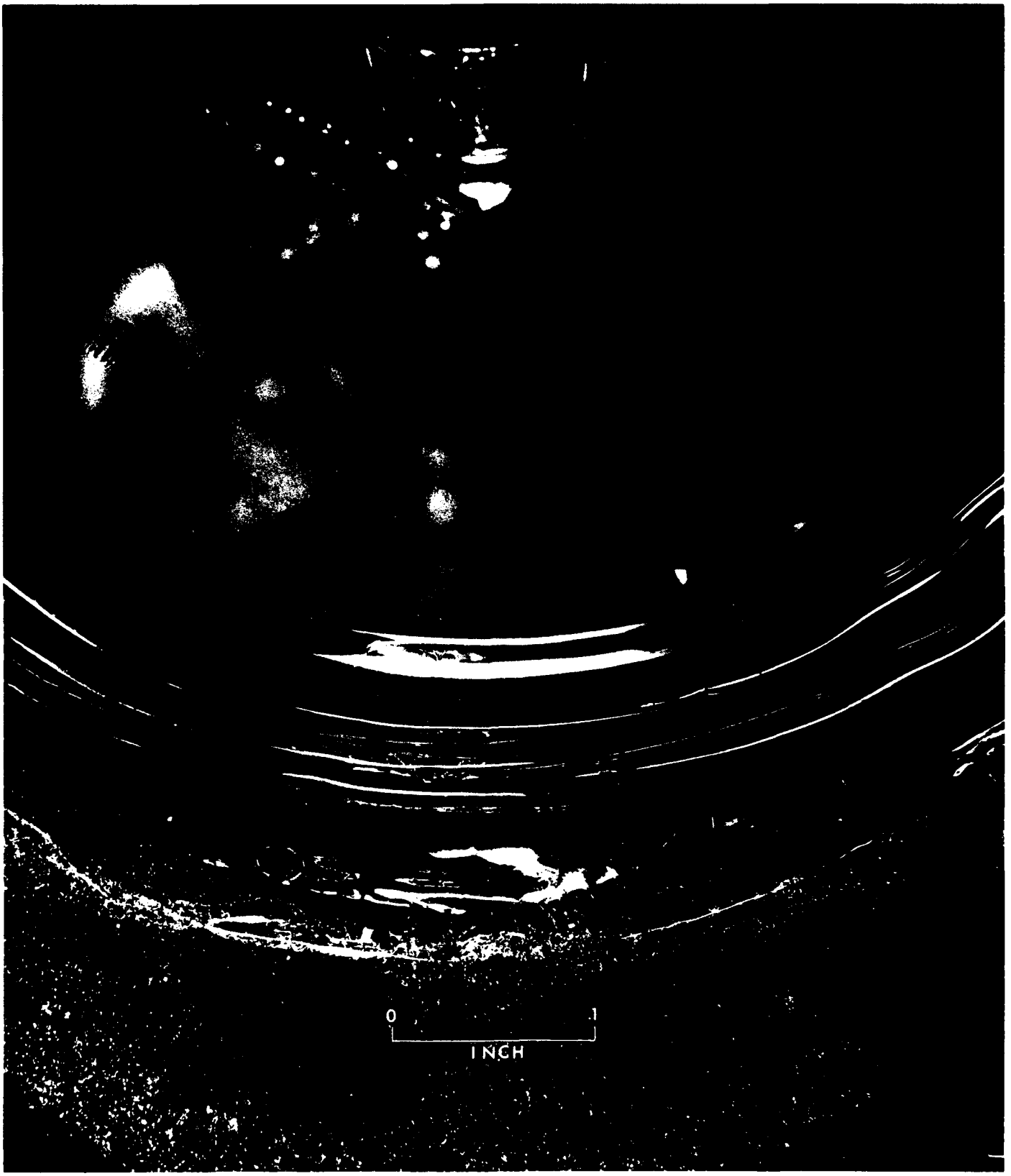


Fig. 2a. Liquid sheet formed by 180-deg impingement of two laminar jets having uniform velocity profiles. Sheet velocity $\simeq$ 39 ft/sec; flow rate = 0.261 lb/sec

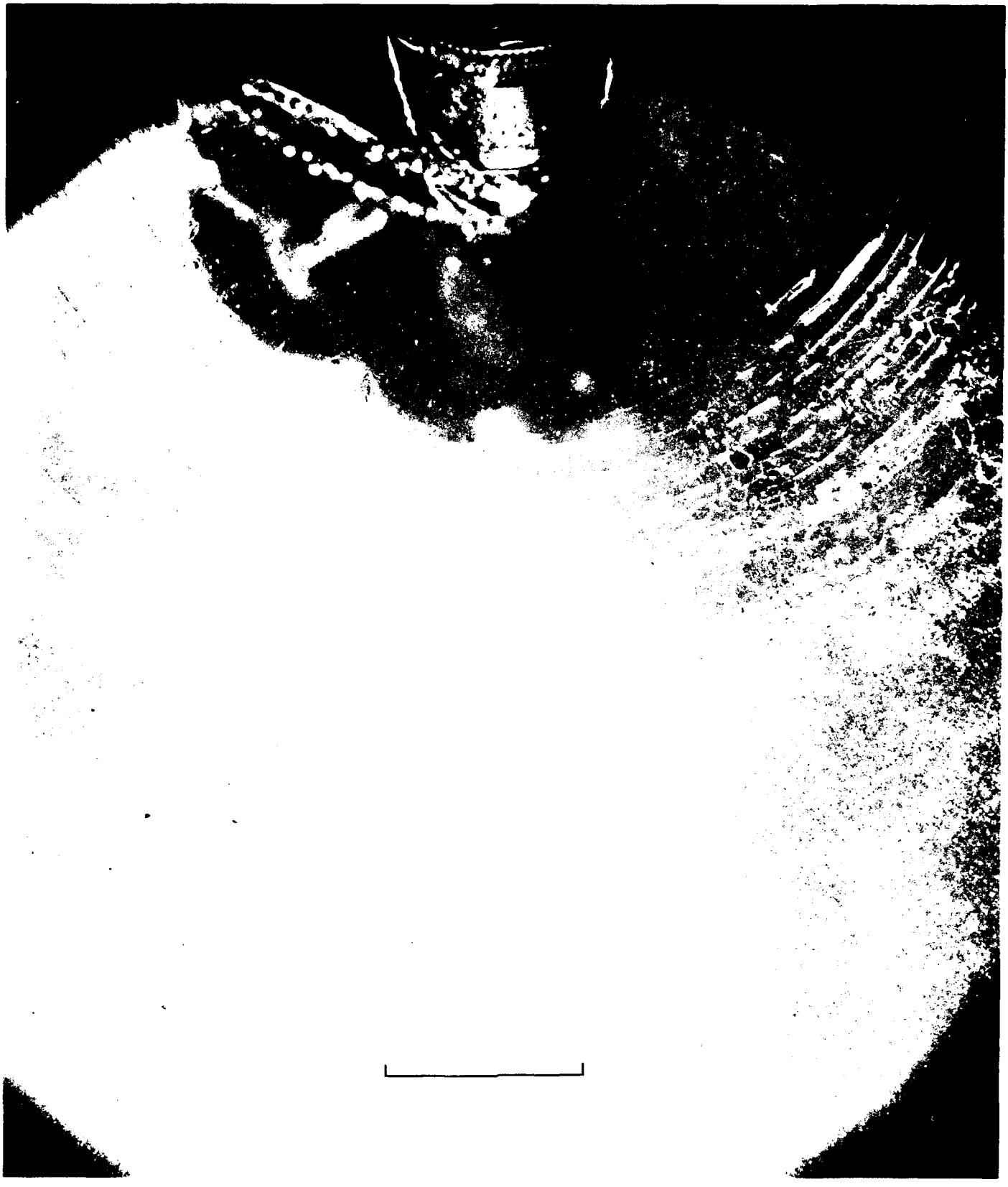


Fig. 2b. Liquid sheet formed by 180-deg impingement of two laminar jets having uniform velocity profiles. Sheet velocity ≃ 102 ft/sec; flow rate = 0.683 lb/sec

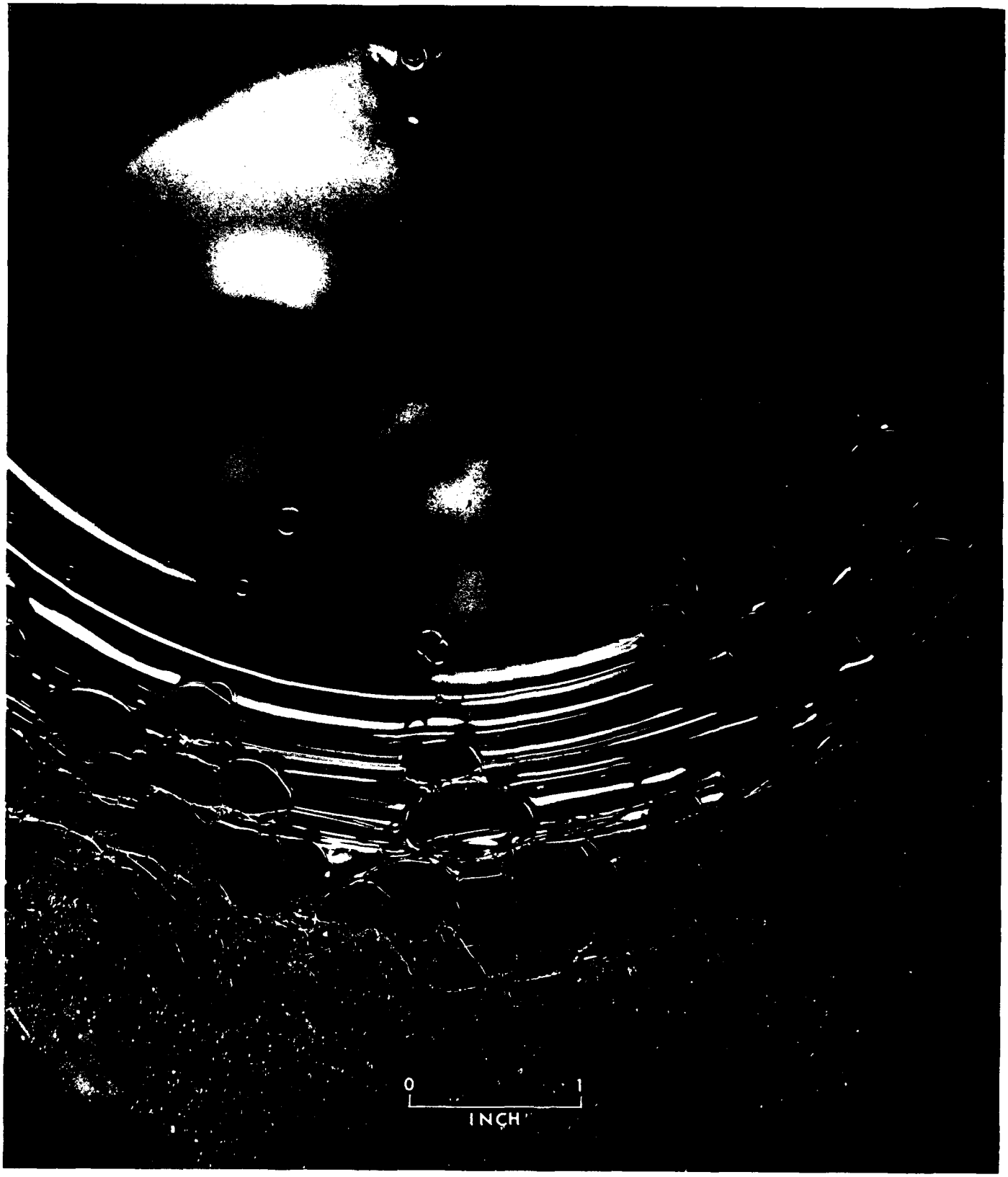


Fig. 3a. Liquid sheet formed by 180-deg impingement of two jets having fully developed laminar velocity profiles. Sheet velocity $\simeq$ 29 ft/sec; flow rate = 0.200 lb/sec

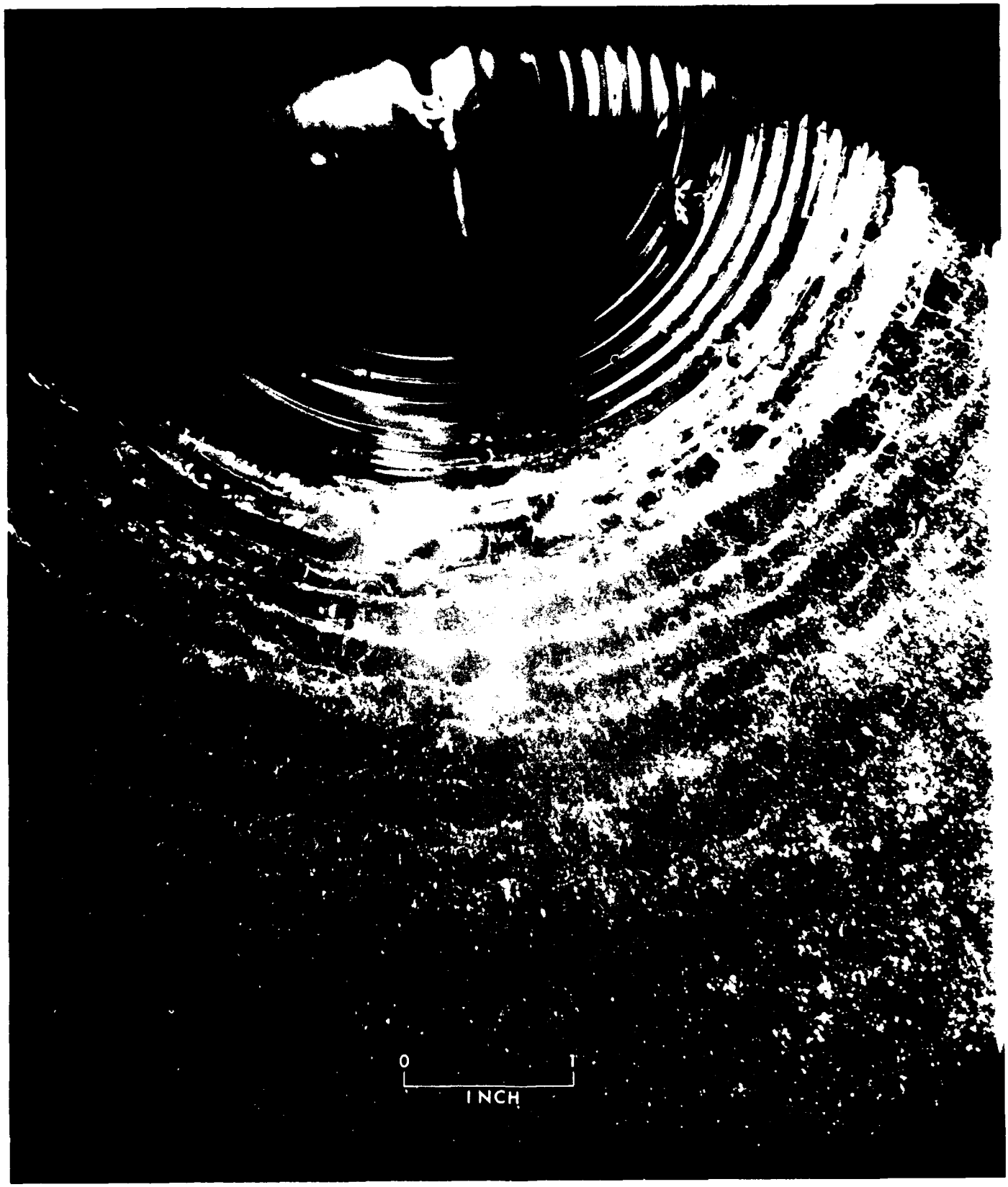


Fig. 3b. Liquid sheet formed by 180-deg impingement of two jets having fully developed laminar velocity profiles. Sheet velocity $\simeq$ 80 ft/sec; flow rate = 0.545 lb-sec

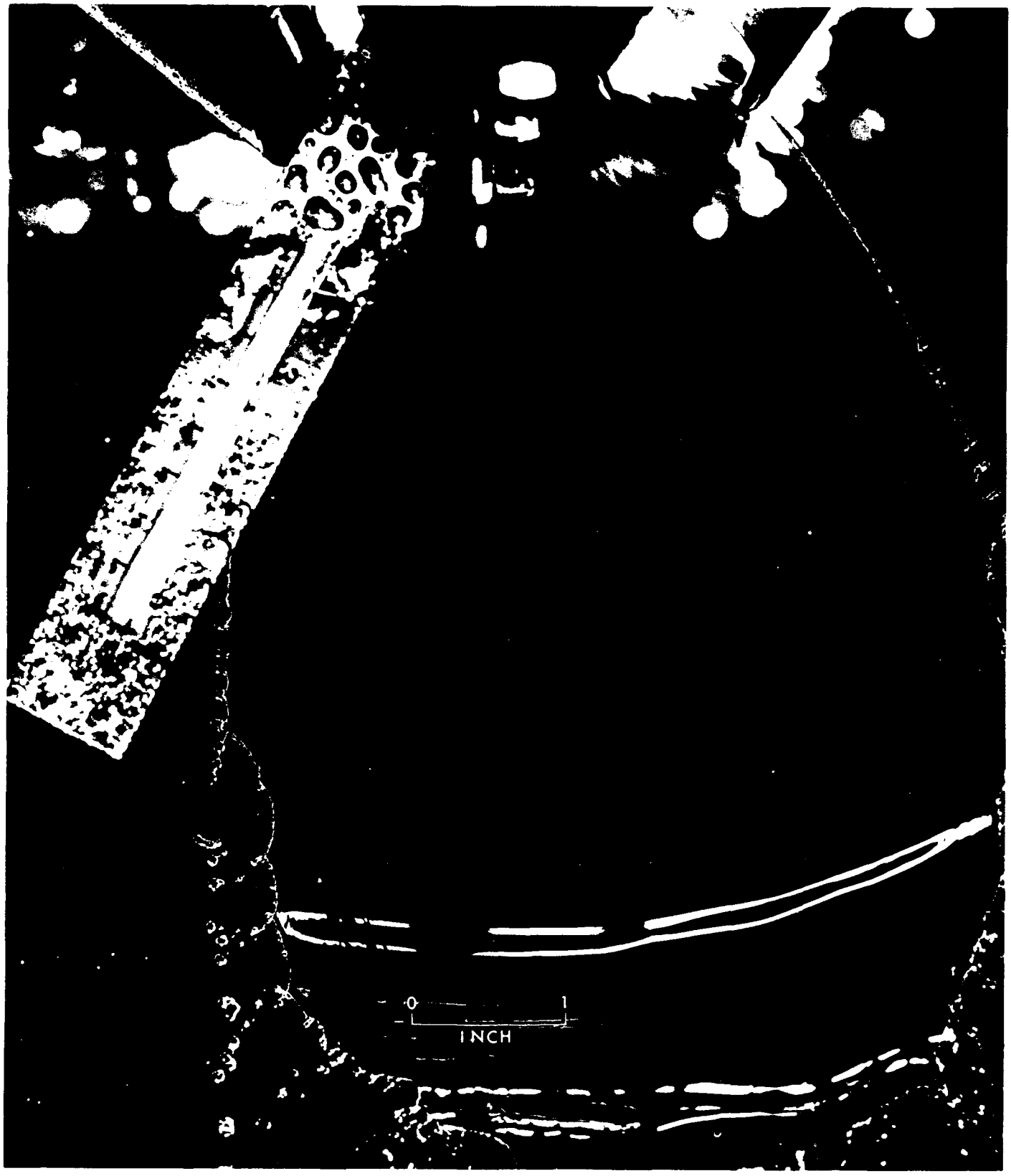


Fig. 4a. Liquid sheet formed by an axially symmetric sharp-edged annular slit.
Sheet velocity $\simeq$ 12 ft/sec; flow rate = 0.295 lb/sec

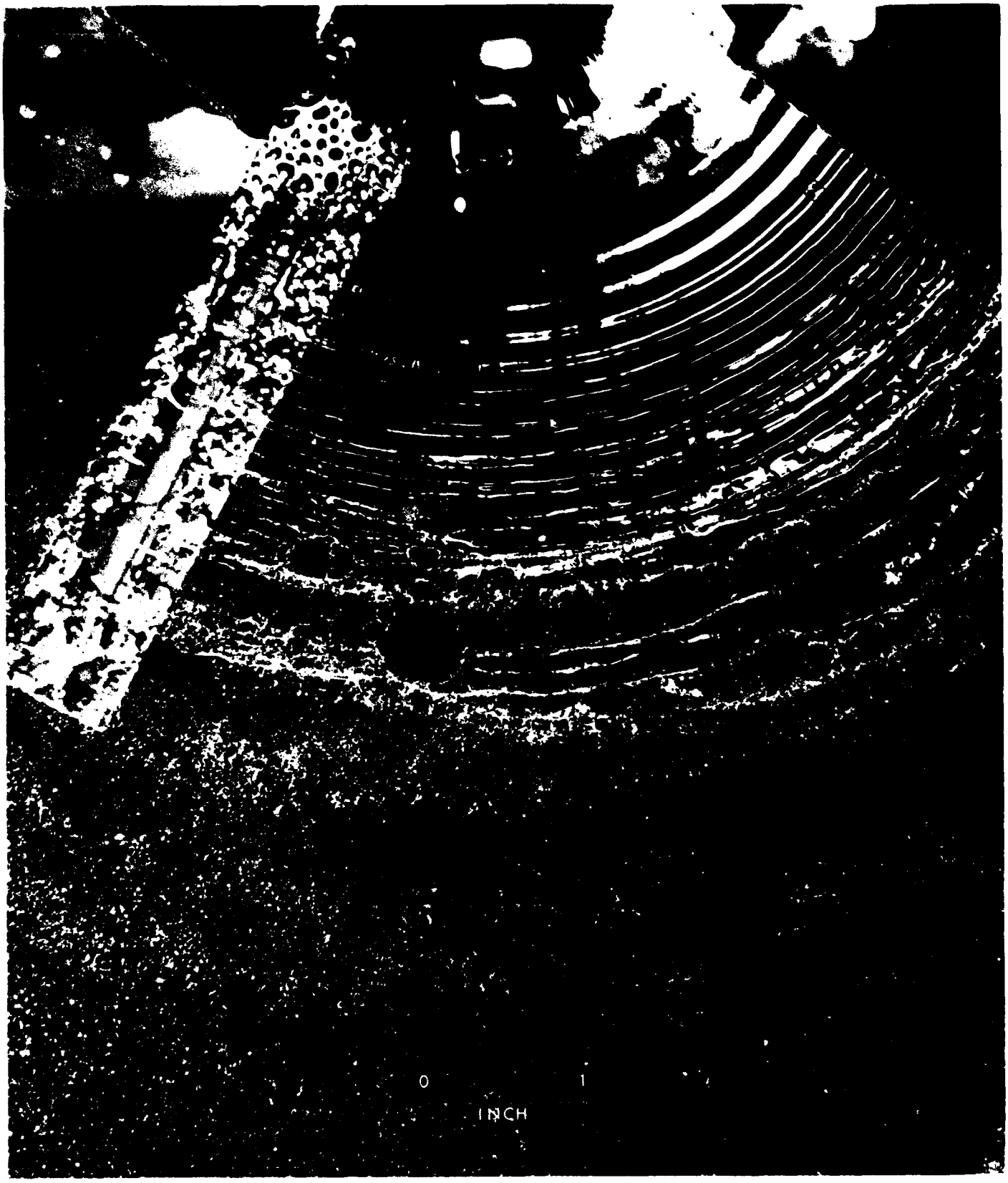


Fig. 4b. Liquid sheet formed by an axially symmetric sharp-edged annular slit. Sheet velocity $\simeq$ 37 ft/sec; flow rate = 0.905 lb/sec

flow velocities. The laminar uniform velocity profile jets produced sheets (Fig. 2) having glassy smooth surfaces which ultimately became wavy, presumably under the influence of aerodynamic forces. The only significant difference between the jets used in TS 1 and TS 2 was the turbulence level of the flow. The velocity profiles are not considered sufficiently dissimilar to produce the noted effects. Hence, it seems clear that certain hydrodynamic parameters such as turbulence, can be important controlling influences on the stability of liquid sheets. Although no quantitative data is currently available, it appears evident from the photos that the mean drop size produced by the turbulent sheet configuration is likely to be larger than that from the others tested. This results in part from the gre heet thickness at breakup for that case.

As previously stated, the liquid sheet resulting from the impingement of two near-uniform velocity profile jets can be considered analogous to the juxtaposition of two similar potential flow solutions. It is not clear to what extent a nonuniform velocity profile modifies the impingement process, but it would seem in such a case the simplified model given above is no longer applicable. It appears likely that the resulting higher sheet centerline stagnation pressure would produce a sheet initially having a velocity profile approaching the degree of nonuniformity of the jets forming it. This hypothesis is to date untested, but it is hoped that future work will determine its validity.

TS 3 utilizes two jets having nonuniform velocity profiles (Fig. 3). It is not possible to accurately estimate sheet thickness for this case. Its visual characteristics as well as its apparent mean droplet size as seen from the photos resemble those of TS 2.

The photos obtained in TS 4 (Fig. 4) illustrate the type of sheet on which prior thickness measurements were made (*RS 36-10*). The sheet tends to be very stable, especially at low flow velocities. Note that in Fig. 4a, for example, the sheet is continuous for more than 7 in. before breakup. This stability is a direct result of the low turbulence level of the flow. The sheet is intercepted by a curved deflector visible in the photographs. The only purpose this device served was to reduce the amount of spray mist in the spray booth in order to improve the picture quality. This sheet configuration was designed to be hydrodynamically similar to that of TS 2. Whether or not this aim has been successfully accomplished cannot be determined from these data alone. The general appearance of the two sets of photos does appear similar, however. The radial striations in the sheet in TS 4 are caused by minute imperfections on the sharp edge of the annular orifice.

Currently work is in progress on the determination of the velocity profile in a liquid sheet. Future work will be directed toward the measurement of the droplet size spectrum produced by sheets such as those discussed herein.

Table 1. Characteristics of liquid sheet test series

Test series	Sheet formed by	Type of flow	Orifice diam., in.	Orifice L/D	Jet Reynolds No.	Distance between orifices, in.	Jet velocity, ft/sec	Test fluid
1	Opposed jets	FDTVPJ[a]	0.0986	210	18,700–70,500	0.789 (8L/D)	15–86	Water
2	Opposed jets	LUVPJ[b]	0.1002	~0	39,900–84,800	1.000 (10L/D)	36–128	Water
3	Opposed jets	FDLVPJ[c]	0.0986	210	710–1160	0.789 (8L/D)	25–85	Glycerol and water mixture (79% glycerol by weight)
4	Annular slit	Laminar	1.000 (slit diam.)	~0	—	0.018 (slit width)	—	Water

[a] Fully developed turbulent velocity profile jets.
[b] Laminar uniform velocity profile jets.
[c] Fully developed laminar velocity profile jets.

Reference

1. Le Clerc, A., *Deflection of a Liquid Jet by a Perpendicular Boundary*, Thesis, Department of Mechanics and Hydraulics, State University of Iowa, Iowa City, 1948.

XIV. Polymer Research

A. Polyoxyalkylenes Terminated with Different Functional Groups: Effect of Catalyst Concentration on the Cyanoethylation of Polyoxyalkylene Glycols

A. J. Havlik and S. H. Kalfayan

Previous reports from this Laboratory (*RS 36-6*, Vol 2; *RS 36-9*, Vol 1) had indicated the formation of solid byproducts during the cyanoethylation of glycols with acrylonitrile. Ethylene, dipropylene, and tetraoxyethylene glycols (EG, DPG, TOEG, respectively) of the low molecular weight class; polyoxyethylene glycols (POEG), Polyglycol E600 and polyoxypropylene glycols (POPG), PPG 1025, of the higher molecular weight class were discussed. The solid byproduct formed with DPG, was shown to be a low molecular weight polyacrylonitrile consisting of six acrylonitrile units per unit of DPG. The solid byproducts from the cyanoethylation of Polyglycol E600 and PPG 1025 were not characterized, but were considered to be block polymers by analogy to the behavior of DPG and to work done elsewhere (Ref 1). In the previous studies (*RS 36-6*, Vol 2; *RS 36-9*, Vol 1), two parameters of catalyst concentration were found to be significant; i.e., C moles of catalyst/liter, and R_s, moles catalyst/equivalent of hydroxyl. When C was kept constant, the low molecular weight glycols were cyanoethylated effectively, but immediate gelation occurred with the higher homologs. When R_s was constant, there was no gelation in the higher glycols, but the yields compared to those of the lower glycols were smaller. Some solid byproduct was formed at all catalyst levels for all the glycols studied. Further studies on the effect of the catalyst concentration have now shown that the byproduct can be eliminated for Polyglycol E600 and substantially reduced for PPG 1025.

Discussion

The results of the present experiments are given in Table 1. For the explanation of symbols used in this discussion. see the footnotes in Table 1. The cyanoethylation of EG (Run 1) was used as a model reaction. The catalyst concentration $R_s = 0.005$ gave high conversions, 98%. If the same R_s value was used with Polyglycol E600 (Run 2), no solid byproduct appeared, but the cyanoethylation was only 35%. When R_s (Runs 4 and 5) was increased to 0.018 ($C_s = 0.064$, $C_f = 0.052$) cyanoethylation increased to 94–95% and still no solid byproduct was formed. Further increase in catalyst concentration (Run 6) caused the appearance of solid byproduct, although the amount was small (<0.5%).

Table 1. Effect of catalyst concentration on the cyanoethylation of Polyglycol E600 and PPG 1025

Run	Cyanoethylation product	Catalyst concentration[a]			[CN] / [OH][b]	Catalyst	Temp, °C	Time, hr	Recovered material,[c] %	Cyanoethylation of recovered material,[d] %	Byproduct found
		R_s	C_s	C_f							
1	Ethylene glycol	0.005	0.180	0.052	1.0	Na	25	24	95	98	no
2	Polyglycol E600	0.005	0.018	0.014	1.0	Na	25	24	67[e]	35	no
3	Polyglycol E600	0.012	0.037	0.028	1.0	Na + CH_3ONa	25	24	70[e]	72	no
4	Polyglycol E600	0.018	0.064	0.052	1.0	CH_3ONa	25	6	82	94	no
5	Polyglycol E600	0.018	0.064	0.052	1.0	CH_3ONa	25	22	87	95	no
6	Polyglycol E600	0.025	0.090	0.072	1.0	CH_3ONa	25	6	89	96	yes
7	PPG 1025	0.030	0.058	0.052	1.0	CH_3ONa	25–30	22	95	80	yes
8	PPG 1025	0.018	0.035	0.031	1.0	CH_3ONa	25	22	95	—	yes
9	PPG 1025	0.018	0.035	0.031	1.0	Na	10, 4 hr 22, 1 hr	5	95	75	yes
10	PPG 1025	0.009	0.017	0.015	1.0	Na	10, 4 hr 22, 1 hr	5	95	82	yes

[a] Concentration:
R_s = [Na]/[OH] = moles catalyst per equivalent of hydroxyl at the start of the reaction
C_s = [RO^-] moles/liter = moles catalyst per liter of solution at the start of the run
C_f = [RO^-] moles/liter = moles catalyst per liter of solution at the end of the run

[b] Ratio of acrylonitrile to hydroxyl.

[c] Percent of material recovered after neutralizing, washing and extracting the reaction mixture, and evaporating the extracting solvent.

[d] Percent cyanoethylation based on the analysis of recovered material; values are obtained from acetylation.

[e] Weight losses high because both glycol and cyanoethylated product soluble in water.

The only difference between Runs 4 and 5 was the duration of the reactions, 22 and 6 hr, respectively. Since both the percent recovered material and the percent cyanoethylation are about the same in these two runs, it can be concluded that the cyanoethylation of Polyglycol E600 at 25°C is over within 6 hr. The final catalyst concentrations C_f in Runs 4 and 5 were the same as in Run 1. However, an effective R_s for cyanoethylation of Polyglycol E600 is about 3–4 times that of EG. Applied, however, to PPG 1025 (Run 7), solid byproduct appeared as a brownish-red colloidal suspension. When $R_s = 0.018$, the solid byproduct was again present (Run 8). Run 9 was performed at the same catalyst level as Run 8, but at a lower temperature (4 hr at 10°C). The amount of cyanoethylation in the two runs was comparable, but this time the colloidal suspension was a light yellow color. In Run 10, R_s was half the amount in Run 9, without any change in the other parameters. Here the suspension appeared later in the course of the reaction, and acetylation values showed some increase in percent cyanoethylation. The amount of solid suspension was estimated to be less than from the former runs.

Compared with Polyglycol E600 the percent cyanoethylation of PPG 1025 was lower in all cases. Lower yields are reported in the literature (Refs. 2, 3, and 4) for the cyanoethylation of secondary hydroxyls as compared to primary. The terminal OH groups of PPG 1025 are mostly secondary, while those of Polyglycol E600 are primary. It has been stated (Ref. 4) that the cyanoethylation of primary and secondary monohydric alcohols is an equilibrium reaction, and the position of equilibrium is more favorable to the addition product with primary than with secondary alcohols. If this is the case, the lower percent cyanoethylations with PPG 1025 are not surprising; neither is the appearance of solid byproducts towards the end of the reaction, when unreacted acrylonitrile would enter into block polymerization with the strongly basic long-chain alkoxide anion.

It is not yet clear whether the C_f values or the R_s value is the more important catalyst parameter. Dilution experiments are in progress whereby the OH concentration per liter of the low molecular weight glycol is reduced to the OH concentrations of the higher molecular weight glycols. Using the same R_s, C_s and C_f values as with Polyglycol E600 and PPG 1025, and determining the percent conversions, a clearer understanding of the relative importance of these catalyst parameters could be obtained. Studies of the relative rates of initiation (cyanoethylation) and propagation (block polymer formation) during the reactions of polyoxyalkylene glycols with acrylonitrile will form part of the future work.

B. Polymer Characterization: Gas-Liquid Chromatography of Polyoxyalkylene Glycols

A. J. Havlik and D. D. Lawson

The application of gas-liquid chromatography (GLC) as a useful tool in the study of the structure of polyoxyalkylene glycols was recently reported (*RS* 36-3, Vol 1, Pt 2). By the use of GLC the isomeric dipropylene and tripropylene glycols were separated on columns prepared from Polyox (Union Carbide Chemicals Company, New York, N. Y.) and Chromosorb (Johns-Manville, Manville, N. J.). Polyoxyethylene glycol *n*-mers up to 5 and polyoxypropylene glycol *n*-mers up to 9 were also separated on this liquid phase. Since Polyox, a high molecular weight polyoxyethylene glycol, showed an upper working temperature of 250°C and gave symmetrical peaks for polar-type samples, it was superior in performance to Carbowax (a low molecular weight polyoxyethylene glycol) and related types of liquid phases.

Additional studies of Polyox as a liquid phase have now been made. In particular, the quantity of liquid phase and its effect on the resolution and retention times of the POPG *n*-mers, POPG positional isomers and the two diastereomers of 1,1′-oxydi-2-propanol were investigated. Some preliminary results on the quantitative analysis of the isomeric DPG's and the *n*-mers in POEG and commercial POPG's were also obtained.

1. Column Study

The thickness of the coating of the liquid phase on the solid support affects the efficiency and resolution which can be attained from a GLC column (Ref. 5). It is desirable to have the largest possible amount of liquid phase per unit cross-section of the column and to coat the support with a sufficiently thin layer of the liquid phase to maximize the mass transfer rate. This ideally represents a layer of a few molecules of liquid phase on the solid support. For this purpose, glass microbeads are an ideal smooth surface as compared to Chromosorb. The lower surface area of the glass microbeads (avg. diam. 100 μ) requires low concentrations of Polyox as the liquid phase. The net effect of a low-loaded glass microbead column is that materials of much higher boiling point may be eluted and studied (Ref. 6). Chromatograms of P-250 and P-400 were obtained (Figs. 1, 2) on columns containing 20%, 10% (Chromosorb support) and 0.25% (glass microbead support) by weight of Polyox as the liquid phase. The glass microbead column gives less tailing, better overall resolution and significantly shorter retention times for the higher *n*-mers than the high-level columns.

2. Quantitative Studies

The area under each peak is related to the quantity of that material eluted from the column. For calculation of the composition from the measured peak areas, the method of internal normalization is ordinarily used (Ref. 7). This method amounts to comparing the proportion of a particular peak area to the summation of all peak areas, assuming of course that all components of the sample are eluted from the column. For accurate analysis the peak areas must first be multiplied by so-called calibration factors before they are added and proportioned. These factors are determined by the analysis of a suitable number of synthetic blends made up of the pure materials in question. When the materials are members of a homologous series, calibrations on the lower member may be extrapolated to higher members (Ref. 5).

Another relationship which may be used to relate compositions to peak areas is that of Eastman (Ref. 8).

$$\frac{m_i}{W} = \frac{A_i (M_i)^{1/2}}{\sum_i A_i (M_i)^{1/2}} \qquad (1)$$

Where m_i/W is the fraction by weight of the ith component, A_i is the integrated peak area, and M_i is the molecular weight of each component. To test this relationship a chromatogram of mono-, di-, tri-, and tetraoxyethylene glycols was obtained. The results summarized in Table 2 indicate an accuracy of only about $\pm 10\%$. However, it is clear from the nature of the equation that precision is required in the determination of areas on the chromatogram. Work is in progress to minimize tailing of peaks which is a source of error.

Table 2. Test of Eastman relation for polyoxyethylene glycols

Compound	Measured area, %	Material, wt %	
		Initially used	Calculated, Eq. (1)
Ethylene glycol	22.2	16.9	15.4
Diethylene glycol	23.5	25.0	21.5
Triethylene glycol	25.9	25.6	28.1
Tetraethylene glycol	28.3	32.2	34.9

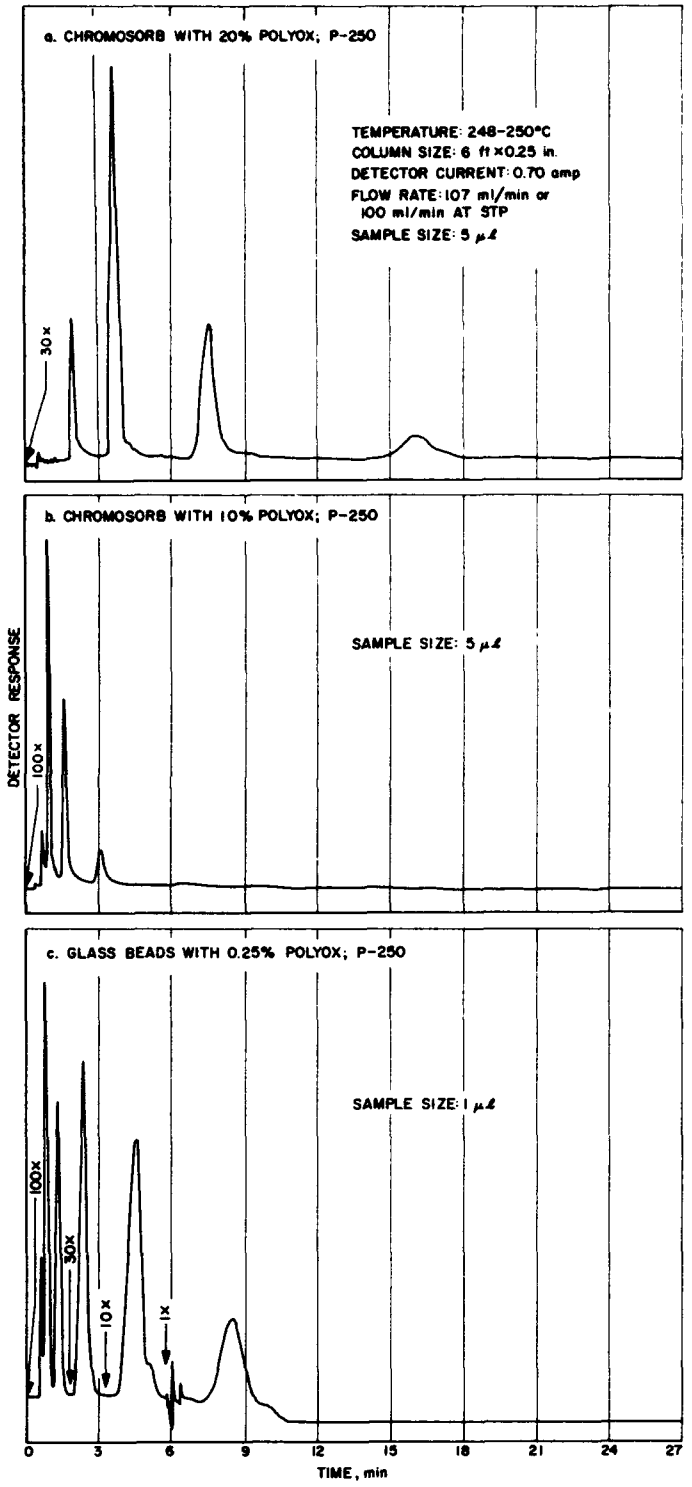


Fig. 1. Chromatograms of P-250

The relative positional isomer contents of commercial DPG and TPG and the *n*-mer distributions in commercial POPG's (PPG-150 to PPG-425) were obtained in terms of percentage peak areas for each component in the mixture. The results are shown in Table 3. Dipropylene glycol was separated into the three possible positional isomers.

The diastereoisomers of disecondary DPG (1,1′-oxydi-2-propanol, (Fig. 3) have been separated by low-temperature crystallization and partially characterized. A mixture of the racemic pair and meso form was not resolvable using either GLC columns prepared with Chromosorb or glass microbead supports and Polyox as the liquid phase.

A sample of P-250 was analyzed under column conditions which were capable of easily resolving the mixture of DPG and TPG positional isomers in commercial materials. A sample of P-250 containing added disecondary DPG (1,1′-oxydi-2-propanol) was analyzed under similar conditions. The disecondary isomer of DPG and one TPG isomer are the only ones which are apparently present in significant amounts (Fig. 4). The peak shapes of the higher *n*-mers suggest that only one positional isomer of each of the higher *n*-mers may be present in signifi-

Table 3. Analysis of commercial glycols

Product and source	Component	Peak area, %
DPG (Union Carbide)	Propylene glycol	1.3 ±0.3
	Disecondary DPG	43.0 ±1.6
	Secondary-primary DPG	46.6 ±0.8
	Diprimary DPG	5.7 ±0.7
	Unknown	4.5 ±0.7
TPG (Dow Chemical)	Low boiling materials	1.4
	Second component	49.5
	Third component	35.1
	Fourth component	8.9
PPG-150 (Union Carbide)	Low boiling materials	1.6
	DPG isomers	47.7
	TPG isomers	51.8
P-250 (Dow Chemical)	Low boiling material	1.3 ±0.5
	Second component	8.9 ±0.6
	Third component	33.2 ±0.6
	Fourth component	31.4 ±0.7
	Fifth component	17.8 ±0.1
	Sixth component	8.6 ±0.5
PPG-425 (Union Carbide)	Low boiling materials	1.05 ±0.005
	Second component	0.5 ±0.04
	Third component	1.5 ±0.3
	Fourth component	5.7 ±0.3
	Fifth component	13.6 ±0.6
	Sixth component	21.6 ±0.9
	Seventh component	35.8 ±1.2
	Eighth component	19.7 ±0.8

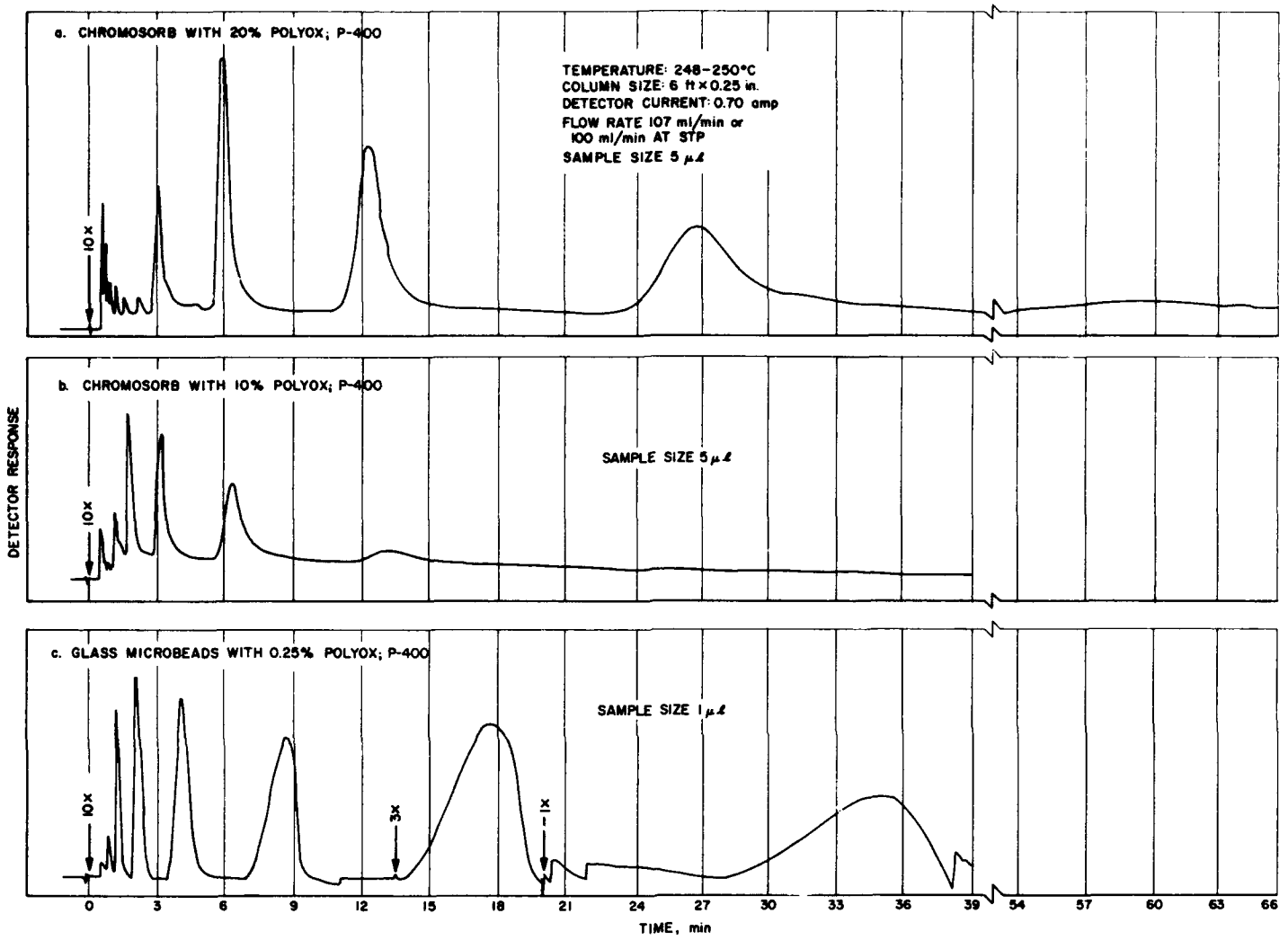


Fig. 2. Chromatograms of P-400

cant quantities in these commercial products (P-250, P-400 and PPG-425). This indicates that DPG and TPG are probably produced under different reaction conditions from those of P-250 and the higher POPG's. Tripropylene glycol contained two large fractions not further identified yet, while (PPG-150 constituted nearly equal amounts of DPG and TPG isomers (Fig. 5). The hexamer could be found in P-250 and the octamer in PPG-425.

Further work on GLC methods for the polyoxyalkylene glycols will be directed toward obtaining the following: (1) quantitative data on pure samples of the *n*-mers in the POEG and POPG systems, (2) the *n*-mer distribution in samples of low molecular weight POPG's prepared under different experimental conditions, and (3) the thermodynamic activity coefficients for selected *n*-mers of the polyoxyalkylene glycols.

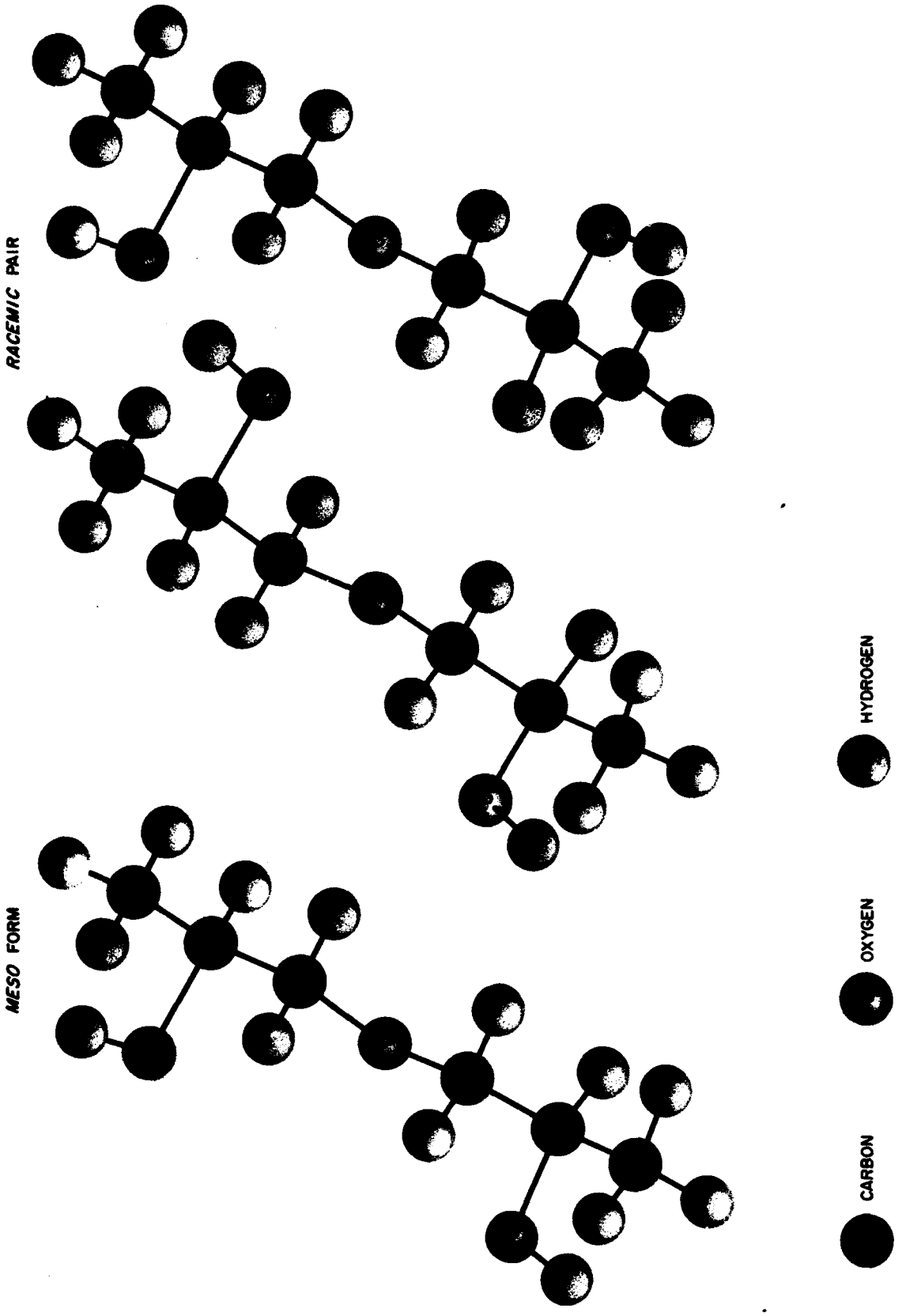


Fig. 3. Configurations of the diastereomers of 1,1'-oxydi-2-propanol

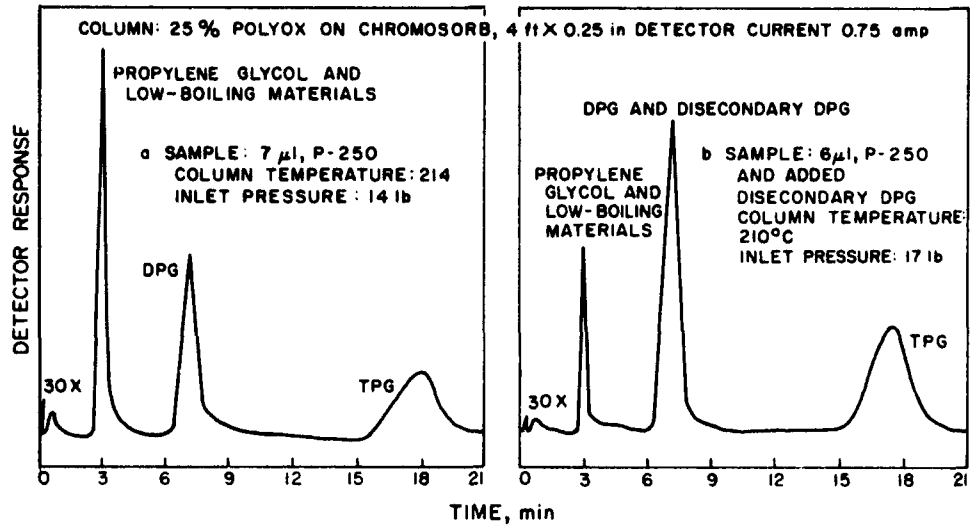


Fig. 4. Chromatograms of P-250 and P-250 with added disecondary DPG

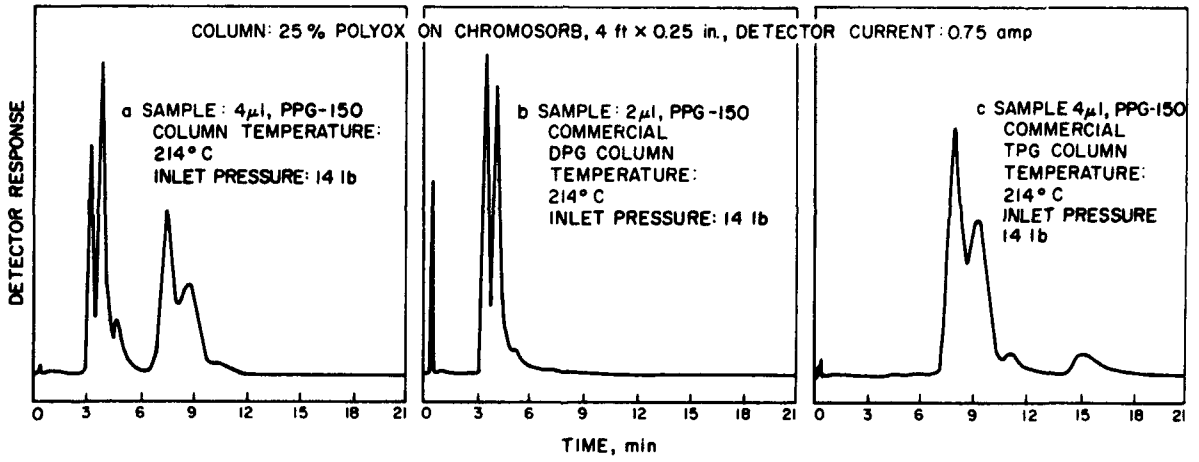


Fig. 5. Chromatograms of PPG-150, DPG and TPG

C. Polymer Degradation Mechanisms

J. D. Ingham and N. S. Rapp

Degradation of polymers is being studied to provide a chemical basis for the development of nondestructive test methods applicable to solid propellants that may undergo chemical changes before use. Thermal degradation in vacuum of polyoxypropylene glycol-toluene diisocyanate (PPG-TDI) is being investigated to determine the mechanism of bond scission.

1. Polymer Column-Elution Fractionation

Fractionation data for several undegraded polymers and one degraded PPG-TDI polymer have been presented and compared with the most probable distributions (*QSR 38-6*). An improved fractionation of degraded polymer that avoids the silicone nonsolvent and accompanying degradation during fractionation has since been completed. The procedure consists of fractionating in the usual way with benzene solvent-isooctane nonsolvent, but at a temperature of 7°C rather than 34°C. The ratio of weight to number average molecular weight (M_w/M_n) was 2.05, conclusively showing that degradation from M_w of 100,000 to 25,000 ($[\eta]$ decreases from 0.68 to 0.27 dl/g) at 200°C in a vacuum takes place by a random scission process.

2. Polymer Weight Loss Measurements

To further clarify the degradation mechanism of PPG-TDI, volatilization measurements at relatively high temperatures (250 to 320°C) in vacuum of <10 μ Hg have been obtained. The balance consisted of a fine pyrex

capillary fixed at one end and positioned horizontally. This was found to be of nearly the same sensitivity, 0.9 mm/mg, as the double pan balance described previously and more reliable. Deflection of the capillary was measured with a steel scale and fixed magnifying glass or, more recently, a cathetometer reading to 0.1 mm. The load is limited to 70 mg, but this was found to be more than adequate for these measurements. The 5 to 15 mg sample was weighed onto an aluminum foil pan which was attached to the free end of the capillary by means of a thin tungsten wire and suspended in the evacuated pyrex tube immersed in the furnace. An experiment showed that identical weight loss rates were obtained using either aluminum or platinum pans. The furnace consisted of a large test tube wrapped with nichrome wire and controlled to within ± 0.5°C with a proportioning controller. The temperature was determined from a thermometer suspended inside the furnace tube near the sample. Although differences in thermal emissivity of the sample and thermometer could lead to errors in absolute temperature, those were probably not more than 2°C and were relatively constant for separate runs as indicated by the straight-line activation energy plots obtained.

Data for polyoxypropylene glycol 2025 (PPG) are shown in Fig. 6 at 275°C and 305°C. Wall, Madorsky, Brown, Straus and Simha have shown (Ref. 9) that a maximum in volatilization rate should occur at approximately 25% weight loss for a random process, followed at higher conversions by a steadily decreasing rate. However, for a polymer of such low initial molecular weight ($\sim$2000), random scission should result in very rapid volatilization of all the sample at low extents of degradation. Since the curves of Fig. 6 do not show a well-defined maximum near 25% volatilization nor the other characteristics of a random process, PPG apparently does not degrade randomly. This is in agreement with results obtained for PPG of much higher initial molecular weight (Ref. 10).

In Fig. 7 are shown a series of volatilization vs time curves for PPG-TDI polymer having an initial molecular weight (M_w) of 122,000. The results, strikingly similar to those for PPG, show substantially zero-order dependence for conversions from 20 to 65% and indicate the same mechanism of degradation at >20% volatilization. Activation energy plots are given in Fig. 8 for rates calculated from the straight-line portions of the curves (25 to 65% conversion). The activation energies are the same within experimental error, 52.0 kcal/mole for PPG-TDI and 48.2 kcal/mole for PPG.

A plot of molecular weight of the residue expressed in percent of the original vs percent volatilization for PPG-TDI is shown in Fig. 9. The molecular weights were estimated from intrinsic viscosity data, using the equation (Ref. 11)

$$[\eta] = 4.13 \times 10^{-4} M^{0.64} \qquad (1)$$

The data points represent separate experiments at different temperatures, as shown in the figure. These results as well as the fractionation data given previously indi-

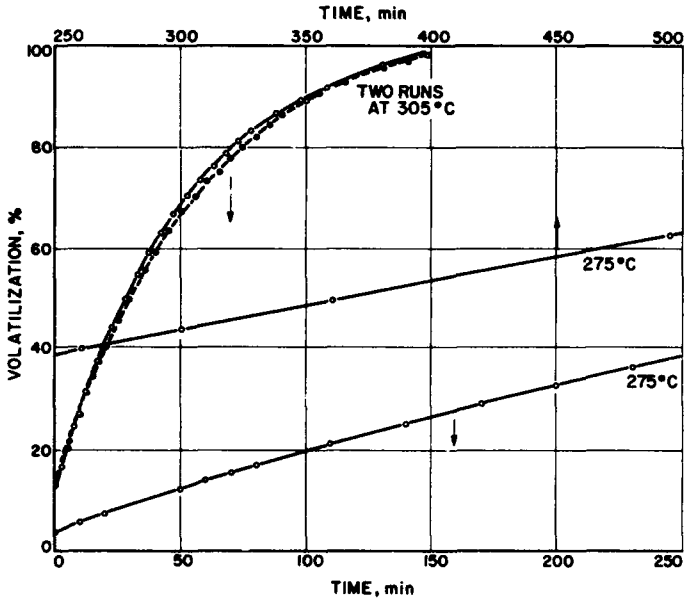


Fig. 6. Volatilization of PPG-2025 vs time

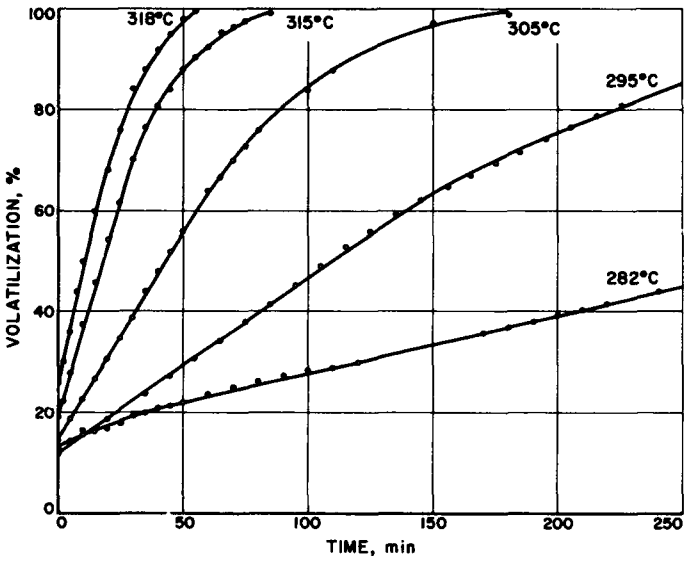


Fig. 7. Volatilization of PPG–TDI polymer vs time

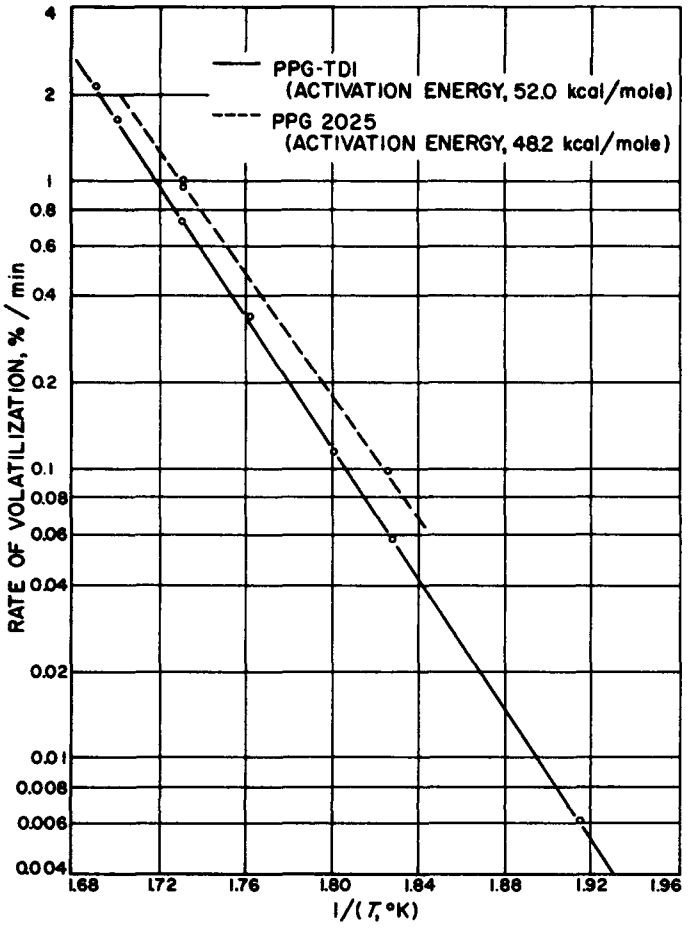


Fig. 8. Rate of volatilization of PPG–TDI polymer and PPG-2025 vs 1/T, °K

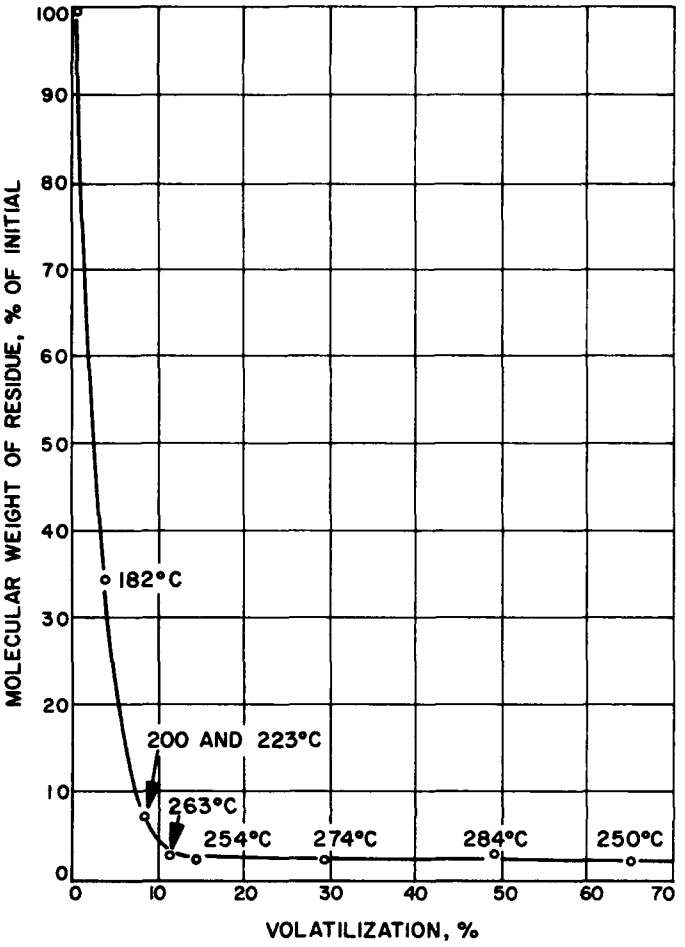


Fig. 9. Decrease in molecular weight vs volatilization of PPG–TDI

cate random scission, probably occurring at the urethane linkages up to 10% conversion. The constancy of the molecular weight (which was ~2400, with one point at 3800) and the volatilization from 10% to 65% indicates free radical chain depolymerization with a minimum kinetic chain length corresponding to a molecular weight of about 2400. However, the approximate zero-order dependence of the rate curves of Fig. 7 suggests a stepwise depolymerization, presumably occurring at the constant number of chain ends or initiating sites. Since the volatilization rates were slightly less for PPG-TDI than for PPG, fewer sites for initiation of depolymerization would be available in the former, indicating the presence of inactivated chain ends, a slightly higher molecular weight, or both.

From the data obtained thus far it appears that PPG-TDI degrades by random scission of most of the urethane linkages at temperatures <250°C, followed by a free-radical depolymerization of the PPG segments from 250 to 320°C. The kinetic chain length has not been determined but is probably less than or roughly equal to that of undegraded PPG. It is not extremely high nor is it a simple depolymerization to propylene oxide, as indicated by the variety of products and range of molecular weights obtained by degrading PPG of molecular weights of 16,000 and higher (Ref. 10). Actually the most abundant volatile products formed were acetaldehyde, propene, and acetone, plus a relatively small amount of propylene oxide.

Because of the diversity of chemically reactive groups in PPG-TDI polymers and the general complexity of urethane and ether chemistry, a large number of simultaneous degradative reactions undoubtedly take place. Therefore, studies of the chemical species obtained from degradation of PPG-TDI and PPG polymers will be carried out, particularly at extents of conversion of <10%.

D. Effect of Pressure on Conductivity of Polyacenequinone Radical Polymers

A. Rembaum, J. Moacanin, and E. Cuddihy

Experimental data obtained on the conductivity of organic compounds fall into four general classes: (a) DC resistivity of powders, (b) AC resistivity of powders, (c) DC resistivity of evaporated films, and (d) DC resistivity of single crystals (Ref. 12). Previously (Ref. 13), it was shown that good correlations can be obtained between DC or AC conductivities of powders and their structure. But correlations with films were poor. Since the polymers currently under study are obtained as powders and in view of the above conclusions, the instrumentation necessary for DC measurements on powders has been developed. AC measurements are also of interest (Refs. 13 and 14) and will be considered later.

To improve contact between adjacent grains, powders are usually highly compressed during measurements. It has been observed (Ref. 15) that the conductivity of some compounds continues to vary with pressure even when the compound has been precompressed. An apparatus was therefore developed to allow conductivity measurements to be made as functions of both pressure and temperature. Using this equipment, an evaluation was made of the properties of four polyacenequinone radical PAQR powders. The conductivities, forbidden energy gap E_g, and pre-exponential factors of these materials were studied as functions of both pressure and temperature.

1. Experimental Procedure

Conductivity measurements were carried out in an apparatus which consisted of a cell, press, and electrometer.

The cell consists of two cylindrical vanadium alloy steel (Vasco Supreme, Vanadium Corp., New York, N. Y.) anvils of 0.25-in. contact diameter held vertically between two hardened steel plates. The outside surfaces of the plates are electrically insulated from the press platens by Micarta. The lower anvil rests in a groove in the lower plate and the top anvil is held to the upper plate by means of an aluminum housing. Guides are provided on the plates to align and center the anvils. Guide pins and Teflon sleeves align the plates yet allow motion normal to their plane. The component parts of the cell assembly are shown in Fig. 10. The sample powder is introduced into a cavity formed by a neoprene O-ring situated on top of the lower anvil. The top plate is then pushed down until the contact area of the top anvil rests on the sample. The cell assembly is placed between the platens of the hydraulic press and surrounded by a 0.5-in. stainless steel safety shield. The cell is heated by a forced hot air stream circulating within the shield. Temperature is monitored by a thermocouple placed close to the sample. The thickness of the sample is determined to within 0.001 cm by using a cathetometer to

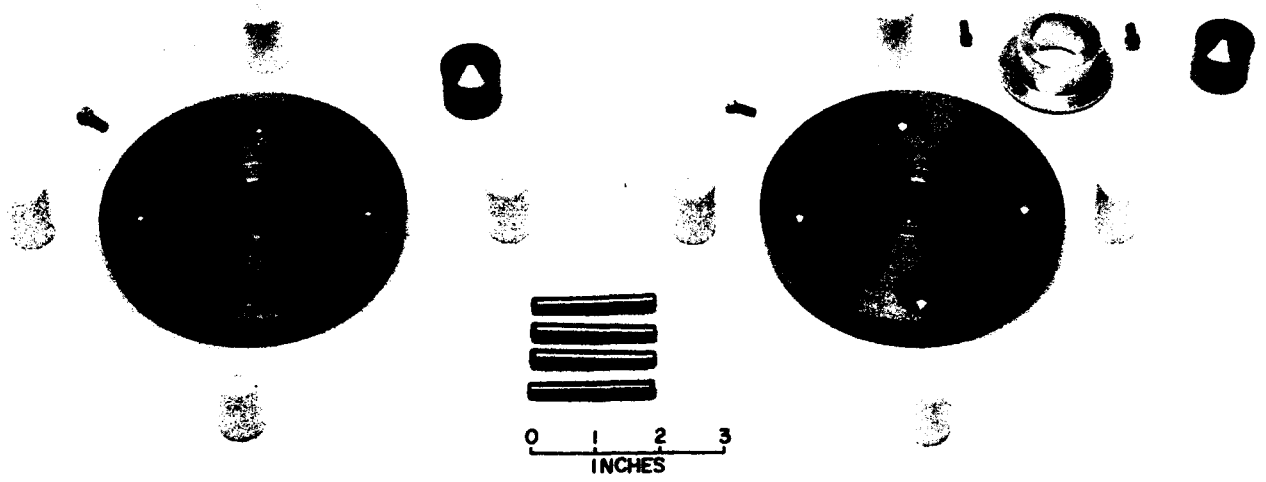


Fig. 10. Component parts of the cell assembly

Fig. 11. Cell assembly and shield mounted in the press

measure the displacement of the anvils. The cell and the shield both mounted in the press are shown in Fig. 11.

The hydraulic press has a full load capacity of 40,000 lb of force (Pasadena Hydraulics Inc., El Monte, California).

Resistance measurements were taken by means of an electrometer which has a variable range of 10 to 10^{14} ohms full scale (Kiethley Model 601A Electrometer, Kiethley Instruments, Inc., Cleveland, Ohio). Electrical connections were made at the cell by using the steel plates as terminals.

The PAQR powders studied were contributed by Dr. Herbert A. Pohl of Princeton University. The polymers are condensation reaction products of

JAB 65—Two parts anthracene to one part phthalyl chloride

JAB 28—Two parts 2-methyl anthraquinone to one part pyromellitic dianhydride

KHO 4—One part phenanthrene to one part pyromellitic dianhydride

KHO 6—One part anthracene to one part pyromellitic dianhydride

2. Results and Discussion

A study was made of the DC conductivities of four of the PAQR powders at eight pressures from 7,000 to 57,000 atm and temperatures of about 22, 60 and 93°C. The experimental results are summarized in Figs. 12, 13, and 14.

The temperature dependence of the conductivity of organic materials at a given pressure may be represented

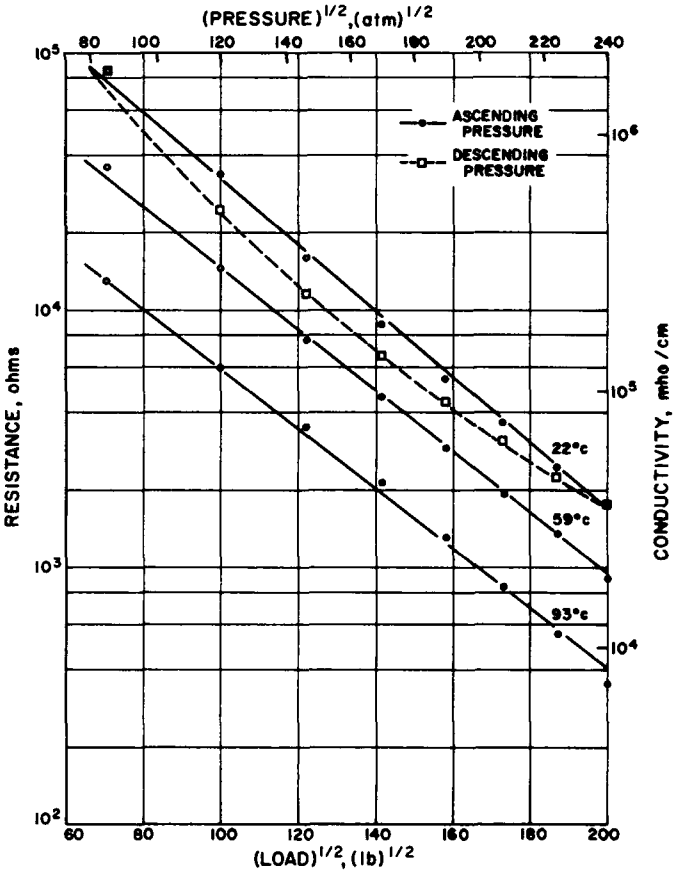


Fig. 12. Log resistance vs the square root of the load for KHO-6

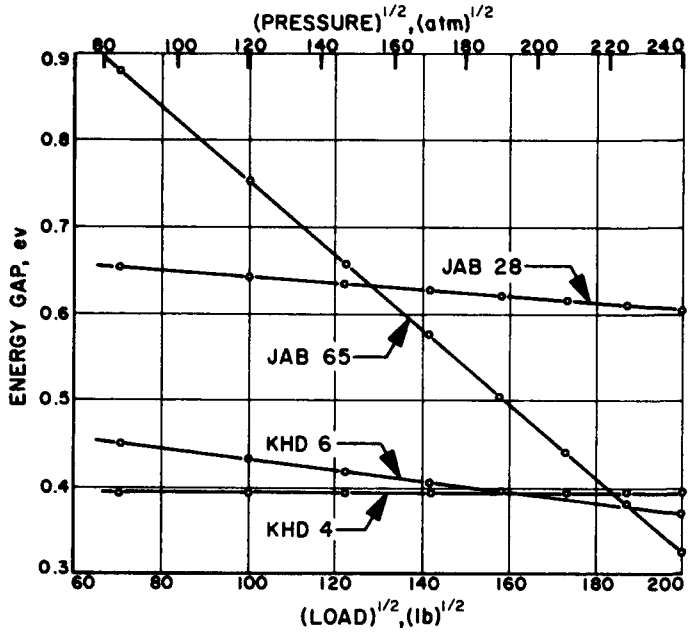


Fig. 13. Energy gap vs the square root of the load for four PAQR polymers

adequately by the standard relation

$$\sigma = \sigma_\infty e^{-E_g/2kT}$$

where σ is the conductivity at temperature T, σ_∞ is the conductivity extrapolated to $T = \infty$, k is the Boltzmann constant, and E_g is the forbidden energy gap. Both σ_∞ and E_g are independent of temperature but functions of the pressure P. By semi-empirical arguments it can be shown (Ref. 15) that E_g should be proportional to $p^{1/4}$, which implies a linear relation between log σ and $p^{1/4}$ at constant temperature. Equivalent relations hold for the experimentally observed resistance R (σ = thickness/$R \times$ area). Data for sample KHO 6 plotted in this manner are shown in Fig. 12. Consistent with the above assumption, the data for ascending pressure are linear. The curvature observed for the descending pressure (shown for 22°C data only) was found to be due to mechanical drag and therefore such data were not used for further analysis. After an initial compression to 28,000 atm, the thickness of the sample remained constant within the accuracy of the measurements over the entire pressure range. The thickness of the samples studied ranged from 0.013 to 0.033 cm.

Conductivities for the four powders were determined from experimentally observed resistances and are, at room temperature and 21,000 atm, 2.49×10^{-7} mho/cm for JAB 65, 3.83×10^{-6} mho/cm for JAB 28, 3.83×10^{-5} mho/cm for KHO 4 and 2.96×10^{-6} mho/cm for KHO 6.

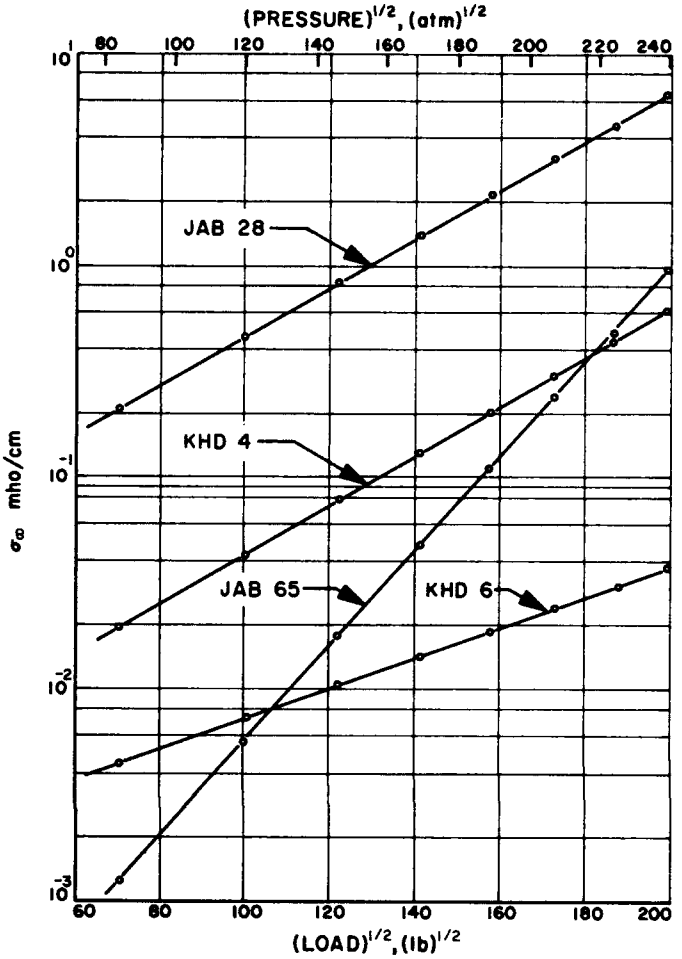


Fig. 14. Log σ_∞ vs the square root of the pressure for four PAQR polymers

For any given polymer at constant temperature, the conductivity increases by about a factor of 100 over the pressure range investigated. At constant pressure, conductivity increases by about a factor of ten over a 70°C temperature range.

From the slope and intercept of log σ vs $1/T$ at constant pressure, E_g and σ_∞ were determined. Plots of E_g and σ_∞ vs the square root of the pressure are shown in Figs. 13 and 14. Over the pressure range investigated, E_g for JAB 65 goes from 0.880 to 0.322 ev with increasing pressure; for KHO 6, from 0.452 to 0.366 ev; for JAB 28, from 0.652 to 0.602 ev; and is a constant 0.393 ev for KHO 4. σ_∞ for JAB 65 goes from 1.24×10^{-3} to 9.57×10^{-1} mho/cm, from 4.47×10^{-3} to 3.67×10^{-2} for KHO 6, from 2.10×10^{-1} to 6.24 mho/cm for JAB 28, and from 1.97×10^{-2} to 6.24×10^{-1} mho/cm for KHO 4. Thus both the rate of change and the magnitude of E_g and σ_∞ depend on the polymer composition.

E. Measurement of Viscoelastic Strains on a Low-Modulus Filled Elastomer

A. San Miguel

The relationships between stresses and strains are known as the mechanical properties of the particular medium and may be defined by any appropriate set of constitutive equations. A problem arises in choosing the proper set of constitutive equations (e.g., elasticity, plasticity, and viscoelasticity) to apply to the medium. When this occurs, the meanings of physical properties such as elastic modulus and Poisson's ratio are obscure. The discrepancy between how a medium is supposed to respond and the way it actually responds is usually accounted for by "appropriate" correction factors. The object of this study at the Laboratory is to review current concepts of, or redefine if necessary, the viscoelastic properties of solid propellants.

The importance of being able to define mechanical properties of solid propellants may best be illustrated by a number of examples. The structural design engineer would like to predetermine the strains in a solid propellant grain for various "working" load conditions. This information must be realistic if the integrity of the motor is to be assured. The need for accurate grain strain analysis is amplified by the trend to high mass ratio motors. The ballistics engineer is interested in burning rate and ignition pressure limitations. These characteristics are functions of the state of deformation in solid propellant grains. Finally, the physical chemist needs to know the measurable physical properties which can be interpreted directly in terms of chemical structure and which can be used as a basis for further material improvement.

In view of the above objectives, a basic investigation into the field of propellant viscoelasticity was initiated. Viscoelasticity will be defined here as the study of the relationships between an environment (e.g., loadings) and a medium response (e.g., deformations) as a function of time. In order to maintain as much generality as possible, the subject was approached using the classical field theories of a continuous medium (Ref. 16). Although the theory is quite elegant and difficult to interpret, an interesting observation can be made. A definite relationship between the strain invariants (I_1, I_2, I_3) and stresses (τ_{ij}) exists for a medium exhibiting an elastic potential. It is noted that a solid propellant, when subjected to any type of deformation, exhibits a time-dependent elastic recovery when physically unloaded. A solid propellant also exhibits a healing phenomenon in that a dewetted propellant will cure itself with time. Therefore one would expect that in the region of operational strains (to 30%) a propellant may be described as a medium exhibiting a time-dependent elastic potential. Accepting these experimental observations, the constitutive equations are of the following form:

$$\tau^{ij} = \frac{1}{2}\left(\frac{g_{mn}}{\bar{g}_{mn}}\right)^{1/2}\left\{\frac{\partial W}{\partial I_1}\left(\frac{\partial I_1}{\partial \gamma_{ij}} + \frac{\partial I_1}{\partial \gamma_{ji}}\right) + \frac{\partial W}{\partial I_2}\left(\frac{\partial I_2}{\partial \gamma_{ij}} + \frac{\partial I_2}{\partial \gamma_{ji}}\right) + \frac{\partial W}{\partial I_3}\left(\frac{\partial I_3}{\partial \gamma_{ij}} + \frac{\partial I_3}{\partial \gamma_{ji}}\right)\right\}$$

where

τ^{ij} = contravariant stress tensor

$g = |g_{mn}|$ = covariant metric tensor (unstrained state)

$\bar{g} = |\bar{g}_{mn}|$ = covariant metric tensor (strained state)

$W(I_1, I_2, I_3)$ = elastic potential per unit volume of the unstrained body; the variable, time t, is implicitly present in the derivation of the strain invariants (I_1, I_2, I_3).

I_1 = first strain invariant
$\lambda_1^2 + \lambda_2^2 + \lambda_3^2$

I_2 = second strain invariant
$\lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_1^2\lambda_3^2$

I_3 = third strain invariant
$\lambda_1^2\lambda_2^2\lambda_3^2$

$\lambda_1, \lambda_2, \lambda_3$ = the principal strain extensions
$(1 + \gamma_{ii})$

γ_{ij} = finite strain components at a point
$\gamma_{11}, \gamma_{22}, \gamma_{12}$, etc.

The problem now is to solve for $W(I_1, I_2, I_3 t)$ in the above equations. This requires an experimental program in which the environment (e.g., loadings) is controlled and the response (e.g., deformations) is observed. Matters are somewhat simplified in that control and observation of the boundary conditions would produce sufficient information to define W. This is because the invariant properties of a particular medium demand that the boundary properties be those of the interior.

In studies of the physics of rubberlike behavior, a few investigators (Refs. 17, 18, 19) have approached the problem by assuming that rubber is incompressible ($I_3 = 1$). Blatz (Ref. 20) has abandoned this hypothesis, however, in favor of retaining the $\partial W/\partial I_3$ term of Eq. (1). Solid propellants cannot be treated as incompressible, and the assumption $I_3 = 1$ is not valid. The first attempt to apply certain concepts of invariant theory to propellants is

found in Ref. 21. In this paper a plot of I_1 vs I_2 for several propellants indicated a linear relationship for a uniaxial loading. Of interest here was that this relationship was linear in the range of operational propellant strains (to 30%). Another interesting observation was that the slope $\Delta I_1/\Delta I_2$ was different for different propellants. This indicated that the slope may take on the aspects of a measurable physical property. The assumptions used to compute $\lambda_1, \lambda_2, \lambda_3$ were realistic for idealized tensile loading but questionable for tensile testing as normally performed, where the actual propellant loading may have bi- and tri-axial components as well. This is because any assumption regarding λ_3 in terms of λ_1 and λ_2 requires the use of the elasticity field equations which are meaningless in view of the above definition of viscoelasticity.

A study program to develop a method to measure principal strains and their corresponding directions on propellants was initiated some time ago at the Laboratory. In *QSR 38-3* and Ref. 22 a photoelastic coating technique was developed. Here it was shown that a low modulus (500 psi) birefringent resin could successfully be bonded to a propellant in order to measure principal strains and their corresponding directions. This coating technique is analogous to the use of an infinite array of surface strain gages. A detailed discussion on how to graphically reduce strain data from photoelastic coatings may be found in Ref. 23.

In order to exploit the technique, biaxial loading experiments of unbonded inert propellant grains were developed. These experiments consisted of pressurizing a rubber bag in two dimensions (*QSR 38-4*). Aside from establishing a standardized biaxial test analogous to the uniaxial tensile test, it was found that strains predicted by infinitesimal elastic theory were quite conservative. Many interesting strain fields for various propellant port configurations were investigated (*QSR 38-5*). Finally (*QSR 38-6*), several bonded live propellant grains were studied for various case wall thicknesses. Here it was shown that, for an optimum mass ratio (lowest propellant strain field), a casing of varying radial thickness is necessary. This observation is contrary to current philosophy regarding motor design; i.e., current rocket cases are designed with a constant radial thickness, and the port configuration is altered as best as possible to reduce strain concentrations.

Since the initiation of the photoelastic coating technique, basic instrumentation has been developed to measure surface strains photoelastically as a function of time. Photoelastic data (principal strains and their directions) as used in the study of continuous media are illustrated in Fig. 15. One observes that the relationship between the invariants is linear for at least strains of the order of 5%. This relationship is encouraging since it holds for any combination of loadings. These data were obtained from twelve propellant grains out of the same mix batch. These grains were subjected to various case reinforcements and loadings. Numerous surface points at each loading were examined for principal strains. These observed strains, along with the assumption that "strain normal to a free surface, in a state of biaxial loading, is zero at the surface but immediately increases in magnitude as one progresses into the interior," enabled the computation of the invariants (I_1, I_2, I_3).

It is noted that, although the relationships between the invariants are linear for the uniaxial and biaxial loadings (tensile specimens and motors), the slope is somewhat different if assumptions are made similar to those in Ref. 21. A further discussion on this subject may be found in *QSR 38-7*.

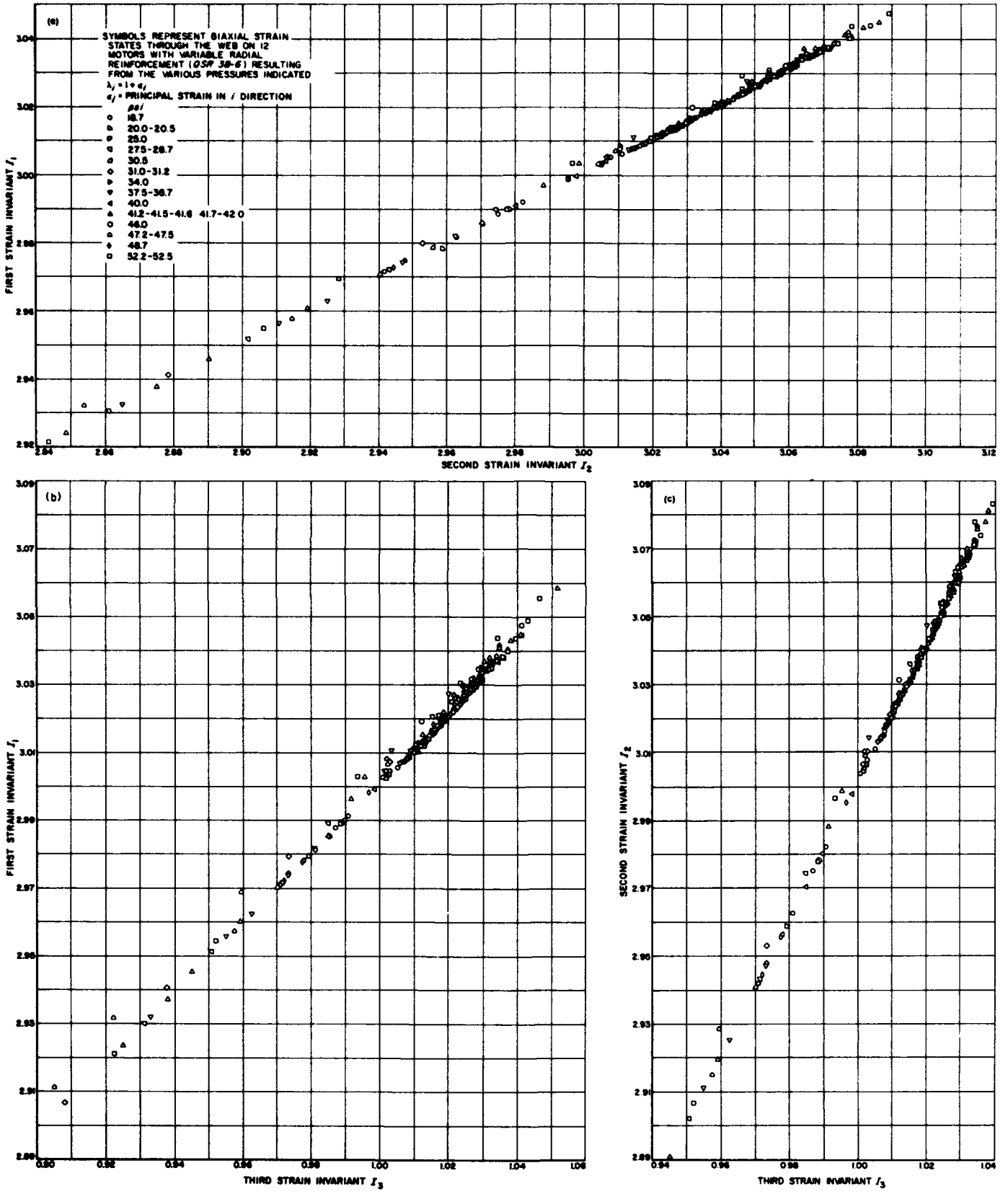


Fig. 15. Invariants I_1 vs I_2, I_1 vs I_3, and I_2 vs I_3 for biaxial loading assuming $\epsilon_3 = 0$

References

1. Furukawa, J., Saigusa, T., and Mise, N., "Synthesis of Block-Copolymer of Ethylene Oxide and Acrylonitrile," *Makromolekulare Chemie,* 38:244–247, 1960.

2. Bruson, H. A., and Reiner, T. W., "The Chemistry of Acrylonitrile IV. Cyanoethylation of Active Hydrogen Groups," *Journal of the American Chemical Society,* 65; 23–27, 1943.

3. Bruson, H. A., *U.S. Patent* 2,401,607, 1946.

4. Utermohlen, W. P, Jr., "Preparation of γ-alkoxy-*n*-propylamines," *Journal of the American Chemical Society,* 67:1505–1506, 1945.

5. Johns, T., *Gas Chromatography Applications Manual,* Beckman Instruments, Fullerton, Calif., pp 41–47, 1959.

6. Hishta, C., et al., "Gas Chromatography of Solid Organic Compound," *Analytical Chemistry,* 32; 880, 1960.

7. Keulemans, A. I. M., Kwants, A., and Rynders, G. W. A., "Quantitative Gas Chromatography," *Analytica Chimica Acta,* 16:29, 1957.

8. Eastman, R. H., "Semi-Quantitative Gas Chromatography," *Journal of the American Chemical Society,* 79:4243, 1957.

9. Wall, L. A., Madorsky, S. L., Brown, D. W., Straus, S., and Simha, R., "The Depolymerization of Polymethylene and Polyethylene," *Journal of the American Chemical Society,* 76:3430–3437, 1954.

10. Madorsky, S. L., and Straus, S., "Thermal Degradation of Polyethylene Oxide and Polypropylene Oxide," *Journal of Polymer Science,* 36:183–194, 1959.

11. Moacanin, J., "Diisocyanate-Linked Polymers. I. Dilute Solution Properties of Toluene Diisocyanate-Extended Polypropylene Glycol," *Journal of Applied Polymer Science,* 1:272–282, 1959.

12. Astronautics Information. "Organic Semiconductors: Properties and Applications," *Literature Search No. 341,* compiled by D. I. Sweitzer, Jet Propulsion Laboratory, Pasadena, Calif., September 1, 1961.

13. Schieler, L., Anderson, H. R., and Nichols, P. L., Jr., *Semiconductivity of Organic Compounds,* Memorandum No. 20-162, Jet Propulsion Laboratory, Pasadena, Calif., July 18, 1958.

14. Eley, D. D., and Parfitt, G. D., "The Semiconductivity of Organic Substances, Part II," *Transactions of the Faraday Society,* 51, 1529–1539, 1955.

15. Pohl, H. A., Rembaum, A., and Henry, A., "Effects of High Pressure on Some Organic Semiconducting Polymers," paper presented at the *141st Meeting of the American Chemical Society,* Washington, D. C., March 20–26, 1962.

16. Truesdell, C., and Toupin, R. A., "The Classical Field Theories," *Handbuch der Physik,* Band III/1, Springer-Verlag, Berlin, 1960.

References (Cont'd)

17. Grumbell, S. M., Mullins, L., and Rivlin, R. S., "Departures of the Elastic Behavior of Rubber in Simple Extension from the Kinetic Theory," *Transactions of the Faraday Society*, 49, 1495, 1953; Rubber Chemistry and Technology 28, 24, 1955.

18. Eirich, F. R., *Rheology*, Vol 1, Ch. 10, "Large Elastic Deformations," Academic Press, Inc., New York, 1956.

19. Mooney, M., "A Theory of Large Elastic Deformations," *Journal of Applied Physics*, 11, 582, 1940.

20. Blatz, P. J., and Ko, W. L., "Mechanical Behavior of Rubber Material," presented at *32nd Meeting of the Society of Rheology*, Madison, Wis., October 1961.

21. Bills, K. W., Jr., Hart, W. D., and Holland, W. E., "Effect of Dewetting and Volume Change on the Tensile Behavior of Solid Composite Propellant," *20th Meeting Bulletin of the Joint Army-Navy-Air Force-ARPA-NASA Panel on Physical Properties of Solid Propellants*, Vol 2, November 14–16, 1961. SPIA, Applied Physics Laboratory, Johns Hopkins University, Silver Spring, Maryland.

22. San Miguel, A., "Propellant Strain Analysis by the Photoelastic Coating Technique," *Bulletin of the 20th Meeting, JANAF SPIA*, the Johns Hopkins University, Silver Spring, Maryland, November 1961.

23. San Miguel, A., *Strain Measurements on a Pressurized Solid Propellant Grain*, Technical Report 32-182, Jet Propulsion Laboratory, Pasadena, Calif., March 15, 1962.