THE ATMOSPHERES

OF

MARS and VENUS

A Report by the

Ad Hoc Panel on Planetary Atmospheres

of the

Space Science Board

Prepared by

William W. Kellogg

and

Carl Sagan

Publication 944

National Academy of Sciences — National Research Council

Washington, D. C.

1961

Library of Congress Catalog Card No. 61-64604

FOREWORD

A Conference on Planetary Atmospheres, with especial reference to Mars and Venus, was held by the Space Science Board of the National Academy of Sciences on June 24, 1960, at Arcadia, California. The purpose of the conference was to discuss the present state of knowledge of planetary atmospheres as known from ground-based observations, to consider the most important characteristics about which additional information is needed, and to discuss space experiments that best promise to yield such information.

The significance of this examination is two-fold. First, the subject is in itself important to science, and definition and assessment of limits of knowledge, areas of ignorance, and the succession of problems is useful in guiding research into the most fruitful directions. Second, the assessment is necessary if we are most wisely to plan the conduct of explorations and investigations by costly spacecraft and space expeditions.

A capability for launching heavy payloads of scientific equipment to investigate other planets of our solar system can be anticipated within the next few years. It is to be expected that the initial investigations of these planets will involve direct measurements of their atmospheres, even before scientific apparatus is landed on their surfaces. Such scientific studies are important not only to the investigation of the planets themselves but because they may also be expected to contribute substantially to our knowledge of the origins and natural behavior of the solar system and of the Earth.

In view of the importance of the subject, and as a sequel to its Arcadia conference, the Board appointed an Ad Hoc Panel on Planetary Atmospheres to study, in greater detail, the state of our knowledge and some of the controversies concerning planetary atmospheres and to explore experimental approaches most likely to lead to the resolution of these controversies. This is the report of that Panel.

The Panel limited its deliberations and its report to the two nearest "terrestrial planets," Mars and Venus, both because of their common qualities of special interest and because they are the objects of man's first approaches to other planets of the solar system.

The Panel held two three-day meetings (December 15-17, 1960, and February 2-4, 1961), both on the campus of the California Institute of Technology, Pasadena, with members of NASA's Jet Propulsion Laboratory making many of the arrangements as well as taking a part in the discussions. Panel members were: William W. Kellogg, chairman (RAND Corporation), Bernard F. Burke (Carnegie Institution of Washington), Von R. Eshleman (Stanford University), Seymour Hess (Florida State University), A. R. Hibbs (Jet Propulsion Laboratory), Lewis D. Kaplan (Massachusetts Institute of Technology), Yale Mintz (University of California, Los Angeles), E. J. Öpik (University of Maryland and Armagh Observatory), Carl Sagan (University of California, Berkeley), John Strong (Johns Hopkins University), Clyde Tombaugh (New Mexico State University), and A. G. Wilson (RAND Corporation). Invited participants included: Bruce Murray (California Institute of Technology), Bradford Smith (New Mexico State University), and Richard W. Davies, Gary Neugebauer, Ray Newburn, and Fumio Yagi (all of the Jet Propulsion Laboratory). The comments of Joseph W. Chamberlain (University of Chicago) and G. de Vaucouleurs (University of Texas) on the study were of great assistance to the Panel; Dr. Chamberlain also contributed an important appendix elucidating the interpretation of planetary atmospheric spectra. George Derbyshire, of the Space Science Board staff, was Secretary-Rapporteur to the Panel.

The five chapters of the report, prepared by William W. Kellogg and Carl Sagan, constitute an expanded summary of the findings of the Panel. They reflect not only the discussions of the group named above but the considered contributions of many of these scientists, who devoted much time to syntheses and assessments of various topics. The Board is grateful to the participants who made this study possible and to their institutions. For the contributions of Drs. Kellogg and Sagan particular acknowledgment is due the RAND Corporation, which supported their work under a grant from the National Aeronautics and Space Administration.

The Panel could not, of course, resolve many of the controversies over Mars and Venus. The report does summarize, however, the state of our knowledge of the atmospheres of Mars and Venus as of 1961 and sharpens and defines areas of uncertainty. The work done by the Panel, the results of which are set forth here, represents an important step forward in our approach to useful understanding of planetary atmospheres—both in its presentation of accomplishments to date and in its definition of future prospects for discovery and understanding. It points the way toward the important and exciting future of the scientific exploration of our neighboring planets.

Lloyd V. Berkner, Chairman
Hugh Odishaw, Executive Director
Space Science Board
National Academy of Sciences

THE ATMOSPHERES OF MARS AND VENUS

Contents

Chapter I

OBSERVATIONAL METHODS

It is evident that any conclusions to be drawn about the atmospheres of Mars and Venus must be based largely on observations. Theoretical reasoning by itself will not carry us very far, since a planet is a complex entity whose history is virtually unknown. Thus, in the discussions to follow, a great deal of emphasis will be placed on observations. Those that have already been made constitute a maze of apparently unrelated facts. Furthermore, in recent years the tools of the astronomer have become more elaborate, and in order to thread this maze of facts it is essential to understand the principles behind the observational methods.

The purpose of this chapter, then, is to explain the principles underlying each type of observation, so that the reader can more easily follow the syntheses of facts and the discussions of interpretation in Chapters III, IV, and V. To one who is already familiar with the various physical principles involved this chapter will be a general review, and he may wish to skip over to the substance of the report.

A. Surface-Based Observations

1. Visual and Photographic

Certainly visual and photographic observations constitute by far the most usual kind of observation, and the features of the planets have been under such surveillance since the invention of the telescope. One approach to the study of planetary atmospheres is to search for characteristic meteorological phenomena; this approach involves noting and intercomparing the brightness, color, and variation in time of visual features. Appendix 1, by Wilson, discusses in some detail the principle of such observations, and it will here suffice merely to summarize his main points. The reader is also referred to Appendix 2, by Tombaugh, for further insight into some of the problems (as well as results) of visual and photographic studies of the planets.

There are essentially five "dimensions" in the photographic (or visual) process: two angular dimensions of the region being photographed, brightness (or density on the film), the spectral dimension (choice of wavelength and wavelength interval, i.e., "color"), and time. Each of these dimensions has a certain range and a threshold which sets the resolving power. These factors and thresholds are given specifically by Wilson (Appendix 1) for a 60-inch telescope at the Earth's surface.

Perhaps the most important threshold in photographic work is the angular resolving power, since this determines the smallest feature that can be distinguished. In theory, the best angular resolution attainable is that set by the diffraction pattern of the telescope aperture. If a point source (e.g., a star) is used, its image will be a bright central point surrounded by concentric dark and light rings, and the angular radius of the first intensity minimum will be $\rho_o = 1.22\lambda/D$ radians, where λ is the wavelength of the light (between about 0.3 and 1.2μ for photography) and D is the diameter of the aperture (assumed to be unobstructed and of perfect optical quality). For two point sources, it is considered that they can just be resolved when only a slight decrease in intensity occurs between the bright centers of their diffraction patterns, and such a decrease occurs when they are separated by 80 to 85 per cent of ρ_o (v., e.g., Russell, Dugan, and Stewart, 1945). Thus, the separation of two barely resolved images of two point sources is about

$$\varepsilon_o \cong 0.8\ \rho_o \cong \lambda/D \text{ radians.}$$

In practice there are variations, depending on contrast, shape of source, etc., but the above expression represents a good practical limit. The corresponding limiting angular resolutions for perfect telescopes having apertures of various diameters are (for the visible at $\lambda = 0.56\mu$) shown in Table I. In practice, for telescopes with larger than 36-inch apertures it is virtually impossible to achieve the ideal resolution.

Table I

Telescope Diameter	*Ideal Resolution*	
	(radians)	*(seconds of arc)*
6 in.	0.36×10^{-5}	0.72
12 in.	$.18 \times 10^{-5}$	.36
24 in.	$.09 \times 10^{-5}$	.18
36 in.	$.06 \times 10^{-5}$	.12
60 in.	$.036 \times 10^{-5}$	.07
200 in.	$.011 \times 10^{-5}$	.02

We now come to the real limitation on resolution for larger telescopes, that inflicted by the atmosphere of the Earth. Due to changing inhomogeneities in the atmosphere the image is always distorted and to some extent moving. No matter how large the aperture, the practical limit to "seeing" is about one second of arc for photographic work. Experienced observers using telescopes of 12-inch aperture or larger claim that they can see down to 0.1 to 0.3 seconds of arc under conditions of excellent seeing. Tombaugh (Appendix 2) maintains that going to telescopes larger than about 24-inch aperture "runs into diminishing returns rapidly . . . from optical aberrations and atmospheric turbulence." These points will be touched on again in the discussion of balloon and aircraft observations.

Thus the practical limit imposed on Earth-bound telescopes by atmospheric seeing can be interpreted in terms of planetary surface features as follows:

Table II

Angular Resolution (seconds of arc)	*Corresponds to Observations Made . . .*	*Smallest Spot Visible with High Contrast (in km)*		
		*Mars at Opposition**	*Venus at*	
			*Dichotomy***	*Inferior Conjunction****
1	Photographically, from the ground	300-400	900	250
0.1	Visually, under ideal seeing conditions, or: Photographically, from a balloon with an ideal 36-inch telescope	30-40	90	25

* "Opposition": Mars, Earth, and the Sun approximately in a straight line, with the Earth in the middle

** "Dichotomy": The Sun-Venus and Sun-Earth lines make a right angle

*** "Inferior Conjunction": Earth, Venus, and the Sun approximately in a straight line, with Venus in the middle.

In spite of the limits imposed by poor angular resolution it is becoming evident, as Wilson points out, that in order to study the temporal changes on Mars (some of which take place in matters of a few hours) it is necessary to keep a relatively constant watch when Mars is favorably situated. Such a watch involves a coordinated effort by observatories throughout the world, so that a continuous record can be kept regardless of local weather or the Earth's rotation. (v. Chapter V.C.1.)

One further point of interest in connection with general planetary photography is the question of the balance of time vs. space resolution, an important subject when we are concerned with meteorological or other atmospheric studies. Arguing by analogy with the Earth, Wilson points out that the "characteristic times" of such phenomena as dust devils, tornadoes, thunderstorms, and cyclones, increase with the size of the disturbance (v. Fig. 1 in Appendix 1). Thus, with the Earth one would ordinarily not expect to see changes which take place in less than two or three hours when observing with a resolution of 100 km. This concept will have to be borne in mind as we approach the planets and get a chance to take more detailed looks.

There is a special sort of observation that can be made with visible radiation on the rare occasions when a star is occulted by a planet, which did occur on July 7, 1959, as Venus occulted the first-magnitude star Regulus (α Leonis). It can be shown that the rays from the star, essentially perfectly parallel, are bent by the density gradient in the planet's atmosphere, and these same rays are caused to diverge as the second derivative of density with respect to height. The effect of such divergence on the star's radiation is to decrease its flux at the Earth, and there is a point at which the star is seen to disappear as a consequence of this effect. The density gradient of an atmosphere can be related to the scale height, $H = \mathrm{kT}/mg$, where k is Boltzmann's constant, T is the absolute temperature ($^\circ K$), m is the mean molecular mass, and g is the acceleration due to gravity. The second derivative of the density can then be related to the derivative of the scale height. Thus, by observing the time of extinction of the starlight it is possible to determine the scale height at some known level in the atmosphere. (The results from the occultation of Regulus by Venus, observed by Menzel and de Vaucouleurs, 1960, are given in Chapter IV.)

There are several other planetary observables that can be determined visually or photographically, including the size of the planet, its shape or "figure," and the scattering of light in its upper atmosphere as it passes close to the Sun (in the case of Venus only for Earth-based observations). A very important measurement is also the planet's albedo, or reflecting power, at various wavelengths, since this determines how much solar heating occurs.

The polarization of the diffusely reflected light is also significant. In polarization studies the percent of the light that is polarized is plotted against the phase angle, i.e., the angle Sun-planet-Earth. In this way a characteristic curve for the planet or part of the planet is obtained, and attempts are then made to compare the curve with characteristic curves of laboratory samples of surface material or suspended particles, and with theoretical curves. For Venus, phase angles between 0 and 180° are tractable, and the complete polarization curve can be obtained. For Mars, however, the maximum phase angle that can be observed from the vicinity of the Earth is about 43°; most of the night hemisphere of Mars is always invisible from Earth. (An extension of the Martian polarization curve beyond 43° is an example of an experiment which can be done only from a space probe.)

Theoretical analyses of polarization curves for cloudy or hazy atmospheres (taking into account multiple scattering) are just becoming practical, because of the advent of large electronic computers. In the past the standard approach to the interpretation of polarization curves has been empirical. In general it is possible to distinguish between transparent and absorbing particles, and also between particles whose size is much less than the wavelength of light and particles whose size is much larger than the wavelength of light. The smaller particles—in the case of visible light these must be air molecules—are much more efficient polarizers. In principle, polarization data can be used to determine atmospheric pressure (based on total optical depth), the chemical composition and size distribution of clouds or haze, and the nature of the surface material (on Mars, the variously colored areas and polar caps). The potentialities of this method have not been fully explored to date.

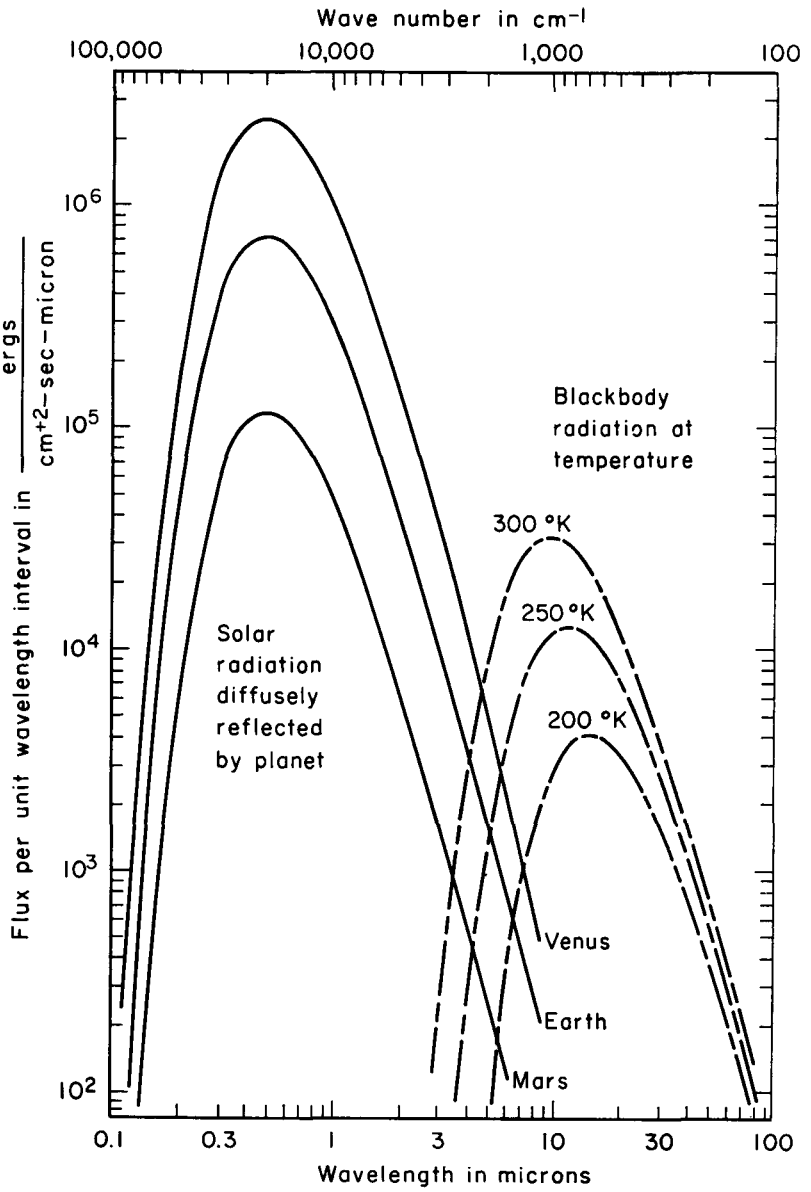


Figure 1

Plot of flux of reflected solar or emitted planetary radiation as a function of wavelength or wave number. The emitted radiation is merely blackbody radiation at the specified temperature (emissivity of one). The reflected or scattered solar radiation takes into account the change of solar constant due to average distance from the Sun, and assumes an albedo (diffuse reflectivity) of 0.7 for Venus, 0.4 for Earth, and 0.15 for Mars. The distribution of the solar spectrum is taken to be that of a 5783°K blackbody. (Calculated by N. T. Divine, RAND Corp.)

2. *Infrared*

When observing any of the terrestrial planets in the visible and *near*-infrared parts of the spectrum one is working entirely with reflected and scattered sunlight; but beyond a certain wavelength the sunlight is overwhelmed by the infrared thermal radiation emitted by the planet itself. The crossover point between reflected and emitted radiation for Mars is at wavelengths from 3 to 5μ, for Earth it is between 4 and 6μ, and for Venus it is between 5 and 7μ.* (v. Fig. 1.) This distinction between reflected (or scattered) solar radiation and planetary thermal emission is an important one, since the interpretation of the spectra obtained in the two regions is quite different.

In the *near* infrared (below about 4μ, say) the spectrum of a planetary atmosphere on the *dark side* will consist of scattered and emitted light, known on Earth as airglow and aurorae. This radiation has never been observed with certainty on Mars or Venus, although a visual phenomenon occasionally reported on Venus as "the ashen glow" may possibly be a faint aurora. However, on the *sunlit side* the near-infrared radiation from the Sun passes down through the atmosphere, is absorbed at certain wavelengths by such molecules as CO_2 and H_2O, is scattered or reflected at the cloud tops or surface, and then passes back out through the atmosphere with further absorption. This emergent radiation is strong enough at certain wavelengths to be detected and analyzed spectrographically at the Earth, and a great deal of our knowledge of the atmospheres of Venus and Mars comes from such spectra.

There are essentially three kinds of deductions to be made from these observations: (a) the presence (or absence) and abundance of molecules with near-infrared absorption spectra, (b) the character of the reflecting layer, be it the surface or the cloud tops, and (c) the temperature at the level where the radiation is backscattered in cases where the scattering occurs in the free atmosphere (e.g., the CO_2 rotational lines in the Venus spectrum at about 8000 Å). The first two kinds of deductions are fairly straightforward, since they involve the identification of gases or solid materials by matching the observed absorption features with known absorption spectra, a very common application of spectroscopy. The last is less easily understood and bears some explanation.

A molecule such as CO_2 or H_2O will have two kinds of discrete energy transitions which result in infrared emission or absorption. The larger energy transitions are those between various vibrational states, and result in a succession of near-infrared spectral bands. Accompanying the vibrational transitions are rotational transitions involving less energy, which in effect cause each vibrational band to be split into a series of lines spaced close together. The relative intensities of these rotational lines are controlled by the gas temperature, since the temperature determines the relative populations in the various rotational energy states. Thus, by measuring the relative intensities of a series of rotational absorption lines one can determine the gas temperature provided energy equipartition applies. (v., e.g., Herzberg, 1950; Chamberlain and Kuiper, 1956.) It must be borne in mind, however, that the temperature so determined is a mean temperature of some atmospheric region or level where most of the absorption took place, and its unambiguous interpretation is difficult. Thus, in the case of Venus there is some argument (as will be pointed out) as to the "level" to which the temperature determined from the 8000 Å bands of CO_2 refers. (v. Appendix 9, by Chamberlain, for a discussion of this problem.)

While in the visible and near infrared the *scattered and reflected* sunlight prevails, in the infrared beyond about 4 to 6μ we are dealing primarily with *emission* from the planet, even on the sunlit side. As shown in Fig. 1, if a planetary surface or atmospheric layer is opaque and has an emissivity of unity it will emit according to Planck's law, and the maximum of this emission is between 10 and 15μ for the temperatures of Earth, Mars, and Venus. A significant factor here is that each of the terrestrial planets contains CO_2 and H_2O in varying amounts, and these gases are effectively opaque at some wavelengths in the far infrared, and are transparent at others. For example, water vapor absorbs very strongly in a band centered at 6.3μ, and the emission from the Earth at 6.3μ comes from the "top of

* In actuality, the situation is somewhat more complicated than this. Strong atmospheric absorption bands will be more effective in cutting out the reflected radiation than in reducing the planetary emission. The result is to shift the cross-over point to somewhat smaller wavelengths.

the water vapor," or from the region of the tropopause. The flux (or specific intensity) of this radiation is then a rough measure of the tropopause temperature (Greenfield and Kellogg, 1960). At 8 to 13μ, on the other hand, the Earth's atmosphere is relatively transparent (except for a narrow absorption band at 9.6μ due to ozone), and this spectral region is known as a "window," since surface radiation can pass through to space. The flux of radiation in a window is a measure of the surface or lower atmosphere temperature (or, perhaps, of the cloud tops). It is only through windows in the terrestrial atmosphere that spectroscopic observations of the planets can be made from the surface of the Earth.

This distinction between atmospheric emission in absorption bands and in windows is a very important one in principle, since it permits a measure of the vertical temperature structure of the atmosphere (Kaplan, 1959; Greenfield and Kellogg, 1960). Unfortunately, this type of measurement has still to be made properly for any of the planets, since to do it well it must be made from *above the Earth's absorbing atmosphere* with telescopes that provide high spectroscopic resolution. Balloons and fly-by vehicles are definitely indicated here, and will be discussed further below.

Although the vertical temperature structure of a planetary atmosphere cannot be measured from the Earth's surface by such infrared techniques, planetary emission in the 8 to 13μ window has been repeatedly used to determine the effective surface (or cloud top) temperature of the planets. After making suitable allowances for the planetary emissivity, residual absorption of infrared radiation in the terrestrial atmosphere and in the detecting system, and the distance of the planet from the Earth, absolute intensity measurements can be performed. The detectors have usually been thermocouples, but more sensitive systems are now coming into use. By comparing the measured flux with the Planck blackbody energy distribution function (as in Fig. 1), the temperature of the emitting body can be determined. A somewhat more reliable method, which to some extent eliminates energy calibration errors, is to compare the flux received at two wavelength intervals, and then to match the relative intensities to the Planck distribution. In the case of Mars, thermocouple bolometry has given temperatures which must be quite close to the true surface temperatures; but for Venus, with its thick cloud cover, the derived temperatures refer to some higher region in the atmosphere.

3. Radio

In the previous section infrared emission from a planet was discussed. Going to still longer wavelengths brings us into the radio spectrum, and the realm of radio astronomy. The large radio telescopes in the United States, Great Britain, Australia, the U.S.S.R., the Netherlands, and in other nations, have detected radio emission from the Moon and the nearer planets.

In the microwave region, from 1 mm to 100 cm wavelength, the emissions from Mars and Venus are probably of thermal origin, and will be our major concern in the present discussion. In the case of Jupiter some radiation at longer wavelengths has also been detected; it is probably not thermal in origin but is thought to originate either from thunderstorm activity on Jupiter or from a Jovian Van Allen belt.

The principles and main results of planetary radio astronomy have been reviewed in Appendix 3, by Burke. The most significant point in connection with Mars and Venus is the fact that, whereas Mars emits radio waves with an effective temperature similar to that deduced from infrared measurements, Venus appears to be considerably "hotter" at centimeter wavelengths than it does from measurements in the infrared. The temperatures deduced from these radio observations suggest a Venus temperature of about 600°K near inferior conjunction. To confuse the issue further, the evidence at 0.8-cm wavelength suggests that the temperature of Venus is several hundred degrees less.

The interpretation of this radiation is the subject of a later chapter, and, as in the case of infrared emission, the important question to be answered is: From which level in the Venus atmosphere do the various wavelengths originate? Certain molecules have absorption features in the microwave region; e.g., H_2O at 0.162 and 1.34 cm, CO at 0.26 cm, and O_2 at 0.5 and 0.25 cm. CO_2

also absorbs to some extent at wavelengths shorter than 1 cm. Thus, the shorter wavelength microwaves may come from the atmosphere if such gases are there in sufficient quantity, and the longer wavelengths may come from near the planet's surface and pass through the atmosphere.

In order to follow the problem, it is important to note that there is another possibility for the origin of the centimeter radio waves from Venus. A very highly ionized and deep ionosphere will absorb and emit radio waves by "free-free" interactions among electrons. Assuming an electron temperature of 600°K and a ground temperature of 265°K, Jones (1961) concludes that the radio observations are consistent with a value of $\int n_e^2\, dz \cong 10^{25}$ cm^{-5} on the unilluminated hemisphere. Here n_e is the electron number density, and the integration is performed over the depth of the ionosphere. The question as to whether an ionosphere on Venus *could* be so highly ionized is discussed below in Chapter IV.

4. Radar

The new science of radar astronomy has only recently reached out to the planet Venus (1958), and at this time no echo has yet been received from Mars or any of the more distant planets. (The first radar return from the Moon was in 1946.) However, its potentialities, which are covered in Appendix 4 by Eshleman, are very great indeed. (v. also Eshleman and Peterson, 1960, and Leadabrand, 1960.)

The most obvious determination by a radar experiment is the *distance* from the Earth to the Moon or planet. This is of value for improving the accuracy of the astronomical unit. Of greater importance to the subject of planetary atmospheres, particularly for Venus, is the ability of radar, in principle, to give the rotation rate and the inclination of the axis of rotation, surface reflectivity and roughness, the heights of scattering cloud layers, and ionospheric characteristics. These remarkable achievements can eventually be made by analyzing the returning pulse in great detail, since, for example, the surface roughness and reflectivity affect the intensity of the return and its degree of polarization, the shape affects the spread in range of the echo, and planetary rotation causes a spread in the frequency of the returning pulse due to Doppler broadening.

B. Planetary Observations from High Altitudes

Sections A.1 and A.2 above described the factors that limit optical observations from the Earth's surface; these factors are mainly the distortion and motion of the image caused by atmospheric turbulence, and the absorption of certain important wavelengths by various atmospheric gases. With regard to the first, distortion and image motion blur a photograph of a planet, and impose a limit of photographic resolution of about one second of arc (v. Table II), regardless of the size of the telescope. A balloon-borne telescope with a perfect 36-inch mirror would improve this resolution by an order of magnitude and permit a resolution of 0.1 second of arc. With regard to the second factor, absorption by ozone draws an impenetrable curtain over the surface for all wavelengths shorter than 3000 Å, and water vapor, ozone, and carbon dioxide combine to make many infrared measurements extremely difficult to interpret because of their complex and overlapping absorption bands.

For these reasons, there have been a number of attempts in the U. S., the U.S.S.R., and France to carry telescopes aloft on balloons, aircraft, rockets, and satellites to make observations from above the main part of the atmosphere. In order to see the radiation from the Sun and stars in the near ultraviolet between 2000 and 3000 Å it is necessary to get above the ozone layer, or above about 50 or 60 km; to see radiation between 1000 and 2000 Å one must climb above most of the atmosphere's molecular oxygen, or above 100 km, and at this level one can also measure X-ray emission from the Sun and stars; and in the far ultraviolet between about 100 and 1000 Å, one must climb still higher, to about 200 km, to observe the radiation from the Sun and stars. However, the planets presumably emit relatively little ultraviolet radiation and the scattered solar ultraviolet radiation has low intensity, so to observe them clearly in the ultraviolet it is not enough to get above the atmosphere—one must get *closer*, with an instrument in a space probe. (This is the subject of Section C, which follows.)

To observe the planets in the visible and infrared, the altitude requirements are much less stringent than for ultraviolet measurements, since the absorption and distortions are greatly reduced at altitudes attainable by balloons and high-altitude aircraft such as the U–2 and the X–15. The technical considerations involved in balloon and aircraft observations are covered in detail by Strong in Appendix 5; some of his main conclusions are summarized here.

First, there are clearly operational problems involved in flying an instrument such as a telescope-spectograph combination. Aircraft are convenient, since they can take off without extensive flight preparations and recovery is no problem. However, an aircraft has some troublesome vibration, and a somewhat limited flight duration. Balloon flying is a complex operation requiring considerable preparation, tracking and recovery crews, and some hazard when the man accompanies his instrument. However, a balloon can provide a very stable platform, it can stay up for long periods, and it can go somewhat higher than an aircraft with a heavy load of instruments. Furthermore, Strong points out that at 80,000 feet or so the sky is dark and the planets are good subjects even in the daytime; accordingly, he notes, high-altitude astronomy can be carried out by day whereas at the Earth's surface astronomy is restricted to the nighttime.

Infrared observations of the planets from a balloon or aircraft with a 12-inch telescope-spectrometer "can yield qualitatively superior spectroscopic results—results that are completely new and utterly unattainable with the most powerful ground observatories ever conceived." Such observations would, of course, not have the spectral resolution attainable with a much larger telescope on the ground (because of the latter's greater light-gathering power), but the advantage would be the greatly reduced telluric absorption in the light path. Strong describes in some detail how one could measure the infrared spectrum from Mars or Venus with such a 12-inch instrument with a spectral resolution of about 2μ at long infrared wavelengths, and with a considerably better resolution at shorter infrared wavelengths.

Plans are already under way at Princeton under the direction of M. Schwarzschild to fly a 36-inch telescope on a balloon, improving the light-gathering power by the square of the ratio of diameters, or a factor of about 9 over the previously cited 12-inch instrument. The primary goal of this 36-inch balloon telescope is high resolution photography of the planets and other astronomical objects (cf. Table II). But its existence also suggests that in the near future we may be able to obtain from the Earth infrared spectra of the nearer planets with better than 0.1μ resolution. These measurements would be of the total radiation from the entire disk, and one could trade spectral resolution for spatial resolution and observe only a part of the planet at a time.

Such spectra would give a great deal more quantitative information about the composition and temperature structure of planetary atmospheres, as pointed out in Section A.2 above.

C. Planetary Observations from Fly-By Space Probes

The ultimate uses of space probes to observe the planets are virtually limitless. In fact, it is fair to say that there are many questions concerning the planets that can *only* be resolved by landing on the planets, exploring their surfaces, sampling and analyzing and searching for the new and unexpected features that are bound to exist on those remote worlds. However, there will probably be a first generation of space probes that will gather most of their data by simply passing by the planet and observing indirectly, and the purpose of this section is to enumerate briefly some of these not-so-remote planetary observations. Later chapters and appendices will return to this subject in more detail.

The great advantage of a fly-by probe is proximity. Proximity permits observations of radiation and fields that are too weak to measure from the Earth, and also permits a greater resolution of the planetary disk with a given size of instrument. Thus, most of the measurements that have been described in the previous sections can be done with more refinement from a space probe, weight and power and telemetry permitting. An additional measurement which can best be made from a probe is that of the planet's magnetic field (and its radiation belt, if a magnetic field exists).

Here, then, in outline form are the planetary observations that can be made from fly-by space

probes; each of these observations would contribute to our knowledge of the planet's atmosphere. Some of the technical problems are touched on briefly, but no attempt will be made to indicate which experiments seem to be the more crucial; this discussion can be found in later chapters and appendices, where the experiments are considered in the light of the various competing theories about the atmospheres of Mars and Venus.

It is pertinent to note the specific experiments that are now being planned for the first two United States planetary-probe programs, Mariner A (to Venus) and Mariner B (to Mars). The tentative programs for these two vehicles are described by Davies, Hibbs, Neugebauer, and Newburn in Appendix 6 and the main planetary experiments are there described. Also in Appendix 6 is a table showing "Planet Availability," i.e., the dates (through 1971) when the planets will be in favorable positions for a space-probe launching and when the subsequent encounters would take place. It will be noted that Venus can be approached from the Earth once in about every 18 months, and Mars once in about every two years.

1. Optical Measurements

a. *Pictures in Visible Light*

It is obvious that, as one approaches a planet with a given telescope, the pictures that one could take would become more and more detailed. From a few hundred kilometers altitude one could take a picture with a resolution of only a few meters, since one second of arc (attainable with a 6-inch telescope—see Table I) at 500 km would subtend 2.5 m.

In practice, however, it turns out that the problem of picture-taking from a planetary probe is not one of optics but of communication, since it requires such a volume of information to convey even one high-resolution picture. At planetary distances of several hundred million kilometers or more, with a directional antenna on the space probe as well as on the ground and using a 10-watt transmitter, one could have a signal-to-noise ratio of about 10 db at the distance of Venus. Engineering estimates, assuming some safety factors in order to assure receipt of the telemetry, predict about 10^6 bits of information to be available during and after the Venus fly-by with such a transmitter. (Hibbs, 1959.) Appendix 6 suggests that communicating with the Mars probe will have the same limitation, since the bandwidth for Mariner B is expected to be about 100 cycles per second, so roughly three hours must be allowed to send one million bits. (A discussion of the fundamental technical problems in space communication has been given by Rechtin, 1961.) For the sake of this discussion, we may compare this nominal one million bits of information for *all* the experiments with the number of bits required to send one picture of a given quality, as given in Table III.

Table III

Kind of Picture	*Resolution (and Picture Size)*	*Number of Picture Elements*	*Number of Bits with 16 Shades of Grey*
Ordinary TV	600 lines	3.6×10^5	1.4×10^6
Aerial photo (ordinary film)	20 lines/mm (10×10 cm)	4×10^6	1.6×10^7
Aerial photo (highest quality)	100 lines/mm (10×10 cm)	10^8	4×10^8
Picture of Mars from a balloon-borne 36-inch telescope*	0.1″ of arc, or 30-40 km on Mars at opposition (see Table II)*	4×10^4	1.6×10^5

* Assume a square picture just containing the disk of Mars at opposition, and a film grain that takes full advantage of the optical resolution.

In view of the practical constraints imposed by the first-generation deep-space communication systems that could be flown in the near future, it is concluded that one could not send a picture from Mars (or Venus) with even the quality of a good TV set, i.e., one could not send a picture like those from the TIROS meteorological satellites. Devoting the entire capacity of the telemetry to it, however, one could send a low-resolution picture of the entire disk of the planet that would have about 2.5 times better resolution than the best picture obtainable from a balloon with a perfect 36-inch aperture. Furthermore, by limiting the field of view one could, of course, get higher resolution at the expense of coverage of the entire disk. For example, if one wanted to see a specified part at a resolution of ten meters, one could transmit a picture of an area on the planet up to 5 km on a side using about 10^6 bits (with 16 shades of grey).

Presumably the state-of-the-art of space communications and power supplies will improve, and larger bandwidths and volumes of information will be available, so the limit of 10^6 bits is not a permanent one. An increase by a factor of 100 in transmitter power, for example, will increase the bandwidth by a factor of 100, all other things being equal. Rechtin (1961) describes a possible communication system for a Mars orbiter with a usable bandwidth of 2.5×10^3 cycles per second. It will therefore be possible, some day, to send a series of pictures from Mars or Venus that is vastly superior to anything obtainable from the Earth.

b. *Infrared Measurements.*

The significance of infrared measurements to understanding a planetary atmosphere has already been stressed in Sections A and B, above. The advantages of infrared observations from a probe lie in the increase in flux that one obtains, and this increase permits a greater spatial and spectral resolution.

It was mentioned that a 12-inch telescope on a balloon could obtain a far-infrared planetary spectrum with better than 2μ spectral resolution. Kaplan (1959) discusses infrared spectrometer requirements for satellite (or balloon) observations of Earth and states that a spectral resolution at 15μ of better than 0.1μ is feasible in a flight instrument, which allows the strong absorption band of CO_2 at 15μ to be analyzed. With this kind of resolution, as shown by Kaplan, one can infer a great deal about the temperature structure of an atmosphere in which the carbon dioxide mixing ratio is constant, which is true for the Earth and is probably so for Mars and Venus as well. In the case of Venus, the larger amount of CO_2 in its atmosphere will result in less stringent resolution requirements for observations made in the vicinity of that planet.

Taking advantage of the increased flux to obtain spatial resolution in the infrared, one could make the kind of infrared intensity maps at various wavelengths that are being made by TIROS II and its successors, maps that reveal the air mass temperature distributions with a resolution on the ground of about 50 km. One could also combine the two approaches, and take separate spectra of different parts of the disk, a feat which has been done from the Earth by Sinton and Strong (1960) with the 200-inch telescope at Mt. Palomar, but which could be done with much more precision and resolution from a nearby space probe. An instrument to make this kind of infrared combined spectral and spatial measurement from a satellite or probe is already being developed under the sponsorship of the Jet Propulsion Laboratory, and it is planned to fly a version of it in Mariner B to Mars (v. Appendix 6). Such a device could obtain very important information on the structure and composition of the Martian atmosphere, and on the presence and distribution of organic molecules on the Martian surface.

c. *Ultraviolet Measurements*

To measure the ultraviolet radiation from another planet, the need for a close approach is even more clear than in the case of the infrared, since (a) the source is very weak, and (b) to go beyond 2000 Å in the ultraviolet requires getting above most of the atmosphere anyway, as discussed in Section B.1. Here balloons will not do. Kaplan (1961), Singer and Wentworth (1957), and others have shown some of the deductions that can be made from satellite or fly-by probe altitudes, and these are briefly summarized.

Ozone has a set of strong absorption bands between 2000 and 3000 Å, with the result that sunlight in this wavelength region will penetrate to a depth determined by the amount of ozone; the intensity of sunlight that is scattered back out of the atmosphere will be proportional to the effective penetration. This means that a sampling of scattered sunlight as a function either of wavelength or of zenith angle will reveal the distribution of ozone in the atmosphere—provided there is more than a trace amount above the highest clouds. The amount of ozone in an atmosphere is directly related in a calculable way to the concentration of free oxygen (and to a lesser extent to oxides such as CO_2), so a measurement in this manner of the ozone content is of great value in determinating the amount of O_2 in the atmosphere. Furthermore, the presence or absence of ozone plays an important role in determining the flux of ultraviolet light reaching the ground, a matter of considerable significance to considerations of life on a planet.

Although we cannot make such far-ultraviolet observations from the ground, our own upper atmosphere emits in the ultraviolet due to bombardment by charged particles from above and to photochemical reactions. These ultraviolet auroral and airglow emissions are expected to be sufficiently strong to be detectable from a satellite or nearby probe. An analysis of the emission spectrum from the dark side of a planet will indicate the relative amounts of various constituents in the upper atmosphere, and from this it may be possible to deduce the constituents in the main body of the planet's atmosphere. Further, since the flux of charged particles causing the auroral excitation will be determined by the magnetic field, the topographic distribution of the emission will reveal the shape and intensity of the magnetic field. (In the case of Earth, a rather definite maximum in auroral excitation occurs in the "auroral zones," zones that encircle each of the magnetic poles with widths of about 18° of latitude.)

It is planned to fly an instrument to observe the ultraviolet light from Venus on Mariner A. This instrument is expected to scan the disk and obtain spectra from 1100 Å to 4200 Å, with a spectral resolution of 10 Å. (v. Appendix 6 for further details.)

d. *Photometric and Polarization Measurements*

The intensity of sunlight diffusely reflected from each point on a planetary disk is the sum of the light reflected from the surface (or upper cloud decks), light scattered by the molecules of the atmosphere (Rayleigh scattering), and the light scattered by particles of dust or thin clouds (Mie scattering). Looking in just one direction, or considering the light from the entire disk, it would ordinarily be impossible to determine which fraction came from which scattering source. However, by mapping the intensity of the light from various parts of the disk at several wavelengths, and also by measuring the polarization, it is possible to fit unambiguously a model of the surface and atmosphere. Sekera and Viezee (1961) have, for example, computed theoretically, taking into account secondary scattering, the intensity and polarization distribution from a variety of hypothetical molecular atmospheres of various depths and surface albedos. It would be possible to match an observed pattern to these computed patterns, and thereby determine the model that best fitted the observations in terms of depth, albedo, and (ultimately) dust content. Dollfus (1957) has obtained such polarization measurements from the Earth, but the resolution was low and the range of phases available for Mars necessarily limited. This could certainly be done better from a probe for both Mars and Venus.

Knowledge of the atmospheric characteristics is not the only parameter to be gained from photometric and polarimetric observations, since, in the case of Mars, something can also be learned from them about the surface. For example, a change in the polarization and intensity of the reflected light with the seasons gives an important clue to the nature of the seasonal changes on the surface of Mars, and indicates whether the color change is accompanied by a change in the sizes of the surface particles or surface roughness (cf. Chapter III. F).

One further potentially valuable measurement of this nature consists in scanning the terminator, or twilight zone. It is evident on Earth that there is scattered sunlight in the sky after sunset, and seen from above this would appear as a gradual shading from the daylight to the darkened hemisphere.

The exact way in which this fading takes place at various wavelengths is determined by the depth and density distribution of the atmosphere and by dust or high clouds. By making polarization measurements at the same time as the intensity distribution is observed, the effects of the dust or clouds can be separated to some extent. Thus, a high-resolution photometric and polarimetric scan of the terminator is another powerful way to obtain atmospheric characteristics from a fly-by probe.

2. *Radio and Radar Measurements*

In Sections A.3 and A.4 the observations that were possible from Earth using active (radar) and passive (radiometric) radio techniques were described. Even the largest radio telescope now in existence cannot resolve the disks of Mars or Venus in angle, since the antenna beamwidths are far too great. However, as one gets closer one could begin to scan the planetary disk with even a moderate sized antenna (say 1 meter in diameter), and thereby obtain variations of signal from point to point on the disk. This would be particularly interesting for Venus, where only radio waves can penetrate the atmosphere and cloud layers and allow us to "see" the surface. As will be pointed out in Chapter IV, the passive radio observations of Venus from Earth present one of the most intriguing scientific puzzles of our time; one good scan across the disk of Venus would resolve a great part of the present controversy. Since we can reach to the surface of Mars more conveniently with passive optical devices working in the visible and infrared, it does not appear as worthwhile to measure the radio emissions from Mars as it is in the case of Venus.

Similar conclusions can be drawn about the use of active radar from a probe. The surface of Venus could be studied in some detail by radar, giving a map of reflectivities of the cloud-covered surface. At shorter wavelengths (less than 1 cm) the clouds would also give some back-scatter, and so with an assortment of long and short wavelengths it might be possible to measure the depth and distribution of the cloud layers by radar. (The possibility of performing this measurement from Earth is discussed below.) For Mars it is not clear that much scientific information could be gained by a probe-borne radar which would not be obtainable more simply by optical means.

Another point of concern is the alternative of performing from the Earth radar experiments that are sufficiently sophisticated to give essentially the same results as a probe-borne radar. In Section A.4 the potentials of Earth-based radar were briefly reviewed, and Appendix 4 by Eshleman goes into the question in more detail. According to Eshleman (Appendix 4, and Eshleman and Peterson, 1960), in addition to the rotation rate, one should be able to determine the radius of the solid planet, and study its surface roughness and large-scale surface features (e.g., plains and mountain ranges), and its ionospheric characteristics. This potential of high-power, high-gain surface-based radars should cast some doubt on the urgency of developing probe-borne radars.

D. Penetration of Planetary Atmospheres

It is certain that probes can be developed to put instruments physically into the atmospheres of Mars and Venus and to land on the surfaces of these planets. The scientific potential of such a probe is boundless, since one can imagine answering almost any question about a planet by properly instrumenting such a vehicle—including the question of the occurrence of life.

In view of the vast assortment of possible experiments that could be done with advanced entry probes, we will defer a discussion of the choice between experiments until the last chapter, where they can be treated in the light of our uncertainties about the planets. However, not that Mariner B to Mars (v. Appendix 6) anticipates an entry capsule from which pictures may be taken and which will make direct measurements of atmospheric temperature and pressure during descent.

References

Chamberlain, J. W., and G. P. Kuiper, Rotational temperature and phase variation of carbon dioxide bands of Venus, *Astrophys. J., 124*: 399, 1956.

Dollfus, A., Étude des planètes par la polarisation de leur lumière, *Ann. d'Astrophys.*, Supp. No. 4, 1957.

Eshleman, V. R., and A. M. Peterson, Radar astronomy, *Sci. Amer., 203*, (2): 50, 1960.

Greenfield, S. M., and W. W. Kellogg, Calculations of atmospheric infrared radiation as seen from a meteorological satellite, *J. Meteorol., 17*: 283, 1960.

Herzberg, C., Molecular Spectra and Molecular Structure, I. Spectral of Diatomic Molecules, van Nostrand, 1950. (See Chap. III, p. 124.)

Hibbs, A. R., ed. Exploration of the moon, the planets, and interplanetary space, Jet Propulsion Lab. Rept. No. 30-1, 1959.

Jones, D. E., The microwave temperature of Venus, *Planetary and Space Sci., 5*: 166, 1961.

Kaplan, L. D., Inference of atmospheric structure from remote radiation measurements, *J. Opt. Soc. Am., 49*: 1004, 1959.

Kaplan, L. D., On the determination of upper atmosphere composition from satellite measurements, *J. Quant. Spectrosc. and Rad. Transfer*, to be published, 1961.

Leadabrand, R. L., Radar astronomy symposium report, *J. Geophys. Research 65*: 1103, 1960.

Menzel, D. H., and G. de Vaucouleurs, *Astron. J., 65*: 351, 1960.

Rechtin, E., Deep-space communications, Part I - Feasibility, Part II - Design considerations, *Astronautics, 6*, (4): 37 *et seq.*, and *6*, (6): 26 *et seq.*, 1961.

Russell, H. N., R. S. Dugan, and J. Q. Stewart, *Astronomy*, Ginn and Co., Boston, 1945.

Sekera, Z., and W. Viezee, Distribution of the intensity and polarization of the diffusely reflected light over the planetary disk, RAND Corporation Report (in press), 1961.

Singer, S. F., and R. C. Wentworth, Determination of ozone distribution from a satellite, *J. Geophys. Research, 62*: 299, 1957.

Sinton, W. M., and J. Strong, Radiometric observations of Mars, *Astrophys. J., 131*: 459, 1960.

Chapter II

THE GENERAL CIRCULATION OF PLANETARY ATMOSPHERES

A. Introduction

The material in this chapter is based primarily on ideas developed by Mintz, contained in Appendix 8. In fact, it is essentially an extended summary of the material in the contribution by Mintz, and the interested reader is referred to the more complete treatment and the references, especially the textbook of Thompson (1961).

It is clear that any atmosphere exposed to sunlight and radiating into space will act as a heat engine. There are two important characteristics of any heat engine: A heat engine converts heat energy to kinetic energy, and does so by heating its working fluid at a high pressure and cooling the fluid at a low pressure; the other characteristic of a heat engine is that it transports heat from the hot source to the cold sink, and in a planetary atmosphere this corresponds to transporting heat from the part of the atmosphere that is most warmed to the part of the atmosphere that is most cooled by radiation.

In order to visualize how an atmosphere behaves as a heat engine, consider the simple case in which a rotating planet is very slightly heated at the equator. Under these gentle conditions the air over the equator will be heated and the isobaric layers will be inflated, producing a pressure gradient force that will make the air flow towards the pole at a high altitude and sink over the poles. The rotation of the planet will, furthermore, produce a Coriolis force on the moving fluid, and this will have the effect of deflecting the purely north-south flow in such a sense that in both hemispheres the poleward flow will also have a west-to-east component (west winds). The picture is therefore that of an atmosphere spiraling cyclonically towards the poles aloft and spiraling away from the poles anticyclonically. (Actually, due to frictional stresses near the ground the easterly winds in the lower part of this system are much weaker than the upper westerlies.) This situation is shown diagrammatically in Fig. 1 of Appendix 8. As long as the heating is steady and gentle, the circulation will follow this pattern without change. Such a circulation, one which can be described by numerical models of the atmosphere or reproduced in a laboratory tank or "dishpan," is called a *symmetrical circulation regime* (sometimes referred to as a "Hadley regime" or, more correctly, a "*lower* Hadley regime." There is an "*upper* Hadley regime" that is achieved in dishpan experiments under conditions of low rotation and/or very large temperature gradients, but this regime is probably not pertinent to the atmospheres of Earth or Mars).

Now consider the same rotating planet and an increased application of heat at the equator and cooling at the pole. The requirement for the atmospheric heat engine to transport heat from equator to pole is now increased, and there comes a point when the symmetrical circulation can not keep pace with the application of heat. At this point the circulation begins to break down into a more irregular circulation pattern and waves begin to form in the atmosphere. At this point the atmosphere has become dynamically unstable. Figure 2, Appendix 8, shows diagrammatically an example of the unstable or *wave regime*. (The unstable regime is often referred to as the "Rossby regime.")

The unstable regime, because of the waves which are able to rapidly carry hot air poleward and cold air equatorward, is in a sense a more efficient heat engine, and as the heat load is increased the atmospheric waves deepen in order to keep up with the transport of heat. The theory developed by Mintz and others predicts, in terms of a simple dynamical model of the atmosphere (a two-layer model),

the number of waves that should be expected to develop under a given set of conditions. However, a much more complex theory involving non-linear equations that must be integrated numerically is required to describe the more detailed conditions in a wave regime, a point to which we will return.

This highly oversimplified description of the principles involved in the atmospheric heat engine is elaborated quantitatively in Appendix 8. If one keeps these general concepts in mind, however, it should be possible to understand intuitively the behavior of various atmospheres.

B. Circulation of the Earth's Atmosphere

Clearly, there are three important factors which determine the character of the atmospheric circulation: the heat which must be transported from equator to pole, referred to here as ΔQ; the rate of rotation of the planet; and the temperature and stability of the atmosphere (an important factor in the case of the atmosphere of Venus where there is considerable uncertainty concerning its thermal structure and depth).

One other factor that must be considered in connection with the circulation is the large scale eddy viscosity of the atmosphere, but for the sake of simplicity this is assumed to have the same value for all the atmospheres.

The flux of heat, ΔQ, that must be carried from equator to pole, as mentioned above, is determined by the difference between the net heat applied at the equator and the net heat that is lost at the pole. The heat applied is just that due to solar radiation; it can be easily calculated, since it is the product of the integrated solar flux at a given latitude over the period of a day, S, times the part of this flux that is absorbed on the average, or 1—A, where A is the average albedo for a given latitude. This is shown plotted in Fig. 5 of Appendix 8. The *net* heat applied at a given latitude is the product S(1—A) *minus* the heat lost by outward infrared radiation, W. An atmosphere loses heat by infrared radiation at all latitudes, and the rate of loss of heat is proportional to the 4th power of the temperature at the effective radiating level. In the case of the Earth, some 80 per cent of the outward radiation comes from the atmosphere itself; due to the fact that the *water vapor* is the primary source of radiation in the atmosphere and is a variable constituent, it turns out that the radiation from the Earth, W, is not a strong function of latitude, as shown schematically in Fig. 5. Put another way, the level from which most of the outgoing radiation originates in the tropics is higher and almost as cold as the level from which the radiation originates in the polar regions, and therefore the effective radiation temperature is relatively insensitive to latitude. (As will be shown, this is quite different from Mars.) The result is that on the average for the Earth the amount of heat that must be transported from equator to pole is about 8 times more than a symmetrical regime can handle. Therefore, as we know, the Earth's atmosphere breaks down into waves at all times of the year.

An important question concerning an atmosphere in a wave regime is the number of waves that exist on the average (by number of waves is meant the number of pairs of troughs and ridges encircling a hemisphere in middle latitudes). The theory tells us that the number of waves is a function of ΔQ, the static stability of the atmosphere, the rotation rate, and the mean temperature. The average number of waves predicted for the Earth's atmosphere is 6, and this is in rough agreement with the real case, as can be seen by looking at a series of weather maps.

In order actually to reproduce the patterns of highs and lows which exist in an atmosphere such as the Earth's, one must resort to an application of the differential equations of motion in a non-linear form. The integration of these equations, known to meteorologists as the construction of numerical circulation models, requires the use of high-speed computers to describe the motion of the atmosphere as it develops. Numerical models have been constructed which can in fact reproduce rather well the observed patterns of semi-permanent highs and lows in the surface pressure field; it can be seen how the low centers tend to move towards the north, and the high centers move towards the south, creating in the lower levels a broad zone of westerly winds at middle latitudes with easterly winds in equatorial

and polar regions. The results of one such numerical exercise is given in Fig. 7, Appendix 8. Although this does not look exactly like a real weather map of the Earth, the general character of the terrestrial atmospheric circulation is seen to be reproduced.

C. The Circulation of Mars

Turning to the probable circulation on the planet Mars, the first thing to consider is the heat, ΔQ, that must be transported from equator to pole. (Recall that Mars has almost exactly the same rotation rate as the Earth, and an atmospheric pressure about 1/10th that of the Earth, i.e., a relatively thin atmosphere.) Mars has virtually no water vapor, and what carbon dioxide exists to make the atmosphere absorptive in the infrared is presumably evenly mixed throughout the atmosphere. Therefore, the surface temperature determines to a large extent the amount of heat radiated from the planet at each latitude, since in the far infrared such a dry atmosphere is 70 to 75 per cent transparent. For this reason, the amount of heat, W, radiated from the equator is very much more than the amount of heat radiated from the poles, and therefore the net gain of heat at the equator and loss of heat at the poles is greatly reduced. Speaking in terms of the atmospheric heat engine, one can say that the heat load, ΔQ, on the Martian atmosphere is relatively light compared to that on the Earth.

Making certain assumptions about the mean stability and viscosity in the Martian atmosphere (actually, assuming that they are the same as on Earth, for lack of better indications), Mintz concludes that the ΔQ on Mars is slightly smaller on the average than the critical value. Thus, *in the mean for the year* the Martian atmosphere is in a symmetrical regime, according to the theory.

There are several interesting sidelights to the calculations of Mintz for Mars. The calculated mean temperature for the planet turns out to be —54°C, and the temperature difference between equator and pole on the average is 42°C. It will be noted that the mean temperature calculated here is considerably less than that indicated by the bolometric data, which are obtained near opposition and therefore refer primarily to the daytime hemisphere (v., e.g., Gifford, 1956). (v. also Chapter III.) However, the calculated mean temperature is in quite good agreement with the radiometric observations of Mars of —62°C $\pm$ 28° (Giordmaine *et al.*, 1959), suggesting that the radiometric observations refer to an average temperature over the disk of a layer far enough below the surface so that it is not greatly affected by the daytime heating.

The previous remarks refer to the *average* for the Martian year. The seasonal change in ΔQ must be more on Mars than on Earth, due to the small heat capacity of the Martian surface which results in an extreme cooling at the poles in winter and presumably a warming of the poles in summer. (Actually, no good bolometric temperatures for the Martian poles have been obtained to date.) For this reason, in the summer hemisphere the ΔQ becomes small, or may even reverse itself, and in winter it reaches its maximum. Therefore the theory would seem to predict that in winter the circulation may become unstable and produce waves in the general circulation. This prediction is in agreement with the observation that in the winter hemisphere (and at the equinoxes) there are occasionally clouds which move from west to east. Hess (1950), one of the few meteorologists to study Martian cloud phenomena carefully, saw indications of moving storm systems in the form of dust clouds in both northern and southern hemispheres during the northern hemisphere winter. This is in apparent disagreement with the theory, which would suggest that storms would only be in the winter hemisphere. However, theory and observation do agree on the preferred number of waves for the Martian unstable regime, which is 3. Each of the three waves on the smaller planet Mars has the same wavelength as on Earth, which has a preferred number of 6.

Mintz suggests that the symmetrical regime would have lower surface winds (about 1 m/sec) than the wave regime (perhaps reaching 10 m/sec in parts of a cyclonic system). It has also been noted that on Earth a wind speed of about 6 m/sec is required to raise dust from the ground, suggesting that dust storms on Mars would be much more likely under a wave regime, i.e., in the winter hemisphere, as

mentioned previously. Upper air winds are generally larger than surface winds in both regimes, so the rate of transport of clouds that reach into the upper levels may be much greater than the speeds mentioned above.

However, at this point we should depart from theoretical reasoning and review some of the observational evidence: Martian dust storms (see Chapter III) are characteristically initiated in small areas, and tend subsequently to expand in area as they progress, sometimes covering most of the planet, as occurred during the 1956 opposition; the most frequent time for dust storms (Focas, 1961) appears to be at perihelion, which is just prior to the summer solstice in the southern hemisphere; dust storms and moving clouds have been observed on occasion in all seasons of the Martian year; finally, most experienced observers of Mars agree that local topography, as seen by the dark and light areas, plays an important role in the formation and movements of the Martian clouds.

It seems that at present our theoretical tools for predicting the state of the Martian weather and climate are still crude, though useful and highly suggestive, and our observations from Earth permit us to draw but a hazy picture of the real situation. There will probably be time to improve both before the first close look at Mars from a space probe.

D. The Circulation of Venus

As will be shown in Chapter IV, there is considerable doubt about the character of the Venus atmosphere and about its rotation. For this reason it is impossible to treat the circulation of Venus with as much assurance as for Mars. One of the main difficulties in the case of Venus is the uncertainty with regard to the depth of atmosphere, since one school of thought assigns a rather deep atmosphere to Venus while another assigns an atmosphere even less massive than that of the Earth.

The indications are that Venus rotates very slowly, and it is probably safe to say that it has a rotation period of greater than 5 days. According to the theory of Mintz, noting the fact that the effective infrared emission temperature of Venus is virtually constant over the entire disk at a value of about 235°K, it can be said that for even a thin atmosphere and such a slow rotation rate the atmosphere would be in a symmetrical regime. The ΔQ for Venus is somewhere intermediate between the values for Earth and for Mars.

Mintz warns against the application of his theory directly to the case of the deep Venus atmosphere, the reason being that it is assumed in the theory that the lapse rate is constant and therefore that any temperature deviation (but not necessarily the pressure deviation) has a vertical axis. For a very deep atmosphere this would probably not be the case, and therefore the simple theory would not apply.

However, one can say intuitively that as the rotation of a planet slows and the atmosphere becomes deeper, the character of the circulation would change from one in which the heat source was the equator and the heat sink the poles to a circulation in which the heat source was the sub-solar point and the heat sink the anti-solar (or midnight) point. If the rotation period of Venus is about ten days, as the Soviet radar results indicate (v. Chapter IV), this transition probably does not take place, but there might be a large diurnal thermal tidal circulation superimposed on the symmetric meridional circulation. However, if the U. S. results are correct and the period is greater than 100 days, then almost certainly the main circulation would be from sub-solar to anti-solar point.

Mintz, in a previous paper (1961), describes a possible way in which a symmetrical regime centered on the sub-solar point in lower levels could change to a symmetrical regime about the poles at higher levels. This suggestion has not been developed to the point where such a possibility can be studied quantitatively, but a transition of this sort is extremely suggestive in view of the fact that Dollfus (1955) sees in the visible a radial cloud pattern centered on the sub-solar point, whereas Ross (1928) and others have identified in the ultraviolet more or less straight and parallel cloud bands. (v. also Appendix 2, by Tombaugh.) It is possible that the different appearance in the visible and

ultraviolet may result from the fact that different levels are being observed. This matter is discussed further in Chapter IV in connection with the various theories concerning the Venus atmosphere.

References

Dollfus, A., Étude visuelle et photographique de l'atmosphère de Venus, *L'Astronomie, 69*: 413, 1955.

Focas, J., private communication, 1961.

Gifford, F., Jr., The surface-temperature climate of Mars, *Astrophys. J., 123*: 154, 1956.

Giordmaine, J. A., L. E. Alsop, C. H. Townes, and C. H. Mayer, Observations of Jupiter and Mars at 3 cm. wavelength, *Astron. J., 64*: 332, 1959.

Hess, S. L., Some aspects of the meteorology of Mars, *J. Meteorol., 7*: 1, 1950.

Mintz, Y., Temperature and circulation of the Venus atmosphere, *Planetary and Space Sci., 5*: 141, 1961.

Ross, F. E., Photographs of Venus, *Astrophys. J., 68:* 57, 1928.

Thompson, P., *Introduction to Numerical Weather Analysis and Forecasting,* Academic Press, New York, 1961.

Chapter III

MARS

A. Introduction

Of the other planets in our solar system, none has held more widespread interest than Mars. It has been the object of extensive speculation, both scientific and popular, and has been more exhaustively observed than any other planet. The reasons for this appeal are apparent: the physical conditions on Mars are, in many respects, similar to those on Earth; and the surface features on Mars can be observed in greater detail than can those of any planet besides the Earth. Speculation has centered on the possibility of life on Mars, and it is clear that the discovery of indigenous Martian organisms would be of the greatest scientific and philosophical significance. But, even if Mars is a lifeless planet, there remain many problems of high interest concerning, for example, the topography and chemistry of its surface, and the physics and meteorology of its atmosphere. If the surface temperature of Venus is 600°K or higher, Mars may be the only other planet on which we can easily land instrumented packages and scientific exploration teams to gather further evidence on the origin and development of the solar system.

We begin with a brief summary of existing knowledge of the Martian physical environment; a more extensive discussion has been made by de Vaucouleurs (1954). We then describe in somewhat more detail four problems that the Panel found of some interest, *viz.*, the structure and photochemistry of the atmosphere, the nature of the blue haze, and the possibility of life on Mars.

B. Summary of Pertinent Observations of Mars and Their Interpretation

1. *Visual and Photographic Detail*

A superficial telescopic view of Mars, even with a large telescope, is generally disappointing. A crude separation of surface features can be made: there are dark areas, predominantly colored neutral gray,* which are localized near the equatorial regions; there are orange ochre or buff-colored bright areas which cover most of the remainder of the disk; and there is a white polar cap. Buff-colored clouds are sometimes seen, originating in the bright areas and later temporarily obscuring some dark areas. This occurrence has led to the generally accepted identification of the bright areas as extensive deserts. The general form and orientation of the features in the dark and bright areas are preponderantly not time-variable. By following the motion of a given surface feature the period of rotation has been determined to high accuracy; it is roughly 24 hours, 37 minutes. The polar caps are time-variable; they diminish in area as local summer approaches and usually disappear entirely before they are reformed the following fall and winter. Mars generally has a polar cap at all times in one hemisphere or the other. Observations of some finer details, and the questions of the nature of the dark areas and the polar caps, will be discussed in later sections. Reference is also made to the list by Wilson of photographically observable Martian phenomena in Appendix 1, and to the discussion of visual coloration by Tombaugh in Appendix 2.

2. *Clouds and Haze*

The fact that the appearance of Mars as seen at various wavelengths changes from day to day—even from hour to hour at times—suggests that the atmosphere of Mars contains changing clouds or

* Considerable debate still exists over the true color of the dark areas. Some observers report vivid greens and blues (v., e.g., Appendix 2, by Tombaugh). But, at least in recent years, most observers report primarily gray coloration (v., e.g., Kuiper, 1957).

layers of aerosols. This should occasion no surprise, since it is clear that there must be some water vapor in the Martian atmosphere (v. discussion of atmospheric composition below), and ice or water clouds can be expected at high levels. One would also expect some dust to be raised at times, and the absence of the cleansing action of rain might lead to the smaller dust particles remaining suspended for substantial periods.

Observers have identified three general categories of Martian clouds or haze, which have different colors and also exhibit somewhat different properties. (v. also Appendices 1 and 2.) The three types usually referred to are:

(a) *Yellow clouds,* generally thought to be composed of dust. At times, they can be seen to form in limited areas, and move with the winds. As they move they may grow in extent, and there have been several occasions when virtually the whole disk of Mars has been obscured by a yellow pall for weeks at a time, as during the opposition of 1956.

(b) *White clouds,* that are not always clearly distinguished from the yellow clouds and the patches of blue haze (see below). However, the existence of distinctly white clouds has been quite well established. These clouds sometimes seem to form preferentially over the dark areas of the surface. They are thought to be high thin clouds, resembling cirrus and probably composed of ice crystals. The possibility that they may be made of CO_2 crystals (dry ice) cannot definitely be ruled out, since we do not know the temperature of the Martian upper atmosphere; it might be less than the 140°K at which the CO_2 would begin to condense. (v. Section C. below.) However, the polarimetric properties of at least some white clouds clearly indicate the presence of frozen water (Dollfus, 1957).

(c) *Blue haze,* a diffuse and variable phenomenon that obscures surface features in the blue and violet. It usually covers most of the planet, but on occasion clearings are observed, allowing the surface features to be seen in the blue—these occasions are referred to as "blue clearings." According to Wilson, who reported to the Panel the preliminary results of his collaboration with R. S. Richardson on a study of the blue haze, there are two types of blue clouds: One type is diffuse and of low contrast, spreading over large areas of the planet, but being associated especially with the sunrise terminator. There are also relatively small and sharp blue clouds that move with the surface features, drifting slowly from day to day. The explanation for the blue haze and associated clouds is still in doubt; a more detailed discussion of this Martian enigma is made in Section E below.

In summary, the yellow and white cloud forms on Mars seem to be similar to our dust storms and cirrus clouds, respectively, and a study of their development and motion across the disk with high-resolution photography would tell a great deal about the Martian circulation patterns. One would use the same techniques now being used with meteorological satellites to study the Earth's atmosphere. The phenomenon of the blue haze, on the other hand, is more obscure, and represents one of the many elusive mysteries of Mars.

3. Albedo

The diffuse spherical reflectivity, or albedo, of the disk of Mars is near 0.3 at 7000 Å, but declines rapidly to about 0.04 below 4500 Å (v., e.g., Woolley, 1953; de Vaucouleurs, 1959). These photometric results are in qualitative agreement with the over-all reddish coloration of Mars. The low albedo in the blue and violet is an atmospheric effect, and is associated with the blue haze which obscures Martian surface detail at wavelengths less than 4500 Å. At any given wavelength, the reflectivity varies with position on the surface; in the yellow, representative albedos of the bright and dark areas are, respectively, 0.15 and 0.05.

Roughly half the energy in the solar spectrum lies in the red and infrared, primarily between 7500 Å and 3 microns. Any computations of surface heating by sunlight therefore critically depend on the reflectivity of Mars in the near infrared. Nevertheless, there are no reliable photometric data for Mars in this wavelength region. This is one important area in which observations from above the terrestrial atmosphere could make major contributions.

There has been only one marginal observation of Mars in the ultraviolet (Boggess and Dunkelman, 1959); this rocket measurement gave an albedo at 2700 Å near 0.24, but with a high probable error. If accepted at face value, this observation is of considerable interest. The terrestrial albedo at the same wavelength is more than an order of magnitude smaller, due to ozone absorption. The high Martian ultraviolet albedo suggests either the presence of ultraviolet fluorescent molecules other than ozone in the atmosphere of Mars; or, alternatively, the absence of much ultraviolet atmospheric absorption, coupled with a surface highly reflecting in the ultraviolet. If the Martian atmosphere is transparent in the ultraviolet—where the terrestrial atmosphere is opaque because of ozone absorption—then a very low upper limit is placed on the oxygen abundance on Mars, since O_3 is formed from interaction of O and O_2 molecules. In addition, the ultraviolet flux at the surface would pose another constraint on the chemistry and physiology of any Martian organisms. Ultraviolet light at 2700 Å is lethal for most unprotected terrestrial organisms. More work is needed in the ultraviolet photometry of Mars, and its interpretation.

4. Surface Pressure

Several methods have been utilized by Dollfus (1957) to determine the surface pressure on Mars polarimetrically. They all depend on the assumptions that the Martian atmosphere has the same scattering coefficient as the terrestrial atmosphere, and that Rayleigh scattering dominates. The mass of dust in the atmosphere can be estimated to be sufficiently small, and the composition of the atmosphere sufficiently similar to the terrestrial (see below), so that these assumptions are not grossly wrong. The derived polarimetric surface pressure of 85 millibars (about 230 gm cm^{-2}) is in good agreement with the value derived photometrically (v., e.g., de Vaucouleurs, 1954). It is unlikely that the true surface pressure differs by as much as a factor of two from 85 millibars.

5. Indications of Composition

a. *Surface*

The reddish coloration of the bright areas is often attributed to a high state of surface oxidation, and it has been difficult to envision likely red chromophores* which do not require high oxidation states. If the surface is more highly oxidized than is that of the Earth, this would help to explain the low atmospheric abundance of oxygen (see below), although the escape of oxygen to interplanetary space may also be contributory.

Attempts to match the polarization curve of Mars with terrestrial samples of various surface materials have led to the conclusion that the surface is largely covered with small opaque granules (Lyot, 1927). There are, in principle, many substances which can approximate the data for the bright areas; however, Dollfus (1957) prefers limonite, an iron oxide polyhydrate. Kuiper (1952), on the basis of infrared reflection spectra of the bright areas, suggests that felsitic rhyolite is a possible candidate. Discussion of the composition of the dark areas is deferred to Section F, which treats the possibility of life on Mars.

Infrared reflection spectra of the polar caps show conclusively that they are not composed of frozen carbon dioxide, the only condensible substance which might be expected besides water; the reflection spectra are also consistent with the assumption that the polar caps are made of ice (Kuiper, 1952). The polarization data indicate that the polar caps are composed of hoar frost (Dollfus, 1957).

b. *Atmosphere*

The only molecule which has been identified unambiguously by spectroscopic means in the Martian atmosphere is carbon dioxide (Kuiper, 1952). An analysis of Kuiper's observations which takes

* A chromophore is a chemical group that preferentially absorbs certain visual wavelengths, and therefore appears colored.

into account the pressure broadening of the spectral lines and the overlapping of CO_2 bands in the Martian and terrestrial atmospheres (v. Chapter I.) gives the fraction of carbon dioxide by volume on Mars as

$$f_{CO_2} = \frac{1.6 \pm 0.5 \times 10^2}{P_s^2}$$

where P_s is the total surface pressure in millibars (Grandjean and Goody, 1955). If $P_s = 85$ mb, as the polarimetric and photometric evidence indicates, then there is about 2.2 per cent carbon dioxide by volume in the Martian atmosphere. In absolute abundance, this is some 3100 cm-atm, or about 14 times more CO_2 on Mars than on Earth. The remaining 97.8 per cent of the atmosphere must be investigated indirectly.

Spectroscopic upper limits have been determined for O_2 and H_2O (250 cm-atm, and 3.5×10^{-2} gm cm^{-2} respectively, both from data of Dunham, 1952), showing that these are not major constituents of the Martian atmosphere; they are, indeed, relatively much less abundant on Mars than on Earth. Upper limits have also been placed on the abundances of O_3, N_2O, CH_4, C_2H_4, C_2H_6, NH_3 (Kuiper, 1952); and of NO_2 and N_2O_4 (Kaplan, 1961; Sinton, 1961). None of the foregoing molecules has been identified spectroscopically on Mars, and, if they are present, their abundance must be very low. Spectra of Mars taken by Sinton (1959) show no strong feature near the 2.35 micron band of CO; from this, one can conservatively estimate that the abundance of CO on Mars is less than 10 cm-atm. The fate of the CO produced photochemically in the upper Martian atmosphere is discussed in a subsequent section. Because of their low masses, H and He escape easily from the Martian exosphere, and cannot be major atmospheric constituents.

Since the polar caps are composed of frozen water, their seasonal retreat directly suggests that there is some water vapor in the Martian atmosphere. Because of the alternating formation of polar caps in opposite hemispheres, the circulation in the lower atmosphere must be such that water vapor is transferred from one hemisphere to the other. Derivation of the water vapor pressure above the sublimating edge of the polar ice cap can be made if the temperature of the cap boundary is known (Kuiper, 1952). Theoretical estimates of this temperature give water vapor abundances in excess of 10^{-3} gm cm^{-2} (Sagan, 1961a). An interpretation of the difference between theoretical and observed surface temperature as due to a carbon dioxide-water vapor greenhouse effect (and neglecting atmospheric circulation) gives values of water vapor between 1×10^{-3} and 2×10^{-2} gm cm^{-2} (Sagan, 1961a). It therefore appears that the abundance of water vapor on Mars is just about one order of magnitude below the present spectroscopic detectability where observations are made from the Earth's surface. But infrared observations from above the terrestrial atmosphere could easily determine the true value. At these partial pressures of water vapor the Martian surface temperatures are too low to permit pure liquid water, except possibly in ground capillaries. The only phase changes in pure water will be between water vapor and ice.

The foregoing discussion has succeeded in eliminating a large number of possible candidates for the unidentified 97.8 per cent by volume of the Martian atmosphere. The only likely remaining candidates are N_2 and Ar. Both are chemically unreactive gases, and are therefore stable over long periods of time. Neither has spectral features in the visible and infrared. Spectroscopic identification of these molecules must be made in the ultraviolet, and from above the Earth's atmosphere. Ar^{40}, the most prevalent argon isotope on Earth, arises from the beta-decay of K^{40} near the planetary surface. On Earth K^{40} is concentrated near the surface, presumably because of processes associated with the differentiation of crust and mantle. The amount of Ar^{40} which is released to the atmosphere is a function of the degree of surface weathering which untraps occluded gases. Thus the abundance of Ar^{40} on Mars depends on the largely unknown "geological" history of the planet. It has been customary to assume, by terrestrial analogy, that the Martian atmosphere is largely composed of N_2, with only trace amounts of Ar. But it is clear that this is at best a guess, and that an unambiguous determination of the primary constituents of the atmosphere of Mars is yet to be performed.

6. *Indications of Temperature*

Thermocouple bolometry in the 8- to 13-micron window in the terrestrial atmosphere (v. Chapter I.) has been used for many decades to determine the surface temperature on Mars. The absence of spectral features on Mars in this wavelength interval (Sinton and Strong, 1960) shows there is probably little atmospheric absorption of the emitted thermal radiation, and that the derived temperatures are indeed representative of the surface.

A crude average of daytime temperatures over the entire planet is about 240°K. Temperatures as high as 300°K have been recorded near noon at summer equatorial latitudes.* The temperature variation of a given region during daylight hours is often of the order of 80K°. Nighttime temperatures are not observed because the dark side of Mars is virtually unobservable from Earth; direct determination of the nighttime temperatures of Mars must be obtained by fly-by or landing probes. The actual diurnal temperature variation is probably well over 100K°, since nocturnal cooling will be very efficient in the absence of appreciable infrared-absorbing water vapor. The seasonal variation in temperature for a given equatorial region in noontime is of the order of 40K°. The steepest seasonal temperature gradient occurs between spring and summer. These temperatures are in rough agreement with temperatures computed for an airless rotating planet, using the appropriate solar constant and gray body albedo.

The average daytime temperature is also in fair agreement with the microwave brightness temperature of 211 ± 28°K (Giordmaine *et al.*, 1959). Actually the microwave emission arises from some subsurface level which has not reached thermal equilibrium with the daytime solar radiation. Thus, the microwave brightness temperature represents a rough average of daytime temperatures and some nighttime temperatures at some level not very far beneath the surface.

C. The Vertical Structure of the Martian Atmosphere

The following method of constructing models of the lower Martian atmosphere has been used by a number of authors: A surface air temperature is selected (usually for noontime at equatorial latitudes.) The atmosphere is assumed in convective equilibrium, so an adiabatic lapse rate for the temperature gradient applies immediately above the surface. By some means, usually by a radiative equilibrium argument, the temperature of the tropopause is derived, the adiabatic lapse rate is caused to break off at that temperature, and at greater altitudes an isothermal stratosphere prevails. In the absence of an ozone layer or other cause of stratospheric temperature inversion, the isothermal regime extends to the region where substantial photodissociation occurs and the thermosphere begins. It is evident that the procedure just described is fraught with uncertainties, especially since no completely satisfactory model of the terrestrial atmosphere has been derived using this method. Nevertheless, a crude approximation to the terrestrial atmosphere can be derived in this way, and one might piously hope that a similarly derived model of the Martian atmosphere would not be too far from the truth. In the following paragraphs we discuss some of the numerical results to be derived from such model atmosphere calculations.

It is natural to use some thermocouple temperature—say 270° to 300°K—as the equatorial starting temperature. However, Mintz (1961) has argued that, by analogy with terrestrial desert regions, the air temperature a few meters above the ground near noon may be very much less than the ground temperature itself. A strictly empirical terrestrial data compilation extrapolated to Mars suggests to Mintz that the maximum near-surface atmospheric temperature on Mars may be 50K° less than the maximum ground temperature at a given equatorial locale. In this case the appropriate near-surface temperature for intermediate latitudes should be 220 to 250°K. The temperature difference between ground and air will be smaller at higher latitudes, where there is less solar heating.

* See also discussion in latter part of Appendix 2, and in Appendix 8.

Because of the low atmospheric water vapor content, the dry adiabatic lapse rate can be used near the equator:

$$\frac{\partial T}{\partial h} = -\frac{g}{C_p}$$

where C_p is the specific heat at constant pressure of the atmosphere, and g is the acceleration due to gravity. With likely models of the atmospheric composition (see above), the lapse rate is quite close to 3.7K°/km. At higher latitudes, where there is heat transported from the equator by advection at high levels, the temperature gradient should be less steep than the dry adiabatic. (v. Mintz, Appendix 8.)

A crude estimate of the equatorial tropopause temperature can be obtained from the Gold-Humphreys theory. The stratosphere is assumed in radiative equilibrium with absorbing gas in the troposphere (and with high clouds). It receives an infrared flux from the troposphere of σT_t^4, where σ is the Stefan-Boltzmann constant, and T_t is the effective tropospheric temperature; it radiates downward and to space a total flux of $2\ \sigma T_s^4$, where T_s is the stratosphere temperature. Thus, the stratosphere temperature is $2^{-1/4}$ the ground temperature. With

$$250°\text{K} \geqslant T_t \geqslant 220°\text{K}$$

we derive for Mars

$$210°\text{K} \geqslant T_s \geqslant 185°\text{K}.$$

The corresponding tropopause altitudes are

$$11\ \text{km} \geqslant h_t \geqslant 9\ \text{km}.$$

Similar procedures for the Earth give tropopause temperatures and altitudes in fair agreement with the observed values. Nevertheless, the entire procedure has been questioned for the Earth, first because the stratosphere cannot be entirely in radiative equilibrium, and second because if the procedure is applied uncritically to other latitudes it leads to a lower stratosphere temperature at high latitudes, contrary to observation.

A much more elaborate equilibrium method has been used by Goody (1957), for two extreme model atmospheres, neither of which corresponds closely to the real Martian atmosphere: (1) with 2.2 per cent CO_2 by volume, and the remainder N_2, and (2) 10^{-2} gm cm^{-2} H_2O and the remainder N_2. Assuming a high surface temperature of 270°K, Goody obtained in the first model a tropopause at 9 km, but the temperature falling off well into the stratosphere, becoming approximately constant at 134°K only above 90 km. In the second model the tropopause is at 25 km, and the stratosphere is approximately isothermal at 153°K. Goody found that a decline in the water vapor concentration in the second model led to a warmer and lower tropopause, but that the highest stratosphere temperature consistent with a ground temperature of 270°K was 227°K, with a tropopause at 12 km. However, if the temperature in the lower atmosphere is less, as Mintz argues, the tropopause temperatures will also be less. The above range of tropopause altitudes is consistent with estimates which have been made for the altitudes of convective clouds on Mars.

A convenient lower limit to the tropopause temperature can be placed by the condition that CO_2 not crystallize out. If the stratospheric temperature dropped below the CO_2 frost point, the carbon dioxide in air rising into the stratosphere would crystallize out in the form of clouds that would obscure surface detail at all visual wavelengths. A haze at the limbs of Mars can be detected polarimetrically (Dollfus, 1957), and—more rarely—visually. But it is not known whether this haze may be composed in part of crystalline CO_2. At an altitude of 20 km the Martian atmospheric pressure is about 30 mb. The partial pressure of CO_2 at this altitude, assuming a constant mixing ratio, is about 1 mb. The temperature at which this is also the vapor pressure above dry ice is about 140°K. It is therefore unlikely that the tropopause temperature is below about 140°K if the tropopause altitude is around 20

km or less. However, decreasing the tropopause temperatures in Goody's two models (as may be required for a lower near-surface temperature) pushes them precariously close to the carbon dioxide frost points. This perhaps suggests that carbon dioxide crystallizes near the tropopause, unless an alternate source of heating exists in Mars' atmosphere. An ozone layer is one possible source. Interaction of solar protons and cosmic rays with the atmosphere (if Mars has a small magnetic field) is another, but is undoubtedly negligible at these altitudes.

Above the tropopause the temperature may continue to decline somewhat, as in Goody's first model. Eventually the atmosphere becomes approximately isothermal, and the barometric equation

$$P = P_0 e^{-h/H}$$

will determine the decline of pressure with altitude. $H = \frac{kT}{mg}$ is the scale height. Because of the smaller value of g on Mars, H is larger, and the pressure falls off with altitude much more slowly than on Earth. Above about 40 km, the atmospheric pressures are higher at a given altitude on Mars than on Earth. Therefore, unit optical depth for solar far-ultraviolet radiation will occur at higher altitudes on Mars than on Earth, and the Martian thermosphere and ionosphere begin at greater altitudes. Photodissociation of O_2 will occur at about 110 km, and of CO_2 at about 150 km on Mars. If there is very little O_2 in the Martian atmosphere, the principal ions probably will be N_2^+, CO_2^+, CO^+, and O^+. Extension of the theory of production of N_2^+ by short wavelength solar ultraviolet light in the terrestrial atmosphere (Chamberlain and Sagan, 1960) suggests that a Martian ionosphere arising from N_2 ionization is localized at about 300 km, and has a mean daytime electron density $\sim 10^4$ cm^{-3}. Yanow (1961) has derived, through other processes, electron densities in excess of terrestrial ionospheric values, but the importance of electron attachment was neglected in this derivation. It is apparent that more theoretical work is needed for a thorough understanding of the Martian atmosphere, and of our own.

D. The Photochemistry of the Martian Atmosphere

In this section we will discuss the dissociation and ionization of Martian atmospheric constituents by solar ultraviolet radiation, and their possible escape into interplanetary space.

For the purposes of this discussion we consider the Martian atmosphere to be composed primarily of N_2 and CO_2, with small amounts of H_2O. In such an atmosphere, the longest wavelength at which ultraviolet light is effectively absorbed is about 1850 Å, where water vapor photodissociation begins:

$$H_2O + h\nu \longrightarrow OH + H, \quad \lambda < 1850 \text{ Å} \tag{1}$$

Some representative water absorption cross-sections are (Watanabe and Zelikoff, 1953)

$$\alpha(\lambda 1850) = 5 \times 10^{-20} \text{ cm}^2;$$
$$\alpha(\lambda 1800) = 7 \times 10^{-19} \text{ cm}^2;$$
$$\alpha(\lambda 1700) = 4 \times 10^{-18} \text{ cm}^2;$$
$$\alpha(\lambda 1450) = 4 \times 10^{-19} \text{ cm}^2;$$
$$\text{and } \alpha(\lambda 1200) = 6 \times 10^{-18} \text{ cm}^2.$$

Carbon dioxide photodissociation,

$$CO_2 + h\nu \longrightarrow CO + O, \quad \lambda < 1692 \text{ Å} \tag{2}$$

has characteristic absorption coefficients of about 10^{-19} cm^2 just on the short wavelength side of the dissociation limit. In the absence of a protective cover of oxygen, it is therefore clear that the photodissociation of water vapor and of carbon dioxide should proceed very rapidly on Mars.

Molecular nitrogen is very difficult to photodissociate. On the long-wave length side of λ796, there are a series of sharp absorption bands but no dissociation continua. Three bands of the Lyman-

Birge-Hopfield system at λ1250, λ1226, and λ1205 participate in a predissociation mechanism (Herzberg and Herzberg, 1948). N_2 has an ionization continuum at $\lambda < 796$ Å; the ionization process is followed by dissociative recombination, the net result being the production of excited nitrogen atoms as follows:

$$N_2 + h\nu \longrightarrow N_2^+, \quad \lambda < 796 \text{ Å};$$
$$N_2^+ + e \longrightarrow N^* + N^{**}. \tag{3}$$

(The asterisks indicate excited states of the nitrogen atom.)

The fate of the photodissociation products is critically dependent on the rates of escape of atoms and molecules from the Martian exosphere. The escape rate is, in turn, critically dependent upon the exosphere temperature. The primary source of heating of the terrestrial exosphere is probably hard ultraviolet and soft X-ray lines emitted by the solar corona, in the spectral range between about 100 and 1000 Å. Absorption of these lines on Mars should be primarily by N_2 alone, rather than by N_2 and O, as on Earth. However, the absorption coefficients of N_2 and O are approximately equal at these wavelengths, so the Martian thermosphere should be heated in the same manner as the terrestrial thermosphere. The intensity of the lines will be diminished by a factor of 2.4 due to the increased distance of Mars from the sun. Thus, the temperature gradient will be less in the Martian thermosphere by the same factor, and therefore in the upper thermosphere the temperature will be correspondingly less.

The exosphere begins where the mean free path, L, equals the scale height H: $L = (n\sigma)^{-1} = H = kT/mg$. In this equation n is the number density of particles, and σ the collision cross-section. Since the acceleration due to gravity is a factor of 2.6 smaller on Mars, thermosphere scale heights should be about the same as on Earth. Accordingly, if we disregard any possible difference in the mean molecular weight of the atmosphere at the exosphere levels on the two planets, and take the collision cross-sections to be equal, the base of the Martian exosphere will have about the same number density of particles as the base of the terrestrial exosphere, but the temperature there will be about one-half as great. The daytime exosphere temperature of the Earth lies near 1800°K (Kallmann, 1961). Thus, on the basis of these very simple arguments, the exosphere temperature on Mars is possibly about 900°K. At these temperatures, hydrogen and helium escape in periods $<10^5$ years (Spitzer, 1952), so the hydrogen released from water vapor by reaction (1) will be lost from Mars. However, with these exosphere temperatures, neutral atomic oxygen and neutral atomic nitrogen will not escape during the lifetime of Mars. The preceding discussion has implicitly assumed that comparable sources of radiative cooling exist in the Martian upper atmosphere as in our own. Since minor atmospheric constituents may be the primary sources of cooling, it is difficult to assess the plausibility of this assumption at the present time. If more efficient cooling exists on Mars, the exosphere temperatures there will be even less; if the cooling is less efficient, the exosphere temperatures may even approach those in the terrestrial atmosphere.

For ions in the exosphere there is another mechanism that enhances the escape. The outward diffusion of electrons from the high Martian atmosphere will reduce the effective gravity of the ions in the exosphere. For a one-component-ion exosphere, the ions will escape as if they had half their molecular weight, as has been emphasized by Öpik and Singer (1960); in a many-component exosphere, the escape of heavy ions will be assisted by the presence of light ions, and, alternatively, the escape of light ions will be inhibited by the presence of heavy ions. On Mars the most abundant ions in the exosphere will probably be N_2^+, CO_2^+, CO^+, O^+, and N^+. Since the ions can escape only from that level above which an escaping particle has a low probability of encountering other particles, charge exchange and collisional recombination will be slow in the exosphere. If the ion population of the Martian exosphere is predominantly hydrogen and helium, as is the case for the Earth, then the escape of heavier ions of lower abundance will be greatly assisted. In particular, if large quantities of CO^+ were produced in the high Martian atmosphere, rapid escape of the ion would ensue, and a low equilibrium population would be established. However, if charge exchange were dominant, this

conclusion would not follow: O^+ has a first ionization potential of 13.6 *eV* compared with 14.1 *eV* for CO^+; thus CO^+ would exchange charge with *O*, and little CO^+ would escape from Mars.

The question of the escape of CO^+ is important for the following reason: In the absence of molecular oxygen on Mars, CO_2 will not be shielded from solar ultraviolet radiation in the Schumann-Runge continuum of O_2 (1250-1760 Å), as is the case on Earth. Thus, reaction (2) would proceed rapidly. The solar flux at wavelengths shorter than 1692 Å on Mars is estimated to be about 10^{12} quanta cm^{-2} sec^{-1}. In the present Martian atmosphere every such photon is absorbed, primarily by CO_2. If these conditions had been maintained during the entire history of Mars, some 10^{29} *CO* molecules would have been produced over each cm^2 of surface. However, this large amount of *CO* could not accumulate, because of the reverse reaction,

$$CO + O + M \longrightarrow CO_2 + M \qquad (4)$$

where *M* is any third body. If we use the Earth as an example, we find p_{CO}/p_{CO_2} less than 10^{-3}. On Mars, because of the lower oxygen abundance, and lower availability of a third body, the reverse reaction should be slower. Detailed calculations on the photochemistry of the Martian atmosphere have recently been performed by Marmo and Warwick (1960, 1961). The calculations were based on reactions (2) and (4) plus the following system of chemical reactions:

$$O(^3P) + CO \longrightarrow CO_2$$
$$O + O + M \longrightarrow O_2 + M$$
$$O + O_2 + M \longrightarrow O_3 + M$$
$$O + O_3 \longrightarrow 2O_2$$

Assuming a CO_2 mixing ratio of 2 per cent, a total abundance of *O* and of *CO* can be derived; the values are approximately 0.03 cm-atm. The maximum *O* abundance occurs between 150 and 200 km. The O_2 maximum lies some 40 km below. The total O_2 abundance is roughly 0.2 cm-atm, well below the spectroscopic upper limit of 250 cm-atm (cf. Section B.5.b). The volume abundance of molecular oxygen near the ground is then less than 10^{-8} per cent. Thus if CO_2 photodissociation is the primary source of molecular oxygen on Mars, its abundance at the ground should be of negligible biological or other significance.

Marmo and Warwick (1960) have also calculated the ozone distribution on Mars as a function of the O_2 abundance. Taking the spectroscopic upper limit on the O_2 abundance of 250 cm-atm, they find that the abundance of O_3 continues to increase with depth all the way down to the surface. On the Earth, an ozone maximum exists, because at greater depths all the ultraviolet light which photodissociates O_2 is absorbed. On Mars the primary source of oxygen atoms is CO_2 photodissociation, and not O_2 photodissociation. With 250 cm-atm of O_2, there is almost as much ozone in the Martian atmosphere as in our own. The spectral region between 2000 and 3000 Å would then be opaque, as is the case on Earth. However, ozone is a highly reactive molecule, and, when in contact with the Martian surface, should chemically react. The possibility that the Martian surface is oxidized (cf. Section B.5.a) by atmospheric ozone was first suggested by Wildt (1934), who also anticipated that the ozone would be localized near the surface. With smaller total abundance of O_2, the ozone abundance is proportionately less. If all the oxygen arises from the photodissociation of CO_2, the total amount of ozone on Mars should be less than 10^{-4} cm-atm. Under these circumstances, ozone will not absorb appreciably in the near ultraviolet; what ozone exists will still be localized near the ground.

The absorption cross-section of O_2 near 1800 Å is about 10^{-19} cm^2 (Watanabe, Inn, and Zelikoff, 1953), and increases for shorter wavelengths. Thus, if there is more than about 0.3 cm-atm of molecular oxygen in the Martian atmosphere, CO_2 and H_2O will be largely protected from photodissociation; but if there is much less than this amount such photodissociation should proceed very rapidly. If all the O_2 arises from carbon dioxide photodissociation, the results of Marmo and Warwick suggest that there will still be some shielding by molecular oxygen.

An infrared spectrophotometric scan of Mars by Sinton (1959) shows no strong feature near the 2.35-micron CO band. It can therefore be estimated that the abundance of CO on Mars is probably less than 10 cm-atm, in agreement with the theoretical photochemistry. This provides direct evidence that large amounts of CO are not accumulating in the Martian atmosphere.

It is possible that sinks other than CO_2 exist for CO in the Martian atmosphere. Since CO is a fairly reactive molecule, the existence of such sinks should occasion no surprise. Because of the low abundance of water vapor, and the chemical inertness of N_2, it is likely that any sink molecule is composed exclusively of C and O. A possible sequence of reactions which is of some interest is as follows:

$$2CO + CO_2 + h\nu \longrightarrow C_3O_2 + 2O \tag{5}$$

$$C_3O_2 + h\nu \longrightarrow C_2 + CO_2 \ (\lambda < 3250 \text{ Å}) \tag{6}$$

The oxygen atoms produced in (5) will react back with CO as in reaction (4). Because so many more photons are available for reaction (6) than for reaction (5), the carbon suboxide produced in (5) will be dissociated almost immediately, and before there is much opportunity for the polymer $(C_3O_2)_n$ to be synthesized. The over-all reaction is

$$2CO + 2h\nu \longrightarrow C_2 + 2O \tag{7}$$

Atomic oxygen would then escape, primarily as the ion. Other possible reaction sequences give over-all results similar to (7). The possibility that C_2 is a sink for CO on Mars is of some interest, since carbon smoke has been suggested as a component of the Martian atmosphere for completely different reasons; *viz.*, to explain the blue haze (Rosen, 1953; Öpik, 1960).

E. The Blue Haze

Photographs of Mars taken in ordinary light usually show, of course, surface detail and the distinction between bright and dark areas. However, as the wavelength in which the photographs are taken is decreased, surface detail becomes increasingly obscured by what appears to be a planet-wide haze. This "blue haze" is already quite apparent at 4500 Å, the extinction increasing rapidly towards shorter wavelengths. It should be emphasized that the haze is not *itself* blue; since it extinguishes the blue it probably would appear reddish. Mars is much less bright in the blue than in the red.

A probably related phenomenon is that discovered by Wright (1925); *viz.*, that the apparent diameter of Mars in blue light exceeds that in red light by an amount of the order of 100 km. However, in part at least, Wright's phenomenon is a photographic effect. Sharanov (1950) has noted that the limb darkening of red images of Mars appears to be more pronounced than the limb darkening of blue images, so that the apparent diameter of the red image will be less.

The standard explanation of the blue haze is a high-altitude layer in the Martian atmosphere which causes extinction of solar blue light reflected from the surface, but which is transparent at longer wavelengths. Such an explanation would also account for any residuum of Wright's phenomenon that cannot be explained as an instrumental artifact. In the following discussion we list some further properties of the blue haze, and then compare the observables with the theories that have been proposed to explain them:

(a) *Blue clearing.* Slipher (1937) discovered that on occasion the blue haze dissipates, and familiar surface details can be discerned. Most of the more dramatic cases of blue clearing have been observed near favorable oppositions. There is an obvious observational selection, since Mars is most thoroughly investigated during favorable oppositions. Nevertheless, de Vaucouleurs has found, from a study of Lowell Observatory plates taken chiefly by E. C. Slipher during the 30 year interval ending in 1956, that when the effects of observational selection

are removed there still remains a correlation of blue clearing with favorable opposition (cited by Wilson, 1959). Cases are also known of blue clearings at unfavorable oppositions (de Vaucouleurs, 1954; Smith and Tombaugh, 1961), of blue clearings several months from opposition (Richardson and Roques, 1959), and of local blue clearings on small topographical scales down to the limit of resolution (Wilson, 1961). (v. also Wilson, Appendix 1.) Any explanation of the blue haze must explain these highly erratic time variations in opacity.

(b) *Limb brightening.* It is a common observation that the blue image of Mars increases in brightness from the center of the disk towards the limb. An optically thin absorbing layer will give limb darkening instead of limb brightening, as has been emphasized by Sharanov (1957), and these observations provide a simple and direct proof that the blue haze is not produced solely by a thin homogenous absorbing layer. However, the haze layer may not be homogenous. Indeed condensation and crystallization of water (and possibly even carbon dioxide) can be expected on the twilight limbs; this would certainly contribute to limb brightening. Sharanov also emphasizes that if the blue haze is due primarily to absorption, regions of increased local haze density—the blue clouds—should appear darker, rather than brighter as observed.

(c) *Spectrum.* The opacity of the blue haze increases with decreasing wavelength at a rate which cannot be explained by Rayleigh scattering (Barabascheff and Semejkin, 1934); similarly, Wright's phenomenon of the increased size of the disk in the blue cannot be explained by Rayleigh scattering alone (Fessenkov, 1926).

Spectra of the blue haze taken by Wilson (1958) give some likelihood that absorption features exist at 4015, 4250, 4495, and 4695 Å. The absorption must occur in the blue haze itself, or in the overlying atmosphere. The presence of these features suggests that there is an absorption component to the blue haze extinction.

In recent years, the most extended theoretical explanation of the blue haze has been in terms of a layer of ice crystals of particle diameter of the order of 0.3 microns (Schatzman, 1951; Kuiper, 1952; Hess, 1958). The extinction in the blue is explained by strong forward scattering of blue light by such particles, so that little light is reflected, and the haze is dark in the blue. At longer wavelengths the haze is transparent. However, Goody (1957) and Öpik (1960) have pointed out that a completely forward-scattering layer cannot possibly explain the blue haze, because light scattered by the layer to the planetary surface will be reflected back by the surface and forward scattered by the layer out of the atmosphere. A purely forward-scattering layer above a reflecting surface is transparent. A predominantly back-scattering layer will have a very high albedo in the blue and ultraviolet, contrary to observation. Particles with strong sidelobes on the scattering diagram occur on Earth with very low frequency, and their large-scale production on Mars seems extremely unlikely. Öpik concludes that a single scattering layer cannot by itself explain the blue haze.

In discussing Öpik's objections, Kuiper (1961) agrees with much of the preceding discussion, but argues as follows: Suppose the blue haze is composed of small particles which, although predominantly forward-scattering in the blue, do have a small back-scattering lobe. Most of the incident solar light is transmitted to the surface, and a small fraction is scattered back to space. The light which reaches the surface is then reflected back up to the haze, and almost all of it is transmitted back to space, again by forward scattering. The critical point, however, is that the blue light reflected from the surface is reflected with a very low albedo, about 0.05. Therefore, if more than 5 per cent of the incident solar radiation is initially back-scattered, the light reflected from the Martian surface in the blue will be swamped by the back-scattered solar radiation in the blue, and surface features will be obscured.

Sharanov (1957) maintains that a pure scattering atmosphere, with both Rayleigh and neutral scatterers, can explain the blue haze; the extinction in the blue can be explained by Mie theory for particles larger than the wavelength of light.

Urey and Brewer (1957) have suggested that the blue haze is caused by solar protons striking the upper Martian atmosphere and producing such ions as CO_2^+, CO^+, and N_2^+, all of which absorb in the blue and violet. The correlation of blue clearing with favorable oppositions might conceivably then be explained by the deflection of solar protons by the terrestrial magnetic field. There are however, several objections to this hypothesis (Sagan, 1961b). The altitude of the blue haze is agreed by all observers to be less than 200 km above the Martian surface; in order to reach such depths solar protons must have energies in the *MeV* range and very high fluxes, contrary to observation. In addition, for the blue clearings at opposition to have the observed durations, impossibly low interplanetary magnetic field strengths are required. However, the possibility should not be overlooked that solar electromagnetic radiation ionizes molecular constituents in the upper atmosphere of Mars, although then there would be no mechanism for control by the Earth during opposition.

A model of the blue haze with two layers, one scattering and the other blue-absorbing, seems acceptable. If the scattering layer is below the absorbing layer, there should be limb darkening in the blue, contrary to observation. However, if the scattering layer is above the absorbing layer, limb-brightening can occur, as shown explicitly for a Rayleigh atmosphere by Coulson (1959). In this case, much more red than blue light penetrates to the absorbing layer, and to explain the low blue albedo, we must suppose that the underlying absorbing layer is very transparent to red light, and very opaque to blue. This conclusion is due to Wright (1927). A recent analysis of the two-layer model by Öpik (1960) shows that the bulk of the extinction in the blue is due to true absorption, and not to scattering. It then remains to identify the absorber, and to explain the blue clearing.

Starting with CO_2, N_2 and H_2O, a number of substances can be produced which absorb in the blue-violet. Among these are CO_2^+, CO^+, N_2^+, C_3O_2, $(C_3O_2)_n$, C_2, C_3, . . . , C_n, and NO_2. Except for the higher polymers of atomic carbon, and of carbon suboxide, each of these substances shows pronounced band structure in its absorption spectrum. At the present time, few spectrophotometric studies have been performed of the blue haze, and the possibility that there exist holes between the bands in the absorption spectra should not be discounted *a priori*. Some of the substances listed also have strong absorption features at wavelengths longer than the widely quoted wavelength at which the opacity of the blue haze becomes very noticeable, 4500 Å; e.g., C_2 has features in the Swan bands at 4737 and 5129 Å. However, Öpik (1960) has computed the transmission coefficients on the two-layer absorbing and scattering model, described above, and finds that there is a continuous trend in absorption, with no sharp discontinuity near 4500 Å. Therefore, substances which absorb strongly longward of 4500 Å cannot be excluded on that basis alone. The original suggestion that carbon smoke was responsible for the blue haze is due to Rosen (1953). In the discussion above on the photochemistry of the Martian atmosphere (Section D), a sequence of reactions which leads to C_2 was suggested. Subsequent polymerization of C_2 can lead to quite large carbon particles, with considerable extinction in the blue. However, a residuum of C_2 precursor should always be present, and strong Swan band absorption would be expected. Since photodissociation of C_3O_2 is more likely than polymerization, more C_2 than $(C_3O_2)_n$ is expected in the Martian atmosphere. From infrared absorption spectra obtained by Sinton (1959), estimates can be made that the NO_2 abundance on Mars is less than 10 cm-atm (Kaplan, 1961; Sinton, 1961), and hence it is clear that NO_2 makes negligible contribution to the blue haze.

If the Urey-Brewer mechanism is untenable, then there are only two remaining explanations of the blue clearing; either it is a gravitational settling of suspended particles, or it is a phase change of a condensible or sublimable substance. Hess (1958) has noted that for suspended particles to settle out in an interval of a few days to account for a clearing, then Stokes' law gives a minimum particle size of about 10 microns. If the blue haze is due to forward-scattering particles, as Hess thought, then this size is far too large. However, if, as now seems equally acceptable, the blue haze is due to a combination of absorption and scattering, particles of 10 micron size are not excluded. They would scatter equally well at all visible wavelengths, and would have to be responsible for the absorption in the blue. Actually, as Wilson has noted, clearing can occur in periods of a few hours, so the sizes deduced by

Hess (10 μ) would have to be still further increased to account for such rapid settling. Nevertheless, it remains to explain why settling occurs preferentially at favorable opposition. At the 1956 favorable opposition, extensive dust storms were observed which obscured much of the Martian disk. Nevertheless some blue clearing was observed in 1956.

For a substance which changes its absorptive properties with phase, similar difficulties ensue. The substance would have to be ordinarily quite close to the temperature at which the phase change occurs, so that small temperature variations could have large absorption consequences. For a blue clearing to occur over the entire Martian disk, the absorbing substance must be sitting poised near the phase change temperature at all points on Mars. Then the temperature change must occur uniformly at all latitudes, multiplying the improbabilities. These and other difficulties were noted first by Urey (1958).

It is clear that our knowledge of the blue haze of Mars is imperfect at the present time, and that no completely satisfactory explanation has been advanced. More theoretical, spectrophotometric, and visual observational work is needed.

F. The Question of Life on Mars

Recent laboratory work and current theories of the origin of the solar system have given some insight into the processes which underlay the origin of life on Earth. It now appears that the physical conditions on the primitive Earth led directly to the formation of complex organic molecules in the early oceans. Interaction of these molecules led to the origin of a molecular system which catalyzed the synthesis of identical molecular systems from the surrounding medium. When occasional errors in copying were made, or alterations in the structure of the parent molecules occurred, identical changes were perpetuated in succeeding generations of these molecular systems. Such conditions provide the requirements for Darwinian natural selection; from that point on, systems of greater complexity and more sophisticated properties were selected, and biological evolution was well under way. On Mars there is every reason to expect that the early physical environment was very similar to the early terrestrial conditions, and, especially if the origin of life requires times short compared with the age of the solar system, life should have arisen on Mars as well. An important present difference between the two planets is the absence of oceans on Mars, but it is not improbable that Mars held oceans in the early stages of its development.* A more detailed discussion of the origin of life may be found elsewhere (v., e.g., Oparin, 1957; Sagan, 1961c).

Early forms of life on the two planets were possibly quite similar. But as time went on, the physical conditions of Mars and Earth diverged increasingly. Because of the lower planetary mass, atomic oxygen—especially as the ion—escaped more readily on Mars than on Earth. The photodissociation products of water did not remain in the atmosphere as easily, and it is quite possible that Mars has never had a period in which the atmosphere contained significant quantities of free oxygen. In order that organisms adapt to these disparate environments, the course of biological evolution must have been very different on the two planets. If there are organisms on Mars today, we must not expect them to be similar to familiar lifeforms.

The low mean temperatures, lack of oxygen, low water abundances, and possible high surface ultraviolet flux would pose serious hazards for most terrestrial organisms. There is, however, reason

* Urey (in press 1961) has pointed out that the present low water content of the Martian atmosphere may be explained in terms of either (a) an equilibrium between the rate of outgassing of water vapor from the Martian interior and the rate of escape of hydrogen, or (b) a transitional state, in which we are now witnessing the final stages of a gradual dissipation of water from a planet that once may have held oceans. In view of the cosmic abundance of hydrogen, the large amount of hydrogen on the Jovian planets, and the presence of oceans on Earth, Urey argues, there is every reason to believe that Mars and Earth both had reducing atmospheres originally, and that water then became abundant in the early stages of planetary development.

to believe that some terrestrial microorganisms might function satisfactorily in the Martian environment. In any case, it is very dangerous to maintain that these somewhat exotic physical conditions preclude all forms of life; the range of adaptation of terrestrial organisms is remarkable, and it is unlikely that extreme evolutionary adaptations were not attempted on Mars. Finally, there are several pieces of direct evidence which are strongly suggestive of life on Mars, and these will now be reviewed briefly.

Early telescopic observations of Mars turned up green coloration in the dark areas, and rectilinear markings in the bright areas. These were interpreted, respectively, as photosynthesizing plants and artificial waterways of intelligent animals. Such uncritical leaps from marginal observations to firm conclusions led to a general unpopularity of planetary studies in astronomical circles, an unpopularity from which we are today only gradually emerging. It appears today that the dominant color of the dark areas is gray, not green, and that previous reports of vivid green coloration were due largely to extrafocal blue light of refracting telescopes and to contrast with the buff-colored deserts. There is still some difference of opinion on the extent of a residual greenish tinge.* Under the very best seeing conditions, the so-called "canals" are broken up into mottled fine detail, which the eye, under somewhat poorer seeing conditions, strings up into rectilinear markings.

In more recent years other observations have suggested, in more convincing manner, that there may be indigenous organisms on Mars. At the time when water vapor must be in transit through the atmosphere from the receding to the forming polar ice cap, seasonal changes are observed in the dark areas. The outlines of the dark areas sharpen, and their albedo decreases. Predominantly gray areas take on brown or, more rarely, green or blue coloration. Except at the rim of the receding ice cap, these colors—at least in recent years—are delicate pastels. As winter approaches, the outlines become more diffuse, the darkness decreases, and the coloration returns to the predominant gray. These color changes can be attributed to the seasonal growth and decay of Martian vegetation.

It has also been suggested that the color changes arise from the response of hygroscopic salts to increased humidity. Although it has been impossible to suggest the precise hygroscopic salts which change color in the appropriate manner at the appropriate humidity, this is not a very convincing objection; after all, no one has described the precise Martian organisms which account for the color changes either. Similar alternative explanations are available for the sharpening of outlines and the darkening of areas.

A second seasonal variation occurs in the polarization of the dark areas (Dollfus, 1957). If the polarization of the dark areas is plotted against phase angle for a given Martian season, a characteristic polarization curve is obtained. The polarization curve for the same areas during a different season has a similar shape but is displaced in the absolute value of the polarization. Curves for the bright areas show no such seasonal variation. In order to reproduce these curves in the laboratory, light must be scattered from small opaque particles about 10^{-2} cm in diameter. In order to reproduce the seasonal variation in the curves, it must be assumed that Mars is covered with small objects 10^{-2} cm in diameter which periodically change in diameter, or absorptivity, or both. The polarization data are therefore amenable to an interpretation involving the seasonal proliferation of microorganisms; but it is also possible that Mars is covered by small nonliving particles which change in size or darkness when the abundance of water vapor varies.

A final observation connected with the possibility of life on Mars is the discovery of absorption features in the 3.4 to 3.7μ range which occur in the reflection spectrum of the dark areas and not of the bright areas (Sinton, 1959). These features have been interpreted as vibrational transitions in hydrocarbon and carbohydrate or aldehyde bonds, although the possibility that they arise from a combination of inorganic substances does not seem to have been explored sufficiently. But even by assuming that the identification is correct, the presence of organic matter on Mars is not necessarily

* See Appendix 2 by Tombaugh, and Kuiper (1957).

evidence for life on that planet. If life never arose on Mars, if the planet never possessed an atmosphere with appreciable free oxygen, and if ultraviolet light has always been absorbed in the Martian atmosphere, then there may still be prebiologically-synthesized organic molecules littering the surface. However, the localization of organic matter in the dark areas—where visual and polarimetric evidence also suggests the presence of life—is most naturally explained by a biological origin.

The evidence taken as a whole is suggestive of life on Mars. In particular, the response to the availability of water vapor is just what is to be expected on a planet which is now relatively arid, but which once probably had much more surface water. The limited evidence we have is directly relevant only to the presence of microorganisms; there are no valid data for or against the existence of larger organisms and motile animals.

The direct testing of the hypothesis that life exists on the planet Mars will soon be possible. The following categories of experiments are currently being studied and are of great relevance (v. also, Chapter V, and Appendix 6 by Davies *et al.*):

(a) Ultraviolet spectroscopy of Mars from Earth satellites and from fly-by vehicles to determine the abundances of O_2 and O_3, and to test the ultraviolet transparency of the Martian atmosphere.

(b) Infrared spectroscopy of the Martian atmosphere from balloons and from fly-by vehicles to determine the abundance and distribution of water vapor on Mars.

(c) Infrared spectroscopy of the Martian surface from balloons and from fly-by vehicles to determine the existence and distribution of surface organic matter. A successful infrared fly-by experiment would tell us which organic molecules are situated in which regions of Mars. It would then be very exciting to correlate this information with the older evidence for biological activity in the dark areas, as cited above.

(d) Direct experiments by landing vehicles to detect Martian organisms. These range from vidicon cameras for detecting large, recognizable, and possibly motile lifeforms, to pH and turbidity monitoring of deposited culture media to test for the proliferation of indigenous microorganisms.

The ancient and exciting question of the possible existence of life on Mars will probably be answered in the next decade.

References

Barabascheff, N., and B. Semejkin, Photometrische Untersuchung der Marsoberfläche und seiner Atmosphäre durch Farbfilter, *Z. Astrophys., 8:* 44, 1934.

Boggess, A., and L. Dunkelman, Ultraviolet reflectivities of Mars and Jupiter, *Astrophys. J., 129:* 236, 1959.

Chamberlain, J. W., and C. Sagan, The origin of nitrogen ionization in the upper atmosphere, *Planetary and Space Sci., 2:* 157, 1960.

Coulson, V. L., Characteristics of the radiation emerging from the top of a Rayleigh atmosphere—I, intensity and polarization, *Planetary and Space Sci., 1:* 265, 1959.

de Vaucouleurs, G., *Physics of the Planet Mars,* Faber and Faber, London, 365 pp., 1954.

de Vaucouleurs, G., Multicolor photometry of Mars in 1958, *Planetary and Space Sci., 2:* 26, 1959.

Dollfus, A., Étude des planètes par la polarisation de leur lumière, *Ann. d'Astrophys.,* Supp. No. 4, 1957.

Dunham, T., Spectroscopic observations of planets at Mount Wilson, in *The Atmospheres of the Earth and Planets,* edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.

Fessenkov, V. G., On the atmosphere of Mars; photometrical analysis of Wright's phenomenon, *Astronomische Nachtrichten, 228:* 25, 1926.

Giordmaine, J. A., L. E. Alsop, C. H. Townes, and C. H. Mayer, Observations of Jupiter and Mars at 3 cm. wavelength, *Astron. J., 64:* 332, 1959.

Goody, R. M., The atmosphere of Mars, *Weather, 12:* 3, 1957.

Grandjean, J., and R. M. Goody, The concentration of carbon dioxide in the atmosphere of Mars, *Astrophys. J., 121:* 548, 1955.

Herzberg, G., and L. Herzberg, Production of nitrogen atoms in the upper atmosphere, *Nature, 161:* 283, 1948.

Hess, S. L., Blue haze and the vertical structure of the Martian atmosphere, *Astrophys. J., 127:* 743, 1958.

Kallmann-Bijl, H. K., Daytime and nighttime atmospheric properties derived from rocket and satellite observations. *J. Geophys. Research, 66:* 787, 1961.

Kaplan, L. D., On the Kiess, Karrer, and Kiess interpretation of planetary spectra, in *Studies of the Physical Properties of the Moon and Planets,* Quarterly Technical Progress Report (3), Contract RM–2769–JPL, The RAND Corporation, p. 62, 1961.

Kuiper, G. P., Planetary atmospheres and their origin, in *The Atmospheres of the Earth and Planets,* edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.

Kuiper, G. P., Visual observations of Mars, 1956, *Astrophys. J., 125:* 307, 1957.

Kuiper, G. P., private communication, 1961.

Lyot, B., Recherche sur la polarisation de la lumière des planètes et de quelques substances terrestres, *Ann. Obs. Meudon, 8:* 66, 1927.

Marmo, F. F., and P. Warwick, Laboratory and theoretical studies in the vacuum ultraviolet for the investigation of the chemical physics of planetary atmospheres, Contract NASw–124, Geophysics Corporation of America, Bedford, Massachusetts, Quarterly Progress Reports of 3 August, 1960, and 3 August, 1961.

Mintz, Y., A note on the temperature of the equatorial troposphere of Mars, in *Studies of the Physical Properties of the Moon and Planets,* Quarterly Technical Progress Report (3), Contract RM–2769–JPL, The RAND Corporation, p. 81, 1961.

Oparin, A. I., *The Origin of Life on the Earth,* Academic Press, New York, 1957.

Öpik, E., The atmosphere and haze of Mars, *J. Geophys. Research, 65:* 3057, 1960.

Öpik, E. J., and S. F. Singer, Escape of gases from the Moon, *J. Geophys. Research, 65:* 3065, 1960.

Richardson, R. S., and P. E. Roques, An example of the blue clearing observed 74 days before opposition, *Publs. Astron. Soc. Pacific, 71:* 321, 1959.

Rosen, B., Origine possible de la couche violette dans l'atmosphère de Mars, *Annales d'Astrophysique, 16:* 288, 1953.

Sagan, C., The abundance of water vapor on Mars, *Astron. J., 66:* 52, 1961*a*.

Sagan, C., Is the Martian blue haze produced by solar protons?, Research Memorandum RM–2832–JPL, The RAND Corporation, 1961*b*.

Sagan, C., On the origin and planetary distribution of life, *Radiation Research, 15:* 174, 1961*c*.

Schatzman, E., Sur les particules diffusantes dans l'atmosphère de Mars, *Compt. rend., 232:* 692, 1951.

Sharanov, V. V., Photographic irradiation and its effect on photographs of planetary disks, (in Russian) *Astron. J. USSR, 27:* 116, 1950.

Sharanov, V. V., On the role of true absorption in the Martian atmosphere, *Astron. J. USSR, 34:* 547, 1957.

Sinton, W. M., Further evidence of vegetation on Mars, *Science, 130:* 1234, 1959.

Sinton, W. M., An upper limit to the concentration of NO_2 and N_2O_4 in the Martian atmosphere, *Publs. Astron. Soc. Pacific, 73:* 125, 1961.

Sinton, W. M., and J. Strong, Radiometric observations of Mars, *Astrophys. J., 131:* 459, 1960.

Slipher, E. C., An outstanding atmospheric phenomenon on Mars, *Publs. Astron. Soc. Pacific, 49:* 137, 1937.

Smith, B., and C. Tombaugh, private communication, 1961.

Spitzer, L., The terrestrial atmosphere above 300 km, in *The Atmospheres of the Earth and Planets,* edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.

Urey, H. C., The blue haze of Mars, *Astrophys. J., 128:* 736, 1958.

Urey, H. C., in *Space Age Astronomy* (Proc. International Symposium, 7 - 9 August, 1961, California Institute of Technology, Pasadena, California), edited by W. B. Klemperer and A. J. Deutsch, Academic Press, New York, in press, 1961.

Urey, H. C., and A. W. Brewer, Fluorescence in planetary atmospheres, *Proc. Roy. Soc. (London), A241:* 37, 1957.

Watanabe, K., and M. Zelikoff, Absorption coefficients of water vapor in the vacuum ultraviolet, *J. Opt. Soc. Am., 43:* 753, 1953.

Watanabe, K., E. L. Y. Inn, and M. Zelikoff, Absorption coefficients of oxygen in the vacuum ultraviolet, *J. Chem. Phys., 21:* 1026, 1953.

Wildt, R., Ozon und Sauerstoff in den Planeten-Atmosphären, *Veröffentlichungen der Universitäts-Sternwarte in Göttingen,* No. 38, 1934.

Wilson, A. G., Spectrographic observations of the blue haze of Mars, The RAND Corporation, Paper P–1509, 1958.

Wilson, A. G., The problem of the Martian blue haze, *Proc. Lunar and Planetary Exploration Colloquium, 1* (4): 33, 1959.

Wilson, A. G., private communication, 1961.

Woolley, R. v.d.R, Monochromatic magnitudes of Mars in 1952, *Monthly Notices Roy. Astron. Soc., 113:* 521, 1953.

Wright, W. H., Photographs of Mars made with light of different colors, *Lick Observ. Bull., 12:* 48, 1925.

Wright, W. H., Photographs of Mars and Jupiter taken by light of different colors during 1926, *Lick Observ. Bull., 13:* 50, 1927.

Yanow, G., A study of the Martian upper atmosphere and ionosphere, Engineering paper No. 974, Missiles and Space Systems Engineering, Douglas Aircraft Co., 1961.

Chapter IV

VENUS

A. Introduction

The planet Venus is, in one respect, the sister-planet of Earth, since it has about the same mass and radius. In nearly all other respects it appears to be different. One is tempted to use analogies drawn from experience with Earth to interpret the facts that we know about Venus, but these analogies usually fail. Next to the Moon, Venus is our closest neighbor in the solar system, and yet we are utterly baffled by many of its attributes.

After an intensive review and discussion of the Venus atmosphere the Panel did not come to any consensus, although it became clearer where the areas of controversy lay. (It was a fortunate circumstance that proponents of most of the rival theories were represented on the Panel, and each had an ample chance to defend his case and show the weaknesses of others.) As a result, it is quite impossible to state any one "best guess" about Venus at this time, but instead we must describe three alternative hypotheses or "models." Each of these models explains most of the facts as we know them, but it is only fair to say that each has weaknesses of one kind or another. The case of Venus in 1961 is an extraordinary and challenging example of a scientific riddle with a variety of partial answers, and it is all the more exciting because in the next few years we may be able to reach out through space and find the one "true" answer.

In presenting these views on Venus we will begin with a brief summary of the observational evidence that will be used as the basis for the alternative models; these are the loosely fitting pieces of the picture puzzle. Reference is made to Chapter I and to the various appendices for the explanations of the physical principles underlying these observations, and to Sagan (1961) for a further discussion of their interpretation, and for more detailed literature citations. Then we will proceed to outline each model for the Venus atmosphere, and for each model we will mention some of the objections that have been raised against it.

B. Pertinent Observations of Venus and Their Interpretation

1. Albedo

The diffuse spherical reflectivity or albedo of Venus is between 0.55 and 0.90 in the visible and photographic infrared, and varies markedly with wavelength. In the ultraviolet it is somewhat lower and topographically more variable; i.e., intensity gradations and distinct cloud patterns are more discernable in ultraviolet than in visible light on the illuminated part of the planetary disk.

2. Clouds

Seen most clearly in ultraviolet light, the cloud patterns are diffuse and changing. There are claims that a definite banded structure exists in these ultraviolet patterns, although there is not general agreement on this point. In the visible very faint markings are seen, and there is some evidence that these radiate from the subsolar point (Dollfus, 1955). (v. also Tombaugh's description, Appendix 2.)

The nature of the visible clouds is still a hotly disputed question. The two most likely possibilities—that the clouds are water droplets or ice crystals, or that they are dust stirred up from the surface—are each associated with different models of the structure of the atmosphere of Venus and

of the origin of the microwave emission. The models are discussed below. No observations of laboratory samples have precisely matched the Venus polarization curve of Lyot (1929); the closest fit in white light is with 2-micron-diameter water droplets. However, it is possible that other substances could also fit the polarization data. Measurements of the polarization as a function of wavelength will help to narrow further the field of choice.

Observations have been made of relatively stationary features in the cloud patterns over periods of weeks. If it is assumed that surface features are being observed through breaks in the clouds, periods of rotation are derived which are quite long and which approach the period of revolution about the Sun, 225 days. However, an alternative explanation of these cloud patterns has been suggested; namely, that the cloud patterns are determined by the position of the Sun and that the planet moves underneath them.

A period of rotation between 10 and 30 days has frequently been advocated. This period was first proposed by Ross (1928) to explain by planetary rotation the banded cloud patterns he photographed in ultraviolet light. The absence of Doppler shift at the limbs of the planet indicates that the period of rotation is greater than about five days. (The most reliable indications of the Venus rotation rate will probably come eventually from radar—see below.)

3. *Indications of Temperature*

Observations of the occultation of Regulus (Menzel and de Vancouleurs, 1960) showed a scale height of 6.8 ± 0.2 km and a logarithmic gradient of scale height $(1/H)(\partial H/\partial z)$ of 0.010 ± 0.002 km^{-1} at an altitude of about 60 ± 10 km above the top of the visual cloud layer at the limb (revised value, de Vaucouleurs, 1961.) The pressure at this level is estimated to be 2.6 ± 0.13 dynes cm^{-2}, and the total mass of gas above it 3.0×10^{-3} gm cm^{-2} (a reduced thickness of 1.7 cm-atm).

The vibration-rotation bands of CO_2 near 8000 Å (0.8μ) were observed in absorption in the spectrum of sunlight scattered from Venus (Chamberlain and Kuiper, 1956), and an "effective" temperature of about 285°K was deduced for the absorbing CO_2. The results showed appreciable variation between observations, and a systematic variation with phase.

The CO_2 rotational temperature in the 8000 Å region is interpreted by Chamberlain and Kuiper (1956; v. also Appendix 9 by Chamberlain) to apply to a region well below the top of the visible cloud layer. This conclusion is based on an assumption made in the radiative transfer model of the atmosphere of Venus from which the rotational temperature of 285° K was extracted, namely, that the *scattering* particles have about the same distribution or scale height as the *absorbing* CO_2 molecules. The visible cloud layer would then be regarded as the altitude at which the optical thickness due to scattering is about unity in the continuum outside the absorption bands in the visible. Chamberlain suggests that the 8000 Å bands are formed several optical thicknesses below this altitude, and therefore show a characteristic temperature greater than that appropriate to the top of the visible cloud layer (it being considered that temperature decreases with increasing height).

The assumption of Chamberlain and Kuiper, that the distribution of scatterers and absorbers is the same, is consistent with the phase variation of intensity of the CO_2 bands. If there were a single scattering layer with a sharp upper boundary, then as the Sun departed from the vertical the optical path above this layer would increase and the absorption by the CO_2 above the layer would be greatest at grazing incidence. However, the observations show minimum absorption at the terminator, a result which is predicted by the radiative transfer treatment of Chamberlain and Kuiper.

On the other hand, the assumption of parallel distribution of scatterers and absorbers is not entirely consistent with observations made of the extensions of the cusps, which demonstrate a sharp upper boundary to the visible cloud with a scale height much less than that appropriate to the gas (Russell, 1899). Therefore, it is interesting that Kaplan (1961a, 1961b) has shown that the phase variation is not necessarily incompatible with a thin scattering layer high in the atmosphere and a more efficient scattering layer below. This model will be discussed further below. On the two-layer model, the rotational temperature deduced from the 8000 Å bands will be considerably greater than the 285° K previously quoted.

In some of the detailed models of the atmosphere of Venus which are illustrated in the following sections, the concept of two scattering layers is adopted. One unsatisfactory feature of two-layer models is that the material responsible for the upper haze layer and its peculiar properties is not identified. According to Kaplan (1961) the thin upper cloud *is* the visible cloud. It is an effective scatterer in the visible, but transparent in the near infrared, including the 1.6μ region, and completely opaque in the ultraviolet and far infrared. The question of this upper layer has prompted a great deal of discussion, and one might remark that the *ad hoc* Panel came up with an *ad hoc* layer.

Sagan (1961) has suggested that CO_2 photodissociation near the occultation level may lead to the formation of a carbon suboxide polymer haze at that altitude. The suggestion that $(C_3O_2)_n$ may be formed at lower altitudes had previously been made by Kuiper (1957) and by Sinton and Strong (1960). Such a haze absorbs strongly in the ultraviolet and far infrared, and is relatively transparent in the visible and near infrared. Weak $(C_3O_2)_n$ absorption in the visible would give a yellowish coloration, and Venus is known to have a lemon yellow tint. However, the albedo of such a cloud is probably too low to explain the high albedo of Venus. Except for this possible difficulty, a carbon suboxide haze has all the required characteristics of the upper haze layer.

The far-infrared (thermal) emission from Venus, observed through the window in the Earth's atmosphere between 8 and 13μ, indicates an effective temperature of the emitting layer of 234°K, which does not vary by more than 5° to 10° over the planetary disk or between dark and sunlit hemispheres, although there is some decrease toward the limbs (Sinton and Strong, 1960). The 8- to 13-micron emission is generally believed to arise from an atmospheric cloud or haze layer on Venus; but whether this level is the same as that of the visible clouds is still an open question.

A major revision of recent thinking about Venus has been forced by the discovery that Venus is a strong source of microwave radio emission (Mayer, McCullough, and Sloanaker, 1958). This fact is the basis on which are founded the three models of the Venus atmospheric and surface conditions which will be described in the following pages.

Some effective brightness temperatures near inferior conjunction deduced from microwave radio emission are as follows (v. Burke, Appendix 3, Table 2, for references and explanations):

Wavelength (cm) (Agency)	*Average Temperature (°K)*
0.80 (Lebedev)	315 ± 70
0.86 (N.R.L.)	410 ± 160
3.37 (N.R.L.-Columbia)	575 ± 58
3.4 (N.R.L.)	580 ± 57
10.3 (N.R.L.)	600 ± 65

There is an increase of about 50°K in the 3 cm effective temperature as the planet moves from inferior conjunction to dichotomy, and perhaps an even larger change with phase at 8 mm. Lilley (1961) has shown that there is also some evidence for an increase of temperature with phase at 10 cm.

4. Indications of Composition

Above the visible clouds it has been estimated (Herzberg, 1952; Kaplan, 1961a; Öpik, 1961) that there is between 100 and 1000 m-atm of CO_2. (1 m-atm = 1 meter of CO_2 reduced to standard pressure and temperature.) This wide variation is due to various interpretations of the observed absorption of the scattered sunlight in the near infrared, below 2μ. Öpik (1961) states that in the ultraviolet (3650 Å) and infrared beyond 1.20μ the depth of CO_2 deduced by measuring the absorption of solar radiation is less by a factor of 5 to 10; these absorption features would then arise from the region above the upper haze layer.

On the other hand, Kaplan (1961) states that these differences in carbon dioxide abundances may be unreal, and suggests that they can be attributed to incorrect interpretations of pressure-broadening effects. He obtains a CO_2 composition of about 20% by volume; the total pressure at the cloud tops is then about 90 mb, if the remainder of the atmosphere is composed of N_2.

In either case the amount of CO_2 on Venus greatly exceeds the amount on Earth (about 220 cm-atm).

One attempt has been made to observe the H_2O content of the Venus atmosphere (Strong, 1960) by measuring the absorption of reflected sunlight at 1.13μ from a high altitude manned balloon (flown by Ross and Moore in November 1959). It was estimated that there were 19μ of precipitable water vapor in the Venus atmosphere above the reflecting clouds. Due to the uncertainty in the correction factors for residual water vapor above the balloon in the Earth's atmosphere and for pressure broadening and other systematic errors, however, this estimate probably should be treated as a rough value, and may be considerably more or less.

5. Radar Observations

In the interval between the meetings of the Panel and the completion of this report several highly successful radar contacts with Venus have been reported in the United States, the Soviet Union, and Great Britain. As pointed out in Chapter I, and in Appendix 4 by Eshleman, from radar observations it is possible to determine several important planetary characteristics, especially roughness, reflectivity, and rate of rotation. Earlier results by the Millstone Hill radar facility of the Massachusetts Institute of Technology showed great variability from year to year in reflectivity—a result of perhaps not very great significance, since the returned signals were very weak. In the 1961 contacts, however, observations by the Goldstone radar facility of the California Institute of Technology were made with larger signal-to-noise ratios (greater than 18 db) and a value for the radar reflectivity at 2388 Mc (12.5 cm) of 0.1 was obtained. (It is understood that M.I.T. also obtained better signals in 1961, but the results have not been released at the time of this writing.) Furthermore, by measuring the degree of polarization of the returning pulse the Venus disk was determined to be effectively "rough," according to the Goldstone results. (Reversing receiver polarization reduced the signal by 11 db, which is about the same as the depolarization of the Moon's disk averaged over its whole surface.)

Scientists in both the United States (Rechtin, 1961) and the Soviet Union (Kotelnikov, 1961a, 1961b) have recently measured the Doppler width of radar pulses returned from Venus. The derived rotation period will depend to some extent on the assumed orientation of the axis of rotation of Venus; it will depend very critically on the assumed surface roughness of the planet at the frequencies used. If Venus is very smooth at the radar frequency, the major pulse return will arise from the center of the disk, where the motion is normal to the line of sight; thus, the period of rotation derived from the specular component will be overestimated. If Venus is very rough at the radar frequency, there will be an appreciable contribution to the returned signal from the limbs, where the motion is parallel and antiparallel to the line of sight. However, in practice, the return from very near the limb is usually lost in the noise, because of the decline in scattered intensity at the limb for a scattering law proportional to $(\cos \theta)^n$, where θ is the angle between the line of sight and the planetary vertical, and n is a positive number. Thus, in both cases, the true period of rotation will be less than the value derived without taking these factors into consideration.

In general, when plotting the amplitude of the returned pulse against the frequency we expect the returned pulse to have a central spike, corresponding to specular reflection from the smooth parts of the surface near the center of the disk, superposed on a broader "mound," corresponding to diffuse reflection from the rougher parts of the surface near the limbs. Both the American and the Soviet results demonstrate these general features. The Goldstone pulse returns at 12.5 cm show a spike of width about 6 cps at the half-power points, superposed on a mound of width less than 30 cps, and of energy about one order of magnitude less than the spike. The Soviet pulse returns at 43 cm show a spike with a width of a few cps, superposed on a mound with a width of about 400 cps and of energy about two orders of magnitude less than the spike. The United States results, based primarily on the spike at 12.5 cm, and the roughness deduced from observations of depolarization, give a period of rotation comparable to the period of revolution—225 days. The Soviet results at 43 cm, based primarily on the "mound," give a period of rotation less than 11 days. While the difference in results might, to some extent, be due to an apparent difference between the surface roughness of Venus as observed at 12.5 cm and at 43 cm, it is unlikely that such a difference can account for

the bulk of the discrepancy. Resolution of some of these difficulties may follow analysis of the 68-cm radar observations of Venus, made at the Millstone facility of M.I.T. at the 1961 inferior conjunction. Otherwise, this curious international disagreement may persist until Venus comes close again in late 1962, unless it can be arbitrated before then.

C. Alternative Models for the Venus Atmosphere

1. The Greenhouse Model

One of the central problems in understanding the Venus atmosphere is the explanation of the high brightness temperature as observed between 3 and 21 cm by radio telescopes. In this first model it is considered that the 600°K deduced by these measurements actually exists at the planet's surface, and that the high temperature is maintained by a very efficient "greenhouse effect," i.e., by direct solar heating. The basic argument, as developed by Sagan (1960), is that Venus may have a cloud layer and atmosphere that is relatively transparent to sunlight in the visible and near infrared, but that its atmosphere has enough H_2O and CO_2 to be nearly opaque to far-infrared (thermal) emission from the surface. Sagan estimates that the required degree of heating can be achieved if the far-infrared absorptivity of the atmosphere is about 0.99 (or has one per cent infrared transmission), and claims that this is achieved in a deep atmosphere with high CO_2 content and from 1 to 10 gm/cm^2 of H_2O. (As will be shown, this model requires a deep atmosphere for other reasons as well.) Such an amount of water, mixed through the atmosphere, would result in ice crystal clouds at a temperature of 220 to 235°K, roughly the observed thermocouple temperatures; it would not result in water or ice clouds lower in the atmosphere. Very recently, Sinton (1961) has obtained spectra of Venus in the 1- to 4-micron region. He finds absorption features characteristic of ice crystals, and concludes that the visible cloud layer is composed in part at least of cirrus, in agreement with the prediction of the greenhouse model.

An atmosphere heated at the surface and without enough water vapor to develop deep clouds will assume a dry adiabatic lapse rate,* and for a CO_2 and N_2 atmosphere and an acceleration due to gravity of 870 cm sec^{-2} this dry adiabatic lapse rate ($\partial T/\partial z = -C_p/g$) is as follows:

$\frac{CO_2}{N_2}$ *Mixing Ratio*	*Adiabatic Lapse Rate*
.80	10.5K°/km
.40	9.2
.20	8.7

For our purposes, a nominal 10K°/km will be adopted, and it is assumed that this neutrally stable condition exists between the surface and the top of the cloud layer. It could hardly be any larger, and if it were smaller (as on Earth, where below the tropopause it is about half this on the average) the atmosphere would be even deeper than indicated in this model. Thus, the layer from which the 234° K thermocouple temperature arises—on the greenhouse theory this is also the visible cloud layer —is $(600 - 234)/10 \simeq 37$ km above the surface.

Öpik (1961) has criticized the greenhouse theory on the grounds that the presumed amount of CO_2 above the surface, plus the small amount of water apparently detected by Strong's (1960) balloon flight, will not give an infrared atmospheric opacity anywhere near the required value. Sagan (1960) maintains that, on the basis of integrated absorptivities available in the boiler and furnace technology literature, the presumed amount of CO_2 plus 1 to 10 gm cm^{-2} of water vapor, and a surface temperature of 600° K together give the required opacities. He states that 1 to 10 gm cm^{-2} of water vapor above the surface is not grossly inconsistent with the observed 2×10^{-3} gm cm^{-2} of water vapor

* An "adiabatic lapse rate" of temperature is one in which a parcel of air carried upward or downward and cooled or heated adiabatically in the process will always have the same temperature as the ambient air at the same level. Therefore, it is the maximum lapse rate that an atmosphere can have without being unstable.

some 37 km above the surface. Detailed computations of the expected infrared spectrum transmitted through the absorbing gases and ice-crystals in the atmosphere of Venus are in progress, and the resolution of this controversy is awaited with interest.

Mintz (1961a) has also criticized the greenhouse theory on the grounds that large-scale circulations will tend to reduce the temperature difference between the high atmosphere and surface. Sagan (1961) pictures a very calm lower atmosphere, where energy transport is primarily radiative. Calculations are planned by Mintz to test the possibility that substantial temperature equalization must occur due to general circulation. (v. Appendix 8 by Mintz.)

Figures 1 and 2 summarize the other features of the greenhouse model more succinctly than words. In Fig. 2 are shown the levels from which various wavelengths of sunlight are supposedly *reflected*, and the levels from which various longer wavelengths (infrared and microwave) are thermally *emitted*.

From the greenhouse model one would expect the surface of Venus to be dry, calm, overcast, and hot as an oven. The Sun would be visible through the high cirrus clouds, and possibly some-

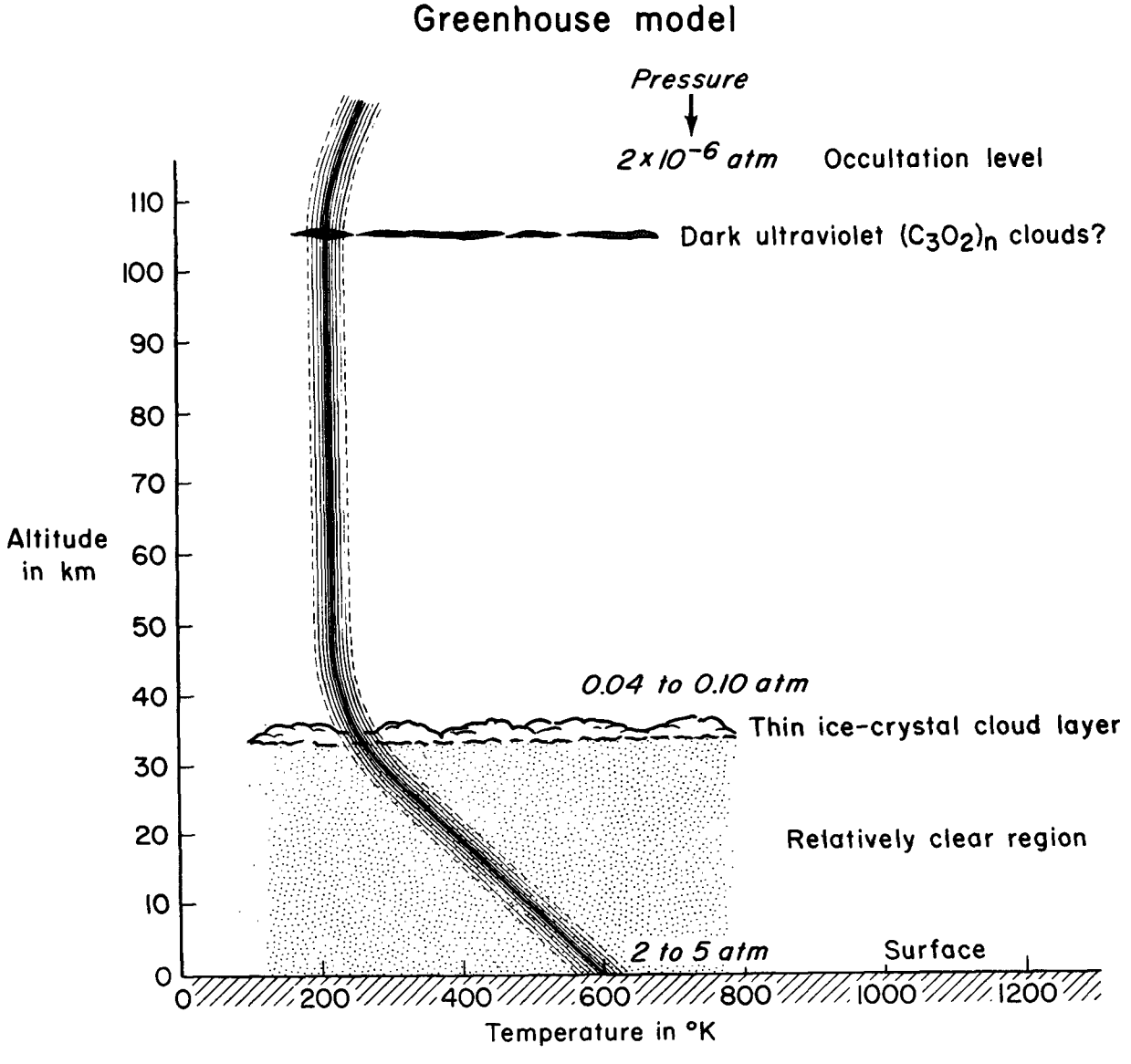


Figure 1

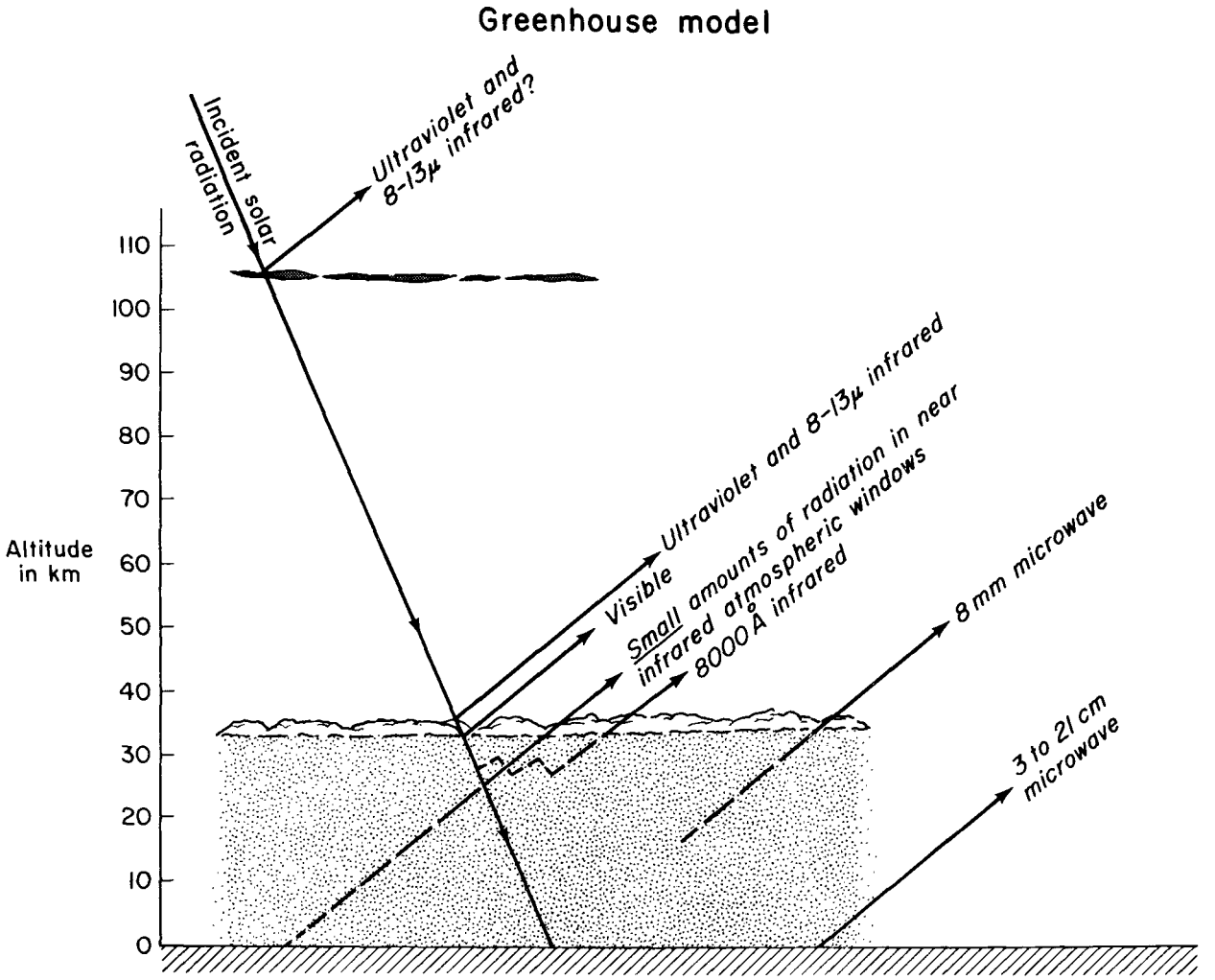


Figure 2

what reddened by dust in the lower atmosphere. The chances for life forms based on familiar terrestrial biochemistry seem very remote indeed under these conditions.

2. The Aeolosphere Model

Öpik (1961) rejects the greenhouse theory, and suggests that the atmosphere below the visible clouds is an extremely dry, dusty region that is kept in motion and stirred by the winds above the clouds. The winds transfer momentum downwards. Since an atmosphere that is stirred will maintain an adiabatic lapse rate, the temperature will increase with increasing depth below the cloud tops. The winds at the surface will produce a small amount of frictional heating, enough to account for the small amount of energy that is lost by radiation. The loss by the surface is small, because the dust in suspension makes the atmosphere a virtually opaque blanket, transparent only to radio waves of wavelength about 3 cm or greater.

The name "aeolosphere" (coined by Öpik) means, loosely translated, "region of the winds," since the surface is heated by the hot dry winds driven from far above. The winds raise a perpetual dust storm, and there are no rains to cleanse the atmosphere, as on Earth. The dust particles, perhaps made of calcium and magnesium carbonates, ground together during the long history of Venus, are as fine as talcum powder, and remain suspended almost indefinitely. At the surface of this incredibly dreary aeolosphere there is no sunlight, only dust and heat and wind.

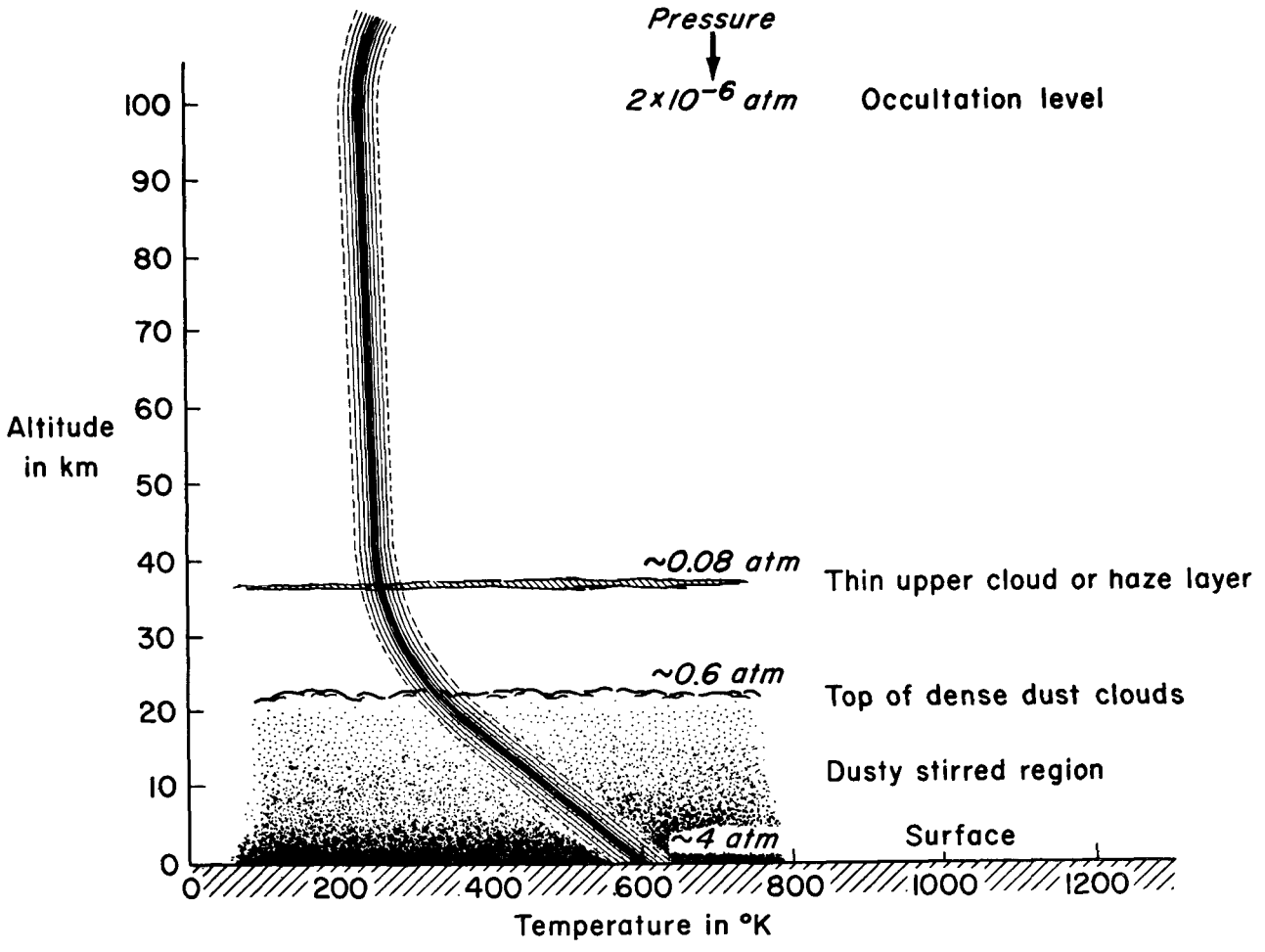


Figure 3

The accompanying Figures 3 and 4 summarize the properties of the aeolosphere model. It will be noted that Öpik's deductions about the structure *above* the cloud layers are not greatly different from that of the greenhouse model. He suggests that the visible radiation comes from a lower dust cloud layer, while the ultraviolet and far-infrared radiation arises from an upper haze layer of unspecified composition. As described in Sections B.3 (and C.1) above, a carbon suboxide polymer cloud layer at the occultation level may have the required characteristics of the haze layer; since the region above 37 km is approximately isothermal on the aeolosphere model (as in the greenhouse model), placing the upper haze layer at the occultation level changes few of its radiative properties.

Öpik (1961) says: "The aeolosphere has no counterpart in the meteorological structure of the terrestrial atmosphere." However, Mintz (1961b) has suggested that perhaps, in one sense, there is a terrestrial counterpart. In the aeolosphere the winds are generated directly in the region above the cloud deck, and so this region is like the troposphere on Earth. These "tropospheric winds" drive, by frictional interaction, a fluid of about eight times more mass than in the equivalent troposphere. Is there an analogy to this on Earth? There may be (though analogies are dangerous), since the Earth's troposphere maintains a stirred region in the oceans known as "the mixing layer," and this mixing layer extends down to the pycnocline, about 100 m below the ocean surface. Thus, the winds on Earth keep a region of roughly 10 times the mass of the atmosphere stirred.

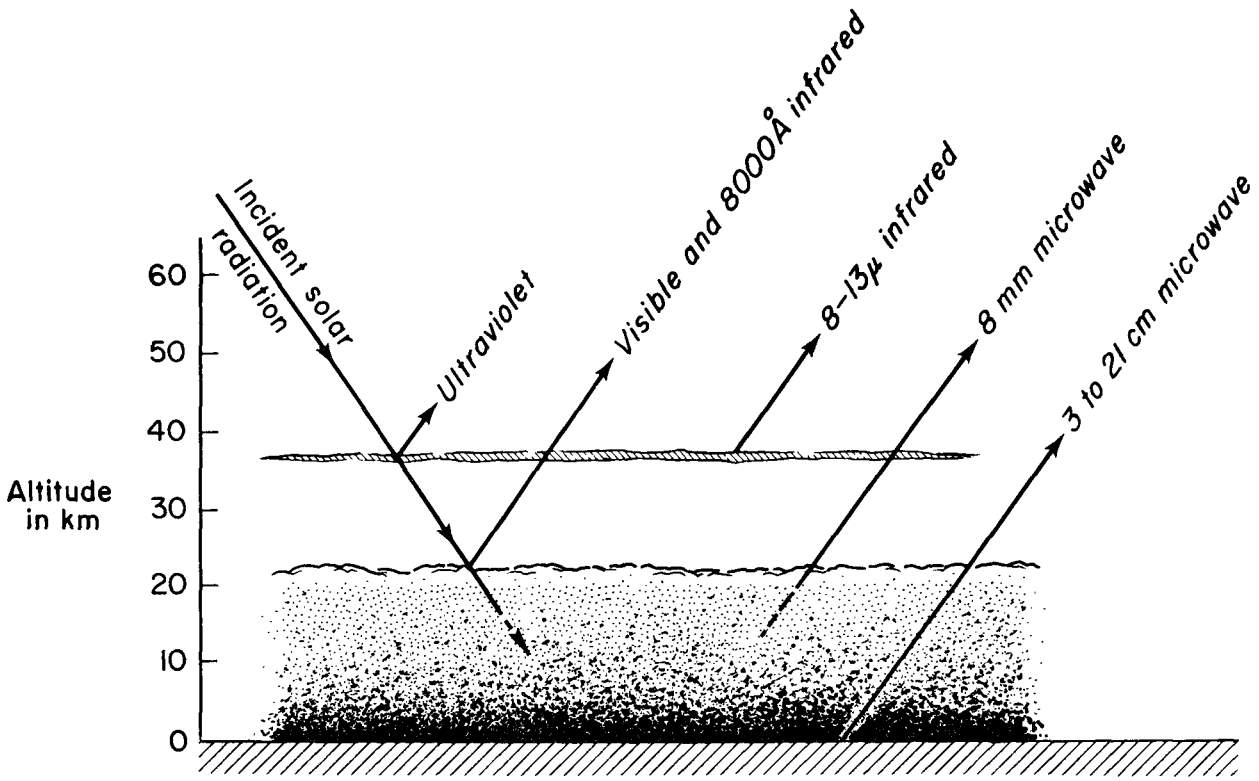


Figure 4

There are a number of unanswered questions raised by the aeolosphere model. First, Öpik (1961) does not offer a description of the atmospheric circulation in the aeolosphere, and a detailed working out of the wind patterns in the thermally-driven upper layers and the frictionally-driven lower layers would be of great interest. Something may be inferred from the general theory of the circulation of planetary atmospheres. (v. Chapter II, and Appendix 8 by Mintz). The winds that drive the aeolosphere—the atmospheric heat engine—are in the region above the visible clouds, a region similar to the terrestrial troposphere. These winds may be fairly regular, since the circulation is in a "Hadley regime" (v. Chapter II) due to the relatively slow rotation rate of Venus. This suggests that there will be a pattern of inflow near the cloud tops in the vicinity of the subsolar point and subsidence on the dark side. This flow must be such that the region of the upper haze layer remains at a constant temperature (though not necessarily at a constant height), as observed by Sinton and Strong (1960). The Panel did not resolve the question of atmospheric circulation, and clearly much more remains to be added to this aspect of the aeolosphere model.

A second question concerns the chemical nature of the clouds of dust stirred up from the surface. The dust clouds must give the observed high albedo on the aeolosphere model. Öpik (1961) suggests that calcium and magnesium carbonates—which have the desired high albedos—are the principal constituents of the cloud layer. Urey (1952) has pointed out that the primary source of carbonates on the Earth is a reaction of silicates with carbon dioxide in the presence of liquid water. In the absence of liquid water, the reaction proceeds much more slowly, and, indeed, the absence of this reaction has been invoked by Urey to explain the great abundance of carbon dioxide on Venus. Öpik (1961) feels that even in the absence of equilibrium conditions, great quantities of silicates will have reacted to form carbonates during geological time. However, if substantial quantities of low albedo

silicates remain, the high albedo of Venus will not be explained. Indeed, Urey and Brewer (1957) have argued that there is no mechanism for preferentially producing high albedo dust clouds; they point out that terrestrial dust clouds in hot arid regions are characteristically dark brown. A way out of this possible difficulty may be mentioned: regardless of the absorption properties of large dust particles, if the particle sizes are smaller than $\lambda/2\pi$, where λ is the wavelength at which the observations are made, the particles will appear white if only their number density is sufficiently great. In a very dry wind-stirred upper atmosphere, it is at least not out of the question that it is primarily populated by very fine dust particles of diameter $\leqq$ 1000 Å. However, the polarization data of Lyot (1929) suggest that this is not the case, and that the particles are micron-sized; but we have already mentioned the possibility of alternative fits to the polarization phase curves.

Sinton's (1961) interpretation of his recent infrared spectrophotometry of Venus has supported the hypothesis that the near infrared radiation from Venus arises from an ice crystal cloud layer; it is in apparent disagreement with that aspect of the aeolosphere model which has the *near*-infrared radiation arise from the lower dust clouds.

A final possible difficulty with the aeolosphere model concerns the phase variation of the microwave emission. The lower atmospheric layers, near the surface, are in almost total darkness on the aeolosphere model. The surface temperature is only indirectly controlled by the solar insolation through the circulation in the upper atmosphere, and the response of the surface temperature to the day-night cycle is expected to be small and extremely sluggish. The near-equality of day and night thermocouple temperatures must be interpreted on similar grounds in the aeolosphere model. It therefore appears that the aeolosphere model may not be consistent with the strong phase variation reported at 8 mm by Kuz'min and Salomanovich (1960), and at centimeter wavelengths by Lilley (1961). The centimeter phase variation is especially relevant, since the centimeter emission must arise from the surface on the aeolosphere model. Observations have been made only for a few months around inferior conjunction, and extrapolation to daytime brightness temperatures is a risky maneuver. Nevertheless, Lilley shows convincingly that some phase effect lies in the NRL observations and that the minimum temperature occurs near inferior conjunction, as is also the case for the 8-mm observations. His least square extrapolation of the centimeter data suggests a brightness temperature at superior conjunction of 1000° K or more. If direct observations of the illuminated hemisphere of Venus at 3 cm or longer confirm this strong interhemispheric temperature gradient, and brightness temperature minima at inferior conjunction, it will tend to cast some doubt on the correctness of the aeolosphere model.

3. *The Ionosphere Model*

Each of the preceding models has several questionable features, as has already been discussed. There is also a lingering hope that, somehow, Venus will turn out to be more like its companion Earth than is suggested by these models.

At present there seems to be only one ray of hope left for low temperatures and a habitable surface, and this is rather dim. There is a possibility that the high apparent temperature (600°K) measured at centimeter wavelengths may refer to the ionosphere. The argument is as follows: If a planetary ionosphere is highly ionized and contains a large concentration of free electrons throughout a considerable depth, then it will be opaque to long radio waves and transparent to shorter radio waves. (Cf. Chapter I. B.3.) Calculations of free-free transition radio emission by Jones (1961) and by Sagan, Siegel, and Jones (1961), in which a moderate surface temperature of about 300°K and a 600°K ionosphere temperature are assumed, show that the microwave spectrum at inferior conjunction can be explained if the integral through the ionosphere $\int n_e^2 dz$ is to 4×10^{25} electrons cm^{-5}. Here n_e is the electron number density. (v. also Appendix 3 by Burke for a discussion of this point.) The observed phase effects can also be explained if the integrated electron density is about an order of magnitude greater in th alluminatd hemisphere, and if the electron tmeperature is greater there.

Kuz'min and Salomanovich (1961) have recently reported rapid day-to-day variations in the Venus microwave emission at 9.6 cm, which, these authors maintain, are free from instrumental and statistical limitations. If true, such observations provide strong support for the ionospheric model: it

can easily be imagined that day-to-day changes occur in the electron density of a planetary ionosphere. But day-to-day changes of several hundred degrees Kelvin in the temperature of the dark side of a slowly rotating planetary surface is much more difficult to explain. However, observations made at neighboring frequencies (Drake, 1961; Mayer, 1961) do not support the observations of Kuz'min and Salomanovich, and the reality of the phenomenon remains unsettled.

If the thickness of the Venus ionosphere is comparable to the Earth's (150-300 km), then we can say that the required electron density in this ionosphere is about 10^9 cm^{-3} on the dark side, and somewhat greater on the bright side. But is such a high electron density reasonable? In the terrestrial ionosphere the electron density in the most highly ionized region, the F_2-region, rarely is much greater than 10^6 cm^{-3}. The basic relationship governing the electron density in such a region when there is a balance between the rate of electron production (q) and the rate of electron recombination (αn_e^2) is

$$\frac{dn_e}{dt} = q - \alpha n_e^2 = 0$$

where α is the recombination coefficient. Thus, we could achieve the large required value of n_e by either a very large value of q or a very small value of α. Some values corresponding to an n_e of 10^9 cm^{-3} are given in Table I,

Table I

q(cm^{-3} sec^{-1})	α (cm^3 sec^{-1})
10^4†	10^{-14}
10^8	10^{-10}
10^{10}	10^{-8}†

where the values marked with † are typical values for the F-region of the Earth's ionosphere. Since the ionization in the terrestrial ionosphere is mostly due to absorption of solar ultraviolet and X-ray radiation, and since the solar flux at Venus is only about twice that on Earth, it seems that the proximity of Venus to the Sun will not account for the difference. We must try to imagine an entirely different and more powerful source of ionization, or a much lower recombination coefficient, α.

While exact values of α for various specific recombination processes are often in doubt, there is little doubt about their order of magnitude for the most likely main constituents of the Venus ionosphere, CO_2^+, CO^+, N^+, N_2^+, O^+, and O_2^+. There are three main types of recombination processes:

Radiative recombination (for atomic ions)

$$e^- + X^+ \rightarrow X + h\nu, \quad \alpha \approx 10^{-12} \text{ cm}^3 \text{ sec}^{-1}$$

Dissociative recombination (for molecular ions)

$$e^- + XY^+ \rightarrow X^* + Y^{**}, \quad \alpha \approx 10^{-8} \text{ to } 10^{-6} \text{ cm}^3 \text{ sec}^{-1}$$

Three-body recombination

$$e^- + X^+ + M \rightarrow X + M' \text{ (rate negligible at ionospheric pressures)}$$

(Note that, for a free electron to unite with a positive ion by either of these three processes, an amount of energy equal to the ionization potential must be released, either as radiation or as kinetic energy (heat).) The important point here is that if molecular ions are present (N_2^+, CO^+, or CO_2^+), even in small quantities, they will recombine with free electrons by the relatively rapid process of dissociative recombination. It is possible but unlikely that the upper atmosphere of Venus is composed entirely of a monatomic gas. A probably more realistic assumption is that the recombination coefficient in the Venus ionosphere is comparable to that in the terrestrial ionosphere, about 10^{-6} to 10^{-8} cm^3 sec^{-1}—possibly 10^{-9} cm^3 sec^{-1}, as a lower limit. (v., e.g., Nawrocki, 1961.)

Turning now to the question of free electron production we must seek an ionization mechanism that is many orders of magnitude more effective than solar ultraviolet and X-rays. A suggestion has been made (Jones, 1961) that possibly solar protons could be a source. We will expand his argument (to give it every benefit of the doubt) and try to maximize it in order to see whether it could possibly be the answer.

On Earth ordinary solar protons cannot penetrate directly into the atmosphere, since they are deflected by the geomagnetic field and are either forced to pass on, are trapped in the outer radiation belt, or are channeled downwards at high geomagnetic latitudes into the auroral zones. Thus, except in the auroral zone, proton streams from the Sun are quite unimportant in establishing the normal terrestrial ionosphere. Suppose, however, that Venus has a very small magnetic field strength (about one thousandth that of the Earth's or less), so that the solar protons would not be deflected appreciably and would stream directly into the upper atmosphere. (Venus *could* have a small field, since, according to some current theories, the Earth's magnetic field is due to a combined effect of rotation and convection in the liquid core. If Venus rotates much more slowly than the Earth, as seems to be the case, then it may have a correspondingly weak magnetic field.)

An upper limit to the solar proton flux—actually several orders of magnitude larger than some would allow—may be that estimated by Bierman (1957), based on observations of the acceleration of comet tails. An upper limit to this deduced proton flux is 10^3 protons cm^{-3} near the Earth traveling at a speed of 1000 km sec^{-1} (an energy of roughly 1 keV, or a flux of 10^{11} protons cm^{-2} sec^{-1}. Suppose this stream of protons pouring into the atmosphere is stopped in a layer 100 km thick, each proton using its 1 keV of energy to cause ionization. The average energy loss per ionization for such particles is 30 eV. The rate of ionization, then, under these highly optimistic conditions, would be

$$q = \frac{E(eV) \cdot \text{Flux}}{30\,\Delta h} = \frac{10^3 \cdot 10^{11}}{30 \cdot 10^7} = 3 \times 10^5 \text{ electrons cm}^{-3} \text{ sec}^{-1}$$

Comparing with values of q required in Table I above, it can be seen that we have come up with an ionization mechanism some 30 times more effective than ultraviolet radiation is in our own ionosphere. With radiative recombination coefficients, the required electron densities can just barely be maintained by the solar proton wind; with the probably more relevant dissociative recombination coefficients, the electron density falls several orders of magnitude below that required to explain the microwave emission. With more realistic values of the solar proton flux, the situation is even worse.

However, it must be pointed out that we are dealing with a complex question in which multiple uncertainties exist—the value of the solar proton (and electron) flux at Venus' distance from the Sun, the composition of the upper atmosphere, the possibility of still another unsuspected ionization mechanism, etc. Thus, the Panel could not definitely rule out the ionosphere model, and it is presented as the third possibility.

If we grant this very dense ionosphere, there are certain other factors that must still be explained. First, if it is optically thick at 3 cm it will certainly be opaque and almost completely absorbing at the 12.5 cm and longer wavelengths that have been used in radar contacts with Venus, since the opacity is proportional to the square of the wavelength. (v. Chapter I.B.4, and Appendix 4 by Eshleman.) How could the ionosphere simultaneously be opaque enough to emit as a blackbody and transparent enough to permit a radar pulse to be reflected from the surface? At this point our delicate flower seems almost wilted; we can keep it alive for a little longer by another *ad hoc* hypothesis. The explanation may lie in the fact that these radar contacts have so far been made only at inferior conjunction, when the radar pulse would be reflected primarily from the antisolar or midnight point of Venus. It would be expected that the ionosphere would decay during the long Venus night, and would only be replenished by horizontal diffusion of free electrons from the daylight side (the side exposed to the bombardment of solar protons). Thus, there could be a hole in the ionosphere near the antisolar point through which the radar pulse could penetrate to the surface. Actually, a reduction of the electron density by a factor of about ten in the hole would be sufficient to account for the Millstone 68-cm radar contact, and by a factor only two for the Goldstone 12.5-cm radar.

If the ionospheric model is correct, then the surface is below the boiling point of water, at least in the night hemisphere and possibly in the day hemisphere as well. The visible clouds would then be only some 15 km above the surface or less, not unlike terrestrial cirrus clouds. On a two-layer model the 8000 Å radiation would arise from very near the ground; on a one-layer model somewhat higher. This model is sketched in Figures 5 and 6.

The great abundance of carbon dioxide must be attributed to the failure of the Urey equilibrium which, on Earth, maintains a steady state partial pressure of CO_2 by reaction of silicates to carbonates. One reason for failure of the Urey equilibrium would be the absence of liquid surface water (Urey, 1952); and this is also the explanation appropriate for the greenhouse and aeolosphere models. An alternative and much more intriguing reason is that set forth by Menzel and Whipple (1955), who suggest that Venus is covered with a planet-wide ocean. The ocean is supposed to prevent contact of outgassed CO_2 with silicates, except for a few carbonate-encrusted islands. Menzel and Whipple showed that such an ocean is still within the spectroscopic tolerances for atmospheric water vapor on Venus.

It will be noted that an important difference between the first two models of the Venus atmosphere (greenhouse and aeolosphere) and the last model (ionospheric) lies in their respective depths, or total mass of atmosphere. It is pertinent that Dole (1956), assuming that outgassed CO_2 remained in the atmosphere of Venus in the failure of the Urey equilibrium, instead of being sedimented as

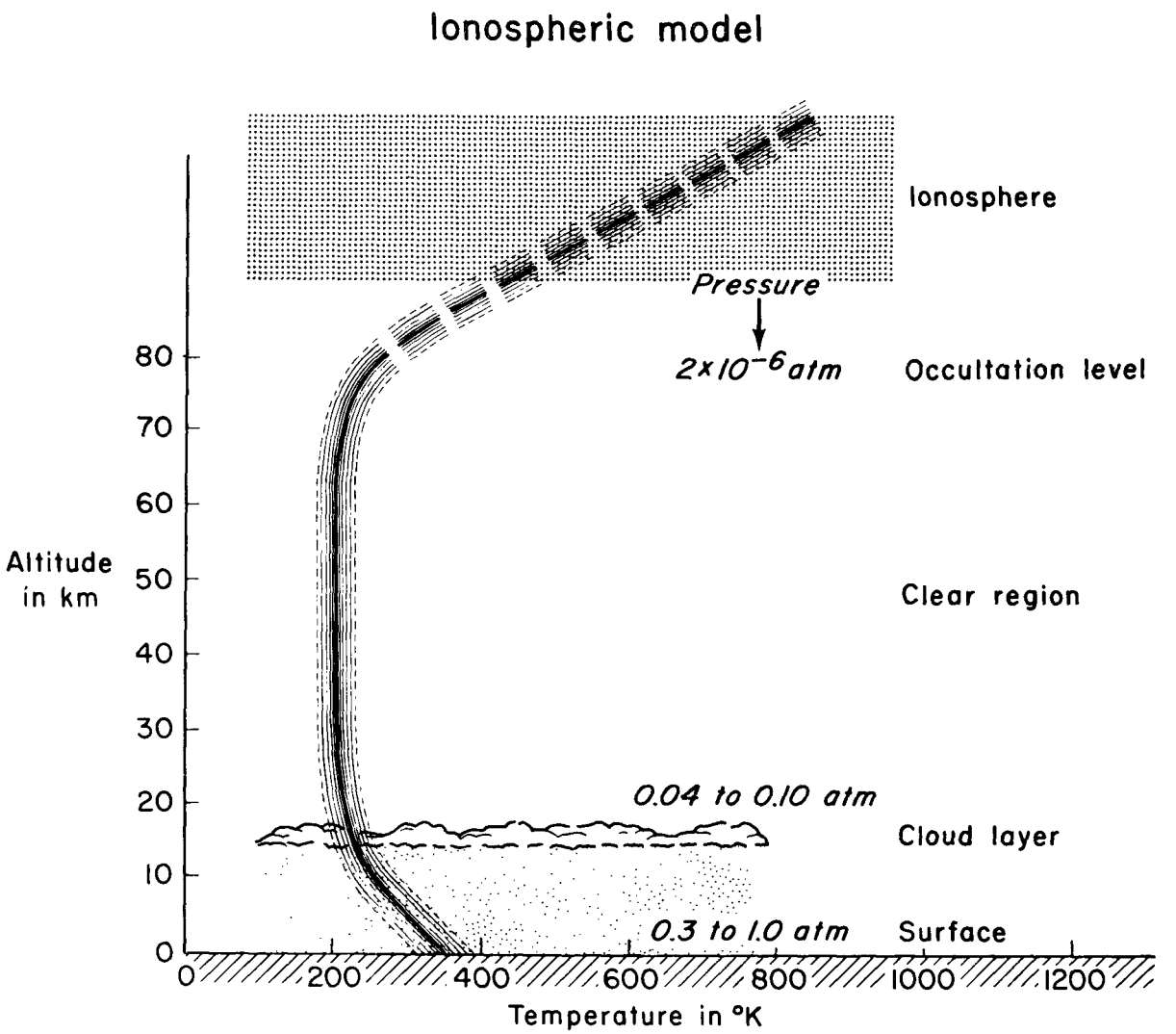


Figure 5

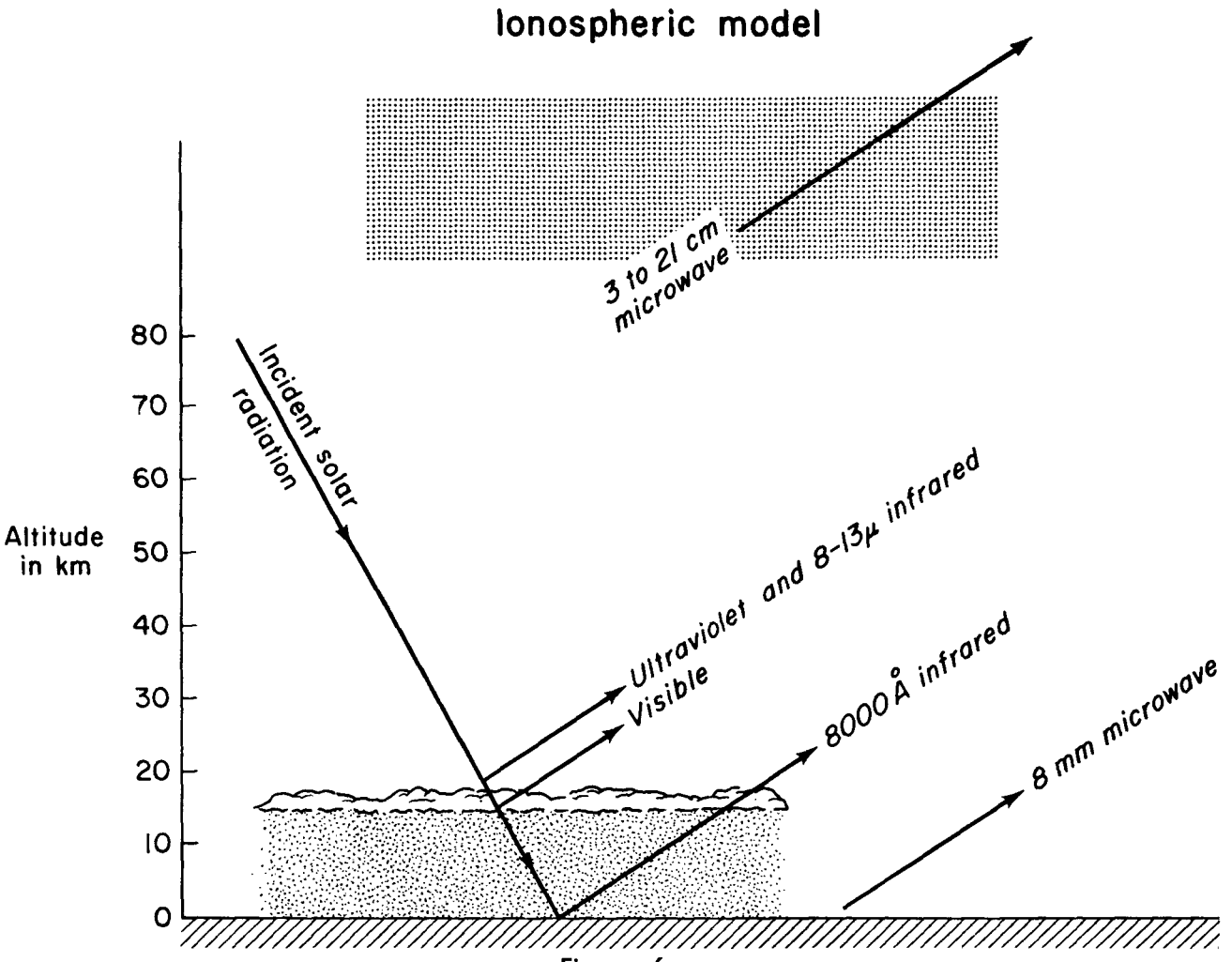


Figure 6

carbonates, concluded that Venus should have a much deeper atmosphere than the Earth. Assuming Venus to have an initial composition and density similar to the Earth's, Dole concludes that the surface pressure should be 8 to 10 times that on Earth. This conclusion would tend to argue in favor of one of the first two models, even though it does not, of itself, explain the high surface temperature.

One final speculative suggestion on the nature of the Venus surface conditions may be mentioned. Hoyle (1955) has suggested that Venus had a great overabundance of abiologically-synthesized hydrocarbons produced in the early history of the solar system. If this were so, then Venus in its early history may have been covered by an ocean with the less volatile hydrocarbons floating on the surface. The action of solar ultraviolet radiation through geological time would have dissociated both the volatile hydrocarbons and most of the amospheric water vapor, leaving CO_2 after the escape of the released hydrogen. The result would be an ocean, with water trapped beneath an unspecified depth of oil. The visible clouds might then be a kind of smog. Mintz (1961) has supported this hypothesis on the grounds that the small day-night temperature change at thermocouple wavelengths indicates a reservoir of high specific heat. However, the atmosphere itself might provide this reservoir (Öpik, 1956), and spectroscopic evidence rules out the possibility that the clouds are composed of simple hydrocarbons (Sagan, 1961).

D. Experiments to Distinguish Among the Alternative Models

There are several experiments that are now becoming feasible and that can decide among the various models of Venus presented above. These experiments illustrate how space vehicles can be used to solve basic astronomical dilemmas.

a. *A scan of the disk of Venus in the 1-cm-wavelength region to determine limb brightening.* On the ionospheric model, the ionosphere must just be approaching optical depth unity near 1 cm. Therefore observations made at the limb of the planet would penetrate through greater paths of electrons, and would give greater brightness temperatures than at the center of the disk. Resolution of the disk of Venus cannot be performed by any existing single radio telescope; and while radio interferometry can perform this feat in principle, in practice the difficulties are extreme. But observations from a fly-by probe at close approach to the planet can be performed conveniently. Observations of limb brightening at 1 cm would provide strong substantiation of the ionospheric model; observations of limb darkening would confirm that the centimeter microwave emission arises from the surface of Venus, and that the surface temperature is 600°K or hotter.

b. *Observations of a strong phase effect in the centimeter microwave emission.* It has already been mentioned that the aeolosphere model is probably incompatible with a centimeter wavelength emission which gives much greater brightness temperatures on the day side than on the night, and for which the temperature phase curve changes slope at inferior conjunction. On the other hand, such emission characteristics are consistent with the greenhouse model. Observations of Venus at microwave frequencies near superior conjunction are impossible with existing radio dishes. But Earth-based antennas now under construction, or antennas carried by a fly-by, would be able to obtain the desired information. If the ionospheric model is untenable, these observations would distinguish between the greenhouse and aeolosphere models.

c. *Observations of twilight airglow in the visible and ultraviolet.* If the ionospheric model is correct, the high density of charged particles on the daylight hemisphere must decay rapidly at twilight and early evening in order to provide the ionospheric hole required on the night hemisphere to explain the radar reflectivities. This recombination—both radiative and dissociative—must lead to intense emission lines. Observations made from Earth have to date been ambiguous, although they have suggested an airglow brighter than the terrestrial. Observations of airglow made from a fly-by probe could settle the question.

d. *Observations of decline in radar reflectivities as a function of phase.* On the ionospheric model, radar pulses incident on Venus should be returned more and more weakly as the ionospheric hole passes farther from view. There should be no radar return from the illuminated hemisphere. Observations of Venus when it is farther from the Earth than at dichotomy are very difficult for present radar antennas. Very large radar dishes on Earth, or an active radar on fly-by probes, should be able to test the existence of the hole directly.

e. *Observations of high-intensity infrared emission through atmospheric and cloud windows.* On the greenhouse model, some radiation from great atmospheric depths must be escaping to space in order to maintain the radiation balance. It is expected that there should be relatively intense emission between the atmospheric water vapor and carbon dioxide bands, especially in the near infrared, and through breaks in the clouds. To obtain high topographical resolution, this experiment would be best performed by an infrared spectrophotometer or filter radiometer in a Venus fly-by.

f. *Detailed observations and interpretations of the polarization-phase data of Venus.* The greenhouse model predicts an ice-crystal cloud layer; the aeolosphere model predicts a carbonate-dust cloud layer. Observations of the polarization curve from Venus at a variety of wavelengths and at high topographical resolution should provide enough information for a distinction between these two models by Mie theory on modern electronic computers. Some preliminary extension of the original polarization curves to a variety of wavelengths has recently been performed by Gehrels (1960), and his observations deserve serious theoretical study. Observations of high topographical resolution are most easily performed by a planetary fly-by.

Of the above experiments, a), b), and e) are tentatively planned for the Mariner R Venus fly-by probe currently scheduled for 1962. The remainder are within the capability of current scientific endeavor, and we may confidently expect that the next decade holds the promise of a much more detailed understanding of the physical environment of our nearest planetary neighbor.

References

Bierman, L., Solar corpuscular radiation and the interplanetary gas, *Observatory, 77*: 109, 1957.

Chamberlain, J. W., and G. P. Kuiper, Rotational temperature and phase variation of carbon dioxide bands of Venus, *Astrophys. J., 124*: 399, 1956.

de Vaucouleurs, G., private communication, 1961.

Dole, S. H., The atmosphere of Venus, Rand Corp. Paper P-978, 1956.

Dollfus, A., Étude visuelle et photographique de l'atmosphère de Venus, *L'Astronomie, 69*: 413, 1955.

Drake, F., private communication, 1961.

Gehrels, T., Measurements of the wavelength dependence of polarization, *Lowell Observ. Bull., 4*: 300, 1960.

Herzberg, G., Laboratory absorption spectra obtained with long paths, in *The Atmospheres of the Earth and Planets,* edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.

Hoyle, F., *Frontiers of Astronomy,* Harper and Bros., New York, 1955.

Jones, D. E., The microwave temperature of Venus, *Planetary and Space Sci., 5*: 166, 1961.

Kaplan, L. D., A new interpretation of the structure and CO_2 content of the Venus atmosphere, *Planetary and Space Sci., 8*: 23, 1961a.

Kaplan, L. D., private communication, 1961*b*.

Kotelnikov, V. A., and I. S. Shklovsky, Radar observations of Venus, *Izvestiya,* Moscow, 12 May 1961*a*.

Kotelnikov, V. A., private communication through Dr. A. G. Massevich, 1961*b*.

Kuiper, G. P., The atmosphere and the cloud layer of Venus, in *The Threshold of Space,* edited by M. Zelikoff, Pergamon Press, London, 1957.

Kuz'min, A. D., and A. E. Salomanovich, Radio emissions from Venus in the 8 mm bandwidth, *Astron. J., USSR, 37*: 279, 1960.

Kuz'min, A. D., and A. E. Salomanovich, Radioemission of Venus at 9.6 cm, (in Russian) *Astronomicheskii Tsirkulyar, 221*: 3, 1961.

Lilley, A. E., The temperature of Venus, paper presented at the 108th meeting of the American Astronomical Society, June, 1961, in Nantucket, Massachusetts, and private communication, 1961.

Lyot, B., Recherches sur la polarisation de la lumière des planètes et de quelques substances terrestres, Thesis, Paris University; *Ann. Obs. Meudon, 8*: 66, 1929.

Mayer, C. H., private communication, 1961.

Mayer, C. H., T. P. McCullough, and R. M. Sloanaker, Observations of Venus at 3.15-cm wavelength, *Astrophys. J., 127*: 1, 1958.

Menzel, D. H., and G. de Vaucouleurs, Results from the occultation of Regulus by Venus, *Astron. J., 65*: 351, 1960.

Menzel, D. H., and F. L. Whipple, The case for H_2O clouds on Venus, *Publs. Astron. Soc. Pacific 67*: 161, 1955.

Mintz, Y., Temperature and circulation of the Venus atmosphere, *Planetary and Space Sci., 5*: 141, 1961*a*.

Mintz, Y., private communication, 1961*b*.

Nawrocki, P. J., *Reaction Rates,* Geophysics Corp. of Amer. Tech. Rept. 61-2-A, Air Force Cambridge Res. Lab. No. 105, 1961.

Öpik, E. J., The surface conditions on Venus, *Irish Astron. J., 4*: 37, 1956; *Armagh Observ. Leaflet 43*, 1956.

Öpik, E. J., The aeolosphere and atmosphere of Venus, *J. Geophys. Research, 66*: 2807, 1961.

Rechtin, E., private communication, 1961.

Ross, F. E., Photographs of Venus, *Astrophys. J., 68*: 57, 1928.

Russell, H. N., The atmosphere of Venus, *Astrophys. J., 9*: 284, 1899.

Sagan, C., The radiation balance of Venus, *Calif. Inst. of Techn. Jet Propulsion Lab.* Tech. Rept. No. 32-34, 1960.

Sagan, C., The planet Venus, *Science, 133*: 849, 1961.

Sagan, C., K. M. Siegel, and D. E. Jones, On the origin of the Venus microwave emission, *Astron. J., 66*: 52, 1961.

Sinton, W. M., Observations of Venus, in *Compendium on Planetary Atmosphere,* edited by Z. Kopal and Z. Sekera, Academic Press, New York, in press, 1961, and private communication, 1961.

Sinton, W. M., and J. Strong, Radiometric observations of Venus, *Astrophys. J., 131*: 470, 1960.

Strong, J., private communication, 1960.

Strong, J., M. D. Ross, and C. B. Moore, Some observations of the atmosphere of Venus and the Earth during the Strato Lab IV balloon flight, (abstract), *J. Geophys. Research, 65*: 2526, 1960.

Urey, H. C., *The Planets: Their Origin and Development,* Yale University Press, New Haven, 1952.

Urey, H. C., and A. W. Brewer, Fluorescence in planetary atmospheres, *Proc. Roy. Soc. (London). A241*: 37, 1957.

Chapter V

FUTURE PLANETARY ATMOSPHERES RESEARCH

A. Introduction

The preceding chapters have dealt with some of the techniques used to observe and study the planets, some results of these studies, and some working hypotheses (or models) of the atmospheres of Mars and Venus. A great deal has been learned about these planets, and we now have some insight into the nature of their atmospheres, but it must be clear that there is still a host of unanswered questions.

Mars is the best understood, and we can describe its lower atmosphere roughly—very roughly. Surface pressure is probably certain to within a factor of two, we can estimate its composition, its surface temperature can be measured, and we can observe and partly explain some of the changes that move across the face of the planet. Yet, we are not sure of the nature of the various yellow and white clouds, the blue haze that comes and goes is still a puzzle, and the vertical temperature structure of the Martian atmosphere is still a good subject for conjecture.

Venus is largely a mystery, and there has been no entirely satisfactory resolution of all the various observations that now exist. We do know something about the atmosphere *above* the clouds, for the temperatures in that region can be estimated, but there are at least three completely different models of the conditions *below* the clouds. It may be a diffusely illuminated desert as hot as molten lead; a dark wind-swept landscape equally hot; or a much more equitable surface, possibly completely covered by an ocean—we cannot certainly rule out any of these possibilities. Furthermore, we are not even sure about the length of the Venus day, whether about ten or more than a hundred Earthdays.

It should be mentioned that the large outer planets (Jupiter, Saturn, Uranus, and Neptune), which in some respects resemble cold stars more closely than the solid terrestrial planets, are harder to observe and even more mysterious. In the case of Jupiter, for example, we know a certain amount about the composition and temperature of its upper atmosphere, the part above the cloud layers, and we can watch its gross circulation. However, we can only speculate about its internal structure, its belted circulation pattern remains to be explained by dynamic meteorologists, and, as if to compound the puzzle further, the famous Red Spot sits in the middle of one of the cloud bands as an apparently permanent feature in an inconstant atmosphere.

This, then, is a brief qualitative review of our knowledge of the planets at the beginning of the first decade of the Space Age. It is clear that Mars and Venus will be the first targets of our space probes. (Venus has already been approached by a Soviet probe, and a U. S. probe has gone roughly half the way to the Venus orbit, but there was no radio communication with either of these probes during their times of closest approach.)

In the course of the Panel's discussions it became clear that a rather large number of crucial experiments remains to be done from the Earth's surface, from balloons or aircraft, and, of course, from space probes. There is also a host of laboratory and theoretical studies to be made. The subsequent sections attempt to describe how the planetary enigmas may be unraveled in the next decade, given enough financial support, wise leadership, and the dedication of capable scientists.

B. Some Guiding Principles

The following general principles for a planetary research program are almost self-evident, but should be stated nevertheless:

(a) It is always cheaper to make an observation from the Earth (or from a balloon or aircraft) than from a space probe. The scientific return from further terrestrial observations for a given investment of effort could be very great if only we can see how to make them properly and encourage competent scientists to make them.

(b) Space probes will fly to the planets with small payloads in the first half of this decade, and with larger payloads later. It requires two or three years to conceive, design, and build a space-probe package. Any pertinent advance in our knowledge in the next few years will greatly assist the planning for these various probes.

(c) Even after space probes have successfully visited the planets there will still be some value to continued Earth-based surveillance of the planets, since they are continually changing their aspects, and these changes must be as much a part of our store of information as the detailed snapshots that we will get from the probes. The planets are obviously far from static, and their "weather" changes as it does on Earth.

(d) Having pointed to the great value of terrestrial observations, we must also mention that some Earth-bound experiments would be inadvisable in view of the advent of probes and other powerful observing tools. Some very elaborate experiments have been suggested which would be superseded even before they could be accomplished.

(e) The same considerations apply to the use of Earth satellites for observing the planets, since by the time the more advanced orbiting astronomical observatories are in operation we should be beginning to get observations from planetary probes. Moveover, it is clear that these new observatories are particularly well suited for solar, stellar, and galactic astronomy, and should be devoted to these purposes.

(f) Much of our knowledge of the planets has been and will continue to be based on lessons learned from studying our own planet. With this in mind, it is clear that no opportunity should be lost to test out planetary-probe experiments from rockets and Earth satellites. In addition to serving as "field tests" for new equipment and techniques, these tests can be valuable scientific experiments in their own right, and in all likelihood will give vital information about our own planet.

(g) Finally, while we fully expect to see manned flights to the nearer planets in our lifetime, we will not mention any manned flights in our program for this decade. The emphasis will be on the first costly steps into space, steps that must of necessity be carefully placed, and by remote control.

C. Terrestrial Observatories

1. The Planet Patrol

Astronomers have by and large been more concerned with the distant parts of the universe and have neglected the planets. In view of the recent burgeoning interest in planetary studies, it is important to note that there are many fine observatories on every continent of the world which are capable of performing significant planetary observations. If a well-conceived international planetary observing program were organized it would undoubtedly receive the support of many of these observatories.

Such a program is being formulated, and reference is made to the report by a JPL-sponsored study group (1961). Some of the primary features of the proposed program are the organization of a worldwide cooperative watch on Mars and Venus, the establishment of communication procedures and uniform observing practices for the observatories involved, and the support of a central organization (probably in the U.S.) to receive the observations (photographic plates and visual

reports), to analyze them, and to make them generally available in a useable form. For the first time a nearly continuous sequence of planetary observations could be obtained, and the daily changes could be determined without a break for long periods. This would clearly be a great step forward, if the program were carefully planned and the observations so obtained were properly collected and processed.

Some of the other recommendations of the JPL-sponsored committee that are pertinent here are the following: (These were reported to the Panel and generally endorsed.)

(a) The international network of observatories would be based on existing observatories, where staff and facilities permitted. Any additional staff needed at these observatories to assist in carrying out the lunar and planetary program would be drawn from nationals of the home country, but interchange of research workers between such observatories would be encouraged.

(b) The central organization, referred to as a Data Reduction Center, in addition to serving as a clearing house and communications center, would also act as a study center where the varied techniques of the related disciplines of *geophysics* and *astronomy* could be brought to bear simultaneously.

(c) If a new site with very good seeing conditions through an appreciable part of the year could be found in the U.S., then consideration should be given to the establishment of a new observatory devoted to lunar and planetary observations at such a site. The exceptional seeing conditions enjoyed at some observatories outside the continental U.S., notably at the Pic du Midi in France, are not found at any of the established U.S. observatories.

The cost of such a program would probably be of the order of a million dollars per year. New observing equipment would be involved if the U.S. wished to provide some for observatories in other countries to help them to join the program, or if an exceptionally good site were found in the U.S. for a new observatory. In this case, about one million would be required for the initial outlay for both overseas and U.S. equipment.

2. *Spectral Observations*

All of our information about the *composition* of planetary atmospheres, such as it is, comes from observing the spectrum of the sunlight reflected from the planets. The near infrared, below about 1 micron, is by far the most commonly used, but the near ultraviolet and infrared beyond 1 micron are also observable. With the continued improvement of astronomical techniques, especially of the detectors used, these spectral observations can and should be refined.

A great deal of our information about planetary *temperatures* comes from observations of the far-infrared emission. Here again, better techniques and detectors may permit improved observations of the total flux in the atmospheric window (8-13 microns) and of the far-infrared spectrum.

Two problems exist here, however:

(a) The Earth's atmosphere contains some of the same molecules as the atmospheres of the other planets (although in different proportions), so absorption and emission spectra must be carefully corrected for telluric absorption and emission. Frequently, it has been necessary to utilize the Doppler shift arising from the relative motion of the source and the Earth to obtain meaningful results. However, some very important gases, notably water vapor and oxygen, are so abundant in our atmosphere that even this technique is not enough to eliminate the effect of the terrestrial atmosphere. We must get above it.

(b) The second problem is one of manpower and interest. The spectral observations of the planets that exist were made by a small handful of highly skilled astronomers who have individually devoted years to the problems involved. Therefore, we cannot expect new observations to be obtained easily, and support and encouragement should be given to this effort over

a period of years to ensure that those who can do it have the best equipment, and to encourage other astronomers to undertake this kind of work.

The total cost of this astronomical observation program would be insignificant relative to other programs in planetary research. The problem is more likely to be that of training more qualified scientists and assuring them continuity of support.

D. Balloon and Aircraft Astronomy

The reasons for wanting to take a telescope high into the atmosphere are twofold, and both have to do with eliminating effects of the atmosphere (v. Chapter I.B.):

(a) With a telescope on a balloon one can take advantage of the much better seeing conditions at high altitude, and begin to achieve the theoretical resolving power of the telescope with pictures free of distortions and blurring (scintillations). If a perfect telescope on a high-altitude balloon has an aperture of 6 inches, it will achieve a photographic resolution as good or better than any existing surface telescope, regardless of size. (A striking demonstration of this increased resolution was performed in the Stratoscope I solar photography program organized by M. Schwarzschild; v. Danielson, 1961.) A 36-inch telescope on a balloon would theoretically give photographs five to ten times better than any obtained from the surface. Thus, the present 100- to 500-km photographic resolution now achievable on Mars at closest approach (v. Chapter I, Table II) would be reduced to the order of 50-km resolution, and for the first time we might photographically distinguish mountains from valleys, see moving cloud systems clearly, follow the detailed color changes in the Martian spring, etc.

(b) The other advantage of going high has already been explained: Spectral measurements in the near and far infrared can be made with greatly reduced absorption by the Earth's gases. The only measurement ever made of water vapor in the Venus atmosphere was from a balloon (Strong, Moore, and Ross, 1960), and there is a need for much more of this kind of observation.

The JPL-sponsored committee (1961) already referred to also urged a balloon-astronomy program. An important point made in their report concerned the need for coordinating the time of balloon flights to coincide with either important planetary events (storms, blue clearings, etc.) or with the close approach of space probes to the planet being observed.

Costs of balloon programs, while small relative to space projects, are not insignificant. One representative program, for example, is estimated to cost 2.5 million dollars over a five-year period, covering the cost of three or more flight packages and several flights per year. We are aware of at least four groups that would be competent to undertake such programs, so about 20 million dollars might be a good estimate for a ten-year program of balloon astronomy.

It should be mentioned that good balloon-launching facilities already exist in the U.S., the Navy has frequently demonstrated the feasibility of shipboard launching, and there are plans to establish at least one more major balloon facility. It therefore seems that the emphasis should be on developing the flight packages and building up competent groups to run the scientific side of the program with an assurance of continuity of support.

The subject of this report is planetary atmospheres, but it is clear that such programs would inevitably result in invaluable solar, stellar, and astrophysical observations as well. It should therefore be considered as "balloon-astronomy" program in the broadest sense.

E. Radio and Radar Astronomy

One of the newest tools of the astronomer, the radio telescope, has already been of great value in observing the Moon and planets. By simply monitoring the microwave emissions from the planets one may deduce the temperature, since every warm body radiates some energy at radio frequencies. By

sending a pulse of radio waves and analyzing the echo one can possibly deduce several other parameters about a planet, such as its radio-reflectivity, its rate of rotation, its surface roughness, and even the locations of surface irregularities. Its ionospheric characteristics may also be determined. (v. Chapter I, and Appendix 4 by Eshleman.) The state of the art of radar astronomy is such that the last three items are still either marginal or unachievable for the planets. But instruments are already being designed and built which may be able to do these things quite well, e.g., the 330-m dish in Puerto Rico, the Navy's steerable 200-m dish in West Virginia, and the new arrays being constructed by Stanford University.

As explained in Chapter IV, the case of Venus is especially interesting from the radio and radar astronomy viewpoint (as is Jupiter, for somewhat different reasons). Since it is covered with clouds, we cannot "see" to the surface except with radio waves. Therefore, the only *surface* temperature observations come from the microwave radiometric measurements. Due to the almost featureless cloud cover it is difficult to determine the rate of rotation of Venus, as we can visually with Mars. The best present method is radar; but at the present writing the only two announced results are in disagreement by more than an order of magnitude (cf. Chapter IV).

There can be no clearer indication of the need for more observations. Specifically, the following programs could be implemented (v. also Chapter IV.C.4):

(a) Continued radiometric observations of Venus at a number of wavelengths, some corresponding to CO_2 and H_2O absorption lines, to observe the change in apparent temperature with changing phase.

(b) Continued radio and radar observations with current instruments as Venus moves away from inferior conjunction to observe the change in brightness temperature, reflectivity, and apparent rotation rate with changing phase.

(c) Improvement in transmitted power, antenna gain, receiver sensitivity and interferometric techniques in order to achieve some of the more detailed planetary observations that are currently only possible in principle. Such techniques are applicable to virtually all the planets in the solar system, and could give results far in advance of any space probes to them.

The cost of these programs is considerable, but fortunately many of the radio telescopes required are already in existence or are under construction with allocated funds. Capital investment is therefore not so much a problem as assuring that these costly instruments will be available part of the time for planetary work. (Some are, for example, tied up with tracking and telemetering space probes and satellites, others are being scheduled almost exclusively for extraplanetary studies.) Many hundreds of millions of dollars are involved in the nation's total radio and radar astronomy programs, so presumably several tens of millions might be shared by the planetary program.

F. Experimental and Theoretical Studies

Meanwhile, there are a host of problems that can be solved in the laboratory and whose solution would contribute to our knowledge of planetary atmospheres, including the Earth's. In fact, this section can be made rather brief by pointing out that an active atmospheric sciences research program would inevitably provide information applicable to Mars and Venus. The converse is certainly true also, for studies of other planets are bound to give new insight into our own.

As one scientist (Harry Wexler) has aptly put it: We try to make liquids in dishpans behave like atmospheres, while the other planets are already there to serve as Nature's own dishpans.

Certainly we must learn to think of the Earth's atmosphere as but one of a family of planetary atmospheres, and a generalized theory for the dynamics of atmospheres will be one ultimate goal of planetary research. (v. Chapter II and Mintz, Appendix 8.) This will require further laboratory studies (in dishpans) of planetary circulations, generalized numerical models of atmospheres, better theories for radiation absorption, emission, and transfer, better values for the many molecular reaction rates and cross sections, etc.

Eventually we will also understand the evolution of planetary atmospheres, and we may gain new insight into the reasons for climatic change—whether these periodic changes in climate are due to changes in the Sun or within our own atmosphere.

This aspect of planetary atmosphere research, really an integral part of atmospheric research in general, is essential to our planetary studies, and deserves much attention and support.

G. Space Probes—Experimental Astronomy

As pointed out earlier in Chapter I, and more specifically in Chapters III (Mars) and IV (Venus), it is clear that ultimately nearly all of our detailed knowledge of the planets will come from instruments close to them or penetrating to their surfaces. At first, due to limitations on payload weight and complexity, most of the observations will be from *fly-by* vehicles, taking advantage of the much greater flux of radiation from the planet and the gain in resolution of features on the planetary disk. Early in the game it will also be possible to put small packages into the atmosphere, and a man-made instrument will send the first brief message from the surface of another planet.

So much attention has already been paid to the possible experiments with space probes, not only in this country but internationally, that we will not endeavor to go into detail here. The outline below reflects as much as possible the results of many planning efforts, especially of those charged with the responsibility of carrying out the U.S. space program (NASA and its Jet Propulsion Laboratory). Implicit are the engineering restrictions imposed on the earlier payloads; implicit also is the consideration that we would not recommend an experiment on a space probe that could be done almost as well on the Earth or from a balloon. More details of the Mariner program may be found in Appendix 6 by Davies *et al.*

1. *Early Fly-By Observations*

a. *Mars*

Infrared spectra of both reflected and emitted radiation, with resolution to a fraction of a micron and with enough spatial resolution to distinguish among the major surface features.

Ultraviolet spectra down to about 1000 Å with better than 10 Å resolution, and with enough spatial resolution to distinguish limb* effects and study the terminator.*

Photoelectric photometry and polarimetry at a number of wavelengths from the near ultraviolet through the near infrared, and providing enough spatial resolution to study limb and terminator effects. Note that phase angles greater than 43° cannot be observed from the Earth, so fly-by observations are essential to obtain the complete characteristic polarization curve for any part of Mars. (v. Chapter I.A.1)

Magnetic fields.

Energetic particles, which would be associated with any radiation belt and would be coupled with the magnetic field.

Micrometeorites.

b. *Venus*

Same experiments as for Mars, plus: Radiometer to measure the microwave emission in the wavelength range from a few millimeters to several centimeters, with enough spatial resolution to observe limb and terminator effects; and active radar. (v. detailed discussion of Venus fly-by experiments and their significance in Chapter IV.C.4.)

2. *Observations with Early Atmospheric Entry Probes During and After Descent*

Pressure.

Temperature.

Density, which may be derived from other measurements during descent.

---

* The "limb" of a planet is the edge of its disk; the "terminator" is the transition zone between the day and night sides.

Radar altitude or, alternatively, a continuous record of acceleration (during descent only, of course).

Composition, e.g., quantitative tests for water vapor, carbon dioxide, oxygen, and ozone. This might be done by direct sampling, or by observing the absorption of sunlight at specific wavelengths as the capsule descends.

Measurements of sky brightness and polarization at several wavelengths in the visible and near infrared, to give total depth of atmosphere and the distributions of clouds and haze.

Wind direction and speed, after landing only.

Succession of low-resolution TV pictures taken downward during descent (to identify the landing place, to indicate wind drift, and to give a hint of the detailed surface conditions in the landing area—may be impractical below the cloud deck of Venus).

Biological experiments.

3. *Observations with Advanced Entry Probes*

Same experiments as above, plus: Continued monitoring of "weather" from the surface, with daily transmission directly back to Earth.

4. *Observations with Advanced Orbiting Vehicles*

Considerable advances will have been made in meteorological satellites and orbiting geophysical observatories in the next two or three years. This experience will serve as the basis for the design of orbiting observatories around the planets. Many of the fly-by observations described above will also be applicable. The purpose of such planetary orbiters would be to monitor the gross changes in the planet's atmosphere, to observe storm systems, map the general circulation, observe seasonal and secular changes in surface features, etc.

H. Epilogue

There have been organisms on our planet for about four thousand million years. By a remarkable stroke of fortune, it is in the next few decades that Man will first discover—rigorously, and in detail—what is happening on the neighboring worlds. The probability of our being alive at the present time, taken on a random basis, is therefore about a millionth of a percent. We are immensely lucky to be living at the dawn of this era of planetary exploration and high scientific adventure.

References

A Recommendation for a Ground-Based and Balloon-Borne Lunar-Planetary Observation Program, *Jet Propulsion Lab* Tech. Memo. No. 33-37, 1961.

R. Danielson, *American Scientist*, in press, 1961.

Strong, J., M. D. Ross, and C. B. Moore, Some observations of the atmosphere of Venus and the Earth during the Strato Lab IV balloon flight, (abstract), *J. Geophys. Research, 65*: 2526, 1960.

Appendix 1

DIRECT PHOTOGRAPHY IN THE EXPLORATION OF PLANETARY ATMOSPHERES

A. G. Wilson

A. Information Spaces

In the systematic study of planetary atmospheres one primary role of direct photography is in the detection and identification of the indigenous meteorological phenomena. In planetary astronomy the detection of an object is accomplished through inter-comparison of areal "information cells" with respect to their brightness, color, and variation in time. The detectable objects are those which emerge above the contrast and resolving power thresholds, and are contained in the light response, spectral, temporal, and angular ranges of the instrumental system. Identification as meteorological phenomena must be through comparison of the detected objects with familiar meteorological objects or events in the Earth's atmosphere.

There exists what may be termed a "similarity threshold," on one side of which phenomena detected on other planets may be identified with familiar terrestrial phenomena or recognized as extrapolations of terrestrial phenomena, but on the other side of which their identification, and even their reality, becomes speculative. As more detailed knowledge of other planets is collected, the base of the familiar against which comparisons are made will be broadened. It is epistemologically fortunate that the first planets to be explored, Mars and Venus, are quite similar to the Earth, allowing ready identification of many phenomena. The exploration of these planets should extend the base of the familiar and provide experience which will create a more advantageous similarity threshold.

Since direct photography must play a basic role in any exploration program whose end is the discovery, observation, and analysis of the meteorological processes on other planets, it will be useful to have a measure of the relative capabilities of various photographic systems for planetary atmosphere studies.

For this purpose the process of direct photography may be considered to be a function of five basic parameters or "dimensions." These are the two linear or angular dimensions of the region photographed, the brightness dimension (which is recorded as photographic density), the spectral dimension, and the temporal dimension. In each of these dimensions there are bounds which define the *range* of a system and a threshold which sets the *resolving power,* as described in Table I.

Although there is not complete symmetry in considering the parameters in this way, the viewpoint is useful in that it allows the construction of a five-dimensional "Instrument Information Space" —the extension of the space is determined by the ranges, the size of the information cells is determined by the resolving powers—which provide a set of figures of merit for evaluating the capabilities of photographic (and other instrument) systems.

The instrument information space for direct photographic observations of planets is defined by the following four factors:

(1) the optical and photographic parameters such as telescopic aperture, emulsion sensitivity, grain, contrast, etc.;
(2) relative motions of the instrument and field being photographed;
(3) the location of the instrument with respect to the planet being explored;
(4) considerations of technical and economic feasibility.

Table I

DIMENSIONS FOR PHOTOGRAPHIC SYSTEMS

Dimension	*Resolving Power*	*Range*	*Effective Range*
Areal (2 dimensions)	Angular resolving power set by telescopic and photographic parameters together with seeing and instrument stability limitations.	Determined by angular field of view.	Field of view, modified by aberrations of the optical system.
Brightness	The contrast, λ	Determined by signal-noise ratio and emulsion saturation.	Determined by a set of exposure times.
Spectral	The filter-emulsion-optical component band pass.	Determined by filter-emulsion combinations and optical and atmospheric transmission limits.	Sums of bands at which exposures are taken.
Temporal	Exposure time and/or frequency of exposure.	Span of observations.	Determined by number of exposures, divided by frequency of exposures.

Item (1) is fully discussed in many texts on photography and optics. (v., e.g., J. Strong, *Procedures in Experimental Physics,* Prentice-Hall, 1945, or G. de Vaucouleurs, *"Planetary Astronomy from Satellite Substitute Vehicles,"* Chap. II, AFMDC–TR–60–6.) Item (2) refers to the dynamic stability of the telescope-camera system, periods of natural oscillations, guidance, relative motions caused by planetary rotations, movement of the instrument carrier, etc. Item (3) takes into account the limiting effects of seeing, sky brightness, and spectral transmission properties of the Earth's atmosphere, also the effects of distance to the planet on the linear field of view and linear resolving power. Item (4) involves the sizes of instruments which may be practically carried in balloons, placed in orbit about the Earth, or carried in fly-by probes and planetary landing capsules. It also involves the cost of constraints of each system and the extent of program economically feasible with each system.

The information space required for the photographic study of a given phenomenon must provide sufficient data for determining the size, structure, position, and movement of the phenomenon (as, say, a storm), the life span, rates of growth, decay, and other changes, the season and frequency of occurrence, and such physical quantities as brightness and color. An information space adequate for this purpose may be called a *description information space* for the phenomenon. Since it is redundant to observe all parts and all features of a phenomenon to the same degree of detail, it is evident that the description information space for the study of the structure and behavior of a phenomenon will best consist of a *set of instrument system information spaces.*

Two basic problems thus arise: First, the defining of the description information space for the phenomenon, and, second, the selection of the set of system information spaces which must economically (with respect to time, energy, dollars) span the description information space.

To provide illustrations of these concepts and to provide an example of the data which determine a description information space, and also what instrument system information spaces would best cover the description space, it is useful to consider the present situation of knowledge concerning the

planet Mars—which is by far the best observed of the planets—gathered over the past 70 years by visual and photographic exploration from the Earth's surface.

In the following section, the phenomena that have been ascertained to exist on Mars are listed together with what is generally known quantitatively about these phenomena.

B. Photographically Observed Phenomena on Mars

One of the first constraints on our knowledge of Mars is the limiting resolving power to which the planet has been observed.

At the most favorable oppositions, Mars' closest approach to the Earth is about 5.6×10^7 km, or 150 times the distance to the Moon. The angular diameter at this time is only 25 seconds of arc compared with 31 minutes for the Moon. Since the average resolving power limit set by the atmospheric seeing is about 1 second of arc and the resolving power of the eye is of the order of 1 minute of arc, the amount of information on an average photograph of Mars is about the same as that which the naked eye receives from the Moon—roughly 10 bits ($\log_2 10^3$). However, there do exist a few photographs of Mars taken at instants of excellent seeing that contain perhaps 100 times this information. Because of the eye's ability to accommodate to seeing effects, it is able to do even better than the photographic plate and many observed Martian phenomena, such as the canals, lie in the region of information space beyond the 16.5-bit ($\log_2 10^5$) level of the best photographs and limited by the eye's capability. Under best conditions there is perhaps another augmentation factor of ten or so.

The present best linear resolution, corresponding to an angular resolving power of 0.1 second of arc, is about 30 km on Mars. This is for point phenomena like oases. For linear phenomena like canals, provided they are long enough, the resolution may be less than 5 km. Visual observers feel that an additional increase in resolution by a factor of 10 would give an information "break through" with regard to knowledge of Martian phenomena similar to the revelation of craters and mountains on the Moon that came with the first telescopes.

A second bound on our present knowledge of Mars is brought about by the temporal resolution with which the planet has been observed. The temporal range and resolving power of information space for Mars is more difficult to estimate. Photographic records go back to about 1890, but the observational coverage of Mars has been limited mostly to the few weeks before and after opposition, when the angular diameter is greater than about 15 seconds of arc. Unfavorable oppositions have for the most part been inadequately observed, partly because Mars is smaller and partly because these oppositions occur during the bad observing season for most observatories. The observations have had adequate temporal resolving power to determine the Martian seasonal changes, (for the southern hemisphere of Mars, spring is 146 Earth days, summer 160 days, autumn 199 days, and winter 182 days), but have been grossly inadequate for synoptic studies of the Martian atmosphere. An aggregation of all photographic observations might give an average temporal resolving power of one day for the six weeks preceding and following the most favorable oppositions since 1909.

During the oppositions of 1954 and 1956 the International Mars Committee organized a worldwide photographic patrol. The 1954 coverage, which was the all-time best, is shown graphically by Table II and Table III taken from the 1954 report of the Mars Committee. (Mitchell, R. I., *The 1954 International Mars Photographic Patrol,* Mars 1954, Report of the International Mars Committee, Lowell Observatory, Flagstaff, 1955.) On the left of the tables are given the Earth dates, on the right are given the Martian seasonal dates, it being southern hemisphere spring.

The photographic coverage of Mars with multicolored filters or color photography has been experimental rather than systematic. Except for the blue-yellow-red patrols of one or two observatories at recent oppositions no complete photographic comparative color record exists.

Table II

OBSERVATIONS OF MARS 1954

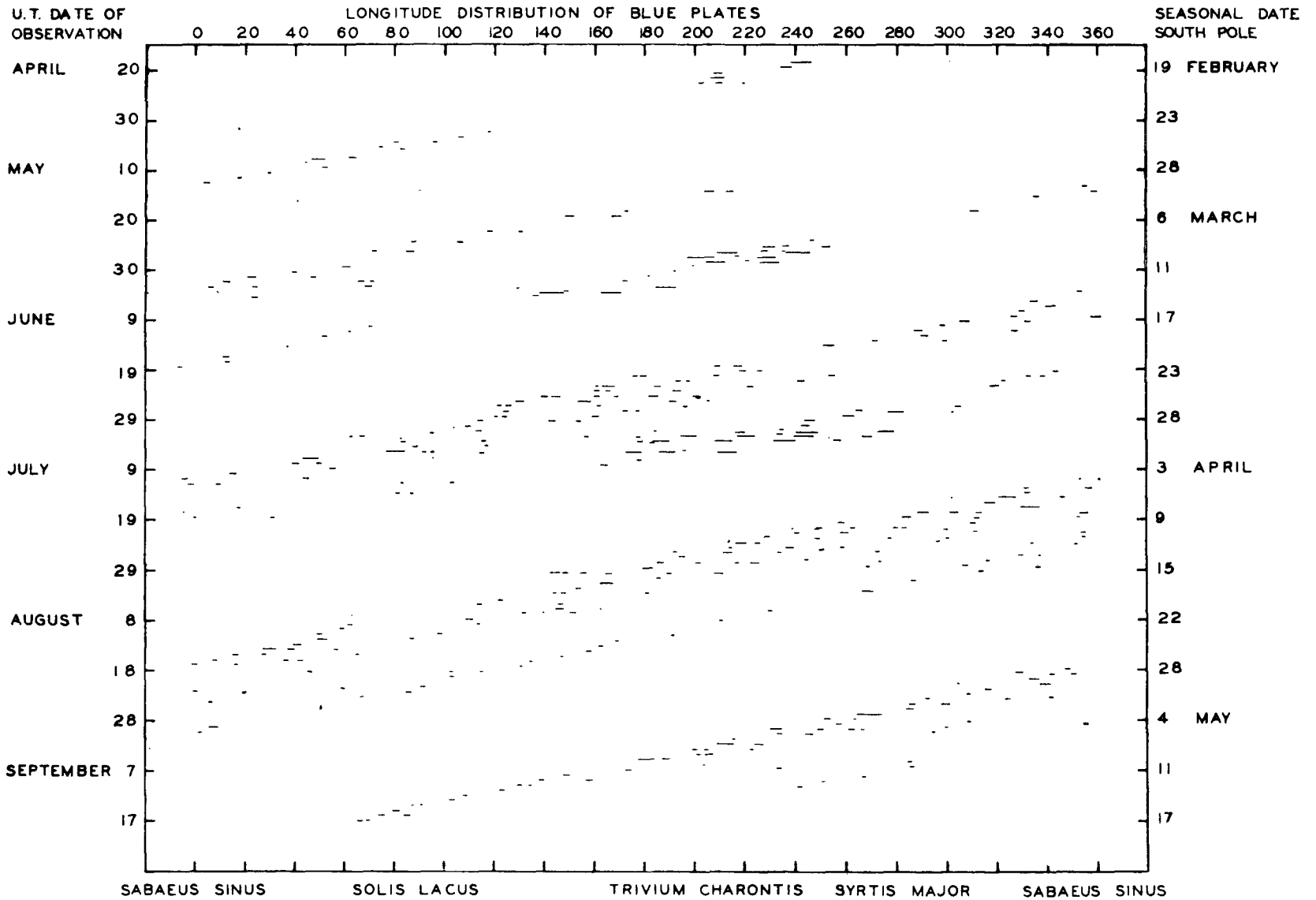
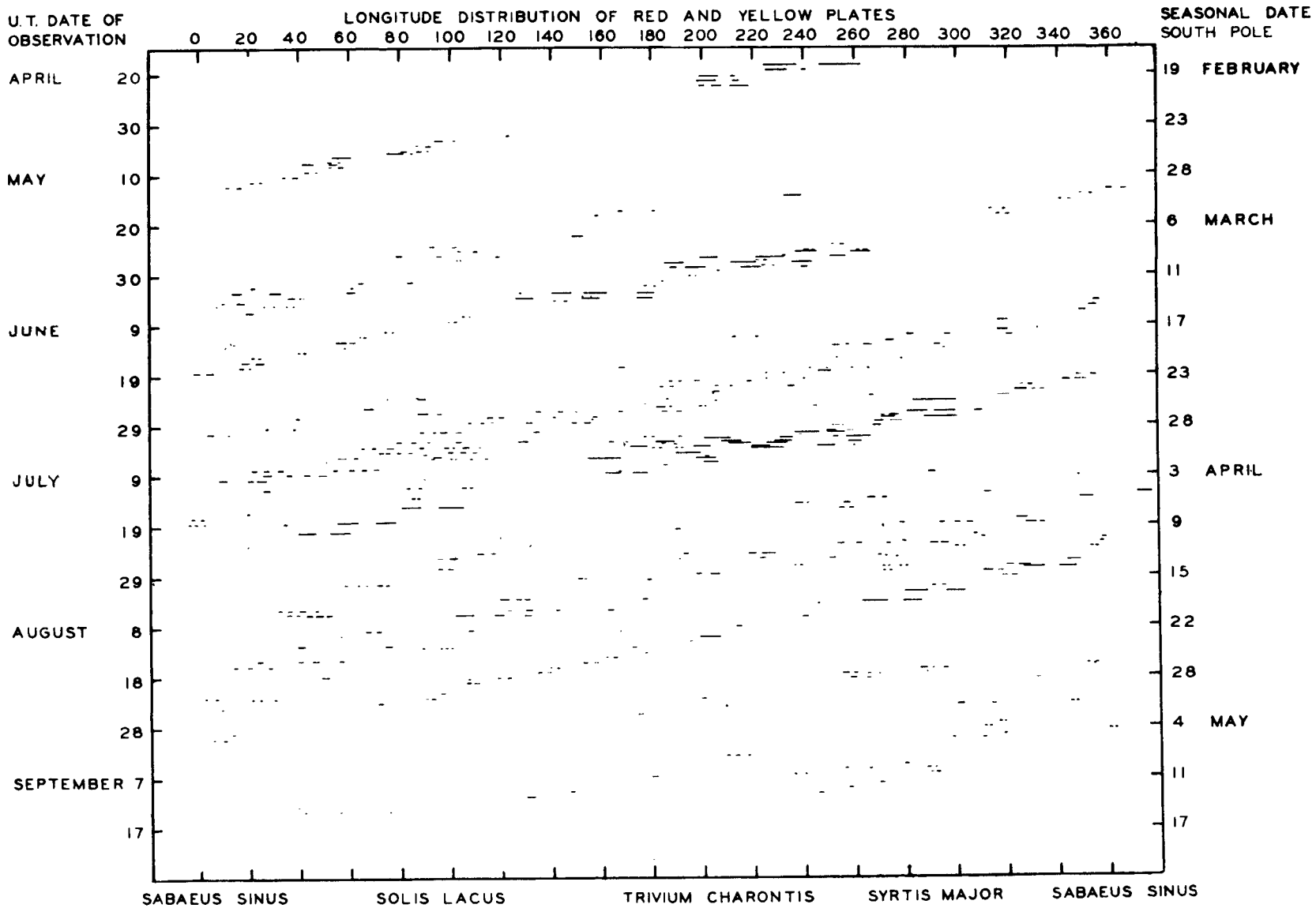


Table III

OBSERVATIONS OF MARS 1954



Faster and finer-grain emulsions have become available in recent years, and experiments with the contrast parameter have been made. Image tubes rather than direct photography, however, show the greatest promise for exploiting what can be done with contrast.

For *direct photographic* information purposes, Martian phenomena can be classified with respect to position, size, color occurrence, duration, and rate of change (with respect to formation, dissipation, size, color, position). The following list is a summary of generally available observational knowledge concerning Martian phenomena, omitting theoretical inferences and interpretations. (v., also, Chapter III on Mars.)

C. List of Photographically Observed Martian Phenomena

1. Surface

(a) *Polar Caps.* The caps appear to be formed in the seasonal autumn while they are largely covered with white clouds or fog. Toward the end of seasonal winter these clouds disappear, and a dark fringe appears on the edge of the caps which then begin to recede in size. The decrease continues through the seasonal spring, the dark fringe being widest when the melting rate is fastest. (Quantitative data on melting rates consisting of the size of four cap areas with corresponding dates, given by Pickering, 1924.) The South Cap is centered on long. 40°, lat. 83°; maximum size: to 45° lat.; minimum size: can disappear completely. It exhibits rifts in spring (Mountains of Mitchell), also occasional bright spots near edge.
The North Cap at maximum size extends to 57° lat.; minimum size: 300 km (1° = 57 km).

(b) *Dark Markings.* Termed maria. For the most part, located in the southern hemisphere, have been carefully mapped and named. Cover about ⅜ of the surface area, are mostly permanent, but additional dark areas appear from time to time lasting for a few years. In 1954 a new area "size of Texas" (± Pecos County) northeast of the Syrtis Major was observed which had developed since the last observations in 1952. Seasonal color changes occur moving from the polar caps toward the equator in seasonal spring. Rate of advance of the change is about 45 km/day. The color change is regarded by most visual observers as from gray (or blue gray) to brown or violet. Dark areas are faint in seasonal winter.

(c) *Bright Areas.* Called "deserts." Are of a general orange or ochre color. Cover about ¾ of the surface area of Mars; are static. Observations of limb indicate that no abrupt heights on Mars exceed 2500 ft. (Lowell).* There are areas that become temporarily whiter from time to time. Hellas for example is usually whiter than the other desert areas.

(d) *Canals.* Controversial network of fine linear markings, many permanent. Some observers have mapped over 400, a fifth of them being double, widths under 25 km. Show color changes similar to dark markings. Seasonal color change moves more slowly in canals (18

* Remark by *C. W. Tombaugh*: Wilson quotes Lowell's value of 2500 feet as the level of abrupt heights of terrain on Mars. I strongly disagree with Lowell's value. It should be remembered that at the time of greatest phase defect (when the best opportunity to see relief occurs) Mars is at twice its oppositional distance. Also, the terminator is, even then, far from the center of the disk, so that any horizontal distance of a cast shadow is foreshortened by a factor of about two. This means that the shortest perceptible horizontal distance in the vicinity of the terminator is 30 × 4 = 120 km on the basis of 0.1 arc-second of resolution. (v. Chapter I, Table 2.) I would say that 0.15 arc-second is a more realistic limit of resolution. Then the smallest perceptible terminator resolution would be 180 km, and the smallest perceptible height would be 4.7 km = 15,500 feet. (Assume a projection, or cliff, casting a shadow on a smooth plain, and the rays at the edge of the shadow tangent to the surface.) But this kind of resolution comes only in fleeting glimpses. An observer would have to confine his attention to a few favored candidate areas on Mars with alert attention, hoping for the superb glimpse to occur before the planet's rotation carries the local area out of the opportune circumstance, which would occur within a few minutes. I would estimate that the amount of time that there is atmospheric seeing of the high quality required to attain this goal is less than one per cent of the time that the planet is within three hours of the observer's meridian. No good telescope of 36-inch aperture or greater is available to such an interested observer. Let us come back to Lowell's case. His 24-inch refractor is an excellent instrument. I have looked at planetary detail with this instrument for a total of some 500 hours.

km/day) than on maria. Largest of canals have been photographed. Some visual observers claim under best seeing conditions canals are resolved into broken linear dark markings.

(e) *Oases.* Roughly circular dark markings usually at intersections of canals, diameters of the order of 150 km. Some 200 have been reported.

2. *Atmosphere*

(a) *Clouds.* Three cloud species, yellow, blue, and white, exist. (White clouds may be distinct species or only thicker blue clouds.) White or blue clouds may be observed on any part of the disk but are concentrated toward the limb. Angular sizes up to 45° in areographic coordinates (about 3000 km) extending along the limb are observed. On the morning side of the disk, clouds may extend almost to the noon meridian, on the afternoon side rarely over 45° from the terminator. During several periods of observation, morning clouds were photographed only over maria, afternoon clouds only over deserts. Clouds were regularly observed on the limb but not on the terminator at 40° phase, indicating that observed limb clouds may be an observational foreshortening appearance of atmospheric haze rather than a distinct physical phenomenon. White (or blue) clouds may form in less than 24 hours and last over two weeks. White clouds are most common at aphelic oppositions. A total of about two dozen measurements of cloud movements available. Speeds up to 35 km/day have been computed. Yellow clouds are usually associated with perihelic opposition. Major storms involving yellow clouds occurred near the 1924 and 1956 perihelic oppositions, resulting in the covering of the entire planet for several days with a yellow pall. After 1956 storm, polar cap reappeared quickly, dark markings more slowly.

(b) *Blue Haze.* Thin haze covering entire planet rendering surface features invisible in photographs taken in blue light ($\lambda <$ 4330 Å). Haze dissipates from time to time particularly near oppositions. Haze can dissipate or form in 3 or 4 hours. Clearings may be planet-wide or cover as small an area as ⅛ of the disk.

3. *Other Phenomena*

Large W-shaped cloud observed in 1926, 1954, and 1958 associated with oasis-canal network in Tharsis region (Tithonius Lacus), rotates with surface. Cloud bands parallel to equator on blue photographs and Y-shaped haze patterns on blue photographs from time to time. Short duration bright flash reported 1954 by Saheki.

From this outline of observed phenomena on Mars, it may be assumed with a fair measure of confidence that all contrasting daytime Martian phenomena photographable in the λ3500 to λ6500 range whose extents are greater than 50 km, which may be observed near opposition, and whose temporal durations exceed two or three days, are now known. This may be taken as the present description information space of Mars.

D. The Scale of Atmospheric Phenomena

It is also evident that the available *quantitative* observations with regard to sizes, rates of motion, rates of growth and decay and time spans of Martian atmospheric phenomena is very sparse. This deficiency is caused by three principal factors: (1) the low resolving power of photographs as limited by the seeing, (2) lack of observations taken often enough over long enough time spans, and (3) incomplete use made of the observations which do exist. This third point is attributable both to the fact that much data has not been published or made generally available and to the fact that a

(Cont.) This telescope is equipped with an iris diaphragm which the observer can conveniently regulate from an aperture of 24 inches down to 6 inches. In my experience, I have never been able to gain in finer detail with apertures larger than 20 inches. The secondary chromatic aberration becomes severe enough to spoil any view the seeing might allow. Lowell and his associates could not possibly have seen any Martian detail smaller than 0.2 arc-second. Therefore, the smallest possible horizontal distance he could have observed in the vicinity of the terminator would be 240 km. The smallest vertical height that could be detected would be 27,500 feet. Lowell was certainly wrong by a factor of 10. Thus, this means of mapping relief is beyond ground-based capabilities.

great many photographs have not been completely reduced. The present situation is such that the data required for synoptic studies of the Martian atmosphere are not available. This means that much of the analytical work which has been done so far has necessarily been top heavy with assumptions and overinterpretations of the available data.

What then is a description information space adequate for synoptic studies of a planetary atmosphere? In order to answer this question it is necessary to know the linear and temporal resolutions necessary to describe meteorological phenomena.

On the Earth the smallest atmospheric phenomena of meteorological significance are perhaps tornadoes and thunderstorms. These have a spatial extent of the order of from two to five km. From atmospheric events of this scale, sizes range up to planet-wide circulation patterns. What spans of observations and temporal revolving power should be used in order to observe adequately these atmospheric phenomena of various sizes? The answer depends on the lifetime of the phenomena and their rates of evolution.

Figure 1 shows a relation between average sizes and lifetimes of four types of terrestrial atmospheric phenomena: tornadoes (A); thunderstorm cells (B); hurricanes (C); and cyclonic storms (D). The approximate average size in kilometers is plotted as the ordinate and one-tenth the average lifetime as abscissa. It is assumed that a generally useful interval between observations of such phenomena is about one-tenth their lifetime. Thus, a tornado should be photographed every three minutes, a hurricane observed every 12 hours, etc.

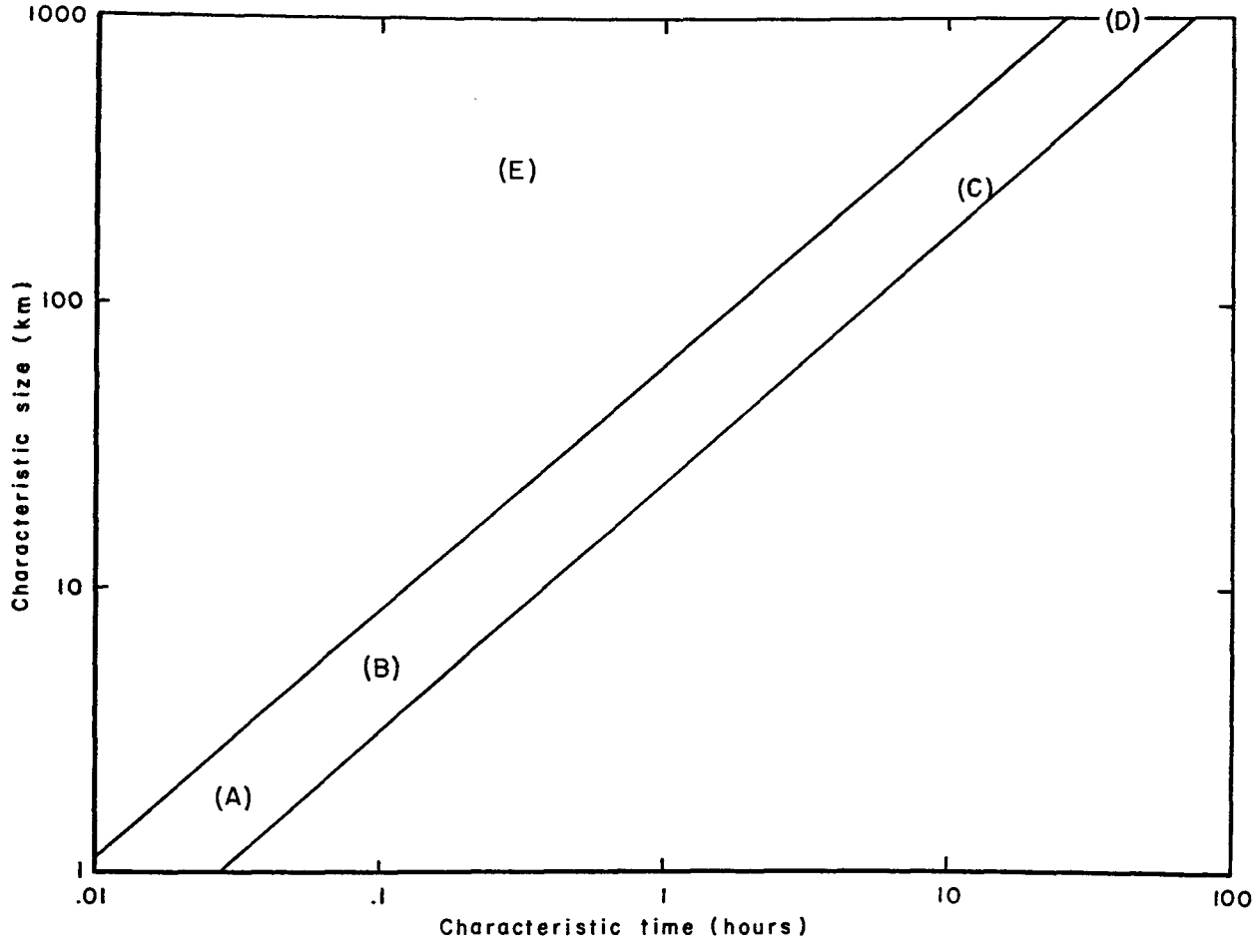


Figure 1

It is thus seen that the frequency with which a phenomenon must be observed to obtain an adequate description information space depends on the spatial extent of the phenomenon, Fig. 1 giving the relation between extents and frequencies for terrestrial atmospheric phenomena. The time span for the observations should be about ten times the plotted characteristic time.

It is reasonable to assume that in the atmosphere of Mars (and other planets) a similar relation of the type

$$s = a\,t^b$$

holds between a characteristic size s and a characteristic time t for several classes of atmospheric phenomena. There is no *a priori* reason for assuming, however, that the coefficients a and b in the equation are the same as for the Earth, though this may prove to be the case. But as a starting point for the determination of the suitable instrument information spaces for Mars, the relations of Fig. 1 can be used to give the temporal resolutions and spans required to describe the meteorological phenomena of various sizes.

An interesting exception to relations of the type illustrated in Fig. 1 is given by (E), the plotted position of the characteristic size and time of a *thunderstorm complex.* Such a complex consists of a great many thunderstorm cells and may be considered as an *aggregate phenomenon* rather than a simple phenomenon.

In the exploration of another planet, exceptions to characteristic time vs. characteristic size relationships may indicate the presence of such aggregate phenomena and suggest areas of investigation which call for observations with higher resolutions. This concept also indicates how, when certain basic relationships have been determined, the epistemological resolving power of the exploration may be extended beyond the actual resolving power of the observational instruments.

To aid in determining the set of instrument information spaces that would optimally span a description information space for synoptic studies of the Martian atmosphere, comparative instrument information spaces are given in Table IV for a 60-inch telescope located on the Earth's surface, a 20-inch telescope in a balloon at 30,000 meters, and a 5-inch telescope in a fly-by probe 40,000 km and 1,000,000 km from Mars.

In Table IV the characterizing parameters are listed on the left column; the other columns show their relative ranges.

The exposure time is limited on the short end by signal-to-noise ratio and on the long end by emulsion saturation, atmospheric turbulence, sky brightness and the relative motions of the object and camera.

The matter of economics of ballon flights is complex. Flights for isolated purposes such as obtaining physical data in spectral ranges inaccessible on the surface of the Earth, or taking photographs with a finer resolving power than can be made from the surface, can be readily justified. But the relative cost of using balloons for patrol purposes, such as the monitoring of atmospheric changes, as compared with surface observations, must be questioned. The atmosphere does not present a filter against detection of changes in phenomena that are already observable from the Earth's surface.

E. Conclusions

About all that can be learned without great effort and expense from the Earth's surface concerning *static* phenomena on the nearer planets has been learned. It would be fruitless, for example, to try to continue to make marginal gains in resolving power against the obstacles of seeing when telescopes above the atmosphere can give orders of magnitude improvements.

Many plates of planets taken at various observatories over the past half century have not been reduced quantitatively to obtain the physical data which they contain. In view of the obvious superiority for planetary studies of observations or physical measurements made from outside the Earth's atmosphere, it would not be worthwhile to reduce most of this plate material.

Table IV

COMPARATIVE SYSTEMS INFORMATION SPACES FOR PHOTOGRAPHY OF MARS

	For Mars at Most Favorable Opposition: Distance 56 × 10^6 km		5-inch Telescope in Fly-By Probe at 10^6 km from Mars	5-inch Telescope in Fly-By Probe at 40,000 km from Mars
	(A) 60-inch telescope at Earth's surface	(B) 20-inch telescope, balloon-mounted, elevation 30,000 meters		
Angular Field	Entire planet = 25 seconds of arc.	Entire planet = 25 seconds of arc.	Entire planet = 20 minutes of arc.	Entire planet = 10° of arc.
Angular Resolving Power	Theoretical optical, 25 km; seeing limited, 140 to 560 km.	Theoretical optical, 75 km; guidance limited, 3 km.	5 km.	0.2 km.
Spectral Range	0.32μ to 0.9μ.	0.29μ to limit of emulsion.	Limited only by Martian atmosphere and emulsion.	
Exposure Time	Signal : Noise limit to seeing and emulsion saturation limits. Rotation of Mars limit, 0.01 sec to 2 min.	Signal : Noise limit to emulsion saturation. Rotation limit, guidance limits, 0.01 sec to 2 min.	Limited by the velocity of the probe relative to the surface of Mars, guidance limits.	
Frequency of Exposures	Any frequency up to reciprocal of exposure time.	Same as surface except for uneconomical frequency intervals.	Up to reciprocal of exposure time.	
Number of Exposures	Limited by economic and data processing factors.	Limited by economic factors governing number of flights.	Limited primarily by data storage and transmission capability; and possibly by frequency of exposures and allowable span of observations.	
Span of Observations	Limited by planetary configurations and sky brightness. Possible up to 45° from Sun.	Daytime observations possible up to 5° from Sun.	Determined by orbital parameters of probe.	

Nonetheless, a large and important role remains for surface observations. (v. Chapter V. C.) Essentially no quantitative work has been done on the *changes* in planetary phenomena. Dynamic planetary phenomena (atmospheric phenomena for the most part) have not been observed frequently enough or over long enough time spans to afford any but the most vague ideas of their properties.

This situation indicates that intensified uniform observations from the surface of the Earth with good instruments, geographically spaced to give a complete coverage of Mars and complemented with an efficient data distribution and reduction facility, is absolutely essential to fill the present gap in our knowledge of planetary atmospheres. It would also be most important to re-examine and reduce according to a standard procedure those existing photographic observations which are suitable for studies of changes. Particularly, this is important for studies of secular changes, for which they are our only source of data.

The following programs for *direct photography* can be recommended:

1. *From the Earth's Surface.*
   (a) Study of existing plate material for data on secular changes.
   (b) Study of existing plate material for *quantitative* data on atmospheric dynamics.
   (c) Continuing program of observations for data on the dynamics of atmosphere, motions of clouds, storms, the formation and decay of phenomena. These observations should include time-lapse photographs, taken with various colors and polarizations, which can be differentially superimposed to study various aspects of changing phenomena.
2. *From Balloons*
   (a) High-resolution direct photographs.
   (b) Short-term time-lapse photographs in color, for observing spans up to 12 hours.
3. *From Probes*
   (a) Reserved for detection of phenomena beyond present resolutions.

Appendix 2

VISUAL AND PHOTOGRAPHIC OBSERVATIONS OF VENUS AND MARS

Clyde W. Tombaugh

A. Venus

A very important requirement in visual observations of Venus is the reduction of excessive light so that irradiation and scattered light will not obliterate difficult markings, whether they be fine or of very low contrast. Fortunately, the angle subtended by the disk of Venus is fairly large, and useful visual observations have been made by the writer using a telescope with only a 4-inch aperture. Especially for Venus, modest apertures work better because it is necessary to reduce the effects of atmospheric turbulence to the utmost. The slightest scattering of light from the brighter adjacent areas bounding a darkish marking of very low contrast will obliterate it.

Since the use of higher magnifying powers is more vulnerable to atmospheric smearing, we have to be content in ground-based observations with seeing the larger markings within the limits of aperture resolution. The writer has found from experience that a magnifying power of 200 diameters is a practical upper limit for Venus, and any aperture larger than 6 inches requires the use of a neutral density filter of the proper opacity to reduce the effects of excessive irradiation.

In a period of 30 years, the writer has looked at Venus at least 2,000 times in all of its phases. He has seen cusp brightenings and weak darkish markings only a few dozen times.

Ultraviolet photographs of Venus yield detail on most days, and the patterns change frequently and suddenly. The most common markings in UV are a few bright stripes, more or less parallel, which cross the disk roughly perpendicular to the line joining the cusps. Sometimes they are broad, other times they are narrow. Occasionally lesser stripes meet or intersect the others at steep angles. Infrequently, bright curved markings are recorded. Often there are drastic changes in the pattern within 24 hours. At other times, the pattern appears roughly the same from day to day; but, because of the Earth's diurnal break in the observations, one is not sure that a bright marking has retained its identity, or whether the old one has dissolved and a new one formed. The bright parallel stripes may indicate some kind of rotating planetary wind system that would serve as an index to the position of the rotational axis. The Venus system appears to be easily modified or even disrupted, quite unlike the behavior of Jupiter's markings in the UV. This behavior suggests that the forces giving direction to the stripes are weak and that the rotation period of Venus is very long.

These bright stripes (some may call them bands, but perhaps "bands" should be reserved for spectroscopic phenomena) in the UV are very weak on the blue plates, and they are invisible on the yellow plates.* Indeed, the disk of Venus is generally featureless in yellow, red, and near infrared. On the hypersensitized Z plates, through UG–8 Schott filter at 1.1 microns, there are suspicions of weak markings.

B. Mars

For effective visual and photographic observations of Mars the minimum telescope apertures are 8 and 12 inches, respectively; the optimum apertures are 16 and 24 inches, respectively. Larger apertures run into diminishing returns rapidly, and severe disadvantages are incurred beyond 60 inches

* Editor's Note: Dollfus has reported a weak pattern in the yellow.

from optical aberrations and atmospheric turbulence. The surface of Mars is much less brilliant than Venus. Nevertheless, for visual observations excessive irradiation is the greatest shortcoming practiced by users of large telescopes. Fine detail on the disk of Mars is best seen through an amber color filter and by employing a magnifying power between 30 and 40 diameters per inch of aperture. For evaluations of the colors of various features on the disk, a reflecting telescope must be used with all selective color filters removed.

In reflecting telescopes, diffraction disturbances to fine definition arises from the secondary mirror and especially from its spider support. This effect can be minimized almost to the point of imperceptibility if the diameter of the secondary (Newtonian or Cassegrain) is kept to one-sixth that of the primary mirror, and the supporting spider legs are strongly curved to distribute the diffraction. In the April 1950 opposition of Mars, the writer, with G. P. Kuiper, found marked improvement in the definition of the 82-inch McDonald reflector by the use of an eccentric diaphragm with an aperture diameter of 27 inches placed between the edges of the primary and secondary mirrors in a radial direction and between the shadow of the spider supports of the secondary.

In visual observations, a good and well-trained eye, much looking, and patience are required to glimpse the finest detail during the infrequent moments of superior seeing, even on good nights. The best training to develop seeing alertness and skill is to attempt drawings of the planet at the eyepiece. Much of the disagreement among observers is due to the lack of sensing spatial proportions on the disk—the shapes, intensities and positions of the markings. Some may see well, but they cannot handle a pencil skillfully. Observers with many years of experience and with intimate familiarity of the many markings in all of their varied appearances may substitute descriptive notes, because any unusual aspect quickly registers in his attention as departing from the normal. Seasonal changes are pretty much repeated each Martian year. Then there are several cases where a conspicuous development occurred to a particular dark marking which was never witnessed before in the lifetime of a veteran observer. A recent instance of this was the extensive darkening of the Thoth and Nubis Lacus area of Mars in 1954, which has persisted through 1960 with some fluctuations. The Ganges, Lucus Lunae, Nilokeras, Pandorae Fretum, and other features are noted for their unpredictable variability. Several dark, ephemeral appendages to the Syrtis Major have been observed.

The reverse phenomena has been observed, also. In the 1870's and 1880's, maps drawn by Flammarion, Green, and Schiaparelli show the Nilosyrtis coming out of the north tip of the Syrtis Major and curving Martian westward for a total distance of a thousand miles. Since then it has declined to a narrow and inconspicuous marking. Certain other maria, such as Sirenum, Cimmerium, Sabaeus, Margaritifer and Erythraeum have not changed in shape. All of this suggests that there are areas on Mars of an intermediate altitude where the supposed vegetation flourishes at certain times and not at others.

As a rule, the various maria have their characteristic colors which undergo regular changes with the seasons. The writer has noted that the maria in the southern hemisphere range from green to gray with the progress of the seasons. There are many fewer maria in the northern hemisphere, and they show very little green; they range from black to gray. However, in one of them, the Mare Acidalium, the writer has seen a distinct chocolate brown color at times. Two equatorial areas, the Dawes' Forked Bay and the northern portion of the Syrtis Major, have ranged from black to a beautiful deep blue color. When the northern part of the Syrtis Major is deep blue, it has a very sharp demarcation with the southern portion, which is grayish green at the time.*

C. Supplementary Remark on Mars

Note: The following communication from C. W. Tombaugh was received following distribution of the text of Chapter III to members of the Panel. It contains an account of some of the author's

* Editors' Note: Several other observers have reported a failure to detect these vivid colorations on the Martian surface during recent oppositions.

visual observations of Mars and some of his conjectures on the ice caps and their interpretation. This material was not discussed by the Panel, and does not necessarily represent a consensus of the opinions of other observers of Mars.

Comment on Chapter III. B.5. and 6.: The highest temperatures on Mars are expected to be in the high southern latitudes soon after summer solstice of that hemisphere, and not on the Martian equator. Integration curves show that the greatest solar insolation received is at the poles of a planet with inclinations like Mars and the Earth, at the time of respective summer solstice. The reason the temperature does not go up on the Earth's polar regions is due to the reservoir of cold created by the thick deposits of ice. On Mars, after the cap recedes to a small remnant, the reservoir of cold is removed and the soil gets pretty warm (Lampland and Coblentz). This prepares the way for my next point.

I have observed Mars through two complete seasonal cycles (each requiring 15 years). I have noted that when either the north or the south polar cap tips toward the Sun in its respective early spring season, the edge of the white is not enhanced by a dark border. At first these polar caps appear to consist of cloud, since they are not intensely bright and no dark terrain detail is seen within their confines; also, they are strong on the blue photographs. Soon after the vernal equinox, the cloud appears to dissipate, apparently by deposition of frost, leaving the polar caps dazzling white. The bright cap starts to shrink, apparently by sublimation. By mid-spring a dark, narrow band appears tightly around the white border. As the season progresses, the dark band around the north cap becomes blue in color and wider than the one that develops around the south cap. I have never been able to see any blue around the south cap border—only black.

Lowell and Slipher also have seen the blue band around the north cap, and Lowell ascribed this to a temporary sea of liquid water. I cannot accept such an explanation because the cap must be much too thin to recede the way it does, and therefore could not yield such a volume of water. Also, I have observed that the dark bluish band is perfectly continuous. This would require a phenomenal flatness in the terrain. I would consider the dark band as merely wet ground with a bluish cast from the atmosphere since we view it quite obliquely. If this is wet ground, it implies that after sublimation had been in progress for some time, a concentration of water vapor locally at the edge of the cap built up the partial pressure enough to permit water to exist in the liquid state, and thus produce wet ground. The triple point for water occurs at 0.0075°C and 4.6 mms of mercury pressure. In Martian May, the north cap is enveloped by a translucent veil which extends a few hundred kilometers beyond the dark band. The bright cap and dark band are visible through this veil, but considerably dimmed, as the veil itself is moderately bright. The ground cap continues to shrink in size. In Martian June, the veil vanishes for good after a few weak recurrences. A very dark, narrow rift invades the cap deeply by a canal running to the Martian northeast from the north edge of Mare Acidalium. The dark surrounding band which has retreated tightly with the shrinking white area now declines and disappears, but the edge of the white area is sharply defined. The small north polar cap remnant never disappears, and is perfectly centered on the geographical pole, unlike the south polar cap remnant. The latter is de-centered six Martian degrees from the geographical pole, always in the exact same position, which I attribute to a local horst plateau. There it continues unchanged for many weeks. The north cap remnant also persists for several weeks with little change. During the shrinking stage of the north polar cap, it has always appeared to be quite circular (after allowing for foreshortening) and of uniform whiteness, unlike the south polar cap in both of these respects. This indicates relatively flat land at about the same altitude. Indeed, the reduced north cap appears to be located in the central area of a tetrahedral face; whereas, there is evidence for a tetrahedral vertex in the vicinity of the south cap. The irregular shape and break-up of the latter suggests irregular terrain which one might expect to find in a "vertex" area. It is improbable that the north polar cap remnant sits on a horst plateau exactly centered on the geographical pole. The reason for the persistence of the remnant may then be caused by a polar cap core of much greater thickness. Aphelion occurs a little before the

northern hemisphere summer solstice, and the decreased solar insolation at this crucial season apparently is sufficient to inhibit the complete dissolution of the cap. Consequently, a core of compacted frost may have attained considerable thickness—much thicker than anywhere else on the planet.

For the south polar cap, perihelion would accelerate the rate of shrinking, which is observed. The warmer temperatures of the surrounding area would hasten the evaporation of the wet ground, causing the dark band to be very narrow.

Appendix 3

RADIO FREQUENCY RADIOMETRY OF THE PLANETS

B. F. Burke

A. Introduction

While there is little difference, in principle, between infrared radiometry and radio frequency radiometry, it has proved in practice that in the radio domain of the electromagnetic spectrum the observed emission does not necessarily arise from a thermal equilibrium process. This has required a classification of radio objects into "thermal" and "non-thermal" emitters, depending on whether it appears likely that the apparent radio temperature refers to a meaningful physical temperature or not. The same object may exhibit both non-thermal and thermal characteristics, and it then becomes necessary, by interpretation of the continuum spectrum, to separate the two components. It was expected that the planets of the solar system (with the exception, of course, of the Earth) would be purely thermal emitters of radio radiation, but this has not proved to be the case. Table I gives a summary of the objects studied so far, and classifies the radiation at various wavelengths into these two general categories:

Table I

Mercury	3 cm	Thermal
Venus	8 mm-10 cm	Thermal
Mars	3 cm	Thermal
Jupiter	10 m-30 m	Non-thermal
	10 cm-75 cm	Mostly non-thermal
	3 cm	Mostly thermal
Saturn	3 cm	Probably thermal

All radio observations are usually made in terms of antenna temperature, T_A, which is the temperature to which an equivalent resistive load would have to be heated to give the same noise power per unit bandwidth. If the antenna is perfect, and if the object being studied is a blackbody at temperature T which completely fills the antenna beam, then $T_A = T$. In practice, (1) the antenna is never perfect, (2) the reference load is never an exact reproduction of the antenna, and (3) the object seldom fills the beam, and as a result uncertainty is introduced into the measurements. All three factors change with wavelength.

In most of the observations taken so far, the disk of the planet subtends such a small angle that it is better to regard the problem as one of measuring source flux, where, instead of antenna beam size, the effective antenna area A is the relevant parameter. Then, for a blackbody at temperature T, subtending solid angle ω at wavelength λ, we can relate all parameters through the Rayleigh-Jeans law:

$$S = \frac{2kT\omega}{\lambda^2}$$

$$\text{and } SA = 2kT_A$$

where S is incident flux per unit bandwidth. Consequently, the equivalent blackbody temperature of the object being studied is

$$T = \frac{\lambda^2 T_A}{\omega A}$$

Measurement of T_A is limited by the available signal-to-noise ratio, and by the precision of the temperature standards. The effective antenna area A can seldom be measured directly, and is probably the greatest contribution to uncertainty. (It is rarely known to better than 10% and has been known to be in error by a factor of 2.) The effective solid angle ω depends on the nature of the emitter, but has usually been taken as the solid angle of the visible disk. The planets, of course, are not perfect blackbodies, and are not of uniform temperature, but it has been hoped that these assumptions are not a bad first approximation. In comparing observations with specific models of planetary atmospheres and surfaces, the reverse procedure is clear: an effective temperature can be computed and compared with the observed effective temperature.

B. Venus: Observations

The observations of Venus at radio wavelengths are summarized in Table II. The effective temperature at opposition has been extrapolated from the data given by the various authors. The uncertainty for this point is consequently much greater than for T_{Av}.

The conclusion seems inescapable that an effective temperature in excess of 500°K exists near inferior conjunction for wavelengths in the range of 3-10 cm. It appears probable that the effective temperature at 8 mm is considerably lower, although judgment of the accuracy of the 8 mm point of Kuz'min and Salomanovich is difficult, since they did not discuss how they determined their effective area. The difficulty of absolute radiometry appears here directly and inescapably, improvement in the reliability of the data resting principally in improved (and independent) determinations of antenna gain. An indirect check is possible at 10 cm, where an independent, absolute measurement of the radio source Cas A has been made by Broton and Medd, a source which has also been measured by the workers at NRL. The flux measurements of the two groups agree to within 8 per cent. Consequently, the 10-cm effective temperature is probably the most certain. The 3-cm values are not independent, but eventually refer to measurements made with equipment that is essentially a scaled version of the 10-cm NRL equipment.

Table II

Observer	*Wavelength (cm)*	T_{Av} (°K)	*Phase angle*	*T (inferior conjunction) (°K)*
M.Mc.S.	10.3	$600° \pm 65°$	$\Phi = 230°$	480°
M.Mc.S.	3.4	$575° \pm 60°$ $585° \pm 53°$	$\Phi = 146°$ $\Phi = 225°$	550°
A.G.M.T.	3.37	$575° \pm 58°$	$\Phi = 275°$	
G.McE.	0.86	$410° \pm 160°$		
K.S.	0.8			$315° \pm 70°$

M.Mc.S. = Mayer, McCullough, Sloanaker (NRL)
A.G.M.T. = Alsop, Giordmaine, Mayer, and Townes (NRL, Columbia)
G.McE. = Gibson and McEwan (NRL)
K.S. = Kuz'min and Salomanovich (Lebedev)

The question of the magnitude of a phase effect is of great importance, and all of the long series of data point to an increase in the effective temperature as the planet moves away from inferior conjunction. There is an annoying experimental complication, for all the best measurements are made near inferior conjunction, when the planet is presenting the largest solid angle. As the illuminated fraction of the disk increases, the distance to the planet increases, and the resulting decrease of solid angle has unfortunate effects on the signal-to-noise ratio. For example, at 3 cm the thermal fluctuations of the Earth's atmosphere limit the sensitivity of receivers to measurements of antenna temperatures of 0.05° to 0.10°. (Exceptional circumstances have been reported when the "atmospheric noise level" decreased to 0.01° or 0.02°, but these conditions do not seem common, at least in the Eastern U.S.) With a 50-foot dish, Venus at inferior conjunction gives an antenna temperature of about 3.5° K, for a signal-to-noise ratio of about 50 to 1. (The results reported in Table II were made with poorer noise levels, however.) By the time the planet is at dichotomy, the antenna temperature is only 0.5°, and at superior conjunction, the antenna temperature is about 0.1°. Errors are thus introduced, from the measurements themselves, which would mask any but the largest effects.

It seems likely, however, that the 3-cm effective temperature is about 50° higher at quadrature than at inferior conjunction, unless the measurements contain systematic errors. The 8-mm measurements suggest a still larger effect.

Several years ago Kraus reported intense radio noise bursts from the planet Venus in the 10-meter range which would have had to be of non-thermal origin. Workers at the Carnegie Institution of Washington, Yale University, the National Bureau of Standards, and the University of Florida have all made unsuccessful efforts to confirm the existence of these bursts, and a retraction has been published by Kraus. It is likely that the observed bursts were of terrestrial origin.

C. Venus: Discussion

(See also Chapter IV.)

The radio temperatures observed at centimeter wavelengths from Venus are so much higher than the accepted infrared temperatures that one wonders if the measurements refer to a surface temperature of the planet. If the effective temperature of the surface is 600° (or higher) a model atmosphere must provide sufficient opacity at 8 mm to diminish the surface contribution and raise the atmospheric contribution. Models of this sort have been proposed by Sagan and Barrett, relying on water vapor and CO_2 to increase the opacity at 8 mm. The contribution to the opacity from CO_2 depends on frequency ν, partial pressure P_{co_2}, and temperature T as $\nu^2 P_{CO_2}^2 T^{-3.7}$. This non-resonant absorption process arises from the collision-induced dipole moment of CO_2. The contribution from H_2O is greatest at the frequencies of the rotational lines at 1.34 and .162 cm; at other frequencies the opacity is determined by the wings of these and other lines in the infrared. Observational tests of these model atmospheres can be made by observing Venus at wavelengths close to 1.34 cm, where the optical depth is great enough (for concentrations of the order of 1% H_2O) to lower the apparent temperature from 600° to 300°. The effect of terrestrial water vapor can be allowed for by referring to the Moon as a relative radiation standard.

Other polar molecules, if present in sufficient quantities, might be detectable. Barrett has mentioned the lines of CO (at 2.6 mm), O_2 (5 and 2.5 mm), NO_2 (1.13 cm), NO (1.99 mm), and N_2O (1.19, .597, .298 and .239 cm). If we assume that an effect of 10% in the effective temperature could be detected, concentrations of the order of .01 to 1% would be required. In all cases, radiometric temperature decrease at the line frequency would be noticed.

The alternative explanation to that of a hot surface covered by a greenhouse-like atmosphere is absorption by free-free electron transitions in a dense ionosphere. For all frequencies far above the

maximum critical frequency, the absorption varies as ν^{-2}; hence, at sufficiently high frequencies, one sees through the ionosphere. At lower frequencies, the optical depth increases, until one eventually sees only the temperature of the ionosphere. With this interpretation, the radio data suggest that the dark side of Venus has an extremely dense ($n_e \sim 10^9 e/\text{cm}^3$) and thick (100-200 km) ionosphere. Sagan has advanced good arguments against the existence of such an ionosphere; but if we ignore the theoretical considerations and seek possible experimental tests, it does appear that such an ionosphere could be either detected or disproved.

Let us, for purposes of simplicity, consider an isotropic isothermal ionosphere, $T = 600°$ K, $n_e = 2 \times 10^9\ e/\text{cm}^3$, and thickness $s = 150$ km. Although this model is oversimplified, we can easily calculate its basic radiation properties, which will in general be close to those of a more physically realizable ionosphere. At both 10 cm and 3 cm, such an ionosphere would have appreciable optical depth (20 and 2, respectively) while at 8 mm, over most of the disk, the optical depth would only be of the order of 0.2. At the rim, however, the optical depth would still be large, and so at 8 mm one would expect to see a 300° disk with a thin 600° rim.

Such an effect is probably detected most easily by a radiometer mounted on a vehicle making a close pass. In principle, the measurements could be made from the surface of the Earth by using interferometric techniques. It is not known yet whether the atmospheric temperature fluctuations that limit the sensitivity of 3-cm radiometers would affect an interferometer—this would depend on the fine structure of the atmospheric water vapor fluctuations. It seems probable that the signal-to-noise ratio would be improved, however. If such is the case, a separated pair of 40- or 50-foot dishes on the Earth might be a powerful tool for studying the planet Venus. Actually, three dishes would be even more valuable, for by using Jennison's method, phase as well as amplitude of the Fourier transform can be measured. So far, such measurements at long wavelengths have been made only to interferometer spacings of 3000 λ, which was ample for the radio sources under study, but there seems to be no fundamental limitation to going to spacings an order of magnitude or more greater. (This would put the dishes about 1 km apart when working at 3 cm.) Such a series of interferometric measurements would not only serve as a way of looking for limb brightening, but it would also give some indication of the temperature gradient across the disk from the sunlit limb. If the dishes can be used North-South as well as East-West, the gradient of temperature from equator to pole could also be measured. Even a simple interferometer (two-element) giving only amplitude information can help in those measurements where 180° reflection symmetry can safely be assumed. It is entirely possible that a constant-spacing interferometer, with rather coarse lobes (say 1 minute of arc) would be a more powerful tool than a single dish in radiometric measurement of the integrated radio brightness of Venus if the interferometer proves less susceptible to "atmospheric noise."

It should be emphasized that a maser receiver of high relative phase stability is necessary for each dish. The other possible sources of difficulty are: 1) ability to align the dishes to within a fraction of a wavelength over baselines of several kilometers—standard geodetic methods should suffice; 2) stability of the atmosphere—probably not a problem, at least to 100,000 λ; and 3) mechanical stability of the dish structures; present day experience with Blaw-Knox and Kennedy dishes indicates no limitation here. The principal limitation undoubtedly is the personal one—only a few people in the world are competent to lead such a project, and all are interested in other things.

Appendix 4

POTENTIALITIES OF RADAR FOR THE STUDY OF PLANETARY ATMOSPHERES

VON R. ESHLEMAN

A. Introduction

Using ground-based radars which are now under construction, and future systems which are technically feasible, it appears that it should be possible to make many planetary radar measurements which would be of importance in determining characteristics of planetary atmospheres. Included in such radar studies might be the determination of: (1) the axis orientation, rotation speed, diameter, sphericity, surface roughness, and large-scale surface features of the terrestrial planets, notably Venus; (2) characteristics of the solar corona, solar streams, and the interplanetary medium; (3) the presence, critical wavelength (maximum electron density), height, integrated electron density, and magnetic field of planetary ionospheres; and (4) the height, radiowave absorption, and large-scale features of planetary cloud covers. Some of these determinations are much simpler than others, some may prove to be impossible, and results, important to the study of planetary ionospheres, may be obtained which are not even mentioned here.

In a sense, the history of the use of radar for the study of planetary atmospheres is as old (about 35 years) as the history of radar itself, since the first target of the first radar was the ionosphere of the Earth. Radar technology developed rapidly during the World War II years, but the last few years have seen even greater advances. Still, of substantial astronomical bodies other than the Earth, only the Moon can be studied in any detail by present-day radars, and with these radars it is only barely possible to detect the Sun and one planet, Venus. The great barrier to the advance of radar astronomy has been distance, because the radar echo intensity varies inversely as the fourth power of the range to the target.

Since we are specifically concerned here with the planets other than the Earth, it might appear from the previous paragraph that radar has little or no role to play in the near future in the exploration of their atmospheres. But such a judgment would be premature on two counts. First of all, with the use of space-probe-borne radars the distance or radar range can be greatly reduced, so that important measurements could be made with radar systems of modest size and power. Secondly, large ground-based radar systems have now bridged the distance barrier to the planets, so that further increases in sensitivity—which are being made very rapidly—can be applied to the study of particular characteristics of the targets.

In information-theory language, the mere detection of Venus represented one bit of information. Using the 10,000-times more powerful systems now under construction, it should thus be possible to obtain a "picture" of certain characteristics of Venus with detail corresponding to 10^4 bits of information. Of significance also is the fact that this same increase in sensitivity should make it possible to detect Venus, Mars, Mercury, and Jupiter not only at closest approach to the Earth, but also at essentially all positions in their orbits. Further increases in radar sensitivity could be used to increase the detail with which these planets were studied, and to increase the range of detection to more distant targets.

The first radar echoes from Venus appear to have provided a significant improvement in the accu-

racy of the astronomical unit in terrestrial standards. In this rather marginal experiment,* a value for the AU of 149.47 $\times$ 10^6 km with an accuracy of about one part in 10^5 was obtained, compared to about one in 10^3 from previous methods. While this result appears to have little direct significance with respect to planetary atmospheres, it is indicative of the complementary role radar can play in space research. Before radar, range was difficult to measure directly, and use had to be made of the laws of celestial mechanics together with the great angular accuracy obtainable with telescopes. In radar, angles are very poorly defined, but range can be determined with great precision. With reasonable radar signal strengths, the accuracy of the AU will be limited only by the accuracy of the value of the velocity of light in vacuum.

In view of the present and potential development of large ground-based radar systems, and the great importance of continuity of study over an extended time period, it is my opinion that the use of radar for studies of planetary atmospheres and surfaces should be relegated primarily to ground-based systems. There may be, of course, certain special problems which require study by probe-borne radars.

B. Possible Radar Experiments of Importance in Studying Planetary Atmospheres

In the first paragraph of the introduction, the possible radar experiments of importance in studying planetary atmospheres were given in the order: (1) surface features and rotation; (2) solar-planetary relationships; (3) ionospheres; and (4) cloud cover. This I believe corresponds to the likely order of achievement. Those experiments under (1) come before (2) and (3), because short-wavelengths will be used for (1), and such systems are more sensitive (because of less cosmic noise and more antenna gain for a given sized antenna) than the long wave systems required for the studies listed under (2) and (3). The studies under (3) and (4) are very difficult to design and discuss in detail, since nothing is known about the detectability by radar of planetary atmospheres and ionospheres. At least we have some evidence that planetary surfaces and the solar corona produce radar echoes.

1. *Surface Features and Rotation*

Several radars are now being built having sensitivities on the order of 10^2 to 10^4 times the sensitivity of the Lincoln Laboratory and Jodrell Bank radars used in the first detection of Venus. These include the 1000-foot zenithal radar telescope in Puerto Rico (Cornell University and the U. S. Air Force), the 600-foot steerable paraboloid of the U. S. Navy, a 150-foot steerable paraboloid (Stanford Research Institute-Stanford University) and a 120-foot antenna for wavelengths as short as 3 cm being built by the Lincoln Laboratory. Transmitters for the first three antennas would have power capabilities measured in megawatts, while the 3-cm transmitter would be on the order of tens of kilowatts. Even larger systems have been proposed.

With such a system it would appear that the rotation of Venus could be measured with great precision, unless all of the surface were very smooth.** Two methods come to mind, one based on the expectation that non-uniformities in the echo strength as the planet rotates would show up as a repetitive pattern, and the other based upon the range-Doppler coordinates of the echo. (Because of the depth of the planet, a short transmitted pulse would be returned as a long echo, and because of planet rotation, the echo spectrum would be spread in frequency. Echoes would thus be spread in range and Doppler frequency, and each range-Doppler coordinate would correspond to unique positions on the planetary surface.) Great precision would be possible from the first method by observing a very large number of rotations. If the planet were "uniformly rough," the first method would not work, but the second method could be used to measure rotation. Note, however, that the measured Doppler fre-

* Since this appendix was written, further determinations by the radars at Goldstone (JPL), Millstone Hill (Lincoln Laboratory), and Jodrell Bank, England, as well as in the Soviet Union, have somewhat modified this value. The U.S. radars now give a value for the AU of 149.5995 $\times$ 10^6 km.

** First attempts to determine the rotation rate of Venus by radar have recently been reported; v. Chapter IV. B.5.

quencies are due both to the rate of rotation and the angle of the planetary axis relative to the line of sight. Since this angle would change over long observing periods, it should be possible to determine both the absolute rotation rate and the axis orientation. The sense of rotation would still not be known, but this could possibly be determined from interferometric measurements, as could also the rate of rotation and axis position.

These same measurements could be used to determine the radius of the planet. Even if there were no relative rotation of the target planet, the spread in range of the echo would give a rough indication of the radius. It appears that the Lincoln Laboratory radar could be used to show the radius of the Moon to an accuracy of at least one part in 100 from the echo shape alone. But a much more accurate measure would be possible from the use of both range and Doppler discrimination. On Venus, it should be possible with the new radar systems to determine a radius of the solid surface which, when compared with the visible angular size, would give a useful measure of the height of the top of the cloud cover.

Other surface characteristics which may be important in studies of planetary atmospheres include large-scale surface features (plains and mountain ranges), average surface roughness, electrical characteristics of the surface material, and departure from sphericity of the surface. The average surface roughness might be determined by noting the changes in echo shape accompanying changes in operating wavelength, while clues to the electrical properties of the surface could be obtained from the changes in the echo strength at different wavelengths. By resolving the echo in range and Doppler frequency, these features could be studied for many different surface areas. Large-scale phenomena would show up in such a range-Doppler radar-map. Since very accurate measures of relative range to the sub-radar point on the planetary surface would be possible, small departures from a spherical shape of the planet could be detected.

2. *Solar-Planetary Relationships*

Several powerful systems operating at the long-wavelengths needed for reflecting echoes from the solar corona will soon be in operation. They include large antenna and transmitter installations in Texas (Lincoln Laboratory field site near El Campo), in Peru (National Bureau of Standards), and at Stanford University. These systems will have from 10 to 10^3 times the sensitivity of the original Stanford system used in the first radar detection of the Sun. With them it should be possible to study in some detail the changing structure of the solar corona at heights determined by the radar wavelength. The relationships between solar and terrestrial events is of prime concern, but with more intense studies of planetary atmospheres we would also be interested in the effects of unusual solar phenomena on these atmospheres.

The same radar systems, used in conjunction with the short wavelength systems, could be used to determine the average electron density in the space between the Earth and a planetary target. This measurement would be based on the difference in propagation velocity of short and long radio wavelengths in an ionized gas. In fact, in order to determine the AU with the greatest precision, it would be important to measure the propagation time at several wavelengths in order to take into account the effects of the interplanetary medium. Thus both the AU and the average gas density would be obtained from the same experiment.

It appears possible that, using radar, one might be able to detect a solar eruption, trace the development of the disturbance out through the corona, monitor the changing interplanetary gas as a solar stream spews out into space, and detect the effects of the stream on a planetary ionosphere.

3. *Ionospheres*

We are almost totally ignorant of the characteristics of any ionospheres which the planets may possess. In order to detect a planetary ionosphere from a ground-based radar, the radar wavelength

must be sufficiently short for the wave to pass through the terrestrial ionosphere, yet long enough to be reflected from the planet's ionosphere.

If the maximum electron density in the ionospheres of Venus and Mars are similar to what the density in the Earth's ionosphere would be at their distances from the Sun, then the maximum density for Venus would be 1.4 times that of the Earth, while that of Mars would be only about 0.6 times the terrestrial value. Day to night variations of maximum electron density in the Earth's ionosphere is on the order of four to one. Thus, we might expect that we could look out through our nighttime ionosphere and detect Mars' daytime ionosphere, but under the above conditions we could not look out through our daytime ionosphere and detect the nighttime ionosphere of Venus. However, these considerations are based on very flimsy arguments, and we should approach this problem with a full realization that the number of variables and unknowns are such that only a good measurement can be definitive.

If the planetary ionospheres are less dense than the minimum terrestrial ionospheric density, then it would appear that probe-borne radar would be the only possibility for the kind of ionospheric studies being considered here.

If the planetary ionosphere is accompanied by a magnetic field, the radar echo will have two components of different delay and polarization. These two components will also have different critical wavelengths; i.e., wavelengths at which the medium changes from reflective to transmissive. Careful measurements of these effects might be used to determine the magnitude of the planetary magnetic field.

By using several wavelengths shorter and longer than the critical value, the height of the ionospheric layer above the solid surface might be determined. Also, by noting the change in delay for several wavelengths longer than the critical value, it might be feasible to determine the manner in which ionization density varies with height above the maximum density of the layer.

If the magnetic field is known, it might be possible to determine the total electron content (through a vertical column) of the ionosphere by noting the effects of the ionosphere on the polarization of the waves (Faraday effect). We noted previously that the total electron content along the Earth-planet path could be determined by measuring the difference in delay time of pulses at different wavelengths.

If the planetary ionosphere differs very greatly from that of the Earth, some of the above discussion might not be applicable. For example, in the terrestrial ionosphere, wavelengths near the critical value are only slightly absorbed due to collisions with neutral atoms at low levels (D-region), with essentially no losses in the layer of reflection (F-region). A preliminary suggestion has been made that perhaps the ionosphere of Venus is sufficiently dense to reflect the 70-cm wavelength used in the radar echo experiments. This is an attractive hypothesis at first glance because it could explain the unity reflection coefficient obtained in 1958, and the complete lack of echoes (by Lincoln Laboratory) in 1959. It might also explain the 600° K temperature measured by radio astronomy techniques at wavelengths of 3 and 9 cm, and the lower temperatures measured at millimeter wavelengths, since such an ionosphere would be optically thick at the longer wavelengths. Thus, the high temperature would apply to the ionosphere, and the lower temperatures to the surface. However, I cannot conceive of a reasonable model for the variation of density with height which would not cause complete absorption of the 70-cm radar wavelength above the level of reflection. As far as the radar results are concerned, I believe we should await more experimental results before trying too hard to explain some apparent difficulties in the scanty information now at hand.

4. *Cloud Cover*

A suggestion was made above of how radar might be used indirectly to help determine the height of the cloud bank on Venus (*viz.*, use radar to determine the radius of the solid body, and visual results

to determine the radius to the top of the clouds). As far as a discussion of direct radar studies is concerned, we are again hampered by our incomplete knowledge of the characteristics of the atmosphere and clouds. However, several possibilities come to mind. The dielectric discontinuity at the top of the cloud bank need not be very great to be detected by the new radar systems that are being built. If it can be detected, then the height can be measured directly. It might also be feasible to use radars at wavelengths that would be absorbed due to the constituents in the cloud or terrestrial atmosphere. By sweeping in wavelength through this critical value, the presence of the absorption band would be detected. Of course there may be problems here of separating effects of terrestrial and planetary atmospheres.

Appendix 5

OBSERVATIONS WITH SATELLITE-SUBSTITUTE VEHICLES

JOHN STRONG

A. Introduction

In order to be effective, one must not only evaluate astronomical projects concerned with planetary atmospheres in terms of the possibility of getting necessary support for them; but, granting adequate support, one must evaluate the possibilities of getting results that are significant enough to be rewarding.

There have been two contrasting approaches to the planetary problem: in one, a theoretically-inclined astronomer invents hypotheses that fit the already assembled observed facts; in the other, the experimentally-inclined astronomer strives to devise new ways to get new facts, as opportunity affords. This report considers opportunities to study planetary atmospheres that are afforded by recent developments of heavy-load-carrying, high-altitude balloons and of high-altitude aircraft, such as the Lockheed U-2. Also, we consider recent experimental developments that will facilitate taking advantage of those opportunities, such as improved star-trackers, electronic-controlled servomechanisms, and new radiation sensors.

As an example of the importance of experimental developments: the process of silvering glass, that made possible substantial extension of telescope apertures, first to the 36-inch Crossley reflector-telescope; then the 60-inch mirror-telescope of Mt. Wilson Observatory; and later the 100-inch mirror-telescope. Then the process of aluminizing made these, and the still later 200-inch apertures, effective in the ultraviolet to the limit of transmission of the atmosphere (about 3000 Å). Many new and significant projects were undertaken and finished as a result of the opportunities that these large mirror-telescopes have afforded.

Here we want to review the recent opportunities for astronomical and meteorological observations that result because balloons and high-altitude aircraft have the capability to transcend enough of the infrared atmospheric absorption to open up, almost completely unsullied, the spectral region from the red to millimeter waves. This broad new infrared region is to be compared with the narrow, though important, ultraviolet extension that resulted from the introduction of aluminizing.

Further plans to exploit the freedom from infrared atmospheric absorption at high altitudes will, for us at Johns Hopkins University, reflect prior involvements. An early involvement in balloon astronomy was a project to determine the amount of water vapor in the Mars atmosphere, stimulated by the existence of the Office of Naval Research "Stratolab" facility. This involvement in a manually-controlled telescope system for acquisition and rough tracking (rather than an unmanned automatic one) was predicated by the deadline to observe Mars at the opposition of the fall of 1958. In addition, shortage of funds precluded the development of a fully automatic telescope system. Preparations for this deadline, to observe Mars, also constituted a commitment for the Venus project, and led to the Moore-Ross flight of November, 1959.

Now this commitment to manned flight is broken. If we plan further manned flights it will be because they are intrinsically attractive. Instead of manual control of acquisition and rough tracking, we will use the Sun for automatic acquisition and the rough tracking of the balloon telescope.

And, furthermore, a balloon "likes" to be launched in the morning, and to come down at sunset. The daytime sky is dark enough at 80,000 ft. to make astronomy, at least of the planets in the infrared, a daytime profession (where at the surface it was formerly a nighttime profession).

Our future plans to make astronomical observations above the Earth's absorbing atmosphere, from what we may call a border-of-space station, will reflect our peculiar astronomical interests. In addition to observation of planets we are interested in a study of the spectrum of the Sun and of the brighter stars; and indeed, some progress has been made in this direction already (Johns Hopkins University, 1960).

Here we will largely ignore, at least for the present, the fact that recently available altitudes have another advantage: they transcend the optical blur due to atmospheric turbulence. Our plans, that do not take advantage of this fact, do not mean that we decry its importance; rather, we recognize that it is being effectively exploited by others.

Aside from our intrinsic interest in infrared spectra, we are planning to exploit improved atmospheric transmission rather than improved resolving power, not only because others are concerned with the latter; but also because we regard it as more difficult to exploit the high altitude freedom from optical turbulence, which sets the limit on astronomical image definition at grounded observatories, than to exploit the high altitude freedom from infrared absorption. Heretofore, water and CO_2 absorption has sullied or completely precluded many observations from grounded observatories that one would like to make, and that can be made now from the border-of-space. In contrast, in order significantly to transcend the performance that astronomers have already achieved, and thus exploit the absence of optical blur due to turbulence, a high-altitude telescope of at least 12-inch aperture is required in daytime, or a 36-inch aperture at night, worked in each case to the diffraction limit of resolving power. This will be very difficult to do, judging from the fact that telescopes are difficult enough to make work at the diffraction limit on the ground. In respect to image quality, programs with high-altitude telescopes can only yield qualitatively superior images. In contrast, when the telescope serves as a starlight-gathering or spectrometer-illuminating device, then a 12-inch aperture for stars and planets, and a 3-inch aperture for the Sun, can yield quantitatively superior spectroscopic results—results that are completely new and utterly unattainable with the most powerful ground observatories ever conceived.

In addition to new astronomical results, programs with satellite-substitute vehicles at the border-of-space altitudes will test observing systems that will serve as precursors to systems designed to go even deeper into space; for, indeed, these border-of-space altitudes provide a hostile environment, and this environment will certainly be used extensively, perhaps in perpetuity, for exploitation of the opportunities that it affords to make precursor studies.

B. Planetary Mysteries

Our interest in the planets stems from a desire to clear up the mysteries which recent and old observations have created. Some of these are outlined below.

Sinton and Strong observed the emission by Venus in the 8-13μ wavelength interval, where the Earth's atmosphere is transparent down to the surface of the land or sea, or to the cloud tops. They have found that local Venus temperatures, as inferred from that emission, did not vary from the subsolar point to its antipode by as much as one would expect. (v. Chapter IV.B.3.) And further, they have found that the spectrum of the Venus emission, over that wavelength range, was smoother than one would observe looking back from space at the Earth's atmosphere. This lack of temperature variation creates a mystery. Because Venus has a slow apparent rotation rate one would expect a greater temperature variation than the 5C° that Sinton and Strong observed. And further, because

of the large CO_2 content of the Venus amtosphere that Dunham and Adams discovered, one would expect a less smooth spectrum than Sinton and Strong found.

Kaplan's (1960) attempt to clarify this combination of facts hypothesizes a cloud of dust that has the peculiar property of being opaque from 8-13μ, where Sinton and Strong found the temperature to be 235°K; and transparent in the near infrared, where Chamberlain and Kuiper found the CO_2 temperature to be $T = 285°K$. The reflection for the near-infrared radiations of Chamberlain and Kuiper occurs, according to this model, at the 300 mb level in the Venus atmosphere; whereas the opaque layer lies at the 100 mb level in the Venus atmosphere, (v., however, Appendix 9 by Chamberlain, and Chapter IV.B.3.) Above the 300 mb layer, the mixing ratio of CO_2 must be 0.2, according to Kaplan.

Another mystery is created by the radio-telescope brightness temperatures of Venus. These observations give 600°K at centimeter wavelengths, and this temperature drops, around 8-mm wavelength, to 300 or 400°K. (v. Appendix 3 by Burke, and Chapter IV.) These high temperatures cannot be easily explained as due to ionospheric emission; and they are almost unacceptable as a planet surface temperature. Although Sagan and Öpik have wrestled to mitigate these mysteries we still cannot be comfortable living with them.

Finally, our knowledge of differences and similarities between Earth and the neighbor planets is not yet extensive enough to give good clues to an acceptable planetesimal hypothesis, and a hypothesis of planetary evolution.

Although, to us, these mysteries of Venus are more provocative than those that relate to Mars, we need new facts about Mars too. For example, we would like to make our knowledge of the pressure and composition of the Mars atmosphere firmer; to understand the nature of atmospheric circulation on Mars; to understand the nature of the blue clearings in the Mars atmosphere (variously attributed to ice crystals, and to carbon particles); and to determine if the observed color changes on the Mars surface are evidence of life.

Abelson (1960) in a recent paper about life on the planets, made the all-but-categorical statement that no life is possible. His strong position leads us to make our own position clear here. We do not *believe* that either Mars or Venus contains life "as we know it"—and for the very same reasons that were advanced by Abelson. And yet, turning the matter around, and because our knowledge of the planets is still so imperfect, we cannot bring ourselves to *believe* that these planets are *not* inhabited by some sort of life. We prefer to establish for ourselves a position of agnosticism. To illustrate our position, let us imagine a race of man, like us, on a planet somewhere, like ours, who might not yet know that fishes or birds exist. In such an instance, if life in water, or in air, was proposed to them we would expect them to think that the fishes might drown, and that the birds might flop rather than fly—because neither water nor air could support "life as they know it."

C. Infrared Transmission from the Border-of-Space

Fig. 1 shows the atmospheric absorption of the solar spectrum at the surface of the Earth, at 35,000 feet altitude, and at 65,000 feet. These spectra were taken in the Fall of 1960 in Alaska with the help of Lockheed's U-2 aeroplane. It is apparent that, in the last spectrum, only ozone absorption remains in substantial amount. And at 80,000 feet we would expect to be above half the residual H_2O and CO_2 that the last spectrum shows. A spectrum envelope, constructed in a manner that is described below, when fitted to the observations, gives the incident solar illumination necessary to determine absorption by water vapor, and by other gases. Using this absorption, one can assay the

amount of water vapor overhead. It will be expressed below in microns (μ) of precipitable water. Dr. Benedict has made the following estimates of the overhead water from the measured absorption:

Table I

Altitude	35,000 ft. 10.7 km	65,000 ft. 19.8 km
Total Pressure	240 mb	60 mb
Water Overhead	23μ ppt H_2O	13μ ppt H_2O
Mixing Ratio	0.9×10^{-5}	2.6×10^{-5}

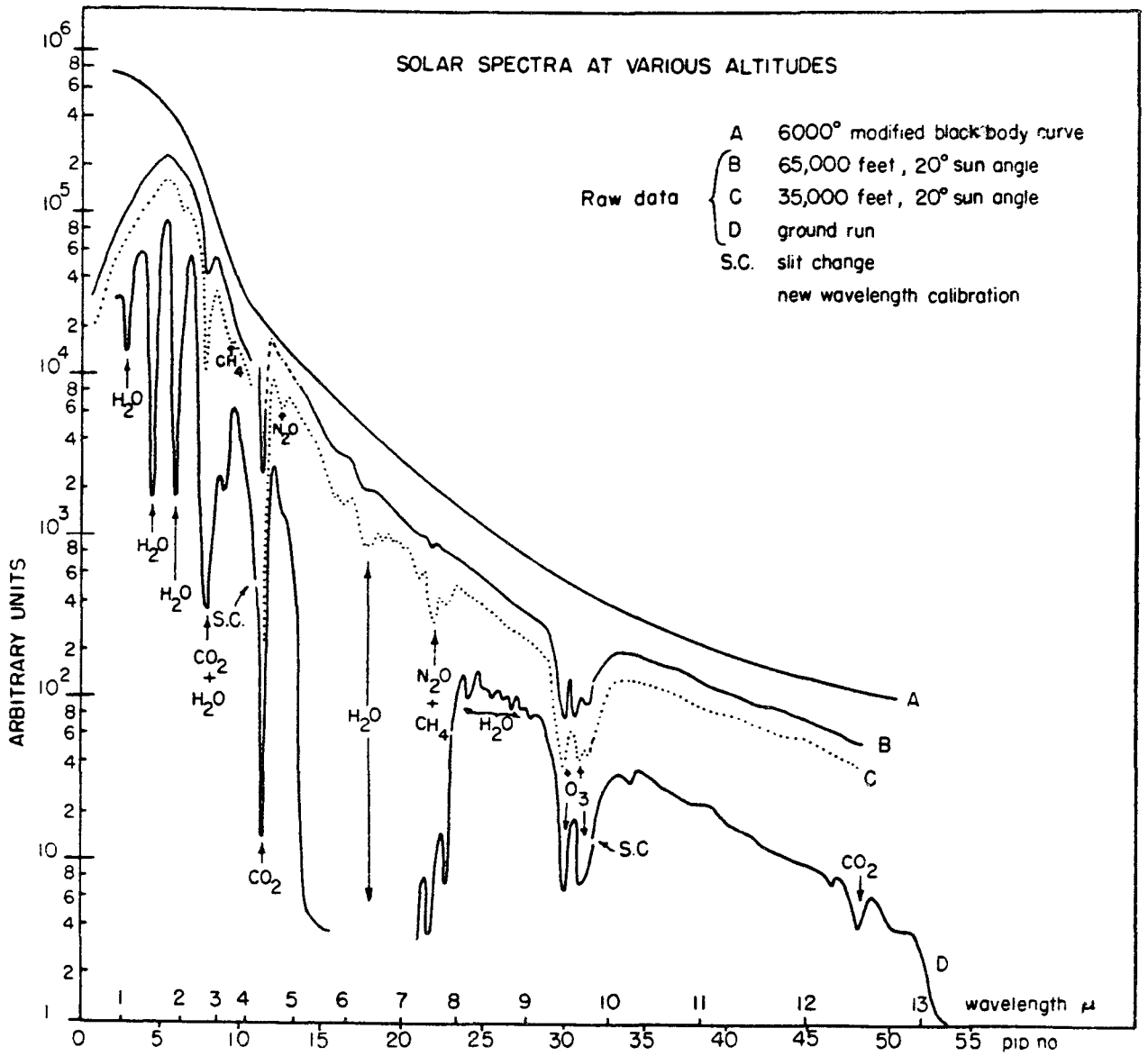


Figure 1

Fig. 2 shows the spectrum envelope referred to above. We constructed the envelope based on the following observed brightness temperatures for the Sun:

Table II

T°K	*λ in μ*	*Authority*
6330°	1.55	(Peytureux)
6045°	2.03	"
5880°	2.31	"
5036°	11.1	(Goody)

Fig. 3 illustrates how the constructed composite spectrum, based on blackbody spectra for the observed temperatures about the observed wavelengths, conforms to our observed spectra, except where we have residual atmospheric absorption.

The infrared transmission around 6μ is estimated for an observing station at a height of 80,000 ft

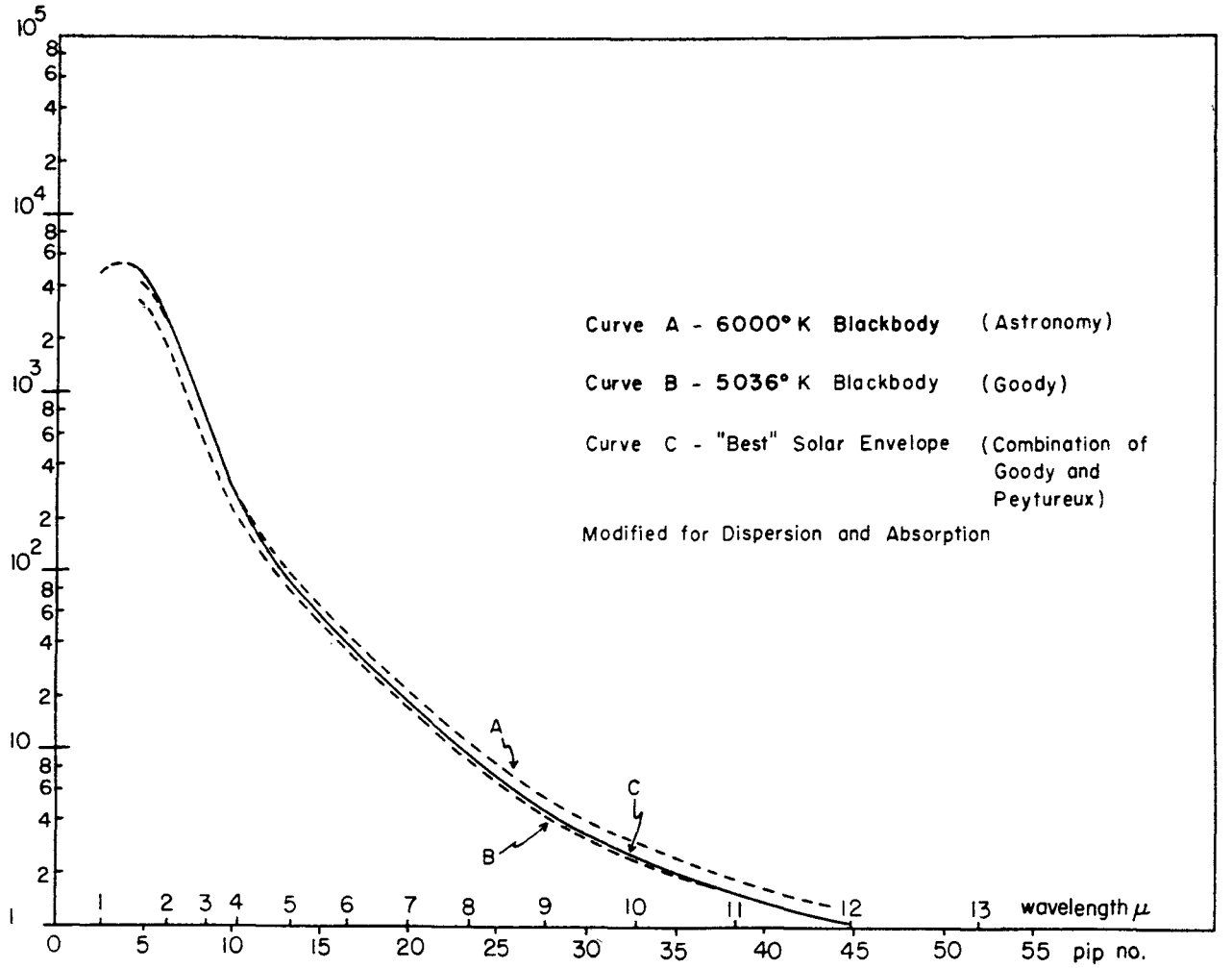


Figure 2

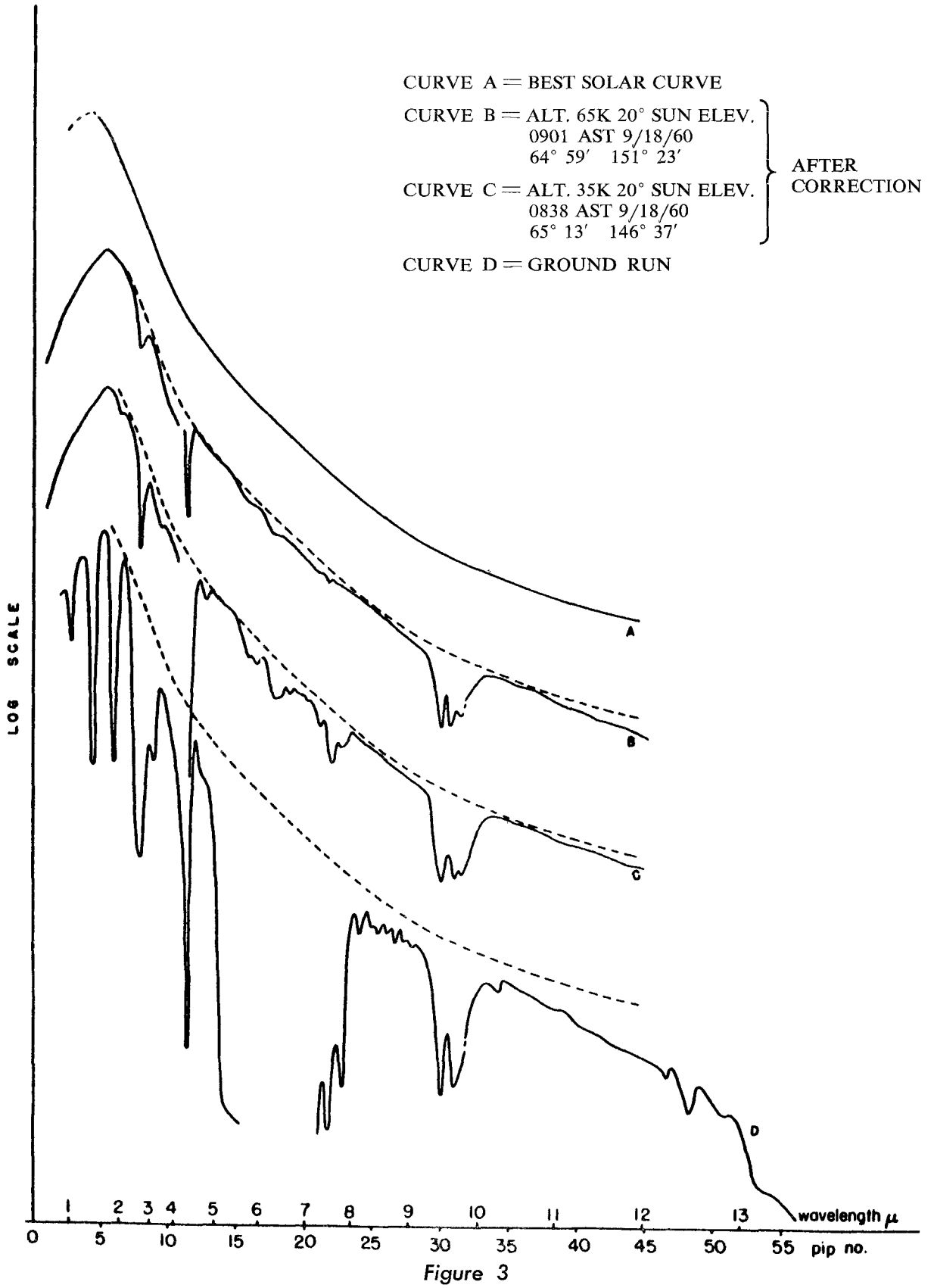

CURVE A = BEST SOLAR CURVE
CURVE B = ALT. 65K 20° SUN ELEV.
0901 AST 9/18/60
64° 59′ 151° 23′
CURVE C = ALT. 35K 20° SUN ELEV.
0838 AST 9/18/60
65° 13′ 146° 37′
AFTER CORRECTION
CURVE D = GROUND RUN
LOG SCALE
A
B
C
D
1 2 3 4 5 6 7 8 9 10 11 12 13 wavelength μ
0 5 10 15 20 25 30 35 40 45 50 55 pip no.

Figure 3

(24 km; 29 mb) to be that due to $13 \times \frac{29}{60} = 7.5\mu$ of water above .029 atmospheres, which is equivalent in absorption to $7.5 \times \frac{1}{2} \times \sqrt{.029} = 0.6\mu$ of water vapor at the surface pressure, after correcting for pressure broadening. Thus the absorption around 6μ wavelength above an altitude of 80,000 feet will be approximately the same as that due to the water vapor in 2 inches of room air. Also, infrared absorption by residual CO_2 will be approximately the same as that due to 66 feet of room air.

The advances in astronomy that will result from infrared observations contrast with those that will come from transcending the optical blur due to atmospheric turbulence much as the advances in astronomy due to the application of the spectroscope and spectrograph, beginning with Wollaston and Fraunhofer, contrasted with the advances due to the improved angular resolution that telescopes brought, beginning with Galileo's observations.

And balloons and high-altitude aircraft promise to add the astronomical information that the infrared wavelengths carry, just as rockets have added that which the ultraviolet wavelengths down to X rays carry; and as radio telescopes have added that which the millimeter and centimeter rays carry.

D. Vehicle Characteristics Contrasted

Rockets are suitable for making studies with ultraviolet radiations where they would not serve for studies with infrared radiations. This is because the short time available for observation that rockets provide at very high altitude above ultraviolet atmospheric absorption is not too restrictive at the shorter wavelengths. Ultraviolet radiations are actinic, so that photography and detectors work quickly at peak altitudes.

In contrast, balloons cannot go to such high altitudes, but they do go high enough to transcend infrared atmospheric absorption. Also, they are suited to carry the heavier instruments for the longer observation times that the less actinic infrared radiations require. Finally, balloons, as vehicles, provide greater motional stability of the telescope than rockets.

If one astronomical telescope has the purpose to achieve critical angular resolving power, above the atmospheric optical turbulence, as contrasted to another telescope which has the purpose only to collect light above the infrared atmospheric absorption for filling the slit and collimator of a spectrometer, then a balloon will be indicated as the proper vehicle for the angular resolving power, while a high altitude aircraft might serve quite as well for the spectroscopic resolving power.

The aeroplane has the advantage that it can launched on schedule. In contrast, balloon launching, balloon flight, and balloon landing is not yet reliable.

In addition to the danger of damaging the balloon or its payload of equipment at launch, there is, even after successful launch, the danger of balloon failure, due to shear winds at the tropopause. This danger, fortunately, can be greatly mitigated when the new balloon materials that are now promised become available. We refer here to Schjeldahl's mylar-with-dacron-scrim balloons. They are, for the same balloon weight, seventeen times as strong as the older polyethylene ones.

High priority must be paid to getting the balloon to altitude in a manner that will meet the requirements of a mission. In particular the balloon must be manageable so that its vertical velocity, relative to the local air, can be caused to vanish.

And finally, balloon landing is currently a matter of luck. This is because ballast facilities fail inexcusably; no provision is made to provide the balloon pilot with an adequate periscope view from within the gondola, as it lands; and even after a safe landing, the peril of being dragged by the parachute remains.

The most urgent requirement for successful balloon astronomy is to make the launching, the flying, and the landing, with heavy loads, all routine and predictable.

Balloon problems have previously been coped with by optimism and expected luck. They can, however, be managed by good engineering practices, based on experience. And then, balloons promise to be reliable vehicles for carrying equipment capable of pointing control to seconds, in contrast to aeroplanes, where vibrations and the air stream preclude this.

Balloons promise further to be more suitable for carrying a heavy large telescope, say one of 50-inch diameter.

E. Programs for the 12-Inch Balloon-Telescope

The Johns Hopkins 12-inch balloon-telescope was provided through ONR-NSF support. In November, 1958, it was mounted and put in operational adjustment in the ONR Stratolab gondola, and acquisition and tracking of Mars was demonstrated. It tracked the planet to five seconds of arc when the gondola was manned and swinging from an 80-foot crane, as it would be on the balloon. The spectroscopic equipment was also shown to yield a signal-to-noise response of 5:1, when worked with the radiations around 1.1287μ that the telescope collected from Mars. This response was achieved with a spectroscopic resolving power of $\Delta \nu < 2\text{cm}^{-1}$ in that region.

This swinging performance of November, 1958, together with the results of star trail photographs, taken on a July, 1958, flight by Ross and Lewis, indicated the feasibility of the proposed use of this telescope and spectrometer at 80,000-feet altitude for a determination of the absorption, around 1.1287μ, by water vapor in the Martian atmosphere.

Unfortunately, just before launch the balloon failed by rupture. This balloon failure aborted all plans to make a water vapor assay of the Martian atmosphere.

The equipment, after being stored for a year, was used in 1959 to assay the water vapor in the atmosphere of Venus.

One of the first programs that we plan for the future with balloons, is to use our 12-inch telescope to repeat our determination of the water vapor in the atmospheres of Venus.

Also, we plan to use this balloon-telescope on separate subsequent flights to measure the high resolving power spectrum of the solar emission from 1μ to about 40μ.

Still later, we will use this balloon-telescope, looking in the 7,600 Å absorption band of oxygen, to determine the oxygen content of the atmospheres of both Venus and Mars, working at altitudes of 80,000 feet or so. At such altitudes it will be necessary to resort to the Doppler shift to avoid the masking effect of even the 3 per cent of telluric oxygen that remains overhead above 80,000 feet altitude. This Doppler shift is not, however, sufficient to separate the planetary absorption lines from the masking telluric lines when working at ground observatories, because of the great line breadth of the telluric lines that arises from the large amount of the absorbing gas that ground observatories must penetrate. But the Doppler shift is expected to suffice at 80,000-feet altitude.

Applications of our 12-inch telescope, with an interferometric spectrometer, are described in a later section.

The 12-inch-aperture equipment is not only adequate to determine the bolometric magnitudes of many of the brighter stars, but also, with it, we can determine the shape of the envelope of the stellar emissions of the very brightest stars. The bolometric magnitudes and stellar spectra that have been made with the 100-inch telescope are characterized by incompleteness, to the extent that the Earth's atmosphere was opaque in the infrared at some wavelengths, and only partially transparent at others. This atmospheric absorption required that awkward extrapolations be made to take account of attenu-

ation where the atmosphere was not completely transparent; and that awkward interpolations be made to account for unknown stellar radiations where the atmosphere was opaque.

Signal-to-noise ratios for detector response, to one-tenth of the total starlight, for the first few brightest stars of the order of 10 are expected by conventional spectrometry. Here we postulate a 12-inch telescope, a detector of 10^{-11} watts ENI, and we assume a filtering efficiency of ½ for separating the 10 spectral components of an assumed blackbody enevelope. It might appear strange that our 12-inch aperture is adequate, being inferior in light-gathering power to the 100-inch telescope by two orders of magnitude. This is because modern semi-conductor detectors adequately compensate this light-gathering deficiency; they are superior to the thermopiles that were used with the 100-inch telescope by an equal factor of two orders of magnitude (or even more in the case of chilled PbS detectors).

Applications of non-conventional interferometric methods of spectrometery to planets and stars are discussed below.

The performance of our balloon-telescope was made possible because star-trackers for aircraft navigation had been developed to a high stage of sophistication. We took advantage of this development in our new telescope. Our telescope trades the difficult task of tracking a heavy telescope, carried on a balloon gondola, for the easier task of only tracking a light optical relay component of that telescope (Griffith Observer, 1959). With our telescope, when we can accomplish manual rough tracking to ± 5 degrees, and keep tracking rates less than 0.1 degree per second, then the automatic star-tracker will keep the image of a planet like Venus or Mars, or a star, on the slit of the telescope to ± 5 seconds of arc, or better.

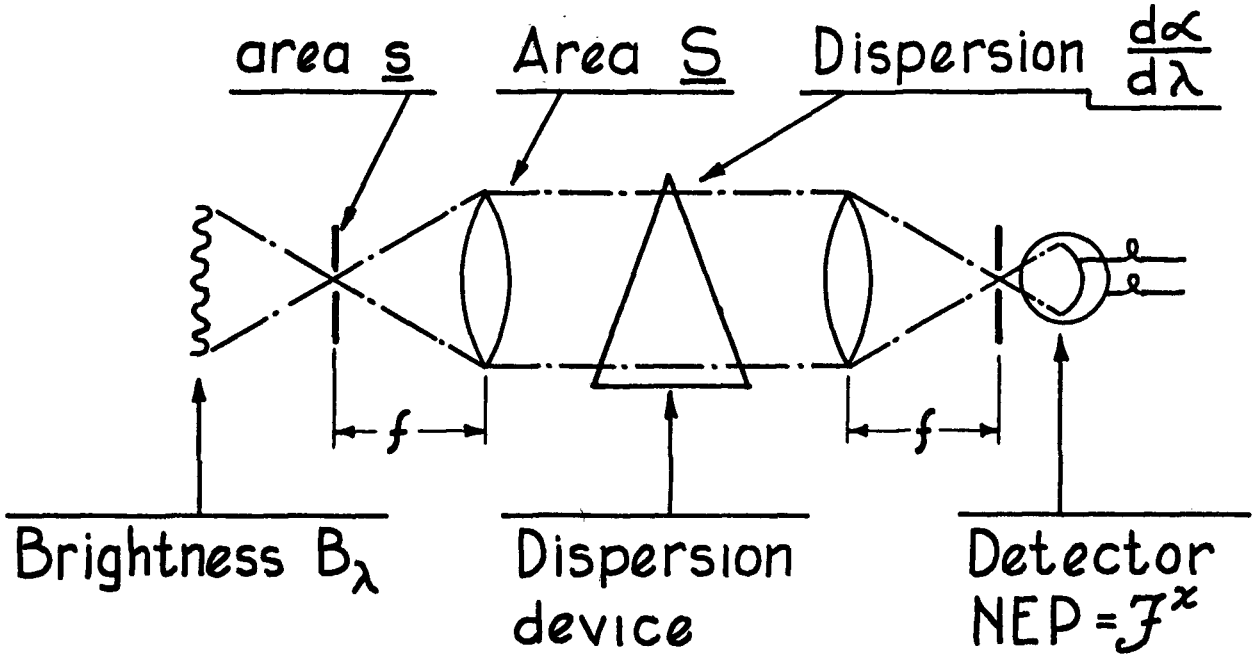
F. Future Programs for a 50-Inch Balloon-Telescope

We have concluded that our 12-inch telescope will not be adequate to measure the variation of temperature over the planetary disk because of its lack of light-gathering power. However, based on the fact that Sinton and this author were able to obtain a spectral resolving power of about 0.1μ at 10μ with the 200-inch telescope (and assuming that the resolving power goes inversely as the square of the diameter), one could achieve a resolving power of about 2μ with a 12-inch telescope on a balloon. The infrared emission spectrum of Venus with 2μ resolving power, free of nearly all atmospheric distortion, would be an exciting thing to see, because one could not only determine what the gases were in the upper atmosphere of Venus, but one could determine something about the thermal structure as well. (v. also Appendix 7, by Kaplan.) However, with a 50-inch telescope, one can not only measure the distributions of planetary temperature, using total infrared emission and no more than 2 per cent of the planetary disk; but using interferometeric methods, one can measure the distribution of spectral components over the planetary disk.

G. Energy Economics of Spectroscopy in the Laboratory and Observatory

Consider a spectrometer with resolving power limited by energy deficiency in its band pass, rather than by aberration or by diffraction, such as the spectrometer shown diagrammatically in Fig. 4. Consider it used in the laboratory, with a source of spectral brightness B_λ; an entrance slit of width e and length l that is filled by radiation from that source; a collimator of focal length f. Let the effective collimator area be S, subtending a solid angle $\Omega = S/f^2$ at the detector, and consider it to be filled with radiation from the source. The spectrometer is assumed to be provided with a dispersive means characterized by a dispersion of $\frac{d\alpha}{d\lambda}$; a telescope mirror (or lens) of focal length equal to the collimator, and subtending an equal solid angle Ω; and an exit slit, also of length l and width $x = e$.

Let ε be the over-all efficiency of this spectrometer, including aggregate transmissions of lenses, filters, and prism faces, reflections of mirrors, and including absorption in prisms—and, in the case



$e \cdot l = s$ $\quad\quad$ τ = transmission

$$\mathcal{F}^x = \tau B_\lambda \Omega s \Delta\lambda$$

$$\frac{e}{f} = \frac{d\alpha}{d\lambda}\Delta\lambda \qquad \Omega = \frac{S}{f^2}$$

In the laboratory

$$\Delta\lambda = \sqrt{\frac{\mathcal{F}^x}{B_\lambda S \tau \frac{d\alpha}{d\lambda}\frac{l}{f}}}$$

Figure 4

of a grating, including the blaze factor. If F is the flux of power necessary to operate a detector with a workable signal-to-noise response, we equate this F to the flux passed by the spectrometer to get,

$$F = \varepsilon B_\lambda \Omega e l \Delta\lambda$$

And, using $\Omega = S/f^2$, and $e = f\dfrac{d\alpha}{d\lambda}\Delta\lambda$, solving for $\Delta\lambda$, we get

$$\Delta\lambda = \sqrt{\frac{F}{B_\lambda S \dfrac{d\alpha}{d\lambda}\varepsilon\dfrac{l}{f}}}$$

Here $\Delta\lambda$ is the energy-limited band pass that will just give the signal-to-noise detector response that is required.

The experimenter who wishes, for example, to decrease $\Delta\lambda$ and get better resolving power, will do as we have here at Johns Hopkins University: He will use the least F that is practical by getting the most responsive detector available. He will increase B_λ, for example, by use of a carbon arc, which is the brightest source of continuous infrared radiation in the laboratory. He will use a grating of maximum areas S—the grating used at Johns Hopkins University has a ruled area of $12'' \times 14''$. This grating has been given the maximum dispersion $\left(\dfrac{d\alpha}{d\lambda}\right)$ and blazing (ε) possible. We use it in an Ebert spectrometer with maximum ratio of slit length to focal length (l/f).

However, when we move from the laboratory into the observatory, we encounter quite different problems. In the laboratory we can usually fill both a large entrance slit and a large collimator lens or mirror with the radiation that is to be studied; although occasionally it may be difficult in the laboratory, as for example in Raman studies. It is, however, usually difficult to fill a slit when one is studying astronomical objects; and the brightness of the object studied is not under control.

Fig. 5 illustrates the situation in the observatory. Consider a telescope of aperture $2h_o$ and focal length f_o, that images an astronomical object of angular diameter δ on the entrance slit of a spectrometer. We assume the image is circular, of diameter $f_o\,\delta$, and of uniform illumination. One usually wants e large enough to take in all the planetary image; but one also wants e small enough to get satisfactorily high spectral resolving power. It can be shown that the fraction of the image that falls between the jaws of the slit is

$$\varepsilon = \frac{2\psi + \sin 2\psi}{\pi}$$

where the angle ψ is defined, as indicated in Fig. 5, where the rim of the image and the slit jaw of the spectrometer intersect, $\sin\psi = \dfrac{e}{f_o\delta}$. Writing $f_c\dfrac{d\alpha}{d\lambda}\Delta\lambda$ for e; and using $\dfrac{h_c}{h_o} = \dfrac{f_c}{f_o}$, by similar triangles, we get the useful relation,

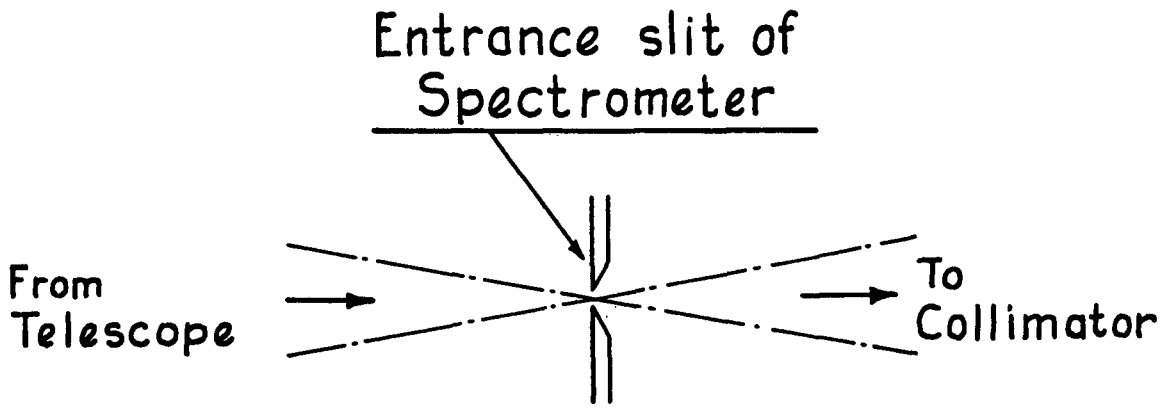
$$\sin\psi = \frac{h_c}{h_o\delta}\frac{d\alpha}{d\lambda}\Delta\lambda.$$

Now the value of δ is predetermined—either at about 1 sec or so, for a star, as set by "seeing"; or at some larger value, say 22 sec, as set by the diameter of Venus. And further, h_o is predetermined by the telescope that one chooses to work with. Thus, after we set an acceptable value on ε such as $\varepsilon = 0.8$ (when $\tan\psi \cong \dfrac{1}{3}$) or $\varepsilon = 1$ (when $\psi = \dfrac{\pi}{2}$), we get

$$h_c\frac{d\alpha}{d\lambda}\Delta\lambda = \text{constant}.$$

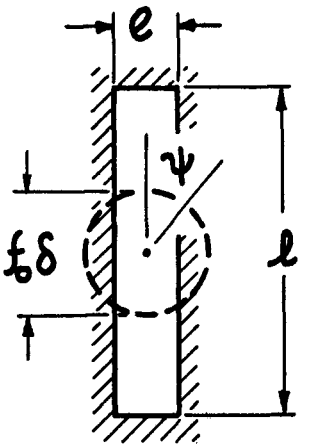
The expression above allows one to make knowledgeable compromises to guide instrument design. For example, when the dispersion $\frac{d\alpha}{d\lambda}$ is made higher, one may use a more modest-sized spectrometer (small h_c) and yet get a satisfactorily small $\Delta\lambda$. In our spectrometer for the Moore-Ross Venus flight, $\Delta\lambda$ was $\cong$ 2 Å around 1.1287μ. We worked as follows:

$$\delta_o f_o = 0.9 \text{ mm}$$
$$e = 0.6 \text{ mm}$$
$$\varepsilon = 0.8$$
$$h_c = 3''$$



ϵ = efficiency

$$= \frac{2\psi + \sin 2\psi}{\pi}$$



$$\sin\psi = \frac{e}{f_o\delta}$$

$$= \frac{f_c \frac{d\alpha}{d\lambda}\Delta\lambda}{f_o\delta}$$

$$= \left(h_c \frac{d\alpha}{d\lambda}\Delta\lambda\right) \Big/ h_o\delta \quad \text{since;} \quad \frac{h_o}{f_o} = \frac{h_c}{f_c}$$

In the observatory (given δ; h_o; ϵ or $\sin\psi$)

$$h_c \frac{d\alpha}{d\lambda} \times \Delta\lambda = \text{constant}$$

Figure 5

The grating surface was 5.6″ × 10″. At 56° angle of incidence this projected to 5.6″ × 10 cos 56° = 5.6″ × 5.6″, which was circumscribed by the beam from a 6″-diameter collimator.

A Fabry-Perot type interferometric spectrometer, worked in the central interference spot, has shown the following advantage over a grating spectrometer (Greenler, 1955): For equals F's:

$$\frac{\Delta\nu_g}{\Delta\nu_i} = 13.5\sqrt{\frac{\varepsilon_i S_i}{\varepsilon_g S_g}}$$

Here the subscript g refers to the grating spectrometer while i refers to the interferometer. Following Greenler, ε_i for an interferometer using evaporated tellurium films will be comparable with ε_g for a blazed grating. Thus, the interferometer has an order of magnitude advantage, as was predicted by Rupert (1952). This advantage arises from the large circular entrance aperture that is tolerable in a Fabry-Perot spectrometer, without loss of resolving power. This is due, in turn, to the infinite dispersion at the center of a Fabry-Perot pattern—and to the large dispersion adjacent, throughout the central spot. The advantages of Fabry-Perot interferometry arise from the small size of equipment to give a satisfactory $\Delta\lambda$.

New methods of interferometry have recently been developed (Vanasse and Strong, 1958, and later papers). In application of the new methods, many wavelengths are observed simultaneously. Each spectral component of the radiation being measured is split into two beams and recombined, after a variable path difference common to all components has been interposed. The record of the intensity that represents the recombined amplitudes of all components, as a function of the interposed path difference, is called an interferogram. The interferogram gives the sum of spectral component intensities at the detector, each being determined by the superposition of amplitudes. Later, the individual strengths of these components are determined from the Fourier transform of the interferogram.

This method of spectral investigation has been fully demonstrated in this laboratory. The interferometric method has, in addition to the advantage of a large entrance aperture, a further advantage of long recording time. And further, the interferometer has a high efficiency, ε, giving good signal. As far as noise is concerned, the long recording time gives integration or read-out times that are one or two orders of magnitude longer than those used in conventional spectrometry. The total scan time gives noise proportional to the square root of the integration time: $N \sim \sqrt{t_o}$. Here t_o may be a hundred times longer than in a conventionally dispersive spectrometer, and in such a case the detector noise would be $\sqrt{100}$ or 10 times less.

These interferometric methods, like the Fabry-Perot interferometric methods, have applications because the astronomical image diameters ($f_o\,\delta$) are usually small, and so small light-weight spectrometers can be used.

H. Detector Response and Energy Budgets of Planets

Table III shows the wavelength limits that divide blackbody emission, for an assumed planetary temperature of 234°K, into ten equal-energy portions, or tithes.

We may use the following formula for thermistor detector sensitivity to estimate the detector's equivalent noise input (*ENI*) in watts:

$$F_{ENI} = 8\sqrt{\frac{s}{\tau\, t_o}} \times 10^{-11} \text{ watts.}$$

Here we take s, the area of the detector, as $\frac{1}{100}$ cm² (for a 1 mm × 1 mm detector); we take τ, the

Table III

234°K Blackbody Radiation

Tithes	*Band λ's in μ*	Δλ	*Remarks*
1st	0 — 8.9	8.9μ	N-band
2nd	8.9—11.9	2.0	CaF_2 Filter
3rd	11.9—13.3	1.4	$\lambda < 13\mu$
4th	13.3—15.3	2.0	
5th	15.3—17.5	2.2	M-band
6th	17.5—20.3	2.8	*MgO* reflection
7th	20.3—23.9	3.6	$14\mu < \lambda < 26\mu$
8th	23.9—29.4	6.5	F-band
9th	29.4—40.0	10.6	CaF_2 Reflection
10th	40.0—∞	∞	$\lambda > 26\mu$

detector time constant, as $\frac{1}{100}$ sec. If we take t_o, the signal read-out time, as 1 sec, we get

$$F_{ENI} = 8 \times 10^{-11} \text{ watts} \cong 10^{-10} \text{ watts for 1 sec. read-out time.}$$

The 50-inch telescope that we referred to above will collect the following approximate tithe fluxes from Venus, or Mars:

$$F' = \frac{1}{10} F = \frac{1}{10} \left(\frac{5.67(0.234)^4}{\pi} \right) \left(\frac{\pi (10^{-4})^2}{4} \right) (4000\,\pi) \text{ watts}$$

The first factor in the expression above is the tithe factor. The first parenthesis is the total spectrum brightness, in $\frac{\text{watts}}{\Omega \text{cm}^2}$, being $\frac{\sigma T^4}{\pi}$, where $\sigma = 5.67 \times 10^{-12}$ watts $\text{cm}^{-2}\ T^{-4}$, and $T = 234°\text{K}$. The second parenthesis is the solid angle, taking $\Omega = \frac{\pi}{4} \Theta^2$, with $\Theta = 10^{-4}$ to represent the angle that Venus or Mars might subtend. The third parenthesis is the area in cm^2 of the 50-inch aperture telescope. Thus

$$F'_{(50'')} = 500\, F_{ENI} \text{ watts}$$

For a 12-inch telescope we would have, counting on a bigger fractional central obstruction, about 20 times less flux, or

$$F'_{(12'')} = 25\, F_{ENI}$$

When we use the methods of interferometric spectroscopy, by averaging over an interferometric scan time of 100 sec, we can plan on gaining a factor of 10 in the reduction of unsystematic errors. We can even plan on gaining another factor of 10 if we can average over three hours of repeated scans. On the other hand, we must expect some attrition of signal, as will be set forth below.

In order to simplify the Fourier transformation of interferograms, we propose separation of the total Venus radiation spectrum, by means of the transmission and reflection filters that are listed in Table III above, into three bands, labeled "N" (for near infrared); "M" (for intermediate infrared); and "F" (for far infrared).

We will record outputs yielded by the three thermistors of the interferometric spectrometers to give three interferograms. In these recorders, the potentiometer current may be monitored by the star-tracker to compensate for erratic guiding. Considering integration time gains, and the three-hour

averaging, the signal-to-noise for tithe bands will be about an order of magnitude better than the 500, or 25, given above respectively for a 50-inch, or 12-inch, telescope.

As far as the spectra obtained with the 12-inch telescope is involved, these calculations indicate that spectra on Venus can give satisfactory rough spectrum envelopes—even using Hulburt's rule that work with much less than 100 *ENI* is next to worthless. Even with only 2 to 4μ resolving power, these spectra should help resolve questions of atmospheric and surface temperatures on the planet as well as composition of the planetary atmosphere.

And, with a 50-inch telescope, using only a fraction of the Mars disk, we expect, later, to be able to measure the distribution of atmospheric temperatures on Mars. Such observations promise to give the wave number for the longitudinal oscillation of the atmospheric circulation of Mars around the planet's axis.

These 12-inch and 50-inch telescope expectations are conservative, for they do not contemplate the well-known gains of thermistor detector response that are available from "immersion" (as under KRS-5, which can give an optical gain of $4\times$ in signal-to-noise at $\lambda > 13\mu$; or under Germanium, which can give an optical gain of $16\times$ for $\lambda < 13\mu$).

Our resolution of the planetary disk is not expected to involve a sampling circle of less than 1 second of arc in diameter. Therefore, it would be advantageous, to help account for observed surface emission (temperature), to have simultaneous photographs in the visible spectrum and monitoring observations of the planet's surface brightness. This points up the need to stimulate greater interest, if not capability, in making surface observations of the planets, whenever weather allows. (v. Chapter V.C and V.D.)

References

Preliminary Report on Solar Observations from a U-2 Observatory. Dec. 22, 1960, Laboratory of Astrophysics & Physical Meteorology, the Johns Hopkins University.

L. Kaplan's remarks at the meeting of the Space Science Board's Ad Hoc Panel on Planetary Atmospheres at California Institute of Technology on Dec. 15-17, 1960.

Philip Abelson's paper at meeting of the American Astronomical Society, New York City, Dec. 27-31, 1960.

"A Manned Observatory in the Stratosphere," *The Griffith Observer, XXIII,* 2 (Jan. 1959).

Greenler, R. G., Interferometry in the infrared, *J. Opt. Soc. Am., 45*: 788 (1955).

Rupert, C. S., "Energy-Limited Resolution with Interferometric Dispersing Elements." From ONR published abstracts of papers presented at *Symposium on Molecular Structure and Spectroscopy* at Ohio State University, June 1952. "Limitations which are inherent in a grating made up of laterally-spaced rulings are avoided in a stratified 'grating-in-depth,' formed by evenly spaced layers of low reflectivity. A similar advantage lies with the Fabry-Perot etalon. Ideally, the energy flux per unit spectral bandpass available with these devices is two orders of magnitude greater than that provided by a laterally ruled grating of the same area. Possibilities and limitations in realizing this gain are considered."

Vanasse, G. A., and J. Strong, Application of Fourier Transformations in Optics: Interferometric Spectroscopy; Appendix F of *Concepts of Classical Optics* by John Strong, W. H. Freeman and Co., San Francisco, 1958. (See also Strong, J., and G. A. Vanasse, *J. Opt. Soc. Am., 49*: 844 (1959); Strong, J., and G. A. Vanasse, *J. Opt. Soc. Am., 50:* 113 (1960).)

Appendix 6

SPACECRAFT EXPERIMENTS ON PLANETARY ATMOSPHERES

R. W. DAVIES, A. R. HIBBS, G. NEUGEBAUER, R. L. NEWBURN

A. Introduction

The accompanying chart shows the dates at which flights to the planets are practical. Plans for experiments to be carried on the first few flights to Venus and Mars are now being assembled.

The structure of the Cytherean atmosphere is an interesting scientific problem in itself and is one of the areas of information required to piece together a history of the solar system. But a knowledge of the atmosphere, e.g., its pressure, temperature profile, and constituents, is also of immediate importance in planning a long-range planetary program. If current estimates of the surface temperature are correct, then designing a spacecraft that can survive a planetary entry will be a very formidable task.

The scientific instrumentation aboard Mariner A has four purposes: (1) the determination of the surface temperature and a rough temperature profile of the Cytherean atmosphere, (2) the identification of some of the constituents of the upper atmosphere, (3) the measurement of the planet's magnetic field and possible radiation belts, and (4) the exploration of the fields and particles of the interplanetary region between the Earth and Venus.

B. Venus Atmospheric Constituents

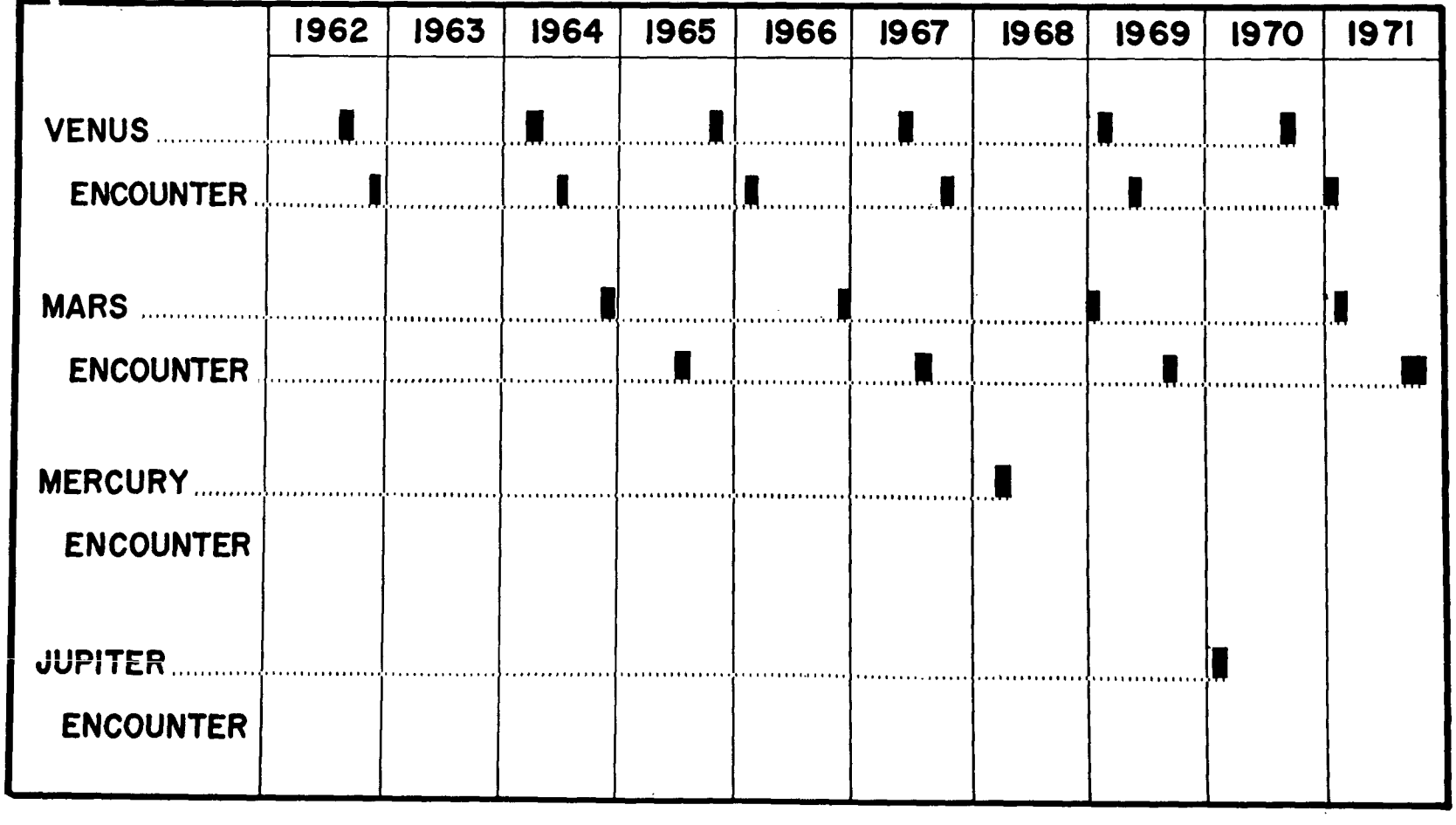
The trajectory of Mariner A,* the first Venus spacecraft, scheduled for flight in the last half of 1962, is a comparatively high-energy path. The planes of the trajectory and of the planetary orbit near Venus will be inclined at perhaps 20°. The approach will be quite close, nominally about 27,000 km from the planet center, which results in a large curvature of the trajectory near Venus. Thus, the instruments on board Mariner A will see both the dark and lighted portions of the planet. The range in phase angle will be from about 145° on approach to 0° as the spacecraft swings around the planet on the sunlit side.

At about 150,000 km, an infrared planet-seeker will locate Venus and center on it. At that time, the microwave radiometer experiment will begin. (See below for a further description of the radiometer system.) A short time later, the exact time being dependent upon the trajectory, the ultraviolet spectrophotometer will begin operation. Both spectrophotometer and radiometer are carried on an articulated head with an axis allowing scans along a line at an angle of 15° to the line through the cusps of Venus. The scans are to be step scans, thus allowing data to be taken from a comparatively localized area rather than integrated over large parts of the disk. A preprogrammed command system controls the actual stepping, the program being designed to give the maximum geographical coverage possible at each point of the trajectory, plus excursions off the planet to study the possible outer envelope of hydrogen (and conceivably other substances) and to calibrate the radiometer.

* As this report was going to press NASA announced that a somewhat smaller spacecraft, designated Mariner R, will be the first to make the Venus attempt in 1962; it will carry the radiometer experiment, a magnetometer, and a number of other instruments for measuring conditions in deep space and in the vicinity of the planet. The more elaborate set of experiments described here will, nevertheless, be flown eventually, and the delay does not detract from its importance.

PLANET AVAILABILITY

QUARTERLY LAUNCH AND ENCOUNTER DATES



Design is fairly well advanced on the spectrophotometer itself. It is a calibrated scanning monochrometer, probably of the Ebert type, fed by a Cassegrain telescope. The wavelength range to be covered is 1100 Å to 4200 Å with a resolution of 10Å. A compact $f/4$ instrument is being built with the area to be viewed on the planet slit-limited to 1/20 radian by 1/200 radian. The detector will be a sodium-salicylate-coated photomultiplier. Logarithmic amplification using a feedback loop to vary the dynode voltages of the photomultiplier will probably give sufficient dynamic range for the spectral interval planned. If necessary, a narrower slit width can be automatically provided for the brighter parts of the spectrum. Every attempt is being made to improve sensitivity by cooling the photomultiplier, by removing those parts of the photocathode not actually used, and by using aluminum with a magnesium fluoride interference film overcoating on all optics. The stray light problem is a very critical one, as attempts will be made to look at the dark side of Venus and reject all light from the lighted side. Absolute intensity and frequency calibrations will be attempted in flight.

The spectrophotometer will scan large parts of both day and night sides of Venus, giving the continuum intensity and absorption features of the former and the emission features (airglow and aurora) of the latter. Below perhaps 2000 Å, emission features should be seen on either side of the planet, whereas scattered sunlight will dominate at longer wavelengths on the sunlit side. The off-planet scans will give data on the outer atmospheric envelope. At least semiquantitative results on the ozone, oxygen, hydrogen, and nitrogen of the upper atmosphere of Venus should result from this study. These additional data, plus those obtained by the radiometer and magnetometer also being carried, should help define the atmosphere of Venus.

C. Venus Temperature Problems

The surface temperature of Venus and the temperature profile of its atmosphere are subjects of considerable speculation and debate. (v. Chapter IV.) The temperature problems are inseparable from the unknowns in the atmospheric constituents, so that speculations on these subjects run parallel.

A set of microwave radiometers on board the Mariner spacecraft will attempt to measure the brightness temperature of the planet at four wavelengths. The location of the Cytherean temperature sources will be found by scanning the disk with pencil beam antenna patterns. The sources in the ionized layers, if present, will be recorded by scanning across the limb of the disk.

The instrument will be pointed at the planet by an external pointing and scanning system. It will begin recording data of the primarily dark planet when it is within 150,000 km. The field of view of the instrument is 2°, which corresponds to a region 5600 km in diameter when the probe is 150,000 km from the planet.

The instrument system consists of four independent radiometers which utilize three parabolic antennas. The crystal audio radiometer is similar to a conventional Dicke type, and the crystal acts as a square-law device. The wavelengths of operation are 4, 8, 13.5, and 19 millimeters. The two longer wavelengths share a single parabolic antenna, whereas the shorter wavelengths utilize separate antennas. The weight of the entire unit is twenty pounds, and the device requires a total of ten watts of power.

D. Mars Experiment

The Mariner B spacecraft will be flown toward Mars at the flight opportunity occurring late in 1964. Present plans for the Mariner B spacecraft are based on present estimates of vehicle performance and current assumptions regarding the availability of money and manpower. The following paragraphs describe the contemplated spacecraft whose design is based on these assumptions.

This spacecraft will consist of a fly-by section (bus) and an instrumented entry capsule. The capsule rides aboard the bus which is sent toward Mars. At 500,000 km from the planet, they are separated, so that the bus can fly by Mars and the capsule can land on the surface. An impulse is

imparted to the bus to establish a miss distance of 8000 km from the surface with a dispersion of 800 km.

The bus serves as a telecommunication-relay station and data-storage system. It will weigh in the neighborhood of 1400 to 1500 pounds, and the capsule will weigh approximately 300 pounds. These weights include from 100 to 180 pounds of scientific instrumentation on the bus and 50 pounds of instrumentation on the capsule, both exclusive of power.

The bus will carry a photographic system with approximately 1-km resolution at closest approach. It will also carry an infrared spectrometer that will examine the surface for the presence of organic molecules as well as investigate the properties of the atmosphere. An ultraviolet spectrometer will search for aurora, airglow, resonance reradiation, and absorption bands in an attempt to identify molecular constituents in the upper atmosphere. All three instruments will require two-dimensional scanning of the planet. In addition, a magnetometer, a particle radiation package, and a cosmic dust package will be flown. The radiation package will probably emphasize the measuring of low-energy particle radiation from the sun and the monitoring of cosmic rays.

The capsule will be covered by a shield of ablative material, which will be used in decelerating the capsule to less than Mach 1. The shield will be jettisoned, a parachute will open, and the capsule will begin its operation as a scientific package.

The capsule will have a one-kilocycle communication bandwidth with the bus, making it possible to take several photographs during the 15-minute descent period. The last of these pictures could have a resolution of ten centimeters with ordinary optics and vidicon tubes.

The capsule will also carry instrumentation for measuring atmospheric temperature and pressure during descent. Measurements of water-vapor content and of atmospheric constituents may also be attempted.

Surface properties and further atmospheric properties will be investigated after impacting on the Martian surface. Emphasis will be placed on biologically important measurements.

All of these operations take place within an hour of the capsule's contact with the Martian atmosphere. After about an hour, the bus goes behind the horizon for two-and-one-half hours. The rotation of Mars brings the bus into view after that time, and communication is resumed at a bandwidth of approximately 100 cycles per second.

One of the major limitations of the system is the ability to store and transmit data. The communication rate from bus to Earth should be in the neighborhood of 100 bits per second.

Some design work has already been completed on infrared spectrometers for the bus portion, designed for the following:

1. To confirm the presence of absorption bands (possibly due to organic substances) observed by Sinton in the 3-4μ region. The probe measurement will have the ability to resolve an area about one hundredth of that covered by Sinton.
2. To look for organic compounds at longer wavelengths (especially in the 4-7μ region).
3. To determine the carbon dioxide and water vapor composition of the atmosphere. It should also be possible to make some determination of the temperature profile of the atmosphere by looking at the shape of the CO_2 and H_2O bands. (v. Appendix 7, by Kaplan.)

There should be no difficulty in performing the first of these; however, the last two depend critically on the amount of water in the atmosphere and the concentration of surface organic matter. The experiments are to a certain degree incompatible.

Two infrared spectrometers were designed for possible use on a 1962 Mars mission and are being extended for the 1964 probe. One instrument is a grating spectrometer using an Ebert configuration, and the other is a Michelson interferometer. Both instruments use lead selenide detectors to cover the 2-5μ region; the detector for the longer wavelengths has not been decided upon, but it is hoped that a cooled detector system can be developed. The pertinent characteristics of both instruments are:

a. Spectral Range: 2-8μ

b. Spectral Resolution: about 300 Å at 4μ

c. Field of View: $2.5° \times 0.25°$ (grating spectrometer),
$1° \times 1°$ (interferometer)

d. Shades of Gray: a minimum of 32

It is hoped to obtain three complete two-dimensional scans of Mars with about forty picture elements to cover the entire planetary surface. In addition to the spectrometer, there will be a crude sensor in the visible to provide some identification of the area being scanned.

Appendix 7

INTERPRETATION OF PLANETARY PROBE MEASUREMENTS

LEWIS D. KAPLAN

A. Introduction

This report is concerned mainly with the interpretation of already-planned measurements, rather than suggestions for other measurements.

The great advantage of a space probe over observations from the surface of the Earth is the gain in signal, which will be of the order of 10^4 even for a fly-by. This will make possible observations that are now impossible and give geographical distribution of atmospheric characteristics that can now just barely be determined for the planet as a whole. Moreover, one would not have to worry about the transmission of radiation by our own atmosphere.

This is not to say that much could not be done from Earth-based observatories, as has been suggested by Kuiper, Edson, Hynek, and others. (v. Chapter V.C.) Much can still be done also on the interpretation of data already taken. Much of this data, unfortunately, is not yet available to the scientific public.

An example of the type of information that can be derived from the small amount of data that is available is given in the recent study of Venus by this author (Kaplan, 1961a). It is a new interpretation of the structure of Venus' atmosphere, and will be assumed to be approximately correct in the following discussion of interpretation of radiation measurements.

B. Infrared Measurements

The infrared measurements from the Mars fly-by (v. Appendix 6) should give valuable information about the atmospheric composition and perhaps the nature of the Sinton bands. It will certainly determine at least whether these bands are due to gaseous absorption or the surface reflective properties. Even small amounts of water vapor should be detectable.

The thermal emission from the 4.3μ CO_2 band, measured with .03 micron resolution, can be interpreted in terms of thermal structure. The principle is the same as that of the sounding of the Earth's atmosphere by measurement of the emission of the 15 micron CO_2 band (Kaplan, 1959, 1961b).

Unfortunately, such measurements are not being planned for the Venus probe, Mariner A. Strong has suggested making such measurements on Venus from a U-2 or balloon with low resolution (v. Appendix 5). The writer believes that this is feasible, and suggests the following wavelength intervals: 8 to 9 microns, 10 to 12 microns, 12 to 14 microns, and 14 to 16 microns. This should cover the range from complete transparency to complete opacity.

The CO_2 concentration in the Venus atmosphere can probably be determined with considerable precision from the measurements of scattered solar radiation in the ultraviolet. The principle of the interpretation is discussed by Singer (1957) and Kaplan (1961b, 1961c), where discussions of the method and its application to the determination of the ozone distribution in our own atmosphere may be found.

The determination of the CO_2 content would probably be made from the measurements in the 2000 to 2500 Å region of the spectrum, but the absorption coefficients there are still unknown. It is important, therefore, that CO_2 absorption at these wavelengths, and their pressure and temperature dependence, be measured in the laboratory.

Below 1600 Å the smallest absorption coefficient of CO_2 is about 1 cm^{-1} in the region of the Lyman-α line. The average emission of Lyman-α is about 4×10^{10} photons cm^{-2} sec^{-1}, and no 100 Å interval below 1600 Å has a larger amount of scattered solar radiation. Even with a CO_2 concentration as low as 10 per cent, the CO_2 reduced path length is about 50 cm per mb. The scattering coefficient at Lyman-α is about 10^{-2} per mb. At .1 mb, therefore, the radiation scattered in the absence of CO_2 is about $4 \times 10^{10} \times 10^{-3} = 4 \times 10^{7}$ photons cm^{-2} sec^{-1}, and the CO_2 transmission is down to e^{-5}. No scattered radiation, therefore, will be detectable at wavelengths less than 1600 Å. The CO_2 absorption coefficient drops off to less than 1 cm^{-1} at about 1730 Å; so the scattered solar radiation should be measurable beginning at some wavelength greater than 1700 Å.

Beyond 1700 Å, the absorption coefficient of O_2 is greater than that of CO_2 by an order of magnitude, and that of H_2O is greater by two orders of magnitude; so appreciable amounts of these gases may be detected. Ozone should be detectable by measurements in the 2500 Å region.

If measurable airglow is expected at wavelengths less than 1700 Å, the above argument implies that there would be no background noise due to scattering of sunlight, even on the daylight side.

There is not enough CO_2 on Mars to use this method for its determination. The method should be applicable, however, to the determination of the ozone distribution and the oxygen content.

References

Kaplan, L. D., Inference of atmosphere structure from remote radiation measurements, *J. Opt. Soc. Am., 49:* 1004, 1959.

Kaplan, L. D., A new interpretation of the structure and CO_2 content of the Venus atmosphere, *Planetary and Space Sci., 8*: 23, 1961a.

Kaplan, L. D., The spectroscope as a tool for atmospheric sounding by satellites, to be published in *Journal of Quantitative Spectroscopy and Radiative Transfer* (May, 1961b).

Kaplan, L. D., On the determination of upper atmosphere composition from satellite measurements, to be published in *Proceedings of Symposium on Chemical Reactions in the Lower and Upper Atmosphere,* Stanford (April, 1961c).

Singer, S. F., A method for the determination of the vertical ozone distribution from a satellite, *J. Geophys. Research, 62:* 299, 1957.

Appendix 8

THE GENERAL CIRCULATION OF PLANETARY ATMOSPHERES

YALE MINTZ

A. Introduction

There are two regimes of large-scale thermal circulation that can develop in a thin layer of fluid which is gravitationally held to a rotating sphere and is heated at the equator of rotation and cooled at the poles. In each of these general circulation regimes the fluid performs two basic functions: (1) it maintains thermal equilibrium by transporting heat from the equatorial heat source to the polar cold source, and (2) it produces the kinetic energy necessary to maintain the circulation against frictional dissipation.

In the *symmetric regime* of general circulation, the fluid, in all longitudes, ascends near the equator of the rotating sphere, flows poleward in the upper levels, descends near the pole and flows equatorward in the lower levels. At the same time the fluid has a zonal component of velocity relative to the sphere, increasing in intensity with height in the direction of rotation of the sphere. As shown schematically in Fig. 1, the meridional and zonal components of the motion are the same in all longitudes (the circulation is symmetric about the pole), and that is why this form of general circulation is called the "symmetric regime."

In the other form of general circulation, called the *wave regime*, the flow is characterized by large amplitude horizontal waves in the middle and upper levels, and by large horizontal eddies (cyclones and anticyclones) in the lower levels of the fluid, as shown schematically in Fig. 2. In each hemisphere, the centers of the cyclones, in the mean, lie closer to pole than the centers of the anticyclones and, as a result, the relative zonal flow at low levels, when averaged over all longitudes, is easterly near the equator and easterly near the pole, but westerly in middle latitudes. In addition, in the wave regime the zonally averaged meridional circulation is reversed in middle latitudes, as shown in the figure.

The symmetric regime of circulation produces a net poleward heat transport only when the vertical lapse rate of temperature is less than the adiabatic rate. Then, as shown in Fig. 1, the northward moving fluid has a higher potential temperature than the southward moving fluid. But regardless of lapse rate, the warmer (and hence lighter) fluid rising near the equator and the colder (and heavier) fluid sinking near the pole converts potential into kinetic energy. When the circulation is in a steady state the individual fluid elements move through the potential temperature surfaces; and with a lapse rate less than the adiabatic and temperature decreasing toward the pole, this requires not only heating of the fluid at the equator and cooling at the poles, but also heating at the lower levels (where the pressure is high) and cooling at the upper levels (where the pressure is low). But the theory of the general circulation is greatly simplified if we are required to specify explicitly only the latitudinal gradient of heating, and can treat the vertical gradient of heating implicity. This can be done by specifying that the vertical lapse rate of temperature is a predetermined constant.

In the wave regime of circulation, unlike the symmetric regime, it is the tongues of warm fluid moving poleward in some longitudes, while at the same elevation tongues of colder fluid move equatorward in other longitudes, which produce the poleward heat transport that balances the differential

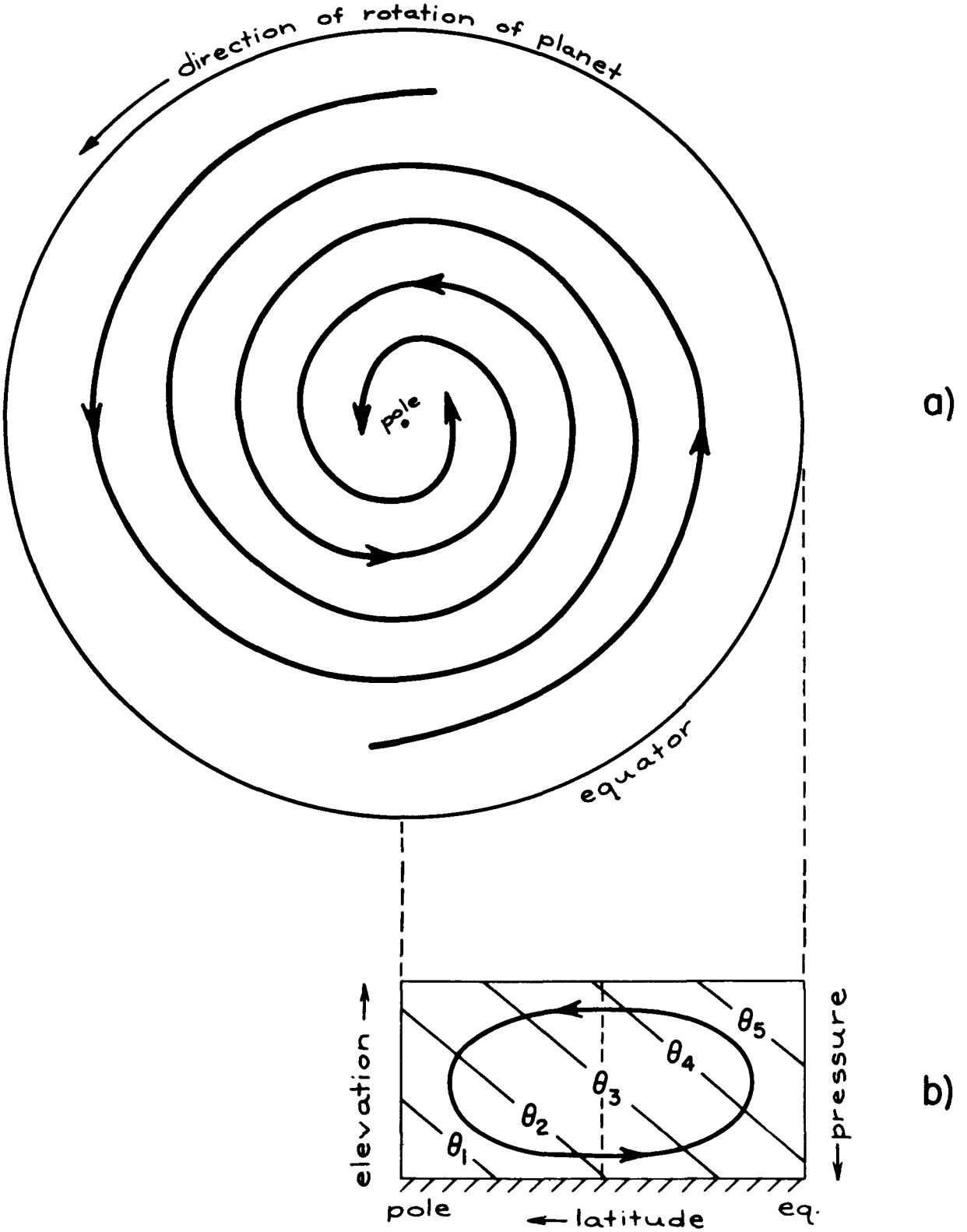


Figure 1. Symmetric Regime of General Circulation

a) Streamlines of the flow at the upper levels.
b) Cross-section showing the meridional projection of the circulation and isotherms of potential temperature θ.

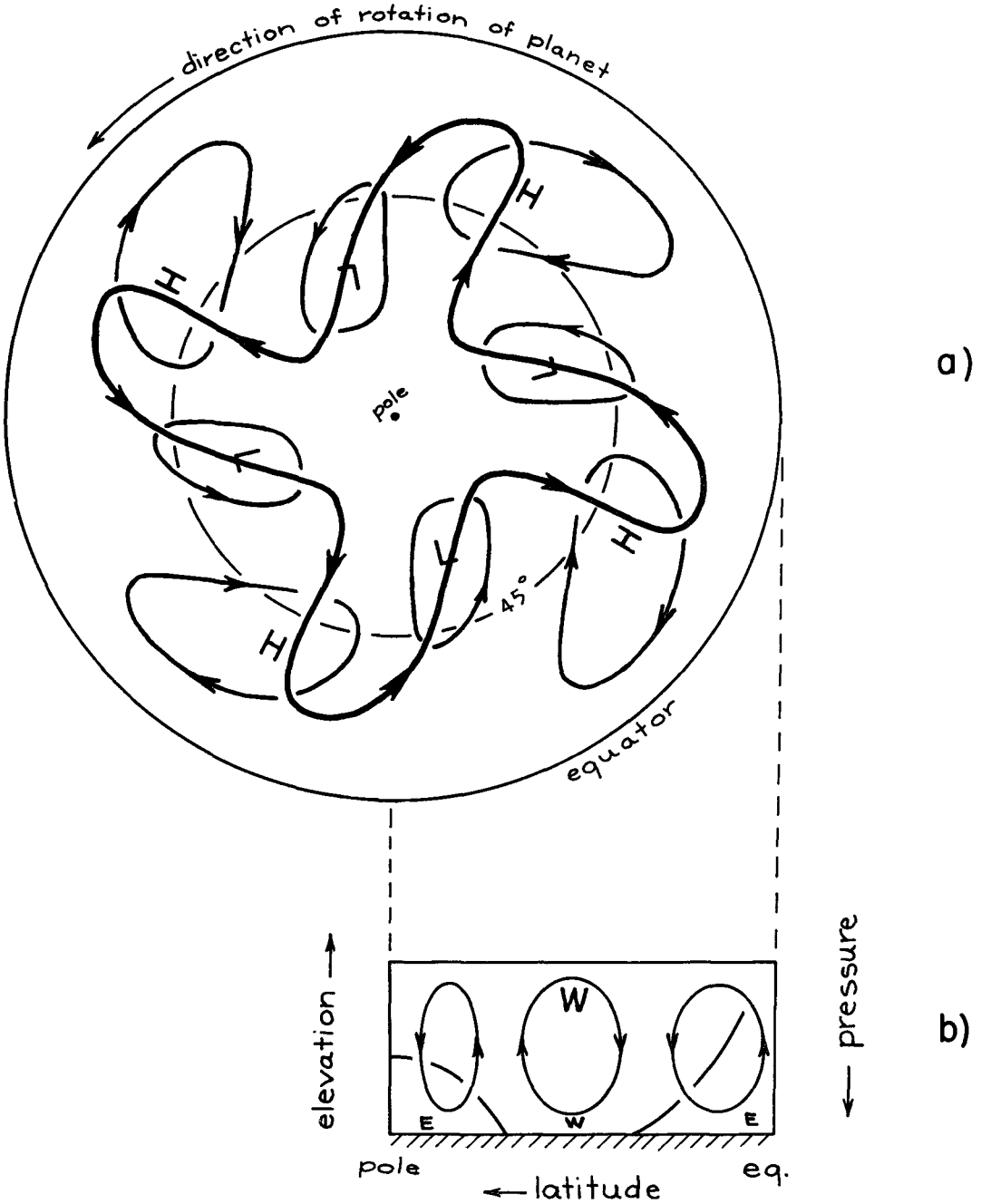


Figure 2. Wave Regime of General Circulation

a) Streamlines of the flow at the middle and upper levels (heavy line) and near the ground (thin line). L = low pressure, H = high pressure.

b) Cross-section showing the zonally-averaged meridional circulation; and the zonally-averaged zonal wind, where W = westerly wind, E = easterly wind.

latitudinal heating. This poleward heat transport by the waves is independent of the vertical lapse rate of temperature. Yet, at the same time the warmer fluid rising at one phase of the waves, while colder fluid sinking at another phase of the waves, produces a net upward transport of heat and converts potential into kinetic energy. But, here also, high level cooling and low level heating is implied if we assume a constant vertical lapse rate of temperature.

With this brief description of the two general circulation regimes, the questions we would like to answer are:

1) What parameters determine which of the two circulation regimes is the dynamically stable one, and hence is the regime of the long term steady state, in any planetary fluid envelope? And,

2) In the light of what we know, or may venture to surmise, about these parameters on *Earth*, on *Mars*, and on *Venus*, which of the regimes ought to be the stable regime of general atmospheric circulation on each of these planets?

To answer these questions, we will show first (in Section B) that a certain poleward temperature gradient, $(-\partial\tilde{T}/\partial\varphi)_{H_s=\Delta Q}$, is required if the regime of symmetric circulation is to transport enough heat poleward to balance the differential latitudinal heating and thereby maintain the planetary fluid envelope in *thermal equilibrium*.

Secondly, we will show (in Section C) that there is a critical limit to the poleward temperature gradient, $(-\partial\tilde{T}/\partial\varphi)_{crit}$, below which only the symmetric regime of general circulation is *dynamically stable*, and above which only the wave regime of general circulation is stable (in the sense of being a "permanently convective" regime) and at the same time able to maintain thermal equilibrium. From this it will be shown that whether a planetary atmosphere has the symmetric regime or the wave regime of general circulation is determined by whether the differential latitudinal heating, ΔQ, is smaller or larger than a certain critical value.

And finally (in Sections D, E, and F), we will try to show, for Earth, for Mars, and for Venus, by comparing in each case the relative magnitudes of ΔQ and ΔQ_{crit} (and an auxiliary parameter λ), what the stable form of general circulation ought to be on each of these planets.

B. The Poleward Temperature Gradient Required for the Symmetric Circulation to Maintain Thermal Equilibrium

Throughout this study we will use the two-parameter (or two-level) model of the atmosphere. As shown in Fig. 3, pressure is used as the vertical coordinate. p_s is the surface pressure, levels 1 and 3 represent respectively the middle of the upper and lower halves of the atmosphere, and the horizontal velocities (both geostrophic and non-geostrophic) are assumed to vary linearly with pressure.

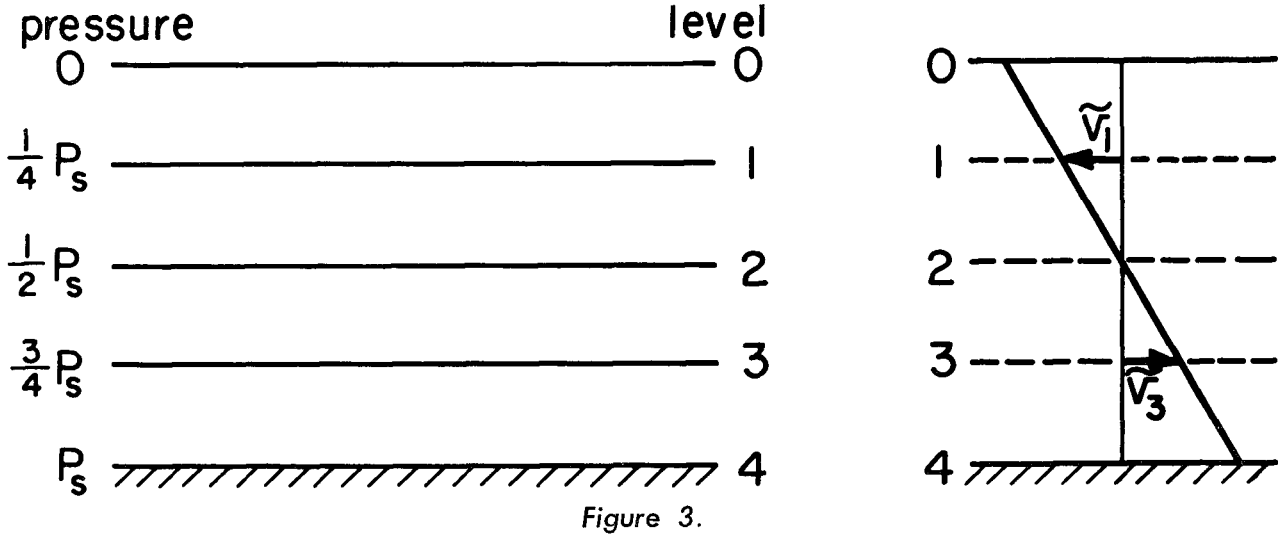


Figure 3.

Reference levels for the two-parameter model atmosphere. p_s is surface pressure. $\tilde{v}$ is the zonally-averaged component of the meridional wind.

We will use the equation of motion in its approximate form,

$$\frac{\partial \mathbf{v}_g}{\partial t} + \mathbf{v}_g \cdot \nabla \mathbf{v}_g = -g\nabla Z - \mathbf{k} \times f\mathbf{v} + \mathbf{F}, \tag{1}$$

where $\mathbf{v}$ is the horizontal velocity, $\mathbf{v}_g = \frac{\sigma}{f}\mathbf{k} \times \nabla Z$ is the geostrophic component of the horizontal velocity, g is acceleration of gravity, Z is the height of the isobaric surface, $f = 2\,\Omega \sin\varphi$ is the coriolis parameter, Ω is the rate of rotation of the planet, φ is latitude, $\mathbf{F}$ is the horizontal frictional force, ∇ is the horizontal del operator on the isobaric surface, $\mathbf{k}$ is unit vertical vector, and t is time.

Equation (1) is called the "quasi-geostrophic" equation of atmospheric motion because the total horizontal velocity is used in the coriolis force term, but only the geostrophic component of the velocity is used in the local and advective acceleration terms. The vertical component of the curl of equation (1) is the "quasi-geostrophic" vorticity equation in one of the forms often used in numerical methods of weather forecasting (see equation 7).

For the case of zonally symmetric flow,

$$v_g \equiv \tilde{v}_g = \frac{g}{f}\frac{\tilde{\partial z}}{\partial x} = 0,$$

where v_g is the meridional (northward) component of the geostrophic wind, x and y are the coordinates toward the east and toward the north, and the superior curved bar denotes the zonal mean (the average with respect to x).

For symmetric flow, therefore, the zonal component of equation (1) reduces to

$$\frac{\partial \tilde{u}_g}{\partial t} = f\tilde{v} + \tilde{F}_x,$$

and, in the steady state, to

$$\tilde{v} = -\frac{1}{f}\tilde{F}_x, \tag{2}$$

where u_g is the eastward (zonal) component of the geostrophic wind, v is the northward component of the total velocity, and F_x is the eastward component of the frictional force.

Equation (2) simply tells us that in the steady-state symmetric circulation the meridional velocity must be such that, at each level, the zonal coriolis force due to the meridional velocity is equal in magnitude and opposite in direction to the zonal frictional force.

Neglecting lateral eddy stresses, the zonal frictional force at level 1 is

$$\tilde{F}_{x,1} = -g\left(\frac{\partial \tau_x}{\partial p}\right)_1 \approx -g\frac{(\tilde{\tau}_{x,2} - \tilde{\tau}_{x,0})}{(p_2 - p_0)} = \frac{-2g}{p_s}\tilde{\tau}_{x,2}$$

$$\approx \frac{-2g}{p_s}\mu_2\left(\frac{\partial \tilde{u}}{\partial z}\right)_2 = \frac{2g^2}{p_s}\rho_2\mu_2\left(\frac{\partial \tilde{u}}{\partial p}\right)_2,$$

where τ_x is the zonal component of the eddy stress, μ is the vertical eddy viscosity, ρ is density, and the numbered subscripts designate the pressure levels as given in Fig. 3.

From the vertical derivative of the geostrophic wind, the hydrostatic equation, and the equation of state, one obtains the thermal wind relation,

$$\left(\frac{\partial u_g}{\partial p}\right)_2 = \frac{1}{af\rho_2 T_2}\left(\frac{\partial T_2}{\partial \varphi}\right), \tag{3}$$

where T is temperature and a is the radius of the planet; and, assuming $(\partial u_g/\partial p)_2 \approx (\partial u/\partial p)_2$, we find

$$\tilde{F}_{x,1} \approx \frac{2g^2}{af p_s \tilde{T}_2}\mu_2\left(\frac{\partial \tilde{T}_2}{\partial \varphi}\right).$$

Substituting into equation (2), and applying the result to the central latitude, where $f = 2\Omega \sin(45°)$, we obtain

$$\tilde{v}_1 \approx \frac{g^2 \mu_2}{a\Omega^2 p_s \tilde{T}_2} \left(-\frac{\partial \tilde{T}_2}{\partial \varphi} \right) . \tag{4}$$

This tells us that in the steady-state symmetric circulation the meridional wind is proportional to the poleward temperature gradient and to the vertical eddy viscosity; it is inversely proportional to the square of the rotation rate of the planet. And given the linear variation of wind with elevation, in the steady state,

$$\tilde{v}_3 = -\tilde{v}_1 .$$

Now, in the symmetric circulation, the atmospheric transport of energy across the central latitude is almost entirely by the transport of sensible heat and potential energy,

$$H_s = \frac{\sqrt{2}\,\pi a}{g} \int_0^{p_s} \tilde{v}(c_p \tilde{T} + gZ)\, dp ,$$

where c_p is the coefficient of specific pressure.

As part of the two-parameter approximation, we will let the winds and temperatures in the upper and lower halves of the atmosphere be represented by their mid-level values. Then

$$\begin{aligned} H_s &= \frac{\pi a p_s}{\sqrt{2}\, g} \left[\tilde{v}_1 (c_p \tilde{T}_1 + gZ_1) + \tilde{v}_3 (c_p \tilde{T}_3 + gZ_3) \right] \\ &= \frac{\pi a p_s}{\sqrt{2}\, g} \tilde{v} \left(\frac{c_p}{g} \frac{\tilde{T}_1 - \tilde{T}_3}{Z_1 - Z_3} + 1 \right) g(Z_1 - Z_3) \\ &= \frac{\pi a p_s R^* \tilde{T}_2 s}{\sqrt{2}\, gm} \tilde{v}_1 , \end{aligned}$$

where $s = \left(1 - \frac{\gamma}{\gamma_d} \right)$ is the static stability, $\gamma = \partial T / \partial Z$ is the vertical temperature gradient, $\gamma_d = -g/c_p$ is the adiabatic vertical temperature gradient, R^* is the universal gas constant, and m is the mean molecular weight.

Substituting from equation (4), we obtain

$$H_s = \frac{\pi g R^* s \mu_2}{\sqrt{2}\, \Omega^2 m} \left(-\frac{\partial \tilde{T}_2}{\partial \varphi} \right), \tag{5}$$

which gives us the relationship between the poleward heat transport and the poleward temperature gradient in the steady-state symmetric circulation. We note that the poleward heat transport, H_s, does not depend on the surface pressure (or columnar mass of the planetary atmosphere).

For thermal equilibrium to be maintained, the poleward heat transport across the central latitude must equal ΔQ, the net heating of the atmosphere on the equatorward side of the central latitude (which, of course, is also equal to the net cooling on the poleward side of that latitude). For thermal equilibrium, therefore, the poleward temperature gradient at the central latitude, in the steady-state symmetric circulation regime, must be

$$\boxed{\left(-\frac{\partial \tilde{T}_2}{\partial \varphi} \right)_{H_s = \Delta Q} = \left(\frac{\sqrt{2}\, \Omega^2 m}{\pi g R^* s \mu_2} \right) \Delta Q .} \tag{6}$$

Given the physical parameters Ω, m, g, s and μ_2, and the differential heating ΔQ (which may, or may not, itself be a function of $-\partial\tilde{T}_2/\partial\varphi$), equation (6) tells us what poleward temperature gradient is required for a symmetric circulation regime to maintain a planetary atmosphere in *thermal equilibrium*. But it does not tell us whether this temperature gradient is one for which the symmetric regime is *dynamically stable*. For that, we must turn, next, to the theory of the dynamic stability of the symmetric circulation.

C. The Critical Poleward Temperature Gradient for Dynamic Stability of the Symmetric Regime

Extensive theoretical studies have been made in recent years of the stability of a symmetric circumpolar vortex and of the form of the wave perturbations that develop when this vortex is unstable. The governing equations used in most of these studies are the quasi-geostrophic vorticity equation [in the form obtained by taking the vertical component of the curl of equation (1), and letting $\mathbf{v}_g \cdot \nabla f = \mathbf{v} \cdot \nabla f$],

$$\frac{\partial \zeta_g}{\partial t} = -\mathbf{v}_g \cdot \nabla(\zeta_g + f) - f\nabla \cdot \mathbf{v} + \mathbf{k} \cdot \nabla \times \mathbf{F}, \tag{7}$$

and the thermodynamic energy equation,

$$\frac{\partial \theta}{\partial t} = -\mathbf{v} \cdot \nabla\theta - \omega \frac{\partial \theta}{\partial p} + \frac{\theta}{c_p T}\dot{q}, \tag{8}$$

where $\zeta_g = \mathbf{k} \cdot \nabla \times \mathbf{v}_g$ is the vertical component of the curl of the geostrophic wind, $\omega = dp/dt$ is the "vertical velocity" in x, y, p-space, θ is the potential temperature, and $\dot{q}$ is the rate of non-adiabatic heating per unit mass.

Making use of the modeling approximation of linear vertical variation of the wind, and applying equation (7) to levels 1 and 3, and equation (8) to level 2 [with substitutions from the geostrophic wind equation, $\mathbf{v}_g = gf^{-1}\mathbf{k} \times \nabla Z$; the equation of continuity of mass, $\nabla \cdot \mathbf{v} = -\partial\omega/\partial p$; Poisson's equation, $\theta = T(p_s/p)^\kappa$; the equation of state, $T = pm/\rho R^*$; and the hydrostatic equation, $\rho g = -\partial p/\partial Z$], we obtain

$$\nabla^2 \frac{\partial Z_1}{\partial t} = -J\left(Z_1, \frac{g}{f}\nabla^2 Z_1 + f\right) + \frac{2f^2}{gp_s}\omega_2 + \frac{f}{g}\mathbf{k} \cdot \nabla \times \mathbf{F}_1, \tag{9}$$

$$\nabla^2 \frac{\partial Z_3}{\partial t} = -J\left(Z_3, \frac{g}{f}\nabla^2 Z_3 + f\right) - \frac{2f^2}{gp_s}\omega_2 + \frac{f}{g}\mathbf{k} \cdot \nabla \times \mathbf{F}_3, \tag{10}$$

and

$$\frac{\partial Z_1}{\partial t} - \frac{\partial Z_3}{\partial t} = \frac{g}{f}J(Z_1, Z_3) + \frac{2R^*s}{mc_p p_s}(Z_1 - Z_3)\omega_2 + \frac{R^*}{gmc_p}\dot{q}, \tag{11}$$

where J is the horizontal Jacobian operator on the isobaric surface.

If Z_1 and Z_3 are given for some initial time, and the force of friction $\mathbf{F}$ and the heating $\dot{q}$ are specified as functions of the fields of Z_1 and Z_3 (i.e., functions of the geostrophic wind, air temperatures and geographic positions); then, with suitable lateral boundary conditions, equations (9) to (11) can be solved for the three unknowns, $\partial Z_1/\partial t$, $\partial Z_3/\partial t$ and ω_2. Integrating over time, these prediction equations yield the future fields of Z_1 and Z_3 (and, implicitly, ω_2) and hence the future three-dimensional fields of temperature, geostrophic wind and vertical velocity of the model atmosphere.

Studies of the general circulation of the earth's atmosphere, by numerical integration of equations (9) to (11), have been made by Phillips (1956) and by Huss and Mintz (1961), and some results from the latter study will be given in Section D of this paper. But one obtains a clearer intro-

duction to the physical processes controlling the stability of the symmetric circulation, as well as a precise mathematical statement of the relationship between the parameters involved, from the analytical solution of the linearized version of these equations.

The linearization consists of neglecting the heating and friction terms and requiring that the zonal component of the geostrophic wind be horizontally uniform and independent of time. A horizontally constant geostrophic zonal wind automatically makes the phase and amplitude of the meridional wind component independent of y.

With this linearization, it has been shown (see, for example, Thompson, 1961) that the behavior of any sinusoidal perturbation of the zonal flow is determined by the parameter

$$\delta = \left(-\frac{\partial \widetilde{T}_2}{\partial \varphi} \right) - \frac{a^2 \Omega^2 m \lambda^2}{R^* n^2 (\lambda^4 - n^4)^{1/2}}, \tag{12}$$

where

$$\lambda = \frac{a\Omega m}{R^*} \left(\frac{2c_p}{T_2 s} \right)^{1/2} \tag{13}$$

and n is the planetary wave numbers, i.e. the number of waves around the hemisphere. When $\delta < 0$, the difference between the phase of the perturbation at the levels 1 and 3 periodically changes sign with time. That is to say, the vertical tilt of the trough and ridge lines of the perturbation alternates with time between an upstream and downstream tilt with respect to the direction of the zonal thermal wind shear. But when $\delta > 0$, regardless of the initial state of the perturbation the difference in phase between the perturbation at the upper and lower levels monotonically approaches a limiting positive value (of about 1/6 the wavelength); that is to say, the wave perturbation acquires and maintains a vertical upstream tilt.

The significance of the vertical tilt of the pertubation is *twofold*:

(1) When the wave perturbation tilts upstream with height, the solution of the governing equations for $\nabla \cdot \mathrm{v} = -\partial\omega/\partial p$, at levels 1 and 3, is such that at both levels there is horizontal velocity convergence on the trough lines of the waves and horizontal velocity divergence on the ridge lines. By equation (7), this signifies an increase in the amplitude of the perturbation at both levels. Conversely, when the wave tilts downstream with height, there is divergence on the troughs and convergence on the ridges at both levels, and the perturbation decreases in amplitude.

Thus, when $\delta < 0$ the perturbation at both levels simply oscillates about its initial amplitude. But when $\delta > 0$, the perturbation acquires a permanent upstream tilt and grows in amplitude without limit (as long as the geostrophic zonal wind speed at the two levels remains constant with time).

Therefore, when $\delta > 0$ *the kinetic energy of the wave perturbation increases.*

(2) The second important aspect of the vertical tilt is that the sign and magnitude of the tilt determine the sign and magnitude of the poleward heat transport by the waves. When the tilt is westward (which is an upstream tilt when the zonal thermal wind shear is from west to east), there is, from the relationship between temperature and $(Z_1 - Z_3)$, a positive correlation at all levels between the temperature and the meridional component of the wind and, as described in the introduction, at each level tongues of warm air move northward while tongues of cold air move southward, producing the poleward heat transport H_w. The opposite heat transport takes place when the vertical tilt is eastward.

It can be shown that the total poleward heat transport across the central latitude by the waves, when the surface winds are weak compared to the upper level winds, is

$$H_w \approx \frac{\pi a^2 \Omega p_s V_2^2 \sin\alpha}{\kappa g n} \tag{14}$$

where $V_2=(V_1+V_3)/2$ is the mean amplitude of the wave in the wind field (the amplitude of the meridional component of the wind at the middle level), α is the phase lag between the waves at levels 1 and 3, and $\kappa = R^*/mc_p$ is Poisson's constant.

Therefore, when $\delta>0$ the wave perturbation acts to reduce the poleward temperature gradient.

If we plot the neutral curve, $\delta=0$, as a function of poleward temperature gradient $(-\partial\tilde{T}_2/\partial\varphi)$ and planetary wave number, n, we obtain the curve shown in Fig. 4.

As the figure shows, there is a critical value of the poleward temperature gradient,

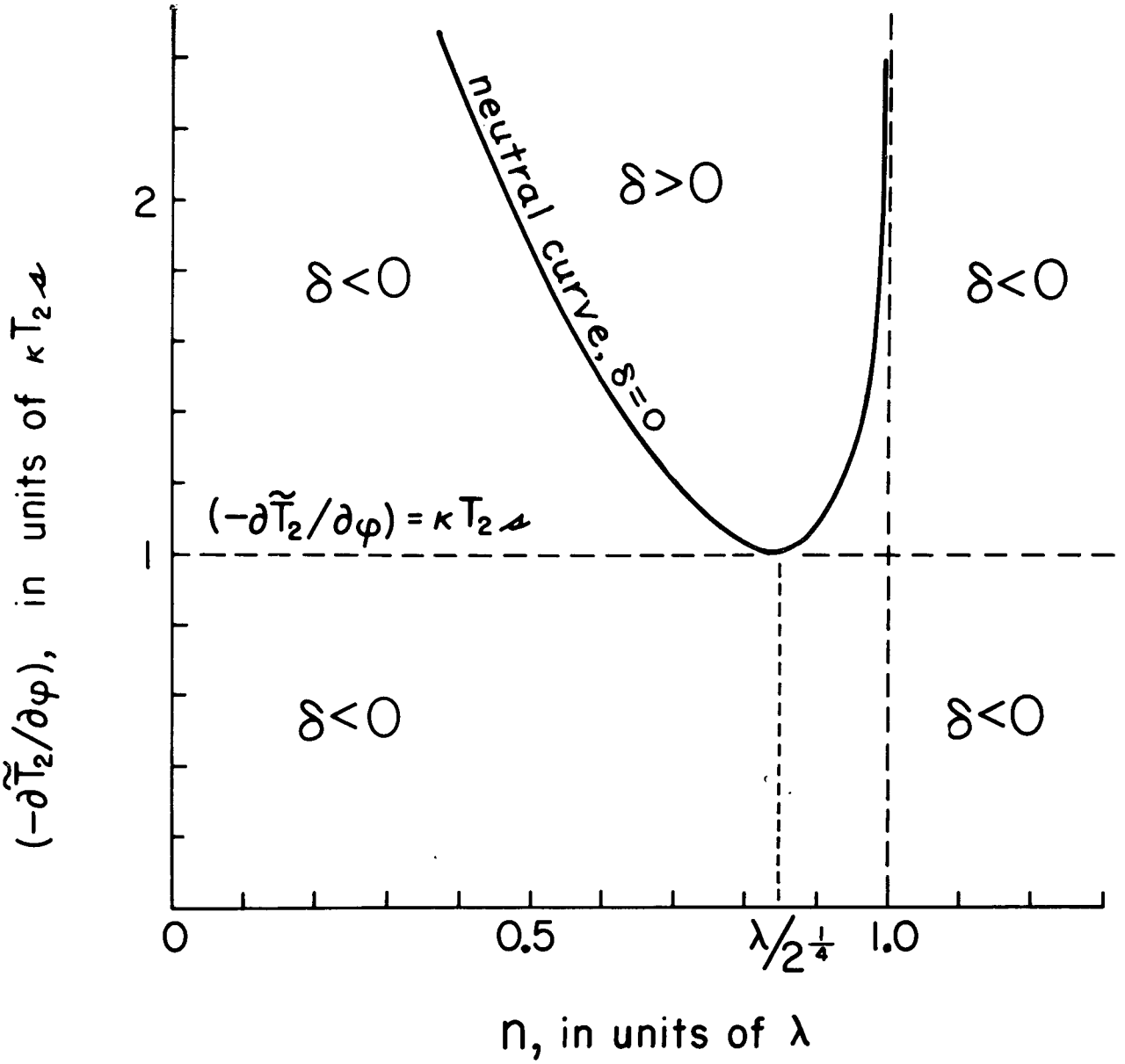


Figure 4.

Stability of the symmetric circulation as a function of the poleward temperature gradient $(-\partial\tilde{T}_2/\partial\varphi)$ and the planetary wave number n.

$$\boxed{\left(-\frac{\partial \widetilde{T}_2}{\partial \varphi}\right)_{\text{crit}} = \kappa T_2 s,} \tag{15}$$

below which $\delta<0$ for perturbations of all wavelengths. When $(-\partial\widetilde{T}_2/\partial\varphi)$ barely exceeds $\kappa T_2 s$, perturbations of planetary wave number $n=\lambda/2^{1/4}$ have $\partial>0$, while longer and shorter waves still have $\delta<0$. For the reason that will be given presently, the wave number for which $n=\lambda/2^{1/4}$ is called the dominant wave number.

$$n_d = \frac{\lambda}{2^{1/4}} \,. \tag{16}$$

Should the poleward temperature gradient be larger than the critical value $\kappa T_2 s$, then δ is positive over a broader range of wavelengths, but it is never positive for wave numbers larger than λ. (This last fact will be important when we apply the theory to the planet Venus.)

The dominant wave number n_d is the important wave number because the critical value of $(-\partial\widetilde{T}_2/\partial\varphi)$ never can be very much exceeded (except when $\lambda \leqq 1$). We may imagine, for example, a differential heating ΔQ which exceeds H_s and thereby makes the poleward temperature gradient $(-\partial\widetilde{T}_2/\partial\varphi)$ gradually increase and approach its critical value from below. Because of irregularities of the surface of the planet, or irregularities in the atmosphere itself, small amplitude perturbations in all wavelengths will always be present, and as $(-\partial\widetilde{T}_2/\partial\varphi)$ passes its critical value perturbations whose wavelengths lie near the dominant wavelength will start amplifying. These waves will transport heat polewards and thermal equilibrium will first be reached when the waves have grown large enough in phase lag and amplitude for the poleward transport to balance the heating, or $H_W=\Delta Q$. But, as the waves are then in the westward tilting phase, they continue to grow. H_W then becomes larger than ΔQ, $(-\partial\widetilde{T}_2/\partial\varphi)$ in time is reduced to less than its critical value, and a complex cycle of changes of zonal wind, meridional wind, and poleward heat transport results, in which thermal equilbrium is maintained by the waves only in the long-term average.

The criterion for the establishment and maintenance of the wave regime, therefore, is whether $(-\partial\widetilde{T}/\partial\varphi)_{H_s=\Delta Q}$ is larger or smaller than $(-\partial T_2/\partial\varphi)_{\text{crit}}$.

But by substitutions from equations (6) and (15), we see that

$$\boxed{\begin{array}{l}(-\partial\widetilde{T}_2/\partial\varphi)_{H_s=\Delta Q} \lessgtr (-\partial\widetilde{T}_2/\partial\varphi)_{\text{crit}} \\ \text{WHEN} \qquad \Delta Q \lessgtr \Delta Q_{\text{crit}} \,,\end{array}} \tag{17}$$

where

$$\Delta Q_{\text{crit}} = \frac{\pi g \kappa R^* \widetilde{T}_2 \mu_2 s^2}{\sqrt{2} m \Omega^2} \,.$$

Therefore, *when* $\Delta Q < \Delta Q_{\text{crit}}$ *the symmetric general circulation regime will maintain thermal equilibrium and be dynamically stable.*

But when $\Delta Q > \Delta Q_{\text{crit}}$ *then the wave regime of general circulation will form.*

In the long term average $H=\Delta Q$, and from equation (14)

$$\overline{H}_W = \Delta Q \approx \frac{\pi a^2 \Omega p_s}{\kappa g n_d} \overline{V_2^2 \sin\alpha} \,,$$

where the superior straight bar denotes the time mean. The time-averaged poleward heat transport by

the waves, $\bar{H}_W$, depends not only on the magnitudes of V_2 and α, but on the correlation between them. But for a rough order-of-magnitude estimate, we will let

$\bar{H}_W = \Delta Q \sim \pi a^2 \Omega p_s \overline{V_2^2} \sin \bar{\alpha} / \kappa g n_d$ and, arbitrarily, $\bar{\alpha} \sim \frac{1}{2} \alpha_{\text{LIMIT}}$, or $\alpha \sim 60°/2$.

Substituting from (16),

$$\overline{V_2^2} \sim \frac{\kappa g \lambda}{2^{1/4} \pi a^2 \Omega p_s (\sin 30°)} \Delta Q . \tag{18}$$

But there is an important exception to the rule given by the criterion (17), and that exception occurs if $\lambda < 1$. [For example, as defined in (13), λ can be less than unity on a very slowly rotating planet.] On such a planet no waves will form and the circulation will stay in the symmetric regime, until $H_s = \Delta Q$, no matter by how far ΔQ exceeds ΔQ_{crit} or $(-\partial \tilde{T}_2 / \partial \varphi)$ exceeds $\kappa T_2 s$.

We may note, here, that the criterion for the change from symmetric to wave regime, described above, is precisely opposite to what is observed in the well known laboratory experiments with differentially heated rotating fluids. According to equation (17), a change from symmetric regime to wave regime is produced by an *increase* of the differential heating ΔQ; whereas in the laboratory experiments, both of the "dishpan" type (Fultz, 1959) and the "annulus" type (Hide, 1958), a change from symmetric regime to wave regime is produced by a *decrease* of the differential heating ΔQ. As we shall show, in the section on Mars, the latter type of transition takes place when the space-averaged static stability is not constant but increases with increasing ΔQ; and this only occurs when the static stability is not greatly influenced by heat of condensation.

D. Circulation of the Earth's Atmosphere

The theoretical determination of the regime of general circulation of the Earth's atmosphere, as of any planet, is obtained by comparing ΔQ with ΔQ_{crit} (with consideration of the magnitude of λ). As we shall see, to a first approximation ΔQ for the Earth is independent of the temperature distribution and therefore easily obtained, in approximate magnitude, from external parameters.

The net heating per unit horizontal area, as a function of latitude, is $[S(1-A)-W]$, where S is the incident solar radiation, A is the planetary albedo, and W is the long-wave radiation emitted by the planet.

Astronomical relations give us the latitudinal distribution of the incident solar radiation S. And for our model of the Earth, we will take the planetary albedo as the constant $A = 0.34$. This gives us the mean annual distribution of $S(1-A)$ shown in the upper part of Fig. 5.

Because the Earth has extensive oceans that are heated by the sun, there is a widespread and persistent upward diffusion of water vapor from the ocean source, and the air in all latitudes of the Earth, roughly speaking, is in a state of near saturation. We will assume, for our model of the Earth, that the relative humidity is uniformly high in all latitudes and that there is sufficient water vapor in the model atmosphere to make all of the radiation come from the water vapor in the air, as shown schematically in the lower part of Fig. 5. Under these conditions, where the temperature is high (as it will become near the equator) the water vapor which emits the outgoing radiation will be at relatively high levels in the atmosphere; where the temperature is low (as it will become near the poles) the emitting water vapor will be at lower levels. But the intensity of the outgoing radiation, which depends essentially only on the temperature of the water vapor and not on its elevation, will be of the same magnitude everywhere. Hence, in the model, the outgoing radiation W will be constant with latitude. For thermal equilibrium, the magnitude of W must be equal to the latitude mean of $S(1-A)$

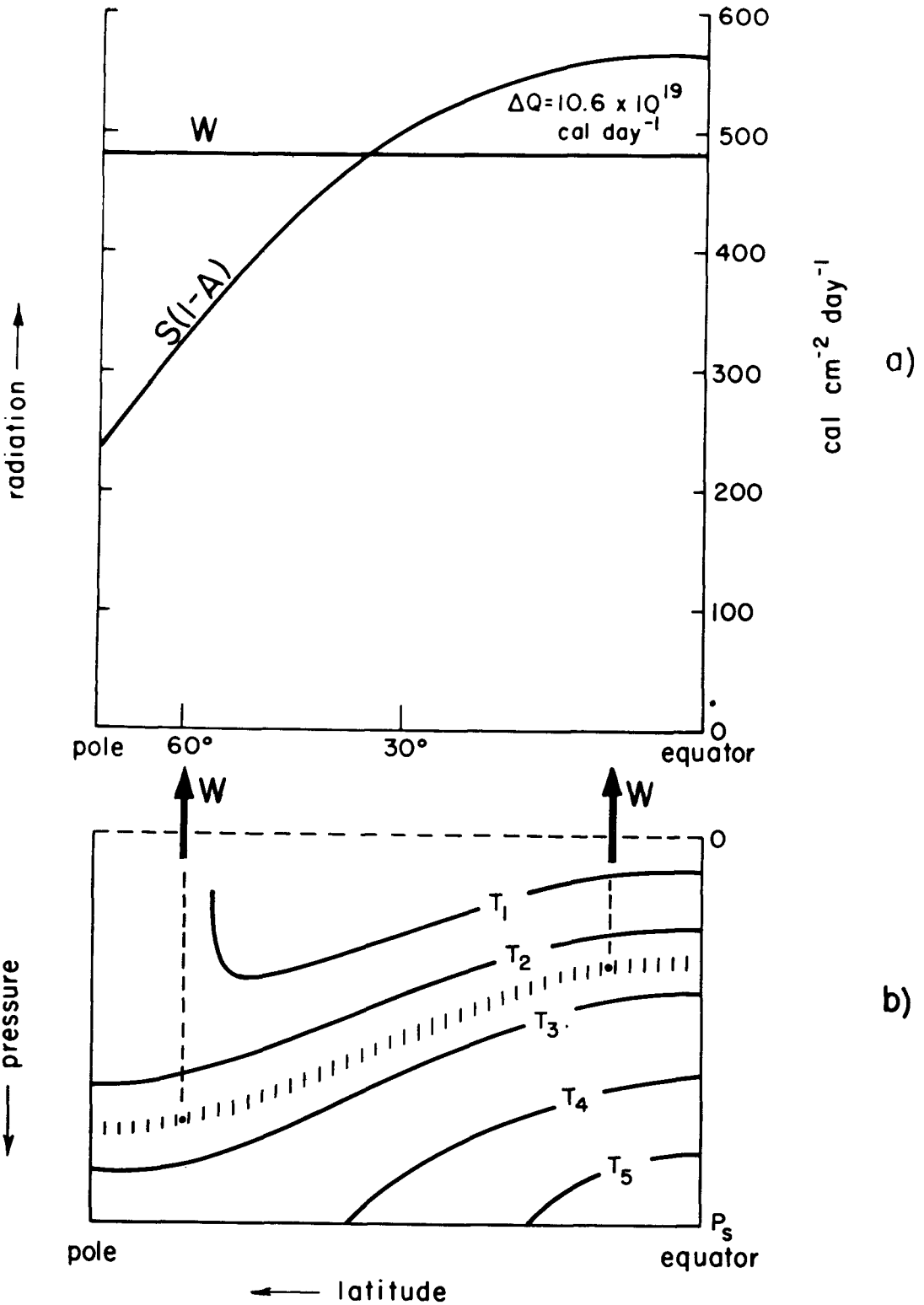


Figure 5. Model of the Radiation Budget for the Planet Earth

a) Energy gain and loss as a function of latitude, where S is the insolation per unit horizontal area per unit time, A is the planetary albedo, W is the infrared emission to space, and ΔQ is the net differential heating.

b) Infrared emission to space when the relative humidity is constant along the isothermal surfaces T. The shaded band represents a layer of water vapor, of constant optical thickness, which contributes to the total infrared emission W.

which is known, and this results in a net heating between the equator and latitude 37°, and a net cooling between latitude 37° and the pole, of 10.6×10^{19} cal day^{-1} (compare footnote*), or

$$\Delta Q \approx 10.6 \times 10^{19} \text{ cal day}^{-1}$$

$$\approx 5.1 \times 10^{12} \text{ kj s}^{-1}.$$

This value of ΔQ, it should be noted, will act to increase the poleward temperature gradient at the central latitude by only about 0.3C° per radian per day.

To obtain ΔQ_{crit}, for the model atmosphere of the Earth, we will use $a = 6.37 \times 10^{6} m$, $g = 9.81 m\,s^{-2}$, $\Omega = 7.29 \times 10^{-5}\,s^{-1}$, $T_2 \sim 260°K$; and a gaseous mixture of 80 per cent nitrogen and 20 per cent oxygen, so that $c_p = 1004\ kj\ t^{-1}\ deg^{-1}$ and $R^*/m = 287 kj\ t^{-1}\ deg^{-1}$.

The vertical eddy viscosity of the free atmosphere of the Earth is not well known, and it undoubtedly varies considerably with space and time. We will use the mean value given by Palmen (1955), which is $\mu \approx 2.2 \times 10^{-2}\ t\ m^{-1}\ s^{-1}$.

To obtain the stability factor s, we must know the adiabatic vertical temperature gradient γ_d and the actual vertical temperature gradient γ. $\gamma_d = -g/c_p = -10C°/km$. But it has not been possible as yet to derive, entirely from first principles, the actual vertical temperature gradient for the Earth's atmosphere (cf. Manabe, 1961). We will use, for our model, the observed mean lapse rate of temperature, averaged with respect to pressure through the entire depth of the atmosphere, which is $\gamma \approx \gamma_d/2$, so that $s = (1 - \gamma/\gamma_d) = 0.5$.

Using the above values of the parameters for our computations, we obtain

$\Delta Q = 5.1 \times 10^{12}\ kj\ s^{-1}$ $\qquad$ $\Delta Q_{\text{crit}} = 0.6 \times 10^{12}\ kj\ s^{-1}$

$\left(-\partial \tilde{T}_2/\partial \varphi\right)_{H_s = \Delta Q} = 310C°/\text{rad}$ $\qquad$ $\left(-\partial \tilde{T}_2/\partial \varphi\right)_{\text{crit}} = 38C°/\text{rad}$

$\lambda = 6.36, \quad n_d = (5.4) \sim 5$ $\qquad$ $L_d = \dfrac{2\pi a \cos(45°)}{n_d} = 5300\ km,$

and

$$\left(\overline{V_2^2}\right)^{1/2}_{H_W = \Delta Q} \sim 13\ m\ s^{-1}, \text{ or } r.m.s.\ (V_2)_{H_W = \Delta Q} \sim 9\ m\ s^{-1}.$$

Thus we see that for our model of the Earth's atmosphere the differential heating ΔQ is many times larger than ΔQ_{crit}, the maximum for a dynamically stable symmetric circulation. For a symmetric circulation which would transport across the central latitude all the heat required for thermal equilibrium, the poleward temperature gradient would have to be 310C° per radian, but the critical dynamical limit of the poleward temperature gradient is only 38C° per radian. Therefore, planetary waves must develop in this atmosphere, with the number of waves around the Earth the nearest whole number to the dominant wave number, 5.4 (corresponding to the dominant wavelength, at latitude

* Houghton (1954), taking into account the observed variation of albedo with latitude on the real Earth, and the observed variation of relative humidity, obtained $\Delta Q = 11.1 \times 10^{19}$ cal day^{-1}. The reason for the close agreement between his empirical computation and that of our simple theoretical model is that the observed increase of relative humidity from subtropics to pole is accompanied by an observed increase of cloudiness. The increase of relative humidity from subtropics to pole makes the radiation come from a lower temperature of the air at the pole and produces a small poleward gradient of outgoing radiation, while the increase of cloudiness from subtropics to pole produces a small increase of albedo and a corresponding increase in the gradient of effective insolation. As a result, the two small modeling errors, introduced by our approximations of constant relative humidity and constant albedo, are errors of opposite sign and act to cancel one another.

45°, of 5300 km). Furthermore, the root mean square amplitude of the meridional wind in these waves will be approximately 13 $m\ s^{-1}$ (which gives a root mean square meridional wind of 9 $m\ s^{-1}$).

But these are the theoretically computed values for a model atmosphere in which there is no release of latent heat of condensation. If we allow condensation to occur as a consequence of the vertical motion ω_2, then the saturated adiabatic process rate γ_s must replace the unsaturated adiabatic process rate γ_d over about one-half of the wave domain. Therefore, after the wave regime has set in and the vertical displacements of the air have produced condensation, we must replace s by s^*, where

$$s^* = \tfrac{1}{2}\left[\left(1 - \frac{\gamma}{\gamma_d}\right) + \left(1 - \frac{\gamma}{\gamma_s}\right)\right].$$

Given $\gamma_s \approx -0.6C^\circ/km$, then $s^* \approx 0.3$.

Replacing s by s^*, after the waves have developed to the extent where they produce condensation, has two important consequences:

(1) It reduces the critical poleward temperature gradient and hence reduces the time-averaged equilibrium value of $(-\partial\widetilde{T}_2/\partial\varphi)$ (but results in only a small change in the dominant wave number, because n_d varies inversely with the square root of s).

(2) Of much greater importance, the change from s to s^*, after the waves have developed, produces a condition of hypercritical instability and leads to a rapid rate of wave amplification in the one-dimensional case—and a rapid rate of cyclogenesis in the two-dimensional system.

To illustrate this phenomenon, imagine an initial condition in which the poleward temperature gradient is sub-critical and the circulation is symmetric. Then $H_s < \Delta Q$, and the poleward temperature gradient increases slowly with time. With no condensation, the circulation is not unstable until $(-\partial\widetilde{T}_2/\partial\varphi)$ exceeds $\kappa T_2 s$, the critical value for unsaturated air. After the temperature gradient exceeds $\kappa T_2 s$, as represented schematically by the point A in Fig. 6, perturbations of the dominant wavelength begin to amplify.

But now, when either before or after $H_W = \Delta Q$ the vertical motion of the amplifying perturbations brings about condensation in the ascending branches of the waves, the stability criterion is changed abruptly from $\kappa T_2 s$ to $\kappa T_2 s^*$. The perturbations at point A, which previously had only a small growth rate because the point lay close to the neutral curve $\delta(s)=0$ (and, in fact, could never move very far above this curve because the given ΔQ can increase the poleward temperature gradient only very slowly), these perturbations now find themselves far from the new neutral curve $\delta(s^*) = 0$. Because $s^*/s \approx 3/5$, about two-thirds of all the potential energy that was slowly stored in the build-up of the temperature gradient to its magnitude $\kappa T_2 s$ now is available for rapid conversion into the kinetic energy of the waves. Therefore, in this model atmosphere, waves which would grow only slowly and to moderate amplitudes ($rms\ V_2 \sim 13\, m s^{-1}$) when the temperature gradient is brought past its critical value from below and no condensation is permitted—these same waves will grow rapidly and to larger amplitudes (for a shorter time) when condensation is allowed. With the larger amplitudes, $H_W >> \Delta Q$, and this makes the poleward temperature gradient fall rapidly to the saturation-critical value $\kappa T_2 s^*$. Only after $(-\partial\widetilde{T}_2/\partial\varphi)$ falls below $\kappa T_2 s^*$ do the waves die out, and the cycle starts again. But when the cycle starts again, it is $\kappa T_2 s$ that is again the determining value for the next release of hypercritical instability.

The following table shows the theoretically predicted values of the critical (and hence the limiting) poleward temperature gradients, the dominant wave numbers, and the root mean square meridional wind for our model of the Earth's atmosphere, compared with the observed values for the real Earth atmosphere. Both sets refer to the yearly averages.

Predicted	Observed
$\left(\frac{-\partial \tilde{T}_2}{\partial \varphi}\right)_{\text{crit, } s=0.5} = 38\text{C}^\circ/\text{rad}$ $\left(\frac{-\partial \tilde{T}_2}{\partial \varphi}\right)_{\text{crit, } s^*=0.3} = 23\text{C}^\circ/\text{rad}$	$\left(\frac{-\partial \tilde{T}_2}{\partial \varphi}\right)_{\text{observed}} = 29\text{C}^\circ/\text{rad}$
$n_{d,\, s=0.5} = (5.4) \sim 5$ $n_{d,\, s^*=0.3} = 7$	$n_{d,\,\text{observed}} = 6$ or 7
$rms(\text{v}_2)_{H_W=\Delta Q} \sim 9\ m\ s^{-1}$	$rms(\text{v}_2)_{\text{observed}} = 10.6\ m\ s^{-1}$

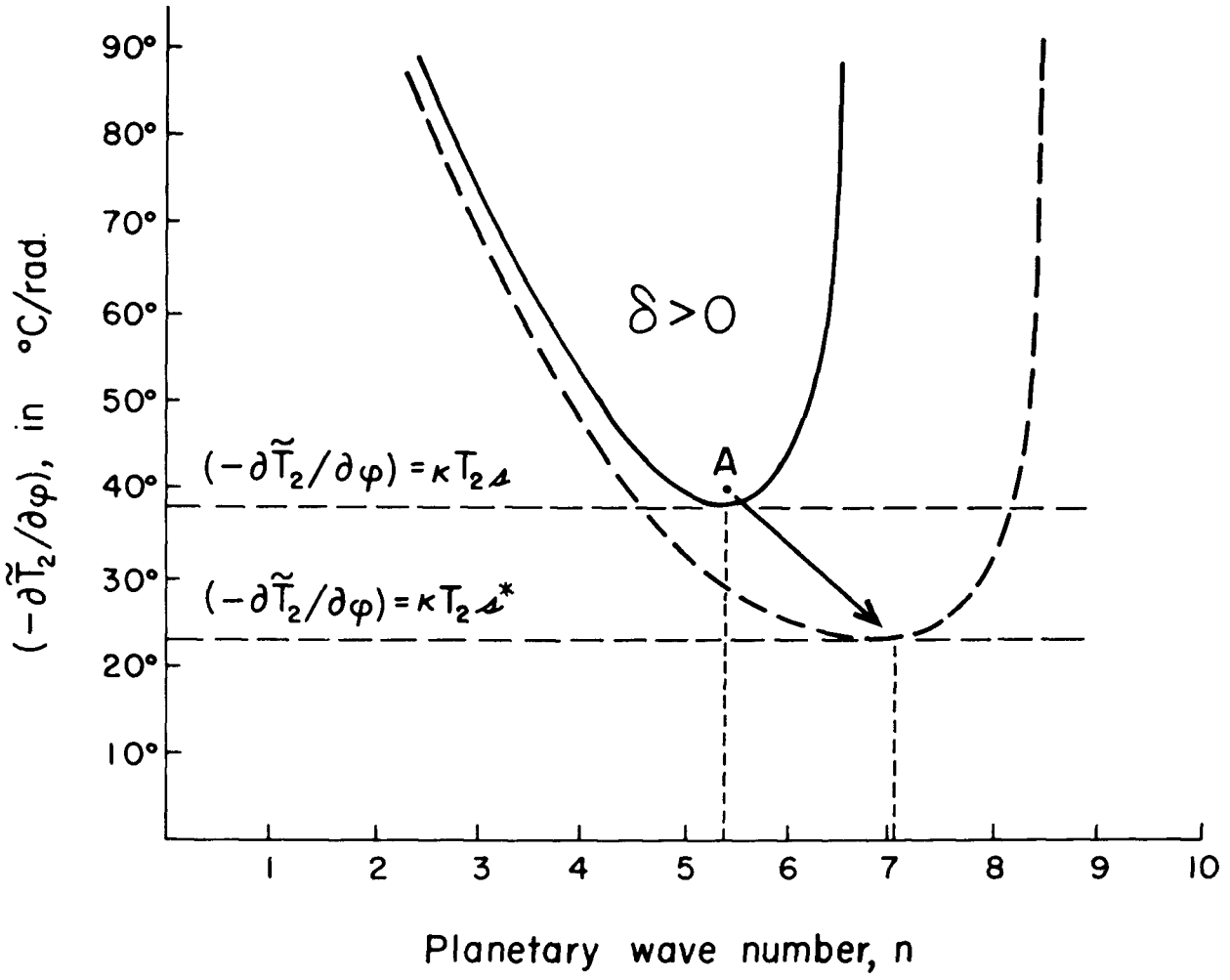


Figure 6.
Illustration of hypercritical instability of atmospheric waves resulting from wave-produced condensation.

The agreement between the theoretically derived quantities for the model atmosphere of the Earth and the observed quantities for the real atmosphere of the Earth is very good. The observed poleward temperature gradient falls almost half-way between the two limiting values, the observed dominant wave number is in close agreement with the theoretical numbers, and the observed root-mean-square meridional wind is of roughly the same value as the theoretical estimate.

With respect to seasonal changes of the general circulation, we note, first, that the intensity of the atmospheric waves depends directly upon ΔQ, but that neither the critical poleward temperature gradient $(-\partial\widetilde{T}_2/\partial\varphi)_{\text{crit}}$ (and hence the actual poleward temperature gradient), nor the dominant wave number n_d depend directly upon ΔQ.

Now, the differential heating of the Earth's atmosphere is approximately the same as the differential heating of the Earth as a whole only in the mean for the year, and therefore the net heating of the atmosphere is approximated by $[S(1-A)-W]$ only in the mean for the year. In the individual seasons we must add or subtract the heat storage in the oceans, and this is not known. Indirectly, however, we can obtain a rough estimate of the seasonal values of the differential heating of the atmosphere from the measured poleward transport of sensible heat at the central latitude, which is about four times larger in winter than in summer (Mintz, 1955). Using this ratio, we obtain

$$\Delta Q_{\text{summer}} \sim 2 \times 10^{12}\ kj\ s^{-1}$$
$$\Delta Q_{\text{winter}} \sim 8 \times 10^{12}\ kj\ s^{-1}.$$

It follows from this (by equation 18) that the root mean square meridional wind must be twice as large in winter as in summer, and this is about what is observed.

On the other hand, the static stabilities, s and s^*, change relatively little with season in the Earth's atmosphere. In the wave regime there is an upward heat transport which is proportional to the poleward heat transport by the waves. This upward heat transport is therefore four times larger in winter than in summer, which by itself should make the space-averaged static stabilities much larger in winter than in summer. But the heat of condensation produced by the waves is almost entirely released in the lower half of the atmosphere, and heating only the lower half of the atmosphere in this way is a destabilizing effect that counteracts the upward transport of heat. Mainly as a result of this vertical distribution of condensation, the static stabilities are only slightly greater in winter than in summer and the critical (and therefore, actual) poleward temperature gradient (by equation 15) is only slightly greater in winter than in summer. Observations show that at the central latitude,

$$\left(-\frac{\partial\widetilde{T}_2}{\partial\varphi}\right)_{\text{summer}} = 22\text{C}^\circ/\text{rad},$$

and

$$\left(-\frac{\partial\widetilde{T}_2}{\partial\varphi}\right)_{\text{winter}} = 36\text{C}^\circ/\text{rad}.$$

This seasonal change of poleward temperature gradient is relatively small compared to the fourfold increase of ΔQ. The dominant wave number (by equations 16 and 13) depends inversely on the square root of the static stability. Therefore, in spite of the large seasonal change in ΔQ, the wave number can decrease only slightly from summer to winter (from 7 or 8 to 5 or 6); and this, in fact, is observed. But in the next section, on Mars, which is a planet without appreciable condensation, we shall see that much larger seasonal changes are to be expected.

The shortcoming of the simple linear theory is that it cannot predict the evolution of the circulation with respect to its variation with latitude. That is why, in applying the linear theory, we dealt only with the circulation at the central latitude and spoke only of amplifying waves and not of amplifying cyclones and anticyclones.

It is the nonlinear terms in the governing equations which produce closed circulation centers—cyclones and anticyclones—at the lower levels and steer the cyclones toward the pole and the anticyclones toward the equator, thereby making the mean zonal surface winds westerly in middle latitudes and easterly in the low and high latitudes. These terms also produce a high-level meandering westerly jet stream, and they reverse the mean meridional circulation in the middle latitudes.

The nonlinear equations, (9) to (11), can be integrated numerically, with the help of a large-capacity high-speed computer. An example of such an integration is shown in Fig. 7 (from Huss and Mintz, 1961). The flow, in this example, is confined in a channel between rigid walls at the north and south, and in the governing equations both f and ∇f are held constant over the channel, which is the so called beta-plane approximation.

Because, in this example, we wished to study only the evolution of a disturbance of the symmetric flow from an initial state in which $(-\partial\widetilde{T}_2/\partial\varphi) > (-\partial\widetilde{T}_2/\partial\varphi)_{crit}$, the heating and friction terms were dropped from the governing equations. And because, in this example, condensation of water vapor was excluded, thereby precluding hypercritical instability, $(-\partial\widetilde{T}_2/\partial\varphi)$ was taken initially as about 40 per cent larger than the critical value $\kappa T_2 s$ in order to obtain a finite growth rate for the disturbance. On the uniform geostrophic zonal current, whose velocity at the ground was zero, an initial small amplitude sinusoidal disturbance of wavelength approximately 6000 km was imposed, and it is the evolution of this small disturbance that is shown in the figure.

In the numerical computations, the history of the motion was carried by the streamfunctions, Z_1 and Z_3, but for convenience of interpretation we show the 1000 mb height field and the temperature field that are derived from the predicted fields of Z_1 and Z_3.

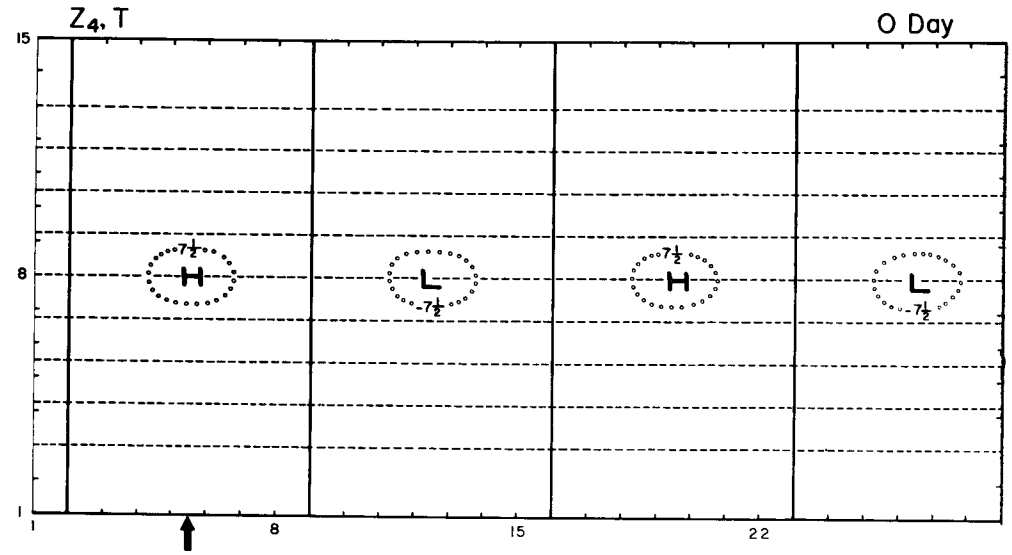
We can interpret the initial change, from the 0th day to the 4th day, as showing (in agreement with linear theory) that the given initial horizontal gradient of temperature was large enough for horizontal advection to exceed vertical advection of potential temperature. As a consequence, the small initial disturbance deformed the initially straight isotherms into a small amplitude temperature wave that lagged 90° behind the disturbance.

To interpret the mechanism of development from this point on, it is helpful to know the horizontal velocity divergence at the surface, $\nabla \cdot \mathbf{v}_4 = -(\partial\omega/\partial p)_4$. By solving the governing equations (9), (10), and (11) for ω_2, we obtain, when the heating and friction terms are dropped,

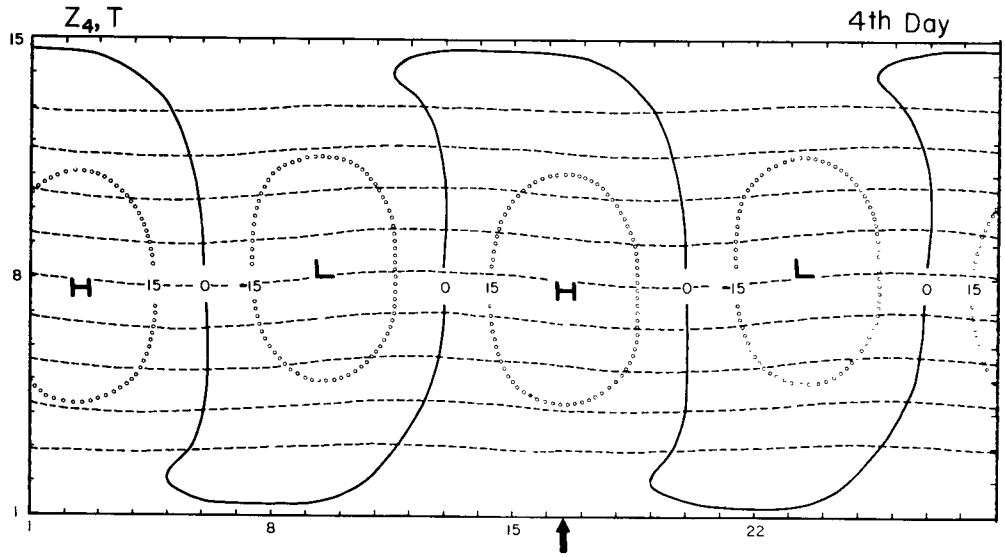
$$(1 - \sigma\nabla^2)\left(\frac{\partial\omega}{\partial p}\right)_4 = -\frac{1}{f}\mathbf{v}_T \cdot \nabla\,(2\zeta_4 + 2\zeta_T + f) - \frac{2}{f}\mathbf{k} \cdot \left[\frac{\partial\mathbf{v}_T}{\partial x} \times \frac{\partial\mathbf{v}_4}{\partial x} + \frac{\partial\mathbf{v}_T}{\partial y} \times \frac{\partial\mathbf{v}_4}{\partial y}\right], \tag{19}$$

where $\sigma = sT_2R^{*2}/2c_pf^2m^2$, $\mathbf{v}_T = (\mathbf{v}_1 - \mathbf{v}_3)$ is a wind shear vector parallel to the isotherms with the warmer air to its right, and $\zeta_T = \mathbf{k} \cdot \nabla \times \mathbf{v}_T$.

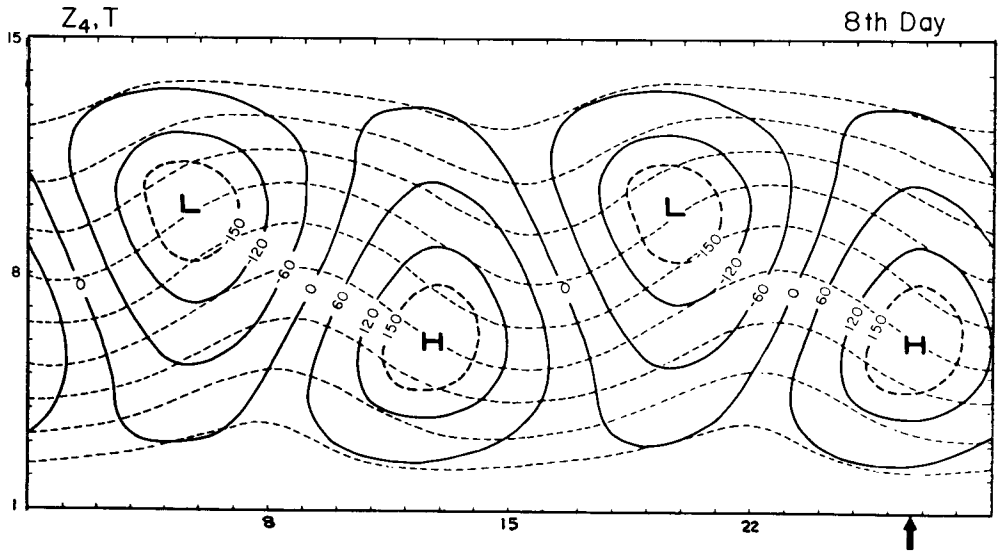
It is easily shown, for a sinusoidal disturbance with wavelength of the order of 6000 km, that the second term on the left in (19) is about half as large as the first term but that both terms are of the same sign. Therefore, although there will be an error in magnitude, there will be no error in the qualitative interpretation of the numerical computation if we drop the second term. In addition, as long as the thermal wind shear vector $\mathbf{v}_T$ has not acquired large horizontal variations, the second group of terms on the right in equation (19) will be smaller than the first group on the right and again no qualitative error will be made if we drop the second group.


Z4, T
0 Day
15
8
1
H
L
7½
-7½
1
8
15
22

a)


Z4, T
4th Day
15
8
1
H
L
-15
0
15
1
8
15
22

b)


Z4, T
8th Day
15
8
1
L
H
0
60
120
150
-60
-120
8
15
22

c)

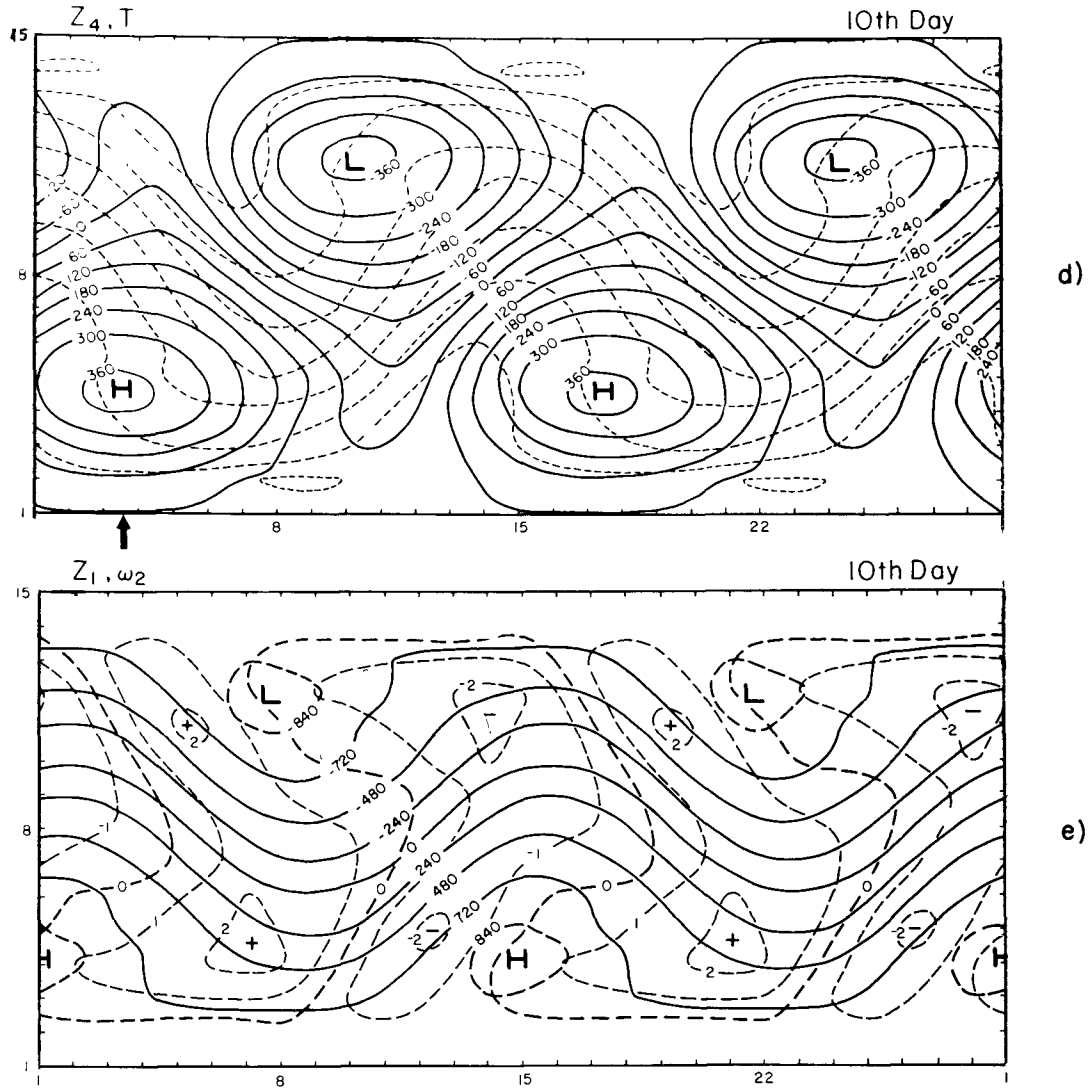
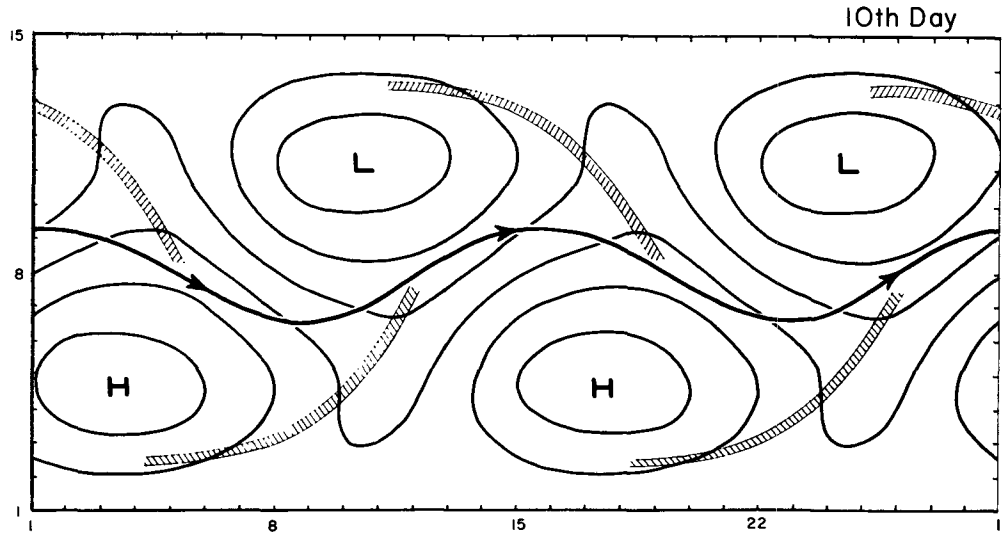


Figure 7. Non-linear evolution of the planetary circulation when $(-\partial\tilde{T}_2/\partial\varphi)>\kappa T_2 s$

a to e) Horizontal maps on which Z_4 and Z_1 (solid lines) are the height contours of the 1000 mb and 250 mb pressure surfaces (as deviations from the mean height of the surface, in meters); T (broken line) is the temperature at any level (as deviation from the mean temperature of the level in intervals of 5C°); and ω_2 (broken line) is the vertical velocity at the 500 mb level (in 10^{-4} cb s^{-1}). (w_2, in cm s^{-1}, is approximately equal to $-1.4 \times 10^4\ \omega_2$.) The arrow shows the successive longitude position of one of the circulation centers, where each scale interval is 444 km.



f)

Figure 7. (Cont.)

f) Summary map for the 10th day. The thin lines are the 1000 mb contours, the shaded bands are the zones of maximum horizontal temperature gradient, and the heavy line is the axis of the wind speed maximum at the 250 mb level.

For the qualitative physical interpretation of the early stages in the evolution of the system, therefore, we can use the simplified relationship

$$\left(\frac{\partial \omega}{\partial p}\right)_4 \sim -\frac{1}{f}\mathbf{v}_T \cdot \nabla\,(2\zeta_4 + 2\zeta_T + f). \tag{20}$$

Equation (20) is the well known Sutcliffe development criterion, and we will use it, in the manner of Sutcliffe (1947, 1950), as a simplified diagnostic tool to interpret the nonlinear numerical integration.

Applying equation (20) to the centers of the disturbance after the initial deformation of the isotherms, we see that $-\mathbf{v}_T \cdot \nabla\,(2\zeta_T + f) > 0$ at the centers of the Lows, and consequently there is horizontal convergence (vertical stretching) at the centers of the Lows. At the centers of the Highs, $-\mathbf{v}_T \cdot \nabla\,(2\zeta_T + f) < 0$, and there is horizontal divergence. As a result, by equation (7), the disturbance intensifies. This deforms the isotherms at an increasing rate, and, in turn, with the increase of $\nabla\zeta_T$, the disturbance grows at an increasing rate, as shown by the change from the 4th to the 8th day.

The first term on the right in equation (20), $-\mathbf{v}_T \cdot \nabla\,(2\zeta_4)$, does not contribute to the intensification of the disturbance, but it causes the Low and High centers to move along the isotherms in the direction of $\mathbf{v}_T$. This produces not only an eastward but also a poleward displacement of the Low centers and an equatorward displacement of the High centers, already visible by the end of the 4th day. As a consequence of the latitudinal displacements, the zonally averaged zonal surface winds, $\tilde{u}_4$, which initially were zero in all latitudes, become easterly in the low and high latitudes and westerly in the central latitudes.

The component of vertical stretching which produces the northeastward and southeastward displacement of the surface Low and High centers is one in which there is ascending motion northeast of the Low centers and northwest of the High centers, and descending motion southwest of the Low centers and southeast of the High centers. Consequently, throughout the stage of latitudinal displacements of the centers, the zonally averaged vertical velocity is one of ascending air north of the central latitude and descending air south of the central latitude.

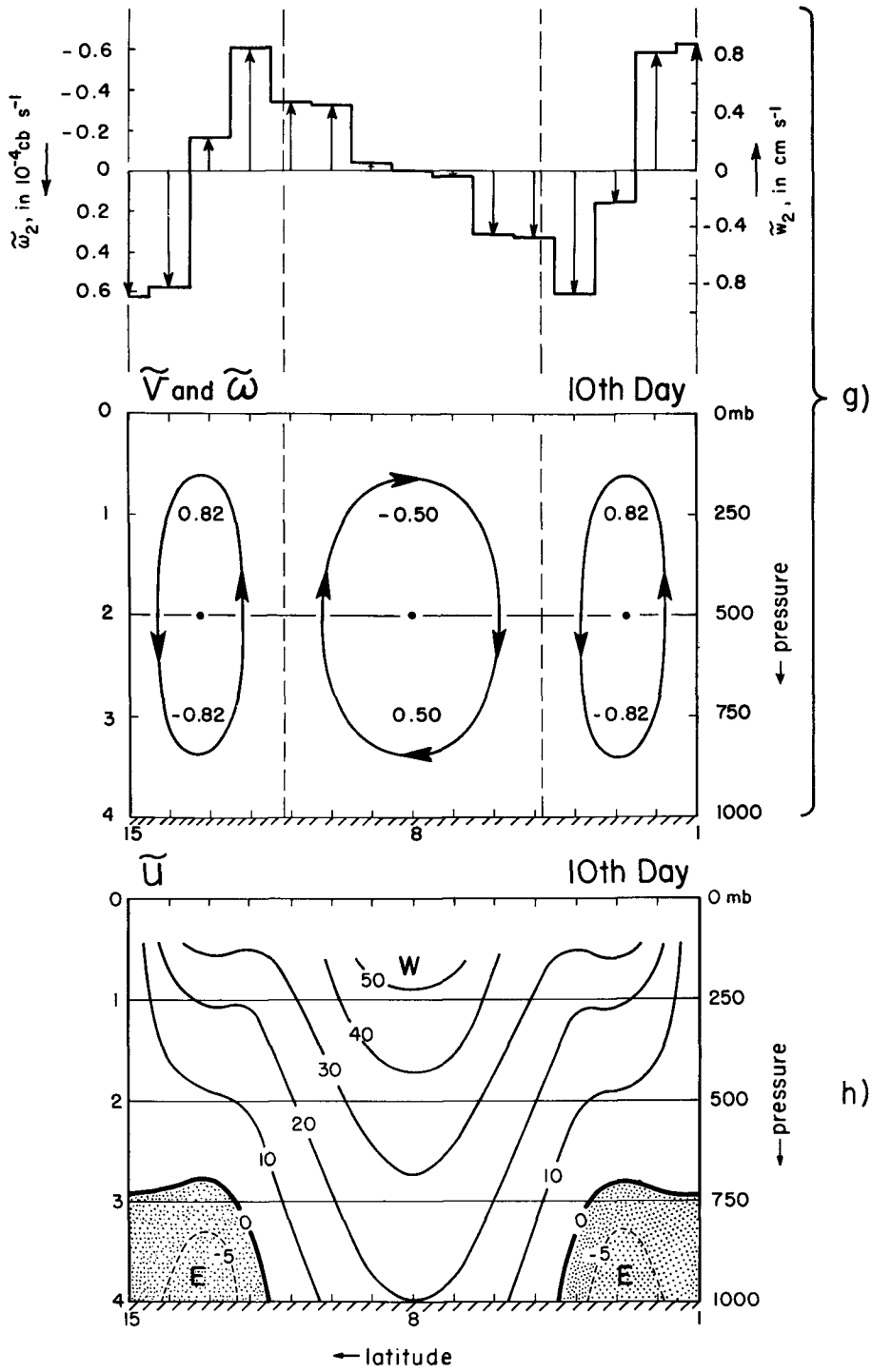


Figure 7. (Cont.)

g) The zonally-averaged vertical velocity and meridional circulation on the 10th day. The numbers on the cross-section give the maximum and minimum values of the zonally-averaged meridional wind $\bar{v}$ (in $m\ s^{-1}$)

h) Cross-section of the zonally-averaged zonal wind $\bar{u}$ (in $m\ s^{-1}$), on the 10th day.

By the 10th day, the intensity of the disturbance has grown very large and the latitudinal displacements of the Low and High centers have become large, with the result that the mean zonal surface winds are of considerable intensity, as may be seen in part *h* of Fig. 7.

The horizontal and vertical advection of potential temperature by the disturbance brings the isotherms together in the form of discontinuous zones of maximum horizontal temperature gradient, as shown on the map for the 10th day. These zones of maximum baroclinicity are not true frontal zones because, with the temperature lapse rate constant in x and y, the zones have no vertical inclination. The effect of the zones, however, is the same as that of true frontal zones in that they contribute to a meandering west-wind maximum or "jet stream" at high levels, as shown by the streamfunction field at 250 mb in part *e* of Fig. 7.

At the same time that the isotherms become crowded together in the pattern described, the latitudinal displacement of the isotherms reduces the zonally averaged temperature difference between the southern and northern boundaries. The initial difference of 48C° is reduced to 36 degrees by the 10th day. This is reflected in the reduction of the vertical wind shear, $(\widetilde{u}_1 - \widetilde{u}_3)$, from the initial value of 28 $m\,s^{-1}$ in all latitudes (except near the boundaries) to 21 $m\,s^{-1}$ at the central latitude, and to 25 $m\,s^{-1}$ at the latitudes of maximum surface easterly winds. But $(\widetilde{u}_1 + \widetilde{u}_3)/2$, which is the vertically averaged wind, increases in middle latitudes (from 22 to 38 $m\,s^{-1}$ at the central latitude), and decreases in the lower and higher latitudes (from 22 to 10 $m\,s^{-1}$ at the latitudes of maximum surface easterly winds).

By the 10th day, the neglected terms of equation (20) are large and Sutcliffe's criterion must now be used with caution when interpreting the divergence and vertical velocity fields. The numerical calculation, which includes all the terms, shows that near the boundaries the vertical velocities are small but of the same sign in all longitudes. Farther in from the boundaries the vertical velocities are larger in magnitude but change sign with longitude. The ascending motion east of the Low center covers a larger area than the descending motion west of the Low center (and there is a corresponding relationship for the High center). Averaged over longitude, the vertical velocity on the 10th day has the distribution shown in part *g* of Fig. 7. Applying the equation of mass continuity and the boundary conditions, we obtain the three-cell mean meridional circulation shown in the figure. This is a three-cell mean meridional circulation, with reversed center cell, produced dynamically without heating or friction.

Part *f* of Fig. 7 shows the 1000 mb contours, the zones of maximum horizontal temperature gradient, and the axis of the wind-speed maximum at the 250 mb level, on the 10th day. Because the surface winds are so strong, the axis of the upper-level wind maximum shows a large displacement from the zones of maximum horizontal temperature gradient. We expect that when surface friction is used in the model the surface winds will be weaker and therefore the upper level wind maximum, or "jet stream," will lie closer to the "fronts".

There are several important ways in which the oceans influence the general circulation of the atmosphere of the Earth. One of these, which was taken into account implicitly in our theoretical calculations, is that evaporation from the oceans keeps the atmosphere in a state of near saturation. With its characteristic temperatures, the order of magnitude of which is given by the effective insolation $S(1-A)$, near-saturation means sufficient water vapor in the Earth's atmosphere to make the infrared emission to space almost constant with latitude. As a result, the differential heating ΔQ exceeds the critical value for a stable symmetric circulation and the wave regime is produced.

A second important effect of water in the Earth's atmosphere was also taken into account implicitly in our theoretical model; namely, that once the static stability s is reduced to its new value s^* by condensation in the waves, the heat of condensation thereafter keeps s^* constant. In the wave regime there is a positive correlation between the ascending motion, $-\omega$, and the temperature, T,

and this produces a net upward transport of heat which, by itself, would make the space-averaged vertical gradient of temperature more stable. But the heat of condensation is almost entirely released in the lower half of the atmosphere (in contrast to the principal cooling term, the divergence of the net vertical infrared radiative flux, which is relatively constant with elevation); and because the heat of condensation is released in the lower half of the atmosphere it opposes the reduction of the space-averaged vertical temperature gradient by the upward heat transport of the waves. In this way, s^* is maintained relatively constant in value. As we shall see in the section on Mars, this property of the Earth's atmosphere, of maintaining its average static stability quasi-constant during the wave regime, will not be present in a planetary wave regime which has no condensation, and this fact has several important consequences.

Besides the fact that evaporation from the oceans keeps the Earth's atmosphere in a state of near-saturation and thereby gives it the properties described above, there is an important influence of the oceans through their seasonal storage of insolation. Very little insolation is absorbed by the Earth's atmosphere directly. By far the larger part of the effective insolation, $S(1—A)$, is absorbed by the underlying surface, whether the surface be land or sea. Because land has a small thermal conductivity, especially when the land is dry, very little heat can be stored in the land as the season changes. As a consequence, if our planet had no oceans at all the sign of ΔQ would change with the season; at the summer equinox, when the insolation has its maximum in the polar region because of the great length of the day there, the sign of ΔQ would be negative. In contrast, the upper layer of the oceans has a high thermal eddy-conductivity and produces a large seasonal storage of insolation. The temperature of the ocean surface therefore changes very little with the season and, consequently, the heat transfer from ocean to overlying atmosphere has relatively little seasonal variation. On a planet that was completely ocean covered, therefore, ΔQ would have only a small seasonal variation. On the true Earth there is apparently sufficient ocean area, even in the northern hemisphere, for ΔQ to maintain its positive sign and to exceed ΔQ_{crit} throughout the year, and thereby maintain the normal wave regime throughout the year.

As a consequence of the storage process and the limited longitudinal extent of the oceans, there is another important effect. This is the effect of the seasonal storage of insolation by the oceans in producing large zonal variations in the heating of the atmosphere. In the winter of the northern hemisphere, for example, in middle and high latitudes there is strong relative cooling of the air over the continents and heating over the oceans as the air of the troposphere flows from west to east. Over the continents the air cools because $S(1—A)$ is smaller than W; while the temperature of the ground adjusts itself to the temperature of the air because the thermal conductivity of the ground is small. But when the air that has been cooled over a continent moves over the ocean again, it is heated because it is colder than the warm ocean surface; while the temperature of the ocean surface remains almost unchanged because the thermal eddy-conductivity of the oceans is large. There are two important consequences of this zonally varying heating of the atmosphere. One is that low-level convergence is produced in the longitudes where there is heating of the air and low-level divergence when there is cooling [the exact form of this effect may be seen by deriving equation (19) with the heating and friction terms retained]. As a result, in the winter season there will be general low-level horizontal convergence over the oceans and divergence over the continents, and, therefore, lower mean surface pressure over the oceans than over the continents. In the summer season, the reverse pattern of heating and cooling and reverse distribution of high and low surface pressures over oceans and continents will occur. At the same time, the heating establishes in the upper levels an ultra-long-wave pattern of opposite phase to the surface pressure distribution. The second important consequence of the zonally varying heating due to the ocean-land contrast is that it affects the horizontal temperature gradients, and hence the baroclinic stability, at the coastlines. Especially near the eastern coastlines of the two northern hemisphere continents in the winter season, the contrast between cold continental air and warm maritime air produces a zone of large horizontal temperature gradient and hence a zone of maximum baroclinic instability. Thus the eastern coastlines become preferred regions of baroclinic wave

development. And as a baroclinic disturbance moves eastward from its source region of maximum horizontal temperature gradient and moves through the region of maximum heating (and therefore surface convergence) of the ultra-long atmospheric wave, the low-level cyclonic part of the moving disturbance is amplified. In this way the baroclinic wave-cyclones, instead of being uniformly distributed around the globe, are concentrated in the oceanic longitudes, especially in the winter season.

Still another important influence of the ocean lies in its own ability to transport heat. Insofar as the mean latitudinal heat transport by the ocean is concerned, this is indirectly estimated to amount to only about 1×10^{19} cal day^{-1} at the central latitude of the northern hemisphere, and hence it has no controlling influence on the form of the atmospheric general circulation. Reducing the required poleward heat transport, from 10.6 to 9.6×10^{19} cal day^{-1}, would not change the equilibrium circulation from the wave to the symmetric regime. But important aspects of the atmospheric circulation, within the wave regime, can be influenced by the oceans' transport of heat. As a hypothetical example, to illustrate this point, let us imagine a long period of negative anomaly of cloudiness in a region of downwelling ocean current. The anomalous excess of insolation, stored in the water and carried downward, might appear at the surface again only in some distant part of the ocean and after some long interval of time. It is conceivable that this anomalous heat, when brought to the ocean surface and transferred to the air, could affect the phase of the ultra-long atmospheric waves. Because of the interaction between the ultra-long-waves and the baroclinic disturbances—in which there may be a sensitive three-way adjustment between the long waves, surface heat sources, and baroclinic disturbances—there could be a shift of the tracks of the rain producing baroclinic disturbances, east or west, resulting in long-period weather anomalies over extensive regions of the globe. The point that is being made, through this hypothetical example, is that the atmosphere's small heat capacity and large velocities gives the atmosphere a "short memory" for anomalies of insolation; but the ocean, with its large heat capacity and small velocities, has a "long memory" for anomalies of insolation.

Any rational theory of long-period weather changes, and hence any direct physical method of long-period weather prediction, must take the interaction between atmosphere and ocean into account. This two-fluid system, containing the incompressible as well as the compressible part of the Earth's fluid envelope, is highly non-linear in its behavior. But the general physical laws which govern the two fluids and their interactions are, to a considerable degree, already known. With the expected advances in computer technology, it will be possible to study the atmosphere-ocean system in the thorough integrated way it deserves.

E. Circulation of the Atmosphere of Mars

The planet Mars rotates on its axis with nearly the same angular speed as the Earth, and therefore the coriolis parameters on the two planets are about the same. Also, the axis of Mars has nearly the same inclination to its orbital plane as the Earth, which makes the seasonal variation of the latitudinal gradient of incident solar radiation about the same on the two planets. Mars is about one and one-half times as far from the Sun, and its solar constant is 0.43 times that of the Earth. This is compensated, to a small extent, by the lower albedo of Mars ($A = 0.15$ as against 0.35 for the Earth), and hence the mean effective isolation on Mars [$S(1 - A)$] is about 0.56 times that of the Earth.* From these external parameters alone one would not expect any great differences in the circulation regimes of the two planets.

But, as we shall see, the absence of oceans on Mars and the lack of any substantial amount of water vapor in its atmosphere have these important consequences: (1) the outgoing infrared radiation W is not constant with latitude on Mars but decreases from equator to pole, thereby making the differential heating ΔQ smaller than ΔQ_{crit} in the mean for the year. (2) The absence of a seasonal heat storage by oceans makes the seasonal change of ΔQ very large, so that in the winter season ΔQ

* de Vaucouleurs (1961) has recently given a new value for the integrated albedo of Mars of $A = 0.26 \pm 0.02$. This will reduce the mean effective insolation on Mars to about 0.49 times that of the Earth.

exceeds ΔQ_{crit}, but in summer ΔQ is *negative* (and the poleward temperature gradient is reversed). (3) The large seasonal variations of ΔQ, and the absence of any substantial heat of condensation, produce large seasonal variations of the static stability s and the parameter λ, thereby causing a large change of the dominant wave number during the winter season wave regime, possibly reaching $\lambda < 1$ and thereby a reversion to the symmetric regime of circulation during the mid-winter maximum of ΔQ.

On Earth the outgoing radiation W is almost constant with latitude because nearly all of it comes from the water vapor in the air. But on Mars, observation and theory indicate so small an upper limit to the water vapor in the atmosphere that, as a first approximation, the water vapor absorption can be neglected. The only infrared absorbing gas of any consequence in the atmosphere of Mars is carbon dioxide.

The absorption spectrum of carbon dioxide has large "windows" in the infrared, and about 70 per cent of the radiation from the surface of Mars, at the characteristic temperatures of the surface, will pass through the atmosphere directly to space. Consequently, if the temperature of the surface of Mars decreases from equator to pole, the radiation from the surface to space will decrease with latitude. In addition, because the carbon dioxide mixing ratio is constant, the radiation from the atmosphere to space in the carbon dioxide bands will also decrease with latitude when the atmospheric temperature decreases with latitude. As we shall see (cf. Fig. 8), the decrease with latitude of the total outgoing radiation, W, makes the differential heating ΔQ relatively small, and (in the mean for the year) smaller than ΔQ_{crit}.

Again, we can write, for the central latitude,

$$\left(\frac{-\partial\widetilde{T}_2}{\partial\varphi}\right)_{H_s = \Delta Q} = \left(\frac{\sqrt{2}\,\Omega^2 m}{\pi g R^* s \mu_2}\right)\Delta Q \ , \tag{6}$$

where, now

$$\Delta Q = 2\pi a \int_0^{\pi/4} [S(1 - A) - (Q_g + Q_o)]) \cos\varphi d\varphi \ . \tag{21}$$

Q_g is the radiation per unit area per unit time that is emitted by the ground to space directly through the "windows" of the carbon dioxide absorption spectrum. Q_o is the radiation to space that is emitted by the carbon dioxide of the atmosphere. $Q_g + Q_o = W$.

For the absorption spectrum of carbon dioxide we will use the simple approximation of complete absorption between wavelengths 12 to 18μ and 3 to 5μ, and zero absorption in all other wavelengths, for the amount of carbon dioxide in the atmosphere of Mars (3100 cm-atm). With this approximation, $Q_g(T)$ and $Q_o(T)$ have the values given in the following table:

	Temperature in °K					
	160°	180°	200°	220°	240°	260°
Q_g	0.049	0.068	0.096	0.137	0.190	0.257 cal cm^{-2} min^{-1}
Q_o	0.004	0.017	0.034	0.053	0.080	0.115 cal cm^{-2} min^{+1}

We will assume that the mean annual ground temperature varies sinusoidally between equator and pole, or

$$T_g = T_{g,\,\pi/4} + \frac{\Delta T_g}{2} \cos 2\varphi \,, \tag{22}$$

where ΔT_g is the difference in ground temperature between equator and pole, and $T_{g,\,\pi/4}$ is the ground temperature at 45° latitude.

We will also assume that, up to a level Z_o, the vertical gradient of temperature varies with latitude from the adiabatic lapse rate at the equator to the isothermal condition at the pole, or

$$\frac{\partial T}{\partial Z} = \frac{\gamma_d}{2} (1 + \cos 2\varphi) \,. \tag{23}$$

The reasoning behind this assumption about the vertical temperature gradient is that in an atmosphere without appreciable condensation the net heating of the air in low latitudes and the net cooling of the air in high latitudes depends upon the radiative transfer and the conductive-convective transfer of heat between the ground and air. The latitudinal gradient of the net heat transfer between ground and air by both of these processes requires that the vertical temperature gradient become more stable from equator to pole. (On Earth it is the latitudinal gradient of the heat of convective condensation that performs this function, and therefore on Earth it is mainly the "conditional stability," the difference between the saturated adiabatic vertical temperature gradient and the actual vertical temperature gradient $(\gamma_s - \gamma)$, rather than γ itself, which becomes more stable from equator to pole.) But the specification of the particular values of $\gamma = \gamma_d$ at the equator of Mars and $\gamma = 0$ at the pole is an arbitrary one.

With γ dependent on latitude, the poleward temperature gradient vanishes at some level, Z_o. We will assume that above Z_o the lapse rate is isothermal in all latitudes.

Furthermore, we will let the level at which we examine the atmospheric poleward temperature gradient, $(-\partial \widetilde{T}_2/\partial\varphi)$, for dynamic stability be midway between the ground and Z_o, or

$$T_2 = T_g + \frac{Z_o}{2}\left(\frac{\partial T}{\partial Z}\right) \tag{24}$$

And finally, for simplicity, because it will have no sizeable effect on the computed value of ΔQ, we will assume that all of the radiation emitted by the carbon dioxide bands directly to space comes from the levels above Z_o, or

$$Q_o(T) \equiv Q_o(T_o) \,, \tag{25}$$

where T_o is the temperature of the isothermal region above Z_o.

From the numerical solution of equations, (6), (21), (22), (23), (24) and (25) (and the table of radiation intensities) we can derive the thermal equilibrium temperatures as a function of latitude and elevation for the symmetric circulation regime, $\widetilde{T}(\varphi, Z)_{H_s=\Delta Q}$, and thereby find $(-\partial \widetilde{T}_2/\partial\varphi)_{H_s=\Delta Q}$ as well as $Q_g(\varphi)$, Q_o and ΔQ.

Using $\Omega = 7.1 \times 10^{-5}\ s^{-1}$, $a = 3.4 \times 10^6$ m, $g = 3.8$ m s^{-2}, $\mu_2 = \mu_2(\text{Earth}) \approx 2.2 \times 10^{-2}$ t m^{-1} s^{-1}, $A = 0.15$, the known values of $S(\varphi)$ in the mean for the year, and (for Nitrogen) $R^*/m = 297$ and $\gamma_d = -g/c_p = -3.8C^\circ$ km^{-1}, we obtain the results shown in the Table below and in Fig. 8:

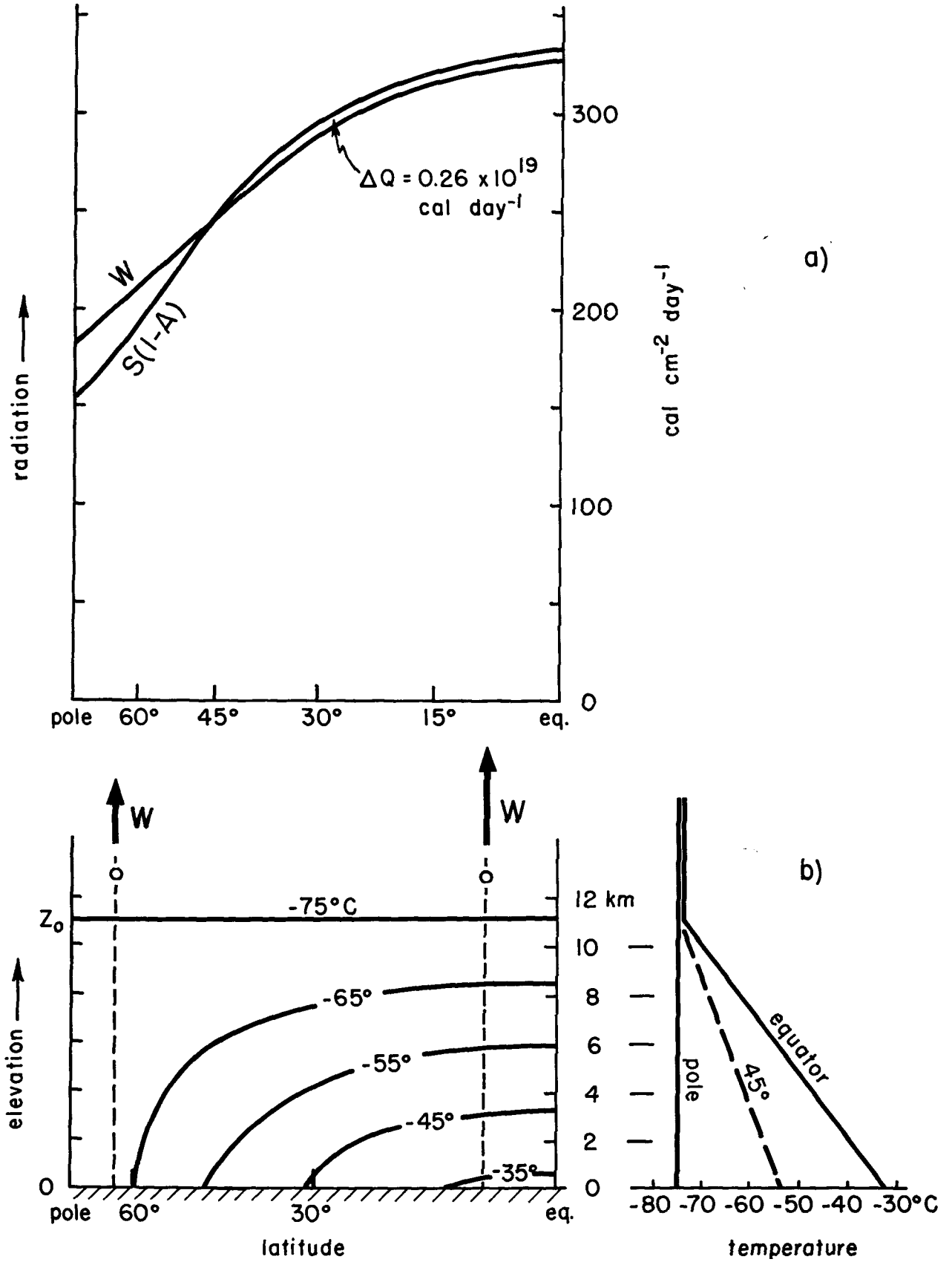


Figure 8. Mean annual radiation budget and temperature field on Mars, as derived from theory

a. Energy gain and loss as a function of latitude, where S is insolation per unit horizontal area per unit time, A is the planetary albedo, W is the infrared emission to space, and ΔQ is the net differential heating.

b. Meridional temperature cross-section; and temperature profiles at the equator, 45° latitude, and pole.

Mean Annual Radiation Budget and Temperatures of Mars

latitude	$S(1-A)$	Q_g	Q_o	$W = Q_g + Q_o$	T_g	T_o
	cal cm^{-2} day^{-1}	cal cm^{-2} day^{-1}	cal cm^{-2} day^{-1}	cal cm^{-2} day^{-1}	°C	°C
0°	332	279	46	325	−33.0°	−75.0°
10	325	274	46	320	−34.3	−75.0
20	315	260	46	306	−37.9	−75.0
30	295	238	46	284	−43.5	−75.0
40	269	210	46	256	−50.4	−75.0
50	232	186	46	232	−57.6	−75.0
60	194	164	46	210	−64.5	−75.0
70	168	148	46	194	−70.1	−75.0
80	158	140	46	186	−73.7	−75.0
90	155	135	46	181	−75.0	−75.0

From $S(1-A)$ and the computed values of W, we obtain

$$\Delta Q = 0.26 \times 10^{19} \text{ cal day}^{-1} ,$$
$$= 0.12 \times 10^{12} \text{ kj s}^{-1} .$$

And, from (17),

$$\Delta Q_{crit} = 0.17 \times 10^{12} \text{ kj s}^{-1} ,$$

for $T_g = 210°K$ and $\kappa = R^*/m\, c_p = 0.296$.

Therefore, in the mean for the year on Mars, $\Delta Q < \Delta Q_{crit}$ and the symmetric circulation will maintain thermal equilibrium and be dynamically stable. Or, expressed in terms of the poleward temperature gradients,

$$\left(\frac{-\partial \widetilde{T}_2}{\partial \varphi}\right)_{H_s = \Delta Q} < \left(\frac{-\partial \widetilde{T}_2}{\partial \varphi}\right)_{crit} ,$$

where, from the above solution,

$$\left(\frac{-\partial \widetilde{T}_2}{\partial \varphi}\right)_{H_s = \Delta Q} = 21C°/\text{rad} ,$$

and, from (17),

$$\left(\frac{-\partial \widetilde{T}_2}{\partial \varphi}\right)_{crit} = 31C°/\text{rad} .$$

The mean surface temperature of Mars, according to this computation, is about −47°C, whereas the blackbody radiative equilibrium temperature is about −56°C for the given mean effective insolation $S(1-A)$. This weak greenhouse-effect of 9 degrees is all that the carbon dioxide can produce because it has such broad windows in its infrared absorption spectrum. [By comparison, the water vapor plus carbon dioxide in the Earth's atmosphere produces a greenhouse-effect increase of the mean surface temperature of the Earth of about 38° (from −23°C to +15°C)]. Our computed mean surface temperature of Mars, of −47°C (which, however, will be reduced to about −55°C if we solve

the above equations with the new albedo value of A = 0.26), is consistent with the measured microwave brightness temperature of Mars, of −62°C ± 28°C (Giordmaine, *et al.*, 1959.)*

It may be worthwhile to digress briefly from our main theme at this point, and to point out that in our model of the Mars atmosphere the stratosphere is not in radiative equilibrium. If the temperature of the troposphere decreases from equator to pole, and if the stratosphere is assumed to have an isothermal lapse rate and to be in radiative equilibrium, then, given *carbon dioxide* as the dominant absorbing gas, the temperature of the stratosphere must *decrease* from equator to pole and the height of the tropopause must *increase* from equator to pole, as shown schematically in Fig. 9. (This is true not only when the tropospheric lapse rate increases in stability from equator to pole, but also, though to a lesser extent, when the tropospheric lapse rate is constant.) The reason for this is that the stratosphere, when in radiative equilibrium, must radiate (in two directions) the same amount of energy that it receives from the troposphere and, when carbon dioxide is the radiating gas, the emission from the troposphere varies with the temperature of the troposphere. By contrast, as the figure shows, when *water vapor* is the dominant radiating gas (and the troposphere has near-constant relative humidity), then the emission from the troposphere is independent of its temperature, and hence the temperature of the stratosphere will be *constant* with latitude and the height of the tropopause will *decrease* with latitude. In both cases, the state produced by pure radiative equilibrium will be altered by vertical convection that originates in the troposphere and penetrates the stratosphere. As shown in the figure, the convection acts to raise the tropopause level. In both cases, the vertical convection, and its effect on the radiatively established tropopause, will be greatest in low latitudes, either because the conditional stability [a function of $(\gamma_s - \gamma)$] is less stable in the low latitudes (as on Earth), or because the actual stability (a function only of γ) is less stable in the low latitudes (as in our model for Mars). Therefore, in the water vapor case, the convection increases the latitudinal tropopause slope established by radiative exchange; whereas in the carbon dioxide case it decreases the slope established by radiation. But these details of tropopause structure, though interesting in themselves, have no large effect on the over-all latitudinal energy budgets of the planetary atmospheres.

Returning to the main theme, we consider now the seasonal variation of ΔQ on Mars, and the seasonal changes of circulation.

On Mars, as well as Earth, relatively little of the effective insolation [S(1 − A)] is absorbed by the atmosphere directly. Instead, most of the effective solar energy is absorbed by the ground. The ground, in turn, transfers energy to the air by infrared radiative exchange and by conductive-convective heat flux (and on Earth, in addition, by evaporation-condensation.) The heating of the air therefore depends on the temperature of the ground.

On Earth, the zonally-averaged temperature of the ground is greatly influenced, in its seasonal variation, by seasonal heat storage in the oceans. As a result, the zonally-averaged temperature of the ground decreases from equator to pole in all seasons of the year, in spite of the fact that in summer the daily mean effective insolation [S(1 − A)] increases from equator to pole. It is for this reason that the differential heating of the atmosphere, on Earth, although decreasing in magnitude by about a factor of 4 from winter to summer, nonetheless maintains its positive sign throughout the year.

But on Mars there are no oceans, and the thermal conductivity of the land is very small. The seasonal storage of insolation is therefore very small on Mars. As a consequence, the latitudinal gradient of ground temperature and the differential heating of the air, ΔQ, should both vary seasonally on Mars in the same way as the effective insolation [S(1 − A)].

The heavy line, in Fig. 10, shows the daily mean effective insolation on Mars, [S(1 − A)], at the time of the northern hemisphere winter solstice, when A = 0.26. From the pole to 65° latitude

* Although only the daytime hemisphere of Mars can be seen from the Earth, the microwave (three-centimeter-wavelength) radiation is emitted from an average depth of several centimeters below the surface. The measured microwave brightness temperature therefore represents, approximately, the *mean diurnal* surface temperature, and not the daytime surface temperature of the planet.

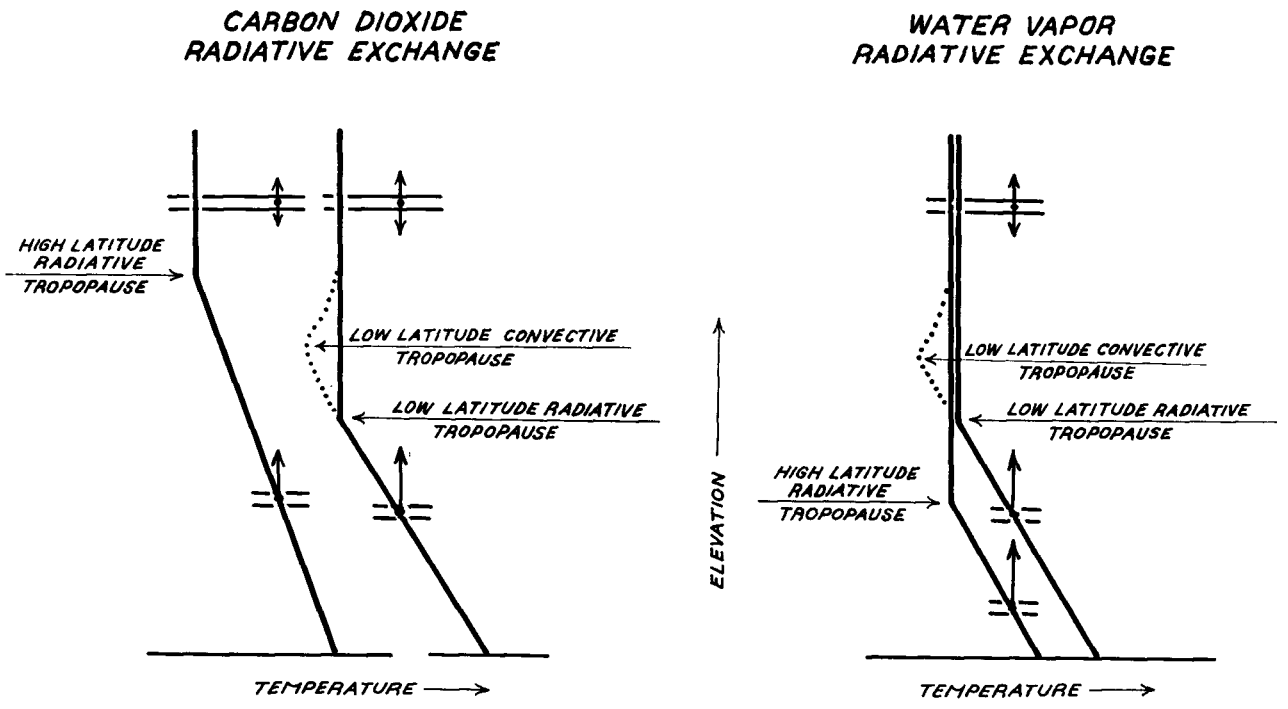


Figure 9. Temperature of the stratosphere and height of the tropopause when the radiative exchange is by carbon dioxide or by water vapor.

The solid lines show an isothermal stratosphere in radiative equilibrium. The upper pairs of arrows represent total radiation emitted by the stratosphere. The lower arrows represent that part of the radiation from the troposphere which the stratosphere absorbs. The dotted lines show the modification of the radiative tropopause by vertical convection, caused either by small absolute stability or small conditional stability at low latitudes.

of the winter hemisphere there is no insolation, but beginning with this latitude the mean daily insolation rises to its maximum at the pole of the summer hemisphere. By contrast, the local noontime insolation has its maximum at 25° latitude of the summer hemisphere; while from 65° latitude to the pole of the summer hemisphere there is some insolation even at midnight.

Because of the low thermal conductivity of dry soil, we can expect that the surface temperature of Mars, in a general way, will follow the effective insolation [S(1 − A)] both seasonally and diurnally. In Fig. 11 an attempt is made to give a first-order estimate of these temperatures, at the time of the northern hemisphere winter solstice, based mainly on radiometric observations of the planet. A theoretical derivation of these temperatures cannot now be given because part of the energy budget consists of the heat transported by the atmosphere. Without the atmospheric heat transport, for example, the temperature at the winter pole would fall close to absolute zero during the long polar night. Therefore, until the atmospheric heat transport, including the large cross-equatorial heat flux, can be computed, we are restricted to the empirical method.

In Fig. 11, the daily mean temperature for the planet as a whole is taken as of the order of −55°C, which, we saw, is the theoretical yearly average surface temperature (for the new albedo value of A = 0.26). This temperature, we remember, is consistent with the measured microwave brightness temperature. The noon maximum temperature on the summer side of the equator, +25°C, is taken from the recent Palomar radiometric observations of Sinton and Strong (1960); and the *temperature difference* between the equator and the summer pole, along the noon meridian, ΔT_g = 45°C, is taken from the older Lowell Observatory "water cell" radiometric scans, summarized by

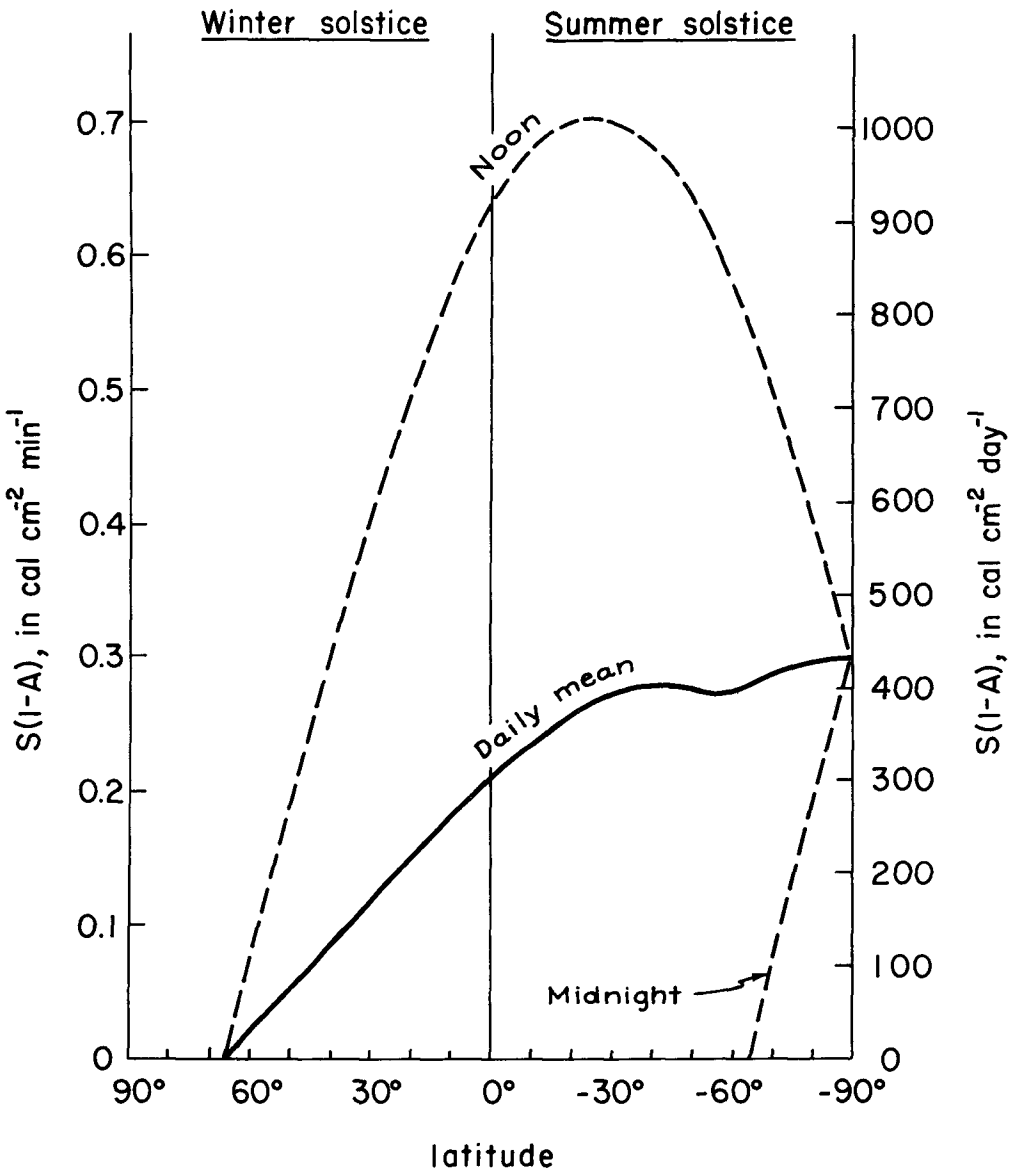


Figure 10. The Effective insolation [S(1 — A)] on Mars, at the northern hemisphere winter solstice

S is the insolation per unit horizontal area per unit time, and $A = 0.26$ is the planetary albedo.

Gifford (1956).* But inasmuch as the insolation has no diurnal variation at the summer solstice pole, the noon surface temperature at the pole, found in this way, −20°C, is also the daily mean surface temperature. Using this value of the surface temperature at the pole, −20°C; the approximate daily mean surface temperature on the summer side of the equator, of about −40°C, from the radiometric measurements of Sinton and Strong (1960); and the planetary average surface temperature, −55°C; the heavy curve of Fig. 11 was sketched freehand to represent a first-order estimate of the daily mean surface temperatures. The sunrise minimum temperature on the summer side of the equator, of about −80°C, indicated by the measurements of Sinton and Strong (1960); plus the fact that the minimum temperature curve, the maximum temperature curve, and the daily mean tem-

* The scans, summarized by Gifford, were taken along the central meridian of Mars, during opposition, and therefore show only noontime temperatures. Because the radiation measured is in wavelengths to which Mars' atmosphere and the Earth's atmosphere are transparent, it is the ground temperature of Mars that is measured.

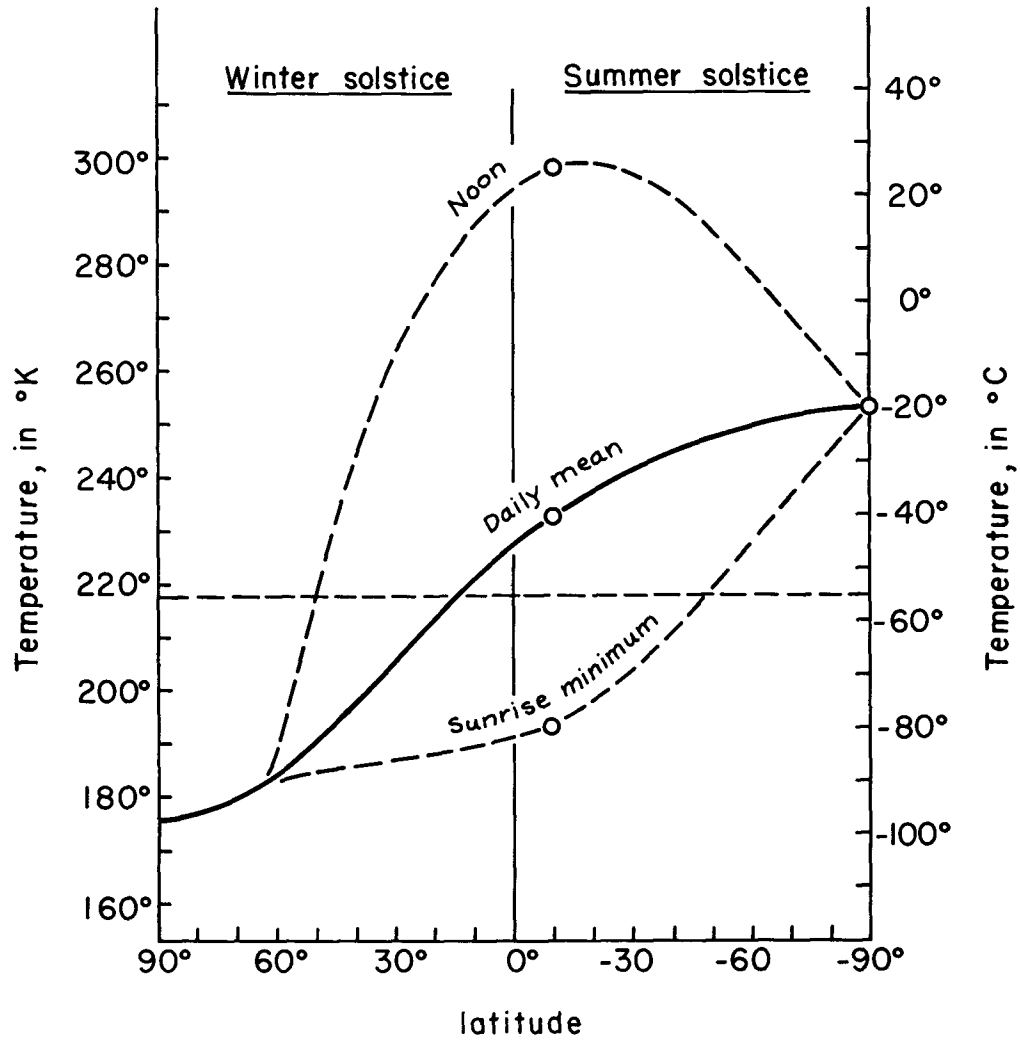


Figure 11. The surface temperature on Mars at the northern hemisphere winter solstice
These temperatures were derived from the observed temperatures near the equator and the observed temperature variation along the noon meridian.

perature curve must all meet at the winter "arctic circle" of zero diurnal insolation, as well as meet at the summer pole of constant diurnal insolation; provide the information necessary to complete the temperature estimates. Although the three insolation curves of Fig. 10 were not used to derive the three temperature curves of Fig. 11, we see that these temperature curves are qualitatively consistent with the given insolation intensities and a ground surface of low thermal conductivity.

According to these empirically derived surface temperatures of Mars at the solstice, the daily mean surface temperature increases all the way from the winter to the summer pole. We expect, therefore, that the daily mean air temperature will also increase from the winter to the summer pole and, as a consequence, the zonal thermal wind shear will be directed from west to east in the winter hemisphere, and from *east to west* in the summer hemisphere. Therefore, the daily mean winds of the middle and upper troposphere, at least in middle and high latitudes, should be westerly in the winter hemisphere and *easterly* in the summer hemisphere of Mars, at the time of the solstice.†

† Essentially the same conclusion about the latitudinal temperature gradient at the solstice was reached by Miyamoto (1960) and related by him " . . . to the predominance of easterlies over middle latitudes in sourthern summer deduced from the drift of great yellow cloud observed in 1958." He also recognized the necessity for a large cross-equatorial heat transport at the solstice.

The large diurnal temperature change, of about 100C° near the equator, is at the ground surface. By terrestrial analogy, we expect that at a few centimeters above the surface the diurnal variation of the air temperature will be of the order of half of this, with a further decrease in amplitude with elevation. But whereas on Earth the diurnal variation of air temperature almost vanishes at one to two kilometers elevation, the smaller mass of the Mars atmosphere may result in a sizeable diurnal temperature wave through the greater part of the troposphere. This could produce a diurnal period, or thermal tide, in the general tropospheric winds on Mars of considerable amplitude.

We will now use the empirically derived daily mean surface temperature to obtain the latitudinal radiation budget and differential heating ΔQ and, from this, the form of the general circulation on Mars at the winter and summer solstice.

The broken line, in Fig. 12, shows the radiation from the surface of Mars directly to space, Q_g, obtained from the daily mean surface temperature at the solstice, shown in Fig. 11, and the table of radiation intensities $Q_g(T)$ given at the beginning of this section. Assuming again, for simplicity, that all of the radiation from the carbon dioxide bands, Q_o, comes from the stratosphere, and that the stratosphere has a constant temperature equal to the theoretically determined mean annual value of $-75°C$, we have $Q_o \equiv 46$ cal cm^{-2} day^{-1}. The sum of the two radiations, $Q_g + Q_o = W$, is shown by one of the solid lines in Fig. 12. The other solid line shows the mean daily effective insolation at the solstice $[S(1 - A)]$.

The difference between $[S(1 - A)]$ and $(Q_g + Q_o)$, integrated from either pole to the latitude φ, gives the net heating over that region and, assuming thermal equilibrium, the poleward heat transport $H(\varphi)$. In Fig. 13 the solid line shows $H(\varphi)$ at the solstice. In the winter hemisphere the heat transport is poleward $(H > 0)$; in the summer hemisphere it is equatorward $(H < 0)$. The maximum flux, 1.96×10^{19} cal day^{-1}, is poleward at latitude 15° on the winter side of the equator. At 45° latitude in the winter hemisphere, $H = \Delta Q = 1.36 \times 10^{19}$ cal day^{-1}. At 45° latitude in the summer hemisphere, $H = \Delta Q = -0.44 \times 10^{19}$ cal day^{-1}. The mean of these two values, 0.46×10^{19} cal day^{-1}, is almost twice as large as the mean annual ΔQ previously obtained, 0.26×10^{19} cal day^{-1}. But considering that the latter was theoretically derived, whereas the solstice values of ΔQ come from the reduction of radiometric measurements of noon surface temperatures on the planet, the discrepancy is not extreme.

In order to understand how the large seasonal change of ΔQ on Mars affects the form of the general circulation, we must first derive a relationship between ΔQ and the static stability s, and hence between ΔQ and λ.

We notice, first, that in the wave regime of circulation the upward heat transport by the waves is proportional to the poleward heat transport. This may be seen by expressing the surface convergence $(\partial\omega/\partial p)_4$, in Eq. (20), in terms of ω_2, multiplying both sides of the equation by the temperature T, and taking the zonal mean for a sinusoidal disturbance at the central latitude (where we can neglect the y-derivatives of the velocity components), which gives us

$$-\widetilde{\omega_2 T_2} \propto u_T \widetilde{v_4 T_4} \ .$$

But as the poleward heat transport $\widetilde{vT}$ is constant with height, when the wind is geostrophic and the vertical temperature gradient is constant in x; and as, in addition, in the state of thermal equilibrium the total poleward heat transport over all elevations is equal to ΔQ, then

$$\begin{aligned} -\widetilde{\omega_2 T_2} &\propto \widetilde{u_T vT} \ , \\ &\propto u_T \Delta Q \ . \end{aligned} \tag{26}$$

Thus, the large scale vertical heat transport by the waves, at the central latitude, is proportional to $u_T \Delta Q$.

Now, as we indicated earlier, the vertical lapse rate of temperature is determined (when there is no condensation) by the complex processes of radiative exchange, small-scale vertical convection, and

the large-scale vertical transport of heat $-\widetilde{\omega_2 T_2}$. *But radiative exchange and small-scale vertical convection are themselves functions of the vertical temperature gradient.* It seems reasonable to suppose, therefore, that the vertical temperature gradient, and hence the static stability s, will depend strongly, in some monotonic fashion, on the large-scale vertical heat transport, $-\widetilde{\omega_2 T_2}$. If we assume, for simplicity, that the vertical static stability would have some positive value, s_0, as a result of radiation and small-scale convection acting alone, and that any change from this value depends upon the large-scale vertical heat transport, we obtain

$$\begin{aligned} s &= s_0 + s_1(-\widetilde{\omega_2 T_2}) , \\ &= s_0 + s_2(u_T \widetilde{vT}) , \\ &= s_0 + s_3(u_T \Delta Q) . \end{aligned} \tag{27}$$

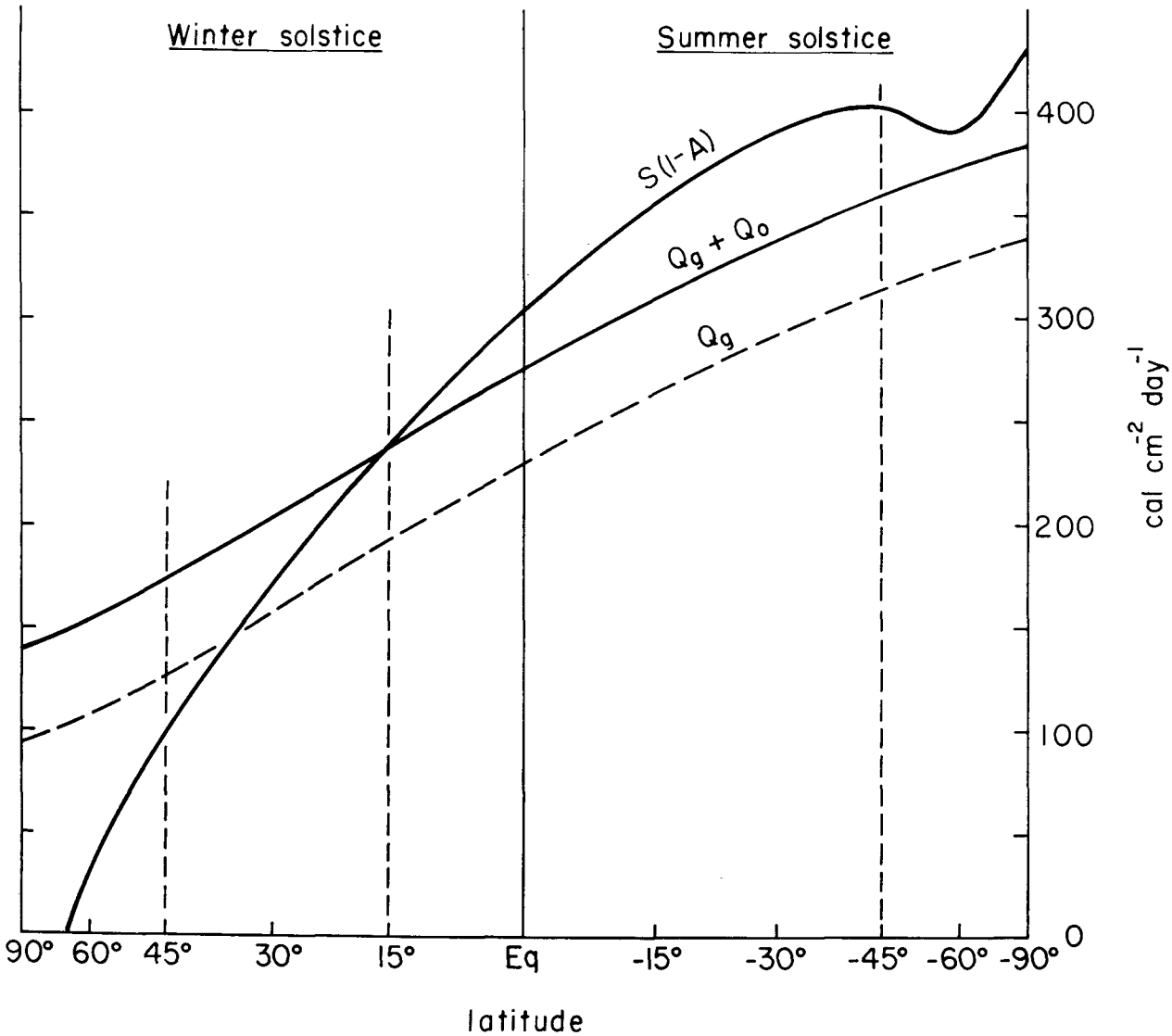


Figure 12. The radiation budget on Mars at the northern hemisphere winter solstice

S(1 − A) is the effective insolation (for A = 0.26). Q_g is the radiation per unit area per unit time emitted by the ground surface directly to space. Q_o is the radiation to space from the carbon dioxide bands.

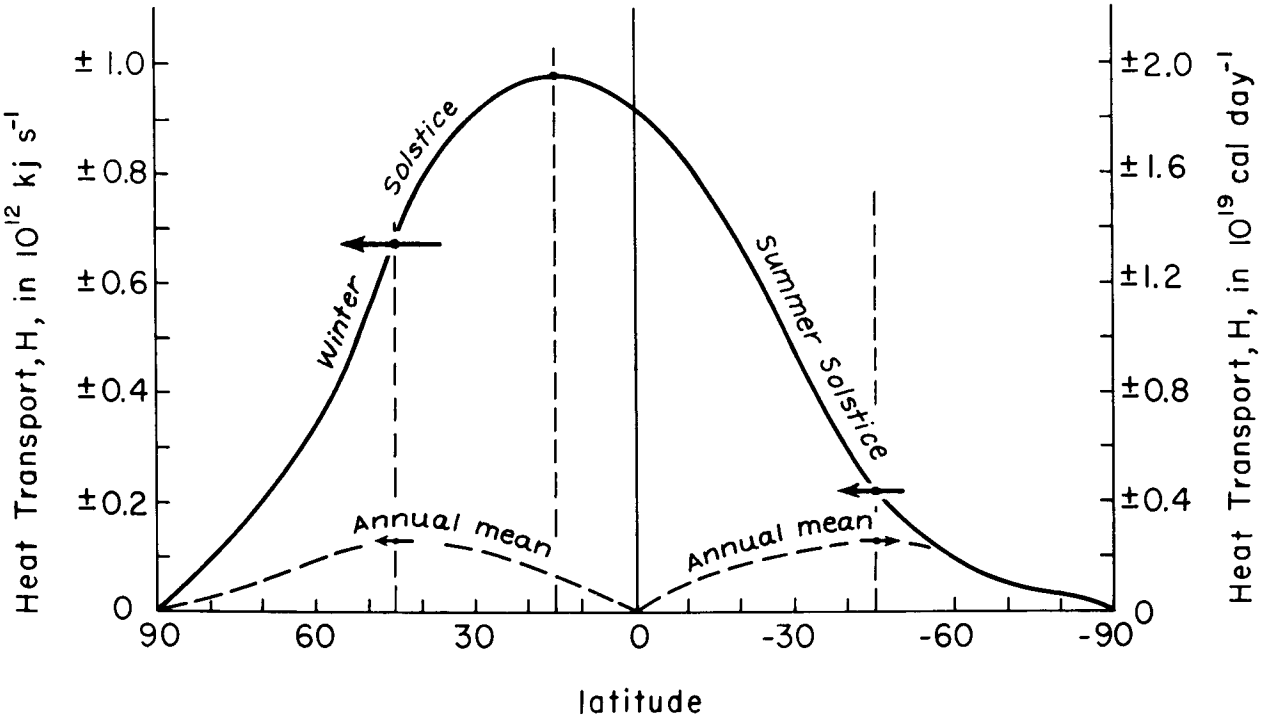


Figure 13. The poleward heat transport on Mars

The solid line shows the transport at the northern hemisphere winter solstice. The broken line shows the theoretically derived annual mean transport.

But inasmuch as in the wave regime the actual zonal thermal wind shear, u_T, is always approximately equal to its critical value, given implicity by Eq. (15), we can write

$$s = s_0 + s(\Delta Q) , \tag{28}$$

where $s(\Delta Q)$ is some monotonic function of ΔQ. This function, $s(\Delta Q) = s_1(-\widetilde{\omega_2 T_2})$, includes the effects of vertical heat transfers by radiation and small-scale convection *implicitly*.

The present state of our knowledge about vertical heat transfer in the atmosphere by radiation and small-scale convection does not enable us to determine the function, $s(\Delta Q) = s_1(-\widetilde{\omega_2 T_2})$, from first principles. We can, however, in an approximate way, obtain this function *empirically* for the atmosphere of the *Earth;* and then, by making a qualitative allowance for the absence of any appreciable heat of condensation in the atmosphere of Mars, we shall be able to see, in a schematic sense, what the consequences are on Mars (as well as what the consequences are in laboratory experiments with rotating fluids that are differentially heated).

For the Earth's atmosphere, we first rewrite (28) as

$$s = s_0 + s_*(\Delta Q) , \tag{28a}$$

where, now, $s^*(\Delta Q)$ is a monotonic function of ΔQ (and hence a monotonic function of $-\widetilde{\omega_2 T_2}$) when condensation takes place in the waves.

For a rough approximation we will assume that

$$s_*(\Delta Q) = C_* \Delta Q , \tag{29}$$

where C_* is a constant.

To determine the constant, C_*, empirically, we substitute from (29) and (28a) into equation (15), and make use of the fact that throughout the year the observed atmosphere of the Earth is in the wave regime. Therefore, both in summer and in winter the actual poleward temperature gradients are approximately equal to the critical poleward temperature gradients, or

$$\frac{1}{\kappa T_2}\left(-\frac{\partial T_2}{\partial \varphi}\right)_{\text{SUMMER}} = s_0 + C_* \Delta Q_{\text{SUMMER}}\,, \tag{30}$$

and

$$\frac{1}{\kappa T_2}\left(-\frac{\partial T_2}{\partial \varphi}\right)_{\text{WINTER}} = s_0 + C_* \Delta Q_{\text{WINTER}}\,. \tag{31}$$

Using the observed quantities:

$$\Delta Q_{\text{SUMMER}} = 2 \times 10^{12}\ \text{kj s}^{-1}\,,$$

$$\Delta Q_{\text{WINTER}} = 8 \times 10^{12}\ \text{kj s}^{-1}\,,$$

$$\left(-\frac{\partial \widetilde{T}_2}{\partial \varphi}\right)_{\text{SUMMER}} = 22\text{C}^\circ/\text{rad}\,,$$

and

$$\left(-\frac{\partial \widetilde{T}_2}{\partial \varphi}\right)_{\text{WINTER}} = 36\text{C}^\circ/\text{rad}\,,$$

we obtain, from (30) and (31),

$$s_0 = 0.234\,,$$

and

$$C_* = 0.031 \times 10^{-12}\ \text{s kj}^{-1}\,.$$

We therefore have $s_{\text{summer}} = 0.30$ and $s_{\text{winter}} = 0.48$, or a mean annual value of static stability of 0.39. This is in good agreement with the limiting values of $s = 0.5$ and $s^* = 0.3$, of section D, which were obtained for the Earth directly from the observed vertical temperature gradient.

The results of this diagnosis of the Earth's atmosphere are summarized in Fig. 14. In this figure, λ, computed from equations (13) and (28a), is shown as a function of $a\Omega$ and ΔQ. The four-fold increase of differential heating, from summer to winter, is accompanied by only a small decrease of λ, and hence only a small decrease of the wave number. This follows from the empirically derived function, $s_1(-\widetilde{\omega_2 T_2}) = s_*(\Delta Q)$. We have already indicated that, in the Earth's atmosphere, the dependence of the static stability on the vertical heat transport should be small, because at the same time that the waves transfer heat from the lower to the upper half of the atmosphere they release latent heat of condensation into the lower half of the atmosphere.

But, on Mars, the characteristic temperatures are so low that there can be no appreciable condensation. Therefore, we can expect that on Mars

$$s(\Delta Q) >> s_*(\Delta Q).$$

Figure 15 is intended to show, in a highly schematic fashion, how λ might vary in the ΔQ and $a\Omega$ plane, when $s(\Delta Q) >> s_*(\Delta Q)$. Also shown, in the figure, are our empirically derived estimates of the seasonal values of ΔQ on Mars.

In Early fall (and late spring), when ΔQ is equal to its mean annual value, ΔQ is smaller than ΔQ_{crit}, according to our theoretical model, and the circulation is in the symmetric regime. As winter comes on, ΔQ increases and the circulation passes into the wave regime. If $s \approx 0.5$ at this time, as

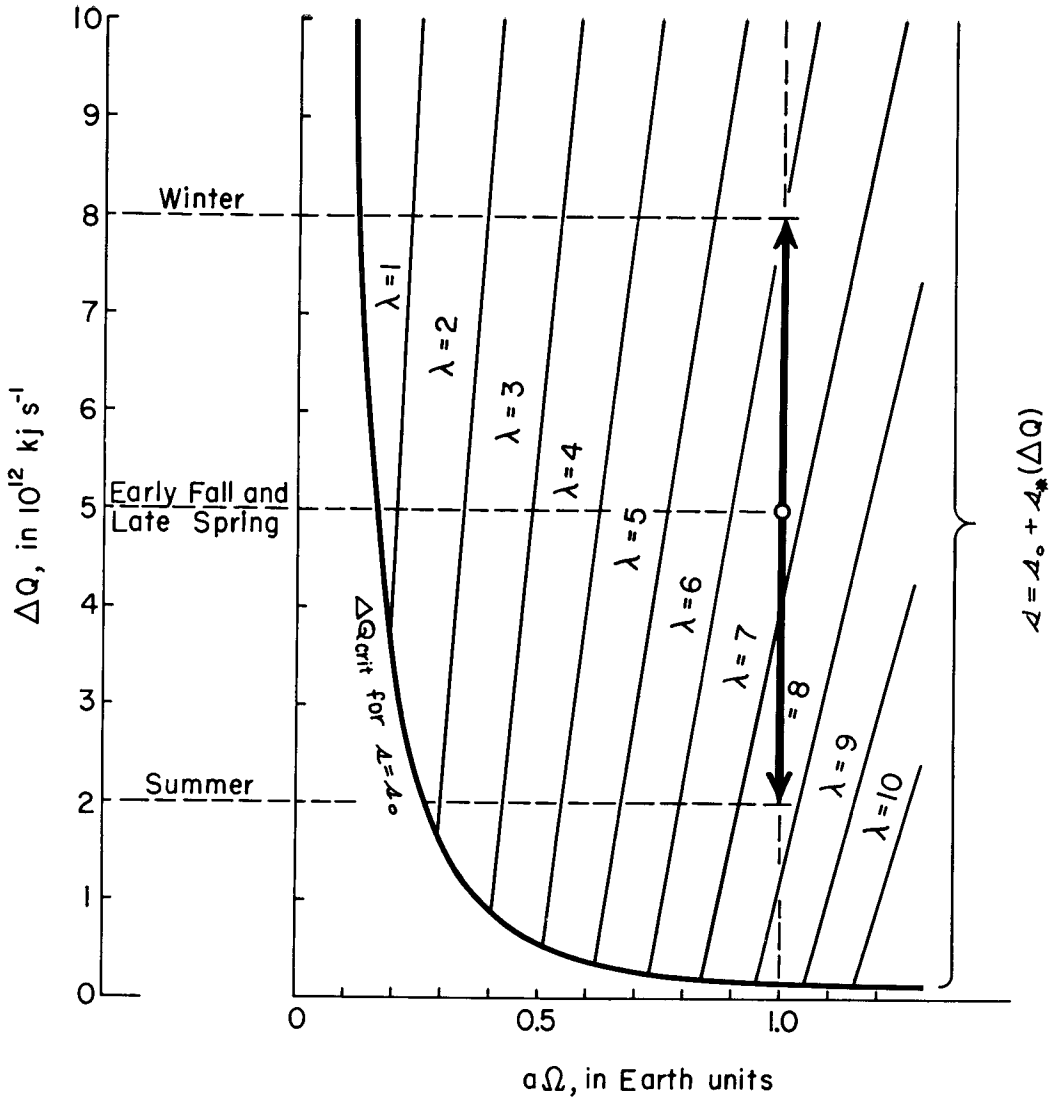


Figure 14. The parameter λ as a function of aΩ and ΔQ

a is the radius and Ω is the speed of rotation of the planet, and ΔQ is the differential atmospheric heating. The dominant wave numbers equals $\lambda/2^{1/4} = 0.84\lambda$, and the maximum wave number equals λ.

assumed for the model, then the dominant wave number is 3. But as ΔQ grows larger, *s* increases and λ decreases, with a consequent reduction of the planetary wave number.

Whether or not λ becomes less than 1 at the time of maximum ΔQ, at the winter solstice, we cannot say until we know the true function $s(\Delta Q)$. But if λ does become less than 1, then the shortest unstable baroclinic wave will be too long for the circumference of the planet, and the circulation will return to the symmetric regime.

As the year proceeds, and ΔQ grows smaller again, the wave number will increase again. But again a change in regime will take place when, in the spring, ΔQ becomes smaller than ΔQ_{crit} and the symmetric regime is reestablished.

With the coming of summer, ΔQ becomes *negative,* and with negative ΔQ there will be a reversed symmetric circulation in the summer hemisphere, and a reversed poleward temperature gradient. The criterion for baroclinic stability, by linear analysis, is independent of the sign of the poleward tempera-

ture gradient, and hence independent of the sign of ΔQ. It is therefore possible that easterly-waves, with their associated low level disturbances, may form in the central latitudes at the time of the summer solstice maximum of $-\Delta Q$.

This schematic treatment of the seasonal changes of circulation regime on Mars is intended mainly to show the kind of changes that can take place on that planet; and which cannot, and do not, take place on Earth because of the water vapor in our atmosphere, the liquid water in our oceans, and our wave-released heat of condensation. Although quantitative studies must still be made to determine the true details of the circulation, it is clear, already, that the circulation on Mars is very different from that of the Earth.

F. Circulation of the Atmosphere of Venus

As we have seen, there are two transitions between the symmetric regime and the wave regime of general circulation. One of these occurs at $\Delta Q = \Delta Q_{crit}$, and at this transition the symmetric regime changes to the wave regime when ΔQ increases. (We are considering, of course, only long term steady-state circulations, and hence only the changes from one steady-state regime to another when ΔQ changes slowly.)

The second transition between symmetric regime and wave regime of circulation occurs when λ becomes less than 1. But λ depends upon s, and s, in an atmosphere without appreciable condensation, depends strongly on ΔQ. As a result, at this second transition the wave regime changes to the symmetric regime when ΔQ increases.

But, by (17), ΔQ_{crit} increases when Ω decreases. And, by (13), λ decreases when Ω decreases. Hence, as shown in Fig. 15, the boundaries between symmetric regime and wave regime approach one another with decreasing Ω. At sufficiently low rotation rates, the circulation will be stable in the symmetric regime for all values of differential heating.

According to some recent radar studies of Venus, the frequency spread of the reflected radar signal indicates that Venus rotates with an angular speed which is about one-tenth the speed of rotation of the Earth, or smaller (v. Chapter III on Venus). As may be seen in either Fig. 14 or 15, with this slow rotation rate the general circulation on Venus must be in the symmetric regime, if we assume that $\mu_2 \sim \mu_2(\text{Earth})$ and $s \sim 0.5$. If Ω is too small for the quasi-geotrophic approximation to be valid, then our criteria for baroclinic instability will not be correct. But as Ω approaches zero the symmetric regime must nonetheless be the stable circulation.

Finally, it is of interest to note that the dependence of regime and wave number upon the heating rate and rate of rotation, as derived in this paper, is in agreement with what is observed in laboratory experiments with differentially heated, rotating fluids.* In these laboratory experiments (both of the simple dishpan type, in which the depth of the fluid is usually less than the radius of the pan; and the annulus type, in which the depth of the fluid, confined between two concentric cylinders, is usually several times its width) (Fultz, 1959; Hide, 1958) it is found that with high rate of heating or low rate of rotation the symmetric regime of circulation is established. But when the heating rate is reduced, or the rotation rate increased, the wave regime is established; with the number of waves increasing as the heating decreases or rotation increases—in qualitative agreement with Fig. 15, where $\Delta Q > \Delta Q_{crit}$.

According to linear baroclinic theory, the *lower* transition from wave regime to symmetric regime depends on the change of coriolis parameter with latitude. But in the dishpan and the annulus, unlike the planetary systems, the coriolis parameter is constant. Therefore in the laboratory experiments, by the linear theory, there should be no *lower* transition to the symmetric regime. Hide (1958),

* Fjortoft (1959, 1960), using a different mathematical technique, has given essentially the same physical explanation for the dependence of wave number on heating rate, and has also derived some of the other properties of the general circulation described in this paper.

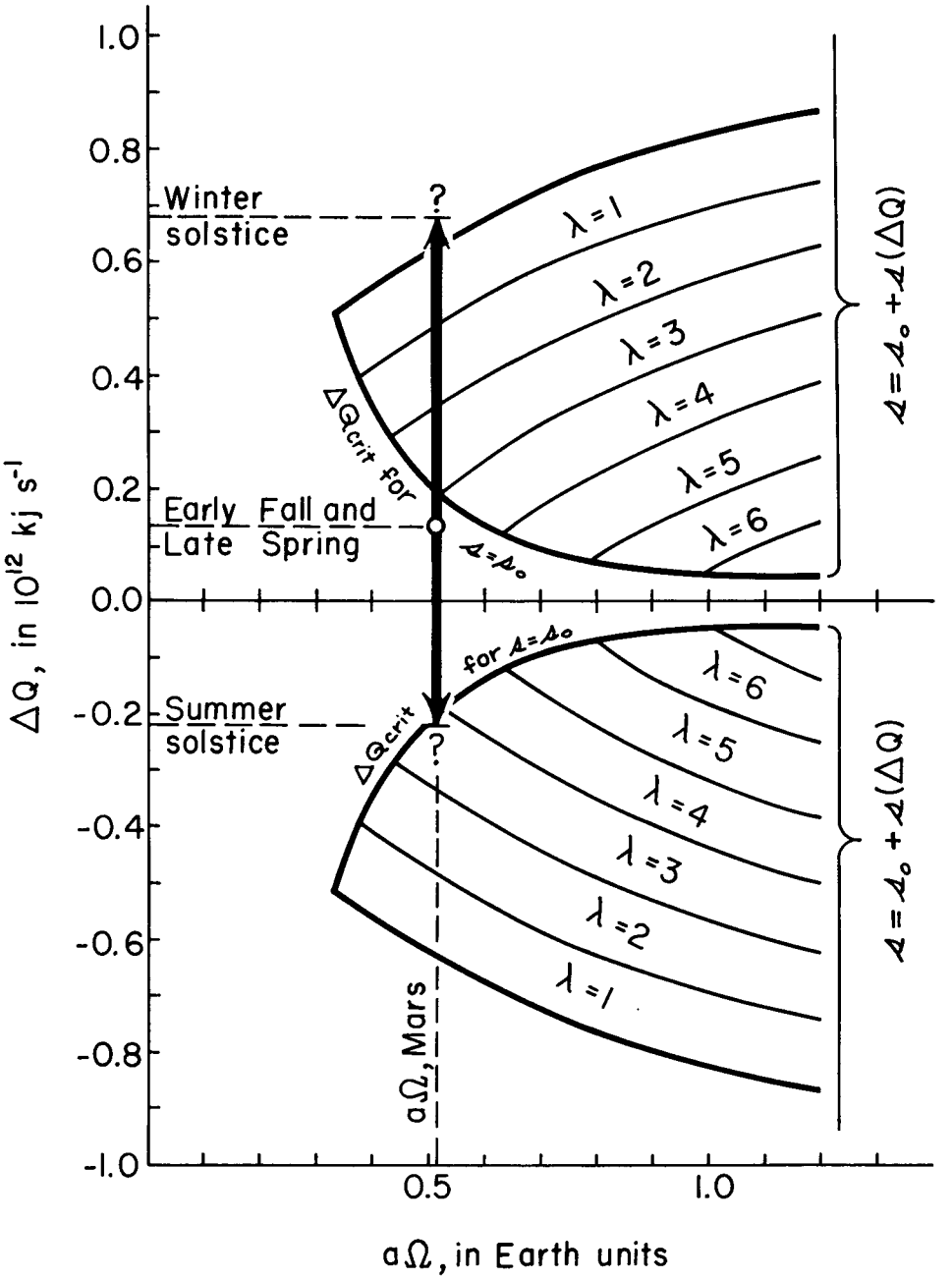


Figure 15.

Schematic representation of the parameter λ as a function of $a\Omega$ and ΔQ, when $s(\Delta Q) >> s_*(\Delta Q)$.

in fact, reports no such lower transition; and Fultz (1959, p. 21) reports the occurrence of the lower symmetric circulation only "under certain circumstances," which he has not yet documented.

We may note, further, that according to the linear theory, by equation (15), the critical poleward temperature gradient, and hence the actual poleward temperature gradient in the wave regime, is proportional to the static stability. Hence the latitudinal slope of the potential temperature surfaces should be constant. A nearly-constant latitudinal slope of the potential temperature surfaces (constant from summer to winter) is, in fact observed in the Earth's atmosphere, and also in the laboratory experiments.

References

de Vaucouleurs, G., Photometric parameters of Mars, *Harvard College Observatory, Report to Jet Propulsion Laboratory,* under contract No. 950014, 1961.

Fjortoft, R., Some results concerning the distribution and total amount of kinetic energy in the atmosphere as a function of external heat sources and ground friction, in *The Atmosphere and the Sea in Motion,* edited by B. Bolin, Rockefeller/Oxford Press, New York, pp. 194-211, 1959.

Fjortoft, R., On the control of kinetic energy of the atmosphere by external heat sources and surface friction, *Quart. J. Roy. Meteorol. Soc., 86:* 437, 1960.

Fultz, D., R. R. Long, G. V. Owens, W. B. Bohan, R. Kaylor, and J. Weil, Studies of thermal convection in a rotating cylinder with some implications for large-scale atmospheric motions, *Meteorological Monographs, 4,* (21) Amer. Meteor. Soc., Boston, 1959.

Gifford, F., The surface-temperature climate of Mars. *Astrophys. J., 123:* 154, 1956.

Giordmaine, J. A., L. E. Alsop, C. H. Townes, and C. H. Mayer, Observations of Jupiter and Mars at 3-cm wave length, *Astron. J., 64:* 332, 1959.

Hide, R., An experimental study of thermal convection in a rotating fluid, *Phil. Trans. Roy. Soc. London, A250:* 441, 1958.

Houghton, G. H., On the annual heat balance of the northern hemisphere, *J. Meteorol., 11:* 1, 1954.

Huss, A., and Y. Mintz, Growth of a baroclinic wave and evolution of the mean zonal and mean meridional circulations, *Proceedings of the International Symposium on Numerical Weather Prediction, Tokyo, November 7-13, 1960,* in press, 1961.

List, R. J., *Smithsonian Meteorological Tables,* 6th ed., Smithsonian Institution, Washington, D. C. (Tables 132-134), 1951.

Mintz, Y., Final computation of the mean geostrophic poleward flux of angular momentum and of sensible heat in the winter and summer of 1949, *Investigations of the General Circulation of the Atmosphere,* Dept. Met., Univ. Calif., Los Angeles, 1955. (Also in: *Selected Meteorological Papers, Meteor. Soc. Japan,* No. 2, August 1960.)

Miyamoto, S., On the general circulation of the Martian atmosphere, *Contributions from the Institute of Astrophysics and Kwasan Observatory, University of Kyoto,* No. 88, 1960.

Palmén, E., On the mean meridional circulation in low latitudes of the northern hemisphere in winter and the associated meridional and vertical flux of angular momentum, *Societas Scientiarum Fennica. Commentationes Physico-Mathematicae, XVII:* pp. 1-33, 1955.

Phillips, N. A., The general circulation of the atmosphere: a numerical experiment, *Quart. J. Roy. Meterol. Soc., 82:* 123, 1956.

Sinton, W. M., and J. Strong, Radiometric observations of Mars, *Astrophys. J., 131:* 459, 1960.

Sutcliffe, R. C., A contribution to the problem of development, *Quart. J. Roy. Meteorol. Soc., 73:* 370, 1947.

Sutcliffe, R. C., and A. G. Forsdyke, Theory and use of upper air thickness patterns, *Quart. J. Roy. Meteorol. Soc., 76:* 189, 1950.

Thompson, P. D., *Numerical Weather Analysis and Prediction,* Macmillan Co., New York, 1961.

Appendix 9

THE INTERPRETATION OF ULTRAVIOLET SPECTRA OF PLANETARY ATMOSPHERES AND THE NEAR-INFRARED CO_2 BANDS OF VENUS*

Joseph W. Chamberlain

A. Introduction

With the advent of space technology it appears that ultraviolet spectra of the atmospheres of the various planets, including the Earth, will soon be obtained from outside those planetary atmospheres. While these spectra will provide a wide array of new and valuable data pertaining to the chemical composition, the atmospheric temperatures, and other characteristics that can be derived from these parameters, the interpretation of such spectra will not always be a straightforward matter. There will be a number of problems involved that have not become apparent so far in the measurement and interpretation of planetary spectra in the visible and near infrared. Therefore, in planning for such experiments, it is essential to bear in mind the interpretational problems that will be encountered and to allow for them insofar as is possible.

It is desirable to have as high a *spectral resolution* as feasible—for more-or-less obvious reasons. The same requirement exists when obtaining spectra in the visible from ground-based observatories. It is perhaps not so apparent that it is also necessary to have good *angular resolution,* i.e., spectra taken over different parts of the planetary disk, for a proper interpretation of spectra in the ultraviolet. This requirement is imposed by the high probability, for spectra in the ultraviolet, that the planetary atmosphere will be optically thick in some of the emissions observed.

A problem encountered in the interpretation of ground-based spectra, which bears some analogies to this one, concerns the near-infrared CO_2 absorption bands of Venus, which are evidently formed in an optically thick, scattering atmosphere. This interpretational problem also does not seem to be widely recognized. In this note an attempt will be made to summarize the interpretational problems that will almost certainly arise in the ultraviolet, and in the final section the absorption bands in Venus will be discussed as an illustrative example.

B. Information to be Obtained from Ultraviolet Spectra in Planetary Atmospheres

In Sections C. and D. we concentrate on the problem of abundance determinations. We shall show that in the cases where the atmosphere is optically thin, or if negligible secondary scattering occurs, one can derive absolute abundances, but that in cases where multiple scattering is important, only relative abundances can be obtained. However, when the problems of continuous absorption in line and band scattering are treated together, it may be possible to derive the absolute abundance of the absorbing substance and then the abundances of other constituents relative to that absorbing substance at known heights in the atmosphere.

* The research reported in this paper was supported in part by the National Aeronautics and Space Administration through Research Grant NsG 118–61.

There are other pieces of information that can be obtained, however, from such spectra. One important datum would be the hydrogen distribution in the atmosphere, which can be obtained by scanning the planet at the Lyman α wavelength, 1210 Å. Such a measurement supposes that one has a good angular resolution, which can be obtained at present only from a close approach to the planet. The hydrogen should, because of its low molecular weight, show a very slow diminution across the disk of the planet, even extending one or two radii away, in the case of the Earth, Venus, and Mars (if hydrogen is present). Direct measurement of the precise rate at which this density decreases away from the center of the plant will give a direct determination of the temperature of the high atmosphere. This kind of information would be of paramount importance to questions on the rate of escape of the planetary atmosphere and the development of that atmosphere into its present state from primordial conditions.

With adequate spectral resolution it will also be possible to measure the rotational distributions in certain molecular bands and thereby ascertain the temperature of the atmosphere at the height at which the lines contained in the bands are formed. A well known example of temperatures given by planetary absorption lines are the CO_2 bands on Venus, where, however, the interpretation involves a knowledge of the mechanism of line formation (as discussed in Section D). For emission lines there will also be an interpretational problem. The maximum intensity within a rotational band will not occur, in the case of an optically thick atmosphere, at the wavelength indicated by the usual Boltzmann distribution. There will, rather, be a shift of the wavelength in accord with the radiative transfer effects. Therefore, to derive the temperature it is necessary to work from the proper model, as derived from the radiative-transfer examination of the problem and a comparison of theory with the relative intensities of the bands in the system and their intensity variation across the disk.

Spectra obtained on the night side of a planet may possibly show airglow emissions. Our expectations on this type of radiation are rather ill-defined. The terrestrial night airglow shows the Herzberg bands of O_2 in the near ultraviolet. But several components of the airglow arising from photochemical reactions (as distinct from ionospheric electron-ion recombinations) are poorly understood, so that we have little theoretical basis on which to make confident assertions about emissions from the high atmospheres of other planets.

We may be slightly more confident about the detection of any planetary aurorae in the ultraviolet, and, indeed, would expect to find auroral emissions, if present, even on the day side. While the daytime fluorescence resulting from incident sunlight (i.e., the day airglow) may exceed the aurora in its ultraviolet intensity, there will be certain band systems that cannot be excited by absorption of sunlight, but which may emit easily from the particle impact of auroral bombardment. Detection of aurorae would give indirect but nevertheless valuable information on atmospheric composition, the planetary magnetic field, and presumably on the presence of a solar-controlled Van Allen radiation belt. (v. also Appendix 7, by Kaplan, concerning an estimate of the intensity of UV *scattered* radiation.)

C. Line Formation in Planetary Atmospheres for Negligible Scattering

There are certain analogies in a planetary atmosphere for the formation of absorption lines (when the solar continuum is "reflected"), continuous absorption which varies with wavelength, and emission lines (when the continuum is totally absorbed, except at line resonances). Let us consider, first of all, cases where secondary scattering—or re-emission following absorption—is negligible.

The line-absorption problem may be considered in the simple fashion of sunlight passing directly through the atmosphere, striking a reflecting layer, and passing back out of the atmosphere. For a specified position of the Sun and the observer, the sunlight follows a well-defined path, and the abundance of the absorbing substance may be derived directly from the total equivalent width of the absorption line (provided that the pertinent atomic constants are known). If the disk as a whole is observed, one may write down an effective zenith angle of entrance and emergence (van de Hulst, 1952).

In the second example, of continuous absorption depending on frequency, one might readily determine the abundance of the substance in the same fashion. If the absolute value of the absorption coefficient of the substance (for example, ozone) is known, the wavelength at which the reflected solar intensity drops to a specified value indicates the amount of the substance in the atmosphere. One might conceive of the ozone abundance on Mars, for example, as being derived in this fashion.

The formation of an emission line similarly follows a fairly simple mechanism. (For line emission we shall assume that the ground albedo is zero, which might be simulated in an actual atmosphere by a line-scattering substance lying above a totally absorbing medium.) Continuous radiation is incident from outside the planet, is absorbed at the resonance frequency of the line, and finally is scattered in all directions. The unabsorbed continuum passes on through the scattering layer and is totally absorbed by the "zero ground albedo," whereas the resonance radiation is scattered in part back into space. From the total intensity of the scattered radiation (and the known atomic constants) one may again derive the total abundance of the scattering substance.

Thus, in the case of *negligible secondary scattering,* one may obtain the abundances without measuring the variation of the absorption lines with distance from the center of the disk or at different phase angles of the planet. If, however, one could obtain the measurements from very close to the planet (probably less than one planetary radius), then the variations of the emission as a function of distance from the center would indicate the height of line formation, according to the common technique for airglow emission as first developed by Van Rhijn.

D. Line Formation in Planetary Atmospheres with Multiple Scattering

If the atmosphere of the planet is optically thick, the whole problem becomes qualitatively altered. For the case of formation of absorption lines, let us now consider in some detail the situation regarding the CO_2 absorption bands in the near infrared on Venus.

The formation of absorption lines in an optically thick, scattering atmosphere was first discussed by van de Hulst (1952), who proposed that it might apply to the CO_2 bands on Venus. In this view the incident continuum is scattered by particles (presumably the Venus cloud particles) until the light emerges from the top of the atmosphere. If the CO_2 and the scattering particles are distributed fairly homogeneously, the radiation in the wavelengths of absorption lines is similarly scattered, but has, in effect, a probability $\tilde{\omega}_\nu < 1$ of being re-emitted at each scattering experience. The line formation is thus treated according to a radiative-transfer theory of scattering with an albedo $\tilde{\omega}_\nu$ (which varies across a line profile).

Subsequently Chamberlain and Kuiper (1956) showed that these bands vary in equivalent width with the phase angle of Venus in rough agreement with radiative-transfer theory for a thick atmosphere. This phase variation is in the opposite sense to that expected if the absorption bands were formed by a simple traversal of light through a CO_2 atmosphere lying above a well-defined reflecting cloud layer.

Moreover, with the multiple-scattering theory the absorption in a weak line varies as the square root of the absorption coefficient, rather than linearly. Hence it also appeared to be significant that the scattering theory gave a consistent interpretation of the rotational-line absorption in terms of the kinetic temperature. Although that CO_2 temperature determination has generally been adopted in theoretical discussions of Venus' atmosphere, some authors have at the same time adopted a CO_2 abundance "above the cloud layer" as derived by Herzberg (1952) on the model of simple reflection of sunlight.

We shall illustrate in the following how the transfer theory places the level of CO_2 formation well below the apparent top of the cloud layer. The CO_2 temperature then also must apply to this fairly low lying level, whereas the smaller temperatures derived from emission in the far infrared evidently pertain to the higher levels of the clouds. (v. discussion in Chapter IV.)

If the multiple-scattering model is the correct one, as present evidence would lead us to believe, our knowledge of the CO_2 abundance becomes quite indefinite. If the scattering coefficient per unit mass is σ, and the absorption coefficient in a line is κ_ν then for weak lines ($\kappa_\nu \ll \sigma$) the amount of absorption varies as $(\kappa_\nu/\sigma)^{1/2}$. This square-root dependence is physically quite different from the square-root law obtained for the total equivalent widths of strong, saturated lines under pressure broadening. The physical reason for this dependence becomes clear from a consideration of the height profile of line formation.

If the continuum were reflected by pure (conservative) scattering, the atmosphere would be illuminated all the way to the ground, the mean intensity approaching a finite limit at great depths. When a fraction $1 - \tilde{\omega}_\nu = \kappa_\nu/(\sigma + \kappa_\nu)$ is lost at each scattering, the intensity deep in the atmosphere decreases exponentially. For example, a solution of the transfer equation in the first approximation* gives a profile for line formation at large optical depth, t_ν varying as $\exp[-t_\nu(1 - \tilde{\omega}_\nu)^{1/2}/\mu_1]$ where $\mu_1 = 0.577$. Thus for a slight increase in $1 - \tilde{\omega}_\nu$, not only does each volume element absorb with greater efficiency, but the incident radiation does not diffuse so deep; the height profile for the formation of the absorption is shifted higher in the atmosphere. The absorption superimposed on the emergent intensity thus has a relative depression proportional to $(1 - \tilde{\omega}_\nu)^{1/2}$ (v. Chamberlain and Kuiper, 1956).

Hence a determination of the abundance of CO_2 on Venus would require a knowledge of the effective height of line formation, which depends on knowing k_ν/σ. We can only say that the amount of CO_2 above the cloud "layer" (defined now by $t_\nu \sim 1$ in the continuum) may be considerably less than the total amount deduced (with simple-reflection theory) from the absorption spectrum. The absorption actually occurs at an effective depth of $t_\nu \backsim 0.577\ (\sigma/\kappa_\nu)^{1/2}$, which may be much greater than unity even for the line centers, and which must be very large in the wings.

In summary, the only way of determining whether the Venus problem involves multiple scattering, with CO_2 mixed with the scattering substance, or whether it involves a CO_2 atmosphere lying above a sharp cloud deck, depends on measurements as a function of phase angle or, even better, measurements over the surface of the disk with high angular resolution. To determine the absolute abundance of the CO_2 it would be necessary to have some idea of the scattering coefficient and distribution with height of the particles producing the continuous opacity.

The problem for interpretation of continuous absorption in the presence of multiple scattering is very similar. For this problem we might conceive of an example where a substance (e.g., ozone) is producing continuous absorption, but where all the atoms are reflecting light by Rayleigh scattering. Then all that one could determine directly from the observations would be the ratio of the absorbing substance (ozone) to the scattering substance ("air"). However, if one can make an assumption or has some independent information regarding the distribution of the "air" as a function of height, then he immediately obtains an abundance estimate for the ozone at a given pressure. With measurements made at different positions on the planetary surface, one sees to different depths or different pressures, and thereby derives the ozone abundance as a function of height in the planetary atmosphere.

Again it is important to note that, without measurements over the planetary surface, one has no legitimate idea of what is being measured. It is only from the variation across the disk that one can ascertain that the reflecting "surface" is not the actual surface of the planet but is only the Rayleigh scattering high in the atmosphere. Without such measurements one might falsely deduce that he was seeing the entire ozone abundance of the planet and derive a grossly incorrect value for that abundance.

---

* The basic theory is given in Chandrasekhar's (1950) treatise. Chamberlain (1954) has examined the height of line formation with the first approximation for a problem similar to the one discussed here. A more general solution has been expressed in terms of a "sink" function by Brandt (1959).

For the formation of emission lines we have the same type of problems, although the number of possibilities involved makes the total situation considerably more complicated. These problems are now being investigated by myself and Mr. Y. Sobouti at the Yerkes Observatory.

Let us consider one or two simple examples. We shall assume, first of all, that the continuously absorbing substance lies well below the scattering substance in the atmosphere. Whether the scattering substance is optically thick or optically thin can, in the case of an emission band system, be determined directly from the relative intensities of the bands in a progression. For the optically thin case the intensities will follow the variation indicated by the transition probabilities. If, however, the scattering substance is optically thick, the band whose transition ends on the ground vibrational level will be greatly reduced in intensity, while the bands from transitions ending on higher vibrational levels will be relatively enhanced. Thus it is possible even from a cursory inspection of band spectra to ascertain whether one is looking at an optically thick or thin scattering atmosphere. This does not indicate, however, the total abundance of the substance in the atmosphere, as one may be observing the scattering substance only down to a level determined by the onset of a continuously absorbing medium. Thus the investigation of such spectra must go hand in hand with the investigation of the continuous absorption of the various constituents (those in the far ultraviolet, as well as ozone).

If, now, the absorbing substance is mixed homogeneously with the scattering molecules, there is again a change in the relative intensities of the bands in the system. In this situation the relative intensities vary across the disk by amounts depending on the relative abundances. In the same fashion that an analysis of absorption lines can indicate the ratio of the absorbing to the scattering substance, so can one, in treating the emission lines, derive the abundance ratio of the constituent scattering to that producing continuous absorption. In the molecular problems one has an admirable leverage on the problem by having emissions in a band system spanning a considerable range in wavelength.

Once again, the important point is that with measurements across the disk of the planet one has sufficient information to allow one to derive the ratio of the desired abundances, and without such information one cannot even ascertain unambiguously which model is most appropriate for the atmosphere.

Acknowledgement

My thinking in these problems has been assisted in large measure by discussions with Dr. Lloyd Wallace.

References

Brandt, J. C., Solar Lyman-β fluorescence mechanism in the upper atmosphere, *Astrophys. J., 130:* 228, 1959.

Chamberlain, J. W., The formation of atmospheric O_2 emission in the airglow, *Astrophys. J., 119:* 328, 1954.

Chamberlain, J. W., and G. P. Kuiper, Rotational temperature and phase variation of the carbon dioxide bands of Venus, *Astrophys. J., 124:* 399, 1956.

Chandrasekhar, S., *Radiative Transfer,* Oxford: Clarendon Press, 1950.

Herzberg, G., Chapter 13, p. 406, in *The Atmospheres of the Earth and Planets,* edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.

van de Hulst, H. C., Chapter 3, p. 49, in *The Atmospheres of the Earth and Planets,* edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.

NATIONAL ACADEMY OF SCIENCES
NATIONAL RESEARCH COUNCIL

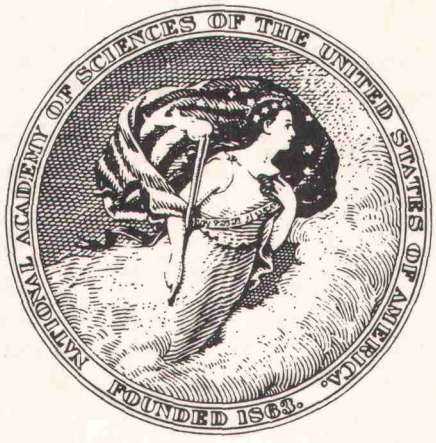
The National Academy of Sciences—National Research Council is a private nonprofit organization of scientists, dedicated to the furtherance of science and to its use for the general welfare.

The Academy itself was established in 1863 under a Congressional charter signed by President Lincoln. Empowered to provide for all activities appropriate to academies of science, it was also required by its charter to act as an adviser to the Federal Government in scientific matters. This provision accounts for the close ties that have always existed between the Academy and the Government, although the Academy is not a governmental agency.

The National Research Council was established by the Academy in 1916, at the request of President Wilson, to enable scientists generally to associate their efforts with those of the limited membership of the Academy in service to the nation, to society, and to science at home and abroad. Members of the National Research Council receive their appointments from the President of the Academy. They include representatives nominated by the major scientific and technical societies, representatives of the Federal Government, and a number of members-at-large. In addition, several thousand scientists and engineers take part in the activities of the Research Council through membership on its various boards and committees.

Receiving funds from both public and private sources, by contributions, grant, or contract, the Academy and its Research Council thus work to stimulate research and its applications, to survey the broad possibilities of science, to promote effective utilization of the scientific and technical resources of the country, to serve the Government, and to further the general interests of science.

The Space Science Board was established by the President of the Academy in 1958. Its primary purpose is to study scientific research opportunities and needs opened up by the advent of rockets and satellites as tools for research; to give advice and recommendations on space science to interested agencies and institutions; to stimulate research in the rocket and satellite fields; and to cooperate with scientists in these fields in other countries, particularly through the Committee on Space Research (COSPAR), established in 1958 by the International Council of Scientific Unions.


NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA.
FOUNDED 1863.