The Structure and Heating of the Chromosphere-Corona Transition Region

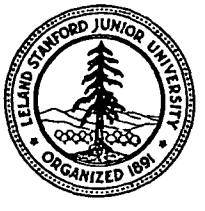
by

R. L. Moore

March 1972

SUIPR Report No. 463

Prepared under
National Aeronautics and Space Administration
Grant NGL 05-020-272 and
Air Force Office of Scientific Research
Contract F44620-69-C-0008



INSTITUTE FOR PLASMA RESEARCH
STANFORD UNIVERSITY, STANFORD, CALIFORNIA

THE STRUCTURE AND HEATING OF THE CHROMOSPHERE-CORONA TRANSITION REGION

by

R. L. Moore

March 1972

SUIPR Report No. 463

Prepared under

National Aeronautics and Space Administration
Grant NGL 05-020-272
and
Air Force Office of Scientific Research
Contract F44620-69-C-0008

Institute for Plasma Research
Stanford University
Stanford, California

ABSTRACT

A broad endeavor of solar physics is to determine from the observed radiation the structure of the solar atmosphere and the nature of the energy balance which this structure implies. A basic feature of the solar atmosphere which is still not well understood is the chromosphere-corona transition region in which, in a height of a few thousand kilometers, the temperature increases from about 10^4 K at the top of the chromosphere to about 10^6 K at the bottom of the corona. The problem investigated in this dissertation is that of the structure and heating (or energy balance) of the transition region and the role of the transition region in the structure and heating of the atmosphere as a whole.

First, we define the problem by summarizing the well-established observed features of the structure of the atmosphere and by reviewing the observed radiative energy losses of the atmosphere. From this study of the observations we find:

1. The atmosphere is approximately static and planar except in active regions during their growth phase.

2. While the corona must be supplied with heat from the inner layers by some nonthermal process such as the propagation and dissipation of mechanical waves, the observations suggest that the heating of the transition region is dominated by conduction heating from the corona.

3. The dominant energy loss of the corona is due to downward heat conduction and is of order 10^6 erg cm^{-2} sec^{-1}. This requires the existence of a high-temperature corona and the steep temperature rise in the transition region.

We then consider a static, planar model of the solar atmosphere which has a temperature profile representative of the actual solar atmosphere. An analysis of wave propagation and dissipation in this model leads to the following conclusions:

1. The corona is heated by nearly vertically propagating compression waves. These waves pass upward to the corona through the chromosphere and transition region from the photosphere where they are generated by the granulation convection cells bobbing to the top of the convection layer.

2. The waves are probably dissipated in the corona by thermal conduction and Landau damping rather than by developing into shock waves.

Having obtained an overall picture of the structure and heating of the atmosphere as a whole, we then develop a static, planar model of the transition region which is heated by thermal conduction from the corona and cooled by radiative losses. From the good agreement between the model and the results of XUV emission-line observations of the transition region, we conclude:

1. The transition region is heated primarily by thermal conduction from the corona rather than by wave dissipation.

2. The balance of conduction heating and radiative cooling in the transition region determines the number density in the transition region and lower corona.

A general conclusion to be drawn from our results is that the temperature profile of the transition region and lower corona results primarily from the energy balance of the corona, while the number density is determined by the energy balance of the transition region.

CONTENTS

CONTENTS (Cont)

ILLUSTRATIONS

ACKNOWLEDGMENTS

This dissertation is the culmination of my doctoral education in solar physics. For this education I am foremostly indebted to Professor Peter A. Sturrock, who introduced me to solar physics and to astrophysics in general, suggested the research topic of this dissertation, and gave helpful guidance and critical advice along the way in addition to financial support. But my true debt to Professor Sturrock lies in the fact that through my association with him--through the example of his own research and due to the freedom he has given me in guiding my research--I have, I believe, achieved the primary objective of a doctoral education, an ability to carry out significant independent research. This is no mean service, and it is a pleasure to acknowledge it and thank him for it.

I am indebted to Professor Peter C. W. Fung with whom I collaborated in developing the model of the chromosphere-corona transition region presented in Chapter 4. I thank Professor I-Dee Chang and Professor R. N. Bracewell for critically reading this dissertation. During the course of my research, I have had several helpful and stimulating discussions with Rich Epstein, Chuck Newman, Geert Dijkhuis, and Spiro Antiochos. I am grateful to my parents for their patient confidence which helped sustain me through several dubious years of graduate school. Finally, I would like to thank my wife-to-be, Barbara Welenc, for typing the first draft and for personal encouragement which greatly aided the writing of the dissertation.

1. INTRODUCTION

The problem of the structure and heating of the chromosphere-corona transition region is part of the general problem of the structure and heating of the outer layers of the solar atmosphere. The heating of the solar atmosphere has been recognized as one of the fundamental problems of the solar atmosphere since about 1945 when it became generally accepted that the temperature of the corona a short distance (a few hundredths of the sun's radius) above the visible surface, or photosphere, is about 10^6 K, more than a hundred times hotter than the photosphere. Due to its proximity to the photosphere, the tenuous, hot corona is essentially in thermal contact with the massive, relatively cold surface layers of the sun. Consequently, the corona would quickly cool to the temperature of the photosphere if the thermal energy of the corona were not continually resupplied by some form of heating. Since the heating causes the temperature of the atmosphere to increase with height, it is clear that the heating of the atmosphere and its structure (e.g., the run of temperature and density with height) are inseparable parts of the same problem.

Our study of the structure and heating of the solar atmosphere has led us to develop a model for the structure and heating of a limited region of the atmosphere, the layer called the chromosphere-corona transition region. The purpose of this dissertation is to present our study of the structure and heating of the transition region by means of this model, and to bring out the position and role of the transition region in the overall structure and heating of the atmosphere. To this end we first summarize the observed structure of the solar atmosphere in Chapter 2, and then give a general discussion and summary of the heating problem in

Chapter 3. With this background, in Chapter 4 the model transition region is developed and compared with observations, and the resulting implications for the structure and heating of the transition region and the atmosphere in general are discussed.

2. THE STRUCTURE OF THE SOLAR ATMOSPHERE

2.1 General Description

The solar atmosphere is the observable part of the sun which surrounds the unobservable interior. The radius of the sun $R_\odot$, which defines the surface of the sun, is defined to be the radius of the edge, or limb, of the sun when observed in white light. The solar atmosphere begins just below the surface at a depth of about 0.0005 $R_\odot$ and extends outward from the sun for hundreds of solar radii. However, we will be almost exclusively concerned here with only the thin inner layer between the bottom of the atmosphere and the level of the temperature maximum in the low corona. The thickness of this layer is probably less than 0.3 $R_\odot$ and may be as little as 0.03 $R_\odot$.[1]

The structural features of the atmosphere separate naturally into two categories: vertical and horizontal. Vertically, the most general feature of the atmosphere below the temperature maximum is the stratification of the atmosphere into four observationally and structurally different layers. In order of position from bottom to top, these are the photosphere, the chromosphere, the chromosphere-corona transition region, and the low corona. Horizontally, the fundamental feature is the existence of active regions and quiet regions: the solar atmosphere consists of a quiet, or undisturbed, relatively uniform background on which there are superposed, like islands on a sea, enhanced, disturbed active regions. The active regions may be thought of as enhancements of the background atmosphere to the extent that the structure of active regions is similar to the structure of quiet regions. Active regions are viewed as disruptions of the background atmosphere to the extent that their structure

differs qualitatively from that of the background atmosphere. Therefore, we will first consider the structure of the quiet solar atmosphere, and then consider the modification of this structure in active regions.

2.2 The Quiet Solar Atmosphere

2.2.1 Average Vertical Structure

Figure 2.1 shows the relative thickness of the basic layers of the quiet solar atmosphere. The approximate height, number density, and temperature at the boundaries of the layers are indicated in the figure.

The photosphere is the surface layer of the sun which is observed in white light. The top of the photosphere is taken to be the limb of the sun in white light; that is, the top of the photosphere is defined to be the surface of the sun. Hence, it is natural to measure the height in the atmosphere from the top of the photosphere. The top of the photosphere is defined more precisely as the level in the atmosphere which has an optical depth of 1.0 when viewed tangentially on the limb in 5000 Å light.[2,3] Viewed at the center of the disk, this level has an optical depth of only 0.004,[3] which demonstrates that deeper levels can be seen at the center of the disk than at the limb. The optical depth increases so rapidly with depth in the atmosphere that only an unobservably small amount of the light emitted from levels below -400 km escapes from the sun directly without being absorbed in the photosphere. Thus, in accordance with the solar atmosphere being defined as the directly observable part of the sun, -400 km may be taken as the bottom of the photosphere.[4]

The radiant energy lost from the photosphere is continually supplied by the conversion of hydrogen into helium near the center of the

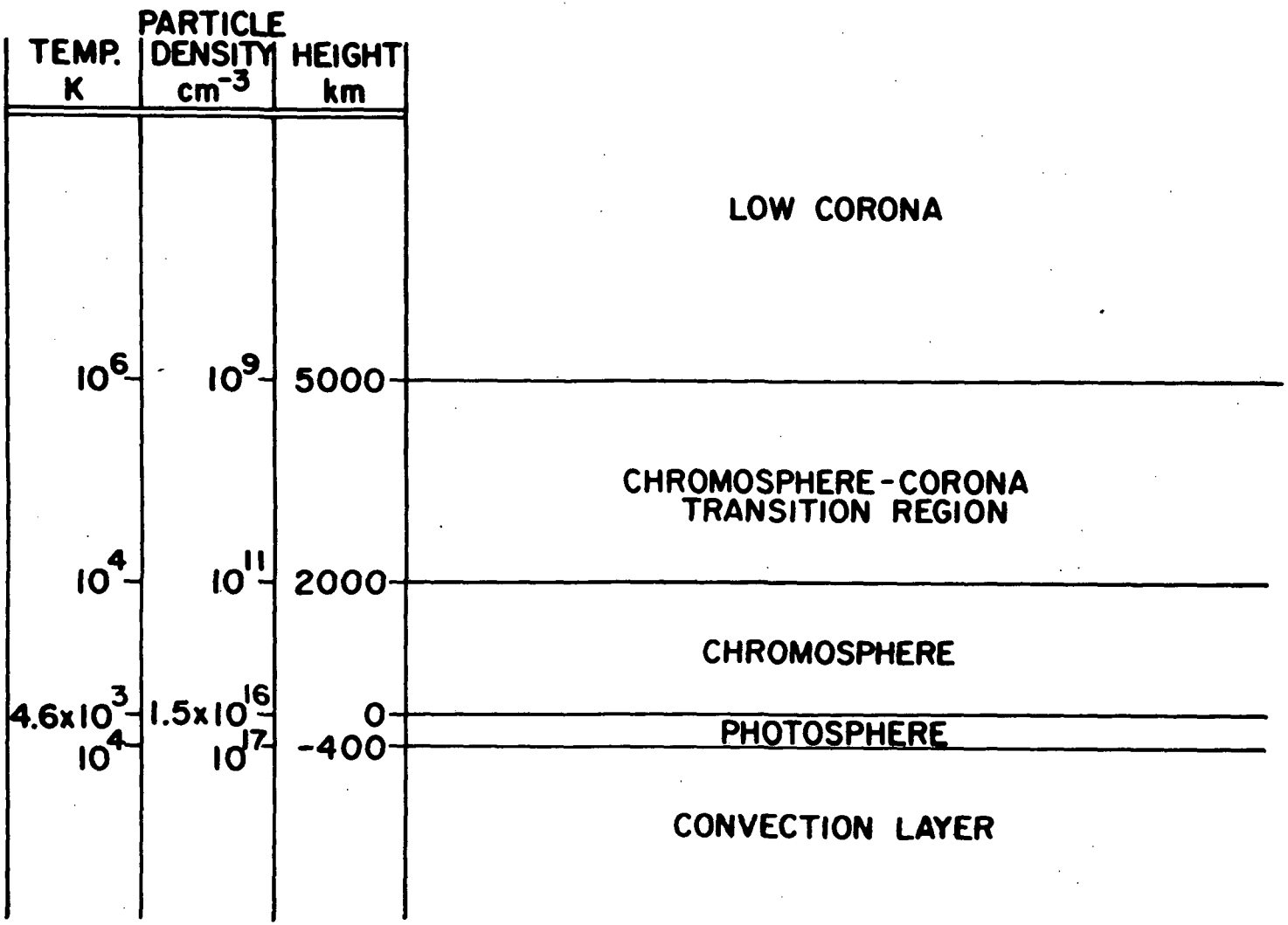


Figure 2.1. BASIC LAYERS OF THE QUIET SOLAR ATMOSPHERE.

sun (within about the inner 0.5 $R_\odot$).[5] Within about 0.9 $R_\odot$ from the center, the sun is convectively stable, and the energy released in the central core flows outward by radiative transfer through the stationary gas. Approximately the outer 0.1 $R_\odot$ of the sun is in a state of free convection. In this convection layer radiative transfer is negligible, and the energy is carried outward by the convective motion of the gas. Just below the surface of the sun, at the base of the photosphere, the sun again becomes stable against convection, so that in the deepest 100 km or so of the photosphere the primary energy flow process changes back to radiative transfer.[6] Thus, the energy which is radiated from the photosphere is supplied to the photosphere from the interior by convection.

During a total eclipse of the sun, a dark-red layer is seen just above the photosphere. Due to its color, this layer is called the chromosphere. Viewed at the limb, "the chromosphere seems composed of a more or less homogeneous layer, from which emerge fine streaks or spikes".[7] These spikes are the spicules which will be discussed in Section 2.2.2.3. Since our purpose here is to consider the <u>average</u> vertical structure of the atmosphere, we define the chromosphere to be the "more or less homogeneous layer" from which the spicules emerge. The height of this layer is about 2,000 km.[8,9]

The outermost layer of the solar atmosphere is the corona. Due to the scattering of the sunlight emitted from the photosphere as it passes through the corona, the corona is visible during a total eclipse as an irregular white envelope surrounding the sun and extending to heights of 1 $R_\odot$ or more above the limb. In addition to (1) scattering

photospheric light, the corona also emits observable amounts of radiation in (2) optical emission lines from highly ionized atoms of heavy elements (e.g., iron, calcium, and nickel), in (3) XUV (far-ultraviolet and X-ray) resonance emission lines from highly ionized atoms of heavy elements, and in (4) radio waves. Quantitative observations of the corona in each of these four types of coronal radiation indicate that the temperature of the corona is between 1 and 2×10^6 K[10,11] (within about 1 $R_\odot$ from the surface[12]). This is the single most important physical feature of the corona.

The red light emitted by the chromosphere is the Balmer α line (Hα) of the neutral hydrogen atom. Since hydrogen rapidly becomes ionized as the temperature increases above 10^4 K, the temperature at the top of the chromosphere cannot be much greater than 10^4 K, two orders of magnitude below the temperature of the corona. Between the chromosphere and the corona there is a transition region in which the temperature increases from about 10^4 K to about 10^6 K. The transition region is observed primarily in XUV resonance lines. The XUV-line observations indicate that the thermal gas pressure in the transition region is about 10^{-1} dyne cm^{-2} (see Section 4.3.2). Considering this result in conjunction with eclipse observations of the chromosphere, Athay[13] concludes that the 10^4 K level at the top of the chromosphere occurs in the vicinity of 2,000 km. Therefore, it is compatible with our above definition of the chromosphere to define the 10^4 K level to be the boundary between the chromosphere and the transition region. We define the 10^6 K level to be the boundary between the transition region and the corona. The XUV-line data imply that the transition region, thus defined, has a thickness of only a few thousand kilometers.

A representative temperature profile for the quiet solar atmosphere is shown in Figure 2.2. This temperature profile was obtained as follows. From the XUV-line data, the temperature gradient can be estimated as a function of temperature in the transition region and low corona (see Section 4.3.2). This allows an estimate of the heat flux flowing down through the transition region from the corona. The data indicate that there is a downward heat flux of about 10^6 erg cm^{-2} sec^{-1}, which remains approximately constant from about the 10^6 K level down to the 10^5 K level. On this basis, the temperature profile above 10^5 K in Figure 2.2 was computed by assuming a constant downward heat flux of 10^6 erg cm^{-2} sec^{-1}. As was pointed out above, it is reasonable to assume a height of 2,000 km for the 10^4 K level. The 10^5 K level was rather arbitrarily placed 500 km above the 10^4 K level. (Some justification for adopting a thickness of this order for the 10^4 to 10^5 K temperature range is given in Appendix B.) The shape of the temperature profile in the 10^4 to 10^5 K region was chosen to fit smoothly with the temperature profile above and below this region. For the chromosphere and photosphere, the temperature profile of the Bilderberg Continuum Atmosphere[14] (BCA) was adopted slightly modified to pass through 10^4 K at 2,000 km above the photosphere. The BCA temperature profile was chosen mainly for its virtue of having to be only slightly modified near 2,000 km to satisfy this condition. Below the photosphere, in the convection layer, the temperature profile was drawn to pass through the point (-1000 km, 14,000 K) given by Allen[3] and to fit smoothly with the temperature profile in the photosphere.

In Figure 2.2, it is seen that each layer has its own characteristic temperature structure. The temperature decreases with height in both the convection layer and the photosphere, but the decrease is more rapid in the

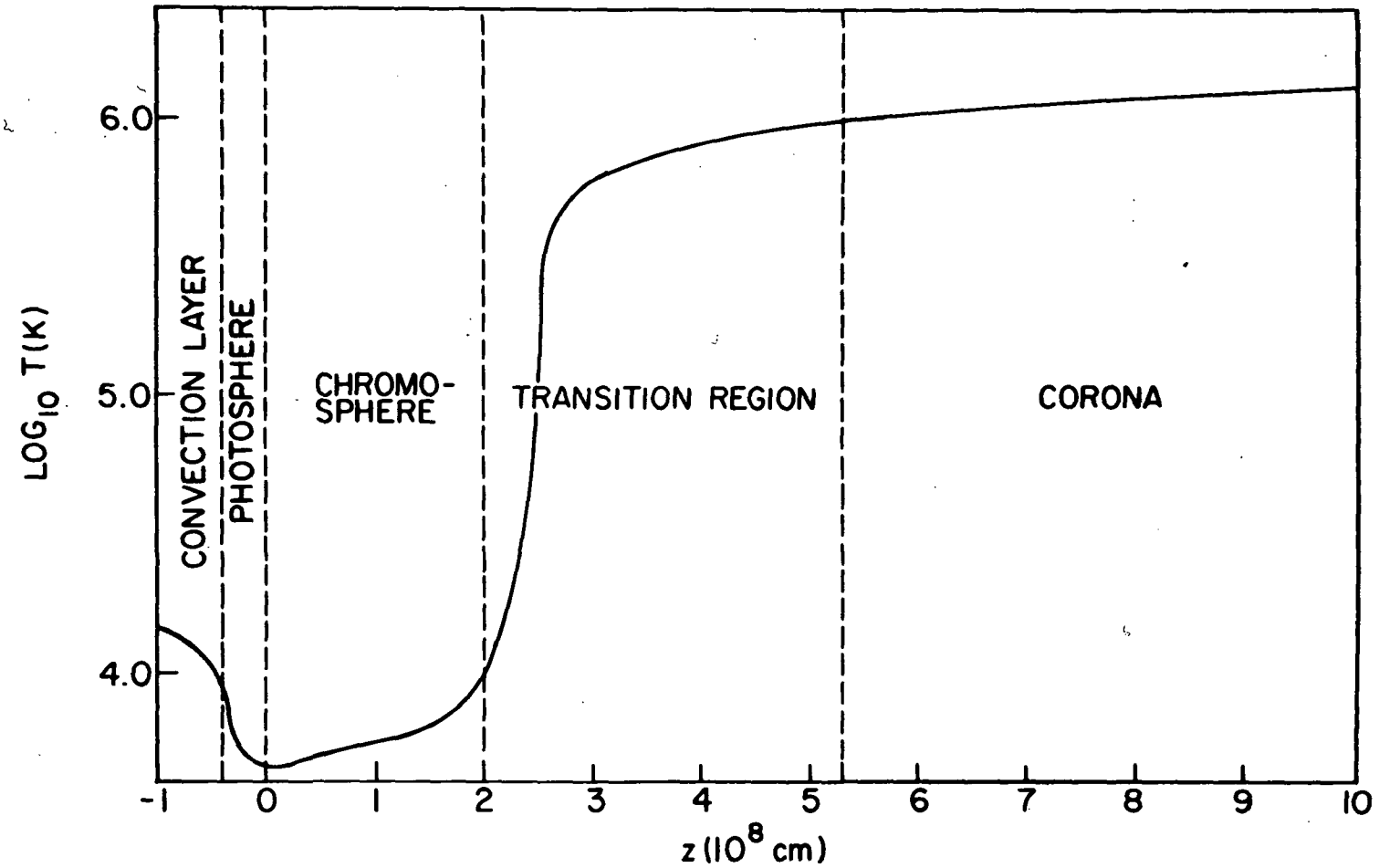


Figure 2.2. REPRESENTATIVE TEMPERATURE PROFILE FOR THE QUIET SOLAR ATMOSPHERE.

photosphere. At the top of the photosphere the temperature reaches a minimum of 4,600 K and then rises slowly through the chromosphere. In the transition region, the profile again steepens before becoming very flat in the corona.

The steepness of the temperature profile in the transition region and its flatness in the corona result from the constancy of the downward heat flux above 10^5 K and the strong temperature dependence of the thermal conductivity. The temperature gradient is given by

$$\frac{dT}{dz} = -\frac{F}{\kappa} \quad , \tag{2.1}$$

where F is the heat flux and κ is the thermal conductivity. At temperatures above 2×10^4 K, the thermal conductivity is given by (see Section 4.2.3)

$$\kappa = 6.6 \times 10^{-6} \, T^{2.36} \text{ erg sec}^{-1} \text{ K}^{-1} \text{ cm}^{-1} \quad . \tag{2.2}$$

This requires the temperature profile to be 230 times steeper at the 10^5 K level than at the 10^6 K level if the heat flux is to be the same at both levels. At 10^6 K and above, the thermal conductivity is so large that only a slight temperature gradient is necessary in order to conduct a heat flux of 10^6 erg cm^{-2} sec^{-1}. Hence, the corona is approximately isothermal.

The average vertical density structure of the quiet solar atmosphere is shown by the density profile in Figure 2.3. This profile is based on the following considerations. From the XUV-line data, it is fairly well established that in the transition region and the first several thousand kilometers of the corona the thermal gas pressure nkT is

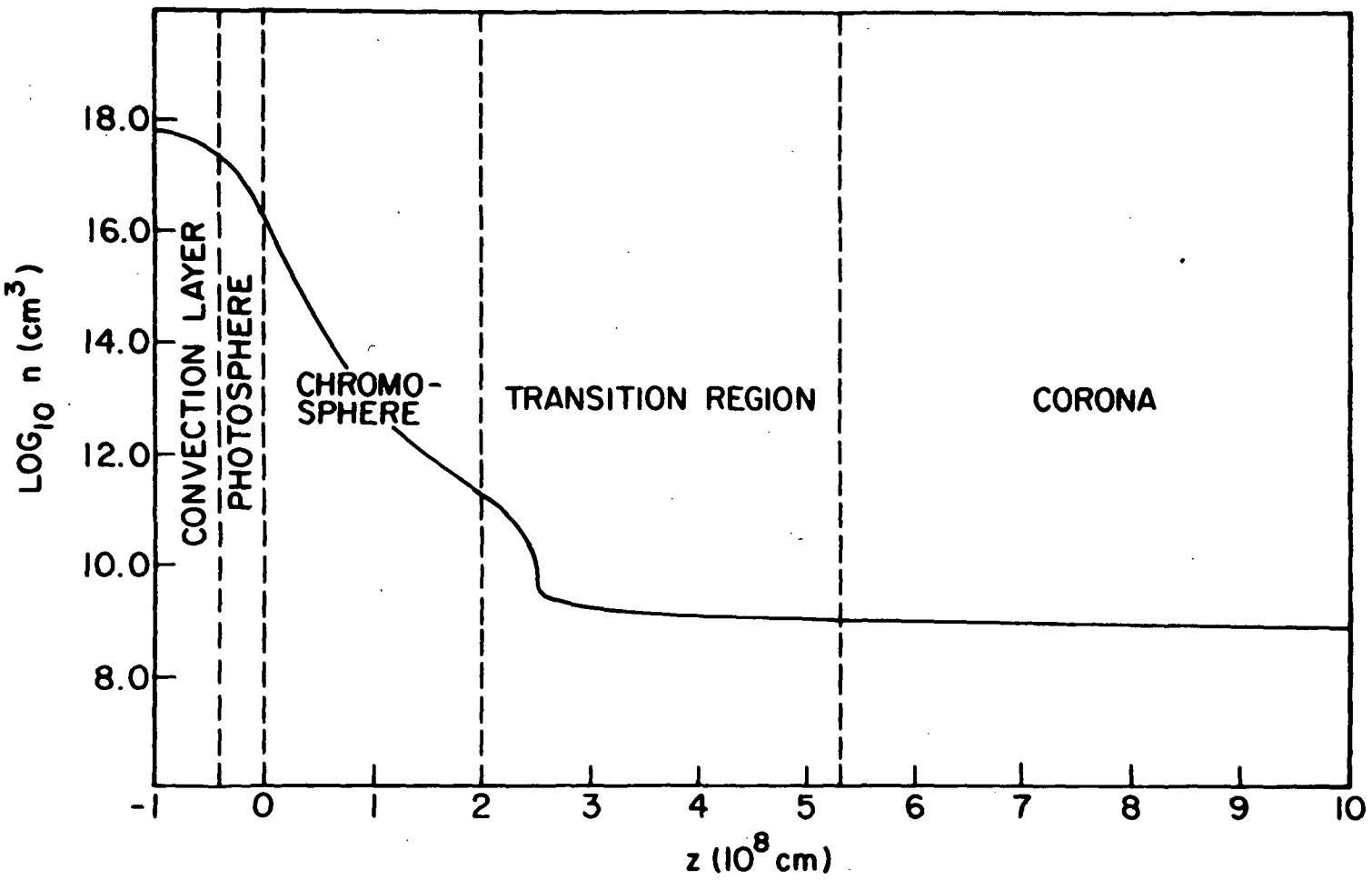


Figure 2.3. REPRESENTATIVE DENSITY PROFILE FOR THE QUIET SOLAR ATMOSPHERE.

approximately constant and that nT has a value of about 10^{15} cm^{-3} K. This constant value of nT was adopted and used to obtain the number density from the temperature profile above the 10^5 K level in Figure 2.2. The value of 2×10^{12} cm^{-3} for the number density at the 10^4 K level results from the plausible assumption (see Appendix B) that the thermal gas pressure is a factor of two greater at the 10^4 K level than at the 10^5 K level. For the chromosphere, a somewhat smoothed version of the profile derived by Athay[13] was adopted. This profile fits smoothly with the profile in the transition region and with the BCA density profile, which was adopted for the photosphere. The value of the density in the convection zone at -1000 km was taken from Allen.[3]

Although velocity oscillations and fluctuations with characteristic periods of several minutes are observed in the photosphere and chromosphere, on the average over longer periods of time, the quiet solar atmosphere may be considered to be in hydrostatic equilibrium. For a hydrostatic atmosphere supported against gravity by thermal gas pressure, the number density gradient is given by

$$\frac{1}{n}\frac{dn}{dz} = -\frac{mg}{kT} - \frac{1}{T}\frac{dT}{dz} \quad , \tag{2.3}$$

where m is the mean particle mass and g is the acceleration of gravity. This equation, in conjunction with the temperature profile in Figure 2.2, explains the general features of the density profile in Figure 2.3. The profile steepens with height in the photosphere because the temperature is decreasing and the temperature gradient is negative and increasing. In the chromosphere, the increasing temperature causes the density profile to become increasingly less steep. The steepening of

the density profile in the transition region is due to the dominance of the temperature-gradient term in equation (2.3) near the 10^5 K level. At higher levels, the density profile becomes increasingly flatter because the temperature increases and the temperature gradient decreases.

The preceding explanation of the density profile is only qualitatively correct. In addition to thermal gas pressure, the atmosphere is supported to some extent by the turbulent pressure of the velocity fluctuations. This effect is probably most pronounced near the base of the transition region, where we find (see Appendix B), in agreement with Athay's model chromosphere[13] which we adopted in Figure 2.3, that the turbulent pressure may be comparable to the thermal gas pressure. Therefore, the turbulent pressure probably contributes significantly to the support of the gas in this region.

2.2.2 Horizontal Structure and Associated Fluid Motion

2.2.2.1 Photospheric Granulation and Oscillations in the Photosphere and Low Chromosphere

The best high-resolution photographs of the photosphere show that in normal, quiet regions the brightness of the photosphere is not horizontally uniform. Instead, the photosphere appears as a mosaic of irregular polygonally-shaped bright cells of various but not greatly different sizes.[15,16] The bright cells are separated, or bounded, by a network of narrow dark lanes. This pattern of bright cells with dark boundaries is called the photospheric granulation. The diameter of a bright cell, or granule, is typically 1000 to 2000 km,[16,17] while the dark boundaries have a width of about 350 km.[16] Some granules are brighter than others; the ratio of the intensity of the light from the

brighter granules to the intensity of the dark intergranule lanes is 1.3, corresponding to a temperature difference of 300 K, while the root-mean-square intensity variation over many granules corresponds to a temperature difference of about 100 K.[17] The granulation pattern is not stationary. Old granules are continually broken up and replaced by new granules so that the average lifetime of a granule is about 6 minutes.[16]

Most of the light emitted from the photosphere comes from an optical depth near 1.0. At the center of the disk optical depth 1.0 occurs about 300 km below the top of the photosphere. Since the convection below the photosphere is believed to persist to about this level in the photosphere,[6] the photospheric granulation is apparently the top of the uppermost layer of convection, each individual granule being a convection cell with hot, bright gas rising in the center and cool, dark gas falling at the boundary. A new granule appears when a fresh blob of hot gas rises to the top of the convection layer. The observed Doppler shifts of absorption lines formed in the vicinity of the -300 km level indicate convective motions consistent with this interpretation of the granulation.[6,18]

The convective motion in the convection layer arises spontaneously because the layer is convectively unstable. Thus, this layer of the atmosphere naturally possesses a nonzero velocity field even though the average velocity is zero and the layer is in overall hydrostatic equilibrium. The convection terminates at about -300 km because at this level and above the atmosphere is convectively stable. Although the upper photosphere and the chromosphere are convectively stable, Doppler shifts in spectra of absorption lines formed in these layers show that

these layers also possess a fluctuating velocity field. However, as is compatible with the convective stability, the basic character of the velocity field in these layers is oscillatory rather than convective. The observations also indicate that this oscillatory velocity field does not arise spontaneously in the photosphere and chromosphere, but is excited in these stable layers by the buffeting of the photosphere by the top of the convection layer.

The oscillatory velocity field in the photosphere and low chromosphere has the following observed features:

1. The motions tend to be in phase over horizontal regions or cells which are comparable in area to the granules.[19]

2. Near the level of the granulation (about -300 km) horizontal velocities and vertical velocities are comparable; at the top of the photosphere the motion is nearly vertical.[19,20]

3. The rms velocity is about 0.2 km sec^{-1} near the level of the granulation, increases to about 0.4 km sec^{-1} near the top of the photosphere, and reaches a value of about 1.6 km sec^{-1} at the level of the core of $H\alpha$ in the chromosphere.[16,19]

4. The lifetime of the oscillatory cells is comparable to the lifetime of the granules.[19]

5. The appearance of a new granule is followed by the onset of oscillations above the granule. These oscillations are initiated by an upward motion.[20]

6. The period of the oscillation increases from about 200 sec at the beginning of the oscillation to a steady period of about 300 sec.[19,21]

7. The phase relation between the brightness and the Doppler shift of the observed lines suggests that the oscillation is initially that of a propagating wave and relaxes to that of a standing wave.[21]

Each of these observations is compatible with the interpretation that the oscillations in the photosphere and chromosphere result mainly from

hot new granules bobbing to the top of the convection layer and impinging on the overlying stable layers of the photosphere.[16,22]

2.2.2.2 Supergranulation and the Chromospheric Network

In addition to the convective motion in the granulation and the associated oscillatory motion in the photosphere, a much larger scale convective velocity field is revealed by the Doppler shifts in the photospheric absorption lines.[16,19] This velocity field has a pattern of closely packed, roughly circular cells similar to the photospheric granulation, but with a scale some 15 to 20 times larger. By analogy, this pattern is termed the supergranulation. A typical supergranule has a diameter of about 30,000 km and a lifetime of about 20 hr. Gas flows upward at the center, radially away from the center, and downward at the boundary of the supergranule. The horizontal flow velocity is about 0.4 km sec^{-1}, and the upward and downward velocities are 0.1 to 0.2 km sec^{-1}. Apparently, the supergranules are the tops of large convection cells which extend deep (10^4 km or so) into the convection layer.[23]

The supergranulation is not apparent on white-light photographs of the photosphere. This indicates that the structure of the photosphere is not much affected by the supergranulation. However, the supergranulation noticeably influences the structure of the overlying chromosphere in normal quiet regions. Observations of the chromosphere in the K line of singly ionized calcium or the Hα line of hydrogen show a network of enhanced emission in the chromosphere which coincides with the network of supergranule boundaries in the photosphere.[19] The enhanced emission implies that the density and/or the temperature is enhanced in the chromospheric network.

The enhanced emission in the chromospheric network is believed to be due to the enhanced magnetic field at the boundaries between the supergranules.[16,19] The average (over areas larger than a supergranule) strength of the magnetic field in the photosphere is observed to be about 1 gauss.[24] The fact that the spicules spurt upward out of the chromosphere more or less vertically[25] indicates that the magnetic field in quiet regions is also approximately vertical since the spicules should move along the magnetic field lines. Since the magnetic field is "frozen in" the highly conducting gas of the photosphere, and since the magnetic pressure in the photosphere in quiet regions is much less than the gas pressure, the horizontal flow in the supergranulation should tend to sweep the field out of the interior of the supergranules and concentrate it in the boundary network.[23] Observations of the magnetic field in the photosphere confirm that the magnetic field is concentrated in the supergranule boundaries and indicate that the field strength is of the order of 50 gauss in these regions.[19]

2.2.2.3 Spicules

When the chromosphere in quiet regions is observed at the limb, especially when it is observed in the Hα line, narrow columns, called spicules, are seen to extend upward out of the relatively homogeneous layer which we have defined as the chromosphere. The spicules are transient; they appear by rising out of the chromosphere and disappear by fading away or by falling back to the chromosphere. Since spicules are observed in chromospheric light and appear to be generated in the chromosphere, they may be considered to be local, temporary extensions of

the chromosphere into the overlying transition region and low corona. Spicules observed on the limb have the following characteristics:[19,25]

1. The typical spicule diameter (about 1000 km) is comparable to the typical photospheric granule diameter.
2. Spicules typically extend to heights from 5000 to 10,000 km.
3. The direction of spicules is predominantly vertical.
4. Upward velocities are typically in the range of 20 to 40 km sec^{-1}.
5. The average spicule lifetime is about 10 minutes.
6. The spicules are not distributed uniformly on the limb but tend to occur in clumps or bushes.
7. Above a height of 2000 km, the spicules occupy less than 0.1 of the total surface area.

The spicules can also be identified on the disk of the sun: narrowband $H\alpha$ filtergrams of the chromosphere on the disk show elongated features whose characteristics correspond to the above characteristics for limb spicules.[19,25] These disk observations show that the spicules are related to the supergranulation. The spicule bushes are found to occur predominantly in the chromospheric network, i.e., at the boundaries of the supergranules. In addition, the spicules may be related to the photospheric granules. Beckers[26] has found that the birth rate of spicules in a spicule bush is approximately equal to the birth rate of the granules in the area of the photosphere beneath the bush. The fact that spicules are restricted to the chromospheric network suggests that the enhanced ($\sim$ 50 gauss) vertical magnetic field at the supergranule boundaries is a necessary condition for the formation of a spicule.[25] The approximate equality of the birth rates of spicules and

underlying granules suggests that the granules cause or stimulate the production of a spicule.[26] These considerations suggest that the disturbance produced in the chromosphere by each new granule may be considered to be a spicule embryo which develops into a full-fledged spicule only in the presence of the special conditions at the supergranule boundaries.

The salient features of the horizontal structure of the quiet solar atmosphere are summarized by the schematic sketch of the atmosphere in Figure 2.4 (after Noyes[19]).

2.3 Active Regions

Photospheric, chromospheric, and coronal observations all show that there are island areas, called active regions, in the atmosphere in which the normal atmosphere is disrupted and enhanced. These regions are distributed randomly in longitude, but are normally found only in the equatorial region between the latitudes of +40 and -40 degrees. An active region is not a permanent feature of the atmosphere, but typically has a lifetime of several months during which it evolves through a rapid growth phase and a slow decay phase. In the growth phase, the active region arises in the normal atmosphere from a small area at the boundary of a supergranule and grows in intensity, complexity and area for about 2 weeks or so.[27,28] During the remaining several months, the area remains more or less constant while the intensity and complexity decrease until the active region dissolves into the surrounding normal atmosphere.[27] At the end of the growth phase, a typical active region may have linear dimensions of the order of 10^5 km,[27] and cover an area corresponding to

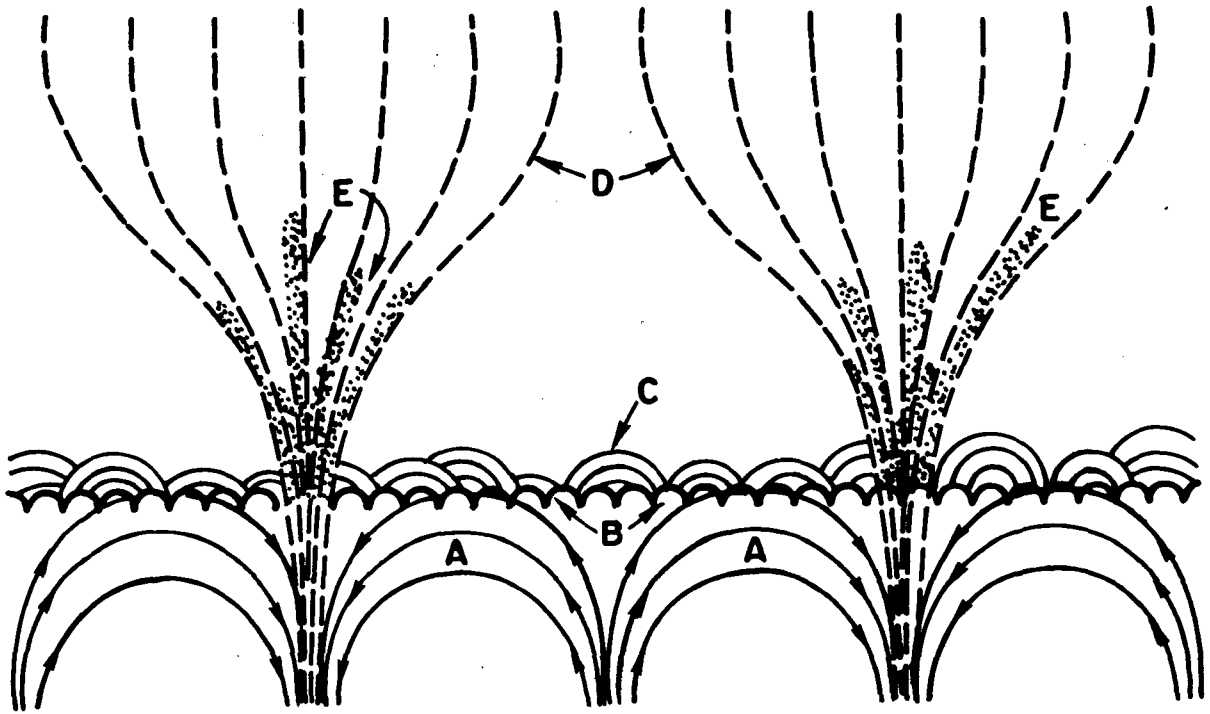


Figure 2.4. SCHEMATIC SKETCH OF THE HORIZONTAL STRUCTURE OF THE QUIET SOLAR ATMOSPHERE (AFTER NOYES[19]). A. Flow lines in a supergranulation convection cell. B. Photospheric granules. C. Oscillations in the photosphere and low chromosphere stimulated by rising granules. D. Magnetic field lines. E. Spicules.

about 10 supergranules, or a few thousandths of the sun's surface.[29] A very large active region may cover a few percent of the sun's surface.

An active region results from the eruption of magnetic fields from below the surface of the sun.[30] Within about a day after the eruption begins, the magnetic field forms a bipolar magnetic region, two adjacent areas of predominantly opposite polarity, which suggests that a kink or loop in a magnetic flux tube is emerging through the surface. In the growth phase of the active region, the flux tube is in the process of emerging through the surface. During this time the magnetic field strength near the center of each half of the bipolar region is of the order of 10^3 gauss, and the average field strength over the area of the active region is perhaps of the order of 10^2 gauss.[31] In the decay phase of the active region, the flux tube has completely emerged, and the magnetic field slowly diffuses out of the active region through the action of the supergranulation convection.[23] Since the field diffuses away into the surrounding quiet region, the field strength remains strongest near the center of the active region and decreases in strength toward the edge. During the decay phase, a field strength of the order of 50 gauss is typical for the interior of the active region,[32,33] while the field at the edge of the active region is typically 10 gauss or greater.[34]

During the growth phase, an active region is more aptly described as a disruption, rather than an enhancement, of the normal background atmosphere. In the central portion of each half of the bipolar region, the 10^3 gauss magnetic field is sufficiently strong to impede the convection at the top of the convection layer.[35] This reduces the supply of energy to the photosphere, and results in the formation of dark sunspots in these regions. Magnetic lines of force with a strong horizontal

component run between these two groups of spots. This field depresses the chromosphere and gives the chromosphere a filamentary pattern reminiscent of iron filings in a bipolar magnetic field.[31,36] Apparently, due to the continuing emergence of the magnetic flux tube, there is a complicated mix of magnetic fields of opposite polarity in the region between the two sunspot groups of opposite polarity. During the growth phase, the magnetic field relaxes to a simpler, more purely bipolar, configuration through a series of sudden events, called flares, in which the higher magnetic energy of the more complex field pattern is explosively released. These explosions give rise to sudden enhancements (flares) of chromospheric and coronal emission, and may result in the violent ejection of matter outward from the active region. Also during the growth phase, complex loops of magnetic field anchored in the spot regions extend upward into the corona. These loops are observed on the limb in enhanced chromospheric and coronal emission, and have a height of the order of 6×10^4 km.[37] Thus, due to the strong, complex, and changing magnetic field during the growth phase of active regions, the layered structure and steady nature of the quiet atmosphere is disrupted at all levels.

During the decay phase, an active region is more aptly described as an enhancement, rather than a disruption, of the normal background atmosphere. Due to the decrease in the magnetic field strength, the sunspots have disappeared. The deep layers of the photosphere are unaffected, while the higher layers have somewhat enhanced emission. The magnetic field of the active region has the general form of an arch (of height of order 10^5 km) connecting the two regions of opposite polarity,[12] and near the surface the field is approximately vertical.[23] Flaring

activity occurs seldomly or not at all. The magnetic field strength in the active region is comparable to the field strength in the chromospheric network in quiet regions, and the chromospheric emission from the active region is similarly enhanced. The low corona is rather uniformly enhanced over the active region. Thus, during the decay phase, an active region is characterized by a general enhancement in the magnetic field and in emission (and, hence, in density and temperature) over that of the normal atmosphere, but retains a layered structure similar to that of the normal atmosphere.

Since an active region is in the decay phase for all but a small fraction of its life, at any one time the average active region on the sun will be in the decay phase. XUV-line observations show (see section 4.3.2) that, as in the normal atmosphere, the downward heat flux is approximately constant in the transition region and low corona in the average active region, and that the downward heat flux in active regions is about five times larger than in quiet regions. At any particular temperature level in the active region the number density is also enhanced by about a factor of five. The maximum temperature of the corona is found from the XUV-line data to be 2×10^6 K for quiet regions and 2.5×10^6 K for the average active region.[11,38]

3. THE HEATING PROBLEM

3.1 Introduction

Every layer of the sun, both in the atmosphere and in the interior, continually loses thermal energy. In the interior, each layer loses energy to the adjacent cooler layer above it. In the atmosphere, part of the energy lost from each layer is transferred to other layers, while the remainder is radiated away directly into space. Excluding active regions in their growth phase, all layers of the sun maintain an average steady state on a time scale of weeks or longer. Since this requires that the thermal energy of each layer remains constant, each layer must be continually resupplied with the thermal energy that it loses. More succinctly, the energy equilibrium of each layer demands that the cooling of each layer be balanced by heating. Thus, the heating problem for any layer of the sun is merely half of the problem of the energy balance of the layer.

The heating problem for any layer may be stated in the form of the following two simple but general questions:

1. How much heating is required, i.e., what is the rate at which thermal energy must be supplied?
2. How is the energy supplied?

The first question asks, "What, quantitatively, needs to be explained?", while the second question asks, "What is the explanation?" The first question is more closely connected with observations than with theory; its answer should be sought from quantitative observations of the energy losses of the sun. The second question is basically theoretical; it asks

for a theoretical explanation or solution of the quantitatively defined heating problem requested by the first question.

The answers to the above questions are well established for the interior of the sun. The energy emitted from the photosphere is continually resupplied from the uppermost layer of the interior. This layer, in turn, gains an equal amount of energy from the adjacent interior layer, and so on down to the central core where the energy is released by nuclear fusion. Thus, for any layer of the interior outside of the fusion core, the answer to the first question is trivial: each layer is heated from below and cooled from above at a rate just equal to the total luminosity of the sun. If the layer is convectively unstable, i.e., if the layer is in the convection zone, it can be shown theoretically that the heating is accomplished by convective transfer. If the layer is in the convectively stable region below the convection zone, the heating can be explained on the basis of radiative transfer. Thus, in the general sense of our defining questions, the heating of these interior layers is well understood.

The heating of the convectively stable photosphere is also well understood in terms of radiative equilibrium, each layer of the photosphere being heated by absorption of radiation from below and cooled by the emission of radiation. Hence, the heating of the photosphere is similar to that of the interior layers.

The purpose of the next section is to define observationally the heating problem for the outer layers of the atmosphere by presenting and discussing observational results which bear on our two defining questions. Although there are uncertainties in the observational data so that the heating of the chromosphere, transition region, and corona

cannot be as well defined as for the photosphere and interior, the observations show that the nature of the heating of the outer atmosphere departs radically from that of the photosphere and interior of the sun.

3.2 Observational Requirements

3.2.1 Amount of Heating

3.2.1.1 Chromosphere

Although spicules may represent a conversion of thermal energy into kinetic energy at the top of the chromosphere (Section 4.4), by far the dominant process by which thermal energy is lost from the chromosphere is the emission of radiation. However, the amount of radiation emitted by the chromosphere cannot be measured by direct observation because, even when the photosphere is covered during an eclipse, most of the light observed from the chromosphere is only photospheric light which has been scattered by the chromosphere, and which, therefore, does not represent a true loss of thermal energy from the chromospheric gas.[39,40] Consequently, the rate of emission of radiation must be computed theoretically from the physical conditions in the chromosphere (such as temperature, density, radiation intensities, and the variation of these quantities with height) implied by the observations. That is, the rate of radiative cooling must be estimated from a model of the chromosphere derived from observations. Using this approach, Athay[40] concludes that in the lower chromosphere, where the temperature is less than 6000 K, the radiative losses are dominated by continuum radiation from the negative hydrogen ion, while in the temperature range 6000 to 10,000 K, the losses are due mainly to the Balmer lines and Balmer continuum of neutral hydrogen.

The radiative loss rates are estimated to be 4×10^6 erg cm^{-2} sec^{-1} for the emission from the negative hydrogen ion, and 1×10^6 erg cm^{-2} sec^{-1} for the Balmer lines and continuum. According to Athay, these values estimate the total radiative loss rate from the atmosphere between the temperature minimum and the 10^4 K level (i.e., the region which we define to be the chromosphere) within a factor of 2 or 3.

3.2.1.2 Transition Region and Corona

A key source of observational information pertaining to the heating of the outer layers of the solar atmosphere is the solar XUV spectrum of far-ultraviolet and X-ray radiation in the wavelength range $3000 > \lambda > 1$ Å. Since this portion of the solar spectrum is completely absorbed in the earth's atmosphere, detailed knowledge of the XUV spectrum has been obtained only in the past ten years or so by means of rockets and satellites. In the range 3000 to 2000 Å, the character of the XUV spectrum is similar to that at visible wavelengths in that it consists of a continuum punctuated by absorption lines. Shortward of 2000 Å there are emission lines in the spectrum, but longward of 1400 Å these are weak, and most of the energy in the spectrum is contained in the continuum. However, shortward of 1400 Å, the spectrum rather abruptly changes over to an emission-line spectrum, so that almost all of the energy in the range 1400 to 1 Å is contained in emission lines.[41]

The fact that the emission lines shortward of 1400 Å are strong, i.e., much more intense than the background continuum, requires that most of the energy in these lines is supplied from the thermal kinetic energy of the atmosphere through collisional excitations rather than by scattering from the continuum. The collisional excitation of most of the lines

shortward of 1400 Å requires electron temperatures in excess of 10^4 K,[41,42] which implies that the XUV spectrum shortward of 1400 Å is emitted predominantly from the transition region and corona. The XUV spectrum longward of 1400 Å can be explained by reasonable models of the chromosphere,[43] and there is no evidence for appreciable energy loss from the transition region and corona at these longer wavelengths. Therefore, it is reasonable to estimate the total radiative energy loss rate for the transition region and corona from the XUV spectrum shortward of 1400 Å.

The observed energy flux in the XUV spectrum of the quiet sun has been tabulated for each interval of about 20 Å from 1775 to 1 Å by Hinteregger et al.[43] The total energy flux in the range 1400 to 1 Å is 4.5×10^5 erg cm^{-2} sec^{-1}. This value is probably accurate to within a factor of 2.[43,45] Therefore, making the reasonable assumption that the transition region and corona radiate as much inward, toward the photosphere, as outward, we conclude that in quiet regions the total radiative cooling rate for the transition region and corona is of the order of 10^6 erg cm^{-2} sec^{-1}.

In addition to radiative losses, there are two other obvious energy losses from the outer atmosphere above the chromosphere: the outward energy loss due to the solar wind, and the inward energy loss due to heat conduction to the chromosphere. At the orbit of the earth, the energy carried by the solar wind is in the form of kinetic energy, due to the flow velocity, and gravitational potential energy, due to the fact that the gas carried by the wind has been lifted through the gravitational field of the sun. Since the solar wind is highly supersonic at the orbit of the earth, the thermal energy density is negligible

compared to the kinetic energy density of the streaming motion. The energy flux of the solar wind can therefore be estimated from the mass density ρ_E and velocity v_E of the solar wind observed at the orbit of the earth. The energy flux F_{SW} at the surface of the sun necessary to maintain the solar wind is given by

$$F_{SW} = \rho_E v_E \left[\frac{1}{2} v_E^2 + \frac{GM_\odot}{R_\odot}\right]\left[\frac{R_E}{R_\odot}\right]^2 , \qquad (3.1)$$

where $GM_\odot/R_\odot$ is (very nearly) the gravitational potential energy per unit mass at the orbit of the earth, and R_E is the radius of the earth's orbit. The average number density of ions (mostly protons) at the earth is about 5 cm^{-3}, and the average velocity is not greater than 500 km sec^{-1}.[46] These values give $F_{SW} = 6 \times 10^4$ erg cm^{-2} sec^{-1}, more than an order of magnitude less than the radiative loss rate.

At any point in the atmosphere, the heat flux F_c carried by thermal conduction is given by

$$F_c = -\kappa \frac{dT}{dz} , \qquad (3.2)$$

where κ is the thermal conductivity. At the base of the transition region, the thermal conductivity is approximately 3×10^5 erg sec^{-1} K^{-1} cm^{-1} (see Section 4.2.3), and from Figure 2.2 we estimate the temperature gradient to be about 2×10^{-4} K cm^{-3}. These values give a downward heat flux of only 60 erg cm^{-2} sec^{-1}.

We conclude that the overall energy loss from the transition region and corona is completely dominated by radiative cooling, and that the amount of the energy loss is of the order of 10^6 erg cm^{-2} sec^{-1}.

Most of the radiation from the atmosphere above the chromosphere is emitted from the transition region rather than from the corona. From identifications of the ions and transitions which produce the lines in the XUV spectrum and from estimates of the temperature at which these lines are most likely to be excited, Nikolsky[47] concludes that less than 10^5 erg cm^{-2} sec^{-1} is emitted from the corona. As is shown in Section 4.3.2, the temperature gradient, deduced from XUV line data, at the base of the corona requires that of the order of 10^6 erg cm^{-2} sec^{-1} is lost from the corona by heat conduction to the transition region. Thus, while radiative cooling is the dominant cooling process for the transition region and corona as a whole, conduction cooling is the dominant process for the corona alone.

3.2.2 Source of Heating

A plausible source for the heating of the low chromosphere is the photospheric radiation.[48,49,50] In both the photosphere and the low chromosphere, the absorption and emission of radiation is dominated by the negative hydrogen ion. Using this fact, Cayrel[49] has shown that if the low chromosphere is in radiative equilibrium, then as the number density decreases with height from 10^{16} cm^{-3} at the top of the photosphere, the temperature should increase to about 5600 K. This estimate is in good agreement with Athay's finding,[40] from eclipse continuum observations, that the continuum emission, presumably from the negative hydrogen ion, dominates the cooling of the chromosphere below the 6000 K level. Thus, we may reasonably assume that the low chromosphere, between the temperature minimum and the 6000 K level, is heated primarily by radiative transfer from the photosphere.

In the temperature range 6000 to 10,000 K the chromosphere radiates mainly in hydrogen Balmer lines and continuum. Since the wavelength of this emission is in the visible and near ultraviolet, and since a gas is a good absorber at the same wavelengths that it is a good emitter, some thermal energy should be supplied to the chromosphere in the 6000 to 10,000 K range by absorption of photospheric radiation at these wavelengths. But it seems unlikely that this could be the dominant energy supply since the effective temperature of the photospheric radiation is only 5800 K. However, the chromosphere receives of the order of 5×10^5 erg cm^{-2} sec^{-1} in the form of XUV radiation from the transition region and corona. This is of the same order as the 10^6 erg cm^{-2} sec^{-1} estimated by Athay[40] for the energy loss of the 6000 to 10,000 K region. Therefore, the observations indicate that a significant, and possibly dominant, portion of the heating of the chromosphere is due to radiative transfer from above and below.

In Section 3.2.1.2 we saw that the observations of the transition region and corona indicate that the rate of radiative cooling of the transition region and the rate at which heat is supplied to the transition region by conduction from the corona are each of the order of 10^6 erg cm^{-2} sec^{-1}. From this we may conclude that much, and possibly most, of the energy lost from the transition region is supplied by heat conduction from the corona. Thus, the observations admit the possibility that the heating of the chromosphere and transition region can be understood in terms of ordinary heat transfer processes, namely radiative transfer and heat conduction, and that the dominant source of the energy supplied to the upper chromosphere and transition region is the thermal energy of the corona. On the other hand, the corona obviously cannot be heated by

ordinary heat transfer from the other layers of the atmosphere simply because the temperature is higher in the corona than in any other layer. Therefore, the fundamental question posed by the observations with respect to the heating of the outer solar atmosphere is "How is the corona heated?"

3.3 Theory

3.3.1 Introduction

The fundamental question of the heating problem for the outer solar atmosphere, "How is the corona heated?", is answered by the fundamental hypothesis of the heating theory for these layers: the corona is heated by the dissipation of mechanical energy which is supplied to the corona in the form of mechanical waves, e.g., sound waves, which are emitted from the top of the convection layer. The function of the convection layer is to transport the total energy lost from the sun [$(6.41 \pm 0.04) \times 10^{10}$ erg cm^{-2} sec^{-1} at the surface] from the inner, convectively stable, layers to the surface layers of the sun. At the top of the convection layer almost all of the convective energy flux is converted into the flux of radiation which is emitted from the photosphere. The basic idea of the heating of the corona is that a minute fraction of the convective energy flux (less than 0.1 percent) goes into mechanical waves which propagate into the outer layers of the atmosphere. Some fraction of this mechanical energy flux reaches the corona where it is dissipated into thermal energy. Thus, the mechanical energy flux emitted from the top of the convection layer is the added ingredient which makes the heating of the outer layers of the solar atmosphere essentially different from the heating of the photosphere and interior.

The problem of the heating of the atmosphere by mechanical waves divides naturally into three main aspects: the generation, the propagation, and the dissipation of the waves. The generation of the mechanical waves by the "turbulent" motion at the top of the convection zone determines the magnitude of the mechanical energy flux entering the base of the atmosphere and the energy spectrum of the waves which carry this energy flux. The wave-propagation properties of the atmosphere govern the flow of the mechanical energy through the atmosphere. In particular, the propagation properties determine which waves are reflected in the atmosphere below the corona and which waves are passed to the corona. The dissipation process transforms the mechanical energy into thermal energy and determines how and where the wave energy is deposited in the atmosphere. The three aspects of the problem are not completely independent. For example, the dissipation in the corona depends on the type of wave to be dissipated, and the type of wave which reaches the corona, in turn, depends on the nature of the generation and propagation.

The structure of the atmosphere is the configuration assumed by the atmosphere in response to the heating. Two of the most outstanding features of the structure are the high temperature of the corona and the sharp inward drop of the temperature in the transition region. The existence of both of these features can be broadly understood by considering the response of the atmosphere to the mechanical-wave heating outlined above.

Although heat conduction is important for the transfer of thermal energy from the corona to the transition region, the heating of the atmosphere is ultimately balanced by radiative cooling; the thermal energy of the atmosphere is converted to radiation through inelastic collisions.

The collision rate is proportional to the product of the number densities of the colliding particles and to the relative velocity of the particles (assuming for simplicity that the collision cross-section is velocity-independent). Therefore, for a given mass of gas at a given temperature, the rate at which the gas can radiate is limited by the number density as well as by the total number of particles, or the quantity of the gas. On the other hand, the rate of emission of radiation should increase without limit if the temperature is increased without limit. In an isothermal atmosphere, the number density and the total number of particles above a given level decrease exponentially with height. All of this implies that the basic reason that the temperature of the corona is much higher than the temperature of the photosphere is that the mechanical waves supply thermal energy to the outer fringes of the atmosphere at a greater rate than the limited number of available atoms could balance by radiative cooling if the temperature of the outer atmosphere was comparable to that of the photosphere. As a result, the temperature rises until the heating can be balanced by radiative cooling or by some other cooling process. In the case of the corona, there are two cooling processes in addition to radiative cooling: cooling by downward heat conduction and cooling due to energy lost to the solar wind. As was pointed out in Section 3.2.1.2, conduction cooling is the dominant process. Thus, the temperature of the outer atmosphere rises to the vicinity of 10^6 K at which temperature the thermal conductivity becomes so high that the 10^6 erg cm^{-2} sec^{-1} supplied by the mechanical waves can be removed by heat conduction back to the cooler layers which contain enough material and are dense enough to radiate away an energy input of this magnitude.

The sharp inward drop of the temperature in the transition region results from the approximate constancy of the downward heat flux in the transition region above the 10^5 K level (see Section 2.2.1). This implies that the density and the total amount of material above the 10^5 K level are too small to radiate away the thermal energy supplied to the atmosphere above this level. Thus, it appears that the existence of both the high-temperature corona and the narrow transition region result from the inability of the outer atmosphere to radiate away the energy supplied to it by the mechanical waves.

The preceding discussion shows that the structure of the atmosphere is a result of the heating. On the other hand, as we shall see in Section 3.3.2, the increase of the temperature with height in the atmosphere causes some waves (depending on the initial direction of the wave and its period) to be refracted toward the vertical while others are refracted away from the vertical and reflected back toward the photosphere and convection layer. Also, the manner in which the density decreases and the temperature increases with height governs the amplitude of the waves as they propagate through the atmosphere, and the amplitude of the waves is directly connected with their dissipation. In the transition region, where conduction heating is important, the heating, locally, depends on the shape of the temperature profile. Thus, although the structure of the atmosphere is appropriately considered to be a result of the heating, the heating of the atmosphere is not independent of the structure.

3.3.2 Propagation

3.3.2.1 General Considerations and Formulation

Before either the emission or the dissipation of the mechanical energy flux can be studied quantitatively, or even qualitatively, it is necessary to have some knowledge of the nature of the waves being emitted or dissipated. The possible types of waves are determined by the propagation properties of the atmosphere. Hence, of the three basic processes of generation, propagation and dissipation, the propagation plays a central role and is fundamental to the other two processes.

The manner in which a disturbance is propagated in a mass of gas depends on what forces act on the gas. In the solar atmosphere the forces are due to the sun's gravity, the gas pressure, and the magnetic field. (Order of magnitude estimates of the viscosity and the velocity gradients in the fluctuating velocity field in the solar atmosphere show that viscous forces should be negligible.) The curve in Figure 3.1 gives the magnetic field strength for which the magnetic pressure would be equal to the gas pressure, i.e., for which

$$\frac{B^2}{8\pi} = nkT \quad , \tag{3.3}$$

as a function of height in the solar atmosphere. The height dependence of the gas pressure was obtained from Figures 2.2 and 2.3. Figure 3.1 shows that for the average magnetic field strength of about 1 gauss for quiet regions, the magnetic pressure is much less than the gas pressure below the transition region. Therefore, in quiet regions, the magnetic field should have only a minor effect on motions below the transition

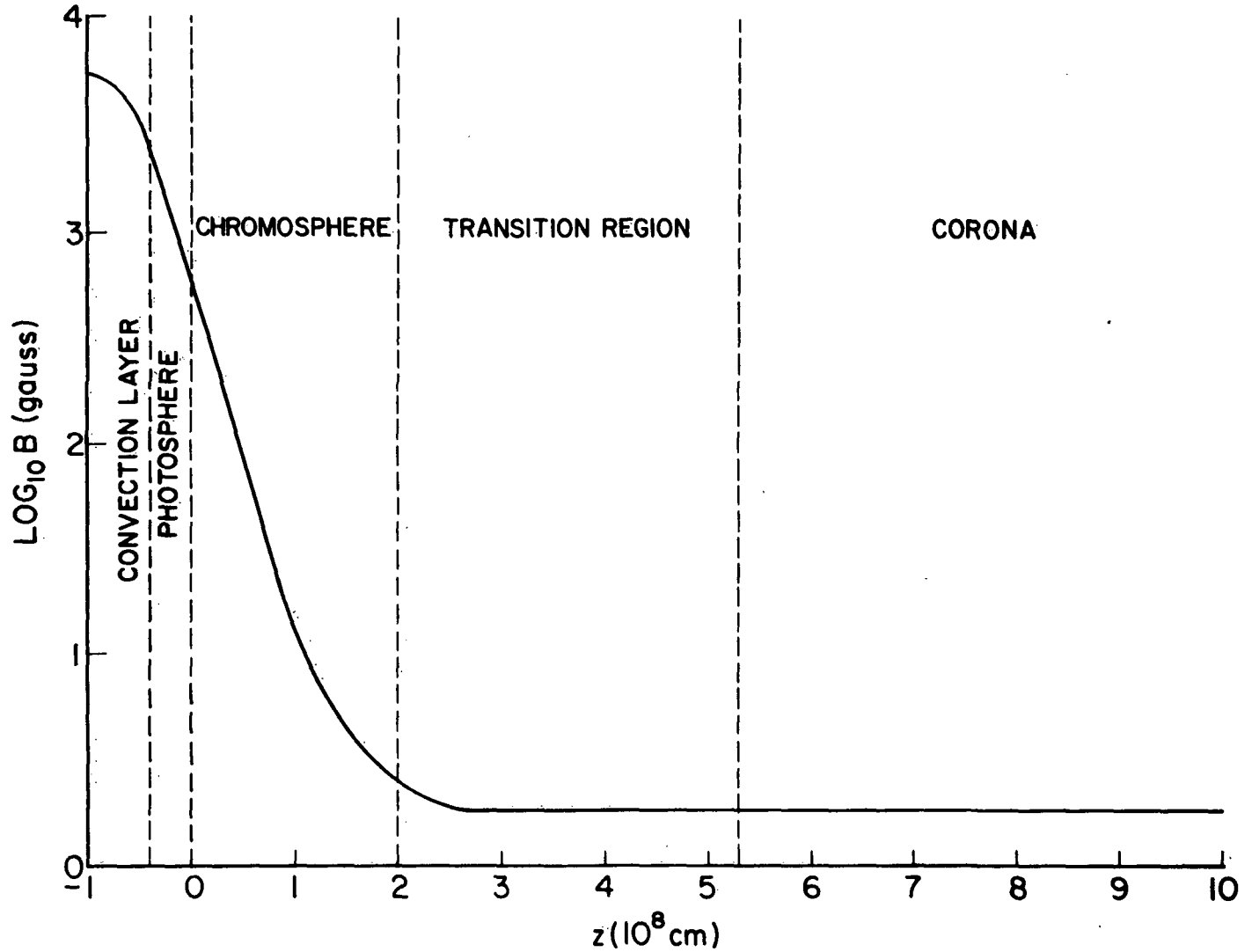


Figure 3.1. MAGNETIC FIELD STRENGTH AT WHICH THE MAGNETIC PRESSURE EQUALS THE GAS PRESSURE IN THE QUIET SOLAR ATMOSPHERE.

region, and neglecting the magnetic field in the equations of motion should not affect the basic character of the propagation in the chromosphere.

In the absence of magnetic forces and viscous forces, or when these forces are negligibly small, the force equation governing the motions in the atmosphere is

$$\rho \frac{d\vec{v}}{dt} = -\nabla p + \rho\vec{g} \quad , \tag{3.4}$$

where ρ is the mass density, $\vec{v}$ is the velocity, and p is the gas pressure. The acceleration of gravity $\vec{g}$ may be assumed to be constant with height since the thickness of the atmosphere between the photosphere and corona is only about 10^{-2} $R_{\odot}$. ρ, $\vec{v}$, and p are functions of the space coordinates x, y, z and time t. $d\vec{v}/dt$ is the total time rate of change of the velocity of a fluid particle:

$$\frac{d\vec{v}}{dt} = \frac{\partial\vec{v}}{\partial t} + (\vec{v}\cdot\nabla)\,\vec{v} \quad . \tag{3.5}$$

Equation (3.4) represents three equations which connect the five variables ρ, p, v_x, v_y, v_z. Hence, we need two more equations in order to have a complete set of equations for the motions of the atmosphere. One additional equation is provided by the conservation of mass, which is expressed by

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\vec{v}) = 0 \quad . \tag{3.6}$$

We now proceed to obtain a final equation connecting p, ρ, and $\vec{v}$ which represents the compressibility of the flow.

By the first law of thermodynamics, the rate of change of the internal energy of a fluid particle is given by[50.1]

$$\frac{de}{dt} = q - \frac{p}{\rho} \nabla \cdot \vec{v} \qquad (3.6.1)$$

where e is the internal energy per unit mass, q is the rate of heat transfer to the fluid particle per unit mass, and the term $-(p/\rho)\nabla \cdot \vec{v}$ is the rate of work done on the fluid particle per unit mass due to the compression of the fluid particle by the surrounding fluid. The general situation which we wish to consider is that of an atmosphere which is perturbed from equilibrium by mechanical disturbances at its base. In the unperturbed atmosphere each term of the energy equation (3.6.1) is zero. Perturbations in the internal energy e are produced through the work term in the energy equation by compressions and expansions in the fluid motions produced in the atmosphere by the disturbances. Since, for a perfect gas, the internal energy is proportional to the temperature, a fluid particle in which the internal energy is perturbed from its equilibrium value will be hotter or colder than required for equilibrium. If the atmosphere is thermally stable, which we assume it is, the heat transfer will act to smooth out the internal energy perturbations produced by the compressions and expansions.

In the limiting case in which the time scale for heat transfer is much longer than the time scale of the compressions and expansions, there is negligible heat transfer. That is, the motions become adiabatic $(q = 0)$ and the energy equation becomes

$$\frac{de}{dt} = - \frac{p}{\rho} \nabla \cdot \vec{v} \qquad (3.6.2)$$

In the opposite extreme, in which the time scale for smoothing the temperature perturbations is much shorter than the time scale of the motion, the heat transfer completely cancels the effect of the compressions and expansions on the internal energy, so that the motion is isothermal and

$$\frac{de}{dt} = 0 \quad . \tag{3.6.3}$$

The equation of state for a perfect gas may be written

$$p = (\gamma - 1)\, \rho e \quad , \tag{3.6.4}$$

where γ is the ratio of specific heats ($\gamma = c_p/c_v$) and has a value greater than 1 ($\gamma = 5/3$ for monatomic gas). In the adiabatic limit, the energy equation can be combined with the equation of state to yield

$$\frac{dp}{dt} = \gamma \frac{p}{\rho} \frac{d\rho}{dt} \quad . \tag{3.6.5}$$

This equation expresses the compressibility of the flow; it specifies the density change which results from a pressure change. In the isothermal limit, the energy equation and the equation of state give

$$\frac{dp}{dt} = \frac{p}{\rho} \frac{d\rho}{dt} \quad . \tag{3.6.6}$$

Since $\gamma > 1$, we see that in rendering the flow isothermal, the heat transfer also results in the flow being more compressible.

In general, in addition to affecting the compressibility of the flow, the heat transfer results in dissipation of the fluid motion. In the limit of adiabatic flow, there is no dissipation because there is

no heat transfer. In the opposite limit of isothermal flow there is no dissipation because the heat transfer takes place isothermally and, hence, reversibly. However, for intermediate cases in which the heat-transfer time is comparable to the compression time, there is dissipation because the flow is not isothermal and the heat transfer is not reversible. (In Section 3.3.3 the manner in which the dissipation of sound waves varies with the ratio of the compression time to the heat-transfer time is shown in detail for the case of heat transfer by thermal conduction.)

Our present purpose is to study the propagation of disturbances in the solar atmosphere rather than their dissipation. As we can see from the limiting cases of adiabatic motion and isothermal motion, the propagation depends on the compressibility of the atmosphere. Hence, in our governing equations we would like to ignore the dissipation resulting from the heat transfer, but take account of the effect of the heat transfer on the compressibility. We can accomplish this aim to a certain extent by adopting the following polytrope law connecting the pressure and density:

$$\frac{dp}{dt} = \alpha \frac{p}{\rho} \frac{d\rho}{dt} , \qquad 1 \leq \alpha \leq \gamma , \tag{3.7}$$

where the polytrope index is assumed to be constant throughout the flow. This equation, along with the force equation and the continuity equation, forms a complete set of equations for the motions of the atmosphere. When $\alpha = 1$ or $\alpha = \gamma$, we have isothermal or adiabatic flow, respectively, and the polytrope law is exact. Since there is no dissipation in these limits, there will be no dissipation in the flow computed from the governing equations using the polytrope law with any other constant

value of α since the form of the equation is not changed. For flows in which there is irreversible heat transfer, the ratio $(\rho/p)(dp/d\rho)$ defined by α in the polytrope law will not actually be constant throughout the flow. For these flows, a constant value of α in the polytrope law is taken to represent an "average value" of this ratio in the flow. For these flows, which are intermediate between adiabatic and isothermal, it is reasonable to expect this average value of $(\rho/p)\cdot(dp/d\rho)$ to be intermediate between 1 and γ, as we have assumed in writing the polytrope law. It should be emphasized that for intermediate flows, the polytrope law is only an approximate, reasonable representation of the effect of the heat transfer on the compressibility. However, in the following we will be mainly concerned with the limiting cases of adiabatic and isothermal propagation in which cases the polytrope law is exact.

The following considerations suggest that it is physically permissible to facilitate the analysis of the propagation by first linearizing the above equations of motion. Since the atmosphere is compressible, one type of wave to be expected is a compression wave or sound wave. If the compressions and expansions in the sound waves obey the polytrope law [equation (3.7)], then the sound speed a is given by

$$a = \left(\alpha \frac{p}{\rho}\right)^{1/2} , \tag{3.8}$$

or, from the equation of state,

$$p = \rho \frac{k}{m} T , \tag{3.9}$$

$$a = \left(\alpha \frac{k}{m} T\right)^{1/2} . \qquad (3.10)$$

In the solar atmosphere, $T \gtrsim 4.6 \times 10^3$ K, and the mass of the hydrogen atom is an approximate upper bound on the mean particle mass $m \lesssim m_H = 1.66 \times 10^{-24}$ g; so

$$a \gtrsim 6 \times 10^5 \text{ cm sec}^{-1} \qquad (3.11)$$

throughout the solar atmosphere, where the approximate equality holds only near the top of the photosphere. The mechanical energy flux F_m carried by one-dimensional sound waves is

$$F_m = \rho \overline{v^2}\, a \ , \qquad (3.12)$$

where $\overline{v^2}$ is the mean square velocity in the compressions and expansions of the waves. As we saw in Section 3.2, the total rate of heating required for the entire atmosphere above the photosphere is of order 10^7 erg cm^{-2} sec^{-1}. For a sound-wave energy flux of 10^7 erg cm^{-2} sec^{-1} at the top of the photosphere, equation (3.12) gives

$$v_{rms} = 2.5 \times 10^4 \text{ cm sec}^{-1} \ , \qquad (3.13)$$

which is of the order of the observed rms velocity at the top of the photosphere (Section 2.2.2.1) and more than an order of magnitude less than the sound velocity. On this basis, it appears that the motions in the photosphere and chromosphere should be fairly well characterized by the linearized equations of motion.

The unperturbed atmosphere is in hydrostatic equilibrium:

$$\vec{v} = 0 \quad , \tag{3.14.1}$$

$$p = p_o(z) \quad , \tag{3.14.2}$$

$$\rho = \rho_o(z) \quad , \tag{3.14.3}$$

$$\frac{dp_o}{dz} = -\rho_o g \quad . \tag{3.14.4}$$

In the perturbed atmosphere,

$$\vec{v} = \vec{v}(x,y,z,t) \quad , \tag{3.15.1}$$

$$p = p_o(z) + \delta p(x,y,z,t) \quad , \tag{3.15.2}$$

$$\rho = \rho_o(z) + \delta\rho(x,y,z,t) \quad . \tag{3.15.3}$$

We linearize the equations of motion in the standard way by expressing them in terms of the perturbation quantities $\vec{v}$, δp, and $\delta\rho$, and then removing all terms which are of second or higher order in these quantities and their space and time derivatives. The resulting linearized equations of motion are:

linearized force equation:

$$\rho_o \frac{\partial \vec{v}}{\partial t} = -\nabla(\delta p) + \delta\rho\vec{g} \quad , \tag{3.16}$$

linearized mass equation:

$$\frac{\partial(\delta\rho)}{\partial t} = -\vec{v} \cdot \nabla\rho_o - \rho_o \nabla\cdot\vec{v} \quad , \tag{3.17}$$

linearized polytrope law:

$$\frac{\partial(\delta p)}{\partial t} = a_o^2 \frac{\partial(\delta\rho)}{\partial t} + a_o^2 \vec{v} \cdot \nabla\rho_o - \vec{v} \cdot \nabla\rho_o \quad , \tag{3.18}$$

where

$$a_o^2 = \alpha \frac{p_o}{\rho_o} \quad . \tag{3.18.1}$$

Upon combining the linearized mass equation and the linearized polytrope law with the linearized force equation to eliminate δp and $\delta\rho$, we obtain the following linear equation for the velocity field,

$$\frac{\partial^2\vec{v}}{\partial t^2} = a_o^2\nabla(\nabla\cdot\vec{v}) + (\alpha - 1)\vec{g}(\nabla\cdot\vec{v}) + \nabla(\vec{v} \cdot \vec{g}) \quad . \tag{3.19}$$

In the absence of gravity, equation (3.19) reduces to the wave equation for ordinary sound waves.

The components of equation (3.19) are:

$$\left.\begin{aligned}
\frac{\partial^2 v_x}{\partial t^2} &= a_o^2\left[\frac{\partial^2 v_x}{\partial x^2} + \frac{\partial^2 v_y}{\partial x\partial y} + \frac{\partial^2 v_z}{\partial x\partial z}\right] - g\frac{\partial v_z}{\partial x} \quad , \\
\frac{\partial^2 v_y}{\partial t^2} &= a_o^2\left[\frac{\partial^2 v_x}{\partial y\partial x} + \frac{\partial^2 v_y}{\partial y^2} + \frac{\partial^2 v_z}{\partial y\partial z}\right] - g\frac{\partial v_z}{\partial y} \quad , \\
\frac{\partial^2 v_z}{\partial t^2} &= a_o^2\left[\frac{\partial^2 v_x}{\partial z\partial x} + \frac{\partial^2 v_y}{\partial z\partial y} + \frac{\partial^2 v_z}{\partial z^2}\right] - g\frac{\partial v}{\partial z} - (\alpha-1)g\left[\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right].
\end{aligned}\right\} \tag{3.19.1}$$

3.3.2.2 Propagation in an Isothermal Atmosphere

In the solar atmosphere, the temperature, and hence the sound speed, varies with height in the atmosphere. However, much can be learned about wave propagation in the solar atmosphere by considering wave propagation in an isothermal atmosphere. In this case, the sound speed is constant, and the partial differential equations (3.19.1) may be Fourier transformed in x, y, z, and t to obtain an algebraic dispersion relation connecting the wave numbers k_x, k_y, k_z, and the radian frequency ω. That is, any velocity field obeying equations (3.19.1) can be considered to be made up of a distribution of sinusoidal plane waves of the form

$$e^{i(k_x x + k_y y + k_z z + \omega t)} ,$$

the distribution being given by the Fourier transform of the velocity field. The dispersion relation given by the transformed version of equations (3.19.1) prescribes the possible waves permitted by the equations of motion, i.e., the possible waves permitted by the propagation properties of the atmosphere.

The following equations are obtained by Fourier transforming equations (3.19.1):

$$\left.\begin{aligned}
&\left(a_o^2 k_x^2 - \omega^2\right) V_x + a_o^2 k_x k_y V_y + \left(a_o^2 k_x k_z + igk_x\right) V_z = 0 \quad , \\
&a_o^2 k_y k_x V_x + \left(a_o^2 k_y^2 - \omega^2\right) V_y + \left(a_o^2 k_y k_z + igk_y\right) V_z = 0 \quad , \\
&\left[a_o^2 k_z k_x + i(\alpha - 1)gk_x\right] V_x + \left[a_o^2 k_z k_y + i(\alpha - 1)gk_y\right] V_y + \left[a_o^2 k_z^2 - \omega^2 + i\alpha g k_z\right] V_z = 0
\end{aligned}\right\} \qquad (3.20)$$

where V_x, V_y and V_z are the Fourier transforms of v_x, v_y and v_z, respectively. In order for these three homogeneous equations to have a solution for V_x, V_y, V_z other than $V_x = V_y = V_z = 0$, the determinant of the coefficient matrix must be zero:

$$\begin{vmatrix} \left(a_o^2 k_x^2 - \omega^2\right) & a_o^2 k_x k_y & \left(a_o^2 k_x k_z + igk_x\right) \\ a_o^2 k_y k_x & \left(a_o^2 k_y^2 - \omega^2\right) & \left(a_o^2 k_y k_z + igk_y\right) \\ \left[a_o^2 k_z k_x + i(\alpha - 1)gk_x\right] & \left[a_o^2 k_z k_y + i(\alpha - 1)gk_y\right] & \left[a_o^2 k_z^2 - \omega^2 + i\alpha gk_z\right] \end{vmatrix} = 0 \quad . \tag{3.21}$$

Expansion of the determinant gives

$$\omega^4 - \left[a_o^2\left(k_h^2 + k_z^2\right) + i\alpha gk_z\right]\omega^2 + (\alpha - 1)\, gk_h^2 = 0 \quad , \tag{3.22}$$

where k_h is the horizontal wavenumber,

$$k_h^2 \equiv k_x^2 + k_y^2 \quad . \tag{3.23}$$

Equation (3.22) is the dispersion relation.

In general, the wave numbers k_x, k_y, k_z and the radian frequency ω of a sinusoidal plane wave may be complex, in which case the amplitude of the wave grows or decays exponentially in space and time. Since we have ignored dissipation in deriving the equation of motion (3.19), the plane waves permitted by this equation should have constant amplitude at a given point in space. Therefore, ω is real in the dispersion relation (3.22). In addition, since the physical quantities in

the atmosphere are constant on horizontal planes, the amplitude of a given plane wave should be independent of x and y. Hence, k_h is real in equation (3.22). On the other hand, since the mass density decreases exponentially with height z, we expect the wave amplitude to increase exponentially with z in order that the energy flux of the wave remain constant. Hence, k_z should have an imaginary component. Therefore, in the dispersion relation (3.22) we set

$$k_z = k_{zr} + ik_{zi} \quad , \tag{3.24}$$

and require k_h and ω to be real. The real part of the dispersion relation is

$$\omega^4 - \omega^2 a_o^2 k_h^2 - \omega^2 a_o^2 \left(k_{zr}^2 - k_{zi}^2 \right) + \alpha g \omega^2 k_{zi} + (\alpha - 1)\, g^2 k_h^2 = 0 \quad , \tag{3.25}$$

and the imaginary part is

$$k_{zi} = - \frac{\alpha g}{2a_o^2} = - \frac{1}{2H} \quad , \tag{3.26}$$

where H is the density scale height in the isothermal atmosphere:

$$\rho = \left(\rho \big|_{z=0} \right) e^{-z/H} \quad , \tag{3.27}$$

$$H = \frac{kT}{mg} \quad . \tag{3.28}$$

Thus, we see that for all plane waves of the form

$$e^{i(k_x x + k_y y + k_z z + \omega t)}$$

permitted by the equation of motion (3.19) in an isothermal atmosphere, the velocity increases exponentially with height, so that the kinetic energy density $(1/2)\ \rho\overline{v^2}$ remains constant.

Additional physical insight about the propagation properties of an isothermal atmosphere can be gained by considering the ω^2, k_h^2 diagram of the dispersion relation. If we use equation (3.26) to replace k_{zi} in equation (3.25) and drop the subscript r on k_{zr}, we obtain for the dispersion relation

$$(2Hk_z)^2 = \left(\frac{\omega^2}{\omega_2^2} - 1\right)(2Hk_h)^2 + \frac{\omega^2}{\omega_1^2} - 1 \quad , \tag{3.29}$$

where

$$\omega_1^2 = \frac{\alpha^2 g^2}{4a_o^2} = \left(\frac{a_o}{2H}\right)^2 \quad , \tag{3.30}$$

and

$$\omega_2^2 = \frac{4(\alpha - 1)}{\alpha^2}\ \omega_1^2 \quad . \tag{3.31}$$

Any sinusoidal plane wave is specified by its values of ω, k_h and k_z. We see from the dispersion relation (3.29) that in order to specify a particular wave which propagates in an isothermal atmosphere it is enough to specify ω and k_h. For if ω and k_h are given, then the value of k_z is dictated by the dispersion relation. In other words, there is a one-to-one correspondence between the waves permitted by the dispersion relation and the points in the ω^2, k_h^2 plane.

The ω^2, k_h^2 diagram obtained from the dispersion relation (3.29) is shown in Figure 3.2. Here we have assumed the wave motions to be adiabatic, i.e., $\alpha = \gamma = 5/3$. The diagram is valid for any temperature, the scale of the axes being determined by the temperature. In the unshaded region separating the two shaded areas, $k_z^2 < 0$, so that k_z is imaginary. Therefore, since the real part of k_z is zero, the direction of propagation is purely horizontal for all waves having combinations of ω and k_h which fall in the unshaded region. Any wave which propagates upward (or downward) in the atmosphere has its (ω^2, k_h^2) point in the shaded regions where $k_z^2 > 0$. In connection with the heating of the solar atmosphere, we are mainly interested in waves which propagate upward.

In the high-frequency limit, the dispersion relation becomes

$$k^2 = a_o^2 \omega^2 \quad , \tag{3.32}$$

where k is the total wave number,

$$k^2 = k_h^2 + k_z^2 \quad . \tag{3.33}$$

Equation (3.32) is the dispersion relation for ordinary sound waves. This means that high frequency disturbances propagate as sound waves in the atmosphere. Hence, a pressure pulse will propagate at the sound speed a_o, which is constant throughout the atmosphere. With this in mind, for each wave period $P = 2\pi/\omega$, we define a length

$$\lambda_s \equiv a_o P \quad , \tag{3.34}$$

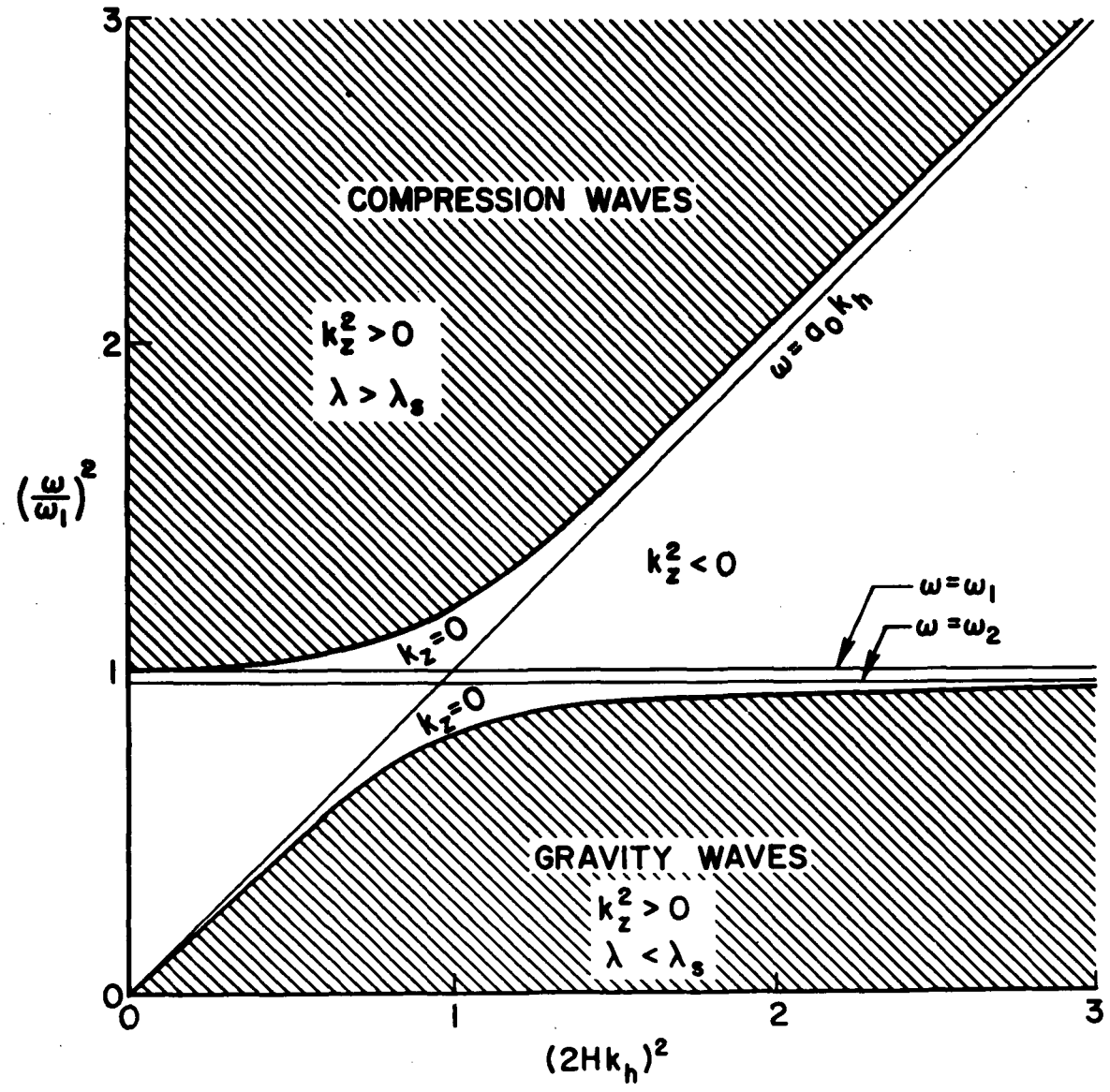


Fig. 3.2. ω^2, k_h^2 DIAGRAM FOR WAVES IN AN ISOTHERMAL ATMOSPHERE.

which is the distance traveled by sound in time P. λ_s is roughly the distance in the atmosphere over which pressure perturbations are smoothed out in time P. That is, a pressure difference can exist across distances much shorter than λ_s only for times much shorter than P, while pressure differences can be maintained across distances much longer than λ_s for times much longer than P. The pressure differences referred to here are due to local compressions and rarefactions in the atmosphere and are in addition to the pressure differences due to the decrease in pressure with height in the undisturbed atmosphere.

Now consider the upward-propagation region below $\omega = \omega_2$ in the ω^2, k_h^2 diagram, Figure 3.2. The dispersion relation (3.29) shows that in this region, for any given k_z, the frequency ω increases to ω_2 as k_h increases without limit. The curve for the particular case of $k_z = 0$ is shown in Figure 3.2. From the definition of the wave number ($k \equiv 2\pi/\lambda$), we know that the wavelength λ goes to zero as k_h increases without limit. Since ω remains finite while λ approaches zero, we see that

$$\lambda \ll \lambda_s$$

for waves with large k_h in the region of upward-propagation below $\omega = \omega_2$. This means that each fluid element remains in pressure equilibrium with its surroundings (i.e., the ambient atmosphere) as it undergoes the oscillatory motions of these waves. This is borne out by the fact that the limiting frequency ω_2 of these waves is just the natural frequency of oscillation of a fluid element which remains in pressure equilibrium with the ambient atmosphere as it oscillates about its equilibrium position.[51] The restoring force for these oscillations

is completely due to buoyancy rather than to pressure forces due to compression. Since the buoyancy force results from the stratification of the atmosphere due to the presence of gravity, the waves having $\lambda \ll \lambda_s$ in the upward-propagation region below $\omega = \omega_2$ are called gravity waves.[52]

From the dispersion relation (3.29) it can be shown that

$$\lambda < \lambda_s$$

throughout the upward-propagation region below $\omega = \omega_2$. Therefore, buoyancy forces are important throughout this region, but compression forces are not completely negligible for the longer wavelengths. Hence, the waves in this region are gravity waves modified by compression.[52]

Throughout the upward-propagation region above $\omega = \omega_1$, we find that

$$\lambda > \lambda_s \quad .$$

Therefore, compression forces are important throughout this region. As ω increases without limit the waves become sound waves with $\lambda = \lambda_s$. For these waves the restoring force is entirely due to compression and the presence of gravity does not affect the oscillations. If gravity were absent, the waves would be sound waves at all frequencies. Hence, λ is greater than λ_s in the upward-propagation region above $\omega = \omega_1$ due to the presence of gravity, and the waves in this region are compression waves (sound waves) modified by gravity. The dispersion relation shows that λ/λ_s increases as ω_1 decreases until at $\omega = \omega_1$, $\lambda/\lambda_s = \infty$ and upward propagation of compression waves is no longer possible.

3.3.2.3 Refraction in an Atmosphere in which the Temperature Varies Slowly with Height

Since ω_1, ω_2, and H depend on temperature, the value of k_z determined by the dispersion relation (3.29) for given ω and k_h depends on the temperature of the atmosphere. This means that the wavelength and the direction of propagation of a given (ω, k_h) wave in an isothermal atmosphere depends on the temperature of the atmosphere. By definition, the temperature varies spatially in a nonisothermal atmosphere. If the temperature variations are gradual enough, then the isothermal dispersion relation may be used at each point in the nonisothermal atmosphere. That is, the wavelength and direction of a wave as it propagates through the nonisothermal atmosphere is given at each point by the isothermal dispersion relation for the temperature at that point in the nonisothermal atmosphere. This is a good approximation if the wavelength changes slowly enough as the wave propagates through the atmosphere. The wavelength (and hence the temperature) varies slowly enough if[53]

$$\left|\frac{d\lambda}{ds}\right| \ll 1 \quad , \tag{3.35}$$

where s is the space coordinate in the direction in which the wave is propagating. In this case, the dispersion relation (3.29) governs the refraction of waves in the nonisothermal atmosphere.

In an atmosphere in which the temperature varies slowly with height but remains constant on horizontal planes, the frequency ω and horizontal wavenumber k_h of a wave remain constant, and only the vertical wavenumber k_z varies as the wave propagates.[54] Hence, the refraction of the wave results from the temperature dependence of k_z

given by the dispersion relation (3.29). To bring out the temperature dependence of k_z, we rewrite the dispersion relation as follows:

$$(2H_1k_z)^2 = -(2H_1k_h)^2 + \left[1 + \frac{4(\alpha - 1)}{\alpha^2}(2H_1k_h)^2\right]\left(\frac{T_1}{T}\right) - \left(\frac{T_1}{T}\right)^2 , \quad (3.36)$$

where T_1 is the temperature at which $\omega = \omega_1$,

$$T_1 = \frac{\alpha mg^2}{4k\omega^2} , \quad (3.37)$$

and H_1 is the scale height at T_1,

$$H_1 = \frac{kT_1}{mg} . \quad (3.38)$$

The dispersion relation (3.36) shows that each wave remains on a curve of constant (ω, k_h) in the k_z^2, $1/T$ plane as the wave propagates through the atmosphere.

The k_z^2, $1/T$ diagram for the dispersion relation (3.36) is shown in Figure 3.3 for the case of adiabatic wave motions $(\alpha = \gamma = 5/3)$. The scale of the axes is determined by the choice of ω, so that the diagram is valid for any ω. Each of the representative curves of constant (ω, k_h) is labeled with its value of $(2H_1k_h)^2$. It is seen that the value of k_h increases counter-clockwise from curve to curve from 0 (vertical waves) to ∞ (horizontal waves). If the temperature T is greater than T_1, then ω is greater than ω_1, and if T is less than $[4(\alpha - 1)/\alpha^2]T_1 = (24/25)T_1$, then ω is less than ω_2. Hence, in the high-temperature region defined by $T_1/T < 1$, the waves are

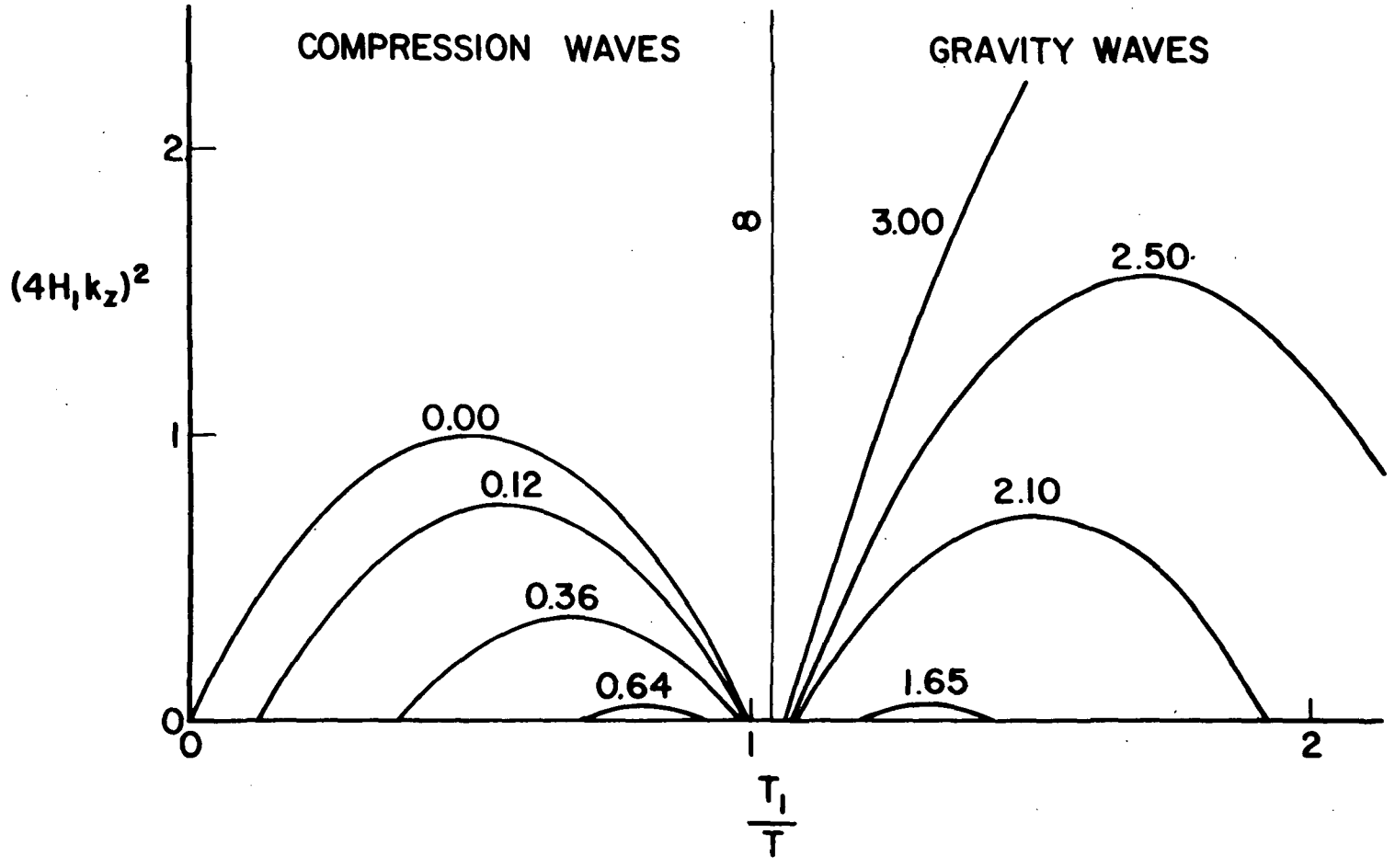


Figure 3.3. k_z^2, 1/T DIAGRAM FOR REFRACTION IN AN ATMOSPHERE IN WHICH THE TEMPERATURE VARIES SLOWLY WITH HEIGHT.

compression waves, and in the low-temperature region defined by $T_1/T > 25/24$, the waves are gravity waves. As a wave propagates upward through an atmosphere with increasing temperature with height, the point in the k_z^2, $1/T$ diagram corresponding to the wave moves to the left along its curve of constant (ω, k_h). If the temperature continues to increase, then at a particular value of the temperature, determined by ω and k_h, the curve intersects the $k_z = 0$ axis, and the wave is reflected. Thus, the k_z^2, $1/T$ diagram conveniently demonstrates the refraction of waves in an atmosphere with slowly increasing temperature with height. We will use this diagram to examine the refraction properties of the chromosphere.

3.3.2.4 Propagation in the Solar Atmosphere

The oscillatory motions observed in the photosphere and low chromosphere provide direct evidence for the existence of mechanical waves in the solar atmosphere. The observed characteristics of these oscillations are listed in Section 2.2.2.1. An obvious and important question bearing on the mechanical-wave heating of the solar atmosphere is whether these oscillations represent compression waves or gravity waves or some combination of the two. The following considerations indicate that these oscillations are composed predominantly of compression modes rather than gravity modes.

The short-period cut-off $P_2 = 2\pi/\omega_2$ below which gravity waves cannot propagate is given by equations (3.30) and (3.31). In the photosphere and chromosphere ($g = 2.74 \times 10^4$ cm sec^{-2}, $m \approx m_H = 1.66 \times 10^{-24}$ gm), these equations give

$$P_2 = 2.1 \, (\alpha - 1)^{-1/2} \, T^{1/2} \text{ sec} \quad . \tag{3.39}$$

Since the polytrope index α is not greater than $\gamma = 5/3$,

$$P_2 \geq 2.6\ T^{1/2} \text{ sec} . \tag{3.40}$$

At the temperature minimum at the top of the photosphere, $T = 4{,}600$ K, and condition (3.40) gives

$$P_2 \geq 177 \text{ sec} . \tag{3.41}$$

Hence, no gravity wave of period shorter than 177 sec can propagate in the solar atmosphere.

We recall that the cut-off period P_2 is the natural period of oscillation of a fluid element which remains in pressure equilibrium with the ambient atmosphere as it oscillates up and down about its equilibrium position. The restoring force which sustains the oscillation is the buoyancy force which results from the fact that the mass density of the fluid element is either higher or lower than that of the surrounding gas when the fluid element is respectively either above or below its equilibrium position. However, if the motion of the fluid element takes place isothermally, then T as well as p are the same inside and outside of the fluid element throughout the motion. In this case, from the equation of state (3.9), the mass density ρ remains the same inside and outside, and there is no buoyancy force on the fluid element. Thus, we see that in the limit of isothermal motions, i.e., as α approaches 1, the buoyancy restoring force goes to zero, and gravity waves are no longer possible. This is the physical reason why ω_2 goes to zero as α goes to 1 in equation (3.31).

In the photosphere, radiative transfer tends to maintain each level at a certain uniform temperature. If the temperature is perturbed locally, by the passage of a wave for example, then the temperature at this point relaxes back to the background temperature through exchange of radiation with the surroundings. For optically thin temperature perturbations, the radiative relaxation time in the photosphere is of the order of 10 sec.[55] Therefore, oscillations in the photosphere with periods of the order of 10^2 sec or longer should occur very nearly isothermally. On the other hand, we found above that gravity waves would have to have periods longer than 177 sec in order to propagate in the photosphere even if the oscillations were adiabatic. Since oscillations of such long periods should take place isothermally in the photosphere, it appears that gravity waves are not possible in the photosphere. This implies that the oscillations observed in the photosphere are compression modes.

In contrast to the case of gravity waves, the ability of an atmosphere to support compression waves is not essentially changed by the condition of isothermal oscillations. For isothermal oscillations ($\alpha = 1$) the sound speed is merely reduced by a factor of $\gamma^{-1/2}$ from that for the adiabatic case ($\alpha = \gamma$). Likewise, the low-frequency cut-off ω_1 for compression waves is reduced by a factor of $\gamma^{-1/2}$. In the photosphere, the corresponding cut-off period P_1 is

$$P_1 = 4.2\ T^{1/2} \text{ sec} \quad . \tag{3.42}$$

We recall that waves with periods $P \geq P_1$ have a purely imaginary k_z; this means that the phase of these waves is constant with height. Hence,

these waves are evanescent or standing waves. Conversely, waves having $P < P_1$ are propagating waves.

The photospheric oscillations are observed above the granulation cells, which stop at about -300 km in the photosphere. The temperature decreases outward through the photosphere and passes through 5000 K at about -150 km. If, in equation (3.42), we adopt 5000 K as a representative temperature for the layer in which the photospheric oscillations occur, we obtain $P_1 = 296$ sec. In good agreement with the observed photospheric oscillations, this shows that waves with periods in the vicinity of 200 sec should be propagating waves, and waves with periods of about 300 sec should be standing waves. Again, this indicates that the oscillations in the photosphere are compression modes.

Since the photospheric velocity field above the granulation cells is composed primarily of compression-mode oscillations, we conclude that the bulk of the mechanical energy flux emitted from the top of the convection layer enters the base of the chromosphere in the form of compression waves. In the following paragraphs we consider the propagation of these waves in the chromosphere and transition region. Our chief aim is to understand the effect of these regions on the transmission of mechanical energy flux to the corona.

The propagation problem in the chromosphere is that of propagation of compression waves in an atmosphere with increasing temperature with height. The temperature increases slowly enough in the chromosphere that we may use the k_z^2, $1/T$ diagram discussed in Section 3.3.2.3 to obtain a qualitative understanding of the wave propagation in this layer of the atmosphere. This can be shown by using the dispersion relation (3.36) in conjunction with the temperature profile in the chromosphere

to evaluate $|d\lambda/ds|$ in the chromosphere. For wave periods in the range 100 to 300 sec, we find that the condition $|d\lambda/ds| \ll 1$ [condition (3.35)] is sufficiently well satisfied for the k_z^2, $1/T$ diagram to be qualitatively valid.

The k_z^2, $1/T$ diagram for compression waves in the chromosphere is shown in Figure 3.4. This diagram is the same as the compression-wave part of Figure 3.3, except that here the curves of constant (ω, k_h) have been extended below $k_z^2 = 0$ in order to include evanescent waves. Above $k_z^2 = 0$ the waves are propagating waves; below $k_z^2 = 0$ the waves are evanescent. In computing the curves in Figure 3.4, we have assumed the wave motions to be adiabatic in the chromosphere rather than isothermal as in the photosphere. This is reasonable because the radiative relaxation time for temperature perturbations increases with height in the solar atmosphere, reaching a value of about 400 sec at a height of 1000 km in the chromosphere.[55] For wave periods in the range 100 to 300 sec, this indicates that the propagation in the chromosphere should be more nearly adiabatic than isothermal. We therefore adopt the k_z^2, $1/T$ diagram for the case $\alpha = \gamma = 5/3$. For any other possible value of α, $1 \leq \alpha \leq 5/3$, the diagram would not be qualitatively different.

The temperature T_1 is determined by the wave period P:

$$T_1 = \frac{\alpha m g^2}{16\pi^2 k} P^2 = 1.0 \times 10^{-1} P^2 \text{ K} \quad . \tag{3.43}$$

So, for a given temperature range, different wave periods define different T_1/T ranges in the k_z^2, $1/T$ diagram. In the chromosphere, the temperature increases from 4,600 to 10,000 K. The shaded regions in Figure 3.4 are the corresponding T_1/T ranges for the representative wave periods of 100, 200, and 300 sec.

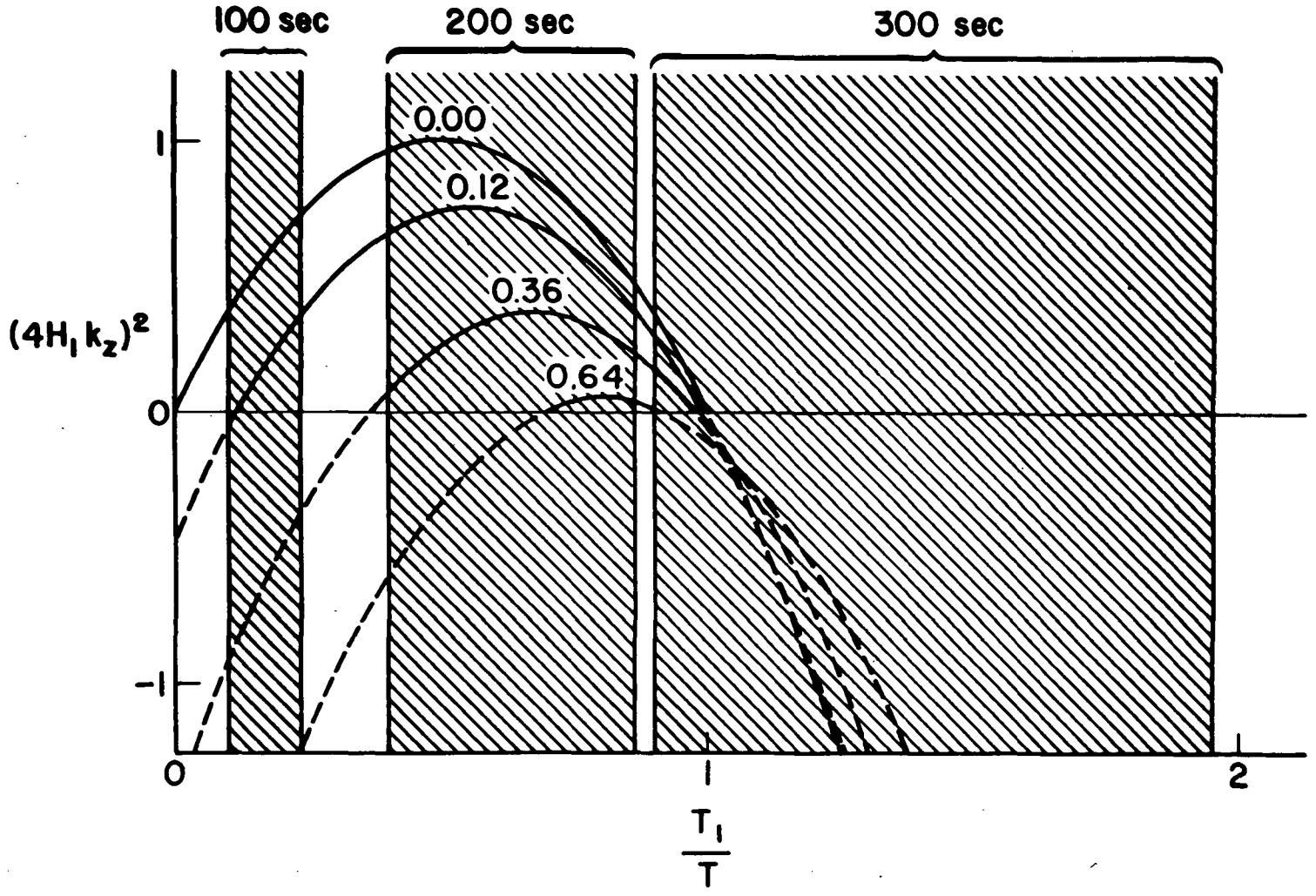


Figure 3.4. k_z^2, 1/T DIAGRAM FOR PROPAGATION OF COMPRESSION WAVES IN THE CHROMOSPHERE.

From equation (3.43) and Figure 3.4 we see that the chromosphere has the following propagation properties:

1. Waves with periods longer than 316 sec are evanescent throughout the chromosphere.

2. Waves having

$$(2H_1 k_h)^2 > 0.64 \tag{3.44}$$

are evanescent throughout the chromosphere. In terms of wave period P and horizontal wavelength λ_h, condition (3.44) is

$$\lambda_h < 5.0 \times 10^3 \, P^2 \text{ cm} \quad . \tag{3.45}$$

Hence, for a wave period of 200 sec only waves having horizontal wavelength longer than 2000 km can propagate in the chromosphere.

3. Waves not satisfying condition (3.44) and having periods less than 214 sec propagate throughout the chromosphere, provided k_h is sufficiently small. If k_h is not sufficiently small, i.e., if the wave does not enter the chromosphere propagating sufficiently vertically, the wave is reflected before reaching the top of the chromosphere.

4. Waves not satisfying condition (3.44) and with periods in the range 214 to 316 sec are evanescent at the bottom of the chromosphere but become propagating waves above some temperature level in the chromosphere.

5. Waves with periods far below cut-off (periods in the vicinity of 100 sec and below) are refracted away from the vertical.

6. Waves with periods near cut-off (periods in the range 200 to 300 sec) are refracted toward the vertical.

7. The farther below cut-off the wave period, the smaller k_h must be in order that the wave not be reflected in the chromosphere. Thus, the chromosphere preferentially passes waves with periods near cut-off.

The observed oscillations at the bottom of the chromosphere are nearly vertical and have periods in the range 200 to 300 sec. Comparing these properties with the above propagation properties of the chromosphere,

we see that conditions are favorable for the transmission of mechanical energy flux through the chromosphere to the transition region.

Next, we consider the transmission through the transition region of the compression waves passed by the chromosphere. The temperature rises steeply at the base of the transition region so that the temperature is much greater than 10^4 K throughout most of the transition region. From equation (3.42) we see that for temperatures much greater than 10^4 K, the cut-off period P_1 for compression waves is much greater than 300 sec. Hence, since the compression waves which pass through the chromosphere should have periods in the range 200 to 300 sec, these waves should behave like sound waves in the transition region.

In the transition region, sound waves of period greater than 200 sec have wavelengths longer than 3000 km, which is about the thickness of the transition region. Since the temperature increases by a factor of 10^2 through the transition region, the sound speed ($a_o \propto T^{1/2}$) increases by a factor of 10. Therefore, it should be a good approximation to treat the transition region as a temperature discontinuity, or interface, between the chromosphere and corona in estimating the transmission of compression waves through the transition region.

Consider a plane interface across which the temperature increases from T_1 to T_2. Let θ be the angle between the normal to the interface and the direction of propagation of a sound wave incident on the interface from the T_1 side. There is a critical angle θ_{crit} for which if $\theta > \theta_{crit}$, there is total reflection of the wave. The critical angle is given by[56]

$$\theta_{crit} = \arcsin\left[\left(\frac{T_1}{T_2}\right)^{1/2}\right] . \qquad (3.46)$$

For the transition region $T_1 = 10^4$ K and $T_2 = 10^6$ K, giving

$$\theta_{crit} = 5.7^\circ .$$

Hence, the waves transmitted to the corona must be propagating very nearly vertically when they reach the base of the transition region. This requirement is compatible with the fact that the oscillations at the base of the chromosphere are nearly vertical, and with our above finding that these waves have periods favorable for the refraction of these waves toward the vertical in the chromosphere.

For $\theta < \theta_{crit}$, a wave incident on an interface is partially transmitted and partially reflected. For normal incidence ($\theta = 0$), the ratio of the transmitted energy flux F_2 to the incident energy flux F_1 is given by[57]

$$\frac{F_2}{F_1} = \frac{\rho_2 a_2}{\rho_1 a_1}\left[\frac{2}{1 + \frac{\rho_2 a_2}{\rho_1 a_1}}\right]^2 , \qquad (3.47)$$

where ρ is the mass density and a is the sound speed. Since the gas pressure remains nearly constant through the transition region while the temperature increases by a factor of 10^2, the density decreases by a factor of 10^2 while the sound speed increases by a factor of 10. Hence, for the transition region,

$$\frac{F_2}{F_1} \approx 0.3 \quad .$$

Since the observations require that a mechanical energy flux of about 10^6 erg cm^{-2} sec^{-1} reach the corona, the above estimate indicates that a mechanical energy flux of at least a few times this amount enters the base of the chromosphere in the form of wave modes which can pass through the chromosphere.

To obtain an order of magnitude estimate of the energy flux F_m carried by the waves composing the nearly vertical oscillations at the base of the chromosphere, we may use equation (3.12),

$$F_m = \rho \overline{v^2}\, a \quad .$$

At the base of the chromosphere, the observed rms velocity of the vertical oscillations is about 0.4 km sec^{-1}, which with equation (3.12) gives

$$F_m \approx 2.5 \times 10^7 \text{ erg cm}^{-2} \text{ sec}^{-1} \quad .$$

This estimate is probably somewhat high since some of the waves represented by the photospheric oscillations are evanescent rather than propagating as equation (3.12) assumes. However, if this estimate is accurate to order of magnitude, it appears that the waves which make up the photospheric oscillations carry enough energy to heat the corona.

In summary, the considerations of this section lead us to two main conclusions:

1. The chromosphere and transition region act as a filter which passes only nearly vertically propagating waves to the corona.

2. The observed amplitude, orientation and period range of the photospheric oscillations indicate the presence of sufficient energy flux in the appropriate wave modes to provide the heating of the corona.

At the beginning of our discussion of propagation (Section 3.3.2.1) we suggested that by understanding the propagation properties of the solar atmosphere, we might gain some insight into the generation and dissipation of waves which heat the atmosphere. From the results of the present section, we see that the propagation properties of the atmosphere, in conjunction with the observed properties of the photospheric oscillations, essentially resolve the generation problem for the heating of the corona. Since the observations indicate that the photospheric oscillations result primarily from the buffeting of the base of the photosphere by rising granules, it appears that this is the basic process which generates the waves which heat the corona. With regard to the dissipation problem for the heating of the corona, we have found that the waves to be considered are upward propagating compression waves with periods in the range 200 to 300 sec. It remains to consider the dissipation of these waves in the corona.

3.3.3 Wave Dissipation in the Corona

The conclusion of the preceding section that the corona is heated by upward propagating compression waves with periods in the range 200 to 300 sec is well accepted in the literature.[58-62] However, three different mechanisms have been proposed for the dissipation of these waves:

1. Dissipation by thermal conduction.[59]

2. Landau damping.[60]

3. Shock-wave dissipation following the steepening of the compression waves to form shock fronts.[61,62]

In this section we consider the dissipation of compression waves in the corona by these mechanisms. Our results indicate that the waves which heat the corona are dissipated by conduction damping and Landau damping rather than by forming shock fronts.

The energy balance of the transition region provides a basic requirement which must be met by the mechanism which dissipates the waves in the corona. In Chapter 4, we find that the dominant source of heating for the transition region is heat conduction from the corona. This requires that the total rate of wave dissipation in the transition region is small compared to the rate of conduction cooling of the corona. Hence, since the conduction cooling of the corona is balanced by heating of the corona by wave dissipation, the waves which pass through the transition region to heat the corona dissipate only a small fraction of their energy flux in the transition region. Therefore, a necessary condition which must be satisfied by the mechanism which dissipates the waves in the corona is that in the transition region the damping length L_D of the waves be longer than the thickness of the transition region, which is about 3000 km.

$$L_D > 3 \times 10^8 \text{ cm} \quad . \tag{3.48}$$

L_D is the distance in which the amplitude of the wave is decreased by a factor of e^{-1} by the damping mechanism.

The conduction damping of the waves can be examined by means of the dispersion relation for sound waves in a uniform medium with thermal conductivity κ,[59]

$$\omega\left(\omega^2 - \frac{\gamma k_B T}{m} k^2\right) = -i \frac{2}{3} \frac{\kappa k^2}{k_B n} \left(\omega^2 - \frac{k_B T}{m} k^2\right) , \tag{3.49}$$

where k is the wave number, k_B is Boltzmann's constant, and i is the imaginary unit. Analysis of this dispersion relation shows that the conduction damping depends only on the dimensionless parameter

$$\delta = \frac{4\pi}{3} \frac{\kappa P}{k_B n \lambda^2} , \tag{3.50}$$

where P is the wave period and λ is the wavelength. The damping parameter δ can be interpreted physically as follows. The temperature T in a sinusoidal sound wave is given by

$$T = T_o + \Delta T , \tag{3.51}$$

where T_o is the ambient temperature and ΔT is the temperature perturbation given by

$$\Delta T = (\Delta T)_{max} \sin\left(\frac{2\pi}{\lambda} x - \frac{2\pi}{P} t\right) . \tag{3.52}$$

The perturbation $\Delta(\rho e)$ in the thermal energy density is

$$\Delta(\rho e) = \frac{3}{2} n k_B \Delta T , \tag{3.53}$$

and the rate of conduction cooling per unit volume is

$$-K \frac{\partial^2 T}{\partial x^2} = K\left(\frac{2\pi}{\lambda}\right)^2 \Delta T \quad . \tag{3.54}$$

Hence, the conduction cooling time τ defined by the ratio of the perturbation thermal energy density to the rate of conduction cooling per unit volume,

$$\tau \equiv \frac{\Delta(\rho e)}{-K \frac{\partial^2 T}{\partial x^2}} \quad , \tag{3.55}$$

is given by

$$\tau = \frac{3}{2} \frac{nk_B}{K} \left(\frac{\lambda}{2\pi}\right)^2 \quad . \tag{3.56}$$

Comparison of this expression for τ with the defining expression (3.50) for δ, shows that

$$\delta = \frac{P}{2\pi\tau} \quad . \tag{3.57}$$

The time $P/2\pi$ is the characteristic time for the compressions and expansions which produce the temperature perturbation ΔT in the sound wave. Thus, we see that the damping parameter δ is the ratio of the compression time in the wave to the conduction cooling time.

The strength of the damping is characterized by the ratio of the wavelength λ to the damping length L_D; the larger λ/L_D, the stronger the damping. From the dispersion relation (3.49) we obtain the following relation between λ/L_D and the damping parameter δ for conduction damping:

$$\left(\frac{\delta\lambda}{2\pi L_D}\right)^2 - \frac{1}{2}\gamma\left[\left(1+\frac{3}{\gamma}\right)\left(\frac{\lambda}{2\pi L_D}\right)^2 + \left(1-\frac{1}{\gamma}\right)\right]\frac{\delta\lambda}{2\pi L_D} + \gamma\left(\frac{\lambda}{2\pi L_D}\right)^2\left[\left(\frac{\lambda}{2\pi L_D}\right)^2 + 1\right] = 0 \ . \tag{3.58}$$

For the case of $\gamma = 5/3$ (fully ionized plasma) this equation gives the curve of λ/L_D vs δ shown in Figure 3.5. We see from this curve that the damping is strongest when $\delta \approx 1$, in which case $L_D \approx \lambda$.

For very small or very large values of δ, the damping is negligible, and the dispersion relation (3.49) may be approximated by

$$\omega^2 - \frac{\gamma k_B T}{m} k^2 = -i\delta\left[\omega^2 - \frac{k_B T}{m} k^2\right] \ . \tag{3.59}$$

For the limiting case of $\delta = 0$, the dispersion relation is

$$\omega^2 = \frac{\gamma k_B T}{m} k^2 \ , \tag{3.60}$$

which is the dispersion relation for adiabatic sound waves. In the limit as δ approaches infinity, the dispersion relation is

$$\omega^2 = \frac{k_B T}{m} k^2 \ , \tag{3.61}$$

which is the dispersion relation for isothermal sound waves. Hence, as is reasonable from the physical meaning of δ, for $\delta \ll 1$ or $\delta \gg 1$, the waves propagate at the adiabatic or isothermal sound speed, respectively, and have negligible dissipation.

For the purpose of examining conduction damping in the transition region and corona, it is convenient to rewrite expression (3.50) for the damping parameter δ in terms of n, T and P. In the corona and at

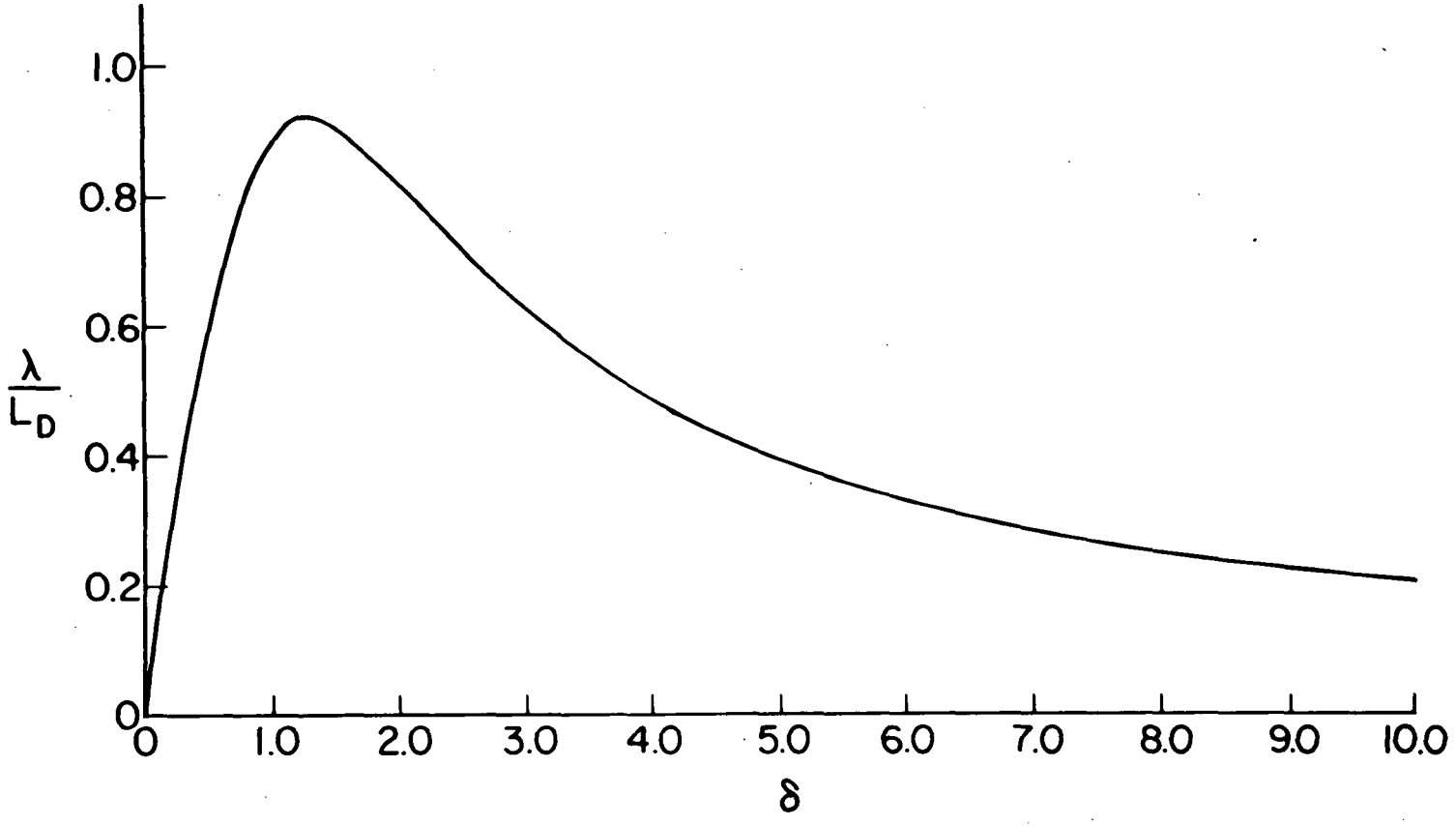


Figure 3.5. VARIATION OF THE DAMPING WITH THE DAMPING PARAMETER δ FOR DAMPING OF SOUND WAVES BY THERMAL CONDUCTION.

temperatures above 2×10^4 K in the transition region, the thermal conductivity κ is given to a close approximation (see Appendix A) by

$$\kappa = 1.0 \times 10^{-6} \, T^{5/2} \text{ erg sec}^{-1} \text{ K}^{-1} \text{ cm}^{-1} \quad . \tag{3.62}$$

In terms of T and P, the wavelength λ is

$$\lambda = \left[\alpha \frac{k_B T}{m} \right]^{1/2} P \quad , \tag{3.63}$$

where $[\alpha(k_B T/m)]^{1/2}$ is the sound speed.

For the value of the polytrope index α we adopt 4/3, the mean between the isothermal value of 1 and the adiabatic value of 5/3. In the transition region and corona, the mean particle mass m is approximately half the mass of the hydrogen atom. These values give

$$\lambda = 1.5 \times 10^4 \, T^{1/2} \, P \text{ cm} \quad . \tag{3.64}$$

Finally, substitution of equations (3.62) and (3.64) into equation (3.50) gives the following expression for the damping parameter in the transition region and corona:

$$\delta = 1.4 \times 10^2 \, n^{-1} \, T^{3/2} \, P^{-1} \quad . \tag{3.65}$$

We may now proceed to check conduction damping against the necessary condition (3.48) that in the transition region the damping length of the waves which heat the corona be longer than the thickness of the transition region. In the transition region $nT \approx \text{const} = 10^{15} \text{ cm}^{-3} \text{ K}$, so that the equation for the damping parameter becomes

$$\delta = 1.4 \times 10^{-13}\ T^{5/2}\ P^{-1} \quad . \tag{3.66}$$

Hence, for a wave of a given period, the damping parameter of the wave increases as $T^{5/2}$ as the wave propagates upward through the transition region. Maximum damping occurs where $\delta \approx 1$, at which point $L_D \approx \lambda$. From these conditions and equations (3.66) and (3.64), we have at maximum damping in the transition region,

$$T \approx 1.4 \times 10^{5}\ P^{2/5}\ K \quad , \tag{3.67}$$

and

$$L_D \approx 5.5 \times 10^{6}\ P^{6/5}\ cm \quad . \tag{3.68}$$

Since $T \leq 10^6$ K in the transition region, we see from equation (3.67) that if $P \gtrsim 140$ sec, the wave does not reach maximum damping until it reaches some point in the corona. From equation (3.68) it follows that the necessary condition (3.48) on the damping length is satisfied by conduction damping if $P \gtrsim 30$ sec. Since the waves which heat the corona are expected to have periods in the range 200 to 300 sec, we conclude that condition (3.48) is well satisfied by the dissipation mechanism of conduction damping.

We can also estimate the conduction damping length near the base of the corona from equation (3.68). In the transition region and corona the pressure scale height H is given by

$$H = 6.0 \times 10^{3}\ T\ cm \quad . \tag{3.69}$$

So, the scale height H_0 at the base of the corona (defined by $T = 10^6$ K) is

$$H_0 = 60{,}000 \text{ km} \quad .$$

Therefore, the condition of constant pressure and the resulting equations (3.67) and (3.68) remain valid for the first several thousand kilometers in the corona. For the period range 200 to 300 sec, equation (3.67) gives

$$1.15 \times 10^6 < T < 1.35 \times 10^6 \text{ K}$$

for the temperature range for which maximum damping of these waves occurs. Figure 2.2 shows that this temperature range occurs within the first several thousand kilometers of the corona, which shows that the use of equations (3.67) and (3.68) is valid for periods of 200 to 300 sec. Using the representative period 250 sec in equation (3.68), we obtain

$$L_{D_0} \approx 4.2 \times 10^9 \text{ cm}$$

for the value of the damping length due to conduction damping near the base of the corona.

As has been pointed out by D'Angelo,[60] Landau damping may also contribute to the dissipation of the waves in the corona. It has been established experimentally that Landau damping of ion-acoustic waves occurs when[60]

$$\frac{\lambda_p}{\lambda} \gtrsim \frac{1}{2\pi} \quad , \tag{3.70}$$

where, for hydrogen plasma, λ_p is the proton mean free path. Hence, Landau damping occurs when the wavelength λ is comparable to or shorter than the proton mean free path. At longer wavelengths Landau damping

does not occur.[60] When condition (3.70) is satisfied, the damping length due to Landau damping is observed to be about 0.5 λ, which is shorter than the damping length due to thermal conduction at maximum damping. So, when Landau damping occurs, it dominates conduction damping.

To examine the relative importance of conduction damping and Landau damping in the corona, expression (3.65) for the conduction damping parameter δ can be put into a more convenient form. Combining equations (3.64) and (3.65), we obtain

$$\delta = 2.1 \times 10^{6}\, n^{-1}\, T^{2}\, \lambda^{-1} \quad . \tag{3.71}$$

Since the proton mean free path λ_p is given by[63]

$$\lambda_p = 8.0 \times 10^{3}\, n^{-1}\, T^{2} \text{ cm} \quad , \tag{3.72}$$

we have

$$\delta = 2.6 \times 10^{2}\, \frac{\lambda_p}{\lambda} \quad . \tag{3.73}$$

With this expression for δ, we see that the condition (3.70) for Landau damping may be written

$$\delta \gtrsim 42.0 \quad . \tag{3.74}$$

When this condition is satisfied, Landau damping dominates conduction damping.

As can be seen from equation (3.65), for a wave of given period P, the damping parameter δ increases with height in the corona. When the critical value of $\delta = 42$ is reached, the remaining energy

of the wave is dissipated by Landau damping. Landau damping is important in the dissipation of the wave in the corona only if $\delta = 42$ is reached before most of the energy of the wave is lost by conduction damping. Therefore, a condition for Landau damping to be important is

$$\delta_{L_{D_0}} \geq 42 \quad , \tag{3.75}$$

where $\delta_{L_{D_0}}$ is the value of δ at one damping length above the base of the corona.

To estimate the value of $\delta_{L_{D_0}}$, we rewrite equation (3.65) as follows:

$$\delta = 1.4 \times 10^2 \ (nT)^{-1} \ T^{5/2} \ P^{-1} \quad . \tag{3.76}$$

Since the low corona is in hydrostatic equilibrium, the value of nT at one damping length above the base of the corona is given approximately by

$$(nT)_{L_{D_0}} = (nT)_0 \exp\left(-\frac{L_{D_0}}{H_0}\right) \quad . \tag{3.77}$$

By definition of the damping length, most of the wave energy is deposited below one damping length above the base. Since most of this energy is conducted back to the transition region, the temperature at one damping length above the base should be nearly the maximum temperature of the corona:

$$T_{L_{D_0}} \approx T_{max} \quad . \tag{3.78}$$

For quiet regions, $T_0 = 1.0 \times 10^6$ K, $n_0 \approx 1.0 \times 10^9$ cm^{-3}, $T_{max} \approx 2.0 \times 10^6$ K, $H_0 = 6.0 \times 10^3$ $T_0 = 6.0 \times 10^9$ cm and, as found above, $L_{D_0} \approx 4.2 \times 10^9$ cm. Using these values in equations (3.77) and (3.76), we obtain

$$\delta_{L_{D_0}} \approx 1.5 \times 10^3 \, P^{-1} \quad . \tag{3.79}$$

Therefore, Landau damping is important ($\delta_{L_{D_0}} \gtrsim 42$) in the corona only for waves with $P \lesssim 35$ sec. Since the waves which carry energy into the corona have periods of order 250 sec, it appears that Landau damping is not important in the corona.

However, Landau damping <u>is</u> important if an appreciable amount of energy in the $P \sim 250$ sec waves is transferred to $P \lesssim 35$ sec waves as the waves propagate from the base of the corona to L_{D_0} above the base. Transfer to shorter period waves occurs when the wave form steepens due to the finite amplitude wave crest overtaking the wave trough. Whether this steepening effect is important can be estimated as follows. The energy flux of a sinusoidal compression wave with velocity amplitude V is given by

$$F_m = \frac{1}{2} nmM^2 a^3 \quad , \tag{3.80}$$

where $M \equiv V/a$ is the Mach number of the wave, and a is the speed of propagation given by equation (3.10):

$$a = \left(\alpha \frac{k}{m} T\right)^{1/2} \quad . \tag{3.10}$$

Substitution of equation (3.10) into equation (3.80) gives

$$F_m = 1.3 \times 10^{-12}\ nT^{3/2}\ M^2\ \text{erg cm}^{-2}\ \text{sec}^{-1}\ , \qquad (3.81)$$

where we have adopted $\alpha \approx 4/3$. At the base of the corona, $T = 1.0 \times 10^6$ K, $n \approx 1.0 \times 10^9$ cm^{-3} and $F_m \approx 1.0 \times 10^6$ erg cm^{-2} sec^{-1}, giving

$$M_0 \approx 0.9\ .$$

Since for $M = 1.0$, the crest overtakes the trough after propagating only half a wavelength, it appears that there should be a sizeable transfer of energy to short-period waves. Therefore, we conclude that, in addition to conduction damping, Landau damping should contribute significantly to the wave dissipation in the corona.

It has often been assumed in the literature,[39,51,61,62] that the above steepening process leads to the formation of shock fronts which account for the dissipation of the waves. However, our results for conduction damping and Landau damping suggest that the waves do not form shock fronts in the transition region and corona. If the waves have shock fronts, then much of the energy of the waves is carried in short-period components. On the other hand, as pointed out above, the waves which heat the corona dissipate only a small fraction of their energy in the transition region. For the dissipation process of thermal conduction, we found that the waves should have periods longer than 30 sec for this to be true. This implies that short-period ($P \lesssim 30$ sec) components are absent in the transition region and hence that there are no shock waves in this region. In the corona, where transfer to shorter

period ($P \lesssim 35$ sec) components probably does occur, these short-period waves will be quickly damped by Landau damping, which indicates that shock fronts do not form in the corona either.

4. A MODEL OF THE CHROMOSPHERE-CORONA TRANSITION REGION

4.1 Introduction

In this chapter we examine the extent to which a static, planar model, heated by thermal conduction from the corona and cooled by radiative losses, describes and explains the structure and energy balance of the chromosphere-corona transition region. In this section we briefly review the salient features of the transition region which motivate such a study.

Observed energy fluxes of XUV resonance lines emitted from the outer solar atmosphere imply (see Section 4.3.2.1) that above the 10^5 K level the atmosphere is approximately planar, and that the flux of heat flowing downward through the transition region from the corona remains roughly constant from the 10^6 K level down to the 10^5 K level. In addition (see Section 3.2.1.2), the rate of radiative cooling of the transition region and the rate at which heat is supplied to the transition region by conduction from the corona are observed to be of the same order (10^6 erg cm^{-2} sec^{-1}). These results are the observational basis for adopting a static, planar model for the transition region, and for suspecting that the dominant processes in the energy balance of the transition region are conduction heating and radiative cooling.

We term the upper part of the transition region, defined by the temperature range 10^5 to 10^6 K, the "constant-heat-flux region". The part of the transition region below the constant-heat-flux region (i.e., with $T < 10^5$ K), we call the "base region". This schematic picture of the transition region is summarized in Figure 4.1.

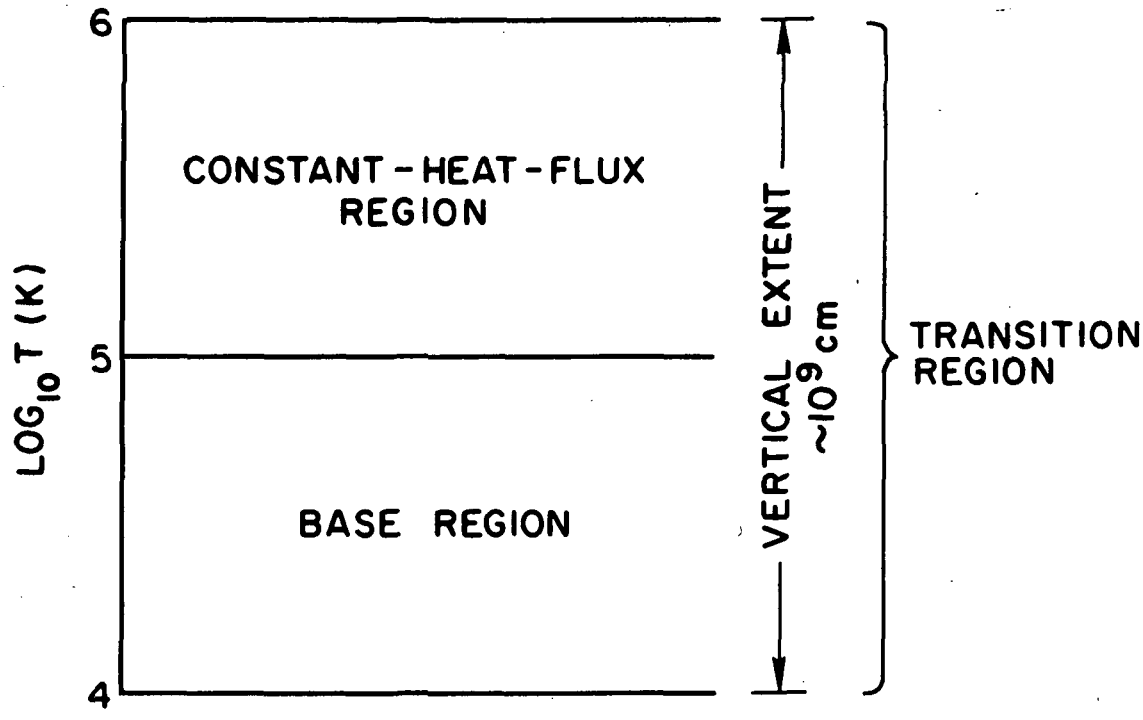


Figure 4.1. SCHEMATIC REPRESENTATION OF THE TRANSITION REGION.

A major function of the base region is to absorb the downward flowing heat which passes through the constant-heat-flux region. The value of this heat flux which enters the base region is not accurately determined by the ultraviolet line data, but a value greater than 10^5 erg cm^{-2} sec^{-1} is indicated (see Section 4.3.2.1). At the 10^4 K level and below, the heat conductivity and the temperature gradient are so small that less than 10^2 erg cm^{-2} sec^{-1} can be conducted out of the bottom of the transition region. Therefore, the heat entering the base region from the constant-heat-flux region must be absorbed in the base region.

Optical eclipse spectra indicate that spicule-like inhomogeneities begin to appear in the chromosphere at heights above about 1,500 km,[9] and spicules often extend to heights of 10,000 km or more. Thus, some spicules extend through the height range of the chromosphere-corona transition region. Since spicules are transient and must have a temperature well below 10^5 K,[25] the constant-heat-flux region cannot be completely static and horizontally uniform. However, above 3,000 km spicules occupy less than a few percent of the horizontal surface area.[17,25] This suggests that the static, planar constant-heat-flux region implied by the ultraviolet emission-line data corresponds to a hot, static background atmosphere which is penetrated here and there by the cooler spicules. Thus, although the constant-heat-flux region appears to be inhomogeneous and fluctuating when viewed optically at the limb in a chromospheric emission line, in terms of overall structure and average heat flow from the corona to the chromosphere, a static, planar model is still reasonable.

In view of the static, planar nature of the constant-heat-flux region (despite the presence of spicules), it is reasonable to consider

a static, planar model for the base region which absorbs the heat which passes through the constant-heat-flux region. A static, planar model of the base region is "possible" if the downward heat flux passing the 10^5 K level can be radiated away above the 10^4 K level. However, Kuperus and Athay[64] have suggested that the base region is so thin that the inflowing heat cannot be balanced by radiation alone, and that the excess energy goes into the kinetic energy of spicules. Hence, a quantitative study of a static, planar model of the transition region may be relevant to the origin of spicules, and their role and importance in the energy balance of the base region.

4.2 Model and Formulation

4.2.1 Basic Physical Assumptions

The model of the transition region which we propose is based on the following assumed physical conditions and approximations.

1. The model is horizontally uniform.
2. The magnetic field is assumed to be vertical.
3. The model transition region is assumed to be in hydrostatic equilibrium.
4. The "turbulent pressure" in the transition region, due to velocity fluctuations (part of which are produced by the mechanical waves which pass through the transition region to heat the corona), is neglected; the model atmosphere is assumed to be supported against gravity entirely by the thermal gas pressure.
5. There is no dissipation of the mechanical waves within the model transition region.
6. The condition of ionization equilibrium is satisfied (for each stage of ionization of each element) at each point in the model.

7. All excitations and ionizations are collisional, and all recombinations and de-excitations are radiative.

8. The presence of elements other than hydrogen is ignored except for their contribution to the radiative losses.

The aim of each of the above adopted conditions is to simplify the mathematical model while hopefully retaining the essential physics involved in the structure and energy balance of the actual transition region. In particular, approximation 4 simplifies the force equation, approximation 5 simplifies the energy equation, while approximations 7 and 8 simplify the ionization equation as well as the force and energy equations. The degree to which the model resulting from our simplifying assumptions retains the essential physics of the transition region is to be judged, in part, from the comparison of the model with observations.

It should be noted that approximations 5 and 7 optimize the removal of heat flux by radiation. In general, the energy radiated from each volume element of the atmosphere may come from three sources: (i) mechanical energy dissipated in the volume, (ii) radiation absorbed in the volume, (iii) heat flux absorbed in the volume. Approximations 5 and 7 require that all of the energy radiated from each volume element be supplied by the absorption of heat flux.

4.2.2 Governing Equations

The equation which expresses condition 6, the condition of ionization equilibrium, is

$$\mathcal{I} = \mathcal{R} \quad , \tag{4.1}$$

where $\mathcal{I}$ is the rate of ionization per unit volume, and $\mathcal{R}$ is the rate of recombination per unit volume. If the vertical coordinate z is taken to be positive upward, the condition of hydrostatic equilibrium requires that

$$\frac{dp}{dz} = -\rho g \tag{4.2}$$

where p is the thermal gas pressure, ρ is the mass density of the gas and g is the acceleration of gravity. Conditions 3, 5, and 7 and the conservation of energy demand that there be a balance of heat conduction and radiation. This is expressed by

$$-\frac{dF}{dz} = \mathcal{L}_r \quad , \tag{4.3}$$

where F is the heat flux (positive upward) and $\mathcal{L}_r$ is the radiative power output per unit volume. We now proceed to express equations (4.1), (4.2) and (4.3) in terms of the total number density n, the electron number density n_e, the temperature T, and their derivatives with respect to the vertical coordinate z, to obtain three equations in three unknowns with which to compute the structure of the model.

Under approximations 7 and 8, equation (4.1) can be written as

$$n_e n_H I_H = n_e n_p R_H \quad , \tag{4.4}$$

where n_H is the number density of neutral hydrogen, n_p is the number density of ionized hydrogen, I_H is the collisional ionization coefficient for hydrogen, and R_H is the radiative recombination coefficient for hydrogen. Now with the simple relations for pure hydrogen,

$$n = n_H + n_p + n_e \tag{4.5}$$

and

$$n_p = n_e \quad , \tag{4.6}$$

the ionization equilibrium equation may be solved for n_e/n:

$$\frac{n_e}{n} = \frac{1}{2 + R_H/I_H} \quad . \tag{4.7}$$

For pure hydrogen, the mass density of the gas is

$$\rho = \left(1 - \frac{n_e}{n}\right) n m_H \quad . \tag{4.8}$$

From this expression and the equation of state,

$$p = nkT \quad , \tag{4.9}$$

the equation of hydrostatic equilibrium can be written in terms of n, n_e and T:

$$\frac{1}{nT} \frac{d(nT)}{dz} = - \frac{(1 - n_e/n)\, m_H g}{kT} \quad . \tag{4.10}$$

The heat flux F is proportional to the temperature gradient:

$$F = -K \frac{dT}{dz} \quad . \tag{4.11}$$

where K is the thermal conductivity. Thus, the left hand side of the energy equation, equation (4.3), may be expressed as

$$-\frac{dF}{dz} = \frac{d}{dz}\left(\kappa \frac{dT}{dz}\right) \quad . \tag{4.12}$$

For the radiative power output density $\mathcal{L}_r$, we adopt the results of Cox and Tucker.[42] For the physical situation specified by conditions 6 and 7, they have computed the radiative cooling coefficient L_r, which is related to $\mathcal{L}_r$ by

$$\mathcal{L}_r = n_e(n_H + n_p)\, L_r \quad , \tag{4.13}$$

as a function of temperature in the range 10^4 K to 10^8 K. In addition to hydrogen, Cox and Tucker have included cosmic abundances of the next eight most abundant elements: He, C, N, O, Ne, Mg, Si and S, in their calculation of L_r. At temperatures above a few times 10^4 K, these additional elements dominate the radiative cooling. Thus, with regard to L_r, we relax our assumption of a pure hydrogen atmosphere. However, since hydrogen is the major component of the solar atmosphere, the elements other than hydrogen may be neglected in the equation of hydrostatic equilibrium, in determining the heat flux and thermal conductivity, and in the ionization equation, which determines the electron, proton and hydrogen number densities in equation (4.13). The expression of the energy equation in terms of n, n_e and T is then

$$\frac{d}{dz}\left(\kappa \frac{dT}{dz}\right) = \frac{n_e}{n}\left(1 - \frac{n_e}{n}\right) n^2 L_r \quad . \tag{4.14}$$

4.2.3 Specification of the Model

The coefficients I_H and R_H in equation (4.7), and κ and L_r in equation (4.14) are functions of temperature which we now specify.

The ionization and recombination coefficients for hydrogen are taken from Cox and Tucker:[42]

$$I_H = 2.34 \times 10^{-8}\ \beta^{-1/2}\ e^{-\beta}\ \text{cm}^3\ \text{sec}^{-1}\ , \tag{4.15}$$

and

$$R_H = 5.20 \times 10^{-14}\ \beta^{1/2}\left[0.4288 + \frac{1}{2}\ell n\ \beta + 0.4698\,\beta^{-1/3}\right]\text{cm}^3\ \text{sec}^{-1}\ , \tag{4.16}$$

where

$$\beta = 158{,}000/T\ . \tag{4.17}$$

Figure 4.2 shows the radiative cooling coefficient curve computed by Cox and Tucker for the temperature range 10^4 K to 10^6 K. We have adopted the following straight-line-segment fit for our model:

$$\left.\begin{aligned}
&4.0 < \log T < 4.2: \log L_r = 8.00 \log T - 55.6 \\
&4.2 < \log T < 4.5: \log L_r = -22.0 \\
&4.5 < \log T < 4.87: \log L_r = 2.44 \log T - 33.0 \\
&4.87 < \log T < 5.5: \log L_r = -21.1 \\
&5.5 < \log T < 6.0: \log L_r = -1.86 \log T - 10.9
\end{aligned}\right\} \tag{4.18}$$

Figure 4.3 shows the thermal conductivity curve which we have obtained from the results of Devoto[65] and Delcroix and Lemaire[66] (see Appendix A). The curve is closely fit by three straight line segments given by

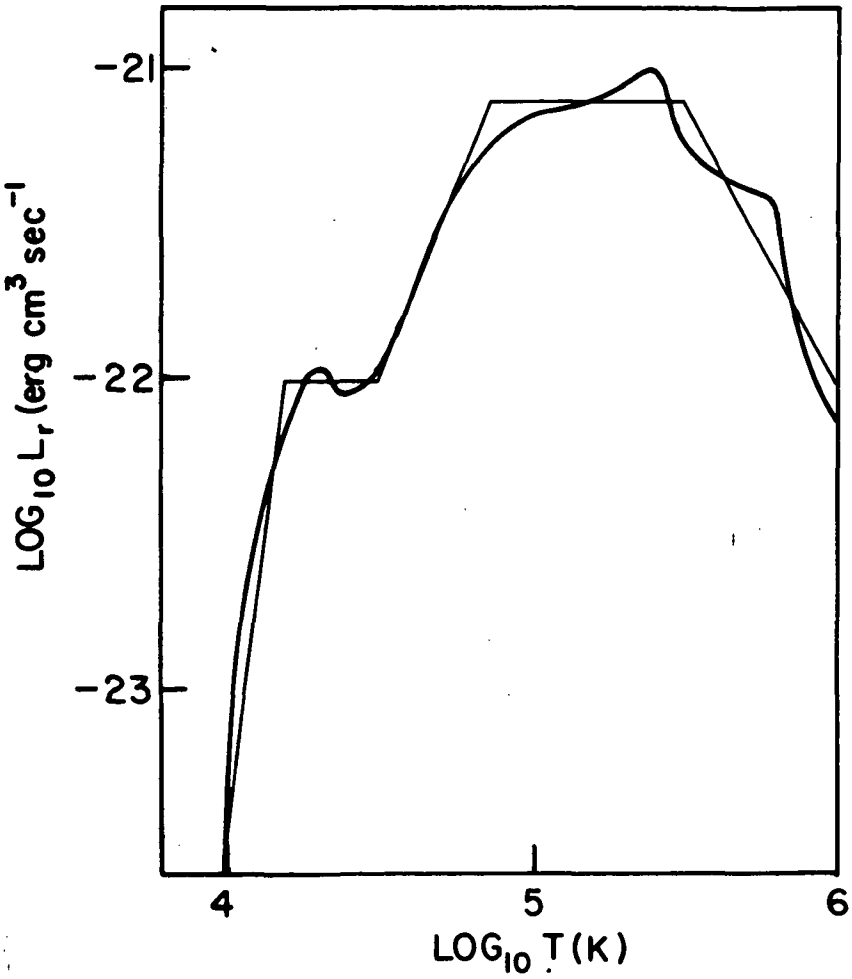


Figure 4.2. RADIATIVE COOLING COEFFICIENT CURVE COMPUTED BY COX AND TUCKER (1969) AND THE STRAIGHT-LINE SEGMENT FIT ADOPTED FOR OUR MODEL.

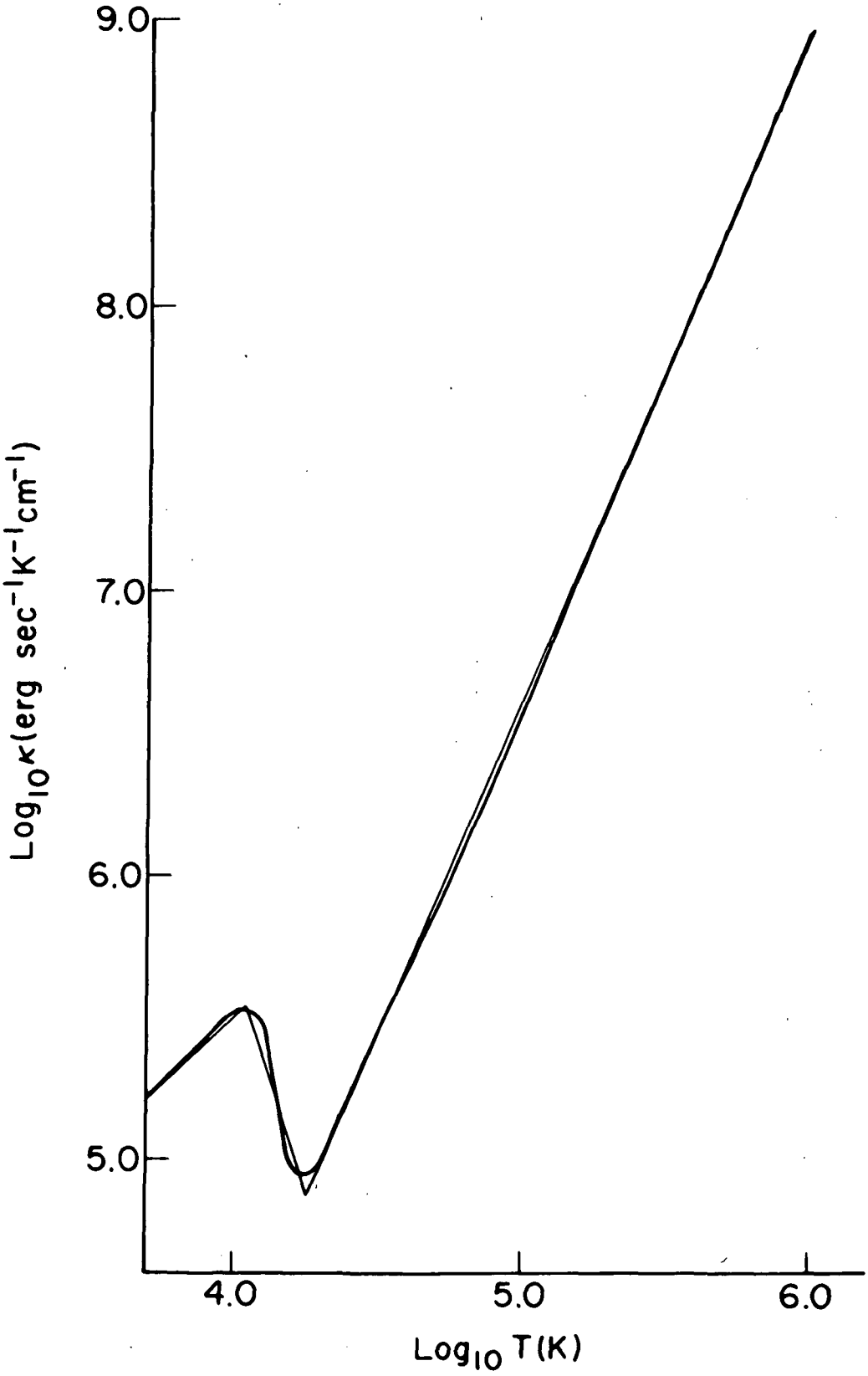


Figure 4.3. THERMAL CONDUCTIVITY CURVE DERIVED FROM THE RESULTS OF DEVOTO (1968) AND DELCROIX AND LEMAIRE (1969) AND THE STRAIGHT LINE FIT ADOPTED FOR OUR MODEL.

$$\left.\begin{aligned} 4.0 < \log T < 4.05&: \log K = \log T + 1.49 \\ 4.05 < \log T < 4.26&: \log K = -3.19 \log T + 18.47 \\ 4.26 < \log T < 6.00&: \log K = 2.36 \log T - 5.18 \end{aligned}\right\} \qquad (4.19)$$

This straight-line-segment fit is the conductivity curve used in computing the structure of the model.

With the above coefficient formulas, we have now completely specified our model. The structure of the model is completely determined by appropriate boundary values of n, T, and F. We choose to specify the values of these quantities at the upper boundary of the model transition region because the value of the number density is best known at the top of the observed transition region.[13] We define the upper boundary of the model to be the 10^6 K level. We therefore fix the temperature T_0 of the upper boundary of the model at 10^6 K and study the model by varying the boundary values of n and F, which are denoted by n_0 and F_0.

4.3 Results

4.3.1 Numerical Results

To study the behavior of the model with variations in n_0 and F_0, we computed the run of temperature and heat flux with height in the model for several values of n_0 and F_0 by numerically solving the governing equations (4.7), (4.10), and (4.14). Examples of the temperature profiles are shown in Figure 4.4, where F_0 is fixed at -1.0×10^6 erg cm^{-2} sec^{-1} and n_0 takes the values 1.00, 1.92, 2.24, and 2.82 times 10^9 cm^{-3}. Curves of the heat flux versus the logarithm of the temperature for the same cases are shown in Figure 4.5.

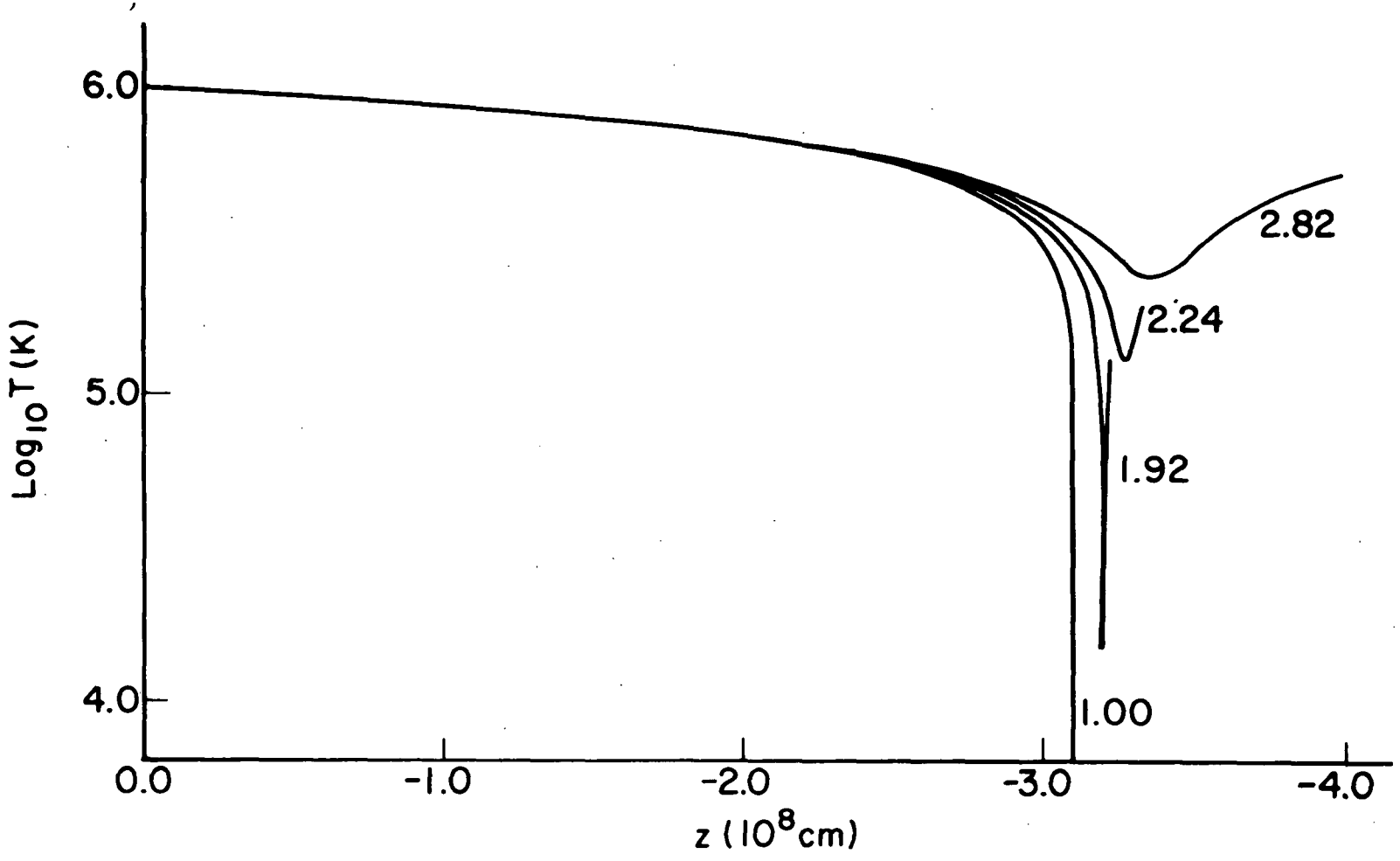


Figure 4.4. COMPUTED TEMPERATURE PROFILES FOR $F_0 = -1.0 \times 10^6$ erg cm^{-2} sec^{-1}. Each curve is labeled with its value of n_0 in units of 10^9 cm^{-3}.

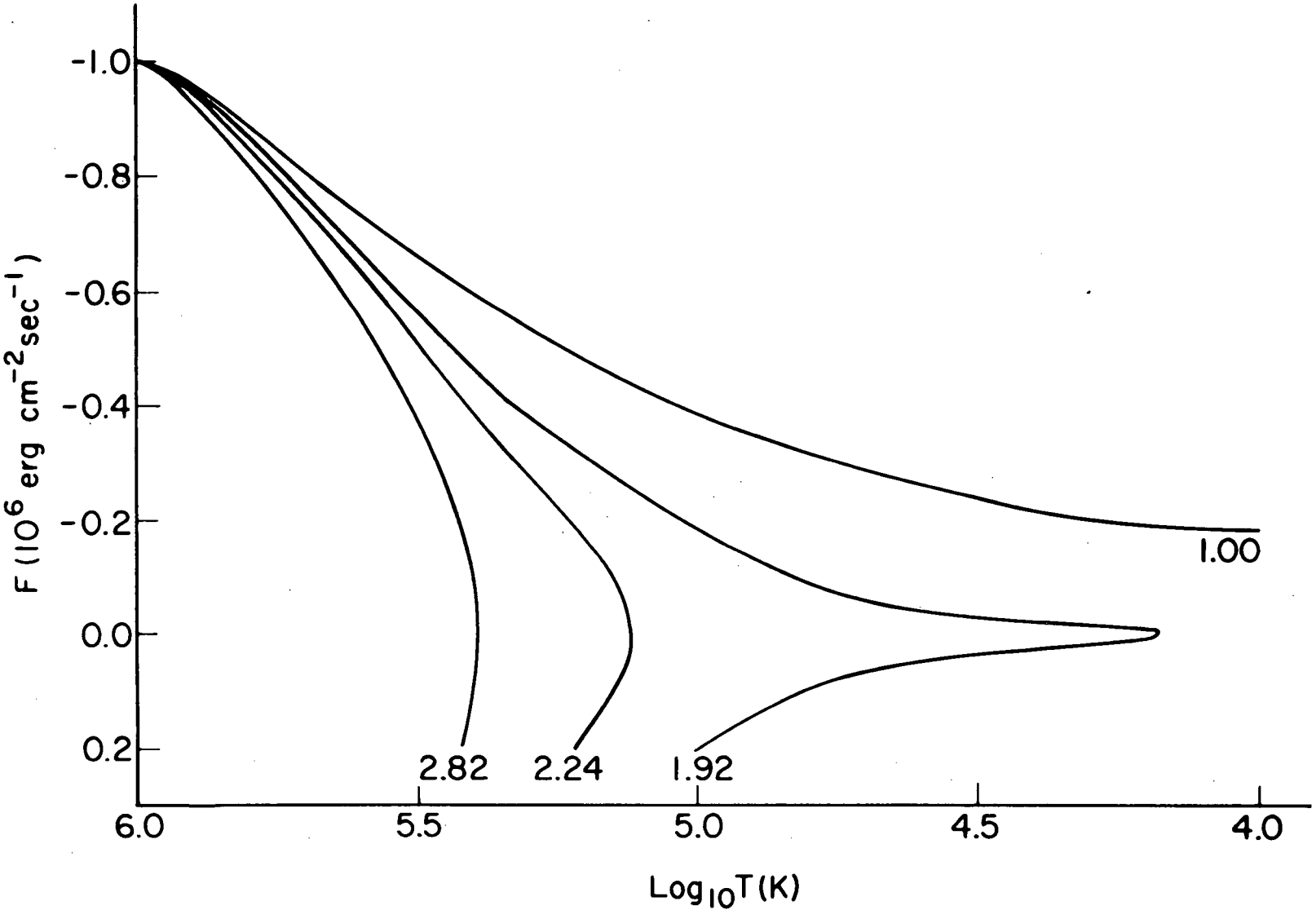


Figure 4.5. F, Log T CURVES FOR EACH OF THE FOUR CASES OF FIGURE 4.4.

The curves in Figures 4.4 and 4.5 may be qualitatively understood as follows. The radiative power output density is proportional to n^2. On the other hand, if there were no radiative losses, the thickness of the transition region would be inversely proportional to the (constant) downward heat flux. Consequently, we expect that for sufficiently large values of the number density n_0 and sufficiently small values of the heat flux F_0 at the upper boundary, the model will radiate away all of the downward heat flux. That is, the heat flux will decrease to zero at some temperature $T_{ZF} > 10^4$ K. Figures 4.4 and 4.5 bear this out: if n_0 is large enough ($n_0 \gtrsim 1.92 \times 10^9$ cm^{-3}) the temperature gradient and the heat flux pass through zero at the minimum temperature ($> 10^4$ K) reached by the curve, but if n_0 is too small (e.g., $n_0 = 1.00 \times 10^9$ cm^{-3}), no temperature minimum occurs because the model is unable to radiate away all of the heat flux.

For a given F_0, the value of n_0 determines the temperature T_{ZF} at which $F = 0$. Thus, from sets of cases such as that represented in Figure 4.5, we may obtain the sets of (n_0, T_{ZF}) points shown in Figure 4.6 for constant values of F_0. A smooth curve has been fitted through each set of points with a common value of F_0. From these curves we then obtain sets of (n_0, F_0) at constant T_{ZF} which we have plotted logarithmically in Figure 4.7. Each set of points for constant T_{ZF} lies along a curve with slope increasing from about 0.9 at $\log n_0 = 9.0$ to about 1.0 at $\log n_0 = 10.0$. Thus, for constant T_{ZF}, F_0 is approximately proportional to n_0.

The fact that F_0 is approximately proportional to n_0 for constant T_{ZF} is readily understood from the governing equations and the fact that the pressure is nearly constant in the model transition region.

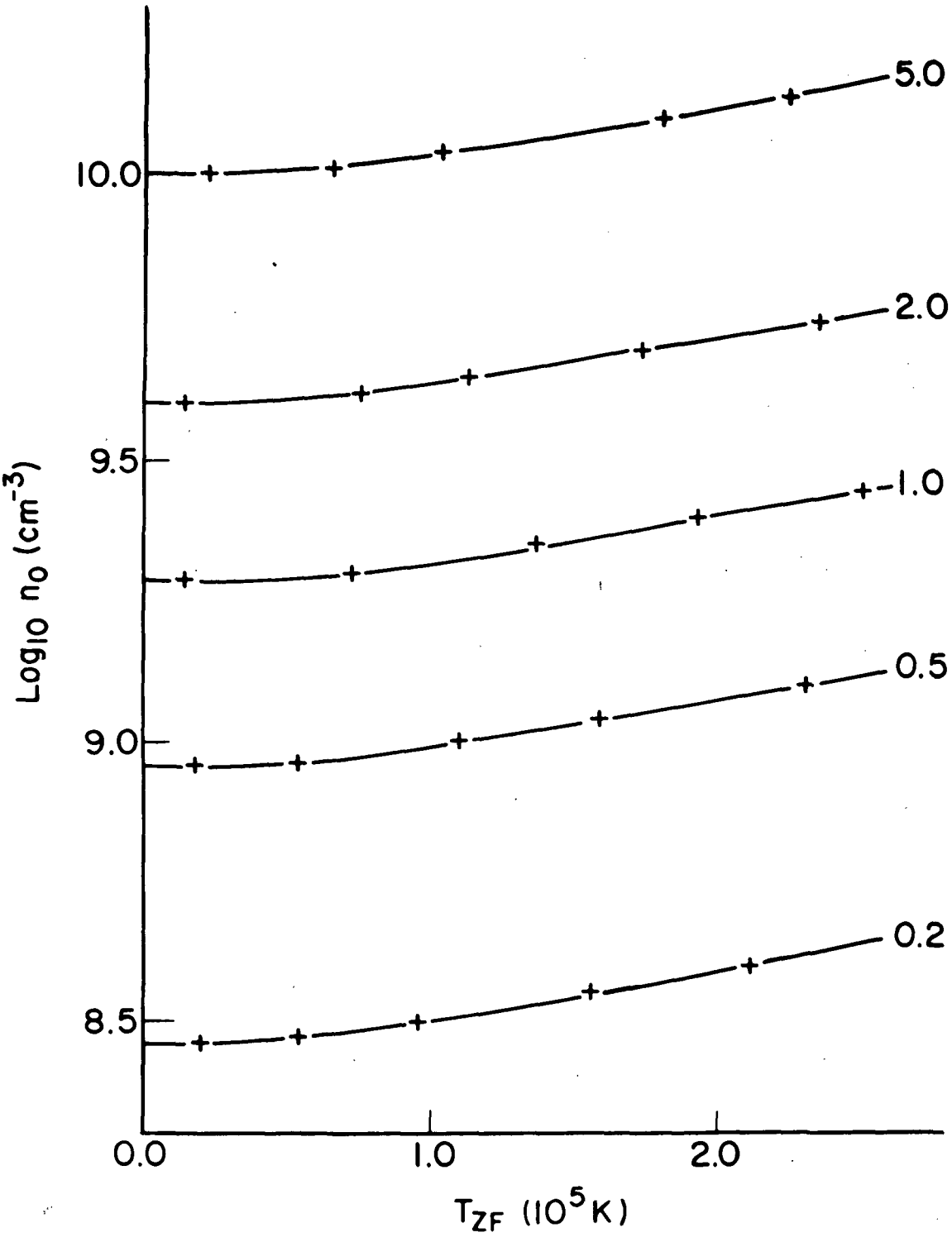


Figure 4.6. CURVES OF Log n_0 VS T_{ZF} FOR CONSTANT VALUES OF F_0. Each curve is labeled with its value of $-F_0$ in units of 10^6 erg cm^{-2} sec^{-1}.

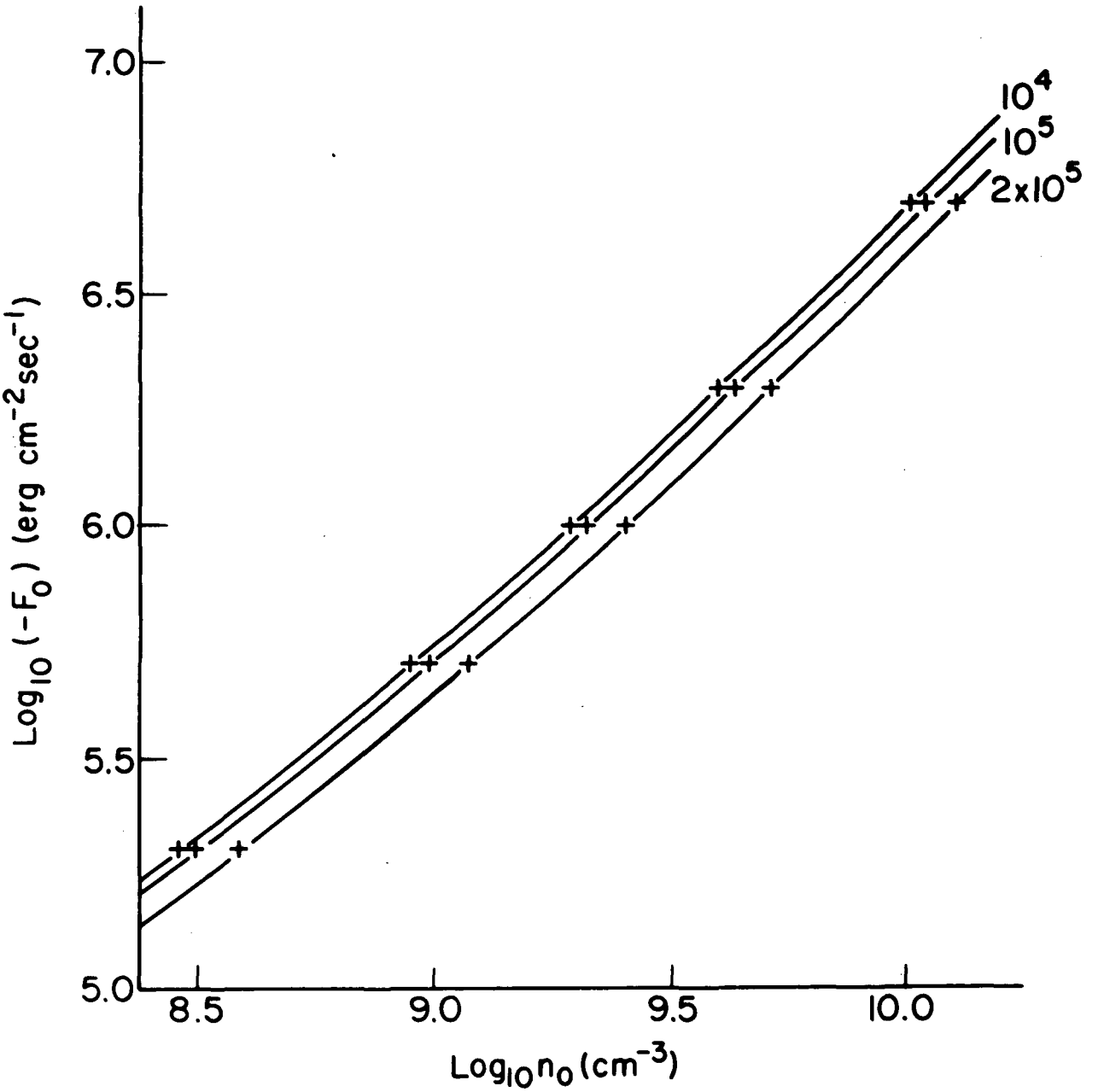


Figure 4.7. Log $(-F_0)$, Log n_0 CURVES FOR CONSTANT VALUES OF T_{ZF}. Each curve is labeled with its value of T_{ZF} in units of degrees K.

With the aid of the heat conduction law, equation (4.11), the energy equation (4.14) may be written as

$$FdF = \left[\frac{n_e}{n}\left(1 - \frac{n_e}{n}\right)\frac{L_r}{T^2}\kappa\right](nT)^2\, dT \quad . \tag{4.20}$$

Equation (4.7) shows that the quantity in square brackets is a function of the temperature alone:

$$\frac{n_e}{n}\left(1 - \frac{n_e}{n}\right)\frac{L_r}{T^2}\kappa \equiv f(T) \quad . \tag{4.21}$$

In conjunction with the heat conduction law, the equation of hydrostatic equilibrium yields the integral

$$nT = n_0 T_0 \exp\left[-\int_T^{T_0} \frac{\kappa}{HF}\, dT\right] \quad , \tag{4.22}$$

where H is the pressure scale height,

$$H = \frac{kT}{\left(1 - \frac{n_e}{n}\right) m_H g} \quad . \tag{4.23}$$

The quantity $-\int_T^{T_0} \kappa/HF\, dT$ is positive due to the fact that F is negative (heat flux downward). Even though F goes to zero, so that $-\kappa/HF$ goes to infinity, as T approaches T_{ZF}, it was found that

$$-\int_{T_{ZF}}^{T_0} \frac{\kappa}{HF}\, dT \ll 1 \tag{4.24}$$

in all cases for which the model was computed. That is, in the ranges of n_0 and F_0 in Figure 4.6, the pressure remains nearly constant through the model.

Equation (4.20) may therefore be written

$$FdF \approx n_0^2 T_0^2 f(T)\, dT \quad . \tag{4.25}$$

Integration of equation (4.25) from T_{ZF} to T_0 gives

$$\frac{1}{2} F_0^2 \approx n_0^2 T_0^2 \int_{T_{ZF}}^{T_0} f(T)\, dT \quad , \tag{4.26}$$

which shows that for constant T_0 and T_{ZF}, F_0 is approximately proportional to n_0. We note that this results from three physical properties of the model: (1) the rate of radiative cooling per unit volume is given by the square of the number density times a function of temperature alone, (2) the thermal conductivity is a function of temperature alone, and (3) the pressure is nearly constant throughout the model.

Integration of equation (4.25) from T_{ZF} to T shows that for constant n_0, T_0, and T_{ZF}, F is approximately a function of temperature alone. This is the reason why the curves in Figures 4.4 and 4.5 are symmetrical about T_{ZF}.

4.3.2 Comparison of the Model with XUV-Resonance-Line Data

4.3.2.1 Observational Evidence for a Planar Constant-Heat-Flux Region

The analyses of Athay[67] and Dupree and Goldberg[68] of the emission of XUV resonance lines from a planar transition region, in conjunction with the observed energy fluxes of several XUV resonance lines,

provide the empirical evidence for a planar constant-heat-flux region in the temperature range 10^5 K to 10^6 K. In these analyses the energy flux E_L observed at the earth in an XUV resonance line is related to the structure of the transition region by

$$E_L = C_L A P_e^2 G^{-1} \quad . \tag{4.27}$$

Here P_e and G are "representative" values of the electron pressure divided by Boltzmann's constant ($P_e \equiv n_e T$) and of the temperature gradient, respectively, in the layer of the transition region in which the line is formed. A is the abundance (element:hydrogen ratio) of the element which emits the line. C_L is a constant of proportionality, which is evaluated for each line by integrating a function of temperature (proportional to the number density of the ion which emits the line and the rate of excitation of the line) over the temperature range in which the ion is produced. This integrand is a sharply peaked function of temperature, contributing significantly to C_L only at temperatures within about a factor of two or less from the temperature $T_{max,L}$ at which it is maximum. Hence, equation (4.27) may be used to obtain an estimate of $A(n_e T)^2 (dT/dz)^{-1}$ at $T_{max,L}$ from the observed energy flux E_L.

The data points in Figure 4.8 (cf. Figure 3 of Dupree and Goldberg) were obtained in this way from the values of C_L and $T_{max,L}$ computed by Dupree and Goldberg. We used the observed values of E_L adopted by them for several XUV resonance lines emitted from oxygen ions and silicon ions formed in the temperature range of the transition region. The value of $(nT)^2 (dT/dz)^{-1}$ for each data point was obtained by assuming that the abundances of oxygen and silicon are the photospheric abundances

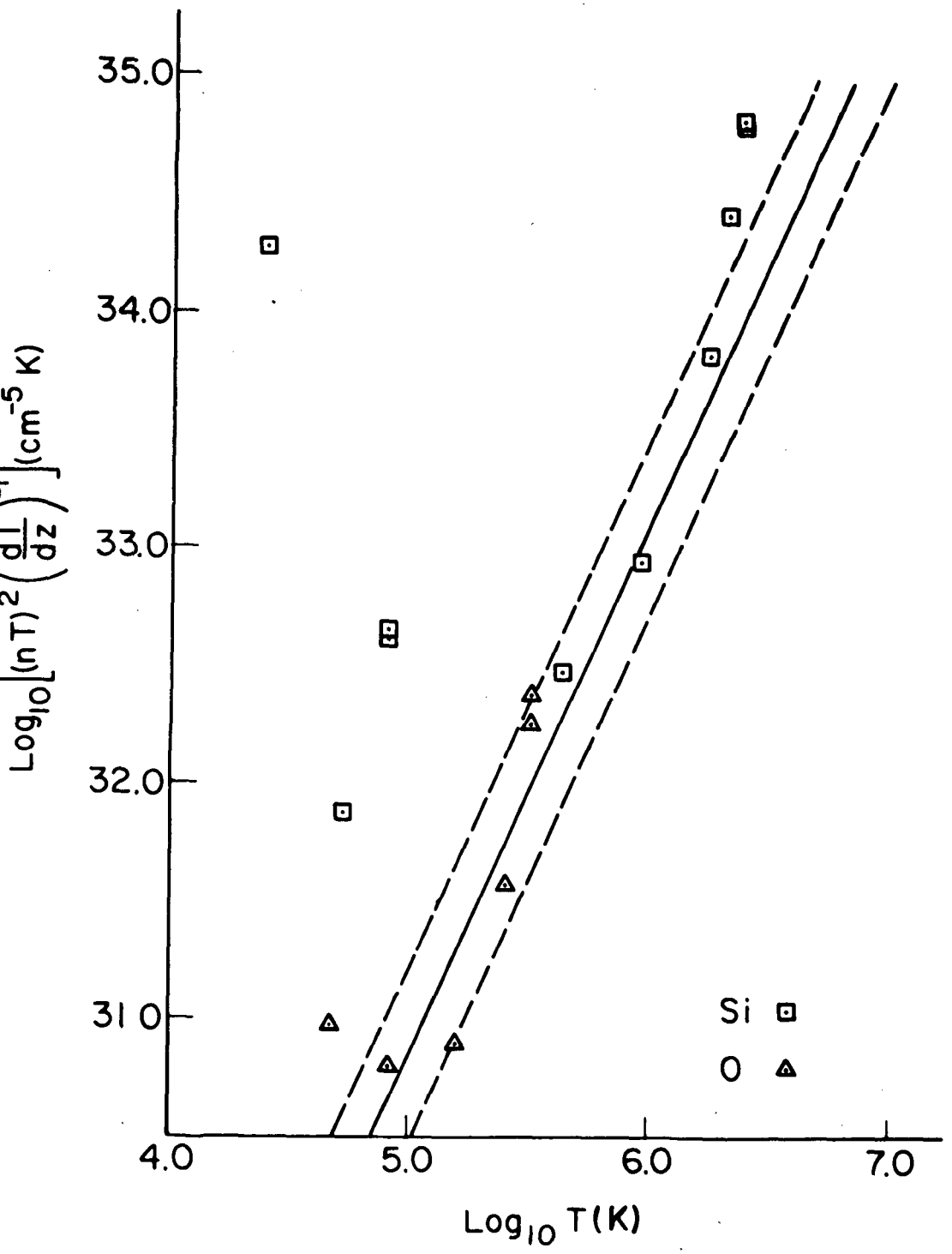


Figure 4.8. REDUCED XUV-RESONANCE-LINE DATA FROM DUPREE AND GOLDBERG (1967). In the temperature range 10^5 K to 10^6 K, the straight lines correspond to constant downward heat flux. The dashed lines show that the data scatter around the "best-fit" solid line by a factor of two.

of Goldberg et al,[69] $A_0 = 9.3 \times 10^{-4}$, $A_{Si} = 3.2 \times 10^{-5}$, and by taking $n = 2n_e$. Since the pressure remains nearly constant, the straight lines of slope 2.36 in Figure 4.8 correspond to constant downward heat fluxes between 10^5 K and 10^6 K, since above 10^5 K the heat conductivity is proportional to $T^{2.36}$. The solid line is the best visual fit to the data in the range 10^5 K to 10^6 K, while the dashed lines show that the data scatter about the solid line by a factor of two. The data thus imply that, between 10^5 K and 10^6 K, the transition region is approximately planar, and that in this region the heat flux varies by less than a factor of four. We see that although the heat flux need not stay completely constant in the constant-heat-flux region, a sizable fraction (25 percent or more) of the heat flux at the top of the transition region should reach the 10^5 K level.

4.3.2.2 Deduction of Values of n_0 and F_0 from the Line Data

The energy fluxes analyzed by Dupree and Goldberg were composed of radiation from all areas of the observable solar hemisphere, including some active regions. However, Withbroe[11] concludes from OSO-IV XUV-resonance-line observations (with spatial resolution of one minute of arc) that "... for XUV lines formed at temperatures less than about 10^6 K the flux radiated by the entire solar disk is characteristic of the equatorial quiet area." On this basis we assume that at 10^6 K and below the data in Figure 4.8 represent the quiet solar atmosphere. We then obtain from Figure 4.8 (at $T = T_0 = 10^6$ K) the following constraint on n_0 and F_0 for quiet regions:

$$\log n_0 = \frac{1}{2} \log (-F_0) + 6.06 \pm 0.15 \quad . \tag{4.28}$$

Withbroe[11] has studied the limb brightening of several XUV lines observed by OSO-IV in the quiet equatorial region. He finds good agreement between the observed limb brightening and that computed for a planar model transition region and lower corona having constant downward heat flux in the region below the 2×10^6 K level. Withbroe's analysis of the limb brightening data allows an estimate of the downward heat flux independent of the number density. Withbroe finds

$$\log \frac{\alpha}{(-F)} = -12.0 \pm 0.3 \quad , \tag{4.29}$$

where F is the constant downward heat flux and α is the coefficient in the thermal conductivity used in his model:

$$\kappa = \alpha T^{5/2} \quad . \tag{4.30}$$

For our adopted thermal conductivity, α has the value 1.0×10^{-6} (cgs units) at the 10^6 K temperature of the upper boundary. Using this value in equation (4.29), we obtain the value of F_0 which we adopt for the quiet regions:

$$\log (-F_0) = 6.0 \pm 0.3 \quad . \tag{4.31}$$

Equations (4.28) and (4.31) define the (n_0, F_0) region required by the data for quiet regions. From equation (4.28), for $\log (-F_0) = 6.0$,

$$\log n_0 = 9.06 \pm 0.15 \tag{4.32}$$

which is a reasonable value for the number density at the base of the quiet corona.

Using the model of Withbroe,[11] Noyes et al[38] have compared the observed XUV resonance line emission from quiet regions with that from active regions. They find that, for a fixed temperature, the total number density and the downward heat flux in active regions are each about five times larger than in quiet regions. For active regions we therefore adopt, in place of equations (4.28) and (4.31),

$$\log n_0 = \frac{1}{2} \log (-F_0) + 6.76 \pm 0.15 \qquad (4.33)$$

and

$$\log (-F_0) = 6.7 \pm 0.3 \quad . \qquad (4.34)$$

In Figure 4.9, the two $(\log n_0, \log F_0)$ regions adopted above for quiet regions and active regions of the sun are compared with the lines of constant T_{ZF} for the model. The data of Figure 4.8 require that $T_{ZF} < 1 \times 10^5$ K in quiet regions, and the results of Noyes et al imply that $T_{ZF} < 2 \times 10^5$ K in active regions. Hence, in the shaded areas of Figure 4.9, the model is compatible with the XUV-resonance-line observations.

4.4 Discussion

Given that $T_{ZF} < 2 \times 10^5$ K in both quiet regions and active regions, our model yields the observational finding that F_0 and n_0 increase by the same factor from quiet regions to active regions. Moreover, although there is some discrepancy, the model is compatible with the magnitudes of F_0 and n_0 required by the XUV observations of quiet regions and active regions. These two favorable results imply that, in a first

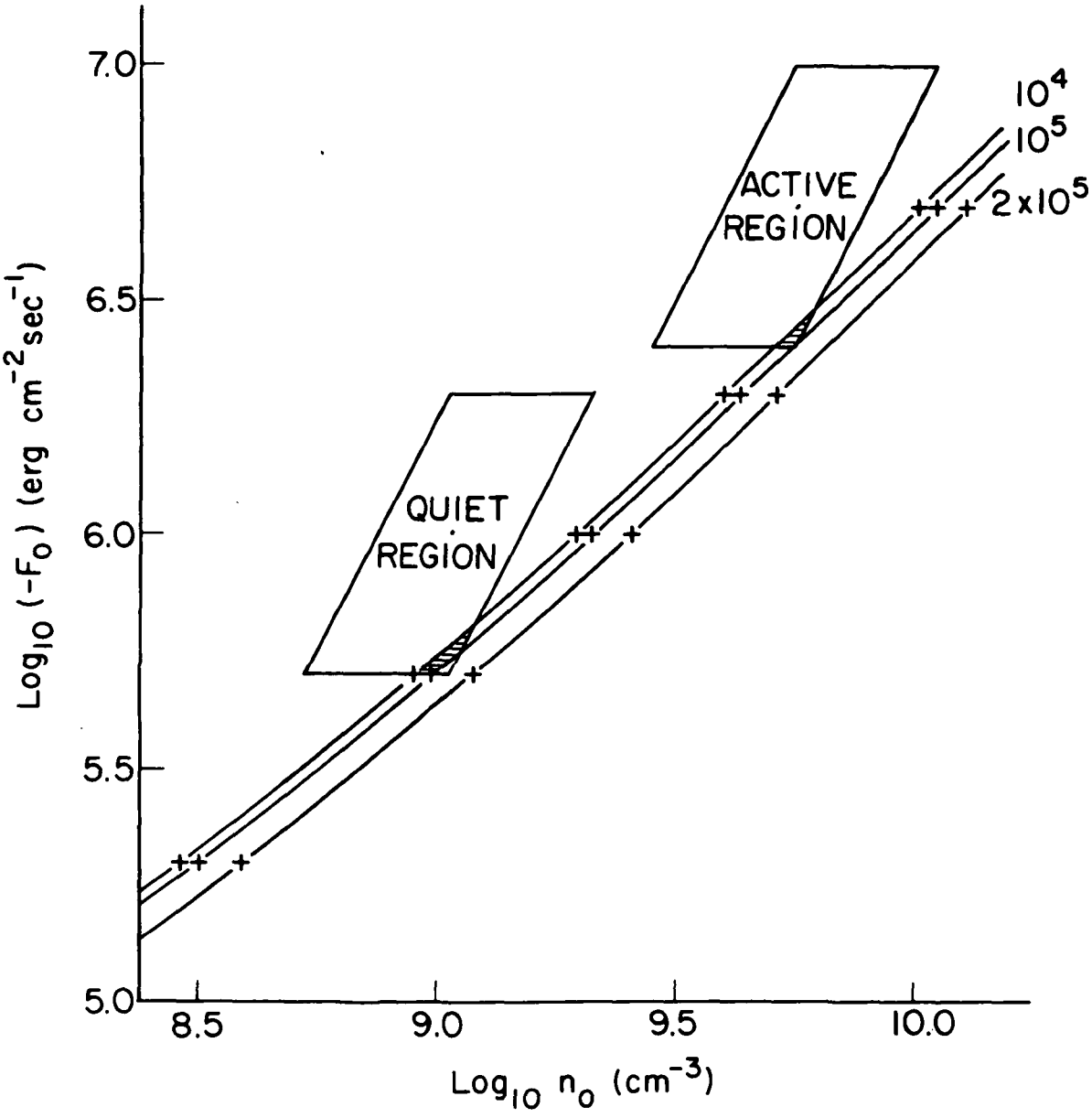


Figure 4.9. COMPARISON OF THE MODEL WITH THE REQUIREMENTS OF THE XUV-RESONANCE-LINE DATA. In the shaded regions the model is compatible with the data.

approximation, the model is a valid description of the structure and the energy balance of the transition region. That is, these results imply that the dominant process in the energy balance of the transition region is the removal of heat flux by radiation, and that, with respect to the overall balance of energy, the transition region may be considered to be static and planar.

Since the downward heat flux F_0, which leaves the base of the corona and enters the top of the transition region, is the dominant energy loss of the corona, the dominant processes in the energy balance of the corona are heating by mechanical wave dissipation and cooling by downward heat conduction. Hence, the value of F_0 is determined by the amount of mechanical-wave heating in the corona. The fact that the model agrees with both quiet regions and active regions implies that when the downward heat flux (and hence the upward mechanical-wave energy flux) increases, the density of the atmosphere automatically increases just enough to radiate away the larger heat flux. The fact that the model ignores the presence of mechanical waves and their dissipation in the transition region implies that the direct cause of an enhancement of the density in the transition region is an increased downward heat flux rather than the presence in the transition region of an increased upward mechanical wave flux. Therefore, it is appropriate to consider the magnitude of the density in the transition region to be the result of the downward heat flux F_0 entering the transition region, rather than considering the density in the transition region to be the cause of F_0.

The preceding paragraphs interpret the agreement between the model and the data. It is also instructive to consider the disagreement between the model and the data. It was found in Section 4.3.2.1 that the

data required that 25% or more of the heat flux which flows into the transition region from the corona be absorbed in the base region. But, as can be seen in Figure 4.5, the model base region is capable of absorbing only about 20% or less of the heat flux F_0 which enters the model transition region. This is the cause of the discrepancy between the model and the data in Figure 4.9. For there would be much better agreement between the model and the data in Figure 4.9, if, for $-F_0 = 1.0 \times 10^6$ erg cm^{-2} sec^{-1}, $T_{ZF} = 10^4$ K were given by $n_0 \simeq 1.0 \times 10^9$ cm^{-3} instead of by $n_0 \simeq 2.0 \times 10^9$ cm^{-3}. In Figure 4.5, we see that this would occur if about 40% (instead of 20%) of F_0 were absorbed by the model base region. That is, for the best observational values of F_0 and n_0, the fraction of F_0 which the model transmits to the base region is in agreement with the observations, but the model base region absorbs only about half of the heat flux which enters it. Thus, we conclude that the discrepancy between the model and the data in Figure 4.9 implies that the base region of the model absorbs only about half as much heat flux as the base region of the actual transition region.

We expect the actual base region to radiate away more heat than the planar model base region (see Appendix B). In Figure 4.4 we see that for $T_{ZF} = 1.5 \times 10^4$ K, the layer between 10^5 K and T_{ZF} is only 10 or 20 km thick. One effect of oscillations (of scale $\approx 10^3$ km) in the base region, due in part to the passage of the compression waves which heat the corona, should be to increase the surface area (and hence the volume) of the thin layer between 10^5 K and 10^4 K over that of the planar model, thus making the base of the transition region a more efficient radiator than if it were strictly planar. Another effect of the velocity fluctuations is to increase the number density in the base region, which

increases the radiative power output density. A factor of two increase in the amount of heat flux radiated away in the base region by these two mechanisms is quite plausible (see Appendix B). These considerations suggest, first, that the dominant process in the energy balance of the base region is the removal of heat flux by radiation and, second, that with respect to the energy balance, the base region is inherently non-static and/or nonplanar.

The preceding discussion suggests that the energy lost from the base of the transition region in the form of spicule kinetic energy is secondary to that lost by radiation. However, the spicule kinetic energy may still be derived from the downward heat flux as suggested by Kuperus and Athay.[64] If, for a given heat flux F_0, the number density n_0 is too low, the point (n_0,F_0) in Figure 4.9 will lie to the left of the $T_{ZF} = 10^4$ K line where the model is unable to radiate away all of F_0. In this case, due to our understanding of the cause of the density enhancement in active regions, we expect the atmosphere to react by increasing n_0 until F_0 can be radiated away. This suggests that the excess heat flux would go into raising material at the base of the transition region to higher levels. Spicules occur over the boundaries of the supergranulation cells where the convective motion of the cells has concentrated the magnetic field. The magnetic field fans out above the boundaries and funnels the downward flowing heat into regions over the cell boundaries, giving the supergranule boundary region heat flux F_0 which is larger than the average value for quiet regions.[70] This suggests that spicules are the manifestation of the atmosphere attempting to maintain an increased density over the supergranule boundary regions.

The transient character of spicules suggests an overstable situation in which the density is alternately higher or lower than the value required to balance the inflowing heat with radiative losses.

5. SUMMARY

The general problem which we have studied in this dissertation is the structure and heating (or energy balance) of the solar atmosphere. The general aim of our investigation has been to determine and understand the interdependence between the structure of the atmosphere and the heating of the atmosphere. The basic and widely accepted hypothesis underlying our study is that the fundamental source of heating for the outer layers of the atmosphere is the dissipation of mechanical waves which are emitted into the atmosphere from the top of the convection layer below the photosphere.

The solar atmosphere has the following salient observed features which are directly relevant to the structure and heating of the atmosphere. These features have been deduced primarily from observations at visible and XUV wavelengths.

1. In quiet regions and in active regions during their decay phase, the atmosphere appears to be approximately static and planar. The atmosphere appears to be most inhomogeneous (nonplanar) and nonstatic (fluctuating) at the base of the chromosphere-corona transition region.

2. Vertical oscillations with periods in the range 200 to 300 sec are observed in the photosphere and low chromosphere. The observations indicate that these oscillations are excited by the buffeting of the base of the photosphere by rising granulation convection cells at the top of the convection layer.

3. The corona, having a temperature in the range of 1 to 2 $\times$ 10^6 K, is the hottest layer of the atmosphere. This requires that the energy transfer to the corona from the inner layers is by some nonthermal process such as the propagation and dissipation of mechanical waves, rather than by thermal mechanisms such as thermal conduction or radiative transfer.

4. The temperature increases slowly through the chromosphere from about 4,600 K at the top of the photosphere to about 10,000 K at the top of the chromosphere at a height of about 2000 km. The temperature then increases very steeply through the chromosphere-corona transition region, passing from 10^4 K to 10^6 K in a vertical distance of about 3000 km. In the corona, the temperature again increases slowly with height from 10^6 K at the top of the transition region to the temperature maximum of about 2×10^6 K some tens of thousands of kilometers above the transition region.

5. The major energy loss of the corona (dominating losses due to radiative cooling and the solar wind) is downward heat conduction to the transition region, and is of order 10^6 erg cm^{-2} sec^{-1}. The corona must be continually resupplied with energy at this rate.

6. In the average active region, the downward heat flux and the number density at the base of the corona are both about five times larger than in quiet regions.

7. In the transition region, the downward heat flux remains approximately constant between the levels of 10^6 K and 10^5 K.

8. The total rate of radiative cooling of the transition region is of the same order as the rate of conduction heating (10^6 erg cm^{-2} sec^{-1}).

9. The entire atmosphere above the photosphere radiates away about 10^7 erg cm^{-2} sec^{-1}. Some of this energy may be supplied by radiative transfer from the photosphere. Hence, 10^7 erg cm^{-2} sec^{-1} is an upper bound on the total rate of heating of the atmosphere by mechanical waves.

10. In the transition region and low corona, the magnetic field is more or less vertical both in quiet regions and in active regions during their decay phase. Therefore, the magnetic field does not strongly affect the downward flow of heat from the corona to the transition region. The average field strength is of order 1 gauss in quiet regions and of order 50 gauss in active regions during their decay phase. In quiet regions, the gas pressure dominates the magnetic pressure in the photosphere and chromosphere, while in the transition region and low corona, the magnetic pressure is comparable to the gas pressure. In active regions in their decay phase, the magnetic pressure is dominated by the gas pressure in the photosphere, but is greater than the gas pressure above the first few hundred kilometers of the chromosphere.

Since, in the first approximation, the solar atmosphere appears to be static and planar, our basic approach to the structure and heating problem has been to study the average (over horizontal area) vertical structure and heating of the atmosphere by means of a static, planar model. The model is assumed to be in hydrostatic equilibrium and to have a temperature profile representative of the actual solar atmosphere. In addition, the magnetic field is neglected or assumed to be vertical, and the wave propagation is assumed to be governed by the linearized equations of motion. With these simplifying approximations, we have studied wave propagation, wave dissipation, heat conduction and radiative cooling in the model, and the connection between these processes and the structure of the model. We have found that the properties of this model are compatible with and aid the physical interpretation of the above observed properties of the solar atmosphere.

The following picture of the structure and heating of the solar atmosphere has emerged from the study of our model in conjunction with the above features of the observed atmosphere. In this picture, the central role is played by the chromosphere-corona transition region, both as the most outstanding feature of the structure and in providing the basic mechanism which connects the structure and the heating of the atmosphere.

1. The photospheric oscillations excited by the granules are composed of a spectrum of compression waves. The oscillations with periods near 200 sec are due to upward propagating waves. The longer period oscillations are due to waves which are evanescent in the photosphere and lower levels of the chromosphere, but which become upward propagating waves higher in the chromosphere.

2. Due to the temperature profile of the chromosphere and transition region, only the components of the photospheric oscillations with very nearly vertical wave-vectors are passed to the corona. Other components are reflected back toward the photosphere by the chromosphere and transition region.

3. The observed amplitude, orientation and period range of the photospheric oscillations indicate the presence of sufficient energy flux in the appropriate wave modes to provide the heating of the outer atmosphere.

4. The vertically propagating compression waves which are passed to the corona are dissipated in the corona by thermal conduction and Landau damping rather than by developing into shock waves.

5. The existence of a high-temperature corona and a narrow transition region is due to the fact that the radiative cooling capacity of the corona is insufficient to balance the mechanical heating. Hence, the temperature of the corona rises until downward conduction cooling balances the heating.

6. Within the transition region, in both quiet regions and active regions, the dominant processes in the energy balance are radiative cooling and heating by thermal conduction from the corona. In particular, conduction heating dominates heating by wave dissipation.

7. The balance of conduction heating and radiative cooling in the transition region determines the number density at the base of the corona. The number density at the base of the corona is approximately proportional to the downward heat flux leaving the corona. Thus, in active regions increased heating of the corona results in increased downward heat flux which, in turn, causes the density enhancement in the transition region and corona in active regions. Spicules may be the manifestation of the atmosphere attempting to maintain an enhanced density over the supergranule boundary regions.

8. The base of the transition region is inherently nonplanar and fluctuating.

Appendix A

THERMAL CONDUCTIVITY IN THE TRANSITION REGION

For the purposes of our model, we need an estimate of the thermal conductivity as a function of temperature which differs from the actual conductivity by less than a factor of two. The temperature range of the transition region is from 10^4 K, where hydrogen is nearly completely neutral, to 10^6 K, where hydrogen is essentially completely ionized. Since the solar atmosphere is composed of about 90% hydrogen by number, the thermal conductivity of the transition region is approximated with sufficient accuracy by the thermal conductivity of partially ionized hydrogen in the temperature range from 10^4 K to 10^6 K.

It was assumed, in computing the radiative energy losses and the degree of ionization in the model, that all excitations and ionizations of the hydrogen atoms are collisional, and all recombinations and deexcitations are radiative. We also assume this condition in estimating the thermal conductivity. Under this condition, only the thermal kinetic energy of the electrons, protons, and hydrogen atoms contributes to the heat conduction; the energy of excitation and ionization of the hydrogen atoms does not contribute.

The thermal conductivity of partially ionized hydrogen due to the transport of thermal kinetic energy has been computed by Devoto[65] from accurate but rather complicated kinetic theory. It appears that his computed values should approximate the actual conductivity with an accuracy of the order of 10%. The curves of Figure A.1 have been obtained from Devoto's table of the conductivity as a function of temperature at constant pressures of 10^4, 10^5, 10^6, and 10^7 times 1.013 dyne cm^{-2} (1.013 dyne $cm^{-2} = 10^{-6}$ atm). None of these curves can be taken for the thermal

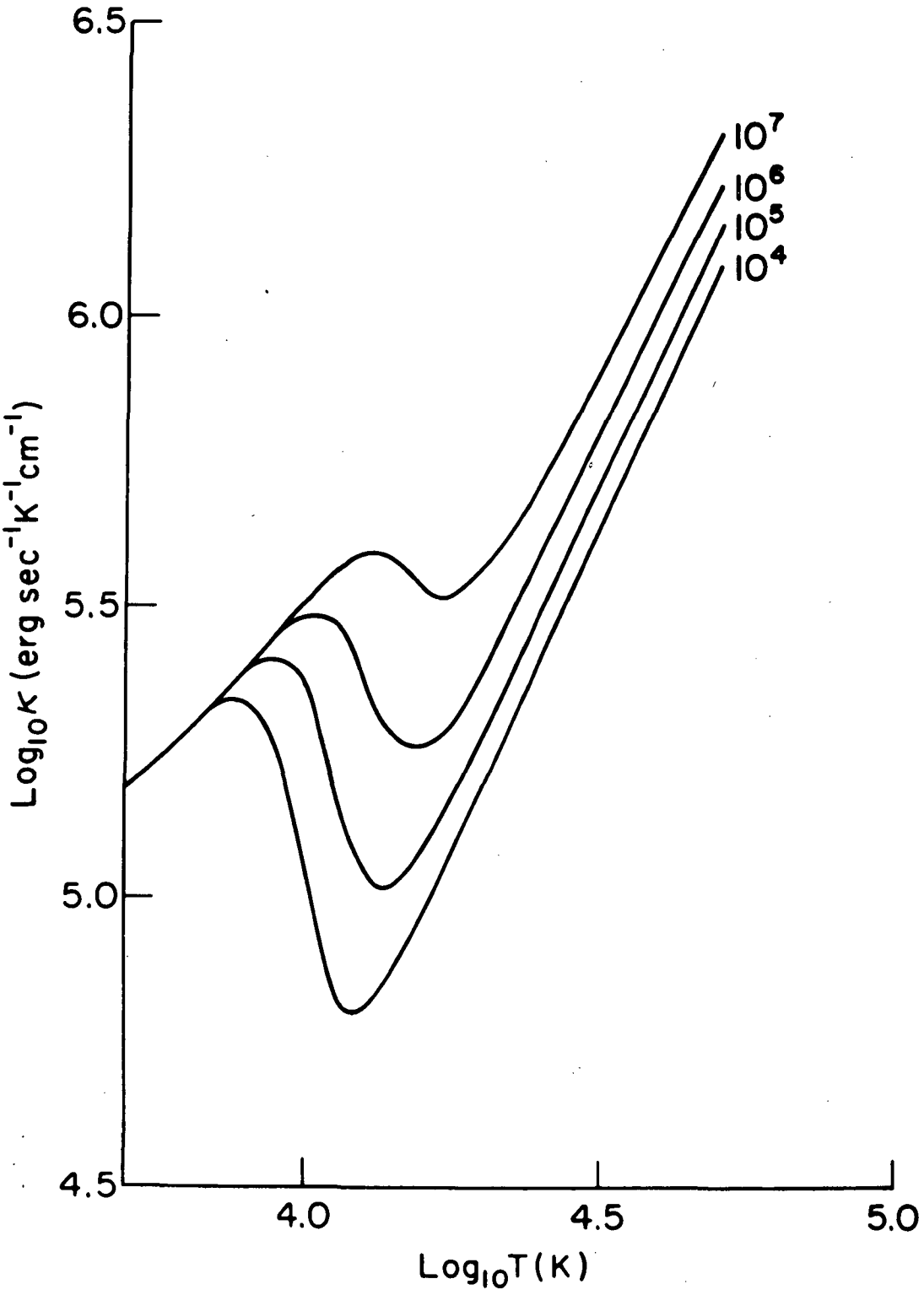


Figure A.1. CONDUCTIVITY CURVES OBTAINED FROM DEVOTO'S (1968) TABLE FOR PARTIALLY IONIZED HYDROGEN AT CONSTANT PRESSURE. Each curve is labeled with its value of the pressure in units of 1.013 dyne cm^{-2}.

conductivity of the transition region for the following two reasons. First, the conductivity computed by Devoto obviously depends on the pressure, and the pressures adopted in Devoto's computations are much larger than the pressure of about 0.2 dyne cm^{-2} [13] in the transition region. Second, the conductivity depends on the degree of ionization. In computing the degree of ionization, Devoto assumed local thermodynamic equilibrium. In this case, the degree of ionization depends on the pressure; an increase of pressure at constant temperature causes a decrease in the degree of ionization. But in the transition region, under the condition of collisional excitation and ionization, and radiative recombination and de-excitation, the plasma is not in local thermodynamic equilibrium; the degree of ionization depends only on the temperature and not on the pressure [see equations (4.7), (4.15), and (4.16)]. Thus, we would not be justified in adopting Devoto's published values of the thermal conductivity of partially ionized hydrogen for the thermal conductivity of the transition region.

However, instead of repeating Devoto's involved and lengthy calculations for the case of the transition region, we can estimate the conductivity of the transition region with sufficient accuracy from Devoto's values by deriving from simple kinetic theory the manner in which the conductivity scales with pressure and degree of ionization. Using simple kinetic theory,[71] we can derive the following formula for the thermal conductivity of partially ionized hydrogen:

$$\kappa = \frac{9}{4} \frac{k^{3/2}}{m_H^{1/2}} \frac{T^{1/2}}{(\sigma_{HH} + n_p/n_H \sigma_{Hp})} + \kappa_e \quad . \tag{A.1}$$

Here σ_{HH} and σ_{Hp} are "collision cross-sections" for energy exchange between neutral hydrogen atoms, and between neutral hydrogen atoms and protons, respectively. Since σ_{HH} and σ_{Hp} are of the nature of cross-sections for binary collisions, we expect them to depend on the temperature but not on the number densities. κ_e in equation (A.1) is the thermal conductivity of fully ionized hydrogen, which may be computed from the formula given by Delcroix and Lemaire:[66]

$$\kappa_e = 1.890 \times 10^{-5} \frac{T^{5/2}}{\ln \Lambda} \text{ erg sec}^{-1} \text{ K}^{-1} \text{ cm}^{-1} \quad , \tag{A.2}$$

where

$$\left.\begin{aligned} \Lambda &= \frac{3}{\sqrt{2\pi}} \frac{k^2}{e^3} \frac{T^2}{p^{1/2}} && \text{for } T < 4.2 \times 10^5 \text{ K} \\ \Lambda &= \frac{3}{\sqrt{2\pi}} \frac{k^2}{e^3} \frac{T^2}{p^{1/2}} \left(\frac{4.2 \times 10^5}{T}\right)^{1/2} && \text{for } T > 4.2 \times 10^5 \text{ K} \end{aligned}\right\} \tag{A.3}$$

Here e is the electron charge is esu and p is the pressure. Equation (A.1), in conjunction with equations (A.2) and (A.3), gives explicitly the dependence of the conductivity on the ion-neutral ratio n_p/n_H and on the pressure p. We use this property of equation (A.1) to estimate the conductivity of the transition region from Devoto's results shown in Figure A.1.

The estimate of the conductivity in the transition region is obtained as follows. First, Devoto's results are used to determine functions of temperature representing σ_{HH} and σ_{Hp} such that equation (A.1) approximately reproduces Devoto's conductivity. This is done by

fitting equation (A.1) to the curves in Figure A.1, using Devoto's values of n_p/n_H and p. The good fit shown in Figure A.2 is given by

$$\sigma_{HH} = 9.12 \times 10^{-14}\ T^{-1/2}\ cm^2 \qquad (A.4)$$

and

$$\sigma_{Hp} = 7.95 \times 10^{-11}\ T^{-1}\ cm^2 \quad . \qquad (A.5)$$

with these functions of temperature for σ_{HH} and σ_{Hp}, equation (A.1) reproduces Devoto's conductivity with an error of not more than about 10%. The thermal conductivity for the transition region is then computed from equation (A.1), keeping these functions of temperature for σ_{HH} and σ_{Hp}, but using values of n_p/n_H and p appropriate for the transition region. We assumed 0.2 dyne cm^{-2} for p, and n_p/n_H was calculated as a function of temperature from equations (4.7), (4.15), and (4.16). The resulting conductivity curve is shown in Figure 4.3.

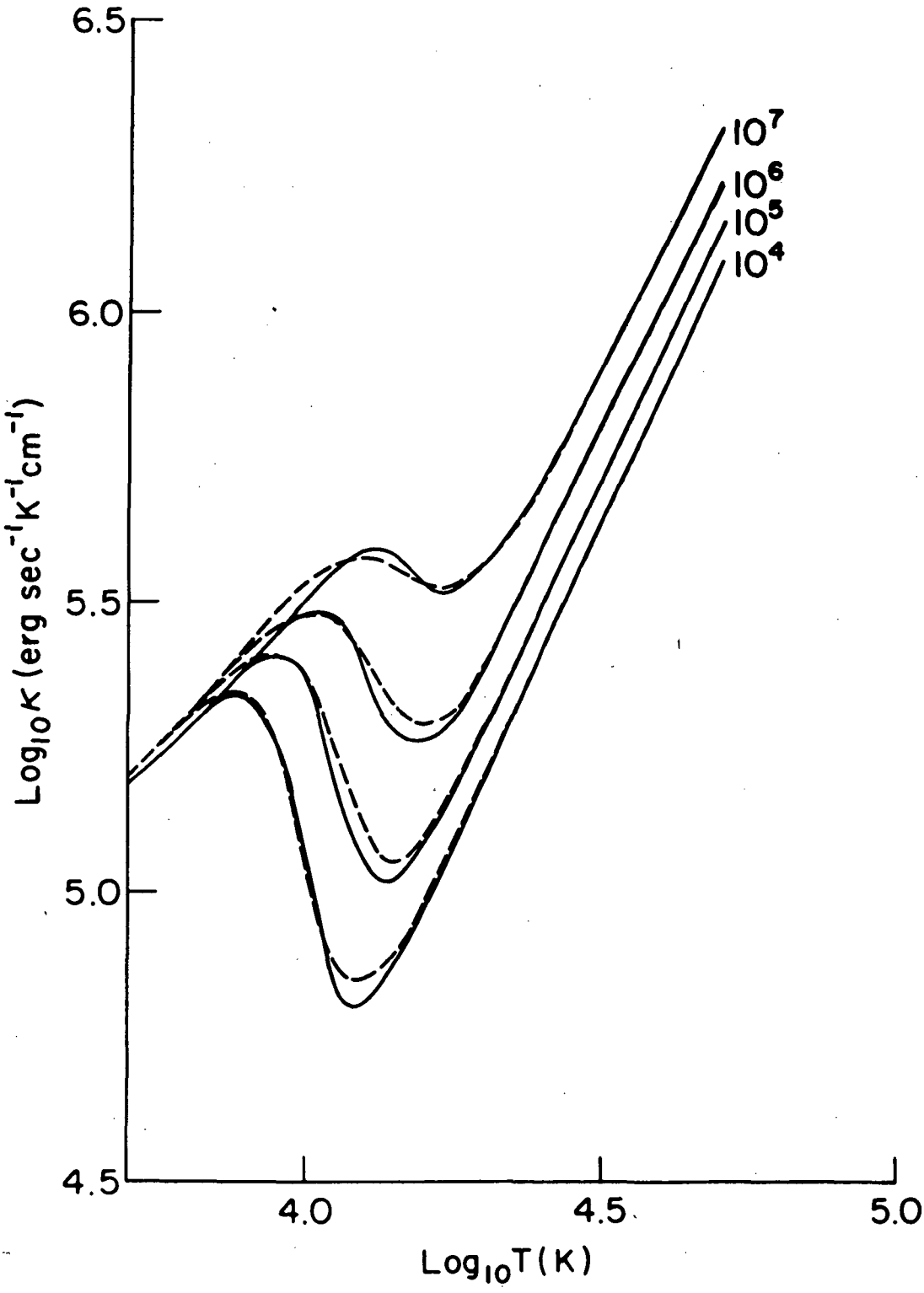


Figure A.2. COMPARISON OF THE CONDUCTIVITY GIVEN BY EQUATION (A.1) (DASHED CURVES) WITH DEVOTO'S CONDUCTIVITY (SOLID CURVES) FOR $\sigma_{HH} = 9.12 \times 10^{-14}\ T^{-1/2}\ cm^2$ and $\sigma_{Hp} = 7.95 \times 10^{-11}\ T^{-1}\ cm^2$.

Appendix B

ENERGY BALANCE IN THE BASE REGION

There are three physically distinct energy fluxes which enter into the energy balance of the solar atmosphere: mechanical energy flux $\vec{F}_m$, radiative energy flux $\vec{F}_r$, and heat flux $\vec{F}_c$ due to thermal conduction. We assume that, on the average over a sufficiently large area and over a sufficiently long time, each of these fluxes is vertical and steady. In other words, we assume that on the scale of sufficiently large horizontal lengths and sufficiently long times, the structure and energy balance of the solar atmosphere may be considered to be horizontally uniform and steady in time. Each of the energy fluxes may vary with height in the atmosphere by depositing thermal energy in the atmosphere or removing thermal energy from the atmosphere. The general form of the steady-state energy equation in terms of the energy fluxes is

$$\frac{dF_m}{dz} + \frac{dF_r}{dz} + \frac{dF_c}{dz} = 0 \quad . \tag{B.1}$$

In our model transition region we have assumed that F_m is constant,

$$\frac{dF_m}{dz} = 0 \quad , \tag{B.2}$$

and that there is no absorption of radiation,

$$\frac{dF_r}{dz} > 0 \quad . \tag{B.3}$$

Under these two conditions, the energy equation requires that the downward heat flux in the model transition region be absorbed by radiative losses.

However, the comparison of our model with XUV-resonance-line data indicates that the actual base region absorbs about twice as much heat flux as the model base region (see Section 4.4). This requires that

$$\int_{\text{base region}} \left(\frac{dF_m}{dz} + \frac{dF_r}{dz} \right) dz$$

for the actual base region be larger than that for the model base region. An equivalent statement is that one or any combination of the following statements is true of the actual base region with respect to the model base region:

1. The actual base region is thicker than the model base region, i.e., the volume of the actual base region is larger than that of the model base region.

2. dF_r/dz is larger in the actual base region than in the model base region, i.e., the radiative power output per unit volume is larger in the actual base region than in the model base region.

3. dF_m/dz is larger in the actual base region than in the model base region, i.e., $dF_m/dz > 0$ in the actual base region.

Each of these statements names a possible mechanism by which the actual base region is able to absorb more heat flux than the model base region. The basic nature of the energy balance of the base region depends upon which of these mechanisms, if any, is dominant in the actual base region. If either of the first two mechanisms is dominant, the dominant process in the energy balance is the removal of heat flux by radiative losses as

is assumed in the model. If the third mechanism is dominant, the dominant process in the energy balance is the conversion of thermal energy to mechanical energy.

The presence of the mechanical energy flux has been neglected in the model transition region. In the following paragraphs we estimate the effect of the mechanical energy flux on the structure of the base region. We find that the presence of the mechanical energy flux can reasonably increase the absorption of heat flux in the base region through mechanisms 1 and 2 sufficiently to account for the amount of heat flux absorbed by the actual base region. This suggests that the dominant process in the energy balance of the base region is the absorption of heat flux by radiation.

Most of the mechanical energy flux in the transition region is carried by upward propagating compression waves. We can estimate the velocity fluctuations produced in the base region by these compression waves from the formula for the energy flux carried by one-dimensional sound waves,

$$F_m = \rho \overline{v^2} a \quad . \tag{3.12}$$

In terms of the temperature T and the pressure $p = nkT$, equation (3.12) is

$$F_m = \left(\gamma \frac{m}{k}\right)^{1/2} \overline{v^2} p T^{-1/2} \quad . \tag{B.4}$$

At the base of the transition region $F_m \approx 10^6$ erg cm^{-2} sec^{-1}, $p \approx 0.2$ dyne cm^{-2}, $T \approx 10^4$ K, $m \approx m_H = 10^{-23.78}$ gm, and $\gamma \approx 5/3$. So, from

equation (B.4), the rms velocity fluctuation produced in the base region by the compression waves is

$$v_{rms} \approx 2.0 \times 10^{6} \text{ cm sec}^{-1} \quad . \tag{B.5}$$

This estimate should be accurate to order of magnitude.

The vertical displacement Δz resulting from the velocity oscillations of a wave of period P is approximately

$$\Delta z = \frac{1}{2} P v_{rms} \quad . \tag{B.6}$$

The compression waves which carry energy to the corona are expected to have periods of the order of 10^{2} sec. We therefore expect that vertical displacements of the order of 10^{3} km are produced at the base of the transition region by the upward propagating compression waves. Assuming that the compression waves are generated by the motions of the photospheric granules, the horizontal scale of the velocity fluctuations at the base of the transition should be comparable to the distance between centers of adjacent granules. The average distance between centers of adjacent granules is about 2,000 km, which length is also representative of the horizontal dimension of vertical oscillations observed in the chromosphere.[16] Thus, it appears that the base region cannot be considered to be horizontally uniform on a scale of 10^{3} km or less.

We are now in a position to estimate the effect of the vertical oscillations on the radiative capacity of the base region through mechanism 1. That is, we may now estimate the factor by which the vertical oscillations increase the volume of the actual base region over that of the planar model base region. We have found from our model transition region

that the thickness of the base region (the distance between the 10^5 K level and the 10^4 K level) is of the order of 10 km. Therefore, the actual base region may be thought of as a thin surface layer coating the 10^4 K "level" which is distorted into a wavy surface by the vertical displacements of the compression waves. The volume of the base region is the product of its surface area and thickness. The thickness is approximately inversely proportional to the heat flux entering the base region, and the heat flux entering the base region is inversely proportional to the surface area of the base region. Hence, the fractional increase V_a/V_p of the volume of the actual base region over that of the planar model transition region is approximately the square of the fractional increase A_a/A_p of the surface area of the actual base region over that of the planar model:

$$\frac{V_a}{V_p} \approx \left(\frac{A_a}{A_p}\right)^2 \quad . \tag{B.7}$$

To estimate A_a/A_p, we consider the hexagonal arrangement of oscillating cells shown in Figure B.1. Each +, -, or 0 marks the center of a cell. The distance between adjacent centers is assumed to be 2,000 km. One third of the centers, those marked +, are assumed to be displaced 1,000 km above the mean level of the 10^4 K surface; another third of the centers, marked -, are assumed to be 1,000 km below the mean level; and the remaining third of the centers, marked 0, are assumed to have no displacement from the mean level of the 10^4 K surface. The 10^4 K surface is assumed to pass smoothly through all of the cell centers. The fractional increase in the surface area of this 10^4 K surface is conservatively estimated by the ratio of the area of the triangle formed by the

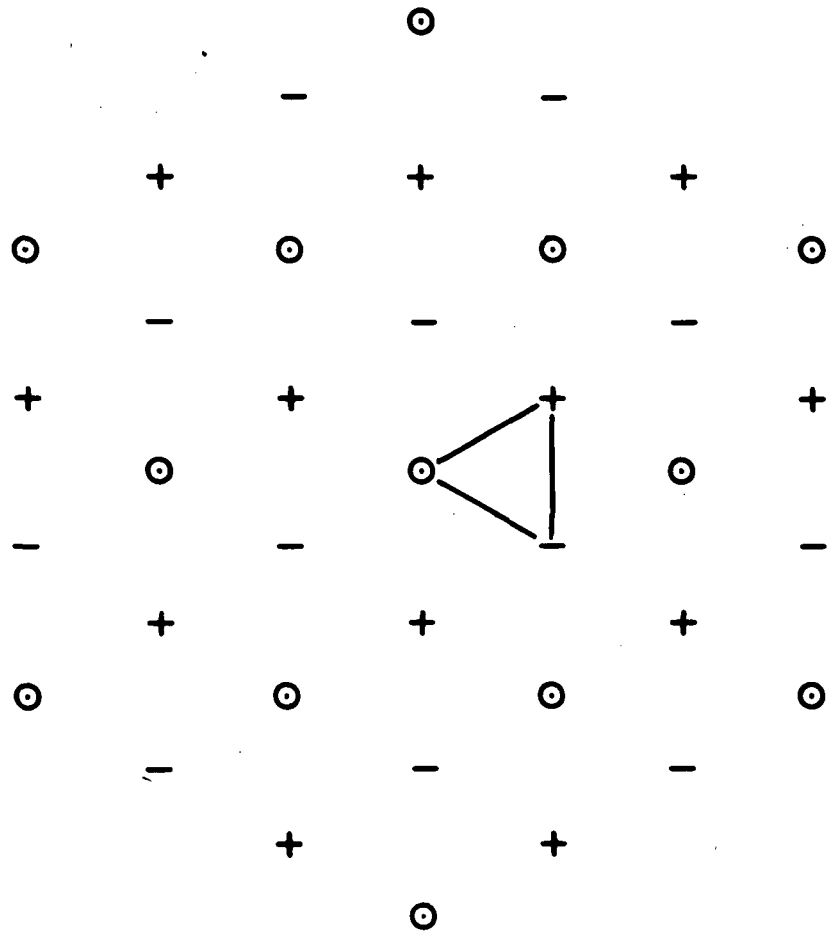


Figure B.1. CONFIGURATION OF CENTERS OF OSCILLATING CELLS ADOPTED TO ESTIMATE THE INCREASE IN THE SURFACE AREA OF THE 10^4 K LEVEL DUE TO THE VERTICAL DISPLACEMENTS OF THE OSCILLATIONS.

centers of any three adjacent +, -, and 0 cells to the projected area of this triangle on the horizontal plane. An example of such a triangle, viewed from above, is shown in Figure B.1. This estimate gives

$$\frac{A_a}{A_p} = \sqrt{2} \quad , \tag{B.8}$$

and

$$\frac{V_a}{V_p} \approx 2 \quad . \tag{B.9}$$

Thus, the nonplanar nature of the actual base region may reasonably be expected to increase the radiative capacity of the actual base region over that of the planar model by a factor of 2, which is sufficient to explain the discrepancy between the model and the XUV-resonance-line data.

We next consider the increase in the radiative capacity of the base region due to the effect of the mechanical energy flux through mechanism 2, i.e., due to the increase in the radiative power output density resulting from the presence of the mechanical energy flux. Under the approximation of collisional excitation and ionization, and radiative recombination and de-excitation, the radiative power output density is proportional to the square of the number density. Hence, we can estimate the increase in the radiative power output density by estimating the increase in the number density in the base region due to the presence of the mechanical flux.

We assume that on the average over a sufficiently long time the base region is in hydrostatic equilibrium, so that

$$\frac{dp}{dz} = -nmg \quad , \tag{B.10}$$

where dp/dz is the total pressure gradient and n is the total number density. The total pressure p is the sum of the thermal pressure $p_{th} = nkT$ and the turbulent pressure p_{turb} due to the velocity fluctuations in the base region. Therefore,

$$\frac{dp}{dz} = \frac{dp_{th}}{dz} + \frac{dp_{turb}}{dz} = -nmg \quad , \tag{B.11}$$

and we may estimate the increase in the number density Δn over what it would be if p_{turb} were absent by setting

$$\frac{dp_{turb}}{dz} = -\Delta nmg \quad . \tag{B.12}$$

The gradient of the turbulent pressure may be estimated as follows. Assuming that the velocity fluctuations are isotropic, the turbulent pressure is given by

$$p_{turb} = \frac{1}{3} \rho v_{rms}^2 \tag{B.13}$$

Assuming that v_{rms} is related to F_m by equation (3.12), we have

$$p_{turb} = \frac{1}{3} \frac{F_m}{a} \quad . \tag{B.14}$$

The condition $dF_m/dz = 0$, which we have assumed for the base region, then gives

$$\frac{dp_{turb}}{dz} = -\frac{F_m}{6aT}\frac{dT}{dz} \quad . \tag{B.15}$$

We may now estimate Δn. From equations (B.12) and (B.15),

$$\Delta n = \frac{F_m}{6mgaT}\frac{dT}{dz} \quad . \tag{B.16}$$

Here,

$$\frac{dT}{dz} \approx \frac{\Delta T}{\Delta z} \quad , \tag{B.17}$$

where ΔT is the temperature change in the base region ($\Delta T \approx 10^5$ K) and Δz is the vertical extent of the region of the atmosphere in which this temperature change occurs. Due to the vertical displacements of the upward propagating compression waves, Δz is of the order of 10^3 km for the temperature range of 10^4 K to 10^5 K. Adopting $\Delta z = 10^3$ km, $\Delta T = T = 10^5$ K, and $F_m = 10^6$ erg cm^{-2} sec^{-1} in equations (B.16) and (B.17), we obtain

$$\Delta n \approx 1.4 \times 10^{10} \text{ cm}^{-3} \quad . \tag{B.18}$$

This estimate is accurate only to order of magnitude, but it indicates that the increase in the number density in the base region due to the turbulent pressure is of the same order as the density in our model base region ($n \approx 10^{10}$ cm^{-3} at $T = 10^5$ K) which has no turbulent pressure. Therefore, it appears that the presence of the velocity fluctuations could easily increase the radiative capacity of the base region as much through density increase (mechanism 2) as through volume increase (mechanism 1). Again, this suggests that most of the heat flux entering the base region is converted into radiation rather than into mechanical energy flux.

REFERENCES

1. D. E. Billings, A Guide to the Solar Corona, Academic Press, New York (1966), p. 257.

2. H. C. van de Hulst, The Sun, ed. G. P. Kuiper, The University of Chicago Press (1953), p. 209.

3. C. W. Allen, Astrophysical Quantities, Athlone Press, University of London (1963), p. 164.

4. E. Tandberg-Hanssen, Solar Activity, Blaisdell Publishing Company, Waltham, Massachusetts (1967), p. 103.

5. J. D. Brandt and P. W. Hodge, Solar System Astrophysics, McGraw-Hill, New York (1964), p. 40.

6. Reference 5, p. 80.

7. Reference 2, p. 212.

8. Reference 4, p. 104.

9. Z. Suemoto and E. Hiei, Publ. Astron. Soc. Japan 14, 33 (1962).

10. Reference 1, p. 238.

11. G. L. Withbroe, Solar Phys. 11, 42 (1970).

12. G. Newkirk, Jr., Ann. Rev. Astron. and Astrophys. 5, 213 (1967).

13. R. G. Athay, Solar Phys. 9, 51 (1969).

14. O. Gingerich and C. DeJager, Solar Phys. 3, 5 (1968).

15. C. DeJager, Encyclopedia of Physics, ed. S. Flügge, Springer-Verlag, Berlin (1959), p. 81.

16. R. B. Leighton, Ann. Rev. Astron. and Astrophys. 1, 19 (1963).

17. Reference 3, p. 185.

18. E. N. Frazier, Astrophys. J. 152, 557 (1968).

19. R. W. Noyes, Aerodynamic Phenomena in Stellar Atmospheres, ed. R. N. Thomas, Academic Press, New York (1967), p. 293.

20. J. W. Evans and R. Michard, Astrophys. J. 136, 493 (1962).

21. J. W. Evans, P. Main, R. Michard, and R. Servajean, Astrophys. J. 136, 682 (1962).

22. F. Meyer and H. U. Schmidt, Trans. International Astron. Union, Volume XII B, 559 (1966).

23. R. B. Leighton, Stellar and Solar Magnetic Fields, ed. R. Lüst, North Holland Publishing Company, Amsterdam (1965), p. 158.

24. R. Howard, Ann. Rev. Astron. and Astrophys. 5, 1 (1967).

25. J. M. Beckers, Solar Phys. 3, 367 (1968).

26. J. M. Beckers, Astrophys. J. 140, 1339 (1964).

27. Reference 15, p. 183.

28. C. Sawyer, Ann. Rev. Astron. and Astrophys. 6, 115 (1968).

29. Reference 3, p. 182.

30. K. O. Kiepenheur, The Sun, ed. G. P. Kuiper, The University of Chicago Press (1953), p. 432.

31. V. Bumba, Plasma Astrophysics, ed. P. A. Sturrock, Academic Press, New York (1967), p. 77.

32. H. Zirin, The Solar Atmosphere, Blaisdell Publishing Company, Waltham, Massachusetts (1966), p. 373.

33. L. H. Aller, Astrophysics--The Atmospheres of the Sun and Stars, Ronald Press Company, New York (1963), p. 482.

34. Reference 4, p. 145.

35. Reference 15, p. 172.

36. Reference 32, p. 382.

37. Reference 5, p. 169.

38. R. W. Noyes, G. L. Withbroe and R. P. Kirshner, Solar Phys. 11, 388 (1970).

39. D. E. Osterbrock, Astrophys. J. 134, 347 (1961).

40. R. G. Athay, Astrophys. J. 146, 223 (1966).

41. S. R. Pottasch, Space Sci. Rev. 3, 816 (1964).

42. D. P. Cox and W. H. Tucker, Astrophys. J. 157, 1157 (1969).

43. L. Goldberg, Ann. Rev. Astron. and Astrophys. 5, 279 (1967).

44. H. E. Hinteregger, L. A. Hall, and G. Schmidtke, Space Research 5, 1175 (1965).

45. C. W. Allen, Space Sci. Rev. 4, 91 (1965).

46. N. F. Ness, Ann. Rev. Astron. and Astrophys. 6, 79 (1968).

47. G. M. Nikolsky, Solar Phys. 6, 399 (1969).

48. B. E. J. Pagel, Ann. Rev. Astron. and Astrophys. 2, 267 (1964).

49. R. Cayrel, Compt. Rend. 257, 3309 (1963).

50. A. Skumanich, Astrophys. J. 159, 1077 (1970).

50.1 H. W. Liepmann and A. Roshko, Elements of Gasdynamics, John Wiley and Sons, Inc., New York (1957), p. 189.

51. M. Kuperus, Rech. Astron. Obs. Utrecht 17, 1 (1965).

52. D. W. Moore and E. A. Spiegel, Astrophys. J. 139, 48 (1964).

53. R. B. Lindsay, Mechanical Radiation, McGraw-Hill, New York (1960), p. 44.

54. M. J. Lighthill, Aerodynamic Phenomena in Stellar Atmospheres, ed. R. N. Thomas, Academic Press, New York (1967), p. 429.

55. E. Schatzman and P. Souffrin, Ann. Rev. Astron. and Astrophys. 5, 67 (1967).

56. Reference 53, p. 226.

57. Reference 53, p. 77.

58. Y. Uchida, Astrophys. J. 147, 181 (1967).

59. P. A. Sturrock, Nature 203, 285 (1964).

60. N. D'Angelo, Solar Phys. 7, 321 (1969).

61. R. A. Kopp, Air Force Cambridge Research Laboratories Scientific Report No. 4, AFCRL-68-0312 (1968).

62. M. Kuperus, Space Sci. Rev. 9, 713 (1969).

63. L. Spitzer, Jr., Physics of Fully Ionized Gases, Interscience Publishers, New York (1962), p. 131.

64. M. Kuperus and R. G. Athay, Solar Phys. 1, 361 (1967).

65. R. S. Devoto, J. Plasma Phys. 2, 617 (1968).

66. A. Delcroix and A. Lemaire, Astrophys. J. 156, 787 (1969).

67. R. G. Athay, Astrophys. J. 145, 784 (1966).

68. A. K. Dupree and L. Goldberg, Solar Phys. 1, 229 (1967).

69. L. Goldberg, E. Muller and L. Aller, Astrophys. J. Suppl. 5, 1 (1960).

70. R. A. Kopp and M. Kuperus, Solar Phys. 4, 212 (1968).

71. W. G. Vincenti and C. H. Kruger, Introduction to Physical Gas Dynamics, John Wiley and Sons, Inc., New York (1965), p. 18.