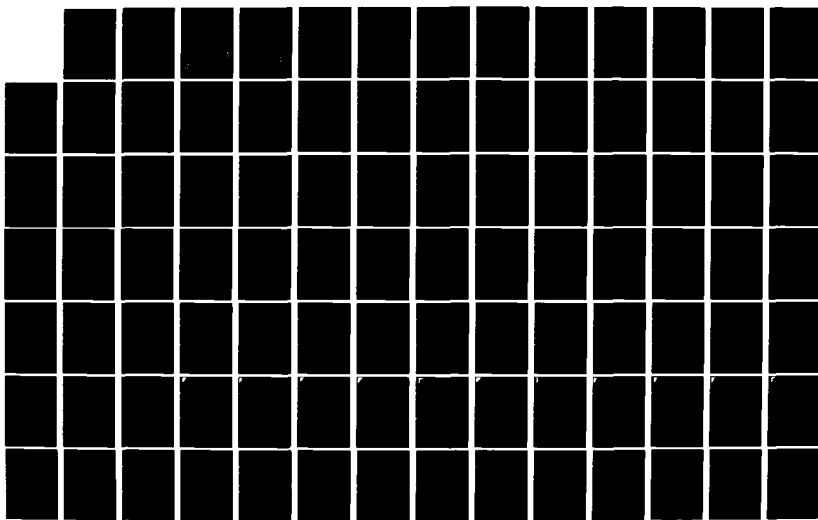


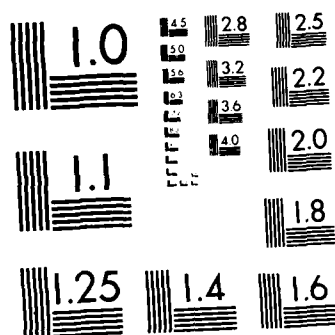
AD-A126 341 GEOPHYSICAL PLASMAS AND ATMOSPHERIC MODELING(U) SCIENCE 1/7
APPLICATIONS INC MCLEAN VA E HYMAN ET AL. FEB 83
SRI-83-144-WA N00014-81-C-2038

UNCLASSIFIED

F/G 4/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

AD A 126341

GEOPHYSICAL PLASMAS
AND ATMOSPHERIC MODELING
FINAL REPORT
SAI-83-144-WA



DTIC
ELECTE
APR 1 1984

SA

SCIENCE APPLICATIONS, INC.

UNCLASSIFIED
DATE 10-10-2001 BY 1045

DTIC FILE COPY

83 03 18 030

GEOPHYSICAL PLASMAS
AND ATMOSPHERIC MODELING
FINAL REPORT
SAI-83-144-WA

DTIC
ELECTE
APR 1 1983



ATLANTA • ANN ARBOR • BOSTON • CHICAGO • CLEVELAND • DENVER • HUNTSVILLE • LA JOLLA
LITTLE ROCK • LOS ANGELES • SAN FRANCISCO • SANTA BARBARA • TUCSON • WASHINGTON

GEOPHYSICAL PLASMAS
AND ATMOSPHERIC MODELING

FINAL REPORT
SAI-83-144-WA

February 1983

Submitted to:
Dr. Sidney Ossakow
Geophysical and Plasma Dynamics Branch
Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

Prepared by:
Ellis Hyman with
Pradeep Chaturvedi, James Chen, Jim Drake, Gurudas Ganguli,
Parvez Guzdar, Y. C. Lee, Richard Lindzen, John Lyon,
Horace Mitchell, Dennis Papadopoulos, Rick Rood,
Harvey Rowland, Punyamurthula Satyanarayana,
George Schmidt, and Robert Smith

Prepared Under:
Contract No. N00014-81-C-2038

SCIENCE APPLICATIONS, INC.

1710 Goodridge Drive
P.O. Box 1303
McLean, Virginia 22102
(703) 734-5840

DTIC
ELECTE
S APR 1 1983

SAI

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution is unlimited

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
I TECHNICAL DISCUSSION	I-1
II REFERENCES	II-1
Appendix A - GLOBAL MHD SIMULATION OF THE MAGNETOSPHERE	A-1
Appendix B - GLOBAL SIMULATION AS A TOOL FOR MAGNETO- SPHERIC RESEARCH	B-1
Appendix C - SATURATION OF THE LOWER HYBRID-DRIFT INSTABILITY BY MODE COUPLING	C-1
Appendix D - ON THE "STABILIZATION" OF THE LOWER HYBRID- DRIFT INSTABILITY IN FINITE β PLASMAS	D-1
Appendix E - FINITE WIDTH CURRENTS, MAGNETIC SHEAR AND THE CURRENT DRIVEN CYCLOTRON INSTABILITY ..	E-1
Appendix F - EFFECT OF ELECTRON BEAMS ON THE CURRENT CONVECTIVE INSTABILITY	F-1
Appendix G - PARAMETRIC EXCITATION AND SUPPRESSION OF CONVECTIVE PLASMA INSTABILITIES IN THE HIGH LATITUDE F-REGION IONOSPHERE	G-1
Appendix H - A DYNAMIC MODEL FOR THE AURORAL FIELD LINE PLASMA IN THE PRESENCE OF FIELD-ALIGNED CURRENT	H-1
Appendix I - LINEAR THEORY OF THE EXB INSTABILITY WITH AN INHOMOGENEOUS ELECTRIC FIELD	I-1
Appendix J - THE MORPHOLOGY OF A MULTI-BUBBLE SYSTEM IN THE IONOSPHERE	J-1
Appendix K - A REVIEW OF DOUBLE LAYER SIMULATIONS	K-1
Appendix L - STRUCTURING STUDIES FOR A URANIUM CLOUD ...	L-1
Appendix M - SIMULATION OF EARLY-TIME URANIUM CLOUD DEVELOPMENT	M-1



Availability Codes	
Avail and/or	Special
Dist	H

Letter 10/1/64

TABLE OF CONTENTS (cont.)

Appendix N - RENORMALIZATION GROUP THEORY OF FRACTALS IN STRIATING IONOSPHERIC PLASMA CLOUDS	N-1
Appendix O - ELF CURRENT GENERATION IN THE IONOSPHERE ..	O-1
Appendix P - MECHANISMS OF STRATOSPHERIC OZONE TRANSPORT	P-1
Appendix Q - A MECHANISTIC MODEL OF EULERIAN LAGRANGIAN- MEAN AND LAGRANGIAN OZONE TRANSPORT BY STEADY PLANETARY WAVES	Q-1
Appendix R - OZONE TRANSPORT BY DIABATIC AND PLANETARY WAVE CIRCULATION ON A β -PLANE	R-1
Appendix S - ANOTHER NOTE ON FINITE DIFFERENCING OF THE ADVECTION DIFFUSION EQUATION	S-1
Appendix T - TURBULENCE ORIGINATING FROM CONVECTIVELY STABLE INTERNAL WAVES	T-1

Section I

TECHNICAL DISCUSSION

The work performed by Science Applications, Inc. (SAI) on this contract, "Geophysical Plasmas and Atmospheric Modeling," Contract Number N00014-81-C-2038, SAI Project Number 1-157-13-705-00, encompasses a wide range of topics in space plasma physics and atmospheric modeling in support of the programs of the Geophysical and Plasma Dynamics Branch of the Naval Research Laboratory (NRL). This report covers the period 1 December 1981 to 17 December 1982. In this section we will summarize the various subjects studied and the results obtained. Details will be included in Appendices which will document reports and publications resulting from our work.

In the following subsections we will describe the major accomplishments in each of the following research efforts: (A) magnetospheric computer simulations, (B) non-linear saturation and transport via the lower hybrid drift instability, (C) the current driven ion cyclotron instability, (D) tearing-mode stability properties of an anisotropic plasma layer, (E) analyses of high latitude irregularity phenomena, (F) auroral plasma transport, (G) the EXB instability, (H) the morphology of a multi-bubble system in the ionosphere, (I) double layer simulations, (J) anomalous resistivity of auroral electrons, (K) energy exchange in the Io plasma torus, (L) uranium cloud modeling, (M) NRL Laser/HANE experiment support, (N) fractal theory applied to striating ionospheric plasma clouds, (O) elf/vlf wave generation, (P) ozone transport in the lower atmosphere and, (Q) stratospheric dynamics.

(A) Magnetospheric Computer Simulations

During this contract year significant results have been obtained both with our 2D and our 3D computer codes. The two codes complement one another, the 2D code being advantageous where higher resolution is required and the 3D code where phenomena that rely on the three dimensional nature of the solar-magnetosphere system cannot be neglected. The 2D code was used to investigate the role of different types of resistivity in controlling reconnection in the magnetotail. Two different types of resistivity, $\eta = \text{constant}$ and $\eta \propto J^2$ were used. In addition, runs were made both with and without Joule heating included for both of these resistivity types. A number of interesting new results were found:

- 1) The form of the resistivity is of secondary importance. What is important, however, is the presence or absence of Joule heating. The absence of Joule heating corresponds to the case in which thermal conduction is so rapid that all the resistive heat is immediately carried away from the x-point. Unless there is a large anomalous thermal resistivity along field lines, the lack of Joule heating corresponds to the actual situation.
- 2) For runs both with and without Joule heating, the x-point tends to move about in the region, $-10R_E > x > -25R_E$, with a typical time scale ~ 20 minutes. This time is typical of either resistive or convective processes, as it is roughly the period required for all the flux in

the tail lobes to be processed through the x-point.

- 3) In the cases with no Joule heating, there were events where the x-point moved tailward to distances $< -60 R_E$ and later a new x-point formed in the near-earth region. These events are similar to the relaxation oscillations observed in our earlier calculations. When these events occur, the behavior of the plasma sheet is similar to that during substorm recovery; the density and pressure gradually build up from near zero to typical values just as the new x-point forms.

The importance of this new work is that it shows the importance of microphysics in determining the macroscopic structure and behavior of the magnetosphere.

This work was presented as an invited paper at The First International School for Space Simulations, Kyoto, Japan in November 1982 and at the Fall 1982 AGU Meeting. The paper presented at the former meeting is reproduced in this document as Appendix A, entitled "Global MHD Simulation of the Magnetosphere". An abstract of the paper presented at the AGU meeting follows:

Reconnection in the Earth's Magnetotail: Resistivity and Thermal Transport Effects*. J. G. LYON, Science Applications, Inc., McLean, VA., 22102, J. D. HUBA and J. A. FEDDER, Naval Research Laboratory, Washington, D.C. 20375--We present results from 2-D global MHD simulations on the earth's magnetosphere which use different dissipation mechanisms. Use of the N-S symmetry plane has allowed us to eliminate numerical

reconnection in the tail. Two different types of resistivity (η) are considered: $\eta = \text{constant}$ and $\eta = \eta_0 J^2$. The effects of thermal transport are also studied by allowing a variable thermal conductivity parallel to field lines. As a limiting case, simulations with no Joule heating were performed. We find that with regard to the initial x-point formation and production of tailward flows the different simulations are similar to one another and to our previously reported results¹. However, the presence of relaxation oscillations in the recombination process (plasmoid formation) depends upon the dissipation of the thermal energy of Joule heating. If the dissipation is sufficiently rapid, relaxation oscillations occur. Otherwise, only changes in the interplanetary magnetic field (shifts from S to N) produce plasmoid formation.

* This work supported by ONR and NASA

1. Lyon et. al., Phys. Rev. Lett., 46, 1038 (1981)

The 3-D simulations during the current contract year have been studied to determine the structure of the electric and magnetic fields and current density during reconnection events. These studies have provided significant new insights into magnetospheric processes during substorms. Among the major insights are

- 1) The electric field from the solar wind does not penetrate to the tail current sheet until strong reconnection takes place. Then the electric field does not cross the entire tail but is localized to a relatively narrow region near the midnight meridian.

- 2) When reconnection takes place the electric field vectors show an arched pattern rather than the purely dawn to dusk of the solar wind field. This should be interpreted as a cylindrical collapse of the magnetotail about the plasma sheet.
- 3) The current shows an intensification and a non cross-tail component during reconnection. The current curves tailward on the dawn side and earthward on the dusk side. This current pattern is associated with strong y components in the tail magnetic field.

This work was presented as invited papers at the Yosemite '82 conference, "Origins of Plasmas and Electric Fields in the Magnetosphere", the Spring 1982 AGU meeting, and the First International School for Space Simulations. The first is reproduced here in Appendix B entitled "Global Simulation as a Tool for Magnetospheric Research". The abstract of the paper presented at the AGU Meeting follows:

Numerical Simulation of Magnetosphere Processes*. J. G. LYON, S. H. BRECHT*, Science Applications, Inc., McLean, VA 22102, J. A. FEDDER, and K. HAIN, Naval Research Laboratory, Washington, D.C. 20375--To illustrate the effectiveness of numerical simulation as a tool for studying magnetospheric processes, we discuss some results of the global MHD simulation effort at NRL. In particular, we look in detail at a 3-D simulation with a southward IMF. This calculation started with a dipole field loaded with plasma. A solar wind was introduced which gradually modified the earth's dipole field to a

magnetospheric configuration. The presence of the southward field in the solar wind produced an open magnetosphere through reconnection on the nose. This open magnetosphere reconnected in the near (15-30 R_e) magnetotail in a series of separate events. Only in the last event seen, had the density in the tail dropped to a realistic level from the initial value and high speed flows accompanied the reconnection. The behavior of this event was quite similar to the neutral line model for substorms, with one exception. An apparent instability caused the reconnection point on the nose to shift duskward. This caused the magnetotail and the reconnection point to shift downward. The reconnection then gave rise to strong fields and velocities in the y direction. Analysis of the electric field in the magnetotail in this event indicate a cylindrical collapse rather than a N/S compression.

* This work supported by ONR and NASA.

** Now at Berkeley Research Associates, Berkeley, CA.

(B) Nonlinear Saturation and Transport by the Lower Hybrid Drift Instability

The lower hybrid drift instability is driven unstable by the relative drift of electrons and ions in an inhomogeneous magnetized plasma. Nonlinearly, the instability causes transport of plasma across the magnetic field, which can be treated as an effective "anomalous" resistivity. This instability is therefore important since it can strongly

influence the microscopic evolution of inhomogeneous plasmas such as in the magnetotail where the "anomalous" resistivity enables the magnetic field lines to reconnect. The resultant dissipation of magnetic energy leads to strong plasma heating.

Our primary emphasis during this contract year has been to understand the nonlinear saturation of the lower hybrid drift instability and the associated plasma transport and anomalous resistivity. A wave equation describing the nonlinear "behavior" of this instability has been derived in two dimensions (perpendicular to the magnetic field $\underline{B}$). This nonlinear equation has been solved by modifying a spectral method code developed by Fyfe, Joyce and Montgomery (1977)(1). The important conclusions of this work are that the instability saturates by transferring energy from growing, long-wavelength modes to damped, short-wavelength modes. The anomalous diffusion coefficient is calculated self-consistently and is found to scale as $(\rho_i/L_n)^2$, where ρ_i is the ion Larmor radius and L_n is the density scale length. This work represents the first completely self-consistent theory of any drift instability. A short paper on this work appears in Appendix C and will appear in the Letters Section of the Physics of Fluids in March (1983). It has also been published as NRL Memo Report 5009, January 1983, entitled "Saturation of the Lower Hybrid-Drift Instability by Mode Coupling". A larger paper is in preparation.

During the course of the work on the nonlinear saturation of the lower-hybrid-drift instability, we also discovered that the effects of finite β on this instability

were not properly understood. It was previously believed that for sufficiently high β , the mode was completely stabilized by ∇B resonant electrons. We have shown, however, both in analytic and numerical computations, that as β increases the fastest growing mode shifts to longer wavelengths and is driven toward marginal stability. This result is important in understanding the nonlinear behavior of the instability since long wavelength modes are always marginally stable rather than damped. This work appears as a report in Appendix D and has been submitted to the Physics of Fluids, under the title "On the 'Stabilization' of the Lower-Hybrid-Drift Instability in Finite β Plasmas".

(C) The Current Driven Ion Cyclotron Instability

The current driven ion cyclotron instability (CDICI) plays an important role in both space and laboratory plasmas. This instability has been widely invoked by space plasma physicists to explain a number of observations. Kindel and Kennel [J. Geophys. Res., 76, 3055 (1971)] were first to study analytically, the electrostatic CDICI in the context of space plasmas. This was subsequently followed by a number of papers both analytical and experimental.

The Kindel and Kennel model, however, is, at best, an idealization of magnetospheric conditions. Their model is based upon the assumption that the background magnetic field is uniform, which is tantamount to ignoring the self-consistent magnetic field generated by the field aligned current. It is important to note that the inclusion of the self-consistent magnetic field (however small) will render the net magnetic field non-uniform due to the introduction of a

magnetic shear in a direction perpendicular to the external magnetic field, $B_0 \hat{z}$. The net magnetic field now becomes a function of position i.e., $B_0 \hat{z} \rightarrow B_0(x) \hat{z}$. Thus the plane wave boundary condition is no longer valid. Now one has to admit a wave packet in the x-direction and impose an outgoing energy flux boundary condition in the x-direction. This is called the nonlocal treatment and was provided by Ganguli and Bakshi [Ganguli and Bakshi, Phys. Flds. 10, 1830 (1982)].

In the nonlocal treatment the drift velocity is considered uniform over all space. This assumption, too, is an idealization since the current observed in space (as well as in the laboratory) always has a spatial profile, i.e., has a finite extent. We have introduced a finite current width (i.e., $V_d \rightarrow V_d(x)$) in the Ganguli and Bakshi model to make the theory more inclusive and pertinent to the actual physical situation [Bakshi, Ganguli and Palmadesso: Submitted to Phys. Flds. Lett. and as an NRL Memo Report]. This work is included in this report as Appendix E. The title is "Finite Width, Currents Magnetic Shear and the Current Driven Ion Cyclotron Instability.

With the introduction of the finite current width, L_c , many interesting features are revealed. Now, there are two length scales i) magnetic shear length L_s , and (ii) the finite current width L_c ; and depending on the ratio L_c/L_s the nonlocal effects due to either finite current width ($L_c/L_s < 0.1$) or due to the magnetic shear ($L_c/L_s > 0.1$) dominate. The transition between the two nonlocal effects and the influence of L_c and L_s on the growth rates (and, hence, the macro parameters) have been studied in detail. Applications of this formalism to both space and laboratory plasmas are now being investigated.

(D) Tearing-Mode Stability Properties of an Anisotropic Plasma Layer

We have studied the tearing-mode stability properties of a field-reversed collisionless plasma layer in which the temperature perpendicular ($T_{\perp}$) and parallel ($T_{\parallel}$) to the equilibrium magnetic field are unequal. The electrons and ions are both described by Vlasov distribution functions of the Maxwellian type. Using simplified particle orbits, we have obtained the dispersion relations for various values of $(T_{\perp}/T_{\parallel})_{\text{ion}}$ with isotropic electrons. It is found that temperature anisotropy has dramatic effects on the growth rate and the stability boundary. In particular, the case with $T_{\perp}/T_{\parallel} > 1$ is more unstable while the case with $T_{\perp}/T_{\parallel} < 1$ is more stable compared with the isotropic case. Moreover, the maximum growth rate increases by one order of magnitude as $T_{\perp}/T_{\parallel}$ is increased by only $\approx 30\%$.

Another new result is that the conventional technique of matching the exterior solution and the interior solution at the distance $d_e \equiv \sqrt{2} a_e \delta$ where a_e = electron Larmor radius and δ = plasma layer thickness has been found to be inadequate, requiring significant modifications on this account alone. The physical reason is simply that the distance d_e is much less than the typical ion Larmor radius so that the forces responsible for tearing that ions experience beyond the distance d_e is neglected in the conventional treatment. We have introduced a new intermediate region,

$d_e < r < d_i/2$, where $d_i = \sqrt{2} a_i \delta$ with a_i = ion Larmor radius. This is the region where the large ion orbits are dominant. It has been found that inclusion of this region is critically important for determining the stability properties of the anisotropic layer.

This work is essentially complete and has already been presented at the American Physical Society, Plasma Physics Division meeting (November, 1982, New Orleans) and at the American Geophysical Union meeting (December, 1982, San Francisco). A more lengthy journal article manuscript is in preparation. The abstract for the APS presentation follows.

IP 4 Tearing-Mode Stability Properties of an Anisotropic Plasma Layer*. J. CHEN, Science Applications, Inc., and P. J. PALMADESSO, Naval Research Laboratory--in an earlier paper [1], temperature anisotropy ($T_{\parallel} \neq T$) was shown to significantly affect the marginal state of collisionless tearing modes. In particular it was shown that temperature anisotropy with $T_{\parallel} < T_{\perp}$ is less stable. In the present paper, we calculate the growth rates of these modes for an anisotropic ($T_{\parallel} \neq T_{\perp}$) collisionless slab of plasma, using kinetic equations for both ions and electrons. No external guide magnetic field is assumed, and the Harris-type equilibrium is used. The results may be relevant to the earth's magnetotail.

* Work supported by Office of Naval Research and National Aeronautics and Space Administration.

1. Chen, J. and Davidson, R. C., Phys. Fluids, 24, 2208 (1981).

(E) Analyses of High Latitude Irregularity Phenomena

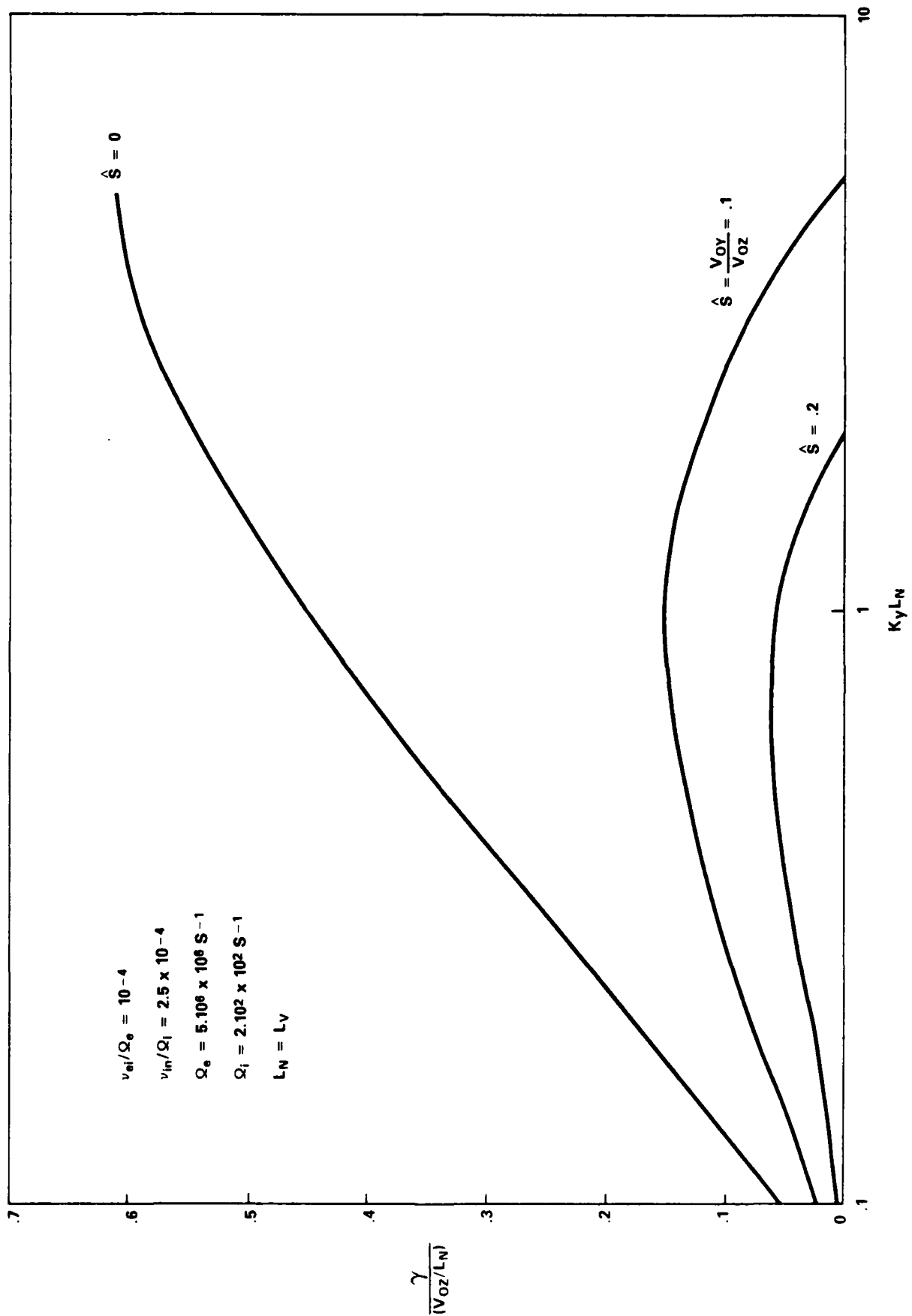
The auroral ionosphere is a region where various sources of free energy co-exist and generate several plasma instabilities. Recently, scintillations observed by the DNA

wideband satellite have been conjectured to be due to density irregularities generated by the current convective instability. The observations seem to suggest the scintillations as being caused by L-Shell field aligned large scale irregularities colocated in the region of soft particle precipitation and horizontal plasma density gradients with the source region confined in latitude. The current convective instability occurs in regions where the currents exceed a certain threshold in the presence of plasma density gradients transverse to the magnetic field. Strong convection electric fields do exist in the auroral ionosphere; these electric fields are strongly inhomogeneous and cause the equilibrium plasma flow velocity to be strongly sheared. During the contract year, we have included the inhomogeneous electric fields along the plasma density gradients and performed a non-local analysis on the current convective instability in the presence of velocity shear. Furthermore, we have also included the self consistent magnetic shear generated by the currents generating the instability itself.

Preliminary results, shown in the figure, indicate that velocity shear suppresses the short wavelength modes beyond a certain kL_n . For stronger shears, we see that the cut-off wavenumber decreases and the mode that is preferentially excited has a longer wavelength. The implications of these results are under investigation. These results were presented at the December 1982 American Geophysical Union meeting in San Francisco. The abstract follows.

Effects of Velocity Shear on the
Current Convective Instability*. P.
 SATYANARAYANA, Science Applications,
 Inc., McLean, VA 22102, and S. L.
 OSSAKOW, Naval Research Laboratory,
 Washington, D.C. 20375--Plasma density
 irregularities with scale sizes around

$\gamma/(V_{O2}/L_N) \text{ Vs } K_Y L_N$



$\frac{\gamma}{(V_{O2}/L_N)}$

1 km, with complicated geometry, have been observed in the auroral ionosphere. The current convective instability (CCI) feeding on the permanent field aligned currents and plasma density gradients transverse to the magnetic field, has been invoked¹ to explain some of the features of these plasma density irregularities. Also considerable data is available² that indicate the existence of strong velocity shears in the auroral region. We incorporate this important source of free energy in addition to the field aligned currents and transverse density gradients. We derive a mode structure equation and solve the differential equation in various domains. We discuss the results and apply them to the auroral ionosphere.

* This work supported by DNA and ONR.

1. S. L. Ossakow and P. K. Chaturvedi, Geophys. Res. Lett., 6, 332, 1979.
2. P. M. Kintner, J. Geophys. Res., 81, 5114, 1976.

We have investigated the effects of an electron beam on the current convective instability (CCI) which is believed to play an important role in generating irregularities at high latitudes. These irregularities cause scintillations of satellite signals such as those from the DNA wideband satellite. The analysis was prompted by the fact that soft electron fluxes are nearly always present at such latitudes. We find that the presence of such electron beams has negligible effect on slow processes such as the current convective instability in diffuse aurora. A more complete analysis is contained in Appendix F. A manuscript is in preparation.

We have also looked into the possibility of parametric excitation and suppression of irregularities at high latitudes by means of a powerful high frequency electromagnetic wave transmitted from the ground. This presents one with the attractive possibility of carrying out controlled experiments in the ionosphere so as to better understand and identify the various physical processes occurring there. In this work, we present threshold estimates for the electromagnetic beam power needed to either trigger the irregularities under subthreshold conditions, or to suppress them when they have been excited naturally just above the marginal stability level. For typical ionospheric parameters, powers below $\sim$ a few mw m^{-2} at the local upper hybrid frequency would be required to excite or suppress the CCI or the ExB instabilities in the ionosphere. In Appendix G, this is described in more detail in a paper entitled "Parametric Excitation and Suppression of Convective Plasma Instabilities in the High Latitude F-Region Ionosphere."

(F) Auroral Plasma Transport

We have developed a one-dimensional multi-moment multi-fluid model of the plasma along an auroral field line in order to study the dynamic behavior of the auroral plasma. This model was used to investigate the plasma dynamics and electrodynamics both with and without anomalous resistivity and to study the effect of multiple ion species in this region.

Initial simulations were performed in the presence of a current parallel to the auroral field line. These simulations indicate that the multiple time scales that exist for interactions between the different plasma species allow large

amplitude transients in density and temperature to exist on time scales comparable to that for variation of the field-aligned current. This suggests that the field line plasma may only rarely be close to a steady state. This work has been published as NRL Memorandum Report 4840, 1982, under the title, "A Dynamic Model for the Auroral Field Line Plasma in the Presence of Field-Aligned Current". It is included in this report as Appendix H, and has been accepted in a revised and extended form for publication in the Journal of Geophysical Research. This work has also been presented at the Theory Institute in Solar-Terrestrial Physics at Boston College (August, 1982), an abstract of which follows.

Electrodynamics and Transport Processes on Auroral Field Lines*. P. PALMEDESSO, Naval Research Laboratory, Washington, D.C. 20375, and H. MITCHELL, Science Applications, Inc.--A one dimensional, dynamic multimoment, multi-fluid model of an auroral field line has been developed in order to study transport and electrodynamic processes involved in ionosphere-magnetosphere coupling. Plasma turbulence effects are included by way of anomalous transport coefficients. In particular, the model is intended to test the dynamic consequences of a number of mechanisms proposed to explain the generation of magnetic field aligned potential drops on auroral field lines. Early results indicate the presence of large amplitude, long lived transients following changes in the magnitude of field aligned currents. The transients decay on a time scale longer than the time scale on which auroral currents typically vary, suggesting that the flux tube is rarely, if ever, close to a steady state. Some specific electrodynamic processes will be discussed in this context.

* This work supported by National Aeronautics and Space Administration and Office of Naval Research.

Simulations have also been performed which include simple models of anomalous resistivity in the auroral plasma. Parameters in these models were varied in order to judge the effect that heating due to such resistivity has on the stability of the plasma distributions and the associated electric fields parallel to the field line. The results of this investigation were presented at the Spring, 1982 meeting of the American Geophysical Union, and the abstract of this talk follows.

A Simulation of Plasma Behavior along an Auroral Field Line in the Presence of Resistivity to Field-Aligned Current*.
H. G. MITCHELL, JR. Science Applications, Inc., McLean, VA 22102, P. PALMADESSO, Naval Research Laboratory, Code 4780, Washington, D.C. 20375--The existence of large and relatively long-lived electric fields parallel to the magnetic field lines above the aurora imply the existence of a form of resistivity along the field in the plasma in this region. We have investigated the response of the auroral plasma to simple models of such resistivity by using a dynamic simulation model of the field line plasma which we have developed. This model is a one-dimensional simulation of a three species plasma (e^- , H^+ , O^+) along about $4 R_E$ of field line above an altitude of 1000 km. The simulation dynamically models the transport of the moments of the species' distribution functions up through heat flow. This allows a reasonable treatment of the energy deposition and transport due to regions of resistivity in the collisionless plasma. Such energy transport effects have a significant effect on the response of the plasma to a field-aligned current and, therefore, on the stability of electric fields parallel to the field line.

* This work supported by ONR and NASA.

As a direct result of our simulations, we have discovered a mechanism for heavy ionospheric ion energization. In a multi-ion plasma such as that present on auroral field lines, low frequency electric fields resulting from resistive friction between electrons and one ion species may freely accelerate any ion species which does not participate in the resistivity. For example, when oxygen ions are a minor species, strong electron-hydrogen ion friction may accelerate the oxygen ions and provide them with the high energies they are observed to possess in the magnetosphere. This work was presented at the Fall, 1982 meeting of the American Geophysical Union, with the following abstract.

Oxygen Ion Acceleration Due to Resistive Momentum Transfer in the Auroral Field Line Plasma*. H. G. MITCHELL, JR., Science Applications, Inc., McLean, VA 22102, and P. PALMADESSO, Naval Research Laboratory, Washington, D.C. 20375--Low frequency electric fields maintained in a plasma by resistivity (either collisional or wave-induced) imply a balance between the momentum transferred to the electrons by the electric field and that transferred by the resistivity from the electrons to other particle species. As a result, the electrons do not experience a net acceleration due to the electric field. If the plasma contains only one ion species, the momentum balance is the same for the ions, and the ion species is not accelerated either. If a second ion species is present which does not participate in the resistive momentum transfer, it may be accelerated by the electric field while the participating ions experience an acceleration opposite to the field. This mechanism may be applied to the auroral field line plasma

in the presence of a current along the magnetic field. Significant resistivity between the electrons and hydrogen ions may accelerate the oxygen ions provided that the resistivity has the appropriate characteristics. We have used an existing auroral field line plasma simulation to dynamically model this effect. This simulation is a one-dimensional multi-moment dynamic model of a three-species plasma (e^- , H^+ , O^+) along about $4 R_E$ of field line above an altitude of 1000 km.

* This work supported by ONR and NASA.

(G) THE ExB INSTABILITY

It has been conjectured that the ExB instability, also known as the gradient drift instability, plays an important role in generating density irregularities in the high latitude F region ionosphere⁽²⁾ and that it is responsible for the early time structuring of barium plasma clouds released into the ionosphere⁽³⁾. This instability is also an interchange instability which can occur in an inhomogeneous magnetized plasma that contains an electric field orthogonal to both the magnetic field and the density gradient.

This collaborative work deals with the generalized theory of the ExB instability with the ambient electric field at an arbitrary angle to the density gradient and incorporates the self consistent inhomogeneous electric field along the density gradient. This work, in fact, generalizes the theory of Perkins and Doles⁽⁴⁾, who have treated a similar equilibrium configuration, for arbitrary collisional and wavelength domains.

Again, we treat the problem non-locally using a differential equation rather than an algebraic equation for the perturbed electrostatic potential. We solve a reduced second order differential equation numerically using a shooting code for eigenvalues and eigenfunctions. The principal results are the following:

- (i) The basic conclusions of Perkins and Doles have been verified. Specifically, the marginal stability criterion they derive agrees well with our numerical results. We also find that the short wavelength modes ($k_y L > 1$) are strongly stabilized by velocity shear.
- (ii) The stabilization mechanism is associated with the x-dependent Doppler shifted frequency, $\omega - k_y V_y(x)$, where $V_y(x) = -cE_x/B$ and not with velocity shear terms proportional to $\partial V_y / \partial x$, $\partial^2 V_y / \partial x^2$.
- (iii) When $E_x(x_0)/E_y > 1$, where x_0 is the position about which the mode is localized, the most unstable modes have $k_y L < 1$.

This work has been accepted for publication in the Journal of Geophysical Research. It has also been published as NRL Memorandum Report 4901 (1982), entitled "Linear Theory of the ExB Instability with an Inhomogeneous Electric Field", and appears in this report as Appendix I.

(H) The Morphology of a Multi-Bubble System in the Ionosphere

The behavior of ionospheric plasma has been extensively studied with the aim of understanding various ionospheric phenomena including equatorial spread F (ESF) and plasma cloud striations. In this regard, the previous theoretical and numerical works have investigated the morphology and evaluation of single-bubble systems or uniform arrays of bubbles, neglecting the intrinsic influence of neighboring bubbles.

We have recently completed an analytical study of the morphology of a multi-bubble system by solving the Pedersen current conservation equation. We have obtained exact solutions without any a priori assumption regarding the separation distances between the bubbles. The qualitative results show significant modification to the single-bubble results in cases where the bubble separation distances are moderately short. A paper describing this work has been submitted to the Journal of Geophysical Research and is included here as Appendix J. An abstract of the paper presented at the American Geophysical Union Meeting follows.

The Morphology of a Multi-Bubble System in the Ionosphere*. S. L. OSSAKOW, Naval Research Laboratory, Washington, D.C. 20375, J. CHEN and P. SATYANARAYANA, Science Applications, Inc., McLean, VA 22102--A multi-bubble model is developed to study the morphology of plasma density depletions in the context of equatorial spread F (ESF) irregularities. The Pedersen current conservation equation with quasi-neutrality is solved analytically using an electrostatic analogy. The solution is exact. It is shown that the influence

of neighboring bubbles is moderately short-ranged and that a small number of bubbles can adequately model a large array of plasma depletions. The $E \times B$ rise velocity is strongly affected and the contours are deformed, significantly deviating from the single-bubble models previously developed. The implications on the stability and dynamic behavior of the bubbles are discussed.

* This work supported by the Defense Nuclear Agency.

(I) Double-Layer Simulations

Continuing work begun under previous contracts, we have developed further results from a continuing series of 1-D Vlasov simulations of double-layer (DL) formations. A unique feature of these simulations is that the DL is part of a self-consistent circuit, in which the circuit equation for the current provides the boundary condition for the DL evolution and is dependent on the evolution itself. Recent progress includes the following:

- (A) A post-processing program has been perfected which provides time histories and or spatial profiles (at selected times) of all field and kinetic quantities which can be constructed from the distribution function $f_{e,i}(x,v,t)$. These quantities include (among others):

- (1) Spatial profiles of
 - (i) species densities;
 - (ii) charge density, electric field, and potential;
 - (iii) current density;
 - (iv) species densities of kinetic energy, kinetic pressure, kinetic temperature, kinetic heat flux, and coarse-grained specific entropy;
 - (v) Joule heating and electrostatic energy density;
- (2) Time histories of system (integrated) quantities:
 - (i) potential;
 - (ii) electrostatic and kinetic energy;
 - (iii) species content;
 - (iv) boundary flux;
 - (v) Joule heating;
- (3) Energy flow diagnostics in the external circuit:
 - (i) battery power;
 - (ii) resistive loss;
 - (iii) inductor discharge.

The availability of these energy diagnostics is due to the use of the circuit model, which conserves energy globally and permits a detailed temporal study of the energy flows.

(B) A boundary condition for a circuit model under "magnetospheric" conditions has been analyzed but not yet implemented. The distinction between the circuit used heretofore and a magnetospheric circuit, besides the basic description of the circuit itself (which is currently under investigation) is that in the former use the distributions of the external source plasmas--necessary to maintain current continuity at the boundaries--may be assumed to be unaffected by the DL itself, while in space these distributions must be modelled such that various constraints (e.g., charge neutrality, momentum balance, etc.) are satisfied taking into account the distributions of particles which have passed through the evolving DL. Development of this model is still in progress, and should lead to the first simulation of DL evolution which may be applicable (in a model sense) to the question of DL evolution in space.

The major goal of this work is to provide a description of DL evolution which may address the following questions:

- (a) Can DL truly form in space, notably in the auroral zone and/or in the magnetotail, and under what conditions?;

(b) What morphology (stability, etc.) would DL manifest in space?;

(c) What is the relation of DL to anomalous resistivity?

It is hoped that such a description, which may ultimately take the form of a dynamic circuit model, may be incorporated into studies of magnetosphere-ionosphere coupling.

An article describing the simulation method and preliminary results has appeared this year (R. Smith, *Physica Scripta* 25, 413, (1982)) and a detailed paper describing the results obtained so far is in preparation. In addition, an invited general review of DL simulations was presented at the 1982 International Conference on Plasma Physics, Göteborg, Sweden, and is in press (*Physica Scripta: proceedings*); this article is attached as Appendix K, and is titled "A Review of Double Layer Simulations."

(J) Anomalous Resistivity of Auroral Electrons

Rowland, Palmadesso and Papadopoulos have continued to study the effects of anomalous dc resistivity on the acceleration of auroral electrons. It was shown⁽⁵⁾ that conditions that exist along auroral field lines ($\delta n_i/n_0 > 0.25$, $\omega_e/\omega_{ce} > 1$) act to break the adiabatic invariance of the electron magnetic moment, cause anomalous dc resistivity, and generate electron distributions in good agreement with observations. The large amplitude ion waves which have been observed in the aurora trap the bulk of the electrons and lead to the formation of a runaway distribution of electrons when an electric field is applied. The

acceleration of the runaways is limited by an instability involving pitch angle scattering of the electrons by lower hybrid waves. Recently we have shown⁽⁶⁾ that the runaway electrons can also be pitch angle scattered by the low frequency ion waves providing another method to generate anomalous resistivity. This work is continuing.

(K) Energy Exchange in the Io Plasma Torus.

We have brought to a successful conclusion a study of energy exchange between electrons and ions in the Io plasma torus (IPT). Ions are created in the torus by electron impact ionization of neutral S and O emitted from volcanoes on Io. In the frame of reference co-rotating with Jupiter, the newly created ions have speeds equal to the co-rotation speed, and thus energies of order 300 eV (for O) and 600 eV (for S); in the lab frame the acceleration mechanism is seen as "pickup" by the rotating magnetic field. The electrons gain less than 0.01 eV by this "pickup" mechanism. From particle observations, however, the temperatures of the various ion species (assuming they are Maxwellian) are inferred to be of order 20-80 eV, while the electron temperature is inferred to be 5-8 eV, based on the rate of impact excitation of ions radiating in the EUV. The problem originally posed was that several authors have concluded that collisional excitation of the electrons by ions was too slow to achieve this thermal balance within the inferred lifetime of $\sim 10^6$ s for the torus plasma.

We originally postulated (Smith, Palmadesso, and Strobel, Trans. AGU (EOS), 62, 999 (1981)) that the ion loss-cone instability driven by the nonthermal pickup ions could effectuate the energy transfer via Landau damping to the electrons. This hypothesis was investigated in detail by

solving a steady-state quasi-linear equation for an "average" ion species (average charge and mass) including the following terms:

- (a) a model Fokker-Planck operator for the ions;
- (b) a Krook-model collision operator for the electrons;
- (c) charge-exchange (CHEX) source and loss terms for the ions;
- (d) a pickup source term;
- (e) an energy-independent loss rate to model diffusion out of the torus;
- (f) a quasi-linear velocity diffusion operator based on the ion-loss-cone instability.

It was found that with the Krook model operator for electron-ion collisions, a Maxwellian core formed in the ion distribution function and the pickup source feature was broadened by quasi-mean diffusion, but energy exchange to electrons was still insufficient ($\leq 10\%$) to account for the EUV radiation loss. Moreover, when the same Fokker-Planck model used for the ion-ion collisions was used for electron-ion collisions, thus more correctly accounting for collisional drag, the Post-Rosenbluth instability was suppressed even though the plasma lifetime in the torus is $\leq \nu_{ei}^{-1}$, the electron-ion collision time.

We then considered a multi-species set of quasi-linear equations for S^+ , S^{++} , S^{+++} , O^+ , and

O^{++} , incorporating ten CHEX reactions among the various ion species as well as the above-mentioned Fokker-Planck models for electron-ion collisions. In this formulation the only free parameters are the densities of neutral O and S. It is found that with these features included in the model, collisional excitation is sufficient to account for the EUV radiation, depending on the densities $n(S)$, $n(O)$. The discrepancy with the conclusions of the authors is apparently due to the inclusion here of the many CHEX reactions in the multi-species set of quasi-linear equations, which is a more detailed set than used by other authors: Charge-exchange provides a net energy input to the torus.

The above results are obtained from moments over the set of simultaneous quasi-linear equations. We have also solved for the distribution functions $f(v_{\perp})$ of the various ion species: These can be used to compare the model to the instrumental response of the MIT Plasma Science Instrument.

The above results were reported in preliminary form in an invited talk at the COSPAR symposium in Ottawa in May 1982: a paper (to be submitted to JGR) is in preparation.

(L) Uranium Cloud Modeling

Of the possible sources of long wave infrared radiation (LWIR) from a high altitude nuclear burst that might severely degrade defense optical detectors, the emission from metal oxides is potentially most important, particularly at late times. Of the metal oxides likely to be associated with a nuclear device the uranium oxides appear to be the most important LWIR radiators. In the previous contract year we performed a simulation of the structuring of

a uranium cloud released at 200 km. This work appeared as an NRL Memo Report #4659, (1981) entitled "Preliminary Numerical Simulation of IR Structure Development in a Hypothetical Uranium Release." This year we have expanded these studies along two fronts: we have (1) developed an analytic model of a structured cloud in an attempt to clarify the relationship between cloud structure and power spectral characteristics and (2) developed a code to simulate the early time hydrodynamics, diffusion, and chemistry of a 2 kgm neutral uranium cloud released at 200 km.

The analytical model we have developed incorporates those properties of the cloud that our previous physics simulations have uncovered. Where the structure and the physics is not well understood, as at small scale sizes near the diffusion limit, we have incorporated parameters enabling us to quantify the effect of our lack of understanding of the cloud physics on the spectral properties of the cloud.

In our studies we exhibit simple cloud shapes and show what are the asymptotic spectral dependences that will result from these clouds. In particular, we distinguish between those asymptotic properties of a cloud that are observed in communications applications from what an optical sensor, which records the line integrated emission, will observe. In preliminary results we show that it is possible in some parameter regimes, and looking along specific directions in a cloud, to observe very low asymptotic spectral indices, ~ 2 . This work is described in a paper presented at the DNA Sponsored Uranium (Oxide) LWIR Meeting at Los Alamos, N.M. in January 1983. A copy of the viewgraphs presented at that meeting including explanatory notes is included in this report as Appendix L entitled

"Structuring Studies for a Uranium Cloud." They will be published in the Proceedings of the above meeting.

The other area in which our uranium cloud studies have been directed is in the early time motion of the cloud of uranium vapor and of the derivative oxides, neutral and ionized, which rapidly are produced. These studies use a 1D hydrocode incorporating the chemistry and realistic diffusion coefficients. Two cases were studied: (a) motion along the magnetic field, in which the ions diffuse. Cartesian-like, and (b) motion perpendicular to the magnetic field, in which the ions do not diffuse. Under these assumptions it is found that the conversion of UO^+ to UO_2^+ can be delayed for several tens of seconds. This result depends strongly on the capability of experimentalists to rapidly vaporize large amounts of the uranium, a capability which presently appears far off. This work was also presented at the meeting at Los Alamos, N.M. A copy of the viewgraphs, including explanatory notes, is included in this report as Appendix M, entitled "Simulation of Early-Time Uranium Cloud Development", and will be published in the Proceedings of the above meeting.

(M) NRL Laser/HANE Experiment

Work has begun and is continuing to provide theoretical support for the DNA laser-coupling experiment designed to test the physics of short-scale-length coupling in HANE. To this end, the following items have been accomplished:

- 1) The KYLSMA code has been run for a set of nominal design parameters of the experiment. Preliminary results show no coupling for these

parameters, for reasons which are now understood;

- 2) Based on existing theory (Lampe et al., NRL Memo Report 3076, 1975), we have attempted to define new nominal design parameters in which coupling may be observed, and a sequence of KLYSMA runs designed to investigate these parameter regimes has been outlined and is in progress.

An active dialogue with the experimenters has been established and is continuing.

In a separate study it has been shown that ions reflected from the coupling shell at early times and returning to it after gyrating can provide an additional source of energetic ions responsible for the debris patches. Their energy can be more than 6-8 times the original debris energy.

(N) Fractal Theory Applied to Striating Ionospheric Plasma Clouds

During this contract period we have begun an investigation to attempt to clarify the relevance of the renormalization group theory of fractals to the behavior of convecting plasma clouds in the ionosphere. These clouds, being unstable under the ExB gradient drift phenomenon, striate at late times. There is also a freezing-in process that has been observed. Our goal here has been to adapt to this problem the theory previously developed to understand Navier-Stokes turbulence. This work is still being

developed and involves essentially a non-perturbative scheme to implement the idea of the renormalization group to study the late-time evolution of striating ionospheric plasma clouds. Further details are contained in a preliminary report included here as Appendix N and titled, "Renormalization Group Theory of Fractals in Striating Ionospheric Plasma Clouds."

(O) ELF/VLF Wave Generation

During this contract period several schemes for ELF/VLF generation were investigated. Two of these involve (1) absorption of the momentum of radio waves generated by a land based antenna by ionospheric electrons and (2) enhancement of the magnetic moment of electrons by cyclotron resonance in the polar region, giving rise to currents due to magnetic mirror effects. This study is described in more detail in a report entitled "ELF Current Generation in the Ionosphere" which has been incorporated in this report as Appendix O.

In addition, analytic and numerical studies of alternative mode coupling schemes for VLF excitation were performed. For the lower frequency range a new regime, the semicollisional regime, was discovered, where pump depletion and efficient downconversion occur for pump power thresholds, much below the collisionless regime. In this regime, the important non-linear force, is due to the ohmic non-linearity. Computations indicated that proof of principle experiments can be performed with reasonable power.

(P) Ozone Transport in the Lower Atmosphere

Ozone transport has been investigated using steady state and time dependent planetary wave models with a prescribed diabatic circulation. NRL Memo Report 4969 (Appendix P) details the transport mechanisms that have been investigated in this study. Two papers have been submitted to the Journal of Geophysical Research (JGR) and one paper has been submitted to Monthly Weather Review (MWR). A description of the content of these papers will serve to describe the completed research. A presentation of some of the results was given at the Fifth International Symposium on Solar-Terrestrial Physics in Ottawa, Canada, May 1982. More results will be presented in March 1983 at the AMS Conference on the Meteorology of the Upper Atmosphere in Boston.

Transport by steady, stationary, dissipative planetary waves is discussed in "A Mechanistic Model of Eulerian, Lagrangian-mean, and Lagrangian Ozone Transport by Steady Planetary Waves" (Appendix Q, co-author M. R. Schoeberl, accepted for publication by JGR). Two transport mechanisms are described: one in which the waves "stir" parcels through regions of highly varying photochemistry and thereby enhance ozone transport, and the other where ozone is advected by the Lagrangian-mean velocity in the lower stratosphere.

The mean field chemistry counteracts most of the eddy transport caused by the stirring with changes on the order of one ppm possible at the polar right boundary and at the bottom of the transition layer. The largest changes in the model domain are caused by advection and occur in the region where ozone is a passive tracer. Steady state

planetary waves do not have vigorous enough north-south transport to explain the springtime polar buildup of ozone; therefore, parameterizations based on steady state planetary waves are unable to model the observed ozone transport.

Transport by the diabatic circulation and time dependent planetary waves is discussed in "Ozone Transport by Diabatic and Planetary Wave Circulations on a β -Plane" (Appendix R, co-author M. R. Schoeberl, submitted to JGR, Dec. 1982). Both circulation regimes are instrumental in the accumulation of ozone at high latitudes during spring. The diabatic circulation transports ozone downward from high altitude source regions during winter to increase the column density of ozone. Planetary waves then transport ozone to very high latitudes during stratospheric warming events.

The planetary wave transport is interpreted with the theory of wave-mean flow interaction at a purely absorbing critical line. The critical line is a region of strong wave transience and there is a jet-like flow of ozone to high latitudes during the initial phases of mid winter sudden warmings. Much of the ozone transported to high latitudes is returned to low latitudes as the mean flow resumes its climatological state after the warming. It can be concluded that the spring ozone maximum is strongly related to the final warming, after which the mean flow evolves into the summer climatological state. The transport associated with the warming events is highly organized and is not conducive to diffusion parameterizations.

A third paper entitled "Another Note on Finite Differencing of the Advection-Diffusion Equation" has been submitted to NWR (Appendix S). This paper discusses some of

the numerical aspects of the advection-diffusion equation with regard to previously reported results.

(Q) Stratospheric Dynamics

A study which investigated the constraints imposed by conservation of potential vorticity and hydrodynamic stability on the amplitude of Rossby waves was completed. The work has been published in the Journal of the Atmospheric Sciences 39, (1982) under the title "A Note on the Limits of Rossby Wave Amplitudes."

Additional studies have been performed to estimate stratospheric mechanisms which generate turbulence. Starting with a theory developed in 1981 and published in J. Geophys. Res. 86, 9707, which described the generation of turbulence by unstable tides and gravity waves, we have extended these considerations to include the possibility that turbulence could be generated by gravity waves that are not convectively unstable. This work has been submitted for publication to J. Geophys. Res. and is included in this report as Appendix T. The title is, "Turbulence Originating from Convectively Stable Internal Waves."

Section II
REFERENCES

1. D. Fyfe, G. Joyce, and D. Montgomery, J. Plasma Phys. 17, 317 (1977).
2. Ossakow, S. L., P. K. Chaturvedi, and J. B. Workman, High altitude limit of the gradient drift instability, J. Geophys. Res., 83, 2691, 1978.
3. Linson, L. M. and J. B. Workman, Formation of striations in ionospheric barium clouds, J. Geophys. Res. 75, 3211, 1970.
4. Perkins, F. W., and J. H. Doles III,, Velocity shear and the ExB instability, J. Geophys. Res., 80, 211, 1975.
5. Rowland, H. L., K. Papadopoulos, and P. J. Palmadesso, Anomalous resistivity on auroral field lines, NRL Report 8537, Jan. 1982.
6. Rowland, H. L., Strong turbulence and anomalous resistivity, invited paper, Theory Conference in Solar Terrestrial Physics, Boston College, Boston, Mass. 1982.

APPENDIX A

GLOBAL MHD SIMULATION OF THE MAGNETOSPHERE

Global MHD Simulation of the Magnetosphere

J.G. Lyon - Science Application, Inc.

J.A. Fedder, K. Hain, and J.D. Huba
Naval Research Laboratory

Global MHD simulation of the magnetosphere is an extremely useful tool for magnetospheric research. It is a tool that has only been utilized in the past few years. However the presence of number of active groups - particularly at UCLA [1] and Los Alamos [2] in addition to NRL - ensure its rapid development. More than any other method global MHD simulation enables one to track the flow of energy through the magnetosphere and study the coupling of the magnetosphere to the solar wind. Topological relationships, such as magnetic field lines, are also exceptionally clear. However, MHD is a limited description of the physics of the magnetosphere and cannot model all the phenomena. Within its limits MHD does provide a consistent physical picture which can be used as a tool quite successfully, as we hope to show below. First, however, we would like to discuss some of the numerical techniques that make global modeling a going proposition. Then we will give some examples from the simulation work at NRL to underline their utility. Finally, we will close with some comments on ionospheric-magnetospheric coupling.

It is particularly important to use the best possible numerical techniques for global magnetospheric simulations. The real magnetosphere has a number of relatively narrow regions where the properties of the medium change discontinuously, e.g., the bow shock, and the magnetopause. Use of techniques that fail to resolve these structures or smear them over too large a region can affect even the gross properties of the model. Refining the mesh spacing until any given method will resolve these structures is simply not feasible. The calculation is inherently three-dimensional. Computing time goes up roughly as N where N is the number of cells along one side of the mesh.

Hyperbolic systems of equations — as in the ideal MHD equations — can have discontinuous solutions. Discontinuities present problems to all accurate finite difference methods. Accurate in this context means of second order accuracy or higher. These schemes generally do not contain sufficient dissipation to damp out the numerical noise caused by the combination of the discontinuity and the inaccuracies of the method. The result is that discontinuities calculated by these methods tend to give rise to a set of violent spatial oscillations in the solution which can extend far from the actual discontinuity. First order methods are free from such oscillations, but at a considerable price. The solution is smooth, but discontinuities in the best of these schemes are now smeared over at least 3 computational cells. Both types of methods have severe drawbacks for magnetospheric simulation.

The approach we have used is that of hybrid techniques. A hybrid scheme is one which switches between an accurate solution and a first order solution depending on which algorithm appears to give the best results at a temporal and spatial location. This idea is to use an accurate method where it gives no spurious oscillations and to use the first order method where dissipation is necessary to prevent oscillations. The methods we use at NRL are flux corrected transport, FCT, [3,4] and the partial donor cell method, PDM [5]. Examples of advection problems show that such techniques have resolutions many times greater than first order schemes. On shock problems, where the underlying physical process is self-steepening, the advantage is reduced to about a factor of two. Because of the N^4 dependence mentioned above, however, the factor of 2 translates to a factor of 16 efficiency. Actually, because the algorithm is more complicated the true increase in efficiency is more like a factor of 8.

Another computational problem is related to the fact that various regions of the magnetosphere have very high Alfvén speeds. This puts a very stringent limit on the time step that can be used for explicit schemes. This can be solved in a number of ways. The method we use is simple, works well, has a physical basis, and to our knowledge has not been generally disseminated. The trick is to keep the displacement current in Maxwell's equations. Physically, this has the effect of limiting the maximum Alfvén speed to the speed of light. The speed of light can then be chosen to be much faster than any speed at physical interest but slow enough so that a reasonable time step can be used. The major effect of this "Alfvén correction" is to add a tensor mass to the momentum equation. Motion along field lines is unchanged. Perpendicular to field lines the magnetic field appears to add a mass density equal to $B^2/4\pi c^2$.

Although the geometry limits the applicability of the results, 2-D simulations can lead to insight into magnetospheric processes. In our simulation of a substorm-like event [6], we found substantial agreement with the neutral line model of substorms. As a southward IMF interacts with the magnetosphere, it merges with the geomagnetic field. The resulting reconnected field lines are convected tailward, increasing the strength of the tail lobe field. After a period of time reconnection commences in the near earth ($x > -20 R_E$) magnetotail. This x-point stays where it formed for approximately 20 minutes while high velocities build up in the anti-sunward direction. Finally the x-point moves tailward at high speed, but is rapidly replaced by another in the near earth region. The process repeats itself in a series of relaxation oscillations. The period of these oscillations is about 20 minutes. The entire sequence of events is strongly suggestive of the neutral line model with one exception. There are no high speed earthward flows on the earthward side of the x-point. We speculate that this is a geometric effect. In 2-D there is no place for the displaced material to go in the sunward direction.

The resistivity in those calculations was purely numerical because of the particular way the equations were solved, and had no Joule heating associated with it. To test the effects of other, potentially realistic, resistivities, we have done a number of simulations using an equatorial plane of symmetry. Numerical reconnection in the tail can quite easily be eliminated for this geometry. We find that there is only one striking difference between the simulations with numerical resistivity and those with some specific type of resistivity including Joule heating. Only the numerically resistive runs show the relaxation oscillations. For a resistivity with Joule heating, the x-point sets up in the near-earth region and stays there, although it moves back and forth over a narrow region in X. Using the same form of resistivity without Joule heating produces the relaxation oscillations. The numerical resistivity can be viewed, with its lack of Joule heating, as a limiting case where thermal conductivity is so high that the heat generated in the x-point is immediately carried away.

The 3-D simulations give some of the most interesting results. Our first 3-D calculation [7] studied the effects of an E-W IMF on the magnetosphere. This field penetrated the magnetosphere and introduces twisting distortion to the magnetotail. Even more interesting are the results of our second simulation [8] where the effects of a southward IMF were studied.

This simulation allowed a southward IMF to reconnect numerically on the nose in a similar fashion to the 2-D calculation. Once again the results bear a strong resemblance to the neutral line model of substorms. After sufficient stress is communicated to the magnetotail an x-point forms in the near-earth region and high speed flows stream away from it in both directions. One important point to be noted is that the speed of the flow away from the x-point appears to depend on a sort of Alfvén speed combining the tail lobe field strength with the plasma density in the plasma sheet. Thus, for magnetotail field strengths, densities in the plasma sheet must drop below 1 cm^{-3} in order to produce speeds greater than 200 km/s . A second point is that the entire reconnection event can be somewhat localized, so much so that a spacecraft in the wrong part of the plasma sheet may see nothing unusual while a dramatic reconnection event is occurring elsewhere in the tail. In contrast with the 2-D results onset of the event seems to be extremely rapid. At 30 Re in the tail flow speeds go from a relatively low value to one in excess of $600 \text{ km}^{-1} \text{ sec}$ in less than 2 minutes.

The change in the field topologies during the course of this event is quite striking. The electric field in the earliest stages reflects the solar wind field, and is an essentially constant, y directed field except within the plasma sheet. There, in the earliest stages the electric field does not penetrate. As the reconnection event reaches its maximum, the electric field in the tail increases in strength and penetrates the plasma sheet in the reconnection region. What is

startling, however, is that the field is no longer simply a Y field. The field vectors now form circular arcs around the reconnection region. This implies that the convection pattern is attempting to collapse the magnetotail in a cylindrical fashion, rather than the commonly visualized planar one. The current patterns are also striking. Early in the event the tail current is quite uniformly E-W across the tail. As the reconnection event goes on, the current density increases and executes a very pronounced tailward jog downward of the reconnection area. Just after the reconnection area it jogs back towards the earth. The picture is reminiscent of the idea of current disruption with the major exception that the disrupted current, rather than closing in the ionosphere, appears to be diverted tailward.

We have concentrated in our work and this talk on the magnetotail. This is for two reasons. First, we have long been interested in the processes leading to the substorms, and the tail seems to be the place where much of the action goes on. Second, the magnetotail is a place where the deficiencies in our simulation would seem to bother us relatively little. Twenty or more R_e in the tail we believe the action is controlled by the solar wind-magnetospheric interaction and that ionospheric effects are at worst a perturbation on the situation. However, substorms are not restricted to the tail. They do produce aurorae where the ionosphere is not just a perturbation. Indeed, modeling of the whole inner magnetosphere depends on doing a good job on the magnetospheric-ionospheric coupling.

We are embarking on a program to simulate the magnetospheric-ionospheric (MI) coupling using an MHD approach. The usual electrostatic equations are contained in the MHD equations. In fact, the electrostatic - J, E - picture is a different, though equivalent formalism to the MHD one. We are choosing an MHD approach for two reasons. First, the coupling of the MI code to the rest of the magnetosphere is much easier if the formalisms used in both codes are similar. Second, with an MHD formalism it is relatively easy to add the effects of plasma inertia and magnetic induction. Our initial attempt will be to simulate with a 3-D MI code the auroral zone and the inner magnetosphere that maps to it. We hope to be able to use this code to study MI coupling during disturbed times.

- [1] C.C. Wu, R.J. Walker, and J.M. Dawson, Geophys. Res. Lett., 8, 523 (1981).
- [2] J.U. Brackbill, "Numerical modeling of magnetospheric reconnection", Los Alamos National Laboratory preprint, LA-UR-82-483, submitted to Geophys. Res. Lett..
- [3] J.P. Boris and D.L. Book, J. Comp. Phys., 11, 38 (1973).
- [4] S.T. Zalesak, J. Comp. Phys., 31, 335 (1979).
- [5] K. Hain, "The Partial Donor Cell Method", Naval Research Laboratory, Memo Report 3713 (1977).
- [6] J.G. Lyon, S.H. Brecht, J.D. Huba, J.A. Fedder, and P.J. Palmadesso, Phys. Rev. Lett., 46, 1038 (1981).
- [7] S.H. Brecht, J. Lyon, J.A. Fedder, and K. Hain, Geophys. Res. Lett., 8, 397 (1981).
- [8] S.H. Brecht, J.G. Lyon, J.A. Fedder, and K. Hain, J. Geophys. Res., 87, A8, 6098 (1982).

APPENDIX B

GLOBAL SIMULATION AS A TOOL FOR

MAGNETOSPHERIC RESEARCH

Global Simulation as a Tool for Magnetospheric Research

S.H. Brecht and J.G. Lyon
Science Applications, Inc.
1710 Goodridge Drive
McLean, Virginia 22102

J.A. Fedder
Geophysical and Plasma Dynamics Branch
Plasma Physics Division
Naval Research Laboratory
4555 Overlook Avenue, S.W.
Washington, D.C. 20375

Numerical simulation occupies a middle ground between theory and experiment. It is theoretical in the sense that only an incomplete set of equations or a limited physical model is studied rather than the magnetosphere itself. However, simulation studies do not have to make the extreme simplifying assumptions which are often necessary to make problems analytically tractable. Numerical simulations are experimental in nature because they have inherent errors exactly analogous to experimental errors arising from instrumental resolution, systematic and random instrumental error, and sampling statistics. Therefore, a researcher performing a numerical simulation must always ask himself if his results are real. One additional similarity exists because the results obtained from both work and simulation are not analytic relations. The determination of physical relationships requires very careful analysis and variation of parameters.

It is precisely the experimental aspect of numerical simulations that make them an effective tool in areas of research, such as space physics, where controlled experiments are difficult to accomplish. Simulations can provide the link between a well considered theory and the

phenomena it is attempting to explain. Traditionally this link comes from precisely designed and controlled experiments. For example, we can do 3-D simulations to see how the magnetosphere responds if the IMF is rotated by 45° for an half hour. Nature is not usually so accommodating.

Global time-dependent simulations of the magnetosphere have been carried out for several years by two main groups, one at UCLA (LeRouef et al., 1978 and Wu et al., 1981), and at NRL (Lyon et al., 1980 and Brecht et al. 1981 (a)). The underlying physics in all of these simulations has been reduced to an MHD description. On large scales considerations of conservation of mass, momentum and energy ensure the approximate validity of the MHD assumption. Granting the validity of MHD, numerical problems still limit the accuracy with which the simulation can be carried out. It is our opinion that the quantitative accuracy of global simulations can be improved greatly in the coming years through judicious use of modern numerical techniques.

However, at the current level of sophistication, global simulations can provide insight into the physics of the magnetosphere. As examples, we refer to some of our own 2-D and 3-D results. At the 2-D level we have studied solar shock interactions with the magnetosphere (Lyon et al., 1980). Here the magnetotail experienced energization through the phenomena of shock focussing. Using the 3-D code we have investigated the effects of rotation in the IMF direction (Brecht et al., 1981) and found modifications to the tail structure that are in agreement with both theory (Cowley, 1981) and observation. Studies of tail reconnection have been performed in 2-D and 3-D (Lyon et al., 1981, and Brecht, et. al., 1981(b)). They resulted in x-point formation near the point predicted by the more or less phenomenological neutral line model for substorms.

The 3-D simulation of a series of reconnection events exemplifies the kind of comparison between simulation and experiment that can occur. Hones and Schindler, 1979 and Lui, 1979 in their studies of the IMP data have produced rather different categorization of the phenomena associated with substorms. In our simulation virtually all of these categories of plasma and magnetic field variation are found. This indicates that perhaps location was as important in determining the phenomena measured as the behavior of the tail during the individual substorms. What is missing from our simulation, and the data reported by these researchers, is the behavior of the solar wind during the events. Experimental data is needed to make the simulation realistic in terms of boundary conditions and initial conditions. In turn the simulation results can provide a global picture of magnetospheric behavior upon which to base analysis of satellite data taken at a given location. Just as clearly, a realistic simulation can be further improved by amplifying the physics embodied in the equations being solved; e.g., physical models to provide anomalous resistivity coefficients the resistive term in the MHD equations.

In the future, we foresee a dramatic increase in the ability of simulations to give quantitative information about the magnetosphere. Progress will come in two areas that are stumbling blocks at the moment - the underlying physics and the treatment of boundary layers. At present we can trace energy flow from the solar wind into and around the magnetosphere in a semi-quantitative fashion. To do better, we need three things: more interaction with experimental data, better theoretical understanding, and better numerical techniques. The last is probably the easiest with which to deal. It is our opinion that the state of the art

of numerics is such that global simulations could be carried out to the order of accuracy of the physics models available. The quality of the physical models depends on the level of the theoretical and experimental effort in the space physics community. Computer simulation can be a unique tool for making sense out of the magnetosphere, but only if it's firmly grounded in theory and observation.

References

- Brecht, S.H., J.G. Lyon, J.A. Fedder, and K. Hain, "A simulation study of east-west IMF effects on the magnetosphere", Geophys. Res. Lett., 8, 397, 1981 (a).
- Brecht, S.H., J.G. Lyon, J.A. Fedder, and K. Hain, "A time-dependent three-dimensional simulation of the earth's magnetosphere: Reconnection events", NRL Memorandum Report 4690, 1981 (b).
- Cowley, S.W.H., "Magnetospheric asymmetries associated with the v-component of the IMF," Planet. Space Sci., 29, 79, 1981.
- Hones, E.W., Jr., and K. Schindler, "Magnetotail plasma flow during substorms: A survey with IMP 6 and IMP 8 satellites," J. Geophys. Res., 84, 7155, 1979.
- LeRouef, J.N., T. Tajima, C.F. Kennel, and J.M. Dawson, "Global simulations of the time-dependent magnetosphere," Geophys. Res. Lett., 5, 609, 1978.
- Lui, A.T.Y., "Observations on plasma sheet dynamics during magnetospheric substorms," Dynamics of the Magnetosphere, 563-597, Reidel Pub. Co., 1979.
- Lyon, J.G., S.H. Brecht, J.A. Fedder and P.J. Palmadesso, "The effect on the earth's magnetotail from shocks in the solar wind," Geophys. Res. Lett., 7, 712, 1980.
- Lyon, J.G., S.H. Brecht, J.D. Huba, J.A. Fedder and P.J. Palmadesso, "Computer simulation of a geomagnetic substorm," Phys. Rev. Lett., 46, 1038, 1981.
- Ku, C.C., R.J. Walker and J.M. Dawson, "A three-dimensional MHD model of the earth's magnetosphere," Geophys. Res. Lett., 8, 525, 1981.

APPENDIX C

SATURATION OF THE LOWER-HYBRID-DRIFT INSTABILITY
BY MODE COUPLING

Saturation of the Lower-Hybrid-Drift Instability

by Mode Coupling

J.F. Drake
University of Maryland
College Park, Maryland 20742

P.N. Guzdar
Science Applications, Inc.
McLean, Virginia 22102

and

J.D. Huba
Geophysical and Plasma Dynamics Branch
Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

Abstract

A nonlinear mode-coupling theory of the lower-hybrid-drift instability is presented. It is found that the instability saturates by transferring energy from the growing, long wavelength modes to the damped, short wavelength modes. The saturation energy, mean square of the potential fluctuations, and diffusion coefficient are calculated self-consistently.

A first principles theory of the nonlinear saturation of the lower-hybrid-drift instability and associated particle transport is presented. A nonlinear equation describing 2-D lower-hybrid-drift turbulence is derived which includes the nonlinear $\underline{E} \times \underline{B}$ and polarization drifts of the electrons, adiabatic ions, and realistic sources and sinks of energy. The equation is solved numerically to obtain the evolution and saturation of the wave spectrum, as well as the self-consistent particle flux which yields the diffusion coefficient.

The lower-hybrid-drift instability is a flute mode ($\underline{k} \cdot \underline{B} = 0$) which is driven unstable in a plasma with strong inhomogeneities, $\rho_i/L_n > (m_e/m_i)^{1/4}$, where ρ_i and L_n are the ion Larmor radius and density scale length, respectively. The linear behavior of the instability is well understood.^{1,2} In the weak drift regime ($V_{di} < v_i$, where $V_{di} = v_i^2/2\Omega_i L_n$ is the ion diamagnetic drift velocity and v_i is the ion thermal velocity) the mode is driven unstable by the resonant interaction of the wave with the drifting ions when $\omega/k_y < V_{di}$. The growth rate maximizes at $k_y \sim \rho_{es}^{-1} \equiv \rho_i^{-1}(m_i/m_e)^{1/2}$. The mode is stable for $k_y \rho_{es} \gg 1$, and approaches marginal stability as $k_y \rho_{es} \rightarrow 0$.³ The wave frequency is such that $\Omega_i \ll \omega \ll \Omega_e$ so that the electrons are strongly magnetized while the ions can be treated as unmagnetized. In a finite β plasma the resonant interaction of VB drifting electrons with the wave is stabilizing.

The nonlinear behavior of the instability is much less well understood. Ion trapping has been observed to quench the growth of the instability in particle simulations.⁴ The onset of stochastic electron heating has also been proposed to explain saturation in these simulations.⁵ On the other hand, ion trapping is not a viable saturation mechanism when a broad 2-D spectrum of waves is excited, and it has not

been observed in recent simulations using realistic mass ratios where such spectra develop.⁶ In this letter we focus on the nonlinear coupling of stable and unstable waves to saturate the instability.

We consider a slab equilibrium of cold electrons and warm ions with a density profile $n_0(x)$ supported by a magnetic field $B_{z0}(x)$. The equilibrium ion velocity distribution is taken to be a Maxwellian with an average drift velocity $\tilde{v}_i = v_{di} \hat{e}_y$. Because of the flute nature of the instability ($\underline{k} \cdot \underline{B} = 0$), we self-consistently limit the spatial variations to the xy plane (as contrasted with universal drift-wave turbulence which is inherently 3-D). In the weak drift regime the ions respond to the perturbed potential $\phi(x,y,t)$ adiabatically to lowest order since $\omega \ll kv_i$. Thus,

$$n_i = n_0 [1 + \sqrt{\pi} v_i^{-1} |\nabla^{-1}| (\frac{\partial}{\partial t} + v_{di} \frac{\partial}{\partial y})] \exp(-e\phi/T_i) \quad (1)$$

where the term proportional to $\sqrt{\pi}$ in (1) is a small correction describing the resonant ion interaction. The perturbed electron motion is simply given by the $\underline{E} \times \underline{B}$ and polarization drifts,

$$\tilde{v}_e = \frac{c}{B} \nabla \phi \times \hat{e}_z - \frac{c}{B\Omega_e} \frac{d}{dt} \nabla \phi \quad (2)$$

where $d/dt = \partial/\partial t + \tilde{v}_e \cdot \nabla$. Using (2), the electron continuity equation can be written as

$$\frac{d}{dt} \ln n_e = - \rho_{es}^2 \frac{d}{dt} \nabla^2 (e\phi/T_i) \quad (3)$$

Finally, invoking charge neutrality ($n_e = n_i$) and combining (1) and (3), we

obtain the nonlinear equation

$$(1 - \hat{v}^2) \hat{\phi}_\tau + \hat{\phi}_y - \gamma_0 |\nabla^{-1}| (\hat{\phi}_{\tau\tau} + \hat{\phi}_{\tau y}) + \gamma_e \hat{\phi} + (\nabla \hat{\phi} \times \hat{e}_z \cdot \nabla) \nabla^2 \hat{\phi} + \gamma_0 \nabla \hat{\phi} \times \hat{e}_z \cdot \nabla |\nabla^{-1}| (\hat{\phi}_\tau + \hat{\phi}_y) = 0 \quad (4)$$

where $\hat{v} = \rho_{es} \nabla$, $\tau = (\rho_{es}/L_n) \Omega_e t$, $\hat{\phi} = (e\phi/T_i)(L_n/\rho_{es})$ and the subscripts on $\hat{\phi}$ denote a derivative with respect to that variable. The quantity γ_e represents the wave damping due to VB resonant electrons in a finite β plasma. Equation (4) is only valid for $\gamma_0 = \pi^{1/2} v_{di}/v_i < 1$ since the adiabatic ion response can only be justified in this limit. Linearizing this equation, we obtain the complex eigenvalue (in our normalized units),

$$\hat{\omega} = \hat{\omega}_k = \frac{\hat{k}_y}{1 + \hat{k}^2} + i\gamma_0 \frac{\hat{k}_y^2 |\hat{k}|}{(1 + \hat{k}^2)^3} - i\gamma_e \frac{1}{1 + \hat{k}^2} \quad (5)$$

In the limit $\gamma_e, \gamma_0 \rightarrow 0$, (4) reduces to the Hasegawa-Mima equation in which the nonlinearity arises from the nonlinear polarization drift.⁷ This equation has two invariants, energy and enstrophy, neither of which is preserved in the more general (4). When γ_0 is finite, the $\underline{E} \times \underline{B}$ nonlinearity also appears in (4). This nonlinearity has been considered by Horton⁸ and, more recently by Waltz,⁹ in studying universal mode turbulence. However, our equation differs from their work in that we do not assume that $\hat{\phi}_\tau + \hat{\phi}_y$ is replaced by its linear counterpart $-i(\hat{\omega}_k - \hat{k}_y)\hat{\phi}$. Thus, our equation is second order in time rather than first order.

The direction of energy flow described by (4) can be understood by calculating the stability of a single large amplitude wave $\hat{\phi}_0$ with $\hat{k}_0$

and $\hat{\omega}_0$ satisfying (5). A perturbation of wavevector $\underline{k}$ is coupled through the pump to modes with $\underline{k} \pm p\underline{k}_0$ ($p = 1, 2, \dots$). For simplicity, we consider only the coupling of $(\hat{\omega}, \underline{k})$ with its nearest neighbors $(\hat{\omega}_{\pm}, \underline{k}_{\pm})$ where $\hat{\omega}_{\pm} = \hat{\omega} \pm \hat{\omega}_0$ and $\underline{k}_{\pm} = \underline{k} \pm \underline{k}_0$, i.e., $p=1$. The dispersion relation for this $(\underline{k}, \underline{k}_{\pm})$ coupled system is

$$\epsilon(\underline{k}, \hat{\omega}) + |\phi_0|^2 \left[\frac{M(\underline{k}_0, \underline{k}_{\pm}) M(\underline{k}, \hat{\omega}_0)}{\epsilon(\underline{k}_{\pm}, \hat{\omega}_{\pm})} + (\underline{k}_0 + -\underline{k}_0) \right] = 0 \quad (6)$$

where $\epsilon(\underline{k}, \omega) = \omega[1 + k^2(1 - i\delta_k)] - k_y + i\gamma_e$, $\delta_k = \gamma_0(k_y - \omega)/k^3$ and $M(\underline{k}_1, \underline{k}_2) = \underline{k}_1 \times \underline{k}_2 \cdot \hat{e}_z [k_1^2(1 - i\delta_{k1}) - k_2^2(1 - i\delta_{k2})]$. When γ_0 and γ_e are neglected in (6) and we take the limit $\hat{\omega} \gg \hat{\omega}_k$, the decay modes are purely growing with a growth rate which peaks around $\underline{k} \cdot \underline{k}_0 = 0$ with $|\underline{k}| < |\underline{k}_0|$. A necessary requirement for instability is that one of the decay waves $(\underline{k}, \underline{k}_{\pm})$ has a longer wavelength than the pump. For this situation, in which (4) reduces to the Hasegawa-Mima equation, the wave energy inevitably cascades to longer and longer wavelength so that no stationary wave spectrum can result.¹⁰ When γ_0 is included, this conclusion no longer remains valid. Taking the limit $|\underline{k}_0| \ll |\underline{k}|$ and again assuming $\hat{\omega} \gg \hat{\omega}_k$, the dispersion relation is given by

$$\hat{\omega}^2 = \frac{2|\phi_0|^2 (\underline{k} \times \underline{k}_0 \cdot \hat{e}_z)^2 k^4}{1 + k^2} \left[1 + i \frac{\gamma_0 \hat{\omega} (2 + k^2)}{|k|^3 (1 + k^2)} \right]. \quad (7)$$

Equation (7) yields a dissipative instability which is driven unstable by the $\underline{E} \times \underline{B}$ nonlinearity. This instability produces a flow of energy from long to short wavelengths and, as will be demonstrated, enables the wave spectrum to reach a steady state.

We solve (4) numerically by decomposing it into two coupled, first-

order (in time) differential equations which are advanced in time. These equations are solved by a spectral method developed by Fyfe et al.,¹¹ based on the work of Orzag.¹² The potential $\hat{\phi}$ is Fourier-decomposed, i.e., $\hat{\phi} \sim \exp(i\mathbf{k} \cdot \mathbf{x})$ where $\mathbf{k} = \mathbf{n}/\lambda$ and $\mathbf{n} = n_x \hat{\mathbf{e}}_x + n_y \hat{\mathbf{e}}_y$ with n_x and n_y integers, and $\lambda = 5$ (which fixes the $|\mathbf{n}|$ for which $|\mathbf{k}| = 1$). The numerical results presented in this letter are nominally computed on a 32×32 mesh.

The electron damping γ_e controls the region of instability in $\mathbf{k}$ space by stabilizing short wavelength modes. We have chosen

$$\gamma_e = \gamma_0 |\hat{\mathbf{k}}|^7 (1 + \hat{k}^2)^{-2} k_m^{-4} \quad (8)$$

with $\hat{k}_m = 1.6$, which yields an unstable wave spectrum which is realistic for the lower-hybrid-drift instability in finite β plasmas. The growth rate peaks at $(\hat{k}_x, \hat{k}_y) = (0, +1)$ and is stable for $\hat{k}_y > 1.6$ and $\hat{k}_x > 1$.

The $\hat{\phi}$ spectrum is initialized with random noise with $\hat{\phi} \sim 10^{-2} - 10^{-3}$, and the system is allowed to evolve until the wave energy, given by

$$W/nT = (m_e/\pi m_i) \gamma_0^2 \sum_k (1 + \hat{k}^2) |\hat{\phi}_k|^2, \quad (9)$$

reaches a steady state value. The mean square of the potential fluctuations,

$$P = \langle (e\phi/T)^2 \rangle^{1/2} = [(2m_e/\pi m_i) \gamma_0^2 \sum_k |\hat{\phi}_k|^2]^{1/2}, \quad (10)$$

as well as the nonlinear and quasilinear diffusion coefficients

$$\hat{D}_{nl} = D_{nl}/v_{cs} \rho_{cs} = (2m_e/\pi m_i)^{1/2} \gamma_0^2 \sum_k (\hat{k}_y/|\hat{\mathbf{k}}|) \hat{\phi}_k^* (\hat{k}_y \hat{\phi}_k - i \frac{\partial \hat{\phi}_k}{\partial \tau}) \quad (11)$$

$$\hat{D}_{q1} = D_{q1}/v_{es} \rho_{es} = (2m_e/\pi m_i)^{1/2} \gamma_0^2 \sum_k |\hat{k}| \hat{k}_y^2 |\hat{\phi}_k|^2 (1 + k^2)^{-1} \quad (12)$$

respectively, are also computed. The nonlinear diffusion coefficient $\hat{D}_{nl}$ is derived by averaging the particle flux in the x direction over y, i.e., $\hat{D}_{nl} \propto \langle nV_x \rangle_y$. The quasilinear expression is then obtained by approximating $\partial/\partial\tau = -i\omega_k$ and using (5) in (11).

In Fig. 1 we show instantaneous 2-D wave spectra in $\underline{n}$ space for $\gamma_0 = .5$. Figure 1a shows the spectrum during the linear phase of the instability, and Fig. 1b shows the spectrum after saturation. In the linear regime, the wave spectrum is strongly peaked around the most unstable modes, $(n_x, n_y) = (0, \pm 5)$. The spectrum begins to saturate as these modes couple through a $n_y = 0$ mode (typically with $n_x \sim 4$) to damped modes with $n_x \gtrsim n_y$. A peak in the wave spectrum appears around (4,0) at this time. This transient phase is completed as the total wave energy saturates and spreads through most of the unstable, weakly damped volume of $\underline{n}$ space as shown in Fig. 1b. The spectrum typically remains peaked around $n_x \sim 0$ with n_y typically somewhat smaller than that of the linearly most unstable mode [see Fig. 1b for which the dominant modes are $(n_x, n_y) = (0, \pm 3)$]. The 2-D wave spectrum exhibits substantial variability in time, even after saturation, as the unstable and stable modes continue to exchange energy in a dynamic manner.

The time evolution of the total wave energy (W), mean square of the potential fluctuations (P), and the nonlinear ($\hat{D}_{nl}$) and quasilinear diffusion ($\hat{D}_{q1}$) coefficients are shown in Fig. 2. All of these quantities exhibit a similar temporal behavior. The initial decay ($\tau < 4$) is associated with the rapid dissipation of energy initialized in the damped

modes, and is followed by a linear growth phase ($4 < \tau < 20$). Subsequently, mode coupling occurs which leads to saturation of the instability, albeit with some initial overshoot ($\tau \sim 24$). The levels of the total wave energy and other parameters of Fig. 2 are quite stationary in time after saturation. Also, the stationary values of all four quantities are relatively insensitive to the initialization of $\hat{\phi}$.

We note that in Fig. 2, the quasilinear diffusion coefficient ($\hat{D}_{ql}$) tracks the actual, nonlinear diffusion coefficient ($\hat{D}_{nl}$) quite well during the entire time evolution of the instability. An important point which must be emphasized with regard to $\hat{D}_{nl}$ is that both species, electrons and ions, continue to exchange both energy and momentum even after a steady state is reached; the electrons through the VB resonance and the ions by direct resonant interaction. If the instability had saturated by ion trapping and the electrons had no resonant interaction with the wave, there would be no diffusion in the steady state since the electrons could not exchange momentum with the ions. Both species must have a dissipative interaction with the waves to have diffusion.

A number of runs have been made for different values of γ_0 , the drift parameter. The saturation level of $\hat{\phi}$ is found to be nearly independent of γ_0 . Thus, from Eqs. (11) and (12), it is found that $\hat{D}_{nl} \propto \gamma_0^2 \propto v_{di}^2/v_i^2$.

Next, we compare our stationary values of P and $\hat{D}_{nl}$ with these obtained from "wave energy bound" calculations.¹³ The "energy bound" results from equating the wave energy with the kinetic energy associated with the relative drift of the electrons and ions, $m_e n (V_{di} - V_{de})^2/2$. This free energy bound is actually an extreme underestimate of the free energy in a finite β plasma, since the instability can also feed off the magnetic free energy, which typically greatly exceeds the drift energy.¹⁴

Nevertheless, this bound has been widely invoked and seems to yield reasonable transport rates when compared with experimental observations. The "drift energy bound" yields a potential $P_{eb} = (m_e/\pi m_i)^{1/2} \gamma_0$, where we have assumed $T_e \ll T_i$. The scaling of P_{eb} is identical with that given in (10) since $\hat{\phi}$ is independent of γ_0 . For the case $\gamma_0 = .5$, $P/P_{eb} = 3$ (from Fig. 2) so that the actual fluctuation level obtained from our code is slightly larger than calculated from the "energy bound". Finally we compare $\hat{D}_{nl}$ with that obtained from the simple formula $\hat{D} = (\gamma/k^2)_m / v_{es} \rho_{es}$, where γ and k are evaluated where γ peaks. We find $\hat{D} = (m_e/\pi m_i)^{1/2} \gamma_0^2$, which again has the same scaling as $\hat{D}_{nl}$ but is a factor of 2.4 smaller for $\gamma_0 = 0.5$ (from Fig. 2).

Measurements of lower-hybrid-drift turbulence in a θ -pinch by CO_2 laser scattering have been reported recently.¹⁵ The observed wave spectra were flute-like ($\underline{k} \cdot \underline{B} = 0$) with clear peaks around $\hat{k}_y \approx 0.7$ (a factor of two smaller than the most unstable linear mode), and $\hat{k}_x < \hat{k}_y$. Data was taken for three different filling pressures, corresponding to three values of V_{di}/v_i . Electron-ion collisions were apparently significant [$v_{ei} \sim (\Omega_e \Omega_i)^{1/2}$] for the two highest filling pressures so we can only compare our results with the data from the lowest filling pressure, which corresponds to $V_{di}/v_i = 0.52$ ($\gamma_0 \approx 0.9$). The measured density fluctuation, $\tilde{n}/n_0 \sim 0.014$, is quite close (smaller by a factor of 2) to our theoretical prediction of $\tilde{n}/n_0 \approx 0.025$. In addition, the shift to long wavelength is consistent with our calculated spectra (compare Figs. 1a and 1b).

Computer simulations of the lower-hybrid-drift instability with realistic mass ratios have also been performed recently using a fluid-kinetic numerical scheme.⁶ Unfortunately, these computations have been carried out for $V_{di}/v_i > 1$, which is outside the range of validity of our

present theory. Nevertheless, a broad spectrum of modes is observed in k space in the simulations which is consistent with our calculations.

In conclusion, the lower-hybrid-drift instability can saturate via a mode coupling process. The basic saturation mechanism involves a transfer of energy from the growing, long wavelength modes to the damped, short wavelength modes. The nonlinear $\underline{E} \times \underline{B}$ drift plays a crucial role in this saturation mechanism by preventing the polarization drift nonlinearity from driving the wave energy towards the undamped, long wavelength modes. The saturation energy and diffusion coefficient associated with the wave turbulence scales as $(v_{di}/v_i)^2$.

Acknowledgments

We are very grateful to Dave Fyfe and Glenn Joyce for providing us with their spectral method code. We also thank Dave Montgomery for several helpful discussions. This research has been supported by ONR, and NASA. One of us (JFD) was partially supported by DOE.

References

1. N.A. Krall and P.C. Liewer, Phys. Rev. A4, 2094 (1971).
2. R.C. Davidson, N.T. Gladd, C.S. Wu, and J.D. Huba, Phys. Fluids 20, 301 (1977).
3. J.F. Drake, J.D. Huba, and N.T. Gladd, submitted to Phys. Fluids (1982).
4. D. Winske and P.C. Liewer, Phys. Fluids 21, 1017 (1978).
5. J.F. Drake and T.T. Lee, Phys. Fluids 24, 1115 (1981).
6. J.U. Brackbill, D.W. Forslund, K. Quest, and D. Winske (private communication)
7. A. Hasegawa and K. Mima, Phys. Fluids 21, 87 (1978).
8. W. Horton, Phys. Rev. Lett. 37, 1269 (1976).
9. R.E. Waltz, submitted to Phys. Fluids (1982).
10. D. Fyfe and D. Montgomery, Phys. Fluids 22, 246 (1974).
11. D. Fyfe, G. Joyce, and D. Montgomery, J. Plasma Phys. 17, 317 (1977).
12. Orszag, S.A., Stud. Appl. Math. 50, 293 (1971).
13. R.C. Davidson, Phys. Fluids 21, 1017 (1978).
14. J.F. Drake, N.T. Gladd, and J.D. Huba, Phys. Fluids 24, 301 (1981).
15. H.U. Fahrback, W. Koppendorfer, M. Munich, J. Neuhauser, H. Rohr, G. Schramm, J. Sommer, and E. Holzhauser, Nuc. Fusion 21, 257 (1981).

Figure Captions

1. Instantaneous 2D wave spectra of lower-hybrid-drift wave turbulence. (1a) Linear stage: The dominant linear modes are $(n_x, n_y) = (0, \pm 5)$. (1b) Nonlinear stage: The dominant modes are $(n_x, n_y) = (0, \pm 3)$. Note the shift to longer wavelength in the nonlinear saturated state.
2. Time evolution of the wave energy (W), mean square of the potential fluctuations (P), and diffusion coefficients (D_{q1} - quasilinear; D_{n1} - nonlinear) for $\gamma_0 = 0.5$.

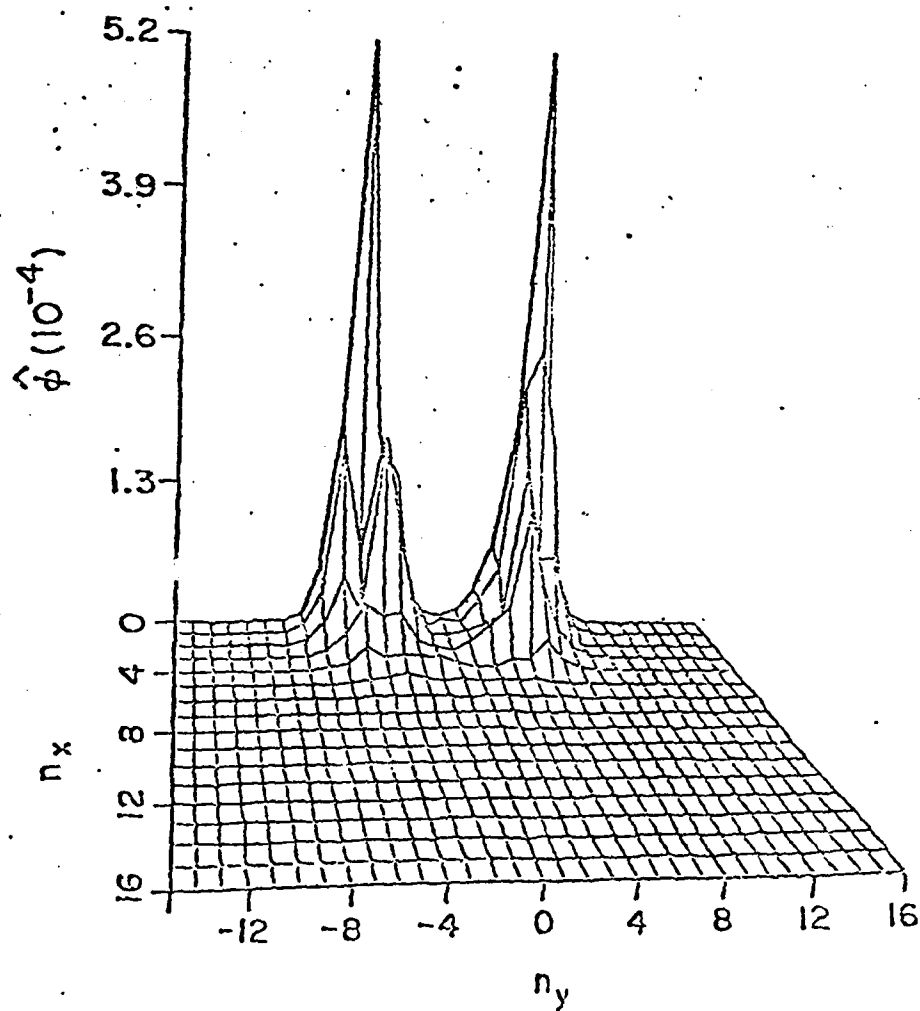


Figure 1a

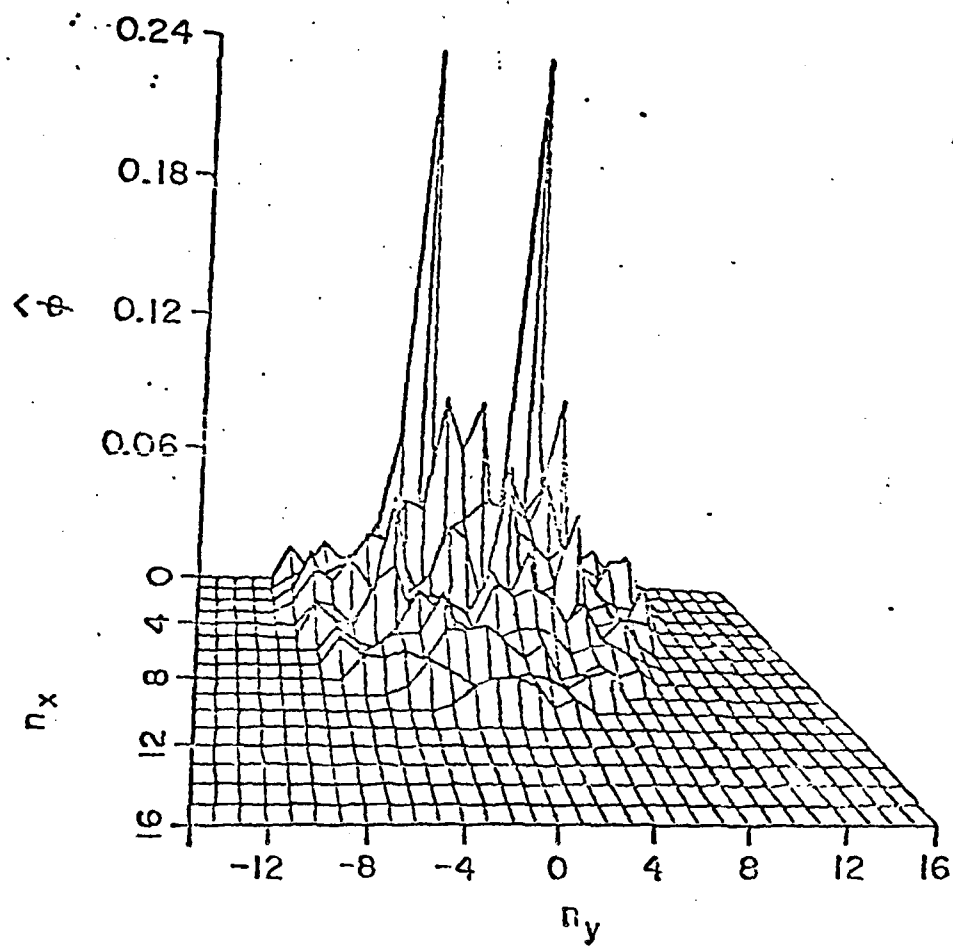
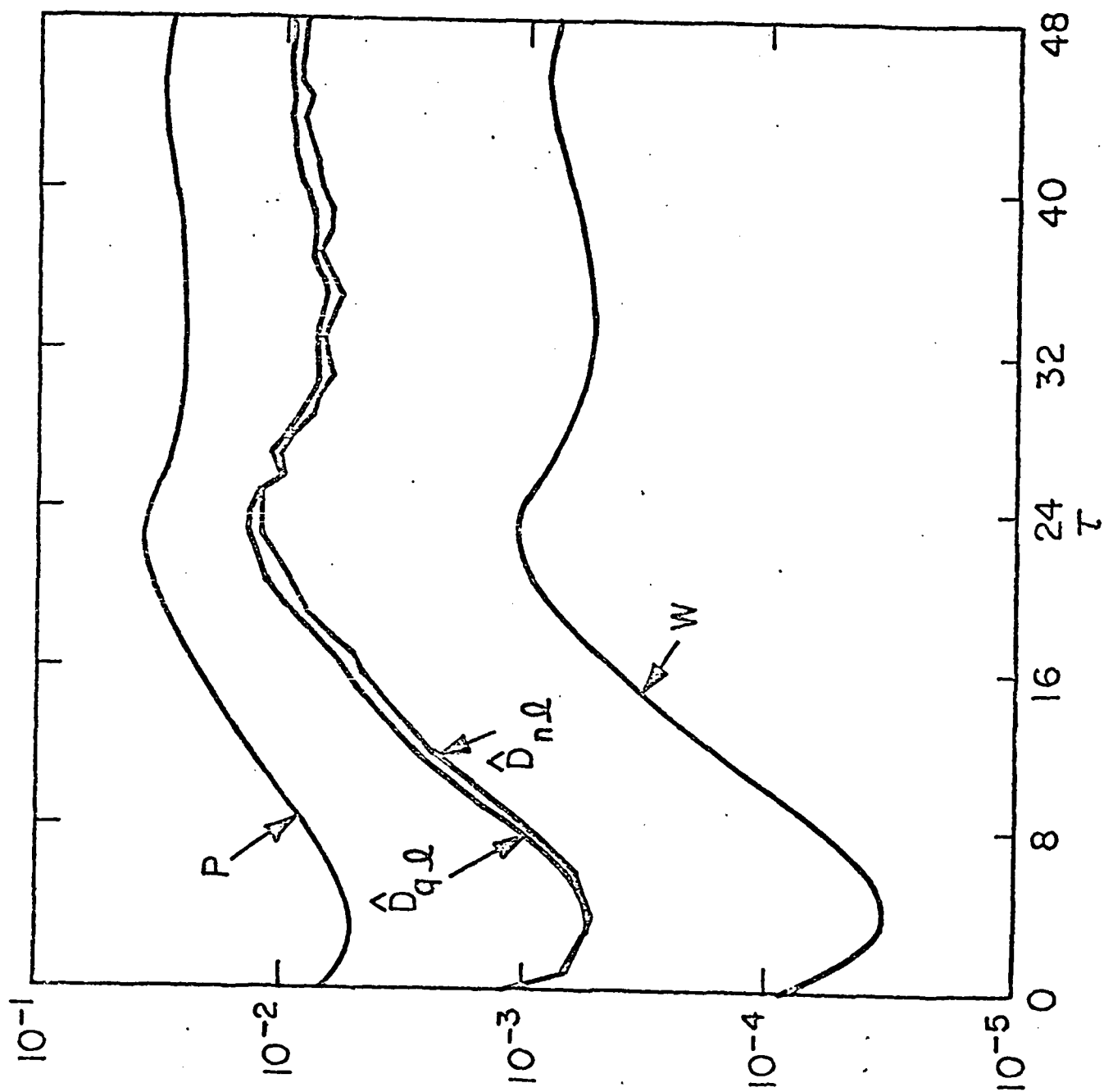


Figure 1b



APPENDIX D

ON THE "STABILIZATION" OF THE
LOWER-HYBRID-DRIFT INSTABILITY IN FINITE β PLASMAS

ON THE "STABILIZATION" OF THE
LOWER-HYBRID-DRIFT INSTABILITY IN FINITE β PLASMAS

J.F. Drake
Science Applications, Inc.
McLean, Va. 22102

J.D. Huba
Geophysical and Space Plasma Branch
Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

and

N.T. Gladd
JAYCOR
San Diego, Ca. 92138

Abstract

The stability properties of the lower-hybrid-drift instability are re-examined for finite β plasmas. In contrast to previous results, it is found that finite β does not stabilize the instability in the sense that the growth rate becomes negative. Rather, as β increases, the fastest growing mode shifts to longer wavelengths and is driven towards marginal stability.

The lower-hybrid-drift instability is a flute mode ($\underline{k} \cdot \underline{B} = 0$) which can occur in magnetized plasmas containing density gradients when $\rho_i/L_n > (m_e/m_i)^{1/4}$, where ρ_i is the mean ion Larmor radius and L_n is the scale length of the density gradient.^{1,2} For density gradients such that $\rho_i/L_n > 1$, the instability arises because of a coupling between a negative energy drift wave ($\omega = \omega_{*i}$ where ω_{*i} is the ion diamagnetic drift frequency) and a positive energy lower hybrid wave ($\omega = \omega_{lh}$ where $\omega_{lh} = \omega_{pi}(1 + \omega_{pe}^2/\Omega_e^2)^{-1/2}$ is the lower hybrid frequency).¹ On the other hand, for density gradients such that

$1 > \rho_i/L_n > (m_e/m_i)^{1/4}$, the instability is driven by an ion-drift wave resonance (i.e., inverse ion Landau damping).² The mode is characterized at maximum growth by $\omega \sim \omega_{*i} < \omega_{lh}$,

$\gamma < \omega_{*i}$ and $k \sim \rho_{es}^{-1} = (m_i/m_e)^{1/2} \rho_i^{-1}$. The electrons are therefore strongly magnetized, while the ions behave as if they were unmagnetized.

In magnetospheric and laboratory plasmas, where the instability is believed to cause anomalous cross-field transport, $\beta \sim 8\pi nT/B^2 \sim 1$ so that the mode is electromagnetic and the ∇B drifts of electrons must be included. Davidson et al. (1977)³ have investigated the effect of finite β on the instability and have concluded that above a critical value of β_e ($\beta_e = 8\pi nT_e/B^2$), ∇B resonant electrons completely stabilize the lower-hybrid-drift instability. In the present note we show that while the dispersion relation derived in Ref. 3 is correct, the conclusion that finite β_e stabilizes this mode is not

entirely valid. Rather, as β_e is increased, the originally most unstable modes with $k\rho_{es} \sim 1$ are stabilized and the unstable k spectrum collapses towards long wavelength. The growth rate of the most unstable mode decreases rapidly when $\beta_e \sim 1$, and falls below the ion cyclotron frequency so that unmagnetized ion approximation becomes questionable. Nevertheless, the long wavelength modes are not stabilized by finite β_e . In the limit $k\rho_{es} \rightarrow 0$, the stabilizing VB resonant contribution from the electrostatic portion of the equation is exactly cancelled by a corresponding resonant contribution from the magnetic portion of the equation. Thus, the well known fluid result that a plasma cannot stabilize itself by digging its own magnetic well^{4,5} also applies to the high frequency limit $\omega \gg \Omega_i$ with resonant particles.

To demonstrate the cancellation of the resonant electron contribution in the dispersion equation for the lower-hybrid-drift mode, we first present the dispersion equation derived by Davidson et al. (1977)³ for plasmas containing gradients along x and B in the z direction. The charge neutrality condition, with magnetized electrons and unmagnetized ions, is given by

$$k^2 \tilde{\phi} + k_{De}^2 (1 - H_{00}) \tilde{\phi} + k_{Di}^2 [(1 + \zeta_i) Z(\zeta_i)] \tilde{\phi} + k_{De}^2 (2H_{10}/k\rho_e) \tilde{B} = 0 \quad (1)$$

where $\tilde{\phi}$ and $\tilde{B}$ are the potential and magnetic (along z) perturbations normalized to T_e/e and B_z , respectively,

$k_{Dj}^2 = 4\pi n_0 e^2 / T_j$, $\zeta_i = (\omega - \omega_{*i}) / |k| v_i$, Z is the plasma dispersion function,

$$H_{ij} = -2 v_e^{-1} \int_0^\infty dv_\perp \left(-\frac{v_\perp}{v_e}\right)^{i+j+1} \exp\left(-\frac{v_\perp^2}{v_e^2}\right) \left(1 - \frac{\omega_{*e}}{\omega}\right) \\ \times \left(1 - \frac{\omega_B}{\omega}\right)^{-1} J_0^i(\mu) J_0^j(\mu), \quad (2)$$

and

$$v_e = (2T_e/m_e)^{1/2}, \quad \mu = kv_\perp/|\Omega_e|,$$

$$\omega_{*j} = kv_{dj} = (cT_j/q_j B) d \ln n/dx$$

$$\omega_B = -(v_\perp^2/2|\Omega_e|) d \ln B/dx,$$

and J_0^j is the j th derivative of the Bessel function J_0 . The term proportional to k_{Di}^2 (k_{De}^2) is the ion (electron) density perturbation. The equilibrium pressure balance relation relates the density and magnetic field gradients via

$$d \ln B/dx = -(\beta/2) d \ln n/dx$$

where the temperature T_j has been taken to be uniform. The dispersion equation is completed with the x component of Ampere's law which yields an equation for $\hat{B}$ in terms of $\hat{\phi}$,

$$\hat{B} + 2\beta_e (H_{11}/k^2 \rho_e^2) \hat{B} - \beta_e (H_{10}/k \rho_e) \hat{\phi} = 0. \quad (3)$$

where $\rho_e = v_e/|\Omega_e|$ is the mean electron Larmor radius. The terms

proportional to β_e are essentially the electron current along x.

To solve the dispersion equation given by (1) and (3), we first consider the limit $T_e \ll T_i$ and $\beta_e \ll 1$. In this limit the velocity integral in the expression for H_{ij} in (2) can be evaluated analytically. Retaining the resonant contributions to these integrals and approximating $Z(\zeta_i) \approx i/\pi$ for $\zeta_i \sim \omega_{*i}/|k|v_i \sim \rho_i/L_n \ll 1$, we obtain the dispersion equation

$$D(\omega, k) = (1 + \beta_i/2)(1 - \omega_{*i}/\omega) + k^2 \rho_{es}^2 (1 + \Omega_e^2/\omega_{pe}^2) \quad (4a)$$

$$+ i\sqrt{\pi} \frac{\omega}{|k|v_i} \left(1 - \frac{\omega_{*i}}{\omega}\right) + i\pi \frac{T_i}{T_e} F s_r^2 \exp(-s_r^2) = 0$$

with

$$F = [J_0(\mu_r) \left(1 - \frac{\omega_{*i}}{\omega}\right) - \frac{\omega}{\omega_{*i}} J_2(\mu_r)]^2, \quad (4b)$$

$s_r^2 = 2\omega/\beta_e \omega_{*i}$ and $\mu_r^2 = (4k^2 \rho_{es}^2 / \beta_i)(\omega/\omega_{*i})$. The real frequency and growth rate of the mode become

$$\omega_r = \omega_{*i} / (1 + k^2/k_0^2) \quad (5a)$$

$$k_0^2 = (1 + \beta_i/2) / [\rho_{es}^2 (1 + \Omega_e^2/\omega_{pe}^2)] \quad (5b)$$

$$\frac{\gamma}{\omega_{*i}} = \left(\frac{\omega_r}{\omega_{*i}}\right)^2 \frac{1}{1 + \beta_i/2} \left[\sqrt{\pi} \frac{\omega_{*i} - \omega_r}{|k|v_i} - \pi \left(\frac{T_i}{T_e}\right) F(\omega_r) s_r^2 \exp(-s_r^2) \right] \quad (5c)$$

The ion resonant contribution, proportional to $\sqrt{\pi}$ in (5c), is destabilizing since $\omega_r < \omega_{*i}$. The resonant electron

contributions from H_{00} , H_{10} and H_{11} in (1) and (3) are all of the same order and, when combined, are stabilizing since $F(\omega_r) > 0$.

In low β plasmas, $s_r^2 \propto \beta_e^{-1} \gg 1$ so that the electron stabilizing contributions are negligible and the mode is unstable. By increasing β_e with $k\rho_{es} \sim \beta_i \sim 1$, we reduce s_r until the electron resonant interaction stabilizes the mode [the second term overcomes the first term in (5c)]. Thus, for $k\rho_{es} \sim 1$, finite β_e stabilizes the lower-hybrid-drift instability.

However, while $F \sim 1$ for $k^2\rho_{es}^2 \sim 1$, for $\mu_r \propto k^2\rho_{es}^2 \ll 1$,

$F \propto k^4\rho_{es}^4$. Even though the individual resonant electron contributions from H_{00} , H_{10} and H_{11} are large, they exactly cancel in the limit $k\rho_{es} \rightarrow 0$. In particular, the destabilizing ion contribution to (5c) scales as $k^2\rho_{es}^2$ for small $k\rho_{es}$, while the stabilizing electron contribution scales as $k^4\rho_{es}^4$. For small enough $k\rho_{es}$, the mode is therefore still unstable.

To show that this conclusion is not a consequence of the assumption $\beta_e, T_e/T_i \ll 1$, we return to (1)-(3) and take the limit $k\rho_{es} \rightarrow 0$. The dispersion equation for the lower-hybrid-drift instability can then be written as

$$D(\omega, k) = 1 - \frac{\omega_{*i}}{\omega} \left[1 + \frac{\beta}{2} \left(\frac{\omega - \omega_{*e}}{\omega} \right) I \right] + \zeta_i Z(\zeta_i) + \frac{(\beta_i/2) I^2 (1 - \omega_{*e}/\omega)^2}{1 + (\beta_e/2)(1-I)(\omega/\omega_{Bt})(1 - \omega_{*e}/\omega)} = 0 \quad (6)$$

where

$$I = \int_0^\infty ds (1 - s\omega_{Bt}/\omega)^{-1} \exp(-s)$$

and $\omega_{Bt} = \omega_B(v_e)$. This equation has been independently derived using a guiding center equation for electrons and the Vlasov equation for unmagnetized ions. Multiplying (6) by the denominator of the last term and combining terms, we reduce (6) to

$$D(\omega, k) = \left(1 - \frac{\omega_{*i}}{\omega}\right) \left(\left(1 - \frac{\omega}{\omega_{*i}}\right) + I \left(1 - \frac{\omega_{*e}}{\omega}\right) \left(\frac{\omega}{\omega_{*i}} + \frac{\beta}{2}\right) \right) + \frac{\omega}{|k|v_i} \left(1 - \frac{\omega_{*i}}{\omega}\right) Z(\zeta_i) \left[1 - \frac{\omega}{\omega_{*i}} + I \left(\frac{\omega - \omega_{*e}}{\omega_{*i}}\right)\right] = 0 \quad (7)$$

with the exact, marginally stable solution $\omega = \omega_{*i}$. The resonant electron contributions therefore exactly cancel in the long wavelength limit.

We present a numerical solution of (1) and (3) for several values of β to further illustrate the shift to long wavelength. Figure 1 is a plot of γ/Ω_i vs. $k\rho_{es}$ for $T_e/T_i = 1.0$,

$\rho_i/L_n = 2.0$, $\omega_{pi}^2/\Omega_i^2 = 2.0 \times 10^5$ and $\beta = 0.0, 1.0, 1.7$ and 2.0 , in a hydrogen plasma ($m_i = 1836 m_e$). Although not shown in Fig. 1, the growth rates for $\beta = 0.0$ and 1.0 maximize at $k\rho_{es} = 1.6$ with growth rates of $\gamma_m/\Omega_i \approx 13.3$ and $\gamma_m/\Omega_i \approx 7.4$, respectively, where the subscript m denotes maximum growth. For $\beta = 1.7$ the growth rate maximizes at $k\rho_{es} \approx 1.0$, and the shift to longer wavelengths is apparent. The mode is stable, i.e., $\gamma < 0$, for $k\rho_{es} > 1.6$. Also, note that $\gamma_m/\Omega_i < 1$ in this case and, therefore, the assumption of unmagnetized ions is no longer valid for $\beta = 1.7$. The lower-hybrid-drift instability is effectively "stabilized" at this value of β (this interpretation is consistent with Fig. 7 of

Ref. 3). However, the plasma may still be unstable to the drift-cyclotron instability since the lower-hybrid-drift instability makes a transition to the drift-cyclotron instability when $\gamma/\Omega_i < 1$.^{6,7} Finally, for $\beta = 2.0$ we still find instability with $\gamma_m/\Omega_i = 9.0 \times 10^{-5}$ at $k\rho_{es} = 0.17$, a significant shift to longer wavelength. The mode is stable for $k\rho_{es} > 0.19$, but approaches marginal stability as $k\rho_{es} \rightarrow 0$.

In conclusion, we have re-examined the stability properties of the lower-hybrid-drift instability in finite β plasmas. In contrast to previous results,³ we find that finite β does not stabilize the instability in the sense that the growth rate becomes negative (i.e., $\gamma < 0$). Rather, as β increases, the instability shifts to longer wavelengths and is driven towards marginal stability (i.e., $\gamma \rightarrow +0$). This has been demonstrated both analytically and numerically. To a certain extent, this result is only of academic interest. In a magnetized plasma, one with B finite everywhere, the lower-hybrid-drift instability is "stable" when $\gamma < \Omega_i$, i.e., it transforms into the drift-cyclotron instability.^{6,7} In this regard, the critical β curve (Fig. 7) presented in Ref. 3 should be interpreted as the β required for this to occur, that is, $\gamma = \Omega_i$ for $\beta = \beta_{cr}$. We also emphasize that this conclusion is based upon local theory. For example, the spatial extent Δx of lower-hybrid-drift modes in a reversed field plasma is given by $\Delta x \sim (L_n/k)^{1/2}$ so that long wavelength modes become global.⁸ Also, in a reversed field plasma, $\rho_e \rightarrow \infty$ in the vicinity of the reversal region, the cancellation of the electron VB resonant damping, which occurs

when $kp_e \rightarrow 0$, should not be expected. The long wavelength modes should therefore be damped by resonant electrons near $B = 0$. We are currently investigating this effect.

Finally, we comment that the stability properties of long wavelength lower-hybrid-drift modes can be quite important in developing models to describe the nonlinear saturation of lower-hybrid-drift turbulence. It has been shown that two dimensional lower-hybrid-drift turbulence ($\mathbf{k} \cdot \mathbf{B} = 0$) is described by the Hasegawa-Mima equation, where the roles of the ions and electrons are exchanged in comparison with universal drift mode turbulence.⁹ This equation describes the inverse cascade of wave energy from short to long wavelength so that the stability of long wavelength modes is important in understanding the ultimate dissipation of the wave energy.

Acknowledgments

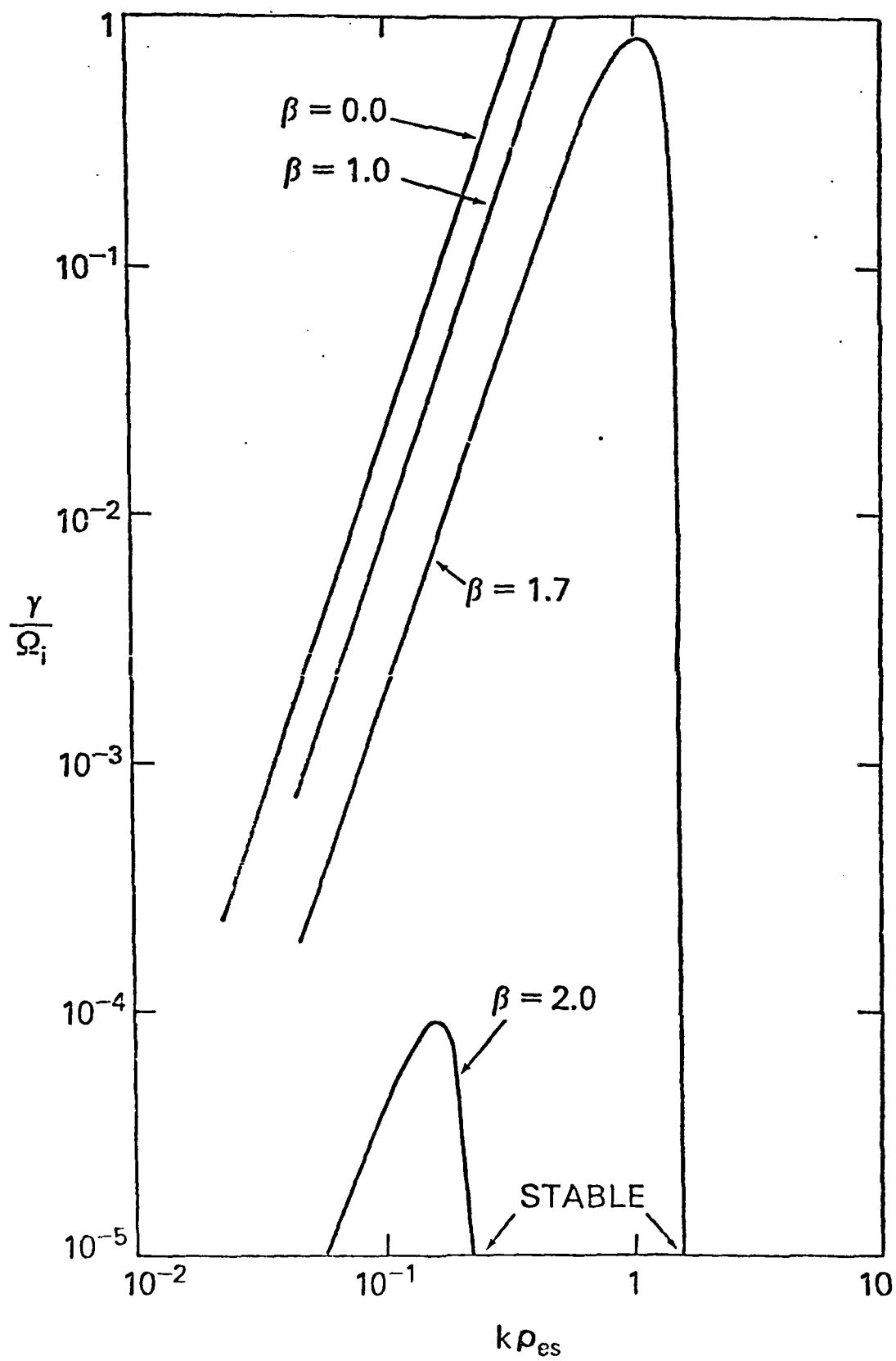
This research has been supported by the Office of Naval Research and NASA.

References

1. N.A. Krall and P.C. Liewer, Phys. Rev. A4, 2094 (1971).
2. R.C. Davidson and N.T. Gladd, Phys. Fluids 18, 1327 (1977).
3. R.C. Davidson, N.T. Gladd, C.S. Wu, and J.D. Huba, Phys. Fluids 20, 301 (1977).
4. N.A. Krall and L.D. Pearlstein, in Plasma Physics and Controlled Nuclear Fusion Research (International Atomic Energy Agency, Vienna, 1965), p. 735.
5. H.L. Berk and R.R. Dominguez, J. Plasma Phys. 18, 31 (1977).
6. J.P. Freidberg and R.A. Gerwin, Phys. Fluids 20, 1311 (1977).
7. N.T. Gladd and J.D. Huba, Phys., Fluids 22, 911 (1979).
8. J.D. Huba, J.F. Drake and N.T. Gladd, Phys. Fluids 23, 552 (1980).
9. J.F. Drake and J.D. Huba, Sherwood Theory Meeting, Austin, Texas, 1981.

Figure Caption

1. Plot of γ/Ω_i vs. $k\rho_{es}$ for $T_e/T_i = 1.0$, $\rho_i/L_n = 2.0$,
 $\omega_{pi}^2/\Omega_i^2 = 2.0 \times 10^5$, $m_i/m_e = 1836$, and $\beta = 0.0, 1.0, 1.7,$
2.0. Note that $\gamma/\Omega_i < 1$ for $\beta = 1.7$, and the significant
shift to longer wavelengths of maximum growth as β increases
from 1.7 to 2.0.



APPENDIX E

FINITE WIDTH CURRENTS, MAGNETIC SHEAR AND THE
CURRENT DRIVEN ION CYCLOTRON INSTABILITY

Finite Width Currents, Magnetic Shear and the
Current Driven Ion Cyclotron Instability

P. Bakshi
Physics Department
Boston College*
Chestnut Hill, MA 02167

and

Naval Research Laboratory
Washington, D.C. 20375

G. Ganguli
Science Applications, Inc.
McLean, VA 22102

and

P. Palmadesso
Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

Abstract

Our earlier results that non-local effects due to even a small magnetic shear produce a significant reduction of the growth rate of the ion cyclotron instability driven by a uniform current are now generalized to finite width currents. Externally prescribed as well as self-consistent shears are considered. If the current width L_c exceeds the shear length L_s , the previous results are recovered. Shear becomes less effective with reduction of L_c , and for typical parameters, the growth rate attains its (shearless) local value for $L_c/L_s \lesssim 10^{-2}$. Non-local effects of the finite current width itself come into play if L_c is further reduced to a few ion Larmor radii and can quench the instability.

In our recent study¹ on the effects of magnetic shear on the current driven ion cyclotron instability (CDICI), we obtained the intriguing result that even a small shear could produce a significant reduction in the growth rate. The transition from the local to the non-local treatment essentially induces a singular perturbation, (the small shear being the coefficient of the highest — here the second — derivative term), and there is also a change in the boundary conditions, from plane waves to a bounded solution. It was thus possible to understand and justify the sharp change in the growth rate from the local to the non-local treatment.

However, it was assumed in the treatment above that both the shear and the current were uniform in space. In laboratory as well as space plasmas on the other hand, one necessarily deals with finite width currents, so that the scale length of the current profile is an additional parameter which could play a role in determining the precise nature of the transition from the local to the non-local treatment. One can see, a priori, that if the current width scale length L_c is much larger than the shear length L_s , the effects due to shear should prevail unhindered. In the other limit, when L_c is much smaller than L_s , it is also clear that the total change in the angle of the magnetic field vector in traversing the slab will be so small that the non-local effects due to shear may not come into play, and the local growth rate should prevail, apart from any non-local effects due to the variation of the current profile itself. Thus a transition from the local to the non-local results will occur as the current width is increased from a value much smaller than the shear length to a value larger than the shear length; this transition can be expected to be primarily governed by the ratio of these scale lengths, and a quantitative study of this phenomenon is the main objective of this Letter.

An additional feature to be considered is the relationship between the current profile and the shear. If the shear is entirely generated by the driving current, we will call it the self-consistent shear. If the shear is generated by (stronger) external currents which do not drive the instability, we will call it the prescribed shear. These two cases could lead to slightly different results. Examples of one situation or the other can be found in various plasmas of practical importance, and there are also situations where the total shear derives from both the mechanisms.

With the introduction of a finite slab, the problem can no longer be translated to a convenient origin where the given wave vector is perpendicular to the local field line. Instead, we will take the center of the slab as the origin and introduce explicitly $k_{\parallel}^0$, the parallel component of the wave vector, as a prescribed parameter. One must maximize over all possible $k_{\parallel}^0$, even in the non-local theory, to obtain the optimal normal mode solution. We introduce a finite current width along the x direction by taking the electron distribution function to be a shifted Maxwellian with a drift velocity parallel to the magnetic field to be

$$V_d(x) = V_d^0 g(x_g/L_c), \quad g(\xi) = e^{-\xi^2}, \quad x_g = x + \frac{v_y}{\Omega_e}. \quad (1a)$$

The Vlasov equation is satisfied since x_g , the guiding centre position is a constant of the motion, v_y being the electron velocity component perpendicular to the magnetic field direction z and Ω_e is the electron gyrofrequency. A very small correction in the constant of motion due to shear can be ignored here. The current profile is then found to be

$$j(x) = n_0 e V_d^0 \exp[-x^2/(L_c^2 + \rho_e^2)] \approx n_0 e V_d^0 g(x/L_c), \quad (1b)$$

where $\rho_e^2 \ll L_c^2$ can be ignored; this is also equivalent to replacing x_g by x in equation (1a). This current profile will generate a self-consistent shear in the magnetic field, given in terms of $\xi = \frac{x}{L_c}$ by

$$B_y(x) = \frac{4\pi}{c} \int_0^x j(x) dx = \frac{4\pi n_0 e v_d^0 L_c}{c} \int_0^{\xi} g(\xi) d\xi, \quad (2)$$

$$B_y(x)/B_z = (L_c/L_s) \int_0^{\xi} g(\xi) d\xi,$$

$$(1/s) = L_s = \frac{c B_z}{4\pi n_0 e v_d^0}, \quad (3)$$

and a corresponding variation of the parallel wavenumber,

$$k_{\parallel}(x) = k_{\parallel}^0 + k_y B_y(x)/B_z, \quad (4a)$$

or

$$u(x) = k_{\parallel}(x)/k_y = u_0 + (L_c/L_s) \int_0^{\xi} g(\xi) d\xi. \quad (4b)$$

The shear length is determined by the physical parameters in eq. (3).

Following the usual procedure for introducing the non-local effects,¹ we let $k_x \rightarrow (1/i)d/dx$, and assuming that the higher derivative terms are less important, we obtain

$$[\rho_i^2 \frac{d^2}{dx^2} + Q(u(x), v_d(x))]\phi = 0, \quad (5)$$

or

$$[(\frac{\rho_i}{L_c})^2 \frac{d^2}{d\xi^2} + Q(u, v_d)]\phi = 0,$$

where ρ_i is the ion Larmor radius. Apart from the change in definitions of $k_{||}$, u and v_d , to include the space variation of the current, Q is given by

$$Q = - \rho_i^2 Q_1 / A \quad (6)$$

where Q_1 and A are described by eqs. (6) and (7) of Ref. 1. In general, it would be necessary to solve (5) by a numerical method such as the Numerov shooting code technique. The case of pure shear, without any effects due to finite current width ($v_d(x) \equiv v_d^0$), was discussed in Ref. 1 by expanding Q to second order around u_0 , the angle of maximum growth in local theory given by $Q(u_0) = 0$. The resulting Weber equation provides a dispersion relation for the determination of the eigenmode frequency. When anharmonic terms become important, the error involved in approximating Q by a quadratic form around u_0 will become significant. An improved procedure is to consider an expansion around the (complex) position ξ_1 where the derivative of Q with respect to ξ vanishes,

$$Q'(\omega, \xi_1) = 0. \quad (7)$$

Then the dispersion relation (analogous to eq. (18) of Ref. 1), reduces to

$$Q(\omega, \xi_1) = (2\ell + 1) (\rho_i/L_c) \left[-\frac{1}{2} Q''(\omega, \xi_1) \right]^{1/2}, \quad (8)$$

where ℓ is the mode number. Equations (7) and (8) are two complex equations for the simultaneous determination of the complex eigenfrequency ω and the complex angle u_1 , corresponding to the complex position ξ_1 . We have tested the results obtained from eqs. (7) and (8) against the numerical shooting code solution of eq. (5) in a variety of cases, and the results are in excellent agreement for the eigenfrequencies as well as the solutions for the potential ϕ .

Considering $Q(u, v)$ as a function of two variables u and v , each a function of ξ , the explicit expressions for the derivatives are

$$Q' = (L_c/L_s)g Q_u - 2\xi(v_d^0/v_e)Q_v, \quad (9a)$$

$$Q'' = (L_c/L_s)^2 g^2 Q_{uu} - 2(L_c/L_s)\xi g Q_u \quad (9b)$$

$$- 4 (L_c/L_s)(v_d^0/v_e)\xi g^2 Q_{uv}$$

$$- 2g(1-2\xi^2)(v_d^0/v_e)Q_v + 4(v_d^0/v_e)^2 \xi^2 g^2 Q_{vv}.$$

We have solved eqs. (7) and (8) for a range of slab widths $(L_c/\rho_i) = 10$ to 10^7 and shear lengths $(L_s/\rho_i) = 10^2$ to 10^6 . The results are

displayed for the $\ell = 0$ mode in Figures 1-5. The basic transition from the local to the non-local results is evident in Figure 1, which depicts the growth rate in units of ion Larmor frequency, (γ/Ω_i) as a function of the width of the current slab expressed in units of ion Larmor radius, (L_c/ρ_i) . Typical parameters are the ion to electron mass ratio $\mu = 1837$, the electron drift to electron thermal velocity ratio $v_d^0/v_e = 0.28$, the ion to electron temperature ratio $\tau = (T_i/T_e) = 0.5$ and the transverse wave number given by $b = (1/2) k_y^2 \rho_i^2 = 0.6$. Three different shear lengths are represented by $(L_s/\rho_i) = 10^6, 10^5$ and 10^4 . In each case, when L_c is sufficiently large, one obtains the non-local growth rate $\gamma/\Omega_i = 0.0127$. Since the shear strengths considered here are all small, we find the same non-local value for γ . For a stronger shear $(\rho_i/L_s) = 10^{-2}$, (not shown in Fig. 1), a lower non-local value is obtained, $(\gamma/\Omega_i) = 0.0115$, and still lower non-local growth rates would be attained if the shear is further increased. In the other limit, when L_c is sufficiently small, the local growth rate prevails, $(\gamma/\Omega_i) = 0.0296$. The transition curves for different shear values look very similar, apart from a shift in L_c , and indeed as shown in the next figure, possess a universal behavior as a function of L_c/L_s . Another interesting feature in Fig. 1 is the decrease in growth rate for $L_c < 10^2 \rho_i$, which arises from the nonlocal effect of the space variation of the current itself. The slab width is too small to obtain any significant variation of the direction of the magnetic field due to shear, and thus this reduction in growth rate is independent of the shear parameter. If the current width is further reduced to just a few ion Larmor radii, the instability is completely quenched. We can call this the

filamental quenching -- current filaments only a few ion Larmor radii in width cannot sustain the instability. This feature may be of practical significance in Q-machine experiments.

Figure 2a represents the same information as in Fig. 1, but now plotted as a function of (L_c/L_s) . Now the transition from the local to the non-local result is seen to be universal, almost independent of the shear parameter. Of course, shear values with $(\rho_i/L_s) \gtrsim 10^{-2}$ will lead to lower non-local values with corresponding, slight, departures from the universal transition curve. For small $(L_c/L_s) < 10^{-2}$, the case of very small shear, $\rho_i/L_s = 10^{-6}$, retains the local growth rate for the entire range considered, but the case of larger shear, $\rho_i/L_s = 10^{-3}$, falls in the domain $L_c/\rho_i < 10$ and the growth rate is reduced due to filamental quenching. Figure 2b represents the same parameters, except for $b = 1.2$ and essentially describes the same phenomena as in Fig. 2a. We note that both the local and the non-local limit growth rates are lower than those for $b = 0.6$.

A typical solution for ϕ is given in Figure 3, for the parameters of Fig. 1 in the infinite slab limit, $L_c/\rho_i \rightarrow \infty$. The shear parameter is $\rho_i/L_s = 10^{-2}$ and the solutions obtained by the analytical method described here or by a direct numerical solution of the differential equation were found to be identical. The solution is centered around $u_{m0} \approx 0.09$ and has a width of $\Delta u \approx 0.01$. The angular width decreases for lower shear, without any significant change in the center of $|\phi|$. Figures 4a and 4b respectively represent the variations in u_1 , the complex angle around which the wave packet for the electrostatic potential ϕ is formed, and u_{m0} the real angle where $|\phi|$ has its maximum along the real angle axis. The angle u_{m0} is experimentally measurable in the sense of normal mode dominance at

that angle as measured from the local magnetic field orientation. For the typical case $b = 0.6$, we note the progression of u_1 from 0.13 in the local limit to $(0.099) - i(0.043)$ in the non-local (shear dominant) limit and a corresponding transition in u_{m0} from 0.13 to 0.09. These curves are drawn for $\rho_1/L_s = 10^{-6}$, but they represent quite accurately the results for higher shears, because of the universality of the transition.

Figure 5 provides the variation of γ as a function of b for various L_c and L_s combinations. The top curve is the local result obtained directly from $Q = 0$; this corresponds to the case of a uniform magnetic field (zero shear). If we introduce a small shear $L_s/\rho_1 = 10^6$ and a comparable current width $L_c/\rho_1 = 10^6$, we essentially reach the shear dominant non-local result discussed in Ref. 1. The transition is shown, with $L_c/\rho_1 = 10^5$ and 10^4 ; the latter case is already good enough to almost approximate the local result. Thus we make the transition from the local to the non-local result as L_c/L_s increases from 10^{-2} to 1. The lowest curve is for $L_c = L_s = 10^2\rho_1$ and the difference from the curve above essentially represents the additional effect of the increased shear.

If we use a prescribed rather than the self-consistent shear, the relation between the angle and the physical distance is simpler,

$$\begin{aligned} u(x) &= u_0 + (x/L_s). \\ &= u_0 + (L_c/L_s)\xi \end{aligned} \tag{10}$$

It can be easily seen that (4b) reduces to (10) if ξ remains small, a situation which is valid when $L_c > L_s$, the case when the shear effects dominate. We have also solved eqs. (7) and (8), using eq. (10), and the results do not significantly differ from the case of self-consistent shear.

We have shown in this Letter that the non-local effect of shear is tempered by the finite size of the current slab: (1) For $L_c > L_s$ the shearing of the magnetic field lines across the length of the slab is sufficient to allow the non-local, singular perturbation effects to take place, and the results of Ref. 1 are recovered; (2) For $L_c < L_s$ on the other hand, the finite slab size is smaller than that needed to provide sufficient angular space to form a parabolic domain for Q_I , the imaginary part of Q , of the type of Fig. 5 of Ref. 1. Thus the shear dominated non-local wave packet formation cannot occur unhindered and for $L_c < 10^{-2} L_s$, the local results are recovered; (3) A new effect occurs for $L_c < 10 \rho_i$, due to the variation of the current itself; thus filamental currents of only a few ion gyroradii may not produce any instability.

The results obtained here are quite significant for space-plasma applications where small shear and finite width current channels obtain in the auroral arc regions. Implications for various space and laboratory plasmas will be described elsewhere.

Acknowledgements

We wish to thank Dr. J. Fedder, Dr. J. Finn, Dr. P. Guzdar, Dr. J. Huba and Dr. H. Mitchell for many valuable discussions during the course of this work. One of us (P.B.) would like to thank Dr. S. Ossakow for the kind hospitality of the Naval Research Laboratory where this work was done under the Navy/ASEE Summer Faculty Research Program. The other two authors acknowledge support of this work by the Office of Naval Research and by the National Aeronautics and Space Administration.

References

* Permanent address.

1. G. Ganguli and P. Bakshi, Phys. Fluids, 25, 1830 (1982).

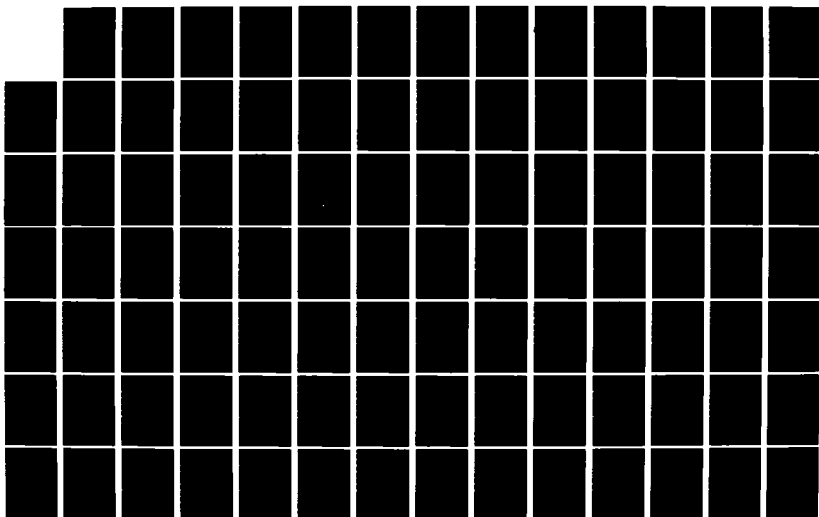
AD-A126 341

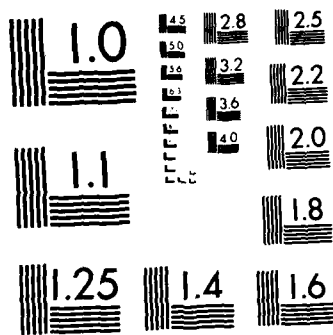
GEOPHYSICAL PLASMAS AND ATMOSPHERIC MODELING(U) SCIENCE 2/7
APPLICATIONS INC MCLEAN VA E HYMAN ET AL. FEB 83
SAI-83-144-WA N00014-81-C-2038

UNCLASSIFIED

F/G 4/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Figure Captions:

Figure 1 A plot of the growth rate $(\frac{\gamma}{\Omega_1})$ against the current channel width (L_c/ρ_i) for three different magnetic shear lengths. Here $V_d^0 = 0.28 v_e$, $\mu = 1837$, $b = 0.6$, and $\tau = 0.5$.

Figure 2a A plot of (γ/Ω_1) against (L_c/L_s) for three different L_s values. Here $b = 0.6$, $\mu = 1837$, $\tau = 0.5$, and $V_d^0 = 0.28 v_e$. The transition from the local to the nonlocal results are universal, almost independent of the shear value.

Figure 2b A plot similar to Figure (2a) except that the value of b here is 1.2.

Figure 3 $\phi(u)$ against u for $L_c/\rho_i = \infty$, and $(\rho_i s) = 10^{-2}$, $b = 0.6$, $V_d^0 = 0.28 v_e$, and $\tau = 0.5$.

Figure 4a A plot for the complex angle u_1 for various values of L_c/L_s . Here $b = 0.6$, $\mu = 1837$, $V_d^0 = 0.28 v_e$, and $\tau = 0.5$.

Figure 4b A plot for the real angle U_{m0} where $|\phi(u)|$ attains its maximum, against L_c/L_s . The rest of the parameters are same as Figure 4a.

Figure 5 A plot of the growth rates against b . Here $\tau = 0.5$, $\mu = 1837$, $V_d^0 = 0.28 v_e$. The top solid line is a solution of the local dispersion relation. The second solid line is for $L_c = L_s = 10^6 \rho_i$. The dotted lines show the transition from local to nonlocal values as L_c approaches L_s . The lowest solid line is for $L_c = L_s = 10^2 \rho_i$.

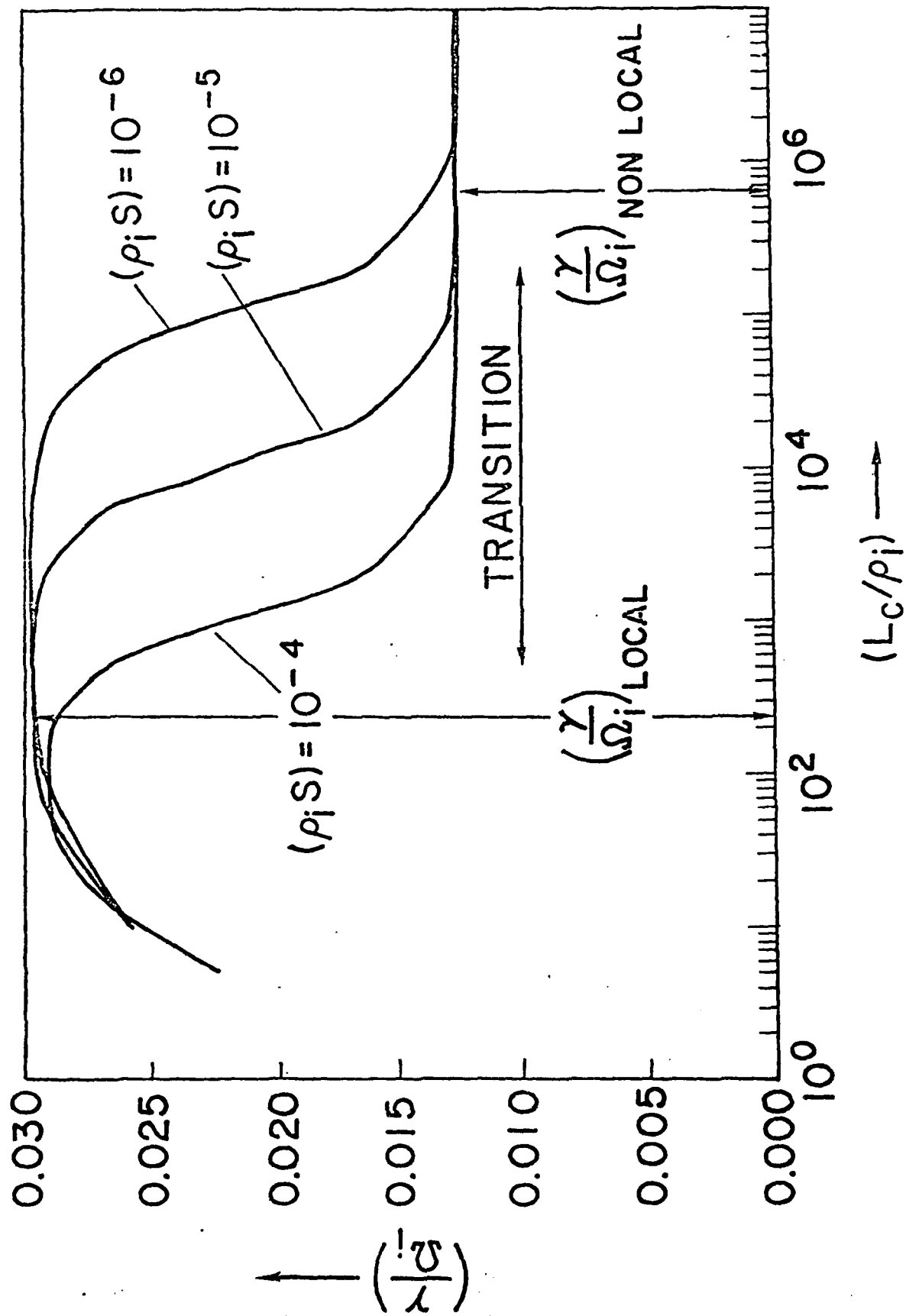


Fig. 1

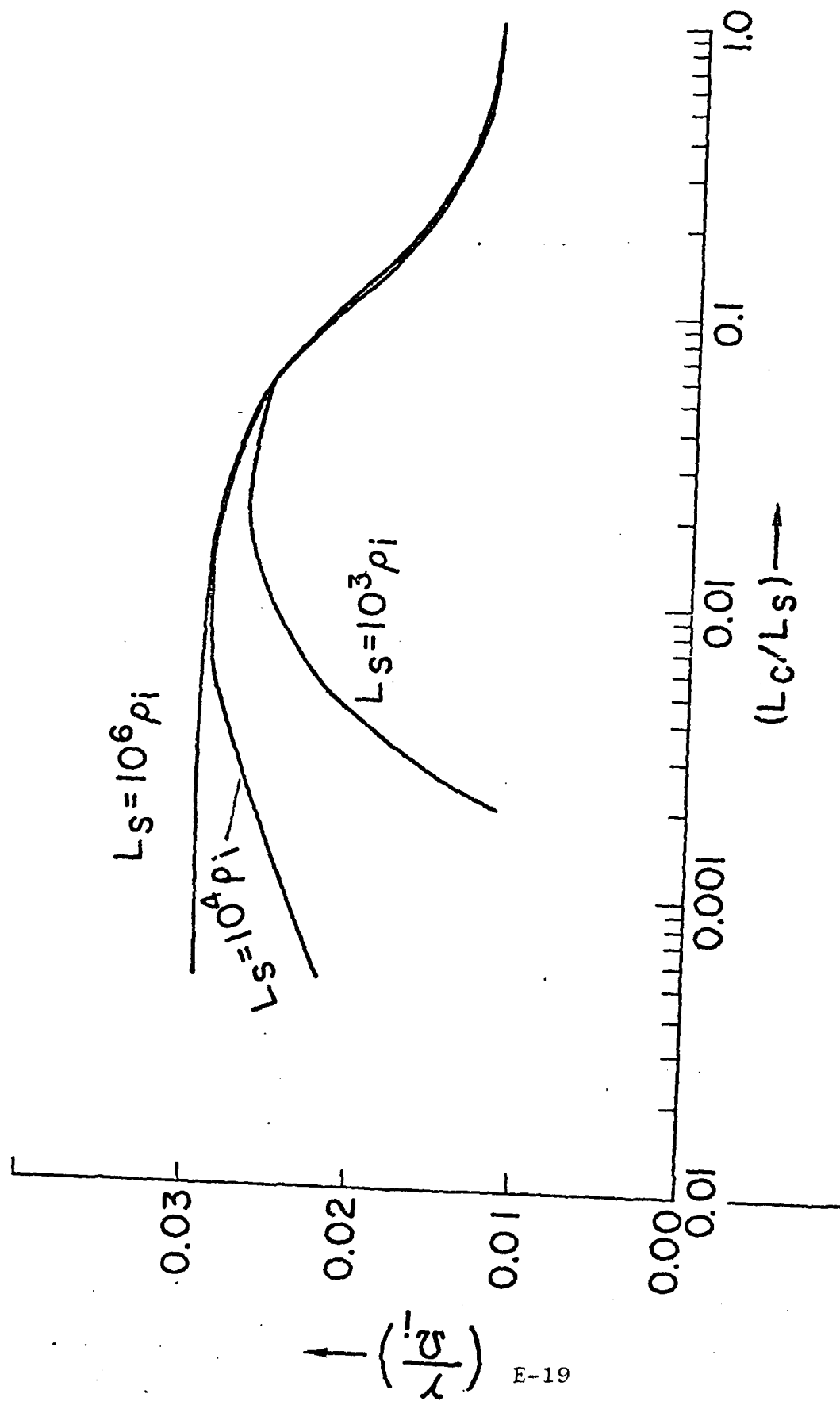


Fig. 2a

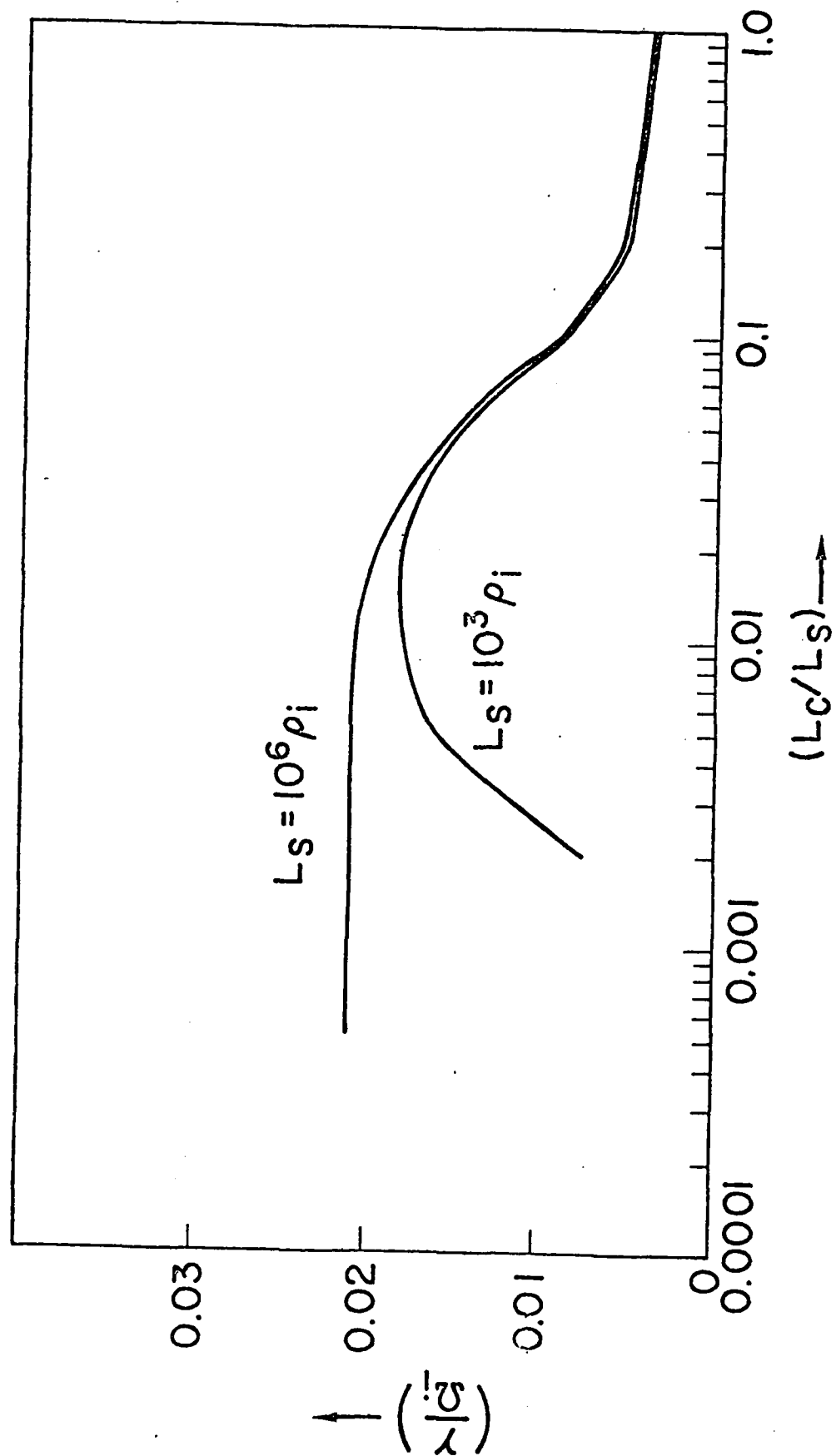
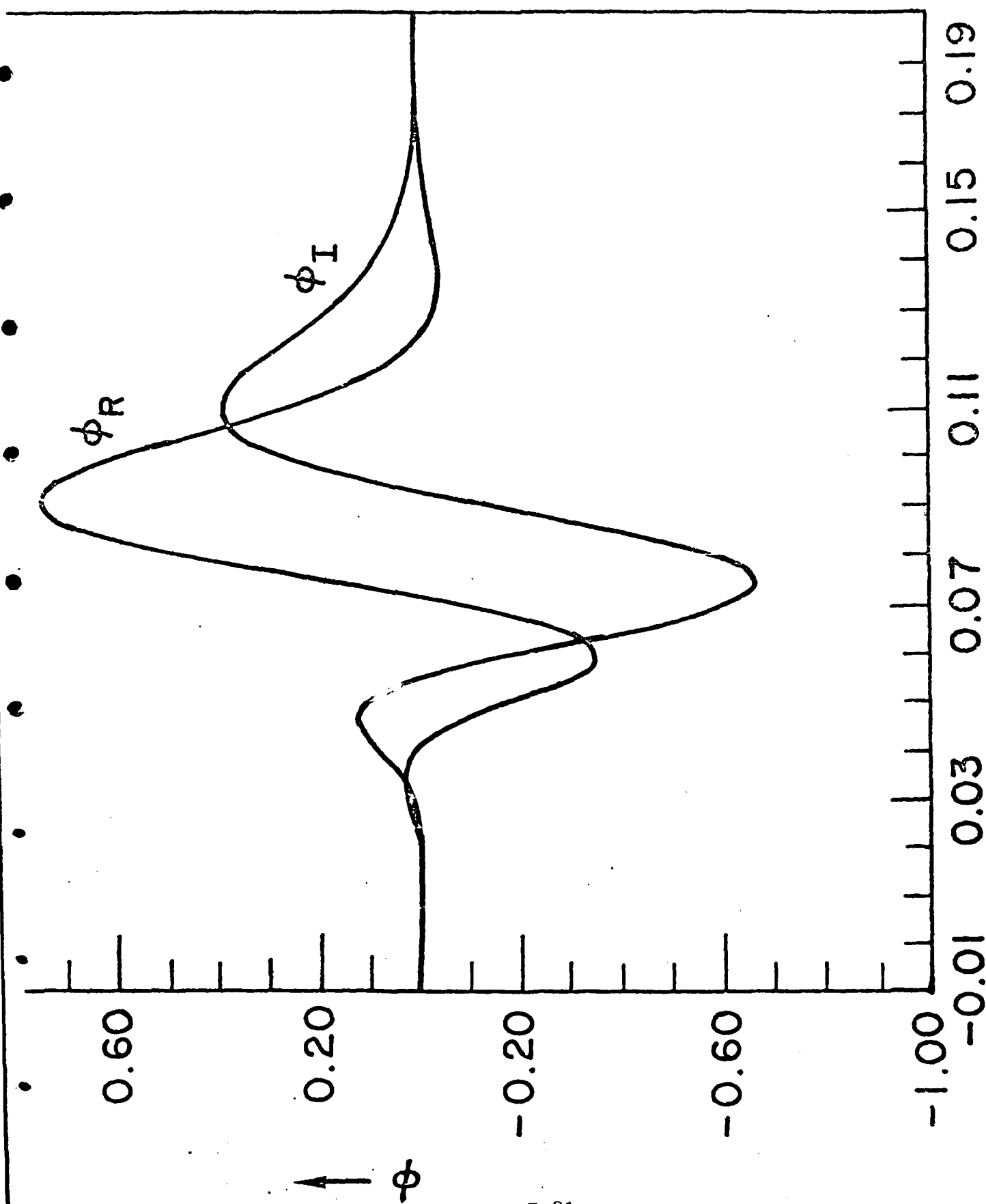


Fig. 2b



U $\longrightarrow$ Fig. 3

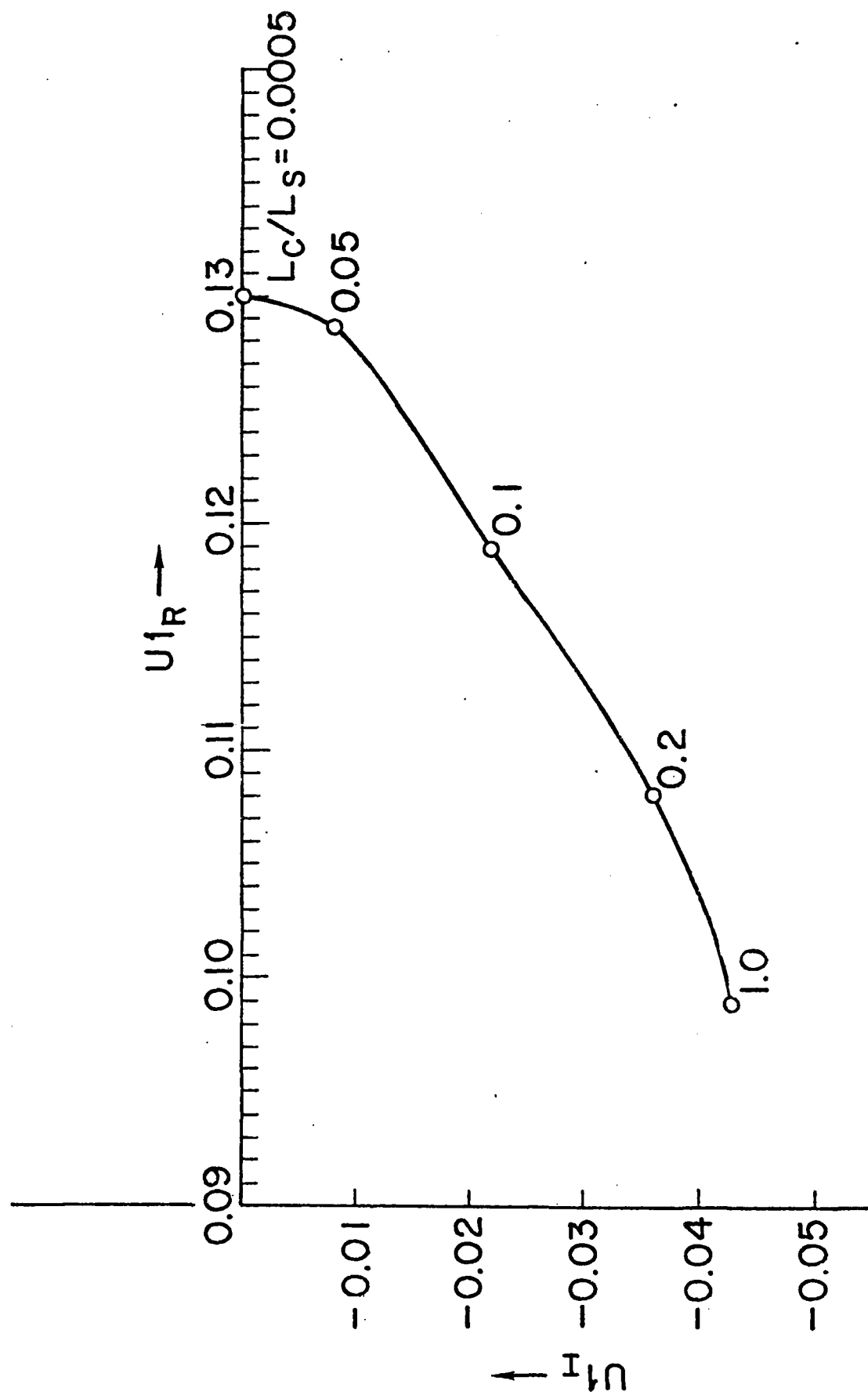


Fig. 4a

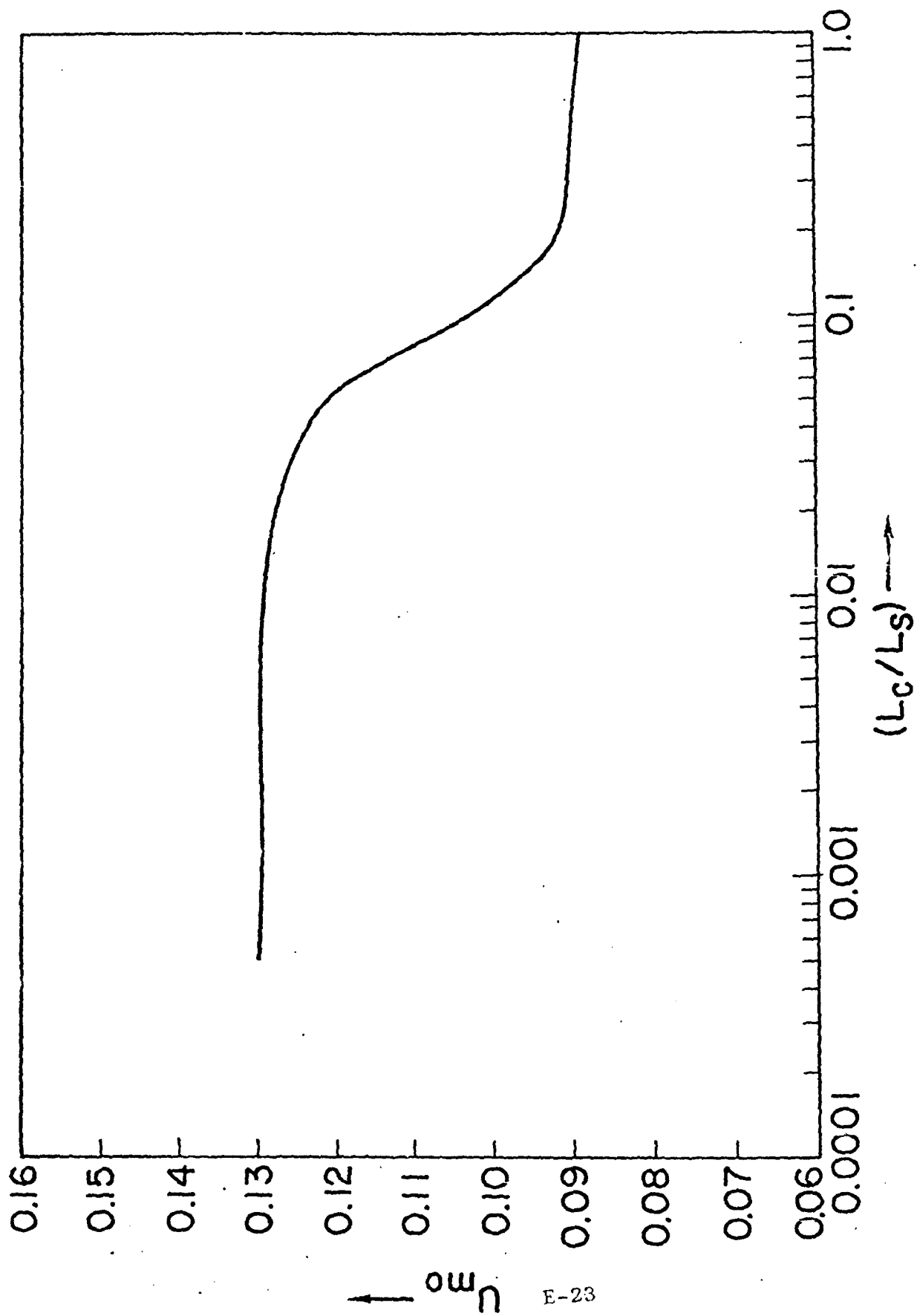


Fig. 4b

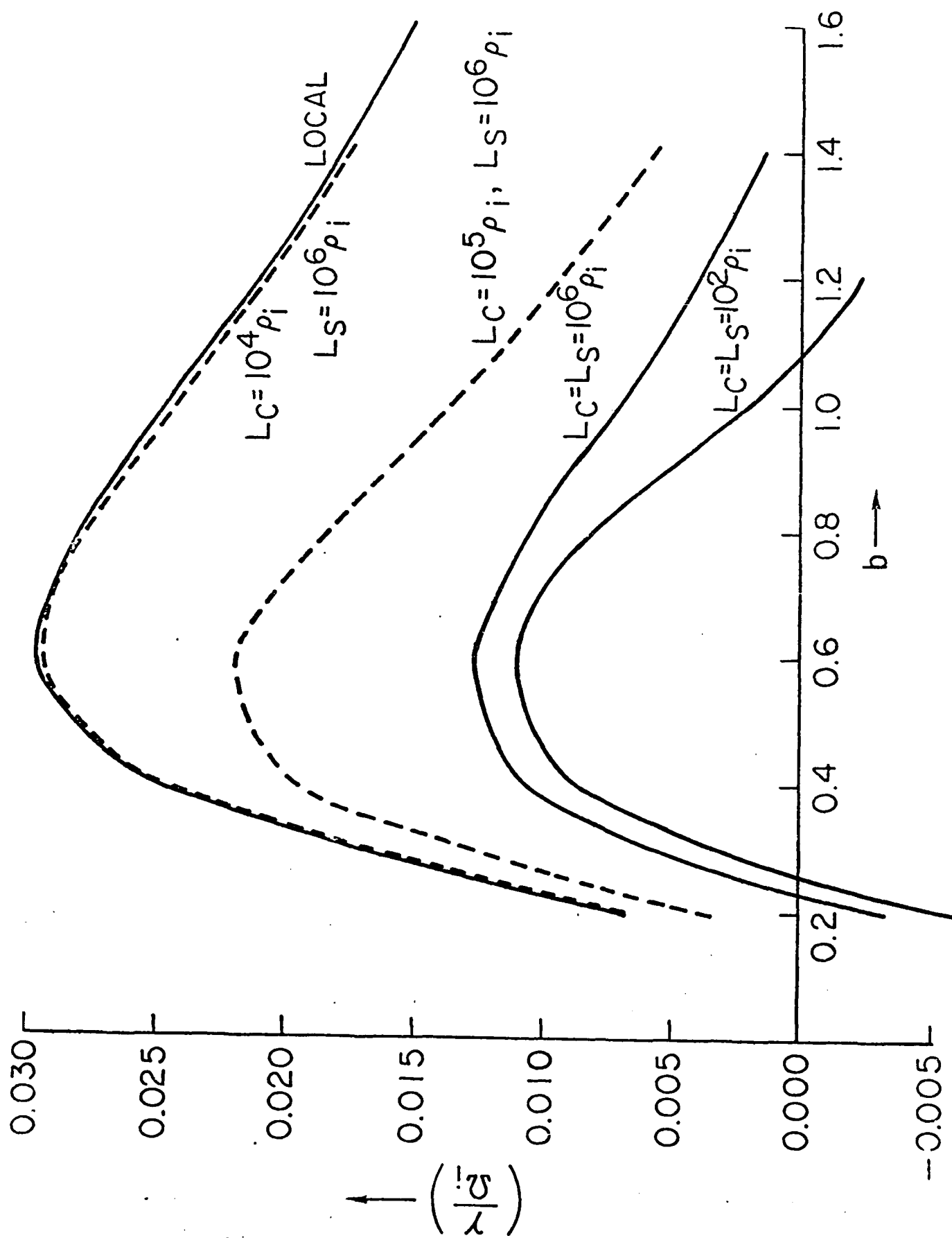


Fig. 5

APPENDIX F

EFFECT OF ELECTRON BEAMS ON
THE CURRENT CONVECTIVE INSTABILITY

Appendix F
EFFECT OF ELECTRON BEAMS ON THE CURRENT
CONVECTIVE INSTABILITY

We have studied the effects of an electron beam on the CCI with application to the diffuse auroral situation. The weak field aligned currents, in association with transverse density gradients, would result in exciting slowly growing large scale size irregularities in the medium. The presence of such irregularities has been associated with scintillations of the DNA wideband satellite. Other mechanisms could also play an important role in the generation of irregularities, in particular the $\underline{ExB}$ instability. It is generally known that high latitudes are characterized by soft particle precipitations during diffuse auroral conditions. Thus, it is of interest to investigate what effects such precipitating particles may have on processes much slower than their own time scales. We have found that, depending on the relative directions of the precipitating particle beams and the current (carried predominately by cold drifting electrons), the growth of the CCI may be enhanced (when the two are in the same direction) or reduced (when the two are antiparallel to each other). But for the ionospheric situation at high latitudes, these effects are essentially negligible.

We consider the general equilibrium as follows: the cold electrons drift with respect to ions with a uniform drift velocity V_d along the magnetic field ($B_0 z$) and there is a density gradient in the ambient plasma along the y-axis (northwards). The homogeneous electron beam (soft particle precipitation) has a velocity $V_b z$. Then, on following the

linear stability analysis of Ossakow and Chaturvedi (1979), one finds that the dispersion relation for the perturbations, now, is

$$\left[\omega \left\{ \frac{k_z^2}{k_\perp^2} \left(\frac{\Omega_i}{v_{in}} + \frac{\Omega_e}{v_{ecn}} \right) + \frac{v_{in}}{\Omega_i} \right\} + k_z v_d \left(i \frac{k_\perp^2 \hat{x} \cdot \nabla_\perp n_0}{k_\perp^2 n_0} - \frac{\Omega_i}{v_{in}} \frac{k_z^2}{k_\perp^2} - \frac{v_{in}}{\Omega_i} \right) \right] [\omega - k_z v_b] \approx - \frac{n_{ob}}{n_{oc}} \omega (\omega - k_z v_d) \left(\frac{k_z^2}{k_\perp^2} \frac{\Omega_e}{v_{ebn}} \right)$$

where ω is the frequency and k the wave number of the mode. Other symbols have their standard meanings, Ω_α is the cyclotron frequency, $v_{\alpha n}$ is the particle collision frequency with neutrals, subscripts z , and $\perp$, refer to parallel and perpendicular with respect to the magnetic field, $\nabla_\perp n_0$ is the density gradient in ambient plasma and n_{ob} is the beam plasma density. In the absence of an electron beam ($n_{ob}=0$, $v_b=0$), it can be readily verified that one recovers the results of Ossakow and Chaturvedi [1979] for the linear growth rate of the current convective instability.

Splitting $\omega = \omega_R + i\gamma$, one finds

$$\gamma = \gamma_c \left[1 + \frac{\frac{n_{ob}}{n_{oc}} \frac{v_d}{v_b} \frac{k_z^2}{k_\perp^2} \frac{v_{ecn}}{v_{ebn}}}{\left\{ \frac{k_z^2}{k_\perp^2} \left(1 + \frac{\Omega_i}{v_{in}} \frac{v_{ecn}}{\Omega_c} \right) + \frac{v_{ecn} v_{in}}{\Omega_e \Omega_i} \right\}} \right]$$

where

$$\gamma_c \approx \frac{-k_z v_d \left(1 + \frac{v_{ecn}}{\Omega_e} \frac{k_z x z \cdot \nabla_{\perp} n_{oc}}{k_{\perp}^2 n_{oc}} \right)}{\frac{k_z^2}{k_{\perp}^2} \left(1 + \frac{\Omega_i}{v_{in}} \frac{v_{ecn}}{\Omega_e} \right) + \frac{v_{ecn} v_{in}}{\Omega_e \Omega_i}}$$

is the current convective instability growth rate in the absence of an electron beam. [Ossakow and Chaturvedi, 1979]. We find that if the beam velocity and cold electron drift are in the same direction, the new growth rate for the modes is enhanced. On the other hand, the reverse is true if the electron beam velocity were anti-parallel to the cold electron drift direction. However, for typical ionospheric parameters at F-region altitudes at high latitudes, $\frac{n_{ob}}{n_{oc}} \sim 10^{-2}$, $v_b \sim 1 \times 10^8 \text{ cms}^{-1}$ (corresponding to electron energy fluxes of energy $\sim 100 \text{ ev}$), $v_d \sim 10^5 \text{ cms}^{-1}$, one finds that

$$\gamma \approx \gamma_c [1 + O(10^{-5})]$$

Thus we find that, for practical purposes, the presence of high velocity electron beams leaves the growth rate of the current convective instability in diffuse aurora unaffected.

References

- S. L. Ossakow and P. K. Chaturvedi, Geophys. Res. Letts, 6, 332, 1979.

APPENDIX G

PARAMETRIC EXCITATION AND SUPPRESSION
OF CONVECTIVE PLASMA INSTABILITIES
IN THE HIGH LATITUDE F-REGION IONOSPHERE

Parametric Excitation and Suppression
of Convective Plasma Instabilities
in the High Latitude F-Region Ionosphere

M.J. Keskinen,¹ P.K. Chaturvedi,² S.L. Ossakow¹

¹Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

²Science Applications, Inc.
McLean, VA 22102

Abstract

Parametric coupling processes of a large amplitude electromagnetic pump wave with convective plasma instabilities ($\underline{E} \times \underline{B}$ gradient drift and current convective) in the high latitude F region ionosphere have been studied. We find that a pump wave, with frequency near the upper hybrid frequency and with electric field component perpendicular to the ambient magnetic field, can stabilize or destabilize the $\underline{E} \times \underline{B}$ and/or current convective instability. For parameters typical of the nighttime high latitude F region ionosphere we find stabilization/destabilization with a vertically incident O mode carrying a free space incident energy flux less than or equal to 10^{-4} W/m^2 .

1. INTRODUCTION

Using a variety of experimental techniques, e.g., satellites [Dyson, 1969; Dyson et al., 1974; Sagalyn et al., 1974; Clark and Raitt, 1976; Phelps and Sagalyn, 1976; Rodriquez et al., 1981], rockets [Olesen et al., 1976; Ogawa et al., 1976; Kelley et al., 1980], scintillations [Aarons et al., 1973; Fremouw et al., 1977; Erukhimov et al., 1981], and radar backscatter [Weaver, 1965; Greenwald, 1974; Hower et al., 1966; Vickrey et al., 1980; Hanuise et al., 1981], it is now known that the high latitude ionosphere, from the auroral zone into the polar cap, is a highly structured and nonequilibrium medium containing irregularities (plasma density fluctuations and structures) with scale sizes ranging from hundreds of kilometers to meters. Aside from being an interesting scientific phenomenon, ionospheric irregularities are of practical interest since they can disrupt transionospheric radio wave communications channels (see recent review by Davies [1981] and references therein).

Several theories, e.g., particle precipitation, plasma instabilities and processes, and neutral fluid dynamics have been proposed to account for high latitude ionospheric irregularities (see recent review by Keskinen and Ossakow [1982a] and references therein). Recently, considerable quantitative progress has been made, especially in the area of ionospheric plasma instabilities, in identifying the physical processes that can lead to high latitude irregularities. In particular, convective plasma instabilities such as the $E \times B$ gradient drift instability [Simon, 1963; Linson and Workman, 1970; Keskinen and Ossakow, 1982b,c] and current convective instability [Lehnert, 1958; Kadomtsev and Nedospasov, 1960; Ossakow and Chaturvedi, 1979; Chaturvedi and Ossakow, 1979; Keskinen et al.,

1980; Chaturvedi and Ossakow, 1981] have been invoked to explain large scale (> 1 km) high latitude density irregularities in and near large scale convecting auroral plasma enhancements [Vickrey et al., 1980]. In addition Chaturvedi [1976] has proposed a long wavelength collisional ion cyclotron instability [Drummond and Rosenbluth, 1962; Kindel and Kennel, 1971] to account for irregularities with scale sizes of hundreds of meters to kilometers perpendicular to the magnetic field.

However, it is well known [Silin, 1965; Dubois and Goldman, 1965; Nishikawa, 1968] that under the influence of high-frequency long wavelength electromagnetic fields electrostatic modes of oscillation in a plasma may become parametrically coupled and may grow exponentially in time or space before saturating at large amplitudes. These results have been used by many investigators [see reviews by Perkins et al., 1974, Fejer, 1979 and references therein] to study the role of parametric instabilities, induced by high power ground based radars, in ionospheric modification. In addition, parametric effects have been used to study the possible stabilization or destabilization of plasma instabilities, e.g., drift waves [Fainberg and Shapiro, 1967; Sundaram and Kaw, 1973] by high frequency electric fields. For ionospheric environments near marginal stability Lee et al. [1972] have shown that high frequency electromagnetic waves launched from ground based radars might lead to stabilization or destabilization of the Farley-Buneman and gradient-drift instabilities in the equatorial electrojet E region ionosphere. In this brief note we show that the $\underline{E} \times \underline{B}$ gradient drift and current convective instabilities in the high latitude F region ionosphere might also be stabilized or destabilized by high power

radio waves. In section 2 we outline the theory of the $\underline{E} \times \underline{B}$ gradient drift and current convective instabilities in the presence of a high frequency electro-magnetic pump wave with frequency near the local upper hybrid frequency. We calculate the incident energy flux required to stabilize/destabilize these instabilities. Finally, in section 3 we summarize and discuss our results.

2. THEORY

For wavelengths greater than the ion mean free path we use fluid equations to describe the ion and electron plasma. The following geometry is used: the y-axis is in the north-south direction, the x-axis points west, and the z-axis is along the magnetic field. The set of equations, in the frame of reference in which the neutrals are at rest, become

$$\frac{\partial N_e}{\partial t} + \nabla \cdot (N_e \underline{V}_e) = 0 \quad (1)$$

$$\frac{\partial \underline{V}_e}{\partial t} + (\underline{V}_e \cdot \nabla) \underline{V}_e = - \frac{\nabla P_e}{m_e N_e} - \frac{e}{m_e} \left(\underline{E} + \frac{\underline{V}_e \times \underline{B}}{c} \right) - \underline{v}_{en} \underline{V}_e - \underline{R}_e \quad (2)$$

$$\frac{\partial N_i}{\partial t} + \nabla \cdot (N_i \underline{V}_i) = 0 \quad (3)$$

$$\frac{\partial \underline{V}_i}{\partial t} + (\underline{V}_i \cdot \nabla) \underline{V}_i = - \frac{\nabla P_i}{m_i N_i} + \frac{e}{m_i} \left(\underline{E} + \frac{\underline{V}_i \times \underline{B}}{c} \right) - \underline{v}_{in} \underline{V}_i - \underline{R}_i \quad (4)$$

$$\nabla \cdot \underline{E} = 4\pi e (N_i - N_e) \quad (5)$$

Here N_α ($\alpha = i$ or e) is the species density, $\underline{E} = \underline{E}_0 + \underline{E}_p \cos \omega_0 t - \nabla \phi$ is the total electric field which includes the ambient electric field $\underline{E}_0$, pump field $\underline{E}_p \cos \omega_0 t$, and self-consistent field $\nabla \phi$. In addition, $\underline{R}_e = -\underline{R}_i = \nu_{ei}(\underline{V}_e - \underline{V}_i)$, $\nu_{\alpha} = N_\alpha T_\alpha$, ν_{ei} is the electron-ion collision frequency, $\nu_{\alpha n}$ is the electron or ion neutral collision frequency, c is the speed of light, and T_α the species temperature. Production and loss terms have been neglected. Since we will be considering perturbations with scale sizes much less than the wavelength of the pump wave we take $\underline{E}_p$ to be spatially uniform.

In equilibrium ($\frac{\partial}{\partial t} = 0$) we have

$$\nabla \cdot N_\alpha (\underline{V}_{\alpha 0} + \underline{V}_{\alpha p}) = 0 \quad (6)$$

$$0 = + \frac{e}{m_\alpha} \left(\underline{E}_0 + \frac{\underline{V}_{\alpha 0} \times \underline{B}}{c} \right) - \nu_{\alpha} \underline{V}_{\alpha 0} \quad (7)$$

$$\frac{d}{dt} \underline{V}_{\alpha p} = + \frac{e}{m_\alpha} \left(\underline{E}_p + \frac{\underline{V}_{\alpha p} \times \underline{B}}{c} \right) - \nu_{\alpha} \underline{V}_{\alpha p} \quad (8)$$

where $\alpha = i, e$ and $\underline{V}_{\alpha 0}$, $\underline{V}_{\alpha p}$ refer to the drifts induced by the ambient and pump electric fields, respectively. Linearizing equations (1)-(5) by assuming $N_\alpha = N_{\alpha 0} + \delta n_\alpha$, $n_\alpha \ll N_{\alpha 0}$, $\underline{V}_\alpha = \underline{V}_{\alpha 0} + \delta \underline{V}_\alpha$, $\underline{E} = \underline{E}_0 + \underline{E}_p \cos \omega_0 t + \delta \underline{E}$ we obtain

$$\frac{\partial \delta n_\alpha}{\partial t} + (\underline{V}_{\alpha 0} + \underline{V}_{\alpha p}) \cdot \nabla \delta n_\alpha + \delta \underline{V}_\alpha \cdot \nabla N_\alpha + N_\alpha \nabla \cdot \delta \underline{V}_\alpha = 0 \quad (9)$$

$$\frac{\partial \delta \underline{V}_\alpha}{\partial t} + (\underline{V}_{\alpha 0} + \underline{V}_{\alpha p}) \cdot \nabla \delta \underline{V}_\alpha = + \frac{e}{m_\alpha} \left(\delta \underline{E} + \frac{\delta \underline{V}_\alpha \times \underline{B}}{c} \right)$$

$$- \frac{\nabla \delta p_\alpha}{n_\alpha N_\alpha} - \nu_\alpha \delta \underline{V}_\alpha - \delta \underline{R}_\alpha \quad (10)$$

$$\nabla \cdot \delta \underline{E} = 4\pi e (\delta n_i - \delta n_e) \quad (11)$$

The equations (9)-(11) can be reduced to a single equation (dispersion relation) for one scalar function in six different ways: one may work in either the lab or oscillating frame; and one may eliminate the electron density, eliminate the ion density, or (using Poisson's equation) eliminate both densities in favor of the electrostatic potential. We will work in the oscillating frame [Silin, 1965; Arnush et al., 1971; Lee et al., 1972] to find the density perturbations, eliminate the electron density, and finally transform back to the laboratory frame to use Poisson's equation. The oscillating frame (denoted by superscripts) is defined by

$\underline{r}_\alpha = \underline{r} - \underline{R}_\alpha(t)$, with $(d/dt) \underline{R}_\alpha(t) = \underline{V}_{\alpha 0}(t)$, $\delta n_\alpha(\underline{r}, t) = \tilde{\delta n}_\alpha(\tilde{\underline{r}}_\alpha, t)$, $\delta v_\alpha(\underline{r}, t) = \tilde{\delta v}_\alpha(\tilde{\underline{r}}_\alpha, t)$, $\delta \underline{E}(\underline{r}, t) = \delta \underline{E}(\tilde{\underline{r}}, t)$. Equations (9) and (10) in the oscillating frame are identical to those in the lab frame except for the substitutions $\underline{V}_{\alpha 0} \rightarrow \underline{V}_{\alpha p} + \underline{V}_{\alpha 0}$, $\delta n_\alpha \rightarrow \tilde{\delta n}_\alpha$, $\delta v_\alpha \rightarrow \tilde{\delta v}_\alpha$, $\delta \underline{E} \rightarrow \delta \tilde{\underline{E}}$, $\underline{r} \rightarrow \tilde{\underline{r}}_\alpha$. Assuming $\tilde{\delta v}_\alpha$, $\tilde{\delta n}_\alpha$, $\delta \tilde{\underline{E}} \propto \exp[i(\underline{k} \cdot \underline{r} - \omega t)]$ with $\underline{k} = k_x \hat{x} + k_y \hat{y} + k_z \hat{z}$, $\omega = \omega_r + i\gamma$, $kL \gg 1$, $L^{-1} \equiv N_0(\nabla N_0)$, one can then define susceptibilities χ_i and χ_e in the oscillating frame by $\tilde{\delta n}_i = (-ik\chi_i/e)\delta \tilde{\underline{E}}$ and $\tilde{\delta n}_e = (ik\chi_e/e)\delta \tilde{\underline{E}}$ giving

$$\chi_i = -\frac{1}{4\pi} \frac{\omega_{pi}^2}{\omega - \omega_i} \frac{1}{\Omega_i} \left[\frac{v_{i1}}{\Omega_i} \frac{k_1^2}{k^2} + i \frac{\tilde{\underline{z}} \times \underline{k}_1 \cdot \nabla N_0}{k^2} \right] \quad (12)$$

and

$$\chi_e = \frac{1}{4\pi} \frac{\omega_{pe}^2}{\omega - \omega_e} \left[\frac{1}{\Omega_e} \frac{\tilde{\kappa}_e}{1 + \tilde{\kappa}_e^2} \frac{k_1^2}{k^2} - \frac{1}{\tilde{v}_e} \frac{k_z^2}{k^2} + i \frac{\tilde{\kappa}_e^2}{1 + \tilde{\kappa}_e^2} \frac{1}{\Omega_e} \frac{\tilde{\underline{z}} \times \underline{k}_1 \cdot \nabla N_0}{k^2} \right] \quad (13)$$

where $\omega_{pa}^2 = 4\pi N_0 e^2 / m_a$, $\bar{\omega}_i = \frac{c}{B} \underline{E}_{01} \times \underline{\hat{z}} \cdot \underline{k} + \underline{k} \cdot \underline{V}_{01z} + i(D_1 k_1^2 + D_i k_i^2)$,
 $D_1 = (v_{in}/\Omega_i)(T_i/m_i \Omega_i) + (v_{ei}/\Omega_e)(C_s^2/\Omega_i)$, $D_i = C_s^2/v_{in}$, $\Omega_a = eB/m_a c$,
 $\bar{\omega}_e = \frac{c}{B} \underline{E}_{01} \times \underline{\hat{z}} \cdot \underline{k} + \underline{k} \cdot \underline{V}_{0ez}$, $\tilde{\kappa}_e = \Omega_e/\tilde{v}_{ei}$, $\tilde{v}_{ei} = v_{ei} - i\omega$, and $k_1^2 = k_x^2 + k_y^2$.

In deriving the above susceptibilities we have used $\Omega_e/v_{ei} \gtrsim \Omega_i/v_i \gg 1$ (F-region approximation). In addition, we have modeled weak cold auroral field aligned currents to a relative drift between electrons and ions

$$\underline{V}_d = \underline{\hat{z}} V_d = \tilde{z}(\underline{V}_{0ez} - \underline{V}_{01z}).$$

Having found the susceptibilities in the oscillating frames we must now transform back to the laboratory frame to use Poisson's eqn. (5). (Since we will be considering high frequency pump waves with $\omega_0 \gg \omega_{pi}$ one can neglect the effect of the pump on the ions.) This transformation and resultant dispersion relation can be found in Arnush et al. [1971] and Lee et al. [1972] and can be written:

$$[1 + 4\pi(\chi_i + \chi_e)] = -J_1^2(\zeta) 4\pi \chi_i (1 + 4\pi \chi_e) \times$$

$$\left[\frac{1}{H_e(\omega + \omega_0)} + \frac{1}{H_e(\omega - \omega_0)} \right] \quad (14)$$

where $J_n(\zeta) = \frac{i^{-n}}{\pi} \int_0^\pi e^{i\zeta \cos \theta} \cos n \theta d\theta$ is the Bessel function of integral order n , ζ is defined by $\underline{k} \cdot \underline{R} = \zeta \sin(\omega_0 t + \beta)$, and $H_e = -(4\pi \chi_e)^{-1} (1 + 4\pi \chi_e)$. In deriving (14), the approximation $\zeta \ll 1$ has been made, i.e., the electron excursion length is small compared to the perturbation wavelength.

In the absence of the pump $E_p = 0$, for low frequency modes,

$\delta n_e \approx \delta n_i$. This gives

$$(\omega - \bar{\omega}_e)^{-1} \left[\frac{c}{B} N_o \left(\frac{\hat{z} \times \underline{k}_1 \cdot \nabla N_o}{N_o} - \frac{i \Omega_e}{v_{ei}} k_1^2 \right) \right] = (\omega - \bar{\omega}_i)^{-1} \left[\frac{c}{B} N_o \left(\frac{\hat{z} \times \underline{k}_1 \cdot \nabla N_o}{N_o} + i \frac{v_{in}}{\Omega_i} k_1^2 \right) \right]. \quad (15)$$

with solution $\omega = \omega_{kr}^L + i \gamma_k^L$ given by the linear growth rate [Keskinen and Ossakow, 1982b]

$$\omega_{kr}^L = \frac{c}{B} E_{o1} \times \hat{z} \cdot \underline{k} \quad (16a)$$

$$\gamma_k^L = \left(\frac{k_z^2}{k_1^2} + \frac{v_{in}}{\Omega_i} \frac{v_{ei}}{\Omega_e} \right)^{-1} \frac{v_{ei}}{\Omega_i} \left(\frac{v_{in}}{\Omega_i} \frac{c}{B} \underline{k}_1 \cdot \underline{E}_o - k_z v_d \right) \frac{\hat{z} \times \underline{k}_1 \cdot \nabla N_o}{k_1^2 N_o} - D_1 k_1^2 - D_1 k_z^2 \quad (16b)$$

$$\text{with } D_1 = \frac{v_{in}}{\Omega_i} \frac{T}{m_i \Omega_i} + \frac{v_{ei}}{\Omega_e} \frac{C_s^2}{\Omega_i} \text{ and } D_1 = \frac{C_s^2}{v_{in}}.$$

For finite E_p we have

$$\begin{aligned} (\omega - \bar{\omega}_i)^{-1} \left[\frac{\hat{z} \times \underline{k}_1 \cdot \nabla N_o}{N_o} - i \frac{\Omega_e}{v_{ei}} k_z^2 \right] - (\omega - \bar{\omega}_e)^{-1} \left[\frac{\hat{z} \times \underline{k}_1 \cdot \nabla N_o}{N_o} + i \frac{v_{in}}{\Omega_i} k_1^2 \right] = \\ J_1^2(\zeta) \Omega_i \omega_{pe}^2 \left[-\frac{1}{\Omega_i} \frac{\hat{z} \times \underline{k} \cdot \nabla N_o}{N_o} + \frac{i}{\Omega_i} \frac{v_{in}}{\Omega_i} k_1^2 \right] \left[-\frac{1}{\Omega_e} \frac{\hat{z} \times \underline{k} \cdot \nabla N_o}{k^2 N_o} - \frac{i}{v_{ei}} \frac{k_z^2}{k^2} \right] \\ \times \left[\frac{1}{H_e(\omega + \omega_o)} + \frac{1}{H_e(\omega - \omega_o)} \right] \quad (17) \end{aligned}$$

Defining $\tilde{\omega}_{UH}^2 = \omega_{pe}^2 + \Omega_e^2 + v_e^2$, $\delta = \tilde{\omega}_{UH}^2 - \omega_o^2$, a measure of the frequency mismatch, $\delta_o = \frac{v_{ei}}{\omega_o} (\omega_{pe}^2 - 2\omega_o^2)$, we find from (17) the modified growth rate

$$\gamma = \gamma_{\underline{k}}^L + 2\zeta^2 \left(\frac{k_1^2}{k_1^2} + \frac{v_{ei} v_{ei}}{\Omega_i \Omega_e} \right)^{-1} \left(\frac{\hat{z} \times \underline{k}_1 \cdot \nabla N_0}{k_1^2 N_0} \right)^2 \left(\frac{\omega_{pe}}{v_{ei}} \right)^2 \frac{\omega_{pe}^2 \delta}{\delta^2 + \omega_0^2} \omega_{pe} \quad (18)$$

$$\text{where } \zeta^2 = \frac{e^2}{m_e^2 (\omega_0^2 - \Omega_e^2)} \left\{ k_x^2 E_{px}^2 + k_y^2 E_{py}^2 + \frac{\Omega_e^2}{\omega_0^2} (k_y^2 E_{px}^2 + k_x^2 E_{py}^2) \right. \\ \left. + 2E_{px} E_{py} [k_x k_y \cos \psi (1 - \frac{\Omega_e^2}{\omega_0^2}) - \frac{\Omega_e}{\omega_0} k_1^2 \sin \psi] \right.$$

$$\left. + \frac{\omega_0^2 - \Omega_e^2}{\omega_0^2} k_z E_{pz} [k_x E_{px} - \frac{\Omega_e}{\omega_0} k_x E_{py} \sin \psi + k_y E_{py} \cos \psi] \right\} \quad (19)$$

Here we have taken [Lee et al., 1972]

$$\underline{E}_p = \frac{E_{px}}{-2i} \exp(-i\omega_0 t) \hat{x} + \frac{E_{py}}{-2i} \exp(-i\omega_0 t - i\psi) \hat{y} + \frac{E_{pz}}{-2i} \exp(-i\omega_0 t) \hat{z} + \text{c.c.} \quad (20)$$

and included E_{pz} since we are considering modes for which $k_y \neq 0$.

When $\delta > 0$, i.e., when $\omega < \tilde{\omega}_{UH}$ the pump induced term produces a destabilizing effect whereas when $\delta < 0$ it produces a stabilizing influence. The physical mechanism responsible for stabilization or destabilization of these convective instabilities can be understood as follows: A nonlinear interaction between $\underline{E} \times \underline{B}$ and current convective perturbation electric fields $\underline{\delta E}$ and the pump field $\underline{E}_p$ gives a new field $\underline{E}'$. The combined beating of the fields $\underline{E}'$ and $\underline{E}_p$ can produce a low frequency ponderomotive force which can in turn affect $\underline{\delta E}$ and modify the dispersion relation. We take $\underline{k} = k_x \hat{x} + k_z \hat{z}$, an ambient electric field $E_0 \hat{x}$, and $N = N_0(y)$, i.e., a configuration unstable to the $\underline{E} \times \underline{B}$ and current

convective instabilities. The growth rate in (18) is maximized for

$\delta_{\max} = \pm \delta_0 = \pm (v_{ei}/\omega_o)(\omega_{pe}^2 - 2\omega_o^2)$. Inserting this value for δ , eq. (18) can be written as

$$\gamma = \left(\frac{k_z^2}{k_x^2} + \frac{v_{in}}{\Omega_i} \frac{v_{ei}}{\Omega_e} \right)^{-1} \frac{v_{ei}}{\Omega_e} (k_{xL})^{-1} \left[\frac{v_{in}}{\Omega_i} \frac{c}{B} k_x E_o - k_z v_d + \frac{\zeta^2}{4} (k_{xL})^{-1} \right. \\ \left. \times \left(\frac{\omega_{pe}}{\Omega_e} \right) \left(\frac{\omega_{pe}}{v_{ei} \omega_o} \right) \omega_{pe} \right] - D k_{xL}^2 - D k_{zL}^2 \quad (21)$$

Here the upper (plus) sign refers to the case when $\delta > 0$ (destabilization) while the lower (minus) sign applies when $\delta < 0$ (stabilization). Eq. (21) gives the approximate minimum required energy flux for destabilization or stabilization

$$\left| \frac{cE_p^2}{8\pi} \right| = \frac{2}{\pi} (kL) \frac{\omega_o^2 \Omega_e v_{ei} \omega_o^2}{\Omega_e^2 \omega_{pe} \omega_o \omega_{pe}^2} \frac{\left(\frac{v_{in}}{\Omega_i} \frac{c}{B} k_x E_o - k_z v_d \right) B^2 (\omega_o^2 - \Omega_e^2)^2}{\omega_{pe} c k^2} \quad (22)$$

where we have assumed a circularly polarized O mode with $\psi = -\pi/2$ in eq. (19).

For typical ionospheric F region parameters at diffuse auroral latitudes in the 300-400 km altitude range, we take $\Omega_e = 10^6 \text{ sec}^{-1}$, $\omega_{pe} = 10^6 \text{ sec}^{-1}$, $v_{ei} = 5 \times 10^2 \text{ sec}^{-1}$, $v_{in}/\Omega_i = 10^{-4}$, $E_o = 10 \text{ mV/m}$, $j_{\parallel} = n_o e v_d = 1 \text{ } \mu\text{A/m}^2$, $kL = 10$, $k^{-1} = 1 \text{ km}$, $k_z/k_{\perp} = 10^{-4}$, we find from (22) the approximate energy flux for stabilization to be $cE_p^2/8\pi = 10^{-4} \text{ W/m}^2$. This is an overestimate since we have not considered electric field enhancements near the reflection point [Ginzburg, 1964]. From the coupling parameter ζ as given by (18) we see that there must be a component of $\underline{E}_p$ perpendicular to $\underline{B}$ since $k_{\parallel}/k_{\perp} \ll 1$. At high latitudes this can be accomplished using a vertically incident O-mode pump wave.

3. SUMMARY

The effects of a large amplitude electromagnetic pump wave on convective plasma instabilities (E x B gradient drift and current convective) in the high latitude F region ionosphere have been studied analytically. These convective instabilities have been invoked [Keskinen and Ossakow, 1982a] to explain naturally occurring density irregularities in and near large scale convecting auroral plasma enhancements [Vickrey et al., 1980]. We find that parametric coupling effects associated with a pump wave oscillating with approximately the upper hybrid frequency can stabilize or destabilize the E x B gradient drift and/or the current-convective instability. For parameters typical of the nighttime high latitude F region ionosphere, we find that these modes can be stabilized/destabilized with a free space incident energy flux less than or equal to 10^{-4} W/m^2 . Since the pump wave should have a component of its electric field vector perpendicular to the magnetic field, a vertically propagating O mode is suggested. Since the aforementioned instabilities most likely occur in and near convecting plasma enhancements, the heater should be turned on so as to illuminate the plasma enhancements.

In our development we have made several approximations. We have ignored heater (pump) induced temperature and density changes. The time scales for these changes, to a first approximation, is longer than the growth times of the instabilities studied here. In addition, we have neglected D region absorption effects. Finally, we have assumed that the instabilities studied here are at or near marginal stability. We reserve these topics for a future report.

Acknowledgements

This work was supported by the Defense Nuclear Agency and the Office of Naval Research.

REFERENCES

- Aarons, J., A descriptive model of F layer high-altitude irregularities as shown by scintillation observations, J. Geophys. Res., 78, 7441, 1973.
- Chaturvedi, P.K., Collisional ion cyclotron waves in the auroral ionosphere, J. Geophys. Res., 81, 6169, 1976.
- Chaturvedi, P.K. and S.L. Ossakow, Nonlinear stabilization of the current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 957, 1979.
- Chaturvedi, P.K. and S.L. Ossakow, The current convective instability as applied to the auroral ionosphere, J. Geophys. Res., 86, 4811, 1981.
- Clark, D.H. and W.J. Raitt, The global morphology of irregularities in the topside ionosphere as measured by the total ion current probe on ESRO-4, Planet. Space Sci., 24, 873, 1976.
- Davies, K., Review of recent progress in ionospheric prediction, Radio Sci., 16, 1407, 1981.
- Drummond, W.E. and M.N. Rosenbluth, Anomalous diffusion arising from microinstabilities in a plasma, Phys. Fluids, 5, 1507, 1962.
- DuBois, D.F. and M.V. Goldman, Radiation-induced instability of electron plasma oscillations, Phys. Rev. Lett., 14, 544, 1965.
- Dyson, P.L., Direct measurement of the size and amplitude of irregularities in the topside ionosphere, J. Geophys. Res., 74, 6291, 1969.
- Dyson, P.L. and J.D. Winningham, Topside ionospheric spread F and particle precipitation in the dayside magnetospheric clefts, J. Geophys. Res., 79, 5219, 1974.

- Erukhimov, L.M., A.M. Lerner, V.I. Kosolapenko, and E.N. Myasnikov, The spectral form of small-scale plasma turbulence in the auroral ionosphere, Planet. Space Sci., 29, 931, 1981.
- Fainberg, Y. and V.D. Shapiro, Drift instabilities of a plasma situated in a high frequency electric field, JETP, 25, 189, 1967.
- Fejer, J.A., Ionospheric modification and parametric instabilities, Rev. Geophys. and Space Physics, 17, 135, 1979.
- Fremouw, E.J., C.L. Rino, R.C. Livingston, and M.D. Cousins, A persistent subauroral scintillation enhancement observed in Alaska, Geophys. Res. Lett., 4, 539, 1977.
- Ginzburg, V.L., The Propagation of Electromagnetic Waves in Plasma, Addison-Wesley, Reading, Mass., 1964.
- Greenwald, R.A., Diffuse radar aurora and the gradient drift instability, J. Geophys. Res., 79, 4807, 1974.
- Hanuise, C., J.P. Villain, and M. Crochet, Spectral studies of F region irregularities in the auroral zone, Geophys. Res. Lett., 8, 1083, 1981.
- Hower, G.L., D.M. Sanz, and C.L. Allison, Comparison of HF radar echoes and high latitude spread-F measurements, J. Geophys. Res., 71, 3215, 1966.
- Kadomtsev, B.B. and A.V. Nedospasov, Instability of the positive column in a magnetic field and the "anomalous diffusion effect," J. Nucl. Energy, Part C, 1, 230, 1960.
- Kelley, M.C., K.D. Baker, C. Rino, and J.C. Ulwick, Simultaneous rocket probe scintillation and incoherent scatter radar observations of irregularities in the auroral zone ionosphere, Radio Sci., 15, 491, 1980.

- Keskinen, M.J., S.L. Ossakow, and B.E. McDonald, Nonlinear evolution of diffuse auroral F region ionospheric irregularities, Geophys. Res. Lett., 7, 573, 1980.
- Keskinen, M.J. and S.L. Ossakow, Theories of high latitude ionospheric irregularities - a review, Radio Sci., submitted, 1982a.
- Keskinen, M.J. and S.L. Ossakow, Nonlinear evolution of plasma enhancements in the auroral ionosphere I: Long wavelength irregularities, J. Geophys. Res., 87, 144, 1982b.
- Keskinen, M.J. and S.L. Ossakow, Nonlinear evolution of convecting plasma enhancements in the auroral ionosphere II: Small scale irregularities, J. Geophys. Res., (to be published) 1982c.
- Kindel, J.M. and C.F. Kennel, Topside current instabilities, J. Geophys. Res., 76, 3055, 1971.
- Lee, K., P.K. Kaw, and C.F. Kennel, External production and control of electrojet irregularities, J. Geophys. Res., 77, 4197, 1972.
- Lehnert, B., Diffusion processes in the positive column in a longitudinal magnetic field, in Proceedings of the Second Geneva Conference on the Peaceful Uses of Atomic Energy, 32, 349, 1958.
- Linson, L.M. and J.B. Workman, Formation of striations in ionospheric plasma clouds, J. Geophys. Res., 75, 3211, 1970.
- Nishikawa, K., Parametric excitation of coupled waves, 2, Parametric plasmon-photon interaction, J. Phys. Soc. Japan, 24, 916, 1968.
- Ogawa, T., H. Mori, and S. Miyazaki, Rocket observations of electron density irregularities in the antarctic auroral E region, J. Geophys. Res., 81, 4013, 1976.

- Olesen, J.K., F. Primdahl, F. Spangsløv, E. Ungstrup, A. Bahnsen, U. Fahlson, C.-G. Falthammer, and A. Pedersen, Rocket borne wave and plasma observations in unstable polar cap E region, Geophys. Res. Lett., 3, 399, 1976.
- Ossakow, S.L. and P.K. Chaturvedi, Current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 332, 1979.
- Perkins, F.W., C. Oberman, and E.J. Valeo, Parametric Instabilities and Ionospheric Modification, J. Geophys. Res., 79, 1478-1496, 1974.
- Phelps, A.D.R. and R.C. Sagalyn, Plasma density irregularities in the high-latitude topside ionosphere, J. Geophys. Res., 81, 515, 1976.
- Rodriguez, P., M. Singh, D.N. Walker, E.P. Szuszczewicz, and J.C. Holmes, The STP/S3-4 satellite experiment: high latitude, Effects of the Ionosphere on Radiowave Systems, ed. J.M. Goodman, U.S. Government Printing Office, Washington, D.C., 1981.
- Sagalyn, R.C., M. Smiddy, and M. Ahmed, High-latitude irregularities in the topside ionosphere based on Isis 1 thermal probe, J. Geophys. Res., 79, 4252, 1974.
- Silin, V.P., Parametric resonance in a plasma, JETP, 21, 1127, 1965.
- Simon, A., Instability of a partially ionized plasma in crossed electric and magnetic fields, Phys. Fluids, 6, 382, 1963.
- Sundaram, A.K. and P.K. Kaw, Parametric excitation and suppression of drift waves by lower hybrid fields, Nucl. Fusion, 13, 901, 1973.
- Vickrey, J.F., C.L. Rino, and T.A. Potemra, Chatanika/Triad observations of unstable ionization enhancements in the auroral F-region, Geophys. Res. Lett., 789, 1980.
- Weaver, P.F., Backscatter echoes from field-aligned irregularities in the F region, J. Geophys. Res., 70, 5425, 1965.

APPENDIX H

A DYNAMIC MODEL FOR THE
AURORAL FIELD LINE PLASMA
IN THE PRESENCE OF FIELD-ALIGNED CURRENT

**A Dynamic Model for the
Auroral Field Line Plasma
in the Presence of Field-Aligned Current**

H. G. MITCHELL, JR.

*Science Applications, Inc.
McLean, Virginia 22102*

P. J. PALMADESSO

*Geophysical and Plasma Dynamics Branch
Plasma Physics Division*

July 23, 1982

This work was supported by the Office of Naval Research and by NASA Contract No. W-14365.



NAVAL RESEARCH LABORATORY
Washington, D.C.

Approved for public release; distribution unlimited

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
NRL Memorandum Report 4840		
4. TITLE (and Subtitle)		5. TYPE OF REPORT & PERIOD COVERED
A DYNAMIC MODEL FOR THE AURORAL FIELD LINE PLASMA IN THE PRESENCE OF FIELD-ALIGNED CURRENT		Interim report on a continuing NRL problem.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)		8. CONTRACT OR GRANT NUMBER(s)
H.G. Mitchell, Jr. and P.J. Palmadesso		
9. PERFORMING ORGANIZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Office of Naval Research National Aeronautics and Space Admin. Arlington, VA 22217 Washington, DC 20546		61153N; RR033-02-44; 47-0884-0-2; NASA W-14365
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE
		July 23, 1982
		13. NUMBER OF PAGES
		23
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)		
Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
This work was supported by the Office of Naval Research and by NASA Contract No. W-14365.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
Ionosphere-magnetosphere coupling Auroral field line current Polar wind		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)		
We have developed a dynamic numerical model of the plasma along an auroral field line in order to provide a vehicle for studying ionosphere-magnetosphere coupling processes. The model is a multi-moment multi-fluid approximation of a gyrotropic plasma consisting of three species (electrons, hydrogen ions, oxygen ions) along a segment of auroral magnetic field line extending from an altitude of 800 km to 18 earth radii. We have performed preliminary simulations for the case of a current-free polar wind equilibrium of the field line plasma and the case in which a large upward field-aligned current is applied (Continues)		

DD FORM 1473

EDITION OF 1 NOV 65 IS OBSOLETE
54-102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Continued)

to the field line. In the former case, the agreement between our model and previous static results is reasonable given the differing boundary conditions inherent in the two cases. In the case of a field-aligned current, we note that the flux tube plasma responds to the current on several time scales. After an initial rapid response due to electron-ion Joule heating, thermal oscillations of the flux tube plasma persist on time scales of the order of one hour, illustrating the complicated nature of the response of a collisionless plasma when heat flow transport is treated in a dynamic manner.

CONTENTS

INTRODUCTION	1
THE FIELD LINE MODEL	2
THE POLAR WIND SIMULATION	7
A FIELD-ALIGNED CURRENT SIMULATION	12
SUMMARY	17
ACKNOWLEDGMENT	18
REFERENCES	19

A DYNAMIC MODEL FOR THE AURORAL FIELD LINE PLASMA IN THE PRESENCE OF FIELD-ALIGNED CURRENT

Introduction

One of the key problems of auroral physics concerns the process by which auroral electrons are accelerated along auroral magnetic field lines. Early theories of the plasma along the field lines above the aurora stated that the collisionless nature of this plasma implied an almost infinite conductivity along the magnetic field so that the field lines could support no parallel electric fields. This conclusion has since been disputed by observational evidence of field-aligned acceleration, first in the "inverted-V" events [reviewed by Frank, 1976] and more directly in the S3-3 satellite data [reviewed by Stern, 1979]. It is now fairly well established that at certain times an anomalously large field is present along auroral field lines at an altitude of several earth radii and is relatively slowly varying in time. This implies the existence of some quasi-steady form of resistivity in the auroral plasma or some other mechanism to maintain a large electric field in a plasma in the absence of collisional resistivity.

A number of mechanisms have been proposed to account for these parallel electric fields. One avenue of research has been devoted to the plasma instabilities which might disrupt the flow of the Birkeland currents parallel to the magnetic field, causing a form of "anomalous" resistivity by wave turbulence. This approach is supported by the observations of terrestrial kilometric radiation emission from this region of the auroral zone [Gurnett, 1974]. Reviews of the theory of plasma instabilities and turbulence in this region can be found in Papadopoulos [1977] and Mozer [1976]. Other mechanisms put forward for auroral acceleration include double layers [Shawhan et al., 1978], electrostatic shocks [Kan, 1975, and Swift, 1975], and magnetic mirror trapping [Lyons, 1980]. However, despite the volume of work which exists on this topic in the form of observation, theory, and numerical simulation, there is no general consensus as to the correct mechanism. One reason for this is the fact that many of these mechanisms are developed and demonstrated for the case of a homogeneous plasma in a constant magnetic field. There are few clear predictions as to their behavior when applied to the auroral plasma. Further, the details of the energy balance are not well understood along the auroral field line even in the absence of current, and these strongly affect the behavior of

Manuscript submitted April 23, 1982.

the particle distributions involved in any resistive mechanism. As a result, it is often hard to say whether a specific process for maintaining the auroral electric field agrees or disagrees with observation.

Clearly, a more global approach to the problem of field line acceleration for the auroral plasma would be useful. The plasma configuration along high latitude field lines has been investigated in the hydrodynamic models of Banks and Holzer [1968] and Schunk and Watkins [1981, 1982] and in the kinetic models of Lemaire and Scherer [1973] and Chiu and Cornwall [1980], among others. These models tend to be static models of steady-state auroral plasma configurations in the absence of an auroral field line current. We will describe in this paper a dynamic model we have been developing for the purpose of understanding the behavior of the auroral field line plasma in the presence of field-aligned current and auroral acceleration mechanisms. This model is a numerical model of a segment of auroral field line beginning in the topside ionosphere and extending well out into the magnetosphere, encompassing the transition of the auroral plasma from collisional to collisionless behavior and employing a multi-moment multi-fluid approximation of the type developed by Schunk [1977]. The details of this model will be given here along with the results of our early simulations of the auroral plasma both with and without field-aligned current. We intend in the future to investigate particular mechanisms of auroral field line acceleration with this model by adjusting the anomalous transport terms in the model's plasma transport equations to simulate the effects of plasma turbulence and by introducing modifications designed to model other proposed acceleration processes. In this way we hope to gain a better understanding of the behavior of these mechanisms in the context of the global behavior of the auroral plasma.

The Field Line Model

The field line model is designed to dynamically simulate the behavior of the plasma in a flux tube encompassing an auroral field line. The field line is assumed to be radial with no curvature and to extend from an altitude of about 800 km in the topside ionosphere out to a distance of 18 earth radii. The actual region of interest in this model is the lower four earth radii or so, the region in which the plasma changes from collisional

to collisionless behavior. The upper region of the flux tube is an extended boundary layer for the purpose of dealing with upper boundary effects. The cross-sectional area of the flux tube diverges as r^3 in the near-earth region, where r is the geocentric distance, and becomes nearly constant in the upper boundary layer. The model is essentially one-dimensional, with all quantities functions of r .

The flux tube plasma consists of three particle species: electrons, hydrogen ions, and oxygen ions. In the present version of the model, the oxygen ions are a static background population at a constant temperature. They are present in the model in order to approximate the behavior of the plasma in the topside ionosphere by providing a thermal reservoir and the correct electron scale height at the lower end of the flux tube. The electrons and hydrogen ions are the dynamic species in the model. The distribution functions for these two species are assumed to be gyrotropic about the field line direction and are characterized by five moments: number density, temperatures parallel and perpendicular to the field line, and species' velocity and heat flow along the field. Including the heat flow as a dynamic quantity rather than calculating it by means of a thermal conduction approximation allows a reasonable treatment of thermal wave effects in the collisionless region of the model.

The moments of the distribution function are treated dynamically using a set of transport equations derived from the 13-moment equations of Schunk [1977]. For a gyrotropic plasma, these equations are:

$$\frac{\partial n_s}{\partial t} = -v_s \frac{\partial n_s}{\partial r} - n_s \frac{\partial v_s}{\partial r} - \frac{n_s v_s}{A} \frac{\partial A}{\partial r} + \frac{\delta n_s}{\delta t} \quad (1)$$

$$\begin{aligned} \frac{\partial v_s}{\partial t} = & -v_s \frac{\partial v_s}{\partial r} - \frac{k}{m_s} \frac{\partial T_{s\parallel}}{\partial r} - \frac{k T_{s\parallel}}{m_s n_s} \frac{\partial n_s}{\partial r} - \frac{k(T_{s\parallel} - T_{s\perp})}{m_s A} \frac{\partial A}{\partial r} \\ & + \frac{e_s}{m_s} E - \frac{GM}{r^2} + \frac{\delta v_s}{\delta t} \end{aligned} \quad (2)$$

$$k \frac{\partial T_{s\parallel}}{\partial t} = -v_s k \frac{\partial T_{s\parallel}}{\partial r} - 2k T_{s\parallel} \frac{\partial v_s}{\partial r} - \frac{6}{5n_s} \frac{\partial q_s}{\partial r} - \frac{2}{5} \frac{q_s}{n_s A} \frac{\partial A}{\partial r} + k \frac{\delta T_{s\parallel}}{\delta t} \quad (3)$$

$$k \frac{\partial T_{s\perp}}{\partial t} = -v_s k \frac{\partial T_{s\perp}}{\partial r} - \frac{2}{5n_s} \frac{\partial q_s}{\partial r} - \left(\frac{4}{5} \frac{q_s}{n_s} + v_s k T_{s\perp} \right) \frac{1}{A} \frac{\partial A}{\partial r} + k \frac{\delta T_{s\perp}}{\delta t} \quad (4)$$

$$\begin{aligned} \frac{\partial q_s}{\partial t} = & -v_s \frac{\partial q_s}{\partial r} - \frac{16}{5} q_s \frac{\partial v_s}{\partial r} - \left(\frac{11}{18} T_{s\parallel} + \frac{8}{9} T_{s\perp} \right) \frac{n_s k^2}{m_s} \frac{\partial T_{s\parallel}}{\partial r} \\ & - \left(\frac{17}{9} T_{s\parallel} - \frac{8}{9} T_{s\perp} \right) \frac{n_s k^2}{m_s} \frac{\partial T_{s\perp}}{\partial r} + \frac{4k^2}{9m_s} (T_{s\parallel} - T_{s\perp})^2 \frac{\partial n_s}{\partial r} \\ & + \left[\frac{n_s k^2}{m_s} (T_{s\parallel} - T_{s\perp}) \left(\frac{1}{3} T_{s\parallel} - \frac{4}{3} T_{s\perp} \right) - \frac{7}{5} v_s q_s \right] \frac{1}{A} \frac{\partial A}{\partial r} + \frac{\delta q_s}{\delta t} \end{aligned} \quad (5)$$

where n_s is the number density, v_s is the velocity, $T_{s\parallel}$ is the parallel temperature, $T_{s\perp}$ is the perpendicular temperature, q_s is the heat flow, m_s is the mass, and e_s is the charge of species s , A is the cross-sectional area of the flux tube, E is the electric field parallel to the field line, k is Boltzmann's constant, G is the gravitational constant, and M is the mass of the earth. For a given moment F of the distribution function, $\delta F/\delta t$ is the change in F due to resistive and plasma turbulence effects. The present version of the field line model includes only resistivity due to Coulomb collisions among the three particle species in the model. Once the behavior of this simulation is understood for this case, the turbulence terms will be altered to reflect the behavior of the particle distribution functions in the presence of plasma microprocesses. The specific resistive terms used in the present simulation are Burgers' [1979] 'linear' collision terms for the case of Coulomb collisions, given by

$$\frac{\delta n_s}{\delta t} = 0 \quad (6)$$

$$\frac{\delta v_s}{\delta t} = \sum_t v_{st} (v_t - v_s) \quad (7)$$

$$\begin{aligned} k \frac{\delta T_{s\parallel}}{\delta t} = & \sum_t \frac{m_s v_{st}}{(m_s + m_t)} \left[\frac{6}{5} k T_{t\parallel} - \left(2 + \frac{4m_t}{5m_s} \right) k T_{s\parallel} + \frac{4}{5} k T_{t\perp} + \frac{4m_t}{5m_s} k T_{s\perp} \right. \\ & \left. + \frac{6}{5} m_t (v_t - v_s)^2 \right] \end{aligned} \quad (8)$$

$$k \frac{\delta T_{s\perp}}{\delta t} = \frac{1}{t} \frac{m_s v_{st}}{(m_s + m_t)} \left[\frac{2}{5} kT_{t\parallel} + \frac{2m_t}{5m_s} kT_{s\parallel} + \frac{8}{5} kT_{t\perp} - \left(2 + \frac{2m_t}{5m_s} \right) kT_{s\perp} \right. \\ \left. + \frac{2}{5} m_t (v_t - v_s)^2 \right] \quad (9)$$

$$\frac{\delta q_s}{\delta t} = \frac{1}{t} \frac{n_s m_s v_{st}}{(m_s + m_t)^2} \left[\frac{27}{10} m_t \frac{q_t}{n_t} - \left(3m_s + \frac{8}{5} m_t + \frac{13}{10} \frac{m_t^2}{m_s} \right) \frac{q_s}{n_s} \right. \\ \left. + \frac{3}{2} \frac{m_t}{m_s} (m_t + m_s) kT_s (v_s - v_t) \right] \quad (10)$$

where each sum includes all charged particles species in the simulation and v_{st} is the Coulomb collision frequency,

$$v_{st} = \frac{(32\pi)^{1/2} e_s^2 e_t^2 n_t}{3m_s^2} \left(1 + \frac{m_s}{m_t} \right) \left(\frac{kT_s}{m_s} + \frac{kT_t}{m_t} \right)^{-3/2} \ln \Lambda \quad (11)$$

($\ln \Lambda$ is the Coulomb logarithm and $T_s = (T_{s\parallel} + 2T_{s\perp})/3$ is the total temperature of species s).

The scale of this model is large compared to the electron Debye length, so the transport equation (1) for electron number density may be replaced by an expression for charge neutrality:

$$n_e = n_p + n_i. \quad (12)$$

(The subscripts e, p, and i represent electrons, hydrogen ions, and oxygen ions respectively.) Further, the assumption will be made that the total flux tube current remains constant at some fixed value I during a simulation, which implies that the electron velocity transport equation (2) is replaced by

$$v_e = \frac{1}{n_e} (n_p v_p - \frac{I}{eA}). \quad (13)$$

This assumption reflects the fact that the auroral current generator is not strongly affected by the behavior of the flux tube plasma. The assumption of charge neutrality and constant current along the field line allows the parallel electric field E to be calculated from a generalized Ohm's law constructed from the electron and ion velocity transport equations (2):

$$E = \frac{m_s}{en_e A} \frac{\partial}{\partial r} (n_p v_p^2 A - n_e v_e^2 A) - \frac{1}{e} \left[\frac{\partial T_{e\parallel}}{\partial r} + \frac{T_{e\parallel}}{n_e} \frac{\partial n_e}{\partial r} + \frac{(T_{e\parallel} - T_{e\perp})}{A} \frac{\partial A}{\partial r} \right] - \frac{n_i m_e G M}{n_e e r^2} + \frac{m_e}{e} \left[\frac{\delta v_e}{\delta t} - \frac{n_p}{n_e} \frac{\delta v_p}{\delta t} \right] \quad (14)$$

where terms on the order of m_e/m_p have been neglected.

The field line model, therefore, uses equations (1)-(5) for hydrogen ions and equations (3)-(5) for electrons to step the values of these distribution moments forward in time, while calculating the electric field and electron density and velocity self-consistently at each time with equations (12)-(14). The simulation calculates transport using the simple partial donor cell method (Hain, 1978) on an unequally spaced grid. The segment of the field line being modelled is divided into about 100 cells. The cell size at the lower end of the segment is small, around 50 km, in order to properly deal with the transport in the presence of the large density gradients due to the small oxygen scale height. The cell size increases with altitude until it reaches roughly 1000 km at an altitude of 6 earth radii (R_E). At this altitude, the upper boundary layer begins, and the cell size begins to increase rapidly, reaching about one R_E at the upper boundary of the model. The time step for this simulation is determined primarily by the transit time of a thermal electron in a cell at the lower end of the field line and is set at about .06 seconds.

Preliminary simulations with the field line model indicated that the behavior of the simulation could become unstable if the species heat flows were allowed to grow too large. A subsequent wave analysis of the model equations (1)-(5) and (12)-(14) revealed that an unstable ion thermal wave exists when the velocity of the ion heat flow exceeds roughly the ion thermal velocity, i.e.,

$$|q_p|/n_p T_p > (kT_p/m_p)^{1/2} \quad (15)$$

This indicates that the thermal velocity is a natural limit on the rate of energy flux relative to the species velocity. The form of a distribution function with a heat flow of this magnitude is highly non-Maxwellian, and it is to be expected that the instability would tend to reduce the heat flow. As a result, the field line model limits the magnitude of the species' heat flow by using the relation

$$|q_s| < n_s T_s (kT_s/m_s)^{1/2} \quad (16)$$

This limit tends to have little or no effect on the electrons because of their large thermal velocity, but can be a relatively stringent condition for hydrogen ions. It slows the response of the hydrogen ion distribution to large temperature gradients, allowing them to persist for longer periods of time.

The Polar Wind Simulation

The first simulations with the field line model were performed to construct a steady-state configuration of the field line plasma in the current-free case. We began by initializing the flux tube in a polar wind configuration similar to that of Banks and Holzer [1968]. In the polar wind, hydrogen ions are accelerated upwards in the flux tube to supersonic velocities due to the flux tube divergence and the small partial pressure of H^+ at the upper end of the field line. The outward H^+ flux depletes the hydrogen ion population at lower altitudes, and, as a result, oxygen ions are the dominant ion species up to an altitude of around 4000 km. The O^+ -e charge separation electric field in this region provides the initial H^+ acceleration up the field line. In the model of Banks and Holzer, the population of H^+ is maintained against depletion by O^+ -H charge exchange at the altitudes below 1000 km, and the hydrogen ion escape flux is limited to about $6 \times 10^8 \text{ cm}^{-2} \text{ sec}^{-1}$ at 1000 km. In the field line model, there are no neutral species at this point, so the hydrogen ion flux is maintained by a fixed density and upward velocity for H^+ at 800 km, the lower boundary of the model. As a result, the hydrogen ion number density tends to be somewhat greater in the lowest few cells of the field line model than would be expected in the presence of O^+ -H charge exchange.

The model of Banks and Holzer assumed that the charged particle species in the flux tube had uniform temperatures and, therefore, heat flow was neglected. We assumed the same, initializing the three flux tube particle species at a constant temperature of 3000 K with no heat flow. To begin our simulations, the transport of temperature and heat flow were suppressed in the field line model in an attempt to duplicate the results of Banks and Holzer. The model achieved a steady-state polar wind under these conditions, with the number density and velocity profiles shown in Fig. 1a,b. Once this equilibrium had been reached, the temperature and heat flow transport were "turned on", and the simulation was run until a new steady-state was achieved. The lower boundary temperature was held fixed at 3000 K during the simulation. This is a somewhat artificial condition, but the actual thermal balance at this altitude involves modelling the upper ionosphere in more detail than this simulation is capable of at the present time. The upper boundary temperature was determined by an outflow condition in that thermal gradients propagating up the field line were allowed to propagate out of the simulation.

The temperature profiles from the new steady-state are shown in Fig. 2. The primary effect displayed by the electron temperatures in this equilibrium is adiabatic cooling due to the expansion of the cross-sectional flux tube area with height. At the lower end of the field line, the high rate of expansion causes rapid cooling, and the resulting temperature gradient is maintained by low thermal conductivity due to high particle density. The electron temperature profiles tend to flatten out near the upper end of the tube as both the rate of expansion and the particle densities decrease. In addition, the electron temperatures exhibit a small anisotropy at the upper end of the tube due to unequal rates of cooling in the perpendicular and parallel directions. This may also be viewed as the fluid model manifestation of the mirror effect, with perpendicular energy being transferred into parallel energy as the electrons are accelerated upwards in the diverging magnetic field.

The hydrogen ion temperatures, on the other hand, increase with increasing altitude at the lower end of the flux tube, before beginning to exhibit adiabatic cooling at an altitude of three-fourths to one R_E . The

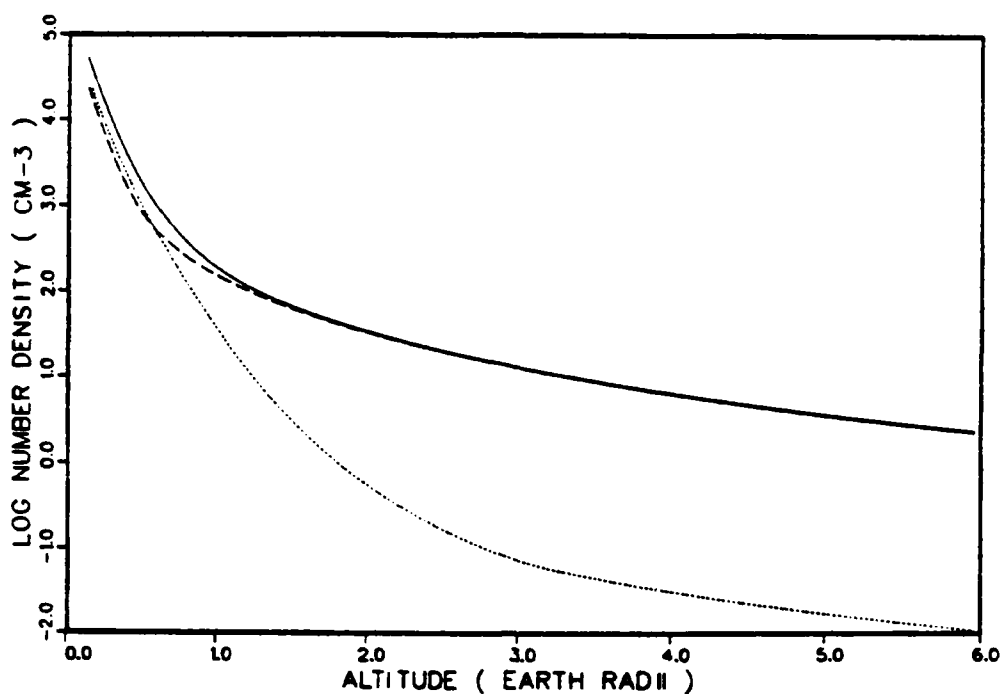


Fig. 1(a) — Number densities for the steady-state polar wind with no field-aligned current: e (solid curve), H^+ (dashed curve), O^+ (dotted curve).

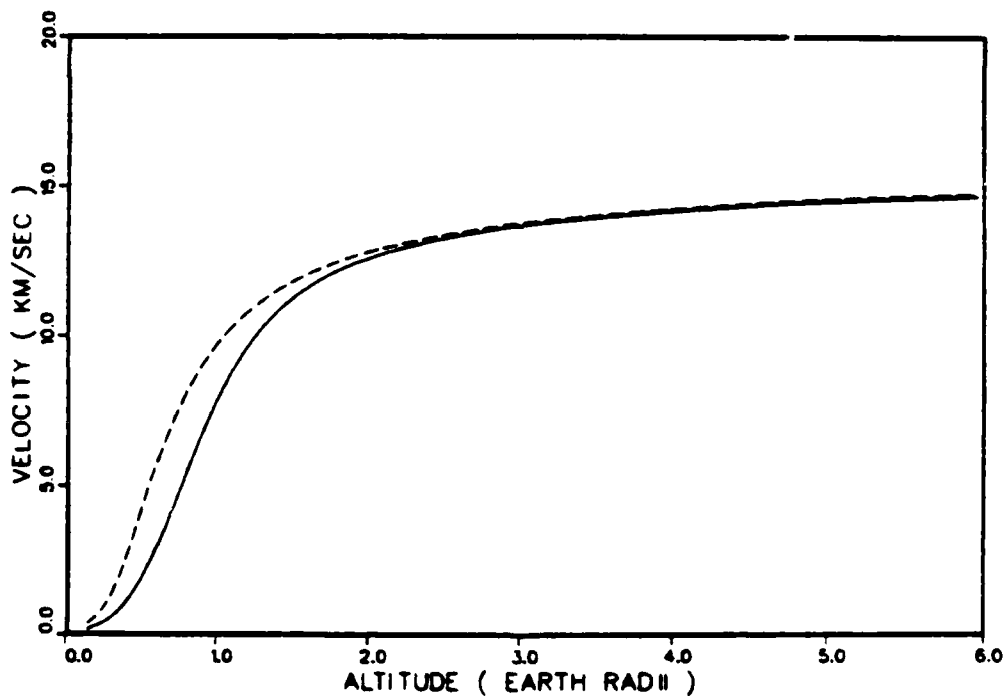


Fig. 1(b) — Velocities for the steady-state polar wind with no field-aligned current: e (solid curve), H^+ (dashed curve). The O^+ velocity is uniformly zero.

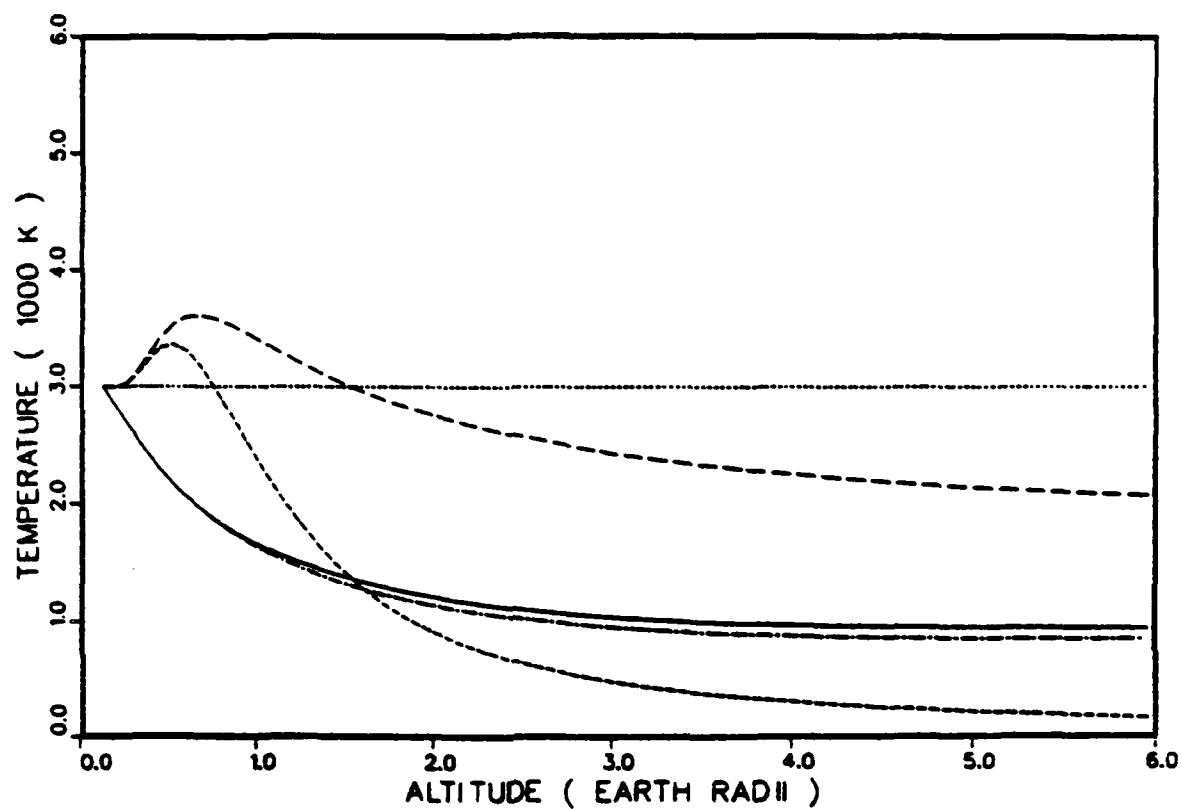


Fig. 2 — Temperatures for the steady-state polar wind with no field-aligned current: $T_{e||}$ (solid curve), $T_{e\perp}$ (dot-dash curve), $T_{H||}$ (long dashed curve), $T_{H\perp}$ (short dashed curve), T_o (dotted curve).

temperature increase is caused by the Joule heating of the hydrogen from collisions with the oxygen ions, a process which heats in the parallel direction more strongly than the perpendicular. The strength of this effect decreases rapidly with altitude due to the small scale height of the oxygen, allowing the adiabatic cooling to become dominant at about 4000 km altitude. The anisotropy of the cooling is much more evident in the hydrogen ions than in the electrons because of the supersonic hydrogen velocities and the fact that, in the collisionless region, heat transport along the field line is less efficient for the hydrogen than for the electrons. As a result, the hydrogen temperature profiles maintain significant gradients at higher altitudes than the electron profiles and exhibit large temperature anisotropies at the upper end of the flux tube.

It is instructive to compare these results with those of Schunk and Watkins [1981, 1982]. These works model current-free steady-state configurations of the auroral field line plasma by integrating up the field line the steady-state form of a set of transport equations for a three species (e , H^+ , O^+) plasma similar to the ones in equations (1)-(5). The steady-state configurations which result are a function of the lower boundary (1500 km altitude) conditions, specifically the values of the species' density, velocity, temperature, and temperature gradient. Comparing the results of our simulations with those of Schunk and Watkins [1982] for the hydrogen ion profiles in the case of supersonic polar wind, a number of similarities can be seen. For low electron temperatures, both exhibit an anisotropic hydrogen ion cooling resulting in a larger parallel than perpendicular temperature at the upper end of the field line. Also, both have a region in which the parallel temperature increases with altitude before adiabatic cooling dominates. The rate of cooling for the hydrogen ions is less in our simulation, most likely because both the oxygen ions and the electrons are hotter at the lower end of the field line (3000 K in our simulation vs. 1000 K in that of Schunk and Watkins). The highly anisotropic cooling of the hydrogen ions is also a feature of another steady-state polar wind model, that of Holzer, Fedder, and Banks [1971], which also attempted to realistically deal with the ion heat flow in the collisionless region. The heat flow for hydrogen ions is upward and monotonically decreasing in these models, implying that the polar wind requires an outward flow of heat from the ionosphere. The magnitude of the

heat flow is roughly an order of magnitude greater in our simulation than that of Schunk and Watkins due to the smaller hydrogen ion velocity at the lower end of the field line in our model. A smaller velocity implies a smaller rate of advective heating, forcing a greater heat flow.

The electron profiles of Schunk and Watkins [1981] cannot be directly compared to our simulation results. By assuming a positive electron temperature gradient at an altitude of 1500 km, they achieve an equilibrium in which the electron temperatures are monotonically increasing with altitude and the electron heat flow is down the field line. Further, the temperature anisotropy is reversed in the upper region of the field line, with the perpendicular temperature being the larger. Thus, the positive temperature gradient boundary condition for the electrons forces the upper boundary of the field line to act as a heat source for the flux tube. The outflow condition in our model, on the other hand, treats the upper boundary as an electron heat sink, resulting in a monotonically decreasing electron temperature with altitude. As long as the electron species velocity is less than the electron thermal velocity, the assumed electron temperature of the upper boundary plasma has a strong effect on the electron temperature profiles in the flux tube. The supersonic outflow of hydrogen ions in the polar wind implies that the energy flux must be outward for the hydrogen ions, with the result that only the ionosphere can provide a heat source for H^+ regardless of the assumed temperature of the upper boundary protons. Schunk and Watkins [1982] note that, for subsonic H^+ outflow, the upper boundary plasma can provide a proton heat source.

A Field-Aligned Current Simulation

Preliminary simulations have been performed with the field line model in the presence of current parallel to the magnetic field and we would like to briefly describe one of these simulations. In this case, a constant upward current of $5.4 \times 10^{-6} \text{ A/m}^2$ at an altitude of 800 km was assumed for the duration of the simulation. This value was chosen to illustrate the effect of a large downward electron energy flux on the flux tube plasma. The initial state of the flux tube is assumed to be the steady-state polar wind described in the previous section.

The behavior of the electron temperature profiles for the first ten

minutes after current onset are shown in Fig. 3a,b. Initially, Joule heating due to the relative electron-hydrogen ion drift causes a rapid temperature increase at the lower end of the field line in both the parallel and perpendicular directions. The outward electron heat flow increases in response to the resulting electron temperature gradient, thus heating the electrons at higher altitudes. Note that the gradients are opposite for the perpendicular and parallel temperatures in the 1 to 4 R_E altitude range, an effect caused by the transfer of parallel to perpendicular energy of the precipitating electrons in the converging magnetic field. Also, the upward electron heat flow has an anisotropic effect on the electron temperature, resulting in a cooling of the perpendicular temperature at higher altitudes as energy is transported upward. Eventually, a rough balance is achieved in the collisionless region between upward heat flow and downward convecting electron energy flux. The location and magnitude of the maximum in the parallel temperature profile for this equilibrium is very sensitive to the rate of flux tube divergence in the collisionless region. As the electron temperature rises in the lower end of the flux tube, the hydrogen ions are also collisionally heated in this region, as shown in the hydrogen ion temperature profiles for the first ten minutes in Fig. 4a,b. Thus, a thermal wave propagating up the field line is also created in the parallel hydrogen ion temperature, although a much slower wave due to the smaller ion thermal velocity. The increased electron temperature also results in an increased ambipolar electric field in the lower regions of the flux tube, causing an increased hydrogen ion outflow velocity (Fig. 5).

During the next twenty minutes of the simulation, the hydrogen ion thermal wave generated previously propagates into the upper boundary layer as the temperature of the hydrogen ions continues to increase in the collisional region below 1 R_E altitude due to $e-H^+$ collisions (Fig. 6). These profiles of the parallel hydrogen ion temperature illustrate the difference between a treatment of heat flow by a thermal conduction approximation and one in which heat flow is a dynamic quantity. In the latter case, the heat flow has a finite response time to thermal gradients. As a result, the flux tube plasma exhibits the large thermal oscillations shown in Fig. 6 rather than having the thermal gradients damped out by the large thermal conductivity of a collisionless plasma. As

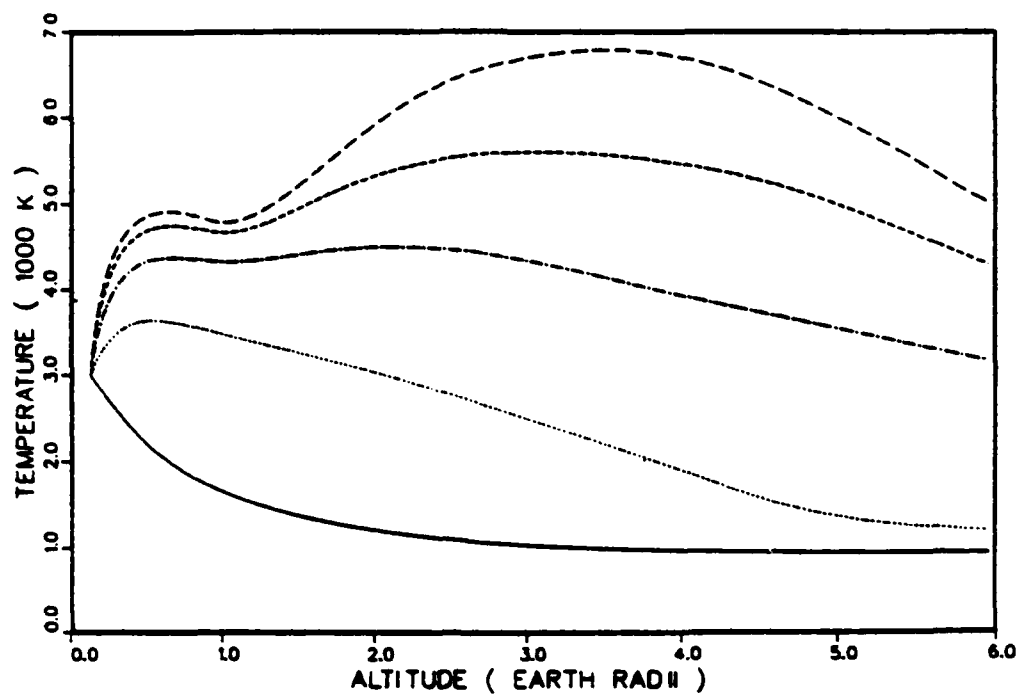


Fig. 3(a) — Parallel electron temperatures for the first ten minutes after the onset of a current of $5.4 \mu\text{A}/\text{m}^2$ at 800 km: 0 min (solid curve), 2.5 min (dotted curve), 5 min (dot-dash curve), 7.5 min (short dashed curve), 10 min (long dashed curve).

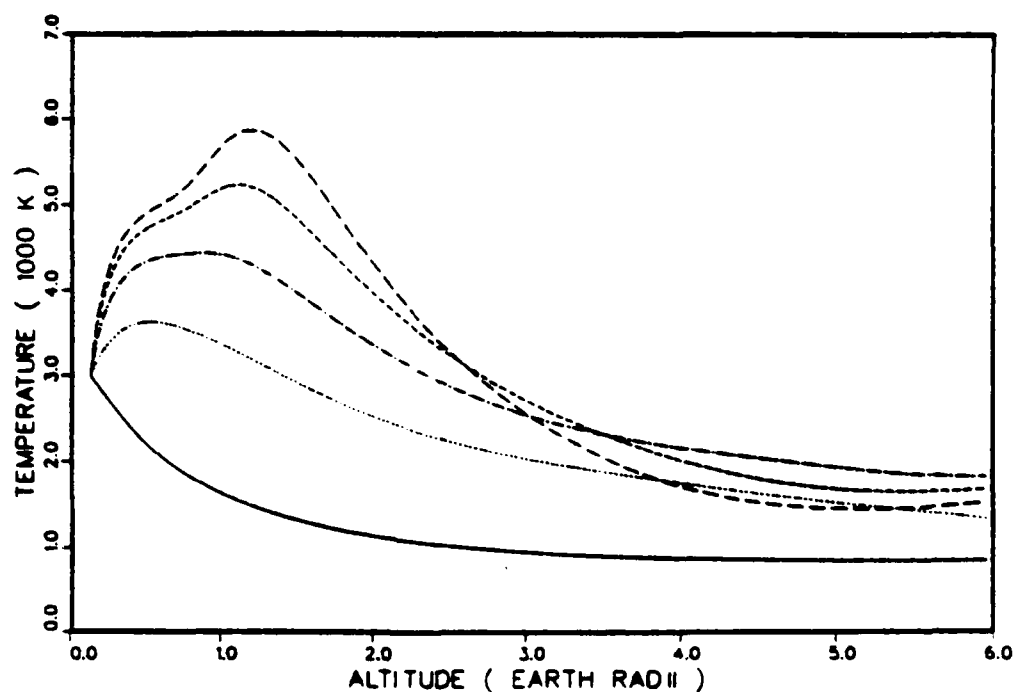


Fig. 3(b) — Perpendicular electron temperatures for the first ten minutes after current onset. (Same legend as Fig. 3a)

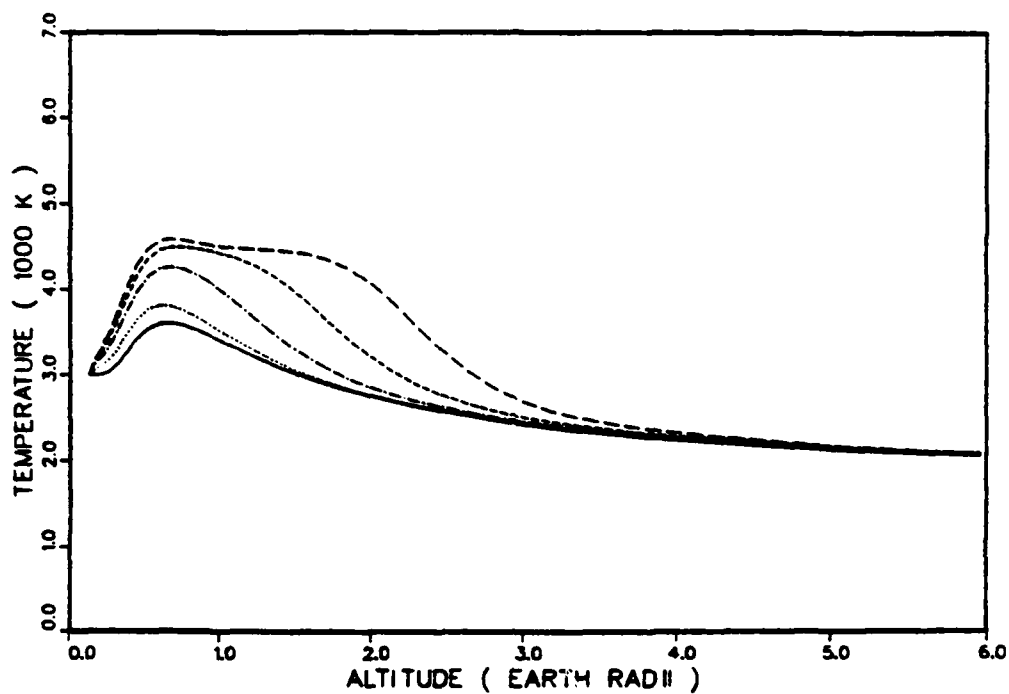


Fig. 4(a) — Parallel hydrogen ion temperatures for the first ten minutes after current onset. (Same legend as Fig. 3a.)

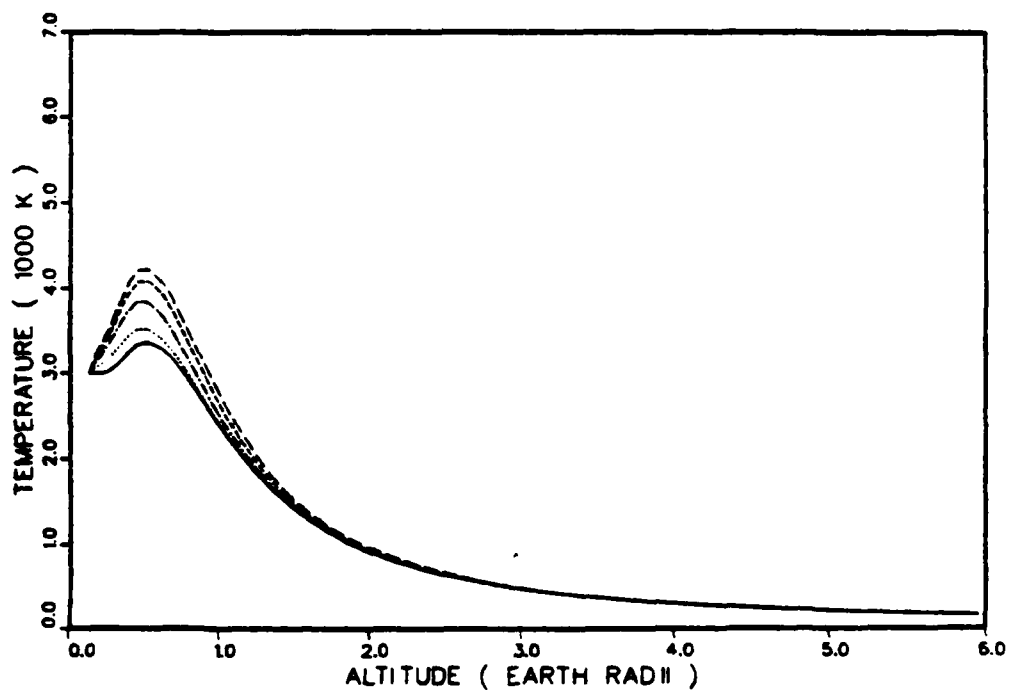


Fig. 4(b) — Perpendicular hydrogen ion temperatures for the first ten minutes after current onset. (Same legend as Fig. 3a.)

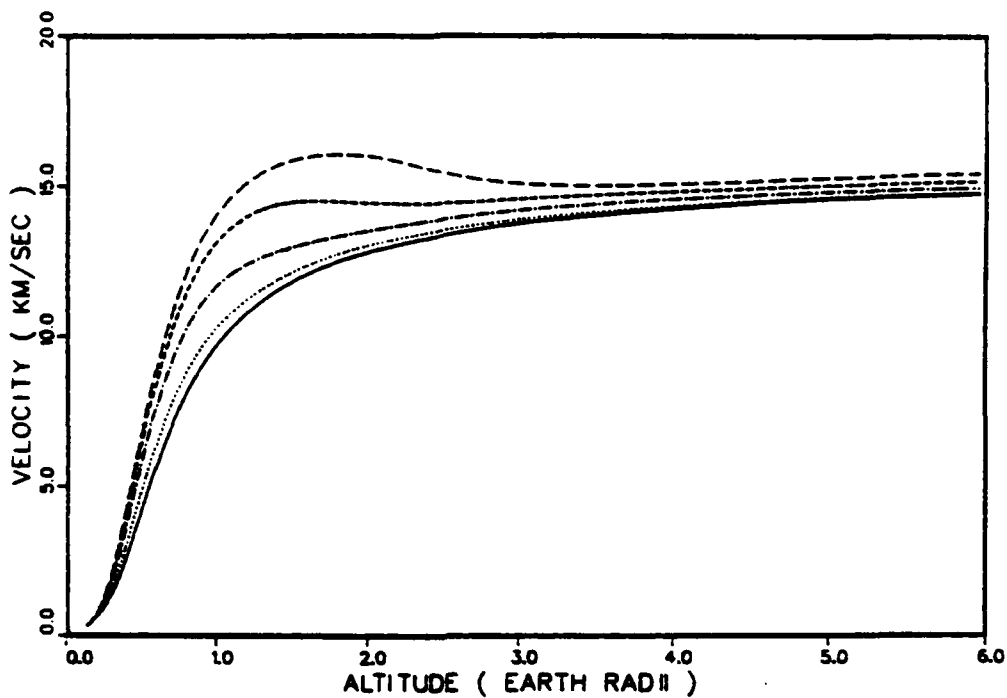


Fig. 5 — Hydrogen ion velocities for the first ten minutes after current onset. (Same legend as Fig. 3a.)

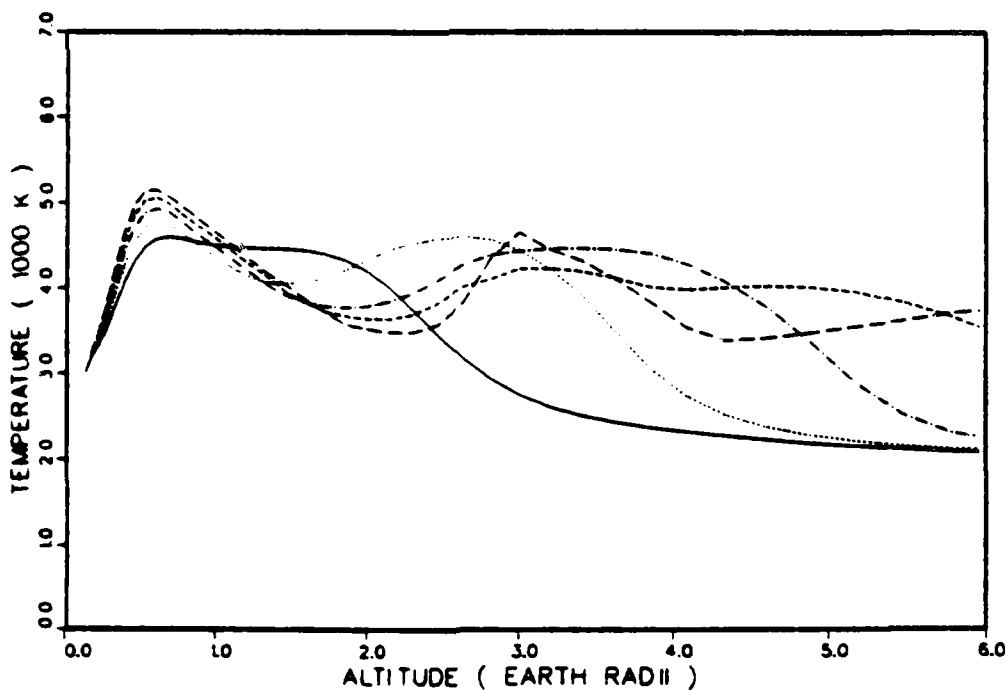


Fig. 6 — Parallel hydrogen ion temperatures for the times from ten minutes to thirty minutes after current onset: 10 min (solid curve), 15 min (dotted curve), 20 min. (dot-dash curve), 25 min. (short dashed curve), 30 min (long dashed curve).

the simulation continues, such thermal oscillations are seen in both the hydrogen ion and electron temperature profiles until, at about one hour into the simulation, the electron thermal waves have propagated through the upper boundary layer and affected the upper boundary cell. At this point the simulation began to exhibit a form of thermal runaway due to the interaction between the precipitation electron distribution and the outward heat flow at the upper boundary, and the simulation was halted. We do not view this failure of the upper boundary condition as a major restriction on the validity of the simulation because the one-dimensional field line plasma and constant current approximations would not be accurate for a simulation which lasted much longer than this.

The point that we wish to stress here is that the flux tube plasma seems to have a rapid initial response to the onset of current, one affecting the electron temperature predominantly. After this response, it is not clear that the plasma "settles down" to some equilibrium configuration in a reasonable time. Large electron and ion thermal waves can persist for relatively long times when the heat flows are treated dynamically. The preliminary nature of our field-aligned current simulations prevents us from describing these results in too much detail. One of our future goals is to use this model to investigate the behavior of these thermal waves for a range of currents and upper boundary layer configurations.

Summary

We have described here some of the preliminary results of our model for the behavior of the auroral field line plasma. Our results for equilibrium in the current-free case are in reasonable agreement with the previous work in this area, illustrating the important role that the upper boundary layer plays in determining the temperature profile of the electron population in contrast to the relative insensitivity of the supersonic hydrogen ion population to this layer. The early simulations with current which we have performed suggest that, although much of the flux tube dynamics depends on the energy transfer among particle species in the collisional regions of the tube, the time scale on which the field line plasma achieves equilibrium is highly dependent on the behavior of the

energy transport in the collisionless region. In fact, there may exist no equilibrium on the field line in the presence of a current varying on the time scale of interest in this problem.

The results presented here are meant to illustrate the kinds of behavior the field line model manifests and to point the way for future improvements in the model. Because of the importance of the collisional regime of the topside ionosphere in the dynamics of the entire field line, one of our first steps will be to investigate the manner in which changes in the details of this regime affect the behavior of the field line simulations. Also, we feel that there is much more to be done in order to understand the effect the upper boundary layer has on the flux tube plasma. It is our intention to proceed in these areas in addition to our main goal, that of investigating the effect of anomalous resistivity and electric field mechanisms in the collisionless region of our model.

Acknowledgement

This work was supported by the Office of Naval Research and by NASA Contract No. W-14365.

References

- Banks, P.M., and T.E. Holzer, The polar wind, J. Geophys. Res., 73, 6846-6854, 1968.
- Burgers, J.M., Flow Equations for Composite Gases, Academic Press, New York, 1969.
- Chiu, Y.T., and J.M. Cornwall, Electrostatic model of a quiet auroral arc, J. Geophys. Res., 85, 543-556, 1980.
- Frank, L.A., Hot plasmas in the earth's magnetosphere, in Physics of Solar Planetary Environments, edited by D.J. Williams, pp. 685-700, AGU, Washington, D.C., 1976.
- Gurnett, D.A., The earth as a radio source: terrestrial kilometric radiation, J. Geophys. Res., 79, 4227-4238, 1974.
- Hain, K., The partial donor cell method, NRL Memorandum Report 3713, Naval Research Laboratory, Washington, D.C., 1978.
- Holzer, T.E., J.A. Fedder, and P.M. Banks, A comparison of kinetic and hydrodynamic models of an expanding ion-exosphere, J. Geophys. Res., 76, 2453-2468, 1971.
- Kan, J.R., Energization of auroral electrons by electrostatic shock waves, J. Geophys. Res., 80, 2089-2095, 1975.
- Lemaire, J., and M. Scherer, Kinetic models of the solar and polar winds, Rev. Geophys. Space Phys., 11, 427-468, 1973.
- Lyons, L.R., Generation of large-scale regions of auroral current electric potentials, and precipitation by the divergence of the convection electric field, J. Geophys. Res., 85, 17-24, 1980.

- Mozer, F.S., Anomalous resistivity and parallel electric fields, in Magnetospheric Particles and Fields, edited by B.M. McCormac, pp. 125-136, D. Reidel, Dordrecht, Netherlands, 1976.
- Papadopoulos, K., A review of anomalous resistivity for the ionosphere, Rev. Geophys. Space Phys., 15, 113-127, 1977.
- Schunk, R.W., Mathematical structure of transport equations for multispecies flows, Rev. Geophys. Space Phys., 15, 429-445, 1977.
- Schunk, R.W., and D.S. Watkins, Electron temperature anisotropy in the polar wind, J. Geophys. Res., 86, 91-102, 1981.
- Schunk, R.W., and D.S. Watkins, Proton temperature anisotropy in the polar wind, J. Geophys. Res., 87, 171-180, 1982.
- Shawhan, S.D., C.-G. Falthammar, and L.P. Block, On the nature of large auroral zone electric fields at 1- R_e altitude, J. Geophys. Res., 83, 1049-1054, 1978.
- Stern, D.P., The electric field and global electrodynamics of the magnetosphere, Rev. Geophys. Space Phys., 17, 626-640, 1979.
- Swift, D.W., On the formation of auroral arcs and the acceleration of auroral electrons, J. Geophys. Res., 80, 2096-2108, 1975.

APPENDIX I

LINEAR THEORY OF THE $E \times B$ INSTABILITY WITH
AN INHOMOGENEOUS ELECTRIC FIELD

Linear Theory of the $E \times B$ Instability With an Inhomogeneous Electric Field

J. D. HUBA AND S. L. OSSAKOW

*Geophysical and Plasma Dynamics Branch, Plasma Physics Division
Naval Research Laboratory, Washington, D. C. 20375*

P. SATYANARAYANA

*Berkeley Scholars, Inc., Springfield, Virginia 22150*P. N. GUZDAR¹*Science Applications, Inc., McLean, Virginia 22102*

A general linear theory of the $E \times B$ instability is developed which considers an ambient electric field that is at an arbitrary angle to the density gradient and allows the electric field component parallel to the density gradient to be inhomogeneous. A differential equation is derived which describes the mode structure of the unstable waves in the direction of the inhomogeneities. The theory (1) includes ion inertia effects, (2) allows for arbitrary density and electric field profiles, and (3) is valid in the long-wavelength regime, i.e., $k_z L < 1$, where L is the width of the boundary layer. The main results of the analysis are as follows. First, the inhomogeneous velocity flow caused by the inhomogeneous electric field can stabilize the instability. Second, short-wavelength modes are preferentially stabilized over longer-wavelength modes. Third, the stabilization mechanism is associated with the velocity shear due to an x -dependent resonance $[\omega - k_z V_z(x)]^{-1}$, where $V_z(x) = -cE_z(x)/B$, and not velocity shear terms explicitly proportional to $\partial V_z/\partial x$ or $\partial^2 V_z/\partial x^2$. Fourth, the marginal stability criterion is weakly dependent on the magnitude of v_{in}/ω . Applications of these results to ionospheric phenomena are discussed, viz., barium cloud striations and high-latitude F region irregularities.

1. INTRODUCTION

An important instability associated with the structuring of ionospheric plasmas (e.g., high-latitude F region and barium clouds) is the $E \times B$ instability, also known as the gradient drift instability. The instability is an interchange instability which can occur in an inhomogeneous, weakly collisional, magnetized plasma that contains an ambient electric field orthogonal to both the ambient magnetic field and the density gradient. A simple physical picture of the instability mechanism is shown in Figure 1. We consider a plasma such that $\mathbf{B} = B\hat{e}_z$, $\mathbf{E} = E\hat{e}_x$, $n = n(x)$ with $\partial n/\partial x > 0$ and $v_{in}/\Omega_i \ll 1$, where v_{in} is the collision frequency between species α and neutrals, and Ω_α is the cyclotron frequency of species α . Upon this plasma we impose a density perturbation $\delta n \sim \delta n \sin(k_y y)$, as shown in Figure 1. The influence of $\mathbf{E}$ on the plasma is to cause (1) the electrons and ions to $\mathbf{E} \times \mathbf{B}$ drift in the x direction and (2) an ion Pedersen drift in the y direction. The latter effect induces a space charge perturbation electric field denoted by δE . The response of the plasma to this perturbed electric field is to drift with a velocity $\delta \mathbf{V} = c\delta E \times \mathbf{B}/B^2$. For the configuration shown in Figure 1, $\delta \mathbf{V}$ causes the 'heavy' fluid perturbation to fall into the 'light' fluid (region I) and the 'light' fluid perturbation to rise into the 'heavy' fluid (region II): the classic interchange phenomenon. Of course, if the direction of $\partial n/\partial x$ or $\mathbf{E}$ were reversed then the density perturbation would be damped.

The original study of the $E \times B$ instability was by Simon

[1963] and Hoh [1963], who applied it to laboratory gas discharge experiments. Subsequent to these first investigations, a considerable amount of research has been devoted to explaining ionospheric phenomena based upon this instability [Linson and Workman, 1970 and references therein; Simon, 1970; Völk and Haerendel, 1971; Perkins et al., 1973; Zabusky et al., 1973; Shiau and Simon, 1972; Perkins and Doles, 1975; Scannapieco et al., 1976; Chaturvedi and Ossakow, 1979; Keskinen and Ossakow, 1982]. Two areas of present interest concerning the instability are barium cloud striations (see, for example, the review papers by Ossakow [1979] and Ossakow et al. [1982] and the references therein) and the structuring of plasma 'blobs' in the high-latitude F region [Vickrey et al., 1980; Keskinen and Ossakow, 1982].

The purpose of this paper is to present a general theory of the $E \times B$ instability which considers an ambient electric field at an arbitrary angle to the density gradient and allows the electric field component parallel to the density gradient to be inhomogeneous. Some aspects of the problem have been treated by Perkins et al. [1973] and Perkins and Doles [1975]. Perkins and Doles [1975] made the important discovery that the sheared velocity flow (resulting from an inhomogeneous electric field parallel to the density gradient) can stabilize the instability. Furthermore, short-wavelength modes are preferentially stabilized over longer-wavelength modes. The work of Perkins and Doles [1975] considered the strong collision limit ($v_{in} \gg \omega$), assumed a specific density profile amenable to analytical theory, and is valid only in the short-wavelength regime, i.e., $k_z L \gg 1$, where L is the scale length of the boundary layer. The present study extends the theory of Perkins and Doles [1975] by removing these restrictions. Namely, we derive a differential equation which describes the mode structure of the $E \times B$ instability. Ion inertia effects are included so that the ratio v_{in}/ω is arbitrary.

¹ Permanently at University of Maryland, College Park, Maryland 20742

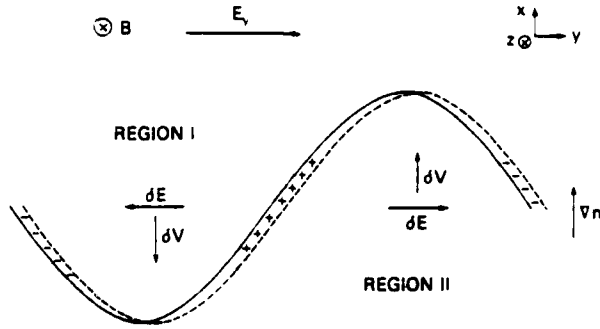


Fig. 1. Schematic of the physical mechanism of the $E \times B$ instability.

Moreover, we solve this equation numerically so that arbitrary density and electric field profiles can be considered and the regime $k_y L < 1$ can be investigated self-consistently.

The principal results of this work are the following:

1. The basic conclusions of Perkins and Doles [1975] are verified numerically. Specifically, the marginal stability criterion they derive analytically agrees well with our numerical result.
2. The marginal stability criterion is weakly dependent upon the magnitude of v_{in}/ω .
3. The stabilization mechanism is associated with the velocity shear due to an the x -dependent resonance $[\omega - k_y V_e(x)]^{-1}$, where $V_e(x) = -cE_x(x)/B$, and not velocity shear terms explicitly proportional to $\partial V_e/\partial x$ or $\partial^2 V_e/\partial x^2$.
4. When $E_x(x_0) \approx E_v$, where x_0 is the position about which the mode is localized, the most unstable modes have $k_y L \approx 1$.

The organization of the paper is as follows. In section 2 we derive the mode structure equation for the $E \times B$ instability. In section 3 we present both analytical and numerical results based upon this equation. Finally, in section 4 we summarize our results and discuss applications to ionospheric phenomena, i.e., barium cloud striations and high-latitude F region irregularities.

2. THEORY

The equilibrium configuration used in the analysis is shown in Figure 2. The ambient magnetic and electric fields are in the z direction and the xy plane, respectively, where $\mathbf{B} = B\hat{e}_z$ and $\mathbf{E} = E_x(x)\hat{e}_x + E_y\hat{e}_y$. The electric field in the y direction is constant, while the electric field in the x direction is allowed to be a function of x . This gives rise to an inhomogeneous velocity flow in the y direction, i.e., $V_y(x) = -cE_x(x)/B$. The density is taken to be inhomogeneous in the x direction ($n = n(x)$), and temperature effects are ignored.

The basic assumptions used in the analysis are as follows. We assume that the perturbed quantities vary as $\delta p \sim \delta p(x) \exp[i(k_y y - \omega t)]$, where k_y is the wave number along the y direction and $\omega = \omega_r + i\gamma$, implying growth for $\gamma > 0$. The ordering in the frequencies is such that $\omega \ll \Omega$, and $v_{in} \ll \Omega$ (the F region approximation), where v_{in} is the ion-neutral collision frequency and Ω is the ion gyrofrequency. We neglect terms of order ω/Ω and v_{in}/Ω , but retain terms of order v_{in}/ω . We ignore finite gyroradius effects by limiting the wavelength domain to $kr_L \ll 1$, where r_L is the mean ion Larmor radius. We neglect perturbations along the magnetic

field ($k_z = 0$) so that only the two-dimensional mode structure in the xy plane is obtained. We retain ion inertial effects, thereby including the ion polarization drift, but ignore electron inertia.

A key feature of our analysis is that a nonlocal theory is developed. That is, the mode structure of the potential in the x direction, the direction in which density and the $E \times B$ velocity are assumed to vary, is determined by a differential equation rather than an algebraic equation obtained by Fourier analysis. This is crucial to the analysis, since Perkins and Doles [1975] have shown that a nonlocal analysis is necessary to demonstrate the stabilizing influence of velocity shear due to the inhomogeneous electric field parallel to the density gradient.

The fundamental equations used in the analysis are continuity and momentum transfer:

$$\frac{\partial n_\alpha}{\partial t} + \nabla \cdot (n_\alpha \mathbf{V}_\alpha) = 0 \quad (1)$$

$$0 = -\frac{e}{m_e} \left(\mathbf{E} - \frac{1}{c} \mathbf{V}_e \times \mathbf{B} \right) \quad (2)$$

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_e \cdot \nabla \right) \mathbf{V}_i = \frac{e}{m_i} \left(\mathbf{E} - \frac{1}{c} \mathbf{V}_e \times \mathbf{B} \right) - \nu_{in} \mathbf{V}_i \quad (3)$$

where α denotes species (e , electrons; i , ions) and other variables have their usual meaning. Note that electron inertia terms are neglected but ion inertia terms are included and that the neutral wind is assumed to be zero. The equilibrium drifts are

$$\mathbf{V}_e = -cE_x(x)/B\hat{e}_y \quad (4)$$

$$\mathbf{V}_i = \left\{ (\nu_{in}/\Omega_i) cE_x(x)/B - [cE'_x(x)/B\Omega_i] cE_y/B \right\} \hat{e}_x + [-cE_y(x)/B - (\nu_{in}/\Omega_i) cE_x(x)/B] \hat{e}_y \quad (5)$$

where we have chosen a reference frame such that $V'_x = V_x - cE_y/B$, $\Omega_i = eB/m_i c$ and $E'_x(x) = \partial E_x/\partial x$. The equilibrium drift $\mathbf{V}_i$ (equation (5)) is valid up to first order in the quantities ν_{in}/Ω and $cE'_x L/\Omega B$, where $E'_x \sim E_x/L$ and L represents the scale length of the inhomogeneity in E_x .

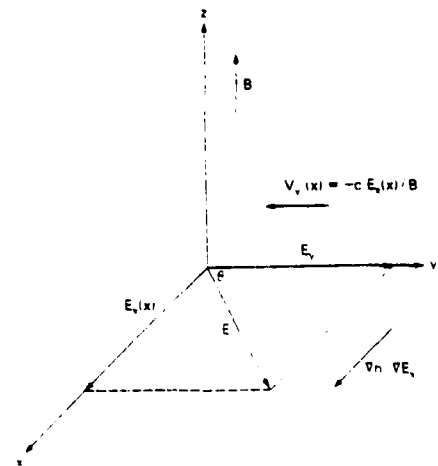


Fig. 2. Plasma and field configuration used in the analysis.

Terms of higher order $[(v_{in}/\Omega_i)^2, (v_{in}/\Omega_i)(cE_x/L\Omega_i B), (cE_x/L\Omega_i B)^2]$ are safely neglected since the time scale for the instability is on a faster time scale, i.e., $\omega \sim cE/B L$.

A relationship between $n(x)$ and $E_x(x)$ can be derived by assuming $\nabla \cdot \mathbf{J} = \nabla \cdot [n(\mathbf{V}_i - \mathbf{V}_e)] = 0$ which reduces to

$$\nabla \cdot (n \mathbf{V}_i) = 0 \quad (6)$$

Equation (6) leads to

$$n(x) [(v_{in}/\Omega_i)(cE_x(x)/B) - (cE_x'(x)/B\Omega_i)(cE_x/B)] = \text{const} \quad (7)$$

where we take the constant to be the left-hand side of (7) evaluated at $x = -\infty$. Thus, by specifying the density profile, the electric field profile $E_x(x)$ can be determined from (7). Of course, if there are sources and/or sinks in the plasma such that $\nabla \cdot (n \mathbf{V}_i) \neq 0$, then (7) is not applicable.

We now consider a linear perturbation analysis of (1)-(3). We assume $n_a = n_0 + \delta n_a$, $\mathbf{V}_a = \mathbf{V}_0 + \delta \mathbf{V}_a$ and $\mathbf{E} = \mathbf{E} - \nabla \phi$, where ϕ is the perturbed electrostatic potential. Using (2), we obtain

$$\delta \mathbf{V}_{e1} = -ik_y \phi c/B \quad (8)$$

$$\delta \mathbf{V}_{e1} = \phi'(c/B) \quad (9)$$

The perturbed ion velocity is derived to first order in the quantities v_{in}/Ω_i , ω/Ω_i , $k_y V_{in}/\Omega_i$, and $V_{in}/L\Omega_i$. We assume $\delta \mathbf{V}_i = \delta \mathbf{V}_{i1} + \delta \mathbf{V}_{i2}$, where $\delta \mathbf{V}_{i1}$ is the lowest-order perturbed velocity and $\delta \mathbf{V}_{i2}$ is the first-order perturbed velocity. Using (3) and assuming $v_{in}/\Omega_i \ll 1$ and $(d/dt)/\Omega_i \ll 1$, we find that

$$\delta \mathbf{V}_{i1} = -\frac{c}{B^2} \nabla \phi \times \mathbf{B} \quad (10)$$

We now iterate on (3) to obtain

$$\delta \mathbf{V}_{i2} = -\frac{1}{\Omega_i} \frac{1}{B} \left(\frac{\partial}{\partial t} + \mathbf{V} \cdot \nabla + v_{in} \right) \mathbf{V} \times \mathbf{B} \quad (11)$$

where $\mathbf{V} = \mathbf{V}_i + \delta \mathbf{V}_{i1}$. Making use of (5) and (10) in (11), we finally arrive at

$$\delta \mathbf{V}_{i2} = [-ik_y(1 - V_{in}'/\Omega_i)\phi - i(\omega/\Omega_i)\phi' - (V_{in}/\Omega_i)\phi''](c/B) \quad (12)$$

$$\delta \mathbf{V}_{i2} = [-k_y(\omega/\Omega_i + iV_{in}'/\Omega_i)\phi - (1 - ik_y V_{in}'/\Omega_i)\phi'](c/B) \quad (13)$$

where $\tilde{\omega} = \omega + i\nu_{in} - k_y V_{in}$, $V_{in}' = \partial V_{in}/\partial x$, $V_{in}'' = \partial^2 V_{in}/\partial x^2$, $\phi' = \partial \phi/\partial x$, and $\phi'' = \partial^2 \phi/\partial x^2$. Substituting (8), (9), (12), and (13) into (1), one can show that for the ions,

$$\begin{aligned} & -\frac{B}{c} \frac{\Omega_i}{\tilde{\omega}} \left(\omega^* \frac{\delta n_i}{n} + iV_{in}'' \frac{\delta n_i}{n} \right) + i \frac{V_{in}}{\tilde{\omega}} \phi''' \\ & - \left(1 - i \frac{n'}{n} \frac{V_{in}}{\tilde{\omega}} \right) \phi'' - \left[-ik_y^2 \frac{V_{in}}{\tilde{\omega}} - \frac{n'}{n} \right. \\ & \left. - i \frac{V_{in}}{\Omega_i} \left(\frac{n'}{n} \right)' \right] \phi - \left[-k_y^2 \left(1 - \frac{V_{in}}{\tilde{\omega}} \right) - \frac{k_y V_{in}''}{\tilde{\omega}} \right. \\ & \left. - \frac{n}{n} \frac{k_y V_{in}}{\tilde{\omega}} - \frac{n}{n} \frac{k_y \Omega_i}{\tilde{\omega}} \right] \phi = 0 \end{aligned} \quad (14)$$

and for the electrons,

$$\frac{\delta n_e}{n} = -\frac{c}{B} \frac{k \phi}{(\omega - k_y V_{e1})} \frac{n}{n} \quad (15)$$

where $\omega^* = \omega - k_y V_{e1} - iV_{e1}'$ and the superscripts (prime, double primes, and triple primes) indicate first, second, and third derivatives with respect to x , respectively. We note again that we have neglected higher-order terms in (14), i.e., terms of order $(v_{in}/\Omega_i)^2$, $(v_{in}/\Omega_i)(cE_x/L\Omega_i)$, and $(cE_x/L\Omega_i)^2$, where L is the scale length of the inhomogeneity in E_x .

We assume quasineutrality and take $\delta n_e = \delta n_i$. The following equation is then obtained from (12) and (13):

$$\begin{aligned} & i \frac{V_{in}}{\tilde{\omega}} \phi''' + \left(1 - i \frac{n'}{n} \frac{V_{in}}{\tilde{\omega}} \right) \phi'' + \left\{ \frac{n'}{n} - i \frac{k_y V_{in}}{\tilde{\omega}} \right\} \left[-k_y \right. \\ & \left. - \frac{\tilde{\omega}}{\Omega_i} \frac{1}{k_y} \left(\frac{n'}{n} \right)' - \frac{\Omega_i}{\omega - k_y V_{e1}} \frac{n'}{n} \right] \phi' - \left[-k_y^2 \left(1 - \frac{V_{in}}{\tilde{\omega}} \right) \right. \\ & \left. - \frac{k_y V_{in}''}{\tilde{\omega}} - \frac{n'}{n} \frac{k_y V_{in}'}{\tilde{\omega}} - i \frac{k_y V_{in}}{\omega - k_y V_{e1}} \frac{\Omega_i}{\tilde{\omega}} \frac{n'}{n} \right. \\ & \left. - i \frac{k_y V_{e1}'}{\omega - k_y V_{e1}} \frac{n'}{n} \frac{v_{in}}{\tilde{\omega}} - \frac{\omega}{\omega - k_y V_{e1}} \right. \\ & \left. - \frac{v_{in}}{\tilde{\omega}} \frac{n'}{n} \frac{k_y}{\omega - k_y V_{e1}} k_y (cE_x/B) \right] \phi = 0 \end{aligned} \quad (16)$$

We simplify (16) by assuming the following ordering scheme: $\partial/\partial x \leq k_y$, $v_{in}/\Omega_i \ll 1$, $V'' = V/L$, $V' = V/L^2$, $k_y L \ll \Omega_i v_{in}$, $k_y L \ll \Omega_i/\omega$, where L is the scale length of the inhomogeneous plasma boundary layer. Equation (16) can now be written as

$$\begin{aligned} & \phi'' + \left[\frac{n'}{n} \left(1 - \frac{i\nu_{in}}{\tilde{\omega}} \frac{k_y V_{e1}}{\tilde{\omega} + i\nu_{in}} \right) \right] \phi' \\ & + \left\{ -k_y^2 - \frac{k_y (cE_x/B)}{\tilde{\omega} + i\nu_{in}} \frac{v_{in}}{\tilde{\omega}} \frac{k_y n'}{n} - \frac{n'}{n} \frac{k_y V_{e1}'}{\tilde{\omega} + i\nu_{in}} \right. \\ & \left. - \frac{k_y V_{e1}}{\tilde{\omega} + i\nu_{in}} \frac{i\nu_{in}}{\tilde{\omega}} \left[\frac{n''}{n} + \frac{V_{e1}'}{V_{e1}} \frac{n'}{n} \frac{\omega}{\tilde{\omega}} \right] \right\} \phi = 0 \end{aligned} \quad (17)$$

where $\tilde{\omega} = \omega - k_y V_{e1}(x) = \omega + k_y [cE_x(x)/B]$. Equation (17) describes the two-dimensional mode structure of ϕ for the $E \times B$ instability in a velocity sheared plasma for arbitrary v_{in}/ω .

3. RESULTS

Analytical Results

In general, (17) requires a numerical analysis for arbitrary density and electric field profiles. However, insight into the nature of the $E \times B$ instability can be gained by first considering several limiting cases.

Local theory. We first reduce the differential equation (equation (17)) to an algebraic equation by making use of local theory. That is, we let $\partial/\partial x \rightarrow ik_y$ and assume $k_y^2 L^2 \gg 1$ and $k_y^2 L^2 \gg 1$, where $L_i = (n/n')^{-1}$ is the scale length of the density inhomogeneity evaluated at $x = x_0$. For simplicity, we also take $E = E_0 \hat{x}$, $E_0 = \text{const}$. In this limit, (17)

becomes

$$k^2 L_n^2 - k_y L_n \frac{\nu_{in} \mathbf{k} \cdot \mathbf{E}(c/B)}{\bar{\omega} \bar{\omega} - i\nu_{in}} = 0 \quad (18)$$

where $\bar{\omega} = \omega - k_y V_{ey}$ and $V_{ey} = -cE_z/B$. Equation (16) has the solution

$$\bar{\omega} = -i \frac{\nu_{in}}{2} \left[1 + 4 \frac{k_y \mathbf{k} \cdot \mathbf{E}(c/B)}{k \nu_{in} k L_n} \right]^{1/2} \quad (19)$$

Instability occurs when

$$\frac{k_y \mathbf{k} \cdot \mathbf{E}}{k k L_n} > 0 \quad (20)$$

- The growth rates of the instability in the strong and weak collisional limits are, respectively,

$$\gamma = \frac{k_y \mathbf{k} \cdot \mathbf{E}(c/B)}{k k L_n} \quad \nu_{in} \gg \omega \quad (21)$$

$$\gamma = \frac{k_y}{k} \left[\frac{\mathbf{k} \cdot \mathbf{E}(c/B)}{k L_n} \nu_{in} \right]^{1/2} \quad \nu_{in} \ll \omega \quad (22)$$

with a real frequency $\omega = k_y V_{ey}$ in each case. Note that instability can occur for $E_z = 0$ as long as $k_y \neq 0$ and $E_x \neq 0$ (equation (20)) [see also Keskinen and Ossakow, 1982]. For $k = k_y$, one obtains the usual $E \times B$ gradient drift instability growth rate [Linson and Workman, 1970] in (21), and the so-called high-altitude limit [Ossakow et al., 1978] of the $E \times B$ gradient drift instability in (22).

Nonlocal theory. In deriving (18) the local approximation is used. That is, the dispersion equation is solved based upon the plasma parameters at a particular value of x , say $x = x_0$, usually where n'/n is a maximum; which leads to maximum growth. If we now assume $E_x = E_r(x)$ then a sheared $E \times B$ velocity flow arises $V_x = V_{ex}(x) = -cE_r(x)/B$. Applying local theory to this situation, one might expect that (16) is still valid with V_{ey} evaluated at x_0 , i.e., $V_{ey} = V_{ey}(x_0)$. Thus, (21) and (22) follow accordingly, but the real frequency is now given by $\omega_r = k_y V_{ey}(x_0)$. However, Perkins and Doles [1975] have shown, both analytically and using numerical simulations, that this is not the case. We do not reproduce their detailed analysis here but rather point out the important result of their work.

Perkins and Doles [1975] consider the strong collision limit ($\nu_{in} \gg \omega$) so that ion inertia terms can be neglected. Furthermore, they assume $\nabla \cdot (n\mathbf{V}) = 0$ which leads to

$$n(x)E_r(x) = n_0 E_{0r} = \text{const} \quad (23)$$

where $n_0 = n(x = -\infty)$ and $E_{0r} = E_r(x = -\infty)$. This is evident from (7) by noting that $E_r'(x) \sim E_r/L_n$ and $\nu_{in} \gg \omega \sim cE_z/B L_n$. In this limit, (17) reduces to

$$\phi'' - \left[\frac{n'}{n} \left(1 - \frac{k_y V_{ey}}{\bar{\omega}} \right) \right] \phi' - \left[-k_y^2 - i \frac{k_y (cE_z/B)}{\bar{\omega}} \frac{k_y n'}{n} - \frac{k_y V_{ey}}{\bar{\omega}} \left(\frac{n''}{n} - \frac{V_{ey}'}{V_{ey}} \frac{n'}{n} \frac{\omega}{\bar{\omega}} \right) \right] \phi = 0 \quad (24)$$

where $\bar{\omega} = \omega - k_y V_{ey}$, $V_{ey} = -(cE_z/B)(n_0/n(x))$ and $V_{ey}' = -(cE_z/B)(n_0/n(x))'$. Perkins and Doles [1975] expand (24) about $x = x_0$, where x_0 is the position of maximum n'/n by

taking

$$n'/n = [1 - (x - x_0)^2/D^2]/L_n \quad (25)$$

Assuming $k_y^2 L_n^2 \gg 1$ and $k_y^2 D^2 \gg 1$ and by making several variable changes, they solve (24) analytically. The important conclusion of their theory is that the $E \times B$ instability is stabilized when

$$E_r(x_0)/E_z > 2/k_y D \quad (26)$$

Thus the influence of velocity shear, i.e., an inhomogeneous E_x , is to stabilize preferentially the short-wavelength modes, those with $k_y D \gg 1$.

Numerical Results

In order to solve (17) numerically and also to gain insight into the nature of the solutions, we transform (17). First, we note that (17) is of the form

$$\phi'' + p(x)\phi' + q(x)\phi = 0 \quad (27)$$

where $p(x)$ and $q(x)$ are the coefficients of ϕ' and ϕ , respectively, in (17). We let

$$\phi = \tilde{\phi} \exp \left[-\frac{1}{2} \int p(s) ds \right] \quad (28)$$

Substituting (28) into (27), we find that the transformed equation is

$$\tilde{\phi}'' - Q(x)\tilde{\phi} = 0 \quad (29)$$

where

$$Q(x) = -q(x) + \frac{1}{2} p'(x) - \frac{1}{4} [p(x)]^2 \quad (30)$$

and

$$p(x) = \frac{n'}{n} \left(1 - \frac{i\nu_{in}}{\bar{\omega}} \frac{k_y V_{ey}}{\bar{\omega} + i\nu_{in}} \right) \quad (31)$$

$$q(x) = \left[\frac{n''}{n} - \left(\frac{n'}{n} \right)^2 \right] \left(1 - \frac{i\nu_{in}}{\bar{\omega}} \frac{k_y V_{ey}}{\bar{\omega} + i\nu_{in}} \right) - \frac{n'}{n} \frac{k_y V_{ey}'}{\bar{\omega} + i\nu_{in}} \frac{i\nu_{in}}{\bar{\omega}} \left(1 - \frac{k_y V_{ey}}{\bar{\omega} + i\nu_{in}} - \frac{k_y V_{ey}}{\bar{\omega}} \right) \quad (32)$$

$$q(x) = -k_y^2 - \frac{k_y (cE_z/B)}{\bar{\omega} + i\nu_{in}} \frac{\nu_{in} k_y n'}{n} - \frac{n'}{n} \frac{k_y V_{ey}'}{\bar{\omega} + i\nu_{in}} - \frac{k_y V_{ey}}{\bar{\omega} + i\nu_{in}} \frac{i\nu_{in}}{\bar{\omega}} \left(\frac{n''}{n} + \frac{V_{ey}'}{V_{ey}} \frac{n'}{n} \frac{\omega}{\bar{\omega}} \right) \quad (33)$$

Equation (29) has a simple form, albeit $Q(x)$ is a complicated function of x , which allows physical insight into the nature of the mode structure. As an example, if Q is real and has $Q > 0$ for $|x| > x_0$ and $Q < 0$ for $|x| < x_0$, then one would expect a bounded solution of $\tilde{\phi}$ in the region $|x| < x_0$ that exponentially decays for $|x| > x_0$. Equation (29) is solved numerically using a finite difference scheme (i.e., Numerov method [Gladd and Horton, 1973]) to obtain eigenvalues and eigenfunctions. The boundary conditions used are

$$\tilde{\phi}(x) = \frac{1}{Q^{1/4}(x)} \exp \left[i \int_{x_0}^x dy Q^{1/2}(y) \right] \quad x \rightarrow \pm \infty$$

The sign of the WKB solution in the above equation is

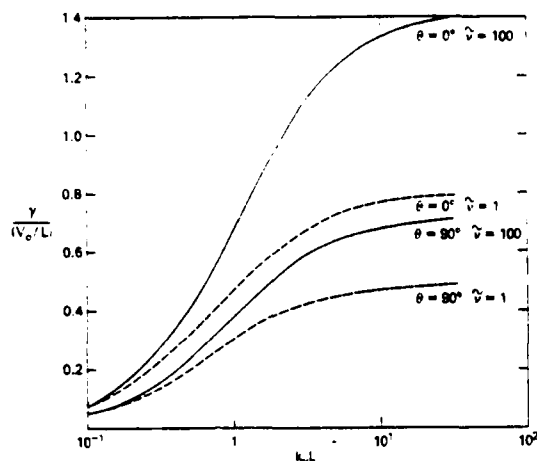


Fig. 3. Plot of $\gamma = \gamma(V_0/L)$ versus $k_y L$ for $\theta = 0^\circ$ and 90° and for $\bar{\nu} = \nu(V_0/L) = 1.0$ and 100.0 . The electric field E_x is assumed to be constant (equation (39)).

chosen such that a damped solution is obtained in the limit $|x| \rightarrow \infty$.

We now solve (29) numerically for a variety of conditions to better understand the influence of an inhomogeneous electric field on the $E \times B$ instability. In all of the cases presented, the following density profile is assumed

$$n(x) = n_0 \frac{1 + \varepsilon \tanh(x/L)}{1 - \varepsilon} \quad (34)$$

where $0 \leq \varepsilon < 1$, L characterizes the width of the boundary layer, and $n_0 = n(x = -\infty)$. By varying ε , the magnitude of the density gradient scale length L_n [$L_n = (n'/n)^{-1}$] can be changed. That is, as $\varepsilon \rightarrow 0$, $L_n \rightarrow \infty$ (a constant density profile); as $\varepsilon \rightarrow 1$, the value of $L_n \rightarrow 0$ (a rapidly changing density profile). We assume $\varepsilon = 0.95$ for the results presented so that

$$(n'/n)_{\max} = 1.45L^{-1} \quad \text{at } x/L = -0.9 \quad (35)$$

The maximum growth rate of the instability is expected to be

$$\gamma_m = 1.45(V_0/L) \quad (36)$$

where $V_0 = cE_0/B$ and we have used (21) assuming $E = E_0 \hat{e}_y$.

The ambient electric field is chosen to be

$$E(x) = E_x(x)\hat{e}_x + E_y\hat{e}_y \quad (37)$$

where

$$E(x = -\infty) = E_0 \sin \theta \hat{e}_x + E_0 \cos \theta \hat{e}_y \quad (38)$$

so that $\theta = \tan^{-1}(E_x/E_y)$ at $x = -\infty$. The influence of the x component of the electric field is then studied by varying θ , the angle between E and $\hat{e}_y$ at $x = -\infty$. Two forms of $E_x(x)$ are considered in the analysis:

$$E_x(x) = E_0 \sin \theta = \text{const} \quad (39)$$

$$E_x(x) = E_0 \sin \theta [n_0/n(x)] \neq \text{const} \quad (40)$$

These allow us to contrast the effects of no velocity shear and velocity shear on the instability. We comment that (40) is an equilibrium solution which satisfies $\nabla \cdot (nV_0) = 0$ in the strong collisional limit $\nu_m \gg \omega$ (i.e., equation (7)). We emphasize that (39) is not a self-consistent electric field

profile since $\nabla \cdot (n(V_0 - V_e)) \neq 0$ but present results based upon it for comparison to the self-consistent case (equation (40)).

In Figure 3 we plot $\gamma = \gamma(V_0/L)$ versus $k_y L$ for $\theta = 0^\circ$ and 90° and $\bar{\nu} = \nu(V_0/L) = 1.0$ and 100.0 , where E_y is chosen to be constant (equation (39)) and $V_0 = cE_0/B$. A general comment on all of the curves shown is that γ is an increasing function $k_y L$, but γ asymptotes to a constant value independent of $k_y L$ for $k_y^2 L^2 \gg 1$. This is consistent with the predictions of local theory. The 'standard' case is $\theta = 0^\circ$, that is, $E = E_y \hat{e}_y$, and there is no component of E parallel to the density gradient. For this case, two values of $\bar{\nu}$ are chosen: strong collisions ($\bar{\nu} = 100.0$) and weak collisions ($\bar{\nu} = 1.0$). As is expected, the growth rate is larger for the larger value of $\bar{\nu}$ in the short-wavelength regime ($k_y L > 1$). Also, the growth rate for $\bar{\nu} = 100.0$ at $k_y L = 30$ is $\gamma = 1.39$ and is still increasing, although slowly, as a function of $k_y L$. This value of γ agrees well with the value obtained from local theory ($\gamma = 1.45$ from (33)). The growth rate for the weak collision case $\bar{\nu} = 1.0$ asymptotes to a somewhat smaller value of γ ($\gamma = 0.79$). However, note that the difference between the growth rates for the strong and weak collisional cases becomes smaller as $k_y L \rightarrow 0$ and that the growth rates are, in fact, comparable for $k_y L \approx 0.1$. The 'nonstandard' case is $\theta = 90^\circ$, or $E = E_x \hat{e}_x$, and the only component of E is along the density gradient. The major result of this limit is simply that the instability can still persist even though $E_y = 0$. The overall influences of $\bar{\nu}$ and $k_y L$ on the instability are the same as in the previous case, $\theta = 90^\circ$.

In Figure 4 we plot γ versus $k_y L$ for $\theta = 0^\circ$ and 70° and $\bar{\nu} = 1.0$ and 100.0 , but consider E_x to be a function of x as in (40) so that velocity sheared flows occur for $\theta \neq 0^\circ$. The curves for $\bar{\nu} = 1.0$ and 100.0 and $\theta = 0^\circ$ are shown for comparative purposes. The important results in this figure are as follows. First, the mode is stable for $k_y L \geq 12$ for both the strong and weak collisional cases when $\theta = 70^\circ$. This is in agreement with the conclusion of Perkins and Doles [1975]: velocity shear effects tend to stabilize the short-wavelength modes, those such that $k_y^2 L^2 \gg 1$. The influence of shear on the long-wavelength modes ($k_y L < 1$) is weakly stabilizing.

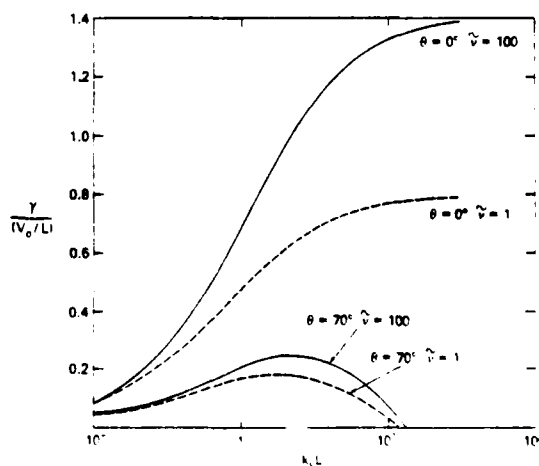


Fig. 4. Plot of $\gamma = \gamma(V_0/L)$ versus $k_y L$ for $\theta = 0^\circ$ and 70° and for $\bar{\nu} = \nu(V_0/L) = 1.0$ and 100.0 . The electric field E_x is assumed to be inhomogeneous (equation (40)).

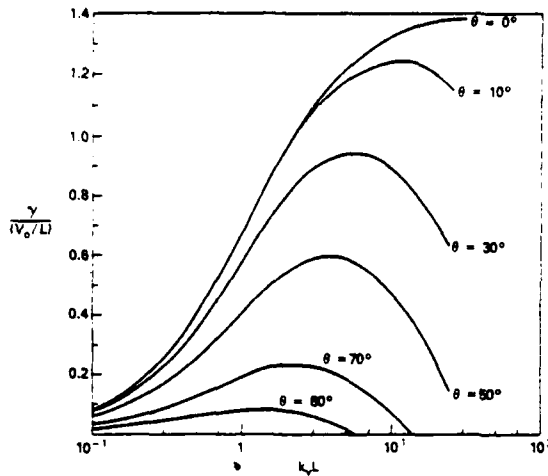


Fig. 5. Plot of $\gamma = \gamma(V_0/L)$ versus $k_y L$ for $\bar{\nu} = \nu_{te}/(V_0/L) = 100.0$ and $\theta = 0^\circ, 10^\circ, 30^\circ, 50^\circ, 70^\circ$, and 80° .

Second, the difference in the growth rate curves for $\bar{\nu} = 1.0$ and 100.0 is much less than that of the case of no shear (i.e., $\theta = 0$). And finally, since velocity shear can stabilize short-wavelength modes before long-wavelength modes, velocity shear causes a preferential wavelength to be excited based upon linear theory. For the case shown ($\theta = 70^\circ$), the linear growth maximizes at $k_y L = 2.0$. Thus velocity shear can lead to growth of long-wavelength modes ($k_y L \approx 1$) rather than short-wavelength modes ($k_y L \gg 1$).

Figure 5 is a plot of γ versus $k_y L$ for $\bar{\nu} = 100.0$ and $\theta = 0^\circ, 10^\circ, 30^\circ, 50^\circ, 70^\circ$, and 80° . It is clear that a peak in the linear growth rate occurs because of velocity shear. Moreover, as θ increases, the position of maximum growth decreases as a function of $k_y L$. For $\theta = 10^\circ$, maximum growth occurs for $k_y L = 20$, while for $\theta = 80^\circ$, maximum growth occurs for $k_y L = 1.0$.

To illustrate further the influence of velocity shear on the $E \times B$ instability we present Figure 6, which is a plot of γ versus θ for $\bar{\nu} = 100.0$, $k_y L = 1.0$ and 10.0 , and E_z given by both (39) and (40). First, the case of $E_z = \text{const}$ shows that instability persists over the entire range of θ ($\theta = 0^\circ$ to $\theta = 90^\circ$) for $k_y L = 1.0$ and 10.0 . Also, γ is a decreasing function of

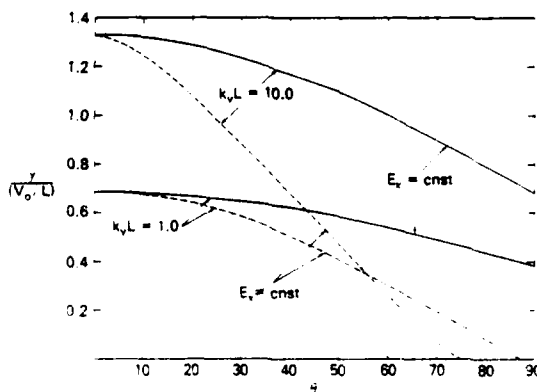


Fig. 6. Plot of $\gamma = \gamma(V_0/L)$ versus θ for $\bar{\nu} = \nu_{te}/(V_0/L) = 100.0$, $k_y L = 1.0$ and 10.0 , and $E_z = \text{const}$ (equation (39)) and $E_z = \text{const}$ (equation (40)).

θ . Second, for the case $E_z \neq \text{const}$, the mode is stabilized at a finite value at θ such that $\theta < 90^\circ$. Note that stabilization of the shorter-wavelength mode ($k_y L = 10.0$) occurs at $\theta = 74^\circ$, while stabilization of the longer-wavelength mode ($k_y L = 1.0$) occurs at $\theta = 88^\circ$. Finally, similar qualitative results are obtained for the weak collision limit ($\bar{\nu} = 1.0$), although quantitative differences occur.

As is evident from the preceding discussion, an inhomogeneous electric field has a dramatic effect on the $E \times B$ instability. However, (29) is rather complicated, and a question that naturally arises is which term (or terms) in (29) is responsible for stabilization of the mode? To shed light on this question, we consider the following simplified equation:

$$\phi'' - \left[k_y^2 - ik_y \frac{k_y (cE_z/B)}{\omega - k_y V_{ex}(x)} \frac{n'}{n} \right] \phi = 0 \quad (41)$$

That is, we consider the limit $\bar{\nu} \gg 1$ and retain the x -dependent, Doppler-shifted frequency $[\omega - k_y V_{ex}(x)]$ as the only contribution of the inhomogeneous electric field profile, i.e., velocity shear. We neglect terms explicitly proportional to V_{ex}' , V_{ex}'' , and n'' . We emphasize that (41) is not the complete mode structure equation, but is solved and contrasted to the correct solution in order to isolate a single effect of the field inhomogeneity, viz. the x -dependent resonance $\omega - k_y V_{ex}(x)$. In Figure 7 we plot γ versus $k_y L$ for $\theta = 70^\circ$, and E_z is given by (40). The solid curve is the solution to (29) for $\bar{\nu} = 100.0$, while the dashed curve is the solution to (41). Although there is a small difference between these curves for $k_y L < 1$, the important point is that the mode is stabilized at $k_y L \approx 13$ in both cases. Thus the stabilization mechanism is related to the x -dependent resonance $[\omega - k_y V_{ex}(x)]^{-1}$, as opposed to velocity shear effects associated with terms explicitly proportional to V_{ex}' and V_{ex}'' . This is a key result of this analysis.

We now turn our attention to the mode structure associated with the $E \times B$ instability, and the influence of an inhomogeneous electric field on its structure. Figure 8 is a plot of the density profile $n(x)/n_0$ (equation (34)) with $\epsilon = 0.95$ and the electric field profile $E_z(x)/E_0$ (equation (40)) with $E_0 = E_0 \sin \theta$ versus x/L . For $\theta = 0^\circ$, the electric field profile is simply $E_z(x)/E_0 = 0$. We present plots of Q and ϕ versus x/L for these profiles. In the subsequent plots of Q and ϕ , the subscript r denotes the real part of Q or ϕ , and i

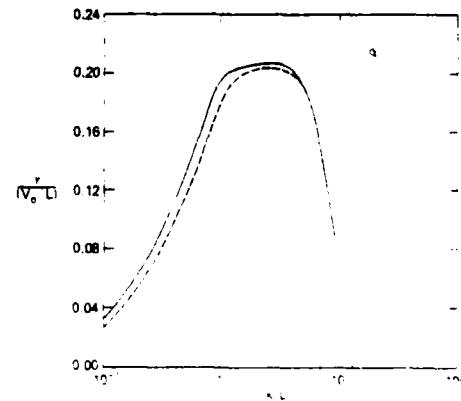


Fig. 7. Plot of $\gamma = \gamma(V_0/L)$ versus $k_y L$ for $\bar{\nu} = \nu_{te}/(V_0/L) = 100.0$ and $\theta = 70^\circ$ using (29) (solid curve) and (41) (dashed curve). The electric field E_z is assumed to be inhomogeneous (equation (40)).

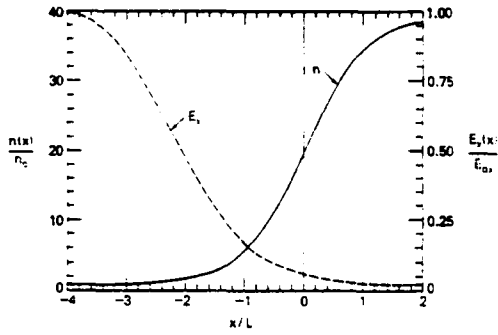


Fig. 8. Equilibrium density (equation (34)) and electric field (equation (40)) profiles for $\epsilon = 0.95$.

denotes the imaginary part of Q or ϕ . The parameters considered for the first set of modes are $k_y L = 10.0$, $\bar{\nu} = 100.0$, and $\theta = 0^\circ$ (Figure 9) and $\theta = 70^\circ$ (Figure 10). These modes are considered to be short-wavelength modes since $k_y L \gg 1$.

Figure 9 is a plot of Q (Figure 9a) and ϕ (Figure 9b) versus x/L for the case of no electric field inhomogeneity ($\theta = 0^\circ$ or $E_r(x) = 0$). The eigenfrequency for the mode is $\bar{\omega}_r = 0.0$ and $\bar{\gamma} = 1.329$. The important points to note are the following. First, the wave potential Q is real and is such that $Q < 0$ for $-1.3 < x/L < -0.5$ and $Q > 0$ otherwise. Second, the wave potential Q achieves a minimum at $x/L \approx -0.9$, the position of maximum L_n (equation (35)). Third, consistent with this form of Q , the wave function ϕ is a bounded mode centered about $x/L \approx -0.9$ that falls off exponentially for $x/L > -0.5$ and $x/L < -1.3$. Finally, the wave function is reasonably broad in that its half width at half maximum (Δx) is comparable to the width of the boundary layer, i.e., $\Delta x \approx L/2$.

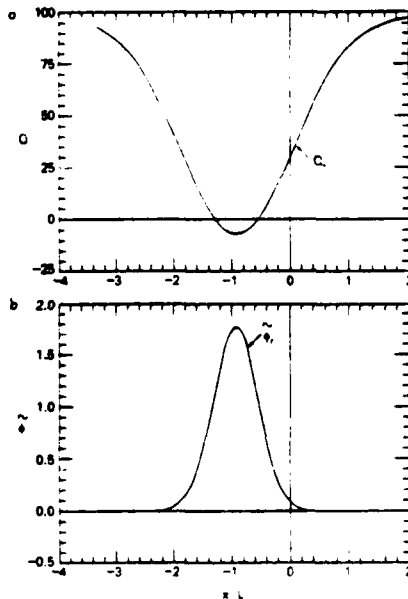


Fig. 9. Wave potential Q and wave eigenfunction ϕ as a function of x/L . The subscripts r and i denote real and imaginary, respectively. The parameters considered are $k_y L = 10.0$, $\bar{\nu} = 100.0$, $(V_{\infty}/L) = 100.0$ and $\theta = 0$ (i.e., $E_r = 0$). The eigenfrequency is $\bar{\omega}_r = 0.0$ and $\bar{\gamma} = 1.329$. (a) Q versus x/L . (b) ϕ versus x/L .

Figure 10 is a plot of Q (Figure 10a) and ϕ (Figure 10b) versus x/L for the case of an inhomogeneous electric field ($\theta = 70^\circ$ and $E_r(x)$ is shown in Figure 8). The eigenfrequency for this case is $\bar{\omega}_r = 0.5307$ and $\bar{\gamma} = 0.0716$. Note that the mode has a real frequency in contrast to the previous case and that the growth rate is smaller. Other important differences between this situation and the previous one are as follows. First, the wave potential Q is shifted to a larger value of x/L . The position of the minimum value of Q is at $x/L \approx -0.12$. Also, note that Q also has an imaginary component. Second, the wave function ϕ is the lowest-order mode and has considerably more structure in x/L than the no shear case. Finally, the spatial extent of ϕ is somewhat narrower with $\Delta x \approx 0.1L$.

A longer-wavelength mode is now considered. We choose $k_y L = 0.1$ so that $k_y L \ll 1$, but still consider $\bar{\nu} = 100.0$, as in the short-wavelength case. Figure 11 is a plot of the density profile $n(x)/n_0$ and electric field profile $E_r(x)/E_0$ for the same parameters as in Figure 8. However, the range of x/L is expanded for comparison to the broadened mode structure. Figure 12 is a plot of Q (Figure 12a) and ϕ (Figure 12b) for the case of no electric field inhomogeneity ($\theta = 0^\circ$ or $E_r(x) = 0$). The eigenfrequency is $\bar{\omega}_r = 0.0$ and $\bar{\gamma} = 0.0930$. The character of Q is considerably different from the short-wavelength case (Figure 9a). The position of the minimum of the potential well is shifted to $x/L \approx 0.0$. Moreover, a 'potential antiwell' exists for $-5.0 \leq x/L \leq 0.0$, which tends to inhibit mode penetration in this region. The corresponding eigenfunction ϕ (Figure 12b) is also substantially different from the short-wavelength case (Figure 9b). First, the wave function has a surface wave character in that $\phi \propto \phi_0 \exp(-\kappa x)$. Second, the wave function is asymmetrical about the position of minimum Q , $x/L \approx 0.0$. The wave function falls off very rapidly in the region $-4.0 < x/L < 0.0$, which is due

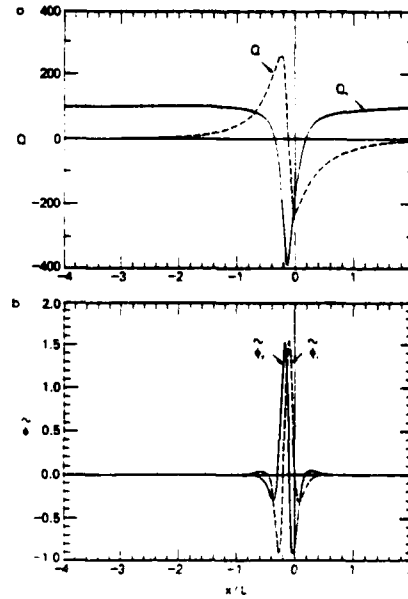


Fig. 10. Wave potential Q and wave eigenfunction ϕ as a function of x/L . The subscripts r and i denote real and imaginary, respectively. The parameters used are $k_y L = 10.0$, $\bar{\nu} = 100.0$, $(V_{\infty}/L) = 100.0$, and $\theta = 70^\circ$ where E_r is given by (40). The eigenfrequency is $\bar{\omega}_r = 0.5307$ and $\bar{\gamma} = 0.0716$. (a) Q versus x/L . (b) ϕ versus x/L .

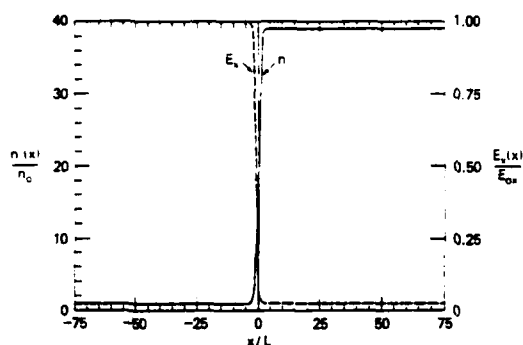


Fig. 11. Equilibrium density (equation (34)) and electric field (equation (40)) profiles for $\epsilon = 0.95$.

to the 'potential antiwell' of Q in this region. For $x/L < -4.0$, ϕ falls off more gradually, similar to its behavior for $x/L > 10.0$. Finally, the wave function is very broad, extending out to $x/L = 50.0$.

Figure 13 is a plot of Q (Figure 13a) and ϕ (Figure 13b) versus x/L for the same parameters as Figure 10, but now we take $\theta = 70^\circ$ so that the electric field is inhomogeneous (see Figure 11). The eigenfrequency is $\omega_r = 0.0057$ and $\gamma = 0.0314$. Although both the wave potential Q and the wave eigenfunction ϕ now have imaginary components, Q and ϕ are quite similar to the no shear case. The wave function is centered about $x/L = 0.0$, has an asymmetrical nature, and extends up to $x/L = 50.0$. Thus the influence of the electric field inhomogeneity on the wave structure in the long-wavelength regime ($k_y L \ll 1$) is much less pronounced than that in the short-wavelength regime ($k_y L \gg 1$). However, the electric field inhomogeneity does reduce the growth rate of the mode significantly.

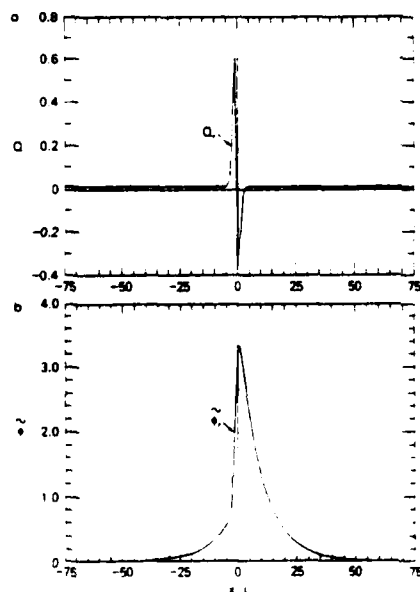


Fig. 12. Wave potential Q and wave eigenfunction ϕ as a function of x/L . The subscripts r and i denote real and imaginary, respectively. The parameters considered are $x/L = 0.1$, $\epsilon = 1.0$, $V_0/L = 100.0$, and $H = 0$ (i.e., $E_1 = 0$). The eigenfrequency is $\omega_r = 0.0$ and $\gamma = 0.0930$. (a) Q versus x/L . (b) ϕ versus x/L .

Finally, we present Figure 14, which is a marginal stability curve (i.e., $\gamma = 0$) of θ versus $k_y L$, where we have taken $\bar{\nu} = 1.0$ (dashed curve) and 100.0 (solid curve). Modes are stable ($\gamma < 0$) and unstable ($\gamma > 0$) above and below each of the curves, respectively. The ratio $E_1(x_0)/E_0$ is for the case $\bar{\nu} = 100.0$ and has the following meaning. It is the ratio of E_1 to E_0 evaluated at $x = x_0$, where x_0 is the position of mode localization. The position of mode localization is defined as the position of the minimum value of Q , which corresponds to the maximum value of ϕ . For the parameters chosen, it is found that $x_0 = 0$. The marginal stability criterion is then given by

$$E_1(x_0)/E_0 \geq 0.05 \tan \theta_{ms} \quad (42)$$

where θ_{ms} is the value of θ at marginal stability and we have used (34) and (40) with $\epsilon = 0.95$. From Figure 14 we find that

$$\theta_{ms} = 88 - 1.4 k_y L \quad (43)$$

where θ_{ms} is measured in degrees. Substituting (43) into (41), we obtain

$$E_1(x_0)/E_0 \geq 0.05 / \tan(1.4 k_y L) \quad (44)$$

or

$$E_1(x_0)/E_0 \geq 2.05 / k_y L \quad (45)$$

for $k_y L \ll 36.0$. Note that (45) is qualitatively consistent with the result of Perkins and Doles [1975] in that there is an inverse relationship between $E_1(x_0)/E_0$ and $k_y L$. Also, (45) is also quantitatively consistent (see (26)) since $D \geq L$ for the profiles used. Finally, as θ approaches 90° , i.e., $E_1 \rightarrow 0$, the wave number of the last unstable mode approaches zero. There is no instability at $\theta = 90^\circ$; this has been demonstrated analytically by Perkins et al. [1973].

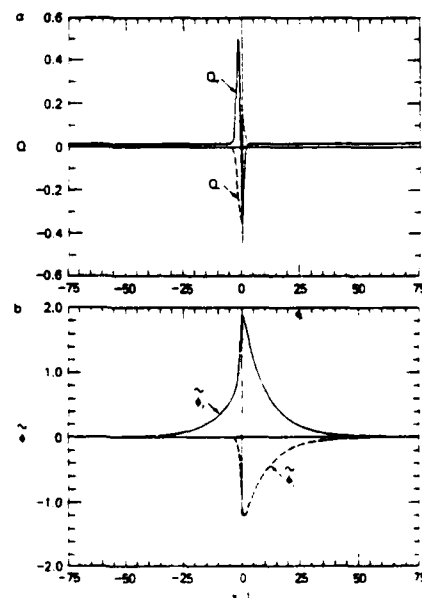


Fig. 13. Wave potential Q and wave eigenfunction ϕ as a function of x/L . The subscripts r and i denote real and imaginary, respectively. The parameters considered are $x/L = 0.1$, $\bar{\nu} = 1.0$, $V_0/L = 100.0$, and $H = 70^\circ$, where E_1 is given by (40). The eigenfrequency is $\omega_r = 0.0057$ and $\gamma = 0.0314$. (a) Q versus x/L . (b) ϕ versus x/L .

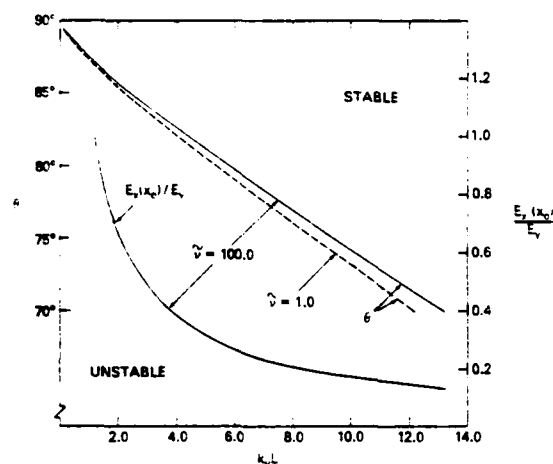


Fig. 14. Marginal stability curve of θ versus $k_{\perp}L$ for $\tilde{\nu} = 1.0$ (dashed curve) and $\tilde{\nu} = 100.0$ (solid curve). The mode is stable ($\gamma < 0$) and unstable ($\gamma > 0$) above and below each of these curves, respectively. The ratio $E_{\perp}(x_0)/E_0$ versus $k_{\perp}L$ is for the case $\tilde{\nu} = 100.0$, where $x_0/L = 0.0$ is the position of mode localization. In all of these curves, (40) has been used for $E_{\perp}(x)$.

4. DISCUSSION

We have presented a general theory of the $E \times B$ instability allowing for an arbitrary (1) density profile, (2) inhomogeneous electric field parallel to the density gradient, and (3) ratio of the collision frequency to the eigenfrequency (i.e., ν_m/ω). A differential equation is derived which describes the structure of the mode in the direction of the inhomogeneity, which we have considered to be the x direction. The theory is restricted to wave numbers such that $k_{\perp}L \ll \Omega/\nu_m$ and $k_{\perp}L \ll \Omega/\omega$; since it has also been assumed that $\nu_m/\Omega \ll 1$ and $\omega/\Omega \ll 1$, this restriction is not important. This work is basically an extension of the analysis of Perkins and Doles [1975], whose theory is restricted to the regime $\nu_m/\omega \gg 1$ and $k_{\perp}L \gg 1$, and considers a specific density and electric field profile. The principal results of this study are as follows.

1. For a constant electric field profile, instability persists even when $E_{\parallel} = 0$ ($\theta = 90^\circ$). In fact, instability also occurs for $E_{\parallel} < 0$ ($\theta > 90^\circ$) when $\partial n/\partial x > 0$; this is contrary to the simple one-dimensional result (i.e., $k = k_z \hat{e}_z$), which requires $E_{\parallel} \partial n/\partial x > 0$ for instability. Thus two-dimensional mode structure (i.e., $k = k_z \hat{e}_z + k_{\perp} \hat{e}_{\perp}$) is crucial to the instability (equation (20)) [Lanson and Workman, 1970]. It is emphasized that $E_{\parallel} = \text{const}$ does not permit a self-consistent equilibrium in that $\nabla \cdot \mathbf{J} = \nabla \cdot [n(\mathbf{V}_i - \mathbf{V}_e)] \neq 0$. For such a situation to occur physically, a source or sink of charge is needed.

2. For an inhomogeneous electric field, an inhomogeneous $E \times B$ velocity occurs ($V_{\perp}(x) = -cE_{\parallel}(x)/B$) which has a stabilizing influence on the mode. Moreover, the short-wavelength modes ($k_{\perp}L \gg 1$) are preferentially stabilized over long-wavelength modes ($k_{\perp}L \leq 1$). This result is consistent with the work of Perkins and Doles [1975].

The functional form of $E_{\parallel}(x)$ is not critical to stabilization of the mode. In the absence of any plasma sources or sinks, the ion continuity equation gives the equilibrium relationship between the density $n(x)$ and the electric field $E_{\parallel}(x)$, as given by (7) (also, see Figure 8). Perkins and Doles [1975] use this relationship in their analysis. However, we have

considered other electric field profiles (e.g., $E_{\parallel}(x) \propto n(x)$ and $E_{\parallel}(x) \propto \tanh(x)/n(x)$). We have found that the instability is still stabilized by the velocity inhomogeneity, again preferentially stabilizing the shorter-wavelength modes but that the marginal stability curves are different from Figure 13.

The mode is stabilized because of the velocity shear contained in the x -dependent resonance $[\omega - k_{\perp}V_{\perp}(x)]^{-1}$ in (17). Terms explicitly proportional to $\partial V_{\perp}/\partial x$ and $\partial^2 V_{\perp}/\partial x^2$ are not important for stabilization.

Perkins et al. [1973] have shown analytically that the mode is stable for $E_{\parallel} = 0$ and $E_{\perp}$ given by (23). The numerical results presented here are consistent with this conclusion. However, we add that as $E_{\parallel} \rightarrow 0$ then $k_{\perp}L \rightarrow 0$. This is clear from Figure 14 by noting that $k_{\perp}L \rightarrow 0$ as $\theta \rightarrow 90^\circ$.

In general, it is found that as ν_m decreases, the growth rate of the mode decreases; this is expected from linear theory (equations (21) and (22)). However, in the case of an inhomogeneous electric field, the difference in growth rates between the strong and weak collisional limits considered is not significant (see Figure 4). Furthermore, the stabilization criterion is not sensitive to ν_m (see Figure 14).

These results are applicable to the development of the $E \times B$ instability in both barium releases and the high-latitude F region ionosphere. First, the important aspects of an inhomogeneous electric field on barium cloud striations has been adequately addressed by Perkins and Doles [1975]. In particular, they note that (1) the back side of a plasma cloud must steepen sufficiently so that it is almost one dimensional to allow the mode to grow (i.e., $\mathbf{E} = E_{\parallel} \hat{e}_{\parallel}$) and (2) the stabilization of the mode due to $E_{\parallel}(x)$ may explain why the sides of a plasma cloud do not become unstable. Furthermore, from our studies, we might hypothesize that the 'freezing' phenomenon in plasma cloud striations [see McDonald et al., 1981] could be due to shear stabilization effects, since shear stabilization acts preferentially on short-wavelength modes, i.e., $k_{\perp}L \gg 1$, but not on long-wavelength modes, i.e., $k_{\perp}L \leq 1$. Second, the role of an inhomogeneous electric field in the $E \times B$ instability can be very important in the structuring of plasma 'blobs' observed in the high-latitude F region. Experimental observations [Vickrey et al., 1980; Tsunoda and Vickrey, 1982] indicate structuring in both the east-west and north-south directions. Moreover, small-scale structuring of the walls of the 'blobs' have also been observed and is attributed to the $E \times B$ instability and/or the current convective instability. The plasma configuration is not well known, but the morphology of the 'blobs' appears to be very complex. Not only are there inhomogeneities anticipated in the electric field, but there are also neutral wind effects, field-aligned plasma currents, and possible coupling effects between the E and F regions. A complete theoretical treatment incorporating these effects is beyond the scope of this paper. However, the results of this analysis strongly suggest that in order for the $E \times B$ drift instability to be a viable candidate for structuring in the high-latitude F region, then the ambient electric field must be orthogonal or nearly orthogonal to the density gradient. A more complete discussion of this problem will be deferred to a later report in which the influences of a field-aligned current and neutral wind are incorporated into this analysis.

Acknowledgments. This work has been supported by the Defense Nuclear Agency and the Office of Naval Research.

The editor thanks J. H. Doles and S. P. Gary for their assistance in evaluating this paper.

REFERENCES

- Chaturvedi, P. K., and S. L. Ossakow, Nonlinear stabilization of the $E \times B$ gradient drift instability in ionospheric plasma clouds. *J. Geophys. Res.*, **84**, 419, 1979.
- Gladd, N. T., and W. Horton, Jr., Critical shear and growth rates for drift waves in a non-uniform current-carrying plasma. *Phys. Fluids*, **16**, 879, 1973.
- Hoh, F. C., Instability of a Penning discharge. *Phys. Fluids*, **6**, 1184, 1963.
- Keskinen, M. J., and S. L. Ossakow, Nonlinear evolution of plasma enhancements in the auroral ionosphere. 1. Long-wavelength irregularities. *J. Geophys. Res.*, **87**, 144, 1982.
- Linson, L. M., and J. B. Workman, Formation of striations in ionospheric plasma clouds. *J. Geophys. Res.*, **75**, 3211, 1970.
- McDonald, B. E., S. L. Ossakow, S. T. Zalesak, and N. J. Zabusky, Scale sizes and lifetimes of F region striations as determined by the condition of marginal stability. *J. Geophys. Res.*, **86**, 5775, 1981.
- Ossakow, S. L., Ionospheric irregularities. *Rev. Geophys. Space Phys.*, **17**, 521, 1979.
- Ossakow, S. L., P. K. Chaturvedi, and J. B. Workman, High-altitude limit of the gradient drift instability. *J. Geophys. Res.*, **83**, 2314, 1978.
- Ossakow, S. L., M. J. Keskinen, and S. T. Zalesak, Ionospheric irregularity physics modelling, paper presented at the AIAA 20th Aerospace Sciences Meeting, Am. Inst. of Aeronaut. and Astronaut., Orlando, Fla., 1982.
- Perkins, F. W., and J. H. Doles III, Velocity shear and the $E \times B$ instability. *J. Geophys. Res.*, **80**, 211, 1975.
- Perkins, F. W., N. J. Zabusky, and J. H. Doles III, Deformation and striation of plasma clouds in the ionosphere. 1. *J. Geophys. Res.*, **78**, 697, 1973.
- Scannapieco, A. J., S. L. Ossakow, S. R. Goldman, and J. M. Pierre, Plasma cloud late time striation spectra. *J. Geophys. Res.*, **81**, 6037, 1976.
- Shiau, J. H., and A. Simon, Onset of striations in barium clouds. *Phys. Rev. Lett.*, **29**, 1664, 1972.
- Simon, A., Instability of a partially ionized plasma in crossed electric and magnetic fields. *Phys. Fluids*, **6**, 382, 1963.
- Simon, A., Growth and stability of artificial ion clouds in the ionosphere. *J. Geophys. Res.*, **75**, 6287, 1970.
- Tsunoda, R. T., and J. F. Vickrey, Evidence of east-west structure in large-scale F region plasma enhancements in the auroral zone. submitted to *J. Geophys. Res.*, 1982.
- Vickrey, J. F., C. L. Rino, and T. A. Potemra, Chatanika/Triad observations of unstable ionization enhancements in the auroral F region. *Geophys. Res. Lett.*, **7**, 789, 1980.
- Völk, H. J., and G. Haerendel, Striations in ionospheric ion clouds. 1. *J. Geophys. Res.*, **76**, 4541, 1971.
- Zabusky, N. J., J. H. Doles III, and F. W. Perkins, Deformation and striation of plasma clouds in the ionosphere. 2. Numerical simulation of a nonlinear two-dimensional model. *J. Geophys. Res.*, **78**, 711, 1973.

(Received June 28, 1982;
revised September 15, 1982;
accepted September 21, 1982.)

APPENDIX J

THE MORPHOLOGY OF A MULTI-BUBBLE
SYSTEM IN THE IONOSPHERE

The Morphology of a Multi-Bubble
System in the Ionosphere

J. Chen and P. Saryanarayana
Science Applications, Inc.
McLean, VA 22101

and

S.L. Ossakow
Plasma Physics Division
Geophysical and Plasma Dynamics Branch
Naval Research Laboratory
Washington, D.C. 20375

29 October 1982

NRL Memo Report

Abstract

A multi-bubble model is developed to study the morphology of a finite array of plasma density depletions (bubbles) in the context of equatorial F-region irregularities during spread F. The Pedersen current conservation equation with quasi-neutrality is solved analytically using an electrostatic analogy. The solution is exact with no a priori assumption regarding the separation distance. A two-bubble system with a piecewise constant density profile is first analyzed and the technique is then applied to multi-bubble systems to calculate the polarization electric field and the rise velocities. It is shown that the influence of the neighboring bubbles is relatively short-ranged and that a small number of bubbles can adequately model the essential physics in a large array of bubbles. For moderately short separation distances, it is found that the $\underline{E} \times \underline{B}$ rise velocity is substantially reduced in comparison with the single-bubble case and that the rise velocity is strongly sheared resulting in the deformation of the contours. The implications of the new morphological results on the stability and dynamical behavior of the bubbles are discussed.

I. INTRODUCTION

The behavior of ionospheric plasma has been extensively studied with the aim of understanding various ionospheric phenomena including equatorial spread F (ESF) and plasma cloud striations (see the following reviews: Ossakow, 1979; Fejer and Kelley, 1980; Ossakow, 1981; Kelley and McClure, 1981; Ossakow et al., 1982). In particular, ESF is thought to be initiated by the Rayleigh-Taylor instability, first proposed by Dungey (1956). In the context of this idea, the plasma density depletions resulting from the instability acquire upward $\underline{E} \times \underline{B}$ drift velocities. A large body of literature has since been developed to describe the linear and nonlinear properties of plasma density depletions ("bubbles") under various assumptions [Haerendel, 1975; Balsley et al., 1972; Chaturvedi and Kaw, 1975a,b; Hudson and Kennel, 1975; Hudson, 1978]. In particular, considerable attention has been given to the morphology and motion [Ossakow and Chaturvedi, 1978; Ott, 1978; Anderson and Haerendel, 1979; Ossakow et al., 1979; Zalesak and Ossakow, 1980; Zalesak et al., 1982] of plasma depletions in the equatorial ionosphere. Moreover, a number of observations [Woodman and LaHoz, 1976; Kelley et al., 1976; Hanson and Sanatani, 1971; McClure et al., 1977; Szuszczewicz et al., 1980, 1981] have indicated the presence of rising plasma bubbles.

Another phenomenon of interest is that of the striations in (artificial) plasma "clouds" (density enhancements). This effect has been attributed to the $\underline{E} \times \underline{B}$ gradient drift instability [Linson and Workman, 1970; Volk and Haerendel, 1971] and appears to be amenable to treatments similar to ESF [Scannapieco and Ossakow, 1976; Scannapieco et al., 1976; Ossakow and Chaturvedi, 1978]. The Rayleigh-Taylor instability and the $\underline{E} \times \underline{B}$ gradient drift instability are both interchange modes that may occur in the leading edge of plasma bubbles and backside of clouds respectively [Ossakow and Chaturvedi, 1978]. In both cases, the resulting density depletions and enhancements are thought to drift by the $\underline{E} \times \underline{B}$ drift. The electric field is produced by the polarization of the plasma across the earth's magnetic field. The essential ingredient is the small but nonzero ion-neutral collision frequency.

As a result of the initial instability such as the Rayleigh-Taylor instability, an array of density depletions are formed with the wave vector perpendicular to the earth's magnetic field. Moreover, the leading edge of a bubble itself is Rayleigh-Taylor unstable resulting in further "bifurcation" [Ossakow and Chaturvedi, 1978]. Similarly, the $\underline{E} \times \underline{B}$ instability causes the backside of an initial cloud to striate, forming "finger-like" structures. Thus, the plasma depletions and enhancements typically occur in multitudes. Indeed, McClure et al., (1977) and Szuszczyewicz et al., (1980;1981) have given evidence for multiple bubbles. In the previous theoretical and numerical simulation works [Ossakow and Chaturvedi, 1978; Anderson and Haerendel, 1979; Zalesak and Ossakow, 1980; Overman et al., 1982], the morphology and evolution have been studied using one-bubble models or a uniformly distributed array of bubbles (clouds). As a result, the intrinsic influence of neighboring bubbles on each other has not been well quantified. Because the bubbles rise due to the polarization induced $\underline{E} \times \underline{B}$ drift, the electric field configuration in and around the bubbles is of central importance.

In the present paper, we study the structure of electrically interacting multi-bubble systems and seek to identify the nature and effects of the mutual interaction. For this purpose, we start with a simple two-bubble configuration in the context of a fluid description and with the emphasis on identifying the basic physics. A simple piecewise constant density profile is used and the electric field inside and outside the bubbles is obtained analytically by the method of images. The method is then applied to multi-bubble configurations.

The scope of the present paper is limited to discussing the morphology of multi-bubble systems and the time-dependent evolution is not explicitly considered. However, the implications of the results will be discussed in the context of the behavior of bubbles consistent with the approximations used in the analysis. Although we primarily discuss the ESF plasma density depletions (bubbles), the technique developed here can be extended straightforwardly to the treatment of plasma density enhancements (clouds).

In Section II, we develop a two-bubble model based on an electrostatic (dielectric) analogy and solve the current conservation equation with

quasi-neutrality for the electric field. Section III describes a multi-bubble model and section IV contains the discussion.

II. A TWO-BUBBLE MODEL

A. Formulation

In the present paper, we consider the electric field configuration of electrically interacting multi-bubble systems imbedded in a uniform background plasma and neutral gas. In order to illustrate the basic physics and the theoretical method, we first develop a simple two-bubble model. For this purpose, we adopt a sharp-boundary density profile in which the plasma density is piecewise constant, being uniform inside (n_1) and outside (n_2) the bubbles and having a discontinuity at the bubble boundaries [Haerendel, 1973; Ossakow and Chaturvedi, 1978; Overman et al., 1982]. The bubbles are modelled by two-dimensional cylinders at the same altitude with circular cross-sections and the axis of the cylinders are aligned with the earth's magnetic field. No neutral wind is included. Figure 1 shows schematically the geometry and the coordinate system. The bubbles are located at $x = -x_0$ and $x = x_0$ so that the inter-bubble separation distance is $2x_0$. For reference purposes, we denote the cylinders by C1 and C2, respectively. The radius of each bubble is a . There is assumed to be a uniform magnetic field B along the $+z$ -direction. In the equatorial and low latitude regions, the x -axis is along the east-west direction and the gravitational force is along the negative y -axis.

The basic equations in our model are the particle conservation, momentum conservation and current conservation equations [see, for example, Ossakow, 1981]. In addition, quasi-neutrality can generally be assumed on the time scales of interest. In the present paper, we study the morphology of two- and multi-bubble systems by solving the current conservation equation to obtain the instantaneous electric field perpendicular to the magnetic field.

The current conservation equation with quasi-neutrality is

$$\underline{\nabla} \cdot \underline{J} = 0 \quad (1)$$

where $\underline{J}$ is the plasma current density due to ion and electron drifts. Neglecting the inertia terms in the momentum equations for cold ion and electron fluids, the current density in the lab frame can be expressed as (see, for example, Ossakow et al., 1979)

$$\underline{J} = en \left[\frac{1}{\Omega_i} \underline{g} \times \hat{\underline{z}} + \frac{\nu_{in}}{\Omega_i} \left(\frac{1}{\Omega_i} \underline{g} + \frac{c}{B} \underline{E} \right) \right], \quad (2)$$

where $\Omega_i = eB/m_i c$ is the ion cyclotron frequency, $\hat{\underline{z}}$ is the unit vector along the earth's magnetic field, $\underline{g}$ is the gravitational acceleration, n is the plasma density, and ν_{in} is the ion-neutral collision frequency. In (2), we have used the fact that $m_e/m_i \ll 1$ and have neglected the electron $\underline{g} \times \hat{\underline{z}}$ contribution. In addition, in arriving at (2), we have made the approximation $\nu_{in}/\Omega_i \ll 1$. In the F region, ν_{in}/Ω_i is typically of the order of 10^{-2} or less.

The second term in the square brackets in (2) gives the force-field ($\underline{g}$ and $\underline{E}$) aligned drift currents due to the finite ion-neutral collisions. It is convenient to separate the electric field $\underline{E}$ according to

$$\underline{E} \equiv \tilde{\underline{E}} - \frac{m_i}{e} \underline{g}.$$

The term $(-m_i/e) \underline{g}$ is the component of $\underline{E}$ cancelling the drift along the gravitational field so that the net drift perpendicular to the magnetic field is described by $\tilde{\underline{E}}$. Then, the current $\underline{J}$ can be written as

$$\underline{J} = \sigma \underline{E}_J \quad (3)$$

where

$$\sigma \equiv \nu_{in} \frac{nec}{B\Omega_i}$$

and

$$\underline{E}_J \equiv \tilde{\underline{E}} + \frac{B}{c\nu_{in}} \underline{g} \times \hat{\underline{z}}. \quad (4)$$

Here, $\underline{B}$ and $\underline{g} = -g\hat{y}$ are assumed to be uniform. Physically, $\underline{E}_J$ can be thought of as the electric field driving the current $\underline{J}$ perpendicular to the magnetic field in the frame moving with the velocity $\underline{V}_d = -\Omega_1^{-1} \underline{g} \times \hat{z}$ relative to the lab frame and σ may be identified as the Pedersen conductivity due to $v_{in} \neq 0$.

Equation (1) can now be written in the equivalent form

$$\underline{\nabla} \cdot (\sigma \underline{\tilde{E}}) = -\frac{B}{cv_{in}} (\underline{g} \times \hat{z}) \cdot \underline{\nabla} n. \quad (5)$$

Perkins et al. (1973) obtained this expression and noted that (5) describes a dielectric immersed in a uniform electric field $\underline{E}_0$ where

$$\underline{E}_0 = \frac{B}{cv_{in}} \underline{g} \times \hat{z}. \quad (6)$$

In an earlier work, Longmire (1970) utilized a similar magnetostatic analogy to treat the motion of isolated ion clouds. In this dielectric analogy, $\underline{\tilde{E}}$ is the polarization (self) electric field of the bubbles in the uniform field $\underline{E}_0$ and $\underline{E}_J$ corresponds to the total electric field ($\underline{E}_0 + \underline{\tilde{E}}$) satisfying the boundary conditions across the bubble boundaries,

$$(\sigma \underline{E}_J)_\perp = \text{continuous} \quad (7)$$

$$(\underline{E}_J)_\parallel = \text{continuous}$$

and at infinity ($x, y \rightarrow \infty$)

$$\underline{E}_J \rightarrow \underline{E}_0. \quad (8)$$

Note that we have implicitly chosen a reference frame in which the electric field of the distant undisturbed ionosphere is $\underline{E}_0$.

The symbols $\parallel$ and $\perp$ refer to the directions parallel and perpendicular to the boundary surfaces, respectively. In the present paper, we also adopt the dielectric analogy and solve the current conservation equation (1) subject to the above boundary conditions (7) and (8). For this purpose, it is illuminating to rewrite equation (1) as

$$\nabla \cdot (\sigma \underline{E}_J) = 0 \quad (9)$$

In the following section, we describe the method of image dipoles used to solve this "dielectric equation". As a matter of notation, in the remainder of the paper, we use $\underline{E}$ without the subscript J to denote the solution of (9).

F. The Method of Image Dipoles

The problem of solving Poisson's equation (9) with multiple disconnected boundaries is generally difficult. However, in the case treated here with circular cross-sections, the dielectric analogy allows us to construct an exact solution. Consider first a single dielectric cylinder of radius a centered at $x = 0$ and a line charge density q located at $x = b$ ($|b| > a$). It is well known (Smythe, 1968) that the induced electric field outside the cylinder is that of a line charge $q' =$

$-q(1-K)/(1+K)$ located at $x = a^2/b$ and a line charge $-q'$ located at $x=0$. The quantity K is the ratio of the dielectric constant (σ_1) inside the cylinder to that outside the cylinder (σ_2)

$$K \equiv \frac{\sigma_1}{\sigma_2} \quad (10)$$

The induced electric field inside the cylinder is that of a single line charge $q'' = 2q/(1+K)$ located at $x = b$. If we replace the line charge q by a line dipole moment $\underline{P}_0 = P_0 \hat{x}$ which is equivalent to two equal and oppositely charged line charges separated by a vanishingly small distance, then we find that the electric field due to polarization of the cylinder is that of a single image dipole $\underline{P}$ given by

$$\underline{P} = - \left(\frac{1-K}{1+K} \right) \frac{a^2}{b^2} \underline{P}_0 \quad (11)$$

located at $x = a^2/b$. Note that no image dipole is present on the axis (to be contracted with the line charge case) and that $\underline{P}_0$ and $\underline{P}$ are colinear, pointing in the opposite directions. The induced electric field inside the cylinder is that due to a dipole moment $\underline{P}^*$ given by

$$\underline{P}^* = \frac{2}{1+K} \underline{P}_0 \quad (12)$$

located at $x = b$. In the remainder of the paper, the dipole moment and the electric field inside the cylinders will be denoted by asterisks.

We now consider two identical dielectric cylinder of radius a centered at $x = -x_0$ and $x = x_0$ (see figure 1) immersed in a uniform electric field $\underline{E}_0$. Suppose, for the moment, that the two cylinders are non-interacting. Then, for the purpose of calculating the polarization electric field outside the cylinders, each cylinder may be replaced by a dipole moment $\underline{P}_0$ located at $x = -x_0$ and $x = x_0$, where

$$\underline{P}_0 = \frac{1}{2} \left(\frac{1-K}{1+K} \right) a^2 \underline{E}_0 \hat{x}, \quad (13)$$

and K is defined by equation (10). The components of the self electric field are

$$E_x = 2P_0 [f(x+x_0, y) + f(x-x_0, y)], \quad (14)$$

and

$$E_y = 2P_0 [h(x+x_0, y) + h(x-x_0, y)], \quad (15)$$

where

$$f(x, y) \equiv \frac{x^2 - y^2}{(x^2 + y^2)^2}, \quad (16)$$

and

$$h(x, y) \equiv \frac{2xy}{(x^2 + y^2)^2}. \quad (17)$$

The electric field $\underline{E}^*$ inside the bubbles is

$$\underline{E}^* = \frac{2}{1+K} \underline{E}_0. \quad (18)$$

In the remainder of the paper, asterisks will be used to denote the electric fields and dipole moments inside the bubbles.

We now allow the two cylinders to interact with each other. In addition to the uniform external field $\underline{E}_0$, each cylinder experiences the

dipole field of the other given by (14) and (15). Iterating the method of image dipoles described above ((8) and (9)), it is straightforward to show that the total field outside the bubbles is given by

$$E_x = -E_o + \sum_{n=0}^{\infty} 2P_n [f(x+x_n, y) + f(x-x_n, y)], \quad (19)$$

and

$$E_y = \sum_{n=0}^{\infty} 2P_n [h(x+x_n, y) + h(x-x_n, y)], \quad (20)$$

where f and h are defined by (16) and (17). Here, for $n \neq 0$,

$$P_n \equiv - \left(\frac{1-K}{1+K} \right) \frac{a^2}{b_n^2} P_{n-1}, \quad (21)$$

$$b_n \equiv x_o + x_{n-1}, \quad (22)$$

and

$$x_n \equiv x_o - \frac{a^2}{x_o + x_{n-1}}. \quad (23)$$

For $n = 0$, we have $x_n = x_o$ and P_o is given by (10).

Similarly, the total electric field inside each cylinder, say, (C2) located at $x = x_o$, is found to be

$$E_x^* = - \frac{2}{1+K} E_o + \frac{2}{1+K} \sum_{n=0}^{\infty} 2P_n f(x+x_n, y), \quad (24)$$

and

$$E_y^* = \frac{2}{1+K} \sum_{n=0}^{\infty} 2P_n h(x+x_n, y), \quad (25)$$

where P_n and x_n are defined above. For the other bubble (C1), the field is obtained by replacing x_o with $-x_o$ in the functions f and h .

It can be seen from (19) and from the field components that the convergence properties of these series depend on the parameter s defined by

$$s \equiv \frac{1-K}{1+K} \frac{a^2}{(2x_o)^2}. \quad (26)$$

Each successive linear dipole moment is reduced by a factor of s and a geometrical factor of order unity from the preceding one. Since $x_0 > a$, we have $|s| < 1/4$ for any K . As a result, the series are generally rapidly convergent except for very small center-to-center separation distances ($x_0 \sim a$).

C. Applications

In order to apply the above dielectric results to the ionospheric bubble and cloud problems in accordance with the dielectric analogy (equations (7) and (8)), we identify the dielectric constant σ with the Pedersen conductivity defined in Section II and recall that $\underline{E}_0$ is given by equation (6). Note that the Pedersen conductivity is proportional to the plasma density so that

$$K = \frac{n_1}{n_2},$$

where n_1 and n_2 are the plasma densities inside and outside the bubbles. With these identifications, equations (19), (20), (24) and (25) describe the electric field of a two-bubble system in the frame moving with $\underline{V}_d = -\Omega_1^{-1} \underline{g} \times \underline{\hat{z}}$ relative to the earth. Note that our formalism guarantees that the field components obtained above exactly satisfy equation (9) and the boundary conditions (7) and (8), as can be verified easily. Thus, the solution is unique. It is also worth noting that all the image line dipole moments are aligned with the x -axis and no higher multipoles such as quadrupole moment arise.

As a result of the polarization electric field, the plasma bubbles (density depletions) $\underline{E} \times \underline{B}$ drift. The drift velocity relative to the distant undisturbed ionosphere where the electric field is $\underline{E}_0$ is given by $\underline{V} = c (\underline{E}^* - \underline{E}_0) \times \underline{B}/B^2$. In particular, the single-bubble rise velocity $\underline{V}_1$ is (see equation 18)

$$\underline{V}_1 = - \left(\frac{1-K}{1+K} \right) \frac{1}{v_{in}} \underline{g}. \quad (27)$$

For bubbles, $K < 1$ and $\underline{V}_1$ is upward. This result has been obtained by Ossakow and Chaturvedi [1978]. Because $\underline{g}$ is assumed to be uniform, a

single isolated bubble maintains its circular cross-section as it rises with the constant velocity V_1 . The presence of a second bubble, however, modifies the rise velocity significantly. In particular, it is no longer uniform and the cross-sections do not remain unchanged. We have numerically carried out the summations indicated in (19), (20), (24), and (25). The results are depicted in figure 2. The solid lines represent the self electric field lines or equivalently the Pedersen current lines while the dashed lines correspond to the instantaneous $\hat{E} \times B$ drift velocity (V_2) for a two-bubble system ($2x_0 = 2.5a$ and $K = 0$ for 100% depletion). This figure shows only one quadrant; the actual system is spatially uniform in the z -direction and is symmetric about the y - z plane and the x - z plane. The solid electric field lines are such that the line density is proportional to the field strength. This figure clearly shows that the electric field inside the bubble is significantly modified from the uniform field of an isolated bubble (see (18)). As a general remark, the interaction vanishes as $K \rightarrow 1$ and the separation distance increases to infinity.

In this example, the electric field strength at the point A ($x = x_0 - a$, $y = 0$) is approximately 1/3 of that at B ($x = x_0 + a$, $y = 0$). The prominently nonuniform electric field inside the bubbles has a number of important implications for the dynamic behavior of the two-bubble system. The drift lines (dashed) in figure 2 show that different regions of a bubble undergo drift in the east-west ($\hat{x}$) direction with respect to the undisturbed ionosphere. This horizontal drift can be a significant fraction of the vertical rise velocity at some points inside the bubble. For example, $|E_y^*/(E_x^* - E_{x0})| \approx 0.55$ at $r = a_-$, and $\theta = \pi/8$ where r is the radial distance from $x = x_0$, a_- is just inside the bubble surface and θ is measured from the point A ($\theta = 0$). Here, $x_0 = 1.25$ and $K = 0$ have been used (figure 2). The reason for the strong divergence in the field lines is that the dipole field due to one bubble opposes the internal field of the other bubble.

Figure 3 shows the two bubble rise velocity (V_y) at a number of points inside the bubble (C2) relative to the uniform rise velocity V_1 of a single bubble given by (27). The ratio $R_2 \equiv V_y/V_1$ is plotted for the interior points A and B (figure 2) as a function of x_0/a . We note that the rise

velocity V_y is strongly affected by the neighboring bubble for center-to-center separation distances ($2x_0$) smaller than $5a$ to $6a$. Moreover, comparing V_y at A and at B, we see that the $\underline{E} \times \underline{B}$ drift velocity is sheared. This, together with the conclusions of the preceding paragraph, implies that an initially circular cross-section would not remain circular so that the two-bubble system described here does not correspond to a steady-state system. This should be contrasted with single-bubble models in which steady-state solutions are possible for piecewise constant density profiles. Because a neighboring bubble generally introduces nonuniformity in the electric field, it seems difficult to construct a steady-state two-bubble system unless x_0/a is large.

In interpreting the results, we note that our results do not include the time-dependence. The above field configuration exists if the two-bubble system as described is created at some time (say, $t = 0$). Thus, it would be appropriate as a consistent initial condition for the purpose of studying the subsequent time-evolution. In practice, the drift lines shown in figure 2 are expected to closely approximate the actual evolution for some period of time after $t = 0$ until the distortion changes the field topology significantly. However, the general feature of the field with the weakest x-component and hence the slowest rise velocity in the region nearest to the neighboring bubble should remain unchanged in time.

III. A MULTI-BUBBLE MODEL

In the preceding section, we have discussed in detail a two-bubble model. In considering a multi-bubble system that may be applicable to the ESF phenomenon, the basic physics and the theoretical treatment remain unchanged. However, as the number of bubbles increases, so does the number of image dipoles. For an N -bubble system, the n^{th} order expression must include $N(N-1)^n$ image dipoles. Fortunately, the influence of a bubble decreases as the inverse square of the separation distance. Thus, the nearest and the second nearest neighbors are expected to have the dominant effects. A consideration of equations (11) and (12) shows that the second nearest neighbors have effects of the order of $4^{-2}E_0 \approx 0.06 E_0$ and that the third nearest neighbors have effects of the order of $6^{-2}E_0 \approx 0.03 E_0$ for a

given value of $a/x_0 < 1$. This means that the inter-bubble (and also inter-cloud) interaction is short-ranged and that the third nearest neighbors and beyond have no significant influence. Thus, only a small number of bubbles are necessary to model an N-bubble system with $N \gg 1$. Note that the influence of the bubbles still vanishes at infinity, differing from systems satisfying periodic boundary conditions.

The theoretical treatment described is exactly applicable to any N. However, in the remainder of this section, we include only up to the second nearest neighbors, a total of 5 bubbles. The mathematical manipulations involved are analogous to those of the preceding section, resulting in series expressions similar to (19), (20), (24), and (25). Since no new insight is to be gained by examining the actual expressions, we give below only the results. In figure 4, we show the electric field and drift configurations of a system with three plasma depletions. The three depletions are again modelled by cylinders of radius a , located at $x = -2x_0$, 0 , and $2x_0$. Only one quadrant is shown. The neighboring bubbles are separated by a distance $2x_0$ as before. The "external" electric field is E_0 given by (6), and the density profile is piecewise constant. The solid lines represent the polarization electric field without E_0 . The quantity $c (\underline{E}^* - \underline{E}_0) \times \underline{B} / B^2$ is then the instantaneous drift velocity relative to the distant undisturbed ionosphere and is represented by the dashed lines. In the example shown in figure 4, we have used $2x_0 = 2.5$ and $K = 0$ (100% depletion). Although the three-bubble system is different from a two-bubble system in that the former has a central bubble about which the system is symmetric, the general features of the field and drift configurations are similar as can be seen by comparing figures 2 and 4. Indeed, a five-bubble and larger systems exhibit similar features to those shown in figure 4. No detailed and separate treatment of such large systems will be given here since, as we will show later, the second nearest neighbors and beyond exert negligible influences. That is, a three-bubble system such as the one shown in figure 4 describes well the basic morphology of N-bubble systems even for $N \gg 1$. As before, only dipole moments, not higher multipoles, are induced.

Figure 4 shows the distortion in the electric field which renders the system non-steady-state, as in the two-bubble case. The drift velocity

(dashed lines) has an east-west (horizontal) component that may be a significant fraction of the vertical velocity. For example, $|E_y^*/(E_x^* - E_0)| \approx 0.43$ at $\theta = \pi/8$ from the point B just inside the boundary of the central bubble and $|E_y^*/(E_x^* - E_0)| \approx 0.55$ at $\theta = \pi/8$ from the point $x = x_0 - a$ and $y = 0$ just inside the boundary of the side bubble. Thus, in the neighborhood of these points, the bubble elements should have significant horizontal drifts.

In addition to the distortion of bubble contours resulting from the non-uniform electric field, the figure also shows that the electric field is substantially reduced from that of a an isolated single bubble. This fact is illustrated in figure 5 which gives $R_3 \equiv |V_y/V_1|$ for the points A, B and C corresponding to $x = 0$, $x = a$ and $x = 2x_0 + a$, all just inside the bubble surfaces. Here, V_y is the vertical drift velocity of the three-bubble system relative to the distant ionosphere. The line D gives the relative vertical rise velocity of a five-bubble system calculated at $x = a_-$. The curve D is to be compared with the curve B. This figure shows that the reduction in rise velocities is increased as the number of bubbles is increased. However, consistent with the previous expectation, the influence of the bubbles beyond second nearest neighbors is small, as confirmed by figure 6. This figure shows the relative vertical velocity $R = V_y/V_1$ as a function of N , the number of bubbles, for several values of x_0 .

Figures 3 and 5 show that the vertical drift velocity of a bubble is a sensitive function of the separation distance ($2x_0$) except for relatively large values ($2x_0/a \geq 6$). In the context of the Rayleigh-Taylor instability this implies that the electric field configuration and the drift velocities may depend sensitively on wavelengths ($2x_0$). In particular, the bubbles with smaller wavelength-to-radius ratio would rise more slowly and be distorted more strongly. This conclusion may be particularly applicable to the initial linear or early-time non-linear stages of the evolution.

Finally, in figure 7, we have plotted the vertical drift velocity versus $K \equiv n_1/n_2$ for a two-bubble ($K < 1$) and two-cloud ($K > 1$) system. The velocity is calculated at $x = \pm (x_0 - a)$ (point A in figure 2) and is upward for bubbles and downward for clouds. The velocity is normalized to

V_1 (equation (27)) for each value of K . Note that the point with $K = 1$ does not exist for each line. For $K = 1$, the ionosphere is not disturbed and there is no bubble or cloud drifting vertically. This is born out by the fact that the drift velocity vanishes for a single-bubble (cloud) and any multi-bubble (cloud) system. Clearly, the velocities vanish differently for different separation distances. Mathematically, the ratio of the vertical drift velocity $V_y = c(E_x^* - E_0)/B_0$ to the single-bubble (cloud) drift velocity V_1 has the limit

$$\frac{V_y}{V_1} = - \left[1 - \frac{2}{1+K} a^2 f(x + x_0, y) \right]$$

as $K \rightarrow 1$. The circled points in figure 6 correspond to the absolute value of the quantity. However, this quantity has no physical meaning.

IV. SUMMARY AND DISCUSSION

In the preceding sections, we have solved the current conservation equation (1) with quasi-neutrality for a two-bubble and three-bubble systems using a dielectric analogy. These two configurations include the dominant near-neighbor interaction and can model the essential morphology of the multi-bubble systems described. This is demonstrated by actually calculating the field including up to five bubbles. The solutions are exact, satisfying the specified boundary conditions on all the multiple disconnected boundary surfaces and at infinity.

Equations (19), (20), (24) and (25) give the solution of the two-bubble system as a superposition of image line dipole moments. Similar expressions are obtained for three- and N -bubble (cloud) systems ($N > 3$). In all cases, the interaction is dominated by the nearest neighbors and is sensitive to the separation distance ($2x_0$) between bubbles. An important result is that the electric field inside the bubbles is generally highly non-uniform so that the multi-bubble and multi-cloud systems are not steady-state configurations even with piecewise constant density profiles (figures 2 and 4). Moreover, the electric field inside the bubbles is significantly weaker than that in a one-bubble system so that the vertical drift velocity relative to the undisturbed ionosphere is slower than the one-bubble case (equation (27) and figures 3 and 5). For moderately small

center-to-center separation distances, $2x_0 \lesssim 6a$, the reduction in the rise velocity is substantial. This implies that the electric field configuration and the $\underline{E} \times \underline{B}$ drift velocity may depend sensitively on the wavelengths of instabilities causing the initial density fluctuations (e.g., the Rayleigh-Taylor instability). It has also been shown that in some regions inside bubbles, the horizontal drift velocity may be comparable to the vertical drift velocity with $|E_y^*/(E_x^* - E_0)|$ as large as 1/2 (Sections II and III).

A corollary that follows from the non-uniform electric field in the bubbles is that the polarization induced $\underline{E} \times \underline{B}$ drift velocity is sheared and that a component of the electric field parallel to the local density gradient is developed in the leading edges of the bubbles. For the simple piecewise constant density profiles used for our analysis, the density gradient is not well defined. In a more realistic density profile, however, some previous model calculations indicate that the electric field component parallel to the density gradient has stabilizing influences on the Rayleigh-Taylor instability [Guzdar et al., 1982] and the $\underline{E} \times \underline{B}$ drift instability [Perkins and Doles, 1975; Huba et al., 1982]. More specifically, these calculations show that velocity shear preferentially stabilizes the short wavelength modes. In light of these results, we suggest that the bifurcation behavior of bubbles and clouds may be inhibited by the presence of nearby bubbles (clouds).

The emphasis of our analysis has primarily been on plasma density depletions (bubbles). However, plasma density enhancements (clouds) can also be treated in a similar fashion [Scannapieco and Ossakow, 1976; Scannapieco et al., 1976; Ossakow and Chaturvedi, 1978]. If a cloud exists in the equatorial or low latitude regions, the preceding results are all applicable with the replacement of $K > 1$ ($K < 1$ for bubbles) where $K = n_1/n_2$. In particular, as the backside of an initial cloud begins to bifurcate, the small x_0/a and small N results may be applicable. One point to note is that the electric field inside a two- and three-cloud system with a piecewise constant density profile is similar to that shown in figure 2 and 4 with weaker field strength in the regions facing the neighboring clouds. However, the boundary condition $K(E_{in})_{\perp} = (E_{out})_{\perp}$ implies that $(E_{in})_{\perp}$ is smaller than $(E_{out})_{\perp}$ by a factor of K^{-1} for

clouds. As a result, the distortion in the electric field lines inside a cloud is less pronounced than in a bubble.

ACKNOWLEDGEMENTS

We would like to thank Drs. P. Chaturvedi, E. Overman, N. Zabusky, and S. Zalesak for useful discussions. This work has been supported by DNA and ONR.

REFERENCES

- Anderson, D.N., and G. Haerendel, The motion of depleted plasma regions in the equatorial ionosphere, J. Geophys. Res., 84, 1251, 1979.
- Balsley, B.B., G. Haerendel, and R.A. Greenwald, Equatorial spread F, recent observations and a new interpretation, J. Geophys. Res., 77, 5625, 1972.
- Chaturvedi, P.K., and P. Kaw, Steady state finite amplitude Rayleigh-Taylor modes in Spread F, Geophys. Res. Lett., 2, 381, 1975a.
- Chaturvedi, P.K., and P. Kaw, Correction, Geophys. Res. Lett., 2, 499, 1975b.
- Dungey, J.W., Convective diffusion in the equatorial F region, J. Atmos. Terr. Phys., 9, 304, 1956.
- Fejer, B.G. and M.C. Kelley, Ionospheric irregularities, Rev. Geophys. Space Phys., 18, 401, 1980.
- Guzdar, P.N., P. Satyanarayana, J.D. Huba and S.L. Ossakow, Influence of velocity shear on Rayleigh-Taylor instability, Geophys. Res. Lett., 9, 547, 1982.
- Haerendel, G., Theory of equatorial spread F, preprint, Max-Planck Inst. fur Astrophys., Munich, 1973.
- Hanson, W.B., and J. Santani, Large N_f gradients below the equatorial F peak, J. Geophys. Res., 78, 1167, 1973.
- Huba, J.D., S.L. Ossakow, P. Satyanarayana, and P.N. Guzdar, Linear theory of the $E \times B$ instability with an inhomogeneous electric field, J. Geophys. Res. (submitted 1982).
- Hudson, M.K., and C.F. Kennel, Linear theory of equatorial spread-F, J. Geophys. Res., 80, 4581, 1975.
- Hudson, M.K., Spread F bubbles: Nonlinear Rayleigh-Taylor mode in two dimensions, J. Geophys. Res., 83, 3189, 1978.
- Kelley, M.C., G. Haerendel, H. Kappler, A. Valenzuela, B.B. Balsley, D.A. Carter, W.L. Ecklund, C.W. Carlson, B. Hausler, and R. Torbert, Evidence for a Rayleigh-Taylor type instability and upwelling of depleted density regions during equatorial spread F, Geophys. Res. Lett., 3, 448, 1976.
- Kelley, M.C., and J.P. McClure, Equatorial spread F: A review of recent experimental results, J. Atm. Terr. Phys., 43, 427, 1981.

- Linson, L.M., and J.B. Workman, Formation of striations in ionospheric clouds, J. Geophys. Res., 75, 3211, 1970.
- Longmire, C.L., On the motion of artificial ion clouds, Los Alamos Nuclear Corp. Report, LANC-N-11, 1970.
- McClure, J.P., W.B. Hanson, and J.F. Hoffman, Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., 82, 2650, 1977.
- Ossakow, S.L., and P.K. Chaturvedi, Morphological studies of rising equatorial spread F bubbles, J. Geophys. Res., 83, 2085, 1978.
- Ossakow, S.L., Ionospheric irregularities, Rev. Geophys. Space Phys., 17, 521, 1979.
- Ossakow, S.L., S.T. Zalesak, B.E. McDonald, and P.K. Chaturvedi, nonlinear equatorial spread F: Dependence on altitude of the F peak and bottomside background electron density gradient scale length, J. Geophys. Res., 84, 17, 1979.
- Ossakow, S.L., Spread F theories - a review, J. Atmos. Terr. Phys., 43, 437, 1981.
- Ossakow, S.L., M.J. Keskinen, and S.T. Zalesak, Ionospheric irregularity physics modelling, AIAA-82-0147, January 1982.
- Ott, E., Theory of Rayleigh-Taylor bubbles in the equatorial ionosphere, J. Geophys. Res., 83, 2006, 1978.
- Overman, E., N.J. Zabusky, and S.L. Ossakow, Ionospheric plasma cloud dynamics via regularized contour dynamics: I. Stability and nonlinear evolution of one contour models, Univ. of Pittsburgh Report, 1982.
- Perkins, F.W., N.J. Zabusky, and J.H. Doles III, Deformation and striation of plasma clouds in the ionosphere, 1., J. Geophys. Res., 78, 697, 1973.
- Perkins, F.W., and J.H. Doles III, Velocity shear and the $E \times B$ instability, J. Geophys. Res., 80, 211, 1975.
- Scannapieco, A.J., and S.L. Ossakow, Nonlinear equatorial spread F, Geophys. Res. Lett., 3, 451, 1976.
- Scannapieco, A.J., S.L. Ossakow, S.R. Goldman, and J.M. Pierre, Plasma cloud late time striation spectra, J. Geophys. Res., 81, 6037, 1976.
- Smythe, W.R., Static and Dynamic Electricity, Third edition, p. 95, McGraw-Hill Book Company, New York, 1968.

- Szuszczeicz, E.P., R.T. Tsunoda, R. Narcisi, and J.C. Holmes, Coincident radar and rocket observations of equatorial spread F, Geophys. Res. Lett., 7, 537, 1980.
- Szuszczeicz, E.P., R.T. Tsunoda, R. Narcisi, and J.C. Holmes, PLUMEX II: A second set of coincident radar and rocket observations of equatorial spread F, Geophys. Res. Lett., 8, 803, 1981.
- Volk, H.J., and G. Haerendel, Striation in ionospheric clouds, J. Geophys. Res., 76, 454, 1971.
- Woodman, R.F. and C. LaHoz, Radar observations of equatorial irregularities, J. Geophys. Res., 81, 5447, 1976.
- Zalesak, S.T., and S.L. Ossakow, Nonlinear equatorial spread F: Spatially initial perturbations, J. Geophys. Res., 85, 2131, 1980.
- Zalesak, S.T., S.L. Ossakow, and P.K. Chaturvedi, Nonlinear equatorial spread F: The effect of neutral winds and background Pedersen conductivity, J. Geophys. Res., 87, 151, 1982.

FIGURE CAPTIONS

- Fig. 1 A schematic drawing of two plasma density depletions and the coordinate system. The depletions have circular cross-sections and are infinite in extent along the z-direction.
- Fig. 2 A drawing of the electric field lines (solid lines) and the $\underline{E} \times \underline{B}$ drift lines (dashed lines) showing one quadrant of a two-bubble system. The system is symmetric about the x-z plane and y-z plane. The separation distance is $2x_0 = 2.5a$ and $K = 0$ (100% depletion). The points A and B are inside the bubbles at $x = x_0 - a$ and $x = x_0 + a$, respectively.
- Fig. 3 The ratio R_2 of the vertical rise velocity V_y of a two-bubble system to the single-bubble rise velocity V_1 (equation (27)) as a function of x_0/a . The separation distance is $2x_0 = 2.5a$ and $K = 0$ (100% depletion). The curve A corresponds to the point A ($x = x_0 - a$) and the curve B to the point B ($x = x_0 + a$) in figure 2.
- Fig. 4 A drawing of the electric field lines (solid) and the $\underline{E} \times \underline{B}$ drift lines (dashed), showing one quadrant of a three-bubble system. The bubbles are placed at $x = \pm 2x_0$ and $x = 0$. The points A, B, and C are inside the bubbles at $x = 0$, $x = a$, and $x = 2x_0 + a$. The separation distance is $2x_0 = 2.5a$ and $K = 0$ (100% depletion).
- Fig. 5 The ratio R_3 of the vertical velocity V_y of a three-bubble system to the single-bubble rise velocity V_1 , plotted versus x_0/a . The curves A, B, and C correspond to the points A, B, and C in figure 4. The curve D (dashed) corresponds to $R_5 = V_y/V_1$ at $x = a$ in a five-bubble system, to be compared with the curve B.
- Fig. 6 $R = V_y/V_1$ versus N , the number of bubbles, evaluated inside at $x = x_0 - a$ for the two-bubble case and at $x = a$ for all others. $K = 0$. The separation distances ($2x_0$) are (a) $2x_0 = 2.4$, (b) $2x_0 = 3.2$, (c) $2x_0 = 4.0$, and (d) $2x_0 = 10$.
- Fig. 7 $R_2 = V_y/V_1$ versus K for two-bubble ($K < 1$) and two-cloud ($K > 1$) systems. The separation distances ($2x_0$) are (a) $2x_0 = 2.4$, (b) $2x_0 = 4$, and (c) $2x_0 = 10$.

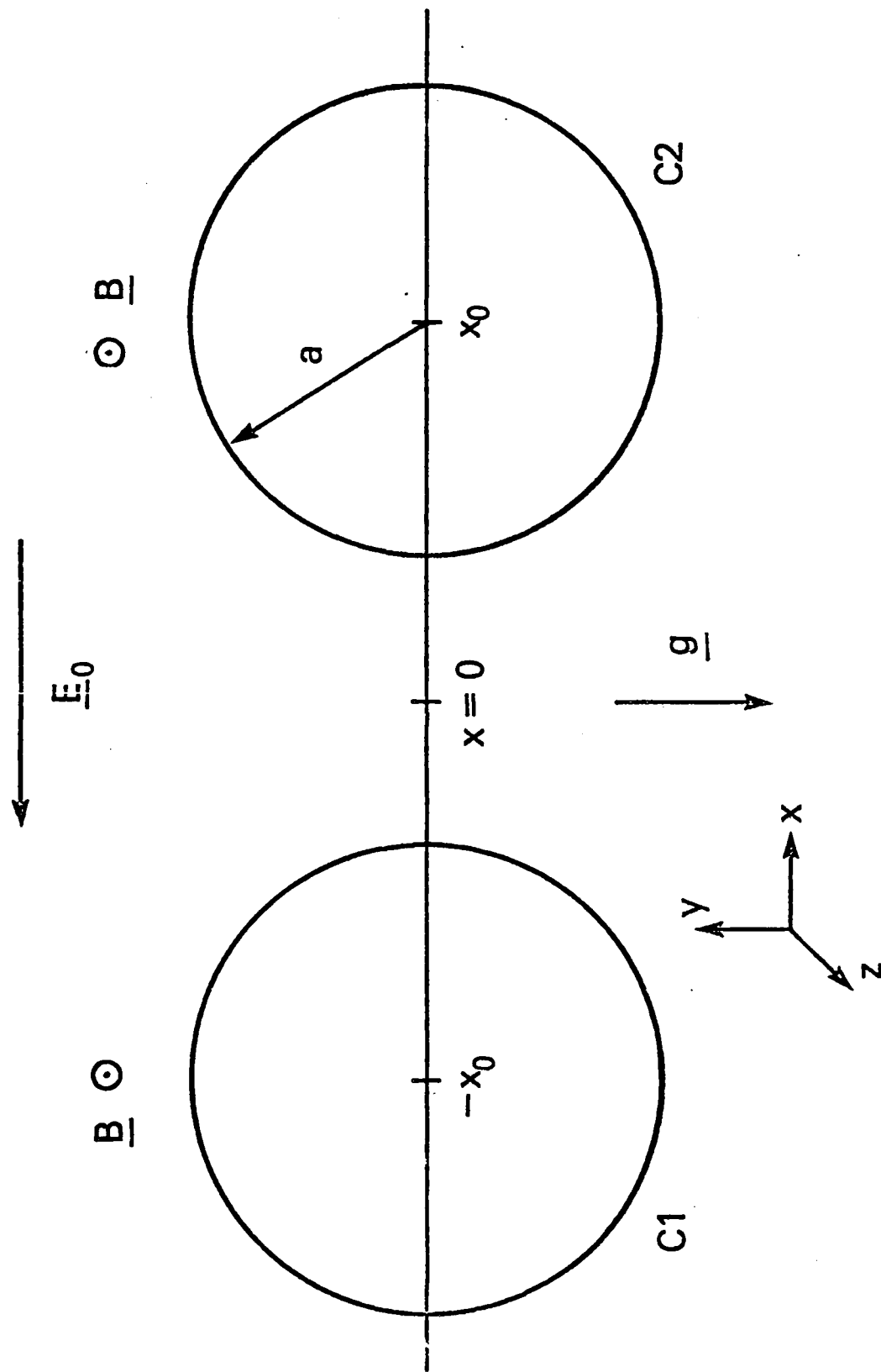
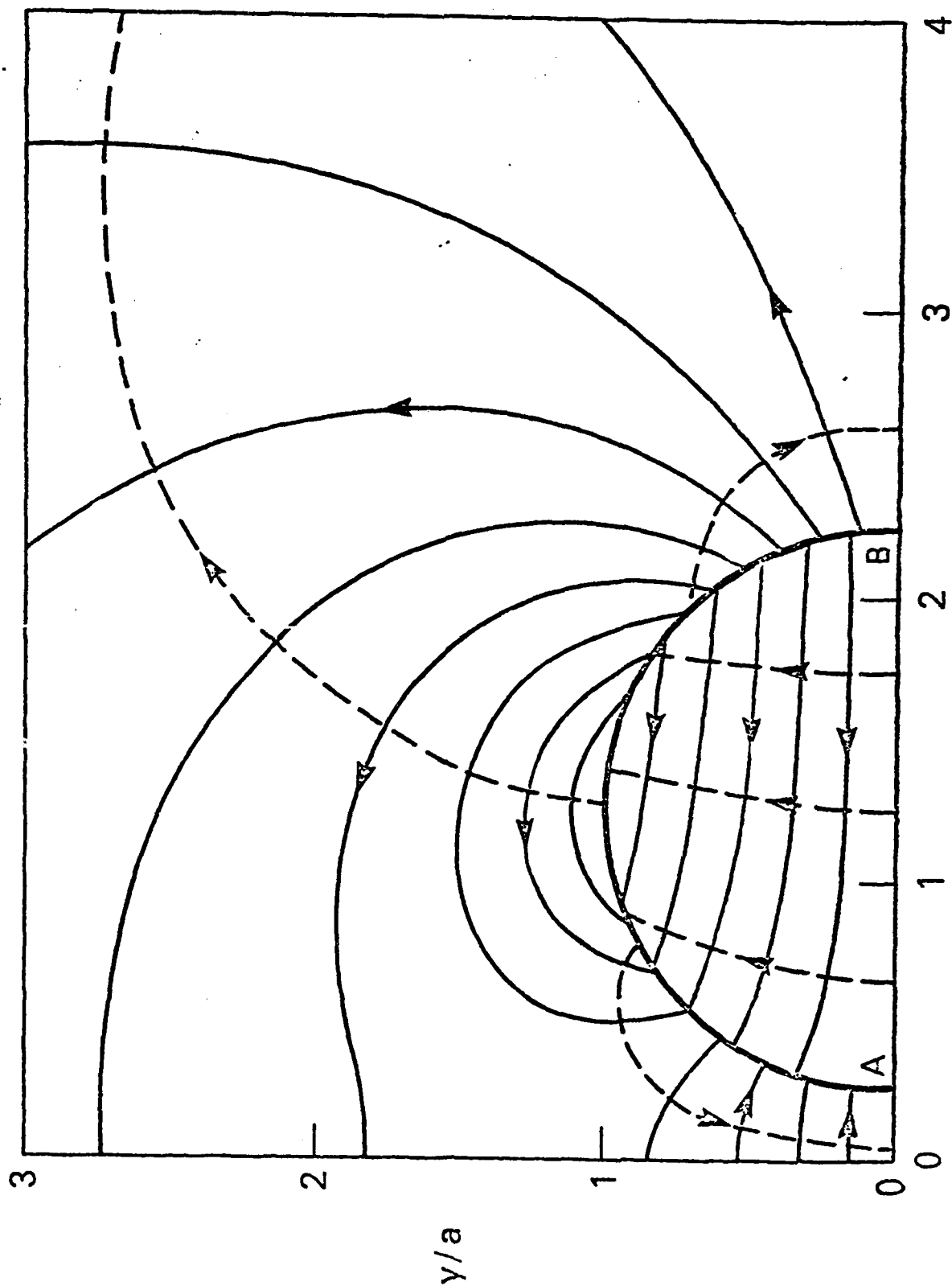
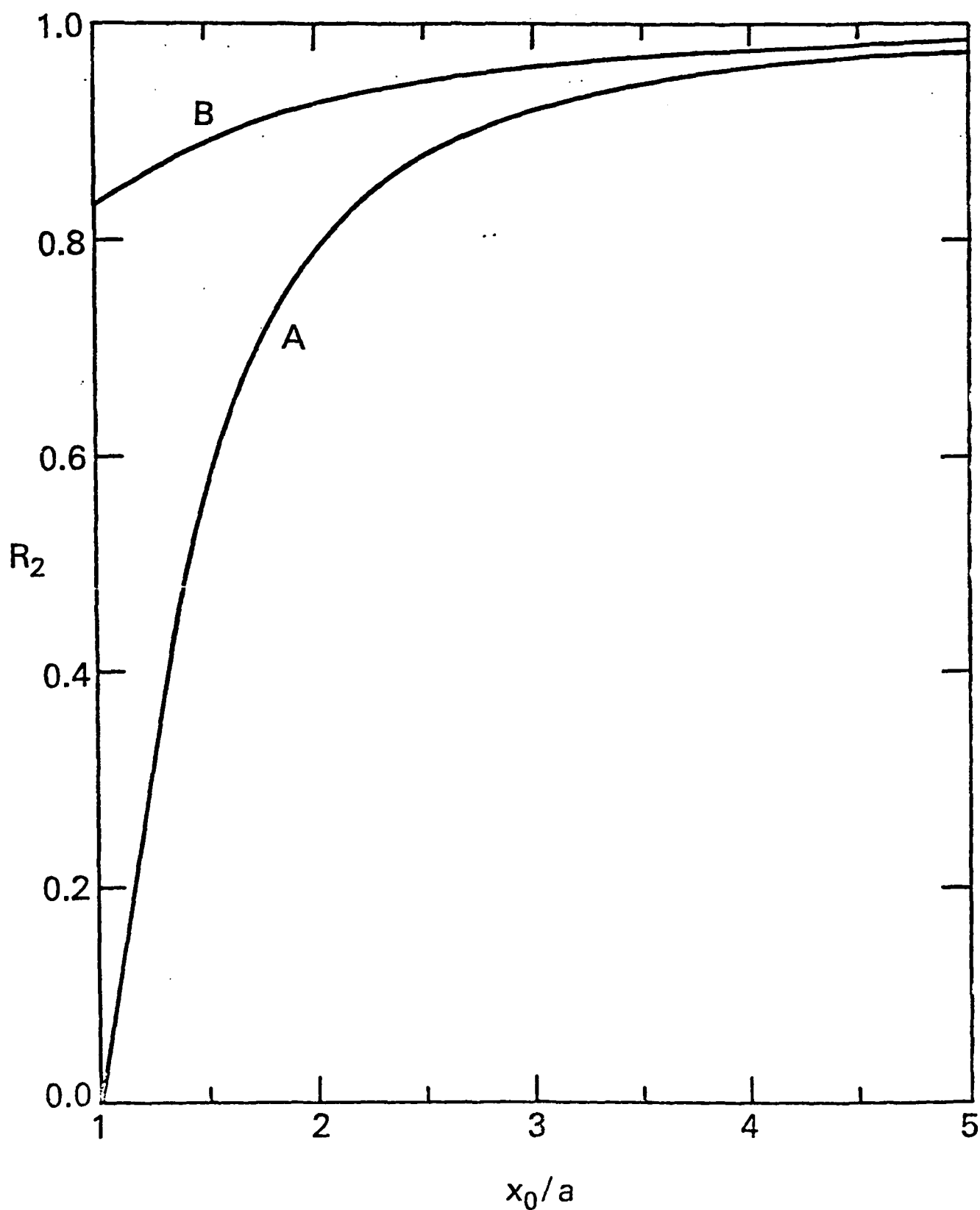


Fig. 1

TWO BUBBLES



x/a
Fig. 2



J-26

Fig. 3

THREE BUBBLES

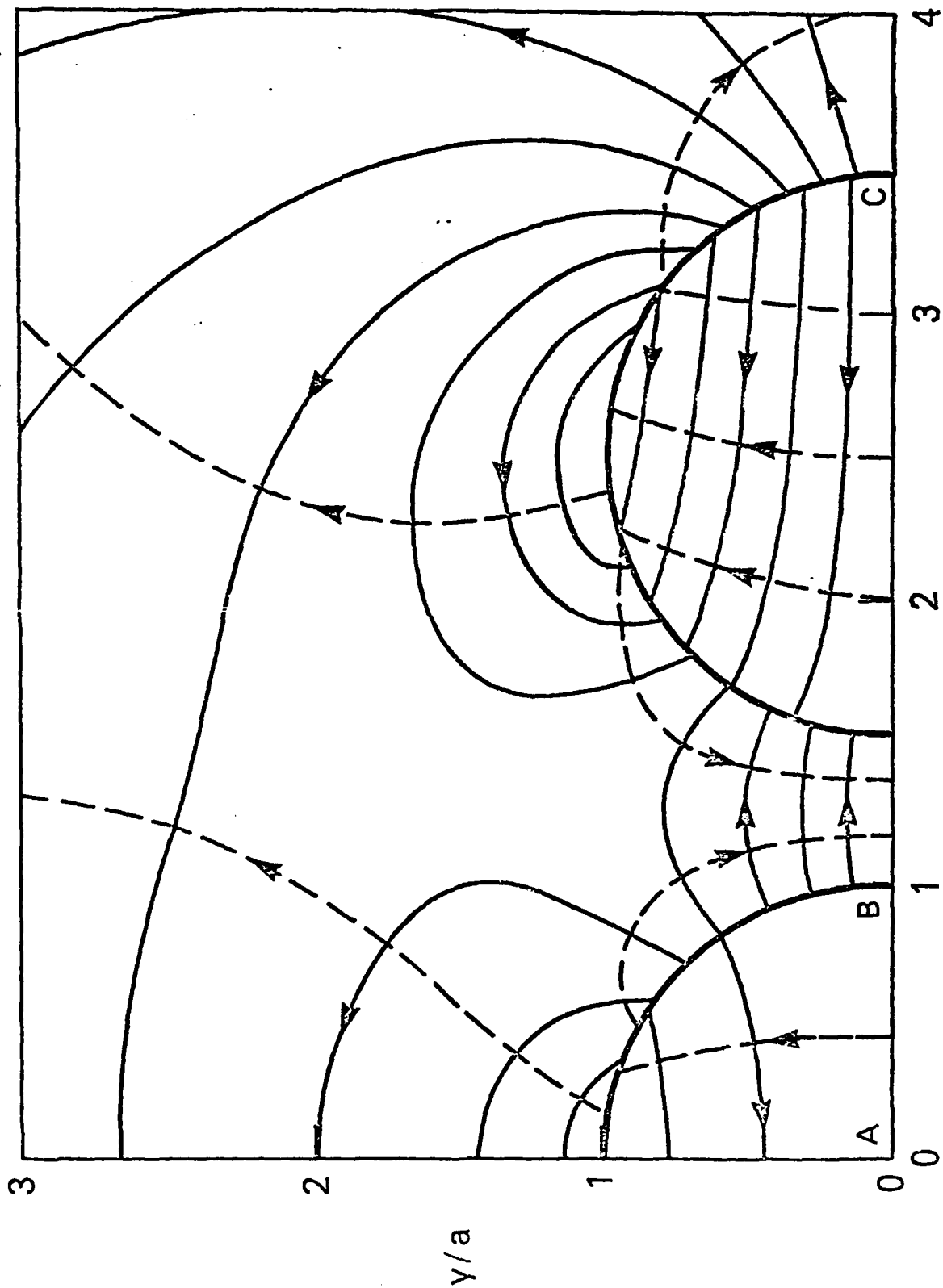
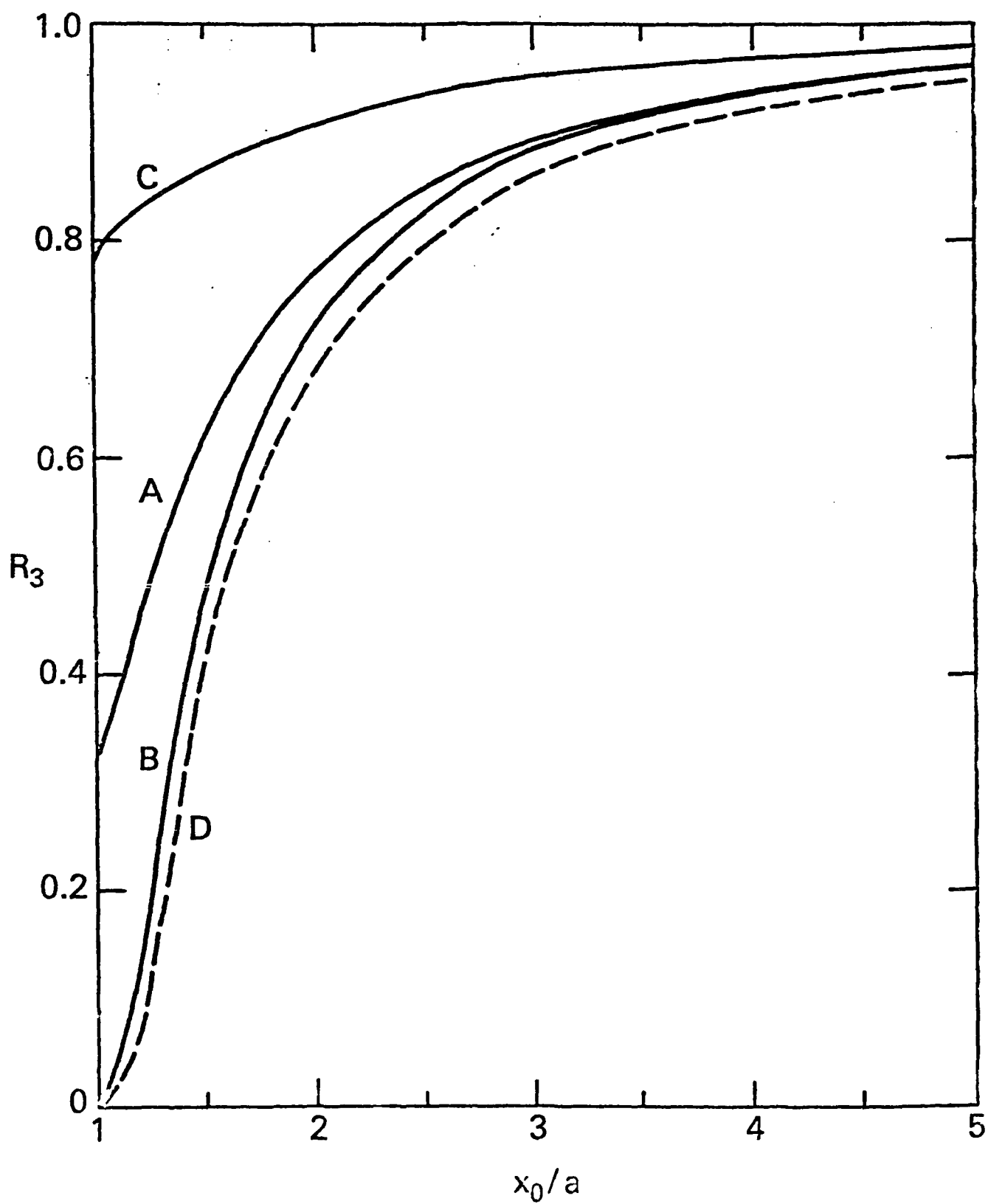
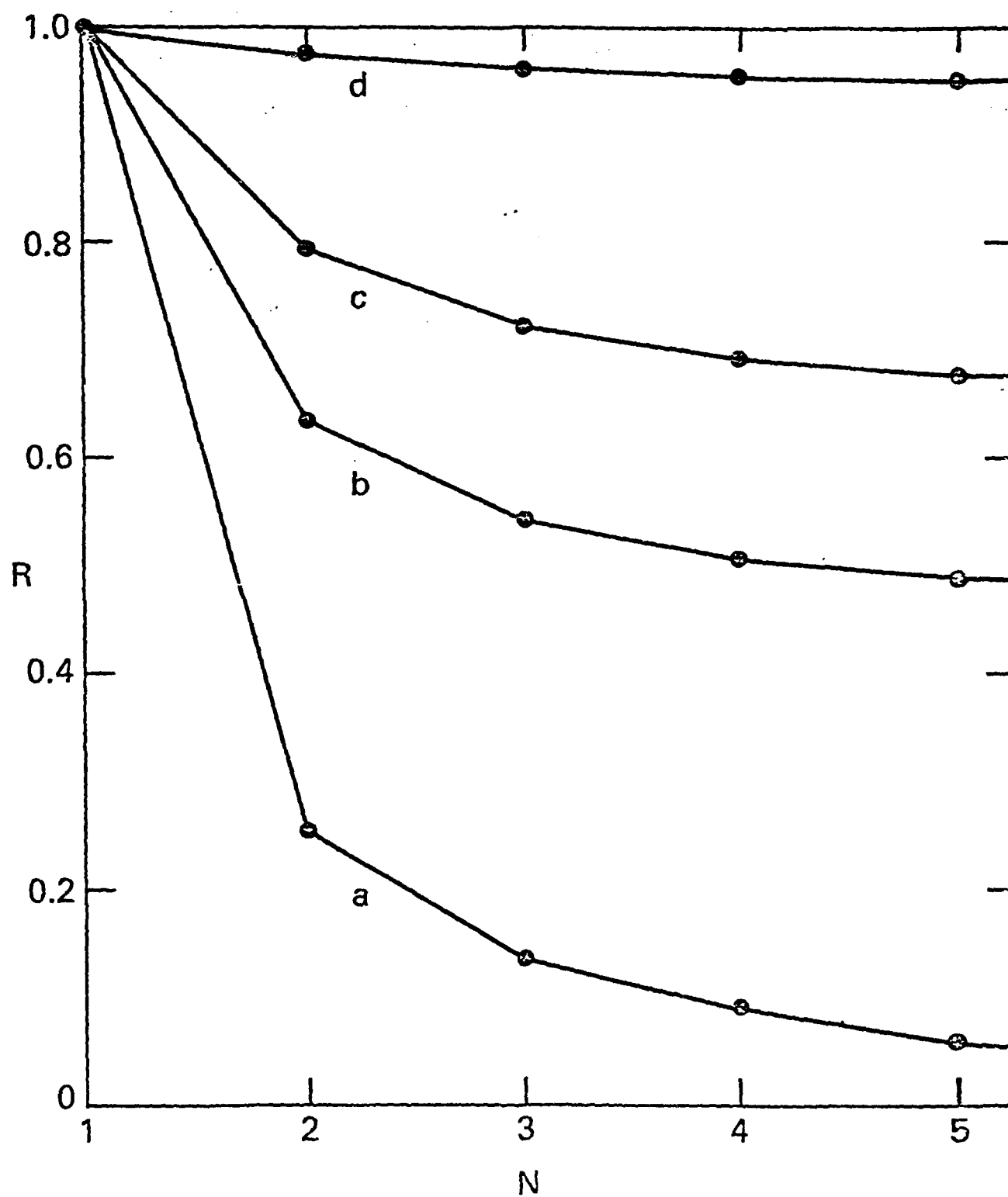


Fig. 4



J-30

Fig. 5



J-31

Fig. 6

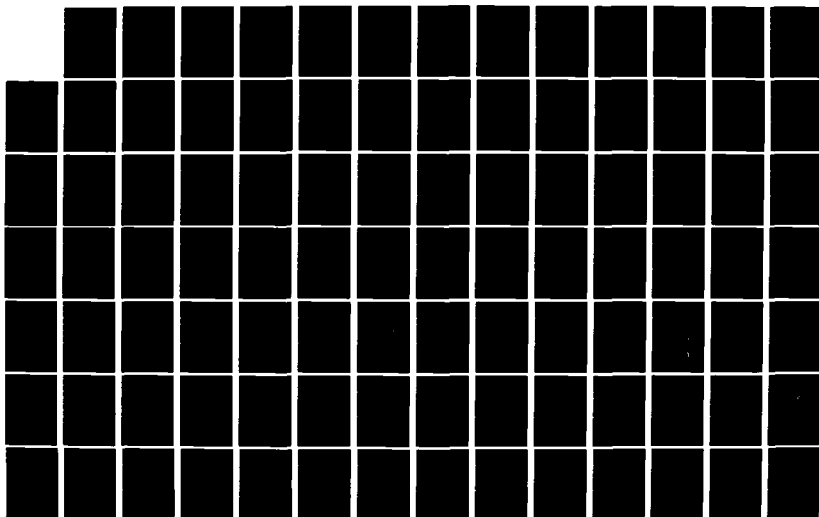
NO-A126 341

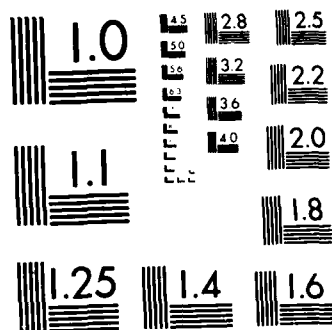
GEOPHYSICAL PLASMAS AND ATMOSPHERIC MODELING(U) SCIENCE 3/7
APPLICATIONS INC MCLEAN VA E HYMAN ET AL. FEB 83
SAI-83-144-WA N00014-81-C-2038

UNCLASSIFIED

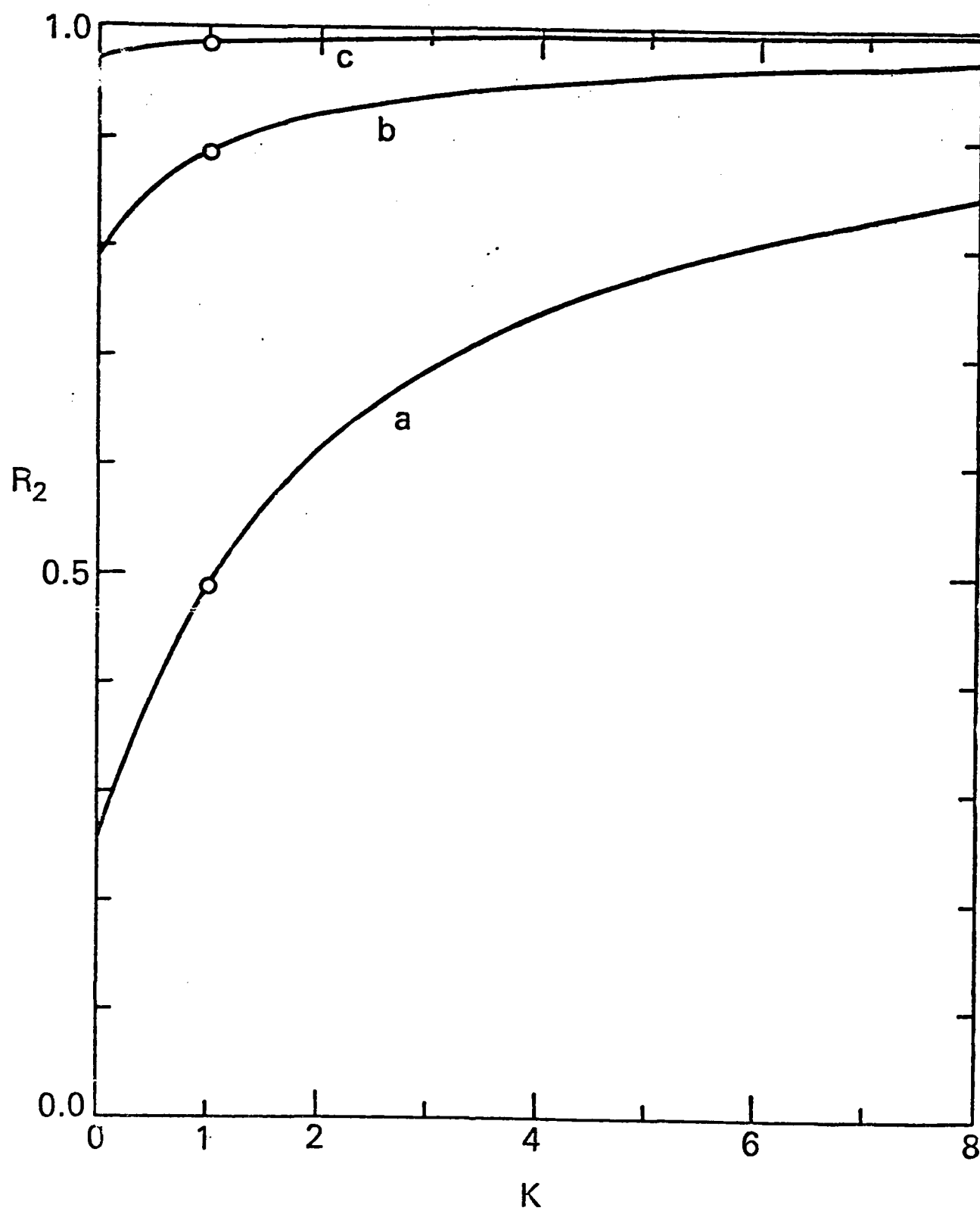
F/G 4/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A



J-32

Fig. 7

APPENDIX K

A REVIEW OF DOUBLE LAYER SIMULATIONS

A REVIEW OF DOUBLE LAYER
SIMULATIONS

Robert A. Smith
Science Applications, Inc.
McLean, Virginia 22102, U.S.A.

Invited review, 1982 International Conference on Plasma Physics, Goteborg,
Sweden: June 9-15, 1982

ABSTRACT

A critical review of numerical simulations of plasma double layers is presented. Physical aspects of the double-layer problem are discussed, together with numerical methods appropriate to treat them. Particular emphasis is given to examining boundary conditions and their effect on the physical dynamics. Some physical processes currently thought to be important for double layers are discussed, and the simulation results are described with particular emphasis on their relation to these concepts.

The review concludes with an interpretive discussion of the results, together with comments on their applicability to laboratory situations and to double layers in space. Some suggestions for future directions in double layer simulation are offered.

I. INTRODUCTION

The past several years have seen a widespread resurgence of experimental, theoretical and numerical investigations of plasma double layers. One may think of several reasons for this. Although double layers (DL) have long been observed in discharge tubes and are closely akin to potential structures in wall sheaths and plasma diodes, much of the current interest stems from suggestions that DL may provide an acceleration mechanism in many energetic phenomena in space, including such diverse examples as auroral discharges, magnetospheric substorms, solar flares, Jovian radio emissions, and extragalactic radio sources. Such suggestions are stimulating and thought-provoking, but are nonetheless premature. Parallel electric fields ($\vec{E} \parallel \vec{B}$) have been observed in the auroral ionosphere [1], and it has been shown [2,3] that if an electric field $\vec{E} \perp \vec{B}$ is applied (by $\vec{V} \times \vec{B}$ motions, e.g.) in a bounded region where a closed circuit contains some portions of current flow parallel to $\vec{E}$, then there must exist regions with parallel electric field components somewhere in the circuit. There is nothing in these arguments to indicate the physical mechanism which generates this parallel electric field, or the spatial structure of the potential. In fact, one of the major unsolved problems in magnetospheric physics is whether the auroral acceleration mechanism involves double layers or anomalous resistivity [4]. Observations have not been definitive on this question; recently, however, observations interpreted as auroral-zone double layers have been reported [5]. Our current knowledge of DL, while growing rapidly, is inadequate to determine the conditions under

which they might form in Nature. Indeed, recent studies have revealed the dynamics of DL formation to be so complex as to provide several stimulating problems in nonlinear plasma theory.

The basic concept of the plasma double layer is that of a free-standing potential structure (not in contact with material surfaces) in which the electric field is maintained by electrostatic charge separation confined in one or more spatial dimensions. In steady state, such DL structures are laminar Bernstein-Greene-Kruskal [6] equilibria, and have been extensively studied in both fluid [7,8] and BGK-Vlasov [9-11] analyses. Such analyses relate the equilibrium state to the imposed boundary conditions, but are unable to address questions about the accessibility of the steady state to a system in dynamic evolution or about its stability. Several experiments [12-14] have provided tests of the steady-state analyses, and others [15-21] have shown the complex dynamics by which double layers evolve. These experiments reveal some fascinating physics questions within the context of the laboratory conditions. But, to the degree that they are influenced by such factors as grids with externally controlled potentials, wall sheaths, and ionization of background neutrals, they provide limited implications about the accessibility of the double layer state under space conditions.

In many cases, and particularly in nonlinear fluid or many-body systems, numerical simulations are useful for investigations at a level intermediate between analytical and experimental modelling. Simulations too, of course, have inherent limitations in the scale and complexity of the descriptions they may economically provide, but subject to these they

offer a means to do "experiments" with controls and diagnostics limited mainly by the ingenuity of the investigator.

A variety of double-layer simulations has already appeared in the literature, and one may confidently expect more to come. The current literature is marked by a diversity of approaches, parameter regimes, and results, and it is outside the scope of this review to discuss each work in detail. Instead, I have attempted to identify some key physical questions, primarily regarding boundary conditions and evolutionary dynamics, and to discuss the various works in the common framework of these questions. Except for brief discussions in Section II intended to help orient readers inexperienced in computer simulation, I shall discuss numerical techniques only in the context of their influence on the interpretation of the results [22]. The emphasis in my discussion will therefore be on basic physical processes, with occasional comments on the applicability of the various simulations to realizable experiments and to double layers in space.

A recent review of experimental work on DL has been written by Torven [23], and reviews of various theoretical aspects have been made by Block [24], Goertz [25], and Carlqvist [26]. In addition, a review of simulations has been given by Goertz and Borovsky [27].

In Section II, I first give a brief overview of simulation methods. The remainder of Section II discusses some theoretical points regarding DL boundary conditions and describes numerical techniques used to treat boundary conditions in DL simulations. Section III discusses some physical concepts currently thought to be relevant to double-layer evolution.

Sections II and III thus provide motivation and background for Section IV, where the simulation results are described. Section V presents a concluding discussion.

Various authors, of course, use different notational and geometrical conventions. As all double layers are topologically equivalent, I adopt a standard Procrustean nomenclature to describe them. Its conventions are:

- all 1-D simulations extend from $x = 0$ to $x = l$. 2-D simulations cover the region $0 < x < l_x$, $0 < y < l_y$;
- The high-potential end of the DL is toward $x = l$; the sense of electron drift is always in the $+x$ -direction. (In 2-D simulations, the periodic coordinate is y .);
- The electric potential ϕ is always equal to zero at $x = 0$, and the potential at $x = l$ is designated as $\phi(l,t) \equiv \phi_{DL}(t)$;
- It is useful in the discussion to distinguish between the macroscopic fluid velocities of a species s , on the one hand, and the drift velocity for a drifting Maxwellian distribution of species s , on the other. The fluid velocity is denoted by U_s , while the drift velocity of a shifted Maxwellian is denoted V_{Ds} .

- The thermal velocity of species s is denoted v_s ; its plasma frequency and Debye length are denoted by ω_s and λ_s , respectively.

Finally, note that the fluid velocity moment of a distribution defined on a half-space in velocity may be very different from the drift velocity of its kinetic description. For example, a half-Maxwellian

$$f = (2\pi v_s^2)^{-\frac{1}{2}} \exp(-v^2/2v_s^2), \quad (v > 0)$$

has a drift velocity $v_{Ds} = 0$ and a fluid velocity moment $U_s = v_s/\sqrt{2\pi}$.

II. SIMULATION METHODS AND THE DOUBLE-LAYER PROBLEM

a) Simulation methods and general constraints.

In general, plasma simulation methods may be categorized into four distinct types: (i) particle-in-cell (PIC) codes, in which the motion of a large number of discrete mass points is followed by applying Newton's laws

$$\frac{d\vec{x}_j}{dt} = \vec{v}_j; \quad \frac{d\vec{v}_j}{dt} = \frac{q_j}{m_j} [\vec{E}(\vec{x}_j) + \frac{\vec{v}_j \times \vec{B}(\vec{x}_j)}{c}] \quad (1)$$

to each particle j ; (ii) Vlasov simulations, in which a distribution function f_s for each species s is advanced in $(\vec{x}, \vec{v})$ phase space by integrating the Vlasov equation

$$\left\{ \frac{\partial}{\partial t} + \vec{v} \cdot \frac{\partial}{\partial \vec{x}} + \frac{q_s}{m_s} \left[\vec{E}(\vec{x}, t) + \frac{\vec{v}}{c} \times \vec{B}(\vec{x}, t) \right] \cdot \frac{\partial}{\partial \vec{v}} \right\} f_s(\vec{x}, \vec{v}, t) = 0 \quad (2)$$

along the phase-space characteristics given by Eq. (1) [where the subscript j is replaced by a set of subscripts, e.g. (i, j) in one dimension, which index the discretized phase space]; (iii) fluid (MHD) simulations, in which macroscopic fluid quantities are advanced by the ordinary MHD equations, in either Eulerian or Lagrangian descriptions; and (iv) hybrid codes, in which one species may be treated as a MHD fluid while the other is described kinetically (in principle by either PIC or Vlasov methods, but usually by PIC).

Because the DL evolution involves microscopic phase-space phenomena such as particle trapping, only the PIC and Vlasov methods are appropriate for studying its dynamics, and both methods have been employed. Although some techniques differ in applying the two methods, there are in principle no essential differences in the physics which may be extracted from them, and the choice of which to use is primarily a matter of taste. Certain practical considerations, however, lend different relative advantages to each. Whatever the method employed, numerical stability requires that the space and time steps be related so as to constrain the speed at which information may propagate on the grid. (In PIC codes, the particle coordinates are continuous, but the fields are calculated on a discrete grid.) The space and time steps are also constrained by the physical scales of the problem; for DL simulations, spatial resolution $\lesssim \lambda_i$ and temporal resolution $\ll \omega_e^{-1}$ are required, where λ_i and ω_e are the ion Debye length and electron plasma frequency, respectively. In PIC codes many particles per Debye "volume" (λ_e in 1-D, λ_e^2 in 2D) are needed to minimize kinetic effects such as those contained in the Balescu-Lenard equation, which describes a weakly interacting gas rather than a collisionless plasma. Vlasov codes are not affected by such phenomena, but instead must limit phase-space graininess by incorporating high resolution in velocity. The result of such considerations is that the Vlasov codes generally manifest finer phase-space resolution at the expense of being limited to shorter physical system lengths. Although the longest DL systems simulated to date have been $4056 \lambda_e$ in 1-D PIC [28] and $400 \lambda_e$ in 1-D Vlasov [42] codes, the more typical range is $\lesssim 1000 \lambda_e$ in PIC and $\lesssim 100 \lambda_e$ in Vlasov simulations; 2-D codes are generally limited to the order of $100 \lambda_e$ in either dimension.

Systems of such size are adequate to explore the essential DL dynamics but this limitation in scale does place special importance on the boundary conditions employed in the simulation. In general, both the definition of the boundary conditions and the way in which they are implemented may affect not only the observed dynamics and their interpretation, but also the applicability of the simulation model to different situations of physical interest. Before discussing such questions in detail, let us consider some current ideas on DL formation.

B) The Bohm and Langmuir criteria; Buneman and ion-acoustic flow regimes.

Consider first the classic steady-state 1-D solution (Figure 1), in which the DL exists in the transition region $0 < x < l$ where the potential varies monotonically between $\phi(x < 0) \equiv 0$ and $\phi(x > l) \equiv \phi_{DL} > 0$. The charge separation in the DL is maintained principally by flux continuity of the "free" particles, those which traverse the entire DL. If the DL is in some sense "strong", i.e., if ϕ_{DL} is sufficiently large, the free particles consist essentially of electrons streaming from $\phi = 0$ towards $\phi = \phi_{DL}$, and counterstreaming ions. Outside the DL, charge neutrality is maintained by "trapped" or reflected populations of ions of density $n_{it}(0)$ on the low-potential end and electrons of density $n_{et}(l)$ at the high-potential end. Supposing for simplicity that the densities $n_{ef}(0)$ and $n_{if}(l)$ are approximately equal (where subscript "f" indicates free particles) we see that as the free particles are accelerated through the double layer, their density decreases

such that $n_{if}(0) < n_{ef}(0)$ and $n_{ef}(l) < n_{ef}(0)$, giving a charge density profile consistent with the potential structure. Kinematic analysis reveals that such a monotonic solution is possible only when the influxes of free particles are sufficiently strong. This leads to a condition known as the Bohm criterion, as it is similar to a condition written by Bohm [30] for wall sheaths. The Bohm criterion is usually written in the form given by fluid analyses [7], in which the free particles of species s are assumed to have temperature T_s and macroscopic drift velocity U_s into the double layer. Then for strong double layers, the Bohm criterion is written as

$$U_e \geq \left(\gamma_e + \frac{T_i}{T_e}\right)^{\frac{1}{2}} v_e; \quad |U_i| \geq (\gamma_i v_i^2 + C_s^2)^{\frac{1}{2}}, \quad (3)$$

where $v_s^2 \equiv T_s/m_s$, $C_s^2 \equiv T_e/m_i$, and where $\gamma_{e,i} = 3$ for adiabatic acceleration.

By definition, the electric field E vanishes at $x < 0$ and $x > l$, so that the DL must be charge-neutral overall.* In an analysis similar to that which yields inequalities (3), this requirement leads to the Langmuir condition

* This is of course an idealization, inspired by the laminar-DL analysis. In actuality, the regions exterior to the DL are probably turbulent due to beam-plasma interactions, and the condition that $E(0,t) = E(l,t) = 0$ should rigorously apply only averaged over the time scale of the oscillations.

$$\frac{n_{e\pm}(z) |U_{\pm}|}{n_{e\pm}(0) U_e} = \frac{n_e}{n_i} \frac{1}{\beta}. \quad (4)$$

The Bohm and Langmuir conditions, in their usually cited forms (3) and (4), are asymptotic conditions in the limit $\phi_{DL} \rightarrow \infty$ [note that ϕ_{DL} does not appear in (3), (4)]. In the context of fluid analysis, somewhat more rigorous conditions, which contain (3) and (4) as asymptotic limits, are given by Levine and Crawford [8]. Equations (47) and (48) of Ref. [8] yield bounds on the influx boundary conditions for finite values of ϕ_{DL} . A result of this modified Bohm criterion is that there is a lower limit to the flow velocities, given by

$$U_s^2 > \gamma_s T_s / m_s, \quad (5)$$

which is generally only slightly less stringent than inequalities (3).

Consideration of the Bohm conditions (3) has led several authors, including the present one [2], to note that plasmas with relative drifts satisfying (3) are unstable to the electron-ion two-stream, or Buneman [31] instability, and to propose that the Buneman instability plays a crucial role in triggering the formation of the double layer, or that the DL is a nonlinear state of the Buneman instability. It was early pointed out, however, that auroral currents are rarely, if ever, strong enough to satisfy (3), and some authors suggested that DL formation could be triggered by ion-acoustic instability when $|U_e - U_i| < v_e$ and $T_e \gg T_i$ [32]. This is indeed the case, although the dynamics and phenomenology of these DL are rather different from the case where (3) is satisfied. In

particular, as is expected from the violation of the Bohm criterion, the potential in these DL is non-monotonic (Figure 2).

We are thus led to distinguish two regimes for the boundary condition on the electron and ion drift velocities:

(i) The Buneman regime $U_s^2 \gtrsim v_s^2$;

and

(ii) the ion-acoustic regime $U_s^2 < v_s^2$, $T_e \gg T_i$.

I shall occasionally speak of DL formed in the two regimes as Buneman double layers or ion-acoustic (IA) double layers, respectively.

C) An Operational Definition of the Double Layer

As I have noted above, DL potential structures in the ion-acoustic regime have the form shown in Fig. 2, rather than the monotonic variation found in the Buneman regime (Fig. 1). As is implied by the discussion of Sec. II-B, however, in the ion-acoustic regime the regions of positive electric field E in Fig. (3b) are a necessary part of the DL structure. As will be clear later on, it is useful to make a conceptual definition of the DL which does not force us to distinguish between the monotonic and tripartite structures in the two regimes. For the purposes of my discussion here, it will be sufficient to define the DL as the region

outside of which $E = 0$ over an extended region. The usefulness of this definition is that it enables us to speak of the DL boundaries unambiguously as the points where the plasma state is implicitly uniform. It will turn out that in simulations, the boundaries of ion-acoustic DL thus defined are coincident with the boundaries of the simulation region, while for Buneman DL they may, but need not, be coincident. When they are not, i.e., when the DL is detached from the system boundaries, it may sometimes be necessary to distinguish between physical boundary conditions at the DL boundaries and numerical boundary conditions at the simulation boundaries.

D. Boundary Conditions for DL Simulations

1. The potential

All DL simulations to date have been done in the electrostatic approximation, in which the electric field is derived from a potential which has an infinite propagation speed on the simulation grid. The boundary conditions (BC) on the potential have generally been of two types:

- (i) fixed potential, in which one boundary ($x = 0$, say) is held fixed at $\phi = 0$, and the other boundary is held fixed at $\phi \equiv \phi_{DL} \neq 0$;
- (ii) floating potential, in which $\phi(0) \equiv 0$ and $\phi(l, t) = \phi_{DL}(t)$, with $\phi_{DL}(t = 0) = 0$.

Fixed-potential BC have been applied in some cases with the simulation

region $[0,1]$ initially empty of plasma, which is subsequently injected from the boundaries [33], or with the simulation region initialized with a uniform plasma state [34,35]. In neither case is the initial state an equilibrium, and the subsequent evolution of the DL shows the dynamics by which the plasma adjusts to the imposed potential.

Fixed-potential BC are most appropriate for simulating experiments, generally in double- or triple-chamber devices, in which the potential is held fixed by a system of grids. Such boundary conditions are not appropriate for studying DL formation under magnetospheric conditions, in which no transient mechanism exists to apply the potential electrostatically along the magnetic field. A parallel potential drop may be applied electrostatically by a mechanism such as that proposed by Goertz and Boswell [36], who investigated the propagation of kinetic Alfvén pulses in which the potential drops in a thin non-neutral sheath travelling at the Alfvén speed. Simulation of such a process is possible in a 2-D formulation with non-radiative electrodynamics [G. Joyce, private communication], but has not yet been attempted.

Floating-potential BC simulate experiments, either in discharge tubes or in triple-chamber devices, in which a current is drawn from the plasma. In principle, such BC are also appropriate to study auroral DL, but no simulations have yet been performed with a complete set of BC properly formulated to apply to the auroral problem, although one paper [37] has attempted to address it with a novel, but not wholly satisfactory, formulation

2) Charge Neutrality and Particle Injection in PIC Codes

In accord with the definition of Sec. II-C, the DL must be charge-neutral overall. This condition may be maintained in simulations by various techniques, all of which require injection of charges into the system by some means. In simulation of a fully ionized system, the only physical method is to inject charges at the boundaries, as is done in most of the cases we discuss below. Some authors, however [38,39], employ a combination of boundary injection and random numerical creation of charges within the simulation volume; this latter method simulates volume ionization of neutrals in a partially ionized plasma such as in the triple-chamber experiments of Coakley and Hershkowitz [13]. For the remainder of the present discussion, we consider only boundary injection.

For simplicity I shall discuss the 1-D case; the arguments generalize immediately to 2-D. For an initially charge-neutral system to remain so for all time, it is necessary and sufficient to have

$$J(0,t) = J(L,t), \quad (6)$$

where J is the current density. A variety of techniques have been used to ensure Eq. (6), but different techniques may have different effects upon the interior dynamics of the DL. It is also important to recognize that a DL is necessarily part of a closed circuit, and the method chosen to implement the boundary condition (6) should be consistent with the relation of the DL to the part of the circuit external to the simulation region. This point is often not discussed explicitly, and some DL simulations have

applied the BC in a manner whose physical relation to an external circuit is dubious.

All methods of boundary injection involve source plasmas external to the simulation region; in general, these sources include both ions and electrons with some assumed velocity distributions. In PIC simulations, a simple method is to count the number of particles of each type leaving the system at each time step, and to replace them by particles drawn from the source plasmas by a variety of prescriptions. These may include:

- (1) A particle exiting at one boundary is replaced at the appropriate boundary by an incoming particle with the same initial velocity as the outgoing one;
- (2) An outgoing particle is replaced by one of the same type and same sense of initial velocity, but with a speed drawn probabilistically from either the flux function or the distribution function of the source plasma at the appropriate boundary;
- (3) Similarly to (2), except that the incoming particle may have either sense of velocity, chosen randomly, and is introduced at the appropriate boundary;
- (4) Periodic injection, in which an outgoing particle is replaced by one of the same velocity at the opposite boundary.

Most of these methods have been employed in the DL literature. Provided that the charge-sharing and placement schemes are such that each particle contributes its total charge to the system, they each conserve both overall neutrality and the total number of particles of each type in the system.*

Before discussing charge injection in Vlasov codes, it is useful to consider the relation of the injection scheme to the concept of the DL in a circuit.

E) The Double Layer as a Circuit Element

A physical double layer is necessarily part of a closed circuit, and the interaction between the DL and the rest of the circuit is manifested in the boundary conditions at the double layer. With one exception [40], DL simulations up to now have not explicitly considered the circuit in their formulation of boundary conditions. With proper implementation, however, most BC can be related to some circuit model; possible examples of such circuits are shown in Figure 3.

*

There are sophisticated algorithms, involving the use of auxiliary cells outside the system boundaries, in which charge neutrality may be preserved by the influx of particles which lie partly in the auxiliary cells. To the best of my knowledge, such algorithms have not been employed in DL simulations in the current literature.

The essential point is that the boundary current density in Eqn. (6) should be determined by the circuit model. In models (a) and (c) of Figure 3, the determination of the boundary current density is affected by the evolutionary dynamics of the DL. In a model circuit containing an applied current source, the boundary current density is not affected by the DL but should be consistent with the current source.

The DL dynamics are affected not only by the value of the boundary current density but also by the plasma distributions which comprise it. Because the current density at each boundary contains contributions from both the interior of the simulation region and the external sources, it is important to consider how the implementation of the BC affects the source distributions. This is especially germane to an interpretation of the interior dynamics, because the sources are generally not considered to be part of the (implicit or explicit) model circuit,* and so the question arises of the physical consistency of a simulation in which the BC implementation significantly affects the source distribution in practice.

Consider for the moment that the circuit, external to the DL, is a known function $J_0(t)$ (it is not necessary here to distinguish between current and current density, as the cross-sectional area of the plasma may

* For this reason, such simulations cannot be directly related to the auroral problem: c.f. Sec. V.

always be absorbed in the circuit equation). Then in the electrostatic approximation, we have

$$\frac{\partial}{\partial t} E(x,t) + J(x,t) = J_0(t), \quad (7)$$

and thus, to maintain $E = 0$ at $x < 0$, $x > l$, Eqn. (6) must be supplemented to read

$$J(0,t) = J(l,t) = J_0(t). \quad (8)$$

(again, for simplicity I am speaking of the 1-D case, and the argument generalizes straightforwardly to 2-D.)

In floating-potential simulations, if the boundary current is not determined by an explicit circuit model, then the implicit model which is physically most consistent with the simulation is one containing a constant current source, such as in Fig. 3(b), with $J_0 = \text{const}$ in Eqn. (8). If we assume for simplicity of discussion that the simulation is initialized with plasma carrying a current (not all simulations are so initialized), then J_0 must be the value of the initial current.

In principle, none of the PIC injection techniques (1)-(4) of Section II-D are capable of satisfying Eqn. (8) throughout the DL evolution. This is because during the evolution, as the conduction current is being interrupted in the interior of the system, the interior contribution to the boundary current varies. Therefore, an injection algorithm which replaces particles one-for-one as they leave the system, while maintaining charge

neutrality, is incompatible with constant-current BC. Because such an implementation in either fixed- or floating-potential models also conserves the total number of charges of each species in the system, it is incompatible too with the evolution of the DL from a uniform plasma state; as the bulk plasma is accelerated through the DL, its density diminishes, and the average density - and thus the total mass - in the simulation region should therefore decrease in time. Evidently, then, the PIC injection schemes (1)-(4) all act to increase the effective boundary current.

This discussion illustrates another point usually overlooked in the interpretation of simulations: even though the plasma sources are not visualized as part of the DL circuit, the flux of plasma the DL must draw from the sources must vary in time in order for the evolution of the DL to be self-consistent with a circuit-type boundary condition. This comment is general; it does not apply only to constant-current BC.

Thus, the evolution of the double layer affects the source plasmas at least up to their apparent normalization, and the applicability of simulations even to idealized versions of such experimental configurations as triple-chamber devices must be considered with this in mind.

One may also ask whether, apart from normalizations, the injection methods affect the apparent form of the source distributions f_j^s ($j = i, e$). In fact, only methods (2) and (3) are capable of preserving the form of f_j^s , and then only providing that the probabilistic selection of particles drawn from the sources is velocity-weighted. In this case, with

the number of injected particles varying in time to satisfy (8), the injected flux function will preserve its form up to normal probabilistic fluctuations. Anticipating the discussion below of results of particular simulations, we may state that method (1) can preserve the form of f_j^s only in the time-asymptotic state of DL evolution, when the potential structure is well-established and the flows are laminar and tending to steady state. This is because the mechanism of current interruption is particle trapping in potential wells, which principally affects particles in some given velocity range. Thus method (1) biases the injected flux towards the entry velocities of untrapped particles.

With floating-potential BC, the periodic injection method (4) clearly tends to accelerate the apparent source distributions of all species as the DL evolves.

In discussing individual simulations below, I shall consider further the possible effects of the injection scheme on the DL dynamics.

F. Injection Algorithms in Vlasov Simulations

Injection algorithms in a Vlasov code involve a prescription for the distribution function of injected plasma, and so must consider both the assumed source distributions f_j^s and, to determine the normalization, some moment over the f_j^s . To preserve overall charge neutrality, this moment must be the injected current density.

Vlasov simulations have been presented by Singh and his co-workers [14,29,33,41,42], Johnson [43], and Smith [40]. As the methods differ among these authors, we discuss them individually.

Smith [40] employs the model circuit of Fig. 3c. At each time, he calculates the circuit current $J_0(t)$ and the interior contributions to the boundary current. In Smith's work the distribution function is defined as constant over a grid cell and is discontinuous at the cell boundaries. Thus an injected distribution function $f_j^{inj} = \alpha f_j^s$ is loaded into an auxiliary cell, with α chosen at each boundary such that (8) is satisfied. The influx occurs as the f_j are advanced in space on the characteristic $\frac{dx}{dt} = v$. Although the f_j , and thus the charge density $\rho = \int dv(f_i - f_e)$, are not summed at the boundaries in Smith's algorithm, his numerical implementation is such that $\partial J / \partial x = 0$ at the boundaries and so formally ρ retains its initial zero value at the boundaries at all times. This is a by-product of the numerical algorithm, however; it is the satisfaction of BC (8) which guarantees overall charge neutrality.

In the work of Singh and his co-workers, fixed-potential BC are employed and no model circuit is considered. Singh instead adopts as an injection criterion that the charge density vanish at the boundaries; i.e.,

$$n_e(0,t) = n_i(0,t) = n(0); \quad n_e(l,t) = n_i(l,t) = n(l). \quad (9)$$

This is a necessary, but not sufficient, condition for DL. As the f_j in Singh's algorithm are calculated explicitly on the cell boundaries, he sets the normalization of f_j^{inj} as a boundary condition to ensure Eqn. (9).

Johnson [43] also uses (9) as the injection condition, but his method of implementing it is unclear from the context.

However implemented, though, the condition (9) does not guarantee overall charge neutrality, and the system does not remain neutral in Refs. [29,33] and [43]. I shall discuss this further below.

III. SOME CURRENT NOTIONS RELATED TO DOUBLE-LAYER EVOLUTION

In this section, I introduce some concepts that have been proposed to be important in DL evolution. In general, the roles of such processes have not been analytically demonstrated, but they help provide a conceptual framework to bridge the gap between steady-state theory and the interpretation of simulations. The principal such ideas concern the roles of (i) linear wave instabilities and their nonlinear evolution, and (ii) intrinsically nonlinear plasma states such as solitary pulses, and their accompanying phase-space signatures.

A) Solitary Waves and Phase-Space Holes

In the steady state, the DL is a particular class of solitary wave, one which is non-periodic in electric potential ϕ . Analysis of such structures often proceeds from use of the pseudo-potential V (sometimes called the Sagdeev potential), where Poisson's equation is written in 1-D, e.g., as

$$\frac{d^2\phi}{dx^2} = -4\pi\rho \equiv -\frac{dV(\phi)}{d\phi} \quad (10)$$

Equation (14) has the first integral

$$\frac{1}{2} \left(\frac{d\phi}{dx} \right)^2 + V(\phi) = 0, \quad (11)$$

where the choice of zero as the integration constant limits the physical

range of V to $V < 0$. Electric potentials and their corresponding pseudopotentials are shown for various cases in Figure 4. One sees that solitary waves in general comprise a class of V such that $V'(0) = 0$, while DL are further restricted by the condition $V'(\phi_{DL}) = 0$. I shall henceforth refer to cases such as Figs. 4b,c, in which $V'(0) = 0$ but $V'(\phi_{max}) \neq 0$, as periodic solitary waves (PSW)*, as they satisfy periodic BC on the potential ϕ .

As seen in Figure 4, PSW may have either positive or negative polarity; the corresponding phase spaces for the plasma contain vortices, or "holes", in the electron or ion distributions, respectively. Such vortices have been extensively discussed in the literature [45-48]. Consider for definiteness the case of an ion hole (Fig. 5) in a negative-potential PSW which results from a negative charge density in some region. The vortex in the ion phase space results because ions from the left are attracted to the excess electron charge and accelerated to higher speeds in the x-direction until they pass the minimum potential of the PSW, then are decelerated. Similarly, ions from the right are first accelerated, then decelerated as they pass over the PSW. For both ion and electron holes, therefore, the "circulation" of the vortices in phase space is clockwise.

* Many authors, including the present one, have sometimes referred to PSW as solitons, but this term is more properly reserved for a subclass of PSW which retain certain self-similarity properties [44].

Existence conditions for both ion and electron holes have been determined by analysis of the pseudopotential [46,47]. For various reasons, however, the pseudopotential method is ill-suited to stability analysis, and the stability of the holes has not been extensively investigated (see, however, [49]). As may be seen from the electron phase space diagram of Fig. 5, however, a negative PSW reflects some of the electrons streaming towards from the left; if the PSW is growing in time, this reflection implies an increasing interruption of the current, which is the fundamental mechanism of DL evolution. Thus, Hasegawa and Sato [46] have proposed that the existence of a negatively-polarized PSW is a necessary intermediate stage in DL evolution; interruption of the electron flux (here taken as being in the + x-direction) leaves an ion-rich region (an electron hole) on the downstream side of the ion hole, raising the potential there. Broadly stated, therefore, the hole concept is a useful one for understanding double-layer dynamics, not necessarily because the plasma passes through a succession of quasi-stable hole equilibria, but rather from the more general possibility that one may recognize in the dynamics localized, persistent structures which correspond to this concept.

B) Wave Instabilities

In principle, wave instabilities are relevant to the DL problem (i) as a mechanism for triggering the DL evolution, (ii) as a possible destabilization mechanism in the evolved DL, and (iii) as consequences of the DL, manifested in the regions exterior to the DL itself.

With regard to aspect (i), we have already noted (c.f. Sec. II) that current-driven instabilities are a seemingly inevitable concomitant of the boundary conditions necessary to support the DL*, and it is thus natural to inquire into what role these instabilities may play in the evolutionary dynamics. As indicated by Hasegawa and Sato [46], the negative PSW that they postulate as a starting point for DL evolution does not arise spontaneously, but must be created "adiabatically". This requires the localization of electron space charges. Wave instabilities may accomplish this in at least two ways:

- (1) In the case of a convective instability propagating away from a defined boundary, spatial growth of the wave envelope leads to a ponderomotive force on each species s ,

$$F_p \sim -\frac{1}{m_s} \frac{\partial}{\partial x} \left\langle \frac{E^2}{8\pi} \right\rangle \quad (12)$$

* We spoke in Sec. II of the Buneman instability, which requires $|U_e - U_i| \geq V_e$ and is independent of the temperature ratio T_e/T_i , and of the ion-acoustic instability, which occurs with $|U_e - U_i| < V_e$ but requires $T_e \gg T_i$. Simulations in the 1A regime but with $T_e = T_i$ show no DL formation. These simulations are 1-D unmagnetized, however, and in principle there seems to be no reason why the electrostatic ion-cyclotron instability, where $|U_e - U_i| < V_e$ but T_e/T_i is arbitrary, may not also lead to DL. Such a case must be investigated in multi-dimensional magnetized simulations, however.

where $\langle \dots \rangle$ indicates averaging over the wave period. Being proportional to n_s^{-1} , the force (12) acts to brake the electron motion if the turbulent envelope grows in the direction of electron drift. This mechanism for triggering DL evolution was proposed by Smith and Goertz [2];

(2) As will be discussed further below, in simulations where the BC

$$E(0,t) = E(l,t) = 0$$

is enforced, the possible wave modes are limited to a discrete set of standing waves with wavenumbers k_n given by

$$k_n l = n\pi \quad (n = 1, 2, \dots).$$

In this case, it is improper to speak of convective growth, but constructive interference between two or more adjacent modes may lead to what appears to be convective growth [40], with attendant particle trapping and apparent ponderomotive force effects. This possibility was first noted by Berk and Roberts [45].

Aspect (ii), the influence of interior wave instabilities on the stability of the DL, has not been directly addressed in the analysis of any simulation but has been addressed theoretically by Knorr and Goertz [10] and by Wahlberg [50]. It appears that, in 1-D at least, the DL is stable against electrostatic instabilities because the sharp density gradients in the interior of the DL mitigate against the growth of unstable waves, the

density gradient scale length being shorter than the growth length.

Aspect (iii) relates most directly to particle heating in the exterior regions and to possible disruption of the double layer by trapping of the current influx. This aspect will be discussed more fully below.

C) The Ion-Acoustic Caviton

Another possible triggering mechanism was proposed by Carlqvist [51] and further investigated by Raadu and Carlqvist [52] and Belova et al. [53]. The concept is similar to the PSW, but differs in that it applies to slow, quasi-neutral motions in which the electrons may be taken to be in steady-state over the (adiabatic) time scale.

The central idea here is that quasi-neutral density perturbations may be driven unstable by a sufficiently large current density. When the electrons carry a steady flux $I_e = n_e v_e$, the momentum equation relates the slow-time-scale electric field to the electron density $n_e (= n_i)$ by

$$E = - \frac{\partial}{\partial x} \left(\frac{m_e I_e^2}{2en_e^2} \right) ; \quad (13)$$

Belova et al. refer to the quantity in parenthesis as a "negative pressure", in the sense that it is increased as the density decreases.

A local perturbation with depressed density may thus be driven into a caviton, as the quasi-static field E forces plasma out of the region (compare Eqns. (12) and (13); the ponderomotive force is the mechanism of

the oscillating two-stream, or modulational, instability). Raadu and Carlqvist performed a Vlasov simulation of the early development of the instability and Belova et al. offer a self-similar analytic solution; both papers find the caviton growth to be explosive. The exact connection to DL evolution is not clear. Carlqvist [51] suggests that as the density is decreased, the quasi-static approximation breaks down and the current is interrupted, charging the downstream side to positive potentials. This suggestion seems to be borne out in recent simulations by Belova et al. [53].

It should be remarked that in general one would expect a threshold depth for the density perturbation to develop into a caviton; this threshold would presumably be a function of the current density and the ion and electron temperatures. If a threshold criterion has been derived, however, I am unaware of it.

IV. SIMULATION RESULTS

Table 1 categorizes the simulations I shall discuss. Because of the diversity of the simulation models, and because some results seem to be influenced by - if not artifacts of - the boundary conditions and numerical techniques employed, I shall discuss various works in distinct groups. These are:

- Group A: 1-D, fixed-potential simulations;
- Group B: 1-D, floating-potential Buneman regime;
- Group C: 1-D, ion-acoustic regime, floating and periodic potential;
- Group D: 2-D simulations.

I remind the reader that I have adopted a standard notation and geometry to describe the simulations (c.f. Sec. I).

A) 1-D, Fixed-potential Boundary Conditions

The works in this group include those of Joyce and Hubbard [34], Hubbard and Joyce [35], Singh [29,33], and Johnson [43]. These simulations share two common features. First, the applied potentials ϕ_{DL} are "strong": $\phi_{DL} \gg T_e/e$. Second, they are initialized in a nonequilibrium state; the transient response of the plasma to the initialization is violently nonlinear and dominates the entire evolution.

The results of Refs. [34-35] are typical. The plasma is initialized with a uniform electric field superimposed on a Maxwellian plasma. In contrast to other works, there is no drift in either the initial plasma or the external-source distribution functions, which are taken to be half-Maxwellians; injection algorithm (3), described in Section II, is employed.

At the start of the run, a large-amplitude positive solitary pulse forms near the low-potential end and propagates toward the high-potential end (Fig 6). The pulse is nonlinear even at very early times, and during its propagation its maximum amplitude may far exceed the applied potential ϕ_{DL} . The pulse forms because the applied electric field induces a current in the plasma; as accelerating electrons leave the system at the high potential end, those injected at the low-potential end are slower on the average than the outgoing particles. Thus the interior electrons tend to separate from the boundary electrons, leaving an electron hole which propagates towards the high-potential boundary as plasma inflow tends to neutralize the low-potential side.

This pulse is the characteristic dynamic feature of all the fixed-potential simulations. The pulse relaxes into a monotonic DL as it traverses the system; the relaxation dynamics are not fully understood. The motion of the pulse is due to current interruption [54], the net force on the pulse being produced by the combined applied and self-consistent electric fields [55]. Thus, the motion stops when the overall charge density in the DL is zero.

In the fixed-potential Vlasov simulations, there are indications that during at least part of the evolution, the entire system is not charge-neutral (c.f.: [43], Fig. 1; [33], Fig. 4; [29], Figs. 5,9); I have already noted (Sec. II) that the Vlasov injection algorithms of Singh and Johnson are incapable of preserving system neutrality. It is unclear whether, and how, system non-neutrality may be affecting these results. It is possible that excess charge may reside in boundary sheaths, which would at least affect the Vlasov injection algorithms but may not affect the qualitative dynamics or structure of the DL itself, provided that it is surrounded by field-free regions between it and the sheaths. It is also possible that in the asymptotic state, charge neutrality tends to be restored [N. Singh, private communication].

A central result in DL simulations is the existence of a scaling law relating the width L and potential of the double layer:

$$\frac{e\phi_{DL}}{T_e} = \alpha \left(\frac{L}{\lambda_{eo}} \right)^2, \quad (14)$$

where λ_{eo} is the upstream electric Debye length. The form of (14) is readily obtained from dimensional analysis; the physical content resides in the constant α , and is essentially determined by the necessity that the distribution function of reflected particles be positive definite [56]. Varying definitions may be used for L , but general agreement exists that α is of order 0.1, and all strong 1-D double layers obey this scaling.

When the system is long enough so that the double layer, whose length L is given by (14), is well detached from the system boundaries, plasma

waves are observed in the static-field-free regions at the high- and low-potential ends. These waves are generated by beam-plasma instabilities of the accelerated electrons and ions with the ambient plasma. To observe them, the field-free region on either side must be larger than some critical length which depends on the wave mode. This is because the beam-plasma resonance conditions determine a range of unstable wavelengths which must be contained within the system. As the boundary conditions reflect waves, the observed waves are standing modes, and so the instabilities are absolute. As expected, on the high-potential side Buneman [34,35] and Langmuir [29,33] modes are identified, and ion modes [29] on the low-potential side. On the high-potential side, these waves may cause turbulent heating of both the beam and trapped electrons [41].

In general, the fixed-potential simulations in long systems ($l > 200 \lambda_e$) observe the DL to be cyclically unstable; the DL is disrupted and replaced by a new DL which forms from a solitary pulse in the same manner as the preceding one. The disruption and regeneration of the DL occurs over several ion plasma periods, as it is the ion inertia which dominate the dynamics. The dynamics of this cyclic instability are not well understood. Hubbard and Joyce [35] suggest that the disruption is triggered by current blockage due to electron trapping in ion-driven oscillations on the low-potential side; such trapping is clearly evident in the electron phase space (Fig. 7). Thus, new electron holes and solitary pulses may form there, driven either by the self-consistent dynamics or by the applied electric field demanded by the fixed-potential BC. Singh [29] suggests that the regeneration is related to the formation of ion-acoustic cavitons, which may be seeded by large-amplitude slow oscillations. These

may be excited either by the ion beam, or by the bounce motion of the trapped electrons [43].

B) 1-D, Floating-potential Buneman Regime

This group includes the papers by Goertz and Joyce [56], Smith [40], and Belova et al. [53], which are quite different from one another both in their approaches and in the observed dynamics.

Goertz and Joyce did a PIC simulation, using injection algorithm (1) of Sec. II, and with the plasma initialized to include stationary Maxwellian ions and drifting Maxwellian electrons, $T_e = T_i$, with $V_{De} = 0.5 V_e$, V_e , and $1.5 V_e$. No DL formation was observed at $V_{De} = 0.5 V_e$, but they obtained DL in the cases $V_{De} = V_e$, $1.5 V_e$. In an infinite plasma, the case $V_{De} = V_e$ would be stable; thus the triggering mechanism in these simulations must be the excitation of an oscillation mode that depends on the finite system length. The evolution of the DL is marked by the appearance of a solitary pulse near $x = 0$, which propagates toward $x = l$ in attempting to neutralize. The potential is exhibited only for $t > 6\omega_i^{-1}$, when the pulse is already nonlinear, and so it is difficult to deduce the pulse components; no wave spectra or phase spaces are shown. The pulse exhibits strong overshoot before settling into a stationary DL.

By varying the system length, Goertz and Joyce first obtained the scaling law (18).

Smith [40] employed the circuit model of Figure 3(c) in a Vlasov simulation. In this model the current density at the boundary is determined by solving the equation for the circuit including the simulation region. In dimensionless units, the boundary current density is thus given by

$$J_o(\tau) = R^{-1} \left\{ -\phi + \int_0^\theta d\theta' \phi_{DL}(\theta') \exp(\theta' - \theta) \right\}, \quad (15)$$

where $\theta \equiv R\tau/L$. Equation (15) includes an important feature: evolution of the DL self-consistently determines the circuit current.

The boundary reservoirs are taken to be drifting Maxwellian for both electrons and ions, with fixed drift velocities $V_{De} = 2V_e$ and $V_{Di} = -2V_i$; $T_i = T_e$. Injection is accomplished as described in Sec. II, guaranteeing charge neutrality (which is monitored) both overall and on the boundaries $x = 0, L$.

The evolution of the DL clearly depends on wave instabilities; the linearly unstable wave mode is the finite-system analog of the Buneman mode, which is unstable for a discrete set of standing modes which are nearly aperiodic. In the linear phase the wave envelope, which is composed mainly of the neighboring most unstable modes, appears to grow in the sense of electron drift. (This is not convective growth per se, however, as the waves are standing modes.) As electrons become trapped, the potential rectifies and quasi-stationary solitary pulses develop, with alternating ion and electron holes. As this happens the k -spectrum changes, so that instead of the four to five linearly unstable wavelengths the system is

dominated by the smaller-k structure of the pulses. This transition is clearly associated with the onset of ion trapping.

The circuit model conserves energy and so is suited to study energy flow and conversion during the evolution. Figure 8(a) shows the time history of ϕ_{DL} and J_0 for a run in which the inductive time constant L/R is equal to an ion crossing time at the drift speed V_{D1} . The time history of the quantity

$$P_J = \int_0^L dx E[J(x,t) - J_0(t)],$$

which is the volume dissipation of the interrupted current, is shown in Figure 8(b). The inductive overshoots and undershoots are marked by strong relaxations between current dissipation (field storage) and heating (field dissipation). In Figure 9 are shown the potential, electric field, and phase spaces during the inductive overshoot and recovery between $\omega_e t = 85$ and $\omega_e t = 125$; the dynamics are dominated by the coalescence of two ion holes and the subsequent destruction of the hole by jetting of the ions toward the high-potential end.

An interesting consequence of the self-consistent determination of the circuit current (15) is that as the DL evolves and the magnitude of the current density decreases, the DL sees an apparent decrease in the density of the plasma sources (this is a consequence of the implementation of the injection keeping the drift velocities of the source distributions constant). For this reason, the scale length of the grid cells changes with respect to the upstream Debye length $\lambda_{e*}(t)$, and the DL eventually

occupies the entire system according to the scaling law (14), with $\alpha = 0.17$, approximately the value found in Refs. [33] and [56]. Thus, with this injection implementation no waves may be observed outside the DL, and it is stable.

Belova et al. initialized the simulation in the ion frame, as did Goertz and Joyce, but with $V_{De} = 1.8 V_e$ and in a slightly longer system (λ_e). They injected a constant influx of both electrons and ions, and maintained system neutrality by random creation of charges within the simulation volume [Yu. Sigov, private communication]. In contrast to the continuous development of the system potential observed in Refs. [40], and [56], Belova et al. observed recurrent explosive development of the potential, the explosions occurring within a few ion plasma periods and recurring on the time scale of $360-400 \omega_e^{-1}$ (Figure 10). They attribute these phenomena to the explosive instability of the ion-acoustic caviton. The explosive potential develops to approximately the level given by Eqn. (14) before the first collapse; subsequent explosions are weaker, probably being modulated by vestigial turbulence.

C) 1-D, Ion-Acoustic Regime

PIC simulations of IA double layers have been performed by Sato and Okuda [28,32], Kindel et al. [57], and Hudson and Potter [58]. The phenomenology and dynamics they exhibit are quite different from the Buneman DL discussed above.

The original IA simulations of Ref. [32] were performed in a periodic

system, with $\phi = 0$ at both ends. Properly speaking, this feature would be inconsistent with the concept of the DL as giving a net potential drop. As we have seen in Section II, however, in the LA regime the potential must have a non-monotonic structure such as in Fig. 2; In the context of the generalized definition of a DL (Sec. II-C), the DL comprises this entire tripartite structure, but all of the authors cited above define the DL as the central region where the potential rises sharply. Sato and Okuda argue that the dynamics of this central region should not be greatly affected by the periodic boundaries, because they are far enough from the central potential step that the dynamics are not contaminated by particles from the boundaries, owing to their long traversal times. By this argument, the periodic BC would simply exaggerate the decreasing slopes in the boundary-adjacent regions. In a run in which the potential at $x = l$ was allowed to float [58], the qualitative features of the DL were not greatly changed from the periodic case. It must be noted, however, that in this case, and in the applied-field cases of Kindel et al. [C. Barnes, private communication], the periodic injection algorithm (4) of Sec. II was used, and this is physically inconsistent with any implicit circuit concept. The exact effect of these various BC cannot be weighed.

In the floating-potential cases of [28], Sato and Okuda also discuss a circuit model for the simulation, where the external elements are a potential source and a resistance. It is unclear what use is actually made of this concept, although some slight differences in the time histories of the total electron field energy density and the current density are noted when the external potential is varied by a factor of 20 [28; Fig. 7]. However, it appears from this discussion [28; p. 3361] that the circuit

model was not exploited to control the simulation; in particular, the boundary current is again controlled by the periodic injection algorithm.

It is found that long systems, of the order of $1000 \lambda_e$, are required in order to observe the formation of IA double layers, and that in contrast to the Buneman DL they require very long times to form; the typical time scale is $1000 \omega_e^{-1}$, about an order of magnitude longer than the Buneman case. A time sequence of potential profiles from the floating-potential ("bounded model") case of Ref. [28] is shown in Figure 11. A profile from an intermediate time in this run, together with ion and electron density profiles, is shown in Figure 12. These illustrate the general phenomenology of the IA regime. A long, gentle potential drop is seen in the leading region, which is interpreted as being due to anomalous resistivity developed as the drifting electrons lose their momentum to the excitation of ion-acoustic waves. All the data are quite noisy, as a broad inherent wave spectrum is developed. Near the middle of the simulation region the central potential step, or DL, is formed. The striking characteristic, which is always evident, is the existence of a sharp negative potential spike immediately preceding the DL; this is an ion hole.

When the system is much larger than $1000 \lambda_e$, periodic DL are observed [28], each evolving on approximately the same spatial scale. In terms of the schematic potential structure of Figure 2, the descending potential ramp downstream of the central potential step serves as the leading region for the succeeding DL. In the floating-potential case, the successive DL give a cumulatively increasing potential use across the system.

The ion-acoustic DL observed in these simulations are all "weak", the central potential jump being of the order $e\phi \lesssim T_e$. They are also unstable; Sato and Okuda [28] interpret their decay as being due to emission of ion-acoustic solitons.

D) -2-D Simulations

Comparatively little work has been published on 2-D simulations; in this brief discussion I shall concentrate on qualitative differences in techniques and physical implications between the 1-D and 2-D cases.

DeGroot et al. [59] considered a periodic model with $V_{De} \lesssim V_e$, which is Buneman-stable but may be unstable to ion-acoustic waves if $T_e \gg T_i$, or to Pierce modes because of the periodic BC on the potential. They find that when $T_e \gg T_i$, IA turbulence leads to anomalous resistivity but prevents the formation of DL; in light of the 1-D IA regime simulations, this is consistent with the fact that DeGroot et al. used a short system ($141 \lambda_e$). In contrast, when they initialized the plasma with a localized density fluctuation, as in the ion-acoustic caviton concept of Raadu and Carlqvist [52], the DL formed in the manner of the long 1-D IA simulations.

Wagner et al. [37] attempted to simulate auroral DL with novel BC. The external sources at the high-potential end were modelled to simulate upstreaming ionospheric ions and backscattered electrons. At the low-potential (plasma sheet) boundary, a current sheet with high-energy ions and neutralizing electrons is introduced. There is a uniform magnetic field $B_0 \hat{e}_x$, and no potential is applied across the two boundaries. When

the current sheet, of thickness $\ll l_y$, enters the simulation region, the ions gyrate outside of the current sheet, while the electrons are tightly magnetized. This causes a potential drop perpendicular to $\vec{E}$ which, because of the conducting boundary at the high-potential end, closes in the U-shaped structure characteristic of 2-D double layers.

Several aspects of this model are unclear. First, the relation of the plasma injection to the Birkeland current system is problematic; the potential is equal to the kinetic energy of the injected ions, which are representative of the distant plasma sheet population, not the auroral acceleration region. Second, although it is stated that there is a background plasma, its effects are not discussed. The physics modelled most closely resembles that of a beam with self-fields injected into a drift tube [60], the principal modifications being that the self-fields are generated by the finite ion Larmor radius, and that the electric field is accelerating due to the presence of the downstream conducting boundary.

Kindel et al. [57] note that in 2-D unmagnetized simulations, DL do not readily form because electrons are free to scatter around localized potential fluctuations instead of being trapped. In the presence of a magnetic field, DL develop independently on neighboring bundles of field lines with a perpendicular correlation length of about twice the DL thickness.

Borovsky and Joyce [38,39] performed a series of 2-D simulations with fixed potentials in the boundaries $x = 0, l_x$ and periodic BC in y . When the high-potential boundary was at constant potential in x , plasma DL

developed with results similar to those in 1-D fixed-potential simulations. When the high potential boundary was of the form

$$\begin{aligned} \phi(l_x, y) = & \begin{aligned} & 0, & 0 < y < a \\ & \phi_{DL}, & a < y < b \\ & 0, & b < y < ly \end{aligned} \end{aligned}$$

U-shaped 2-D structures formed. An interesting result of these structures is that both parallel and normal to the magnetic field, and regardless of the magnetic field strength, they scale with the Debye length and not with ion gyroradius as might be expected.

V. SUMMARY AND DISCUSSION

In the preceding discussion I have concentrated on describing the essential physical processes revealed by simulations of DL formation, and have indicated to some extent the general influence of numerical methods on the interpretation of these phenomena. In this Section, I shall briefly discuss some interpretive conclusions, and shall conclude with some suggestions for future exploitation of simulations.

In the simulations utilizing fixed-potential boundary conditions, all the observed dynamics are controlled by the intrinsically nonlinear initial response of the plasma to the applied potential. The strong double layers developed in these simulations are driven by the boundary condition, although in the asymptotic state they obey the scaling law (14) found in Ref. [56].

In the floating-potential Buneman regime, the results of Ref. [40] clearly show evolution of the DL progressing through a linear phase of wave instability, followed by nonlinear phases dominated by trapping, first of electrons and then of ions. The DL development is marked by highly dynamic interactions of nonlinear electron and ion hole states, again with the ion holes persisting longest and dominating the final evolution. Detailed studies of these dynamics are in progress.

Belova et al. [53] find recurring, explosively unstable DL. Their results are in marked contrast to those of Ref. [40]. The principal reason for this is that their treatment of the ions does not provide sufficient influx for stability; there may be other reasons as well.

The ion-acoustic double layers offer some important and difficult questions. Most authors [57,58,59,61] interpret their results at least partly in terms of anomalous resistivity. I have avoided this term in describing the simulation phenomenology. The most elementary concept of anomalous resistivity is in terms of an anomalous collision frequency arising from turbulent wave-particle interaction; in general, there is no simple relation implying a constitutive relation between current and electric field, and definitions of the resistivity may be based on different considerations and may involve spatial averages, global energy dissipation, and other situation-dependent factors. It seems to me that in the context of the highly complicated IA double layer, this concept needs clarification.

The same basic physical phenomena, involving current interruption by particle trapping, seem to be important in both the strong Buneman DC and the weak ion-acoustic DC. The very different phenomenology between the two types indicates important differences in the underlying dynamics, however. The principal reason seems to be that in the Buneman cases of Ref [40], the inertia of the drifting ions causes them to be trapped later than

the electrons; the nonlinear phase of the evolution is controlled first by the electrons, while the intermediate and final stages are dominated by the ions. In the ion-acoustic regime, the ion inertia dominates the dynamics throughout.

It is also quite interesting to note with Sato and Okuda [28,32] that periodic simulations of short systems have generally been used to study anomalous resistivity from ion-acoustic instability, and that the IA double layers appear in the same simulation codes when the system is made longer. This implies that the essential state of ion-acoustic turbulence may be one involving recurring DL, and that it is necessary to consider its description in terms of statistics of inhomogeneous turbulence. A major question involves the conditions for development of the ion hole state.

It is regrettably outside the scope of this review to comment in detail on comparisons between simulations and the many experiments that have been and are continuing to be performed. Some comparisons have already been made [14]. Much fruitful work could be done in this regard. As one example, there are striking similarities between the cyclic instability of the fixed-potential simulations and the fixed-potential experiment of Iizuka et al. [19]. Novel experiments, for example that of Stenzel et al. [20], would be interesting subjects for simulation.

Many fundamental questions could also be addressed by further exploitation of circuit-model simulations. The circuit model has a particular advantage in being globally energy-conserving, so that energetic constraints appear naturally.

There has as yet been no simulation truly applicable to double layers in space. The principal reason concerns the plasma reservoirs, which may be assumed to be unaffected by the DL in present simulations in which current closure is implicitly at the system boundaries. In space, however, the true plasma reservoirs are far from the immediate vicinity of the DL, and so the distributions of the injection sources must be modelled in such a way that they depend on the evolution of the DL itself.

In addition, much attention should be paid to aspects of double layers which have observable consequences in space; some work of this sort has begun [42]. The magnetosphere-ionosphere coupling model of Goertz and Boswell [36] would also be an interesting subject for simulation.

Finally, of course, simulations may eventually be exploited to study some of the fundamental paradigms which seem to be important for double layers, such as the development of ion-acoustic cavitons and ion holes.

ACKNOWLEDGMENTS

I am grateful to G. Joyce for reading and commenting on the manuscript. In addition, I thank him, R.F. Hubbard, C.K. Goertz, N. Singh, and Yu. Sigov for useful conversations regarding their work. This work was supported by and performed at the Naval Research Laboratory, under Contract No. N0014-81-C-2038.

REFERENCES

1. Mozer, F.S., Geophys. Res. Lett. 8, 823 (1981); Temerin, M., Boehm, M.H., and Mozer, F.S., Geophys. Res. Lett. 8, 799; Temerin, M., Cattell, C., Lysak, R., Hudson, M., Torbert, R.B., Mozer, F.S., Sharp, R.D. and Kintner, P.M., J. Geophys. Res. 86, 11278 (1981).
2. Smith, R.A. and Goertz, C.K., J. Geophys. Res. 83, 2617 (1978).
3. Lyons, L.R., J. Geophys. Res. 85, 17 (1980).
4. Rowland, B.L., Palmadesso, P.J., and Papadopoulos, K., Geophys. Res. Lett. 8, 1257 (1981).
5. Temerin, M., Cerny, K., Lotko, W. and Mozer, F.S., Phys. Rev. Lett. 48, 1175 (1982).
6. Bernstein, I.B., Greene, J.M. and Kruskal, M.D., Phys. Rev. 108, 546 (1957).
7. Block, L.P., Cosmic Electrodyn. 3, 349 (1972).
8. Levine, J.S. and Crawford, F.W., J. Plasma Phys. 23, 223 (1980).
9. Smith, A., J. Plasma Phys. 4, 511 (1970a); 549 (1970b).
10. Knorr, G. and Goertz, C.K., Astrophys. Space Sci. 31, 209 (1974).
11. Swift, D.W., J. Geophys. Res. 84, 6427 (1979).
12. Torven, S. and Babić, M., Proc. 12th Int. Conf. on Phenomena in Ionized Gases (Ed. J.G.A. Houlscher and D.C. Schram, American Elsevier Publ. Comp., New York, 1975), p. 124; Torven, S. and Lindberg, L., J. Phys. D.: Appl. Phys. 13, 2285 (1980).
13. Coakley, P. and Herskowitz, N., Phys. Rev. Lett. 40, 230 (1978); Phys. Fluids 22, 1171 (1979).

14. Baker, K.D., Singh, N., Block, L.P., Kist, R., Kampa, W., and Thiemann, H., J. Plasma Phys. 26, 1 (1981).
15. Quon, B.H. and Wong, A.Y., Phys. Rev. Lett. 37, 1393 (1976).
16. Leung, P., Wong, A.Y. and Quon, B.H., Phys. Fluids 23, 992 (1980).
17. Iizuka, S., Saeki, K., Sato, N. and Hatta, Y., Phys. Rev. Lett. 43, 1404 (1979).
18. Saeki, K., Iizuka, S. and Sato, N., Phys. Rev. Lett. 45, 1853 (1980).
19. Iizuka, S., Michelsen, P., Rasmussen, J.J. and Schrittwieser, R., Phys. Rev. Lett. 48, 145 (1982).
20. Stenzel, R.L., Ooyama, M. and Nakamura, Y., Phys. Fluids 24, 708 (1981).
21. Stenzel, R.L., Gekelman, W. and Wild, N., UCLA Preprint PPG 610 1982).
22. For a lucid introduction to numerical simulation, see Potter, D., "Computational Physics", Wiley, New York (1973).
23. Torven, S., in "Wave Instabilities in Space Plasmas" (Ed. P.J. Palmadesso and K. Papadopoulos, Reidel, Dordrecht, 1979), p. 109.
24. Block, L.P., Astrophys. Space Sci. 55, 59 (1978).
25. Goertz, C.K., Rev. Geophys. Space Phys. 17, 418 (1979).
26. Carlqvist, P., in "Wave Instabilities in Space Plasmas" (Ed. P.J. Palmadesso and K. Papadopoulos, Reidel, Dordrecht, 1979), p. 83.
27. Goertz, C.K. and Borovsky, J.E., to be publ. (Proc. 1982 Nobel Symposium, Kiruna).
28. Sato, T. and Okuda, H., J. Geophys. Res. 86, 3357 (1981).
29. Singh, N., to be published (Plasma Phys.).
30. Bohm, D., in "The Characteristics of Electrical Discharges in Magnetic Fields" (Ed. A. Guthrie and R.K. Wakerling, McGraw-Hill, New York, 1949), p. 77.

31. Buneman, C., Phys. Rev. Lett. 1, 8 (1958); Phys. Rev. 115, 503 (1959).
32. Sato, T. and Okuda, H., Phys. Rev. Lett. 44, 740 (1980).
33. Singh, N., Plasma Phys. 22, 1 (1980).
34. Joyce, G. and Hubbard, R.F., J. Plasma Phys. 20 391 (1978).
35. Hubbard, R.F. and Joyce, G., J. Geophys. Res. 84, 4297 (1979).
36. Goertz, C.K. and Boswell, R.W., J. Geophys. Res. 84, 7239 (1979).
37. Wagner, J.S., Tajima, T., Kan, J.R., Leboeuf, J.N., Akasofu, S.-I. and Dawson, J.M., Phys. Rev. Lett. 45, 803 (1980); Wagner, J.S., Kan, J.R., Akasofu, S.-I., Tajima, T., Leboeuf, J.N. and Dawson, J.M., in "Physics of Auroral Arc Formation" (Ed. S.-I. Akasofu and J.R. Kan, American Geophysical Union, Washington, 1981), p. 304.
38. Borovsky, J.E. and Joyce, G., U. Iowa Prep. 81-31 (1981).
39. Borovsky, J.E. and Joyce, G., U. Iowa Prep. 81-34 (1981).
40. Smith, R.A., Physica Scripta 25, 413 (1982).
41. Singh, N. and Thiemann, H., Geophys. Res. Lett. 7, 737 (1980); Phys. Lett. 76A, 383 (1980).
42. Singh, N. and Schunk, R.W., to be published (J. Geophys. Res.).
43. Johnson, L.E., J. Plasma Phys. 23, 433 (1980).
44. Miura, R.M., in "Solitons in Action" (Ed. K. Lonngren and A. Scott, Academic Press, New York, 1978), p. 1.
45. Berk, H.L. and Roberts, K.V., Phys. Fluids 10, 1595 (1967); Berk, H.L., Nielsen, C.E. and Roberts, K.V., Phys. Fluids 13, 980 (1970).
46. Hasegawa, A. and Sato, T., Phys. Fluids 25, 632 (1982).
47. Schamel, H., Physica Scripta 20, 336 (1979); Schamel, H. and Bujarbarua, S., Phys. Fluids 23, 2498 (1980); Bujarbarua, S. and Schamel, H., J. Plasma Phys. 25, 515 (1981); Schamel, H., Phys. Rev. Lett. 48, 481 (1982).

48. Saeki, K., Michelsen, P., Pécseli, H.L. and Rasmussen, J.J., Phys. Rev. Lett. 42, 501 (1979); Lynov, J.P., Michelsen, P., Pécseli, H.L., Rasmussen, J.J., Saeki, K. and Turikov, V.A., Physica Scripta 20, 328 (1979).
49. Bloomberg, H.W., Astron. Astrophys. 30, 59 (1974).
50. Wahlberg, C., J. Plasma Phys. 22, 303 (1979).
51. Carlqvist, P., Cosmic Electrodyn. 3, 377 (1972).
52. Raadu, M. and Carlqvist, P., Astrophys. Space Sci. 74, 189 (1980).
53. Belova, N.G., Galeev, A.A., Sagdeev, R.Z., and Sigov, Yu.S., JETP Lett. 31, 518 (1980).
54. Singh, N., Phys. Lett. 75A, 69 (1979).
55. Bergeron, K.D. and Wright, T.P., Phys. Fluids 21, 1578 (1978).
56. Goertz, C.K. and Joyce, G., Astrophys. Space Sci. 32, 165 (1975).
57. Kindel, J.M., Barnes, C. and Forslund, D.W., in "Physics of Auroral Arc Formation" (Ed. S.-I. Akasofu and J.R. Kan, American Geophysical Union, Washington, 1981), p. 296.
58. Hudson, M.K. and Potter, D.W., in "Physics of Auroral Arc Formation", op. cit., p. 260.
59. DeGroot, J.S., Barnes, C., Walstead, A.E. and Buneman, O., Phys. Rev. Lett. 38, 1283 (1972).
60. Lawson, J.D., "The Physics of Charged Particle Beams", Oxford Univ. Press (1977).

Note added in proof

At the 1982 International Conference on Plasma Physics (Göteborg), results of a new simulation, by G. Chanteur, J.C. Adam, and R. Pellat, were presented. This simulation is 1-D IA with periodic boundary conditions on the potential. The plasma is initialized with a density depression, similar to Ref. [59]. A potential depression forms and propagates in the sense of electron drift at about $0.8 C_s$ and decelerates when ion trapping occurs. The authors interpret the initial growth of the potential as being due to electron reflection by the potential well.

FIGURE CAPTIONS

1. Schematic of the classical double layer concept: (a) the potential profile; (b) the ion phase space; (c) the electron phase space. After Smith and Goertz [2]).
2. Schematic of the potential when the Bohm criterion is violated.
3. Possible models of circuits containing double layers. The double layer is in the simulation region between the dashed lines, which represent grids connecting the plasma to lumped external elements. External plasma sources (see text) are symbolized by the arrows: (a) fixed potential source; (b) fixed current source; (c) potential source with external resistance and inductance. In all the diagrams, the double layer may be pictured as a nonlinear, leaky capacitor between the grids.
4. Electric potential $\phi(x)$ and corresponding pseudopotential $V(\phi)$ in various cases.
5. Potential and phase spaces of an ion hole.
6. Initial propagation of a pulse driven by fixed potential boundary conditions. Top left: the potential profile. Top right: Electron distribution function at the high-potential boundary. Bottom left:

electron phase space. Bottom right: ion phase space. (After Hubbard and Joyce [35]).

7. Same as Fig. 6, but at later time. The double layer has evolved, with the potential modulated by downstream waves at the high-potential end. Note wave trapping of the incoming electron flux. (After Hubbard and Joyce [35]).
8. Time histories of the potential and current density (a) and dissipated current (b) in a double layer in the inductive circuit of Fig. 3(c). The points A,B,C of an inductive overshoot are at times $\omega_e t = 85, 100, 125$ (After Smith [40]).
9. The potential, electric field, and phase spaces corresponding to points A,B,C of Figure 8. (After Smith [40]).
10. Recurrent explosive development of the double-layer potential in a Buneman-regime simulation. (After Belova et al [53]).
11. Time sequence of the potential in a floating-potential ion-acoustic double layer (After Sato and Okuda [28]).
12. Same run as in Fig. 11, at $t = 960 \omega_e^{-1}$: (a) potential; (b) electron density; (c) ion density. Note the sharp density spike at the foot of the potential rise.

Table 1 - Double Layer Simulation Characteristics

Ref.	Dim	Type	System length (λ_e)	Boundary Flow Regime	Boundary Condition on Potential
34,35	1	PIC	30-360	N/A	Fixed
29,33	1	V	50-100	Buneman	Fixed
43	1	V	50	Buneman	Fixed
56	1	PIC	37-100	Buneman	Floating
40	1	V	20-57	Buneman	Floating
53	1	PIC	100	Buneman	Floating
32	1	PIC	512-1024	Ion-acoustic	Periodic
28	1	PIC	1024-4096	Ion-acoustic	Periodic, Floating
57	1,2	PIC	1024 (1-D) 128 x ? (2-D)	Ion-acoustic	Periodic, Fixed
58	1	PIC	1024	Ion-acoustic	Periodic
59	2	PIC	141 x 141	Buneman	Periodic
37	2	PIC	128 x 32	N/A	Floating
38,39	1,2	PIC	32x60 (1-D) <i>planar</i> 64 x 40 (2-D)	N/A	Fixed

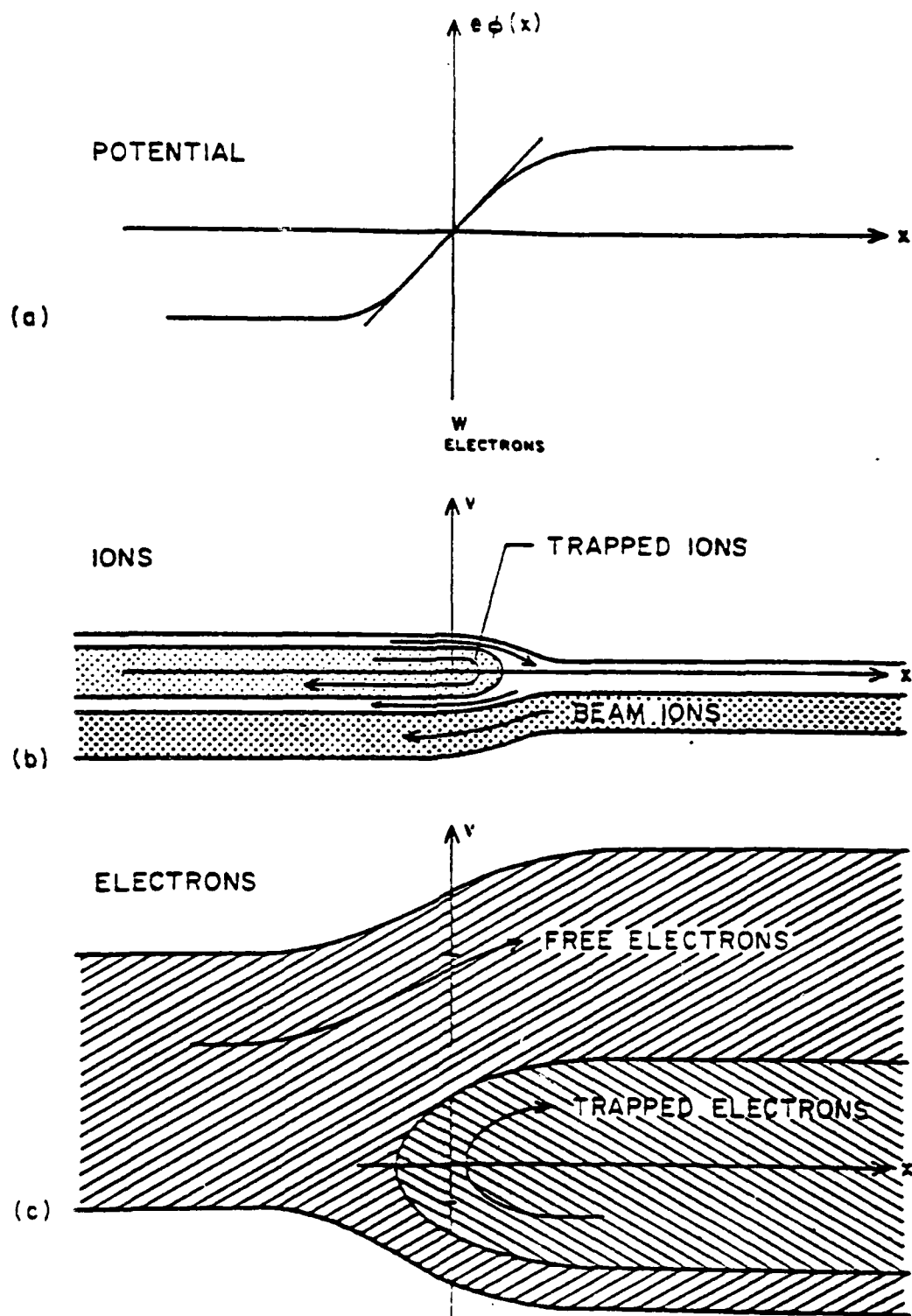


Fig. 1

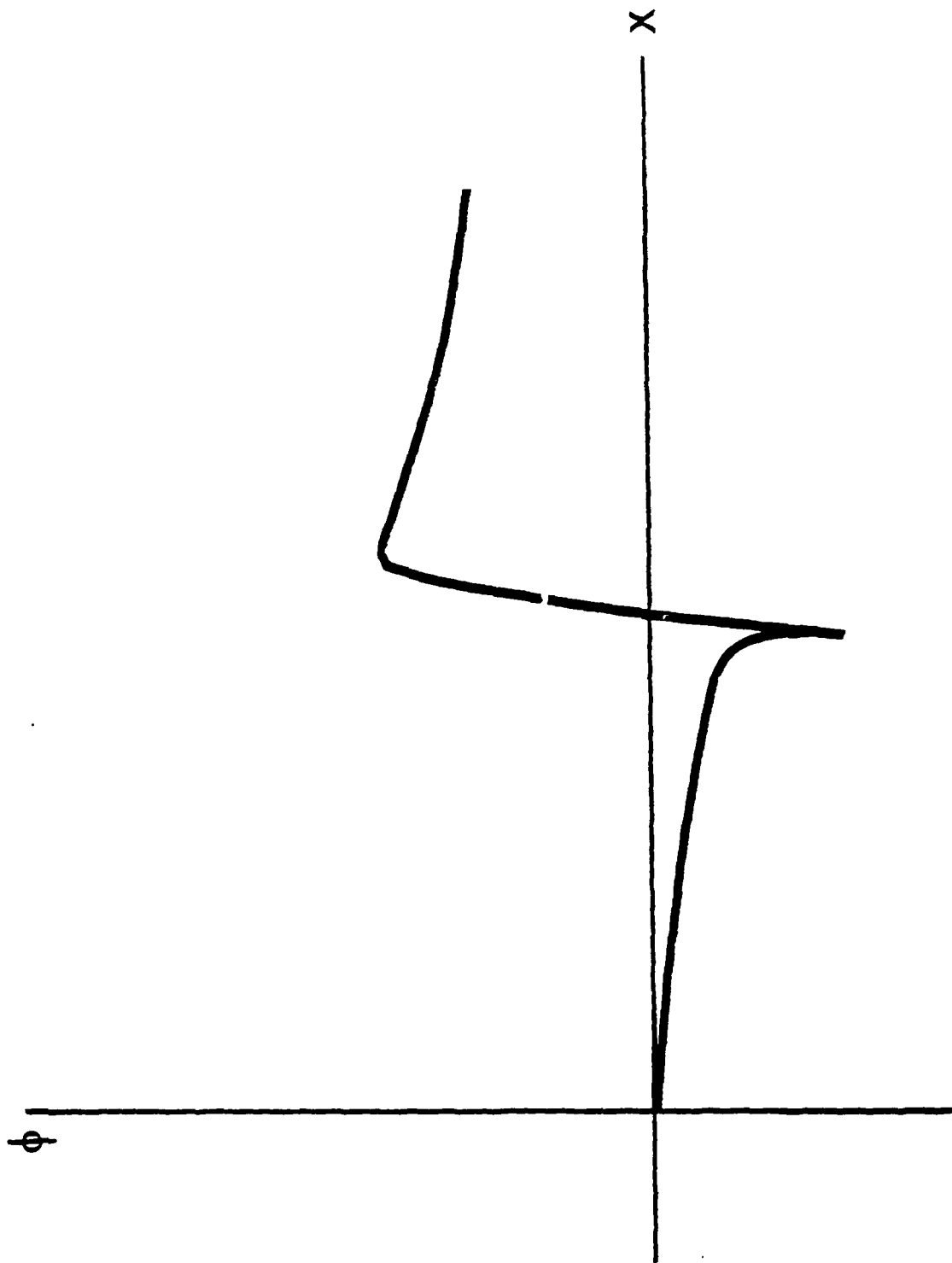
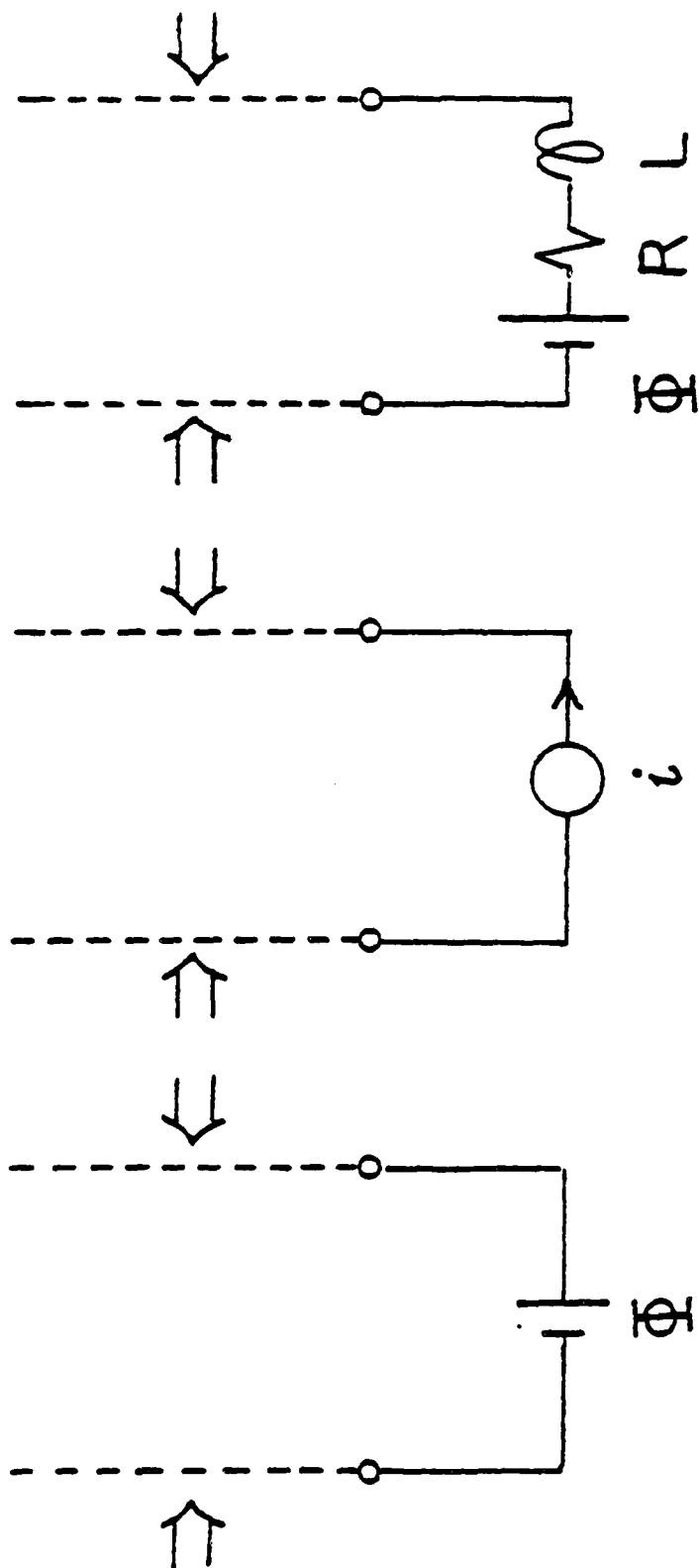


FIG. 2



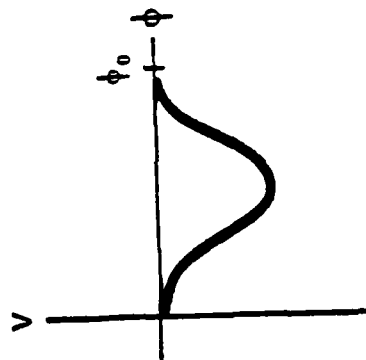
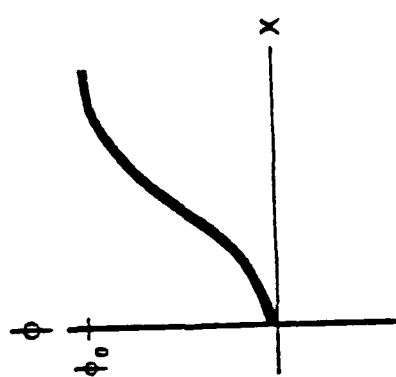
(C)

(B)

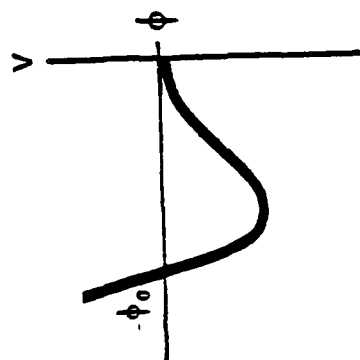
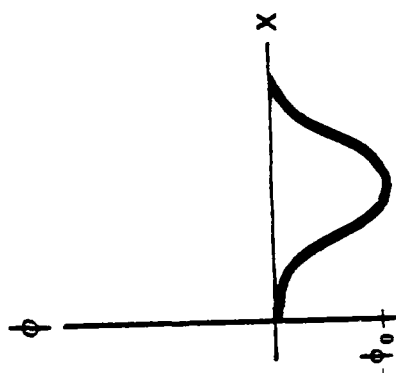
(A)

FIG. 3

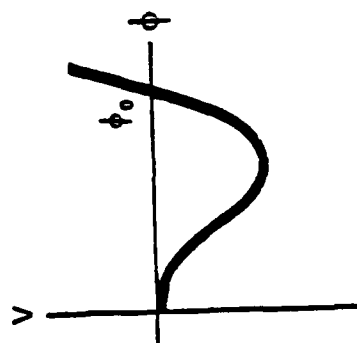
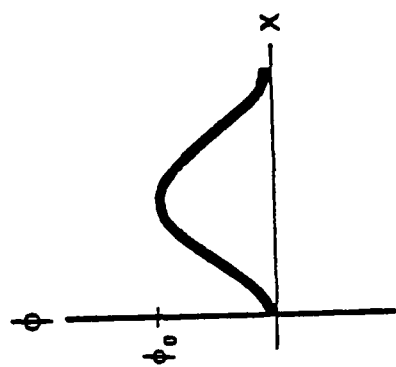
DOUBLE
LAYER



ION
HOLE



ELECTRON
HOLE



SINUSOID

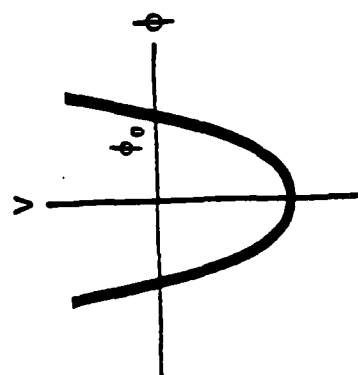
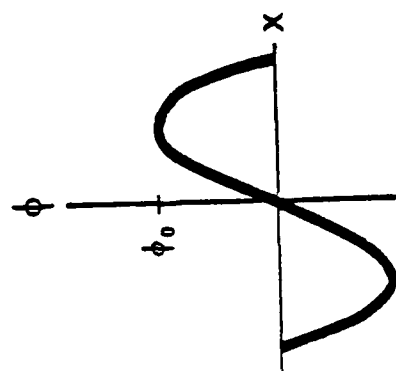


FIG. 4

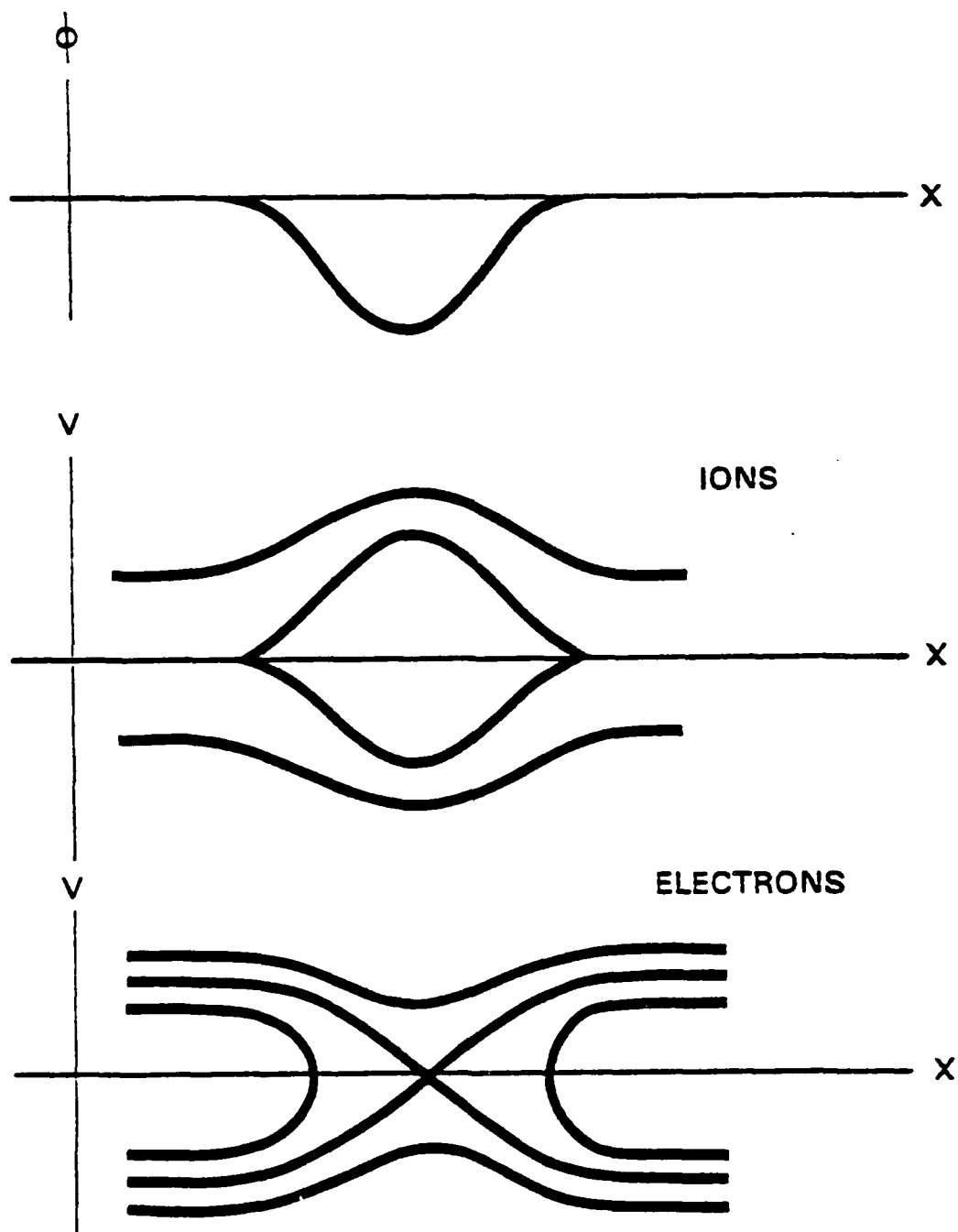


Fig. 5

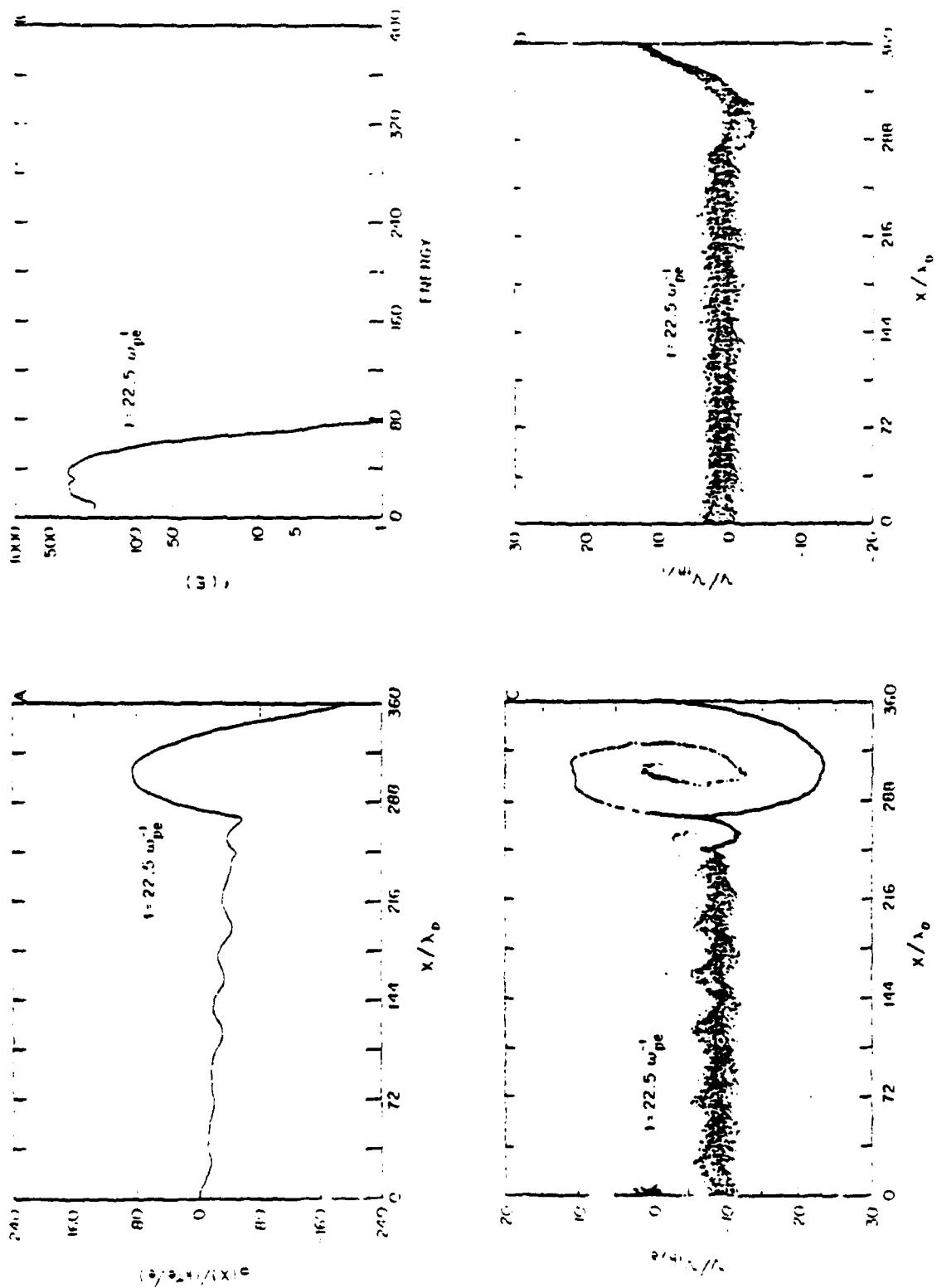


FIG. 6

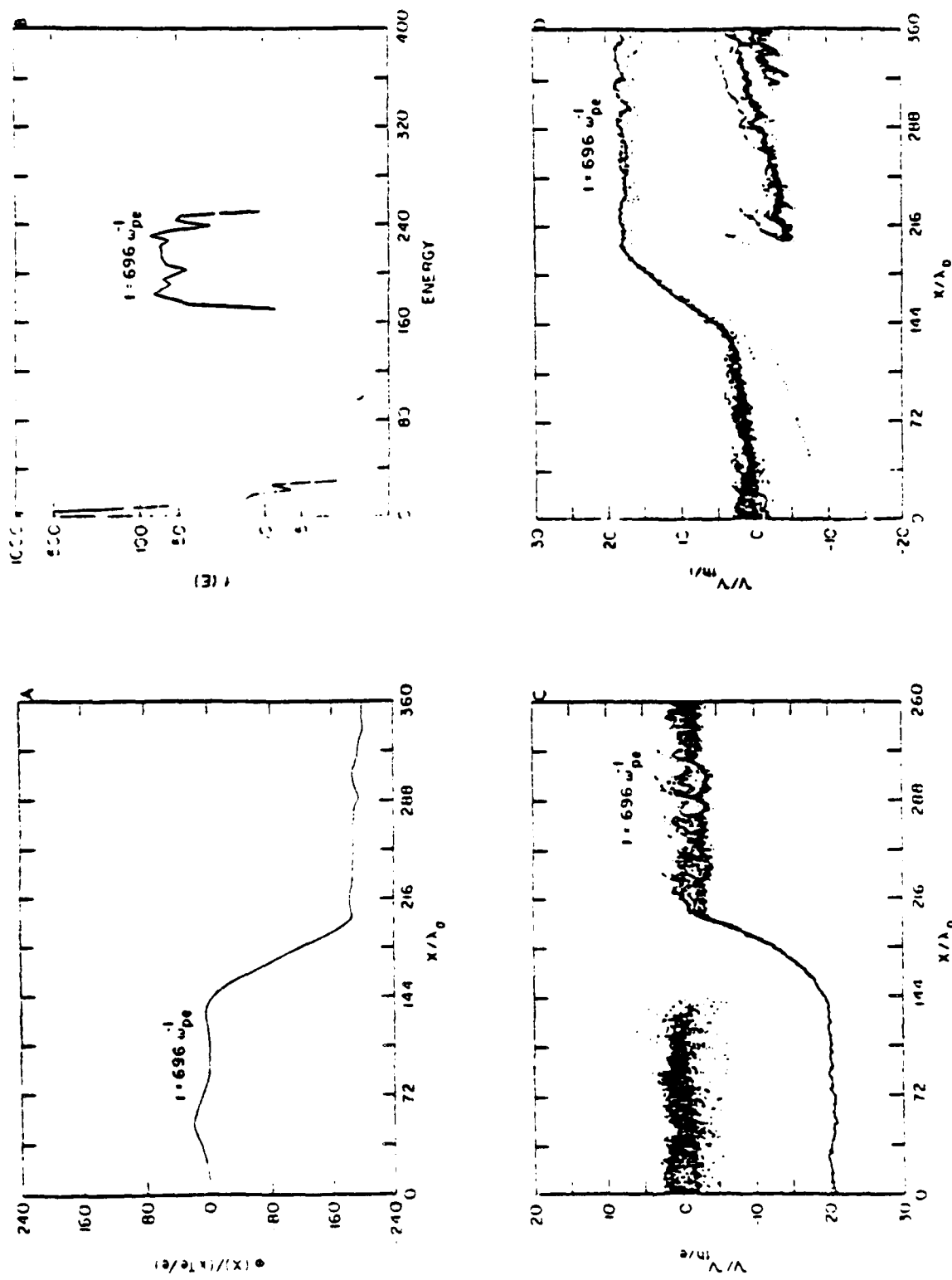


Fig. 7

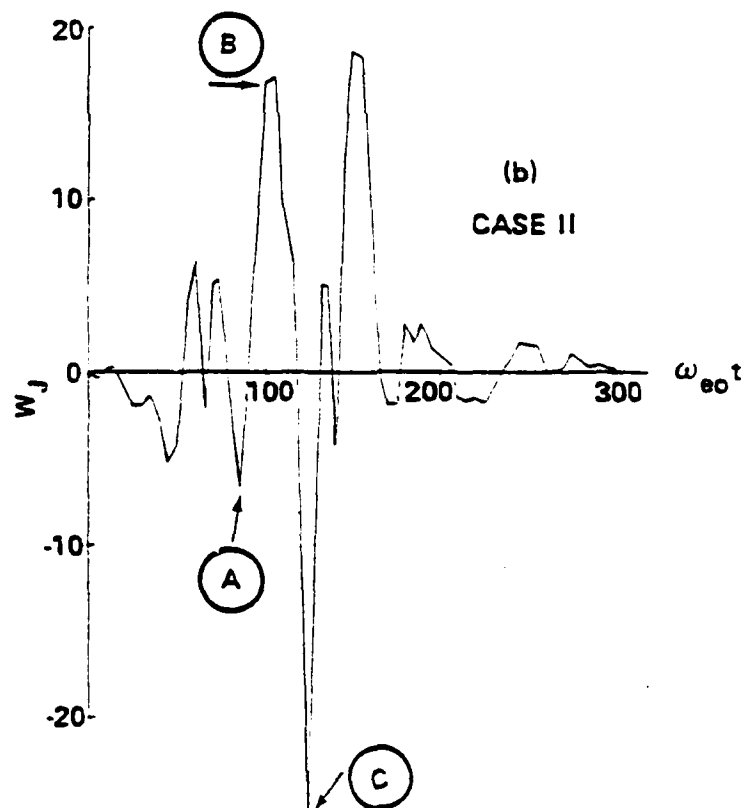
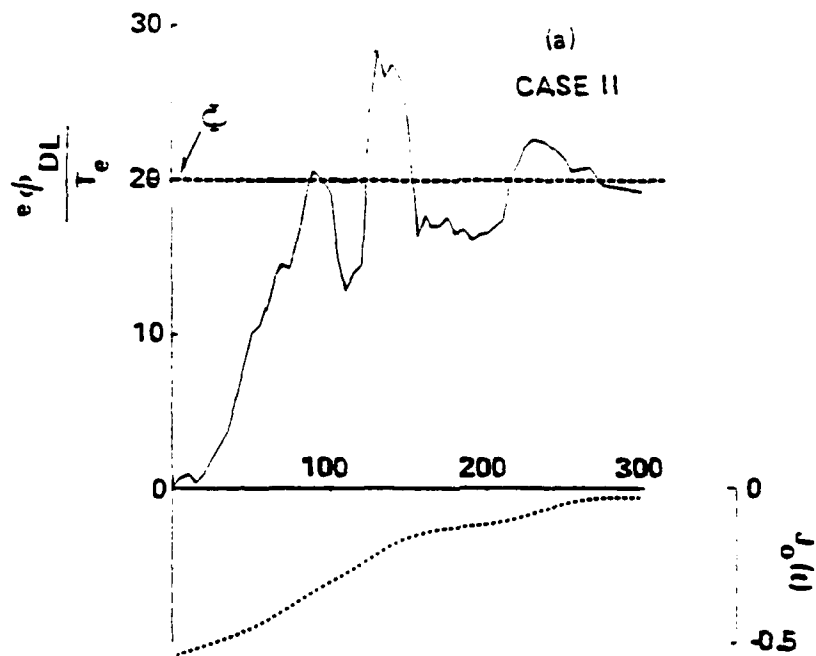


Fig. 8

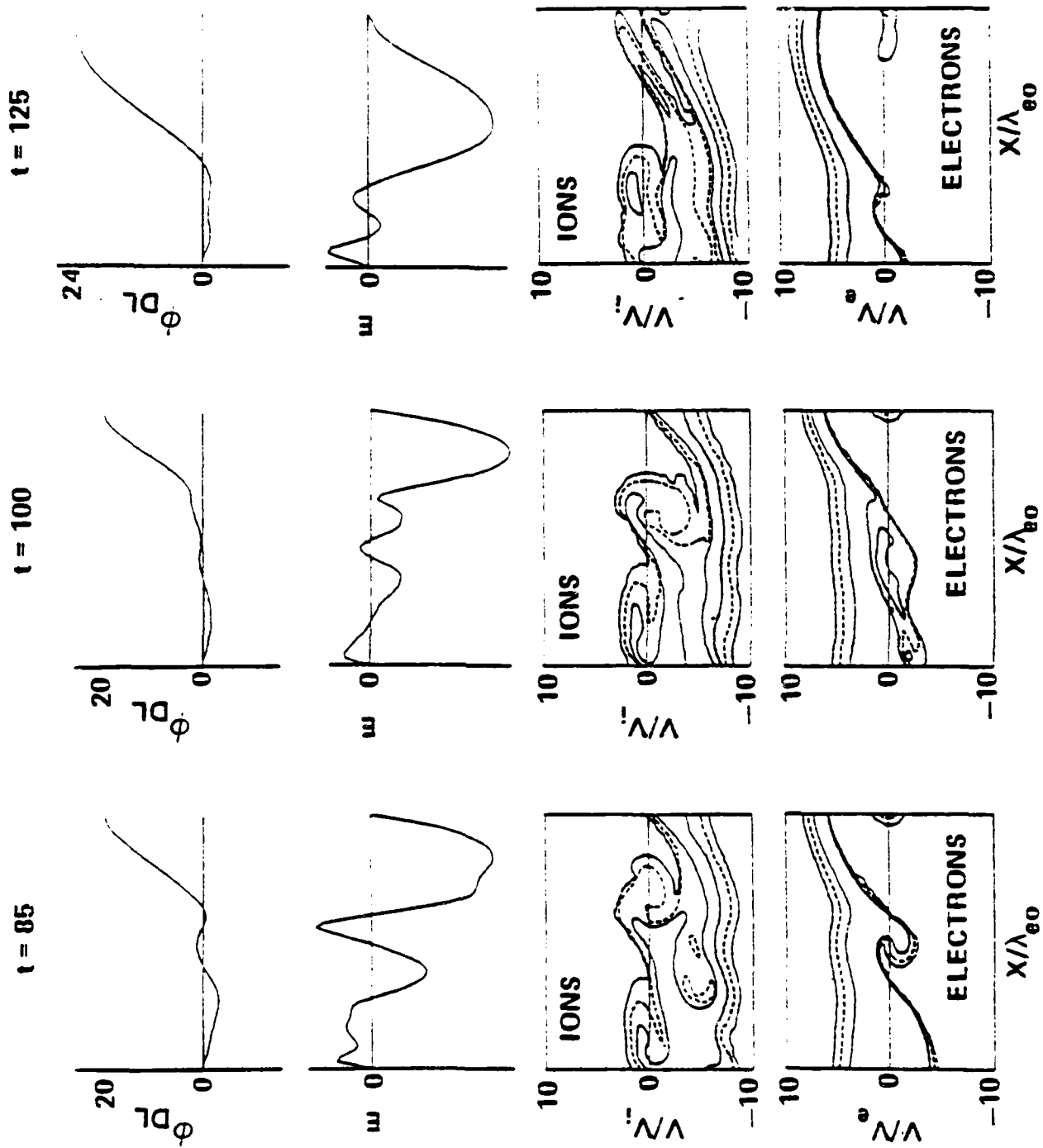


Fig. 9

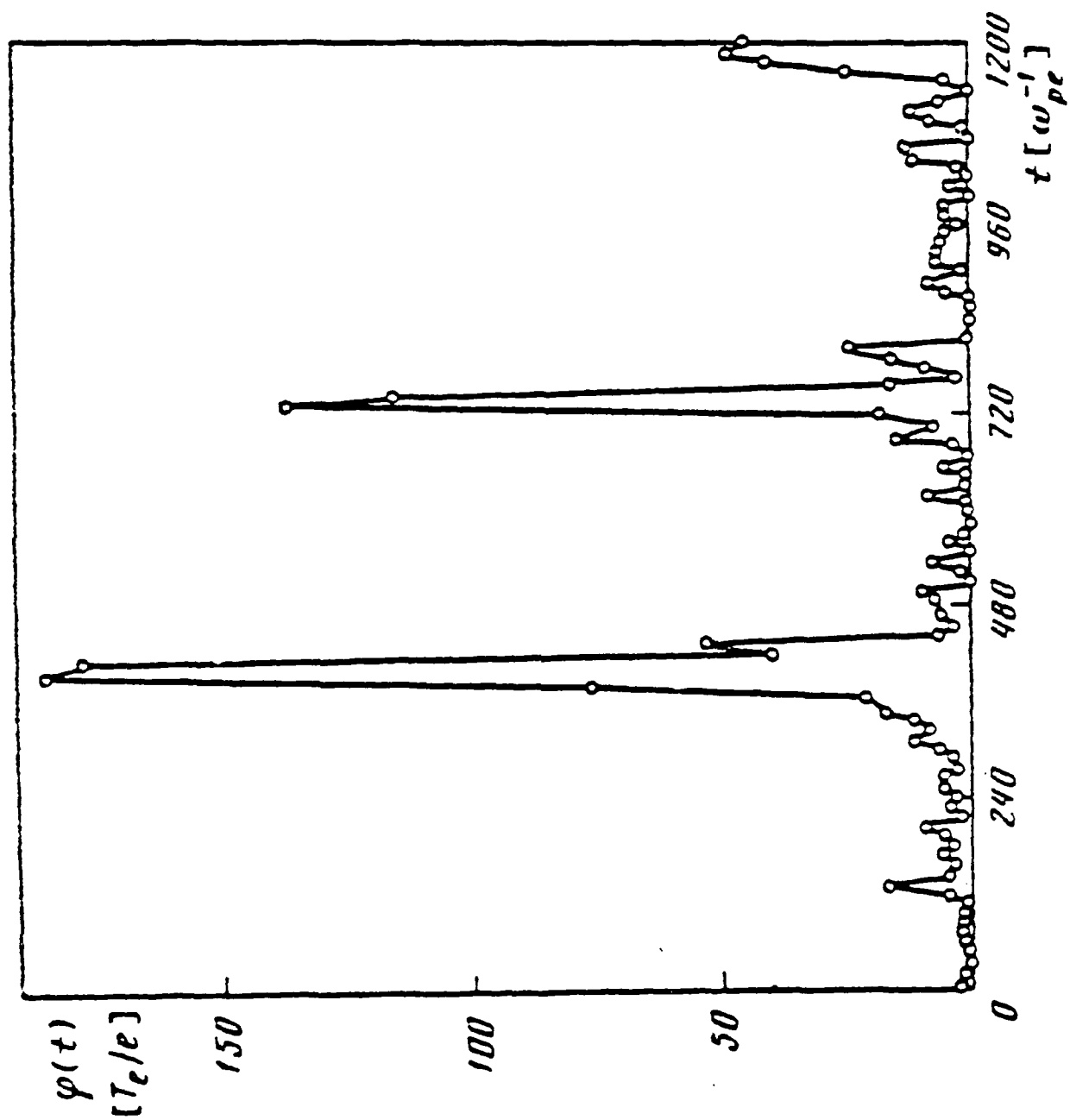


Fig. 10

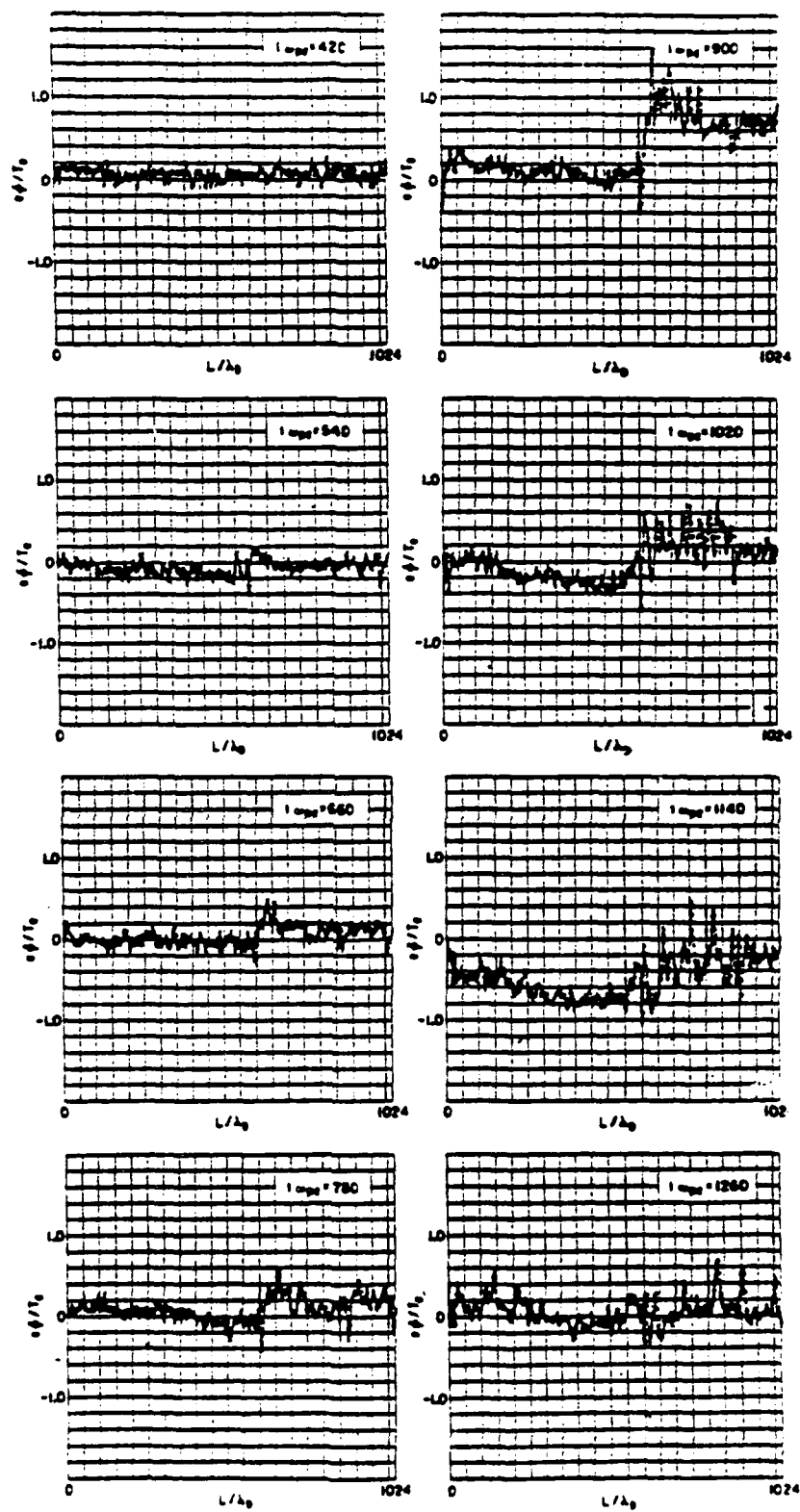


Fig. 11

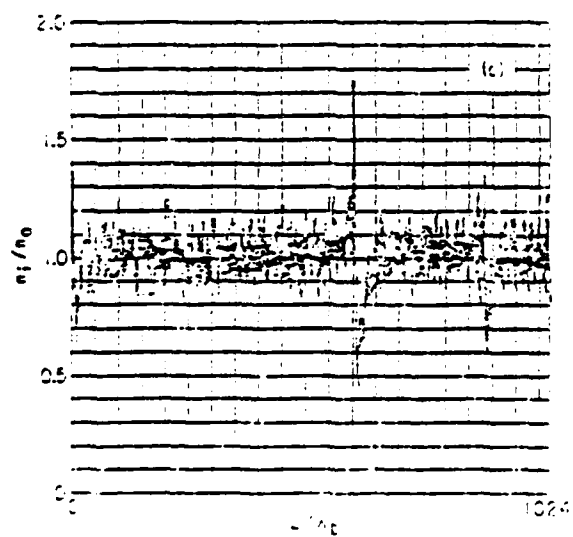
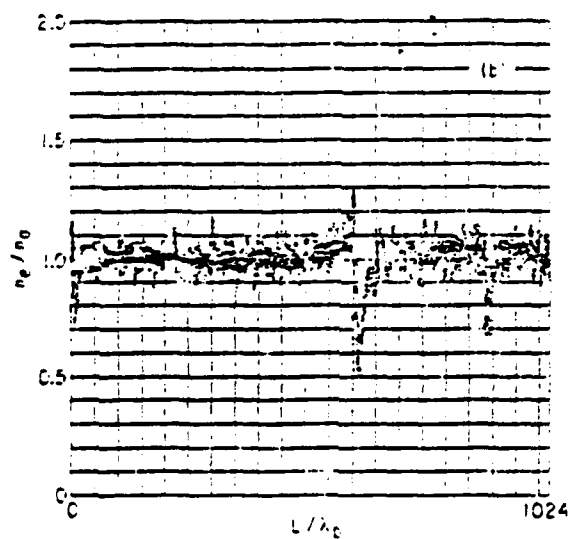
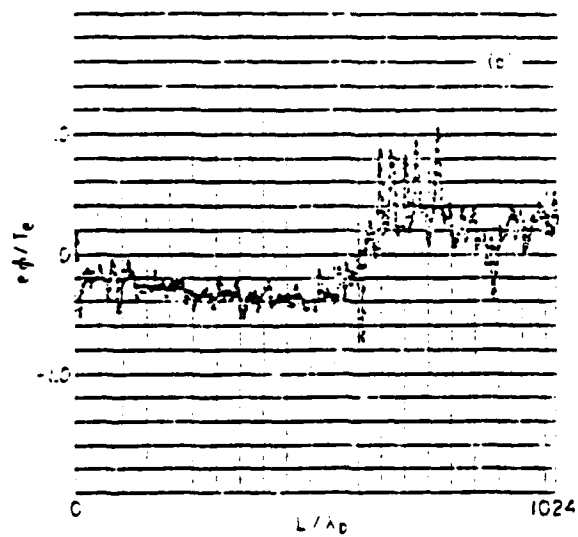


Fig. 12

K-73

APPENDIX L

STRUCTURING STUDIES FOR A URANIUM CLOUD

STRUCTURING STUDIES
FOR A URANIUM CLOUD

Presented at the DNA Sponsored
Uranium (Oxide) LWIR Meeting
11-12 January 1983
Los Alamos National Laboratory

NRL

CONTRIBUTORS:

Ellis Hyman*
Margaret Mulbrandon
Norman Zabusky**
Sidney Ossakow

*Science Applications, Inc., McLean, Virginia

**Fluid Sciences, Inc., Pittsburgh, Pennsylvania

Viewgraph 1

Structuring Studies for a Uranium Cloud

The following description provides some elaboration on the material illustrated in the viewgraphs. Viewgraphs 1 and 2 are self-explanatory.

Viewgraphs 3, 4, and 5 are extracted from "Structure-Models: Definition and Comparison", by R.D. Hake, Jr., and J.E. Casper, Tech. Memo #1, January 1982, SRI International. They are that document's Figures 4, 5, and 6. These viewgraphs are shown to exhibit recent studies that have been carried out with motivation similar to ours, though with a more simplistic model for striations than is predicted by physical theory. Viewgraph 3 shows the specific model realization involving 'lucite rod' striations of various sizes as viewed perpendicular to the magnetic field. By a lucite rod is meant a cylinder of constant emission/area inside the rod, zero outside and a sharp edge at the surface. All rods are assumed to have this same emission density. Viewgraph 4 shows the surface radiance that would be observed by a sensor at $z = -\infty$, directed along a line of constant x , (i.e., parallel to z), as the sensor is moved in x over the central 30 km region. Viewgraph 5 shows the power spectral density (PSD) of viewgraph 4, which has an asymptotic dependence $\sim k^{-3}$. Now, this dependence is a direct result of the model - uniform density cylinders with infinitely sharp edges.

In viewgraph 6 we illustrate the development in time of striations in a uranium cloud as suggested both by theory and observation. The development of fingering and the pinching off of little blobs can be seen. Neither the fingers nor the blobs have cylindrical symmetry. They are characterized by very steep gradients on the 'backside' of the cloud and more gradual gradients elsewhere. Precisely how steep the gradients are is a subject of current research at the Naval Research Laboratory (NRL) and cannot now be satisfactorily predicted by theory. The steepness, in turn, influences the high k cutoff of the PSD, the so-called inner scale size, which is an important parameter for systems designers.

Viewgraphs 7 and 8 are self-explanatory. Viewgraph 9 differentiates between 2 kinds of power spectral calculations; those useful in communications studies, where we deal with transmission through a structured medium, and those useful in emission calculations, as here, where LWIR is emitted from all portions of the cloud. In the first case, the relevant PSD is obtained by

doing a 2 dimensional Fourier Transform, squaring the results, and summing over one or the other of the mode parameters. We have called this a 2D power spectrum. In the second case, we sum over one of the spatial parameters first, calculate the Fourier Transform of the other, and square the results. We call this a 1D power spectrum, or a scan power spectrum. In this latter case, the initial summing introduces a smoothing which often, but not always, results in a larger (numerical) value (smaller when you account for the minus sign) for the asymptotic index of a scan power spectrum.

The distinction between these two types of asymptotic dependences is illustrated in Viewgraph 10. Consider first, the first four columns. The top row shows four different idealized clouds in which the emission is a constant inside and zero outside, with an infinitely sharp boundary. All four will give rise to a k^{-2} dependence for the 2D PSD. Each will have a different asymptotic dependence in the scan PSD. The differences arise from the different scan radiance profiles (row 3). In the first column we observe a very sharp edge, and this gives a k^{-2} dependence for the scan PSD. The third column illustrates a linear profile which gives k^{-4} . By varying the curvature between the jump in column 1 and the linear in column 3 we can obtain any scan PSD index in the range $-2(2 + \epsilon)$, where $-1 < \epsilon < 0$. The curvature arising from a conic surface (column 2) corresponds to $\epsilon = -1/2$. Column 4 corresponds to a curvature more gradual than linear and gives an index $-2(2 + \epsilon)$, with $\epsilon > 0$.

Columns 5 and 6 illustrate idealized clouds with non-sharp boundaries. In the figure in column 5, row 1, the emission is maximum at the center and falls off linearly to zero at the edge. The 2D asymptotic index is -4 and the scan index is -5 . Column 6 shows a slightly more complex situation, the frustum of a cone, where the emission is constant in a central region, and then falls off linearly to zero at the edge. Here there will be two distinct asymptotic regions. For the 2D PSD there will be a k^{-2} falloff followed by a region where the falloff is k^{-4} . For the scan PSD, each region has an index one higher (numerically). That is, there is a k^{-3} region followed by a k^{-5} region.

In these studies we are concerned with LWIR emission and are therefore interested in the scan power spectral indices that may occur in a real situation. Viewgraph 11 illustrates the fact that even with complex clouds,

where fingering has occurred, if the finger edges are sufficiently sharp, a scan power spectral index at or near 2 will occur if we happen to be looking along an edge.

In viewgraph 12 we exhibit the analytic model that we have constructed to incorporate the essential results of NRL numerical simulations. We currently have limited our studies to a 2D model corresponding to a detector an infinite distance away, and looking along the magnetic field. Later we will generalize this to a 3D model, where parallax effects can be studied. The figure shows a model with 4 fingers but this can be varied from none on up. Region I on the left is the frontside and is characterized by an elliptical shape. Region II simulates the backside and has the cosine dependence shown. There is an intermediate region which blends the two smoothly, along a parabolic curve. The radiance from a particular point x, y is determined by the Heaviside function shown. The parameter Δ determines the sharpness of the cloud edge. It is taken to be a very small constant on the backside (sharp edge), but to increase with angle on the frontside, giving a smoothed edge in this region. In addition to what is illustrated in viewgraph 12, additional smoothing is incorporated in the model, by performing a weighted average of each mesh value with nearest neighbors. In preliminary studies the smoothing factor has been varied from a value of 2, corresponding to a high degree of smoothing, to 10, corresponding to very little smoothing. Also, additional flexibility can be obtained by applying the smoothing repetitively.

Starting with the above model we can obtain a radiance profile as would be obtained from a detector at an infinite distance. We can scan along x with a detector at $y = \pm \infty$ (X SCAN) or along y with a detector at $x = \pm \infty$ (Y SCAN) or, more generally, at some intermediate angle. Viewgraph 13 illustrates a typical radiance scan obtained, scanning along x . There is a fairly sharp rise in radiance on the right as we move into the finger region, and a more gradual fall on the left side. How much smoothing has been applied is not clearly evident in these plots due to lack of plotter resolution. But it shows up clearly in a PSD plot by the rate of falloff of power at high k values.

Viewgraph 14 shows a similar profile as we scan along y . Here the effect of the fingers is clearly evident. Viewgraph 15 shows a scan in which the detector is rotated 9° off of a true y scan. Distortion of the fingers from

looking off-angle is obvious. More important, at 9° we happen to be looking very nearly along edges of the fingers as was illustrated in Viewgraph 11.

In Viewgraph 16 we present some preliminary numerical results illustrating the variation in the power spectral index that results as a function of variation of the cloud smoothness and as a function of variation of off-angle axis of the detector. There are 3 groups of 3 conditions, each for both an (approximately) x power spectral index and an (approximately) y power spectral index. The first group of 3 fixes the angle at 0° but increases the smoothing from almost none to very smoothed. The y scan has, in all cases, a (numerically) large spectral index which increases with smoothing. The x scan shows a very low spectral index which increases somewhat with smoothing.

The second group of 3 again maintains a constant detector angle, here of 9° , and again varies the smoothing as in the first group. Now the y scan gives a very low spectral index when the cloud is sharp but increases appreciably as the cloud becomes smoother. The x scan is characterized by a relatively large and insensitive index. Finally, the third group shows the variation with angle at a fixed smoothness. Here the y scan varies from a large value at 0° to a relatively small value at 9° , even with a fairly smoothed model. The x scan varies in the opposite way from a very low index at 0° to a quite high value at 9° .

In the future our plans are to do some careful studies with a few carefully chosen cloud shapes, varying cloud smoothness and detector angle. Then, we plan to generalize to a 3D cloud to study off-angle (of the magnetic field) effects and parallax, varying the detector cloud separation. Some more detailed theoretical studies of the effect of specific curvature shapes on observed spectral indices are also in process.

Bibliography

1. Hake, R.D., and Casper, J.E., "Structure Models: Definition and Comparison", Tech. Mem. #1, January 1982, SRI International.
2. Hyman, E., Mulbrandon, M., Ossakow, S.L., and McDonald, B.E., "Preliminary Numerical Simulation of IR Structure Development in a Hypothetical Uranium Release", NRL Memo. Report 4659, (1981).
3. Sachs, D.L., Linson, L.M., Baxter, D.C., "Characteristics of Striation Structure and Effects on Scintillation", DNA4612F, (1978).
4. McDonald, B.E., Keskinen, M.J., Ossakow, S.L., and Zalesak, S.T., "Computer Simulation of Gradient Drift Instability Processes in Operation Avefria", J. Geophys. Res., 85, 2143 (1980).
5. Zabusky, N.J., and Block, J., "Estimating the Spectrum of Electron Density Fluctuations from Simulations of Ionospheric Plasma Clouds", NRL Memo. Report 3586, (1978).

Uranium Cloud Structuring Simulations

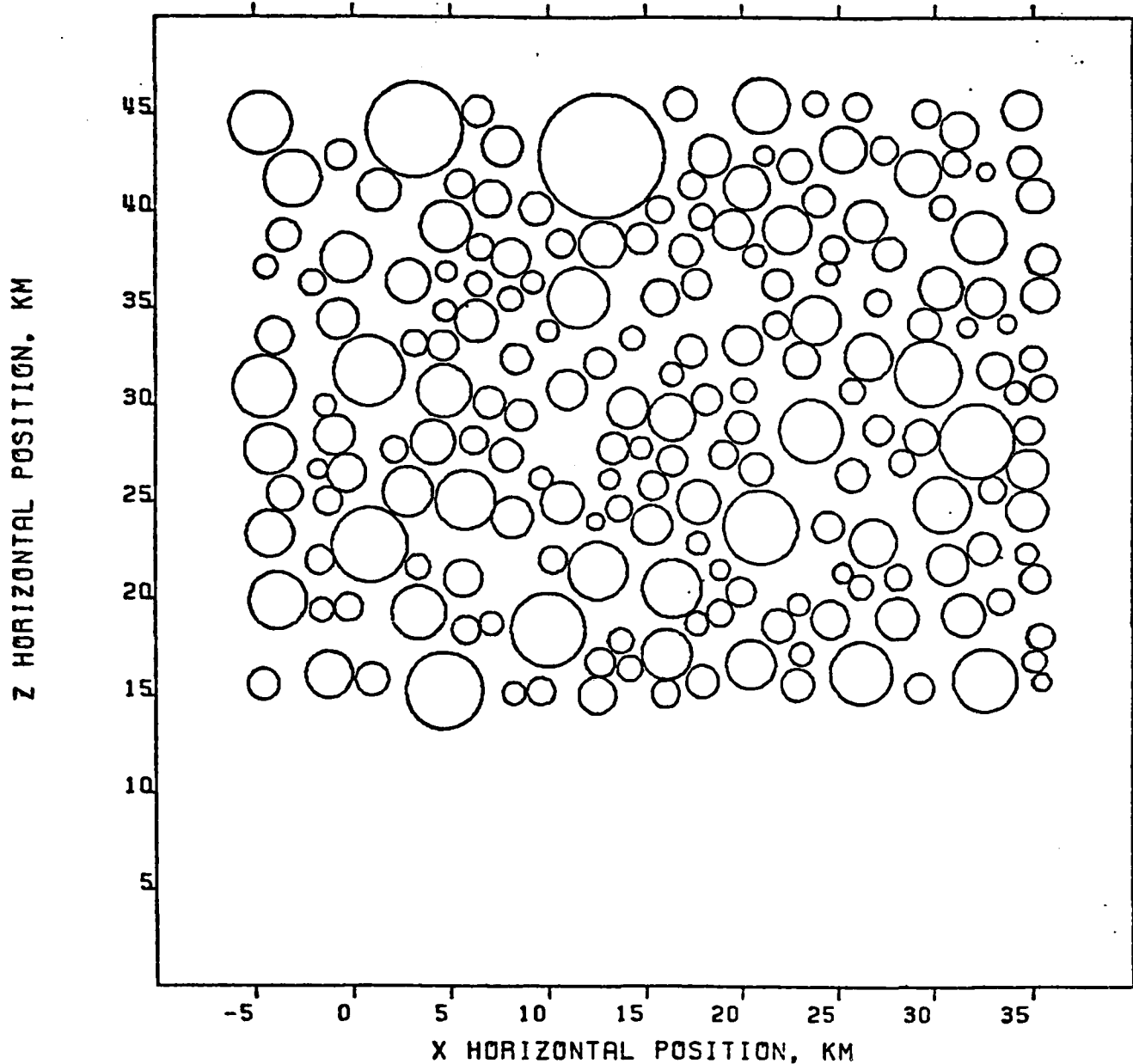
Goal: To predict the spatial power spectral characteristics of ir emission from an ionized uranium cloud.

Method: Construct a 3D analytic model which simulates the essential features of a striating cloud, based on the results of NRL physics simulations.

Features: The model will have very steep gradients on the backside with temporal growth of fingering, but gradual gradients on the frontside. Striations will not have cylindrical symmetry.

Viewgraph 2

STRIATION SIZE AND LOCATION



SEED= 17 200 STRIATIONS ACROSS 41X 31 KM
CIRCLE SHOWS LUCITE ROD STRIATION FILL FACTOR= .5027
DEC 21, 1981, 14:23:23.250

Figure 4

L-11

Viewgraph 3

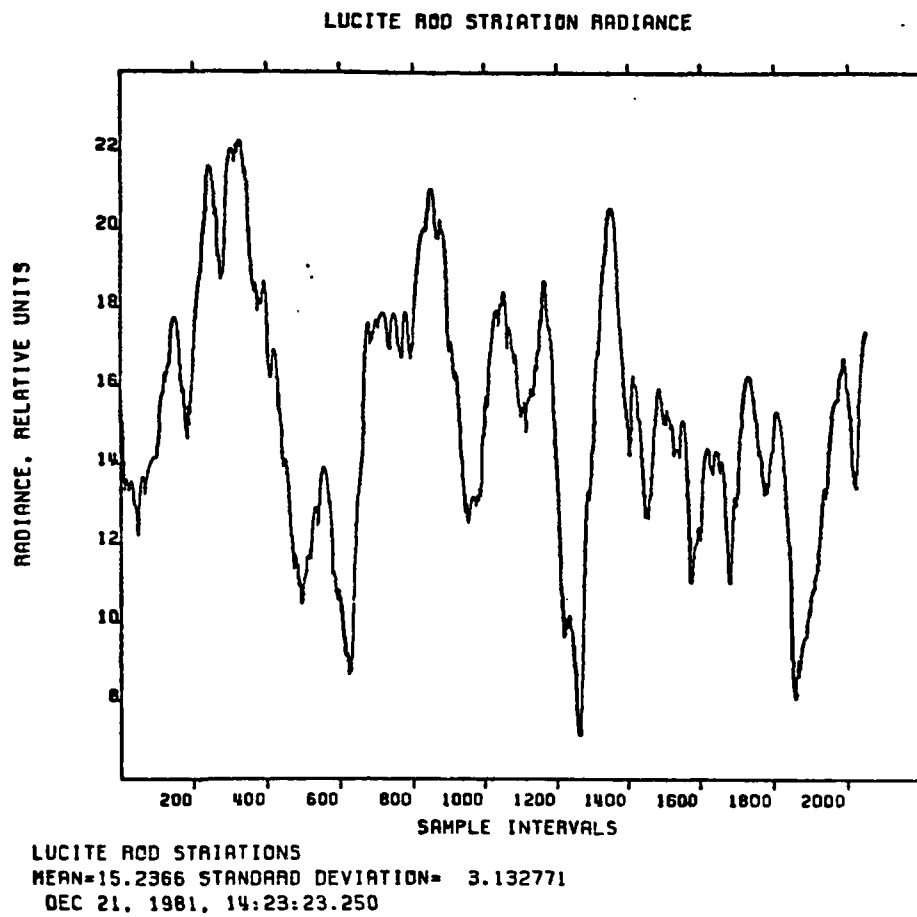
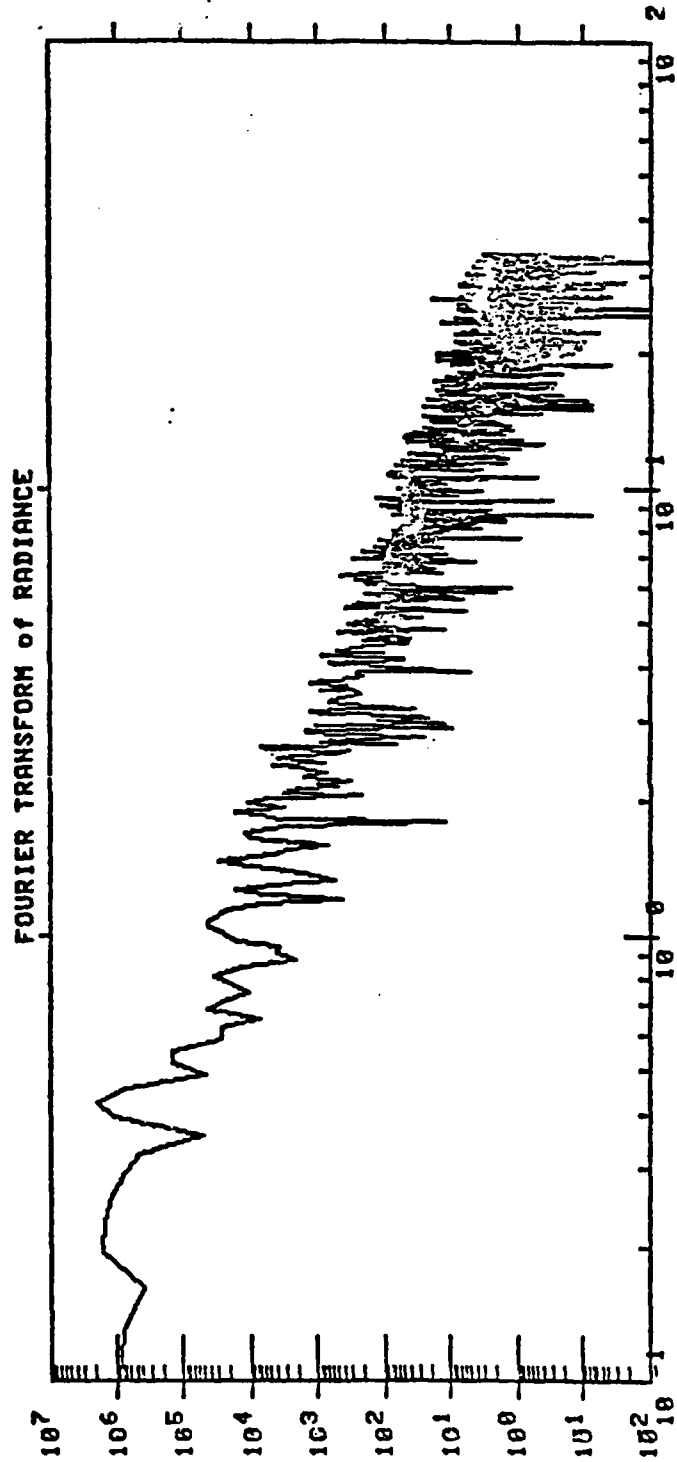


Figure 5

L-12

Viewgraph 4



SEED= 17 200 STRIATIONS ACROSS 41 km
 2848 SAMPLES ACROSS CENTER 31 km
 DEC 21, 1981, 14:23:23.250

RELATIVE POWER

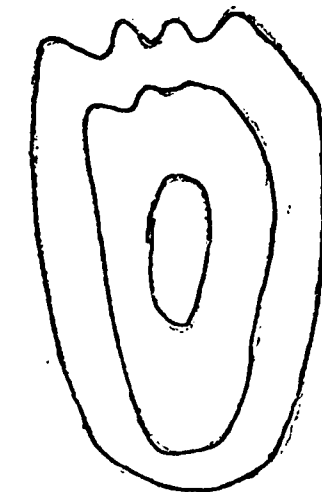
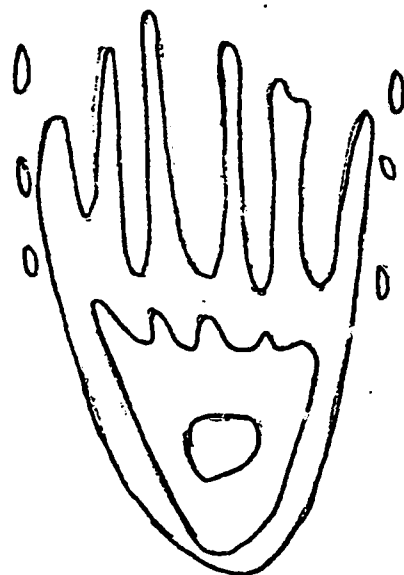
Figure 6

L-13

Viewgraph 5

STRIATION DEVELOPMENT

B



TIME



n

steep
backside

cross section

Why not a physics model for these studies?

- Our knowledge of the physics processes required to proceed to smaller scales is incomplete at this time.
- Increasing the resolution of the physics simulation to determine parameters such as the inner scale size will not work until we have incorporated the correct physics to simulate these small scale processes.
- Current NRL simulations do not incorporate a complete front side. Thus, view directions are limited.
- Increasing resolution significantly is not possible in any case because of computer core limitations and cost.

What advantages are there to an analytic model?

- Resolution can be greatly increased.
- When viewing off-angle, no interpolation between grid points is needed. (Interpolation results in a numerical, non-physical smoothing.)
- We can incorporate in the analytic model the properties that the physics model has established and, where the results of a physics calculation are not yet determined, we can introduce parameters in the model.
- By performing parameter studies we can establish which parameters and parameter ranges are important, so that future physics simulations can address these problem areas.

Power Spectral Calculations

Starting with a 3D cloud with volume emission at $x, y, z = U(x, y, z)$, calculate the radiance $R(x, y)$ (where for simplicity of discussion we have assumed we are viewing along z)

We define 2D power spectra, familiar from communications studies and 1D power spectra (scans) useful here

$\mathcal{F}$ means Fourier Transform

$$\underline{2D} \quad P(k_x) = \sum_{k_y} \left\{ \mathcal{F}_{x,y} [R(x, y)] \right\}^2$$

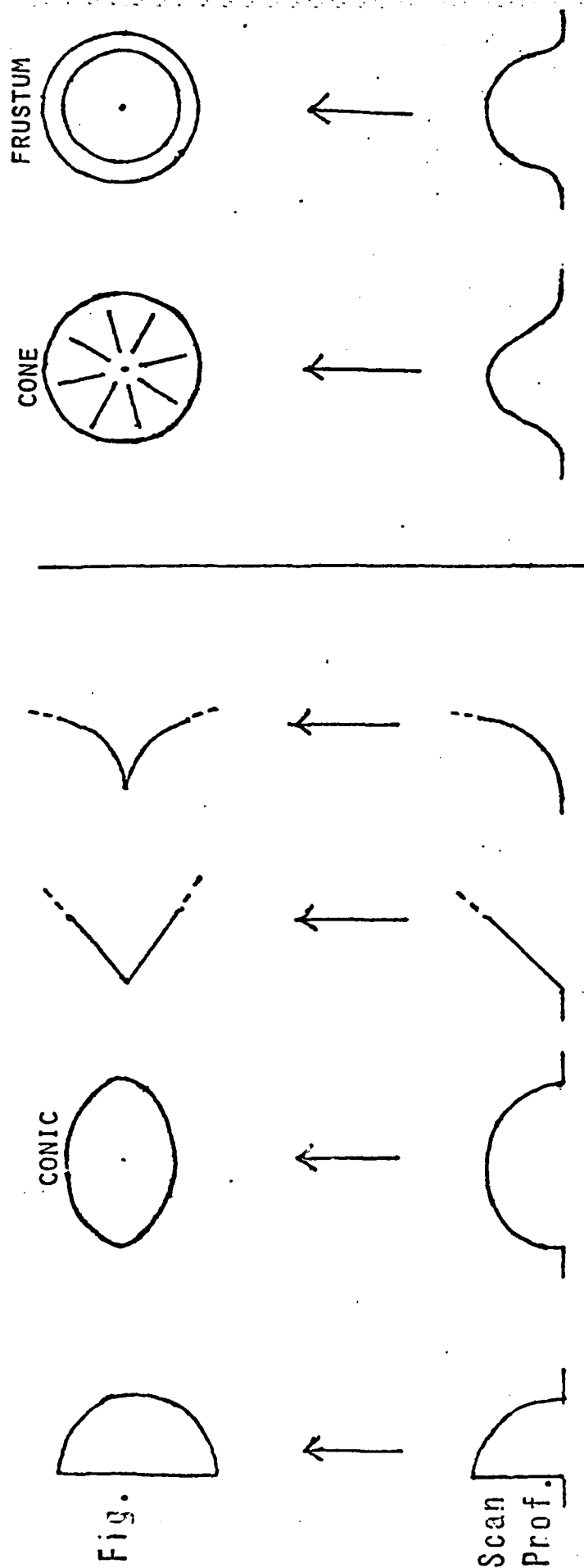
$$P(k_y) = \sum_{k_x} \left\{ \mathcal{F}_{x,y} [R(x, y)] \right\}^2$$

$$\underline{Scan} \quad P'(k_x) = \left\{ \mathcal{F}_x \left[\sum_y R(x, y) \right] \right\}^2$$

$$P'(k_y) = \left\{ \mathcal{F}_y \left[\sum_x R(x, y) \right] \right\}^2$$

In a scan we sum first, which introduces smoothing prior to the Fourier transform. Consequently, P 's generally have larger spectral indices than P 's.

Asymptotic Wavenumber Dependences of Simple Clouds



L-18

Asymptotic Dependences

P_{2D}	k^{-2}	k^{-2}	k^{-2}	k^{-4}	$k^{-2} \rightarrow k^{-4}$
P_{scan}	k^{-2}	k^{-4}	$k^{-2}(2+ E)$	k^{-5}	$k^{-3} \rightarrow k^{-5}$

1

2

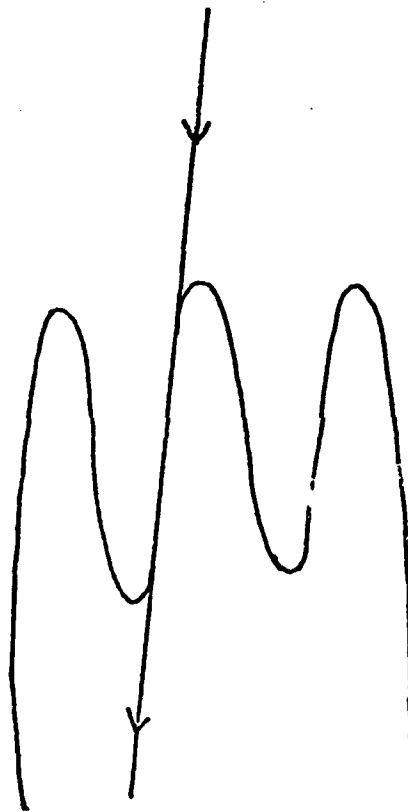
3

4

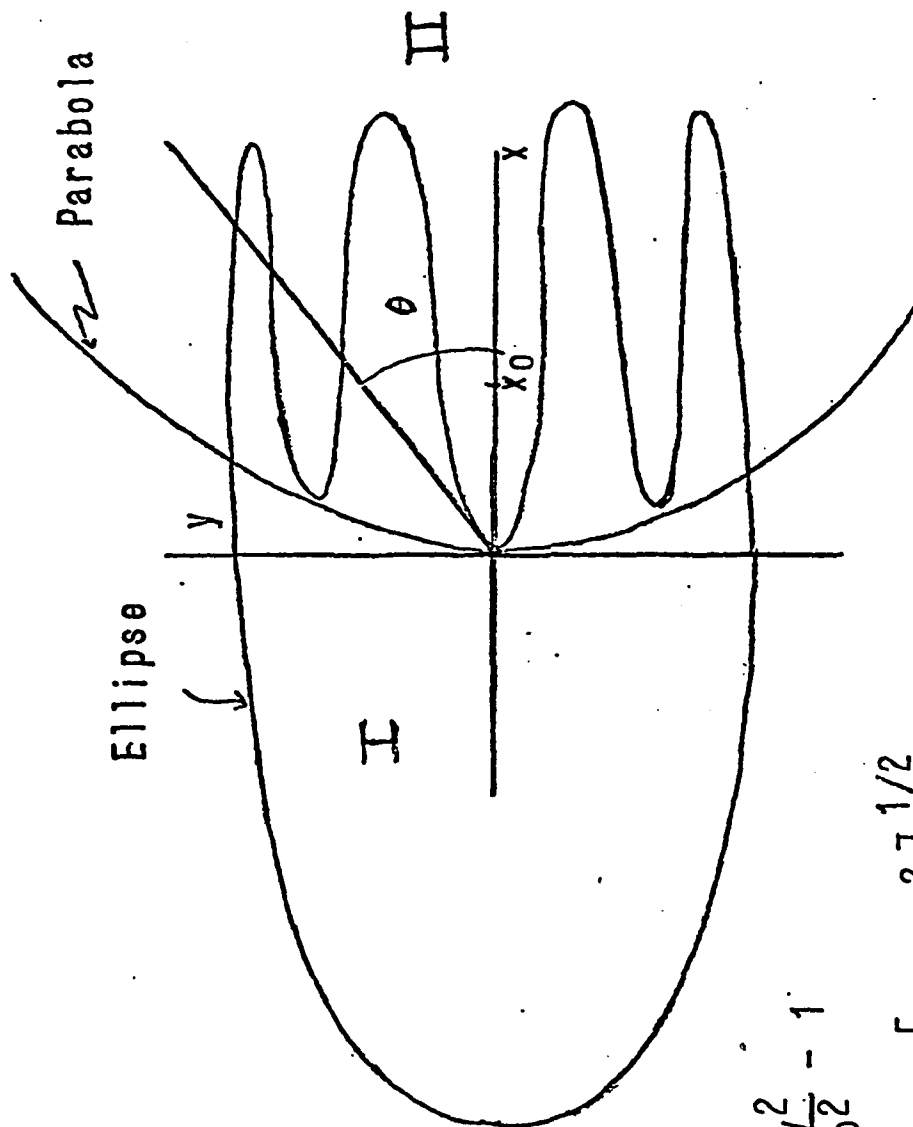
5

6

Summary: Asymptotic dependences in P_{scan} of k^{-2} with simple figures only occur if you are looking directly along an edge with a very sharp boundary. However, with more complex figures, such as a cloud with fingers, it is possible to find directions in which the P_{scan} will have k^{-2} dependence. This will occur when you are looking along a region in which the cloud curvature goes through zero.



Analytic Model (x, y Plane)

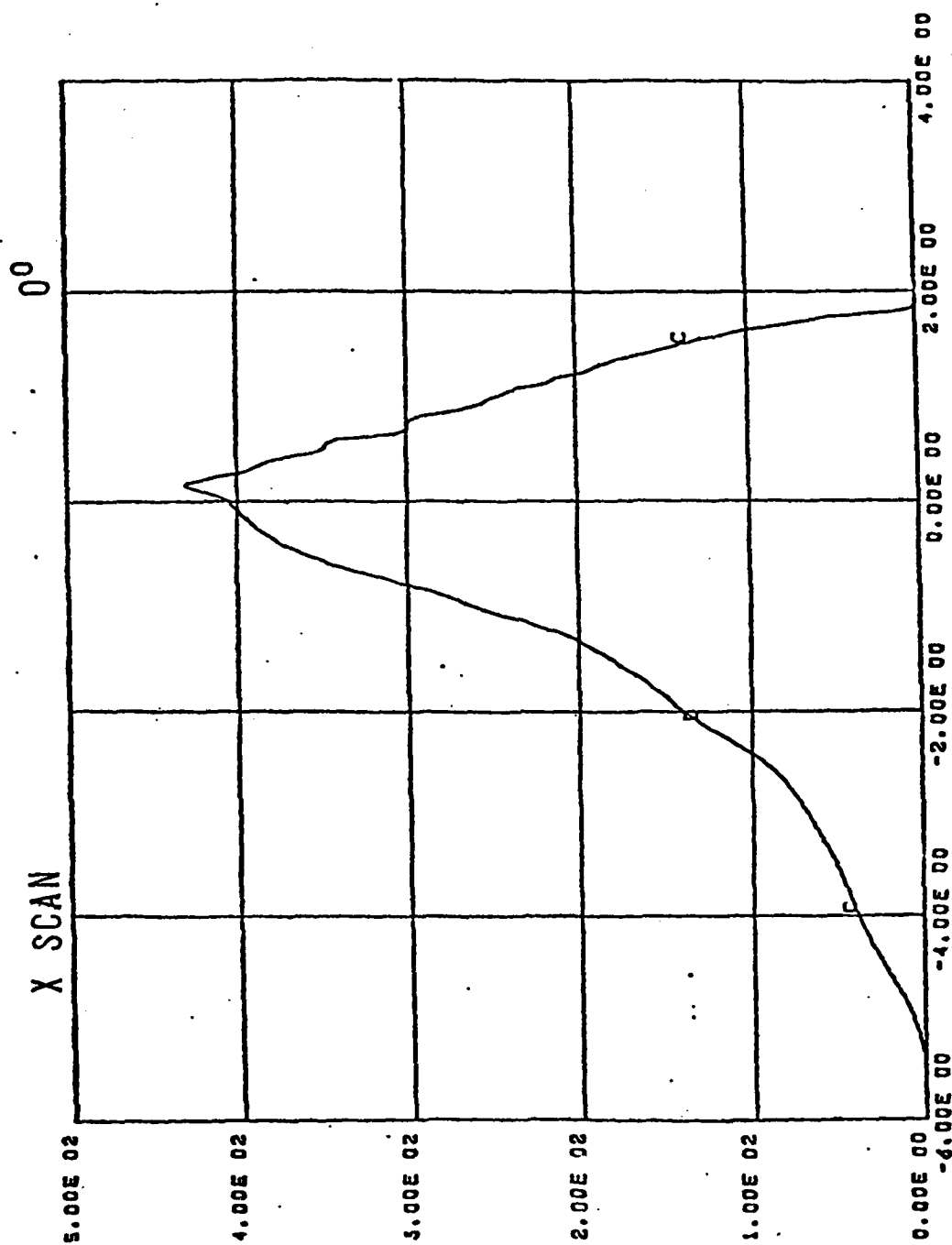


$$F_I(x, y) \sim \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$$

$$F_{II}(x, y) \sim x - x_0 + c [\cos(dy^2)]^{1/2}$$

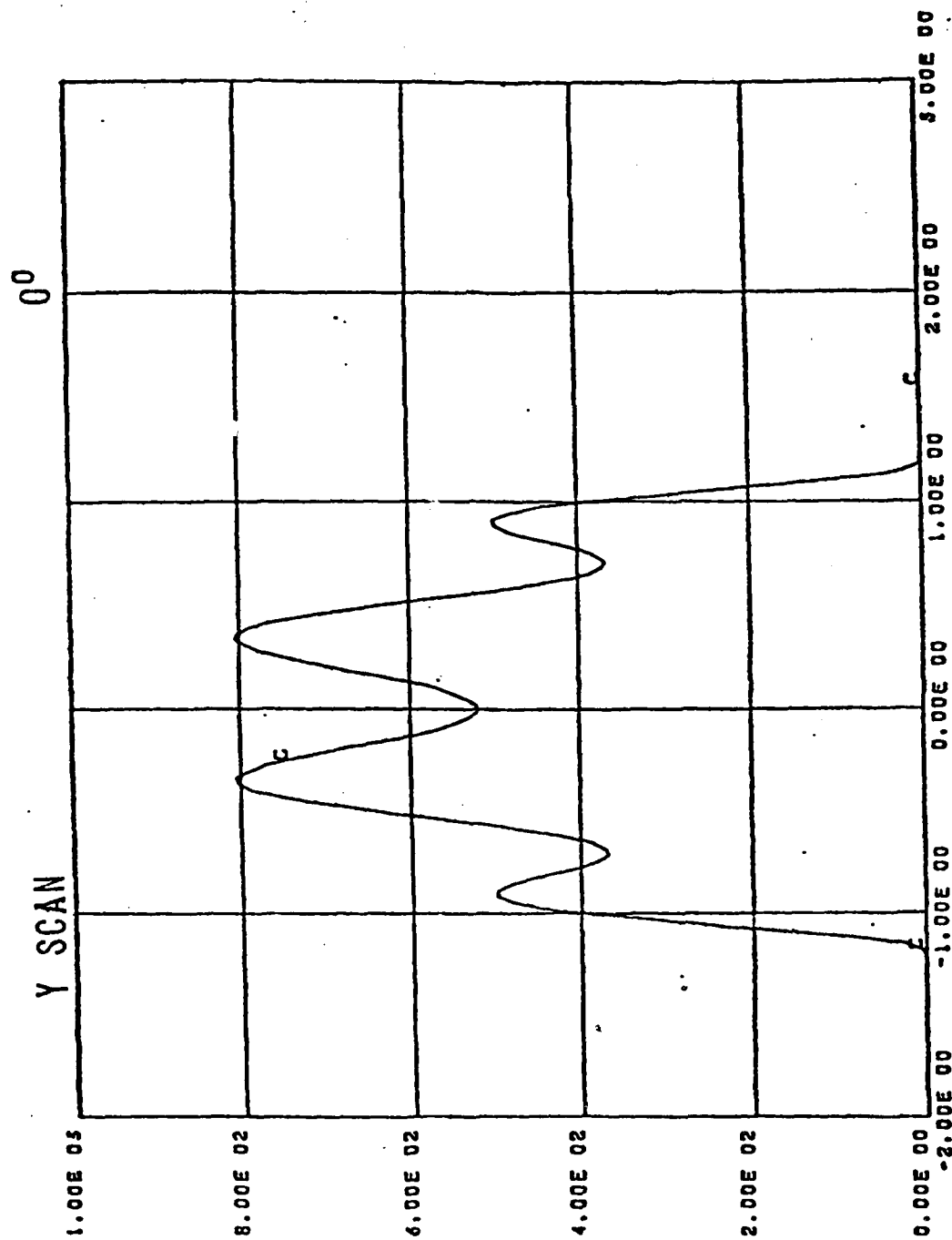
$$\text{Radiance} \quad R(x, y) \sim \frac{1}{1 + \exp[F^2(\text{sign of } F)/\Delta^2]}$$

$$\Delta_{II} = \text{constant}; \Delta_I(\theta) = \alpha + \epsilon\theta^2$$

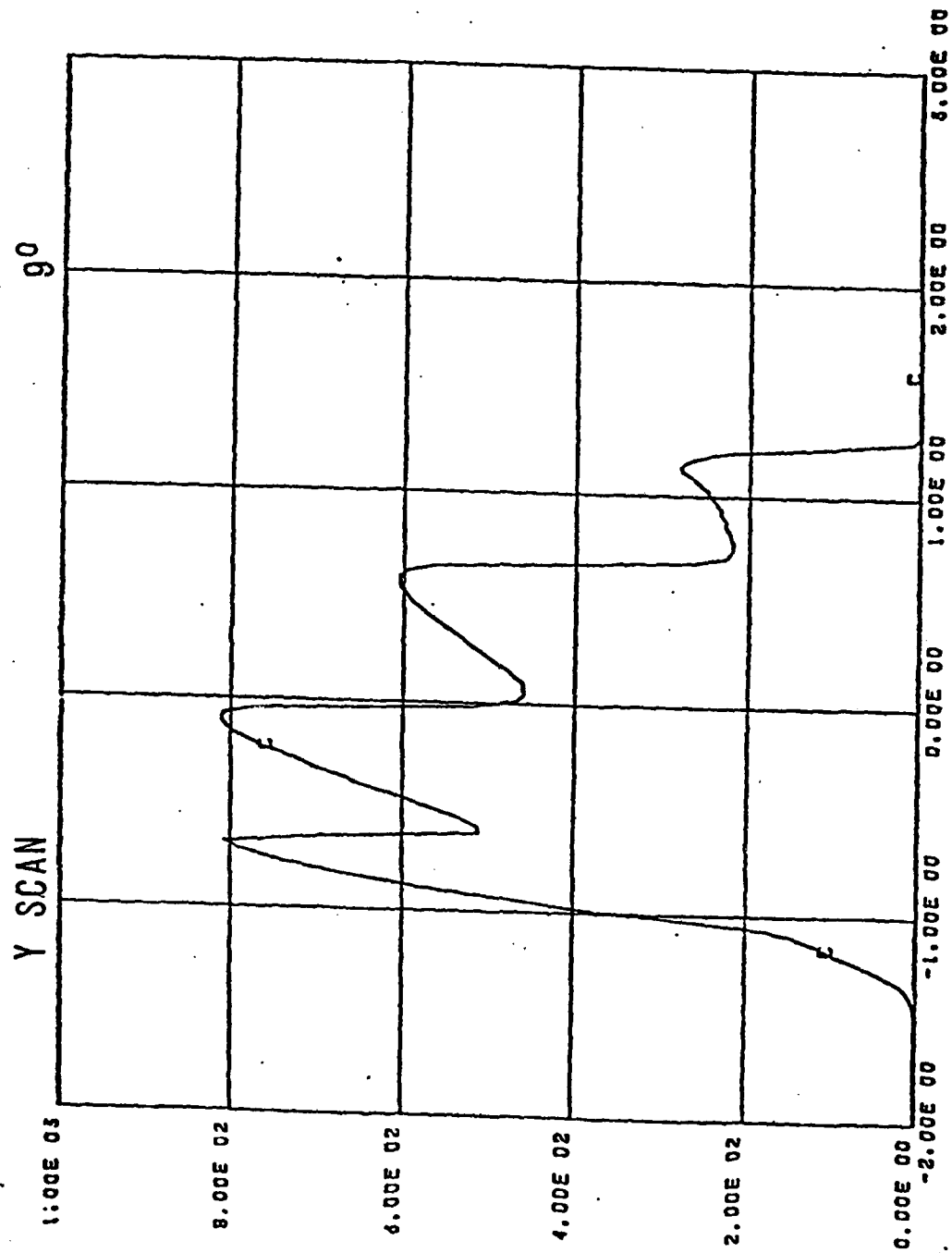


L-21

Viewgraph 13



L-22

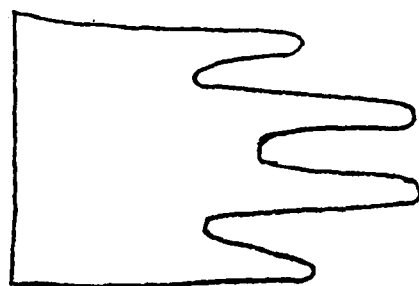


L-23

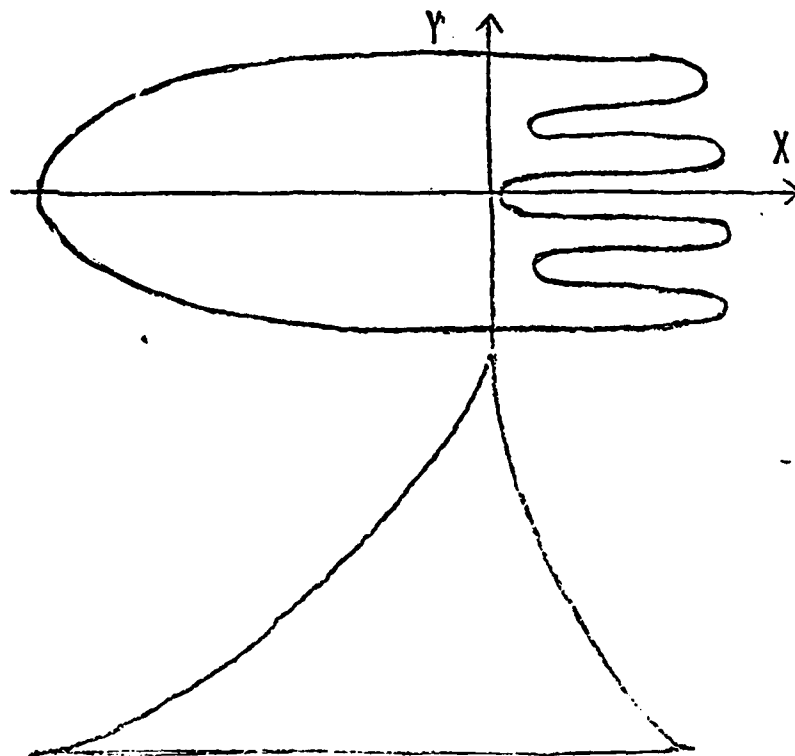
Viewgraph 15

Power Spectral Indices of Scans ~ Parallel to Cloud Fingers

Angle (Deg)	Smoothing		n (Y)	n (X)
0°	10	twice	5.8	2.0
0°	2	twice	6.1	2.2
0°	2	4 times	8.6	2.4
9°	10	twice	1.9	4.7
9°	2	twice	2.7	5.0
9°	2	4 times	3.8	5.4
0°	2	twice	6.1	2.2
5°	2	twice	3.5	3.5
9°	2	twice	2.7	5.0



Y SCAN



X SCAN

L-24

APPENDIX M

SIMULATION OF EARLY-TIME
URANIUM CLOUD DEVELOPMENT

Simulation of Early-Time
Uranium Cloud Development

NRL

J. Fedder

K. Hain

E. Hyman

Viewgraph #1 describes the physical model used at NRL to simulate a large uranium vapor release at 200 km altitude. We follow the species N_2 , O_2 , O , U , UO , UO_2 , UO^+ and UO_2^+ in their temporal evolution as influenced by hydrodynamics, diffusion and chemistry via the first equation. The hydrodynamic transport assumptions and equations are described. On 1b the treatment of diffusion is shown. We choose K in the last equation to lead to a nominal diffusion coefficient at 200 km altitude of $0.1 \text{ km}^2/\text{sec}$. For a gradient scale of 100 m the diffusion velocity would be 100 m/sec in the ambient atmosphere. This value is about 17% of the sound speed at that altitude. We have also studied the evolution of a release for diffusion coefficients 1/3 and 3 times the nominal value. The slower diffusion rates lead to very slow oxidation of the cloud and are therefore uninteresting. In the following graphs we show results for nominal diffusion and faster diffusion.

Viewgraph #2 presents chemical reaction rates and the properties of the ambient atmosphere and uranium vapor release. It is assumed that the vapor release takes place quickly (within 1 sec).

Viewgraph #3 shows initial conditions for the neutral gas perpendicular to B . The vertical axis is logarithm to the base 10 of the number density in cm^{-3} . The horizontal axis is distance from the center of the release in km. The following slides show the evolution of the release perpendicular to B for a nominal diffusion.

Viewgraph #4 Ion densities at 2 secs.

Viewgraph #5 Neutral densities at 2 secs.

Viewgraph #6 Ions at 5 secs.

Viewgraph #7 Neutrals at 5 secs.

Viewgraph #8 Ions at 10 secs.

Viewgraph #9 Neutrals at 10 secs.

Viewgraph #10 Ions at 20 secs.

Viewgraph #11 Ions at 30 secs.

Viewgraph #12 Ions at 40 secs.

Viewgraph #13 Ions at 50 secs.

Viewgraphs #14, #15, and #16 show contours of the densities of UO^+ , UO_2^+ and O_2 . The vertical axis is time after release; the horizontal axis is radius from release point.

Comment: The above results show that such a large release initially burns out the O and O_2 of the neutral atmosphere. The O is able to fairly rapidly diffuse into the U vapor and oxidize it to UO^+ . However, O_2 takes much longer

AD-A126 341

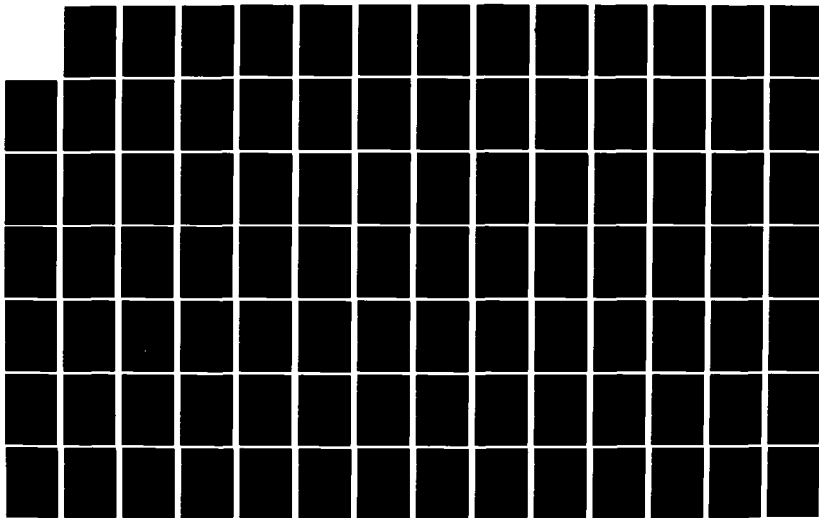
GEOPHYSICAL PLASMAS AND ATMOSPHERIC MODELING(U) SCIENCE
APPLICATIONS INC MCLEAN VA E HYMAN ET AL. FEB 83
SRI-83-144-WA N00014-81-C-2038

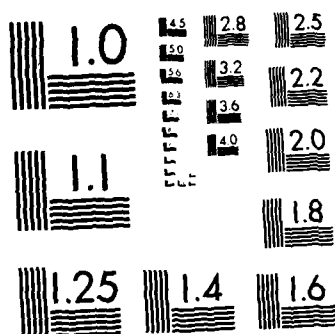
4/7

UNCLASSIFIED

F/G 4/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963 A

to oxidize the UO^+ to UO_2^+ . This slower oxidation results because O_2 has a smaller ambient density than 0 and is chemically consumed as it tries to diffuse into the release region.

Viewgraph #17 shows the initial conditions for the simulations parallel to B. The diffusion is nominal for the following set of graphs. The horizontal axis here covers a range to 15 km.

Viewgraph #18 Ions at 2 secs.

Viewgraph #19 Neutrals at 2 secs.

Viewgraph #20 Ions at 5 secs.

Viewgraph #21 Neutrals at 5 secs.

Viewgraph #22 Ions at 10 secs.

Viewgraph #23 Neutrals at 10 secs.

Viewgraph #24, #25, and #26. Contours of UO^+ ; UO_2^+ and O_2 number densities in time parallel to B.

Comments: The evolution parallel to B is quite similar to that perpendicular to B with the exception that the ion cloud becomes considerably larger as it diffuses outwards. Again ambient O_2 is burnt out in the release region and oxidation of UO^+ to UO_2^+ is allowed to take place only in a shell around the release.

Viewgraphs #27 and #28 show results perpendicular to B for diffusion 3 times nominal. Comparison of these results to #8 and #9 show that oxidation is taking place more rapidly. However even for this very rapid diffusion ambient O_2 is burnt out in the release region and most of the oxidation takes place in a shell.

Viewgraphs #29 and #30 show results parallel to B for fast diffusion. They can be compared to #22 and #23. The comments about shell oxidation and O_2 burn out still apply.

Conclusions: These simulation results demonstrate clearly that for a large, rapid uranium vapor release in the atmosphere at about 200 km altitude, the oxidation of U to UO_2^+ can take a fairly long time (a few 10s of seconds). The long oxidation time is caused by the dense U vapor burning out the atmospheric molecular oxygen. Although the O_2 attempts to diffuse into the release region it continues to be chemically consumed by the uranium vapor. Only after oxidation is complete in a shell surrounding the release point can O_2 finally reenter the central region of the release. For smaller releases or for slower releases where the uranium vapor is spread along the rocket trajectory, these conclusions may not apply. In particular for releases where the uranium vapor density is less than the ambient molecular oxygen density complete oxidation would take place promptly. Moreover UO_2^+ would not be formed in a shell-like structure around the release point.

Physical Model

$$\frac{\partial n_a}{\partial t} = \frac{\partial n_a}{\partial t} \text{ hydro} + \frac{\partial n_a}{\partial t} \text{ diff} + \frac{\partial n_a}{\partial t} \text{ chem}$$

Transport ions and neutrals separately

Two cases

parallel to B

spherical transport of
neutrals - cartesian
transport of ions

perpendicular to B

spherical transport of
neutrals - fixed ions

Transport Equations

$$\frac{\partial \rho_s}{\partial t} = - \nabla \cdot (\rho_s \underline{v}_s)$$

$$\frac{\partial \underline{v}_s}{\partial t} = - (\underline{v}_s \cdot \nabla) \underline{v}_s - \frac{1}{\rho_s} \nabla (\rho_s + q_s) - \sigma_{it} (\underline{v}_s - \underline{v}_t)$$

$$\frac{\partial p_s}{\partial t} = - \nabla \cdot (p_s \underline{v}_s) - (\gamma - 1) (p_s + q_s) \nabla \cdot \underline{v}_s$$

$$- \sigma \rho_s \rho_t (T_t - T_s) + (\gamma - 1) \frac{\sigma \rho_s \rho_t}{(\rho_s + \rho_t)^2} \rho_s (\underline{v}_s - \underline{v}_t)^2$$

Diffusion Equations

$$\sum_{\beta} \rho_a \rho_{\beta} \sigma_{a\beta} (\underline{v}_a - \underline{v}_{\beta}) = T (\nabla \rho_a - \rho_a / \rho \nabla \rho)$$

$$\text{for } \sigma_{a\beta} = \sigma \approx \text{constant}$$

$$\frac{\partial \rho_a}{\partial t} = - \nabla \cdot (k \nabla \rho_a / \rho)$$

Chemical Reaction Rate Coefficients

(1)	$U + O \rightarrow UO^+ + e^-$	$5 \times 10^{-10} \text{ cm}^3/\text{sec}$
(2)	$U + O_2 \rightarrow UO_2^+ + e^-$	$3.8 \times 10^{-12} \text{ cm}^3/\text{sec}$
(3)	$U + O_2 \rightarrow UO + O$	$1.8 \times 10^{-10} \text{ cm}^3/\text{sec}$
(4)	$UO + O \rightarrow UO_2^+ + e^-$	$4 \times 10^{-10} \text{ cm}^3/\text{sec}$
(5)	$UO + O_2 \rightarrow UO_2 + O$	$2 \times 10^{-9} \text{ cm}^3/\text{sec}$
(6)	$UO^+ + O_2 \rightarrow UO_2^+ + O$	$4 \times 10^{-9} \text{ cm}^3/\text{sec}$

Ambient Atmosphere - Altitude = 200 km.

$$[O]_0 = 4.4 \times 10^9 \text{ cm}^{-3} \qquad T_0 = 0.1 \text{ ev}$$

$$[N_2]_0 = 4.2 \times 10^9 \text{ cm}^{-3}$$

$$[O_2]_0 = 2.6 \times 10^8 \text{ cm}^{-3}$$

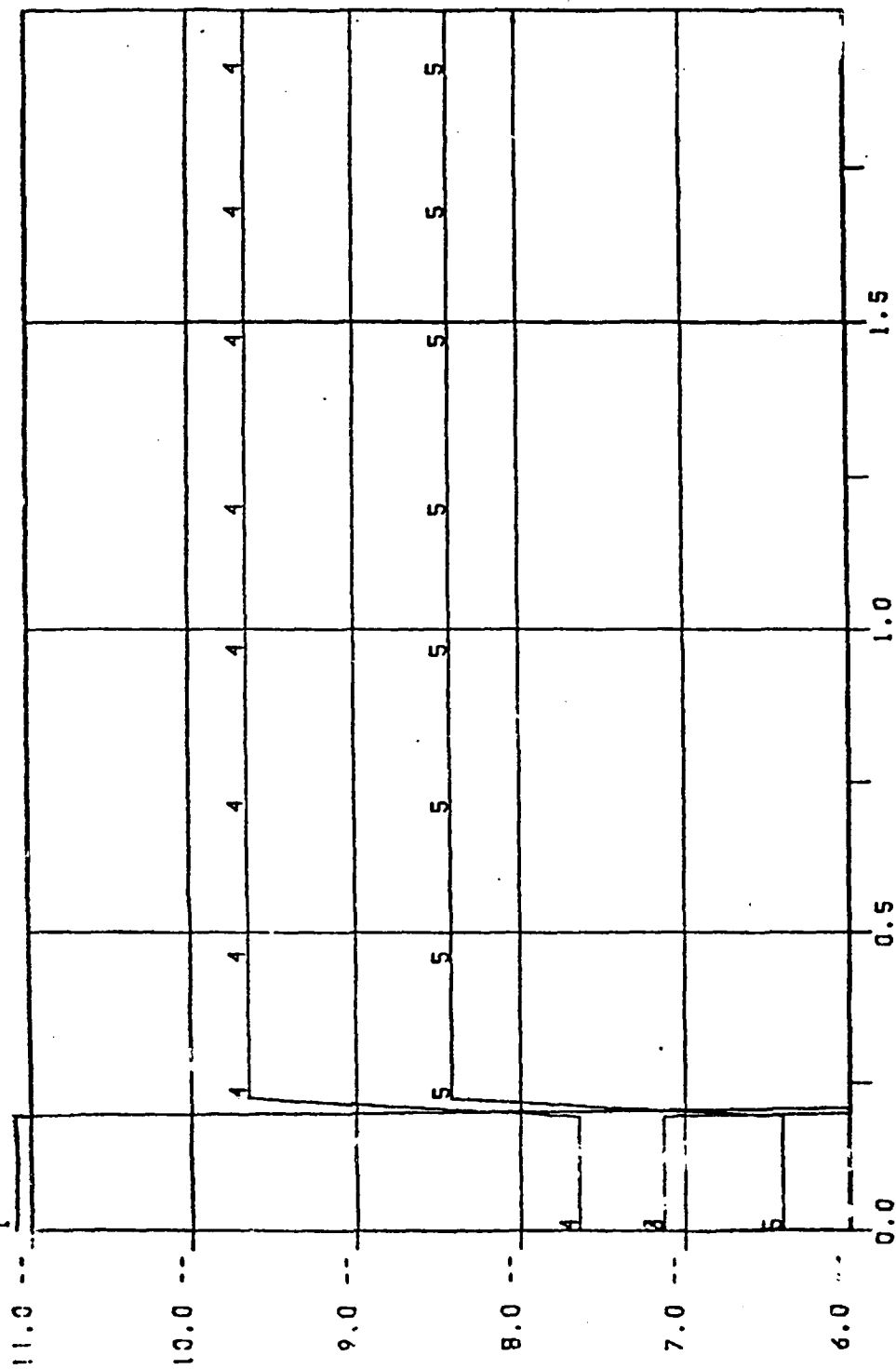
Release of 2 kg of atomic uranium $\sim 5 \times 10^{24}$ atoms
at a temperature of 0.276 ev.

01/04/83

TIME 0.00E 00

DENS.LOG10

NUM.



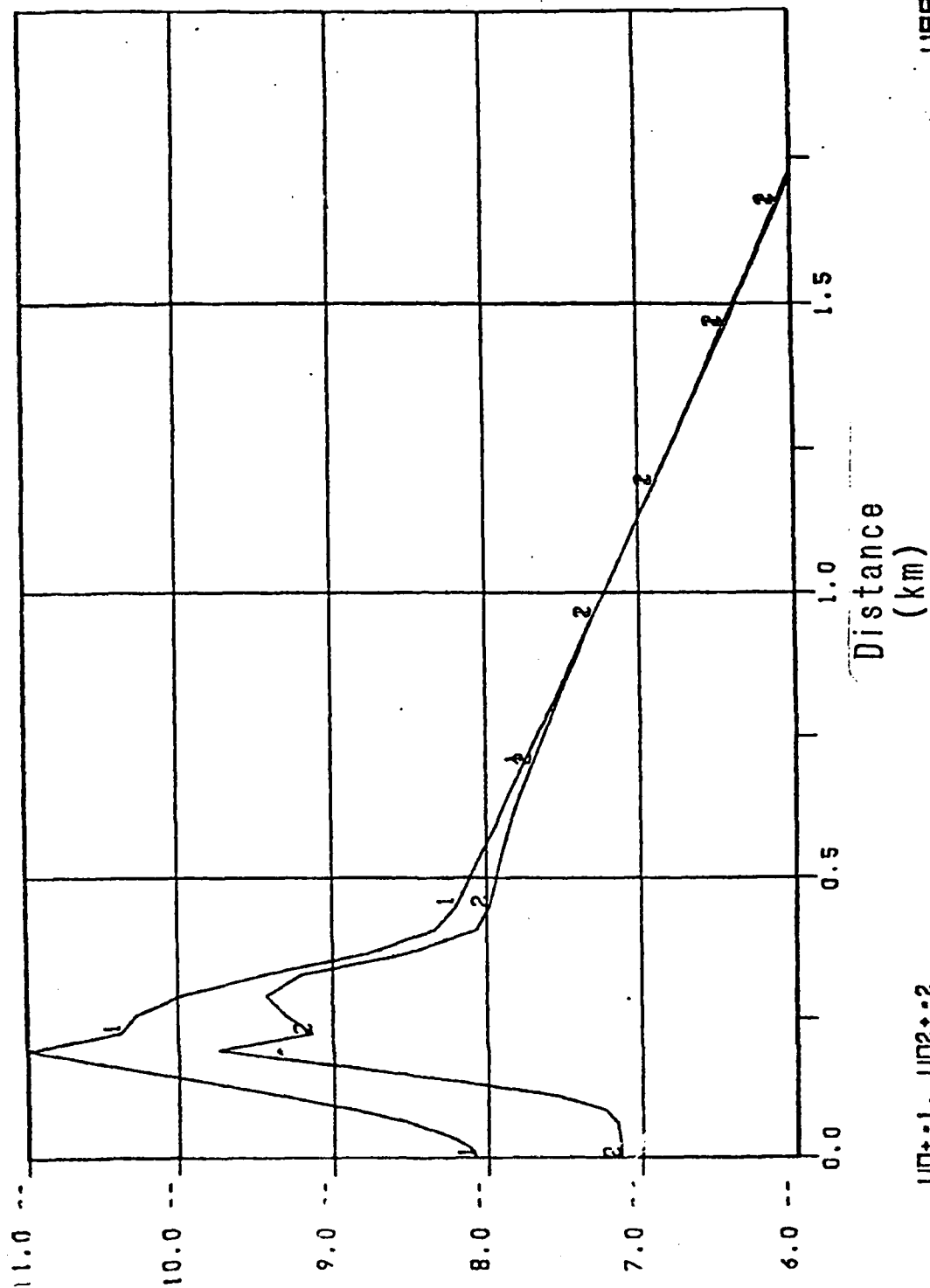
Distance
(km)

U-1. UD-2. UD2-3. D-4. D2-5

URR0

01/04/83

ION. NUM. DENS.LOG10 TIME 2.01E 00



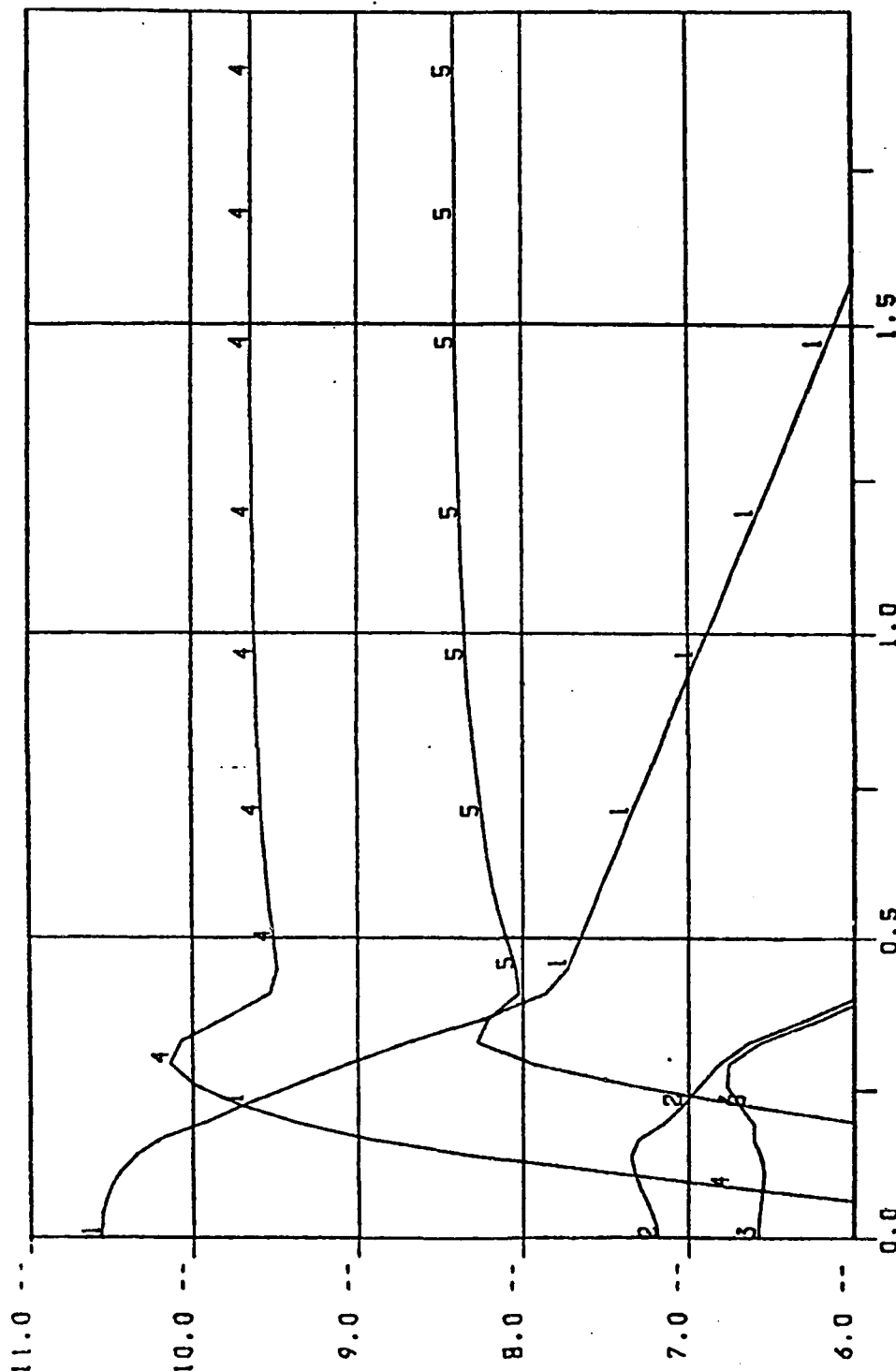
UD+1, UD2+2

URR0

01/04/83

TIME 2.01E 00

NUT. NUM. DENS.LOG10



Distance
(km)

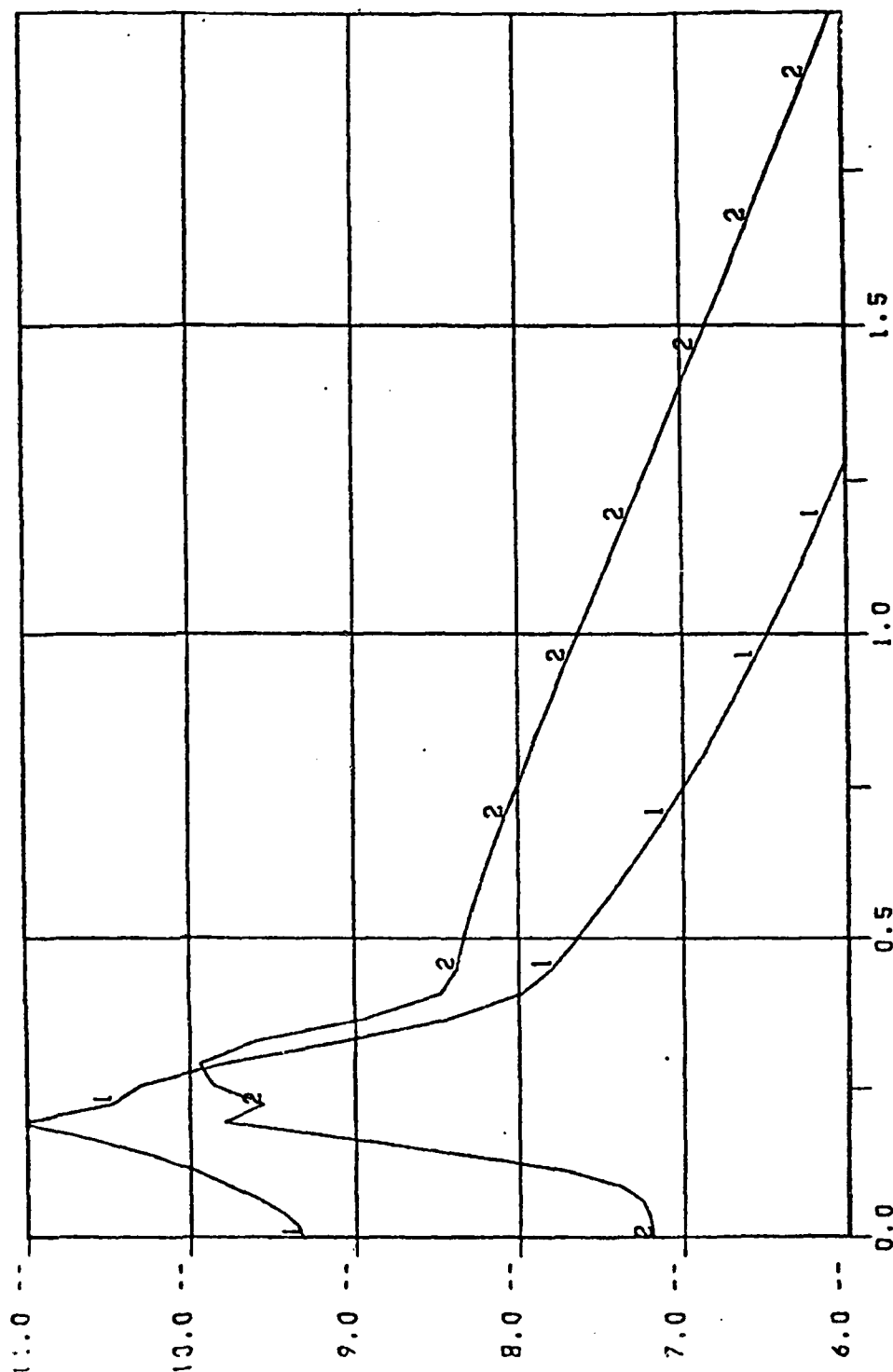
U-1, UD-2, UD2-3, D-4, D2-5

URRO

01/04/83

TIME 5.02E 00

ION. NUM. DENS. LOG10



Distance
(km)

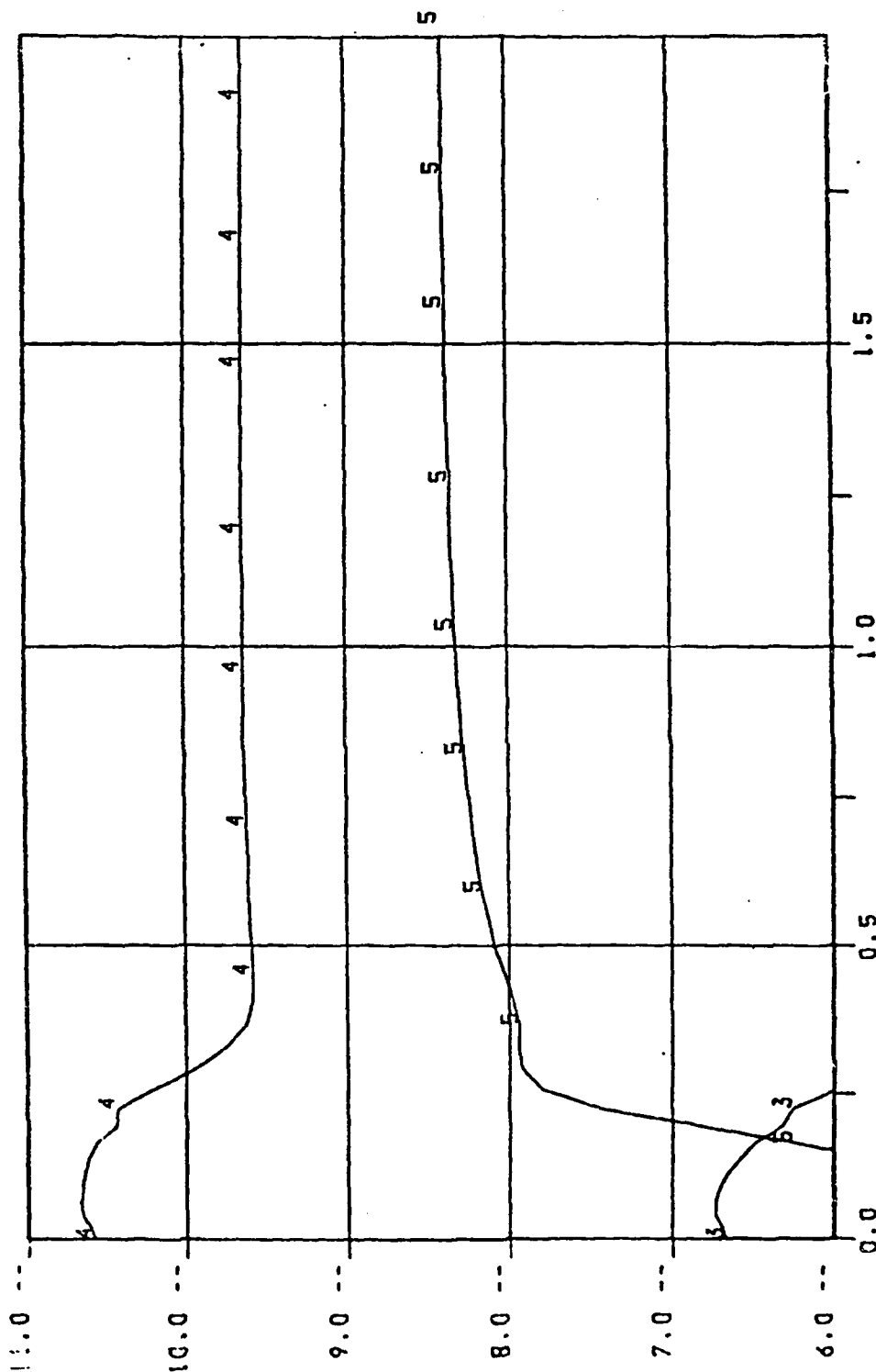
UD01, UD02+02

UFR0

01/04/83

TIME 5.02E 00

NUT. NUM. DENS.LOG10



Distance
(km)

U-1. U02-2. U02-3. 0-4. 02-5

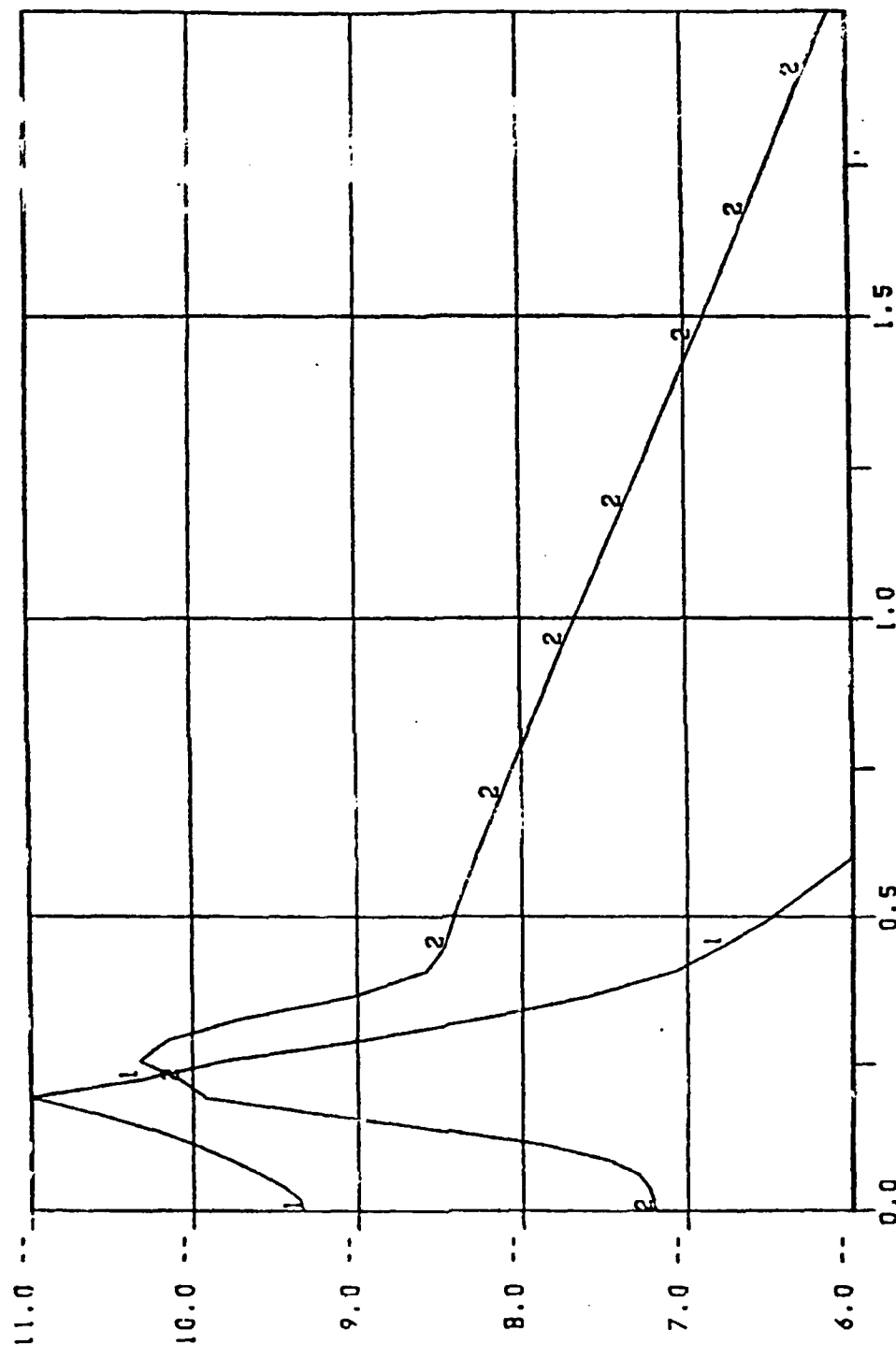
UHR0

M-13

01/06/83

TIME 1.00E 01

ION. NUM. DENS. LOG10



Distance
(km)

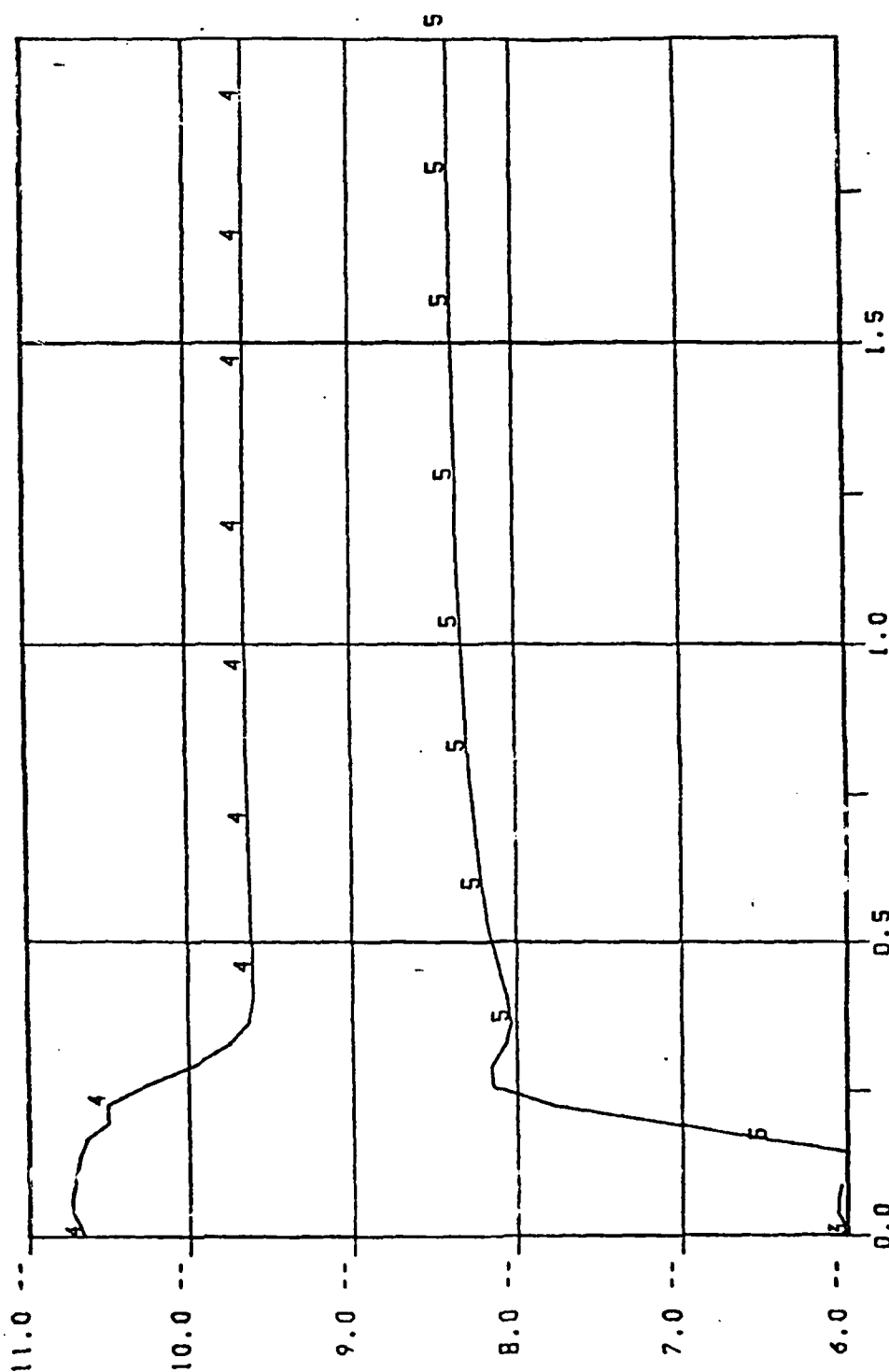
UD-1, UD2-2

UR00

01/06/83

TIME 1.00E 01

NUT. NUM. DENS. LOG10



Distance (km)

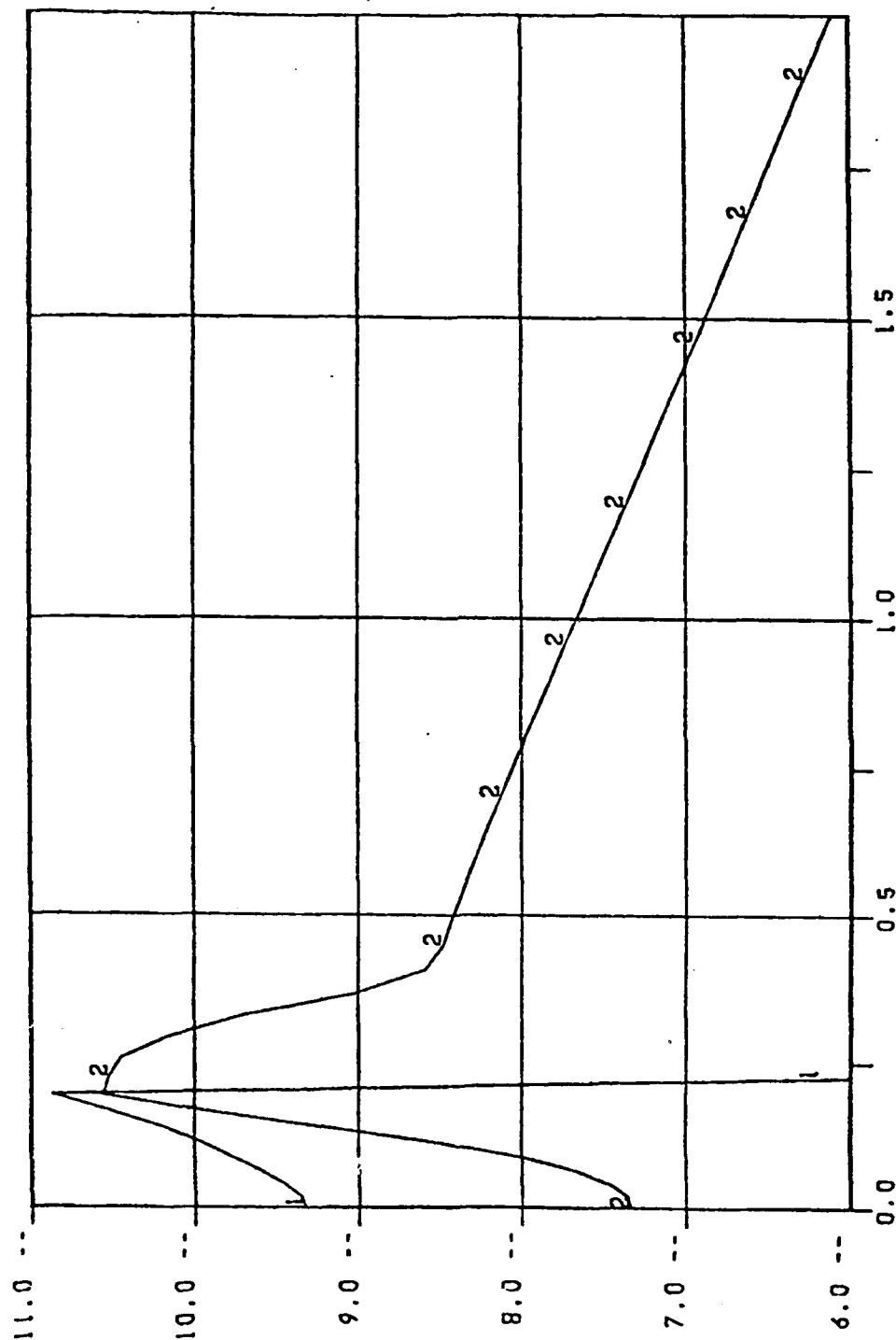
U-1, U0-2, U02-3, D-4, D2-5

UFGO

01/06/83

TIME 2.00E 01

IDN. NUM. DENS.LDG10

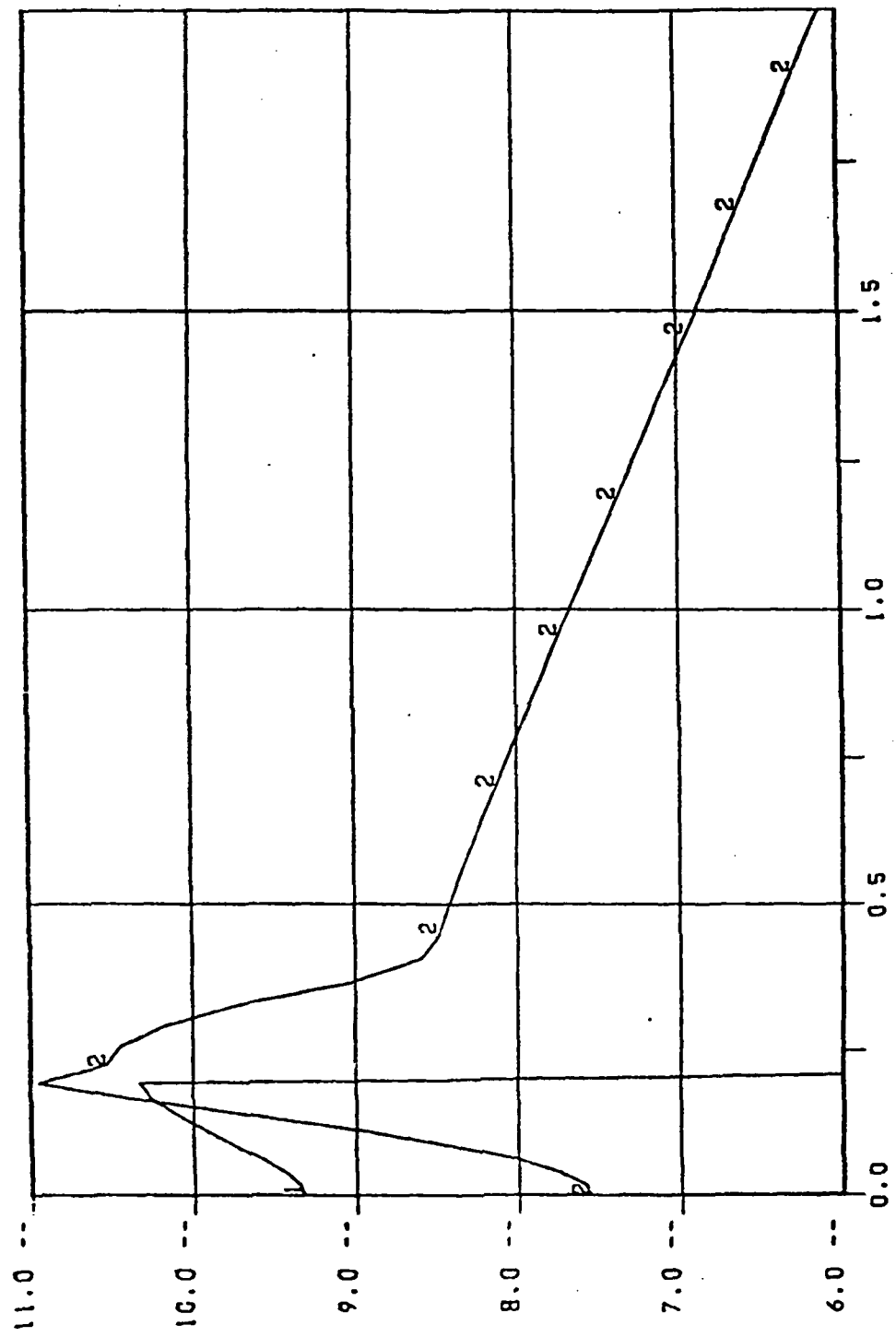
Distance
(km)

UD*-1. UD2*-2

UR00

01/06/83

IGN. NUM. DENS.LOG10 TIME 3.00E 01



Distance (km)

UR00

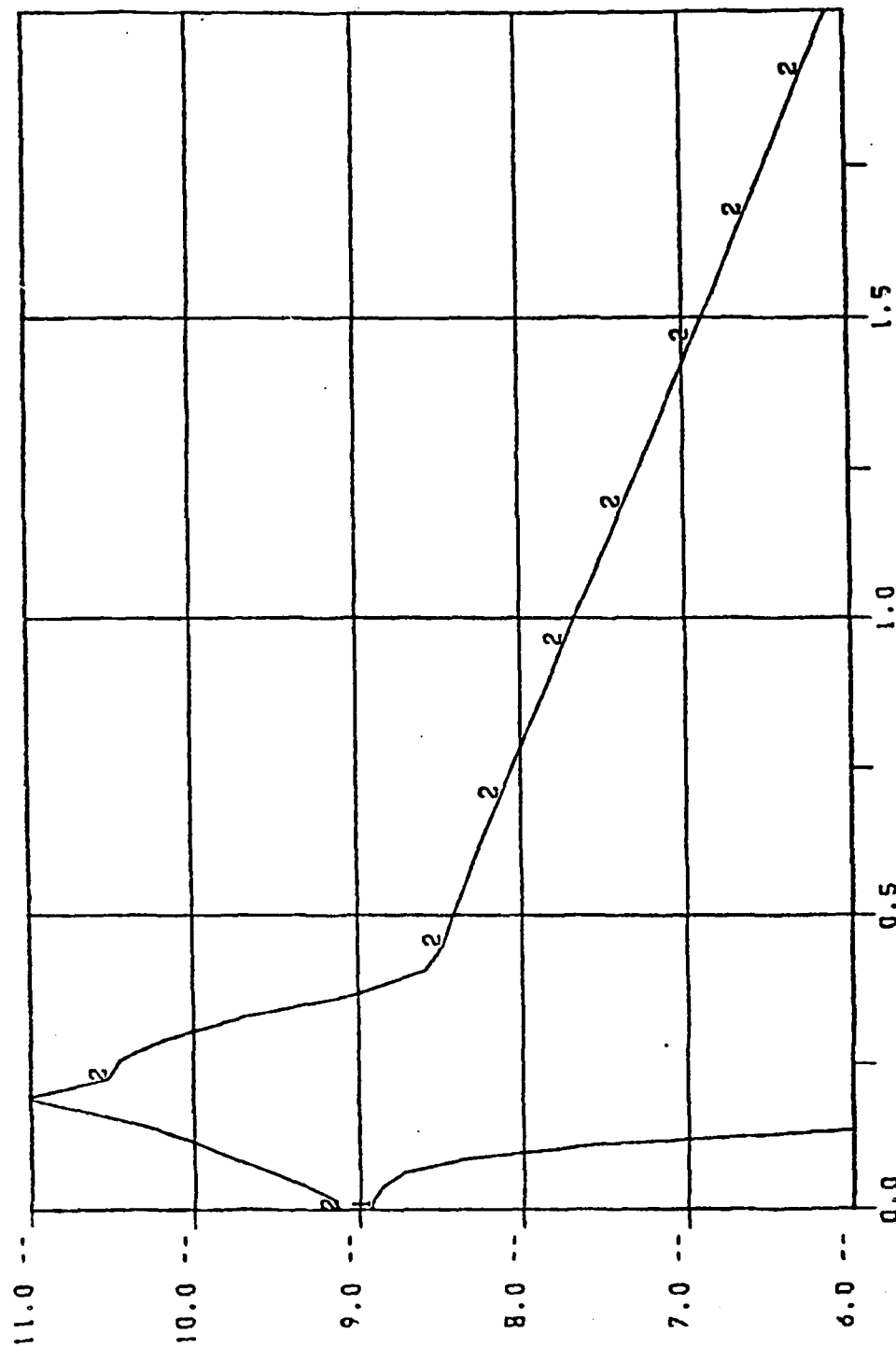
UD+1. UD2+2

M-17

01/06/83

TIME 4.00E 01

IGN. NUM. DENS. LOG10



Distance
(km)

UN+1. UD2+2

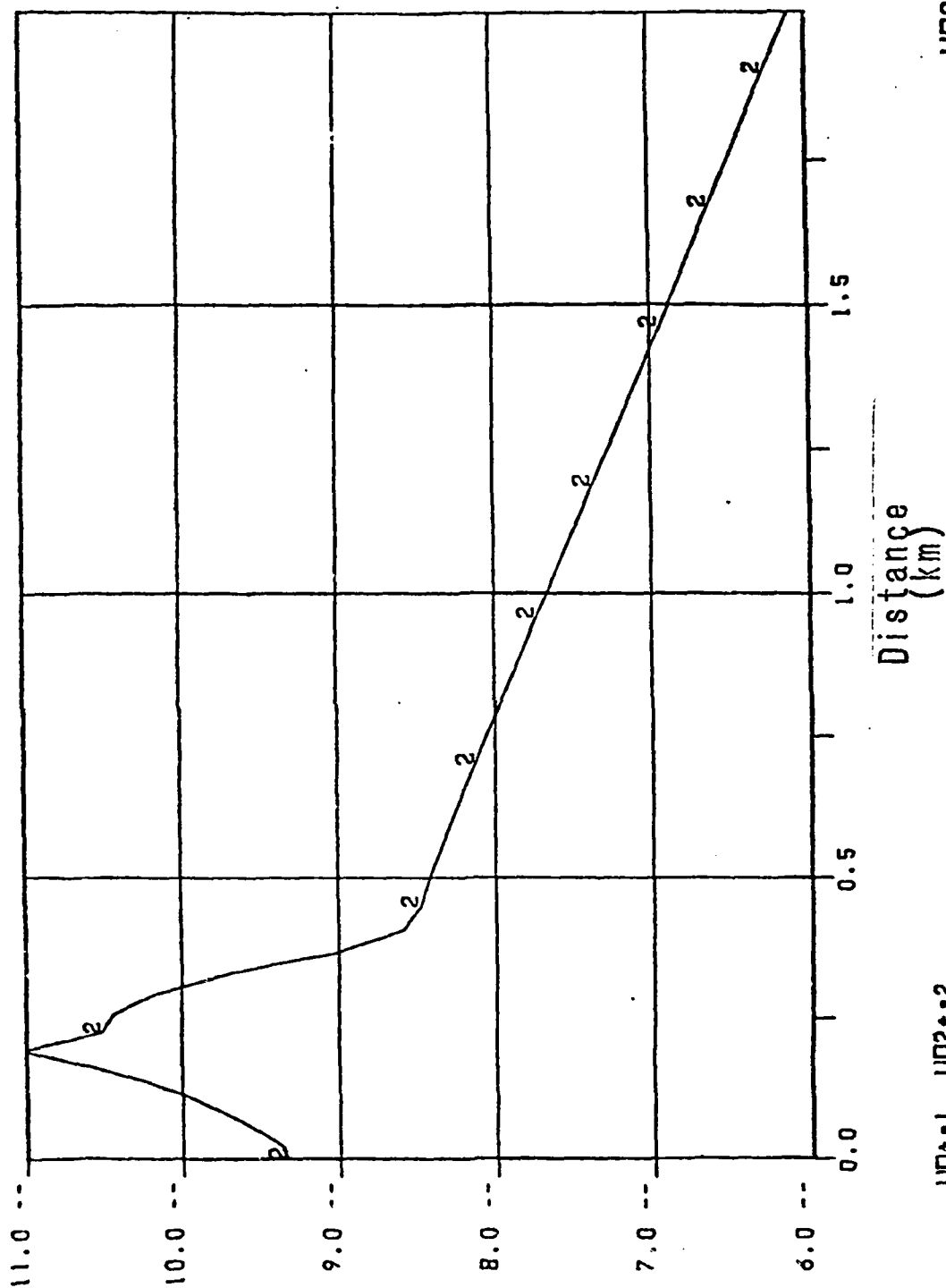
UR00

M-18

01/06/83

TIME 5.00E 01

ION. NUM. DENS. LOG10

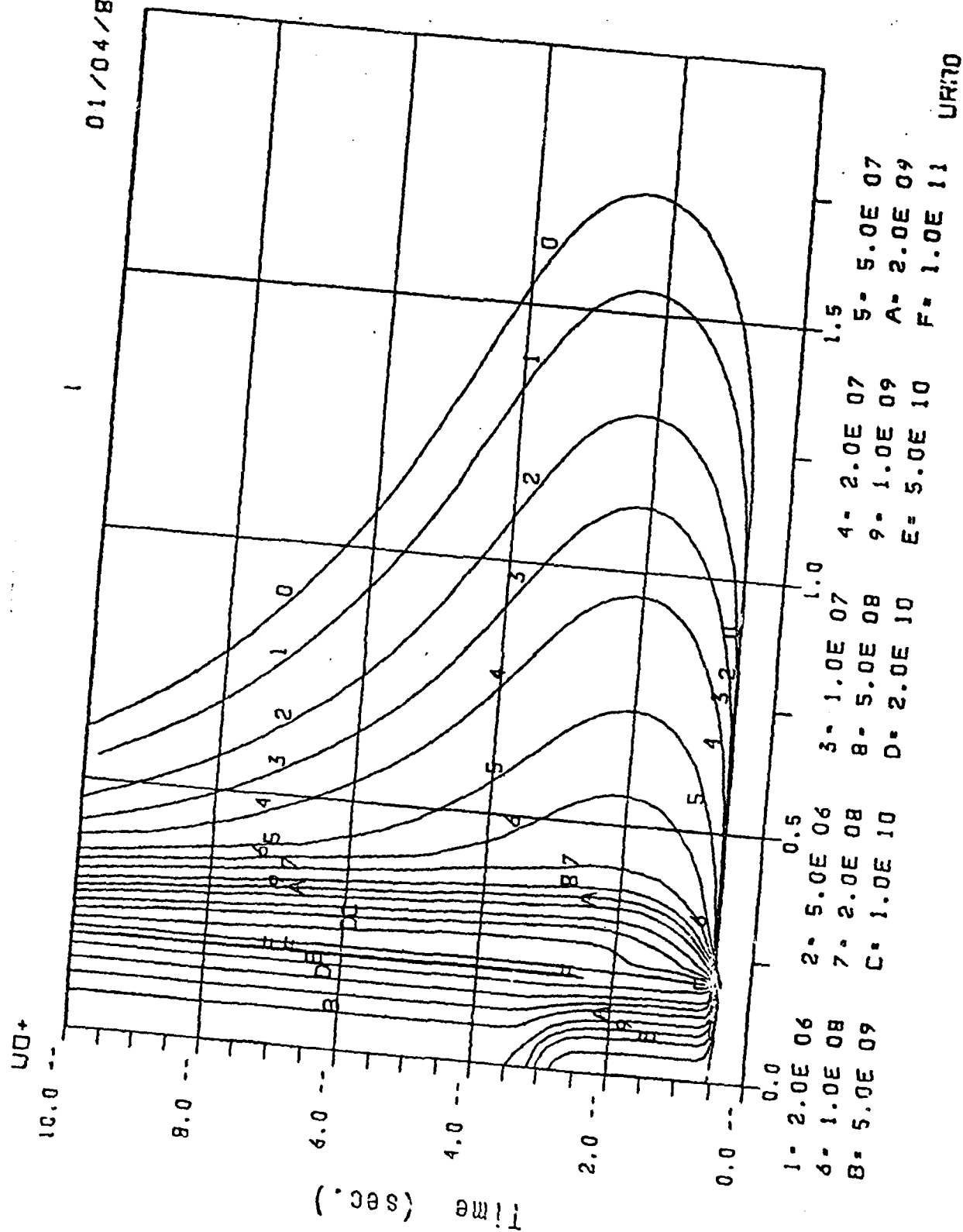


UR00

UR00-1, UR00-2

M-19

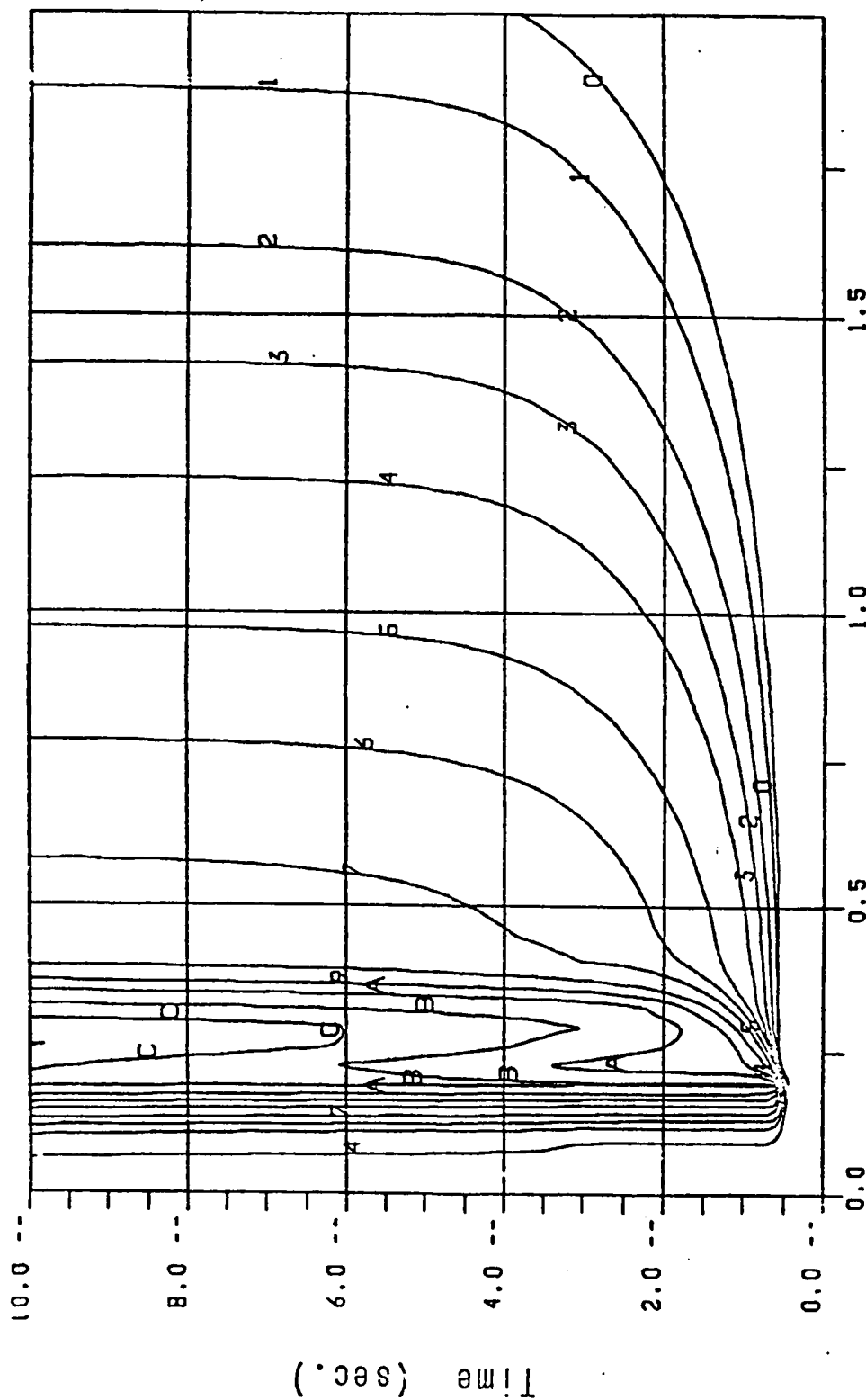
—



02:57

01/04/83

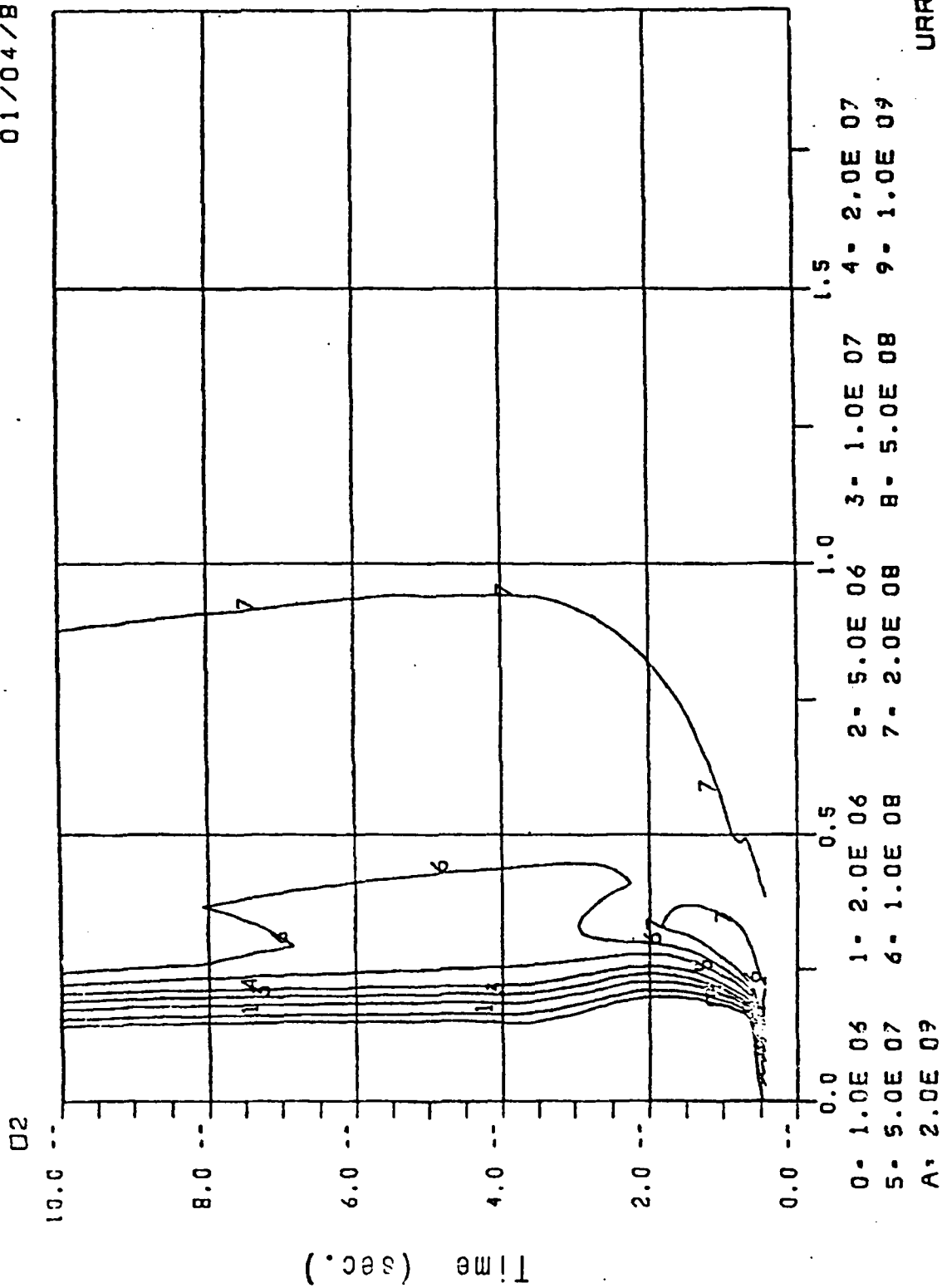
UD2+



URR0

M-21

01/04/83

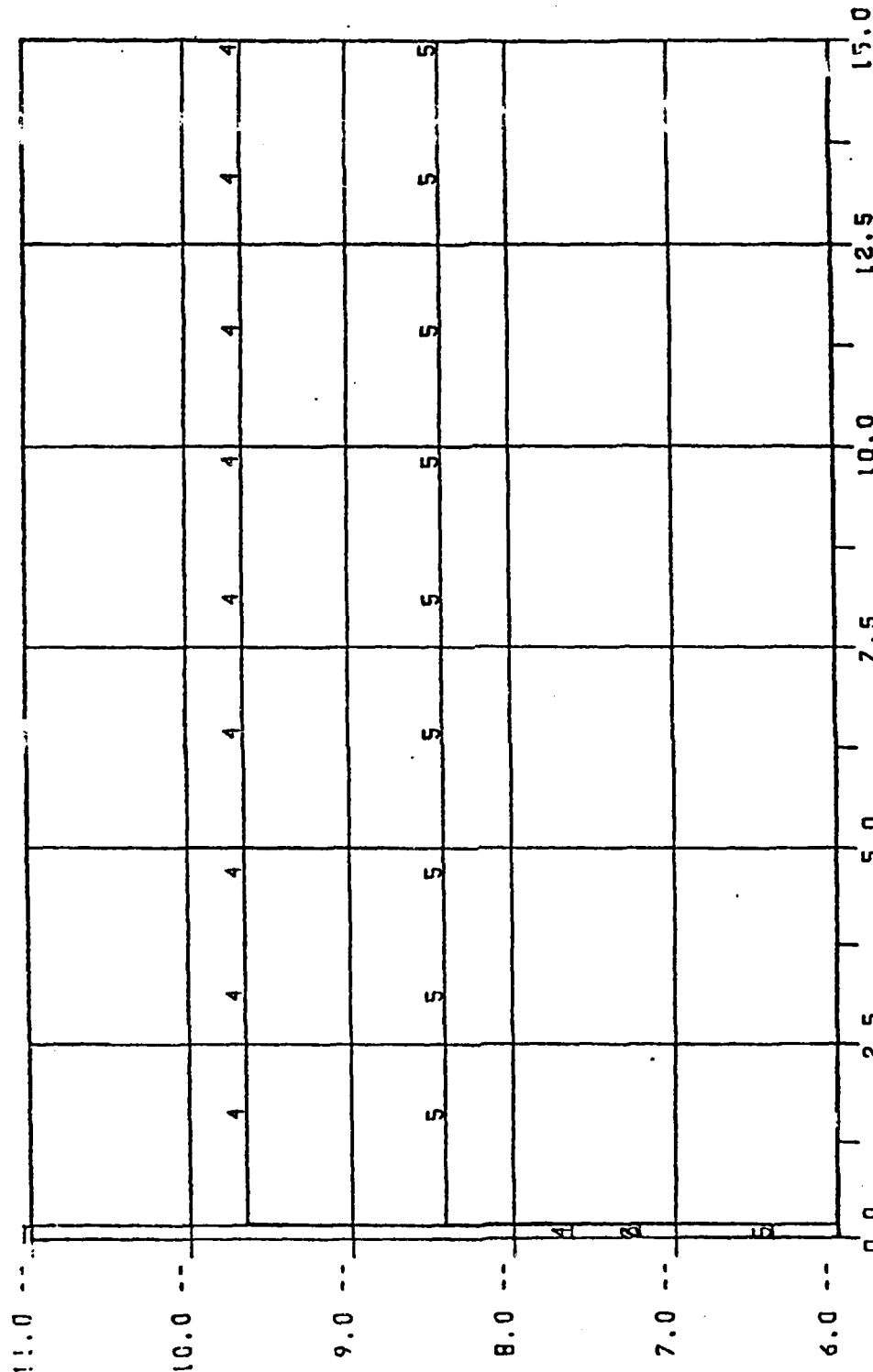


01/04/83

TIME 0.00E 00

DENS.LOG10

NUM.



Distance
(km)

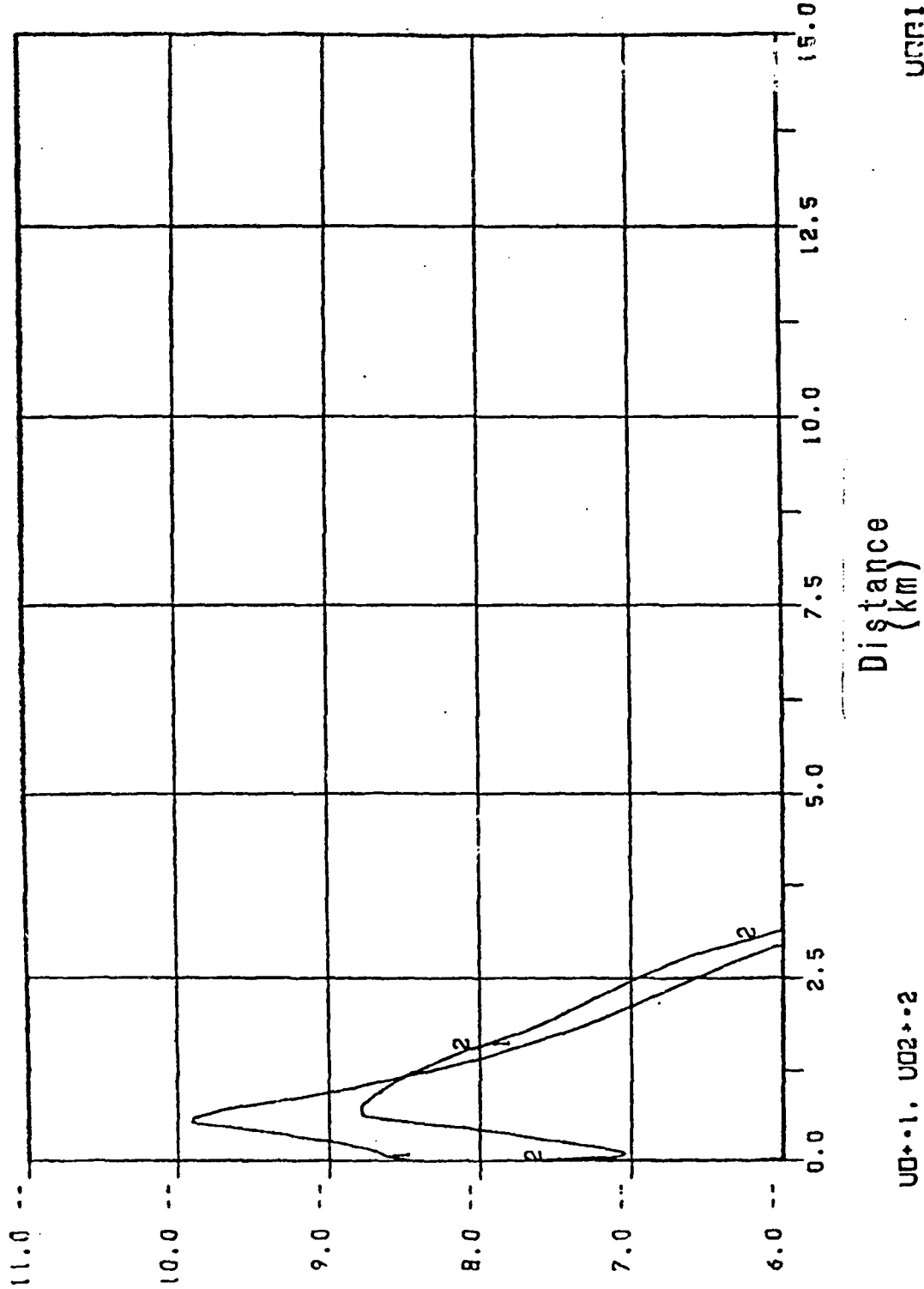
U-1. U0-2. U02-3. 0-4. 02-5

URR-1

M-23

01/04/83

IDN. NUM. DENS.LOG10 TIME 2.00E 00

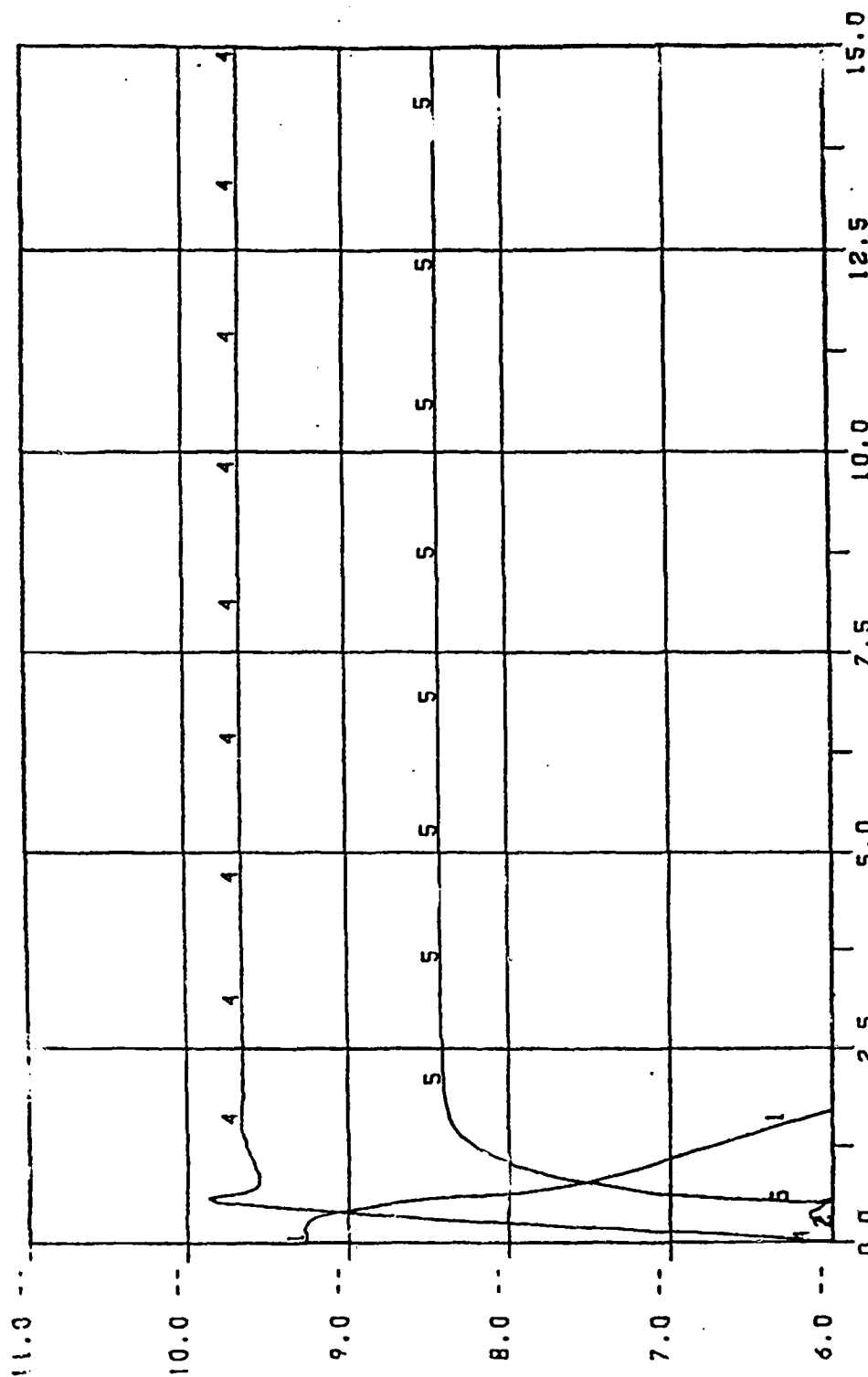


M-24

01/04/83

TIME 2.00E 00

NUM. DENS. LOG10



Distance
(km)

U-1, U0-2, U02-3, 0-4, 02-5

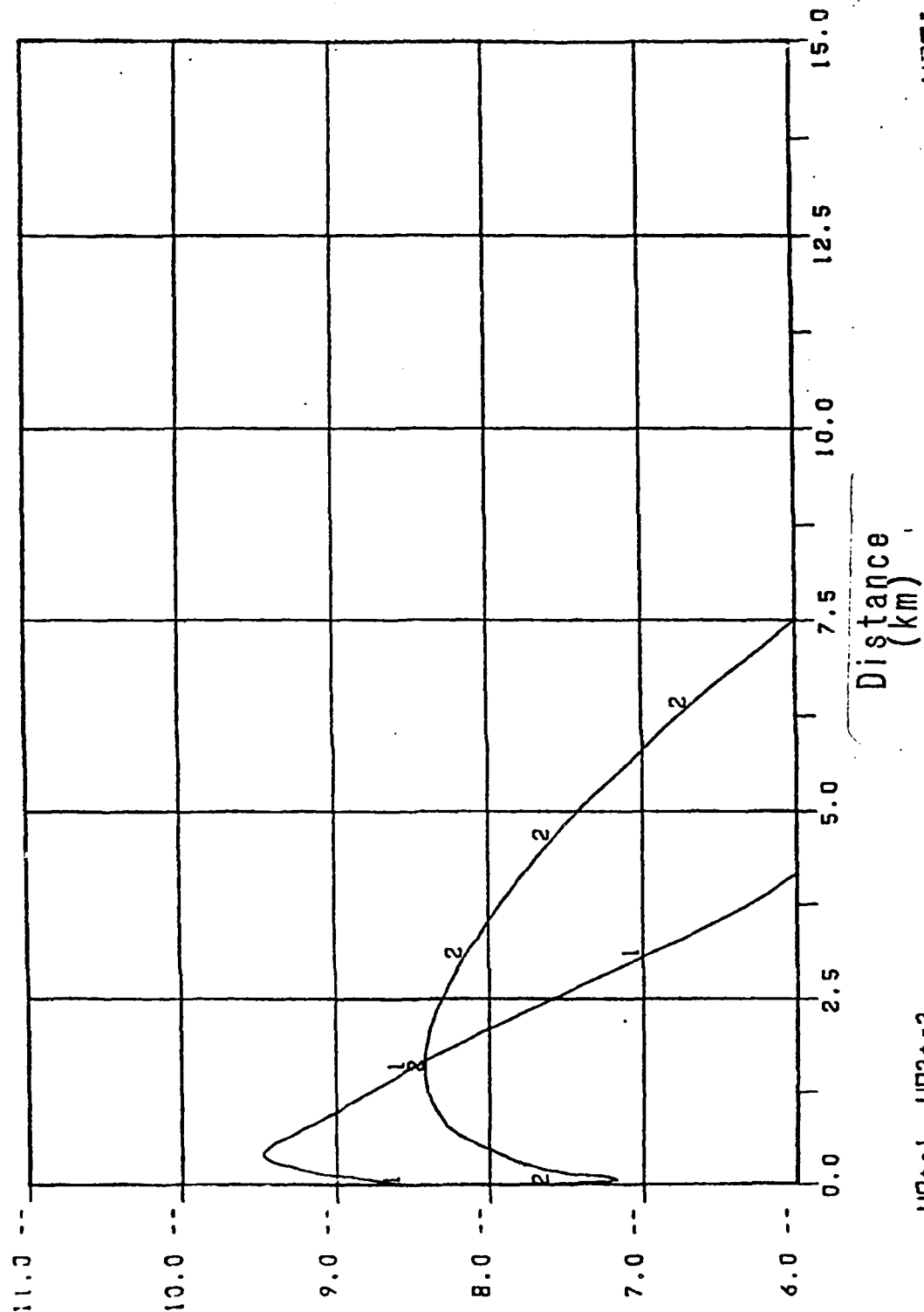
URR-I

M-25

01/04/83

TIME 5.01E 00

ION. NUM. DENS.LOG10



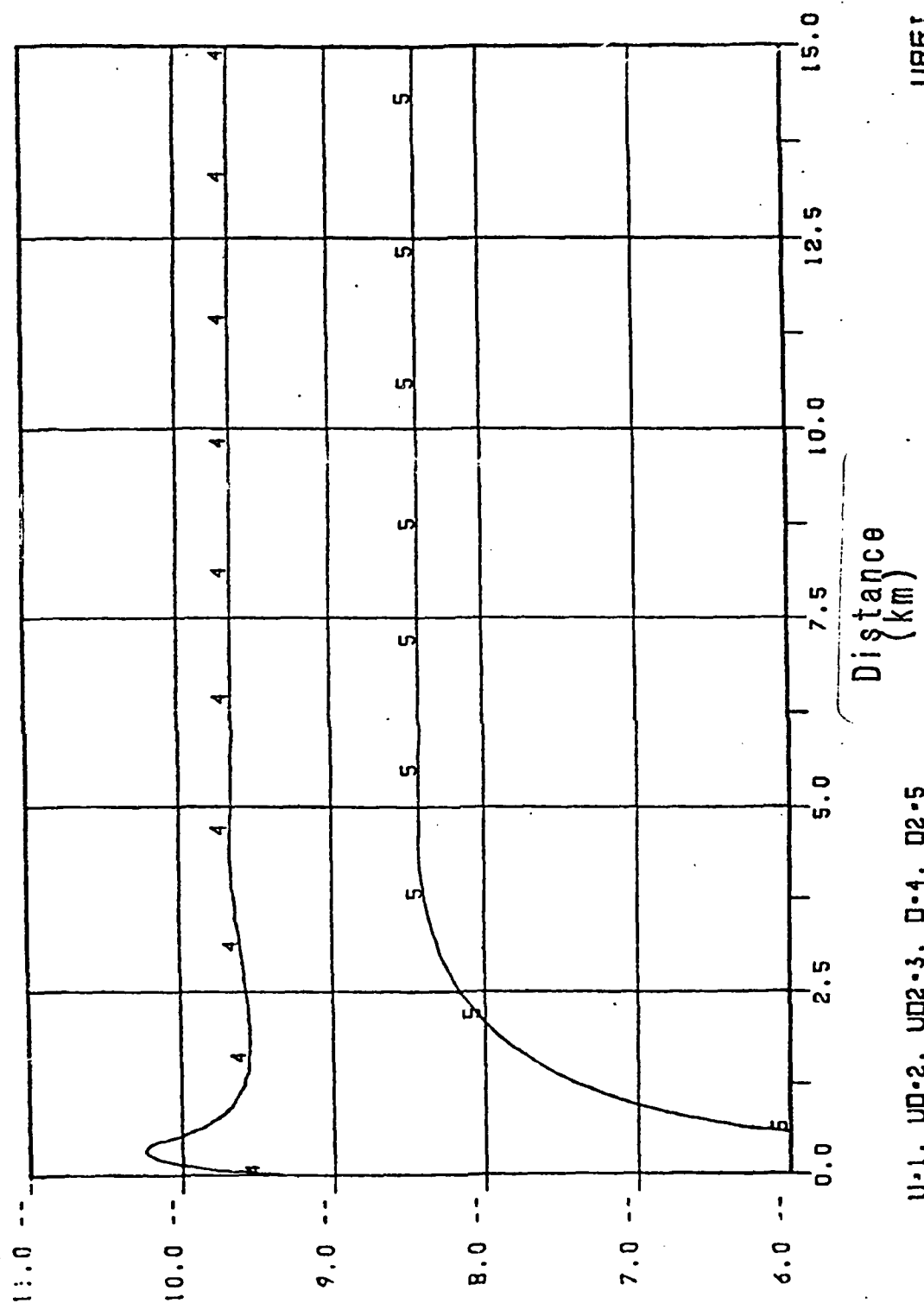
UD+1. UD2+2

URRI

01/04/83

TIME 5.01E 00

NUT. NUM. DENS.LOG10



URRI

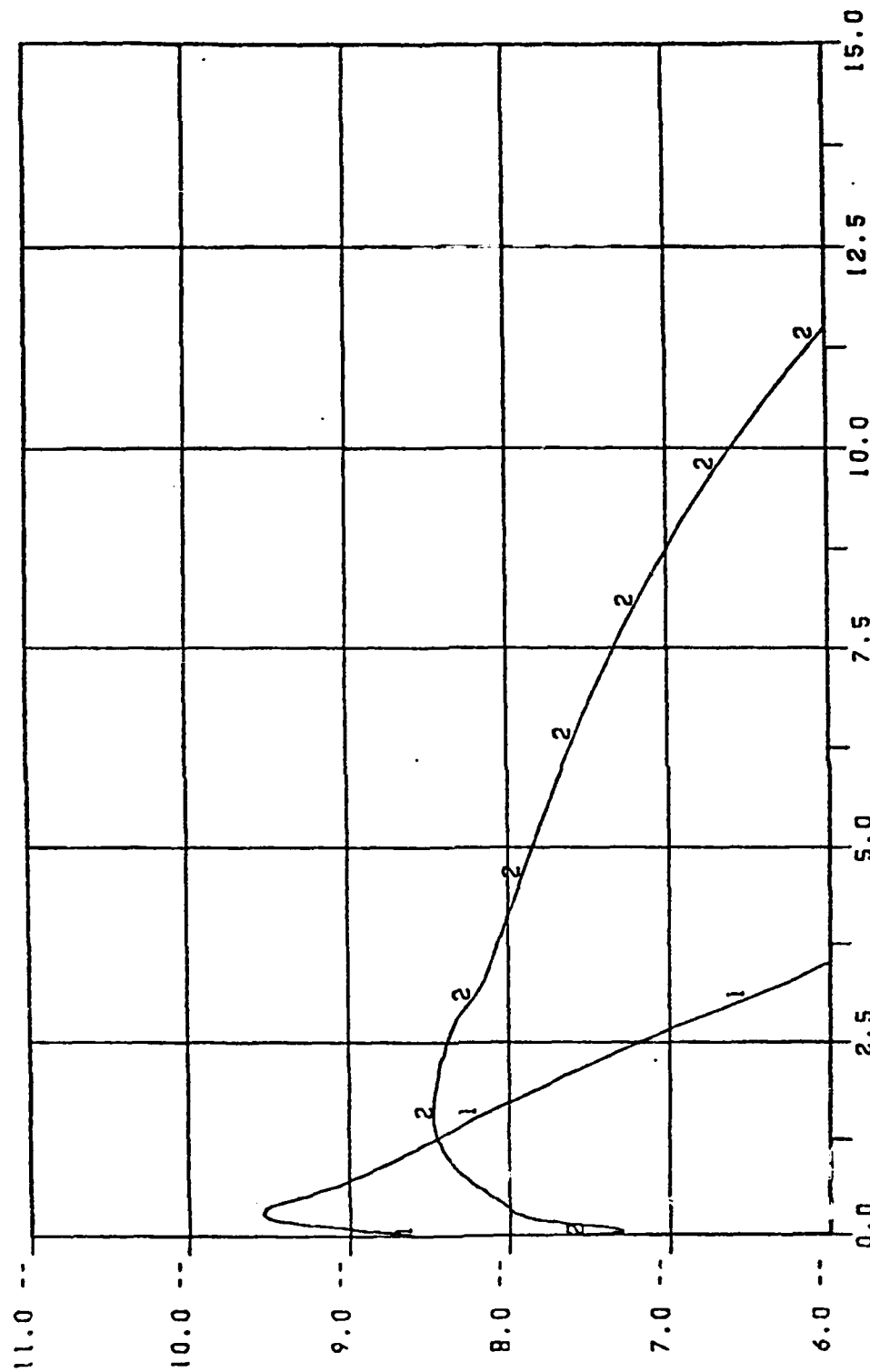
U-1. UD-2. UD2-3. D-4. D2-5

M-27

01/04/83

TIME 1.00E 01

ION. NUM. DENS. LOG10



Distance
(km)

UD0+1, UD2+2

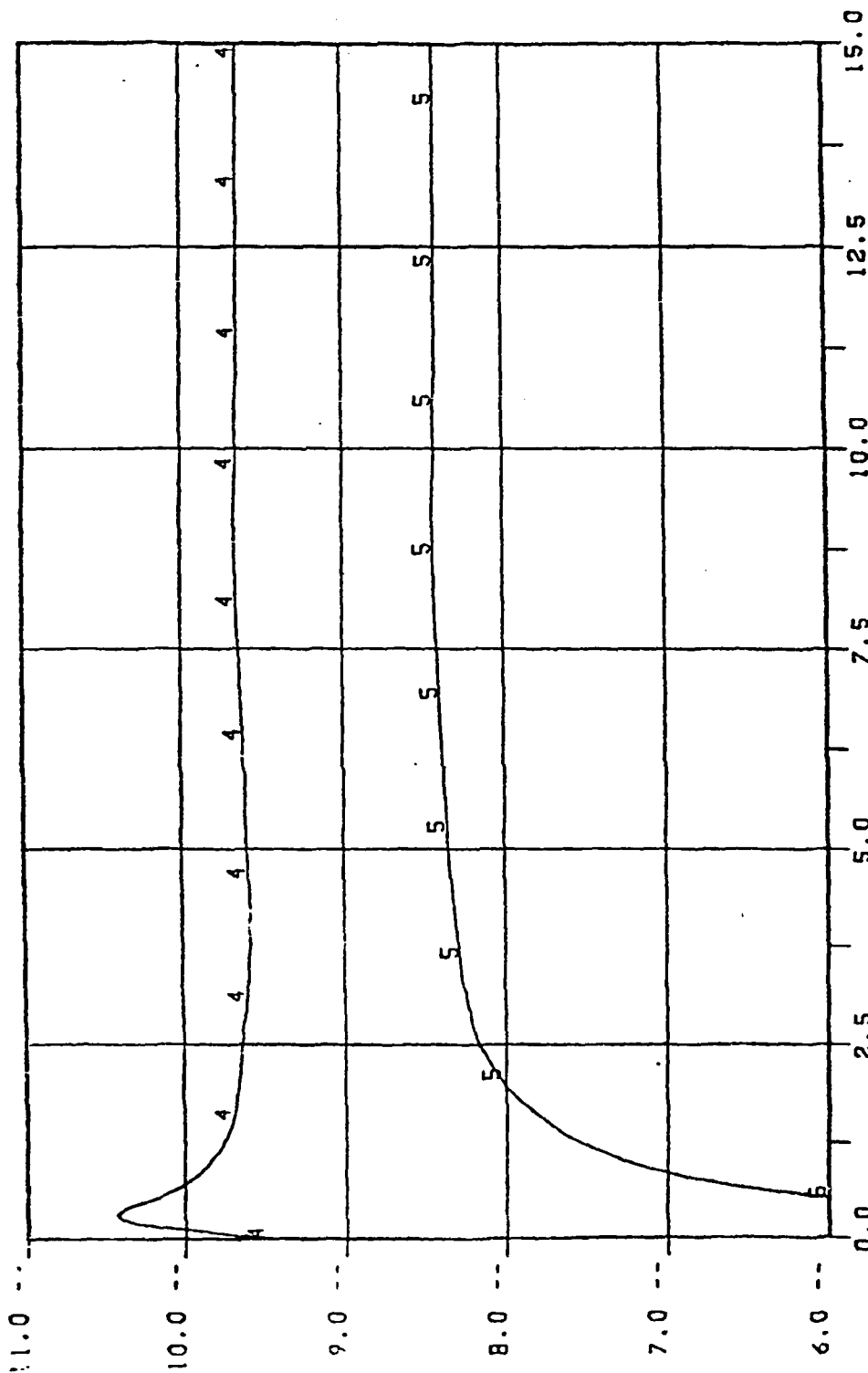
URR1

M-28

01/04/83

TIME 1.00E 01

NUT. NUM. DENS. LOG10

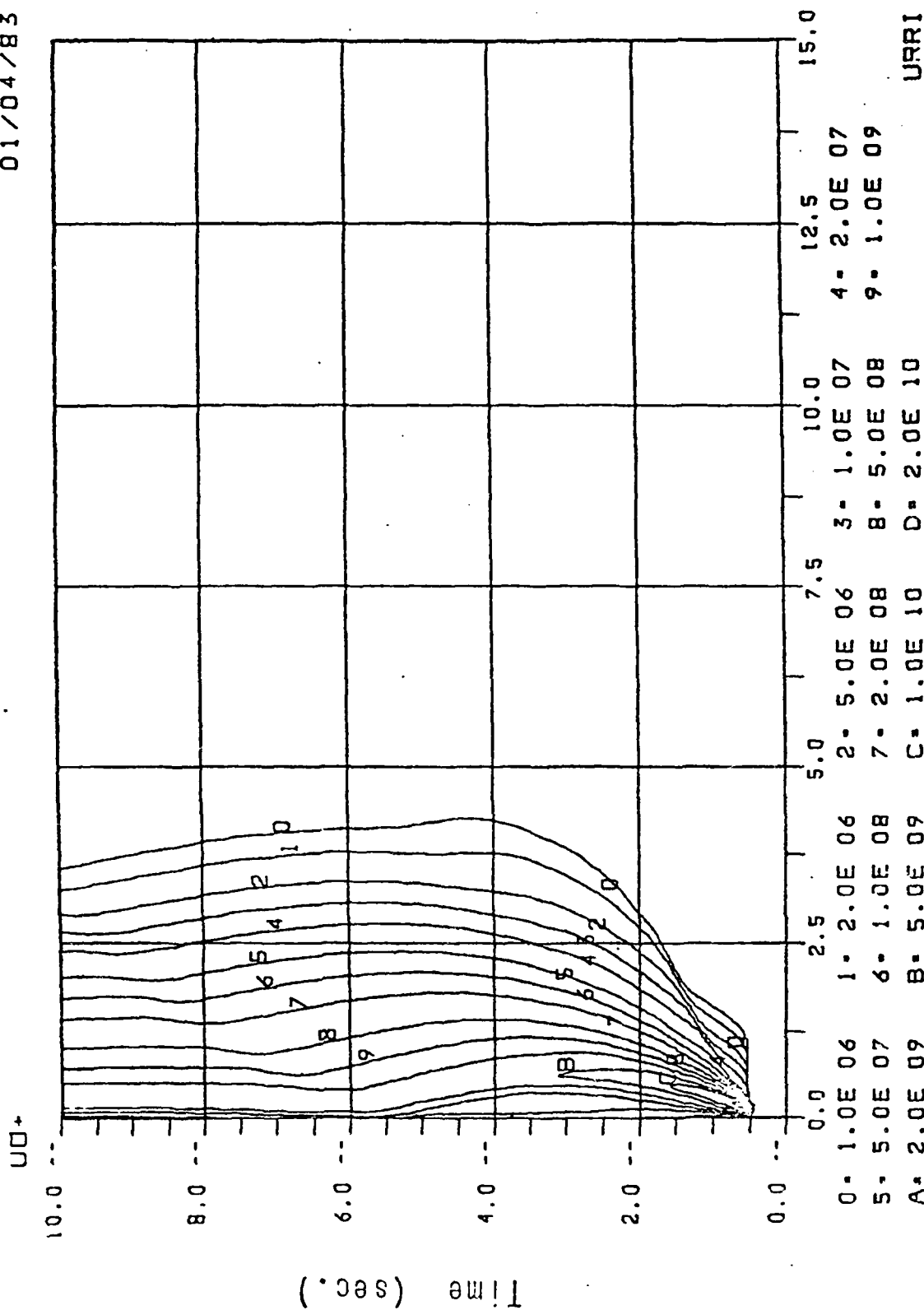


Distance
(km)

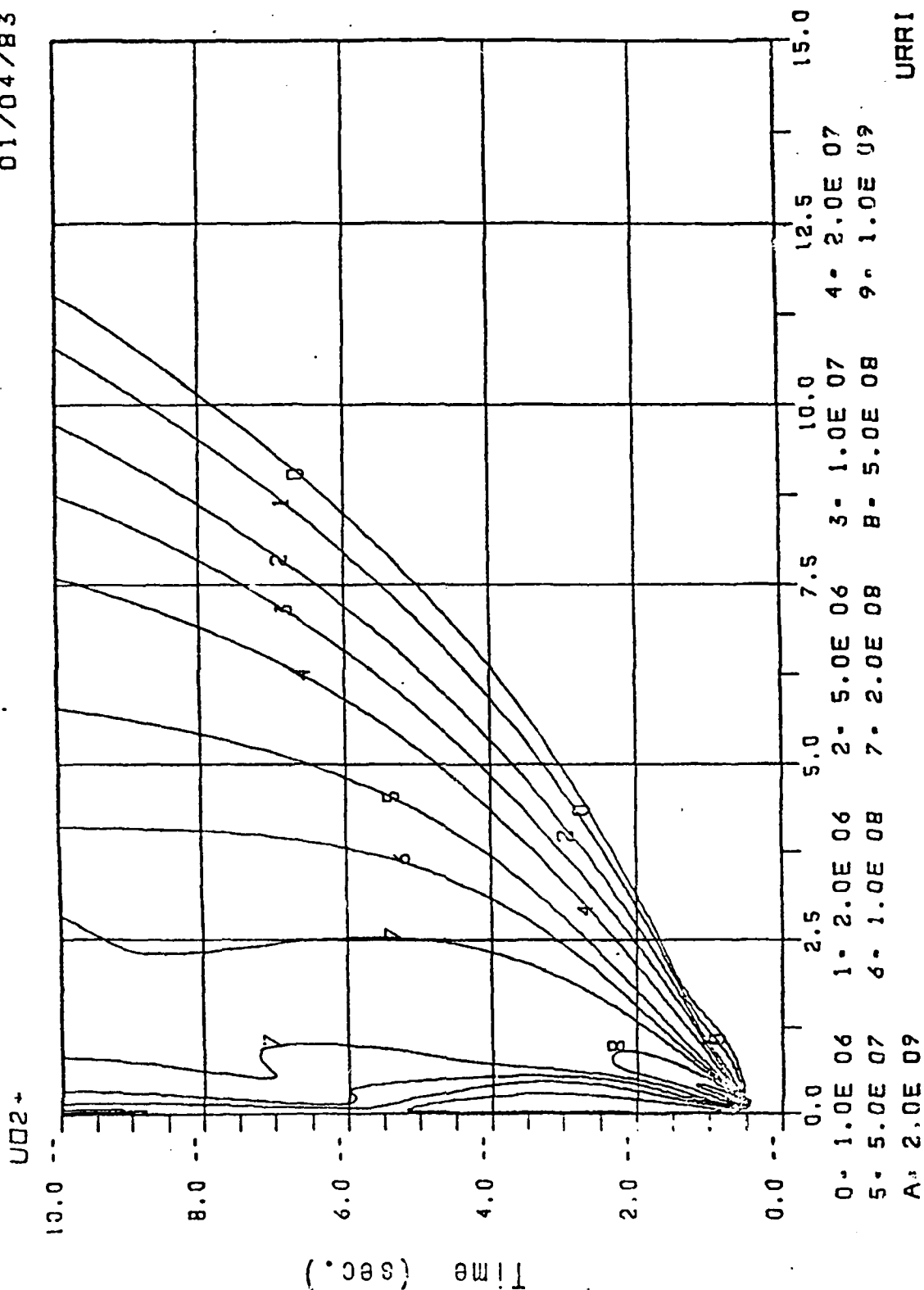
U-1. U0-2. U02-3. 0-4. 02-5

URRI

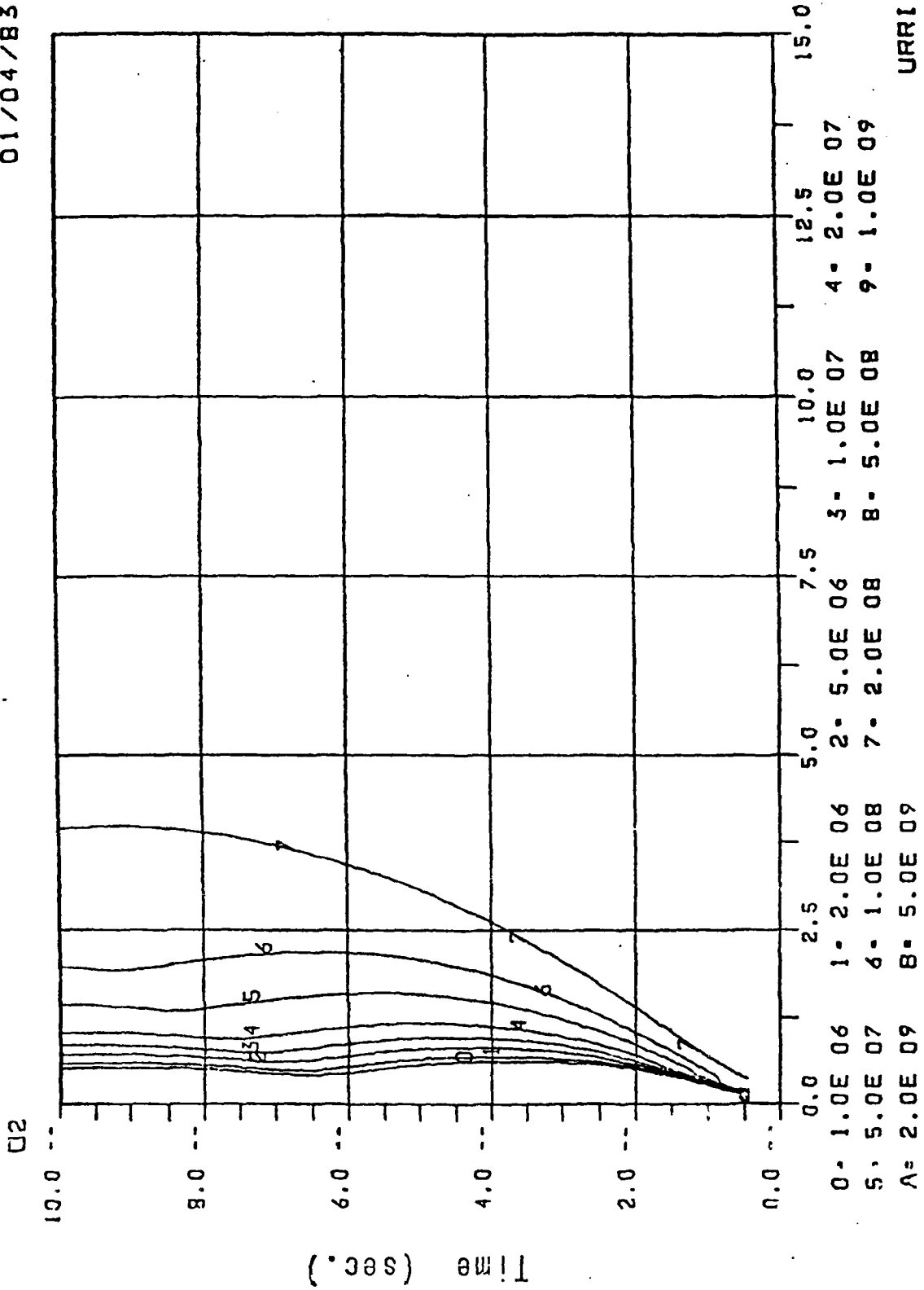
01/04/83



01/04/B3



01/04/B3

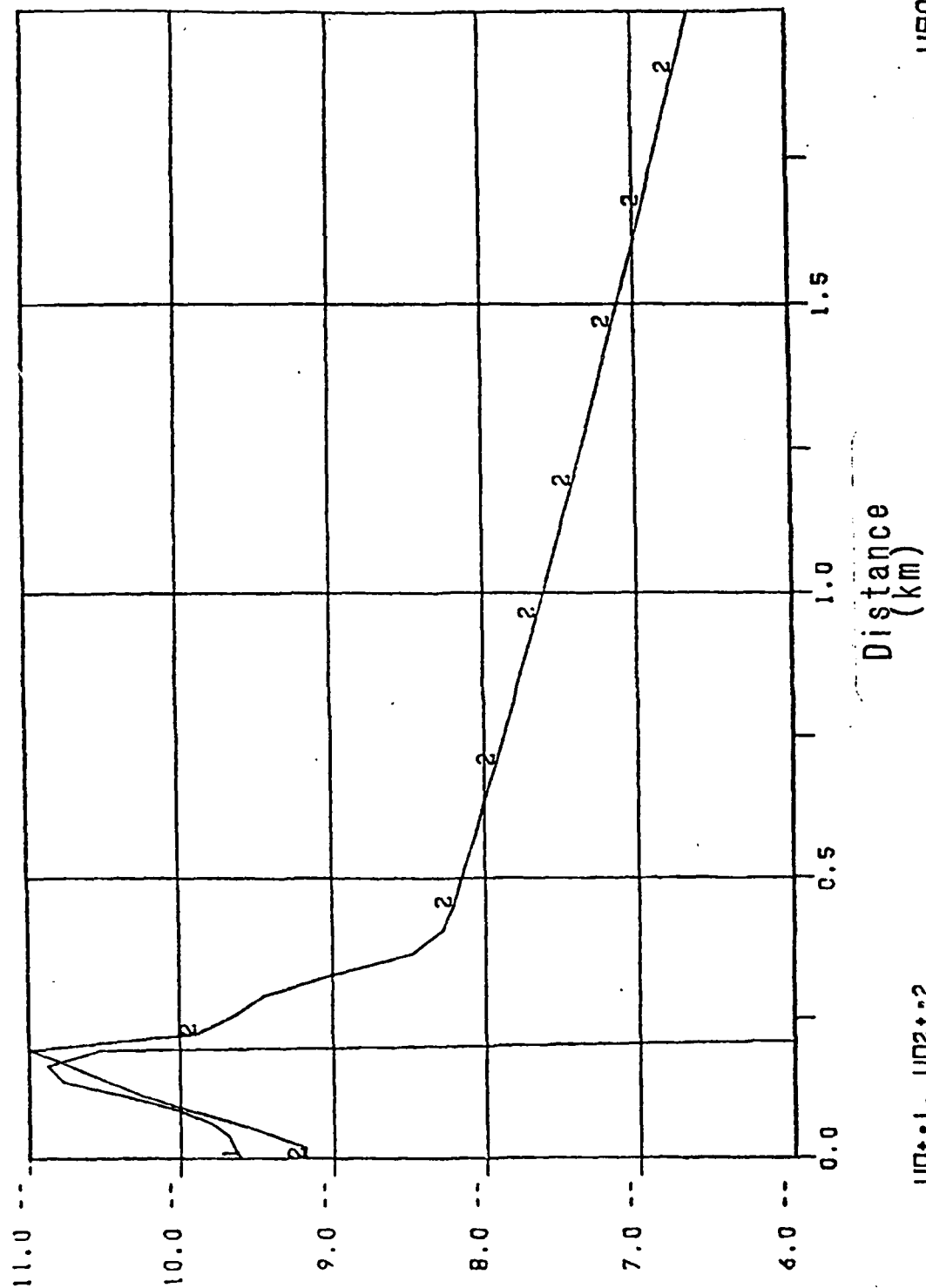


M-32

01/05/83

TIME 1.00E 01

ION. NUM. DENS. LOG10



UD+1, UD2+2

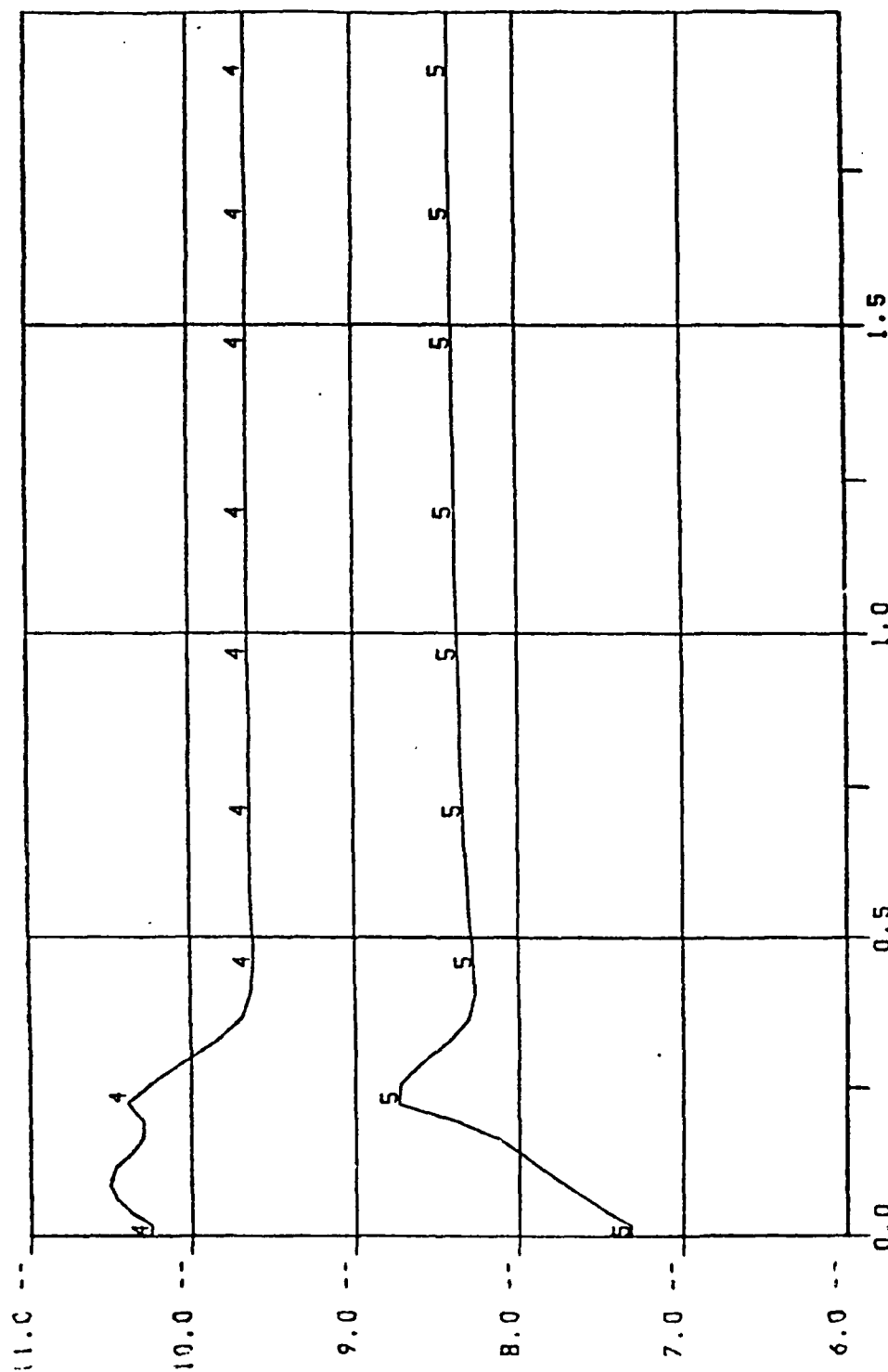
UR03

M-33

01/05/83

TIME 1.00E 01

NUT. NUM. DENS.LOC10



Distance
(km)

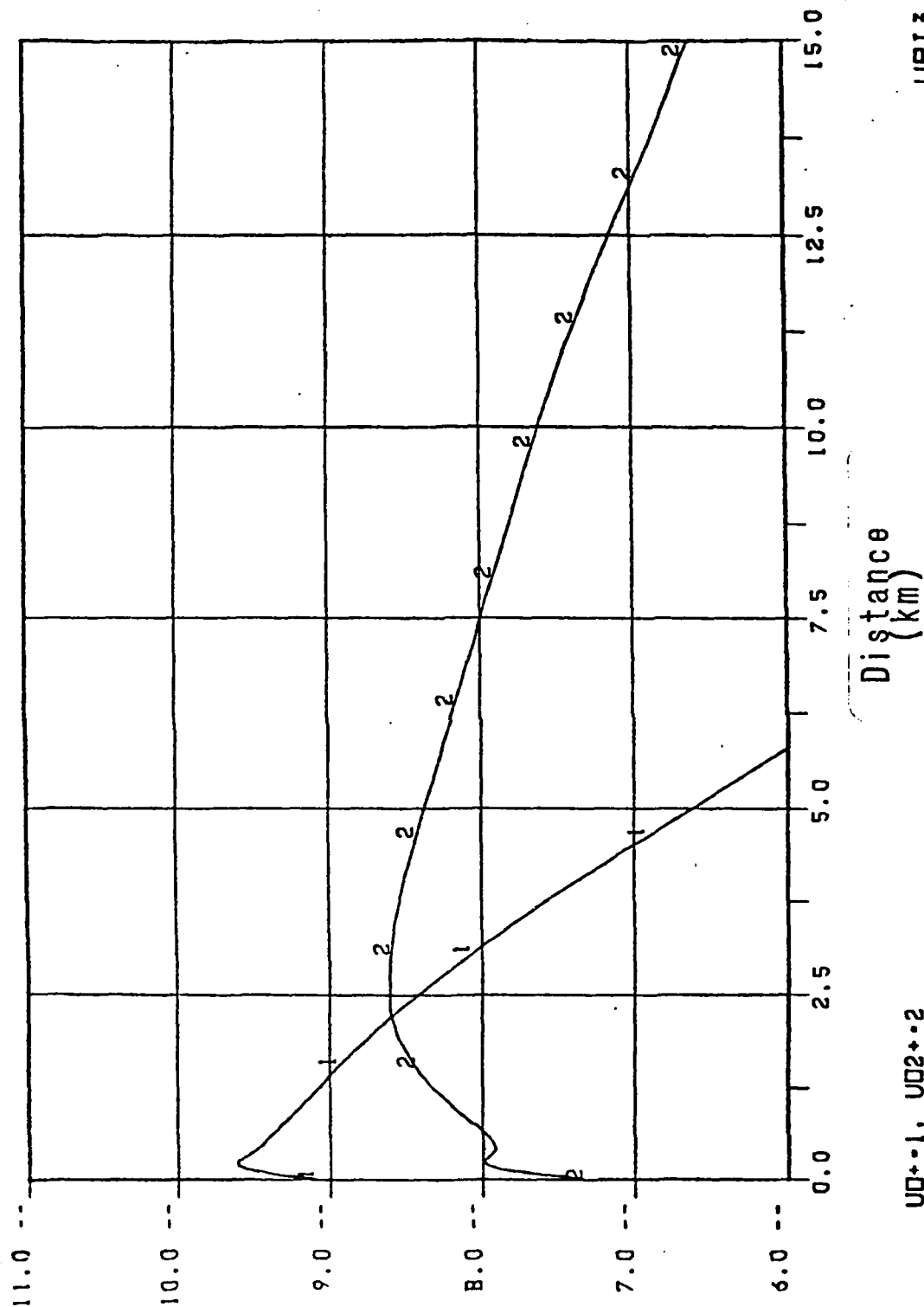
U-1, UD-2, UD2-3, D-4, D2-5

UR03

01/05/83

TIME 1.00E 01

ION. NUM. DENS. LOG10



UR13

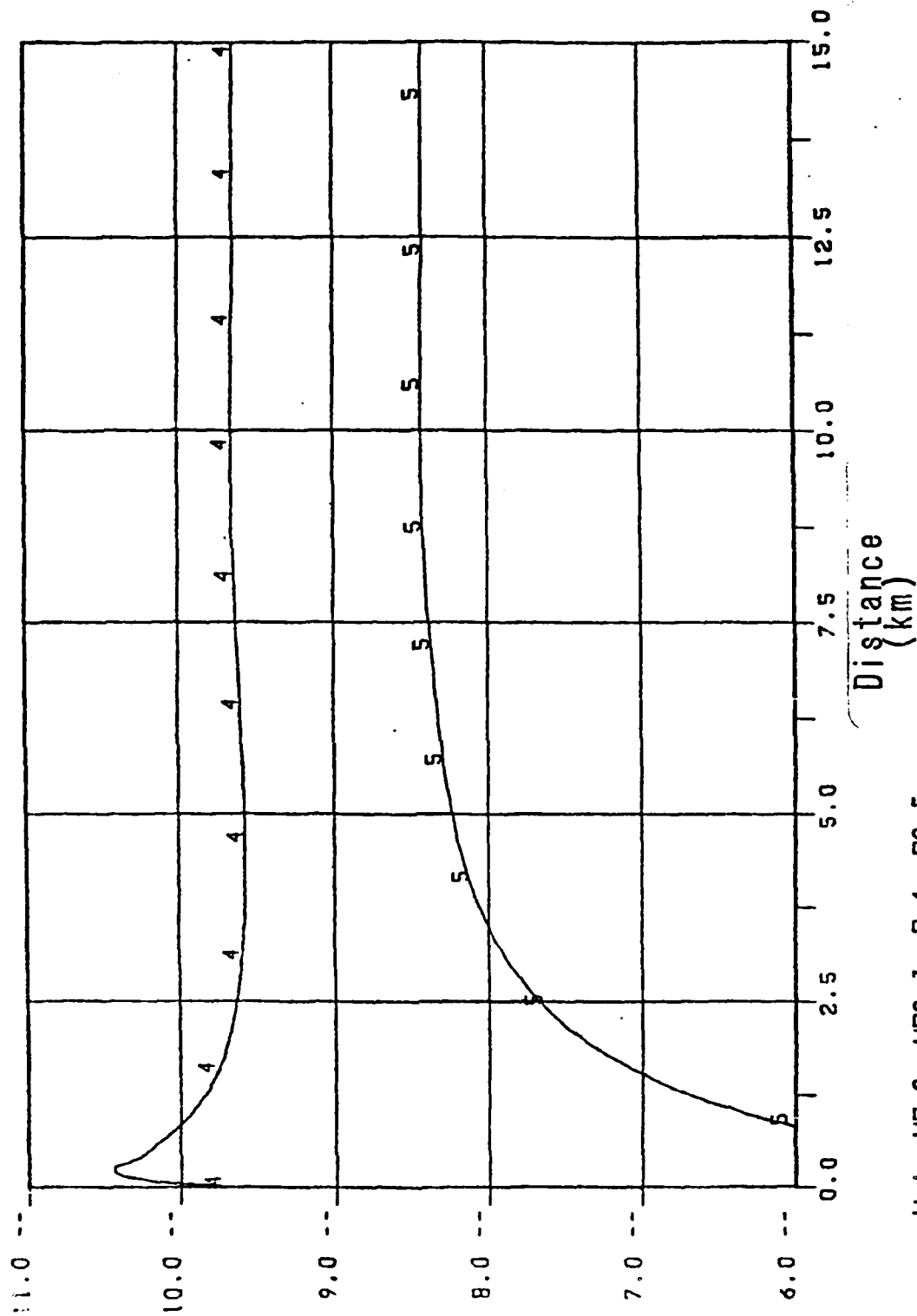
UD+-1, UD2+-2

M-35

01/05/83

TIME 1.00E 01

NUT. NUM. DENS. LOG10



U-1. U0-2. U02-3. 0-4. 02-5

UR13

APPENDIX N
RENORMALIZATION GROUP THEORY OF FRACTALS IN
STRIATING IONOSPHERIC PLASMA CLOUDS

Appendix N

RENORMALIZATION GROUP THEORY OF FRACTALS IN STRIATING
IONOSPHERIC PLASMA CLOUDS

In trying to understand late-time striation phenomena that occur in ionospheric clouds, it has been suggested¹ that convecting plasma clouds, being unstable to the $E \times B$ gradient drift instability, nonlinearly fractalize through multiple bifurcation processes. In an entirely different physical context, fractals have been used by Mandelbrot² and by Frisch, Sulem and Nelkin³ to explain the intermittency in self-similar cascades of fully developed Navier-Stokes turbulence. Loosely speaking, a regular fractal is a contoured structure whose shape remains unchanged locally by a rescaling of spatial scales by a fixed proportion. By its very nature any contour curve in a finite region should have infinite length. It is highly unlikely, however, that the fractal structure of either the striating ionospheric plasma cloud or a fully developed fluid turbulence could be regular. Much more likely fractal structures in both cases are of random irregular character, such that a rescaling of spatial scales reveals a new structure which is statistically similar to the original one but not identical in details. This kind of self-similar structure is closely related to the multiple scale-length structures in statistical systems near critical transition. The mathematical tool required to study critical phenomena has been introduced by Wilson. Known as the renormalization group technique, it was first developed by Gell-Mann and Low to study the scaling symmetry that occurs in Quantum Electrodynamics. The renormalization group addresses the rescaling symmetry of a

statistical system. It therefore describes the symmetry of a random fractal. As a result, it would seem that the renormalization group technique would be the ideal tool to study both self-similar striations in ionospheric plasma clouds and self-similar cascades in fully developed fluid turbulence. Unfortunately, while no work on the application of the renormalization group approach to plasma clouds has been done, attempts to study fluid turbulence using the renormalization group in a dynamical setting have not been very successful. The lack of success in the renormalization group approach to turbulence can be attributed to the fact that the renormalization coefficients appearing in the "renormalization group equation" can only be calculated by perturbation theory. In critical phenomena where the correlation length goes to infinity as the critical transition is approached the renormalization group transformation is performed by integrating over large wavenumber degrees of freedom. Since, in general, the renormalized coupling constants decrease as more and more large wavenumber degrees of freedom are eliminated, the perturbation expansion actually improves with repeat applications of the renormalization group transformation. A similar procedure is also applicable for fully developed fluid turbulence. In fact, the effective coupling usually goes to zero as one goes to larger and larger spatial scales. This property is called infrared stability of the renormalization group transformation and it makes the perturbation treatment essentially exact in the long wavelength, long time limit. From the point of view of inertial range spectral cascades in fluid turbulence, however, one is interested in the physics which governs small spatial scales. Since the effective coupling increases with the decrease of spatial scales, the renormalization group transformation is possibly ultraviolet

divergent, rendering the perturbation expansion approximation totally ineffective. It is unclear at this moment whether a perturbative renormalization group analysis would be useful in the study of small scale fractal behavior of ionospheric plasma clouds at late times.

We shall describe, in the following, a nonperturbative scheme to implement the idea of the renormalization group to study the late-time evolution of striating ionospheric plasma clouds. Our starting point is the Fourier representation of the nonlinear $\underline{E} \times \underline{B}$ equations:

$$\nabla \cdot (N \nabla \phi) = \underline{E}_0 \cdot \nabla N \quad (1)$$

$$\partial_t N + \frac{c}{B} \nabla \cdot [N \hat{z} \times \phi] = 0 \quad (2)$$

where N is the plasma cloud density, ϕ is the polarization potential, $\hat{z} = \frac{B}{B}$ is the direction of the magnetic field, and $\underline{E}_0$ is the applied bias zeroth order electric field, which is assumed to be spatially uniform. It has been shown that the approximation that N is nearly constant on the LHS of Eq. (1) leads to the following equations in Fourier space

$$\partial_t \tilde{N}_k + \frac{icE_0}{BN_c} \sum_q \frac{\hat{z} \cdot [\underline{q} \times (\underline{k}-\underline{q})](\underline{k}-\underline{q}) \cdot \hat{E}_0}{(\underline{k}-\underline{q})^2} \tilde{N}_q \tilde{N}_{k-q} = 0 \quad (3)$$

where $\hat{E}_0 = \underline{E}_0/E_0$, N_c is the constant unperturbed cloud density, and $\tilde{N}_k$ is the Fourier transform of the perturbed density.

Because Eq. (3), being diffusion free, is invariant under dilatations, it is convenient to make the following change of the independent variables $\underline{k}$:

$$\underline{k} = e^s \hat{k} \quad (4)$$

where $\hat{k}$ is a unit vector pointing in the direction of $\underline{k}$ and $s = \ln k$. Eq. (3) now becomes

$$i\partial_t U_{s,\hat{k}} = \sum_{s',\hat{q}} \frac{e^{(s-s')}\hat{z} \cdot [\hat{q} \times (\hat{k} - e^{(s'-s)}\hat{q})] \hat{x} \cdot (\hat{k} - e^{(s'-s)}\hat{q})}{(\sqrt{1+e^{2(s'-s)} - 2\hat{k} \cdot \hat{q} e^{(s'-s)}})^3} U_{s',\hat{q}} U_{s+1,\hat{p}} \quad (5)$$

where

$$U_{s,\hat{k}} \equiv e^s \frac{cE_0}{BN_c} \tilde{N}_{\underline{k}}, \quad s'\hat{q} = \hat{q}, \quad \text{and} \quad \hat{p} = \frac{\hat{k} - e^{(s'-s)}\hat{q}}{\sqrt{1+e^{2(s'-s)} - 2\hat{k} \cdot \hat{q} e^{(s'-s)}}}$$

is a unit vector pointing in the direction of $\hat{k} - e^{(s'-s)}\hat{q}$. Eq. (5) can be further rewritten as

$$i\partial_s U_{s,\hat{k}} = \sum_{\ell,\hat{q}} M(\hat{k},\hat{q}|\ell) U_{s+\ell,\hat{q}} U_{s+\ell} n f(\hat{k} \cdot \hat{q}, \ell), \quad \hat{p}(\hat{k},\hat{q}|\ell) \quad (6)$$

$$\text{where } \ell = s' - s, \quad f(\hat{k} \cdot \hat{q}, \ell) = \sqrt{1+e^{2\ell} - 2\hat{k} \cdot \hat{q} e^{\ell}},$$

$$\hat{p}(\hat{k},\hat{q}|\ell) = \frac{\hat{k} - e^{\ell}\hat{q}}{f(\hat{k} \cdot \hat{q}, \ell)} \quad \text{and}$$

$$M(\hat{k},\hat{q}|\ell) = \frac{e^{\ell}\hat{z} \cdot [\hat{q} \times (\hat{k} - e^{\ell}\hat{q})] \hat{x} \cdot (\hat{k} - e^{\ell}\hat{q})}{[f(\hat{k} \cdot \hat{q}, \ell)]^3} \quad (7)$$

Eq. (6) clearly displays the translational symmetry, as can be seen from the fact that (6) is invariant under the transformation $s \rightarrow s + \text{constant}$. The advantage of going to s

variables is immediately evident, because the dilatation transform $k \rightarrow \lambda k$ is just a translation $s \rightarrow s + \ln \lambda$. Hence the translational symmetry of Eq. (6) simply corresponds to the dilatational symmetry of the original Equations (3). Note that $k \rightarrow \lambda k$ induces a transformation $\underline{N}_k \rightarrow \frac{1}{\lambda} \underline{N}_k$ which leaves $U_{\lambda, k}$ invariant.

Since we are dealing with a $\underline{k}$ -space which is continuous rather than discrete, it is more convenient to work with Fourier density $\underline{N}(\underline{k}, t)$ instead of $\underline{N}_k$ and replace the summation $\sum_{\underline{k}}$ by integration $\int d^2 \underline{k}$. The resulting equation is

$$i \dot{U}_{s, \hat{k}} = \iint m(\hat{k}, \hat{q} | \lambda) U_{s+\lambda, \hat{q}} U_{s+\lambda n f(\hat{k} \cdot \hat{q}, \lambda), \hat{p}(\hat{k}, \hat{q} | \lambda)} d\lambda d\theta_{\hat{q}}, \quad (8)$$

where $m(\hat{k}, \hat{q} | \lambda) = M(\hat{k}, \hat{q} | \lambda) / f^2(\hat{k} \cdot \hat{q}, \lambda)$, $U_{s, \hat{k}} = e^{-2s} U_{s, \hat{k}}$, and $\theta_{\hat{q}}$ is the angle associated with the unit vector $\hat{q}$. Eq. (8) displaces the mathematically desirable property that the mode-coupling is essentially local in s . To see this, we note that $m(\hat{k}, \hat{q} | \lambda)$ has the following asymptotic behavior:

$$m(\hat{k}, \hat{q} | \lambda) = \begin{cases} -\hat{z} \cdot \hat{q} \hat{x} \hat{k} \hat{x} \cdot \hat{q} e^{-3\lambda} \rightarrow 0, & \lambda \rightarrow \infty \\ \hat{z} \cdot \hat{q} \hat{x} \hat{k} \hat{x} \cdot \hat{k} e^{\lambda} \rightarrow 0, & \lambda \rightarrow -\infty. \end{cases} \quad (9)$$

Hence couplings to $|\lambda|$ significantly greater than one are unimportant. However, Eq. (9) does allow direct coupling to very long wavelength modes ($s \rightarrow -\infty$) through the zero of $f(\hat{k} \cdot \hat{q}, \lambda)$, which occurs for $\lambda=0$ and $\hat{k}=\hat{q}$. The reason for the direct long range coupling is clear: the zeroth order uniform electric field E_0 represents a $\underline{k}=0$ ($s \rightarrow -\infty$) energy

source for the $\underline{ExB}$ gradient drift instability whose linear growth rate for short wavelength perturbations is essentially independent of the magnitude of the wavenumber. More generally, any long wavelength electric field perturbation could be considered as a pump for shorter wavelength $\underline{ExB}$ gradient drift instabilities.

We now consider the possibility that the energy cascade from long wavelength to short wavelength, i.e., from $\lambda \rightarrow -\infty$ to $\lambda \rightarrow \infty$, can result in a steady state flow with $\partial_t = 0$, which, for Eq. (8), becomes

$$\int m(\hat{k}, \hat{q} | \lambda) U_{s+\lambda, \hat{q}} U_{s+\lambda n f(\hat{k} \cdot \hat{q}, \lambda), \hat{p}(\hat{k}, \hat{q} | \lambda)} d\lambda d\theta_{\hat{q}} = 0. \quad (10)$$

Without attempting to find the most general solution to the above equation, (which seems impossible in any case). We note that Eq. (10) is no longer scale invariant in s , but instead, has an internal scale λ_s which is of order 1. This opens up the possibility that Eq. (10) admits solutions of the form

$$U_{s, \hat{k}} = e^{\alpha s} p(s, \hat{k}), \quad (11)$$

where $p(s, \hat{k})$ is a periodic function of s with period $T \sim 1$, i.e., $p(s+T, \hat{k}) = p(s, \hat{k})$. Transforming back to k -space representation, we find that $\phi(k, \hat{k}) \equiv \frac{BN}{cE_0} \tilde{N}(\underline{k})$ satisfies the relationship

$$\phi(\lambda k, \hat{k}) = \lambda^{\alpha-3} \phi(k, \hat{k}), \quad (12)$$

where $\lambda \equiv e^T$. In other words, the function $\lambda^{(3-\alpha)m} \psi(\lambda^m k, \hat{k}) = \psi(k, \hat{k})$ is independent of m for any integer m . Hence, $\psi(k, \hat{k})$ is self affine with respect to the variable k .

The function $\psi(k, \hat{k})$ is closely related to a class of functions known as Weierstrass functions. In fact, if we take $\psi(k, \hat{k})$ to be

$$\psi = \sum_{n=0}^{\infty} \lambda^{(\alpha-3)n} \delta(k - \lambda^n k_0) f(\hat{k}), \quad (13)$$

then the Fourier transform of ψ becomes

$$\hat{\psi}(\underline{x}) = \sum_{n=0}^{\infty} \lambda^{(\alpha-3)n} \int e^{i\lambda^n k_0 \hat{k} \cdot \underline{x}} f(\hat{k}) d\theta_{\hat{k}}. \quad (14)$$

Eq. (14) is a two-dimensional generalization of Weierstrass functions, it has the property of being continuous but nowhere differentiable provided that $2 < \alpha < 3$. Physically, $\psi(x)$ describes a fractal structure.

The function $\psi(k, \hat{k})$ given by Eq. (13) only serves as an example of the kinds of functions satisfying Eq. (12). It does not, however, solve the time-independent Equation (10). It might be possible to find a solution to Eq. (10) which is a slightly modified form of Eq. (13). (For example, by making k_0 a function of $\hat{k}$). This possibility is still under investigation. It should be pointed out that we have found model dynamical equations in one dimension which admit of time-independent solutions of the Weierstrass variety.

The existence of time-independent fractalized solutions to the scale-free Equation (3) can obviously explain the phenomena of shape freezing. However, time-independent solutions are interesting only if they can be approached time-asymptotically. The simplest way to include time dependence is to take advantage of the translating invariance of Eq. (8) and assume that $U_{s,\hat{k}}$ is of the form

$$U_{s,\hat{k}} = U(t - \beta s, \hat{k}). \quad (15)$$

(15) describes a travelling wave in s-space. The corresponding equation is

$$-i\beta \frac{\partial}{\partial s} U = \int m(\hat{k}, \hat{q} | \ell) U(t - \beta s - \beta \ell, \hat{q}) U[t - \beta s - \beta \ell n f(\hat{k} \cdot \hat{q}, \ell), \hat{p}(\hat{k}, \hat{q} | \ell)] d\ell d\theta \hat{q}. \quad (16)$$

To see the physical meaning of (15), we note that in k-space, the perturbed charge density $\tilde{N}_{\underline{k}}$ has the form

$$\tilde{N}_{\underline{k}} \propto k^{-3} U(t - \beta \ln k, \hat{k}). \quad (17)$$

If we define k_c to be the characteristic wavenumber for short-wavelength substructures, then we see that k_c evolves in time according to

$$\frac{dk_c}{dt} = \frac{k_c}{\beta}. \quad (18)$$

Thus, the cloud develops a spectrum of finger sizes which become smaller and smaller as time progresses, and it does this at an increasing pace. This cannot, however, happen in a finite time, as predicted by Cornwall. Instead, it still takes an infinite time to approach the infinitely detailed bifurcation structure.

In a physically more realistic situation, one will have to take into account the effects due to particle discreteness, large-scale deformation of plasma clouds, and other random disturbing forces, all of which could be regarded as noises, as well as viscous, diffusive effects. The result will be a statistical dynamic process toward a statistical time-asymptotic equilibrium rather than the more coherent processes alluded to before. A simple model which takes into account both the random noise (force) and the viscous diffusion is

$$\partial_t \tilde{N}_k + \nu_0 k^2 \tilde{N}_k + i \lambda_0 \sum_q \frac{\hat{z} \cdot [\underline{q} \times (\underline{k} - \underline{q})] (\underline{k} - \underline{q}) \cdot \hat{x}}{(\underline{k} - \underline{q})^2} \tilde{N}_q \tilde{N}_{k-q} = \underline{f}_k \quad (19)$$

where ν_0 is the (unrenormalized) viscosity, $\underline{f}_k$ is the random force function, and λ_0 is a parameter which measures the strength of the nonlinear coupling. λ_0 will eventually be set to unity. For simplicity we will assume that $\underline{f}_k$ obeys Gaussian statistics and consists predominantly of long wavelength fluctuations. Rewritten in $s \equiv \ln k$ space, Eq. (19) becomes

$$i(\partial_t + \nu_0 e^{2s}) U_{s, \hat{k}} = \lambda_0 \int \int m(\hat{k}, \hat{q} | \ell) U_{s+\ell, \hat{q}} U_{s+\ell, \hat{k}-\hat{q}} n f(\hat{k} \cdot \hat{q}, \ell) \hat{p}(\hat{k}, \hat{q}, | \ell) d\ell d\theta_q^\wedge + \underline{F}_{s, \hat{k}}, \quad (20)$$

where $\underline{F}_{s, \hat{k}} = e^{-3s} \underline{f}_{\hat{k}}$ and is accordingly even more peaked at small $\underline{k}$ ($s \rightarrow -\infty$) than $\underline{f}_k$. The presence of ν_0 and $\underline{F}_{s, \hat{k}}$ destroys the translational invariance enjoyed by Eq. (8). For our interest, we will consider the limiting behavior of $\nu_0 \rightarrow 0$ and $\underline{F}_{s, \hat{k}} \rightarrow 0$ for any finite s , corresponding to the inertial range having an infinite extent. In the inertial

range, the translational invariance of the equation is preserved.

As explained by Ma and Mazenko, the dynamic renormalization group procedure consists of basically two parts. First, the modes $U_{s,k}^{>\Delta}$ such that $\Lambda - \Delta < s < \Lambda$ (Δ being a finite number) are eliminated from Eq. (20) by formally solving the equations for $U_{s,k}^{>\Delta}$ as a power series in λ_0 , whose coefficients are functions of the remaining modes $U_{s,k}^{<\Delta}$ as well as the forcing function $F_{s,k}^{>\Delta}$. These formal solutions are then inserted into the equations for $U_{s,k}^{<\Delta}$ to eliminate their explicit dependence on $U_{s,k}^{>\Delta}$. Subsequently, the reduced set of equations is averaged over the part of the force $F_{s,k}^{>\Delta}$ that acts in the shell $\Lambda - \Delta < s < \Lambda$, partially renormalizing the coefficients which enter the reduced equations of motion. The remaining fluctuating part is added to the noise forcing function $F_{s,k}^{<\Delta}$, again partially renormalizing the coefficients characterizing its spectrum. Second, a rescaling of space (translation in s), time, viscosity, remaining $U_{s,k}^{<\Delta}$ and forces is done to make the new set of equations look as much like the original equations as possible. The resulting equations are differential equations for the renormalized coefficients as functions of Δ which can be integrated to eliminate a larger and larger part of the high- k degrees of freedom.

As has been demonstrated by Forster, Nelson and Stephen, the above procedure works well for the large-distance, long-time behavior of velocity correlations generated by the Navier-Stokes equations with random stirring force. However, running the recursion equations backwards to attempt to study large wave-number properties leads to an

increase in the renormalized mode coupling strength $\lambda(\Delta)$, invalidating weak-coupling perturbation schemes which are crucial for the success of the dynamic renormalization group procedure.

In our case, we note that the random stirring force acts primarily in the small wavenumber regime, whereas the viscous diffusion term tends to smear out the large wavenumber modes. Furthermore, the perturbative renormalization theories of the DIA-type tend to support the physical intuition that short wavelength fluctuations contribute to renormalization of viscous diffusion, while long wavelength disturbance generates random Galilean transformation as well as adiabatic pumping of energy to shorter wavelength modes. It follows from these intuitive ideas that one should introduce both an upper wavelength and a lower wavelength cutoff, instead of eliminating only the high- k spectrum, as in Ma and Mazenko's original formalism. Specifically, let us introduce Λ and Δ such that the s -space is divided into three parts:

$$s < \Lambda - \frac{\Delta}{2}, \quad \Lambda - \frac{\Delta}{2} < s < \Lambda + \frac{\Delta}{2}, \quad s > \Lambda + \frac{\Delta}{2}. \quad (21)$$

Likewise we separate $U_{s,k}$ into three corresponding groups: $U_{s,k}^{<}$, ($s < \Lambda - \frac{\Delta}{2}$); $U_{s,k}^{>}$, ($s > \Lambda + \frac{\Delta}{2}$) and $U_{s,k}$, ($\Lambda - \frac{\Delta}{2} < s < \Lambda + \frac{\Delta}{2}$). $U_{s,k}^{<}$ and $U_{s,k}^{>}$ are subsequently eliminated following the procedure outlined by Ma and Mazenko. By making $\Delta \gg 1$, only modes with wavelength much larger or much smaller than the wavelength of interest are eliminated. As a result, the elimination of $U_{s,k}^{<}$ and $U_{s,k}^{>}$ can be done separately and independently, since the interaction between $U_{s,k}^{<}$ and $U_{s,k}^{>}$ can be expected to be small. As a further approximation, we will

assume that only the mean squared spectrum of $U_{s,k}^{>\wedge}$ and the fluctuating part of $U_{s,k}^{<\wedge}$ contributes to renormalization of viscous diffusion and the fluctuating forcing function, respectively. Corrections to the above approximation can be evaluated in a perturbative manner.

One obvious advantage of the approach taken here over the popular DIA scheme is that only partial renormalization of the random stirring force together with viscous diffusion is performed, rendering low-order perturbation expansion more accurate. In addition, the absence, in lowest order, of contributions to viscous diffusion from modes with k smaller than the k retained, as well as from modes with comparable k , removes the long-standing difficulty associated with the non-diffusive random-Galilean - convective effects due to long wavelength fluctuation. The effects due to random - Galilean - convection, as well as adiabatic pumping, are restored by renormalizing the fluctuating force through the nonlinear coupling due to the coupling singularity at $f(\hat{k} \cdot \hat{q} | l) = 0$.

The renormalization group procedure just outlined gives rise to two compatible sets of Pfaffian differential equations in Δ and Λ , respectively. Their properties are still under investigation.

Finally, it may be possible to get around the technical difficulty inherent in the semi-perturbative approach we have just described by solving Eq. (20) directly in the viscous-free limit using the following numerical scheme: First, we assume that $\underline{F}_{s,k}^{\wedge}$ is only important at $s \rightarrow -\infty$ and can hence be treated as simply a boundary condition at $s \rightarrow -\infty$; likewise the diffusion term v_0 can be treated

as an energy sink at $s \rightarrow \infty$. Second, we can take advantage of the translational symmetry to assume that, apart from some simple rescaling, different s modes essentially have the same statistical properties. Thus we can solve the Equations (20) in the range $\Lambda - \frac{\Delta}{2} < s < \Lambda + \frac{\Delta}{2}$ where Δ is sufficiently large that there is hardly any statistical correlation among modes separated by more than Δ in the s space. Any modes needed in Eq. (20) which lie outside of the range defined above can be approximated by the following "pull-back" formula:

$$U_{s,k}^{\wedge} = e^{\pm \alpha \Delta} U_{s \pm \Delta, k}^{\wedge}. \quad (22)$$

α is an eigenvalue determined in the following way: pick an arbitrary α and run the numerical code until $U_{s,k}^{\wedge}$ settles down to a stable-temporal behavior. If $U_{s,k}^{\wedge}$ is growing steadily in time change α in the direction so as to reduce the growth rate to zero; likewise, for damped modes. When $U_{s,k}^{\wedge}$ reaches a steady statistical state, the corresponding α is the eigenvalue sought. To relate α to the power spectrum of $\tilde{N}(k)$, we note that Eq. (22) corresponds to

$$\langle |\tilde{N}(k)|^2 \rangle \propto k^{2(\alpha-3)}. \quad (23)$$

From the argument immediately following Eq. (14), it is expected that α will be between 2 and 3, making the spectrum somewhat sharper than k^{-2} .

REFERENCES

1. M. Cornwall, "Studies of the effects of striations on radio communications", JASON Technical Report JSR-81-31, V-1, Dec. (1981).
2. B. Mandelbrot, In "Turbulence and Navier-Stokes equations", (ed. R. Temam). Lecture Notes in Math., Vol. 565, p. 121, Springer, (1976).
3. U. Frisch, Pierre-Louis Sulem and M. Nelkin, "A simple dynamical model of intermittent fully developed turbulence", J. Fluid Mech., Vol. 87 part 4, pp. 719-736, (1977).
4. S. K. Ma and G. F. Mazenko, "Critical dynamics of ferromagnets in $6-\epsilon$ dimensions: General discussion and detailed calculation". Phys. Rev. B., Vol. 11, p. 4077, (1975).
5. D. Forster, D. Nelson, and M. J. Stephen, Phys. Rev. A., 16, 732, (1977).

APPENDIX O

ELF CURRENT GENERATION IN THE IONOSPHERE

ELF Current Generation in the Ionosphere

George Schmidt

Abstract

Two mechanisms for current generation in the ionosphere are investigated. In the first, momentum of waves generated by a land based antenna are absorbed by electrons. In the second the magnetic moment of electrons is enhanced by cyclotron resonance in the polar region, giving rise to currents due to magnetic mirror effects. The two effects are additive with the first one dominating in the highly collisional ionosphere. Order of magnitude estimates indicate that with a land based antenna of 100 MW power, significant currents can be generated.

Introduction

In order to produce ELF electromagnetic waves for communication, the use of an ionospheric antenna has been suggested. Various mechanisms for low frequency current generation in the ionosphere have been proposed, in particular parametric excitation of such waves from a land based antenna structure¹, and electrojet current modulation².

Here we wish to point out two other possible avenues to reach the same goal.

1) A land based antenna radiating into the ionosphere produces waves that can easily be absorbed by ionospheric plasma electrons. Regardless of the absorption mechanism, the momentum of photons is converted into the momentum of electrons giving rise to a current³. If the wave is pulsed this gives rise to current pulses, that can be used as an antenna. If the wave is reflected rather than absorbed a factor of two is gained in the current.

2) The earth's magnetic field forms a magnetic mirror. If the energy of the electromagnetic wave in the polar region is absorbed by electrons via cyclotron resonance, the magnetic moment of these electrons is increased, resulting in increased acceleration of these particles along the magnetic field toward the equatorial plane. Now the background magnetic field provides the source of momentum to produce the desired current.

Momentum Absorption

Consider electromagnetic waves propagating along the earth's magnetic field carrying energy ϵ , momentum $p = \epsilon k / \omega$. For simplicity we take $\omega / k = c$,

so $p = \epsilon/c$. This momentum is absorbed by a cylindrical column of plasma, cross sectional area A , length L . If the free electron density is n , the current generated is

$$I = - enA\delta v \quad (1)$$

where δv is the velocity increment of the electrons due to photon absorption.

From momentum conservation

$$\epsilon/c = ALn\delta v \quad (2)$$

Substituting δv from (2) into (1) results in

$$IL = - \frac{e\epsilon}{mc} = - 586 \epsilon \quad (3)$$

Note that (3) does not depend on the size of the absorption volume, electron density, or any other details of the absorber. A more elaborate calculation, integrating over a nonuniform plasma yields the same result.

If the wave, instead of being absorbed, is reflected, the electrons take up twice the wave momentum $p = 2\epsilon/c$, and one has

$$IL = - 1172 \epsilon \quad (4)$$

So far we have not considered the interaction of the electrons with the environment. The effect of electron-neutral, electron-ion collisions is easy to account for. If the collision frequency is ν , the effective acceleration time is ν^{-1} , so in (3) and (4) $\epsilon = p\nu^{-1}$, where p is the radiated power. Since

ν is sensitively dependent on light over the ground, this results in a strong dependence on the number of Amp meters generated. Good results can be obtained above an altitude of 100 kms, where $\nu < 10^4 \text{ sec}^{-1}$. With a power of 100 MV, $\nu = 10^4 \text{ sec}^{-1}$, one obtains from (3) $IL = - 5,86 \cdot 10^6 \text{ Am}$, which is quite respectable. The drawback is the higher altitude than the one used in the electrojet generation scheme.

The electromagnetic interaction with the background magnetoplasma is more difficult to calculate. The push on the electrons due to photon absorption is equivalent to an electromotive force, driving the current. One expects that this current will close outside the absorption region, creating a dipole type current configuration. So the problem boils down to the investigation of the following model.

An electromotive force is generated in a closed region. Calculate the current pattern in a tensor dielectric with $\nabla \cdot \vec{J} = 0$ ($\vec{J}$ is the current density). This calls for a numerical calculation, similar to the one by Chang and Papadopoulos² for the electrojet modulation problem.

The equivalent electric field E_0 , due to power absorption is easily estimated

$$E_0 \approx \frac{p}{c\nu v} \quad (5)$$

where p is the power absorbed in volume v .

Mirror Acceleration

Consider now the situation where an electron in the polar region

picks up the energy $\delta W_{\perp}$, perpendicular to the magnetic field B by cyclotron acceleration. This results in an increase of its magnetic moment $\delta \mu_m = \delta W_{\perp} B^{-1}$. In consequence there will be an incremental acceleration along the B field $\delta \dot{v}_{||} = -\frac{1}{m} \delta \mu_m \frac{dB}{dx}$, where B is the coordinate along the magnetic field. The particle (without interaction with others) arrives at the equatorial plane with the incremental velocity

$$\delta v_{||} = -\frac{\delta W_{\perp}}{m} \int \frac{1}{B} \frac{dB}{dx} \frac{1}{v_{||}} dx \quad (6)$$

For an order of magnitude estimate we replace $v_{||}$ with some average $\langle v_{||} \rangle$, and integrate to obtain

$$\delta v_{||} \approx \frac{\delta W_{\perp}}{m \langle v_{||} \rangle} \ln \frac{B_1}{B_0} \quad (7)$$

where B_1 is the magnetic field at the polar, B_0 at the equatorial region.

For n electrons per unit volume, this leads to the current in the equatorial plane, using (1) and (7)

$$I \approx -enA \frac{\delta W}{m \langle v_{||} \rangle} \ln \frac{B_1}{B_0} \quad (8)$$

The energy absorbed by all electrons is $\epsilon = LAN\delta W$, to yield

$$IL \approx -\frac{e\epsilon}{m \langle v_{||} \rangle} \ln \frac{B_1}{B_0} \quad (9)$$

Since $\ln \frac{B_1}{B_0}$ is of order unity, it seems that the current generated is $c/\langle v_{||} \rangle$ times larger than the one obtained by momentum absorption, an enormous increment.

The fly in the ointment is our neglect of collisions. In reality electrons collide many times in their passage from the polar to the equatorial region.

A more realistic approach is the calculation of the equivalent electric field from $-eE_1 = -\delta\mu m \frac{dB}{dx}$, or

$$E_1 = \frac{\delta W_1}{Be} \frac{dB}{dx} \quad (10)$$

This is to be compared to the equivalent electric field from momentum absorption.

Since $\frac{p}{v_n} \approx \frac{\delta W}{\tau}$, where τ is the mean collision time, from (5)

$$E_0 \approx \frac{\delta W}{c\tau e} \quad (11)$$

So finally

$$\frac{E_0}{E_1} \approx \frac{L}{c\tau} \approx 7.1 \cdot 10^{-3} v \quad (12)$$

where L is the characteristic scale length of the earth's magnetic field. This ratio is typically larger than one so the first process is more effective.

It should however be emphasized that these two mechanisms are not competing but cooperating. As energy is being absorbed by the electrons to increase their magnetic moment, momentum is simultaneously absorbed.

The next task is a numerical study on collective effects in an anisotropic lossy dielectric to determine the current pattern.

References

1. K. Papadopoulos, Nonlinear Production of ELF Waves by two High Frequency Waves, March 1975 (NRL, Internal Report).
2. C. L. Chang, K. Papadopoulos, ELF Communication by Electrojet Current Modulation, July 1981 (SAI Internal Report).
3. G. Schmidt, Phys. Letters, 74A, 222 (1979).

APPENDIX P

MECHANISMS OF STRATOSPHERIC OZONE TRANSPORT

Mechanisms of Stratospheric Ozone Transport

R. B. ROOD

*Science Applications, Inc.
McLean, Virginia 22102*

M. R. SCHOEBERL

*Geophysical and Plasma Dynamics Branch
Plasma Physics Division*

December 3, 1982



NAVAL RESEARCH LABORATORY
Washington, D.C.

Approved for public release; distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NRL Memorandum Report 4969	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MECHANISMS OF STRATOSPHERIC OZONE TRANSPORT	5. TYPE OF REPORT & PERIOD COVERED Interim report on a continuing NRL problem.	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) R.B. Rood* and M.R. Schoeberl	8. CONTRACT OR GRANT NUMBER(s)	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Research Laboratory Washington, DC 20375	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N-33; RR033-02-42; RA403728; 47-0912-0-2; 47-0887-0-2	
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research NASA Arlington, VA 22217 Washington, DC 20546	12. REPORT DATE December 8, 1982	
	13. NUMBER OF PAGES 176	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES *Present address: Science Applications, Inc., McLean, VA 22102		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Ozone transport Stratospheric warmings		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Two β -plane planetary wave models are used to study ozone transport in the stratosphere. In the first model, ozone transport is calculated for steady, dissipative planetary waves using the Eulerian, Lagrangian-mean, and residual circulations. A Lagrangian model of parcel dynamics is used to interpret planetary wave-photochemistry interaction. In chemically active regions the mean field ozone changes are found to be significant only where there are large gradients in chemical sources and sinks along particle trajectories. The largest changes in the mean field are found in the lower stratosphere and are due to the Lagrangian-mean advection. (Continues)		

DD FORM 1473
1 JAN 73EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Continued)

In the second model, ozone transport is calculated for the combined diabatic and time dependent planetary wave circulations. Both circulations are instrumental in the formation of the polar spring ozone maximum. The diabatic circulation transports ozone into the lower stratosphere, and planetary waves transport large amounts of ozone northward during sudden and final warmings. Using the transport mechanisms revealed in these models, a scenario is deduced to explain observed large scale ozone transport phenomena.

Contents

Chapter 1: INTRODUCTION TO THE OZONE TRANSPORT PROBLEM

1.1. Introduction.....	1
1.2. Diffusion Models.....	3
1.3. General circulation models.....	5
1.4. Lagrangian-mean models.....	8
1.5. Objectives.....	10

Chapter 2: MODEL DESCRIPTIONS

AND THE OBSERVED OZONE DISTRIBUTION

2.1. Transport model	12
2.2. Chemistry model.....	18
2.3. Dynamic models.....	21
2.3a. Steady state model.....	21
2.3b. Time dependent β -plane model	25
2.4. Observations of ozone.....	28

Chapter 3: TRANSPORT BY STEADY STATE PLANETARY WAVES

3.1. Introduction.....	40
3.2. Constituent dynamics.....	40
3.3. Results.....	47
3.3a. Dynamic quantities.....	47
3.3b. Eulerian transport.....	54
3.3c. Lagrangian-mean transport.....	58
3.3d. The "stirring" term.....	61
3.3e. Lagrangian model.....	64
3.3f. The effects of eddy transport on the mean field.....	70
3.4. Conclusions.....	75

Chapter 4: TRANSPORT BY TRANSIENT PLANETARY WAVES
AND THE DIABATIC CIRCULATION

4.1. Introduction.....	78
4.2. Dynamics.....	78
4.2a. 600 m topography.....	79
4.2b. 900 m topography.....	83
4.2c. The diabatic circulation.....	84
4.2d. Initialization.....	85
4.3. Transport by the diabatic circulation.....	88
4.4. Planetary wave transport.....	94
4.4a. Conservative transport.....	94
4.4b. Chemical transport.....	97
4.5. Transport by both waves and the diabatic circulation.....	101
4.6. Discussion.....	107
4.6a. Diabatic circulation.....	109
4.6b. Planetary waves.....	109
4.7. Summary.....	124

Chapter 5: SUMMARY AND A POSSIBLE OZONE TRANSPORT SCENARIO

5.1. Introduction.....	126
5.2. Low latitude ozone.....	127
5.2a. The diabatic circulation and the initial ozone distribution.....	127
5.2b. Geometric effects.....	128
5.2c. Photochemistry.....	130
5.2d. Planetary waves.....	134

5.3. Vertical structure.....	134
5.3a. Small scale dynamics.....	134
5.3b. Photochemistry.....	135
5.4. Unstable waves.....	136
5.5. Summary and discussion.....	137
5.5a. Steady state waves.....	137
5.5b. Time dependent waves.....	141
5.5c. A possible transport scenario.....	148
ACKNOWLEDGMENTS.....	152
Appendix A: NOMENCLATURE AND CONVENTIONS.....	153
Appendix B: NUMERICAL PROPERTIES OF THE	
ADVECTION-DIFFUSION EQUATION.....	155
References.....	162

MECHANISMS OF STRATOSPHERIC OZONE TRANSPORT

Chapter 1

INTRODUCTION TO THE OZONE TRANSPORT PROBLEM

1.1 Introduction

Due to the complex nature of ozone photochemistry and atmospheric dynamics it has been difficult to discern the mechanisms most important for the maintenance of the ozone distribution. To simultaneously model both the photochemistry and the dynamics is a prodigious task as each system involves a variety of temporal and spatial scales. One goal of atmospheric scientists is to understand transport dynamics so that they may be simplified and then meaningful predictions of ozone perturbations by pollutants can be generated for a considerably reduced computational cost.

The oxygen photochemistry (Chapman, 1930; Craig, 1965) predicts that the maximum amount of ozone should be observed in the tropical stratosphere and the minimum amount observed in the polar stratosphere. However, even as early as 1930 it was known that the maximum amount of ozone occurs at extratropical latitudes (Dobson, 1930), and it thus was postulated that atmospheric transport redistributes ozone from its tropical source regions to polar latitudes.

Manuscript approved October 7, 1982.

Simple mean cell circulations were proposed by Brewer (1949) and Dobson (1956) which were capable of transporting ozone from tropical to polar regions. Planetary wave mechanisms were also suggested that were capable of producing poleward, downward transport (Newell, 1961, 1963). As more data became available and dynamical theory developed, it became obvious that the transport of minor constituents was a complex interaction of planetary waves and the mean circulation (Mahlman, 1969). An excellent review of the history of stratospheric circulation and transport studies is given in Hsu (1980).

Since the classic study of Craig (1950), the photochemical models of ozone have become progressively more complex. The remeasurement of rate coefficients by Benson and Axworthy (1957) indicated that oxygen-only photochemical models would predict more ozone than is observed (Hunt, 1966). Catalytic destruction cycles involving hydrogen oxides, nitrogen oxides, and halogen oxides have been added to the photochemistry models. With the addition of these reactions it became obvious that anthropogenic pollutants might cause serious degradation of the ozone layer (WMO, 1981). Complete photochemical models now require the use of over 100 chemical reactions (Johnson and Podolske, 1978). Use of the complete set of photochemical equations requires that the dynamics be simplified in order to make the problem tractable. The most common parameterization has been to treat atmospheric transport as if it were diffusive.

1.2 Diffusion models

The dynamics in photochemistry models have been parameterized with both one-dimensional (1-D) and two-dimensional (2-D) schemes. The 1-D models represent transport by vertical diffusion from high altitude source regions to the lower stratosphere where ozone is observed to accumulate (McConnell and McElroy, 1973). Hunten (1975) pointed out that vertical diffusion models are useful because of the tendency of the atmosphere to be horizontally stratified. Upward motions are always associated with compensating downward motions; therefore, the atmosphere "looks like just another big eddy" (Hunten, 1975). Strobel (1981), however, has shown that 1-D diffusion models are only strictly valid when the horizontal gradient of the constituent is small, a condition not valid for ozone.

Two-dimensional models attempt to model the atmosphere more realistically by approximating both vertical and latitudinal transport (Hidalgo and Crutzen, 1977; Harwood and Pyle, 1977; Miller et al., 1981). In 2-D models a mean advective field is usually prescribed and wave motions are modeled by a diffusion scheme. Such diffusion parameterizations are often based on the empirical theory of Reed and German (1965) which has been shown to be inconsistent with the hydrodynamics of planetary wave transport (Matsuno, 1980).

Hidalgo and Crutzen (1977) and Harwood and Pyle (1977) use an Eulerian mean formulation where ozone is advected with a zonal mean velocity and the zonal mean effects of the planetary waves are modeled by diffusion. In another 2-D model Miller et al. (1981) use

the concepts of the Lagrangian-mean circulation, as discussed by Dunkerton (1978), to derive a Lagrangian-mean circulation for the stratosphere. Miller et al. use this Lagrangian-mean field instead of the Eulerian velocity for the mean field advection.

The problems of applying diffusion models to atmospheric dynamics has been discussed by Mahlman (1975), Tuck (1979), and Pyle and Rogers (1980a). These authors have pointed out the inadequacies of diffusion parameterizations when trying to model countergradient transport, the spatial variation of chemical sources and sinks, and the diffusion of different chemical species. These processes are important for ozone in some parts of the atmosphere and should be properly represented in ozone models.

Matsuno (1980), Clark and Rogers (1978), Dunkerton (1980), and Plumb (1979) have shown that the nature of planetary waves is advective rather than diffusive for an inert tracer. This means that the diffusion tensor should be antisymmetric rather than symmetric as is generally assumed. Matsuno (1980), Strobel (1981), and Holton (1981) show that there are also diagonal components in the diffusion tensor which are associated with chemical source terms and wave growth and decay.

The manner in which diffusion parameterizations have been applied to the atmosphere is a source of uncertainty in the results of chemical models. Pyle and Rogers (1980a), however, argue that a dynamically consistent reformulation of the diffusion coefficients might allow for the continued use of diffusion models. Strobel (1981) has formulated such a consistent theory for linear waves with

complex frequencies. This theory demonstrates the strong coupling between the magnitude of the diffusion coefficients and photochemistry. Because of this strong coupling, a simple diffusion model can adequately predict ozone dynamics only under special circumstances.

Even with all the shortcomings, diffusion can improve the results of chemical models. A recent study by Pyle and Rogers (1980b) shows that the addition of diffusion generates more realistic ozone distributions than those produced by the diabatic circulation alone. Since it is difficult to fully model both photochemistry and dynamics, such parameterizations are desirable and necessary. In the next section attempts to study ozone with an accurate representation of the dynamics and a simplified chemistry will be discussed.

1.3 General circulation models

Several general circulation models have been used to study ozone transport and the interaction of radiation, dynamics, and photochemistry (Clark, 1970; Cunnold et al., 1975; Schlesinger and Mintz, 1979; and Mahlman et al., 1980, and the references therein). These models emphasize an accurate representation of the dynamics, but the chemistry is modeled in a simplified manner. General circulation studies have met with some success in modeling the large scale features of the ozone distribution, but no general circulation model has ever reproduced all of the large scale features of the ozone distribution.

For example, the quasigeostrophic model of Clark (1970) established the usefulness of the Chapman photochemistry in modeling ozone and demonstrated that the stratospheric circulation was able to perturb the radiative equilibrium ozone distribution to a distribution that approximated observations. In particular, planetary waves were found to transport large amounts of ozone from equatorial source regions to high latitudes. By increasing the horizontal gradient of temperature, Clark was able to generate high latitude warmings; however, these model warmings do not accurately represent the observed characteristics of stratospheric warmings. Therefore, while the dynamics in this model do indeed cause an increase in ozone at high latitudes, the mechanisms that lead to this increase are in doubt due to the inability of the model to accurately represent observed dynamical features.

Using the same dynamical equations of Clark (1970), Cunnold et al. (1975) generated more accurate zonal mean circulations than Clark; however, the model did not generate stratospheric warmings. In their formulation of the ozone continuity equation Cunnold et al. used a large vertical eddy diffusion coefficient, based on diffusivities of 1-D chemistry models, which dominated the vertical transport in the model. Although Cunnold et al. did calculate a polar spring ozone maximum, the accumulation of ozone at high latitudes was underestimated.

Mahlman et al. (1980) used a 3-D primitive equation model to perform two ozone transport experiments. In the first experiment the ozone was specified at 10 mb (the upper boundary) and was not

allowed to deviate from this climatological value. In the second experiment, ozone was calculated at 10 mb using a simple oxygen, hydrogen, nitrogen photochemistry. Mahlman et al. predicted the vertical structure in the lower stratosphere quite accurately using either of the specified photochemistries. This indicates that the ozone distribution in the lower conservative region of the stratosphere is to a large degree independent of the details of the photochemistry.

The most serious discrepancy encountered by Mahlman et al. was in the calculation of the total ozone field. Their spring maximum was too large, occurred too far south, and too late in the spring. The model of Mahlman et al. (1980) did not produce midwinter stratospheric warmings and, indeed, none of the general circulation models accurately represent warming events. It was suggested in Mahlman et al. (1980) that the transport during a stratospheric warming might be important in improving their results.

General circulation studies are so complex that transport mechanisms are difficult to isolate and describe. The dynamics are a combination of diabatically forced circulations and wave regimes, and it is difficult to pinpoint the dominant transport mechanisms. Therefore, the use of simple models to study the transport associated with specific dynamic mechanisms might be useful not only for application to chemical models but for interpreting general circulation model results.

1.4 Lagrangian-mean models

In general the constituent and dynamics fields in atmospheric studies are formulated as a mean field and a deviation field. The mean fields are generally formed by taking averages around a latitude circle in what is called the Eulerian formulation. An alternative to the Eulerian zonal mean and deviation method is the Lagrangian-mean method (Andrews and McIntyre, 1978; McIntyre, 1980), in which averages are taken over particular fluid tubes. As the averaging operator is defined, the Lagrangian-mean velocity is the center of mass velocity of a particular ensemble of fluid particles and is zero when the conditions of the noninteraction theorem (Boyd, 1976; Andrews and McIntyre, 1978, Dunkerton, 1980) are met.

With the above attributes it may seem natural to study tracer transport in the Lagrangian-mean system. However, it turns out the Lagrangian-mean transform is complicated and is easily formulated only for the most simple flows (McIntyre, 1980). Furthermore, dispersion of the particles about their center of mass may render the location of the center of mass of little use in the actual location of the tracer (Hsu, 1980). McIntyre (1980) has suggested that some sort of continuous reinitialization of the Lagrangian-mean model might solve this problem, but no practical scheme is currently available.

Much of the work concerning the application of the Lagrangian-mean to the transport of tracers in the stratosphere is based on the concepts of Dunkerton (1978). Dunkerton shows that for steady, non-dissipating planetary waves, the Lagrangian-mean circulation of the stratosphere is to first order the zonal mean circulation first

estimated by Murgatroyd and Singleton (1961). This circulation is specifically the radiatively forced circulation (diabatic circulation).

Planetary waves are dissipative and not steady and have a significant effect on tracer transport; therefore, the diabatic circulation is not sufficient to describe all the large scale transport processes. Schoeberl (1981a) has calculated the Lagrangian-mean circulation associated with steady, dissipative planetary waves. Schoeberl found that planetary waves transport ozone from southern to northern regions, and that strong downward transport can be obtained with "realistic" mean zonal wind profiles. It is found that in the lower stratosphere the Lagrangian-mean flow associated with planetary waves is dominate over the diabatic circulation.

The 2-D model of Miller et al. (1981) used a modified version of Dunkerton's diabatic circulation and a diffusion tensor based on the coefficients of Luther (1974) and Hunten (1975). Since the Lagrangian-mean circulation as calculated by Dunkerton (1978) does not include the effects of dissipative planetary waves, the use of the Luther and Hunten diffusion coefficients does not accurately represent planetary wave effects. The large scale transport by waves in this model are advective, and any diffusion parameterization must include an antisymmetric component.

In another approach to the diffusion problem Holton (1981) and Strobel (1981) use the residual circulation defined by Andrews and McIntyre (1976) to calculate tracer transport. Holton's model

incorporates both planetary wave and diabatic effects. A diffusion model is then used to calculate the small scale correction to the advection by the residual circulation. This formulation involves an antisymmetric tensor that represents an advective correction to the residual circulation which is assumed to be negligible. If this antisymmetric component is indeed small then the diffusive part of the model should truly represent subplanetary wave effects.

1.5 Objectives

The global scale transport of ozone can be divided into two circulation regimes: transport by the diabatic circulation and transport by large scale waves. Each of these transport regimes is important and both must be modeled in a complete study. There are, of course, important smaller scale processes that significantly affect the ozone distribution. These small scale processes are particularly important in the lower stratosphere and in the troposphere. No attempt will be made here to study these small scale processes.

In this dissertation steady state and time dependent planetary wave models are used to investigate ozone transport by the diabatic circulation and the planetary waves. The steady state model is formulated in both the Eulerian and Lagrangian-mean systems and results from the two formulations compared. The Lagrangian-mean circulation is also compared to the residual circulation. Using the Lagrangian-mean displacement fields a Lagrangian parcel model is developed to study the interaction of photochemistry and dynamics.

The simplicity of the steady state model makes it easier to use than the time dependent model to understand the fundamental transport mechanisms; however, the steady state model has the disadvantage that potentially important transient effects are neglected.

In the time dependent model quasigeostrophic waves are calculated on a midlatitude β -plane. This model is the same as used in Schoeberl (1982a, 1982b) to study vacillation and wave saturation. A diabatic circulation based on the flow calculated by Dunkerton (1978) is superimposed on the planetary wave circulation. With this model it is possible to identify the transport characteristics of the planetary waves and the diabatic circulation and, therefore, measure their relative importance.

The use of these simple models aids in the understanding of the specific mechanisms involved in the interaction of photochemistry and dynamics. With this understanding it is possible to make a rational approach to the development of dynamic simplifications appropriate for chemical models.

In Chapter 2 the dynamic and transport models are described with particular attention given to the transport scheme. In Chapter 3 the transport by steady, dissipative waves is covered and the Eulerian, Lagrangian-mean, and residual circulations are compared. The transport by the diabatic circulation, the planetary wave circulation, and the two circulations combined is presented for the time dependent β -plane model in Chapter 4. Finally, the results of the various models are discussed and interpreted in Chapter 5.

Chapter 2

MODEL DESCRIPTIONS AND THE OBSERVED OZONE DISTRIBUTION

2.1 Transport model

A general transport model has been developed that is capable of interfacing with a variety of dynamic models or a specified time dependent data set. The transport model runs in an off-line fashion; that is, the constituent distribution does not affect the dynamics. Since ozone is radiatively active, with the use of an off-line model it has been tacitly assumed that the change in radiative forcing associated with the changes in the ozone field is small. This assumption is valid as most of the transport in this model takes place during winter when insolation is small, and in the infrared, ozone cooling is much smaller than carbon dioxide cooling (Apruzese et al., 1982). Mahlman and Moxim (1978) discuss in detail the differences between on-line and off-line models.

For the transport model, the constituent continuity equation (Eq. 3.1) is written in discrete form using fourth order centered finite differences in the meridional and vertical directions. The longitudinal dependence is represented by zonal harmonics, and the time derivatives are estimated with second order centered differences (leapfrog). The spatial differences are written in the general form

$$\frac{\partial \mu}{\partial y} = \frac{C_1(\mu_{i+1} - \mu_{i-1})}{\Delta y} + \frac{C_2(\mu_{i+2} - \mu_{i-2})}{\Delta y} \quad (2.1)$$

where $C_1 = 2/3$ and $C_2 = -1/12$ for fourth order accuracy (Ames, 1977; Haltiner and Williams, 1980). Further details of the numerical scheme are given in Appendix B.

Mahlman and Sinclair (1977) review in some detail the difficulties of modeling fluid advection. They tested several transport algorithms using a one-dimensional model with a steady, uniform flow and cyclic boundary conditions. The most serious deficiency encountered by Mahlman and Sinclair was the generation of nonphysical negative mixing ratios which had to be removed. This removal process is called "filling" (Mahlman and Moxim, 1978). Filling is a subjective process that usually involves borrowing mass from surrounding grid points in a mass conserving fashion so that negative mixing ratios are increased to zero.

Mahlman and Sinclair found that the pseudo-spectral method produced the best results when a large number of waves and a short time step were used. However, the pseudo-spectral scheme requires a large amount of computer time and memory. Of the other schemes tested the fourth order derivative scheme worked well and is computationally efficient (see also Kreiss and Oliger, 1972).

Clancy (1981) has suggested that the advection-diffusion equation can be efficiently integrated for a wide range of parameters using second order spatial differences and a forward time scheme. A comparison of the current model with Clancy's scheme illustrates the computational gain of the fourth order differences.

Following Clancy, the advection-diffusion equation,

$$\frac{\partial S}{\partial t} = -u \frac{\partial S}{\partial x} + D \frac{\partial^2 S}{\partial x^2}$$

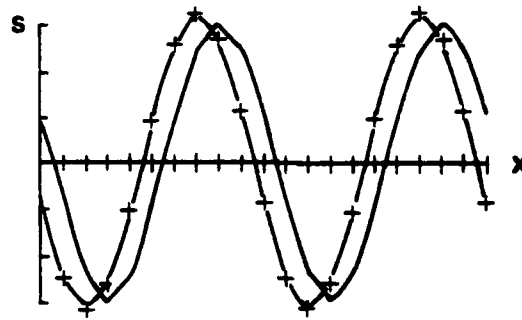
is integrated with $\Delta t = 3600$ sec, $\Delta x = 10^5$ m, $u = .6$ m sec⁻¹, and $D = 10^3$ m² sec⁻¹. The initial condition is

$$S = S_0 \sin(kx) \quad (t = 0)$$

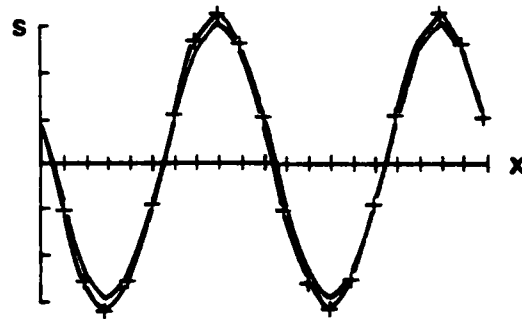
and the boundary conditions are

$$S = -S_0 \exp(-k^2 Dt) \sin(kx - ukt)$$

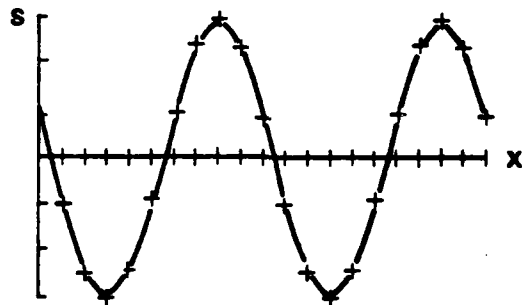
for $x = 0$ and 10^9 m, where $k = 2\pi(10^8 \text{ m})^{-1}$. The results after 30 days using second order spatial differences and forward time differences are shown in Fig. 2.1a (same as Clancy Fig. 2). In Fig. 2.1b forward time differences are still used, but the second order differences are replaced with the fourth order estimates. The phase error in Fig. 2.1b has been reduced to nearly zero, but as in Fig. 2.1a the amplitude is overestimated.



a. Forward in time, second order in space
(same as Clancy (1981), Fig. 2)



b. Forward in time, fourth order in space



c. Centered in time, fourth order in space

Figure 2.1. A comparison of the analytic and numerical solution of the advection-diffusion equation using different numerical schemes.

(+ on the numerical solution)

In Fig 2.1c the same problem has been redone using centered time differences and fourth order spatial differences. Both the phase error and the amplitude error have been reduced to the point that the numerical solution and the analytic solution are virtually indistinguishable. The use of forward time differences, as in Fig. 2.1a and 2.1b, leads to an unstable solution when the diffusion is zero (Haltiner and Williams, 1980, p. 130). The incorporation of diffusion by Clancy serves to stabilize his scheme. A more detailed analysis of this problem is presented in Appendix B.

Using second order differences without diffusion to model ozone transport leads to unrealistic results because $2\Delta x$ waves generated in regions of large gradients propagate throughout the domain and dominate the computed transport. Only with the incorporation of very large diffusion can this effect be countered for long integrations. The inferiority of second order transport models is well illustrated in Mahlman and Sinclair (1977).

Besides filling, there are other nonphysical processes that must often be incorporated into transport models in order to obtain reasonable numerical results. Diffusion is commonly introduced to represent subscale effects (Mahlman and Moxim, 1978). As indicated above, diffusion generally increases the numerical stability of a finite difference scheme and reduces the occurrence of negative mixing ratios.

At the boundary some technique must be used in order to close the finite difference scheme. In this model either the ozone density or flux is specified at the boundary, and the derivative at

the first interior point is represented by second order centered differences ($C_1 = 1/2$, $C_2 = 0$ in Eq. 2.1). With this closure system the boundary points and the first two interior points are directly affected by the prescription of the boundary.

The closure method chosen generally affects the ability of the model to conserve mass due to the uncompensated difference terms at each boundary. In a problem where there is no flux through the boundary and no source or sink terms, it is possible to calculate the mass change caused by these uncompensating terms and then distribute an appropriate mass correction over the model domain. In a problem with boundary fluxes and sources, mass correction becomes more complex.

Closure at the boundaries has another effect which particularly manifests itself in the ozone problem. Since the mixing ratio estimates at or near the boundaries are less accurate estimates than at the interior points, transport into the polar region, a region of special interest, will be inaccurate.

The philosophy taken in this study is to incorporate as little subjectivity as possible, so that important transport mechanisms will be revealed. Neither diffusion nor mass correction have been incorporated into this model, and it has not been necessary to incorporate filling.

The omission of diffusion means that the lower regions of the stratosphere and troposphere are expected to be poorly modeled. In particular there is no mechanism for tropospheric-stratospheric interaction and there are no subplanetary scale dynamics to produce

small scale mixing effects. As a result, there is sometimes a tendency for ozone to build up to unrealistic levels in the lower stratosphere. In certain instances an ozone "shock front" develops and the finite difference estimates of the derivatives become very poor, ultimately leading to the failure of the model in all regions. Strong ozone fronts do not develop with the dynamics used here.

The ability of the model to conserve mass has been studied using an inert tracer. The model does not conserve mass exactly and the entirety of the mass correction can be attributed to noncompensating finite difference terms. The mass discrepancy remains small for the integration times considered here. When ozone chemistry is included the total ozone in the model domain never differs unrealistically from the initial total ozone. For these reasons it was judged that the incorporation of any mass correcting techniques would contaminate the model results at least as badly as the uncorrected mass. Therefore, the results presented here have no mass correction added.

2.2 Chemistry model

The chemistry used in this study is extremely simple, namely:

$$Q = - \lambda \mu \quad (\lambda > 0) \quad (2.2)$$

where Q is the photochemical source term (see Eq. 3.1), λ is the photochemical relaxation rate, and μ is the constituent mixing

ratio. This formulation requires that the perturbations relax back to zero, while changes in the mean field relax to the mean equilibrium. The rate of relaxation is inversely proportional to λ , which is assumed to be a function of height only. The λ values are based on those of Hartmann and Garcia (1979).

The chemical reactions that the λ values are based on are the oxygen-only reactions (Chapman reactions) with a modified ozone destruction rate to represent catalytic destruction. Ozone chemistry is temperature dependent; however, the temperature dependence is weak (Hartmann and Garcia, 1979; Garcia and Hartmann, 1980) and will be neglected here.

It is convenient to divide the atmosphere into three regions in order to discuss ozone transport (Fig. 2.2). The lower region is the region in which ozone is assumed to be a conservative tracer ($\lambda \ll k\bar{u}$). In the upper region ozone photochemistry is so fast ($\lambda \gg k\bar{u}$) that the transport is unimportant. The middle region is defined by the ratio $.1 < \lambda/\bar{u}k < 10$ and is called the transition region (Hartmann and Garcia, 1979). The interaction of the photochemistry with the dynamics is expected to be greatest in the transition region.

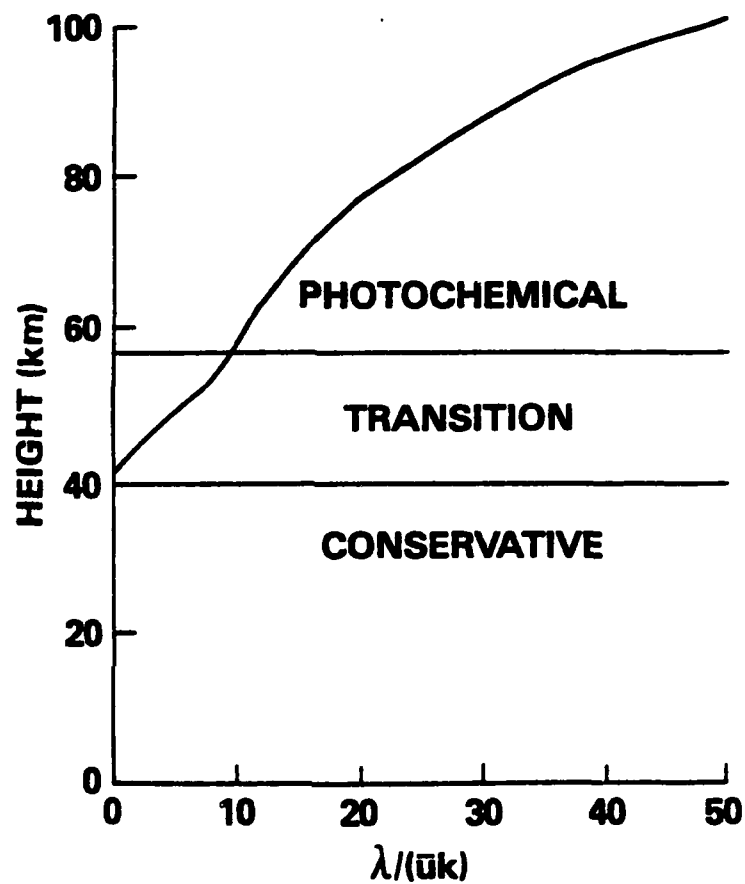


Figure 2.2. Schematic representation of the atmosphere divided into photochemical, transitional, and conservative regions based on the ratio of the dynamical to the photochemical time scale.

2.3 Dynamic models

The planetary waves are represented in all the dynamic models by the quasigeostrophic potential vorticity equation. The governing equation is

$$\left(\frac{\partial}{\partial t} + a + \bar{u} \frac{\partial}{\partial x} \right) q' + \frac{1}{f} \frac{\partial \bar{q}}{\partial y} \frac{\partial \phi'}{\partial x} = F' \quad (2.3)$$

where q' is defined as in Holton (1975). The mean flow equation is

$$\frac{\partial \bar{q}}{\partial t} = - \frac{\partial(\bar{q}'v')}{\partial y} - a(\bar{q} - \bar{q}_0) + \bar{F} \quad (2.4)$$

$\bar{q}$ is the wave induced change in the background potential vorticity $\bar{q}_0$. Damping is modeled as Newtonian cooling and Rayleigh friction, which are assumed to be equal, and is represented by a . F is diabatic forcing other than Newtonian cooling. Other nomenclature is given in the Appendix A.

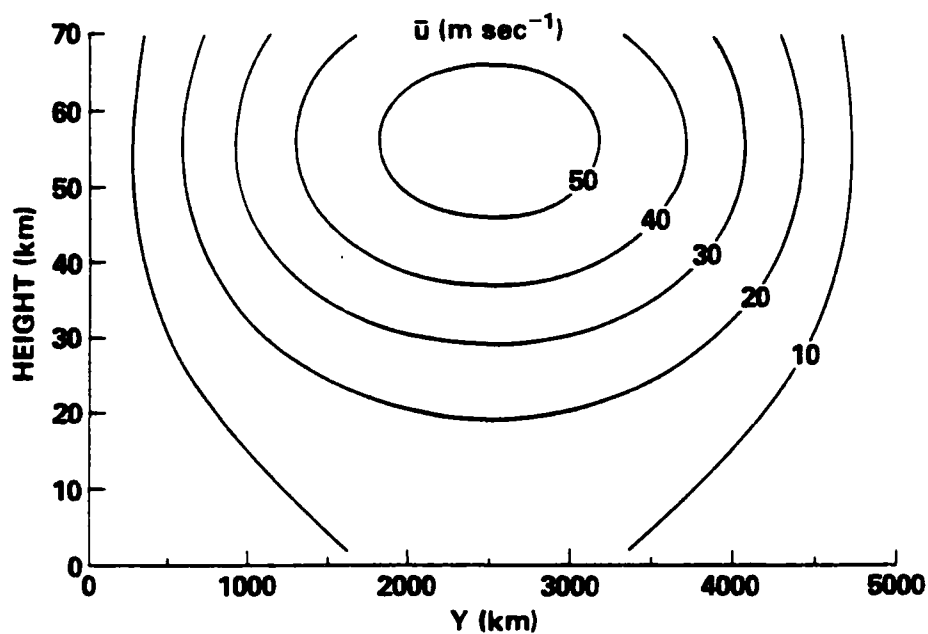
a. Steady state model

In the steady state model $\frac{\partial}{\partial t}$ is assumed to be zero and β -plane geometry is used. The damping is a slowly varying function of height except in the uppermost regions where it is assumed to increase at an exponential rate. This region of increase provides a sponge layer to reduce reflections from the upper boundary and is well above the region where ozone transport is being studied. The time scale of the damping below the sponge layer is approximately 10

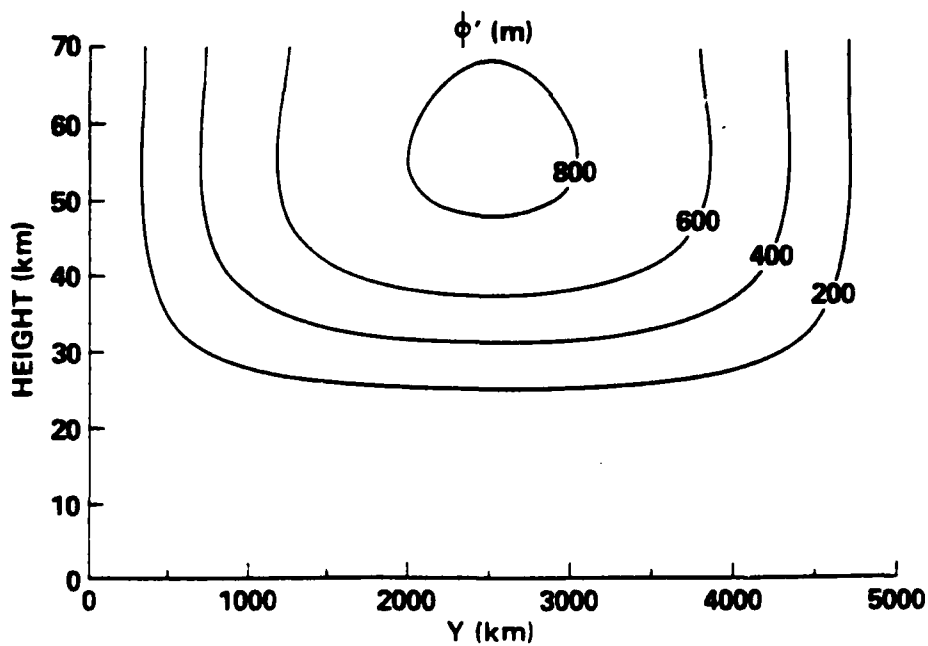
days, which is representative of the physical processes in the middle or upper stratosphere (Schoeberl and Strobel, 1978; Dickinson, 1973). The diabatic forcing, F , is assumed to be zero; therefore, only the steady state planetary wave transport is modeled.

Equations (2.3) and (2.4) are solved for a wavenumber one perturbation with forcing analogous to a mean zonal wind of 10 m sec^{-1} flowing over a 1 km mountain. The mean zonal wind profile and the geopotential amplitude are shown in Fig. 2.3. The mean zonal wind is presumed to approximate the polar night jet. The geopotential reaches a maximum amplitude of about 1 km in the mesosphere and does not exceed the steady state saturation limit (Schoeberl, 1982a). The magnitude of the geopotential is in rough agreement with observation (van Loon et al., 1973). The eddy geopotential field is used to calculate the wave forced mean meridional and vertical velocity.

It is assumed in the steady state model that wave effects on the mean flow are not so large that the wave field needs to be recomputed with the perturbed mean flow. This restricts the validity of the steady state model to small amplitude waves and is consistent with the transport approximations to be made later (Section 3.2). This model is the Eulerian analogue to the Lagrangian-mean model of Schoeberl (1981a).



a. Mean zonal wind profile for the steady state model.



b. Magnitude of the steady state geopotential.

Figure 2.3.

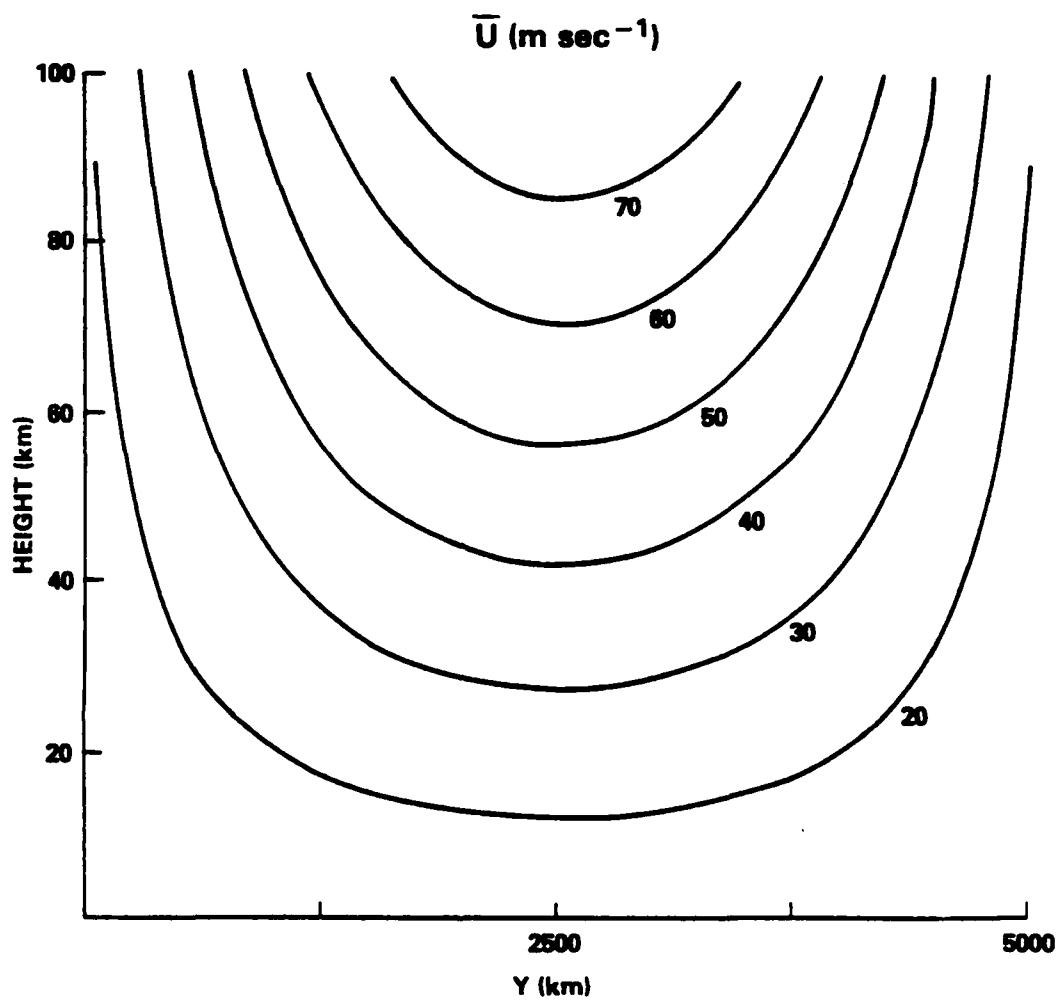


Figure 2.4. Initial mean zonal wind profile
for the time dependent β -plane model.

b. Time dependent β -plane model

In the time dependent β -plane model the diabatic forcing, F in Eqs. (2.3) and (2.4), is assumed to be zero, and the Newtonian cooling and Rayleigh friction are similar to that of the steady state model. Diabatic effects are modeled by adapting a circulation from Dunkerton (1978) and Murgatroyd and Singleton (1961) and superimposing this circulation on the calculated planetary wave circulation.

The waves are forced by zonal flow over a mountain. The basic state mean zonal wind profile is shown in Fig. 2.4. This model is quasilinear in that after every time step the wave perturbed mean field is calculated; the waves are then recalculated using the updated mean fields. The mountain height is a model parameter and the behavior of the model varies depending on the mountain height. The planetary wave model is described in Schoeberl (1982b).

For small mountains equilibrium is reached in which westerlies are present everywhere. For mountains in an intermediate range the model produces at least one stratospheric warming and then is observed to reach a steady state with a permanent layer of easterlies. For large topography the model produces a series of warmings, i.e. it vacillates. For the transport experiments two mountain heights will be used: 600 m and 900 m. Both of these mountain heights are in the intermediate range where there is a warming and the eventual onset of a permanent layer of easterlies.

The β -plane transport model is centered at 60° N and has a horizontal extent of 5000 km. This corresponds to a latitudinal

extent from 37.5° to 82.5° . The domain extends from the ground to 100 km. There are 20 grid points in the horizontal and 40 grid points in the vertical and a perimeter of points around the grid at which the boundary values are prescribed.

At the top and bottom of the model the constituent flux is set equal to zero. This boundary condition is accurate at the top since the ozone is in photochemical equilibrium. At the northern and southern boundaries the flux associated with the planetary waves is assumed to be zero. At the southern boundary, where ozone is observed to have relatively little annual variation, the flux by the diabatic circulation is calculated assuming that ozone is constant. The flux associated with the diabatic circulation at the northern boundary is assumed to be zero.

The closure of the finite difference scheme directly involves the prescribed boundary point and the first interior point. The second and third interior point are directly associated with the first interior point at each time step. Because of this possible direct contamination by the closure system, results will be presented at horizontal grid points 4 (46°) and 17 (74°) as representative of the northern and southern regions of the model. Table 2.1 lists the grid points, the corresponding distance from the southern boundary of the β -channel, and the latitude for the time dependent model.

Table 2.1
Horizontal Distance Scales for the Time
Dependent β -plane Model

Grid Point	Y (km)	Latitude °
1	235	39.6
2	476	41.8
3	714	43.9
4	952	46.1
5	1190	48.2
6	1428	50.4
7	1667	52.5
8	1905	54.6
9	2143	56.8
10	2381	58.9
11	2619	61.1
12	2857	63.2
13	3095	65.4
14	3333	67.5
15	3571	69.6
16	3810	71.8
17	4048	73.9
18	4285	76.1
19	4523	78.2
20	4761	80.4

2.4 Observations of ozone

In order to evaluate the success of the transport model it is necessary to compare the calculated results to the observed ozone distribution. In this section a brief review of the observed behavior of ozone will be given. Some mention should be made concerning the units used in reporting the observations. The transport calculations will be done using the mixing ratio (ppm) as the basic unit of measure. Many of the observations have been made in other units such as ozone partial pressure or absolute density. Due to the variation of atmospheric density, the ozone distribution appears quite different in the various units. Kreuger and Minzner (1976) present conversion factors in their Table 2. Much of the general review of the ozone distribution is taken from Craig (1965) and Dütsch (1974).

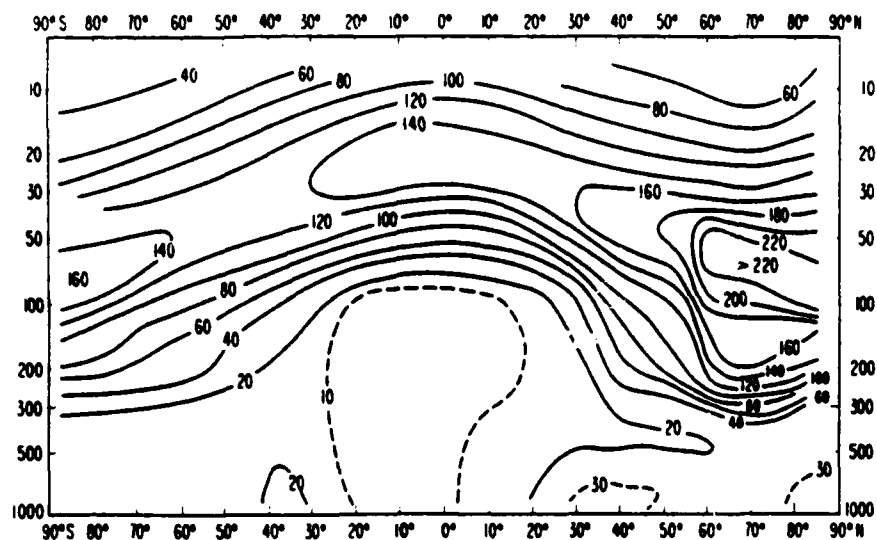
The ozone distribution in the Southern Hemisphere is less well known and considerably different from that in the Northern Hemisphere. The distribution in the Southern Hemisphere will not be discussed here since the dynamics are more appropriate for the Northern Hemisphere. There will be some comment made about transport mechanisms in the Southern Hemisphere based on model results in Chapter 5.

At extratropical latitudes the ozone density reaches a maximum in the lower stratosphere between 15 and 30 km. The mixing ratio maximum occurs at somewhat higher altitudes between 35 and 45 km. Above and below the ozone maximum the average state shows a monotonic decrease of ozone concentration; though, at any time an

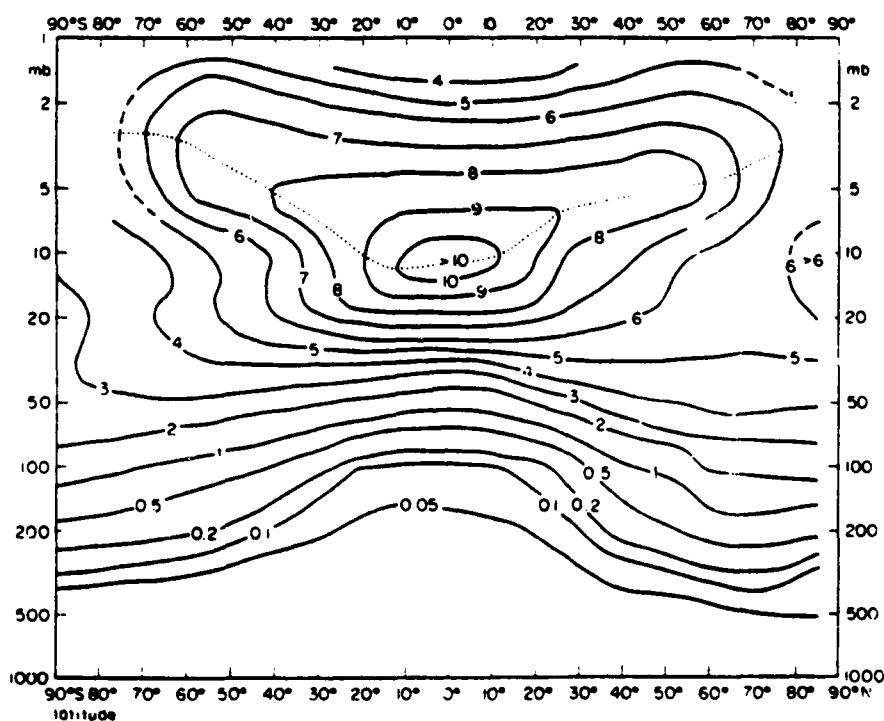
ozone sounding might show considerable vertical structure due to small scale transport. There is ozone of stratospheric origin in the troposphere, but the amounts are quite small; and there is no attempt made here to model tropospheric ozone.

The absolute maximum amounts of ozone in the Northern Hemisphere occur at polar latitudes in the lower stratosphere during late winter and early spring. This is in complete disagreement with photochemical theory which predicts maximum amounts in the tropics at high altitudes in the middle and upper stratosphere. Figure 2.5 shows a March/April pole to pole cross section of ozone in both ozone density and ozone/air number mixing ratio. The predominant maximum seen at high latitudes in the density field is not clearly seen in the mixing ratio field since this maximum occurs in a region of relatively high atmospheric density. Two important features to note in the mixing ratio field are the positive horizontal gradient at high latitudes in the lowest part of the stratosphere and the negative horizontal gradient at high latitudes in the middle and upper stratosphere (Dütsch, 1974).

In the photochemical region the variation of ozone is quite small. In the lower stratosphere, however, the variation is large on both daily and seasonal time scales. Much of the daily variation is due to the infusion of synoptic size weather systems into the stratosphere and will not be modeled here. The seasonal variation is due to the diabatic circulation and planetary wave transport and will be investigated in this dissertation.



Partial Pressure



Mixing Ratio

Figure 2.5. Latitude-height cross section of ozone for March/April in partial pressure and ozone/air number mixing ratio (from Dütsch, 1974).

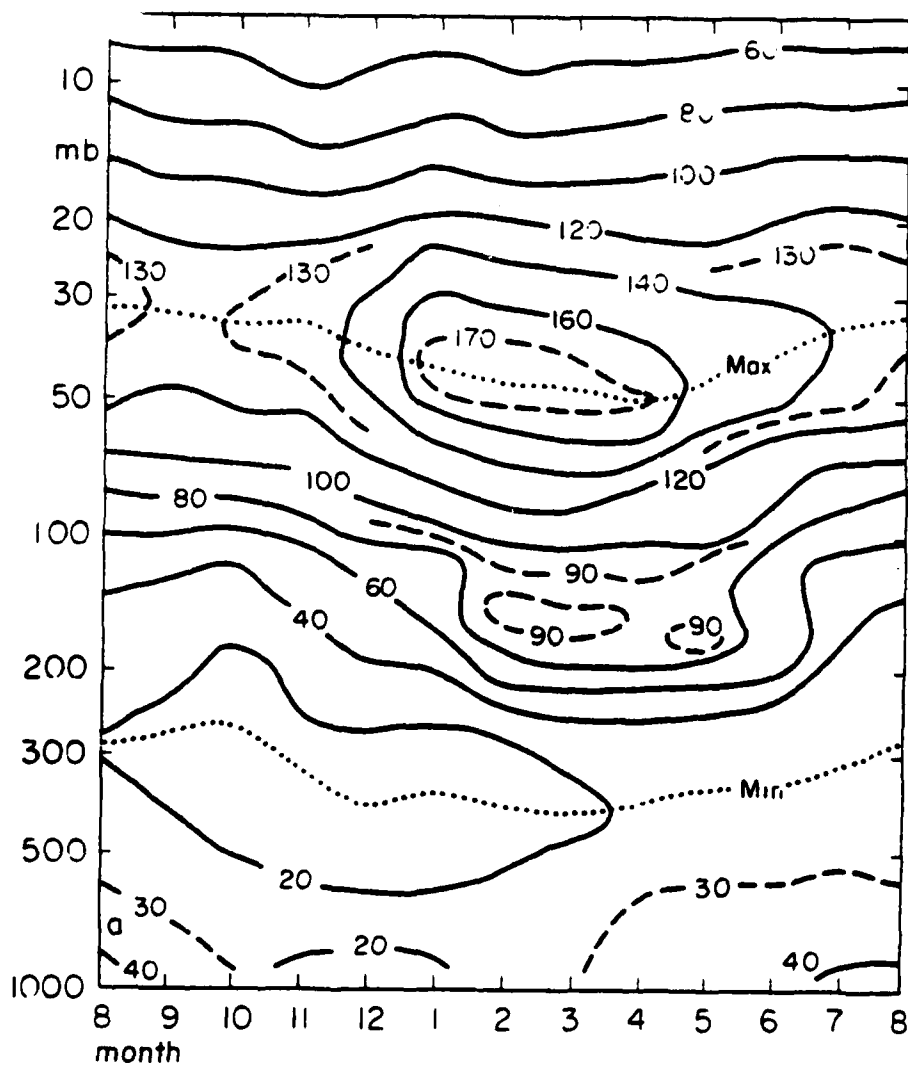
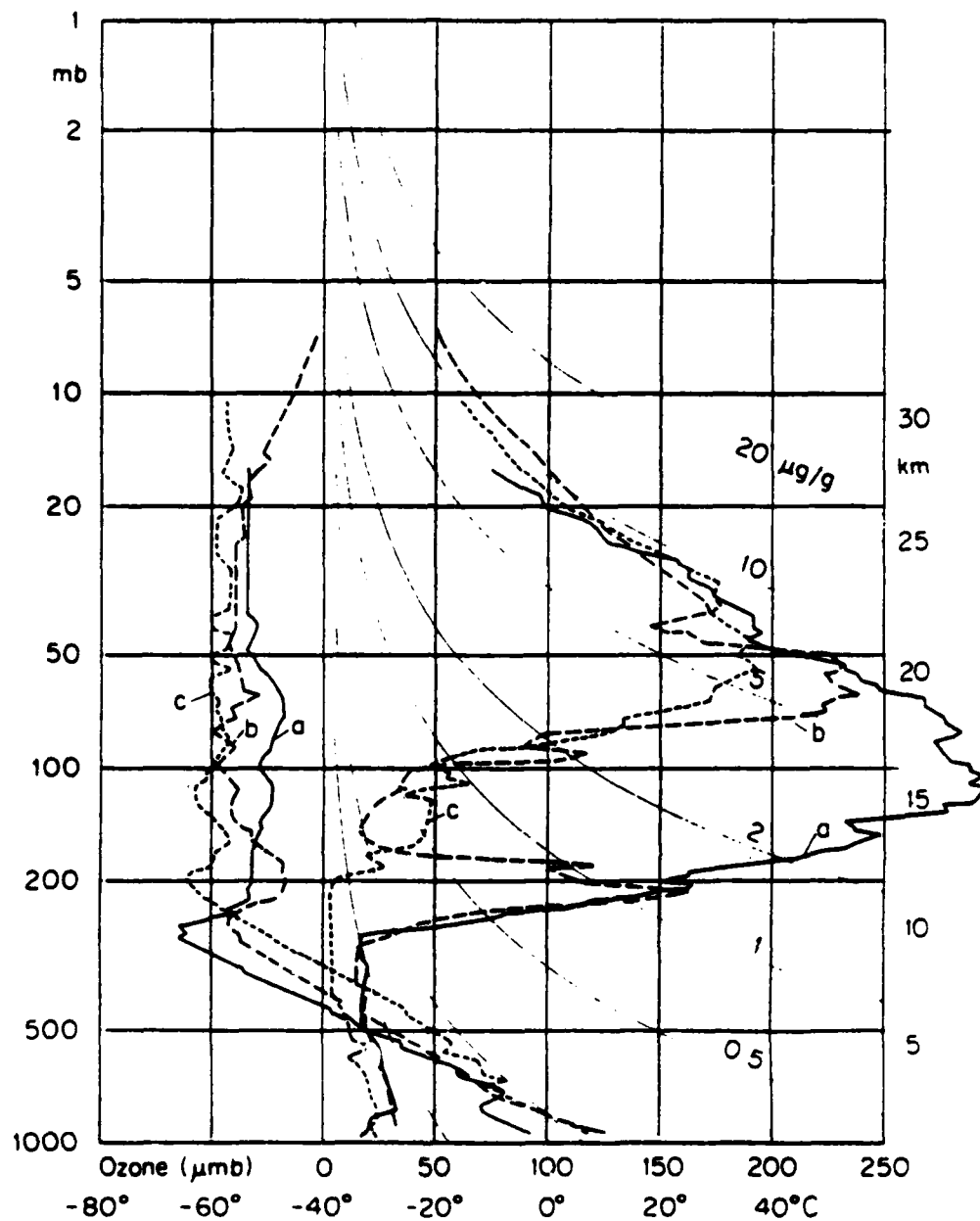


Figure 2.6. 6 year average time-height cross section of ozone partial pressure at Arosa, Switzerland (from Dütsch, 1974).



a. Arctic type

b. Polar type

c. Subtropical type

Temperature on left in °C.

Figure 2.7 Vertical distribution of ozone from 3 soundings in late winter and early spring (μmb) (from Dütsch, 1974).

Figure 2.6 shows a time cross section of vertical ozone over Arosa, Switzerland. The most important feature to note is the strong increase between 30 and 50 mb (21-25 km) which begins in December and reaches a maximum in late winter. By the first of April there is a definite decreasing trend. Since this maximum appears in the conservative portion of the stratosphere, it can be deduced that transport is responsible for the ozone buildup. Another notable feature is the minimum concentration that occurs at the level of the tropopause.

In Figure 2.7 is shown the vertical distribution from three soundings in late winter and early spring. In the polar type sounding (b) there is a pronounced secondary density maximum in the lowest part of the stratosphere and a primary maximum around 21 km. In the Arctic type (a) there is just one broad maximum centered at 100 mb. In the subtropical type profile (c) there is only one large maximum, and there is relatively little ozone in the lower stratosphere.

Retrieval of vertical profiles in general requires the inversion of remote sensing observations which involve strongly overlapping contribution functions. Consequently, there are significant inaccuracies in the determination of the vertical profile. Ozonesondes have been used to directly determine the vertical structure, but these observations are limited in both space and time. Total ozone has been studied longer and is known with greater accuracy than the vertical structure; therefore, total ozone will be used in many of the comparisons of the model results to

observation. Using total ozone for comparison has the disadvantage that model errors might be masked during the integration to form the total amounts.

Many of the features of the behavior of total ozone have been known since the early days of ozone observations. Using a network of only six instruments, Dobson (1930) deduced that both total ozone and the temporal variation of total ozone increased with increasing latitude. Perhaps the most extensive documentation of analyzed total ozone observations is the atlas of London et al. (1976) which contains ten years of data from 1957-1967. The average latitude-time cross section of total ozone in Dobson Units (DU) is shown in Fig. 2.8 and is not greatly different in the Northern Hemisphere from the cross section published by Götz (1951). This figure shows a rapid buildup of ozone at middle and high latitudes during winter. The maximum is reached at polar latitudes around the first of March. During early winter the maximum amount of ozone is found to between 60° and 70° N. There is a lack of data north of 80° so that the results in the extreme north are of dubious quality.

The regularity of the total ozone field is demonstrated in Fig. 2.9 taken from Hilsenrath et al. (1979). These data from the NIMBUS 4 BUUV experiment show some fluctuation in the annual variation of total ozone, but the large scale features remain quite constant during the time series.

Figure 2.10 from Zullig (1973) further demonstrates the extent of the interannual variation in the total ozone field. In this figure the average amount of total ozone in the Northern Hemisphere

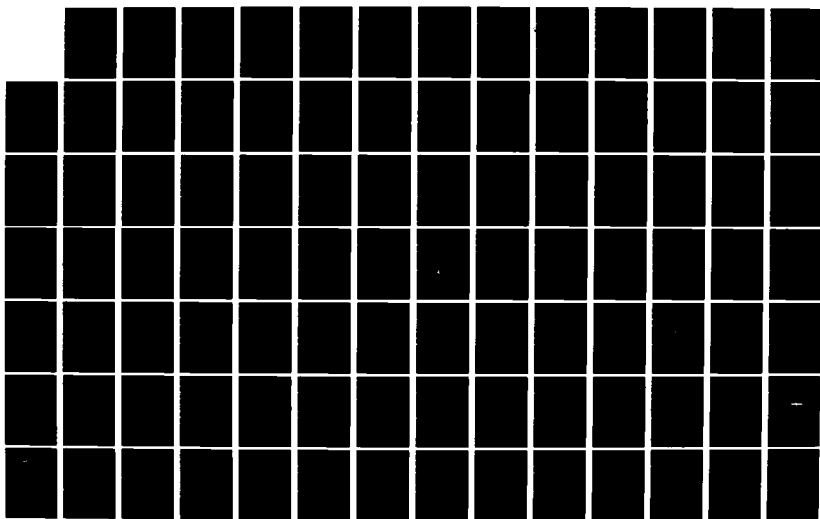
AD-A126 341

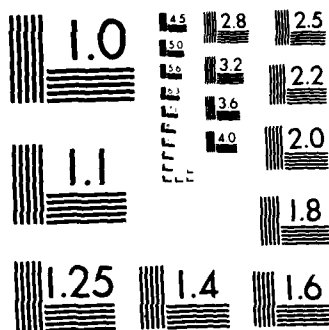
GEOPHYSICAL PLASMAS AND ATMOSPHERIC MODELING(U) SCIENCE 5/7
APPLICATIONS INC MCLEAN VA E HYMAN ET AL. FEB 83
SAI-83-144-WA N00014-81-C-2038

UNCLASSIFIED

F/G 4/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

north of 40° for three winters is shown. Even with the large amount of averaging involved in the formation of these graphs, there is an obvious temporal wave evident in the data. In the two winters during which warmings occur (1962/63, 1967/68), the maximum amount of ozone is approximately 6% larger, and the increase is more sudden, than in the winter during which no major warming occurs. Even though Zullig's observations are a rather limited amount of data on which to base transport theories, it is safe to conclude that major stratospheric warmings have a strong influence on the ozone distribution.

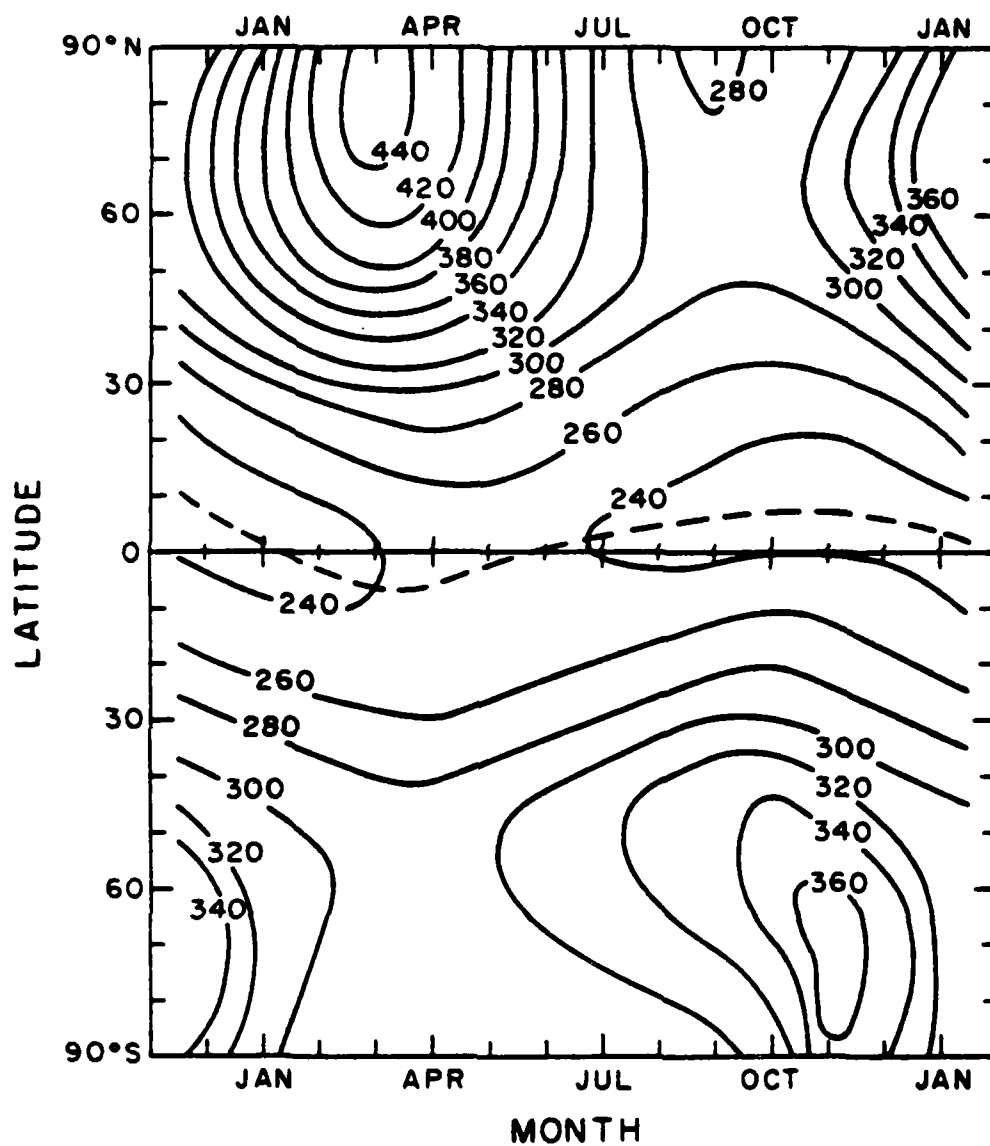


Figure 2.8. 10 year average latitude-time cross section
of total ozone in Dobson Units (DU)
(from London et al., 1976).

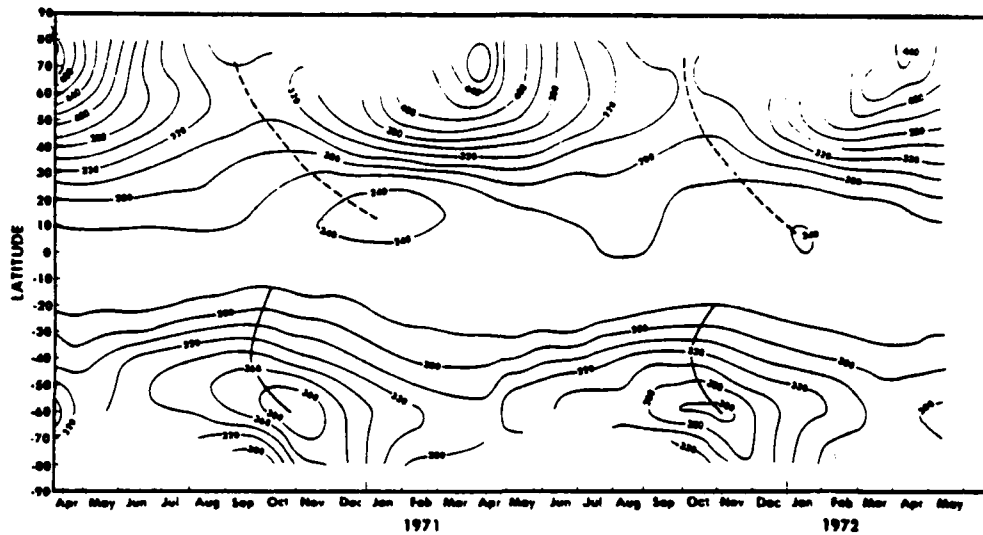


Figure 2.9. Latitude-time cross section of total ozone
data from Nimbus 4 BUUV (DU)
(from Hilsenrath et al., 1979).

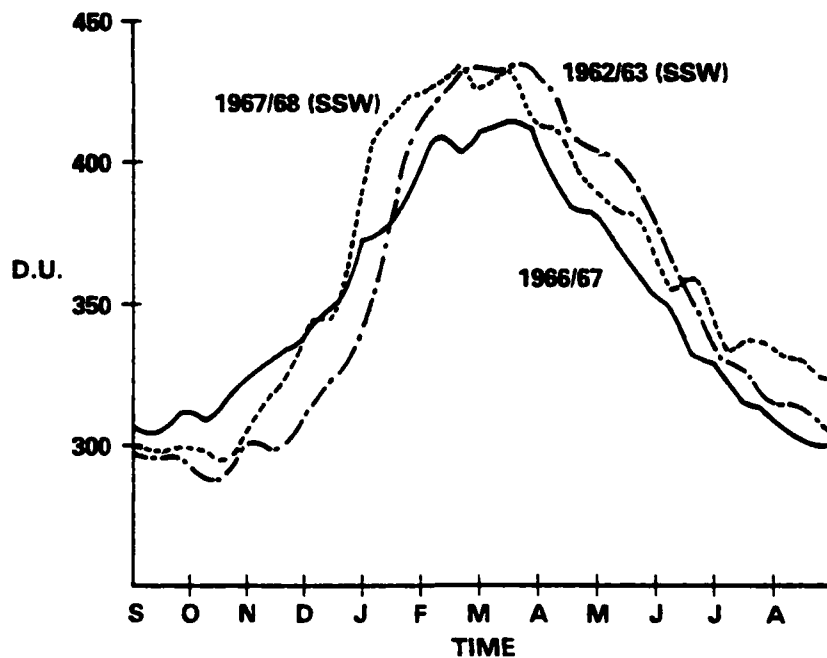


Figure 2.10. Mean total ozone north of 40°
(from Zullig, 1973).

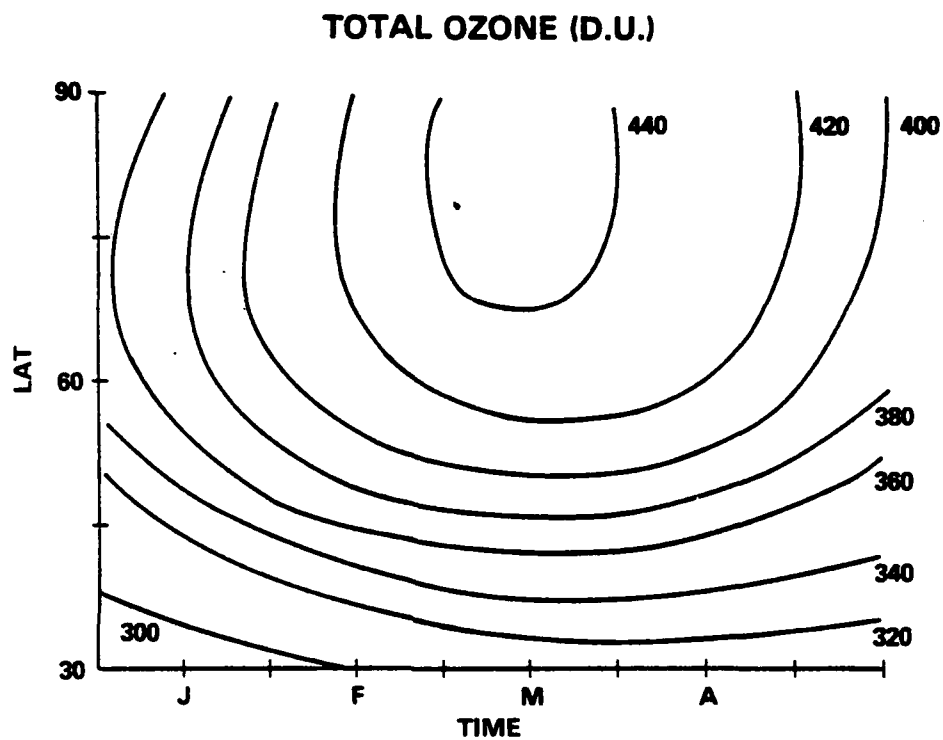


Figure 2.11 Reproduction of Figure 2.8 north of 30°
from mid-December through April.

The basic features of the ozone variability that will be studied here are the total ozone amount at extratropical latitudes and the transport of ozone into the lower stratosphere during winter. Total ozone increases at all latitudes during winter with the magnitude of the increase being largest near the pole. During early winter the maximum is in subpolar latitudes. Much of this increase in total ozone is associated with the injection of large amounts of ozone between 21 and 25 km. There is little change in the ozone above 40 km and the observed changes in this region can be explained photochemically. The total ozone reaches a maximum in early March at polar latitudes and decreases uniformly until it reaches a minimum in the fall.

The time dependent β -plane transport model will be used to run simulations from mid-December until April. In order to facilitate comparisons to observations a reproduction of the appropriate portion of the average total ozone field given by London et al. (1976) is presented in Fig. 2.11.

Chapter 3

TRANSPORT BY STEADY STATE PLANETARY WAVES

3.1 Introduction

In this chapter steady state transport by stationary planetary waves on a β -plane is presented. In Section 3.2 the continuity equation is developed for small amplitude waves in both the Eulerian and Lagrangian-mean formulations, and the relation between the residual circulation and the Lagrangian-mean velocity is derived. In Section 3.3 model results are presented and a Lagrangian model based on particle trajectories is given. The results are discussed and summarized in Section 3.4.

3.2 Constituent dynamics

The ozone continuity equation is formulated in both the Eulerian and Lagrangian-mean systems. In general, the constituent continuity equation is

$$\frac{Du}{Dt} = Q \quad (3.1)$$

where Q represents the source term. By decomposing the fields into

zonal averages and deviations and linearizing, the Eulerian formulation is obtained:

$$\frac{\partial \bar{\mu}}{\partial t} + \bar{u}_i \frac{\partial \bar{\mu}}{\partial x_i} + \overline{u_i' \frac{\partial \mu'}{\partial x_i}} = \bar{Q} \quad (3.2)$$

and

$$\frac{\partial \mu'}{\partial t} + u_i' \frac{\partial \bar{\mu}}{\partial x_i} + \bar{u}_i \frac{\partial \mu'}{\partial x_i} = Q' \quad (3.3)$$

where repeated indices indicate summation. It is pertinent to note that the zonal mean equation contains eddy flux terms which represent the effects of wave transport. The tendency for the eddy transport to be counteracted by the advection by the mean flow has been discussed extensively, and represents a major difficulty of atmospheric transport problems (Dunkerton, 1980; Mahlman and Moxim, 1978).

Using the Lagrangian-mean formulation (Andrews and McIntyre, 1978) it is possible to redefine the averaging operator such that wave forcing does not appear in the mean equation. Using this concept the mean continuity equation is written as

$$\left(\frac{\partial}{\partial t} + u_i^L \frac{\partial}{\partial x_i} \right) u^L = Q^L \quad (3.4)$$

In this system the time rate of change of the Lagrangian-mean ozone distribution is equal to the advection by the Lagrangian-mean velocity plus a source term.

The difference between the Lagrangian-mean and the Eulerian zonal mean is defined as the Stokes correction

$$\mu^S \equiv \mu^L - \bar{\mu}$$

Relating the Lagrangian-mean transform to the Eulerian zonal mean by the premise that the unperturbed fluid would be described by the Eulerian flow, it is possible to define the Stokes correction to second order as

$$\mu^S = \overline{\xi_i \frac{\partial \mu}{\partial x_i}} + \frac{1}{2} \overline{\xi_i \xi_j} \frac{\partial^2 \mu}{\partial x_i \partial x_j} \quad (3.5)$$

where $\xi_i = (\xi_1, \xi_2, \xi_3) = (\xi, \eta, \zeta)$ are the displacements from the equilibrium position of the fluid.

The displacement fields are related to the Eulerian perturbation velocity fields by the following equation (Matsuno and Nakamura, 1979):

$$\left(\frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} \right) \xi_i = u'_i + \xi_j \frac{\partial}{\partial x_j} \bar{u}_i + \dots \quad (3.6)$$

The mean meridional circulation is forced by the waves and is assumed to be second order in wave amplitude ($O(a^2)$). Therefore, the changes in the mean ozone field are $O(a^2)$ and the terms involving $\bar{v}$, $\bar{w}$ in (3.3) and (3.6) can be ignored. The perturbation equation can now be written to $O(a^2)$ as

$$\frac{\partial \mu'}{\partial t} + \bar{u} \frac{\partial \mu'}{\partial x} + \bar{v} \frac{\partial \mu_o}{\partial y} + \bar{w} \frac{\partial \mu_o}{\partial z} = -\lambda \mu' \quad (3.7)$$

Using the mass continuity equation, (3.2) may be written as

$$\frac{\partial \bar{\mu}}{\partial t} + \frac{\partial}{\partial y} (\bar{v} \bar{\mu}_o + \overline{v' \mu'}) + \frac{1}{\rho} \frac{\partial}{\partial z} \rho (\bar{w} \bar{\mu}_o + \overline{w' \mu'}) = -\lambda \bar{\mu} \quad (3.8)$$

where $\bar{\mu}$ now represents changes in the mean field from the equilibrium value $\bar{\mu}_o$.

The relation between the perturbation velocity fields and the displacement fields (Eq. 3.6) can be rewritten for steady waves as

$$ik\bar{u}\xi = u' + \eta \frac{\partial \bar{u}}{\partial y} + \zeta \frac{\partial \bar{u}}{\partial z}$$

$$ik\bar{u}\eta = v', \quad ik\bar{u}\zeta = w'$$

These are the equations used to calculate the Lagrangian-mean displacements.

Under the assumptions above, the following relationships between the Eulerian and the Lagrangian-mean quantities can be written for steady waves:

$$\frac{\partial \mu^L}{\partial t} = \frac{\partial \bar{\mu}}{\partial t} + O(a^4) \quad (3.9a)$$

$$v^s = \overline{\epsilon_1 \frac{\partial v}{\partial x_1}} + O(a^4) \quad (3.9b)$$

$$w^s = \overline{\epsilon_1 \frac{\partial w}{\partial x_1}} + O(a^4) \quad (3.9c)$$

$$\bar{Q}^L = \bar{Q} + \overline{\epsilon_1 \frac{\partial Q}{\partial x_1}} \quad (3.9d)$$

Therefore (3.4) can be written to the same accuracy as (3.8) as

$$\frac{\partial \bar{\mu}}{\partial t} = -u_1^L \frac{\partial \bar{\mu}_0}{\partial x_1} - \overline{\epsilon_1 \frac{\partial \lambda \mu}{\partial x_1}} - \lambda \bar{\mu} \quad (3.10)$$

The first term on the right hand side of (3.10) will be referred to as the advective term; the second term is the "stirring" term; and the third term is the mean chemistry. The stirring term represents the effects of the planetary waves moving ozone parcels through a region of varying photochemistry. This term will be discussed in detail later. Equations (3.7), (3.8), and (3.10) are the ozone continuity equations used in this study.

A number of authors have suggested that the Lagrangian-mean velocity may be a suitable quantity to approximate the large scale transport of ozone in chemical models (Dunkerton, 1978; Miller et al., 1981; McIntyre, 1980). However, since the Lagrangian-mean velocity can be difficult to calculate, it has also been suggested that the residual circulation defined by Andrews and McIntyre (1976)

might be an adequate approximation for the Lagrangian-mean velocity (Holton, 1981). Strobil (1981) has derived a general linearized 2-D transport model using the residual circulation for waves with complex frequencies. The relation between the residual velocity and the Lagrangian-mean velocity for steady, dissipating planetary waves is outlined below.

The eddy mass continuity equation can be written in terms of the parcel displacement fields as

$$\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} + \frac{1}{\rho} \frac{\partial \rho \zeta}{\partial z} = 0 \quad (3.11)$$

Using (3.11) and the definition of the Stokes corrections it is possible to write

$$v^L = \bar{v} + \frac{1}{\rho} \frac{\partial}{\partial z} \overline{\rho v' \zeta} \quad (3.12)$$

and

$$w^L = \bar{w} - \frac{\partial}{\partial y} \overline{v' \zeta} \quad (3.13)$$

The residual circulation for the current problem is

$$v^* = \bar{v} - \frac{1}{\rho N^2} \frac{\partial}{\partial z} \rho \overline{v' \phi_z'} \quad (3.14)$$

and

$$\bar{w}^* = \bar{w} + \frac{1}{N^2} \frac{\partial}{\partial y} \overline{v' \phi_z'} \quad (3.15)$$

Using the eddy thermodynamic equation

$$(a + \bar{u}ik) \phi_z' - f\bar{v}' \frac{\partial \bar{u}}{\partial z} + \bar{u}ikN^2\zeta = 0$$

and multiplying by v' and averaging, it can be shown that

$$\overline{v' \phi_z'} = - \frac{N^2 i k \bar{u} \overline{v' \zeta} - f \bar{v}'^2 \frac{\partial \bar{u}}{\partial z}}{(a + i k \bar{u})} \quad (3.16)$$

Using (3.16) and comparing (3.12) and (3.13) with (3.14) and (3.15) it is clear that the Lagrangian-mean velocity and the residual velocity are equal only when $a = 0$. When $a = 0$, however, both the Lagrangian-mean and the residual circulation are zero and the transport by the waves and the mean flow exactly counterbalance each other. Since a is an order of magnitude or more smaller than $k\bar{u}$ below the sponge layer in this model, the Lagrangian-mean and residual circulations should be similar. When transience or external heating and cooling processes are allowed, the relationship between the two circulations becomes more complex. In Holton's (1981) parameterization, the difference between the Lagrangian-mean and residual circulations is contained in the antisymmetric part of the diffusion tensor and is assumed to be negligible.

In the next section the Eulerian velocities, the Lagrangian-mean quantities, and the residual circulation will be presented for

steady, dissipative planetary waves. Zonal mean ozone changes can then be calculated and the various transport methods can be compared.

3.3 Results

The zonal mean ozone distribution used in the steady state model is shown in Fig. 3.1. This distribution is based on the October/November values given by Dütsch (1969) and is similar to the distribution used by Hartmann and Garcia (1979).

a. Dynamic quantities

The $\bar{v}$ and $\bar{w}$ fields are shown in Fig. 3.2. These fields, as expected, show rising motion in the northern regions of the β -channel and sinking motion in the southern regions. The meridional velocity is from north to south. This flow is forced by the northward eddy heat transport of the planetary waves.

The Stokes drifts as calculated with (3.9b) and (3.9c) are shown in Fig. 3.3. These drifts can be viewed as the effective velocity fields associated with the wave. The Stokes drifts are directed in an opposite sense to the Eulerian mean velocity and are of approximately the same magnitude. This demonstrates the tendency for the wave fields to counterbalance the mean fields.

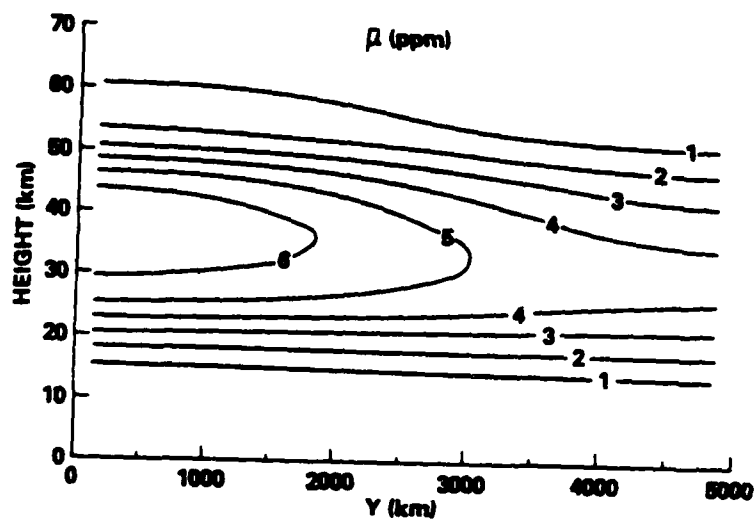


Figure 3.1. Zonal mean ozone distribution
for the steady state model (ppm).

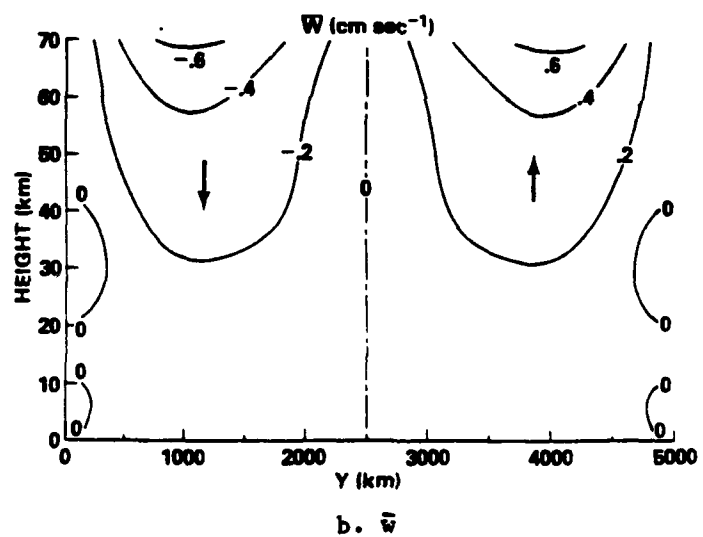
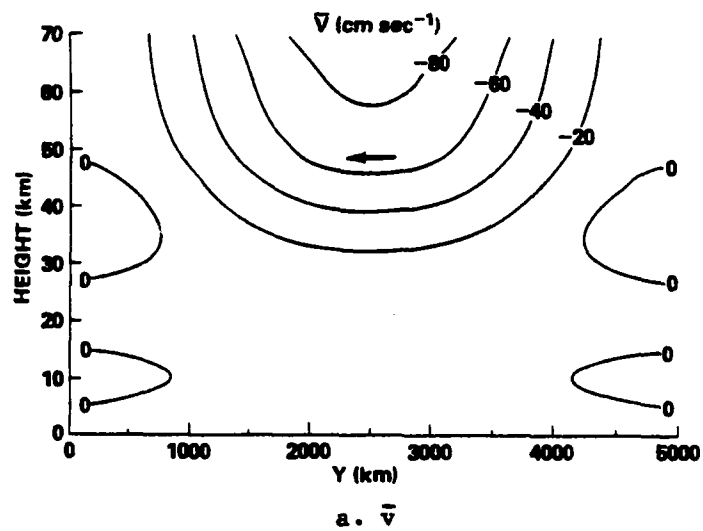
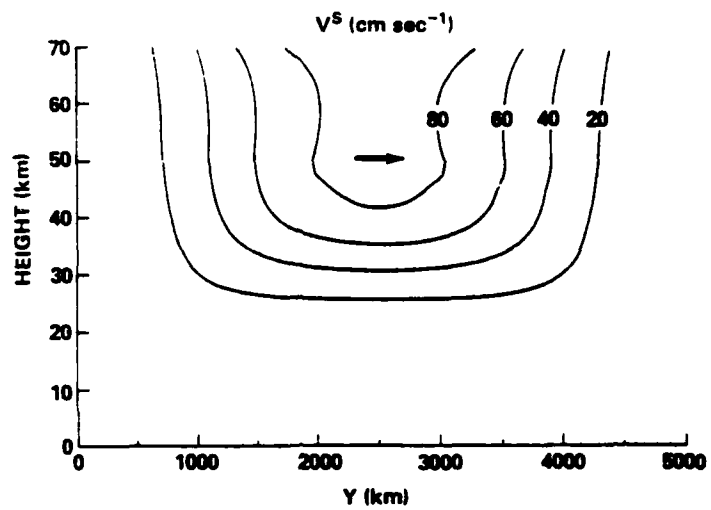
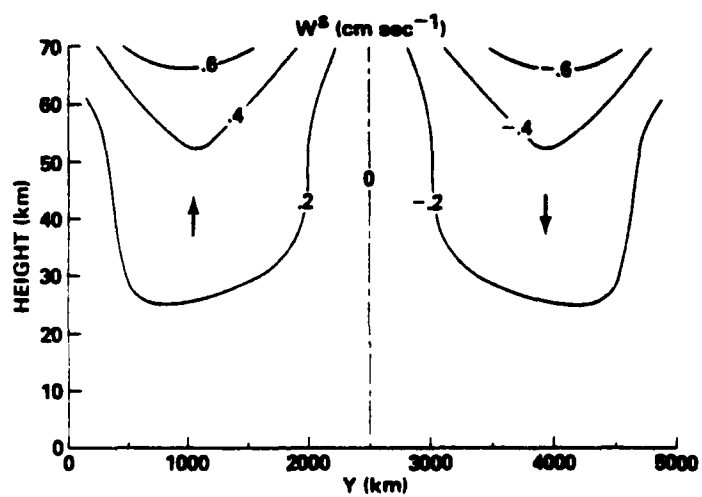


Figure 3.2. Eulerian zonal mean velocities (cm sec^{-1}).



a. V^S

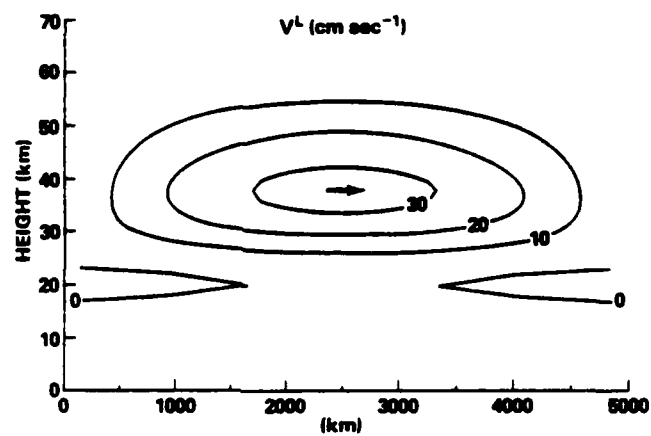


b. W^S

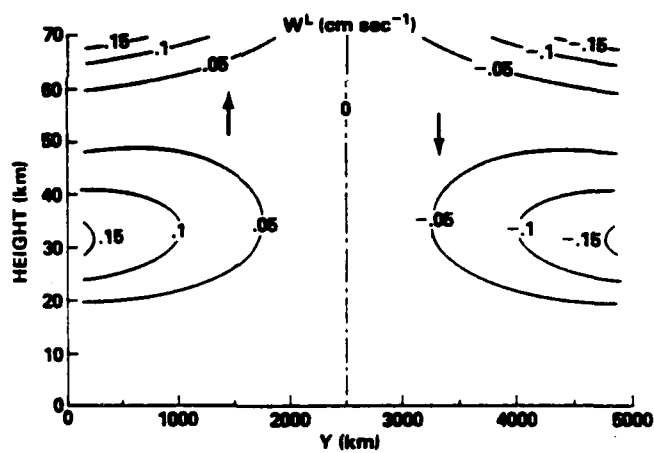
Figure 3.3. Stokes drifts (cm sec⁻¹).

The Lagrangian-mean velocity fields, which are the sums of the Eulerian zonal mean and the Stokes drifts, are shown in Fig. 3.4. The sense of the circulation is rising motion in the southern regions and sinking motion in the polar regions. This motion is the same as indicated by Schoeberl (1981a). Furthermore, this motion would tend to enhance the diabatic circulation as calculated by Dunkerton (1978). The Lagrangian-mean velocity is of much smaller magnitude than either the Stokes drifts or the Eulerian mean velocities.

The residual circulation is shown in Fig. 3.5. It is geometrically similar to the Lagrangian-mean velocity but the meridional component is approximately 30% larger than the Lagrangian-mean meridional velocity. The largest differences between the two circulations occur in the region where $\partial \bar{u} / \partial z$ is largest. Similar increases are calculated in the vertical residual velocity field.

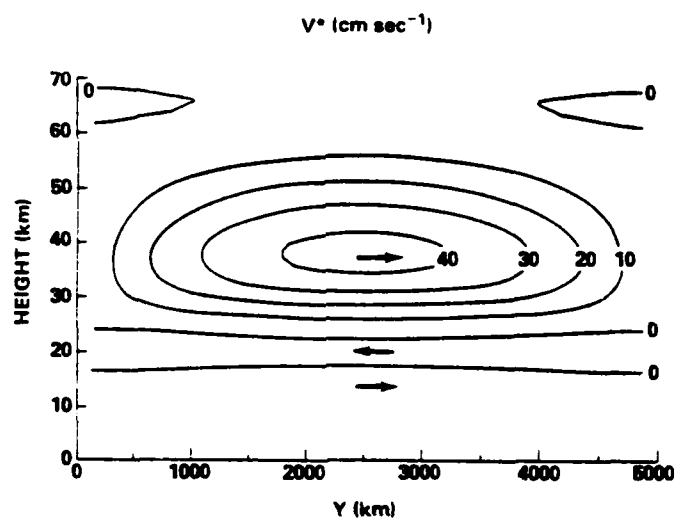


a. V^L

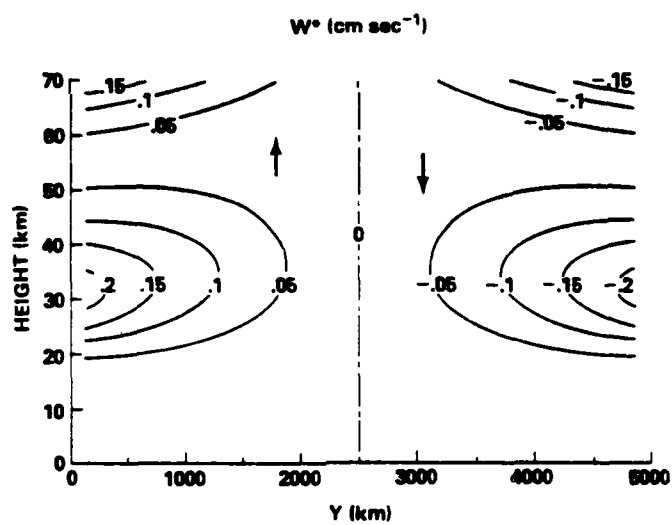


b. W^L

Figure 3.4. Lagrangian-mean velocities (cm sec⁻¹).



a. V^*



b. W^*

Figure 3.5. Residual velocities (cm sec^{-1}).

b. Eulerian transport

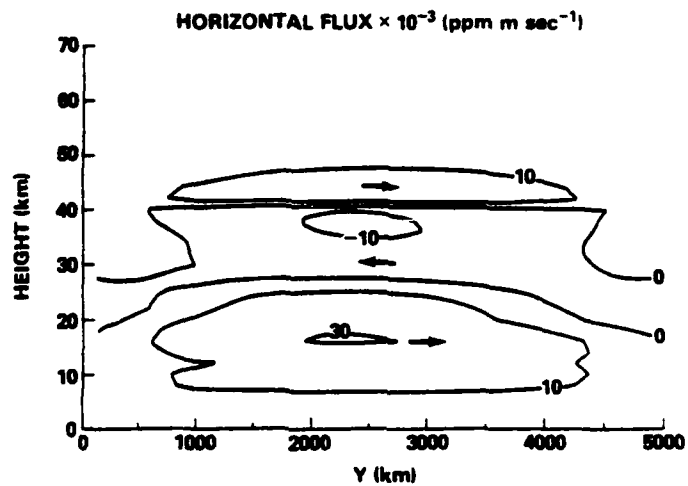
The continuity equations (3.7), (3.8), and (3.10) are solved using the appropriate dynamic quantities. Since the mean chemistry appears in the same form in both (3.8) and (3.10) it is ignored in order that it might not obscure the differences between the two formulations. The effects of the mean chemistry term will be considered separately.

The horizontal ($\overline{v}\mu + \overline{v'}\mu'$) and vertical ($\overline{w}\mu + \overline{w'}\mu'$) fluxes as calculated by the Eulerian method are shown in Fig. 3.6. The vertical flux is upward in the northern region of the model and downward in the southern region. The vertical flux is confined primarily to the conservative region and reflects the Eulerian-mean vertical velocity field.

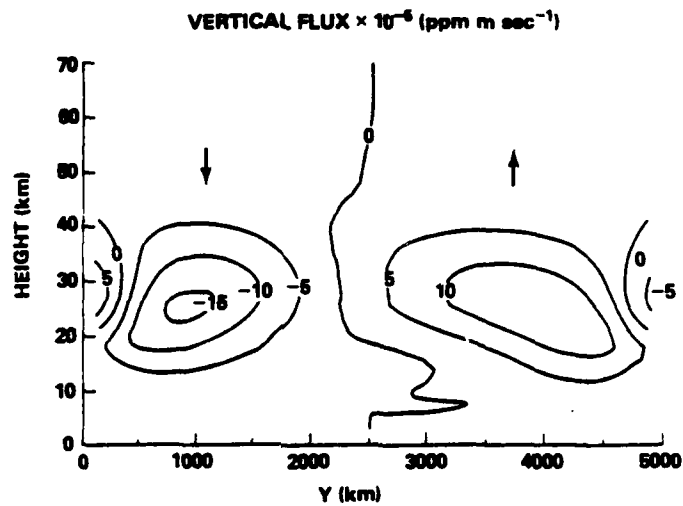
The horizontal flux field is quite complicated showing two regions of northward transport separated by a region of southward transport. Once again the flux is largest in the conservative region. The lower region of northward flux is dominated by eddy flux as described by Clark and Rogers (1978). The region of southward flux is dominated by the mean meridional velocity. The upper region of northward transport is once again dominated by the eddy flux. In the upper most regions of the model both the horizontal and vertical fluxes are very small due to rapid photochemical relaxation.

The rate of change of the mean ozone field is calculated as the divergence of the fluxes. This field is shown in Fig. 3.7. There is an increase of ozone in the northern section and a decrease in

the southern section. Despite the fluxes' being largest in the conservative region, the mean field change is concentrated almost entirely in the transition region in agreement with the results of Hartmann and Garcia (1979).



a. horizontal flux



b. vertical flux

Figure 3.6. Eulerian fluxes of ozone (ppm m sec $^{-1}$).

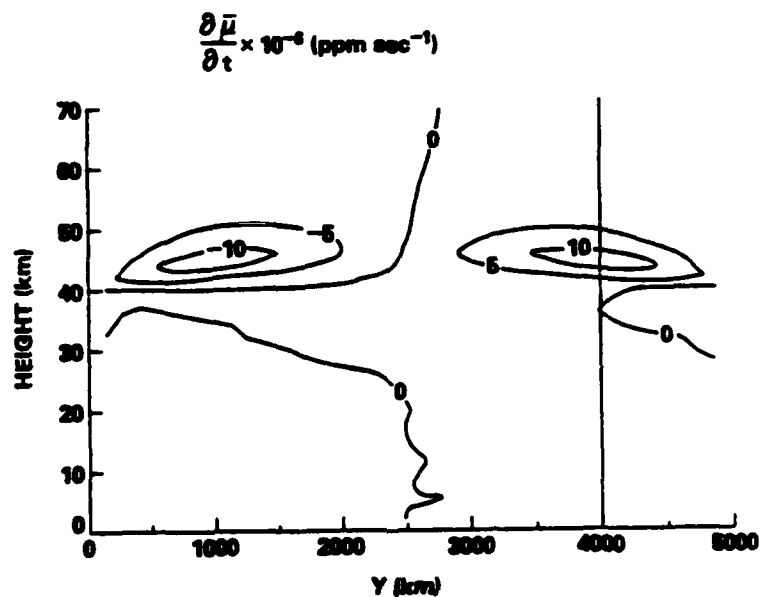


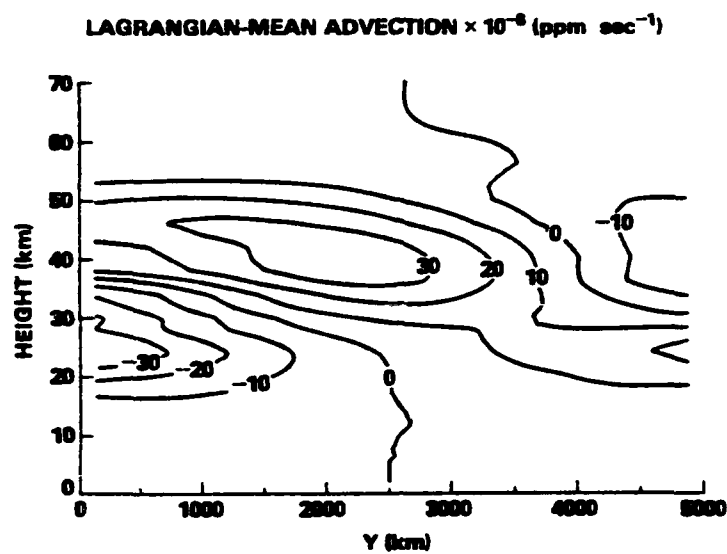
Figure 3.7. Mean ozone tendency, $\frac{\partial \bar{\mu}}{\partial t}$, as calculated with the Eulerian formulation (ppm sec⁻¹).

c. Lagrangian-mean transport

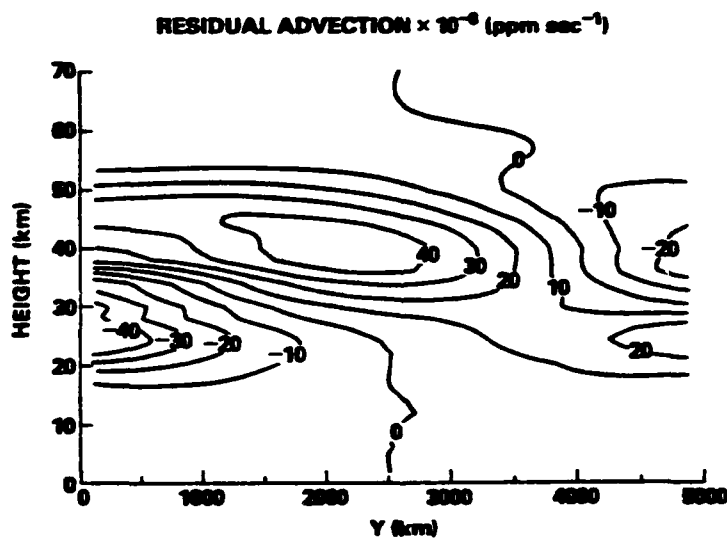
The Lagrangian-mean results are shown in Figures 3.8-3.10. The advective term as defined in Eq. (3.10) is shown in Fig. 3.8. In the extreme polar region there is a decrease of ozone above the ozone maximum and a buildup of ozone in the lower stratosphere. In the southern regions there is a decrease in the lower stratosphere and an increase in the upper stratosphere. These changes are primarily due to the vertical advection. There is a strong transport of ozone northward by the meridional Lagrangian-mean velocity in the midlatitudes. The Lagrangian-mean velocity associated with planetary waves advects ozone poleward and downward.

The advective term calculated using the residual circulation shows increases and decreases in the mean ozone field in the same regions as the Lagrangian-mean calculations. In all regions the magnitude of advection by the residual circulation is larger. Differences as large as 30% are calculated reflecting the increased magnitude of the residual circulation.

Comparison of Figures 3.10 and 3.7 show that the mean field changes are the same when calculated by the Eulerian or Lagrangian-mean methods. Substituting the residual velocity for the Lagrangian-mean velocity causes no significant changes in the calculation of the mean field changes.



a. Lagrangian-mean velocity



b. residual velocity

Figure 3.8. Mean ozone tendency
caused by advection (ppm sec $^{-1}$).

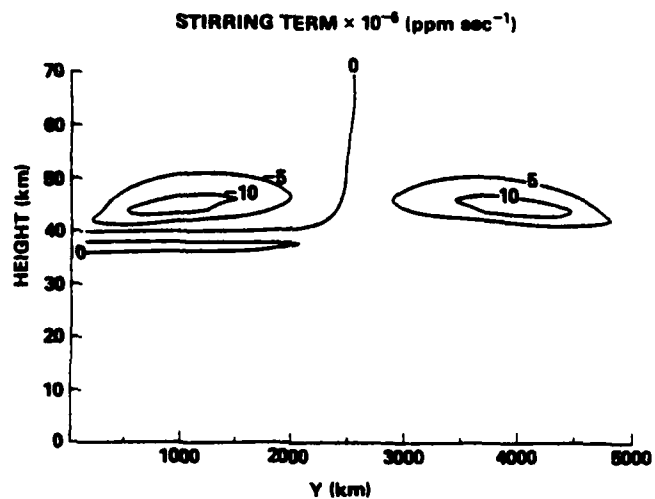


Figure 3.9. Total mean ozone tendency as calculated with the Lagrangian formulation (ppm sec^{-1}).

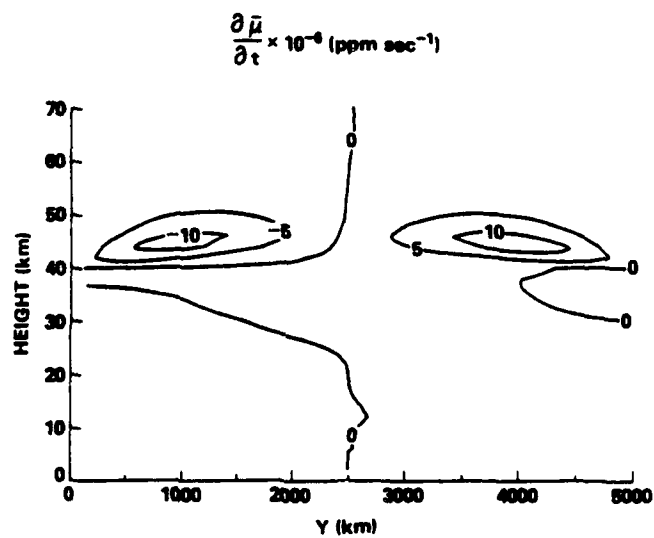


Figure 3.10. Mean ozone tendency caused by "stirring" (ppm sec^{-1}).

d. The "stirring" term

The magnitude of the stirring term is shown in Fig. 3.9. This field quite accurately represents the ozone time rate of change shown in Fig. 3.10. The advection by the Lagrangian-mean velocity is less than 1% of the stirring term in regions where ozone is photochemically active. The stirring term is identically zero when the photochemistry is zero.

In order to understand the physical meaning of the stirring term and its analogy to the Eulerian problem, it is helpful to compare the stirring term to the eddy flux. Assuming that the wave is steady and stationary (3.7) becomes

$$\mu' = - \frac{\lambda - i k \bar{u}}{\lambda^2 + (k \bar{u})^2} u_1' \frac{\partial \bar{\mu}_o}{\partial x_1} \quad (3.17)$$

The change in the zonal mean ozone field due to the divergence of the eddy flux (DEF) is

$$\begin{aligned} \text{DEF} &= - \frac{1}{\rho} \frac{\partial}{\partial x_j} \rho \overline{u_j' \mu'} \\ &= \frac{1}{\rho} \frac{\partial}{\partial x_j} \rho \left\{ \frac{\lambda}{\lambda^2 + (k \bar{u})^2} \overline{\text{Re}(u_j') \text{Re}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} + \frac{k \bar{u}}{\lambda^2 + (k \bar{u})^2} \overline{\text{Re}(u_j') \text{Im}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} \right\} \end{aligned} \quad (3.18)$$

The stirring term (ST) is

$$\begin{aligned}
 ST &= - \overline{\epsilon_j \frac{\partial \lambda \mu}{\partial x_j}} = - \frac{1}{\rho} \frac{\partial}{\partial x_j} \overline{\rho \lambda \epsilon_j \mu} \\
 &= \frac{1}{\rho} \frac{\partial}{\partial x_j} \frac{\rho \lambda}{k\bar{u}} \left\{ \frac{\lambda}{\lambda^2 + (k\bar{u})^2} \overline{\text{Im}(u_j^*) \text{Re}(u_1^*)} \frac{\partial \bar{\mu}_0}{\partial x_1} + \frac{k\bar{u}}{\lambda^2 + (k\bar{u})^2} \overline{\text{Re}(u_j^*) \text{Re}(u_1^*)} \frac{\partial \bar{\mu}_0}{\partial x_1} \right\}
 \end{aligned}
 \tag{3.19}$$

ST is functionally very similar to DEF. In order to investigate the precise similarities it is useful to consider the limits of λ relative to the dynamic frequency $k\bar{u}$.

For $\lambda/k\bar{u} < 1$,

$$\frac{1}{\lambda^2 + (k\bar{u})^2} \approx \frac{1}{(k\bar{u})^2} \left(1 - \frac{\lambda^2}{(k\bar{u})^2} \right)
 \tag{3.20}$$

Using (3.20), (3.18) and (3.19) can be rewritten to lowest order as:

$$\begin{aligned}
 DEF &= \frac{1}{\rho} \frac{\partial}{\partial x_j} \rho \left\{ \frac{\lambda}{(k\bar{u})^2} \overline{\text{Re}(u_j^*) \text{Re}(u_1^*)} \frac{\partial \bar{\mu}_0}{\partial x_1} + \frac{1}{(k\bar{u})} \overline{\text{Re}(u_j^*) \text{Im}(u_1^*)} \frac{\partial \bar{\mu}_0}{\partial x_1} \right\}
 \end{aligned}
 \tag{3.21}$$

and

$$ST = \frac{1}{\rho} \frac{\partial}{\partial x_j} \frac{\rho \lambda}{(k\bar{u})^2} \overline{\text{Re}(u_j') \text{Re}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} \quad (3.22)$$

Thus for small λ the divergence of the eddy flux can be divided into a purely chemical and purely conservative part. The conservative part, the second term on the RHS of (3.21) is the conservative flux described by Clark and Rogers (1978). The chemical part of DEF is identical to ST; therefore, for small λ it can be explicitly shown that the stirring term represents the chemical contribution of the divergence of the eddy flux.

For large λ DEF approaches zero, while ST approaches some finite value. In fact, as λ becomes large ST loses its explicit chemistry dependence and becomes

$$ST = \bar{u}_1^s \frac{\partial \bar{\mu}_c}{\partial x_1} \quad (3.23)$$

At this limit both the Eulerian mean, (3.8), and Lagrangian-mean, (3.10) continuity equations become

$$\frac{\partial \bar{\mu}}{\partial t} = - \bar{u}_1 \frac{\partial \bar{\mu}_o}{\partial x_1} - \lambda \bar{\mu} \quad (3.24)$$

Eq. (3.24) indicates that, when the photochemistry is strong, the wave effects become small so that the mean field changes are governed by the mean chemistry and the advection by the Eulerian-mean velocity.

e. Lagrangian model

While the analysis above reveals the relationship of the stirring term to the divergence of the eddy flux, the physical nature of the stirring term remains obscured in the mathematics. Hartmann (1981) showed how pure oscillatory motion of a fluid parcel could produce changes in the mean ozone density in a fluid tube. A similar Lagrangian approach can be used here to investigate the effects of the planetary wave and to further coalesce the Lagrangian-mean and Eulerian viewpoints.

The Lagrangian model is based on the particle trajectories as determined by the Lagrangian-mean displacement fields. Using (2.3) and assuming that $\bar{u}$ is constant and F' is zero, it can be shown that for harmonic disturbances the horizontal displacement field, η , and the vertical displacement field, ζ are given by

$$\eta = \eta_0 \cos(\alpha z + kx) \quad (3.25)$$

and

$$\zeta = \zeta_0 \cos(\alpha z + kx) + \zeta_1 \sin(\alpha z + kx) \quad (3.26)$$

where α is the vertical wavenumber. The coefficients, η_0 , ζ_0 , and ζ_1 are functions of height, meridional distance, and the dynamic parameters. For a particular value of y and z , the projection of η and ζ on the y, z plane is an ellipse (Fig. 3.11). These displacement fields are similar to those calculated by Matsuno (1980) and Danielsen (1981).

The displacement fields provide a model of the parcel motion associated with a planetary wave. The planetary wave displaces a particular parcel from its equilibrium position and causes the parcel to move in a helical trajectory around the globe. The Lagrangian-mean velocity is the velocity at which the axis of the helix moves. For the small amplitude, steady state waves of this model the vertical extent of the ellipse is on the order of 1-3 km and the horizontal extent is on the order of hundreds of kilometers.

First consider a parcel in the absence of chemistry. The equilibrium density of the parcel, μ_p , is $\bar{\mu}_0(0,0)$. If the parcel is displaced to (η, ζ) , then the perturbation density μ' is equal to the parcel density minus the background density. If the displacements are small, then

$$\mu' = -\eta \frac{\partial \bar{\mu}_0}{\partial y} - \zeta \frac{\partial \bar{\mu}_0}{\partial z}$$

Assuming that $\bar{\mu}_0$ is a function of z only and that it increases with height, then positive ozone perturbations are generally associated with northward motion and negative ozone perturbations are associated with southward motion (see Fig. 3.11). This is equivalent to northward eddy flux of ozone. Hence this is the Lagrangian description of the conservative flux as defined by Clark and Rogers (1978). This conservative flux is largely cancelled by the wave induced mean circulation.

In the case when ζ_1 in (3.26) is zero, which corresponds to an evanescent wave, the trajectories then become lines instead of

ellipses. Then the perturbations are equally correlated with both northward and southward velocity and the conservative eddy flux is identically zero.

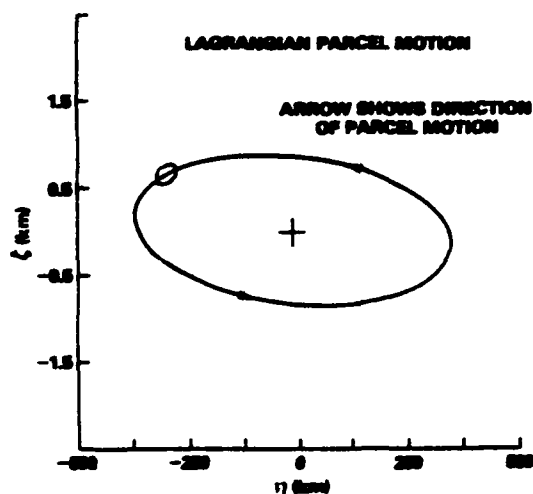


Figure 3.11. Projection of displacement fields η, ζ on the y, z plane (km).

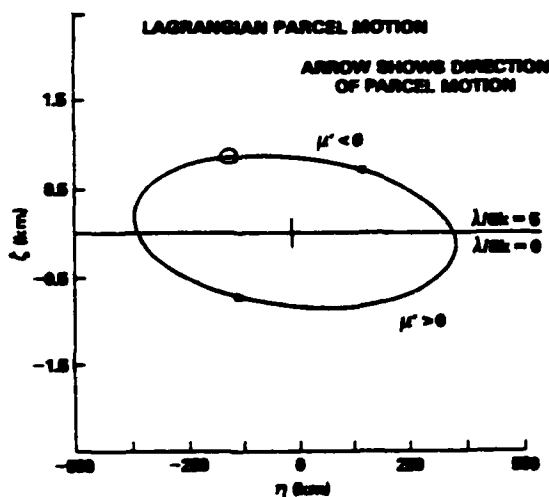


Figure 3.12. Parcel orbiting through a region of discontinuous photochemistry $\frac{\partial \bar{\mu}}{\partial z} > 0$.

When chemistry is present μ' still changes due to the varying background field, but the chemistry tries to bring the parcel's ozone density towards the background value. The constituent density for the parcel is determined by

$$\frac{d\mu'}{dt} = S - \lambda\mu' \quad (3.27)$$

where S represents the effective source caused by the parcel moving through the varying background density. For elliptical trajectories S is a harmonic function in time, proportional to the magnitude of the velocity and the background constituent gradients. In general, as the parcel orbits the ellipse, it goes through varying photochemistry so that λ is also a function of time. In a reference frame moving with the zonal velocity of the parcel, (3.27) can be rewritten as

$$\frac{d\mu'}{dt} = S_0 \cos \omega t - \lambda(t)\mu' \quad (3.28)$$

where $\omega = k\bar{u}$.

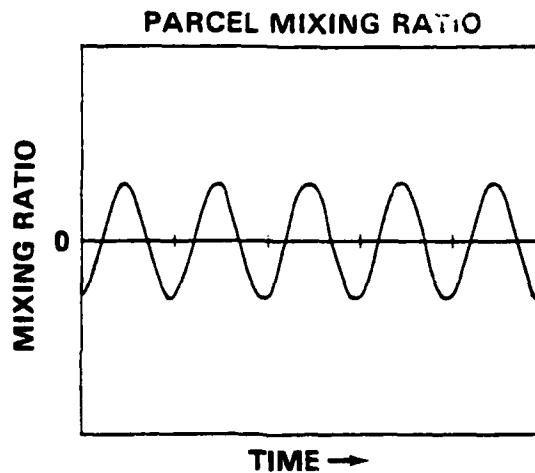
If λ is constant, then the solution to (3.26) for large t can be written as

$$\mu' = \frac{S_0}{\lambda^2 + \omega^2} (\lambda \cos \omega t + \omega \sin \omega t) \quad (3.29)$$

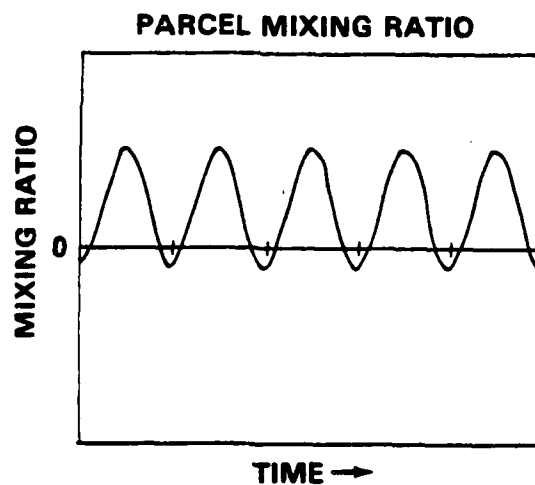
For large λ , μ' approaches zero which is equivalent to the previous result that the eddy fluxes disappear in the fast photochemical

region. For $\lambda = 0$ the solution reduces to the ozone density appropriate to the conservative problem. Note that the average of μ' over one period is zero.

In the regions where λ varies strongly over the extent of the ellipse, the perturbation can take on a different character. Such regions would be expected to exist in the lower parts of the transition layer and at the polar night boundary where the chemistry is nearly discontinuous. Figure 3.12 shows an extreme example where the chemistry is discontinuous across an ellipse. As the parcel travels through the conservative region the perturbation density varies as S . In the transition region ($\lambda/\bar{k}u = 5$) the perturbation is decreased by $\sim 80\%$. Assuming the background density increases with height, this behavior means that the negative perturbations are selectively damped while the positive perturbations are not. Therefore, the average value of μ' over one period, which was zero when λ did not vary, may become nonzero in the presence of extreme chemical gradients. Such regions would be expected to have a great effect upon transport calculations. This effect is analogous to mean density changes in a fluid tube as calculated by Hartmann (1981).



a. constant photochemistry



b. photochemistry at the bottom of
the transition layer

Figure 3.13. Solution, at large t , of perturbation ozone density for an elliptically orbiting parcel.

Figure 3.13 demonstrates this effect for a more realistic problem. Figure 3.13a shows (3.29), the steady state solution to (3.28). This is a harmonic function with zero average. Figure 3.13b shows the numerical solution to (3.24) for an ellipse situated at the bottom of the transition layer with the model photochemistry. The steady state solution has a non-zero time average, as expected.

f. The effects of eddy transport on the mean field

Since the changes in the mean field have been assumed to be $O(a^2)$, the calculation of the steady state forcing by the eddies on the mean field is independent of the mean constituent density. Therefore, at any point in space it is possible to write the mean field equation (3.8) or (3.10) as

$$\frac{\partial \bar{\mu}}{\partial t} = -\lambda \bar{\mu} + F \quad (3.30)$$

where F is the eddy forcing which represents the divergence of both the horizontal and vertical fluxes. F is independent of time; therefore, the solution to (3.30) can be written as

$$\bar{\mu} = \frac{F}{\lambda} (1 - e^{-\lambda t}) \quad (3.31)$$

where it has been assumed that $\bar{\mu}$, the change in the equilibrium field, is zero at time zero. As $t \rightarrow \infty$ the change in the mean field is F/λ .

In the conservative region the mean continuity equation can be written as $\frac{\partial \bar{\mu}}{\partial t} = F$ which yields

$$\bar{\mu} = Ft \quad (3.32)$$

In order to investigate the effect of the steady state waves on the mean field the appropriate equation (3.31) or (3.32) was solved at $t = 90$ days. This calculation would represent the effects of the waves if they persisted for the entire winter. The forcing is taken from the line in Fig. 3.7 through the region of maximum ozone increase.

Figure 3.14 shows the change in $\bar{\mu}$. The large rates of change that are calculated in the transition region are almost completely counterbalanced by the mean chemistry. At the altitude of maximum increase in Fig. 3.7 the actual change in the mean field is only .5 ppm for 90 days.

The largest changes in the mean field are in the lower stratosphere where changes of 1.4 ppm are seen at 90 days. These changes are comparable to those observed in the northern hemisphere. Immediately below the transition region is a region of depletion caused by the advection. This vertical structure of the changes in the mean field is unrealistic and due largely to the simplified representation of photochemistry and dynamics in the lower stratosphere.

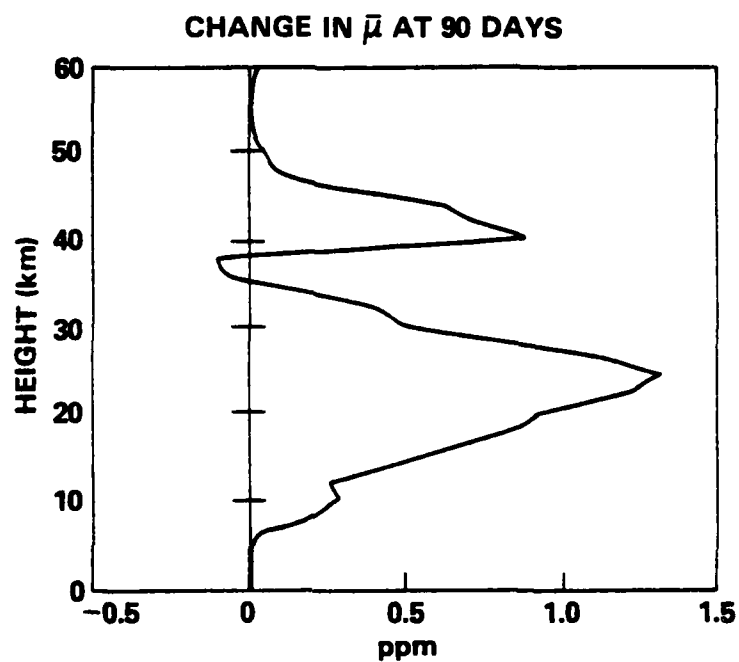
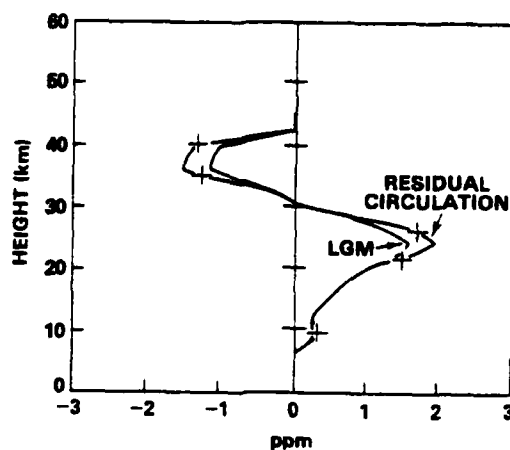


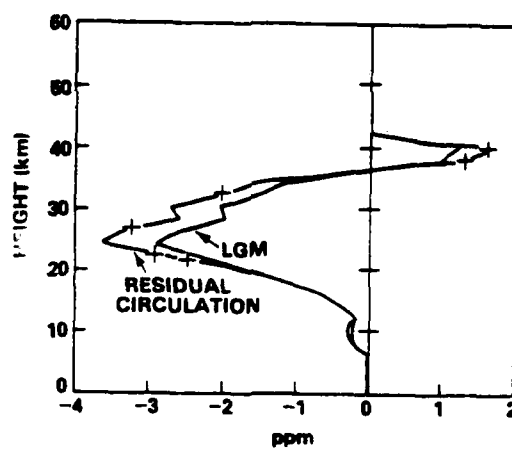
Figure 3.14. 90 day zonal mean ozone change, with the mean chemistry included, at the latitude of the maximum ozone increase in Fig 3.7 (ppm).

CHANGE IN $\bar{\mu}$ AT 90 DAYS



a. Southern

CHANGE IN $\bar{\mu}$ AT 90 DAYS



b. Northern

Figure 3.15. 90 day zonal mean ozone change, with the mean chemistry included, at the northern and southern extremes, as determined by Lagrangian-mean advection (LGM) and residual velocity advectons (+) (ppm).

The relatively large positive changes in the mean field at the bottom of the transition layer are in the region where the photochemistry gradients are large. This is the region where a parcel takes on a non-zero average during a parcel orbit. The equilibrium values at this altitude are reached after 5 days of integration. This region of large increase just above a region of decrease demonstrates an important model sensitivity to the chemical structure.

Since the advection term is responsible for all of the change in the mean field below the transition region it is worthy of further investigation. Mean field tendencies for 90 day periods were calculated 400 km from both the northern and southern boundaries using both the residual and Lagrangian-mean advection (Fig. 3.15). The large changes in the midlatitudes caused by meridional transport are largely negated by the photochemistry.

In the southern region the change in the mean field by advection is unlike that observed in the stratosphere. There is a depletion of nearly 4 ppm at 24 km and an enhancement of almost 2 ppm at 40 km. In the northern region the changes in the lower stratosphere are not unreasonable showing a 1.5 ppm increase at 24 km. However, at 36 km a 1 ppm depletion occurs. It is obvious that the modeling of the lower stratosphere is inadequate. Both small scale dynamics and photochemistry which may act to counter the tendencies shown here have been neglected. The advection calculated using the residual circulation produces larger changes in the mean field than that calculated using the Lagrangian-mean velocity. This

means, that in this model, the antisymmetric part of Holton's (1981) diffusion tensor is not negligible.

3.4 Conclusions

The changes in zonal mean ozone as forced by steady state planetary waves have been calculated using both a Lagrangian-mean and Eulerian formulation. Although the calculated results from the two methods are the same, the use of the two methods allows for increased interpretation of the results.

In the Eulerian formulation the tendency for the eddy and mean fluxes to cancel is well illustrated. Both the horizontal and vertical flux fields are complex and largest in the conservative region. However, the changes in the mean field in the absence of the mean photochemistry are largest in the transition region as found by Hartmann and Garcia (1979).

The problem of compensation by two large terms which appears in the Eulerian framework also appears in the calculation of the Lagrangian-mean velocity. If the Lagrangian-mean flow could be calculated directly, then it might be possible to avoid the numerical difficulties of compensation. There seems to be no computational advantage of the Lagrangian-mean method to calculate planetary wave transport as it requires the formulation of Eulerian quantities, their conversion to Lagrangian-mean quantities, and then the calculation of the transport.

In the Lagrangian-mean formulation the stirring term is dominant in the absence of mean chemistry. The stirring is equal to

the chemical contribution of the divergence of the eddy flux when $\lambda/\bar{u}k < 1$. When λ is large the stirring term is equivalent to the Stokes advection. Using particle displacement fields it is possible to predict the type of eddy transport that is to be expected in a certain region. It was found that μ' has a non-zero time average in the presence of strong chemical gradients. Calculation of the mean field changes shows that the mean chemistry is able to compensate for the eddy chemistry changes except where the chemical gradients are large. This indicates that the interaction of steady planetary waves and photochemistry need only be modeled in the presence of large chemical gradients.

For seasonal integrations the largest changes in the mean field are caused by advection by the Lagrangian-mean velocity in the conservative region. The advection is so weak that in the chemical regions its effects are negligible. Using the residual circulation instead of the Lagrangian-mean circulation overestimates the advection by the planetary waves. In this particular model the largest differences occur between the Lagrangian-mean and residual circulations in chemically active regions. Hence, the difference between the ozone field changes calculated with Lagrangian-mean and residual circulations is small. This would not necessarily be true for constituents with different chemistry.

The transport mechanisms investigated here are probably not adequate to model planetary wave effects in the stratosphere. Sudden warmings, which are highly transient, will certainly have a great effect upon the ozone distribution (Garcia and Hartmann,

1980). Therefore, steady state planetary wave transport models are probably only relevant in the late fall and early winter. The planetary wave transport would have to be supplemented by transport due to the diabatic circulation and small scale dynamics in the lower stratosphere. Transport by planetary waves during sudden warmings, transport by the diabatic circulation, and the relative importance of the two circulations will be presented in the next chapter.

Chapter 4

TRANSPORT BY TRANSIENT PLANETARY WAVES AND THE DIABATIC CIRCULATION

4.1 Introduction

In this chapter the transport of ozone by time dependent planetary waves on a β -plane is discussed. In Section 4.2 the results of the dynamic model for both 600 m and 900 m topographic forcing are presented. In Sections 4.3, 4.4, and 4.5 the transport of ozone by the diabatic circulation alone, by planetary waves alone, and by the combination of the diabatic and planetary wave circulations is discussed. The results are analyzed in Section 4.6.

4.2 Dynamics

Equations 2.3 and 2.4 are solved on a β -plane centered at 60° N for zonal wavenumber one perturbations. The wave is forced by zonal flow over a wavenumber one topography with a height of either 600 m or 900 m. With forcing of this magnitude there is at least one sudden warming accompanied by a reversal in the zonal mean wind field. For long integrations a steady state solution is reached with permanent easterlies at some level. Schoeberl (1982b) presents a detailed analysis of the model dynamics.

During winter it is common to see distinct pulses of wave

amplitude and either major or minor warmings (Labitzke and Coratzki, 1982). After a major or minor warming the stratosphere usually returns to its winter climatological state. If the warming occurs late in the winter, then the zonal mean circulation after the warming evolves into the summertime regime where easterlies are present throughout the entire stratosphere. This type of warming is called a final warming. The dynamical model used here does not attempt to represent all of the features of the winter stratosphere only to simulate warming events and their corresponding circulation patterns.

The details of the numerical scheme were given in Section 2.3b. Results will be presented as time-height cross sections at the center of the channel and at a northern and southern point, corresponding to 60° N, 46° N, and 74° N on the β -plane.

a. 600 m topography

Time series of 60 days for $\bar{u}$, the zonal mean wind, and the magnitude of ϕ , the planetary wave geopotential, are shown at the channel center in Figures 4.1 and 4.2 for the 600 m orography. The wave amplitude increases at all altitudes until day 22 when it reaches a maximum of about 1.6 gpk. After day 22 the wave amplitude decreases rapidly until day 30 and continues to decrease more slowly until day 47. On day 47 the wave amplitude starts to increase at high altitudes and is still increasing at the end of the time series.

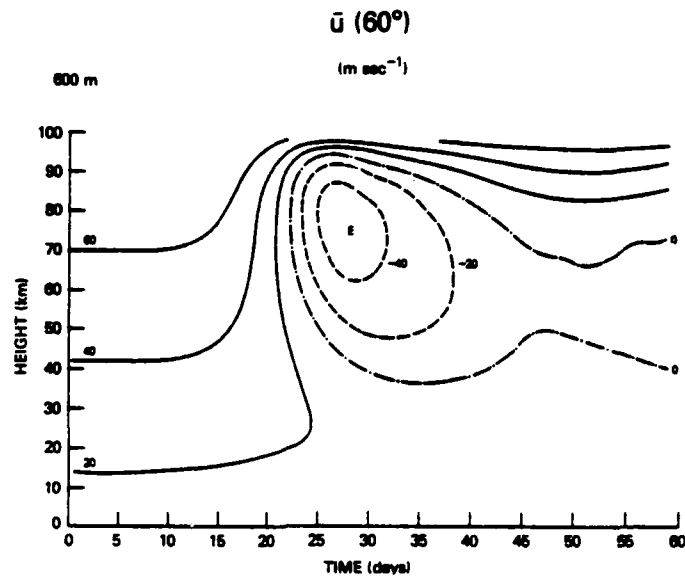


Figure 4.1. Zonal mean wind for 600 m topography
at the channel center (m sec⁻¹).

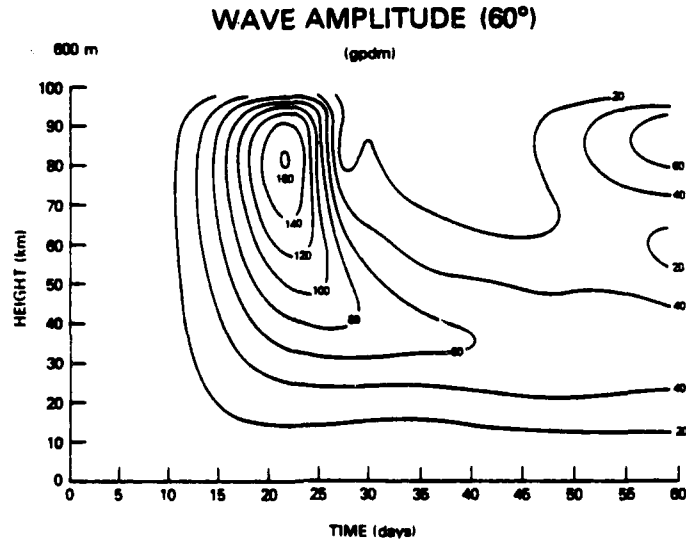


Figure 4.2. Geopotential amplitude for 600 m topography
at the channel center (gpdm).

Coincident with the maximum in the wave amplitude field is the appearance of easterlies in the zonal mean wind field. The presence of easterlies prohibits further vertical propagation of the stationary wave (Charney and Drazin, 1961) and, hence, the rapid decrease of the wave amplitude. The easterly zonal wind reaches a maximum velocity of 53 m sec^{-1} on day 29 and then decreases in intensity until day 45 when westerlies begin to appear again at the edges of the channel. The reappearance of westerlies permits vertical planetary wave propagation subsequently leading to the second increase in wave amplitude at day 47.

Since it has been predicted that transport in the vicinity of critical lines might be large (Schoeberl, 1981b; Matsuno and Nakamura, 1979), the descent of the zero wind line (critical line for stationary waves) is of particular importance to the study of tracer transport. For the 600 m topography the critical line in the center of the channel has descended to 48 km on day 25 and reaches a minimum altitude of 36 km on day 35. The critical line remains in the neighborhood of 36 km and then starts to rise on day 40. This rise is related to the reappearance of westerlies at high and low latitudes. The critical line ascends to 50 km on day 45 and then starts to descend again as the wave amplitude increases.

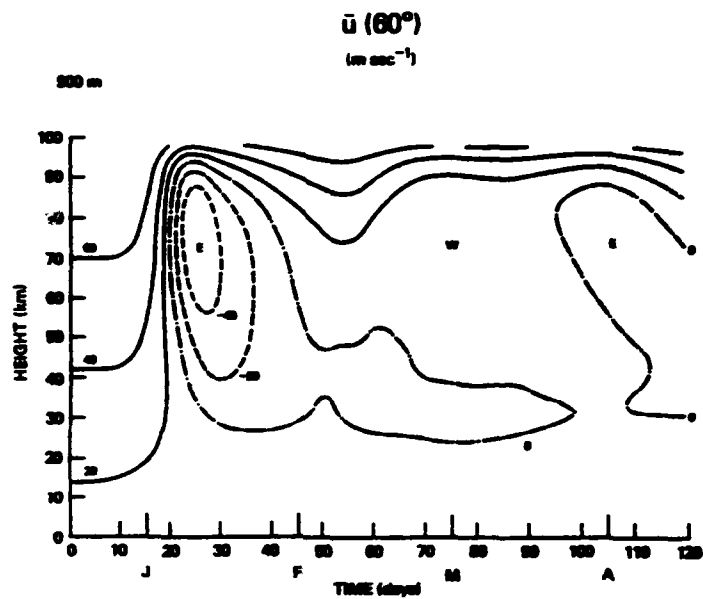


Figure 4.3. Zonal mean wind for 900 m topography
at the channel center (m sec⁻¹).

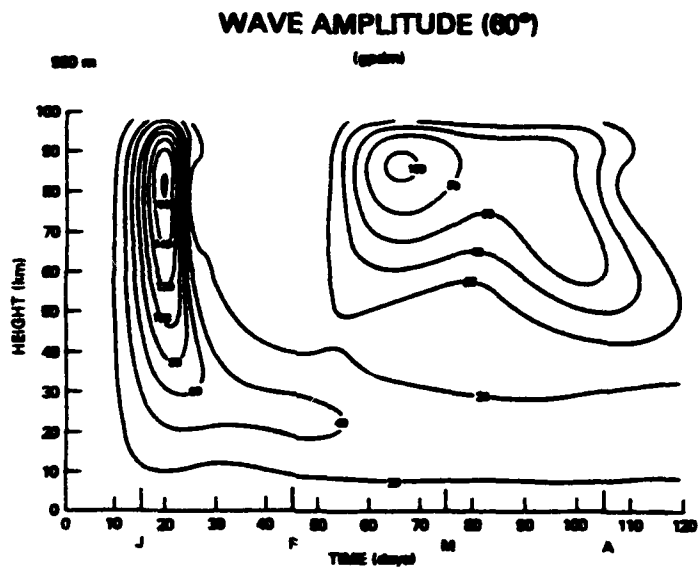


Figure 4.4. Geopotential amplitude for 900 m topography
at the channel center (gpdm).

b. 900 m topography

The mean zonal wind field and the wave amplitude field associated with the 900 m topography are shown in Figures 4.3 and 4.4 for a 120 day integration. These fields are similar in many ways to the 600 m topography fields. The wave amplitude increases rapidly until about day 20 where it reaches a maximum of 1.6 gpkm. The fact that the forcing has been increased by 50% but the wave amplitude reaches the same maximum amplitude is consistent with the wave saturation theory of Schoeberl (1982a). The main effect of the increased forcing is to cause the wave maximum to be reached somewhat earlier. The wave amplitude rapidly decreases between days 20 and 30 after which it slowly decreases until day 50. There is another rapid increase in wave amplitude after day 50 which maximizes on day 65. The wave amplitude decrease following this maximum is not as rapid as that following the initial warming event.

Coincident with the first wave maximum is the rapid deceleration of the mean zonal wind and the formation of easterlies which reach a maximum velocity on day 26 of 57 m sec^{-1} . The second wave maximum on day 65 is not instantaneously associated with the onset of high level easterlies. The wave forcing of the mean flow from days 55-65 decelerates the mean zonal wind restoring the low level easterlies which then prohibit propagation of the wave into the stratosphere. The low level easterlies ultimately disappear about day 100 which allows the wave to once again propagate vertically. The zonal mean westerlies are so weak at this time that even the relatively weak forcing associated with the wave pulse at

day 100 is capable of producing easterlies throughout a deep portion of the atmosphere.

The zero wind line reaches minimum altitudes of 26 km on day 37 and 24 km on day 76. Between days 45 and 55 the critical line rises and the wave propagates and rebuilds the low level easterlies. The low level critical line is absent for a brief period beginning at day 100.

c. The diabatic circulation

As mentioned in the introduction, early theories of stratospheric circulation postulated a direct cell with rising motion in the equatorial region and sinking motion in the polar region. This circulation was thought to describe the dryness of the stratosphere and the buildup of ozone at high latitudes (Brewer, 1949; Dobson, 1956). Observations during the 1960's showed, however, that the northern hemisphere stratospheric mean circulation consisted of two cells with rising motion over the pole and the equator. Dunkerton (1978) pointed out that in the absence of waves the diabatically driven zonal mean field is a good approximation to the Lagrangian-mean velocity field and is consistent with the Brewer-Dobson circulation. Hence, the large scale mass transport associated with radiative forcing is described by a one cell circulation as originally postulated.

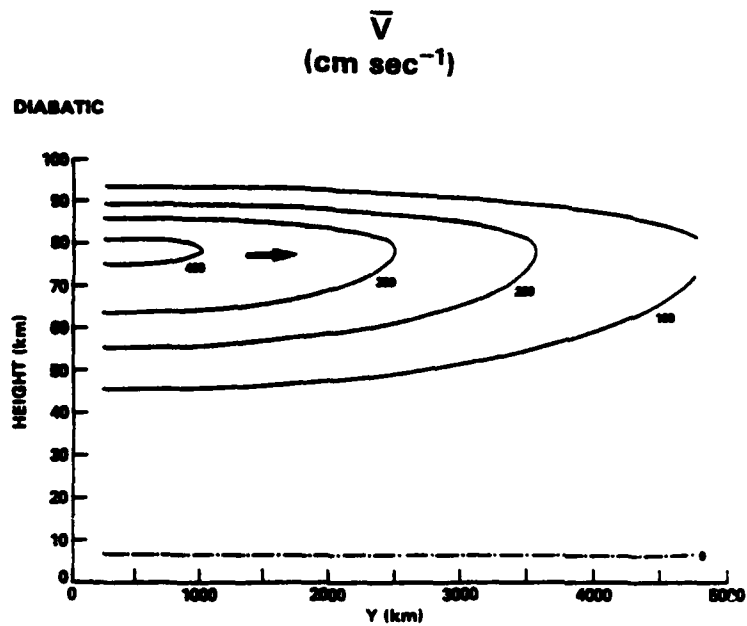
The diabatic circulation used here is derived from Dunkerton (1978) and the more detailed calculations of Murgatroyd and Singleton (1961). The circulation appropriate for the winter

solstice is shown in Fig. 4.5. The time variation is assumed to be harmonic such that the intensity of the flow reaches a maximum shortly after the winter solstice and then decreases and reaches a minimum in the spring.

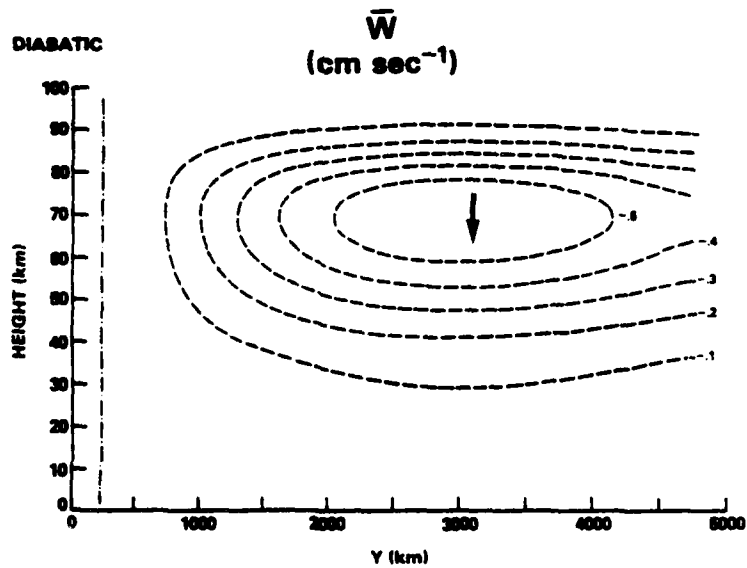
d. Initialization

The zonal mean ozone field used to initialize the transport model is shown in Fig. 4.6. This field is derived from the data of Dütsch (1969) and is representative of mid-November. Data north of 70° and above 45 km are scarce and of low quality. In these regions the prescribed ozone field is based on the vertical profile derived by Kreuger and Minzner (1976) and the ozone field used by Fels et al. (1980). The eddy ozone field is initially assumed to be zero.

A review of stratospheric warming data indicates that there is either a major or minor warming during the first or second week of January. Since the model warmings occur about three weeks after the planetary waves are turned on, December 15 has been chosen as the initialization date. In order to properly correlate the ozone field with the planetary wave dynamics the initial ozone field is advected with the diabatic circulation until Dec 15.



a. meridional component



b. vertical component

Figure 4.5. Diabatic circulation
at the winter solstice (cm sec⁻¹).

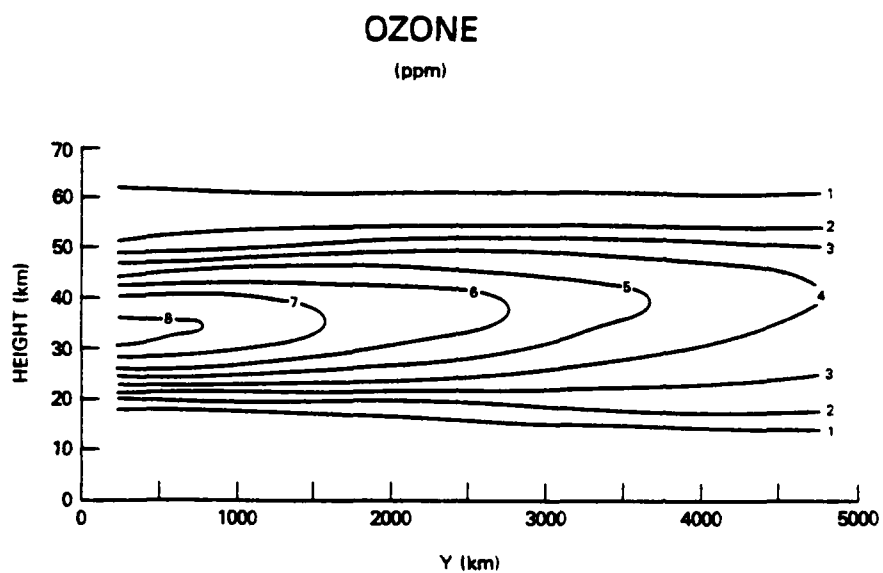


Figure 4.6. Initial ozone profile (ppm).

4.3 Transport by the diabatic circulation

The change in the mean ozone field (anomaly) due solely to the diabatic circulation is presented in one year time series in Figures 4.7 and 4.8. In general there is a positive anomaly in the lower stratosphere (20-30 km) for much of the year and a negative anomaly in the autumn. Since the ozone is in near equilibrium above 35 km, a positive (negative) anomaly is associated with an increase (decrease) in total ozone. Using this correlation it is seen that the total ozone reaches a maximum in early April and a minimum in October as opposed to an observed March maximum and October minimum.

There is a strong latitudinal dependence of the anomaly. The wintertime increase is larger and reaches lower altitudes at high latitudes. In the northern region (Fig. 4.8) a maximum increase of 3.5 ppm is centered around 26 km, and increases greater than 1 ppm are seen as low as 15 km. At the center of the channel (Fig. 4.7) an anomaly of 3 ppm is centered at 30 km, while in the southern region (not shown) a maximum increase of .5 ppm is centered around 32 km. The largest increase in the entire domain, 4.25 ppm, occurs at 70° .

The computed total ozone field (Fig. 4.9) differs considerably from observations. The time series for Dec.15-Apr.15 shows a maximum of 538 DU, an overestimate of 20%, which is reached about the first of April at 70° . Observations of ozone in the stratosphere (Fig. 2.11) show a maximum near the pole about March 1 of approximately 450 DU.

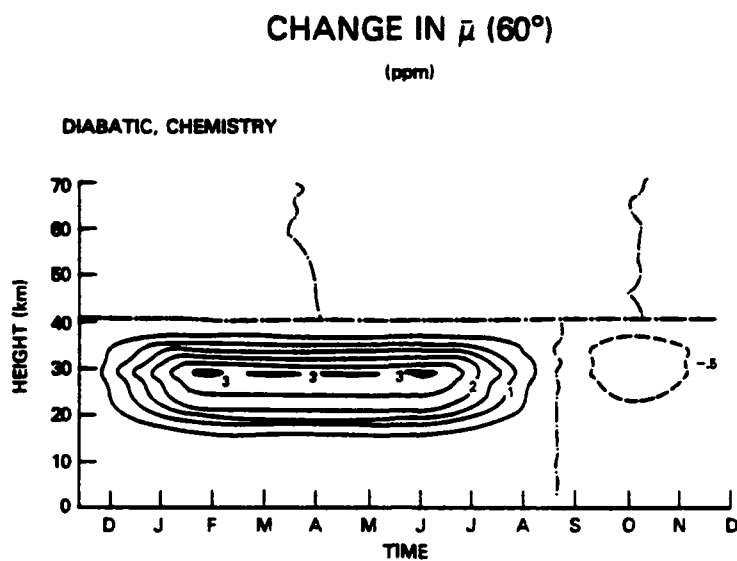


Figure 4.7. Change in ozone at the center of the channel (60°)
for transport by the diabatic circulation alone (ppm).

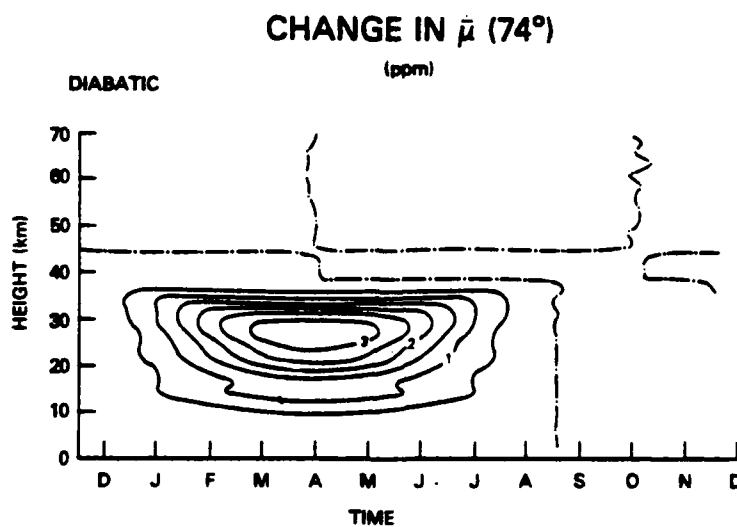


Figure 4.8. Same as Fig. 4.7 but at 74°.

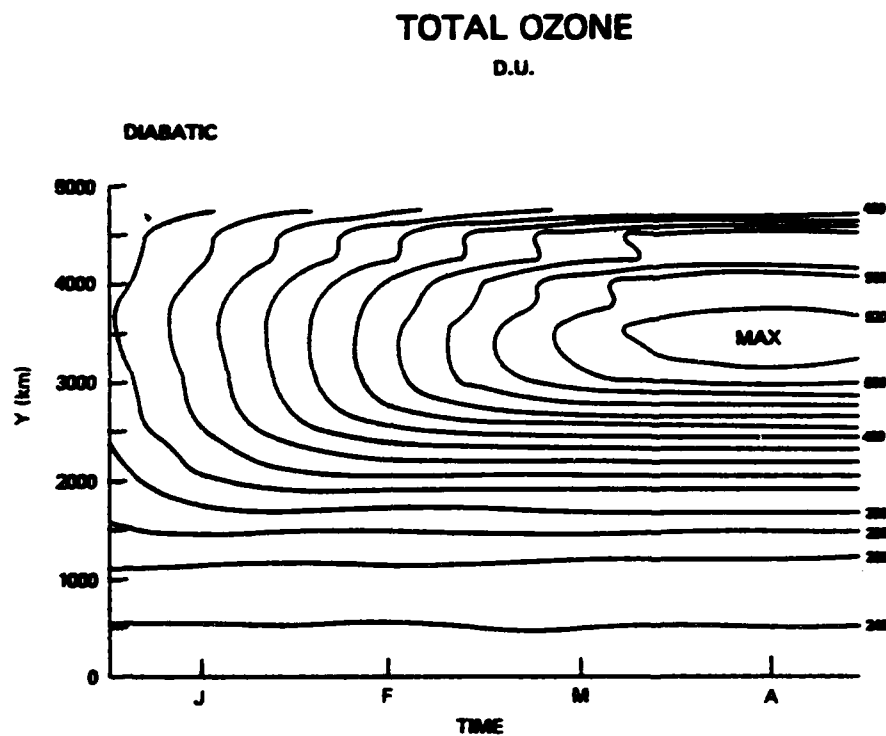


Figure 4.9. Latitude-time cross section of total ozone
as calculated with the diabatic circulation (DU).

The changes in ozone above 45 km in Figures 4.7 and 4.8 are negligible due to the strong photochemistry. Therefore, where the diabatic circulation is strongest no changes in the ozone field are found. The transport effects by the diabatic circulation are most notable in the transition and conservative regions and are due primarily to the vertical velocity which is downward at all latitudes and constantly advecting ozone into the lower stratosphere. The maximum anomaly occurs about 500 km north of the maximum vertical velocity indicating the relative ineffectiveness of the meridional velocity ($< .3 \text{ m sec}^{-1}$) at transporting ozone northward.

Because insolation is weak during the winter the ozone photochemistry is much slower than modeled here at high latitudes. The overestimation of the photochemical source term may be responsible for the unrealistic buildup in the northern region. The results from two experiments where latitudinal variation of the chemistry is included are shown in Fig. 4.10. In the first experiment the photochemistry is given a cosine of latitude dependence such that the chemical relaxation weakens and disappears at the northern boundary. This is the same latitudinal dependence as used by Garcia and Hartmann (1980). Comparing Figures 4.8 and 4.10a indicate there is no significant difference in the results from the case with no latitudinal variation of the chemistry. Apparently, the diabatic circulation is so weak that even with much slower photochemistry the ozone never varies much from its equilibrium value.

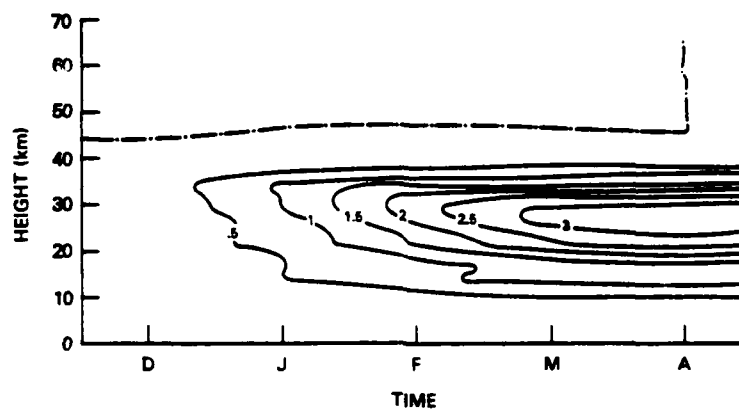
In the second experiment (Fig 4.10b) the photochemistry is set exactly equal to zero north of 67.5° for the entire time integration. The positive anomaly in the lower stratosphere remains the same as in the other experiments, but there is a negative anomaly around 50 km of more than 1 ppm. This region of depletion is due to the inability of the ozone to return to its equilibrium state. Since the increase in the lower stratosphere remains the same, some other process besides the latitudinal variation of the photochemistry must be responsible for prohibiting the excessive buildup of ozone at 70° .

In summary, the diabatic circulation during winter is directed such that transport is poleward and downward. The northward component of the velocity is too small to effectively transport ozone to polar regions; however, the vertical component acts over a considerable depth of the stratosphere and transports significant amounts of ozone into the lower stratosphere. The total ozone calculated using the diabatic circulation alone is too large in magnitude. In addition, the maximum amounts occur too far south and about 1 month too late.

CHANGE IN $\bar{\mu}$ (74°)

(ppm)

DIABATIC, cos (LAT)

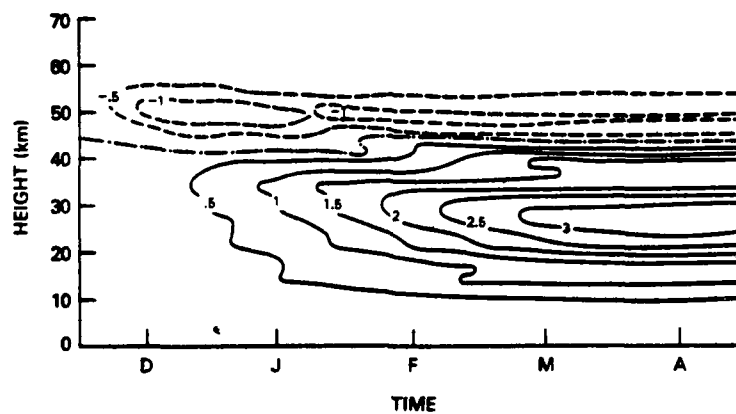


a. cosine of latitude dependence

CHANGE IN $\bar{\mu}$ (74°)

(ppm)

DIABATIC, POLAR NIGHT



b. $\lambda=0$ north of 67°

Figure 4.10. Change in ozone at 74° with horizontally varying chemistry (ppm).

4.4 Planetary wave transport

Transport by steady, stationary, dissipative planetary waves was discussed in Chapter 3. It was found that in the conservative region of the stratosphere planetary waves tended to advect ozone from low latitudes towards the pole. In the transition region it was found that stirring of parcels by the waves across regions of widely varying photochemistry could produce changes in the mean field that the mean chemistry could not completely counterbalance. In this section the results of transport by transient planetary waves will be presented. In order to study transport mechanisms it is useful to study both conservative transport (chemically inactive) and chemical transport.

a. Conservative transport

The changes in the mean ozone field for conservative transport by planetary waves forced with 600 m orography is shown in Figures 4.11 and 4.12 for the southern and northern points. Centered around 40 km is a zone of poleward ozone transport and just above this zone, centered at 50 km, is a region of equatorward transport. With the 900 m orography (Fig. 4.13), similar structure is evident with the northward transport centered at 35 km and the southward transport centered at 50 km. The maximum positive change at the northern point for the 900 m case is 2.2 ppm and occurs on day 32; for the 600 m case a maximum of slightly more than 1 ppm occurs on day 35. For both cases there is relatively little change at the center of the channel.

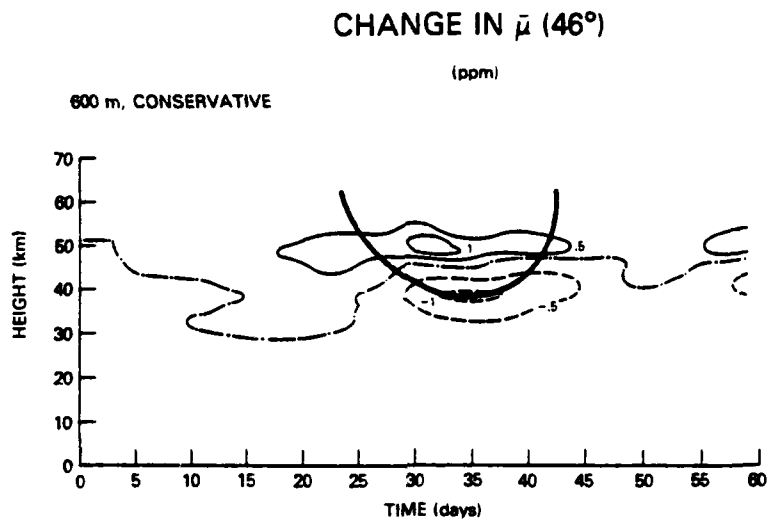


Figure 4.11. Change in ozone for 600 m topography at 46° with no chemistry (ppm).
(bold solid line is the critical line)

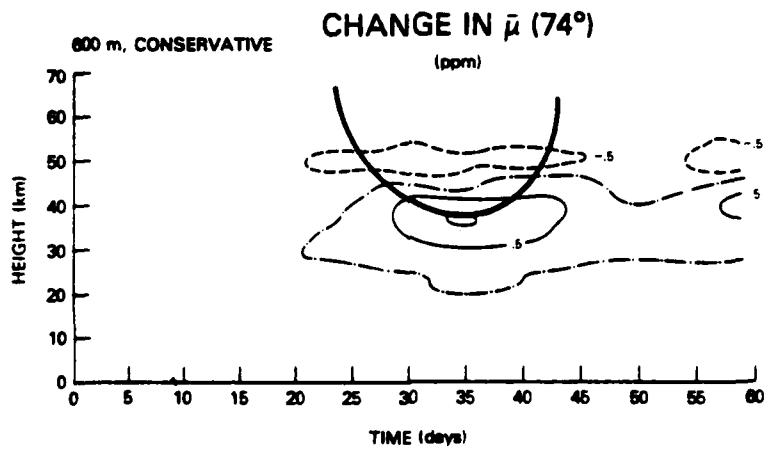


Figure 4.12. Same as Fig. 4.11 but at 74°.

CHANGE IN μ (74°)
(ppm)

900 m, CONSERVATIVE

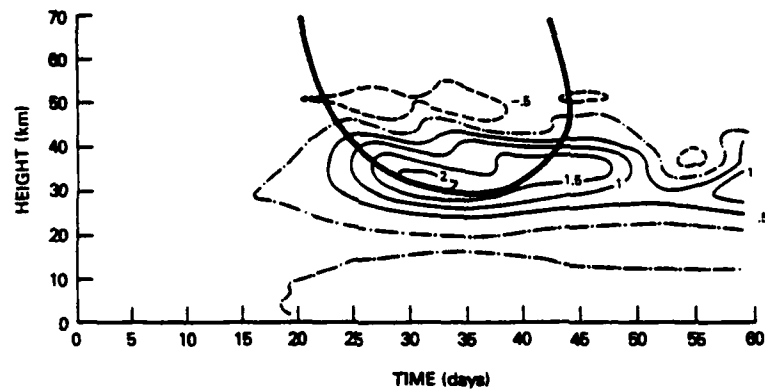


Figure 4.13. Same as Fig. 4.12 but for 900 m topography.

Superimposed on Figures 4.11-4.13 is the location of the critical line at the appropriate latitude. For all cases the level of maximum northward transport is coincident with the minimum descent of the critical line. The layer of southward transport is above the critical line. The transport calculated here supports the theory of Schoeberl (1981b) which predicts that a wave energy absorbing critical line will produce a jet of northward transport at the critical line. This transport mechanism will be discussed further in Section 4.6.

b. Chemical transport

The transport of ozone for both 600 m and 900 m orography with photochemistry is presented in Figures 4.14-4.16. There is the general tendency for the upper region of equatorward transport to be significantly reduced in all cases due to the photochemistry. With photochemistry the transport appears to occur in a thin layer in the conservative region and at the bottom of the transition region.

For the 600 m orography at the southern point (Fig 4.14) the magnitude of the depletion at 40 km has decreased by 35% with the addition of photochemistry (compare to Fig. 4.11). This might be expected since the chemistry would tend to moderate large differences from equilibrium. The magnitude of the positive anomaly at the northern grid point (Fig. 4.15) on the other hand has not moderated with the addition of chemistry but has increased by 14%.

At the northern point, with the 900 m orography (Fig 4.16), the anomaly field is once again more jet like than without chemistry,

but the magnitude has remained the same. The transport calculated at the southern point shows similar behavior with the low level negative anomaly maintaining the same magnitude with the addition of photochemistry. The level of maximum northward transport for the 900 m case occurs significantly lower in the stratosphere than with the 600 m forcing; therefore, the addition of chemistry would be expected to have relatively little effect on the transport.

It is of interest to examine the extended time series of the 900 m forcing (Fig. 4.16). In this figure the solid heavy line is the critical line at the latitude of the figure and the heavy dashed line is the critical line at the center of the channel. While the critical line at high and low latitudes ascends and dissipates, there is still intense south to north transport centered around 35 km. This transport is obviously associated with the critical line at the center of the channel. The anomaly decreases and increases coincident with the ascent and descent of the critical line.

The largest northern anomaly during the entire time integration occurs on day 70 when the critical line reaches its lowest altitude at the center of the channel. At this time there is no critical line present at any altitude at the northern point. This result shows that the horizontal extent of the transport associated with the critical line at the center of the channel is significant.

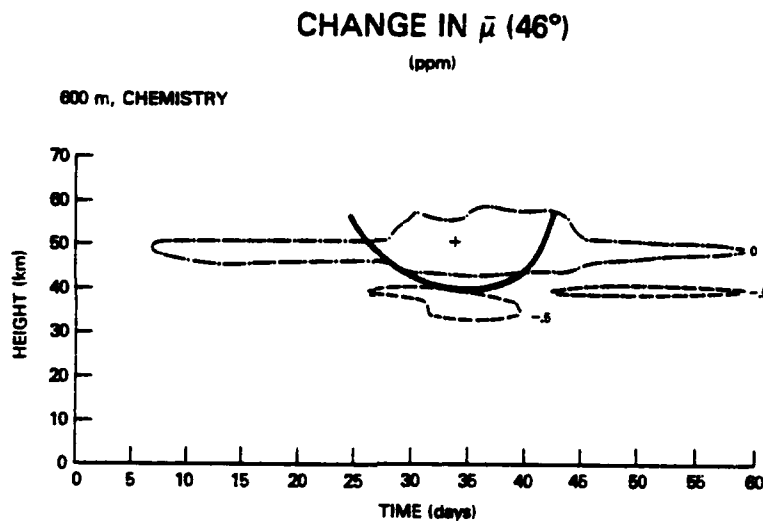


Figure 4.14. Change in ozone for 600 m topography at 46° with chemistry (ppm). (bold solid line is the critical line)

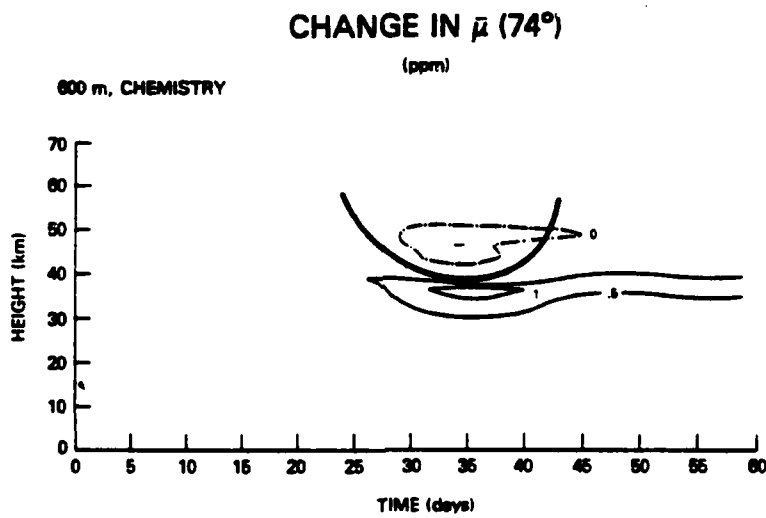


Figure 4.15. Same as Fig. 4.14 but at 74°.

CHANGE IN $\bar{\mu}$ (74°)
(ppm)

900 m, CHEMISTRY

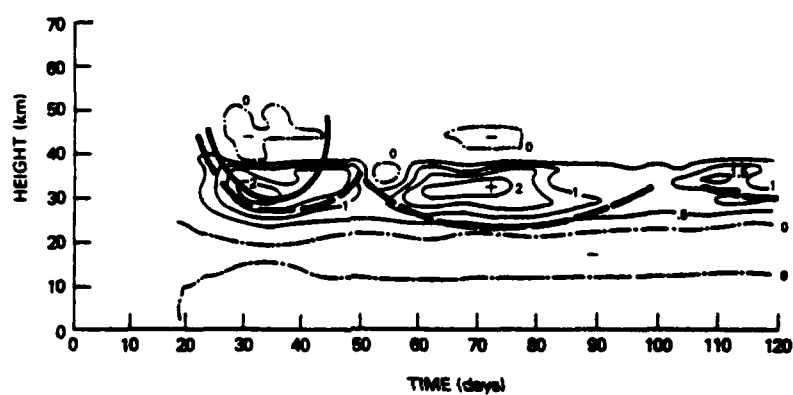


Figure 4.16. Same as Fig. 4.15 but for 900 m topography.
(bold dashed line is the critical line at the center of
the channel)

4.5 Transport by both waves and the diabatic circulation.

The change in the zonal mean ozone field associated with the diabatic circulation superimposed on the 900 m topography dynamics is shown in Figures 4.17-4.19. The 900 m topography was chosen for these experiments since this amplitude forcing is more consistent with the terrestrial topography (Berkofsky and Bertoni, 1955), and the critical line for this case descends below 32 km. One of the WMO criteria for the classification of a major stratospheric warming is that the wind reversal descends below 10 mb which is approximately 32 km. The transport by the combined circulations is much different from the transport by each individual type of circulation indicating that both the wave and diabatic transport regimes are important in the lower stratosphere.

At the center of the channel (Fig. 4.18) there is a maximum increase between 25 and 30 km of 1.25 ppm in late March. There is a series of short term decreases around 40 km with the most intense decrease of 1.25 ppm also occurring in late March. The region of steady increase is due to the downward transport by the diabatic circulation. The intermittent nature of the ozone depletion at 40 km is associated with the wave transport. With the combined circulation the increase in ozone at the channel center is reduced by more than 50% when compared to the changes caused by the diabatic circulation alone (Fig. 4.7), which indicates that the planetary waves are transporting ozone out of the center of the channel.

CHANGE IN $\bar{\mu}$ (46°) (ppm)

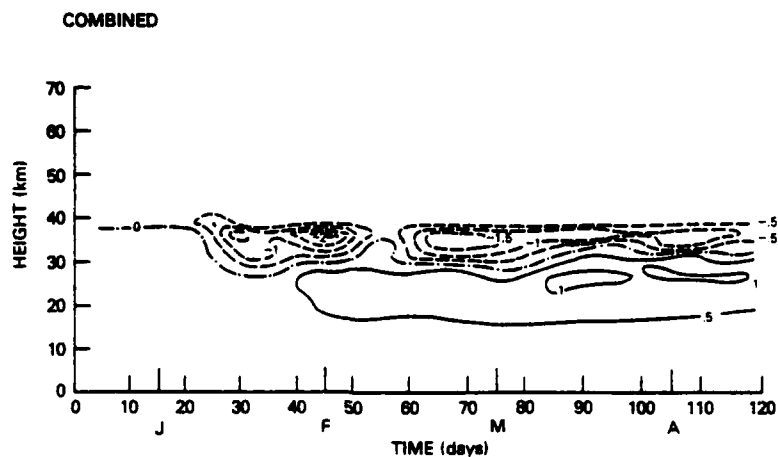


Figure 4.17. Change in ozone for the combined planetary wave (900 m) and diabatic circulations at 46° with chemistry (ppm).

CHANGE IN $\bar{\mu}$ (60°) (ppm)

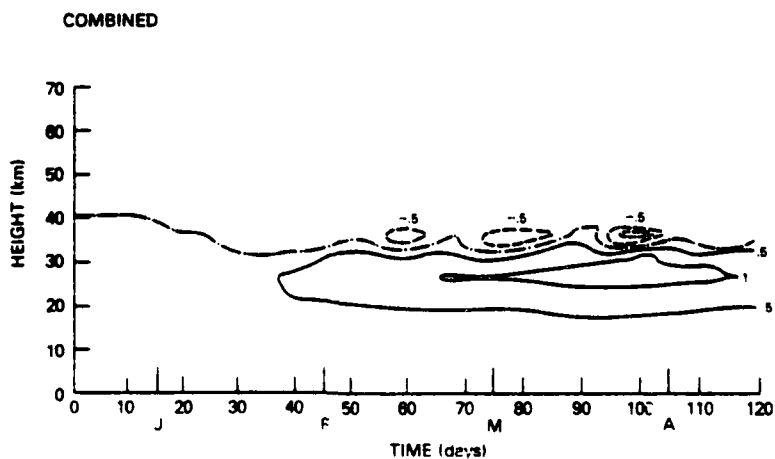


Figure 4.18. Same as Fig. 4.17 but at 60°.

CHANGE IN $\bar{\mu}$ (74°)
(ppm)

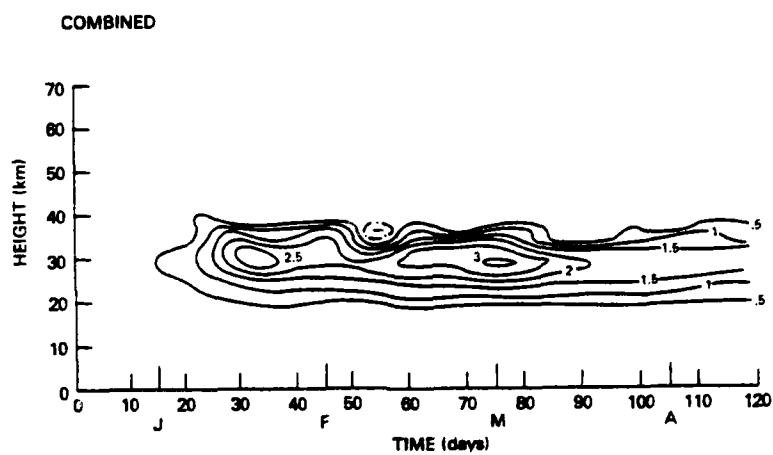


Figure 4.19. Same as Fig. 4.17 but at 74° .

At the northern latitude (Fig. 4.19) there is a uniform increase of ozone at all levels with the maximum increase of 3.12 ppm occurring close to 30 km, about 4 km lower than the maximum associated with the planetary waves alone (Fig. 4.16) and 4 km higher than that associated with the diabatic circulation alone (Fig. 4.8). An anomaly of 1 ppm is evident at 25 km, which is 10 km higher than a similar anomaly calculated using only the diabatic circulation (compare to Fig. 4.8). The largest anomaly occurs in early March and is coincident with the lowest altitude of the critical line. Following the maximum in early March the positive anomaly remains relatively stable at a maximum value somewhat greater than 2 ppm.

At the southern latitude (Fig. 4.17) there is a region of relatively strong poleward transport between 35 and 40 km and below this layer is a region of increase. As at the center of the channel, the region of increase is associated with the diabatic circulation and the decrease is associated with the planetary wave. Such large depletions are not observed at this latitude (44°) in the stratosphere.

The total ozone time series (Fig. 4.20) shows much better agreement with observation at high latitudes than the total ozone field calculated with the diabatic circulation (Fig 4.9). A late winter maximum of almost 440 DU occurs during early March at the most northern latitudes. There is an earlier maximum of slightly more than 420 DU coincident with the initial descent of the critical line. This earlier maximum produces a wave like feature that is

similar to the wave structure shown in Fig 2.10. Both maxima are associated with the critical line descent after the first warming event (Fig. 4.3).

In the southernmost section of the model there is too little ozone from the beginning of the combined planetary wave-diabatic integration (Dec. 15). Furthermore, during the initial descent of the critical line there is excessive transport out of the south. The excessive transport occurs at the time that the critical line descends to low altitudes at the northern and southern points.

TOTAL OZONE

D.U.

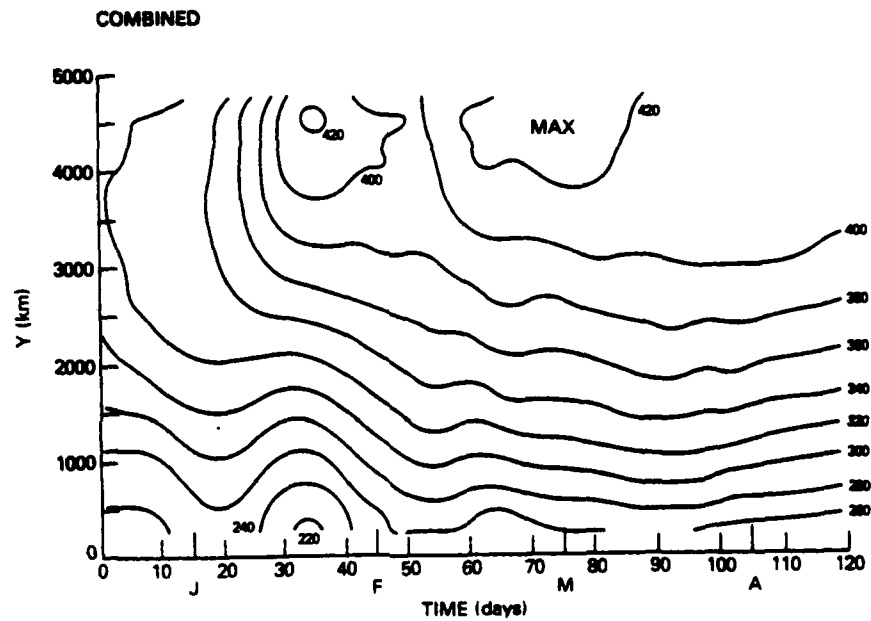


Figure 4.20. Time-latitude cross section of total ozone as calculated with the combined planetary wave-diabatic circulation (DU).

4.6 Discussion.

The transports associated with the diabatic and planetary wave circulations are fundamentally different from each other and of equal importance. The diabatic circulation is primarily responsible for transport of ozone downward from the photochemical source region while the planetary wave circulation is responsible for north-south transport. The combined circulations do a particularly good job of reproducing the features of the total ozone distribution at high latitudes.

The vertical structure of the winter ozone field is not as well simulated as the total ozone distribution. The computed increase of more than 3 ppm at high latitudes is concentrated at 30 km which is larger than observed at this altitude. The increase between 20 and 25 km is significantly smaller than is observed. The increase in ozone at high latitudes is due largely to the downward transport of ozone out of the photochemical region by the diabatic circulation at 70° and the subsequent northward transport of this ozone to high latitudes by the planetary waves.

Enough ozone is transported out of the source region to high latitudes to model the total ozone maximum; however, redistribution of ozone within the conservative region is needed to accurately reproduce the ozone distribution. Synoptic scale disturbances are probably instrumental in determining the vertical structure of ozone within the lower stratosphere.

In the southern region of the channel and to a lesser degree in the middle of the channel the vertical structure is more poorly

represented. There is excessive transport out of the south by the planetary waves between 35 and 40 km which leads to unrealistic vertical profiles. This excessive transport is also evidenced by the underestimation of total ozone at low latitudes during mid-January. The overestimation of transport out of the southern region might be due to an inadequate representation of chemistry at low altitudes; the underestimation of the transport by the diabatic circulation into the south from either high altitudes or lower latitudes; geometric effects of the β -plane transport; and/or to the neglect of small scale dynamics.

In order to understand the particular nature of the transport, it is necessary to further analyze the transport mechanisms involved. It was mentioned in Chapter 3 that even for steady, stationary, dissipative planetary waves, there were computational difficulties in calculating Lagrangian-mean quantities. In the presence of highly transient dynamics these difficulties are more severe, and the calculation of Lagrangian-mean velocities is time consuming with the result being uncertain (McIntyre, 1980). Furthermore, in the vicinity of the critical lines, the displacement vectors are nearly singular and their calculation is difficult. Dunkerton et al. (1981) also experienced difficulty in the calculation of Lagrangian-mean quantities with realistic dynamical models. Therefore the transport will be analyzed in the Eulerian framework with comments made on the Lagrangian-mean interpretation where it is appropriate.

a. Diabatic circulation

At all times during the winter at extratropical latitudes the flux due to the diabatic circulation is northward and downward. The northward transport is quite weak and does little to change the ozone distribution. The vertical transport is effective at transporting ozone from the photochemical and transition regions into the upper part of the conservative region.

While the diabatic circulation does transport ozone northward and downward it is inadequate alone to explain the wintertime polar ozone buildup. The meridional component of the diabatic circulation is large enough to cause significant transport only at high altitudes where the photochemistry is dominate. Therefore Brewer-Dobson circulations, which involve the upward transport of ozone at low latitudes; northward transport at high altitudes in the midlatitudes; and subsidence at high latitudes, cannot explain ozone transport.

b. Planetary waves

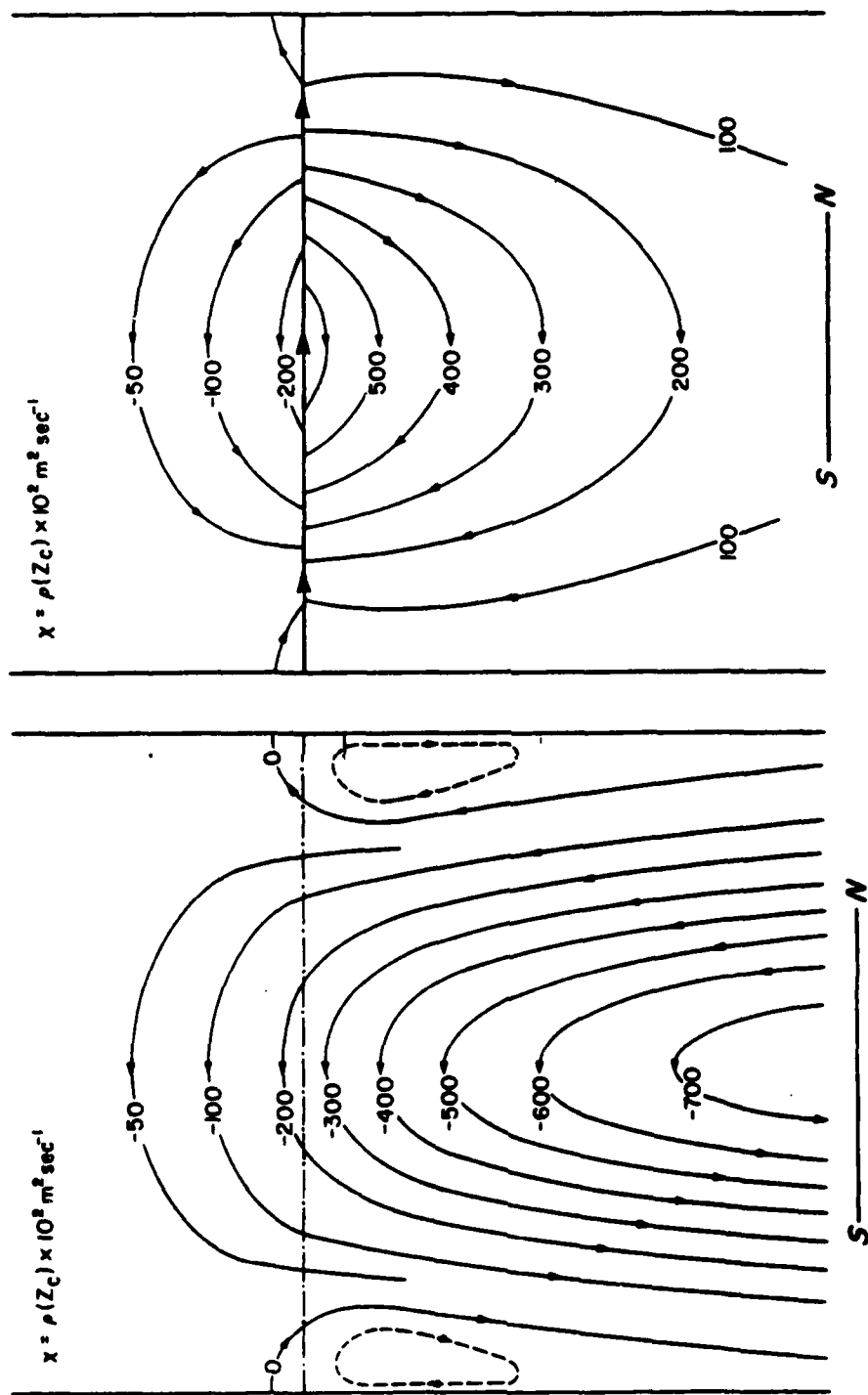
The planetary wave transport is more complex than that presented in Chapter 3 for steady-state transport. Examination of Fig. 4.16, for instance, shows that the transport is highly time dependent with strong pulses of northward transport. The planetary wave transport was demonstrated to be directly related to the descent of the critical line in the previous section.

According to theory (Charney and Drazin, 1961), it is impossible for stationary planetary waves to propagate vertically in

an easterly mean zonal flow; that is, through a critical line. Therefore, when a propagating wave is incident upon a critical line the wave energy must be absorbed, reflected, or possibly overreflected if the critical line is in a region of negative $\bar{q}_y$ (Lindzen et al., 1980). In general, the vicinity of the critical line is a region of strong wave transience and dissipation (Matsuno and Nakamura, 1979).

In this model the critical region is generally a region of partial if not full absorption of wave energy. This fact is demonstrated by the persistence of the critical line. The Rayleigh friction incorporated into the model tends to force the mean flow back to its equilibrium state. Since the critical line remains persistent for such long periods, as is evidenced in Fig 4.3, some wave energy must be absorbed at the critical level in order to maintain the critical level.

Matsuno and Nakamura (1979) and Schoeberl (1980, 1981b) have discussed the Eulerian and Lagrangian-mean circulations associated with an idealized absorbing critical line oriented at various angles with the vertical. In Fig 4.21 are shown the mass stream functions associated with the two different formulations as taken from Matsuno and Nakamura (1979) for a horizontal critical line. In the Eulerian mean formulation there is rising motion in the north with descending motion in the south and a general southward transport. This flow regime is viewed as the mean field response to the northward heat flux of the planetary waves.



b. Lagrangian-mean

a. Eulerian mean

Figure 4.21. Stream functions for a horizontal critical line

(from Matsuno and Nakamura, 1979).

The Lagrangian-mean stream function is much different with singular northward flow at the critical line and weaker return flow both above and below the critical line. Schoeberl (1981b) found that, except for vertical critical lines, there is always singular flow at a purely absorbing critical line. These authors use these idealized models to explain stratospheric warming phenomena and commented on the consequences of critical line flows on the distribution of minor constituents in the stratosphere. Schoeberl (1981b) predicted that where the critical line changes its orientation with the vertical, the constituent flux will be highly divergent and these regions will be an area of large constituent density changes.

In Schoeberl's study it was assumed that the critical line movement was so slow that it was not necessary to recalculate the zonal mean wind field. Nevertheless, it is possible to calculate the acceleration of the mean zonal wind which was found to be negative. That is, for the case of an absorbing critical line there is a tendency for the critical line to descend. Therefore the transport described above occurs when the mean zonal flow is being decelerated at the critical line and the critical line is descending.

For the case of an ascending critical line the flow is accelerating rather than decelerating at the critical line. This being the case the sign of Eq. (7) in Schoeberl (1981b) can be changed to indicate that the critical region is forcing the mean flow to become more westerly. With this sign change a strong

southward Lagrangian-mean jet is found at the critical line.

Both the models of Schoeberl (1980, 1981b) and Matsuno and Nakamura (1979) assume the interaction of the wave with the critical line can be modeled as a step function so that the Eliassen-Palm (EP) flux is discontinuous across the critical line. With such a functional dependence of the wave forcing, the EP flux divergence is proportional to a delta function and, therefore, the Lagrangian-mean flow is singular at the critical line. The delta function dependence is equivalent to putting all of the wave transience at the critical line. In the real atmosphere and in numerical models the forcing is not so abrupt; hence, the effects of the critical line are spread out over a thin layer of the atmosphere.

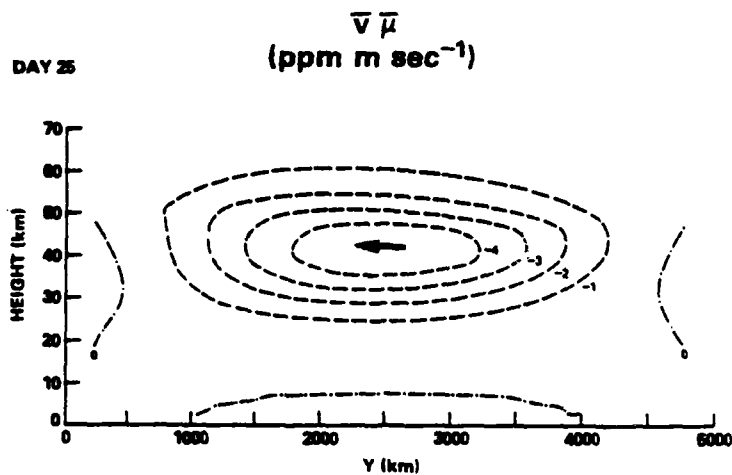
In the Lagrangian-mean framework one would expect to see large northward fluxes at the critical line with the changes in the mean field being proportional to the divergence of the flux. In the Eulerian formulation, since the mean velocity is generally southward, the strong transport at the critical line should manifest itself as strong eddy flux oriented with the critical line. The direction of the eddy flux will depend on the background constituent density.

In the following, the transport for 900 m planetary waves will be discussed for both chemical and conservative transport. It is desirable to study the transport for conservative tracers because the chemistry tends to counter any changes in the upper stratosphere and might conceal transport mechanisms which would be important or incorrect for other chemical species. The transport will be

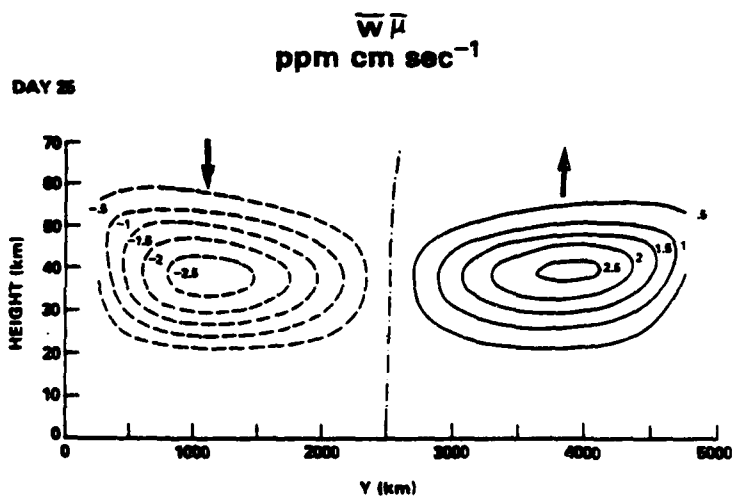
discussed with regard to the time scale in Fig. 4.3 measured from the beginning of the model integration.

The mean flux is shown at day 25 in Fig 4.22. At all times the Eulerian mean flux is from north to south with upward flux at high latitudes and downward flux at low latitudes. This flux field is consistent with that expected for a wave interacting with a critical line.

The eddy horizontal flux at day 25 is shown in Fig. 4.23 for conservative transport by the planetary waves alone. At this time the critical line has descended to approximately 40 km and is still descending. Centered at the critical line is a region of strong northward transport which has a strong meridional derivative in both the north and the south. This region of northward transport is responsible for the positive anomaly centered at 40 km seen in Fig. 4.13. The region of southward transport centered at 50 km is responsible for the negative anomaly at 50 km in Fig. 4.13. The transport shown here is consistent with the transport as predicted by Schoeberl (1981b) and Matsuno and Nakamura (1979) with strong northward transport at the critical line and return flow above.



a. horizontal



b. vertical

Figure 4.22. Eulerian mean flux at day 25 (ppm m sec⁻¹).

$$\overline{v'\mu}$$

(ppm m sec⁻¹)

DAY 25. CONSERVATIVE

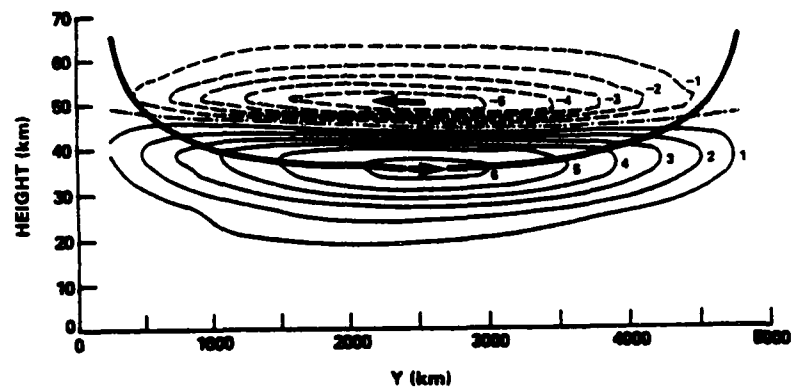
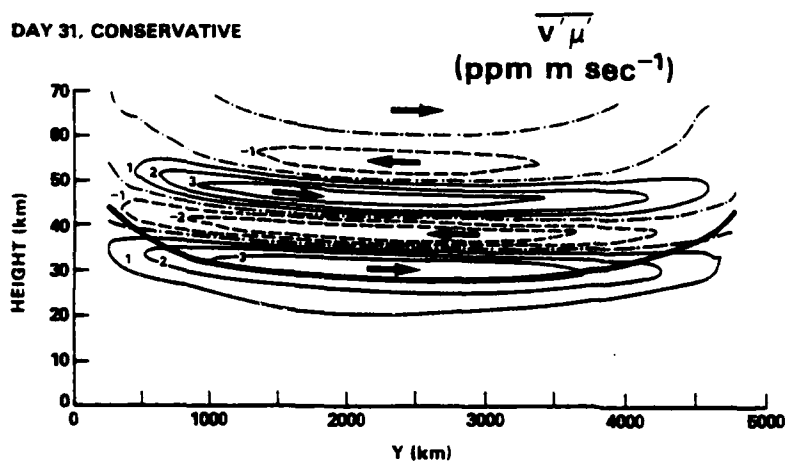


Figure 4.23. Horizontal eddy flux at day 25
conservative transport (ppm m sec⁻¹).

(bold solid line is the critical line)

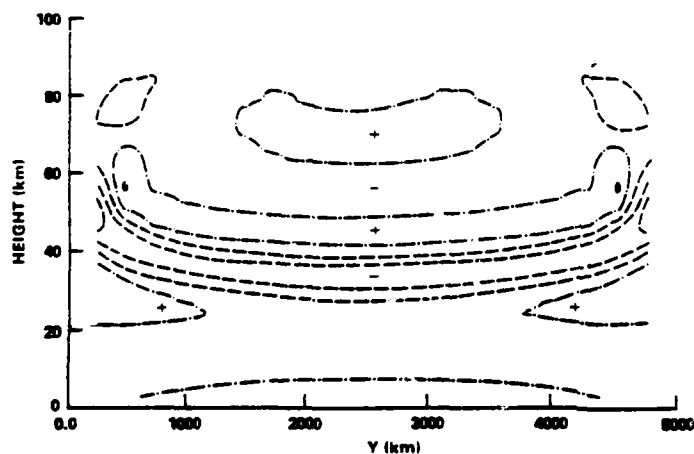
DAY 31. CONSERVATIVE



a. Same as Fig. 4.23
but at day 31.
(chain-dot = 0,
dash < 0, solid > 0)

DAY 31

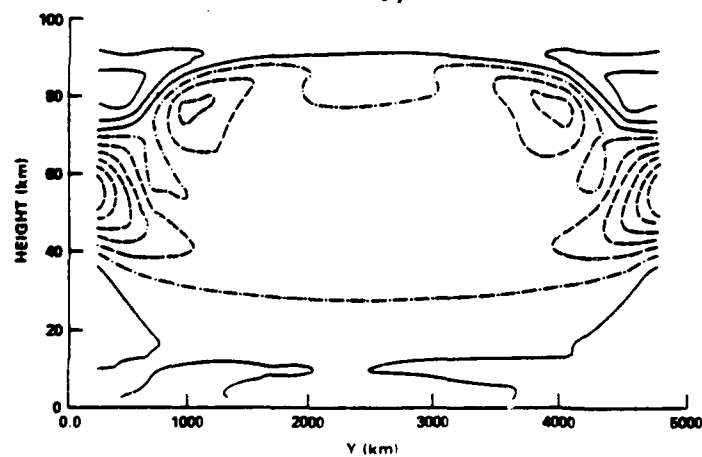
WAVE TRANSIENCE



b. Wave transience at day 31.

DAY 31

$\frac{\partial \bar{q}}{\partial y}$



c. $\bar{q}_y$ at day 31.

Figure 4.24.

By day 31 (Fig 4.24a) the eddy transport has changed considerably with the formation of several layers of northward and southward flux above the two layers that existed on day 25. The lowest level of northward transport has decreased in magnitude by half and is still centered at the critical line which is now near its maximum descent. At 50 km where there was strong southward transport on day 25 there is now predominately northward transport. This reversal of transport is responsible for the decrease in the magnitude of the anomaly at 50 km around day 30 in Fig. 4.13. It is obvious from Fig. 4.13 that the eddy flux at high altitudes has relatively little effect on the mean constituent field, indicating that the eddy and mean fluxes are nearly compensating away from the critical line.

In order to understand the mechanisms that are responsible for the formation of this layered transport structure it is useful to consider the following form of the horizontal eddy flux which can be derived from the perturbation continuity equation (Eq. 3.6) assuming that the perturbation experiences exponential growth or decay and that the chemistry is zero.

$$\begin{aligned} \overline{v'u'} = & \frac{-c_1}{c_1^2 + (k\bar{u})^2} \left\{ \overset{A}{\overline{Re(v')}}^2 \frac{\partial \bar{u}}{\partial y} + \overset{B}{\overline{Re(v')Re(w')}} \frac{\partial \bar{u}}{\partial z} \right\} \\ & - \frac{\bar{k}u}{c_1^2 + (k\bar{u})^2} \left\{ \overline{Re(v')Im(w')} \frac{\partial \bar{u}}{\partial z} \right\} \end{aligned} \quad (4.1)$$

C

This equation is useful for diagnostics away from the critical line.

At the critical line many of the terms in the denominator are nearly zero so that (4.1) would be singular. c_1 is positive for a growing wave.

The transience at day 31 is shown in Fig. 4.24b. Positive transience is associated with a growing wave and negative transience is associated with a decaying wave. There is a strong correlation of the sign of the transience with the correlation in term C in Eq. 4.1. When the wave is growing (decaying) the correlation of $\overline{\text{Re}(v')\text{Im}(w')}$ is negative (positive); the vertical gradient of ozone is negative above 40 km, and the sign of $\bar{k}_u$ in the region of the layering is negative. The horizontal eddy transport in the region where it is layered is dominated by term C. The horizontal eddy flux in this region is proportional to the vertical constituent derivative. This dependence is proven by an experiment where the horizontal gradient of ozone was reversed and the vertical gradient held constant. In this experiment the layering structure remained the same while the eddy flux at the critical line reversed direction.

The meridional gradient of the zonal mean potential vorticity, $\bar{q}_y$, is shown in Fig 4.24c. A necessary condition for baroclinic instability is that this gradient changes sign (Charney and Stern, 19 2). The potential vorticity gradient is negative in a large area above the critical line, but there are three regions where it changes sign within the zone of easterlies. The positive transience in Fig. 4.24b is localized near these gradient changes. Thus, the layered structure appears to be due to the

transport by an unstable mode in the zone of easterlies.

The positive anomaly in the ozone field is persistent until about day 47 when the anomaly begins to decrease and nearly disappears (Figures 4.13 or 4.16). This feature is coincident with the rise in the critical line which occurs around day 50. The horizontal eddy flux shown in Fig. 4.25 indicates that there is large southward flux associated with the rising critical line. As the critical line descends again after day 50 strong northward flux is once again associated with the critical line.

In actuality the critical line rises slowly for two weeks before rising rapidly on day 47. During the time the critical line is nearly stationary and rising slowly, there is a small flux in the southward direction. This explains why there is not enormous buildup in the northern region while the critical line is nearly steady.

The forcing of the mean flow in the vicinity of the critical line is shown in Fig. 4.26 where the EP flux divergence and critical line at the center of the channel is shown for a dynamics run similar to the one used in the transport study. Different dynamics were used here merely to conserve computer resources and all of the essential features of the transport dynamics are represented. While the critical line is rising there is convergence in the critical region which tends to accelerate the mean zonal flow in a westerly fashion. When the critical line is at its peak on day 42 the EP flux divergence is small and as the critical line descends the EP flux divergence becomes positive indicating absorption of wave

energy and a deceleration of the mean flow.

The effects of photochemistry are clearly exhibited in Fig. 4.27 which shows the horizontal eddy flux at day 31, the same time as Fig 4.24a. The multilayer structure which was present in the conservative experiment is not present due to the strong photochemistry above 40 km. This indicates that caution must be exercised if the current dynamics were to be applied to other gases with different chemistries. If, for instance, the transition layer were located in the region where the unstable mode was present large unrepresentative mean field changes might be calculated.

The strong poleward transport at the critical line is enhanced by the photochemistry because the weaker equatorial return transport is largely countered by the photochemistry. The photochemistry limits the transport largely to the conservative region and the lower part of the transition region for the 900 m topography. The importance of the critical line descending below the transition region is demonstrated by the transport with the 600 m topography which is only a third of that associated with the 900 m topography. The mechanism of transport when photochemistry is included is consistent with the critical line transport discussed above. The high latitude maximum in Fig. 4.20 on March 1 is associated with jet like transport of ozone as the critical line descends on day 70.

$$\overline{v \mu}$$

(ppm m sec⁻¹)

DAY 47. CONSERVATIVE

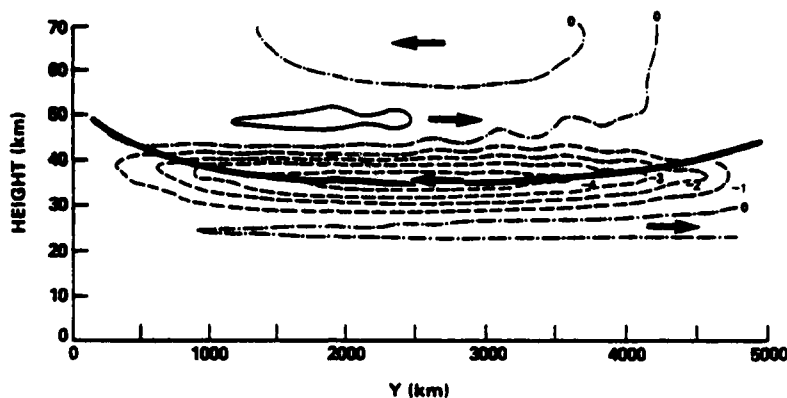


Figure 4.25. Eddy horizontal flux at day 47 for conservative transport during the rise of the critical line (ppm m sec⁻¹).

(bold solid line is the critical line)

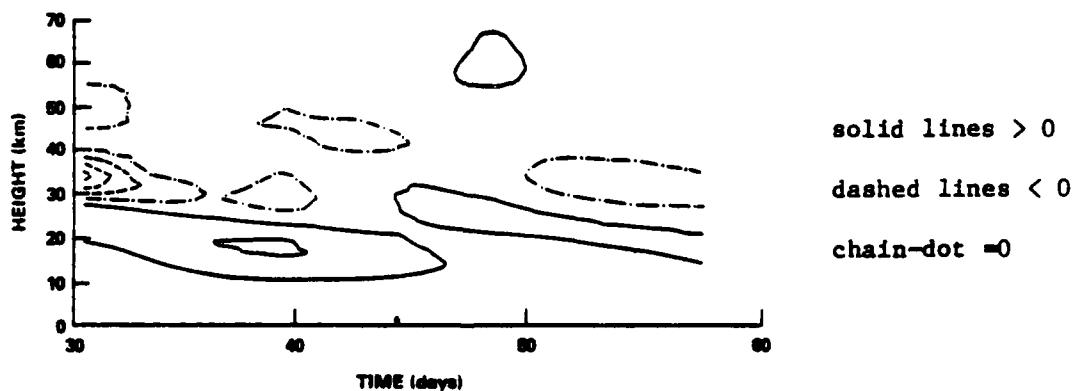


Figure 4.26. Eliassen-Palm (EP) flux divergence during the rise of the critical line. (negative values correspond to westerly acceleration)

$$\overline{v'\mu}$$

(ppm m sec⁻¹)

DAY 31, CHEMISTRY

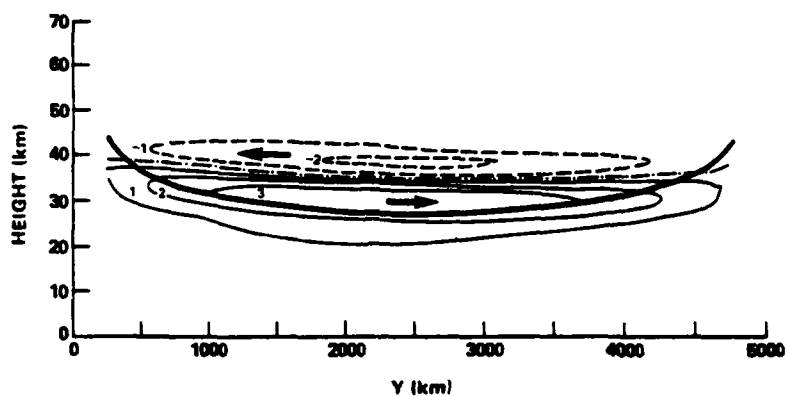


Figure 4.27. Horizontal eddy flux on day 31
for chemical transport (ppm m sec⁻¹).
(bold solid line is the critical line)

4.7 Summary

The effect of the planetary waves during a sudden warming is to transport ozone poleward. This transport is explained by the mechanisms revealed by Matsuno and Nakamura (1979) and Schoeberl (1981b). The transport by such strongly transient waves is much more vigorous and confined to a smaller vertical area than transport by steady state waves.

Northward transport coincides with a descending critical line. Southward transport occurs when the critical line ascends. The magnitude of the transport varies with the rate of ascent or descent, an indicator of the wave absorption properties of the critical line. When the critical line is stationary the transport is small.

Photochemistry enhances the buildup of ozone near the pole by reducing the return flux associated with the Lagrangian-mean jet at the critical line. The magnitude of the poleward transport is highly dependent on the critical line descending from the photochemical region into the transition and conservative regions.

The diabatic circulation transports ozone northward and downward but the northward transport is relatively ineffective. The downward transport produces increases in ozone in the lower stratosphere and replenishes the ozone in the southern region that is transported northward.

The β -plane model predicts a total ozone distribution at high latitudes consistent with observations but is less accurate at reproducing the vertical structure of ozone. Much of the discrepancy

can be attributed to the lack of small scale dynamics in the model lower stratosphere. In the next chapter the effects of the simplifying assumptions on the calculated transport as well as the applicability of the steady state and time dependent model to the real atmosphere will be discussed.

Chapter 5

SUMMARY AND A POSSIBLE OZONE TRANSPORT SCENARIO

5.1 Introduction

In the preceding chapters ozone transport by the diabatic circulation and both steady state and time dependent planetary waves on a β -plane has been described. The results simulate many of the large scale features of the wintertime ozone distribution with surprising accuracy. In this chapter the effects of some of the simplifying assumptions incorporated into the model will be evaluated, and the results will be interpreted with regard to their applicability to the atmosphere. It should be remembered that the purpose of the study is primarily to investigate transport mechanisms, not to accurately model all aspects of the ozone distribution.

Many simplifications were made both tacitly and explicitly in order to facilitate both the execution and the interpretation of the models. These specifically include the omission of small scale dynamics in the lower stratosphere; the lack of temporal and realistic latitudinal variation in the photochemistry; the specification, rather than the calculation, of the diabatic circulation; and the many assumptions incorporated in applying

quasigeostrophic planetary waves on a β -plane to a spherical atmosphere. The effects of these assumptions on the final results are varied and difficult to evaluate quantitatively but in some instances their qualitative influences can be described.

5.2 Low latitude ozone

One of the most serious inadequacies of the model is the overestimation of planetary wave transport out of the southern region of the model during sudden warmings. This excessive transport is exacerbated by the underestimation of the total ozone south of 50° on Dec. 15, the time when the planetary waves are turned on.

a. The diabatic circulation and the initial ozone distribution

The problem of too little ozone on Dec. 15 might be attributed to the initial ozone distribution being in error either with regard to total amount or the specification of mean gradients (see Section 4.2d). If the initial distribution is assumed to be correct, then the underestimation of ozone on Dec. 15 might be due to the improper designation of either the horizontal or the vertical diabatic velocity in the southern region. If the vertical velocity is too small, the transport from high altitude source regions to low altitudes would be underpredicted.

At the southern boundary there is a northward flux of ozone associated with the meridional component of the diabatic velocity during fall and winter. If this component is too small, then the

flux across the southern boundary may be underestimated. Furthermore, the ozone has been assumed to be constant at the southern boundary point (37.5°); when in fact, there is a 10% increase of total ozone mixing ratio during the winter (see Fig. 2.8). The specification of a constant mixing ratio at the southern boundary is, therefore, not strictly correct.

The inclusion of planetary waves in forming the Dec. 15 ozone field would tend to increase the total ozone error at low latitudes since both steady state and transient waves cause a depletion of ozone in the southern part of the β -channel. Even with an improved estimate of the ozone field, the transport of ozone out of the south during the initial warming is too strong. There is a large poleward transport out of the southern region associated with the initial warming which causes a more than 20 DU decrease in total ozone. Any decrease of ozone at 40° N during winter is unrealistic.

b. Geometric effects

The geometric effects of the β -plane would also contribute to the overestimation of transport out of the south. Figure 5.1 shows a spherical atmosphere divided into cells and the corresponding cells on a β -plane. The high latitude cells on the sphere contain less volume than the low latitude cells; in fact, a latitude band 2.15° wide (the approximate horizontal grid length) at 82.5° contains less than a third the volume of a similar cell centered at 60° . Likewise, a cell at 60° contains only 64% of the volume of a cell at 37.5° . Since the β -plane geometry assumes a constant cell

volume, the constituent divergence at low latitudes is overestimated while the constituent convergence at high latitudes is underestimated. In other words, the ozone transported north should be distributed over a smaller volume than it is, and the ozone removed from the south should be drawn from a larger volume than it is.

In order to understand the nature of this error consider the meridional term of the mean constituent divergence

$$\frac{1}{a \cos \theta} \frac{\partial}{\partial \theta} (\bar{v} \mu \cos \theta) = \frac{1}{a} \frac{\partial}{\partial \theta} \bar{v} \mu - \frac{\bar{v} \mu}{a} \tan \theta \quad (5.1)$$

In the β -plane approximation $\cos \theta$ is replaced by $\cos \theta_0$, where θ_0 is the latitude of the center of the channel, and (5.1) becomes

$$\frac{1}{a \cos \theta} \frac{\partial}{\partial \theta} (\bar{v} \mu \cos \theta) = \frac{\partial}{\partial y} \bar{v} \mu \quad (y=a\theta) \quad (5.2)$$

So it is seen that the β -plane estimate is exact only at the center of the channel. The error due to the geometric effect is related to the second term on the right hand side of (5.1), and is proportional to the tangent of the latitude and the magnitude of the flux. At very high latitudes the tangent term would become very large. In this model both the mean and eddy horizontal fluxes approach zero at the northern boundary (82.5°); so that, the errors remain small. At low latitudes, the tangent term is smaller than that at the channel center; therefore, the divergence would be overestimated on the β -plane. The fluxes associated with the planetary wave dynamics

are small at the southern boundary because the wave amplitude is assumed to be zero. However, the flux associated with the diabatic circulation at the southern boundary is relatively large and could contribute to the underestimation of the ozone in the south.

c. Photochemistry

The photochemistry is not truly conservative in the lower stratosphere, and sufficiently rapid photochemistry might exist at southern latitudes to affect the transport. Figure 5.2 from Gille et al. (1980) shows an observational estimate of the transition region based on the phase between the geopotential and the ozone wave (Hartmann and Garcia, 1979). In the southern parts of this figure the transition region descends to nearly 20 km. Inclusion of chemistry between 28 and 30 km would help to alleviate the severe reduction calculated in the southern part of the model by generating new ozone to replace that transported northward.

AD-A126 341

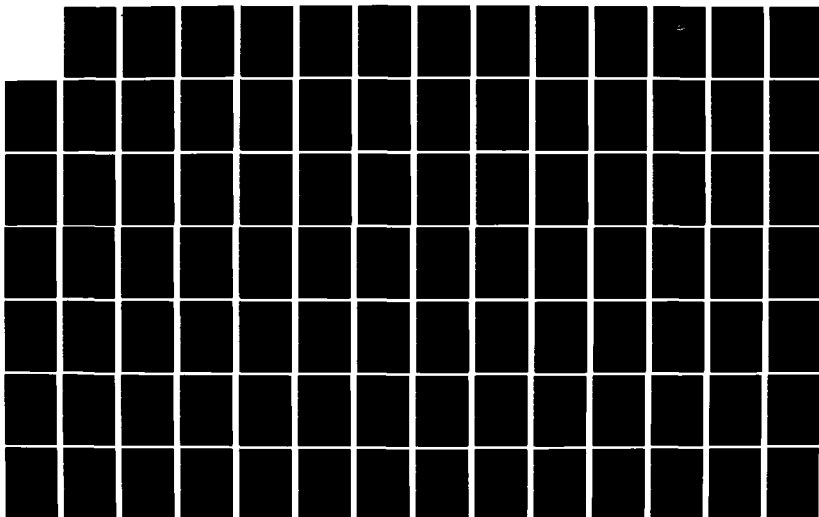
GEOPHYSICAL PLASMAS AND ATMOSPHERIC MODELING(U) SCIENCE
APPLICATIONS INC MCLEAN VA E HYMAN ET AL. FEB 83
SAI-83-144-WA N00014-81-C-2038

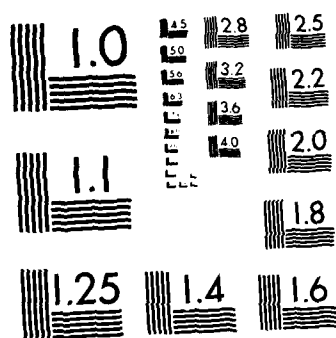
6/7

UNCLASSIFIED

F/G 4/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963 A

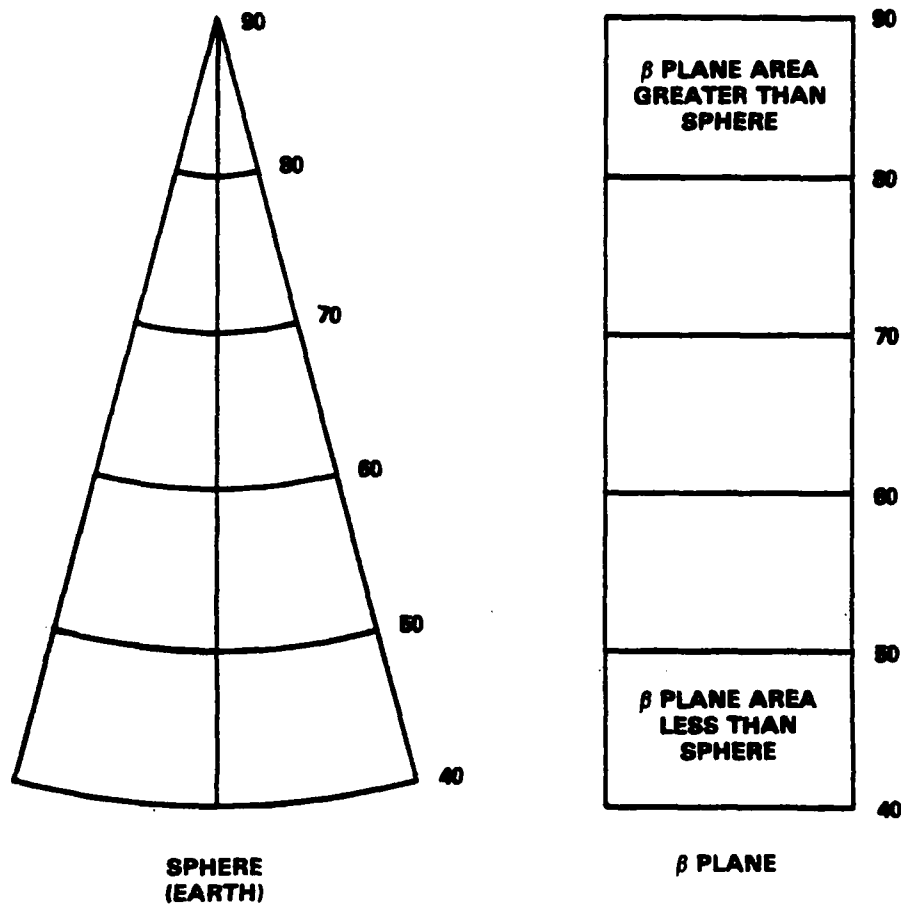


Figure 5.1. Comparison of spherical and β -plane geometry.

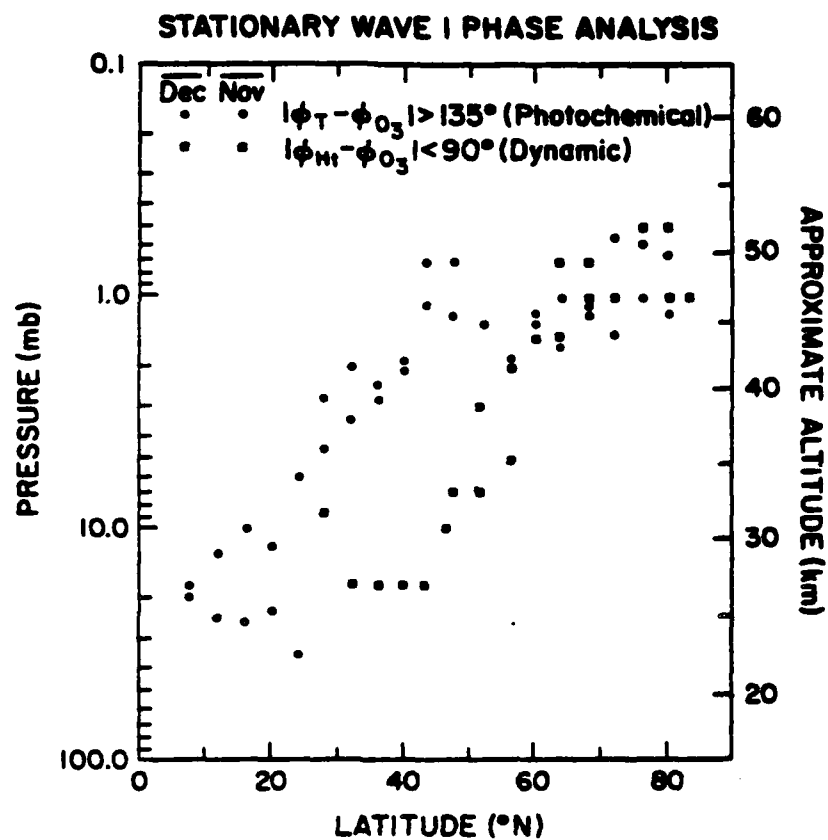
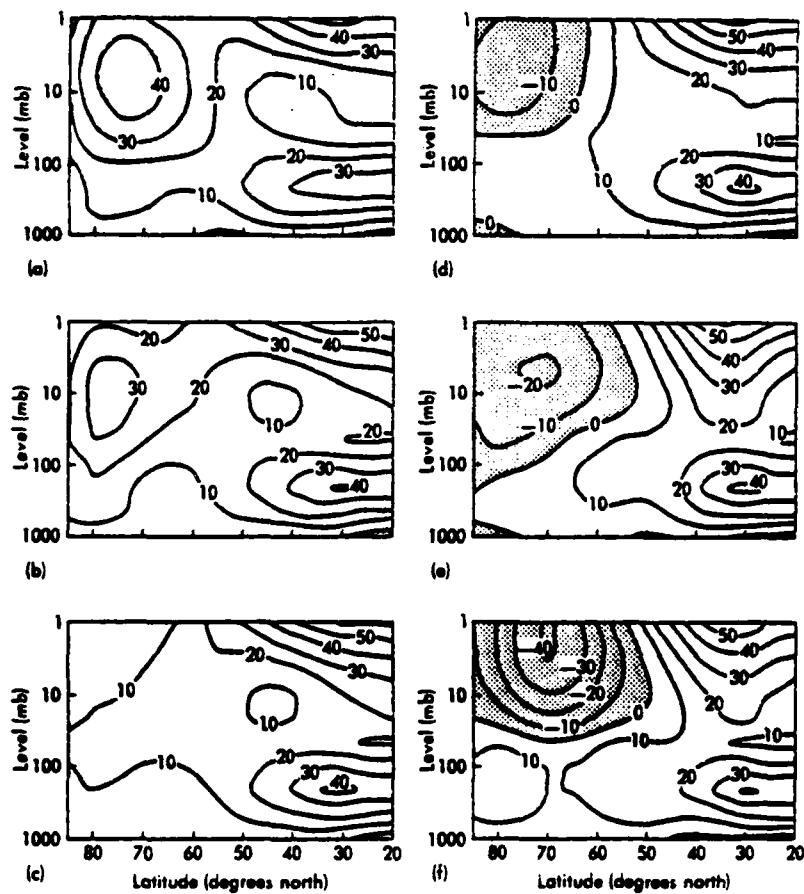


Figure 5.2. Latitudinal dependence of the lower boundary of the photochemically controlled region (circles) and the upper boundary of the dynamically controlled region (squares), for November and December (from Gille et al., 1980).



a. 17 February 1979

d. 21 February

b. 19 February

e. 23 February

c. 20 February

f. 27 February

Figure 5.3. Meridional cross section of the zonal mean wind velocity (m sec^{-1}). (from Palmer, 1981)

d. Planetary waves

The excessive transport out of the southern regions of the model during sudden warmings is influenced by all of the processes discussed above; however, the problem is exaggerated by the simplicity of the planetary wave dynamics. The critical lines in many of the figures in Chapter 4 (Fig 4.23 for instance) are nearly horizontal and extend from the northern boundary to the southern boundary. During stratospheric sudden warmings, the critical line generally does not extend south of 50° , which is close to the center of this model (Fig. 5.3). Furthermore, the critical line is rarely perfectly horizontal and is likely to be in the photochemical region at midlatitudes and in the conservative region at high latitudes. Therefore, much of the excessive transport out of the south exists because the model distorts the horizontal scale of the warming event. More realistic initial wind profiles would help alleviate this problem.

5.3. Vertical structure

a. Small scale dynamics

In Chapter 4 the high latitude total ozone was well estimated, but the polar buildup occurred higher than observed. This was attributed to the lack of subplanetary scale dynamics in the lower stratosphere. Since ozone is nearly conservative in the lower stratosphere, transport here would serve only to redistribute ozone and not change the total amount. Planetary waves which are not capable of propagating to very high altitudes (wavenumber 2-5) are

often present in the lower stratosphere as are the remnants of tropospheric weather systems. Also, it is well known that intense baroclinic zones in the upper troposphere cause excursions of ozone into the troposphere and may be instrumental in the downward transport of ozone (Kida, 1977; Shapiro, 1980). Thus, it appears that a variety of mechanisms exist which can redistribute ozone downward, while maintaining the column density, and that the inability of the current model to properly simulate the vertical distribution does not vitiate the importance of the observed transport mechanisms.

It might be possible to model the subplanetary processes in the lower stratosphere related to the redistribution of ozone by diffusion. However, for chemically active constituents other than ozone, many of the problems that appear when applying diffusion models to ozone might become important. In order to evaluate that effects of pollutants on ozone, the dynamics of the pollutants must be accurately calculated; therefore, any diffusion model must be used judiciously.

b. Photochemistry

Many aspects of the chemistry of ozone as used in this model are inadequate and might contribute to incorrect results. The chemistry was deliberately simplified so that attention could be focused on the transport and interaction of the dynamics and photochemistry. Rather than speculate on the particular effects of the omission of certain chemical species, it is more pertinent to

comment on the spatial variation of the simple chemistry included.

Garcia and Hartmann (1980) found that the latitudinal variation of photochemistry (the effect of the polar night) was important in enhancing the polar buildup of ozone. In Section 4.3 the latitudinal variation of photochemistry was shown to have little effect on transport by the diabatic circulation into the lower stratosphere. When ozone is assumed to be conserved in the polar night, a significant depletion develops over the course of the winter above 40 km. Very weak chemistry, though, will significantly moderate this depletion since the advection time scales are so long.

As the planetary wave transport is modeled here for time dependent waves, the latitudinal variation of photochemistry would have relatively little effect at high latitudes. Most of the transport occurs below the transition layer where ozone is nearly conservative at all times. The reduction of chemistry would have a more noticeable effect on transport during minor warmings when the critical line does not descend very low.

5.4 Unstable waves

During the transient transport an unstable wave exists within the critical level which causes the layered transport phenomenon discussed in Section 4.6. For ozone this layering had little effect, because it occurred in the photochemical region. For a conservative tracer or a compound with its transition region located in the unstable region, the unstable wave may have considerable transport consequences.

5.5 Summary and discussion

Despite the simplifications discussed above, the transport as calculated in this model reproduces many of the features of the wintertime ozone distribution. Furthermore, due to the simplicity of the model, specific mechanisms which are responsible for transport can be identified.

a. Steady state waves

In Chapter 3, transport by steady state dissipative planetary waves was examined from both an Eulerian and Lagrangian-mean viewpoint. In the lower stratosphere these waves were found to cause ozone depletion at low latitudes and an ozone enhancement toward the pole. The increase in ozone in the northern region is consistent with the observations; however, there is not an observed decrease at low latitudes as is calculated here. The ozone transport in the lower stratosphere was due to advection by the Lagrangian-mean velocity associated with the planetary waves.

In the lower part of the transition region there is a increase of mean ozone at high latitudes and a decrease at low latitudes (see Figures 3.7 and 3.14). This behavior in the transition region was shown to be due to the strong photochemical gradients in the region (Sec. 3.3d). The particle trajectories associated with the planetary waves are such that parcels move through a region of widely varying photochemistry. As a result the parcels experience a photochemical forcing different from the mean field chemistry at the parcel's equilibrium position. Thus the perturbation chemistry

cannot be balanced by the mean field chemistry, and changes in the mean ozone mixing ratio occur.

Where the photochemical gradients are small there is little or no effect of the planetary waves on the mean field. The steady state model indicates that transport by steady planetary waves might be adequately represented by advection by the Lagrangian-mean velocity in the lower stratosphere, and some scheme which includes the parcels orbiting through regions of strong photochemical gradients. Indeed, a model based on observed trajectories has been recently proposed by Garcia and Solomon (1982).

An accurate estimate of the Lagrangian-mean velocity for a realistic model of the stratosphere is difficult to calculate since the Lagrangian-mean velocity is the residual between two large terms. In atmospheric data since there are large uncertainties in the estimates of the quantities needed to produce the Lagrangian-mean displacement fields, it would be even more difficult to form meaningful estimates of the Lagrangian-mean velocity from observations.

In Section 3.3c the steady state ozone transport using the Eulerian residual circulation was calculated. The results were similar to those found using the Lagrangian-mean velocity, but the advection was consistently overestimated. The greatest differences between the residual velocity and the Lagrangian-mean velocity occur in the photochemical region so that the chemistry obscures the errors in transport. For a constituent, with a different photochemistry from ozone, there might be significant differences

between the Lagrangian-mean and residual transport. It appears that the Eulerian formulation is best for carrying out transport calculations in realistic models, and the other formulations are useful for interpreting the results.

Over the course of the winter, the computed ozone increases in the lower stratosphere at high latitudes show reasonable agreement with observation. Using the "realistic" zonal wind profile of Schoeberl (1981a) too much ozone was transported into the northern region. The depletion of ozone in the southern lower stratosphere, could be partially alleviated with the inclusion of the diabatic circulation. The diabatic circulation would serve to transport ozone into the lower stratosphere and help replenish the depleted areas. However, even with the inclusion of the diabatic circulation the steady state ozone model would not describe the winter ozone transport and would exhibit many of the failures of the total ozone field calculated with the diabatic circulation alone (Sec. 4.3). The north-south transport associated with the planetary waves is larger than that associated with the diabatic circulation but is still not adequate to supply the required northward transport observed in the winter.

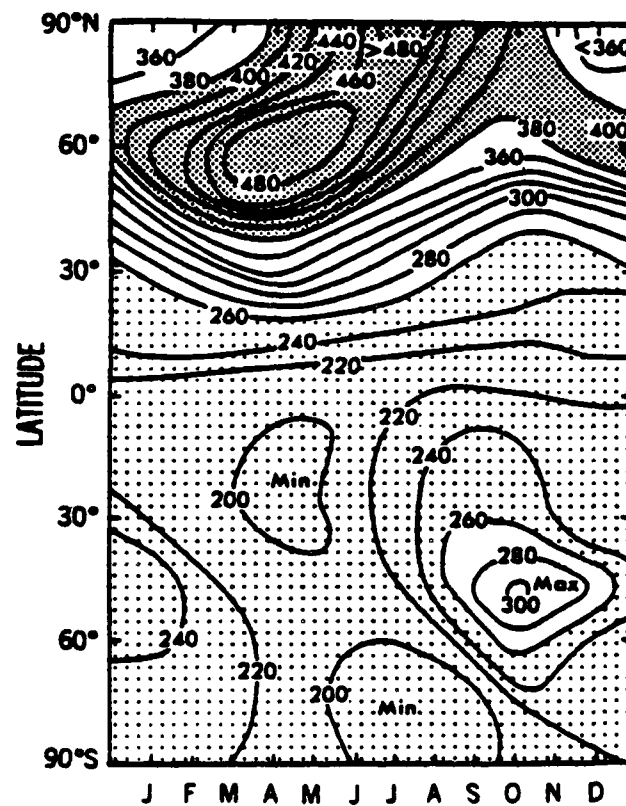


Figure 5.4. Simulated zonal mean total ozone (DU)
(from Mahlman et al., 1980).

It is interesting to consider the results of Mahlman et al. (1980) within the context of the simple transport models used here. Figure 5.4 taken from Mahlman et al. shows their predicted total ozone field. Comparison with Fig. 2.8 shows that they predict a spring ozone maximum which occurs too far south, too late in the year, and that is too large. There is an intrusion of ozone into the polar region during the springtime reversal of the mean circulation. These errors are consistent with the errors associated with transport by the diabatic circulation alone or the diabatic circulation with steady state planetary waves. The results from this study suggest that it is the inability of Mahlman et al. model to simulate stratospheric warmings which is responsible for much of the error in their total ozone field.

b. Time dependent waves

The transport by transient planetary waves (Chapter 4) was found to be much different from steady state transport, but it is conceptually quite simple. Intense poleward transport is associated with the deceleration and reversal of the zonal mean flow during warmings, and equatorward transport is associated with the reacceleration of the flow after the warming. The transport by highly transient planetary waves is largely confined to the horizontal plane with only a small vertical component. More realistic zonal wind profiles might generate critical lines which have a larger vertical transport associated with them (Schoeberl 1981a, 1981b).

With the transport being so much different for transient and steady state waves, the question arises as to the consistency between the two models. To compare the two models an experiment was run in which the time dependent model was forced with a 200 m mountain. Such forcing causes large wave-mean flow interaction initially but there is never a reversal in the zonal mean wind field, and an equilibrium state is reached with westerlies present in the entire domain. At an altitude of 40 km the wave amplitude assumes a near steady value of approximately 350 gpm.

The anomaly field is shown at 74° in Fig. 5.5. There is a positive change in the lower part of the transition region and a negative change just below the transition region which is in qualitative agreement with the steady state model (Fig 3.14). The buildup found in the lower stratosphere at high latitudes is not reproduced in this experiment. The changes in the mean field for the 200 m waves are an order of magnitude or more smaller than those calculated with the 600 and 900 m waves and are also much smaller than those calculated with the steady state model.

The difference in the lower stratosphere between the steady state and time dependent models is due to wave-mean flow interaction in the upper stratosphere. In the time dependent model the wave induced mean velocity fields are strongest in the region of large wave-mean flow interaction. These fields are active at a significantly higher altitude than in the steady state model so that the largest advection by the planetary waves does not occur in the lower stratosphere. This implies that the assumption in the steady

state model (Section 2.3a) that the changes in the mean field are small may significantly affect the results.

It is interesting to compare the eddy flux field from the 200 m experiment with those actually observed by Gille et al. (1980) (Fig 5.6). The model results agree well with the observations suggesting that the 200 m experiment might be appropriate for the early winter.

The steady state model may be useful for modeling transport in the late fall and early winter; however, the time dependent model more accurately represents the atmospheric transport in late winter and spring. Figure 5.7 schematically summarizes the different transport mechanisms studied here. Both planetary wave models deplete ozone in the south and enhance it in the north. The effects of the transition layer are not obvious for transport during stratospheric warmings as is expected because the transport at the critical line is so strong that it obscures this relatively weak effect.

The different topographies used to generate the time dependent dynamics in Chapter 4 allow for the study of transport during all types of warming events: major (900 m), minor (600 m), and final. The critical line remains above 32 km in the 600 m experiment and is analogous to a minor warming while the 900 m critical line extends below 32 km as it does during a major warming. The persistence of the critical line for such a long period of time after the 900 m warming allows for the study of transport during a final warming.

CHANGE IN $\bar{\mu}$ (ppm)

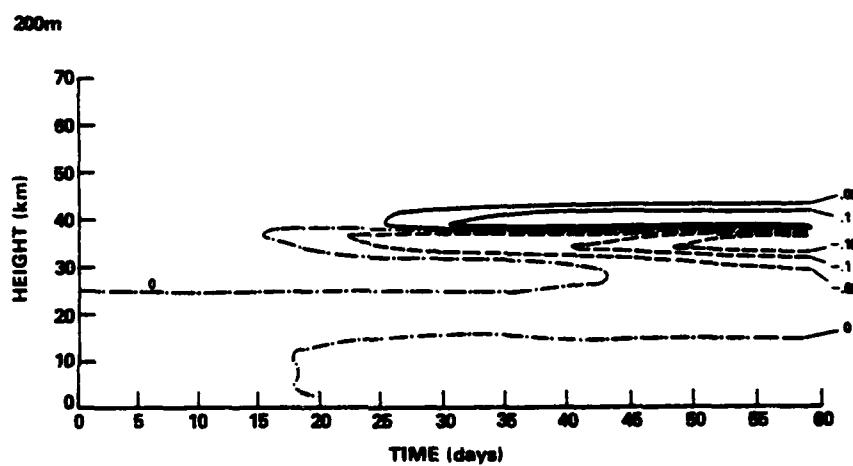
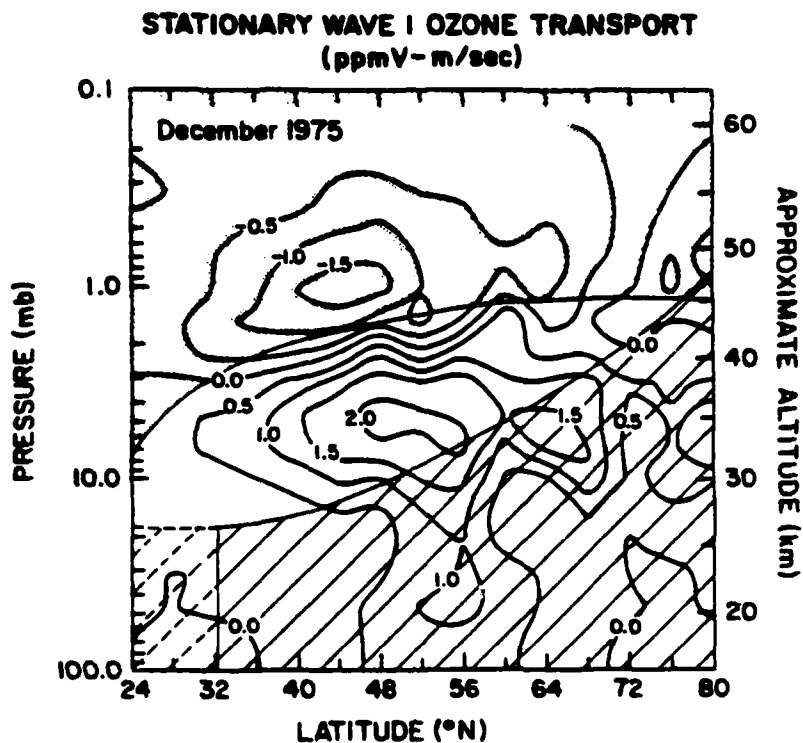
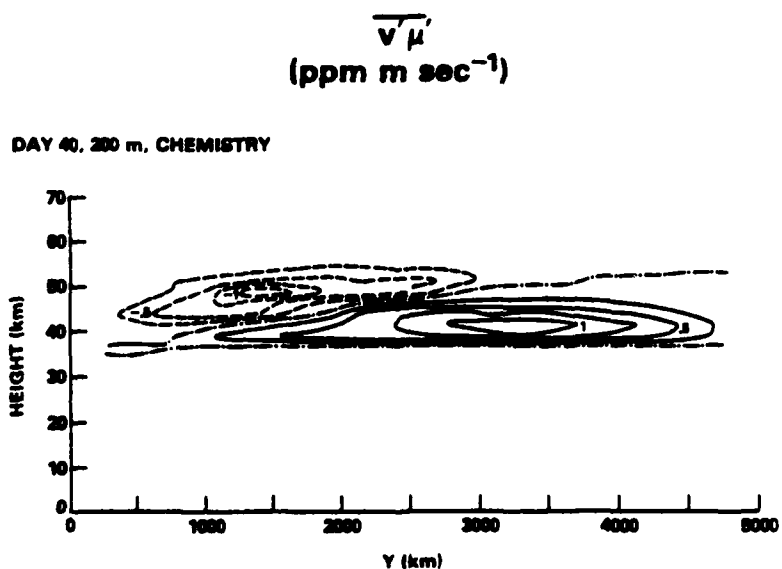


Figure 5.5. Change in ozone at 74° for 200 m topography (ppm).

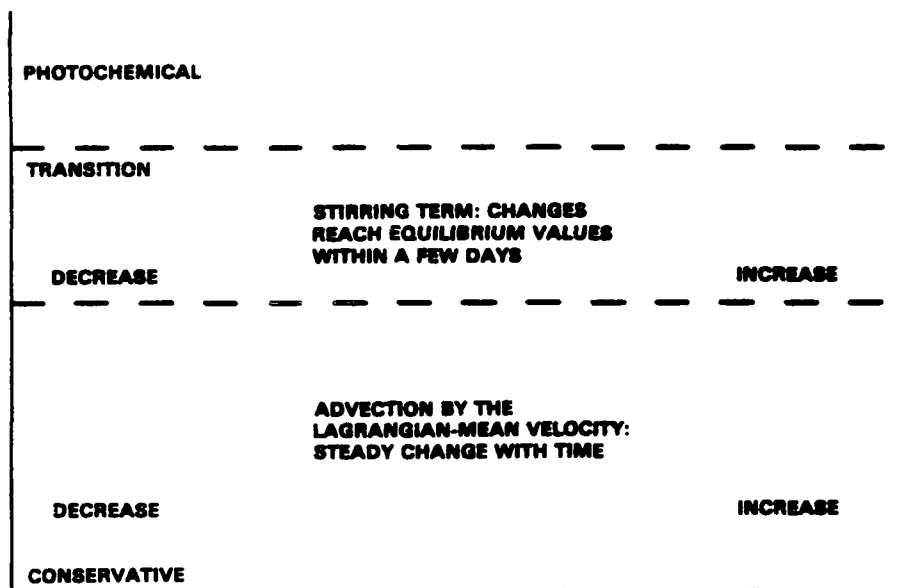


a. measured by Gille *et al.* (1980)



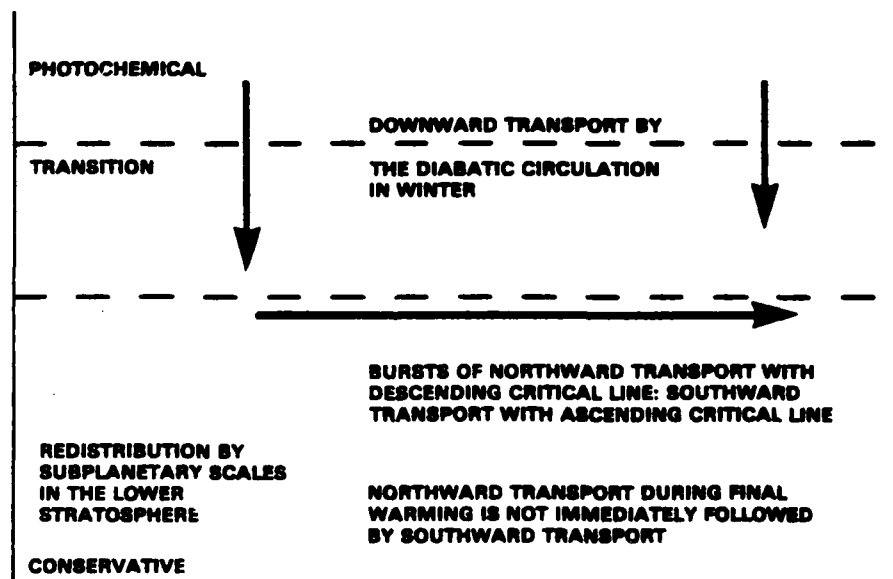
b. for 200 m topography

Figure 5.6. Horizontal eddy flux of ozone (ppm m sec⁻¹).



STEADY STATE TRANSPORT

a. steady state planetary waves



DIABATIC-CRITICAL LINE TRANSPORT

b. diabatic circulation and transient waves

Figure 5.7. Schematic summary of transport mechanisms.

The transport for the 900 m topography is a factor of 3 or more greater than that for the 600 m topography. The increased transport associated with the 900 m topography arises because the critical line descends through the transition region into the conservative region where the transport is not countered by the chemistry. This suggests that transport during major warmings should be more pronounced than during minor warmings. This behavior is consistent with the observations in Fig. 2.10 where in the two years that major warmings occur the ozone maximum at high latitudes is 6% greater than the year in which no major warming occurs.

After both major and minor warmings there is a tendency for westerlies to develop and the stratosphere to return to its climatological state. In this event the critical line rises and much of the ozone transported into the north is transferred back to the south. This is clearly seen by the slight rise in the critical line that occurs in the 900 m experiment on day 47 (Fig. 4.3). Such behavior might be responsible for the oscillatory structure seen in Fig. 2.10.

The rise in the 900 m critical line (Fig 4.3) followed by the descent and persistence of the critical line is representative of what happens in a winter when there is a major warming followed by a final warming. The descent of the critical line after Feb. 1 is instrumental in the formation of the sharp high latitude total ozone maximum calculated with the combined circulation (Fig 4.9).

The suggestion is made from this model that transport during major sudden warmings is large and may have significant effect on

the zonal mean ozone distribution, but that the high latitude spring maximum is due largely to transport during the spring reversal. Following the spring reversal the stratosphere remains in an easterly state so that there is not a strong reverse flow. It might also be suggested that the fall minimum is associated with the reformation of westerlies in the stratosphere.

c. A possible transport scenario

Using the transport mechanisms described in this dissertation, and the observed behavior of ozone in both the stratosphere and general circulation models; it is possible to derive a possible scenario for large scale ozone transport in the stratosphere. First it is necessary to form an analogue of the dynamics in the different models used here to the atmosphere.

Diabatic forcing determines the direction of the mean zonal wind, which in turn controls the propagation of waves into the stratosphere; therefore, the planetary wave transport and the diabatic transport are not completely independent as modeled here. The changeover from winter westerlies to summer easterlies after the final stratospheric warming is related the insolation at high latitudes forcing the zonal mean wind in a easterly direction. It is possible that major and minor warmings in the winter cause the springtime reversal to occur somewhat earlier than if diabatic forcing alone were responsible for the reversal. This possibility is supported by the reversal of the zonal westerlies in the Southern Hemisphere which occur somewhat later in the winter than in the

Northern Hemisphere. Stratospheric warmings are not as a predominate feature of the Southern Hemisphere circulation as of the Northern Hemisphere circulation. Furthermore, the model of Mahlman et al. (1980) shows the intrusion of ozone into high latitudes coincident with the spring reversal. This intrusion occurs later than is actually observed. Since this general circulation model does not simulate stratospheric warmings, the idea that planetary wave activity is important in the formation of the summertime easterlies is further supported.

During summer, with continual diabatic forcing at the pole, easterlies dominant the stratospheric circulation and planetary waves become unimportant. In the fall as the insolation decreases at high latitudes westerlies reform and planetary wave activity once again increases.

In Figure 4.8 the diabatic circulation produced a fall minimum in October in good agreement with observation. The minimum occurs at the time that diabatic circulation is changing over from upward motion at high latitudes during the summer to the downward motion present in the winter. This minimum is also coincident with the change over from zonal mean easterlies to zonal mean westerlies in the stratosphere.

After the ozone minimum in October, downward advection by the diabatic circulation tends to build up ozone in the lower stratosphere. As ozone is increasing, planetary wave activity is becoming more vigorous and warming events begin to occur. "Steady" waves constantly transport ozone poleward, but during major

stratospheric warmings large amounts of ozone are transported northward. Due to small scale "diffusive" processes, some of the ozone transported to high latitudes during warmings remains there leading to the enhanced extratropical total ozone evidenced in Fig. 2.10. Also, as ozone is being advected into the conservative region, small scale dynamics conservatively redistribute it into the lower stratosphere and troposphere. All of these features are evident in Fig. 2.6 where there is first an increase of ozone at 30 mb and then downward propagation from mid-January through March.

At the time of the final warming, which is determined not only by planetary wave forcing but also by insolation at high latitudes, there is strong northward transport. The total ozone maximum is coincident with the breakdown of the polar vortex; this maximum occurs somewhat earlier than the maximum calculated using the diabatic circulation in Fig. 4.9. The transport during the final warming is strong because of the interaction of the well developed winter planetary waves with the critical line. During fall when the westerlies reappear, such dramatic transport does not occur because the wave-mean flow interaction is much smaller. After the spring maximum, during summer, the diabatic circulation slowly reduces the total ozone as it transports it upward into the photochemical region where the ozone then assumes its equilibrium value.

In the Southern Hemisphere the intrusion of ozone into polar latitudes occurs during the reversal of the mean circulation associated with the diabatic forcing. This indicates the reduced

importance of major stratospheric warmings in Southern Hemisphere dynamics.

Transport as described by this scenario suggests that the diabatic circulation (Brewer-Dobson circulation) is responsible for many of the large scale features of the observed ozone distribution. However, the diabatic circulation cannot effectively transport ozone into polar regions, and transport to very high latitudes is accomplished by planetary waves during sudden and final warmings. Subplanetary scale and evanescent planetary scales redistribute ozone in the lower stratosphere and transport ozone into the troposphere.

Finally, some comment must be made about the implications of the modeled transport on diffusion models. The transport by the diabatic circulation is advective and simple to model. The advective transport by steady planetary waves does not seem to be adequate to model the northward transport of ozone during winter. The transport during sudden warmings is large, coherent, rapid, and not diffusive.

The model results indicate that the large scale ozone transport is dominated by vertical diabatic transport and rapid pulses of horizontal transport during warmings. The planetary waves are not responsible for much vertical transport; though, subplanetary waves in the lower stratosphere might cause significant vertical transport.

The diffusion coefficients as calculated by Strobel (1981) clearly indicate the complexities involved in accurately calculating

diffusion coefficients. They are a strong function of source terms, mean field gradients, and wave transience. The type of transport calculated here is highly transient and indicates that rapidly varying diffusion coefficients are required to properly simulate transport processes. If these coefficients are to be calculated from dynamic and chemical quantities then the entire purpose of diffusion parameterizations, computational simplicity, is defeated.

ACKNOWLEDGMENTS

This research was supported by the Office of Naval Research and the Upper Atmospheric Research Office of NASA.

Appendix A

NOMENCLATURE AND CONVENTIONS

a Newtonian cooling and Rayleigh friction coefficients

f Coriolis parameter ($1 \times 10^{-4} \text{sec}^{-1}$)

H scale height (7 km)

k zonal wave number ($6.28 \times 10^{-9} \text{cm}^{-1}$)

N^2 Brunt-Väisälä frequency squared ($4 \times 10^{-4} \text{sec}^{-2}$)

p pressure; p_0 reference pressure

$\bar{q}_y$ horizontal derivative of the mean potential vorticity

Q photochemical source term

$(u,v,w) = (u_1, u_2, u_3) = (\text{zonal, meridional, vertical})$

velocity

$(v^*, w^*) = (\text{meridional, vertical})$ residual velocity
(Andrews and McIntyre, 1976)

t time

$(x, y, z) = (x_1, x_2, x_3) = (\text{east-west, north-south,}$
 $z = -H \ln \frac{p}{p_0})$ coordinates

β meridional variation of Coriolis parameter, f

λ photochemical time constant

μ ozone mixing ratio

$(\xi, \eta, \zeta) = (\xi_1, \xi_2, \xi_3) = (\text{longitudinal, meridional,}$
vertical)
displacement fields

ϕ geopotential

CONVENTIONS:

For a given T

$\bar{T} \equiv$ zonal average (Eulerian)

$T' = T - \bar{T} \equiv$ zonal deviation

$T^L \equiv$ Lagrangian-mean (Andrews and McIntyre, 1978)

$T^S = T^L - \bar{T} \equiv$ Stokes correction (see Eq. (3.5))

Appendix B

NUMERICAL PROPERTIES OF THE ADVECTION-DIFFUSION EQUATION

The one dimensional advection-diffusion equation is

$$\frac{\partial S}{\partial t} + u \frac{\partial S}{\partial x} = D \frac{\partial^2 S}{\partial x^2} \quad (B.1)$$

Using second order, centered time differences and generalized spatial differences (B.1) can be written in discrete form as

$$\begin{aligned} S_i^{n+1} = S_i^{n-1} - \frac{2u \Delta t}{\Delta x} \{ A_1(S_{i+1}^n - S_{i-1}^n) + B_1(S_{i+2}^n - S_{i-2}^n) \} \\ + \frac{2D \Delta t}{\Delta x^2} \{ A_2(S_{i+1}^n - 2S_i^n + S_{i-1}^n) + B_2(S_{i+2}^n - 2S_i^n + S_{i-2}^n) \} \quad (B.2) \end{aligned}$$

where Δx is the spatial grid increment and Δt is the time step. The superscripts are time indices and the subscripts are space indices. Both second order and fourth order differences will be considered, and the values of the A and B coefficients in (B.2) are given in table B.1.

Table B.1
Coefficients for Spatial Difference Scheme

	2 nd order	4 th order
A ₁	1/2	2/3
B ₁	0	-1/12
A ₂	1	4/3
B ₂	0	-1/12

Assuming the solution can be written as

$$S = B e^{n\Delta t} e^{imk\Delta x} \quad (B.3)$$

where m is the spatial index, then (B.2) can be written following Haltiner and Williams (1980, Chapter 5) as

$$B^{\Delta t} = B^{-\Delta t} - 1 \frac{4u \Delta t}{\Delta x} \{A_1 \sin(k\Delta x) + B_1 \cos(2k\Delta x)\}$$

(B.4)

$$+ \frac{4D \Delta t}{\Delta x^2} \{A_2(\cos(k\Delta x)-1) + B_2(\cos(2k\Delta x)-1)\}$$

In the transport model there is no diffusion so first consider (B.4) with $D = 0$. It can then be shown that for second order differences

$$\frac{u \Delta t}{\Delta x} < 1 \text{ for stability, and for fourth order differences}$$

$$\frac{u \Delta t}{\Delta x} < .73 \text{ for stability.}$$

Thus, for a given Δx , the fourth order scheme requires a smaller time step than the second order scheme to assure numerical stability. However, as demonstrated in Fig. 2.1, the phase error for the fourth order differences is considerably reduced (see Haltiner and Williams, 1980, Table 5.1). Haltiner and Williams (1980) discuss in detail the numerical properties of the second order and fourth order advection equation. Kreiss and Oliger (1972) argue that, by comparing the error in the phase speed with the number of computations, the fourth order scheme is "optimal."

Clancy (1981) suggests using forward time differences and second order spatial differences to integrate the advection-diffusion equation. Clancy derives a new stability criterion,

$$\Gamma R_c < 1$$

where

$$\Gamma = \frac{u \Delta x}{\Delta t} \quad \text{and} \quad R_c = \frac{u \Delta x}{D} \quad (\text{B.5})$$

Using this stability criterion R_c can be infinitely large if Γ is small enough to keep the inequality (B.5) true. It had been shown previously that R_c must be less than or equal to 2, if this scheme is to be numerically stable. In the derivation of the stability criterion (B.5), however, Clancy effectively divided by zero in the formation of his equation (15), and the usefulness of (B.5) is questionable.

If the first term on the right hand side of (B.4) is replaced by 1 and the second and third terms are divided by 2, then (B.4) is the appropriate equation for forward time differences instead of centered time differences. If $D = 0$, this finite difference scheme becomes the Euler scheme, and it can be shown that the Euler scheme is unstable for almost all wavelengths (Haltiner and Williams, 1980, p.130). The addition of diffusion to this method serves to stabilize an otherwise unstable numerical scheme.

This stabilizing effect is illustrated in Fig. B.1 where the growth rate, G , is plotted against the diffusion parameter, D . If G is greater than 1 then numerical instability exists. As D increases from 0, G decreases until it is less than 1 and the scheme is stable. As D gets very large G once again increases and the method becomes unstable. So while it is possible to make the Euler scheme

stable by adding diffusion, the physical meaning of the diffusion is obscure; because the diffusion coefficient must not only represent physical processes, but must also counteract the inherent numerical instability.

While it is possible to produce a finite difference formulation that appears to be stable using Clancy's criterion 3.5, this is done at great computational expense. For Clancy's ocean case (the case presented in Fig. 2.1) $R_c = 60$ and $\Gamma = .022$. With Γ of this magnitude it takes almost 50 time steps for a parcel to propagate through one grid increment. Using centered spatial differences the time step can be increased by more than an order of magnitude and better results are obtained than using the Clancy algorithm. Using fourth order spatial differences, with a 10 times larger time step, the numerical results and the analytic results are very nearly equal (Fig. B.2). Any additional computations required by the fourth order scheme are more than compensated for by the increased accuracy and the larger time step.

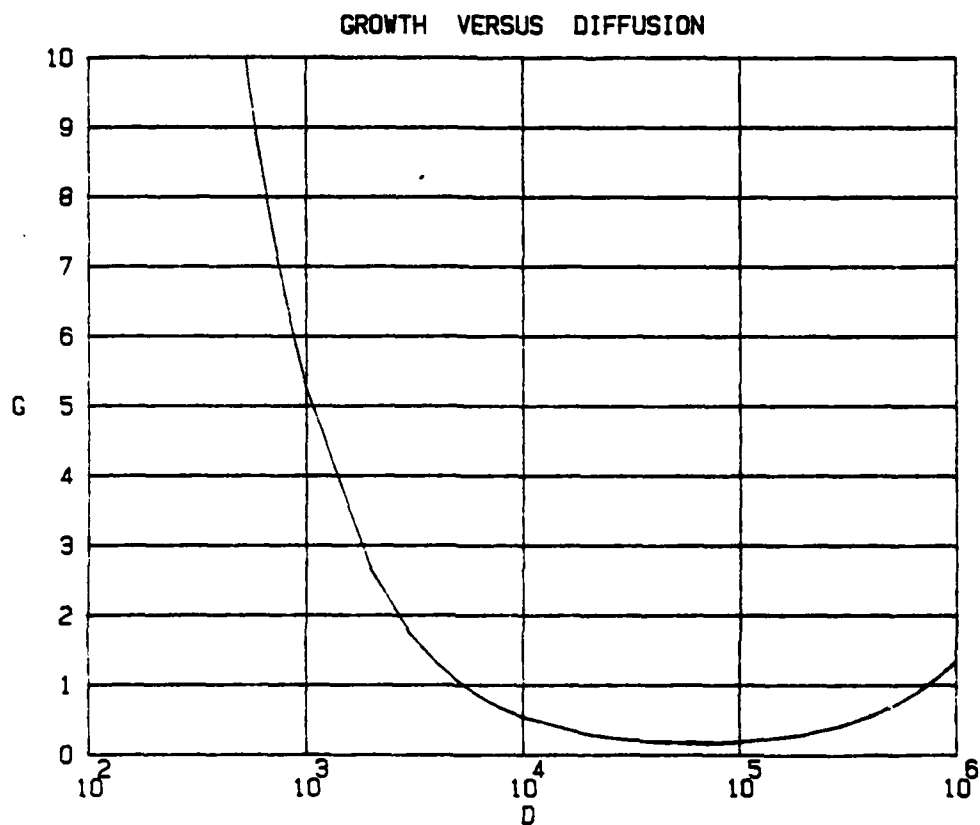


Figure B.1. Growth rate of the numerical solution for forward time differences as a function of diffusion.

For $G > 0$, the solution is unstable.

($\Delta x = 100$ km, $\Delta t = 10$ hours, $u = .6$ m sec⁻¹)

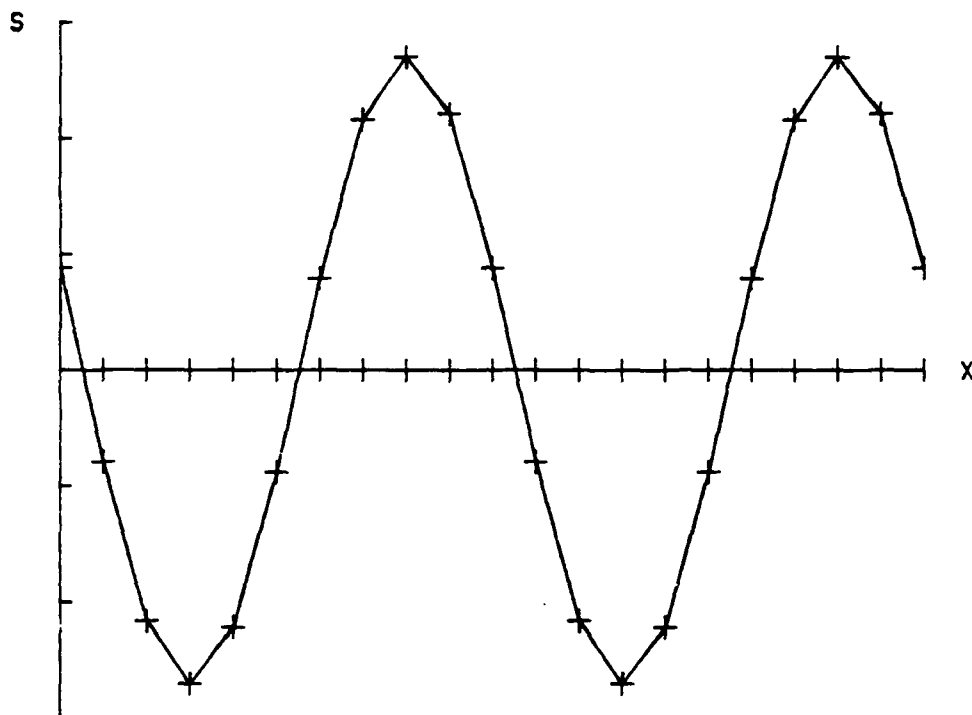


Figure B.2. Integration of Clancy's (1981) oceanic case with $\Delta t = 10$ hours. Same as Figure 2.1c but with the time step increased an order of magnitude. The numerical solution (+) is virtually indistinguishable from the analytic solution

References

- Ames, W.F., Numerical Methods for Partial Differential Equations, 2nd Ed., Academic Press, New York, pp.365, 1977.
- Andrews, D.G. and M.E. McIntyre, Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration, J. Atmos. Sci., 33, 2031-2048, 1976.
- Andrews, D.G. and M.E. McIntyre, An exact theory of nonlinear waves on a Lagrangian-mean flow, J. Fluid Mech., 89, 609-646, 1978.
- Apruzese, J.P., M.R. Schoeberl, and D.F. Strobel, Parameterization of IR cooling in a middle atmosphere dynamics model, J. Geophys. Res., in press, 1982.
- Berkofsky, L. and E.A. Bertoni, Mean topographic charts for the entire earth, Bull. Amer. Meteor. Soc., 36, 350-354, 1955.
- Benson, S.W. and A.E. Axworthy, Jr., Mechanism of the gas phase, thermal decomposition of ozone, J. Chem. Phys., 26, 1718-1726, 1957.
- Boyd, J.P., The noninteraction of waves with the zonally averaged flow on a spherical earth and the interrelationships of eddy fluxes of energy, heat, and momentum, J. Atmos. Sci., 33, 2285-2291, 1976.

- Brewer, A.W., Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, Quart. J. Roy. Meteor. Soc. 75, 351-363, 1949.
- Chapman, S., A theory of upper-atmospheric ozone, Mon. Roy. Meteor. Soc., 3, 103-125, 1930.
- Charney, J.G. and P.G. Drazin, Propagation of planetary-scale disturbances from the lower into the upper stratosphere, J. Geophys. Res., 66, 83-109, 1961.
- Charney, J.G. and M.E. Stern, On the stability of internal baroclinic jets in a rotating atmosphere, J. Atmos. Sci., 19, 159-172, 1962.
- Clancy, R.M., A note on finite differencing of the advection-diffusion equation, Mon. Wea. Rev., 109, 1807-1809, 1981.
- Clark, J.H.E., A quasi-geostrophic model of the winter stratospheric circulation, Mon. Wea. Rev., 98, 443-461, 1970.
- Clark, J.H.E. and T.G. Rogers, The transport of conservative trace gases by planetary waves, J. Atmos. Sci., 35, 2232-2235, 1978.
- Craig, R.A., The observations and photochemistry of atmospheric ozone and their meteorological significance, Meteor. Monogr., 1, No. 2, 1-50, 1950.
- Craig R.A., The Upper Atmosphere: Meteorology and Physics, Academic Press, New York, pp. 509, 1965.
- Cunnold, D., F. Alyea, N. Phillips, and R. Prinn, A three-dimensional dynamical-chemical model of atmospheric ozone, J. Atmos. Sci., 32, 170-194, 1975.

- Danielsen, E.F., An objective method for determining the generalized transport tensor for two-dimensional Eulerian models, J. Atmos. Sci., 38, 1319-1339, 1981.
- Dickinson, R.E., Method of parameterization for infrared cooling between altitudes of 30 and 70 kilometers, J. Geophys. Res., 78, 4451-4457, 1973.
- Dobson, G.M.B., Observations of the amount of ozone in the earth's atmosphere and its relation to other geophysical conditions, part IV, Proc. Roy. Soc. London, A129, 411-433, 1930.
- Dobson, G.M.B., Origin and distribution of polyatomic molecules in the atmosphere, Proc. Roy. Soc. London, A236, 187-193, 1956.
- Dunkerton, T., On the mean meridional mass motions of the stratosphere and mesosphere, J. Atmos. Sci., 35, 2325-2333, 1978.
- Dunkerton, T., A Lagrangian-mean theory of wave, mean-flow interaction with applications to nonacceleration and its breakdown. Rev. Geophys. Space Phys., 18, 387-400, 1980.
- Dunkerton, T., C.-P.F. Hsu, and M.E. McIntyre, Some Eulerian and Lagrangian diagnostics for a model stratospheric warming, J. Atmos. Sci., 38, 819-843, 1981.
- Dütsch, H.U., Atmospheric ozone and ultraviolet radiation, World Survey of Climatology, Vol. 4, D.F. Rex, Ed., Elsevier, 383-432, 1969.
- Dütsch, H.U., The ozone distribution in the atmosphere, Can. J. Chem., 52, 1491-1504, 1974.

- Fels, S.B., J.D. Mahlman, M.D. Schwarzkopf, and R.W. Sinclair, Stratospheric sensitivity to perturbations in ozone and carbon dioxide: Radiative and dynamic response, J. Atmos. Sci., 37, 2265-2297, 1980.
- Garcia, R.R. and D.L. Hartmann, The role of planetary waves in the maintenance of the zonally averaged ozone distribution of the upper atmosphere, J. Atmos. Sci., 37, 2248-2264, 1980.
- Garcia, R.R. and S. Solomon, A numerical model of the zonally-averaged dynamical and chemical structure of the middle atmosphere, submitted to J. Geophys. Res., 1982.
- Gille, J.C., G.P. Anderson, W.J. Kohri, and P.L. Bailey, Observation of the interaction of ozone and dynamics, Proceedings of the Quadrennial International Ozone Symposium, J. London, Ed., 1007-1011, 1980.
- Götz, F.W.P., Ozone in the atmosphere, in Compendium of Meteorology, American Meteorology Society, Boston, 275-291, 1951.
- Haltiner, G.J. and R.T. Williams, Numerical Prediction and Dynamic Meteorology, 2nd Ed., John Wiley and Sons, Inc., New York, pp. 477, 1980.
- Hartmann, D.L., Some aspects of the coupling between radiation, chemistry, and dynamics in the stratosphere, J. Geophys. Res., 86, 9631-9640, 1981.
- Hartmann, D.L. and R. Garcia, A mechanistic model of ozone transport by planetary waves in the stratosphere, J. Atmos. Sci., 36, 350-364, 1979.

- Harwood, R.S. and J.A. Pyle, Studies of the ozone budget using a zonal mean circulation model and linearized photochemistry, Quart. J. Roy. Meteor. Soc., 103, 319-343, 1977.
- Hidalgo, H. and P.J. Crutzen, The tropospheric and stratospheric composition perturbed by NO_x emissions of high altitude aircraft, J. Geophys. Res., 82, 5833-5866, 1977.
- Hilsenrath E., D.F. Heath, B.M. Schlesinger, Seasonal and interannual variations in total ozone revealed by the Nimbus 4 backscattered ultraviolet experiment, J. Geophys. Res., 84, 6969-6979, 1979.
- Holton, J.R., The Dynamic Meteorology of the Stratosphere and Mesosphere, Meteor. Monographs, Vol. 15, #37, pp. 218, 1975.
- Holton, J.R., An advective model for two dimensional transport of stratospheric trace species, J. Geophys. Res., 86, 11989-11994, 1981.
- Hsu, C.-P., F., Air parcel motions during a numerically simulated sudden stratospheric warming, J. Atmos. Sci., 37, 2768-2792, 1980.
- Hunt, B.G., Photochemistry of ozone in a moist atmosphere, J. Geophys. Res., 71, 1385-1398, 1966.
- Hunten, D.M., Vertical transport in atmospheres, in Atmospheres of Earth and the Planets, B.M. McCormac, Dordrecht, Holland, 59-72, 1975.
- Johnston, H.S. and J. Podolske, Interpretations of stratospheric photochemistry, Rev. Geophys. Space Phys., 16, 491-520, 1978.

- Kida, H., A numerical investigation of the stratospheric general circulation and stratospheric tropospheric mass exchange. II. Lagrangian motion of the atmosphere, J. Meteor. Soc. Japan, 55, 71-88, 1977.
- Kreiss, H.O. and J. Oliger, Comparison of accurate methods for the integration of hyperbolic equations, Tellus, 24, 199-215, 1972.
- Kreuger, A.J. and R.A. Minzner, A mid-latitude ozone model for the 1975 U.S. Standard Atmosphere, J. Geophys. Res., 81, 4477-4481, 1976.
- Labitzke, K. and B. Goretzki, A catalogue of dynamic parameters describing the variability of the middle stratosphere during the northern winters, in Handbook for MAP, 5, Middle Atmosphere Program, C.F. Sechrist, Ed., 1982.
- Lindzen, R.S., B. Farrel, and K.K. Tung The concept of wave overreflection and its application to baroclinic instability, J. Atmos. Sci., 37, 44-63, 1980.
- London, J., R. Bojkov, S. Oltmans, and J. Kelley, Atlas of the Global Distribution of Total Ozone July 1957-June 1967, NCAR Technical Note, TN113, Boulder, pp. 276, 1976.
- Luther, F.M., Large scale eddy transport, in Second Annual Report, UCRL-51336-74, pp. 66-73, Lawrence Livermore Lab., Calif., 1974.
- Mahlman, J.D., Heat balance and mean meridional circulations in the polar stratosphere during the sudden warming of January 1958, Mon. Wea. Rev., 97, 534-540, 1969.

Mahlman, J.D., Some fundamental limitations of simplified transport models as implied by results from a three-dimensional general circulation/tracer model, Proc. Fourth Conference Climatic Impact Assessment Program, T.M. Hard and A.J. Broderick, ed. DOT-TSC-OST-75-38, U.S. Dept. of Trans., Washington, D.C., 132-146, 1975.

Mahlman, J.D. and W.M. Moxim, Tracer simulation using a global general circulation model: Results from a midlatitude instantaneous source experiment, J. Atmos. Sci., 35, 1340-1374, 1978.

Mahlman, J.D. and R.W. Sinclair, Tests of various numerical algorithms applied to a simple trace constituent air transport problem, Fate of Pollutants in the Air and Water Environments, Part 1, 8, I.H. Suffet, ed., John Wiley and Sons, Inc., 1977.

Mahlman, J.D., H. Levy II, and W.J. Moxim, Three-dimensional tracer structure and behavior as simulated in two ozone precursor experiments, J. Atmos. Sci., 37, 655-685, 1980.

Matsumo, T., Lagrangian motion of air parcels in the stratosphere in the presence of planetary waves, Pure Appl. Geophys., 118, 189-216, 1980.

Matsumo, T. and K. Nakamura, The Eulerian- and Lagrangian-mean meridional circulations in the stratosphere at the time of a sudden warming, J. Atmos. Sci., 36, 640-654, 1979.

McConnell, J.C. and M.B. McElroy, Odd nitrogen in the atmosphere. J. Atmos. Sci., 30, 1465-1477, 1973.

- McIntyre, M.E., Towards a Lagrangian-mean description of stratospheric circulations and chemical transports, Phil. Trans. Roy. Soc. Lond., A296, 129-148, 1980.
- Miller, C., D.L. Filkin, A.J. Owens, M.J. Steed, and J.P. Jesson, A two dimensional model of stratospheric chemistry and transport, J. Geophys. Res., 86, 12039-12065, 1981.
- Murgatroyd, R.J. and F. Singleton, Possible meridional circulation in the stratosphere and mesosphere, Quart. J. Roy. Meteor. Soc., 87, 125-135, 1961.
- Newell, R.E., The transport of trace substances in the atmosphere and their implications for the general circulation of the stratosphere, Geofis. Pura Appl., 49, 137-158, 1961.
- Newell, R.E., Transfer through the tropopause and within the stratosphere, Quart. J. Roy. Meteor. Soc., 89, 167-204, 1963.
- Palmer, T.N., Diagnostic study of a wavenumber-2 stratospheric sudden warming in a transformed Eulerian-mean formalism, J. Atmos. Sci., 38, 844-855, 1981.
- Plumb, R.A., Eddy fluxes of conserved quantities by small-amplitude waves, J. Atmos. Sci., 36, 1699-1704, 1979.
- Pyle, J.A. and C.F. Rogers, Stratospheric transport by stationary planetary waves - The importance of chemical processes, Quart. J. Roy. Meteor. Soc., 106, 421-446, 1980a.
- Pyle, J.A. and C.F. Rogers, A modified diabatic circulation model for stratospheric tracer transport, Nature, 287, 711-714, 1980b.

- Reed, R.J. and K.E. German, A contribution to the problem of stratospheric diffusion by large-scale mixing, Mon. Wea. Rev., 93, 313-321, 1965.
- Schlesinger, M.E. and Y. Mintz, Numerical simulation of ozone production, transport and distribution with a global atmospheric general circulation model, J. Atmos. Sci., 36, 1325-1361, 1979.
- Schoeberl, M.R., The secondary circulation associated with a vertically aligned planetary wave critical line, Geophys. Res. Lett., 7, 153-156, 1980.
- Schoeberl, M.R., A simple model of the Lagrangian-mean flow produced by dissipating planetary waves, J. Atmos. Sci., 38, 1841-1855, 1981a.
- Schoeberl, M.R., The secondary flow near a baroclinic planetary wave critical line, J. Atmos. Sci., 38, 630-638, 1981b.
- Schoeberl, M.R., Vacillation, sudden warmings and potential enstrophy balance in the stratosphere, to appear in J. Atmos. Sci., 1982a.
- Schoeberl, M.R., A study of stratospheric vacillations and sudden warmings on a β -plane Part I: Single wave-mean flow interaction, submitted to J. Atmos. Sci., 1982b.
- Schoeberl, M.R. and D.F. Strobel, The zonally averaged circulation of the middle atmosphere, J. Atmos. Sci., 35, 577-591, 1978.
- Shapiro, M.A., Turbulent mixing within tropospheric folds as a mechanism for the exchange of chemical constituents between the stratosphere and troposphere, J. Atmos. Sci., 37, 994-1004, 1980.

- Strobel, D.F., Parameterization of linear wave chemical transport in planetary atmospheres by eddy diffusion, J. Geophys. Res., 86, 9806-9810, 1981.
- Tuck, A.F., A comparison of one- two- three- dimensional model representations of stratospheric gases, Phil. Trans. Roy. Soc. Lon., A290, 477-495, 1979.
- van Loon, H., R.L. Jenne and K. Labitzke, Zonal harmonic standing waves, J. Geophys. Res., 78, 4463-4471, 1973.
- WMO, The Stratosphere 1981: Theory and Measurements, WMO Global Ozone Research and Monitoring Project Report No.11, 1981.
- Zullig, W., Relation between the intensity of the stratospheric circumpolar vortex and the accumulation of ozone in the winter hemisphere, Pure Appl. Geophys., 106-108, 1544-1552, 1973.

APPENDIX Q

A MECHANISTIC MODEL OF
EULERIAN, LAGRANGIAN-MEAN, AND LAGRANGIAN OZONE
TRANSPORT BY STEADY PLANETARY WAVES

A Mechanistic Model of
Eulerian, Lagrangian-mean, and Lagrangian Ozone
Transport by Steady Planetary Waves

Richard B. Rood*
Science Applications, Inc.
McLean, Virginia 22102

and

Mark R. Schoeberl
Geophysical & Plasma Dynamics Branch
Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

February 1982

* Current address: NASA/Goddard Space Flight Center, Code 964,
Greenbelt, MD 20771.

Abstract

Ozone transport is calculated for steady, dissipative planetary waves using the Eulerian, Lagrangian-mean, and residual circulation. A Lagrangian model of parcel dynamics is used to interpret planetary wave-photochemistry interaction. In chemically active regions the mean field ozone changes are found to be significant only where there are large gradients in chemical sources and sinks along parcel trajectories. The largest changes in the mean field are found in the lower stratosphere and are due to the Lagrangian-mean advection. When the Lagrangian-mean advection is approximated by the residual circulation, errors in the transport velocities as large as 30% may occur.

1. Introduction

The atmospheric ozone distribution is maintained by a complex interaction of photochemical, dynamical, and radiational processes. In order to make dynamical modeling of ozone tractable, it has been the practice to severely parameterize either the photochemistry or the dynamics. For instance, 1-D chemistry models have used vertical diffusion to model atmospheric dynamics (Hunten, 1975 and the references therein), while 2-D models have involved a combination of a mean advective field with parameterized eddy dynamics (Harwood and Pyle, 1977; Miller et al., 1981). The wave transport by planetary waves in 2-D models has usually been parameterized as a symmetric diffusion tensor. However, planetary wave effects have been shown to be largely advective (Matsuno, 1980; Dunkerton, 1980; Pyle and Rogers, 1980; Danielsen, 1981; Miller et al., 1981; Strobel, 1981); therefore, the parameterized dynamics have often been very unrealistic (NASA, 1979). Since the transport of ozone away from source regions is important in the maintenance of the ozone distribution it is necessary to use realistic dynamic processes in ozone models.

In order to identify the mechanisms responsible for ozone transport, and therefore the proper dynamical processes to be included in chemical models a series of studies involving simple dynamics and photochemistry has been undertaken. In this paper the effects of steady, dissipating planetary waves on the ozone distribution will be discussed. The problem will be formulated in both the Eulerian and Lagrangian-mean systems and the relation of the Lagrangian-mean to the residual circulation will be investigated. In a subsequent paper the effects of time dependent planetary wave-mean flow fields on the ozone distribution will be examined.

In Section 2 the dynamical model is described. In Section 3 the continuity equation is developed for small amplitude waves in both the Eulerian and Lagrangian-mean formulations. Also the relation between the residual circulation and the Lagrangian-mean velocity is derived. In Section 4 model results are presented and a Lagrangian model based on particle trajectories is given.

2. Dynamic Model

The planetary waves are modeled by solving the steady, linear quasigeostrophic potential vorticity equation on a midlatitude β plane. The governing equation is

$$(a + \bar{u} \frac{\partial}{\partial x}) \{ \nabla^2 \phi' + \frac{f^2}{\rho N^2} \frac{\partial}{\partial z} \rho \phi'_z \} + \bar{q}_y \frac{\partial \phi'}{\partial x} = 0 \quad (1)$$

It is assumed that the mean flow is steady and governed by

$$a \left\{ \frac{\partial^2 \bar{\phi}}{\partial y^2} + \frac{f^2}{\rho N^2} \frac{\partial}{\partial z} \rho \bar{\phi}_z \right\} = f \frac{\partial^2}{\partial y^2} \overline{u'v'} - \frac{f^2}{\rho N^2} \frac{\partial}{\partial z} \rho \frac{\partial}{\partial y} \overline{v' \phi'_z} \quad (2)$$

The mean meridional circulation ($\bar{v}$ and $\bar{w}$) is forced by the wave field.

Nomenclature is given in the Appendix. It is assumed that wave effects on the mean flow are not so large that the wave field needs to be recomputed with the perturbed mean flow. This restricts the validity of the model to small amplitude waves and is consistent with the transport approximations to be made later. The model is essentially the Eulerian analogue to the Lagrangian-mean model of Schoeberl (1981).

Damping (a in Eqs. (1) and (2)) is modeled by Newtonian cooling and Rayleigh friction which are assumed to be equal. The damping is a slowly varying function of height except in the uppermost regions where it is assumed to increase at an exponential rate. This region of increase provides a sponge layer to reduce reflections from the upper boundary. The time scale of the damping below the sponge layer is approximately 10 days, a number representative of the middle or upper stratosphere (Schoeberl and Strobel, 1978; Dickinson, 1973).

3. Constituent dynamics

The chemistry used in this study is extremely simple, namely:

$$Q = -\lambda(\mu - \bar{\mu}_0) \quad (\lambda > 0) \quad (3)$$

where $\bar{\mu}_0$ is the zonal mean basic state. This formulation requires that the perturbations relax back to zero, while changes in the mean field relax to the mean equilibrium. The rate of relaxation is inversely proportional to λ , which is assumed to be a function of height only. The λ values are based on those of Hartmann and Garcia (1979).

The chemical reactions that the λ values are based on are the oxygen only equations with a modified ozone destruction rate constant to represent catalytic destruction. The chemistry is temperature dependent; however, the temperature dependence is small and will be neglected. Hartmann and Garcia (1979) and Garcia and Hartmann (1980) found the temperature dependent effects on ozone transport to be small.

In order to discuss the transport of ozone by stationary waves the atmosphere is divided into three regions as shown in Fig. 1. The lower region is the region in which ozone is assumed to be a conservative tracer. In the upper region ozone photochemistry is so fast (λ large) that the transport is unimportant. The middle region is defined by the ratio $.1 < \lambda/\bar{u}k < 10$ and is called the transition region (Hartmann and Garcia, 1979).

Figure 2 shows the ozone mixing ratio used in this model which is based on the ozone distribution used by Hartmann and Garcia (1979). The use of Hartmann and Garcia ozone distribution and chemistry facilitates comparison with their results.

The ozone continuity equation is formulated in both the Eulerian and Lagrangian-mean systems. In general, the constituent continuity equation is

$$\frac{D\mu}{Dt} = Q \quad (4)$$

where Q represents the source term. By decomposing the fields into zonal averages and deviations and linearizing, the Eulerian formulation is obtained:

$$\frac{\partial \bar{\mu}}{\partial t} + \bar{u}_i \frac{\partial \bar{\mu}}{\partial x_i} + \overline{u_i' \frac{\partial \mu'}{\partial x_i}} = \bar{Q} \quad (5)$$

and

$$\frac{\partial \mu'}{\partial t} + u_i' \frac{\partial \bar{\mu}}{\partial x_i} + \bar{u}_i \frac{\partial \mu'}{\partial x_i} = Q' \quad (6)$$

where repeated indices indicate summation. It is pertinent to note that the zonal mean equation contains eddy flux terms which represent the effects of wave transport. The tendency for the eddy transport to be counteracted by the advection by the mean flow has been discussed extensively, and represents a major difficulty of atmospheric transport problems (Dunkerton, 1980; Mahlman and Moxim, 1978).

Using the Lagrangian-mean formulation (Andrews and McIntyre, 1978) it is possible to redefine the averaging operator such that wave forcing does not appear in the mean equation. Using this concept the mean continuity equation is written as

$$\left(\frac{\partial}{\partial t} + u_i^L \frac{\partial}{\partial x_i} \right) u^L = Q^L \quad (7)$$

In this system the time rate of change of the Lagrangian-mean ozone distribution is equal to the advection by the Lagrangian-mean velocity plus a source term. Two features of the Lagrangian-mean velocity as defined by Andrews and McIntyre are:

- 1) For the premises of the non-interaction theorem (Boyd, 1976; Dunkerton, 1980) the Lagrangian-mean velocity is zero.
- 2) The Lagrangian-mean velocity is the center of mass velocity of a particular ensemble of fluid particles.

With the second feature above, it may seem natural to study tracer transport in the Lagrangian-mean system. However, it turns out the Lagrangian-mean transform is complicated and is easily formulated only for

the most simple flows (McIntyre, 1980). Furthermore, dispersion of the particles about their center of mass in nonlinear and time dependent problems may render the location of the center of mass of little use in the actual location of the tracer (Hsu, 1980).

The difference between the Lagrangian-mean and the Eulerian zonal mean, for an arbitrary quantity T , is defined as the Stokes correction

$$T^S \equiv T^L - \bar{T}$$

Relating the Lagrangian-mean transform to the Eulerian zonal mean by the premise that the unperturbed fluid would be described by the Eulerian flow, it is possible to define the Stokes correction to second order as

$$T^S = \xi_i \frac{\partial \bar{T}}{\partial x_i} + \frac{1}{2} \xi_i \xi_j \frac{\partial^2 \bar{T}}{\partial x_i \partial x_j}$$

where $\xi_i = (\xi_1, \xi_2, \xi_3) = (\xi, \eta, \zeta)$ are the displacements from the equilibrium position of the fluid.

The displacement fields are related to the Eulerian perturbation velocity fields by the following equation (Matsuno and Nakamura, 1979):

$$\left(\frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} \right) \xi_i = u'_i + \epsilon_j \frac{\partial}{\partial x_j} \bar{u}_i + \dots \quad (8)$$

The mean meridional circulation is forced by the waves and is assumed to be second order in wave amplitude ($O(a^2)$). Therefore, the changes in the mean ozone field are $O(a^2)$ and the terms involving $\bar{v}$, $\bar{w}$ in (6) and (8) can be ignored. The perturbation continuity equation can now be written to $O(a^2)$ as

$$\frac{\partial \mu'}{\partial t} + \bar{u} \frac{\partial \mu'}{\partial x} + \bar{v} \frac{\partial \bar{\mu}_0}{\partial y} + \bar{w} \frac{\partial \bar{\mu}_0}{\partial z} = -\lambda \mu' \quad (9)$$

Using the mass continuity equation, (5) may be written as

$$\frac{\partial \bar{\mu}}{\partial t} + \frac{\partial}{\partial y} (\bar{v} \bar{\mu}_0 + \overline{v' \mu'}) + \frac{1}{\rho} \frac{\partial}{\partial z} \rho (\bar{w} \bar{\mu}_0 + \overline{w' \mu'}) = -\lambda \bar{\mu} \quad (10)$$

where $\bar{\mu}$ represents changes in the mean field from the equilibrium value $\bar{\mu}_0$.

The relationship between the perturbation velocity fields and the displacement fields can be rewritten as

$$iku\xi = u' + \eta' \frac{\partial \bar{u}}{\partial y} + \zeta' \frac{\partial \bar{u}}{\partial z}$$

$$iku\eta = v', \quad iku\zeta = w'$$

These are the equations used to calculate the Lagrangian-mean displacements.

Under the assumptions above the following relationships between the Eulerian and the Lagrangian-mean quantities can be written for steady waves:

$$\frac{\partial \bar{\mu}^L}{\partial t} = \frac{\partial \bar{\mu}}{\partial t} + O(a^4) \quad (11a)$$

$$v^S = \xi_1 \frac{\partial v'}{\partial x_1} + O(a^4) \quad (11b)$$

$$w^S = \xi_1 \frac{\partial w'}{\partial x_1} + O(a^4) \quad (11c)$$

$$\bar{Q}^L = \bar{Q} + \xi_1 \frac{\partial Q'}{\partial x_1} \quad (11d)$$

Therefore (7) can be written to the same accuracy as (10) as

$$\frac{\partial \bar{\mu}}{\partial t} = -u_1^L \frac{\partial \bar{\mu}_0}{\partial x_1} - \xi_1 \frac{\partial \lambda \mu'}{\partial x_1} - \lambda \bar{\mu} \quad (12)$$

The first term on the right hand side of (12) will be referred to as the advective term; the second term is the "stirring" term; and the third term is the same as the Eulerian mean chemistry. The stirring term represents the effects of the planetary waves moving ozone parcels through a region of

varying photochemistry. This term will be discussed in detail later. Equations (12), (10), and (9) are the ozone continuity equations used in this study.

A number of authors have suggested that the Lagrangian-mean velocity may be a suitable quantity to approximate the large scale transport of ozone in 2-D chemical models (Dunkerton, 1978; Miller et al., 1981; McIntyre, 1980). However, since the Lagrangian-mean velocity can be difficult to calculate, it has also been suggested that the residual circulation defined by Andrews and McIntyre (1976) might be an adequate approximation for the Lagrangian-mean velocity (Holton, 1981). Strobel (1981) has derived a general linearized 2-D transport model using the residual circulation for waves with complex frequencies. The relation between the residual velocity and the Lagrangian-mean velocity for steady, dissipating planetary wave is outlined below.

The eddy mass continuity equation can be written in terms of the parcel displacement fields as

$$\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} + \frac{1}{\rho} \frac{\partial \rho \zeta}{\partial z} = 0 \quad (13)$$

Using (13) and the definition of the Stokes corrections it is possible to write

$$\mathbf{v}^L = \bar{\mathbf{v}} + \frac{1}{\rho} \frac{\partial}{\partial z} \overline{\rho \mathbf{v} \zeta} \quad (14)$$

and

$$\mathbf{w}^L = \bar{\mathbf{w}} - \frac{\partial}{\partial y} \overline{\mathbf{v} \zeta} \quad (15)$$

The residual circulation for the current problem is

$$\mathbf{v}^* = \bar{\mathbf{v}} - \frac{1}{\rho N^2} \frac{\partial}{\partial z} \rho \overline{\mathbf{v} \phi_z} \quad (16)$$

and

$$w^* = \bar{w} + \frac{1}{N^2} \frac{\partial}{\partial y} \overline{v' \phi'_z} \quad (17)$$

Using the eddy thermodynamic equation

$$(a + \bar{u}ik) \phi'_z - f v' \frac{\partial \bar{u}}{\partial z} + \bar{u}ik N^2 \zeta = 0$$

and multiplying by v' and averaging it can be shown that

$$\overline{v' \phi'_z} = \frac{-(\bar{k}\bar{u})^2 N^2 \overline{v' \zeta} + a \left[f \frac{\partial \bar{u}}{\partial z} \overline{v' v'} - N^2 \overline{v' w'} \right]}{a^2 + (\bar{k}\bar{u})^2} \quad (18)$$

where it has been assumed that only the correlations of real parts are considered. Using (18) and comparing (14) and (15) with (16) and (17) it is clear that the Lagrangian-mean velocity and the residual velocity are equal only when $a = 0$. When $a = 0$, however, both the Lagrangian-mean and the residual circulation are zero and the transport by the waves and the mean flow exactly counterbalance each other. Since a is an order of magnitude or more smaller than $\bar{k}\bar{u}$ below the sponge layer in this model, the Lagrangian-mean and residual circulations should be similar. When transience or external heating and cooling processes are allowed the relationship between the two circulations becomes more complex. In Holton's (1981) parameterization the difference between the Lagrangian-mean and residual circulations is contained in the antisymmetric part of the diffusion tensor and is assumed to be negligible.

In the next section the Eulerian velocities, the Lagrangian-mean quantities, and the residual circulation will be presented for steady, dissipating planetary waves. Zonal mean ozone changes can then be calculated and the various transport methods can be compared.

4. Results

a. Dynamic quantities

Equations (1) and (2) are solved for a wave number one perturbation with forcing analogous to a mean zonal wind of 10 m/sec flowing over a 1 km mountain. The mean zonal wind profile and the geopotential amplitude are shown in Fig. 3. The mean zonal wind is presumed to approximate the polar night jet. The geopotential reaches a maximum amplitude of about 1 km in the mesosphere and does not exceed the steady state limit (Schoeberl, 1982). The magnitude of the geopotential is in rough agreement with observation (van Loon et al., 1973). The eddy geopotential field is used to calculate the wave forced mean meridional and vertical velocity.

The $\bar{v}$ and $\bar{w}$ fields are shown in Fig. 4. These fields, as expected, show rising motion in the northern regions of the β channel and sinking motion in the southern regions. The meridional velocity is from north to south. This flow is forced by the northward eddy heat transport of the planetary waves.

The Stokes drifts as calculated with (11b) and (11c) are shown in Fig. 5. These drifts can be viewed as the effective velocity fields associated with the wave. The Stokes drifts are directed in an opposite sense to the Eulerian mean velocity and are of approximately the same magnitude. This demonstrates the tendency for the wave fields to counterbalance the Eulerian mean fields as has been discussed by Matsuno (1980).

The Lagrangian-mean velocity fields, which are the sums of the Eulerian zonal mean and the Stokes drifts, are shown in Fig. 6. The sense of the circulation is rising motion in the southern regions and sinking motion in the polar regions. This motion is the same as indicated by Schoeberl (1981). Furthermore, this motion would tend to enhance the diabatic circulation as calculated by Dunkerton (1978). The Lagrangian-mean velocity is of much smaller magnitude than either the Stokes drifts or the Eulerian mean velocities.

The residual circulation is shown in Fig. 7. It is geometrically similar to the Lagrangian-mean velocity but the meridional component is approximately 30% larger than the Lagrangian-mean meridional velocity. The

largest differences between the two circulations occur in the region where $\partial \bar{u} / \partial z$ is largest (see Eq. 18). Similar increases are calculated in the vertical residual velocity field.

b. Eulerian transport

The continuity equations (9), (10) and (12) are solved using the appropriate dynamic quantities. Since the mean chemistry appears in the same form in both (10) and (12) it is ignored in order that it might not obscure the differences between the two formulations. The effects of the mean chemistry term will be considered separately.

The horizontal ($\bar{v}\bar{\mu} + \overline{v'\mu'}$) and vertical ($\bar{w}\bar{\mu} + \overline{w'\mu'}$) fluxes as calculated by the Eulerian method are shown in Fig. 8. The vertical flux is upward in the northern region of the model and downward in the southern region. The vertical flux is confined primarily to the conservative region and reflects the Eulerian mean vertical velocity field.

The horizontal flux field is more complicated showing two regions of northward transport separated by a region of southward transport. Once again the flux is largest in the conservative region. The lower region of northward flux is dominated by the eddy flux which is proportional to the vertical derivative of the mean constituent distribution as described by Clark and Rogers (1978). The region of southward flux, which occurs where the vertical constituent density is small, is dominated by the mean meridional velocity. The upper region of northward transport is once again dominated by the eddy flux. In the upper most regions of the model both the horizontal and vertical fluxes are very small due to rapid photochemical relaxation.

The rate of change of the mean ozone field is calculated as the divergence of the fluxes. This field is shown in Fig. 9. There is an increase of ozone in the northern section and a decrease in the southern section. Despite the fluxes being largest in the conservative region the mean field tendency is concentrated almost entirely in the transition region in agreement with the results of Hartmann and Garcia (1979).

c. Lagrangian-mean transport

The Lagrangian-mean results are shown in Figures 10-12. The advective term as defined in Eq. (12) is shown in Fig. 10. In the extreme polar region there is a decrease of ozone above the ozone maximum and a build up of ozone in the lower stratosphere. In the southern regions there is a decrease in the lower stratosphere and an increase in the upper stratosphere. These changes are primarily due to the vertical advection. There is a strong transport of ozone northward by the meridional Lagrangian-mean velocity in the midlatitudes. The Lagrangian-mean velocity associated with planetary waves advects ozone poleward and downward.

The advective term calculated using the residual circulation shows increases and decreases in the mean ozone field in the same regions as the Lagrangian-mean calculations. In all regions the magnitude of advection by the residual circulation is larger. Differences as large as 30% are calculated reflecting the increased magnitude of the residual circulation.

Comparison of Figures 11 and 9 show that the mean field changes are the same when calculated by the Eulerian or Lagrangian-mean methods. Since the advective term is much smaller than the stirring term, substituting the residual circulation for the Lagrangian-mean velocity shows only minute changes in the tendency field.

d. The "stirring" term

The magnitude of the stirring term is shown in Fig. 12. This field quite accurately represents the ozone time rate of change shown in Fig. 11. The advection by the Lagrangian-mean velocity is less than 1% of the stirring term in regions where ozone is photochemically active. The stirring term is identically zero when the photochemistry is zero.

In order to understand the physical meaning of the stirring term and its analogy to the Eulerian problem, it is helpful to compare the stirring term to the eddy flux. Assuming that the wave is steady and stationary (9) becomes

$$\mu' = - \frac{\lambda - i k \bar{u}}{\lambda^2 + (k \bar{u})^2} u_1' \frac{\partial \bar{\mu}_0}{\partial x_1} \quad (19)$$

The change in the zonal mean ozone field due to the divergence of the eddy flux (DEF) is

$$\begin{aligned} \text{DEF} &= -\frac{1}{\rho} \frac{\partial}{\partial x_j} \rho \overline{u_j \mu'} \\ &= \frac{1}{\rho} \frac{\partial}{\partial x_j} \rho \left\{ \frac{\lambda}{\lambda^2 + (\bar{k}\bar{u})^2} \overline{\text{Re}(u_j') \text{Re}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} + \frac{\bar{k}\bar{u}}{\lambda^2 + (\bar{k}\bar{u})^2} \overline{\text{Re}(u_j') \text{Im}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} \right\} \end{aligned} \quad (20)$$

The stirring term (ST) is

$$\begin{aligned} \text{ST} &= -\xi_j \frac{\partial \lambda \mu'}{\partial x_j} = -\frac{1}{\rho} \frac{\partial}{\partial x_j} \rho \lambda \xi_j \mu' \\ &= \frac{1}{\rho} \frac{\partial}{\partial x_j} \frac{\rho \lambda}{\bar{k}\bar{u}} \left\{ \frac{\lambda}{\lambda^2 + (\bar{k}\bar{u})^2} \overline{\text{Im}(u_j') \text{Re}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} + \frac{\bar{k}\bar{u}}{\lambda^2 + (\bar{k}\bar{u})^2} \overline{\text{Re}(u_j') \text{Re}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} \right\} \end{aligned} \quad (21)$$

ST is functionally very similar to DEF. In order to investigate the precise similarities it is useful to consider the limits of λ relative to the dynamic frequency $\bar{k}\bar{u}$.

For $\lambda/\bar{k}\bar{u} < 1$,

$$\frac{1}{\lambda^2 + (\bar{k}\bar{u})^2} \approx \frac{1}{(\bar{k}\bar{u})^2} \left(1 - \frac{\lambda^2}{(\bar{k}\bar{u})^2} \right) \quad (22)$$

Using (22), (20) and (21) can be rewritten to lowest order as:

$$\text{DEF} = \frac{1}{\rho} \frac{\partial}{\partial x_j} \rho \left\{ \frac{\lambda}{(\bar{k}\bar{u})^2} \overline{\text{Re}(u_j') \text{Re}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} + \frac{1}{(\bar{k}\bar{u})} \overline{\text{Re}(u_j') \text{Im}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} \right\} \quad (23)$$

and

$$\text{ST} = \frac{1}{\rho} \frac{\partial}{\partial x_j} \frac{\rho \lambda}{(\bar{k}\bar{u})^2} \overline{\text{Re}(u_j') \text{Re}(u_1')} \frac{\partial \bar{\mu}_o}{\partial x_1} \quad (24)$$

Thus for small λ the divergence of the eddy flux can be divided into a purely chemical and purely conservative part. The conservative part, the second term on the RHS of (23), is the conservative flux described by Clark and Rogers (1978). The chemical part of DEF is identical to ST; therefore, for small λ it can be explicitly shown that the stirring term represents the chemical contribution of the divergence of the eddy flux.

For large λ DEF approaches zero, while ST approaches some finite value. In fact, as λ becomes large ST loses its explicit chemistry dependence and becomes

$$ST = \bar{u}_1^s \frac{\partial \bar{\mu}_o}{\partial x_1} \quad (25)$$

At this limit both the Eulerian mean, (10), and Lagrangian-mean, (12), continuity equations become

$$\frac{\partial \bar{\mu}}{\partial t} = - \bar{u}_1 \frac{\partial \bar{\mu}_o}{\partial x_1} - \lambda \bar{\mu} \quad (26)$$

Eq. (26) indicates that, when the photochemistry is strong, the wave effects become diminishingly small so that the mean field changes are governed by the mean chemistry and the advection by the Eulerian-mean velocity.

e. Lagrangian model

While the analysis above reveals the relationship of the stirring term to the divergence of the eddy flux, the physical nature of the stirring term remains obscured in the mathematics. Hartmann (1981) showed how pure oscillatory motion of a fluid parcel could produce changes in the mean ozone density in a fluid tube. A similar Lagrangian approach can be used here to investigate the effects of the planetary wave and to further coalesce the Lagrangian-mean and Eulerian viewpoints.

The Lagrangian model is based on the particle trajectories as determined by the Lagrangian-mean displacement fields. Using (1) and assuming that $\bar{u}$ is constant then it can be shown that for harmonic disturbances the horizontal displacement field, η , and the vertical displacement field, ζ are given by

$$\eta = \eta_0 \cos(\alpha z + kx) \quad (27)$$

and

$$\zeta = \zeta_0 \cos(\alpha z + kx) + \zeta_1 \sin(\alpha z + kx) \quad (28)$$

where α is the vertical wavenumber. The coefficients, η_0 , ζ_0 , and ζ_1 , are functions of height, meridional distance, and the dynamic parameters. For a particular value of y and z , the projection of η and ζ on the y,z plane is an ellipse (Fig. 13). These displacement fields are similar to those calculated by Matsuno (1980) and Danielsen (1981).

The displacement fields provide a model of the parcel motion associated with a planetary wave. The planetary wave displaces a particular parcel from its equilibrium position and causes the parcel to move in a helical trajectory around the globe. The Lagrangian-mean velocity is the velocity at which the axis of the helix moves. For the small amplitude, steady state waves of this model the vertical extent of the ellipse is on the order of one km and the horizontal extent is on the order of hundreds of kilometers.

First consider a parcel in the absence of chemistry. The equilibrium mixing ratio of the parcel, μ_p , is $\bar{\mu}_0(0,0)$. If the parcel is displaced to (η, ζ) , then the perturbation mixing ratio μ' is equal to the parcel mixing ratio minus the background mixing ratio. If the displacements are small then

$$\mu' = -\eta \frac{\partial \bar{\mu}_0}{\partial y} - \zeta \frac{\partial \bar{\mu}_0}{\partial z}$$

Assuming that $\bar{\mu}_0$ is a function of z only and that it increases with height, then positive ozone perturbations are generally associated with northward motion and negative ozone perturbations are associated with southward motion (see Fig. 13). This is equivalent to northward eddy flux of ozone. Hence this is the Lagrangian description of the conservative flux as defined by Clark and Rogers (1978). This conservative flux is largely cancelled by the wave induced mean circulation.

In the case when ζ_1 in (28) is zero, which corresponds to a trapped (non-propagating) wave, the trajectories then become lines instead of

ellipses. The perturbations are then equally correlated with both northward and southward velocity and the conservative eddy flux is identically zero. An evanescent wave has the same property.

When chemistry is present μ' still changes due to the varying background field, but the chemistry tries to bring the parcel's ozone mixing ratio towards the background value. The constituent density for the parcel is determined by

$$\frac{d\mu'}{dt} = S - \lambda\mu' \quad (29)$$

where S represents the effective source caused by the parcel moving through the varying background density. For elliptical trajectories S is a harmonic function in time, proportional to the magnitude of the velocity and the background constituent gradients. In general, as the parcel orbits the ellipse it goes through varying photochemistry so that λ is also a function of time. In a reference frame moving with the zonal velocity of the parcel (29) can be rewritten as

$$\frac{d\mu'}{dt} = S_o \cos \omega t - \lambda(t)\mu' \quad (30)$$

where $\omega = k\bar{u}$.

If λ is constant then the solution to (30), for large t , can be written as

$$\mu' = \frac{S_o}{\lambda^2 + \omega^2} (\lambda \cos \omega t + \omega \sin \omega t) \quad (31)$$

For large λ , μ' approaches zero which is equivalent to the previous result that the eddy fluxes disappear in the fast photochemical region. For $\lambda = 0$ the solution reduces to the ozone density appropriate to the conservative problem. Note that the average of μ' over one period is zero.

In the regions where λ varies strongly over the extent of the ellipse the perturbation can take on a different character. Such regions would be expected to exist in the lower parts of the transition layer and at the polar night boundary where the chemistry is nearly discontinuous. Figure 14 shows an extreme example where the chemistry is discontinuous across an

ellipse. As the parcel travels through the conservative region the perturbation density varies as S . In the transition region ($\lambda/\bar{k}\bar{u} = 5$) the perturbation is decreased by $\sim 80\%$. Assuming the background density increases with height, this behavior means that the negative perturbations are selectively damped while the positive perturbations are not. Therefore, the average value of μ' over one period, which was zero when λ did not vary, may become nonzero in the presence of extreme chemical gradients. Such regions would be expected to have a great effect upon transport calculations. This effect is analogous to mean density changes in a fluid tube as calculated by Hartmann (1981).

Figure 15 demonstrates this effect for a more realistic problem. Figure 15a shows (31) the solution at large t to (30). This is a harmonic function with zero average. Figure 15b shows the numerical solution to (26) for an ellipse situated at the bottom of the transition layer with the model photochemistry. The steady state solution has a non-zero time average, as expected.

f. The effects of eddy transport on the mean field

Since the changes in the mean field have been assumed to be $O(a^2)$ the calculation of the steady state forcing by the eddies on the mean field tendency is independent of the mean constituent density. Therefore, at any point in space it is possible to write the mean field equation (10) or (12) as

$$\frac{\partial \bar{\mu}}{\partial t} = -\lambda \bar{\mu} + F \quad (32)$$

where F is the total forcing which represents the divergence of both the horizontal and vertical fluxes. If F is assumed independent of time, the solution to (32) can be written as

$$\bar{\mu} = \frac{F}{\lambda} (1 - e^{-\lambda t}) \quad (33)$$

where it has been assumed that $\bar{\mu}$, the change in the equilibrium field, is zero at time zero. As $t \rightarrow \infty$ the change in the mean field is F/λ .

In the conservative region the mean continuity equation can be written as $\frac{\partial \bar{\mu}}{\partial t} = F$ which yields

$$\bar{\mu} = Ft \quad (34)$$

In order to investigate the effect of the steady state waves on the mean field the appropriate equation (33) or (34) was solved at $t = 90$ days. This calculation would represent the effects of the waves if they persisted for the entire winter. The forcing is taken from the line in Fig. 9 through the region of maximum ozone increase in the northern region of the domain.

Figure 16 shows the change in $\bar{\mu}$. The large rates of change that are calculated in the transition region are almost completely counterbalanced by the mean chemistry. At the altitude of maximum increase in Fig. 9 the actual change in the mean field is only .5 ppm for 90 days.

The largest changes in the mean field (Fig. 16) are in the lower stratosphere where changes of 1.4 ppm are seen at 90 days. These changes are comparable to those observed in the northern hemisphere. Immediately below the transition region is a region of depletion caused by the advection. This vertical structure of the changes in the mean field is unrealistic and due largely to the simplified representation of photochemistry and dynamics in the lower stratosphere.

The relatively large positive changes in the mean field at the bottom of the transition layer are in the region where the photochemistry gradients are large. This is the region where a parcel takes on a non-zero average during a parcel orbit. The equilibrium values at this altitude are reached after 5 days of integration. This region of large increase just above a region of decrease demonstrates an important model sensitivity to the chemical structure.

Since the advection term is responsible for all of the change in the mean field below the transition region it is worthy of further investigation. Mean field tendencies for 90 day periods were calculated 400 km from both the northern and southern boundaries using both the

residual and Lagrangian-mean advection (Fig. 17). The large changes in the midlatitudes caused by meridional transport are largely negated by the photochemistry.

In the southern region the change in the mean field by advection is unlike that observed in the stratosphere. There is a depletion of nearly 4 ppm at 24 km and an enhancement of almost 2 ppm at 40 km. In the northern region the changes in the lower stratosphere are not unreasonable showing a 1.5 ppm increase at 24 km. However, at 36 km a 1 ppm depletion occurs. It is obvious that the modeling of the lower stratosphere is inadequate. Both small scale dynamics and photochemistry which may act to counter the tendencies shown here have been neglected. The advection calculated using the residual circulation produces larger changes in the mean field than that calculated using the Lagrangian-mean velocity. This means, that in this model, the antisymmetric part of Holton's (1981) diffusion tensor is not negligible.

5. Conclusions

The changes in zonal mean ozone as forced by steady state planetary waves have been calculated using both a Lagrangian-mean and Eulerian formulation. Although the calculated results from the two methods are the same, the use of the two methods allow for increased interpretation of the results.

In the Eulerian formulation the tendency for the eddy and mean fluxes to cancel is well illustrated. Both the horizontal and vertical flux fields are complex and largest in the conservative region. However, the changes in the mean field in the absence of the mean photochemistry are largest in the transition region as found by Hartmann and Garcia (1979).

The problem of compensation by two large terms which appears in the Eulerian framework also appears in the calculation of the Lagrangian-mean velocity. If the Lagrangian-mean flow could be calculated directly then it might be possible to avoid the numerical difficulties of compensation; however, the Lagrangian-mean quantities can be easily calculated only for simple flows. There seems to be no computational advantage of the Lagrangian-mean method to calculate planetary wave transport as it requires the formulation of Eulerian quantities, their conversion to Lagrangian-mean quantities, and then the calculation of the transport.

In the Lagrangian-mean formulation the stirring term is dominant in the absence of mean chemistry. The stirring is equal to the chemical contribution of the divergence of the eddy flux when $\lambda / \bar{u}k < 1$. When λ is large the stirring term is equivalent to the Stokes advection. Using particle displacement fields it is possible to predict the type of eddy transport that is to be expected in a certain region and it was found that μ' has a non-zero time average in the presence of strong chemical gradients. Calculation of the mean field changes shows that the mean chemistry is able to compensate for the eddy chemistry changes except where the chemical gradients are large. This indicates that the interaction of steady planetary waves and photochemistry need only be modeled in the presence of large chemical gradients.

For seasonal integrations the largest changes in the mean field are caused by advection by the Lagrangian-mean velocity in the conservative

region. The advection is so weak that in the chemical regions its effects are negligible. Using the residual circulation instead of the Lagrangian-mean circulation overestimates the advection by the planetary wave. The largest differences between the Lagrangian-mean and residual circulations occur in chemically active regions; hence, the difference between the ozone field changes calculated with the Lagrangian-mean and residual circulations are small. This would not necessarily be true for constituents with different chemistry.

The transport mechanisms investigated here are probably not adequate to model planetary wave effects in the stratosphere. The use of time independent forcing without changing the mean advective field is not strictly correct. Sudden warmings, which are highly transient, will certainly have a great effect upon the ozone distribution (Garcia and Hartmann, 1980). Therefore, steady state planetary wave transport models are probably only relevant in the late fall and early winter. The planetary wave transport would have to be supplemented by transport due to the diabatic circulation and small scale dynamics in the lower stratosphere. Transport by planetary waves during sudden warmings, transport by the diabatic circulation, and the relative importance of the two circulations will be presented in a subsequent paper.

Acknowledgements

This research was supported by the National Aeronautics and Space Administration and the Office of Naval Research. We would like to thank the two anonymous reviewers for many helpful comments.

References

- Andrews, D.G. and M.E. McIntyre, Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration, J. Atmos. Sci., 33, 2031-2048, 1976.
- Andrews, D.G. and M.E. McIntyre, An exact theory of nonlinear waves on a Lagrangian-mean flow, J. Fluid Mech., 89, part 4, 609-646, 1978.
- Boyd, J.P., The noninteraction of waves with the zonally averaged flow on a spherical earth and the interrelationships of eddy fluxes of energy, heat, and momentum, J. Atmos. Sci., 33, 2285-2291, 1976.
- Clark, J.H.E. and T.G. Rogers, The transport of conservative trace gases by planetary waves, J. Atmos. Sci., 35, 2232-2235, 1978.
- Danielsen, E.F., An objective method for determining the generalized transport tensor for two-dimensional Eulerian models, J. Atmos. Sci., 38, 1319-1339, 1981.
- Dickinson, R.E., Method of parameterization for infrared cooling between altitudes of 30 and 70 kilometers, J. Geophys. Res., 78, 4451-4457, 1973.
- Dunkerton, T., On the mean meridional mass motions of the stratosphere and mesosphere, J. Atmos. Sci., 35, 2325-2333, 1978.
- Dunkerton, T., A Lagrangian-mean theory of wave, mean-flow interaction with applications to nonacceleration and its breakdown. Rev. Geophys. Space Phys., 18, 387-400, 1980.
- Garcia, R.R. and D.L. Hartmann, The role of planetary waves in the maintenance of the zonally averaged ozone distribution of the upper atmosphere, J. Atmos. Sci., 37, 2248-2264, 1980.
- Hartmann, D.L., Some aspects of the coupling between radiation, chemistry, and dynamics in the stratosphere, J. Geophys. Res., 86, 9631-9640, 1981.
- Hartmann, D.L. and R. Garcia, A mechanistic model of ozone transport by planetary waves in the stratosphere, J. Atmos. Sci., 36, 350-364, 1979.
- Harwood, R.S. and J.A. Pyle, Studies of the ozone budget using a zonal mean circulation model and linearized photochemistry, Quart. J.R. Met. Soc., 103, 319-343, 1977.
- Holton, J.R., An advective model for two dimensional transport of stratospheric trace species, J. Geophys. Res., 86, 11989-11994, 1981.

- Hsu, C-P,F, Air parcel motions during a numerically simulated sudden stratospheric warming, J. Atmos. Sci., 37, 2768-2792, 1980.
- Hunten, D.M., Vertical transport in atmospheres, in Atmospheres of Earth and the Planets, B.M. McCormac, Ed., Dordrecht, Holland, 59-72, 1975.
- Mahlman, J.D. and W.M. Moxim, Tracer simulation using a global general circulation model: Results from a midlatitude instantaneous source experiment, J. Atmos. Sci., 35, 1340-1374, 1978.
- Matsuno, T., Lagrangian motion of air parcels in the stratosphere in the presence of planetary waves, Pure Appl. Geophys., 118, 189-216, 1980.
- Matsuno, T. and K. Nakamura, The Eulerian- and Lagrangian-mean meridional circulations in the stratosphere at the time of a sudden warming, J. Atmos. Sci., 36, 640-654, 1979.
- McIntyre, M.E., Towards a Lagrangian-mean description of stratospheric circulations and chemical transports, Phil. Trans. R. Soc. Lond., A296, 129-148, 1980.
- Miller, C., D.L. Filkin, A.J. Owens, M.J. Steed and J.P. Jesson, A two dimensional model of stratospheric chemistry and transport, J. Geophys. Res., 86, 12039-12065, 1981.
- NASA, The Stratosphere: Present and Future, NASA Reference Publication 1049, 1979.
- Pyle, J.A. and C.F. Rogers, Stratospheric transport by stationary planetary waves- The importance of chemical processes, Quart. J. Roy. Meteor. Soc., 106, 421-446, 1980.
- Schoeberl, M.R., A simple model of the Lagrangian-mean flow produced by dissipating planetary waves, J. Atmos. Sci., 38, 1841-1855, 1981.
- Schoeberl, M.R., Vacillation, sudden warmings and potential enstrophy balance in the stratosphere, in press, 1982.
- Schoeberl, M.R. and D.F. Strobel, The zonally averaged circulation of the middle atmosphere, J. Atmos. Sci., 35, 577-591, 1978.
- Strobel, D.F., Parameterization of linear wave chemical transport in planetary atmospheres by eddy diffusion, J. Geophys. Res., 86, 9806-9810, 1981.
- van Loon, H., R.L. Jenne and K. Labitzke, Zonal harmonic standing waves, J. Geophys. Res., 78, 4463-4471, 1973.

Appendix

NOMENCLATURE and CONVENTIONS

a	Newtonian cooling and Rayleigh friction coefficients
f	Coriolis parameter ($1 \times 10^{-4} \text{sec}^{-1}$)
H	scale height (7 km)
k	zonal wave number ($6.28 \times 10^{-9} \text{cm}^{-1}$)
N^2	Brunt-Väisälä frequency squared ($4 \times 10^{-4} \text{sec}^{-2}$)
p	pressure; p_0 reference pressure
$\bar{q}_y$	horizontal derivative of the mean potential vorticity
Q	photochemical source term
(u, v, w)	$(u_1, u_2, u_3) = (\text{zonal, meridional, vertical})$ velocity
(v^*, w^*)	$(\text{meridional, vertical})$ residual velocity (Andrews and McIntyre, 1976)
t	time
(x, y, z)	$(x_1, x_2, x_3) = (\text{east-west, north-south, } -H \ln \frac{p}{p_0})$ coordinates
β	meridional variation of Coriolis parameter, f
λ	photochemical time constant
μ	ozone mixing ratio

$(\xi, \eta, \zeta) = (\xi_1, \xi_2, \xi_3) =$ (longitudinal, meridional, vertical)
displacement fields

ϕ geopotential

CONVENTIONS:

For a given T

$\bar{T} \equiv$ zonal average around latitude circles (Eulerian)

$T' = T - \bar{T} \equiv$ zonal deviation

$T^L \equiv$ Lagrangian-mean (Andrews and McIntyre, 1978)

$T^S = T^L - \bar{T} \equiv$ Stokes correction (see Eq. (8))

Figures

- Figure 1. Ratio of dynamical to photochemical time scales, $\lambda/\bar{u}k$; showing the atmosphere divided into conservative, transitional, and photochemical regions.
- Figure 2. Zonal mean ozone distribution (ppm).
- Figure 3. a) Mean zonal wind, $\bar{u}$ (m/sec).
b) Amplitude of wavenumber 1 geopotential (m)
- Figure 4. Eulerian zonal mean velocities (cm/sec).
a) $\bar{v}$
b) $\bar{w}$
- Figure 5. Stokes drifts (cm/sec).
a) v^S
b) w^S
- Figure 6. Lagrangian-mean velocities (cm/sec).
a) v^L
b) w^L
- Figure 7. Residual velocities (cm/sec).
a) v^*
b) w^*
- Figure 8. Eulerian fluxes of ozone (ppm m/sec).
a) horizontal flux
b) vertical flux
- Figure 9. Mean ozone tendency, $\frac{\partial \bar{u}}{\partial t}$, as calculated with the Eulerian formulation (ppm/sec).

- Figure 10. Mean ozone tendency caused by advection (ppm/sec).
 a) Lagrangian-mean velocity
 b) residual velocity
- Figure 11. Total mean ozone tendency, $\frac{\partial \bar{\mu}}{\partial t}$, as calculated with the Lagrangian formulation (ppm/sec).
- Figure 12. Mean ozone tendency caused by "stirring" (ppm/sec).
- Figure 13. Projection of displacement fields η, ζ on the y, z plane (km).
 For $\frac{\partial \mu}{\partial z} > 0$ northward moving parcels are associated with positive perturbations.
- Figure 14. Parcel orbiting through a region of discontinuous photochemistry $\frac{\partial \mu}{\partial z} > 0$.
- Figure 15. Steady state solution of perturbation ozone density for an elliptically orbiting parcel.
 a) constant photochemistry
 b) highly varying photochemistry at the bottom of the transition layer.
- Figure 16. 90 day zonal mean ozone change, with the mean chemistry included, at the latitude of maximum ozone increase (ppm).
- Figure 17. 90 day zonal mean ozone change, with the mean chemistry included, at northern and southern extremes, as determined by Lagrangian-mean advection (LGM) and residual velocity advection (+) (ppm).
 a) Southern
 b) Northern

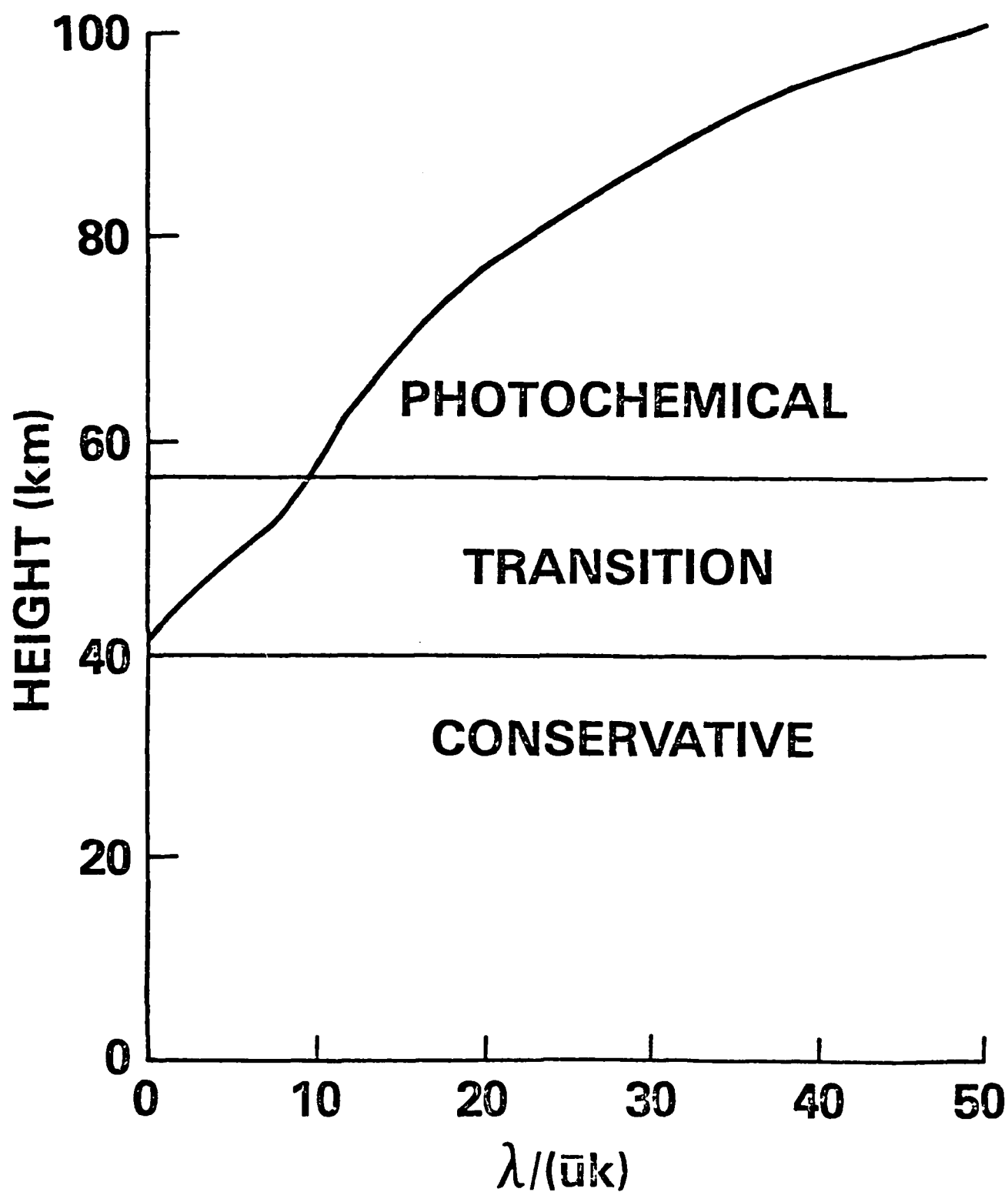
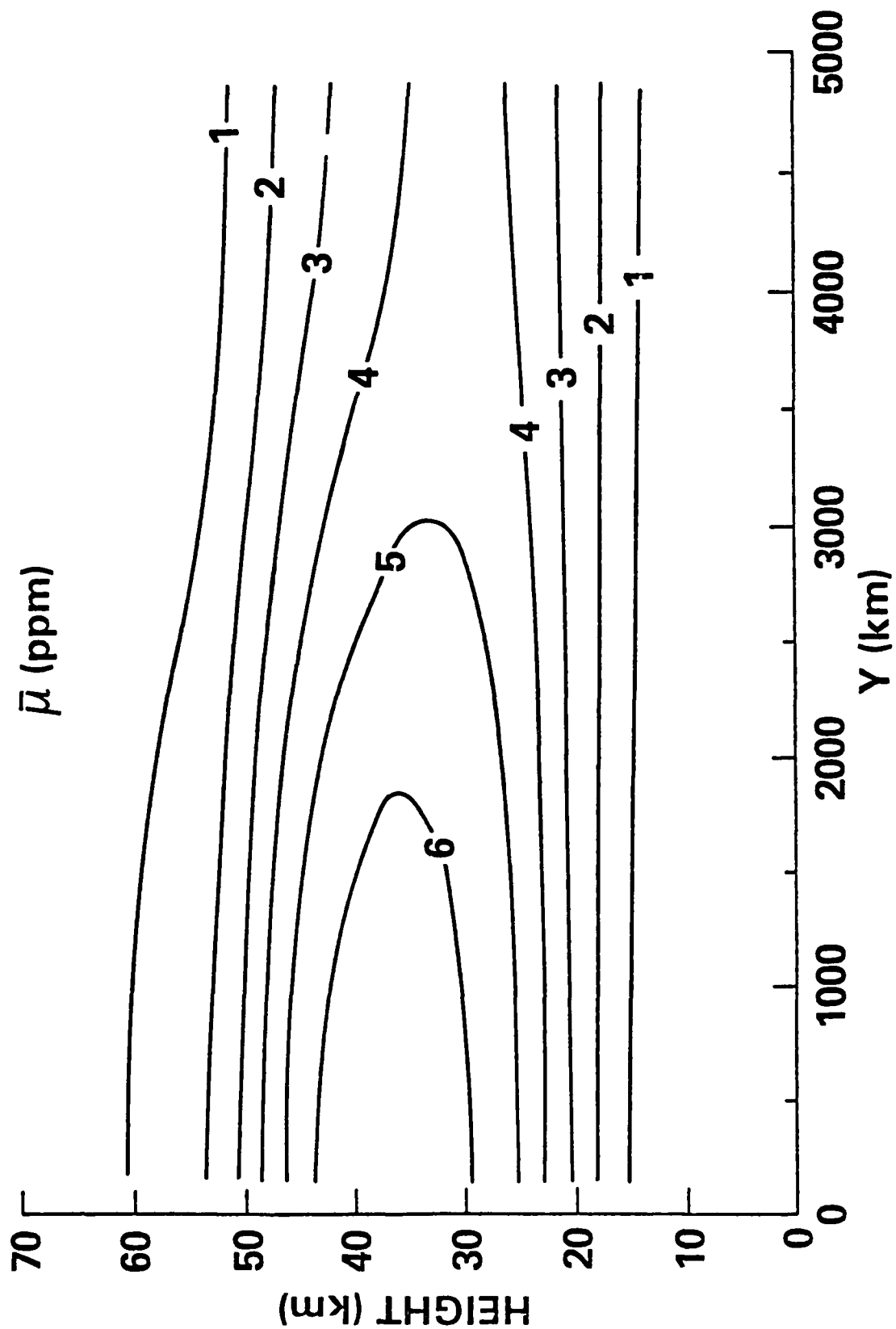


Figure 1



Q-32

Figure 2

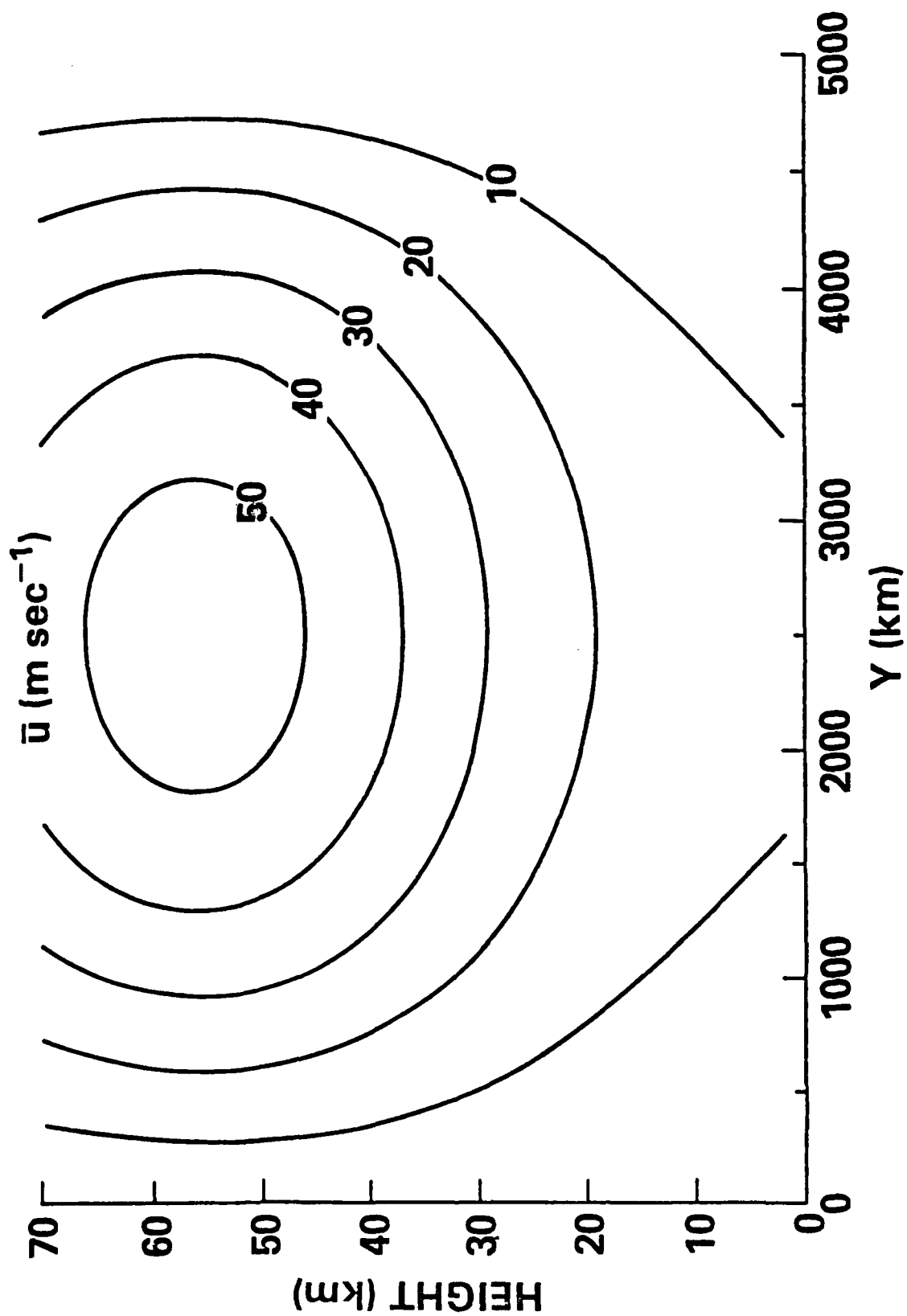


Figure 3a

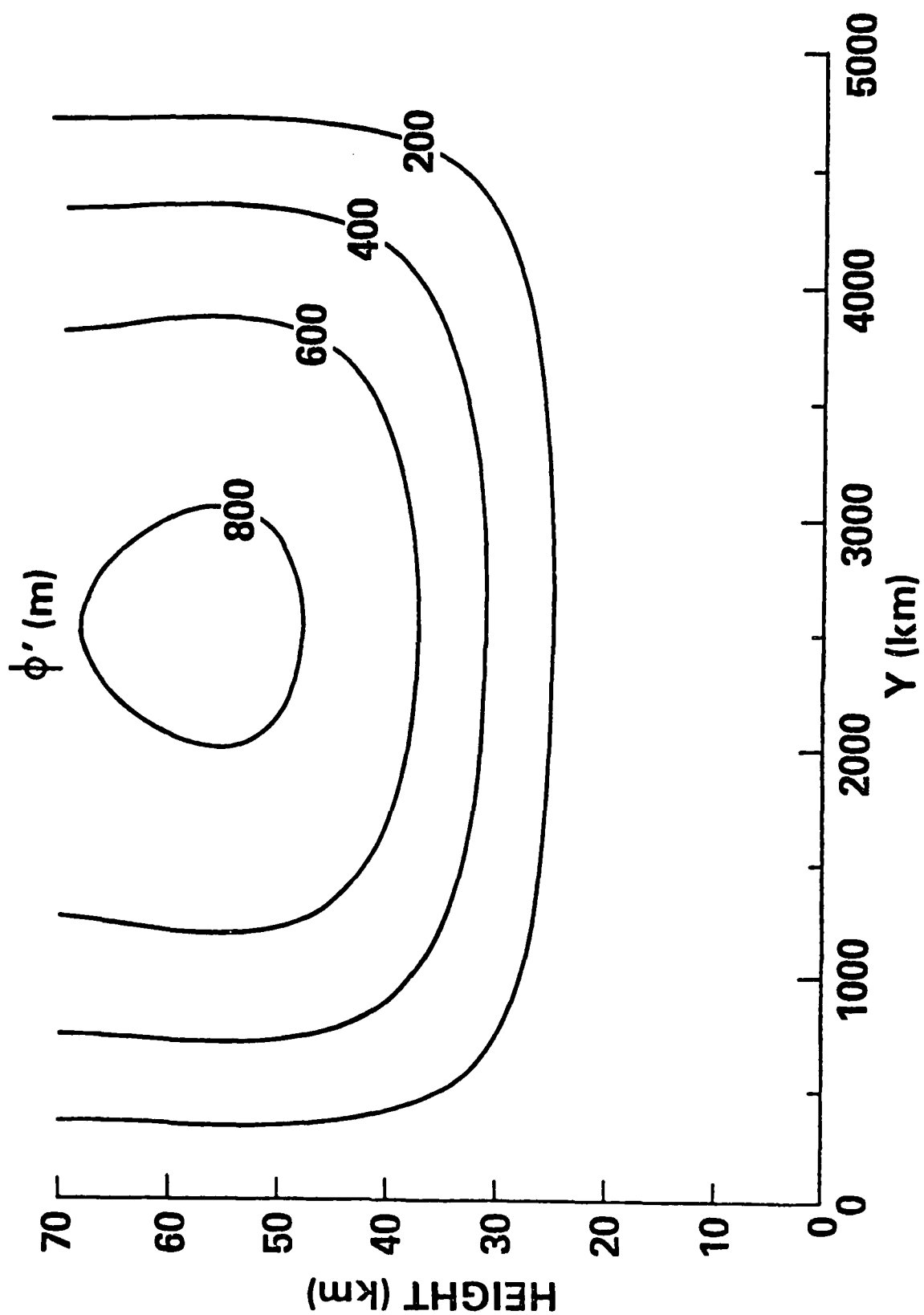


Figure 3b

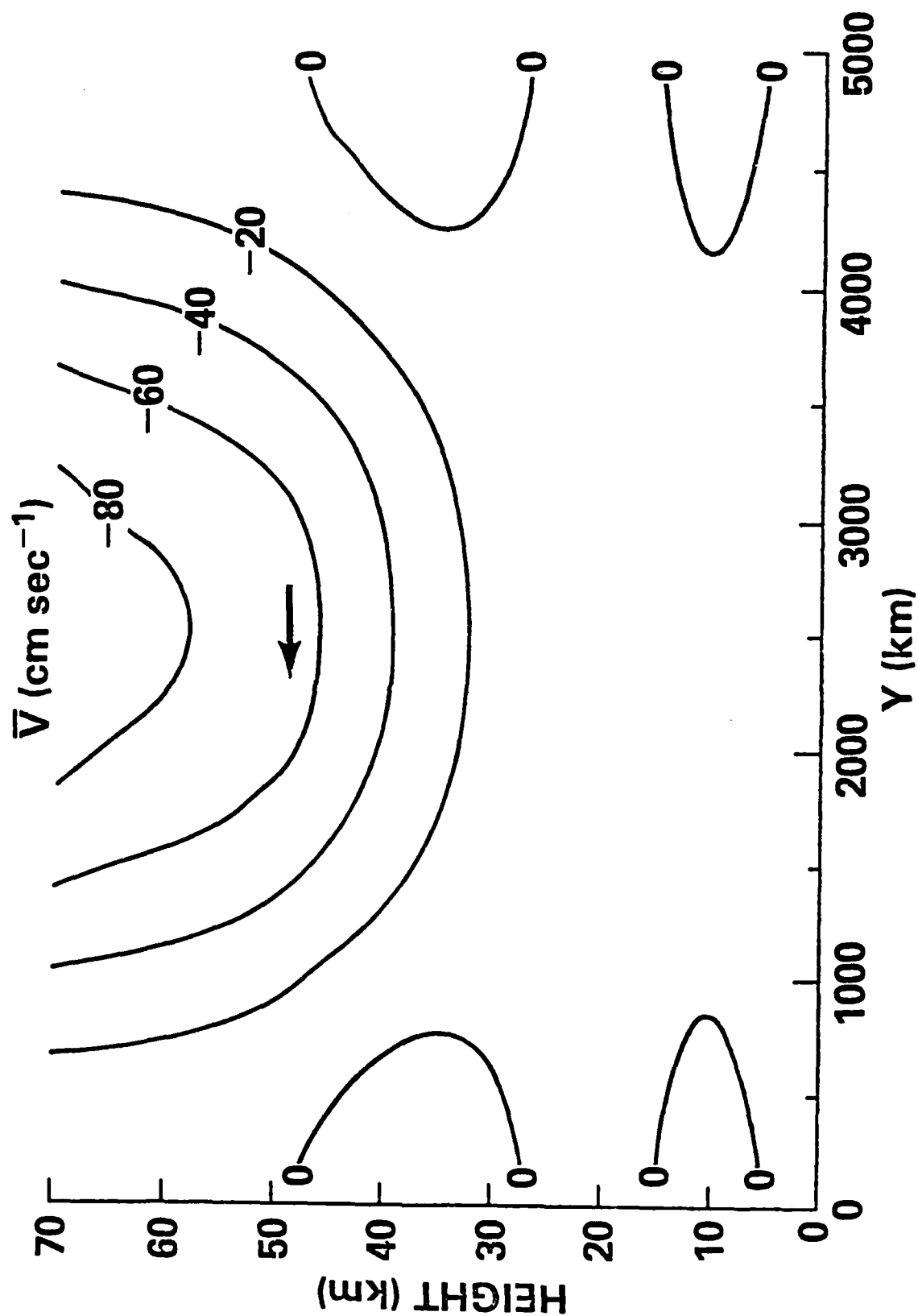


Figure 4a

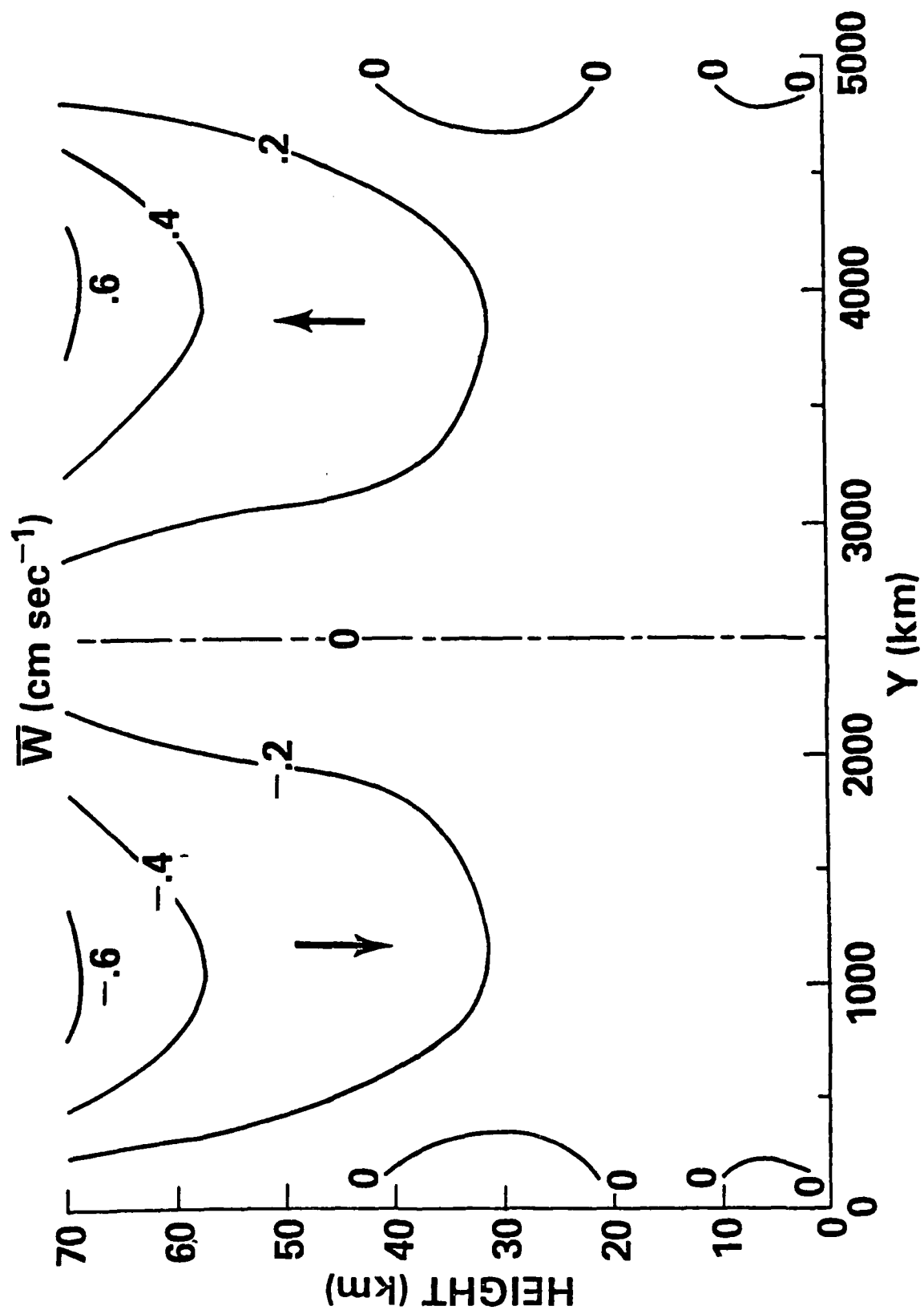


Figure 4b

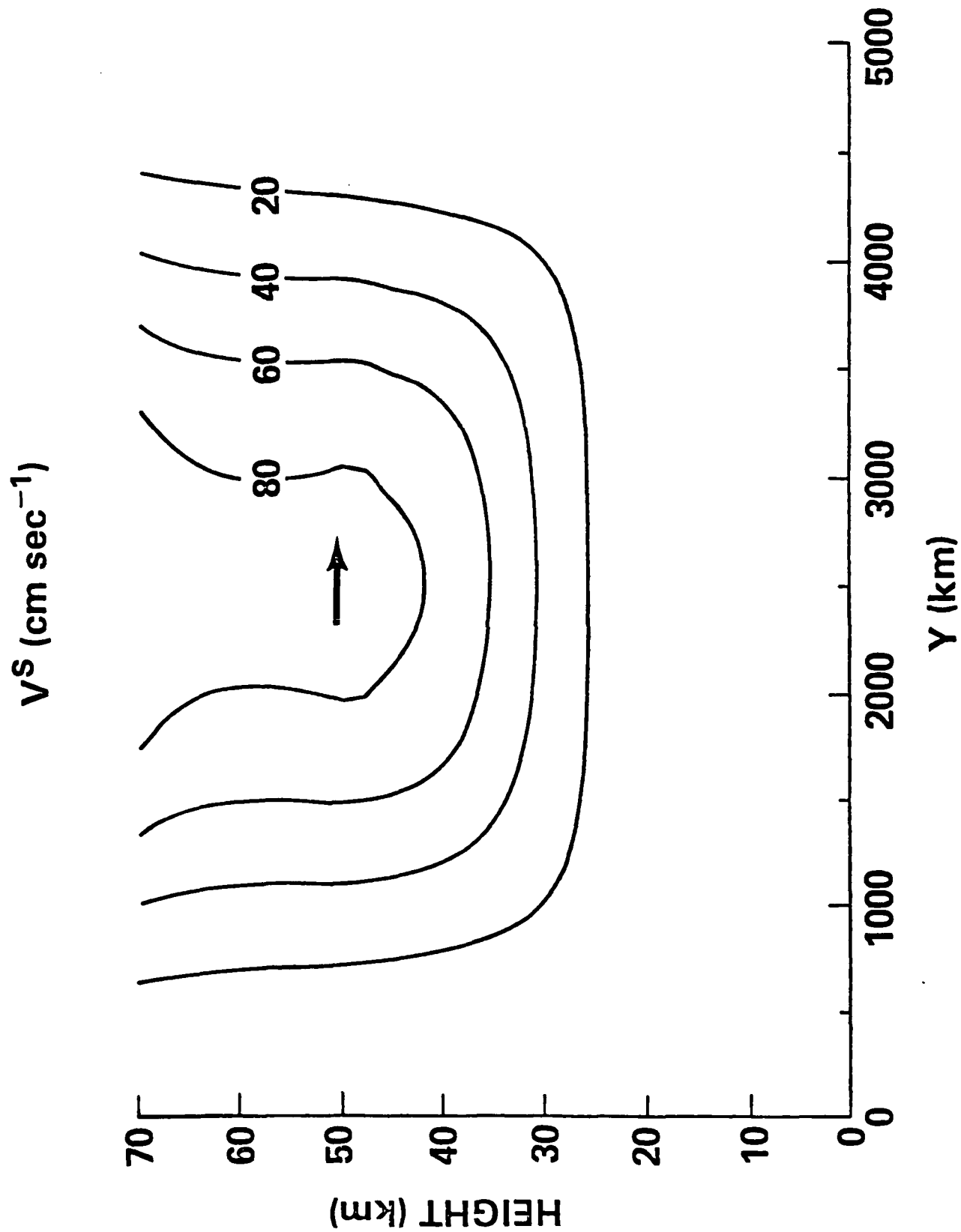


Figure 5a

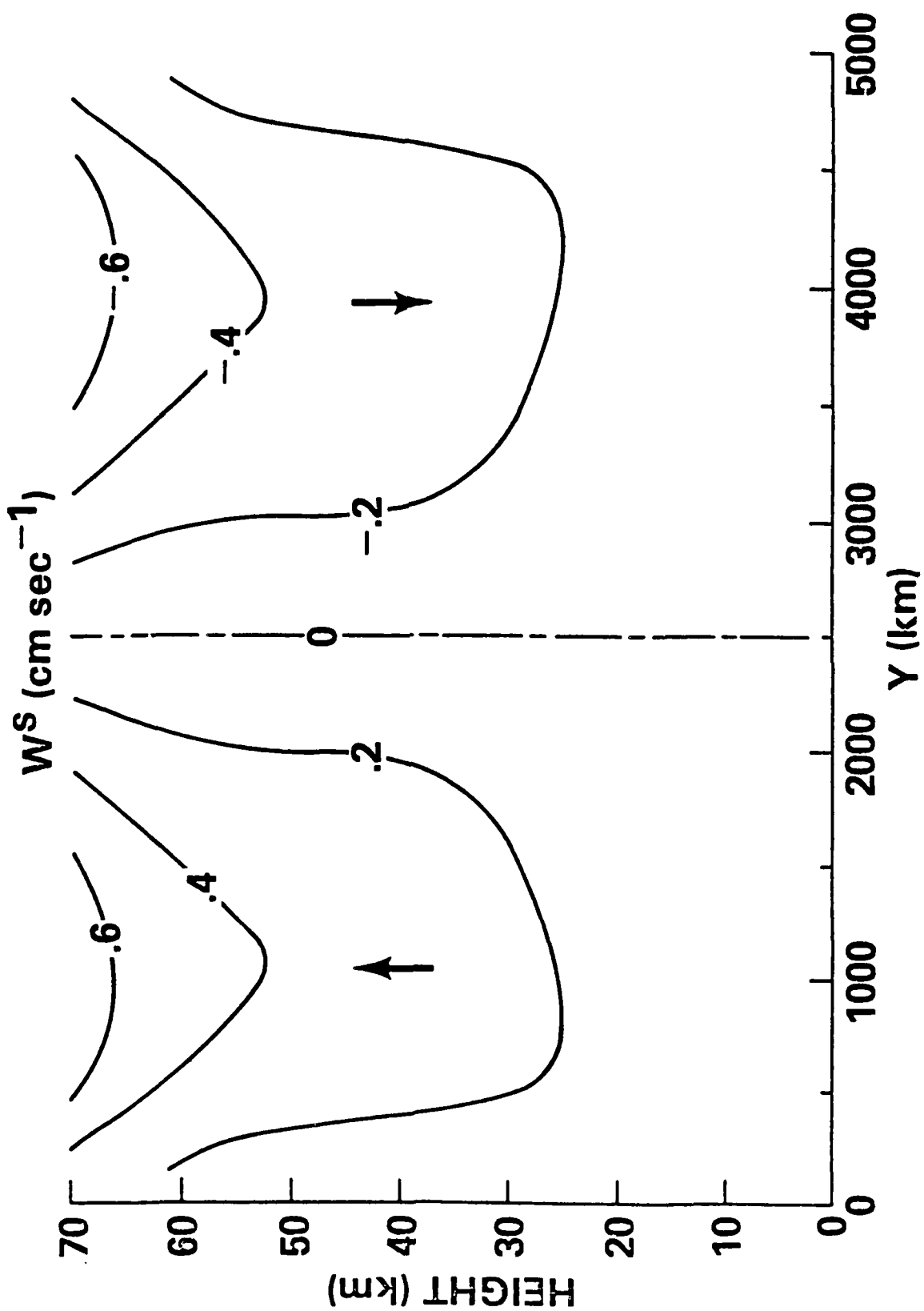


Figure 5b

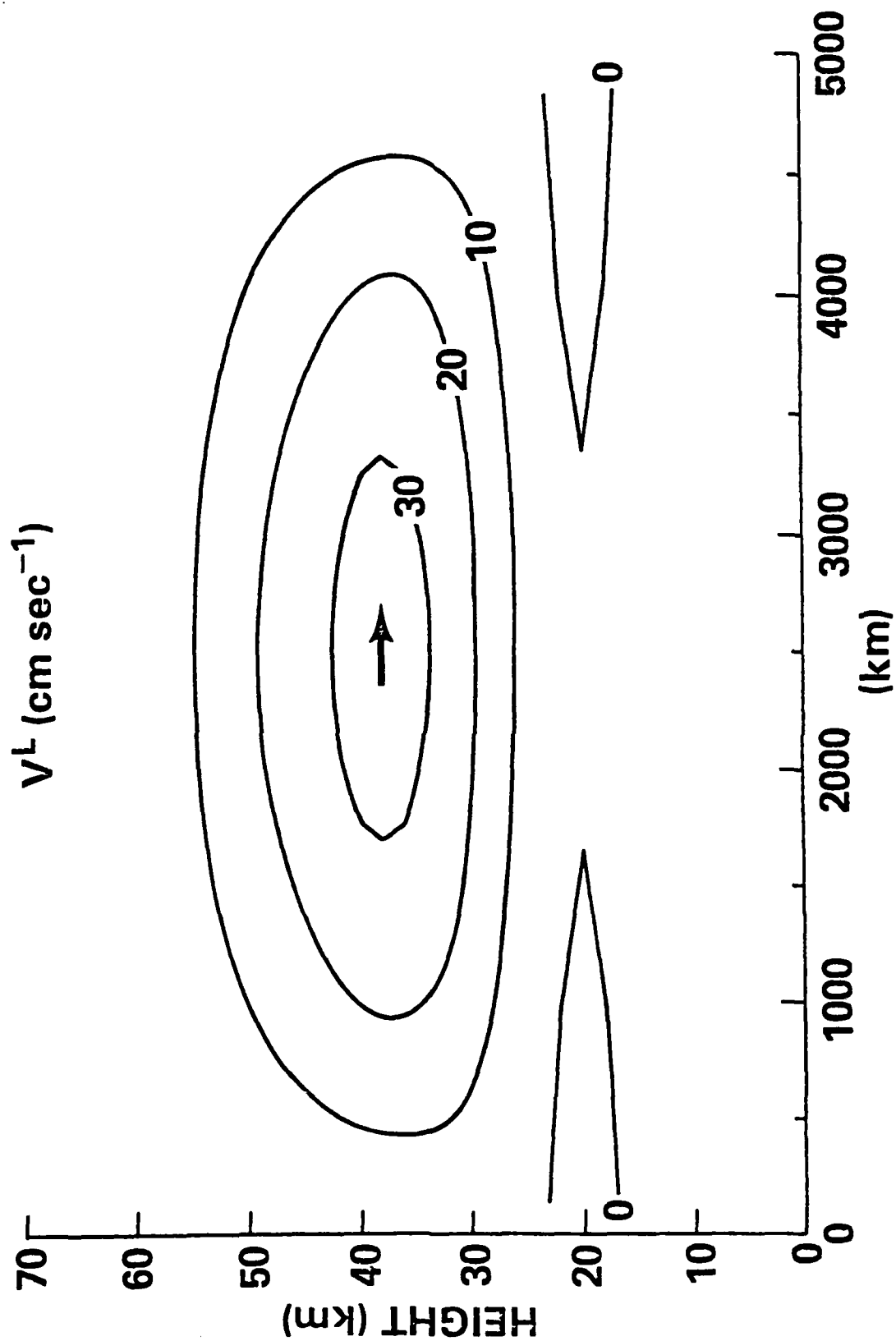


Figure 6a

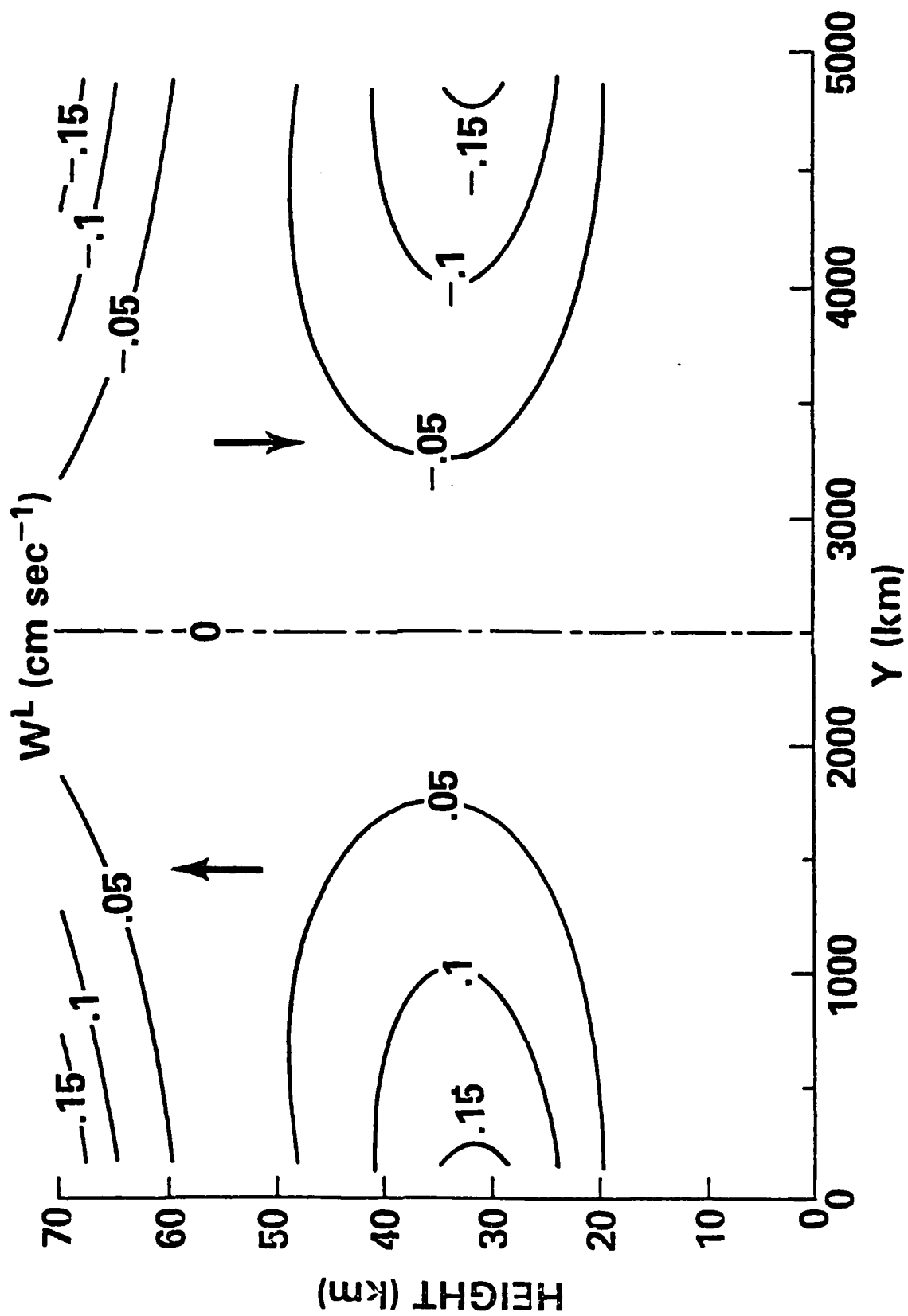


Figure 6b

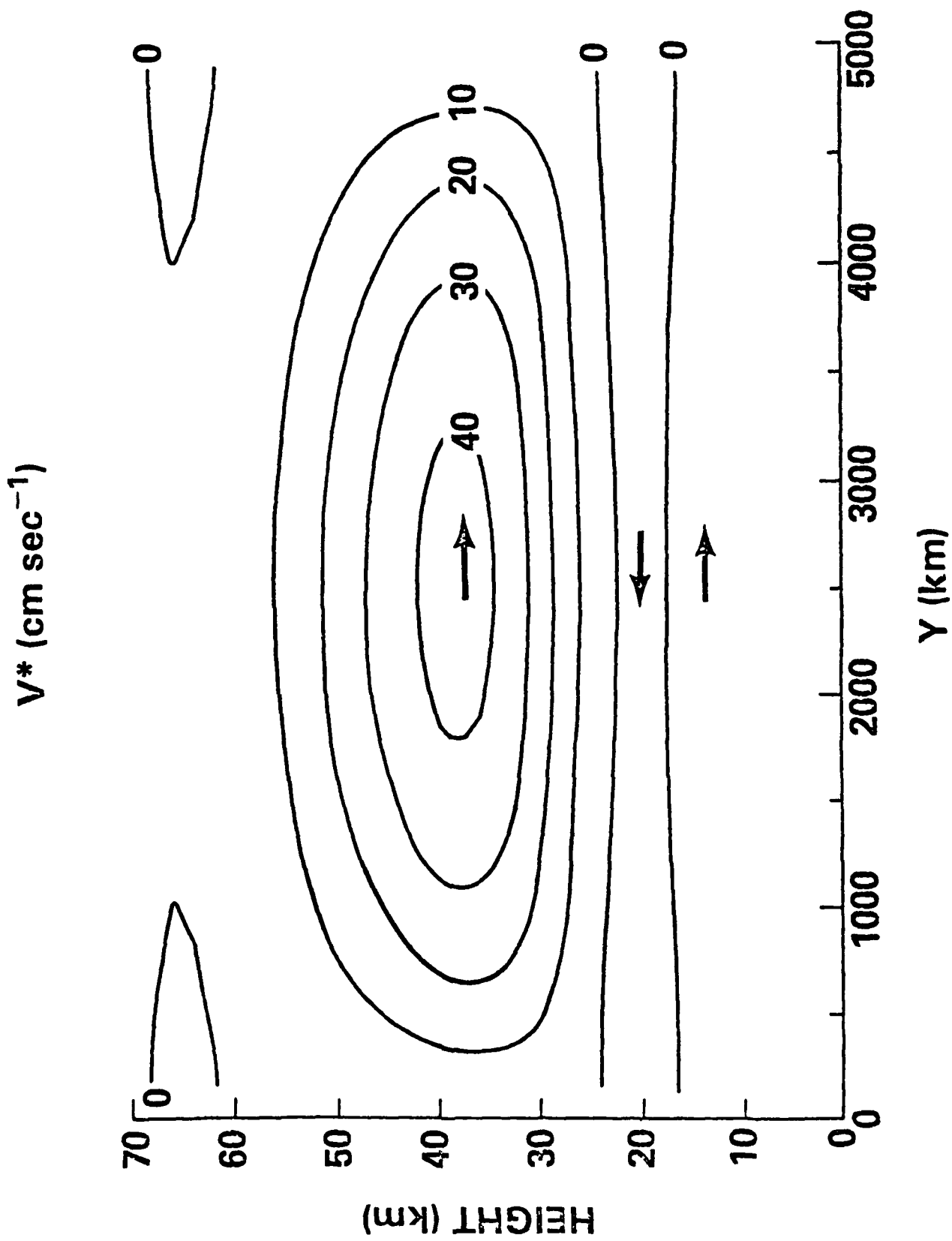


Figure 7a

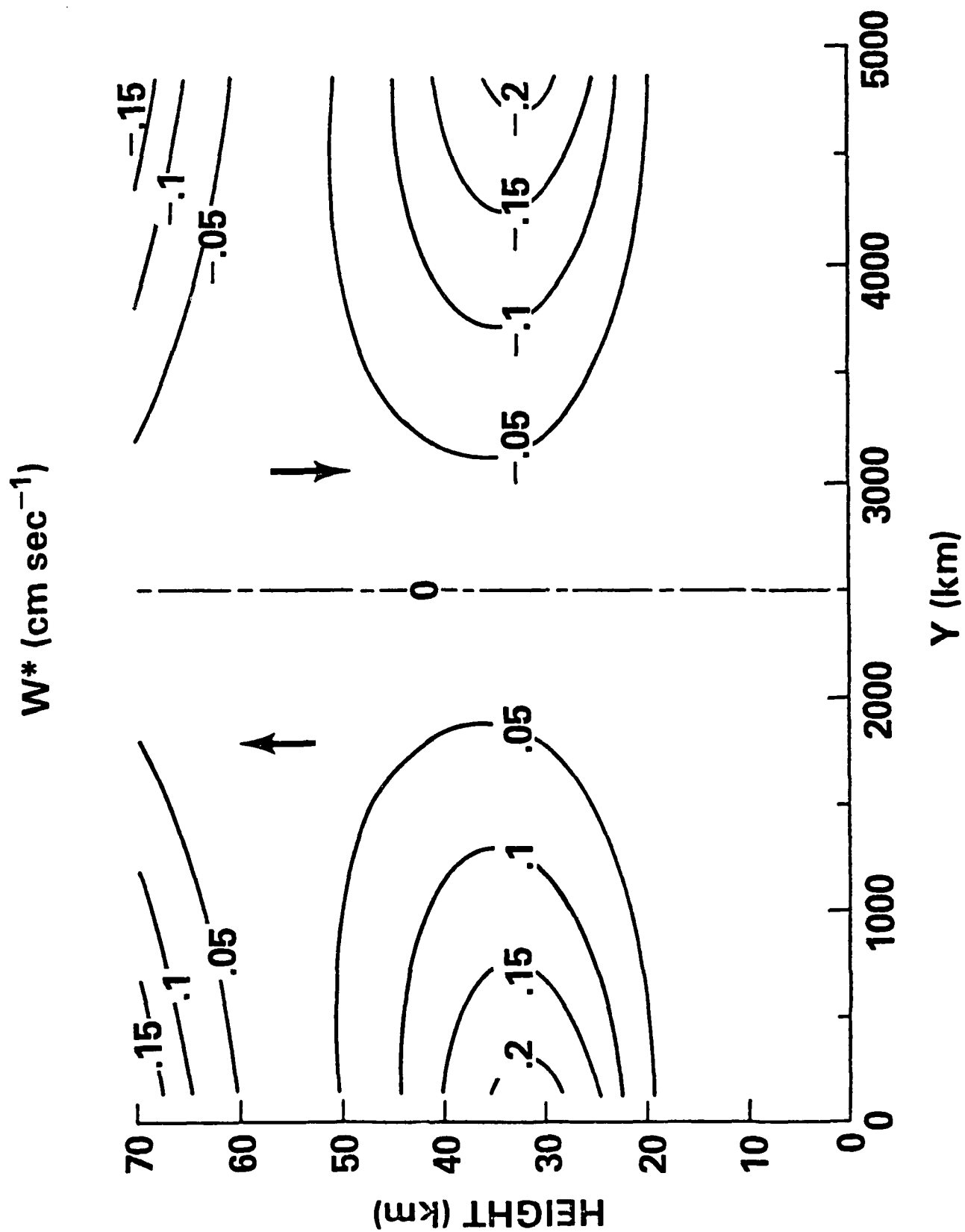


Figure 7b

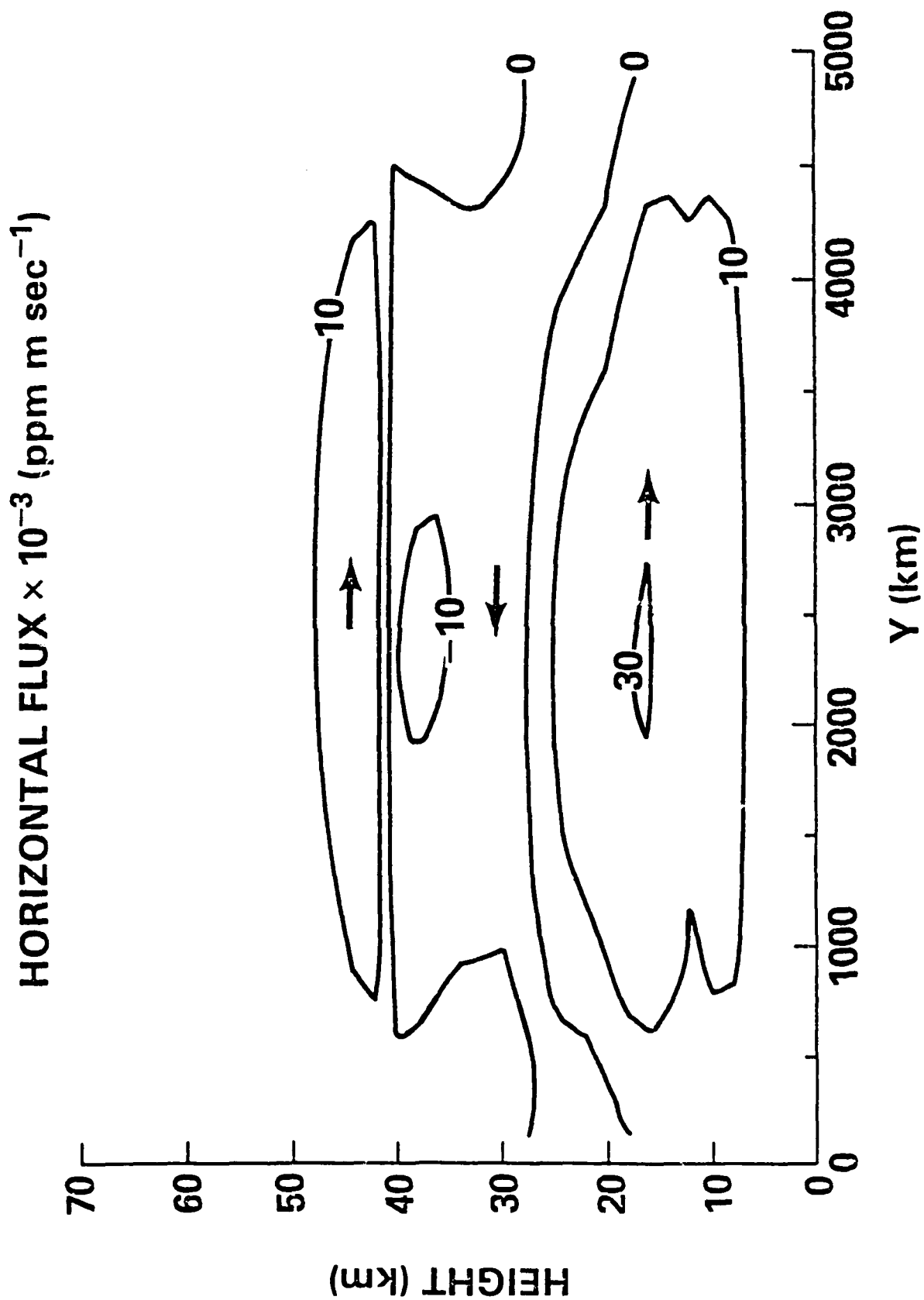
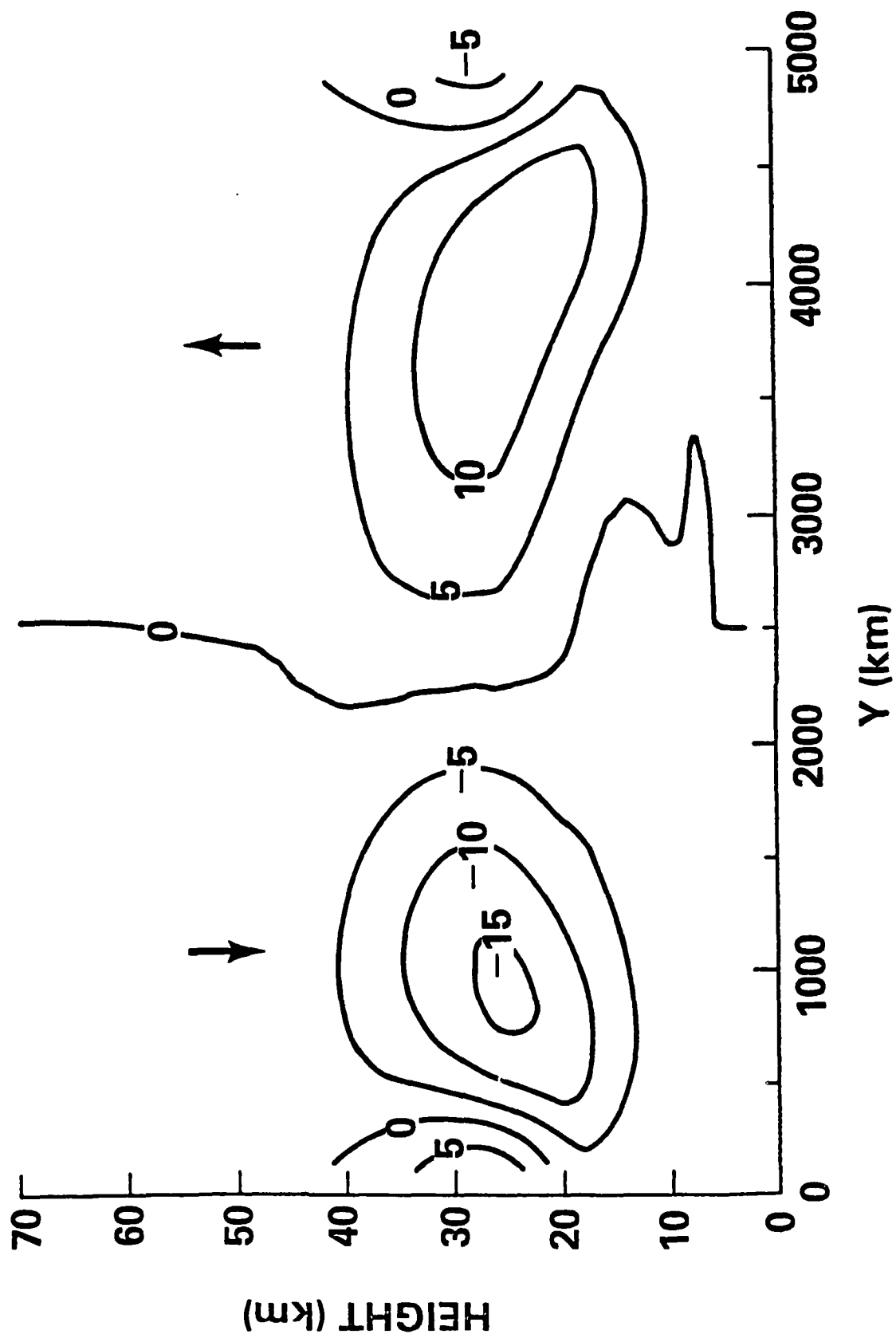


Figure 8a

VERTICAL FLUX $\times 10^{-5}$ (ppm m sec $^{-1}$)



$$\frac{\partial \bar{\mu}}{\partial t} \times 10^{-6} \text{ (ppm sec}^{-1}\text{)}$$

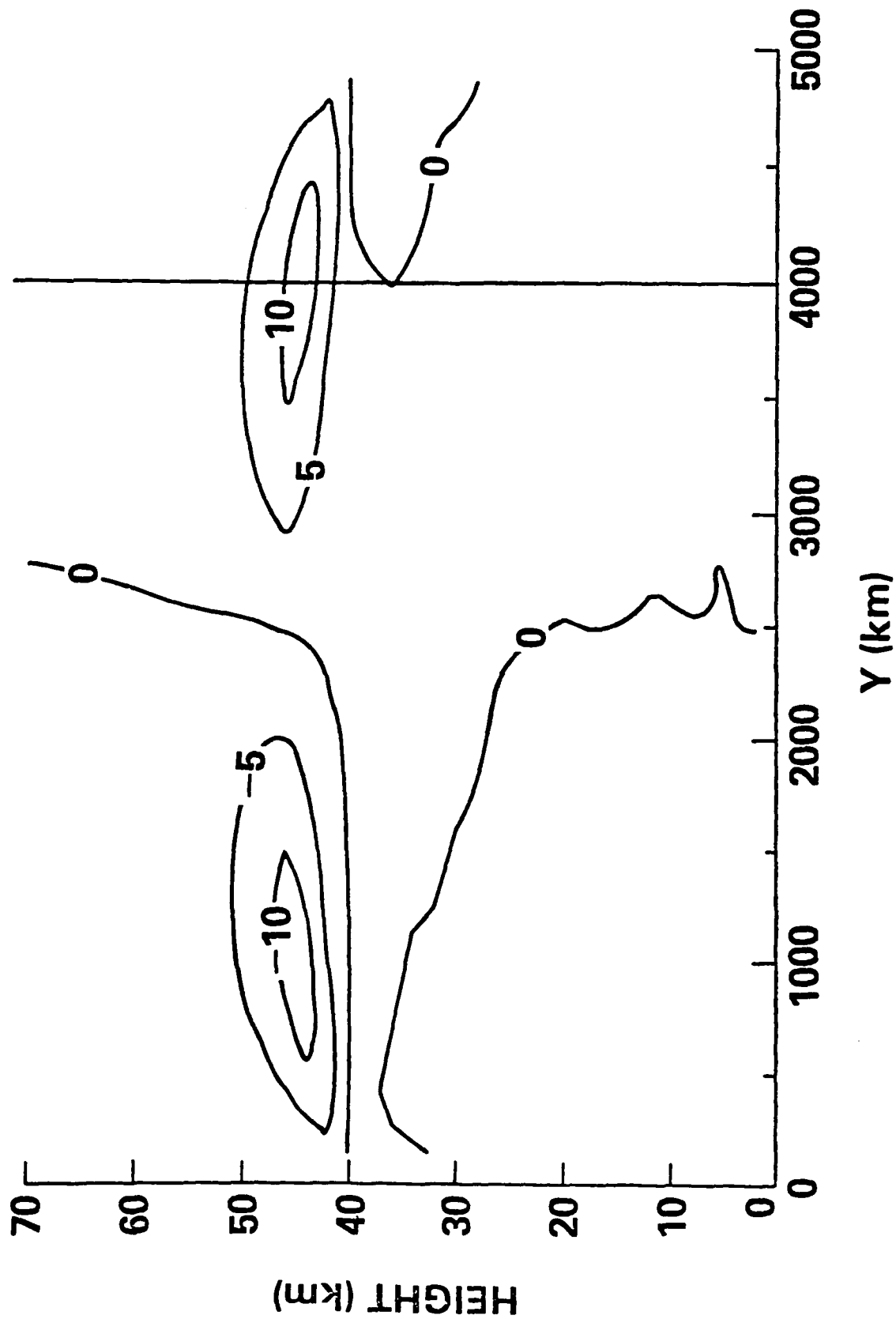


Figure 9

LAGRANGIAN-MEAN ADVECTION $\times 10^{-8}$ (ppm sec $^{-1}$)

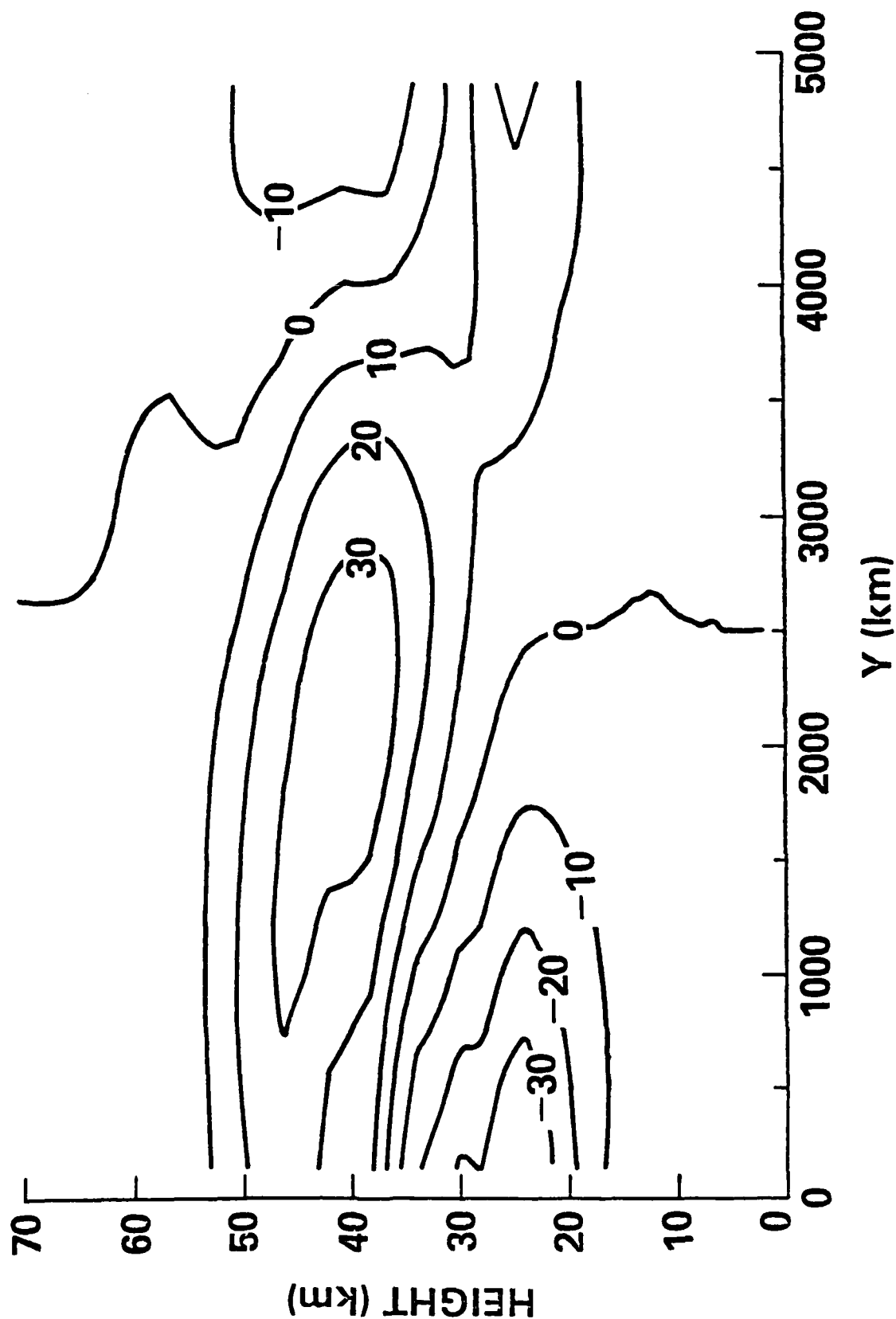


Figure 10a

RESIDUAL ADVECTION $\times 10^{-8}$ (ppm sec $^{-1}$)

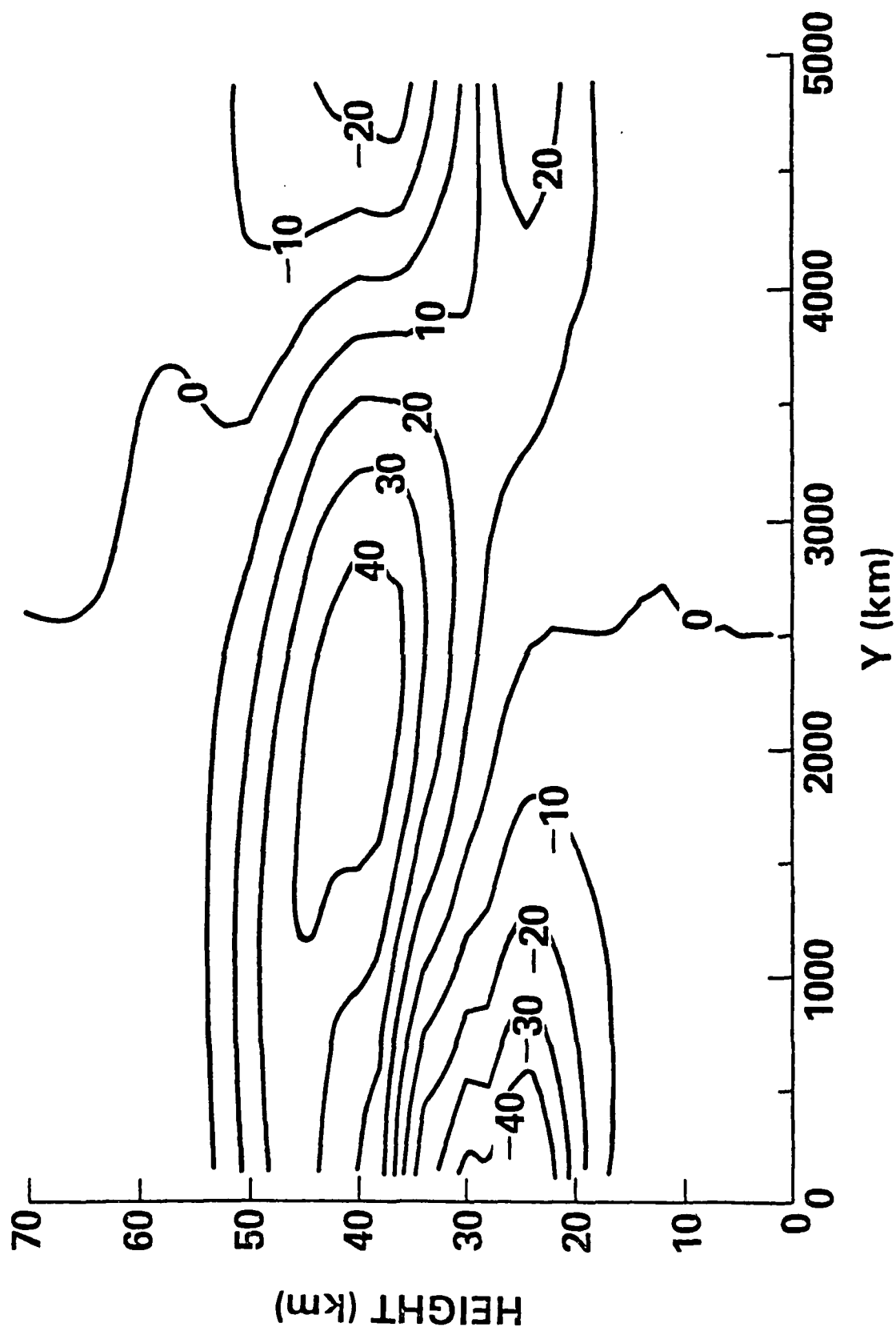
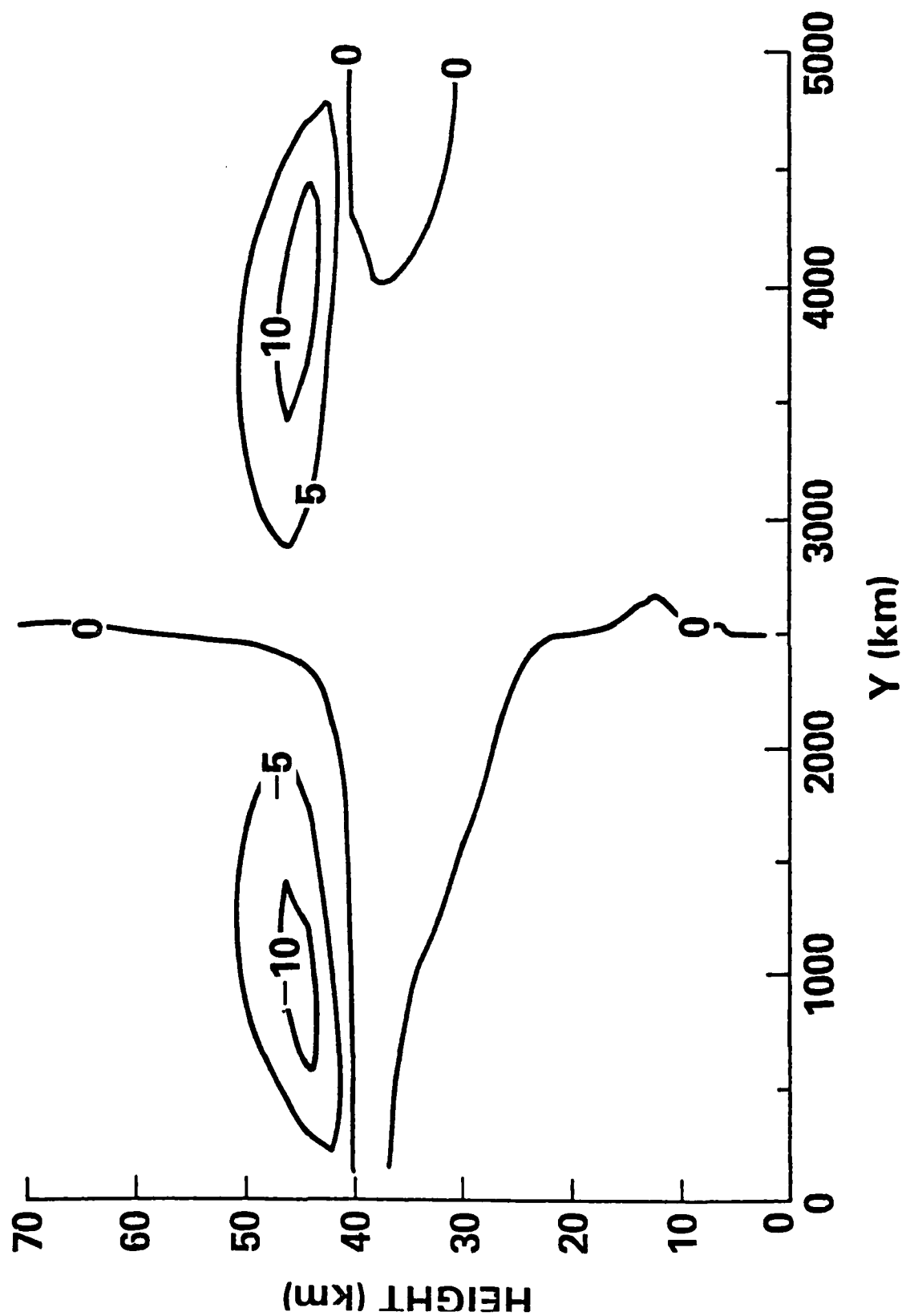


Figure 10b

$$\frac{\partial \bar{\mu}}{\partial t} \times 10^{-6} \text{ (ppm sec}^{-1}\text{)}$$



STIRRING TERM $\times 10^{-6}$ (ppm sec $^{-1}$)

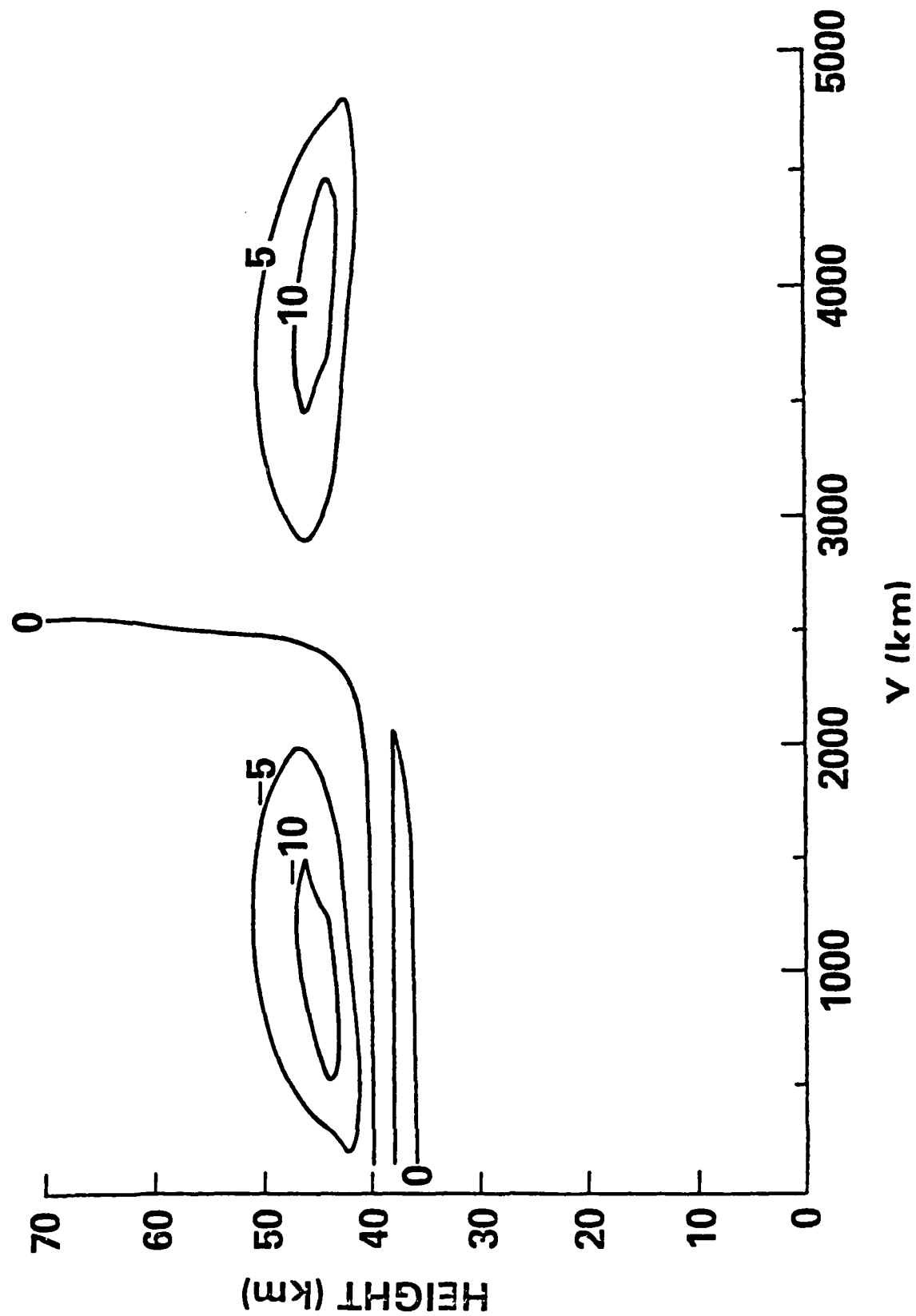


Figure 12

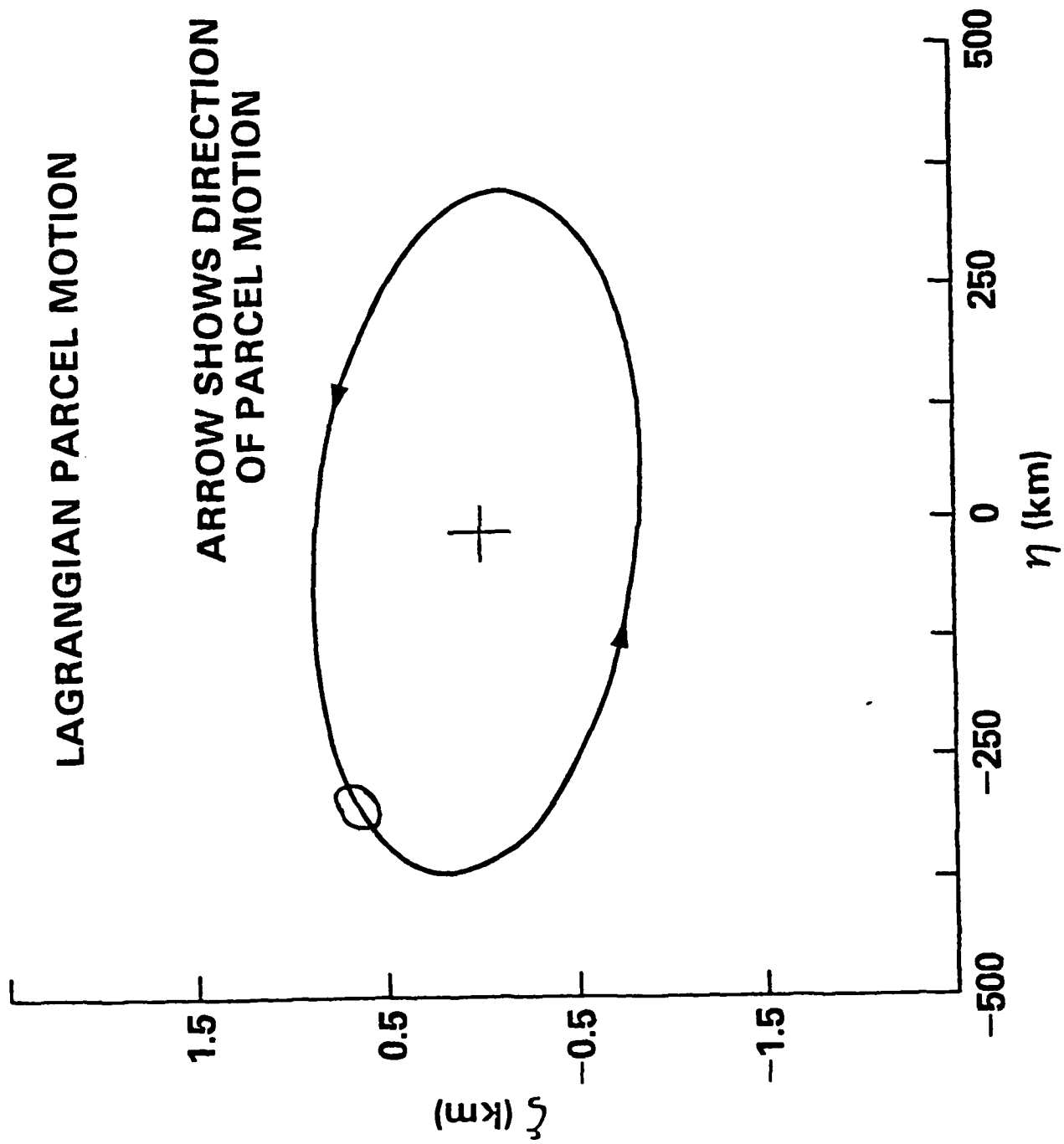


Figure 13

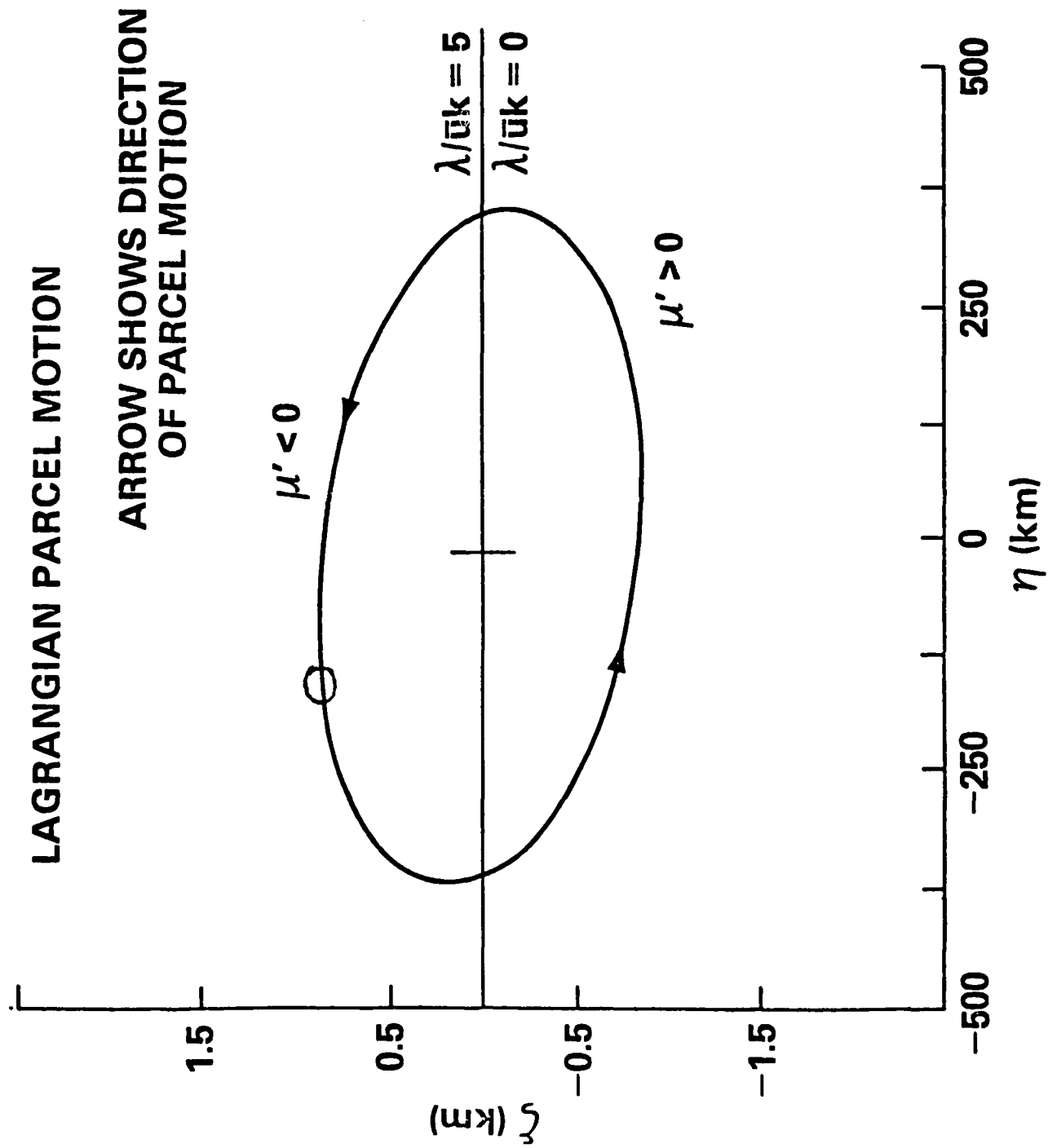
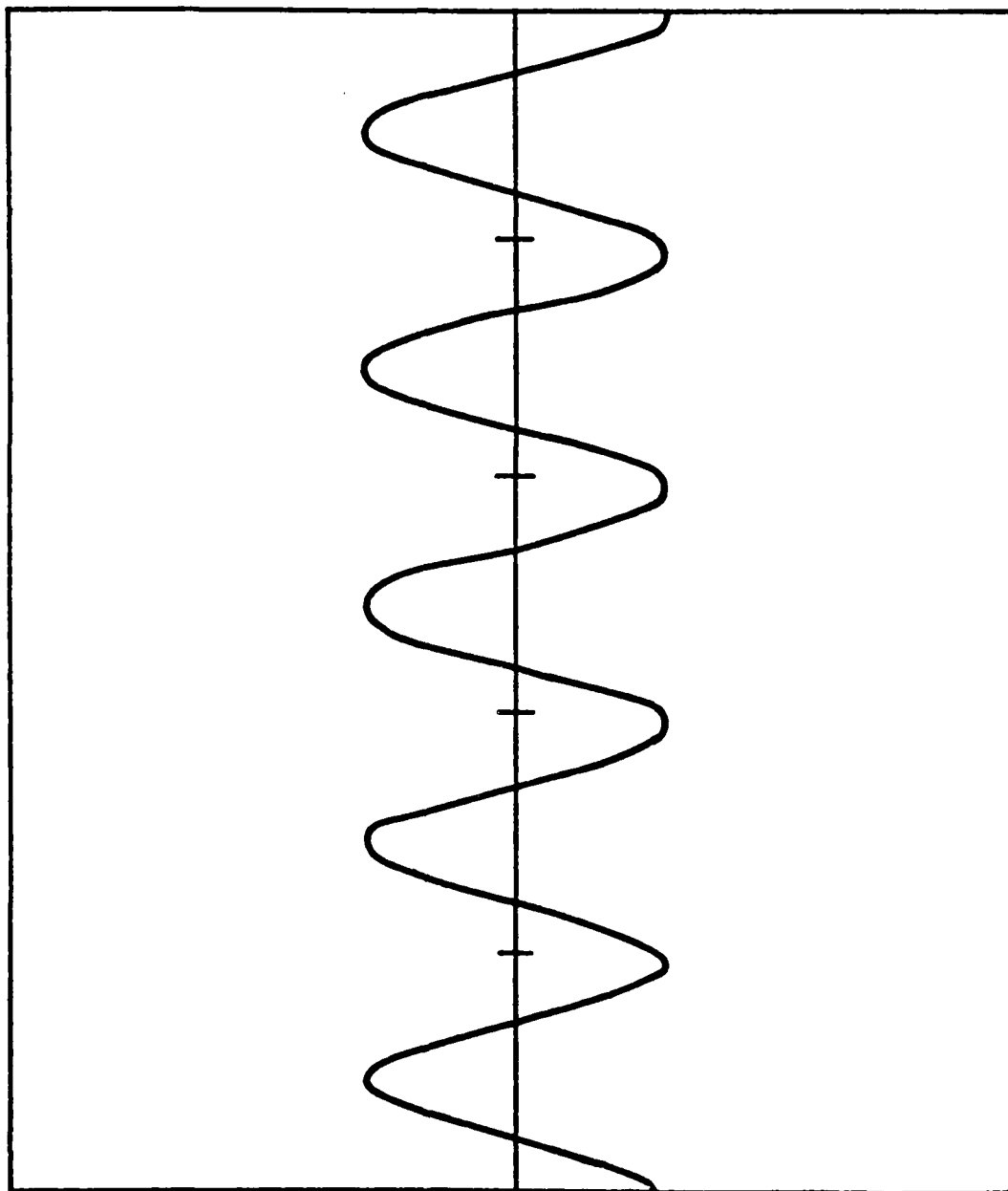


Figure 14

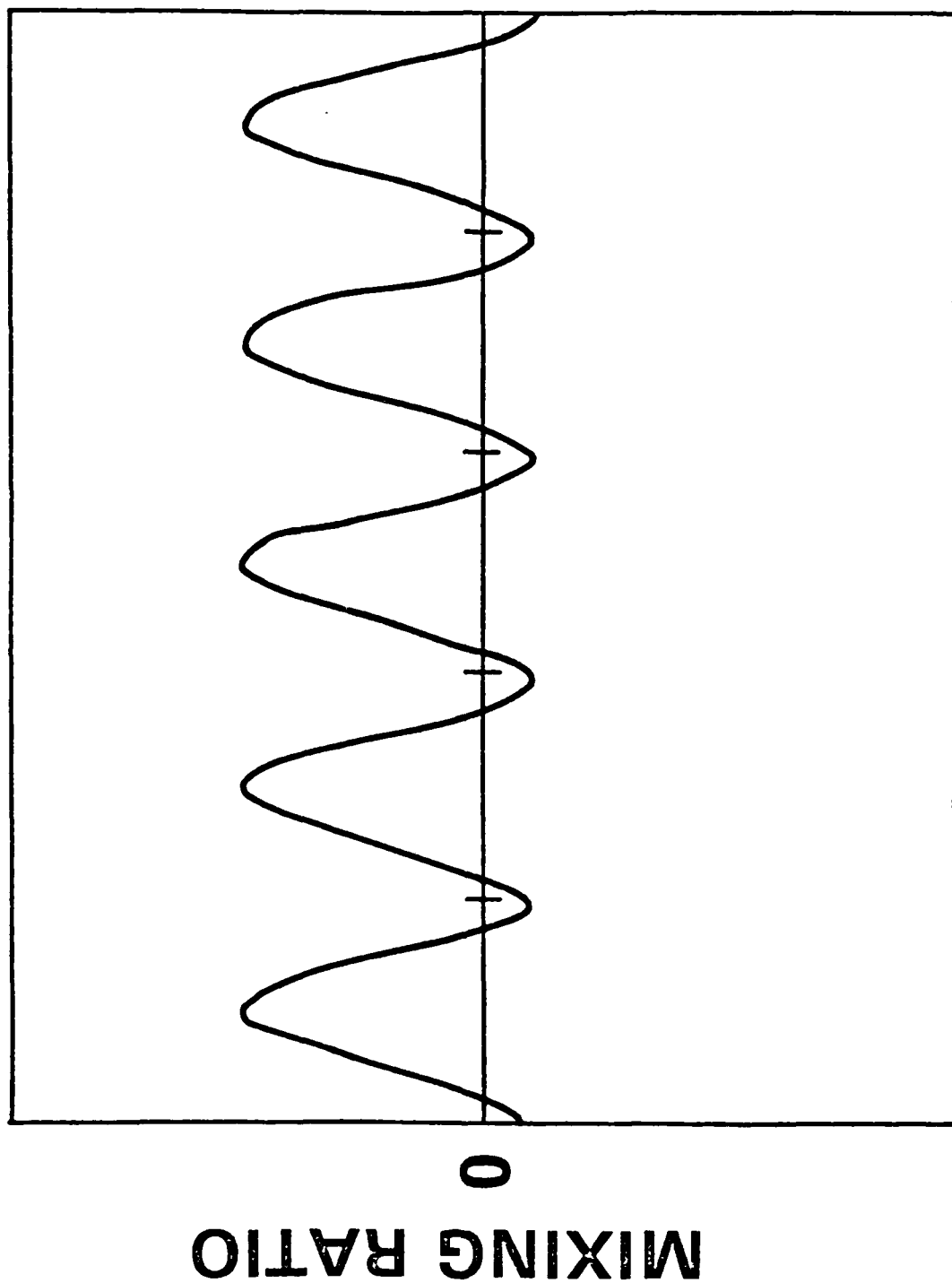
PARCEL MIXING RATIO



TIME →

MIXING RATIO

PARCEL MIXING RATIO



TIME →

Figure 15b

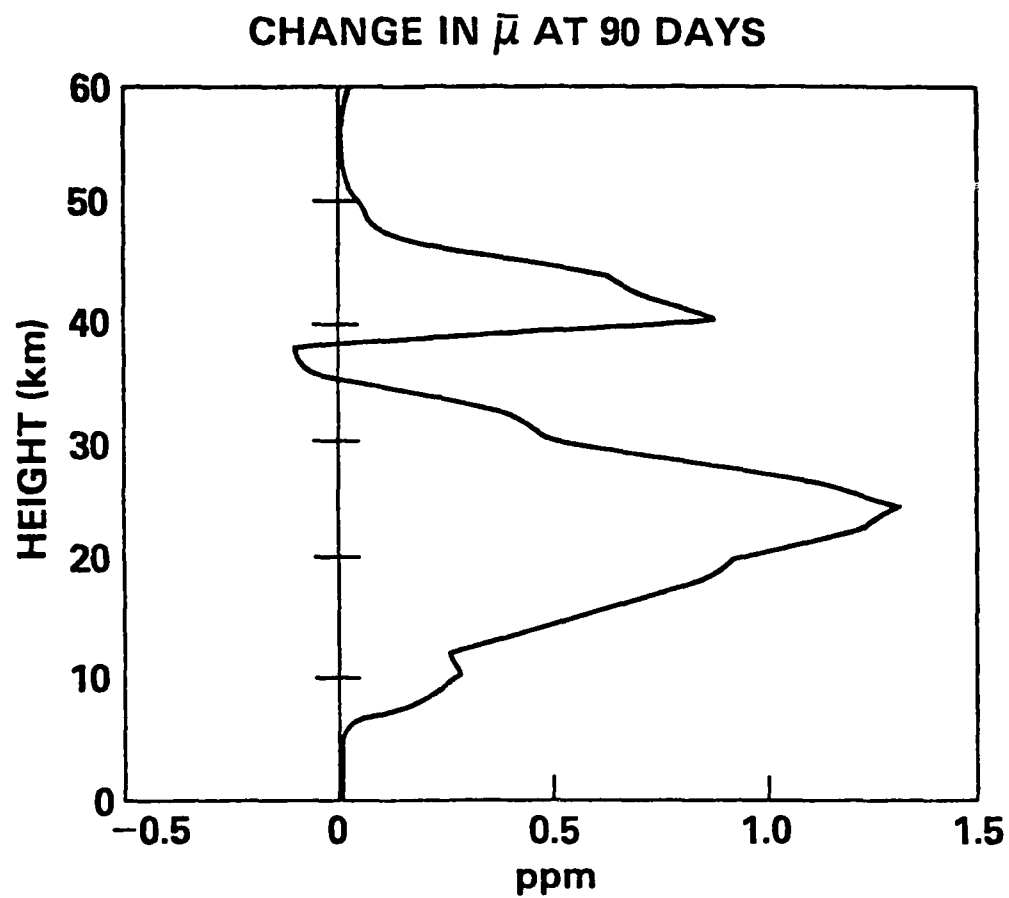


Figure 16

CHANGE IN $\bar{\mu}$ AT 90 DAYS

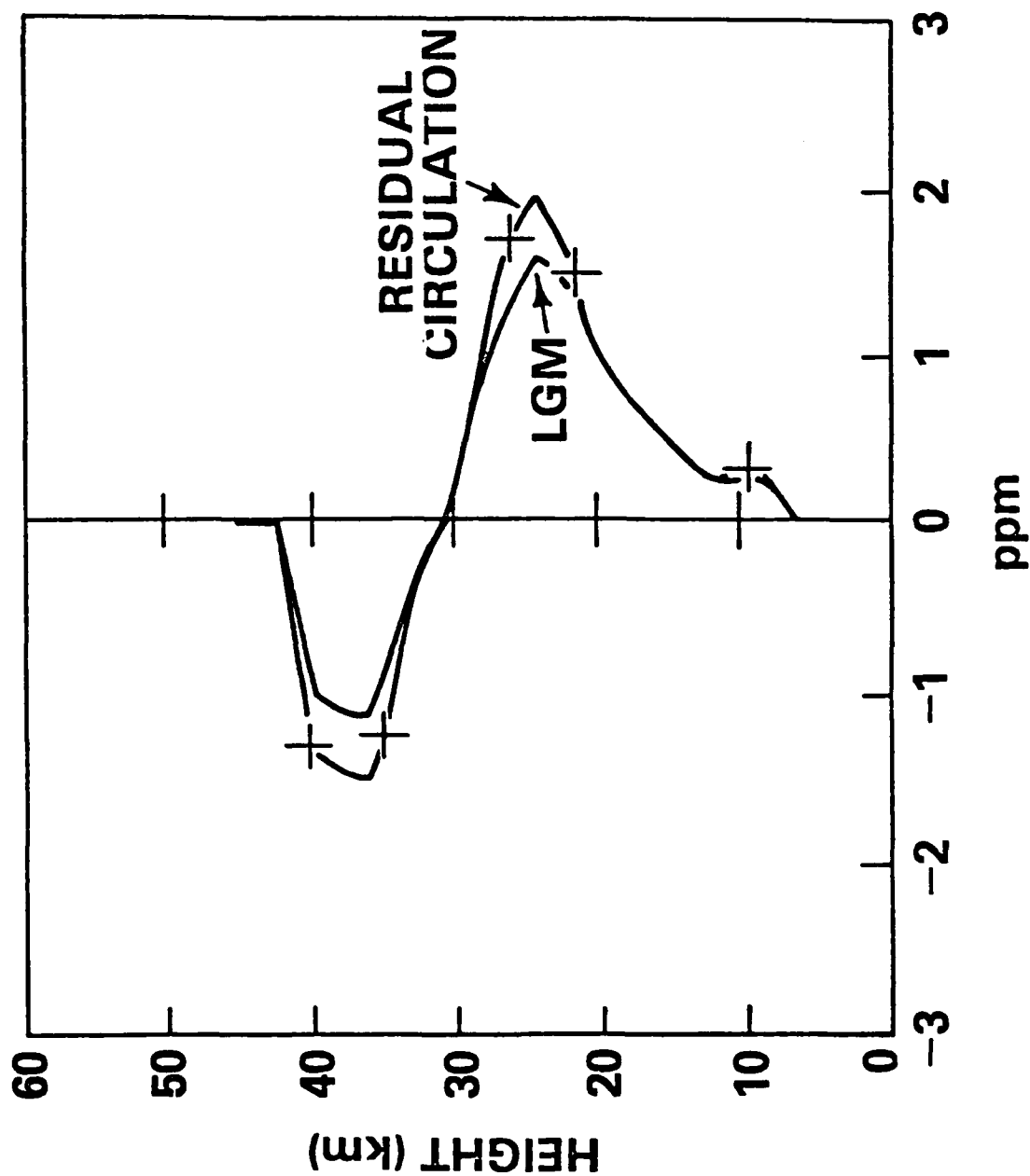


Figure 17a

CHANGE IN $\bar{\mu}$ AT 90 DAYS

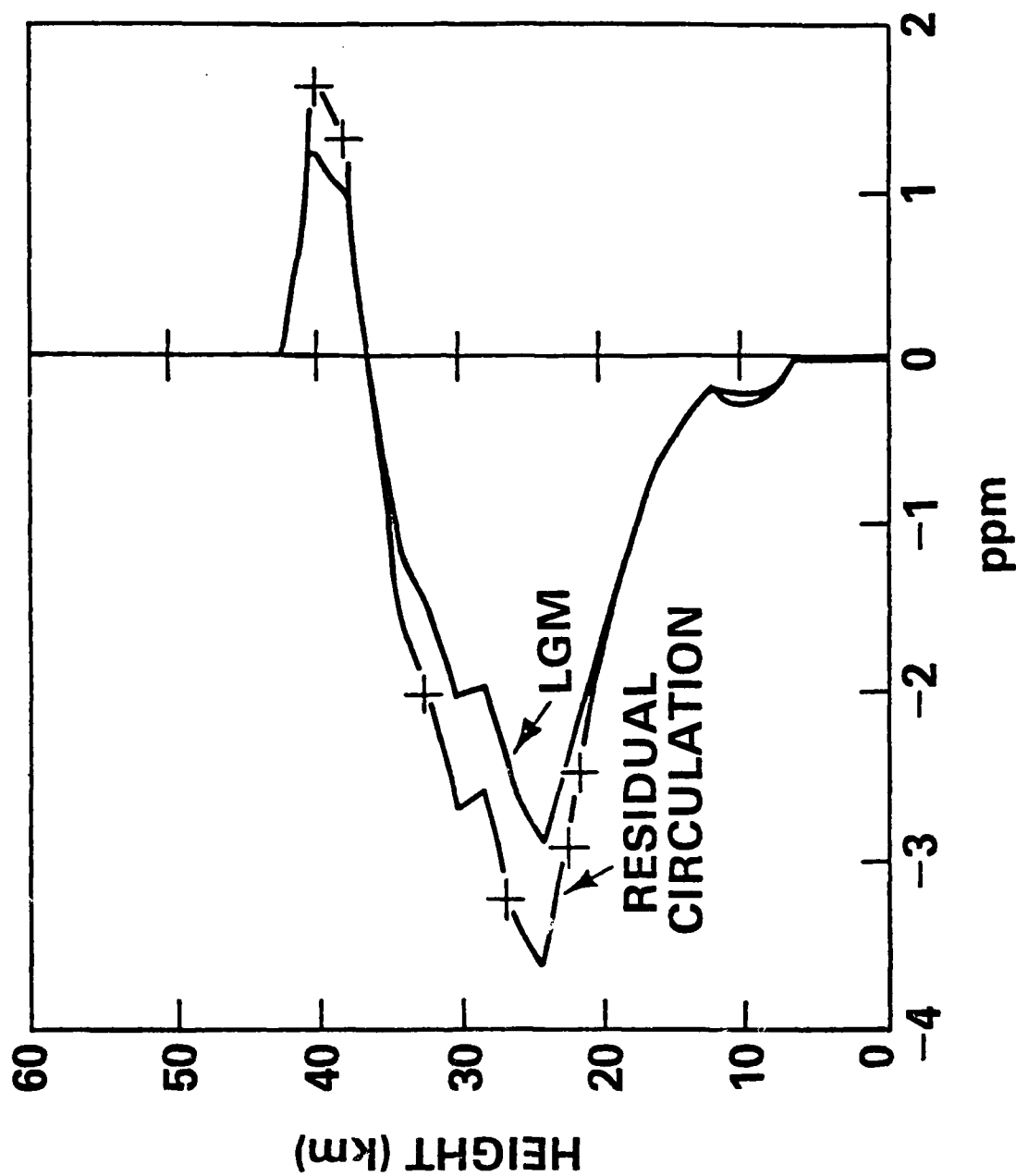


Figure 17b

NO-A126 341

GEOPHYSICAL PLASMAS AND ATMOSPHERIC MODELING(U) SCIENCE 7/7
APPLICATIONS INC MCLEAN VA E HYMAN ET AL. FEB 83
SAI-83-144-WA N00014-81-C-2038

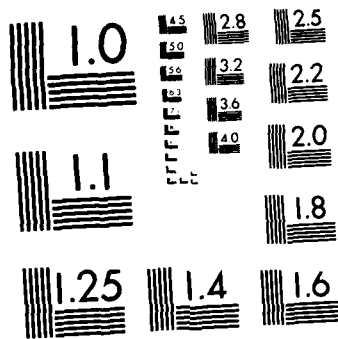
UNCLASSIFIED

F/G 4/1

NL

END

FILED
SEP 83
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

APPENDIX R

OZONE TRANSPORT BY DIABATIC AND PLANETARY WAVE

CIRCULATIONS ON A β -PLANE

Ozone Transport by Diabatic and Planetary Wave
Circulations on a β -Plane

Richard B. Rood^{*}
Science Applications, Inc.
1710 Goodridge Drive
McLean, VA 22102

and

Mark R. Schoeberl^{*}
Geophysical and Plasma Dynamics Branch
Plasma Physics Division
Naval Research Laboratory
Washington, D.C. 20375

* Current affiliation
Code 964
Goddard Space Flight Center
Greenbelt, Maryland 20771

Abstract

A β -plane model is used to study the transport of ozone and a tracer distributed like ozone. In the first experiment, ozone is transported with a time dependent diabatic circulation in a β -channel. Increase in the ozone column density occurs principally through downward advection from the high altitude source region into the lower stratosphere. Compared with observations, the maximum in the total ozone field occurs about one month too late and is about 20% too large.

Transient planetary waves produce large temporary increases in total ozone near the pole during sudden warming events, but permanent increases are much smaller. Furthermore, only warmings which penetrate below the chemical transition region are effective in changing the zonal mean ozone distribution.

Combining the sudden warming and diabatic circulation produces a quite realistic simulation of the changes in total ozone at high latitudes. The diabatic circulation provides the major enhancement of the ozone column density while the transient planetary waves move ozone to high latitudes. This result suggests sudden warmings can play an important role in our understanding of the fluctuations in total ozone, and the final warming may be the key event which determines the date and magnitude of the spring maximum.

consensus of the general circulation studies is that planetary waves and the diabatic circulation (Dunkerton, 1978; Mahlman et al., 1980) are instrumental in producing the observed annual variation of stratospheric ozone. In order to investigate further the transport characteristics of these individual processes, simplified or mechanistic models can be used. Such models are easier to interpret than the complex chemical and general circulation models, and their use may result in the implementation of more realistic dynamics in chemical models. An example which shows the utility of the mechanistic model approach is the study of Hartmann and Garcia (1979). Using a quasi-one-dimensional model of planetary wave propagation and wave-mean flow interaction, they were the first to show that planetary waves act to enhance ozone transport in the region where there is a transition from photochemical to dynamical control of the ozone distribution.

The research to be presented here involves the use of a highly simplified ozone photochemical scheme and a time dependent β -plane planetary wave model. A circulation that approximates the radiatively driven (diabatic) circulation is superimposed on the planetary wave circulation. Using this model, the transport by both planetary waves and the diabatic circulation of ozone and an ozone-distributed tracer will be studied.

In a previous paper (Rood and Schoeberl, 1982) the transport effects of steady, stationary, dissipative planetary waves were investigated. Two major mechanisms were found: The first was associated with the stirring of air parcels in the region of highly variable photochemistry and is essentially the transition region transport discussed by Hartmann and Garcia (1979). The second mechanism was due to the advection of ozone by the wave induced Lagrangian-mean velocity in the lower stratosphere (Schoeberl, 1981a). The large eddy fluxes produced by the transition region transport were found to be largely counterbalanced by the zonal mean chemistry so that only small changes in the zonal mean ozone field were the result of this effect. For 90 day integrations, the advection by the wave forced Lagrangian-mean velocity was found to cause an ozone buildup in the lower stratosphere at high, but subpolar, latitudes on the order of 1.5 ppm. However, the steady planetary waves did not transport enough ozone to high latitudes to simulate the spring ozone maximum. Even more unrealistically, an ozone depletion appeared at low latitudes which has no observational counterpart. Thus steady dissipative

1. Introduction

As concern over anthropogenic destruction of the ozone layer has increased, an intensive effort has been made to model and understand the transport of ozone and other chemically active trace species in the stratosphere. Numerical models, due largely to computational considerations, have emphasized either an accurate representation of the photochemistry (chemical models) or stratospheric dynamics (dynamic models) but not both.

Chemical models use a complex set of photochemical reactions and a one- or two- dimensional diffusive parameterization of planetary and smaller scale wave transport processes (Hidalgo and Crutzen, 1977; Harwood and Pyle, 1977; Johnston and Podolske, 1978; Miller et al., 1981). The diffusion coefficients are computed from the observed dispersion of neutral trace species such as radioactive debris. However, the transport properties of large scale waves have been recently shown to be largely advective rather than diffusive (Matsuno, 1980; Dunkerton, 1980; Danielsen, 1981; Miller et al., 1981), and both the diffusive and advective properties of these waves have been shown to be a function of the constituent photochemistry (Pyle and Rogers, 1980; Strobel, 1981). Therefore, diffusive transport models of the stratosphere are inherently limited and their validity in approximating multi-species transport is in doubt.

Dynamic modelers approach quantitative modeling of stratospheric constituents by using general circulation models, which usually have limited resolution in the stratosphere (Clark, 1970; Cunnold et al., 1975; Schlesinger and Mintz, 1979; Cunnold et al., 1980; Mahlman et al., 1980, and references therein), or by using models with severely truncated horizontal resolution but with high vertical resolution in the middle and upper atmosphere (Holton, 1982; Hartmann and Garcia, 1979). Both types of models operate with a highly simplified set of photochemical reactions to save computational cost.

The general circulation models have met with some success in reproducing the observed ozone distribution but usually underestimate the Northern Hemisphere spring maximum of total ozone at polar latitudes as well as placing the maximum too far south. Because the general circulation models are so complex, it is often difficult to isolate and separate the important phenomena which generate the observed structure of the ozone layer. Nevertheless, the

planetary waves alone cannot be entirely responsible for the large scale transport of ozone. Furthermore, the steady state planetary wave assumption is inappropriate for the entire winter stratosphere since potentially important transient effects are neglected (Schoeberl, 1982b).

The planetary wave model used in this study is the same as used by Schoeberl (1982a) to study vacillation and wave saturation. The diabatic circulation superimposed on the β -channel is based on that calculated by Dunkerton (1978). It should be emphasized that the purpose of this study is to isolate transport mechanisms, not to provide a quantitative simulation of the dynamics of the stratosphere. However, to test the viability of particular transport mechanisms it is necessary to compare the results to observations. These comparisons are performed with the intent to verify transport mechanisms.

In the next section the planetary wave model, the model of the diabatic circulation, the photochemistry, and the transport models are described and discussed. In Section 3 the transport of ozone by the diabatic circulation is presented. In Section 4 the transport of ozone and a tracer distributed like ozone is discussed for the planetary wave circulation in the absence of the diabatic circulation. The transport of ozone by the combined diabatic and planetary wave circulations is given in Section 5. Finally, the results are summarized and discussed in Section 6.

2. Models

a. Planetary Waves

The planetary wave propagation and wave-mean flow interaction are modeled using the quasilinear, quasigeostrophic potential vorticity equations. Newtonian cooling and Rayleigh friction are simulated using a single potential vorticity dissipation coefficient which is taken to be a weak function of height except at the top where a sponge layer exists. In the stratosphere the dissipation time scale is 23.5 days a value typical of the lower stratosphere. Since the model is the same as Schoeberl (1982a), but with twice the horizontal resolution, the numerical methods and governing equations will not be repeated here.

The waves are forced by zonal flow over a sinusoidal mountain with zonal wavenumber one and a meridional half wavelength of 5000 km. The mountain is brought up to its maximum height on a time scale of 10 days. Its height is a model parameter and the behavior of the model varies depending on its value. For small mountains (< 300 m), equilibrium is reached in which westerlies are present everywhere. For mountains in an intermediate range (300-900 m) the model produces at least one stratospheric warming and then is observed to reach a steady state with a permanent layer of easterlies at some altitude. For large topography (> 900 m) the model produces a series of warmings, i.e., it vacillates. The model dynamics are discussed extensively in Schoeberl (1982a). The basic state mean zonal wind profile is shown in Fig. 1.

During winter it is common to see distinct planetary wave amplitude pulses associated with either major or minor warmings (Labitzke and Goretzki, 1982). After a major or minor warming the stratosphere usually returns to its winter climatological state. If the warming occurs late in the winter, then the zonal mean circulation after the warming remains in the summertime regime where easterlies are present throughout the entire summer stratosphere. This type of warming is called a final warming. The dynamical model used here does not attempt to represent all of the features of the winter stratosphere but is used only to simulate warming events and their corresponding circulation patterns.

Time series for the mean zonal wind field and the wave amplitude field are shown in Figures 2 and 3 for the 900 m topography. The 900 m topography

was chosen for most of experiments since forcing of this amplitude is in line with the terrestrial topography (Berkofsky and Bertoni, 1955), and the critical line for this case descends below 32 km (a criterion for a major warming). Transport experiments were run for forcing heights of 200 m, 600 m, and 900 m.

The wave amplitude (Fig. 3) increases rapidly until about day 20 when it reaches a maximum of 1.6 gpkm and then decreases rapidly between days 20 and 30 and more slowly until day 50. There is another increase in wave amplitude after day 50 which maximizes on day 65. The wave amplitude decrease following this maximum is not as rapid as that following the initial warming event.

Coincident with the first wave maximum is the deceleration of the mean zonal wind and the formation of easterlies which reach a maximum velocity on day 26 of 57 m sec^{-1} (Fig. 2). The second wave maximum on day 65 is not immediately associated with the onset of high level easterlies. The wave forcing of the mean flow from days 55-65 decelerates the mean zonal wind, restoring the low level easterlies which then prohibit propagation of the wave into the stratosphere. The low level easterlies ultimately disappear about day 100 which allows the wave to once again propagate vertically. The zonal mean westerlies are so weak at this time that even the relatively weak forcing associated with the wave pulse at day 100 is capable of producing easterlies throughout a deep portion of the atmosphere.

The zero wind line (critical line for stationary waves) descends to 26 km on day 37 and 24 km on day 76. Between days 45 and 55 the critical line rises and the wave propagates and rebuilds the low level easterlies. The low level critical line is absent for a brief period beginning at day 100.

b. The Diabatic Circulation

Early theories of stratospheric circulation postulated a direct cell with rising motion in the equatorial region and sinking motion in the polar region in order to explain the dryness of the stratosphere and the buildup of ozone at high latitudes (Brewer, 1949; Dobson, 1956). Observations during the 1960's showed, however, that the Northern Hemisphere, lower stratospheric zonal mean circulation consisted of two cells with rising motion at high latitudes and at the equator. Dunkerton (1978) pointed out that in the absence of strong wave transience and dissipation the diabatically driven

zonal mean field is a good approximation to the Lagrangian-mean velocity field and is consistent with the Brewer-Dobson circulation. Hence, the large scale mass transport associated with radiative forcing is described by a one cell circulation as originally postulated.

The diabatic circulation used here is derived from Dunkerton (1978) whose calculations were based on those of Murgatroyd and Singleton (1961). The circulation appropriate for the winter solstice is shown in Fig. 4, and is derived by estimating the meridional component of the velocity and then using the continuity equation to compute the vertical component. Due to the lack of sphericity on the β -plane, the polar maximum calculated in the vertical velocity field by Dunkerton (1978) is not reproduced. The time variation of the diabatic circulation is assumed to be harmonic with a one year period such that the intensity of the flow reaches a maximum shortly after the winter solstice and then decreases and reaches a minimum in the spring near the equinox and is reversed during summer.

The stratospheric heating used by Murgatroyd and Singleton (1961) circulation is based on observed temperature fields. These fields include the redistribution of heat by the stratospheric standing eddies; so, the diabatic circulation computed by Dunkerton partially incorporates the effects of the eddies. Dunkerton's (1978) calculation assumes that wave transience and dissipation is small. To test the sensitivity of the ozone transport to the diabatic circulation, the circulation calculated by Apruzese et al. (1982) is also used. This circulation is based on a calculated radiative equilibrium temperature which only includes eddy effects parameterized as Rayleigh friction.

c. Chemistry Model

The chemistry used in this study is the same as used in Rood and Schoeberl (1982), namely:

$$Q = -\lambda(\mu - \bar{\mu}_0) \quad (\lambda > 0) \quad (1)$$

where Q is the photochemical source term, λ is the photochemical relaxation

rate, μ is the ozone mixing ratio, and $\bar{\mu}_0$ is the zonally averaged basic state of ozone. This formulation requires that the wave perturbations relax back to zero, while changes in the mean field relax to the mean equilibrium. The rate of relaxation is inversely proportional to λ , which is assumed to be a function of height only.

The λ values are based on those of Hartmann and Garcia (1979) and are derived from the oxygen-only reactions (Chapman reactions) with a modified ozone destruction rate to represent catalytic destruction. Ozone chemistry is temperature dependent; however, several model experiments indicate that the effect of the temperature dependence is small and does not change the character of the transport. Therefore, the temperature dependence will be neglected here.

It is convenient to divide the atmosphere into three regions in order to discuss ozone transport (Figure 5). The lower region is the region in which ozone is assumed to be a conservative tracer ($\lambda \ll k\bar{u}$, k is the zonal wavenumber and $\bar{u}$ is the mean zonal wind). In the upper region ozone photochemistry is so fast ($\lambda \gg k\bar{u}$) that the transport is unimportant. The middle region is defined such that λ and $k\bar{u}$ are of comparable magnitude and is called the transition region (Hartmann and Garcia, 1979). The interaction of the photochemistry with the dynamics is expected to be greatest in the transition region. There are nonzero values of λ below 40 km, but they are so small that the resolution in Fig. 5 does not accurately represent them.

The chemistry used here contains many inadequacies, most notably the assumption of a purely conservative tracer in the lower regions of the model. However, the chemistry does approximate to first order an ozone-like constituent, and should be adequate for the identification of important transport mechanisms.

d. Transport model

The transport model runs in an off-line fashion; that is, the constituent distribution does not affect the dynamics. Since ozone is radiatively active, with the use of an off-line model it has been tacitly assumed that the change in radiative forcing associated with the changes in the ozone field can be neglected. This assumption is valid as most of the transport in this model

takes place during winter when insolation is weak. Mahlman and Moxim (1978) discuss in detail the differences between on-line and off-line models.

The constituent continuity equation is divided into zonal means and perturbations and appears as:

$$\frac{\partial \bar{\mu}}{\partial t} + \bar{u}_i \frac{\partial \bar{\mu}}{\partial x_i} + \overline{u_i' \frac{\partial \mu'}{\partial x_i}} = \bar{Q} \quad (2)$$

and

$$\frac{\partial \mu'}{\partial t} + \bar{u}_i \frac{\partial \mu'}{\partial x_i} + u_i' \frac{\partial \bar{\mu}}{\partial x_i} = Q' \quad (3)$$

where bars and primes represent zonal averages and deviations, a repeated indices indicate summation. $u_i = (u_1, u_2, u_3) = (u, v, w)$ are the velocities in the $(x_1, x_2, x_3) = (x, y, z)$ directions where x is longitude, y is the β -plane analogue to latitude, and z is height. The linear perturbation equation is solved at each time step and is used to update the mean field for subsequent time steps.

For the transport model, the constituent continuity equation is written in discrete form using fourth order centered finite differences in the meridional and vertical directions. The longitudinal dependence is represented by zonal harmonics, and the time derivatives are estimated with second order centered differences (leapfrog). The spatial differences are written in the general form

$$\frac{\partial \mu}{\partial p} = \frac{C_1(\mu_{i+1} - \mu_{i-1})}{\Delta p} + \frac{C_2(\mu_{i+2} - \mu_{i-2})}{\Delta p} \quad (4)$$

where $C_1 = 2/3$ and $C_2 = -1/12$ for fourth order accuracy (Ames, 1977; Haltiner and Williams, 1980) and p is either the vertical or horizontal coordinate. No diffusion other than that inherent in the finite difference scheme is used, and filling (Mahlman and Moxim, 1978) has not been necessary.

The transport and dynamic models are centered at 60° N and have a horizontal extent of 5000 km. This corresponds to a latitudinal extent from 37.5° to 82.5° . The domain extends from the ground to 100 km. There are 20 interior grid points in the horizontal and 40 grid points in the vertical.

Equations (2) and (3) are solved in flux form, and the constituent flux

is set equal to zero at the top and bottom of the model. This boundary condition is accurate at the top since the ozone is in photochemical equilibrium. At the northern and southern boundaries the flux associated with the planetary waves is zero since the wave amplitude is zero. At the southern boundary, where ozone is observed to have relatively little annual variation, the flux by the diabatic circulation is calculated assuming that ozone mixing ratio is a constant function of time. The horizontal flux associated with the diabatic circulation at the northern boundary is assumed to be zero.

The closure of the finite difference scheme directly involves the prescribed boundary point and the first interior point. The second and third interior point are directly associated with the first interior point at each time step. Because of this possible direct contamination by the closure system, results will be presented at horizontal grid points 4 (46°) and 17 (72°) as representative of the northern and southern regions of the model.

e. Initialization

The zonal mean ozone field used to initialize the transport model is shown in Fig. 6. This field is derived from the data of Dütsch (1969) and is representative of the mid-November zonally averaged ozone mixing ratio. Data north of 70° and above 45 km are scarce and of low quality. In these regions the prescribed ozone field is based on the vertical profile derived by Kreuger and Minzner (1976) and the ozone field used by Fels et al., (1980). The eddy ozone field is initially assumed to be zero.

A review of stratospheric warming data indicates that there is often a major or minor warming during the first or second week of January. Since the model warmings occur about three weeks after the planetary waves are turned on, December 15 has been chosen as the initialization data. In order to properly correlate the ozone field with the planetary wave dynamics the initial ozone field is advected with the diabatic circulation until Dec 15.

3. Transport by the diabatic circulation

The results will be reported as the change from the initial mean ozone field which is defined as the anomaly. The anomaly due solely to the diabatic circulation is presented as a one year time series for two latitudes in Figures 7 and 8. In general there is a positive anomaly in the lower stratosphere (20-30 km) for much of the year and a negative anomaly in the autumn. Since the ozone is in near equilibrium above 35 km, a positive (negative) anomaly is associated with an increase (decrease) in total ozone. Using this correlation it is seen that the total ozone reaches a maximum in early April and a minimum in October as is evident in Fig. 9. The October minimum is observationally consistent, but the April maximum appears too late.

As is apparent in the figures there is a strong latitudinal dependence of the anomaly. The high latitude wintertime increase is larger and reaches lower altitudes. In Fig. 8 (74°) a maximum increase of 3.5 ppm is centered around 26 km, and changes greater than 1 ppm are seen as low as 15 km while in Fig. 7 (60°) an increase of 3 ppm is centered at 30 km. In the southern region (not shown) a maximum increase of 0.5 ppm is centered around 32 km. The largest increase in the entire domain, 4.25 ppm, occurs at 70° N.

The changes in ozone above 45 km in Figures 7 and 8 are negligible due to the fast photochemistry. Therefore, where the diabatic circulation (Fig. 4) is strongest no changes in the ozone field are found. The transport effects by the diabatic circulation are most notable in the transition and conservative regions and are due primarily to the vertical velocity which is downward at all latitudes and constantly advecting ozone into the lower stratosphere during winter. The maximum anomaly occurs about 500 km north of the maximum vertical velocity indicating the relative ineffectiveness of the meridional velocity ($<.3 \text{ m sec}^{-1}$) at transporting ozone northward.

The total ozone time series from Dec 15-April 15 shown in Fig. 9 better illustrates the overall effects of transport by the diabatic circulation. A maximum of 538 DU is reached about April 1 at 70° . Total ozone climatology (Fig. 10) shows a maximum of 450 DU near the pole about March 1.

Because insolation is weak during the winter the ozone photochemistry is much slower than modeled here at high latitudes. Since the increase of ozone column density by vertical motions requires the existence of a source region, the overestimation of the photochemical relaxation may be responsible for the

unrealistic buildup in the northern region evident in Fig. 9. The results from two additional experiments where latitudinal variation of the chemistry is included are shown in Fig. 11. In the first experiment (Fig. 11a) the photochemistry is given a cosine of latitude dependence such that the chemical relaxation coefficient decreases to zero at the northern boundary. This is the same latitudinal dependence as used by Garcia and Hartmann (1980). It is apparent comparing Figures 8 and 11a that there is no significant difference in the results from the case with no latitudinal variation of the chemistry. The vertical transport is so slow that even with much slower photochemistry the ozone never varies much from its photochemical equilibrium value.

In the second experiment (Fig. 11b) the photochemistry is set exactly equal to zero north of 67.5° for the entire time integration which is an exaggerated representation of the polar night. The positive anomaly in the lower stratosphere remains the same as in the other experiments, but there is a negative anomaly around 50 km of more than 1 ppm. This region of depletion exists because there is no chemical relaxation to restore the ozone transported downward, and the ozone is not significantly replenished by meridional transport from the southern source region. The ozone increase in the lower stratosphere is roughly the same as in the other experiments, but there is no increase in total ozone at high latitudes due to the lack of the high altitude source region. Therefore, incorporation of the polar night into the calculation of the ozone transport does not rectify the inability of the diabatic circulation to simulate the ozone distribution.

To check the sensitivity of the ozone distribution to the specification of the diabatic velocity field an experiment was run using the diabatic velocity calculated by Apruseze et al. (1982). This diabatic circulation is qualitatively similar to the circulation found by Dunkerton (1978) but is based on a calculated temperature field that does not include the redistribution of heat by planetary waves. To preserve the polar maximum, the vertical velocity field was estimated and the meridional component was then calculated by continuity.

With the Apruseze et al. (1982) circulation, the meridional component of the velocity was still found to be too small to produce significant poleward transport of ozone. An ozone maximum of greater than 550 DU occurred in April at high latitudes ($>80^{\circ}$). The size of this maximum is strongly related to the

magnitude of the vertical component of the diabatic velocity between 30 and 50 km. The Apruzese et al. (1982) vertical velocity is larger and does not show the distinct midlatitude maximum at these altitudes that the Dunkerton (1978) circulation shows. The variation between estimations of the diabatic circulation depends on differences between atmospheric radiation models, and this experiment demonstrates the sensitivity of transport experiments to these differences.

In summary, the diabatic circulation during winter is directed such that transport is poleward and downward. The northward component of the velocity is too small to effectively transport ozone to polar regions; however, the vertical component acts over a considerable depth of the stratosphere and transports significant amounts of ozone into the lower stratosphere from high altitude source regions. If there is no photochemistry, then the vertical velocity only moves ozone up and down and does not change the column density; however, the diabatic velocity is so small that even very weak photochemistry at high altitudes causes significant increases in the ozone column density. The spring maximum in total ozone calculated using the diabatic circulation alone is too large in magnitude and occurs about one month too late. In addition, for the diabatic circulation based on Dunkerton (1978), the maximum in total ozone occurs too far south.

4. Planetary Wave Transport

In this section the transport of ozone and an ozone-distributed passive tracer (conservative transport) by planetary waves and the wave induced mean circulation forced with 900 m topography is discussed (see Figures 2 and 3). Superimposed on the time height cross sections discussed in this and later sections is the location of the critical line at the latitude of the figure and in the center of the channel. The transport of the planetary waves and the diabatic circulation together is discussed in Section 5.

a. Conservative transport

The changes in the mean ozone field for conservative transport by planetary waves forced with the 900 m orography is shown at 74° in Fig. 12. There is a zone of ozone enhancement centered at 35 km and a zone of depletion centered at 50 km. At the southern end of the channel (not shown) similar structure is seen, but with depletion at 35 km and enhancement at 50 km. The maximum positive change at the northern point for the 900 m case is 2.2 ppm and occurs on day 32. With 600 m topography a maximum change of slightly more than 1 ppm occurs on day 35. For both cases there is relatively little change in $\bar{u}$ at the center of the channel. Figure 12 clearly shows that the largest changes in ozone are coincident with the descent of the critical line.

The transport of constituents along a critical line (CL) has been investigated using analytical models by Matsuno and Nakamura (1979) and Schoeberl (1981b). In these studies, the Eliassen-Palm (EP) flux is modeled as a step function across the CL. This gives rise to a singular, poleward Lagrangian-mean flow at the CL with weaker return flow outside the region of the critical line. The Lagrangian-mean mass stream function computed by Matsuno and Nakamura for such an idealized CL interaction is illustrated in Fig. 13a.

In the analytic models the singular convergence of the EP flux at the CL imbeds all of the wave transience, which is proportional to the eddy potential vorticity flux, in the singularity. Since the wave is not propagating above the CL, the return Lagrangian-mean flow in that region is due entirely to the induced Eulerian mean circulation which is shown in Fig. 13b. Schoeberl (1981b) showed that where the CL shifts its orientation with respect to the horizontal, the strength of the Lagrangian-mean flow changes and a constituent flux convergence will occur at the cusps.

In this numerical model, the CL shows many of the properties predicted by the analytical studies. For example, the strongest eddy constituent flux centers on the CL as predicted by the analytic models (Fig. 14). However, the EP flux is not absorbed precisely at the CL because of dissipation, time dependence, and numerical approximations. Therefore both the wave forcing of the mean flow and the wave induced flux of constituents are smeared out around the CL. The eddy flux constituent convergence/divergence is largest at the poleward/equatorward side of the channel where the CL has turned upward and the wave forcing of the mean flow is weaker (the mountain height falls off towards the edges of the channel and there is not very much lateral propagation of the wave (Schoeberl, 1982a)).

Above the CL in Fig. 12, a moderate equatorward eddy flux of the constituent occurs. This reverse flux is not predicted by the steady state CL theories because it is due to the collapse of the wave field in the easterly wind zone as the CL moves downward. This collapse produces negative wave transience and equatorward flux of constituent. The reverse flux in this region is responsible for the weak negative anomaly at 50 km in Fig. 12. The negative wave transience is also partially responsible for the reacceleration of the mean zonal flow after day 25 (see Fig. 2).

The Eulerian mean constituent flux shown in Fig. 15 compares quite well with the Eulerian circulation above the CL predicted in Fig. 13b. This indicates that the forcing of the mean flow by the negative wave transience above the CL is overshadowed by the forcing at the CL.

The positive ozone anomaly shown in Fig. 12 at high latitudes disintegrates rapidly after day 42. This effect is almost entirely due to the reversal of the ozone flux associated with the reacceleration of the mean flow and the wave amplitude decrease. Fig. 16 shows that the eddy flux of ozone center around the critical line on day 47 is equatorward and located around the region of maximum mean zonal flow acceleration which is, in this case, the CL.

To summarize, the ozone transport associated with the sudden warming from day 20 to 45 is highly transient with a strong poleward eddy flux during the mean flow deceleration phase. This results in an increase in ozone at high latitudes and a depletion in ozone at lower latitudes. The reacceleration of the mean flow after day 35 which ends the warming produces a reversal of the

poleward eddy flux and a disappearance of the anomaly at high latitudes. The net increase in ozone over the warming cycle is only 0.5 ppm, about 20% of the maximum buildup just after the peak of the warming. The net accumulation of ozone is due to the irreversible transport associated with wave dissipation.

At this point it should be clear that the wave transport of potential vorticity and the wave transport of the tracer are closely linked. The differences in the sign of the transport at the CL (potential vorticity is transported southward during the warming while the constituent is transported northward) is due to the fact that their background gradients with respect to latitude have opposite signs. However, one important difference between the potential vorticity transport and tracer transport is that while the vertical flux of potential vorticity is ignored in quasigeostrophic theory it is retained in the constituent transport equations and could be very important.

b. Chemical transport

The transport of ozone with the 900 m orography and photochemistry is presented in Fig. 17. With the addition of photochemistry there is a general tendency for changes in $\bar{\mu}$ in the upper regions to be significantly reduced. Transport is confined to a thin layer in the conservative region and at the bottom of the transition region.

As with the conservative results discussed above, the buildup of ozone is closely correlated with the CL position in the middle of the channel. The presence of chemistry suppresses the buildup of ozone in and above the transition region. These effects are more clearly exhibited in Fig. 18 which shows the horizontal eddy flux at day 25. The net effect of the poleward eddy transport at the CL is greater because the weaker equatorial return transport is largely countered by the photochemistry (compare with Fig. 14). The photochemistry limits transport largely to the conservative region and the lower part of the transition region. The importance of the CL descending below the transition region is illustrated by noting that the 600 m topography produces only a third of the ozone buildup associated with the 900 m topography.

The total ozone field for the planetary wave circulation alone (Fig. 19) reveals a rapid poleward buildup of ozone coincident with the descent of the CL from high altitudes and depletion of high latitude ozone as the mean flow returns to the prewarming state. Comparison of this total ozone field with

the observations in Fig. 10 shows that the planetary wave transport cannot accurately reproduce the total ozone field. While there is significant change in the total ozone field, the total ozone at high latitudes is much smaller than is observed. This is because the planetary waves are largely redistributing ozone already present within the lower stratosphere and not transporting ozone out of the high altitude source region.

5. Transport by both waves and the diabatic circulation.

The change in the zonal mean ozone field associated with the diabatic circulation added to the 900 m topography dynamics is shown in Figures 20-22. The transport by the combined circulations is much different from the transport by each individual type of circulation indicating that both the wave and diabatic regimes are important in the lower stratosphere.

At the center of the channel (Fig. 20) there is a maximum increase in $\bar{\mu}$ between 25 and 30 km of 1.25 ppm in late March. There is a series of short term decreases around 40 km with the most intense decrease of 1.25 ppm also occurring in late March. The region of steady increase is due to the downward transport by the diabatic circulation. The intermittent nature of the ozone depletion at 40 km is associated with the wave transport. With the combined circulation the increase in ozone at the channel center is reduced by more than 50% when compared to the changes caused by the diabatic circulation alone (Fig. 7).

At the northern latitude (Fig. 21) there is a uniform increase of ozone at all levels with the maximum increase of 3.12 ppm occurring close to 30 km, about 4 km lower than the maximum associated with the planetary waves alone (Fig. 17) and 4 km higher than that associated with the diabatic circulation alone (Fig. 8). An anomaly of 1 ppm is evident at 25 km, which is 10 km higher than a similar anomaly calculated using only the diabatic circulation (compare to Fig. 8). The largest anomaly occurs in early March and is coincident with the lowest altitude of the CL. Following the maximum in early March the positive anomaly remains relatively stable at a maximum value somewhat greater than 2 ppm.

At the southern latitude (Fig. 22) there is a region of relatively strong poleward transport between 35 and 40 km and below this layer is a region of increase. As at the center of the channel, the region of increase is associated with the diabatic circulation and the decrease is associated with the planetary wave. Such large depletions are not observed at this latitude (46°) in the stratosphere, but in the stratosphere easterlies will not generally extend this far south; therefore, such vigorous transport is unrealistic.

The total ozone time series (Fig. 23) shows much better agreement with

observation (Fig. 10) at high latitudes than the total ozone field calculated with the diabatic circulation (Fig. 9). A late winter maximum of almost 440 DU occurs during early March at the most northern latitudes. There is an earlier maximum of slightly more than 420 DU coincident with the initial descent of the CL. Both maxima are associated with the CL descent after the first warming event (Fig. 2). In the southernmost section of the model there is too little ozone from the beginning of the combined planetary wave-diabatic integration (Dec. 15). Furthermore, during the initial descent of the CL there is excessive poleward transport. This excessive transport occurs at the time that the CL descends to low altitudes at the northern and southern points.

The vertical structure of the winter ozone field is not as well simulated as the total ozone distribution. The computed increase of more than 3 ppm at high latitudes is concentrated at 30 km which is larger than observed at this altitude. The increase between 20 and 25 km is significantly smaller than is observed. The increase in ozone at high latitudes is due largely to the downward transport of ozone out of the photochemical region by the diabatic circulation at 70° and the subsequent northward transport of this ozone to high latitudes by the planetary waves.

In the southern region of the channel and to a lesser degree in the middle of the channel the vertical structure is more poorly represented. There is excessive transport out of the south by the planetary waves between 25 and 40 km which leads to unrealistic vertical profiles. This excessive transport is also evidenced by the underestimation of total ozone at low latitudes during mid-January. The excessive poleward transport of ozone is largely caused by the unrealistic descent of the CL at these latitudes, and may also be due to an inadequate representation of chemistry at low altitudes; the underestimation of the transport by the diabatic circulation into the south from either high altitudes or lower latitudes; geometric effects of the β -plane transport; and/or to the neglect of small scale dynamics.

The transports associated with the diabatic and planetary wave circulations while fundamentally different from each other are of equal importance. In the combined circulation study the diabatic circulation is primarily responsible for transport of ozone downward from the photochemical

source region while the planetary wave circulation is responsible for north-south transport. Enough ozone is moved out of the source region to high latitudes to reproduce the total ozone maximum; however, redistribution of ozone within the conservative region is required to accurately reproduce the morphology of ozone. Synoptic scale disturbances are probably instrumental in determining the vertical structure of ozone within the lower stratosphere.

6. Summary and Discussion

At all times during the winter at extratropical latitudes the zonal mean ozone flux due to the diabatic circulation is northward and downward. The diabatic vertical velocity transports ozone from the photochemical and transition regions downward into the upper part of the conservative region. This is the primary mechanism through which the ozone column density increases at high latitudes during winter. Reverse transport takes place during summer and the column density decreases as ozone is brought upward into the photochemical region and destroyed. Meridional advection of ozone by the diabatic circulation appears to be unimportant for the two models studied here.

While the diabatic circulation does produce a high latitude build up of ozone it is not consistent with the climatology of ozone in that the total ozone maximum produced by the diabatic field is too large and occurs too late in the spring. Thus the Brewer-Dobson circulation, as quantitatively computed by Dunkerton (1978), is not a complete explanation of ozone transport. However, much different Brewer-Dobson type circulations than those investigated here could probably be concocted to explain the ozone distribution, but these circulations would be at odds with the radiative - dynamic structure of the winter stratosphere from which the Dunkerton (1978) and Apruzese et al. (1982) circulations are derived.

The effects of planetary waves during sudden warmings is to transport ozone northward during the first half of the warming when wave transience is positive and the CL descends. Ozone is transported southward during the last half of the warming when the wave transience is negative and the mean flow reaccelerates. This transport mechanism is consistent with the idealized CL models of Matsuno and Nakamura (1979) and Schoeberl (1981b).

The transport by strongly transient waves is much more vigorous, confined to a smaller vertical region (mostly around the CL), and is conceptually more complex than the steady-state transport presented in Rood and Schoeberl (1982). For example Fig. 12 shows that the wave transport is highly time dependent and confined to a narrow altitude region. Transport associated with wave transience (in this case the growth and decay of a sudden warming) is completely reversible (Plumb, 1979) which is why the anomaly field shown in Fig. 12 shows distinct pulses. The irreversible ozone transport by waves is

proportional to either dynamical or photochemical "dissipation". Therefore, the amount of ozone which remains at high latitudes after a warming event must be related to the dissipation rate. If the sudden warming cycle is short compared to the dissipation time scale then relatively little ozone will remain at high latitudes. If the other hand there is a high degree of wave variance over the same period then a significant amount of ozone might build up at high latitudes.

As a final point, the transport of ozone by sudden warmings must take place in the conservative region to be effective. Transport by warmings which occur at higher altitudes (minor warmings) is nearly nullified by the zonal mean photochemistry. For example, in the warming with 600 m orography the CL does not descend below 38 km and the polar ozone anomaly is less than one half of that calculated for 900 m forcing.

The effects of midwinter sudden warmings on the ozone distribution suggested above have been observed. For example, Zullig (1973, Fig. 24) shows total ozone north of 40° for three winters. During the year which mid winter major warmings occurred there is almost 10% more ozone than in the winter when no warming occurred. These data, though limited, do indicate a role for midwinter warmings in the formation of the high latitude ozone distribution. Furthermore, the results presented here suggest that the spring maximum in early March is not associated with a midwinter stratospheric warming but with the final warming when there is a reversal of the circulation from winter westerlies to summer easterlies. After this reversal easterlies are present in the stratosphere until autumn so that the reverse transport through negative wave transience would not occur.

To summarize, the model results indicate that both the vertical motion field associated with the diabatic circulation and planetary wave transport during sudden warmings are instrumental in the formation of the polar spring ozone maximum. The diabatic circulation transports ozone out of the high altitude source regions while the planetary waves transport the ozone poleward. Neither circulation appears sufficient by itself to produce the observed structure of the ozone distribution.

Acknowledgements

This research is supported by the National Aeronautics and Space Administration and the Office of Naval Research. We would like to thank Dr. Darrel Strobel for many useful comments.

References

- Ames, W.F., Numerical Methods for Partial Differential Equations, 2nd Ed., Academic Press, New York, pp.365, 1977.
- Apruzese, J.P., M.R. Schoeberl, and D.F. Strobel, Parameterization of IR cooling in a middle atmosphere dynamics model, J. Geophys. Res., in press, 1982.
- Berkofsky, L. and E.A. Bertoni, Mean topographic charts for the entire earth, Bull. Amer. Meteor. Soc., 36, 350-354, 1955.
- Brewer, A.W., Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, Quart. J. Roy. Meteor. Soc. 75, 351-363, 1949.
- Charney, J.G. and P.G. Drazin, Propagation of planetary-scale disturbances from the lower into the upper stratosphere, J. Geophys. Res., 66, 83-109, 1961.
- Clark, J.H.E., A quasi-geostrophic model of the winter stratospheric circulation, Mon. Wea. Rev., 98, 443-461, 1970.
- Cunnold, D., F. Alyea, N. Phillips, and R. Prinn, A three-dimensional dynamical-chemical model of atmospheric ozone, J. Atmos. Sci., 32, 170-194, 1975.
- Cunnold, D.M., F.N. Alyea and R.G. Prinn, Preliminary calculations concerning the maintenance of the zonal mean ozone distribution in the Northern Hemisphere, Pure Appl. Geophys., 118, 329-354, 1980.
- Danielsen, E.F., An objective method for determining the generalized transport tensor for two-dimensional Eulerian models, J. Atmos. Sci., 38, 1319-1339, 1981.
- Dobson, G.M.B., Origin and distribution of polyatomic molecules in the atmosphere, Proc. Roy. Soc. London, A236, 187-193, 1956.
- Dunkerton, T., On the mean meridional mass motions of the stratosphere and mesosphere, J. Atmos. Sci., 35, 2325-2333, 1978.
- Dunkerton, T., A Lagrangian-mean theory of wave, mean-flow interaction with applications to nonacceleration and its breakdown. Rev. Geophys. Space Phys., 18, 387-400, 1980.
- Dütsch, H.U., Atmospheric ozone and ultraviolet radiation, World Survey of Climatology, Vol. 4, D.F. Rex, Ed., Elsevier, 383-432, 1969.

- Fels, S.B., J.D. Mahlman, M.D. Schwarzkopf, and R.W. Sinclair, Stratospheric sensitivity to perturbations in ozone and carbon dioxide: Radiative and dynamic response, J. Atmos. Sci., 37, 2265-2297, 1980.
- Garcia, R.R. and D.L. Hartmann, The role of planetary waves in the maintenance of the zonally averaged ozone distribution of the upper atmosphere, J. Atmos. Sci., 37, 2248-2264, 1980.
- Haltiner, G.J. and R.T. Williams, Numerical Prediction and Dynamic Meteorology, 2nd Ed., John Wiley and Sons, Inc., New York, pp. 477, 1980.
- Hartmann, D.L. and R. Garcia, A mechanistic model of ozone transport by planetary waves in the stratosphere, J. Atmos. Sci., 36, 350-364, 1979.
- Harwood, R.S. and J.A. Pyle, Studies of the ozone budget using a zonal mean circulation model and linearized photochemistry, Quart. J. Roy. Meteor. Soc., 103, 319-343, 1977.
- Hidalgo, H. and P.J. Crutzen, The tropospheric and stratospheric composition perturbed by NO_x emissions of high altitude aircraft, J. Geophys. Res., 82, 5833-5866, 1977.
- Holton, J.R., An advective model for two dimensional transport of stratospheric trace species, J. Geophys. Res., 86, 11989-11994, 1982
- Johnston, H.S. and J. Podolske, Interpretations of stratospheric photochemistry, Rev. Geophys. Space Phys., 16, 491-520, 1978.
- Kreuger, A.J. and R.A. Minzner, A mid-latitude ozone model for the 1975 U.S. Standard Atmosphere, J. Geophys. Res., 81, 4477-4481, 1976.
- Labitzke, K. and B. Goretzki, A catalogue of dynamic parameters describing the variability of the middle stratosphere during the northern winters, in Handbook for MAP, 5, Middle Atmosphere Program, C.F. Sechrist, Ed., 1982.
- Lindzen, R.S., B. Farrel, and K.K. Tung The concept of wave overreflection and its application to baroclinic instability, J. Atmos. Sci., 37, 44-63, 1980.
- London, J., R. Bojkov, S. Oltmans, and J. Kelley, Atlas of the Global Distribution of Total Ozone July 1957-June 1967, NCAR Technical Note, TN113, Boulder, pp. 276, 1976.
- Mahlman, D., H. Levy II, and W.J. Moxim, Three-dimensional tracer structure and behavior as simulated in two ozone precursor experiments, J. Atmos. Sci., 37, 655-685, 1980.

- Mahlman, J.D. and W.M. Moxim, Tracer simulation using a global general circulation model: Results from a midlatitude instantaneous source experiment, J. Atmos. Sci., 35, 1340-1374, 1978.
- Matsuno, T., Lagrangian motion of air parcels in the stratosphere in the presence of planetary waves, Pure Appl. Geophys., 118, 189-216, 1980.
- Matsuno, T. and K. Nakamura, The Eulerian- and Lagrangian-mean meridional circulations in the stratosphere at the time of a sudden warming, J. Atmos. Sci., 36, 640-654, 1979.
- Miller, C., D.L. Filkin, A.J. Owens, M.J. Steed, and J.P. Jesson, A two dimensional model of stratospheric chemistry and transport, J. Geophys. Res., 86, 12039-12065, 1981.
- Murgatroyd, R.J. and F. Singleton, Possible meridional circulation in the stratosphere and mesosphere, Quart. J. Roy. Meteor. Soc., 87, 125-135, 1961.
- Plumb, R.A., Eddy fluxes of conserved quantities by small-amplitude waves, J. Atmos. Sci., 36, 1699-1704, 1979.
- Pyle, J.A., and C.F. Rogers, Stratospheric transport by stationary planetary wave-The importance of chemical processes, Quart. J. Roy. Meteor. Soc., 106, 421-446, 1980a.
- Rood, R.B. and M.R. Schoeberl, A mechanistic model of Eulerian, Lagrangian-mean, and Lagrangian ozone transport by steady planetary waves, to appear in J. Geophys. Res., 1982.
- Schlesinger, M.E. and Y. Mintz, Numerical simulation of ozone production, transport and distribution with a global atmospheric general circulation model, J. Atmos. Sci., 36, 1325-1361, 1979.
- Schoeberl, M.R., A simple model of the Lagrangian-mean flow produced by dissipating planetary waves, J. Atmos. Sci., 38, 1841-1855, 1981a.
- Schoeberl, M.R., The secondary flow near a baroclinic planetary wave critical line, J. Atmos. Sci., 38, 630-638, 1981b.
- Schoeberl, M.R., A study of stratospheric vacillations and sudden warmings on a β -plane Part I: Single wave-mean flow interaction, to appear in J. Atmos. Sci., 1982a.
- Schoeberl, M.R., Wave-mean flow statistics, J. Atmos. Sci., 39, 2363-2368, 1982b.

Strobel, D.F., Parameterization of linear wave chemical transport in planetary atmospheres by eddy diffusion, J. Geophys. Res., 86, 9806-9810, 1981.

Zullig, W., Relation between the intensity of the stratospheric circumpolar vortex and the accumulation of ozone in the winter hemisphere, Pure Appl. Geophys., 106-108, 1544-1552, 1973.

Figure Captions

- Figure 1. Initial mean zonal wind profile for the time dependent β -plane model.
- Figure 2. Zonal mean wind for 900 m topography at the channel center (m sec^{-1}).
- Figure 3. Geopotential amplitude for 900 m topography at the channel center (gpdm).
- Figure 4. Diabatic circulation at the winter solstice (cm sec^{-1})
a. meridional component
b. vertical component
- Figure 5. Schematic representation of the atmosphere divided into photochemical, transitional, and conservative regions based on the ratio of the dynamical to the photochemical time scale.
- Figure 6. Initial ozone profile (ppm)
- Figure 7. Change in ozone at the center of the channel (60°) for transport by the diabatic circulation alone (ppm).
- Figure 8. Same as Fig. 7 but at 74° .
- Figure 9. Latitude-time cross section of total ozone as calculated with the diabatic circulation (DU).
- Figure 10. 10 year average latitude-time cross section of total ozone in Dobson Units (DU) (from London et al., 1976).
- Figure 11. Change in ozone (anomaly) at 74° with horizontally varying chemistry (ppm)
a. cosine of latitude dependence
b. $\lambda = 0$ north of 67°
- Figure 12. Ozone anomaly for 900 m topography at 74° with no chemistry (ppm)
(bold solid line is the critical line at the latitude of the figure)
(bold dashed line is the critical line at the channel center)

Figure 13. Stream functions for a horizontal critical line (from Matsuno and Nakamura, 1979).

- a. Lagrangian-mean
- b. Eulerian mean

Figure 14. Horizontal eddy flux at day 25 for conservative transport (ppm m sec^{-1}).

(bold solid line is the critical line)

Figure 15. Eulerian mean flux at day 25 (ppm m sec^{-1})

- a. horizontal
- b. vertical

Figure 16. Same as Fig. 14 but at day 47.

Figure 17. Same as Fig. 12 but with chemistry.

Figure 18. Same as Fig. 14 but for chemical transport.

Figure 19. Time-latitude cross section of total ozone using 900 m topography planetary waves alone (DU).

Figure 20. Ozone anomaly for the combined planetary wave and diabatic circulations with chemistry at 60° (ppm).

Figure 21. Same as Fig. 20 but at 74° .

Figure 22. Same as Fig. 20 but at 46° .

Figure 23. Time-latitude cross section of total ozone as calculated with the combined planetary wave and diabatic circulations (DU).

Figure 24. Mean total ozone north of 40° (from Zullig, 1973).

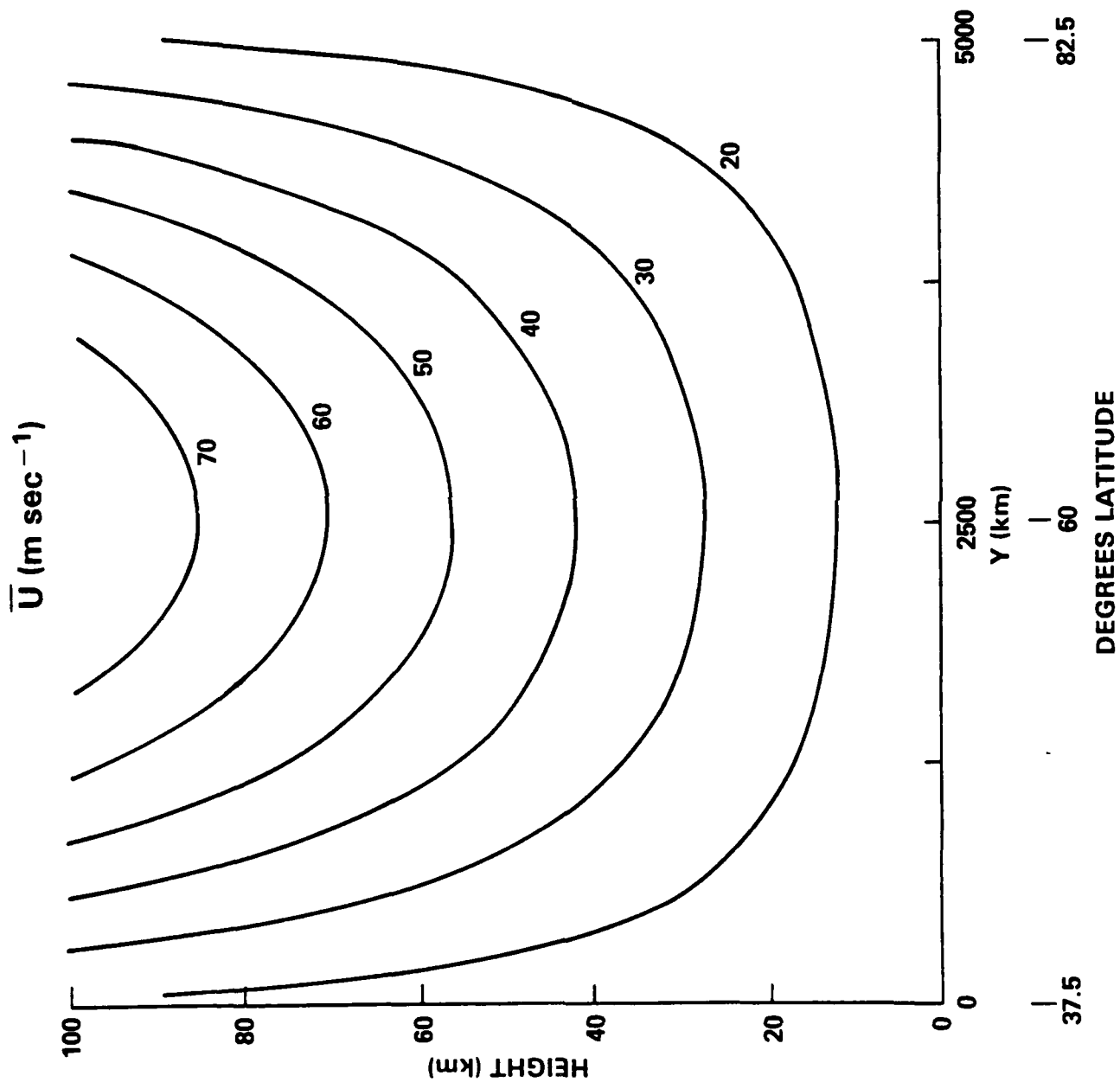
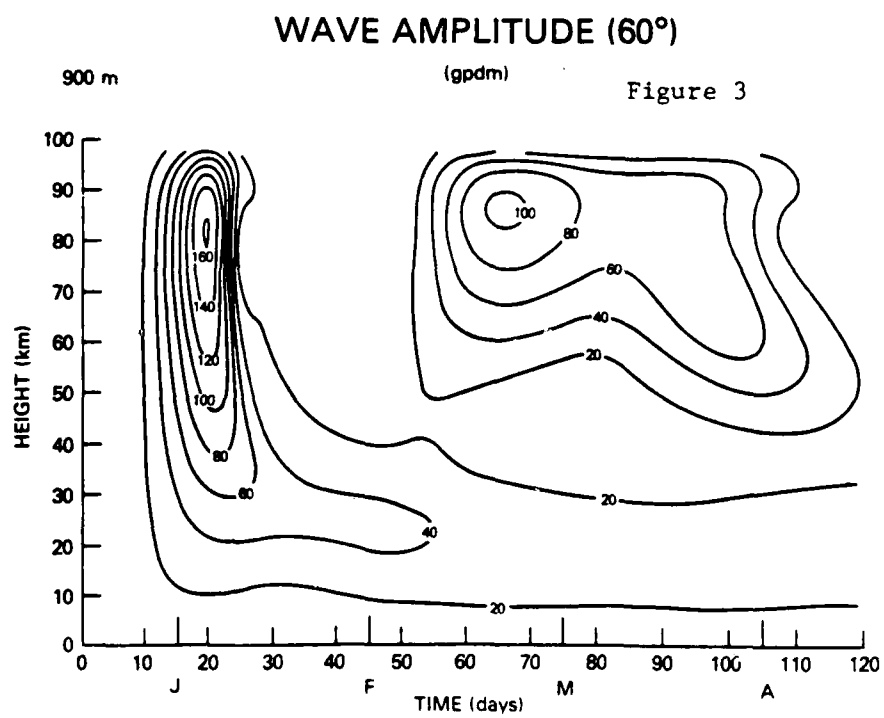
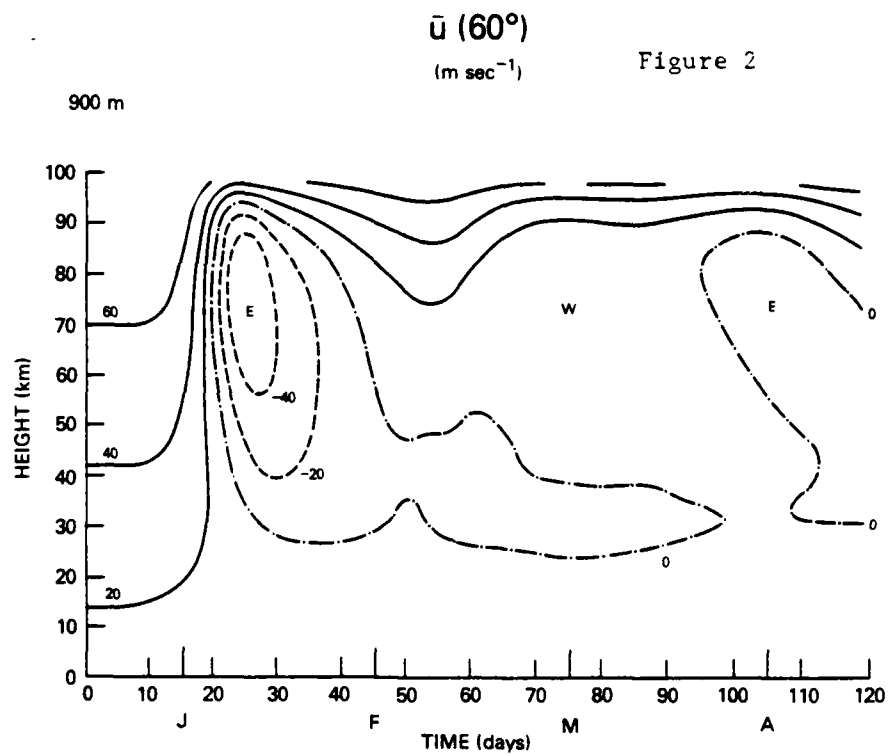


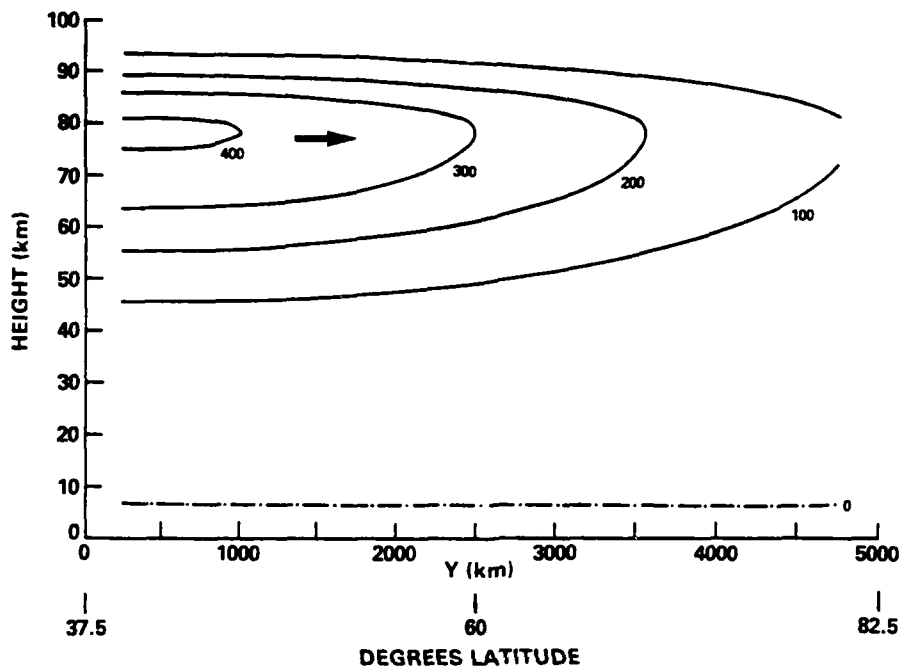
Figure 1



$\bar{V}$
(cm sec⁻¹)

DIABATIC

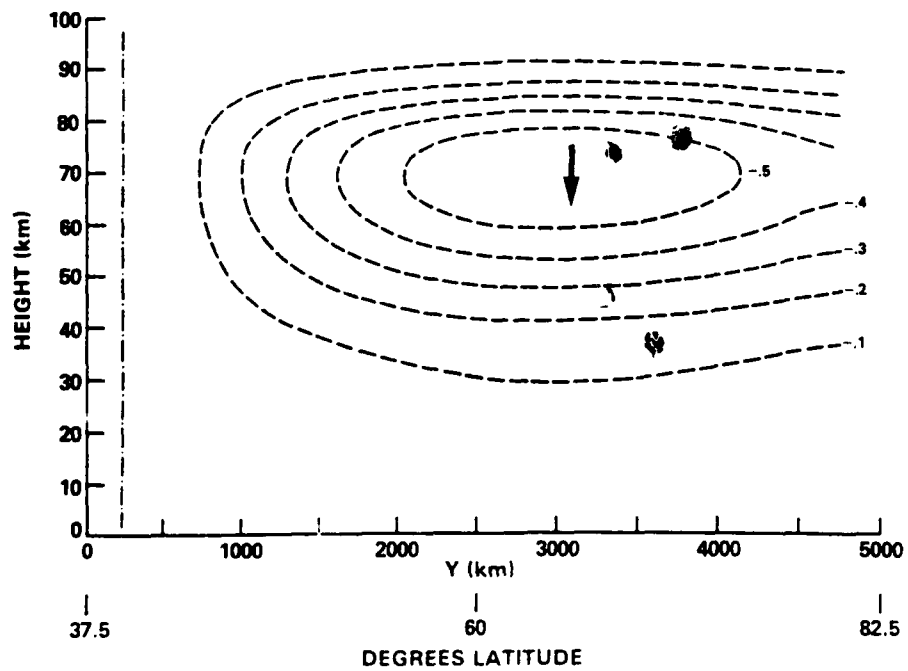
Figure 4a



$\bar{W}$
(cm sec⁻¹)

DIABATIC

Figure 4b



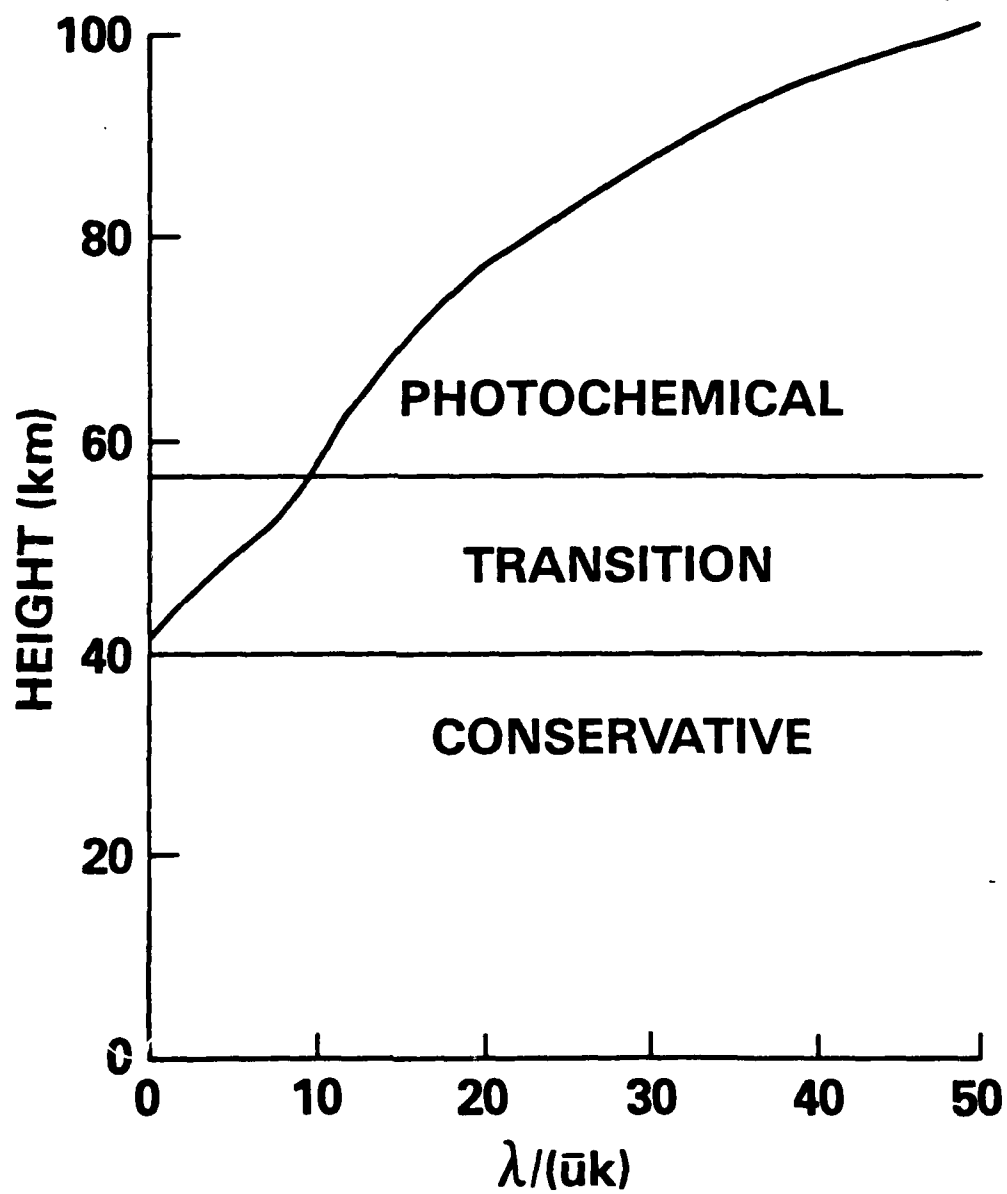


Figure 5

OZONE

(ppm)

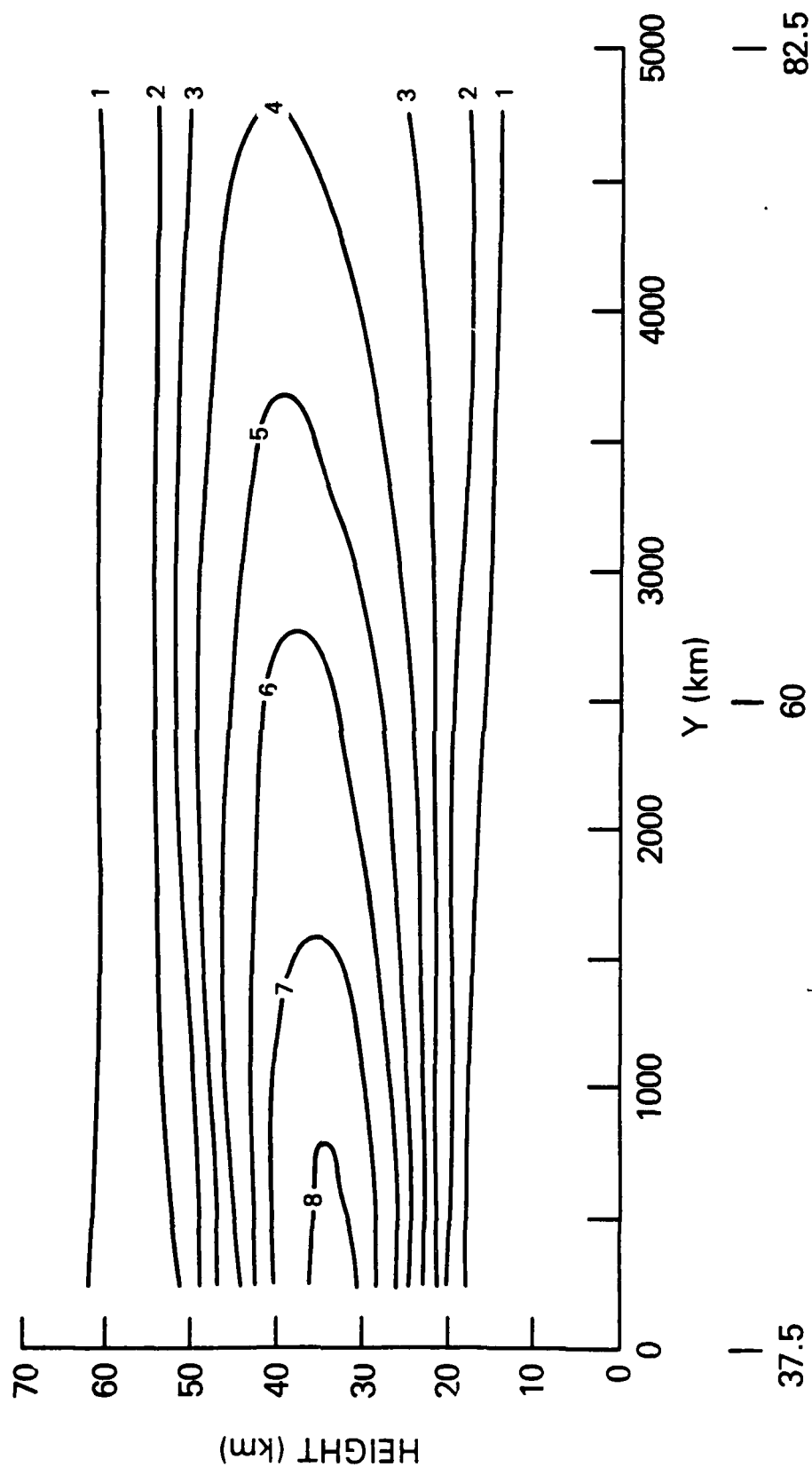


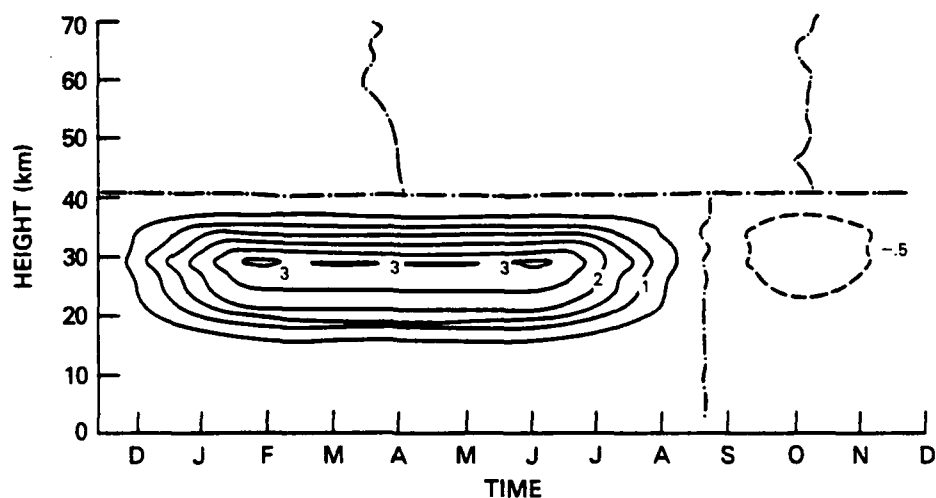
Figure 6

CHANGE IN $\bar{\mu}$ (60°)

(ppm)

Figure 7

DIABATIC, CHEMISTRY

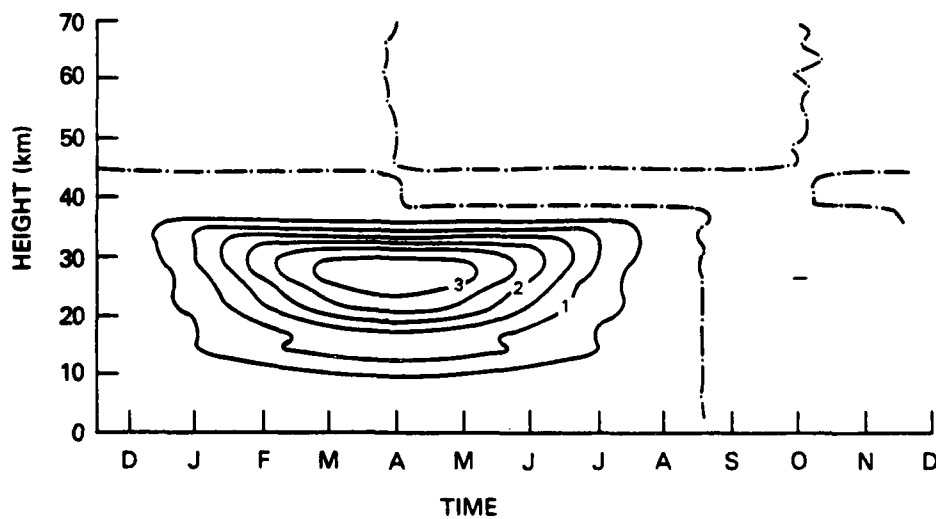


CHANGE IN $\bar{\mu}$ (74°)

(ppm)

Figure 8

DIABATIC



TOTAL OZONE

D.U.

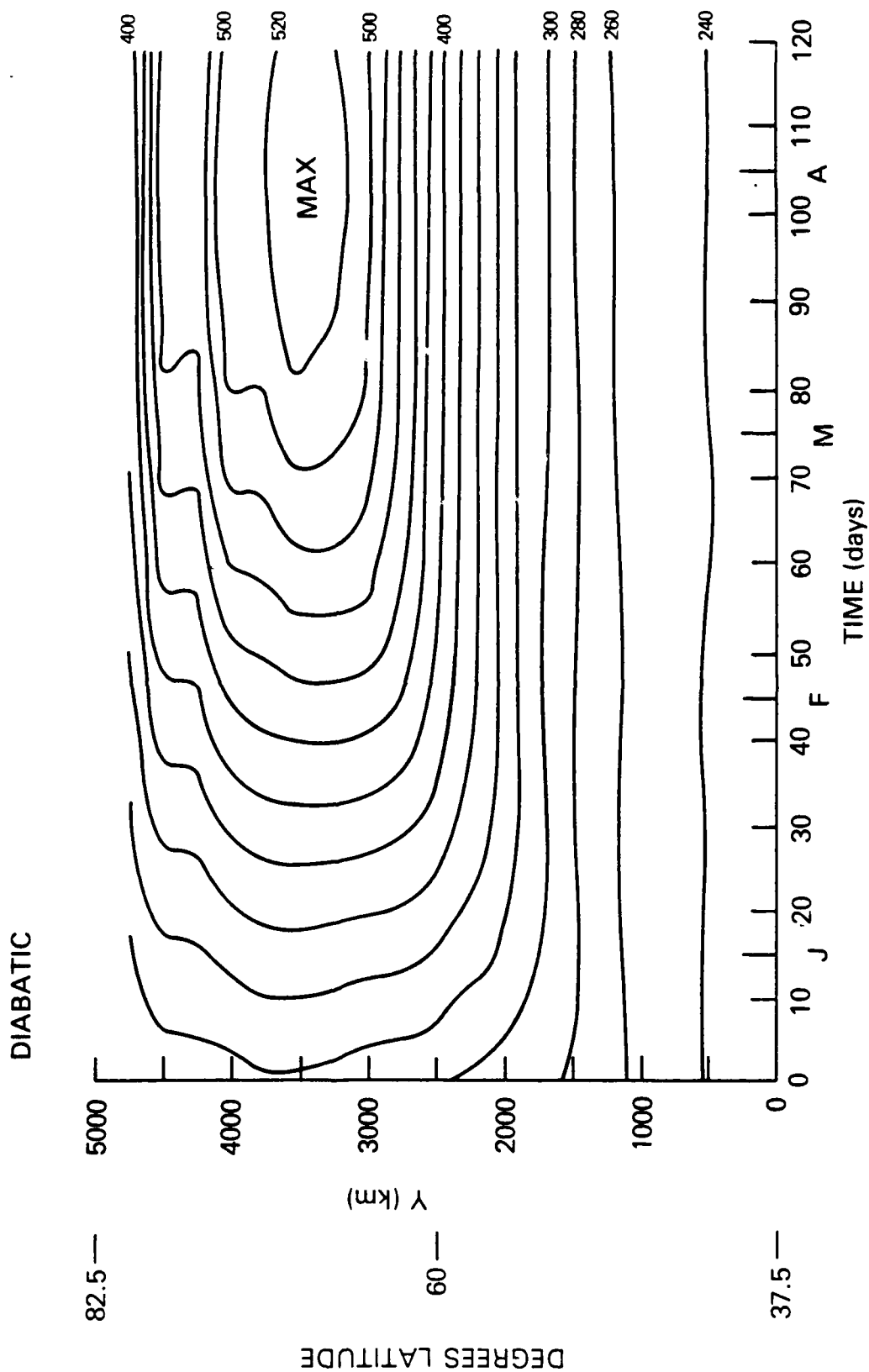


Figure 9

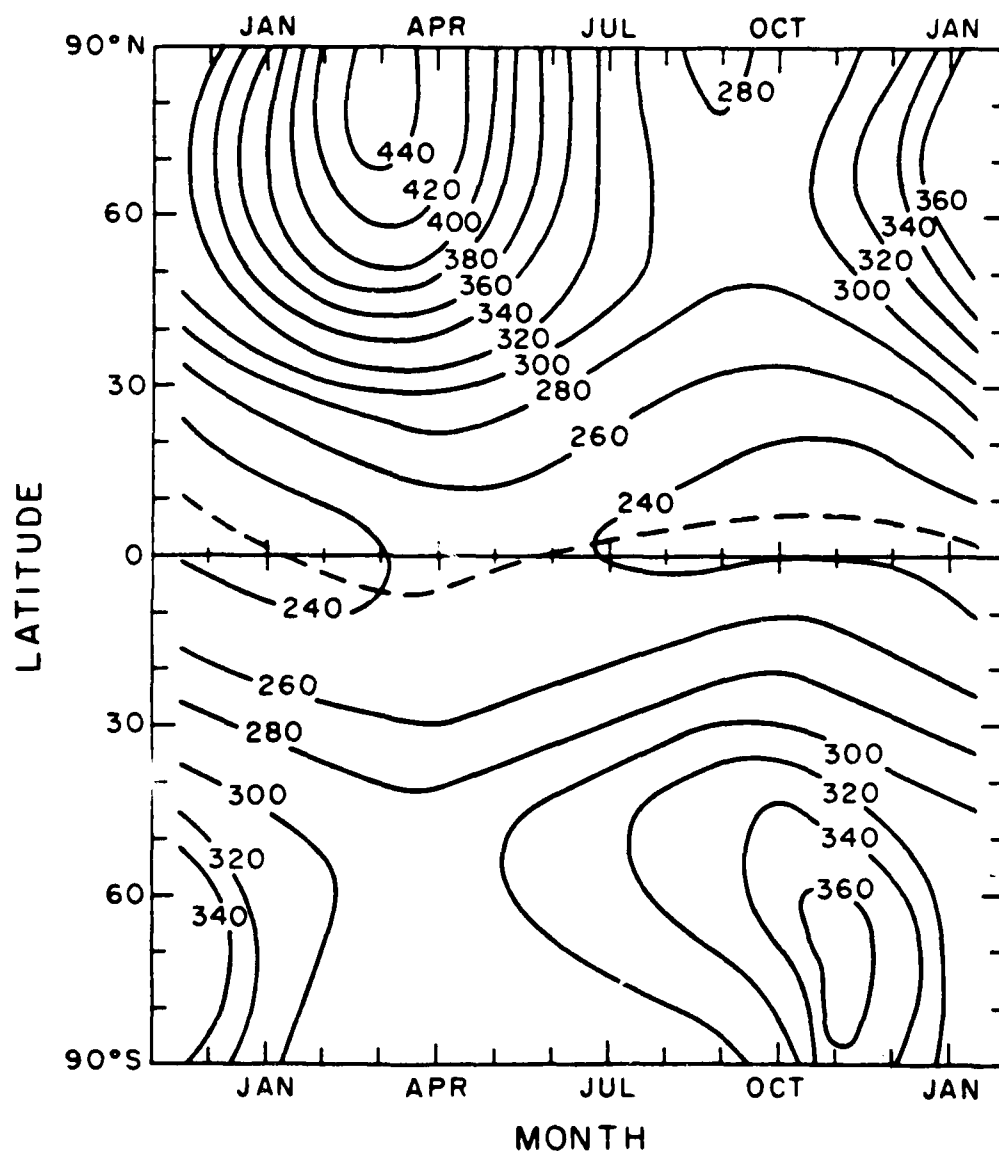


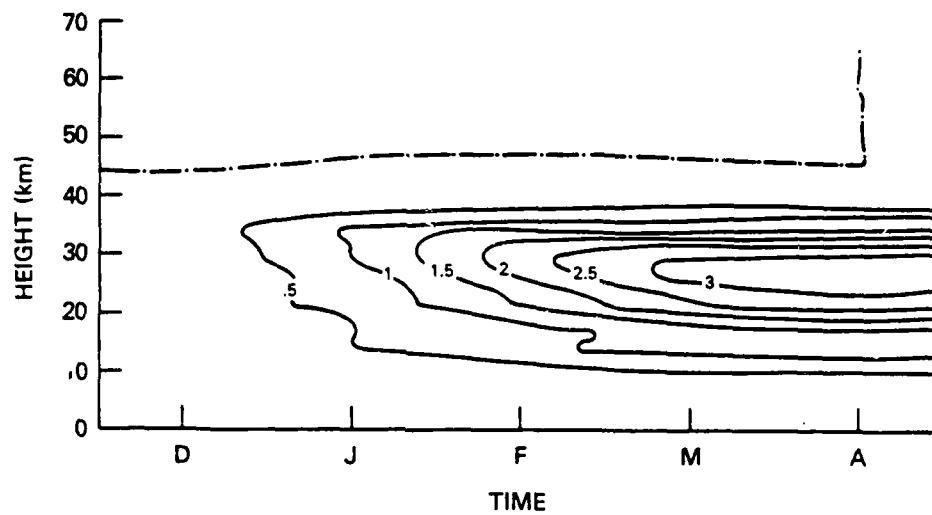
Figure 10

CHANGE IN $\bar{\mu}$ (74°)

(ppm)

Figure 11a

DIABATIC, $\cos(\text{LAT})$

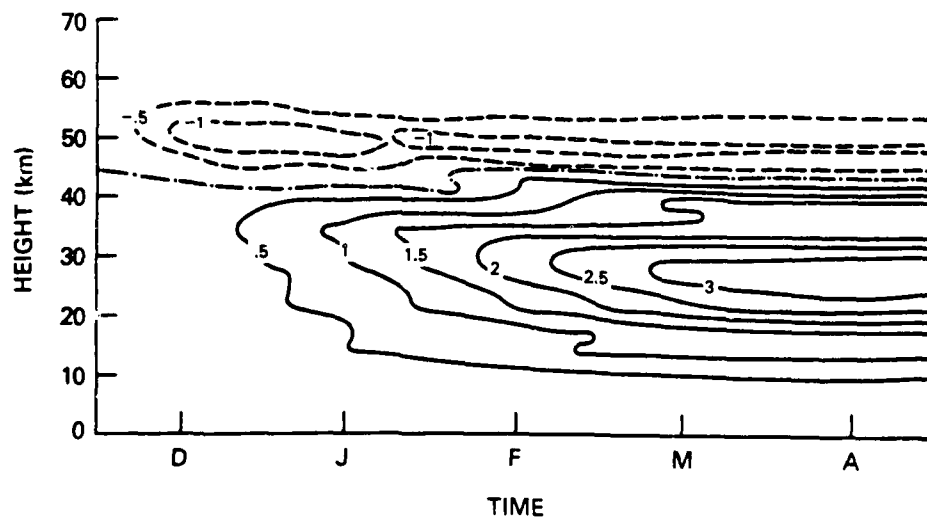


CHANGE IN $\bar{\mu}$ (74°)

(ppm)

Figure 11b

DIABATIC, POLAR NIGHT



CHANGE IN $\bar{\mu}$ (74°) (ppm)

900 m, CONSERVATIVE

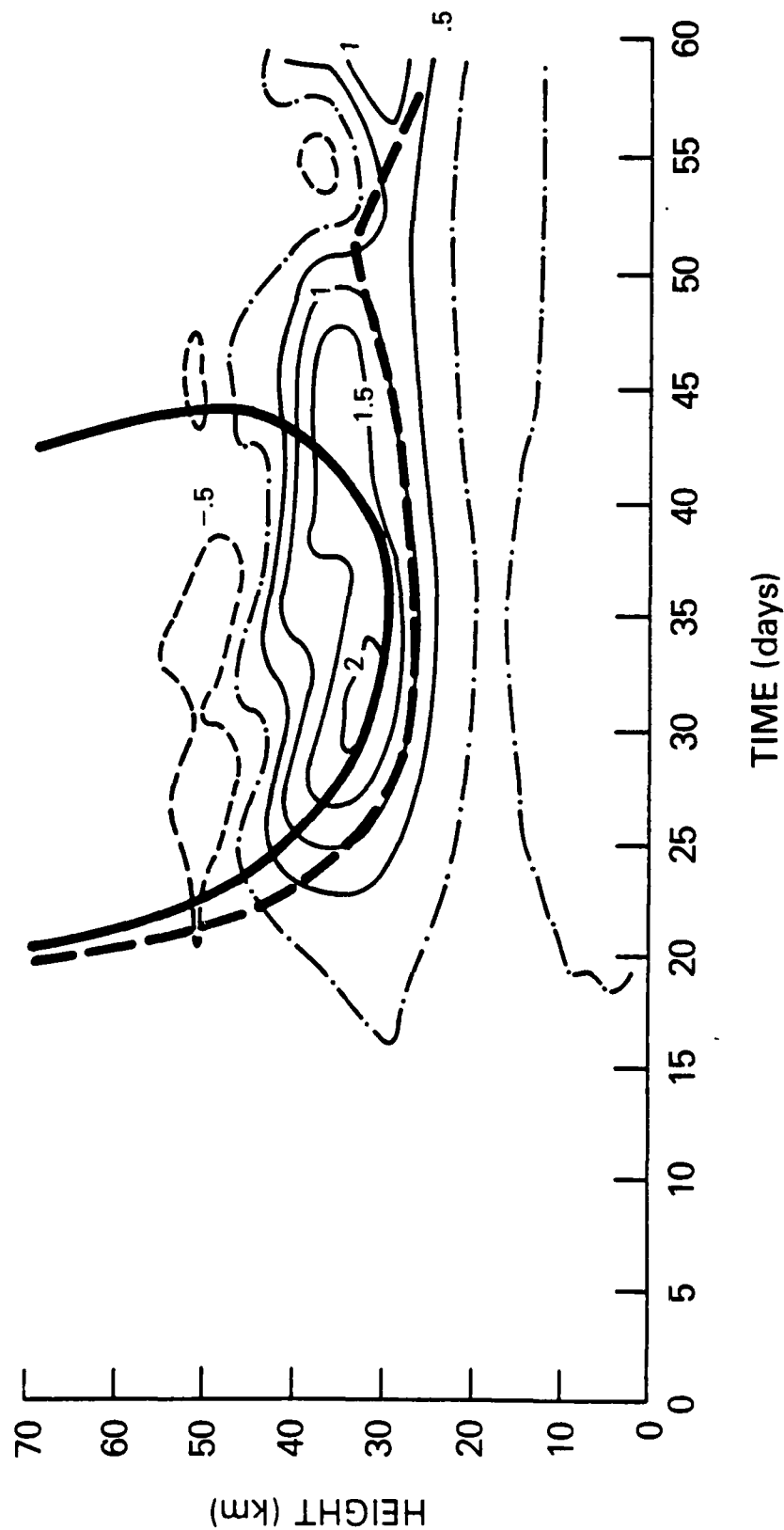


Figure 12

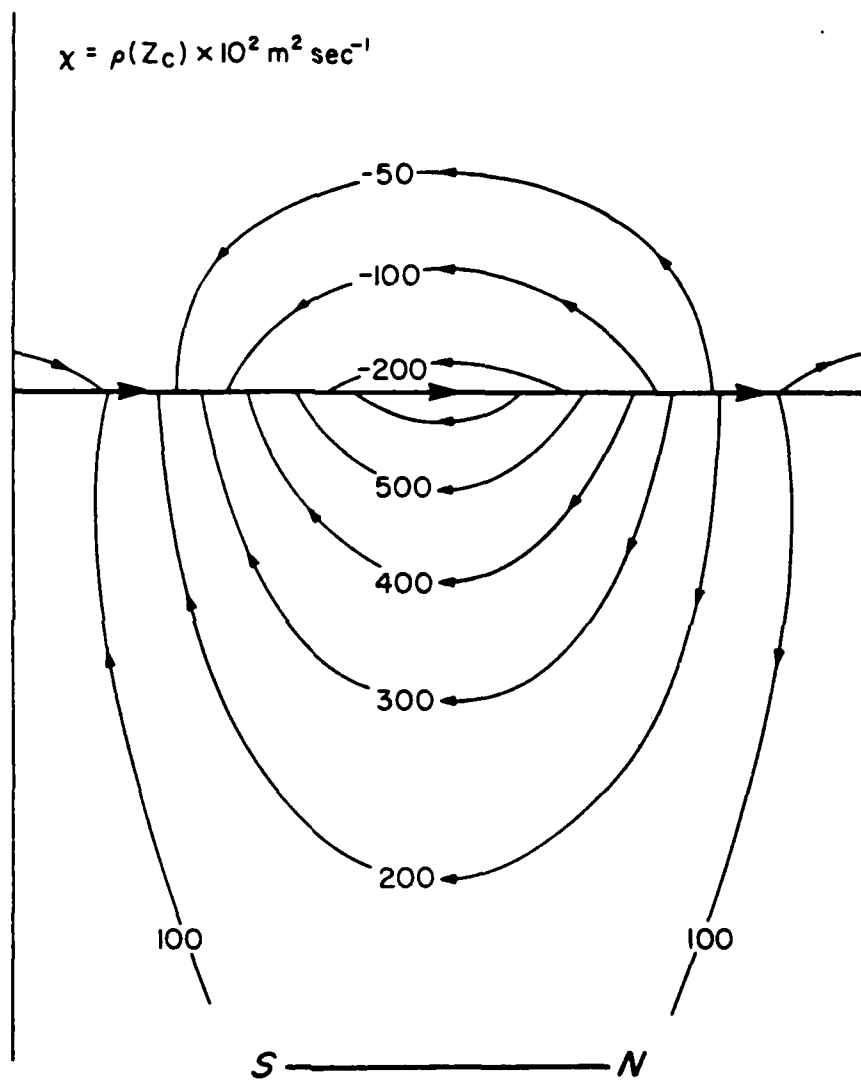


Figure 13a

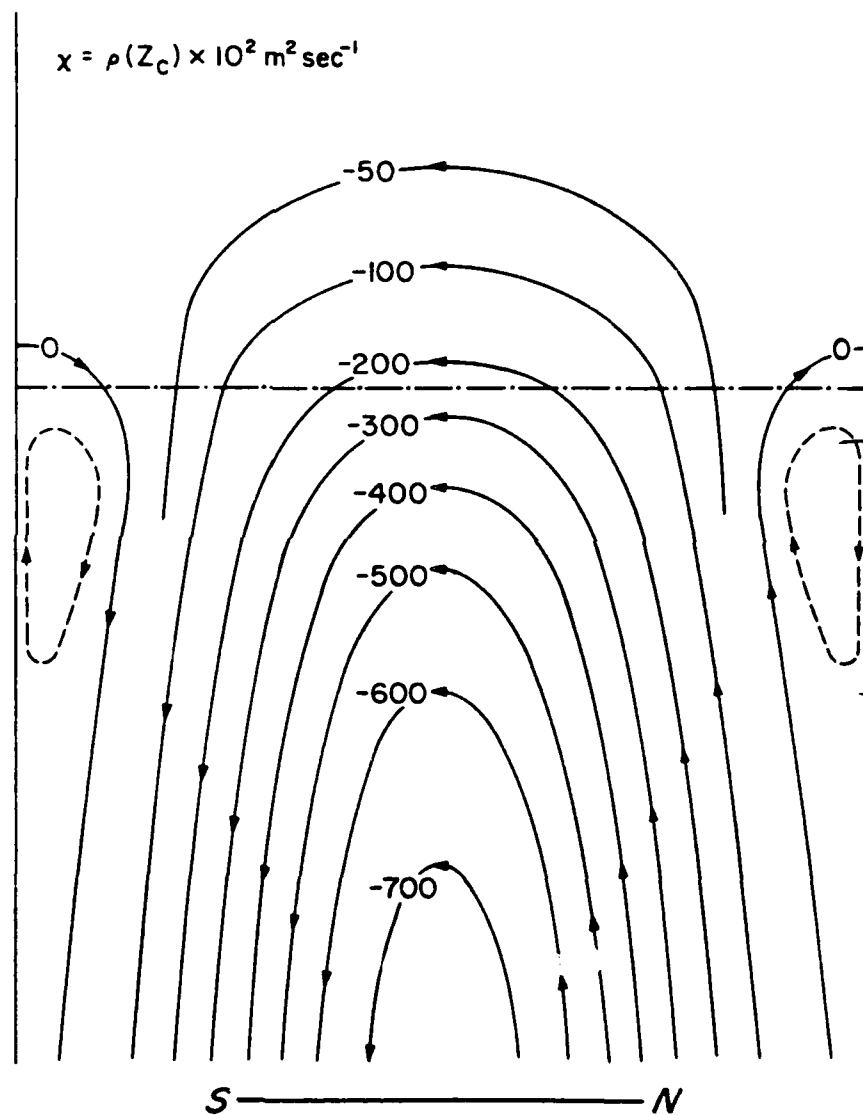
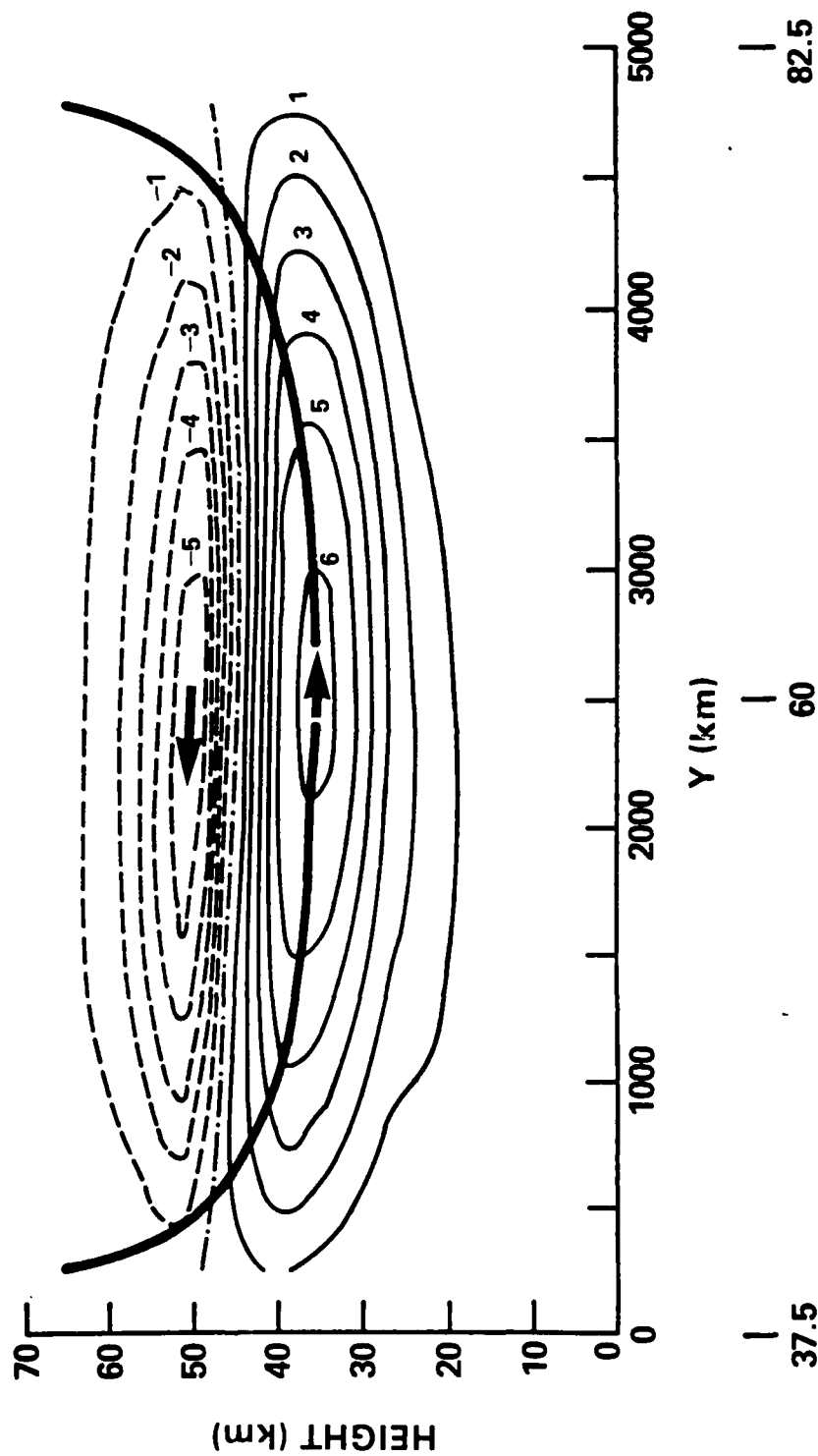


Figure 13b

$\overline{v' \mu'}$
(ppm m sec⁻¹)

DAY 25, CONSERVATIVE



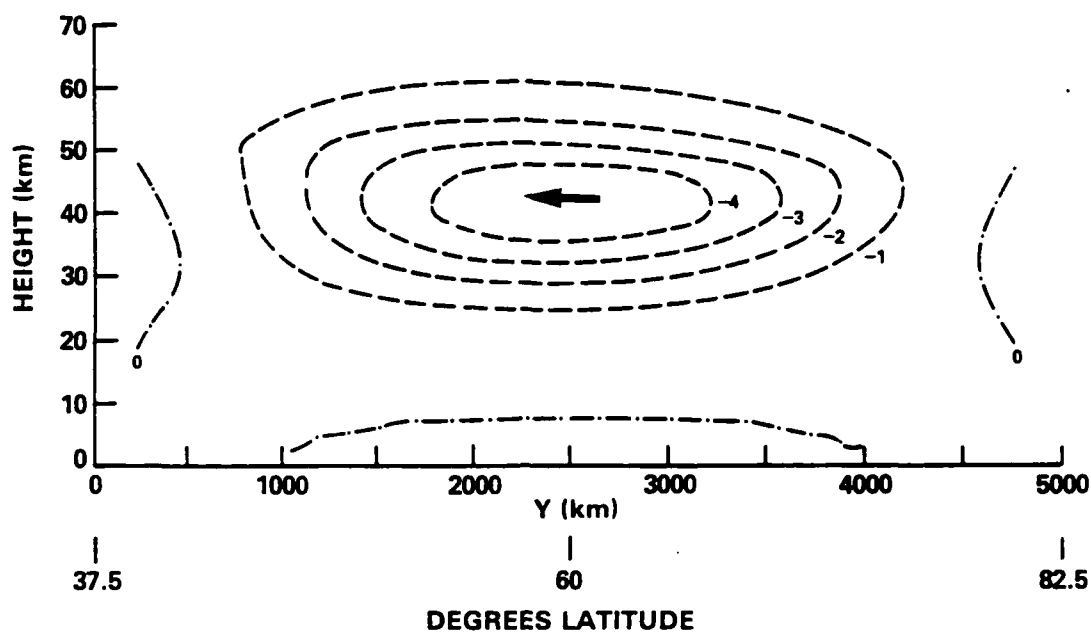
R-45

Figure 14

DAY 25

$\bar{v} \bar{\mu}$
(ppm m sec⁻¹)

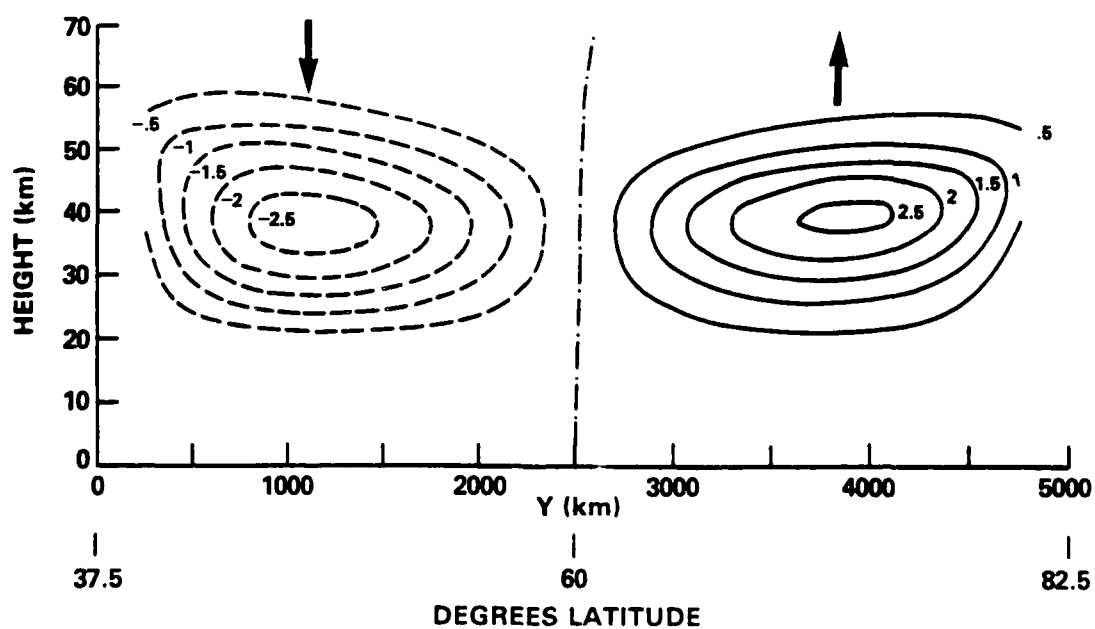
Figure 15a



DAY 25

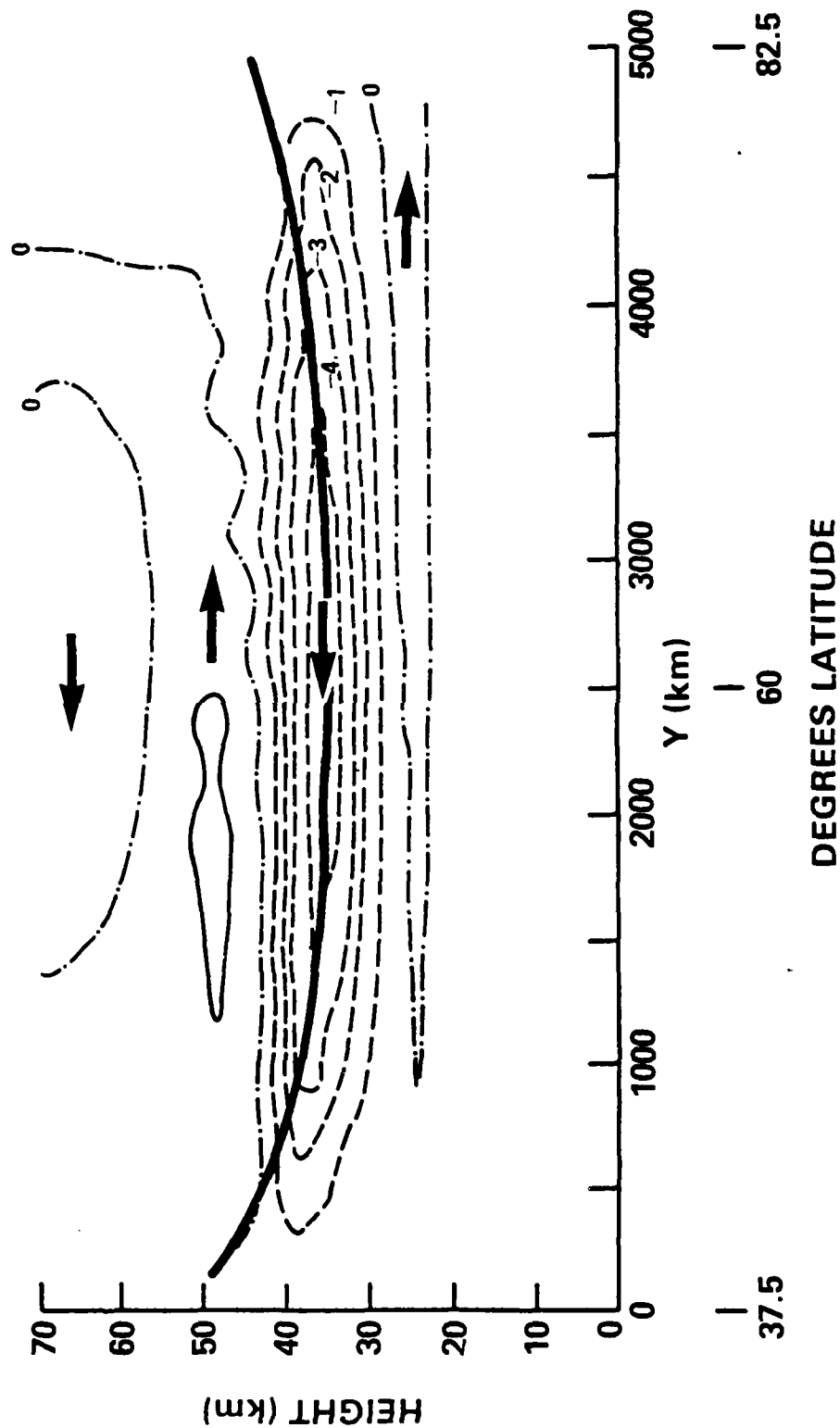
$\bar{w} \bar{\mu}$
ppm cm sec⁻¹

Figure 15b



$\overline{v'\mu'}$
(ppm m sec⁻¹)

DAY 47, CONSERVATIVE



R-47

Figure 16

CHANGE IN $\bar{\mu}$ (74°) (ppm)

900 m, CHEMISTRY

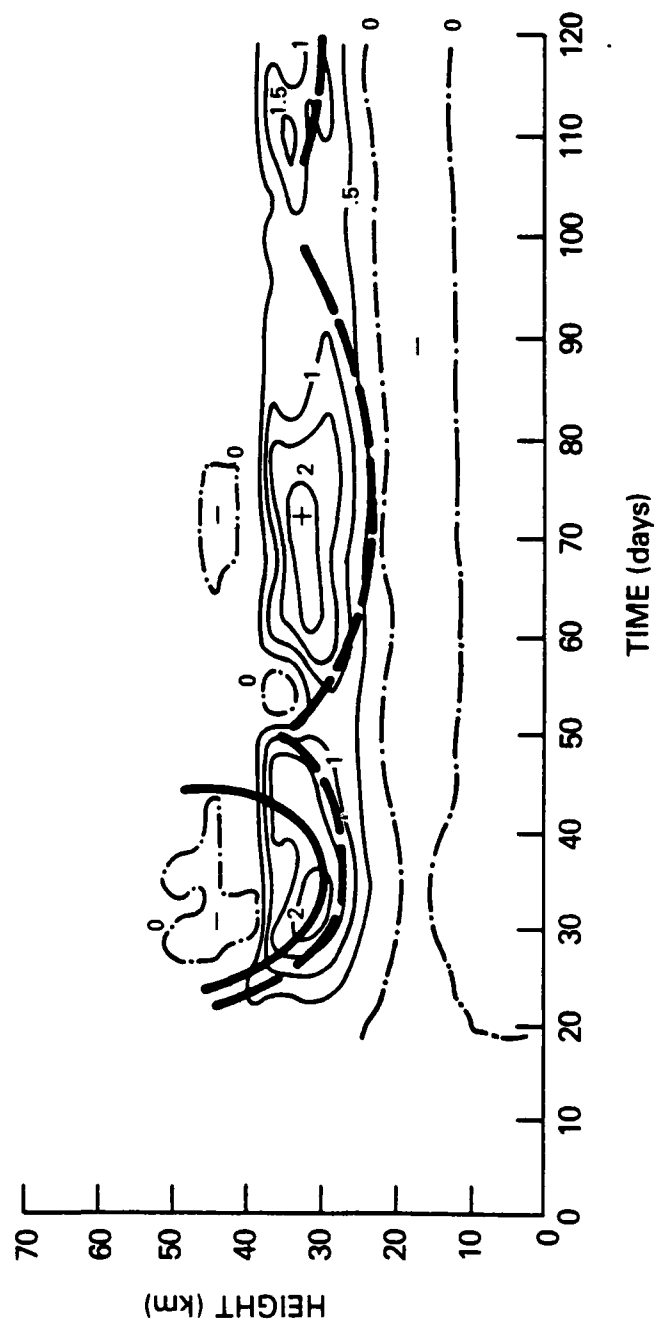


Figure 17

$\overline{v' \mu'}$
(ppm m sec⁻¹)

DAY 25, CHEMISTRY

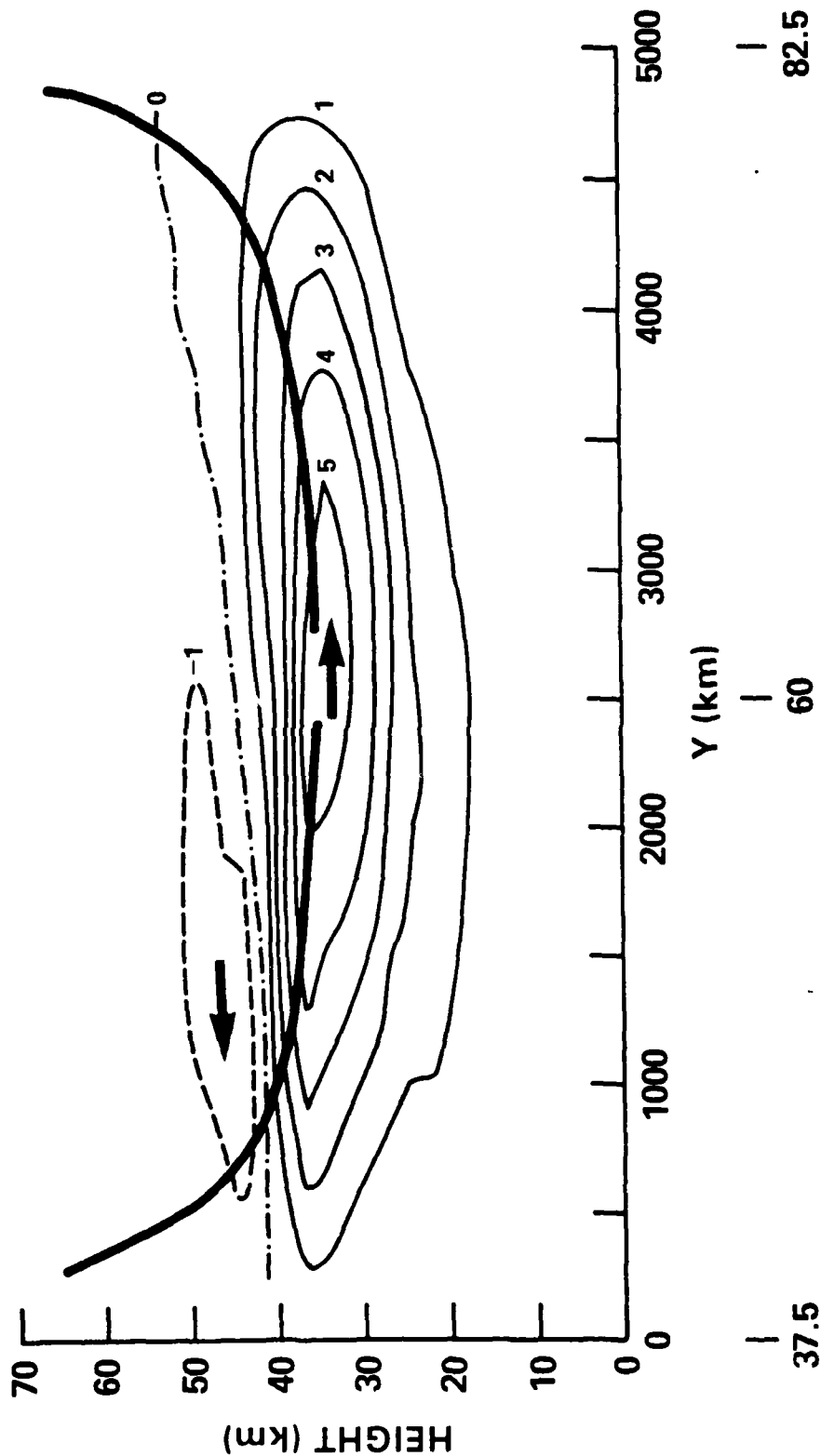


Figure 13

TOTAL OZONE

D.U.

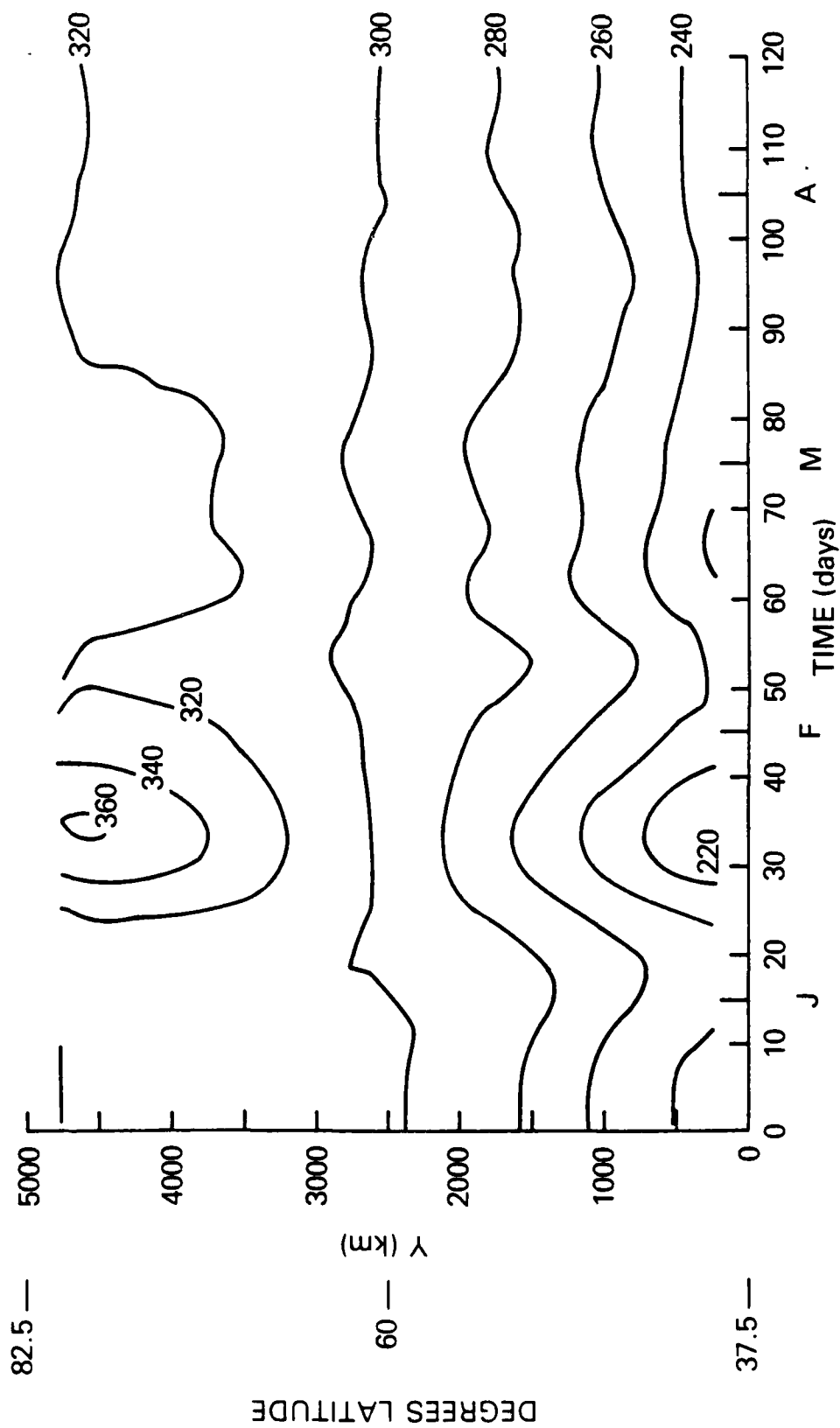


Figure 19

CHANGE IN $\bar{\mu}$ (60°) (ppm)

COMBINED

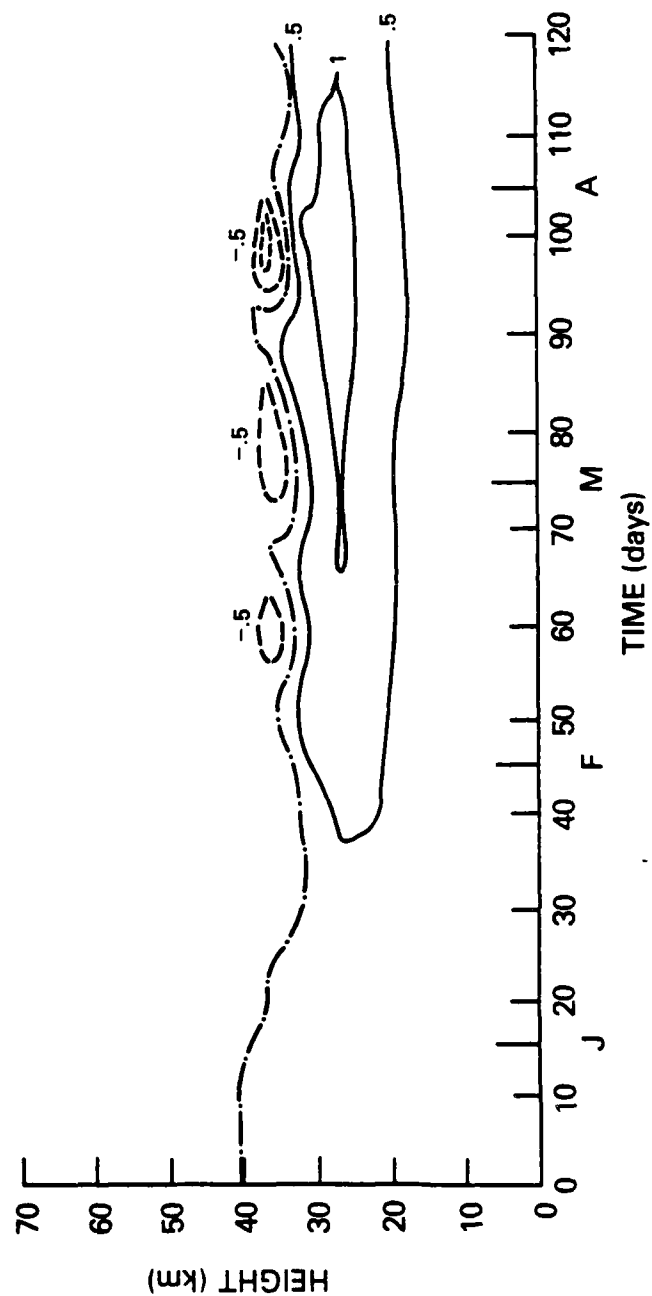


Figure 20

CHANGE IN $\bar{\mu}$ (74°) (ppm)

COMBINED

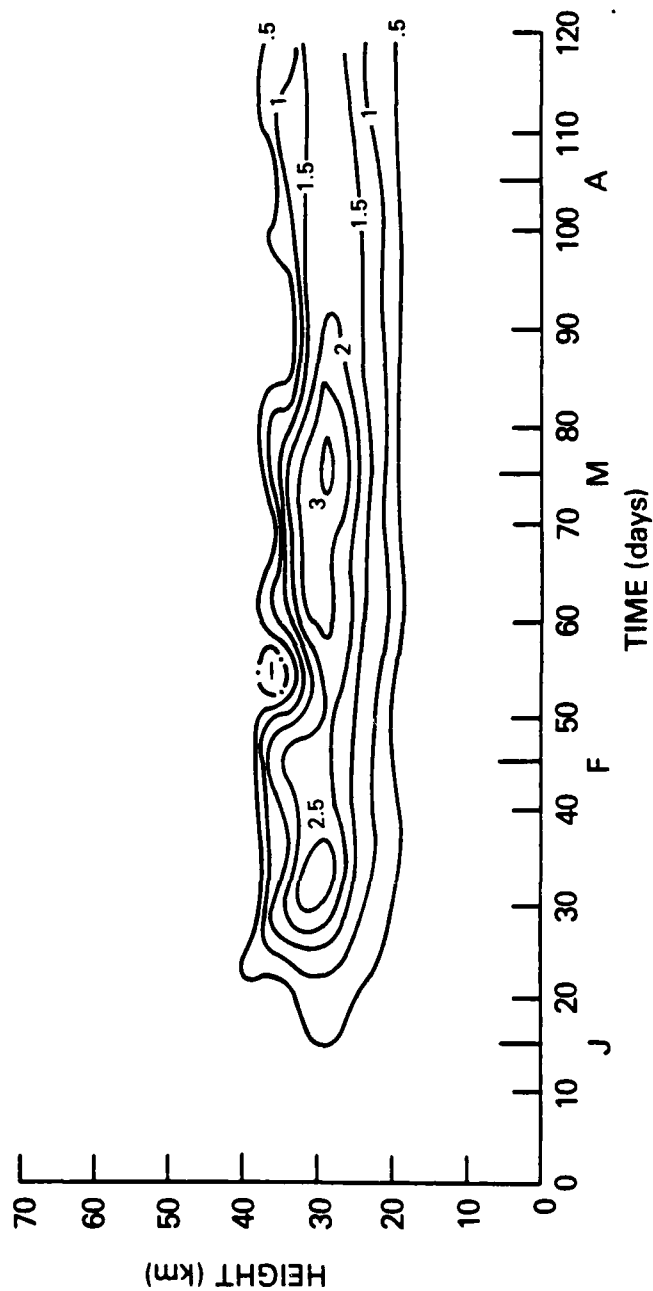


Figure 21

COMBINED

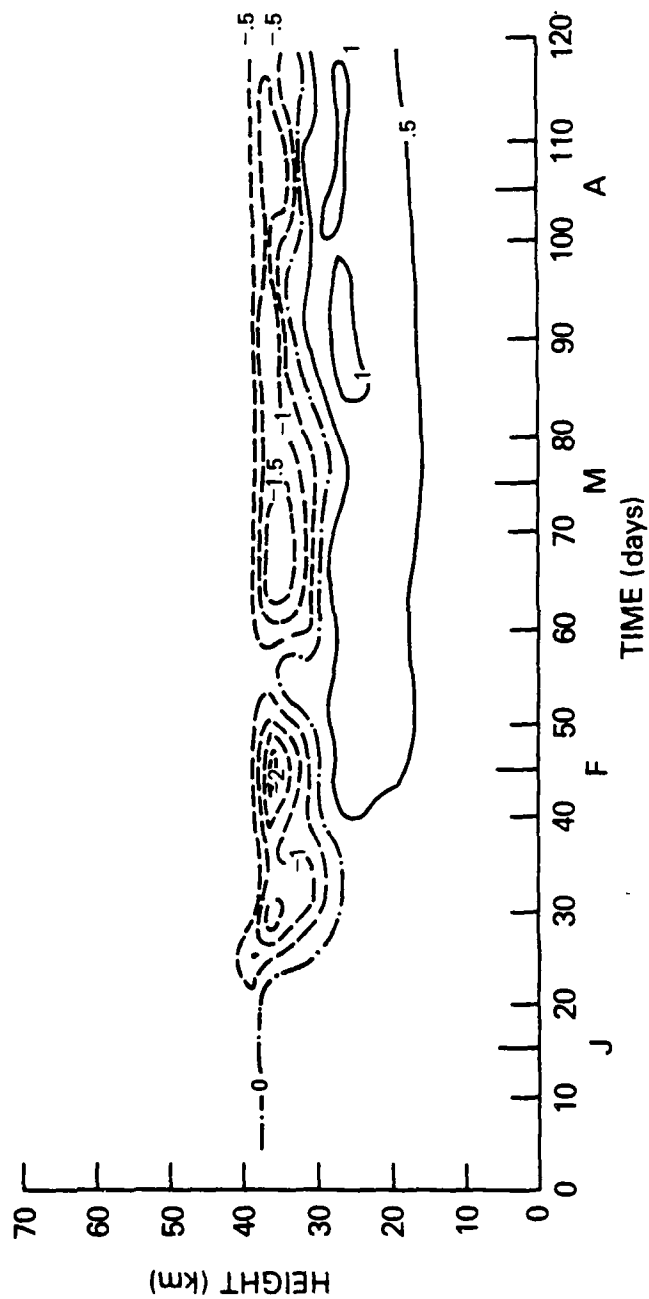


Figure 22

TOTAL OZONE

D.U.

COMBINED

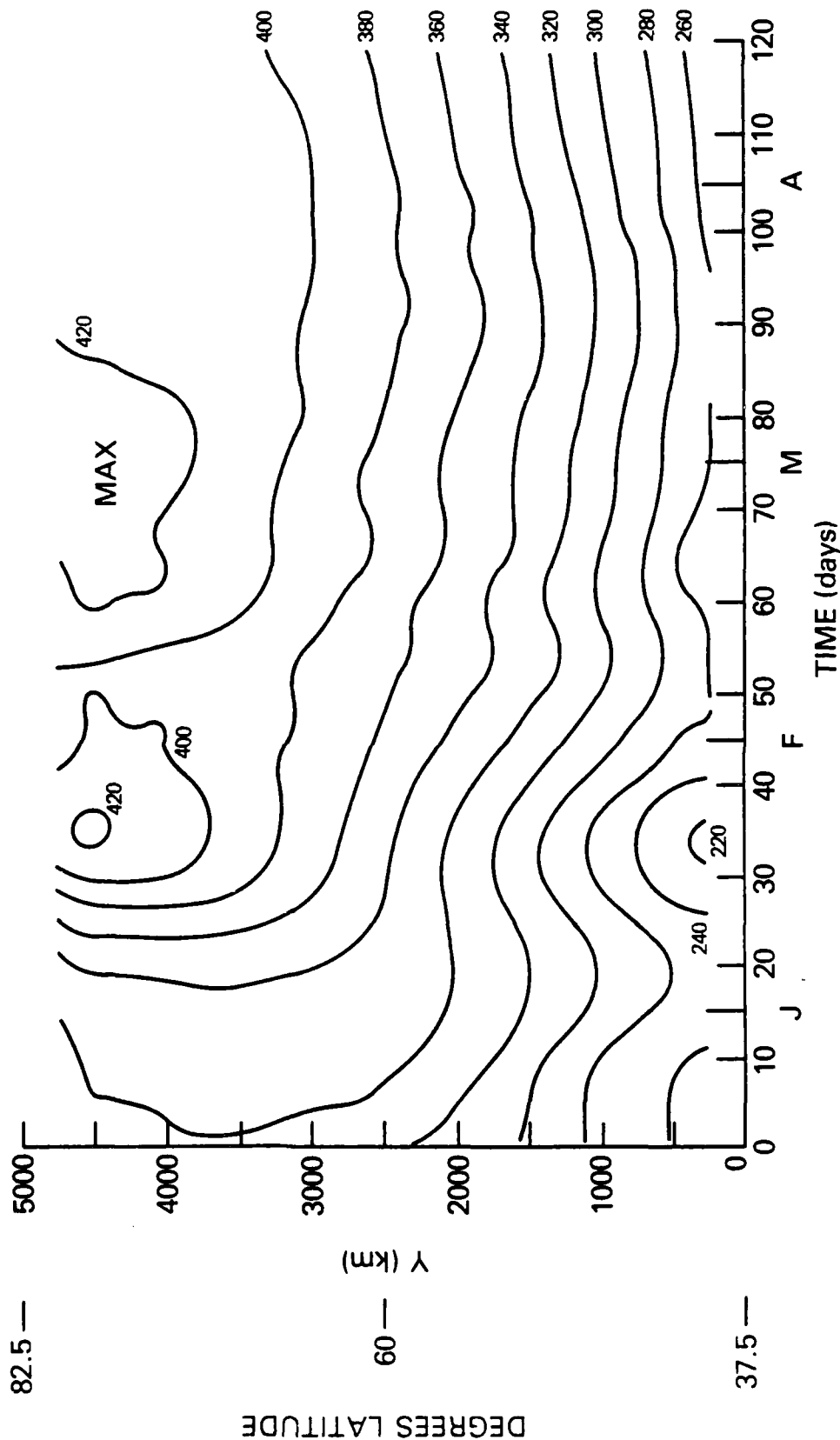


Figure 23

MEAN TOTAL OZONE NORTH OF 40°

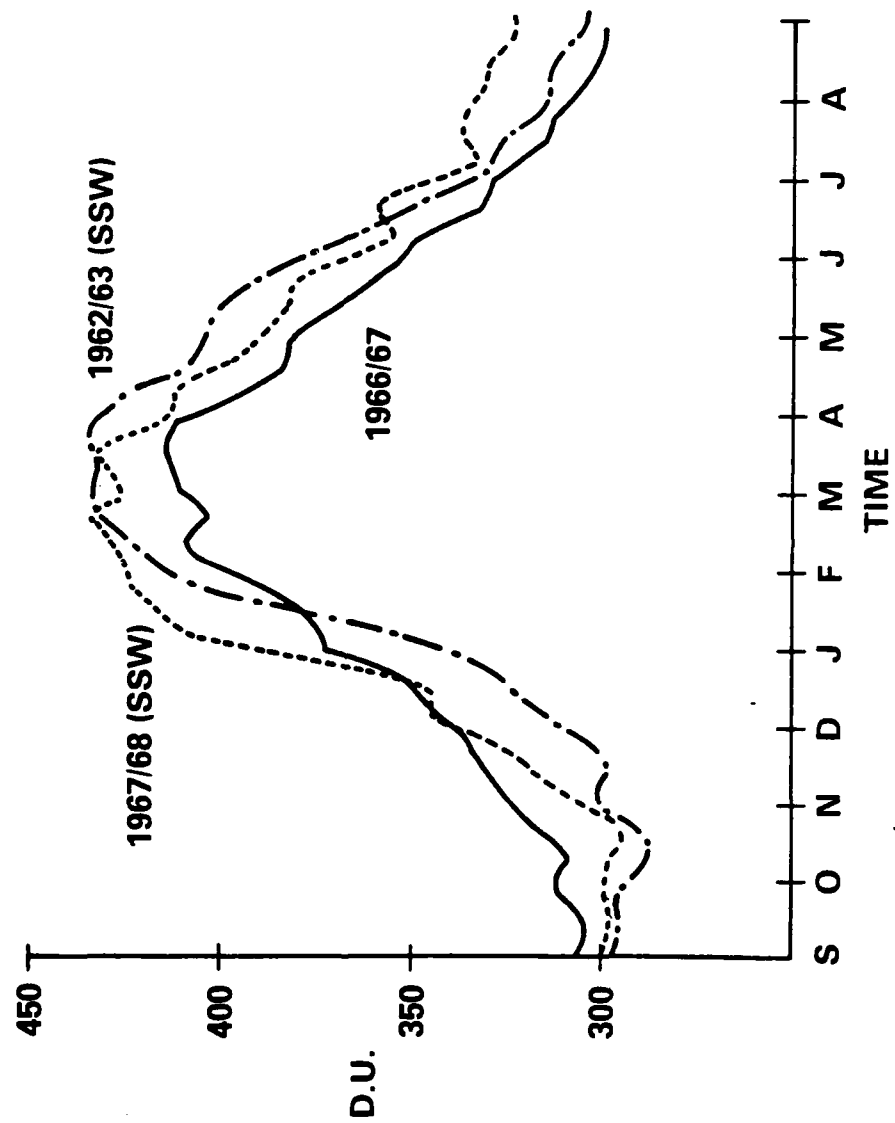


Figure 24

APPENDIX S

ANOTHER NOTE ON FINITE DIFFERENCING OF THE
ADVECTION-DIFFUSION EQUATION

Another Note on Finite Differencing of the
Advection-Diffusion Equation

Richard B. Rood

Science Applications, Inc.
1710 Goodridge Drive
McLean, Virginia 22102*

November 1982

* Current Address
Code 964 NASA/GSFC
Greenbelt, Maryland 20771

Abstract

Clancy (1981) derives stability criteria for forward-in-time, centered-in-space finite difference integration of the advection-diffusion equation. It is shown here that it is more efficient and physically realistic to use a time scheme centered with respect to the advective term and forward with respect to the diffusive term. The time scheme proposed by Clancy is inherently unstable, and the stability criterion derived in Clancy (1981) locates the smallest value of diffusion that stabilizes the scheme.

Clancy (1981) derives stability criteria for the forward-in-time, centered-in-space, finite difference solution to the advection-diffusion equation. These stability criteria are "less restrictive" than those given earlier by Fromm (1964) and Roache (1976) and would allow for the use of this finite difference scheme in a broad range of problems. In this note it will be shown that for the values of the parameters considered by Clancy (1981) it is more efficient to integrate the advection-diffusion equation with centered time differences.

The advection-diffusion equation for a quantity S is

$$\frac{\partial S}{\partial t} + u \frac{\partial S}{\partial x} = D \frac{\partial^2 S}{\partial x^2} \quad (1)$$

where t is time, x is a spatial coordinate, u is velocity in the x direction, and D is a diffusion coefficient. Using second order, centered time differences and generalized spatial differences (1) can be written in discrete form as

$$\begin{aligned} S_i^{n+1} = & S_i^{n-1} - \frac{2u \Delta t}{\Delta x} \{A_1(S_{i+1}^n - S_{i-1}^n) + B_1(S_{i+2}^n - S_{i-2}^n)\} \\ & + \frac{2D \Delta t}{\Delta x^2} \{A_2(S_{i+1}^{n-1} - 2S_i^{n-1} + S_{i-1}^{n-1}) + B_2(S_{i+2}^{n-1} - 2S_i^{n-1} + S_{i-2}^{n-1})\} \end{aligned} \quad (2)$$

where Δx is the spatial grid increment and Δt is the time step. For stability, the diffusive term is written so that it appears to be forward in time with a time step of $2\Delta t$ (Haltiner and Williams, 1980). The superscripts are time indices and the subscripts are space indices. Both second order and fourth order differences will be considered, and the values of the A and B coefficients in (2) are given in Table 1.

The solution to (2) for the "oceanic case" of Clancy (1981) is presented in Fig. 1. In this case $\Delta x = 10^5$ m, $\Delta t = 3600$ sec, $u = .6$ ms⁻¹, and $D = 10^3$ m²s⁻². The advection-diffusion equation is solved on the domain $x=0$ to $x = 10^8$ m with $S = \sin(kx)$ at $t=0$ and the boundary conditions $S = -\exp(k^2Dt)$ $\sin(ukt)$. For these assumptions the analytic solution to (1) is written as

$$S = \exp(-k^2Dt) \sin(kx-ukt).$$

(Note that Clancy's S_0 has been set to 1 and that the boundary conditions are negative to those of Clancy's since there is a sign error in Clancy's Eq. (22).)

Figure 1a is the same as Clancy (1981) Fig. 2; that is, the forward in time, second order centered in space finite difference solution to (1). There are both phase and amplitude errors as discussed by Clancy. Figure 1b shows the forward in time, fourth order spatial finite difference solution to (1). The phase error in this case is reduced nearly to zero. In Fig. 1c fourth order spatial differences and time differences centered with respect to the advective term are used and both the phase and amplitude errors are small. The analytic and numerical solutions are coincident in Fig. 1c. Clearly, with fourth order spatial differences and centered time differences with respect to the advective term and forward time differences with respect to the diffusive term, a much better estimate of the analytic solution is produced than with the Clancy scheme.

Calculating the CFL condition, $u\Delta t/\Delta x$, for the above parameters yields

$$u\Delta t/\Delta x = .0216$$

With these values it takes almost 50 time steps for a disturbance to propagate through one spatial grid. Figure 2 shows the solution to (1) using centered time differences and fourth order spatial differences with Δt increased by a factor of 10 to 36000 s. Once again the analytic and numerical solutions are indistinguishable. This increase in time step and increased accuracy of the solution compared to the forward time scheme more than compensates for any computational advantage the forward time scheme might possess.

In order to analyze the problems of the forward time scheme consider equation (1) written in finite difference form for second order spatial differences and forward time differences:

$$S_i^{n+1} = S_i^n - \frac{u\Delta t}{2\Delta x} (S_{i+1}^n - S_{i-1}^n) + \frac{D\Delta t}{\Delta x^2} (S_{i+1}^n - 2S_i^n + S_{i-1}^n) \quad (3)$$

If $D=0$ then this is the Euler scheme and is known to be unstable (Haltiner and Williams, 1980, page 130). If D is nonzero then the "old" stability criteria presented in Clancy (1981) are:

$$\frac{u\Delta x}{D} < 2 \quad \text{and} \quad \frac{D\Delta t}{\Delta x^2} < 1/2 \quad (4)$$

and the "new" stability criteria derived by Clancy (1981) are:

$$\frac{D\Delta t}{\Delta x^2} < 1/2 \quad \text{and} \quad \frac{u^2\Delta t}{2D} < 1 \quad (5)$$

Both sets of stability criteria are derived from Clancy's Eq.(6):

$$-4 \frac{D \Delta t}{\Delta x^2} (1 - \cos \theta) + 4 \frac{D^2 \Delta t^2}{\Delta x^4} (1 - \cos \theta)^2 + \frac{u^2 \Delta t^2}{\Delta x^2} (1 - \cos \theta) (1 + \cos \theta) \leq 0 \quad (6)$$

where θ is the grid size, Δx , multiplied by the wave number. The first set of stability conditions (Eq. 4) is derived by assuming that $\theta \neq 0$, dividing by $(1 - \cos \theta)$, and solving the resulting inequality. The second set of stability criteria (Eq. (5)) are derived by dividing (6) by $(1 - \cos \theta)$ and then considering the limit of the resultant equation as θ approaches zero (Clancy, 1981; Eq.(15)). The stability criterion obtained by such manipulation namely:

$$\frac{u^2 \Delta t}{2D} = \frac{u \Delta t}{\Delta x} \frac{u \Delta x}{2D} \leq 1 \quad (7)$$

has the somewhat dubious quality that D may be quite small, thereby violating the first inequality in Eq. (4); nevertheless, the equation is still integrable if Δt is reduced or Δx is increased such that (7) is satisfied. For the parameter values considered above, this procedure generates a numerical scheme where many time steps are required for a disturbance to propagate across one grid increment.

It is useful to reconsider the derivation of Clancy (1981) in order to understand the dubious nature of the "new" stability criteria.

Assuming $\theta \neq 0$, (6) can be rewritten as

$$\frac{D \Delta t}{\Delta x^2} + \frac{u^2 \Delta t}{4D} + \cos \theta \left(-\frac{D \Delta t}{\Delta x^2} + \frac{u^2 \Delta t}{4D} \right) \leq 1 \quad (8)$$

If it is then assumed that the coefficient multiplying the $\cos\theta$ term is negative, then:

$$\frac{u^2 \Delta x^2}{D^2} < 4 \quad \text{which yields} \quad \left| \frac{u \Delta x}{D} \right| < 2 \quad (9)$$

Under the condition of (9), (8) reaches a maximum if $\cos\theta = -1$; therefore, (8) is always satisfied if:

$$\frac{D \Delta t}{(\Delta x)^2} < 1/2 \quad (10)$$

Equations (9) and (10) do not constitute two stability criteria, but should be read as: given the conditions that make (9) true, (10) must be true if the growth is to be less than 1.

To derive the "new" stability criteria Clancy assumes that the coefficient of the $\cos\theta$ term in Eq. (8) is > 0 . This requires that

$$\frac{u^2 \Delta t^2}{\Delta x^2} > 4 \frac{D^2 \Delta t^2}{\Delta x^4} \quad \text{which implies} \quad \left| \frac{u \Delta x}{D} \right| > 2 \quad (11)$$

Given that (11) is true then (8) is a maximum when $\cos\theta = 1$; therefore, Clancy investigates Eq. (8) as $\theta \rightarrow 0$. This step must be carefully considered since (6) was divided by $(1-\cos\theta)$ to yield (8). In any event, letting θ approach zero in (8) yields Clancy's new stability criterion

$$\frac{u^2 \Delta t}{2D} < 1 \quad (12)$$

Therefore, the conditions (11) and (12) can be read: Given that (11) is true,

then (12) must be true if the growth is to be less than 1.

The statement by Clancy that if both conditions in Eq. (4), the "old" stability criteria, are true then (12) is necessarily satisfied is a false statement. In fact, (12) is derived under the assumption that (11) is true which directly contradicts (4). It is important to note that neither (9) or (11) (the same as Clancy's (9) and (17)) are stability criteria, but are in fact assumptions made during the derivation.

The stability characteristics of the forward in time, centered in space finite difference scheme can be better investigated by considering the magnitude of the function (G) on the left hand side of (8). A graph of G versus diffusion (D) is given in Figure (3). For $G > 1$ the scheme is unstable. It is seen that when using forward time differences, diffusion acts to stabilize the otherwise unstable Euler scheme. For small D the scheme is unstable; as D increases the scheme becomes stable; and as D further increases the scheme once again becomes unstable.

It is useful to consider the curve in Fig. (3) in light of the derivation of the stability criteria. This curve is drawn for a value of θ such that $\cos\theta$ is neither maximum or minimum. With $\Delta t = 36000$ s and all other parameters as before, the bold line in Fig. 3 is where the sign of the coefficient multiplying the $\cos\theta$ term in (8) changes. What Clancy's (1981) stability criterion (15) does is locate the minimum value of diffusion for which the numerical scheme is always stable. The stability criteria, in fact, are Eqs. (10) and (12), with these equations in general defining the upper and lower values of diffusion that stabilize the scheme for all θ .

Despite the fact that the forward in time, centered in space advection-diffusion equation can be made numerically stable by a judicious choice of the diffusion coefficient, its use is not advisable for transport problems. A

certain amount of the diffusion is always serving to stabilize the advective portion of the problem; therefore, it is difficult to incorporate physically derived diffusion coefficients into the model in a meaningful fashion.

Finally, for the parameters considered here, the time step can be increased by more than a order of magnitude and greatly improved results are obtained when using fourth order spatial differences, with the time derivative centered with respect to the advective term and forward with respect to the diffusion term.

Acknowledgements

This research is supported by the National Aeronautics and Space Administration and the Office of Naval Research. The author thanks M.R. Schoeberl and H.R. Schneider for their helpful comments.

Table 1
Coefficients for Spatial Difference Scheme

	2 nd order	4 th order
A ₁	1/2	2/3
B ₁	0	-1/12
A ₂	1	4/3
B ₂	0	-1/12

List of Figures

Figure 1. A comparison of the analytic and numerical solution of the advection-diffusion equation using different numerical schemes.

(+ on the numerical solution)

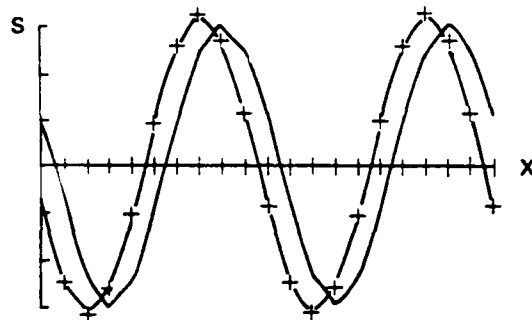
- a. Forward in time, second order in space (same as Clancy (1981), Figure 2)
- b. Forward in time, fourth order in space
- c. Centered in time, fourth order in space

Figure 2. Integration of Clancy's (1981) oceanic case with $\Delta t = 10$ hours. Same as Figure 1c but with the time step increased an order of magnitude. The numerical solution (+) is virtually indistinguishable from the analytic solution.

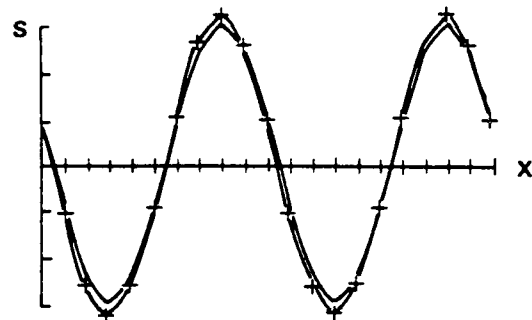
Figure 3. Stability characteristics of the numerical solution for forward time differences as a function of diffusion. For $G > 1$ the solution is unstable.
($\Delta x = 100$ km, $\Delta t = 10$ hours, $u = .6$ m sec⁻¹)

References

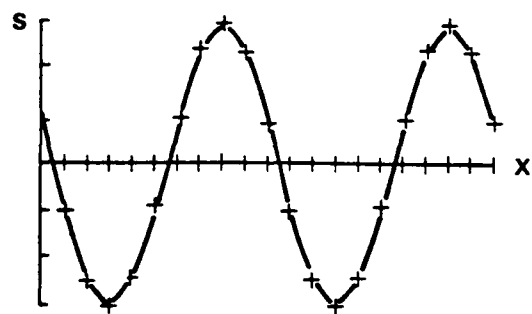
- Clancy, R.M., 1981: A note on finite differencing of the advection-diffusion equation, Mon. Wea. Rev., 109, 1807-1809.
- Fromm, J.E., 1964: The time dependent flow of an incompressible viscous fluid, Methods of Computational Physics, Vol. 3, Academic Press, 345-382.
- Haltiner, G.J. and R.T. Williams, 1980: Numerical Prediction and Dynamic Meteorology, 2nd Ed., John Wiley and Sons, Inc., New York, pp 477.
- Roache, D.J., 1976: Basic Computational methods for incompressible flow, Computational Fluid Dynamics, Hermosa Publishers, 15-48.



a. Forward in time, second order in space
(same as Clancy (1981), Fig. 2)



b. Forward in time, fourth order in space



c. Centered in time, fourth order in space

Figure 1. A comparison of the analytic and numerical solution of the advection-diffusion equation using different numerical schemes.
(+ on the numerical solution)

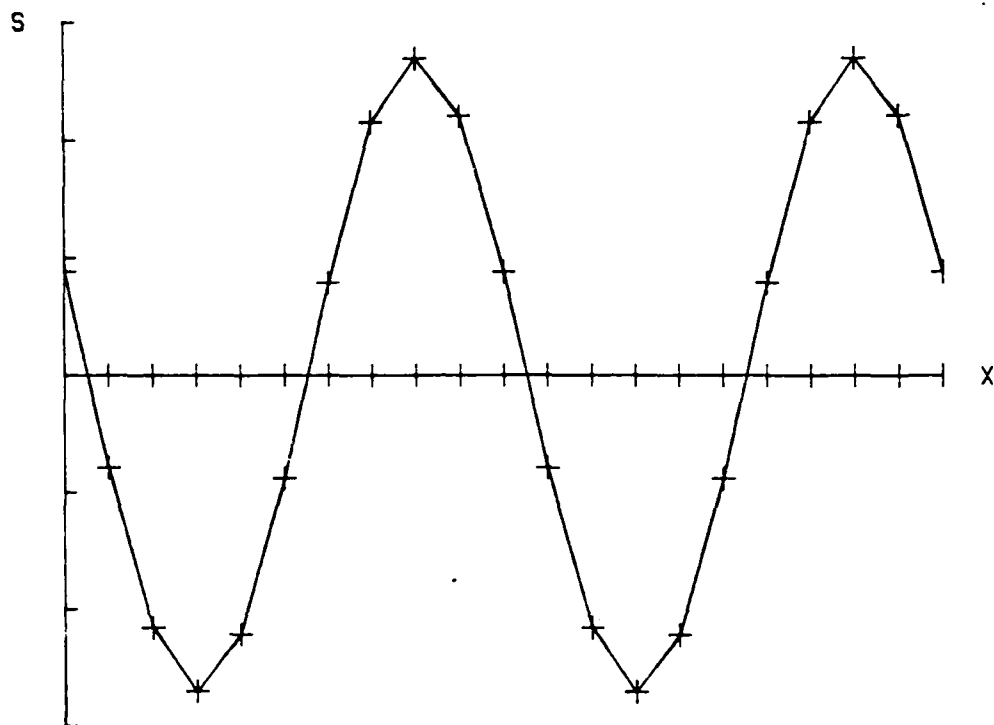


Figure 2. Integration of Clancy's (1981) oceanic case with $\Delta t = 10$ hours. Same as Figure 1c but with the time step increased an order of magnitude. The numerical solution (+) is virtually indistinguishable from the analytic solution.

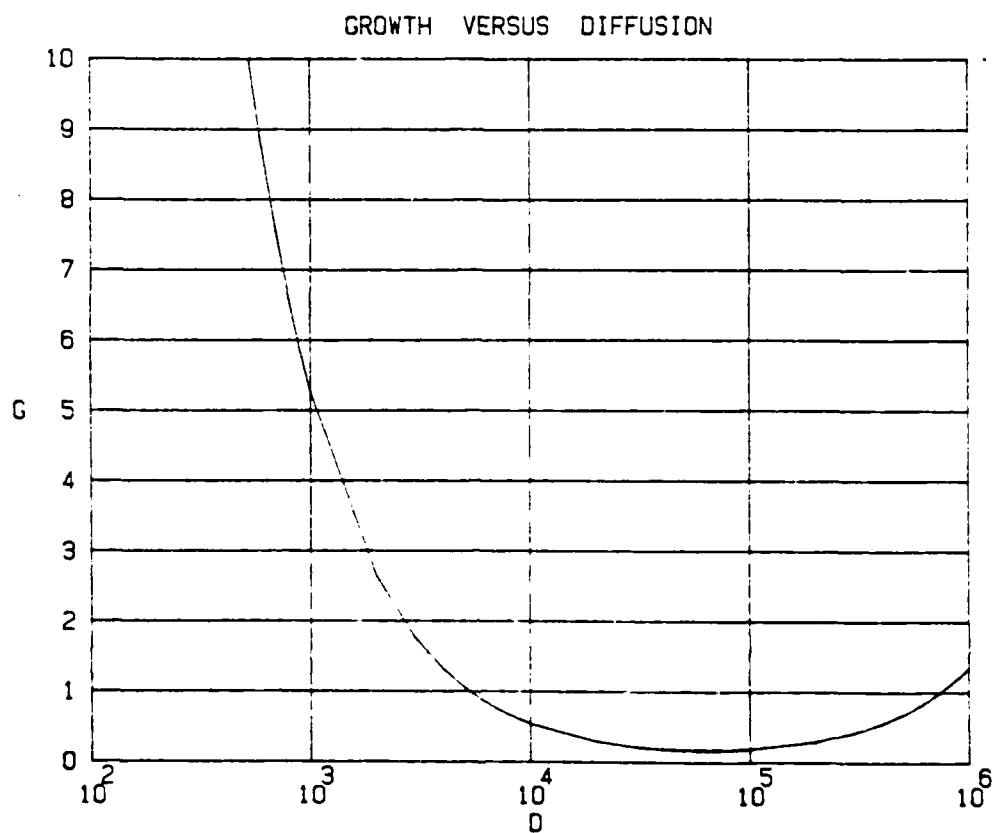


Figure 3. Stability characteristics of the numerical solution for forward time differences as a function of diffusion. For $G > 1$ the solution is unstable.
 ($\Delta x = 100$ km, $\Delta t = 10$ hours, $u = .6$ m sec⁻¹)

APPENDIX T

TURBULENCE ORIGINATING FROM
CONVECTIVELY STABLE INTERNAL WAVES

Turbulence Originating from Convectively Stable Internal Waves

R. S. Lindzen

Science Applications, Inc.
McLean, VA

and

Center for Earth and Planetary Physics
Harvard University
Cambridge, MA 02138

and

J. Forbes

Department of Physics
Boston College
Chestnut Hill, MA 02167

and

Center for Earth and Planetary Physics
Harvard University
Cambridge, MA 02138

Revised December 1982

Abstract

A theory has recently been developed for the generation of turbulence by unstable tides and gravity waves (Lindzen, 1981). Omitted from this theory was consideration of the possibility that turbulence could be generated by gravity waves which were not convectively unstable (i.e., breaking). The possibility of such generation was suggested by a result of McComas and Bretherton (1976) showing that internal gravity waves are indeed unstable to other gravity waves with shorter wavelengths. The present paper estimates the maximum turbulence which might be generated by such a process and shows that even this turbulence would not greatly alter earlier results.

1. Introduction

It is generally recognized that as internal gravity waves reach amplitudes for which they are unstable, they will break down -- generating turbulence sufficient to inhibit further growth (Hodges, 1969; Lindzen, 1981). The question arises as to whether smaller amplitude waves might generate turbulence as well. The analysis of Weinstock (1974) appears to suggest that this might be possible.¹ Apart from Weinstock's analysis, several studies suggest that the fields of gravity waves will cascade to higher wavenumbers (Dewan, 1979; McComas and Bretherton, 1977). In connection with this possibility, McComas and Bretherton (1977) showed that internal gravity waves were unstable to higher wavenumber gravity waves.² This strongly suggests that there will be some cascade to smaller scales. However, we have little further theory on this matter. Nevertheless such cascade might lead to turbulence by generating waves whose vertical wavelengths are short enough to permit breaking. The resulting turbulence might prove important for both mixing and for attenuating the original sub-breaking gravity wave.

¹Weinstock (1974) has a criterion for saturation which does not appear to require breaking. We do not understand Weinstock's analysis sufficiently to comment further on it.

²McComas and Bretherton (1977) study resonant triad interactions among gravity waves. However, one of their results concerning parametric subharmonic instability amounts to a proof that an internal gravity is unstable to a pair of internal gravity waves -- each having $1/2$ the vertical wavelength of the original wave -- one travelling upward, the other downward.

At present, we do not actually know whether convectively stable gravity waves cascade so as to form waves at scales where breaking can occur. Despite this ignorance, we will attempt to show, in the present brief paper, that such cascades, if they do exist, will not significantly modify earlier results which omitted such cascades (Lindzen, 1981). The additional eddy diffusion attributable to such cascade will be shown to be small -- compatible with current observations. Our approach is to assume the existence of a cascade and to specifically consider a cascade wherein smaller scales have the same amplitudes as larger scales (i.e., cascade without loss of amplitude). Such a cascade is intuitively more effective than any real cascade is likely to be. The fact that the eddy diffusion resulting from even this cascade model is still of very limited importance allows us to reach the above conclusions despite our ignorance.

2. Cascade Model

Dimensionally, we expect vertical eddy diffusion coefficients to be of the form

$$D \sim W' \cdot h \quad (1)$$

where W' is a characteristic eddy vertical velocity and h is a characteristic vertical scale. Of course, internal gravity waves have vertical velocity amplitudes which are simply related, at least approximately, to more readily measured temperature amplitudes through the relation

$$W' \approx \frac{\omega T'}{\Gamma} \quad (2)$$

where ω = frequency (doppler shifted when mean flow is present), T' temperature amplitude, and $\Gamma = \frac{d\bar{T}}{dz} + \frac{g}{c_p}$ where $\bar{T}$ = mean temperature (see Lindzen, 1981).

In addition, in the absence of shear in the mean flow, gravity waves are associated with two vertical scales: λ^{-1} (where λ = vertical wavenumber) and $2H$ (where H , here, refers to the pressure scale height), associated with growth of amplitude with height. For simplicity we will, in this note, ignore effects due to shear in the basic flow. Now one might conceivably use the above magnitudes in eq. (1), to obtain

$$D \sim f(\lambda H) W' \lambda^{-1} \quad (3)$$

but in general we do not expect a gravity wave (or even the superposition of many gravity waves) to stir the environment (Andrews and McIntyre, 1976). An exception to this is a breaking gravity wave where

$$\lambda T' \sim \Gamma \quad (4)$$

In this case Hodges (1969) and more recently Lindzen (1981) have shown that

$$D \approx \frac{1}{2H\lambda} \frac{\omega T'^2}{\Gamma^2} \quad (5)$$

In the absence of breaking one might expect $D = 0$. However, if due to nonlinear cascade vertical scales smaller than λ^{-1} are generated, then it is conceivable that despite smaller amplitudes, λ might be sufficiently large to cause breaking at the smaller scale -- and, therefore, diffusion.

Unfortunately, at present, there is limited direct observational theoretical evidence for such a cascade (van Zandt, 1982) though some theoretical arguments

³ (4) assumes $\lambda \gg \frac{1}{2H}$. If $\lambda < \frac{1}{2H}$, then λ should be replaced by $\frac{1}{2H}$.

⁴ In the limit of no mean flow Lindzen (1981) actually obtains

$$D = \frac{1}{2H\lambda} \frac{u}{\lambda^2}$$

at breaking. Substitution of (4) leads to (5).

have been put forth (Dewan, 1979). However, some interesting insights can be obtained by hypothesizing a cascade and investigating the consequences.

Assuming our basic wave is sub-breaking with amplitudes T , $W = \frac{\omega T}{\Gamma}$, and vertical wavelength λ . We will assume that at some larger λ' ,

$$T_{\lambda'} \sim T \left(\frac{\lambda}{\lambda'} \right)^\alpha \quad (6)$$

and

$$W_{\lambda'} \sim \frac{\omega T}{\Gamma} \left(\frac{\lambda}{\lambda'} \right)^\beta \quad (7)$$

where α , and β are for the moment assumed cascade laws. If $\alpha < 1$, then there will exist a convectively unstable (breaking) λ' where

$$\lambda' T_{\lambda'} \sim T \lambda^\alpha \lambda'^{1-\alpha} \sim \Gamma$$

i.e.,

$$\lambda' \sim \left(\frac{\Gamma}{T \lambda^\alpha} \right)^{\frac{1}{1-\alpha}} \quad (8)$$

For the breaking scale we might expect

$$\begin{aligned} D &\approx f(\lambda H) W_{\lambda'} \frac{1}{\lambda'} \\ &\approx f(\lambda H) \frac{\omega T^2}{\Gamma^2} \left(\frac{\lambda T}{\Gamma} \right)^{\frac{\alpha+\beta}{1-\alpha}} \end{aligned} \quad (9)$$

If we further require that when $\lambda' = \lambda$, (9) reduces to (5) then we take

$$f(\lambda H) = \frac{1}{2H\lambda} \quad (10)$$

and

$$D \approx \frac{1}{2H\lambda} \frac{\omega T^2}{\Gamma^2} \left(\frac{\lambda T}{\Gamma} \right)^{\frac{\alpha+\beta}{1-\alpha}} \quad (11)$$

The maximum value of D does correspond to $\lambda T = \Gamma$, in which case

$$D = D_{\max} = \frac{1}{2H\lambda} \frac{\omega}{\lambda^2} \quad (12)$$

(11) can be rewritten

$$D \approx D_{\max} \left(\frac{\lambda T}{\Gamma} \right)^n \quad (13)$$

where

$$n \equiv \frac{2-\alpha+\beta}{1-\alpha} \quad (14)$$

The choice of α and β is of course our major problem. Indeed, as already mentioned, the existence of α 's and β 's is even in question. However, it seems reasonable to assume that even if there is cascade, α and β are unlikely to be less than zero. Otherwise cascade products would dominate the original waves and flow fields would always be dominated by breaking waves. This is certainly contrary to observations. Thus results for $\alpha = \beta = 0$ should provide an upper bound on the turbulence result: from the cascade of

⁵ A priori, it might seem more reasonable to use $\frac{1}{2H\lambda'}$. However, this form assumes that D is exactly sufficient to prevent growth with height -- the cause of breaking for the primary wave. For $\lambda' < \lambda$ the direct cause of breaking is cascade instead. Moreover, at small scales we expect dependence on $H\lambda'$ to become asymptotically unimportant. Thus the choice in (10) seems more appropriate.

sub-breaking waves. For $\alpha = \beta = 0$, $n = 2$. Thus, even in this case $D < D_{\max}$ for $|\lambda T| < \Gamma$. This is because the "mixing length" associated with breaking is markedly reduced.

It is interesting to note that a $\lambda^{-5/3}$ power spectrum would lead to $\alpha = \beta = 1/3$ and $n = 3$. As we shall see, the consequences of such a choice do not differ very much from the results for $n = 2$.

3. Some Estimates

Before proceeding to our main question (namely, whether turbulence generated via cascade could significantly attenuate the original wave -- perhaps even preventing its breaking), it may prove interesting to use our formulae in order to check some orders of magnitude.

3a. Turbulent Diffusion in the Tropical Stratosphere

Short period wave motions in the tropical lower stratosphere appear to be dominated by mixed gravity-Rossby waves and by Kelvin waves (Lindzen and Tsay, 1975). Typical values associated with these waves and with the lower stratosphere are

$$T \sim 3^\circ\text{C}$$

$$\Gamma \sim 11^\circ/\text{km}$$

$$\omega \sim 2\pi/10 \text{ days}$$

$$\text{and } H \sim 7 \text{ km.}$$

Substituting these values in (13) and taking $n = 2$ we get

$$D_{\max} \approx \frac{1}{2H^2} \frac{u}{\lambda} \approx 2.1 \frac{\text{m}^2}{\text{s}}$$

and

$$D \approx D_{\max} \left(\frac{\lambda T'}{\Gamma} \right)^2 \approx 6.15 \times 10^2 \text{ cm}^2/\text{s} \quad (15)$$

(15), given our cascade assumption ($\alpha = 0 = \beta$), is almost certainly an overestimate. It is interesting, therefore, to note that the value given in (16) is quite modest. Current estimates are that $D \lesssim 1-5 \times 10^2 \text{ cm}^2/\text{sec}$ -- not far from (16) (Lilly et al., 1974). Note that

$$\left| \frac{\lambda T}{\Gamma} \right| \approx .17$$

so that the observations are compatible with a choice of n between 2 and 3.

It has been argued that sources of turbulence other than wave cascade exist in the lower stratosphere. A frequently cited example is wave breaking due to wave amplification in the neighborhood of critical layers (Geller et al., 1975). Such a mechanism is likely to occur in only limited regions and is likely to generate only weak diffusion (Lindzen, 1981). It is, therefore, far from obvious that cascade is not a competitive possibility.

3b. Relative importance of diurnal and semidiurnal tides in the generation of turbulence.

A major part of the ambient wave field in the tropical mesosphere is due to tides: diurnal ($\omega = \frac{2\pi}{\text{day}}$) and semidiurnal ($\omega = \frac{\pi}{\text{day}}$) (viz Chapman & Lindzen, 1970; Lindzen, 1979). From (12) we see that D_{max} is proportional to ω and inversely proportional to λ^3 , and thus strongly favors the semidiurnal mode.⁶

⁶ $\lambda \approx 1/3.5 \text{ km}$ for the main propagating diurnal tidal mode; for the main semidiurnal mode λ is so small that it is replaced by $1/2H$. As an aside, (12) implies that the maximum D producable by any internal wave will be limited to $D_{\text{supermax}} \approx \frac{1}{2H} \frac{\omega_B}{1/(2H)^2}$, where ω_B = Brunt-Vaisala frequency. $D_{\text{supermax}} \approx 4.7 \times 10^6 \text{ m}^2/\text{s}$. These results are, of course, independent of assumed cascades.

However, near 80 km the diurnal mode is believed to be unstable, while the semidiurnal is not. From (4) we see that this is associated with

$$T'_{\text{diur}} \sim 30^\circ \text{ K},$$

and

$$D \approx D_{\text{max}} \approx \frac{1}{2H \left(\frac{1}{3.5 \text{ km}} \right)^3} \frac{2\pi}{86,400 \text{ s}} \approx 2 \times 10^2 \frac{\text{m}^2}{\text{sec}}$$

consistent with the results of Lindzen (1981).

The semidiurnal tide at this height is generally believed to be associated with $T'_{\text{semidiur}} \sim 10^\circ \text{ K}$ and from (13) we get (taking $n = 2$)

$$D \approx 1.2 \times 10^2 \text{ m}^2/\text{s}.$$

Despite our assumption of unrealistically efficient cascade, this is still less than diffusion from the breaking diurnal tide. However, above 95 km, it is possible that the semidiurnal mode might be the major source of diffusion -- and might, in fact, significantly attenuate the propagating diurnal mode.

4. Self-Consistent Wave Breaking Calculations

Once one considers the possibility that stable, non-breaking internal waves can generate turbulence, the question arises as to whether this turbulence might actually prevent the breaking of the primary wave. We have chosen to study this question in the context of the main propagating diurnal mode. Using the EGM (equivalent gravity mode) formalism of Lindzen (1970), one may calculate the vertical structure of the main propagating mode forced by daily variations in insolation absorption by O_3 and H_2O in the presence molecular viscosity (indicated by the — • — • line in fig. 2). The resulting

vertical structure is shown by the solid line in fig. 1. The solid line indicates breaking above about 82 km. For the following calculations we have taken $n = 3$. (13) gives an eddy diffusion coefficient associated with the solid curve as shown by the solid curve in figure 2. One may next recompute the vertical structure using the sum of this eddy diffusion and molecular diffusion; recompute the eddy diffusion; and keep on repeating this whole procedure until convergence is obtained. The result is shown by the dashed curve in fig. 1. Below 80 km there is little difference between the initial and the converged solutions. Above 80 km the converged solution is, of course smaller, but breaking still occurs -- albeit at a somewhat greater altitude. Consistent with the above is the fact that converged eddy diffusions are about the same as the initial determination below 80 km, but somewhat less between 80 km and the final breaking level (viz fig. 2). Interestingly, the converged eddy diffusion above 80 km is very insensitive to the choice of cascade law (viz fig. 3). Thus, the particular choice of n in figures 1 and 2 is not important. Below 80 km, eddy diffusion depends strongly on the assumed cascade law; however, at these levels, the eddy diffusion has little effect on the tide itself. Note that eddy diffusion above 110 km is irrelevant insofar as molecular diffusion is much larger.

5. Observational Possibilities

The above calculations are able to bound the rôle of cascade to shorter vertical wavelengths without in any way demonstrating that such a process actually exists. However, as observational methods with high vertical resolution become available, the possibility of observational confirmation or rejection becomes possible. Assuming $\alpha > 0$ and $\beta > 0$, eq. (4) tells us

$$\lambda' \geq \frac{r}{T} . \quad (16)$$

Evaluation of (16) allows some estimate of the amount of resolution which would be necessary.

For the parameters in section 3a, we get

$$\lambda' \geq \frac{11^\circ/\text{km}}{3^\circ} \approx 3.7 \text{ km}^{-1}$$

or vertical wavelength $\leq \frac{2\pi}{3.7 \text{ km}}^{-1} \approx 1.7 \text{ km}$. Thus, in the tropical stratosphere, resolution better than $1.7 \text{ km}/2\pi \approx 270 \text{ m}$ would be needed.

Table 1 gives similar estimates for the main propagating diurnal mode. It is evident that breaking due to cascade of this mode will be very difficult to observe below 50 km.

The remaining question is how one distinguishes turbulence from waves. Two characteristics may be amenable to observation:

- i) Internal waves are characterized by horizontal scales which are much larger than vertical scales, whereas for convective instabilities horizontal and vertical scales are comparable. Thus, as one observes smaller and smaller vertical scales, one should begin to see a breakdown in horizontal coherence as λ exceeds λ' .
- ii) Similarly, convective elements are generally associated with shorter time scales than are internal waves. Presumably a time-vertical space spectral analysis could reveal such a transition.

Finally, if independent estimates of D are available, measurements of temperature variance, characteristic frequency, and vertical scales should permit the determination of whether a cascade parameter exists and if so what its value is.

7. Concluding Remarks

Underlying the present conjectural work is the notion introduced in Lindzen et al. (1980) that turbulent mixing can only arise from mass redistributing instabilities such as convective and/or inertial instability. Thus, for stable internal waves to generate turbulence there must first be a transfer of energy into convectively unstable scales. In this paper we have put forth a simple parameterization for such a cascade. Despite the ad hoc nature of the parameterization, it permits one particular concrete result to be developed: namely, that turbulence due to stable waves does not diminish the amplitude of waves below the level at which they would have broken in the absence of turbulence. It should be noted that our "worst case" cascade ($\alpha = \beta = 0$) would itself strongly modify the basic wave regardless of ultimate breaking. Observations of most internal gravity waves do not have sufficient time and height resolution to unambiguously discount this possibility. However, observations of tides (summarized in Chapman and Lindzen (1970) and Lindzen (1979)) do display the characteristics of the main modes including the anticipated growth with height -- at least below 80 km. Thus, cascade for these modes is almost certain to be less effective than our "worst case".

In addition to a variety of speculative estimates, we suggest some ways by which observations can help make our ideas of how internal waves generate turbulence more concrete. Finally, it should be noted that the sudden onset of diffusion at the breaking level as proposed in Lindzen (1981) can cause difficulties when incorporated into numerical models of the large scale mesospheric circulation. Eq. (13) provides a plausible expression for smoothing the onset, and section 4 suggests that this smoothing will not greatly affect the primary wave. The deposition of momentum will, however, be spread somewhat.

Acknowledgements

This work was supported by the National Science Foundation under Grant No. ATM-78-23330, by NASA under Grant No. NGL-22-007-228 and by the Office of Naval Research. J. Forbes gratefully acknowledges support under Grant No. AFOSR-81-0090 to Boston College.

References

- Andrews, D. G. and M. E. McIntyre, 1976: "Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration," J. Atmos. Sci., 33, 2031-2048.
- Chapman, S. and R. S. Lindzen, 1970: "Atmospheric Tides," D. Reidel, Hingham, Mass.
- Dewan, E., 1979: "Stratospheric wave spectra resembling turbulence," Science, 204, 832-835.
- Geller, M. A., H. Tanaka and D. C. Fritts, 1975: "Production of turbulence in the vicinity of critical levels for internal gravity waves," J. Atmos. Sci., 32, 2125-2135.
- Hodges, R. R., Jr., 1969: "Eddy diffusion coefficients due to instabilities in internal gravity waves," J. Geophys. Res., 74, 4087-4090.
- Lilly, D. K., D. E. Waco and S. I. Adelfang, 1974: "Stratospheric mixing estimated from high-altitude turbulence measurements," J. Appl. Meteor., 13, 488-493.
- Lindzen, R. S., 1970: "Internal gravity waves in atmospheres with realistic dissipation and temperature: Part I. Mathematical development and propagation of waves into the thermosphere," Geophys. Fl. Dyn. 1, 303-355.
- Lindzen, R. S., 1979: "Atmospheric Tides," An. Rev. Earth & Plan. Sci. 7, 199-225.
- Lindzen, R. S., 1981: "Turbulence and stress owing to gravity wave and tidal breakdown," J. Geophys. Res., 86, 9707-9714.
- Lindzen, R. S., B. Farrell and K. K. Tung, 1980: "The concept of wave overreflection and its application to baroclinic instability," J. Atmos. Sci., 37, 44-63.
- Lindzen, R. S. and C. Y. Tsay, 1975: "Wave structure of tropical stratosphere over the Marshall Islands during 1 April - 1 July, 1958," J. Atmos. Sci., 32, 2009-2021.
- McComas, C. H. and F. P. Bretherton, 1977: Resonant interaction of oceanic internal waves," J. Geophys. Res., 83, 1397-1412.
- van Zandt, T. E., 1982: "A universal spectrum of buoyancy waves in the atmosphere," Geophys. Res. Letters, 9, 575-578.
- Weinstock, J., 1976: "Nonlinear theory of acoustic gravity waves, I, Saturation and enhanced diffusion," J. Geophys. Res., 81, 633-652.

Table 1. Maximum wavelengths associated with characteristic amplitudes of the main propagating diurnal tide at different altitudes, Z (assuming $n = 2$).

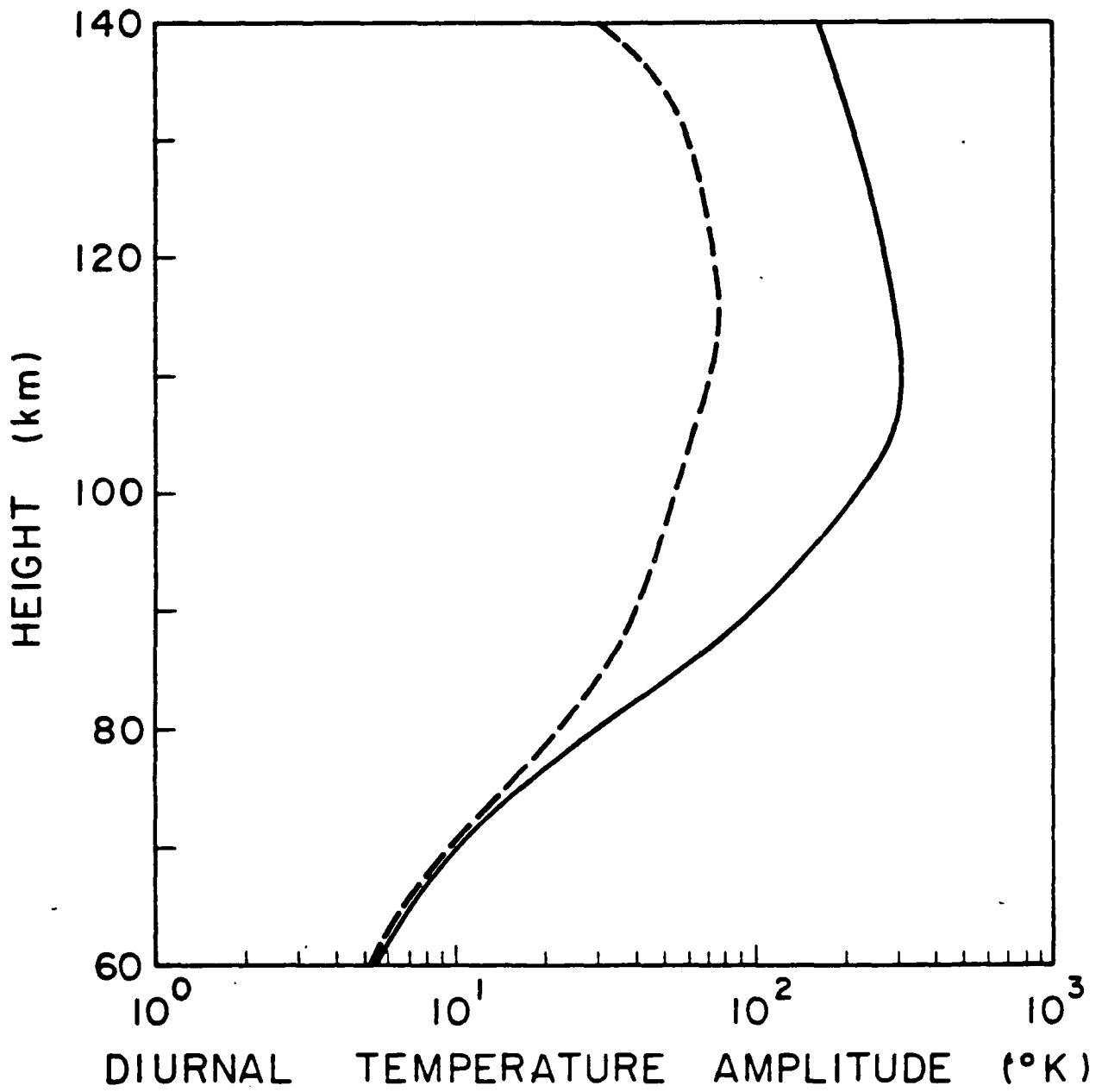
Z (km)	δT ($^{\circ}\text{C}$)	lower bound	upper bound
		for λ' (km^{-1})	for $2\pi/\lambda'$ (km)
30	.5	22	.28
40	1	11	.57
50	2	5.5	1.1
65	5	2.2	2.9
75	10	1.1	5.7

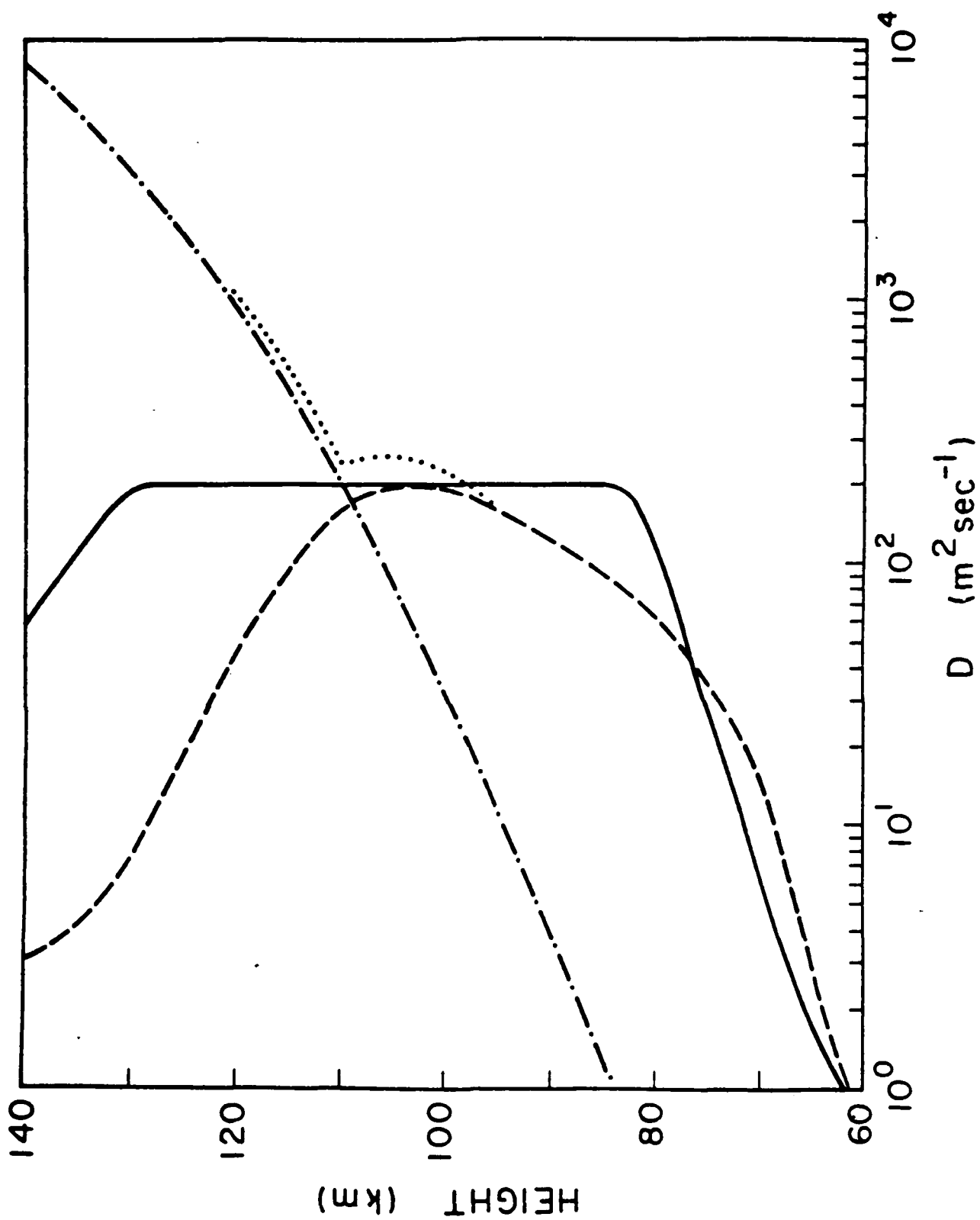
Figure Captions

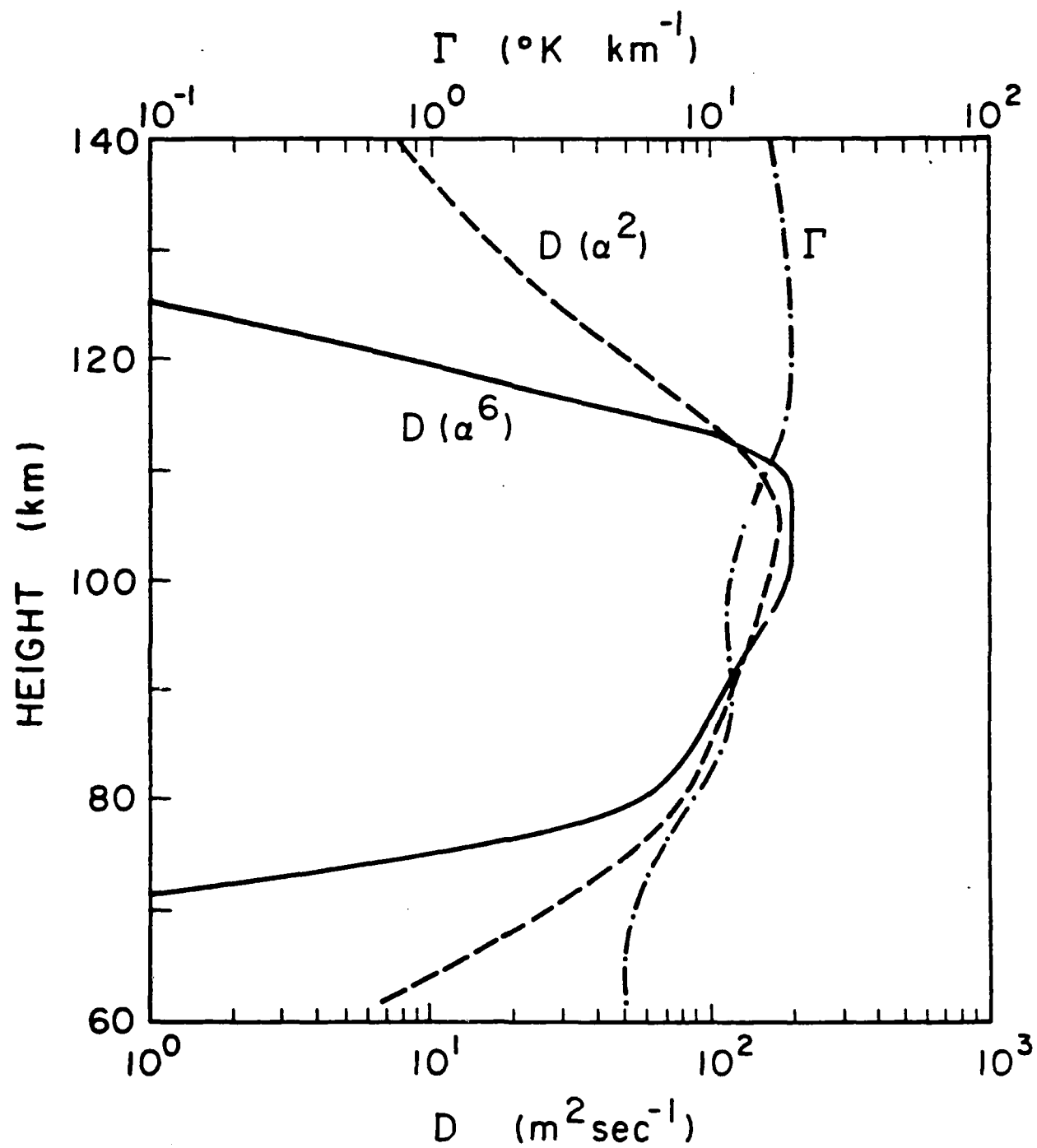
Figure 1. Equatorial diurnal temperature amplitudes for zeroth iteration (—) and converged solution (— — —) corresponding to $n=3$ power law.

Figure 2. Profiles of eddy diffusion coefficient for zeroth iteration (—) and converged solution (— — —) corresponding to $n=3$ power law, profile of molecular diffusion coefficient (— • —), and sum of molecular and converged eddy diffusion coefficient profiles (.....).

Figure 3. Stability function (Γ) and converged eddy diffusion coefficient profiles corresponding to $n=2$ (— — —) and $n=6$ (—) power laws.







END

FILMED

4-83

DTIC