MET O 11 TECHNICAL NOTE NO 184

AN EXTENDED LAGRANGIAN THEORY OF SEMI-GEOSTROPHIC FRONTOGENESIS

by

M J P CULLEN AND R J PURSER

This is a revised version of a paper submitted to the Journal of Atmospheric Sciences.

Met O 11 (Forecasting Research Branch)
Meteorological Office,
London Road,
Bracknell,
Berkshire,
ENGLAND

February 1984

AM

AN EXTENDED LAGRANGIAN THEORY OF SEMI-GEOSTROPHIC FRONTOGENESIS

M J P CULLEN AND R J PURSER

Meteorological Office, Bracknell, Berkshire RG12 2SZ UK

ABSTRACT

The Lagrangian conservation law form of the semi-geostrophic equations used by Hoskins and others is studied further in two and three dimensions. A solution of the inviscid equations containing discontinuities corresponding to atmospheric fronts is shown to exist for all time under fairly general conditions, and to be unique if the potential vorticity is required to be non-negative. Computational results show that this solution agrees with high resolution solutions of the viscous semi-geostrophic equations. The solution, however, disagrees with that obtained from the two dimensional viscous primitive equations. An important aspect of the difference is that the semi-geostrophic solutions allow the front to propagate into the interior of the fluid while the primitive equation solutions do not. This is discussed. If correct, it may indicate a tendency for a separation effect in the atmosphere where frictional effects are present.

1. Introduction

This paper considers an attempt to model mature atmospheric fronts by extending the frontogenesis model originally developed by Hoskins and Bretherton (1972). Their model uses the semi-geostrophic equations, in particular a Lagrangian conservation form of them. In a previous paper, Cullen (1983), referred to henceforward as C, one of the authors proposed an extension to this model by continuing to solve the Lagrangian conservation laws after the initial formation of the front. In this paper we develop and examine these ideas further.

The theory of Hoskins and Bretherton is now widely accepted as a mechanism for producing atmospheric fronts. However, comparison of their predictions with observation is difficult because their solutions only continue up to the initial formation of a disontinuity. At this stage frontogenetic effects dominate frontolytic effects. Observations of mature fronts, for instance those described by Sanders (1955), show a balance between the effects. Thus authors who have attempted to compare the Hoskins and Bretherton model with the atmosphere, such as Blumen (1980) and Ogura and Portis (1982), have found disagreements. Despite this a number of qualitative features of observed fronts were successfully predicted. Attempts to resolve the disagreements by adding frictional effects to the model were not very successful. In particular it was still found to be impossible to reproduce an observed vertical velocity maximum near the ground. There is thus a strong incentive to try to extend the theory to cover mature fronts.

The difficulty in using semi-geostrophic theory to describe fronts beyond their initial formation is that it is clear that the scale analysis used to derive the equations is no longer valid everywhere. In C it was suggested that the Lagrangian conservation laws derived from the equations might still be valid for most fluid particles. This is plausible because no fluid crosses a front, by definition, and thus only a small proportion of the fluid is likely to come close to the frontal surface. The solution is constructed geometrically, assuming that certain volumes of fluid have specified values of absolute momentum and potential

temperature. Thus changes to the values of these quantities for a small proportion of the fluid particles caused by local breakdown of the conservation laws may only have a small effect on the solution. If organised convection were present, however, the conservation laws could be violated for a large proportion of the fluid and the solution would not be expected to be very useful.

The likely validity of the model proposed in C thus depends on the proportion of the fluid volume in which the scale analysis leading to semi-geostrophic theory breaks down and the ageostrophic accelerations are large. According to Hoskins and Bretherton, this should only happen when the frontal vorticities, ζ , reach about 25f. They also suggest that three dimensional turbulence would have set in before then, because the local Richardson number is of the order f/ζ . However, computations by Orlanski and Ross (1977) and Ross and Orlanski (1982) using the primitive equations suggest that the scale analysis is not valid for vorticities larger than about 2f. After this stage the ageostrophic wind parallel to the front became significant. A possible explanation for this disagreement is the different method of determining the cross frontal circulation. In the semi-geostrophic model the cross-frontal flow is determined implicitly by an omega equation. The implied accelerations can be estimated from the rate of change of the forcing terms in the omega equation, as was done by Hoskins. In a primitive equation model the circulation can only be changed explicitly by an ageostrophic wind parallel to the front. Orlanski and Ross found that the resulting circulation was similar to that obtained using the omega equation method with the addition of an inertio-gravity oscillation. It is possible that the larger accelerations were associated with this oscillation. The calculated pattern of the ageostrophic wind and the separation they obtained between the line of maximum wind shear and temperature gradient are on scales close to their model grid length and will require further investigations using higher resolution models. At the present time there must therefore be uncertainty about the range of validity of semi-geostrophic models.

Though the method described in C of extending the solution to cover mature fronts seems plausible, it may not be the only way of doing so. In particular, the results reported in C suggest that a different type of solution is obtained by solving the primitive equations with artificial viscosity. It is only possible to choose between such solutions by incorporating the assumptions into realistic three dimensional models and testing against real data.

This paper continues the work started in C and is divided into two parts. The first shows how the geometrical method which was applied in an ad-hoc manner in C can be shown rigorously to give a unique solution for data with non negative potential vorticity under certain conditions on the physical domain. The method is also shown to be applicable to the three dimensional equations derived by Hoskins and Draghici (1977) using the geostrophic momentum approximation. The method allows a number of additional properties of fronts to be deduced, for instance that a discontinuity in temperature and velocity cannot be formed even at a finite discontinuity of potential vorticity. This is a stronger result than that obtained by Hoskins and Bretherton. The results also mean that such flows are predictable of type 1 or 2 in the sense of Lorenz (1969). This means that the solutions are determined for all time by the initial and boundary conditions, and are continuously dependent on the data. The latter requirement is essential if the neglected ageostrophic accelerations are to be small.

The second part presents further computational results. A model in which the two dimensional semi-geostrophic equations are solved using finite differences is described. In this model the ageostrophic winds have to be determined iteratively. This has two purposes. It allows a direct comparison with finite difference integrations of the primitive equations. Some of the differences reported in C may have simply represented the difference between two types of computational method, despite the use of very high resolution. Secondly, since artificial viscosity has to be added to the semi-geostrophic equations when solved by finite differences, we can check the agreement between the inviscid Lagrangian solution

described in C and a slightly viscous solution. This is important because an inviscid solution can only be regarded as physically useful if it is the limit as the viscosity tends to zero of a viscous solution.

Comparisons between solutions obtained by the Lagrangian construction and the two finite difference models are presented. The same values for the artificial viscosity are used in both cases. The results show closer agreement between the finite difference semi-geostrophic solution and the Lagrangian solution than between the two finite difference solutions. The largest vertical velocities are concentrated nearer the boundary in the semi-geostrophic solution.

The most important qualitative difference between the semi-geostrophic and primitive equation solutions is in the handling of the upper and lower boundary conditions. The Lagrangian theory gives a unique solution if the rigid wall boundary conditions are interpreted as meaning only that no fluid can cross the boundary. If the requirement is that fluid initially in contact with the boundary remains there, then no solution exists. Therefore fluid has to be squeezed away from the boundaries by lateral motions; in effect the boundary is sucked into the fluid. In the primitive equation model the vertical velocity is zero at the boundary and so fluid cannot move away from it. Therefore at least one of the Lagrangian conservation properties cannot be satisfied.

In reality, this need not be such a contradiction. A real fluid would not contain an actual discontinuity and the conservation laws would not be valid for every fluid particle. If the discontinuity is replaced by a shear layer, the kinematic boundary condition can be satisfied. The semi-geostrophic results suggests a tendency for a separation effect at the boundary. The physical correctness of this prediction can only be checked by tests against real data. It is tempting but speculative to relate it to the occlusion process.

2. Basic theory for the two-dimensional deformation model.

a. Basic equations

Consider the dimensionless Lagrangian form of the equations used for the deformation model introduced by Hoskins and Bretherton (1972) and studied in C. These are:

$$\frac{DM}{Dt} + \alpha M = 0$$

$$\frac{D\theta}{Dt} = 0 \qquad (2.1)$$

$$\frac{DA}{Dt} + \alpha A = 0$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + w\frac{\partial}{\partial z}$$

$$M \equiv v + x$$

α is a given deformation rate.

M is called the potential momentum and A is the area of a fluid element in the (x,z) plane. The coordinate z is a function of pressure as defined in the references above. The equations for cross-front geostrophic balance and the hydrostatic equation give

$$\frac{\partial \phi}{\partial x} = v$$

$$\frac{\partial \phi}{\partial z} = \theta \qquad (2.2)$$

The problem is to be solved in a closed convex region Ω with rigid upper and lower boundaries at z = 0, 1

Since fluid elements shrink with the basic deformation rate it is convenient to consider a domain in the (x z) plane which also shrinks at this rate, so we suppose that the boundary conditions in x are

rigid walls at

$$x = \pm e^{-\alpha t} \tag{2.3}$$

b. Smooth solutions

It is convenient to start by recalling some standard analysis of the solutions of (2.1) assuming that they are differentiable. In this case the potential vorticity

$$q = \frac{\partial(M, \theta)}{\partial(x, z)} \tag{2.4}$$

is conserved following the motion:

$$\frac{Dq}{Dt} = 0 \tag{2.5}$$

Introduce a modified pressure variable

$$P \equiv \phi + \tfrac{1}{2} x^2 \tag{2.6}$$

Then

$$(M, \Theta) = \left(\frac{\partial P}{\partial x}, \frac{\partial P}{\partial z} \right) \tag{2.7}$$

Define the Hessian matrix

$$\underset{\sim}{Q} \equiv \begin{bmatrix} \frac{\partial^2 P}{\partial x^2} & \frac{\partial^2 P}{\partial x \partial z} \\ \frac{\partial^2 P}{\partial z \partial x} & \frac{\partial^2 P}{\partial z^2} \end{bmatrix} \tag{2.8}$$

Then

$$q = \det \left(\underset{\sim}{Q} \right) \tag{2.9}$$

This matrix is of special importance in determining the ageostrophic circulation (u, w) and the pressure tendency $\Gamma \equiv \partial\phi/\partial t$.

For example, we have

$$\underset{\sim}{\nabla} \Gamma + \underset{\sim}{Q} \cdot \underset{\sim}{u} = \underset{\sim}{b} \tag{2.10}$$

where

$$\underset{\sim}{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial z} \right), \quad \underset{\sim}{u} = (u, w), \quad \underset{\sim}{b} = \left(-\alpha \frac{\partial P}{\partial x}, 0 \right)$$

The continuity equation:

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} + \alpha = 0 \qquad (2.11)$$

can be used to eliminate the velocity components to give:

$$\underset{\sim}{\nabla} \cdot \left(\underset{\sim}{Q}^{-1} \cdot \underset{\sim}{\nabla} T \right) = \alpha + \underset{\sim}{\nabla} \cdot \underset{\sim}{Q}^{-1} \cdot \underset{\sim}{b} \qquad (2.12)$$

This equation can be solved; subject to suitable boundary conditions, provided that it is elliptic, implying that $\underset{\sim}{Q}$ is positive definite. In some circumstances (2.12) can be solved when $\underset{\sim}{Q}$ is singular by using an integrated form of the continuity equation (integrated along the line of the characteristics between opposite boundaries), but the problem is ill-posed when $\underset{\sim}{Q}$ has a negative eigenvalue. In the latter case the physical system would be subject to convective, inertial, or symmetric instability; it would not be properly describable by the semi-geostrophic equations because acceleration components then became of comparable magnitude to the corresponding pressure gradient forces.

The following geometric picture of this constraint is important. Regard $P(x, z)$ as the elevation of a surface above the (x, z)

plane. Then $\underset{\sim}{Q}$ gives the curvature components of the surface, in the case where P varies slowly with x and z. Thus the potential vorticity q is essentially proportional to the Gaussian curvature of the surface. The condition that $\underset{\sim}{Q}$ has no negative eigenvalues, assuming that P is twice differentiable, is equivalent to stating that the surface $P(x,z)$ is convex when viewed from below. It will be shown that this geometrical interpretation can be used in cases where the solutions are not differentiable.

c. Discontinuous solutions

In C it was demonstrated that, for simple piecewise constant initial data, a solution of (2.1) and (2.2) could be constructed by geometry. The construction is illustrated in Fig.1. The values of M and θ are constant on each element, and the slope of the line separating the elements in the (x,z) plane is given by the geostrophic and hydrostatic relations by the formula first pointed out by Margules in 1906:

$$\frac{[M]}{[\theta]} \qquad (2.13)$$

The solution then appeared to be uniquely specified by the areas of the elements, the values of M and θ within them, and the condition that the solution is statically stable.

The modified pressure P defined by (2.6) can be used for piecewise constant data also. The surface $P(x,z)$ is now made up of flat faces on which $\frac{\partial P}{\partial x}$ and $\frac{\partial P}{\partial z}$ are constant. If P is continuous, the the jump condition (2.13) is automatically satisfied for the edges where faces join. The surface corresponding to the solution in Fig. 1 is shown in Fig. 2. The condition that the solution in physical space is statically stable becomes a condition that the

surface $P(x,z)$ is convex viewed from below. This suggests that the convexity condition can be used as a condition for dynamical stability for solutions not differentiable everywhere, instead of the potential vorticity which is difficult to define. By using convexity to define dynamical stability it is then possible to prove a rigorous existence theorem for piecewise constant solutions of (2.1) and (2.2). In this paper a somewhat intuitive version of the proof is given, a more rigorous treatment will be given elsewhere (Purser and Cullen (in preparation)).

d. Definitions and consequences

A solution of the semi-geostrophic problem (2.1), (2.2) is called dynamically stable if the associated surface $P(x,z)$ is convex (viewed from below), i.e. the region in $\underset{\sim}{R}^3$ with coordinates (x, z, s) defined by

$$s \geqslant P(x,z) \tag{2.14}$$

is a convex subspace of $\underset{\sim}{R}^3$

Write the convex subspace defined by (2.14) as V. Then V has the following properties (e.g. see Rockafellar (1970)).

(i) V can be represented as the intersection of half spaces

$$L(\underset{\sim}{x}) \geqslant 0 \tag{2.15}$$

where L is a linear function of $\underset{\sim}{x}$. The planes

$$L(\underset{\sim}{x}) = 0 \tag{2.16}$$

are called tangent planes of V.

(ii) Every point on the boundary of V has at least one tangent plane touching it.

e. Theorem 1

The problem (2.1) (2.2) in Ω has a unique dynamically stable solution for all time given bounded piecewise constant initial data for M and Θ at t = 0. The solution is continuously dependent on the initial data.

Proof - basic construction

This theorem is proved by explicit construction of the surface $P(x,z)$. The essential step is shown in Fig. 3. Given the region Ω in the (x,z) plane, consider an infinite extension of it in the s-coordinate to a convex cylindrical surface V in R^3 (Fig.3(a)). Now construct a convex surface Y as the successive intersection of planes W_i with V, subject to the condition that the planes are not perpendicular to the plane S = 0. (Fig.3(b)). The angle of each plane W_i with the S axis is specified. At each stage the lower boundary of Y, denoted Y_i, is a piecewise flat convex surface. If the planes W_i are moved up and down, the areas of those parts of Y_i which have a given tangent plane change. (Fig.3(c)).

Suppose the initial data are given in the form of a list (M_i, Θ_i, A_i) of values of M and Θ on finite elements with areas A_i in the (x,z) plane. Suppose that

$$\sum_{i=1}^{n} A_i = \text{area of } \Omega \qquad (2.17)$$

Then at any future time t, we seek to arrange the elements with values of (M, Θ) given by $(M_i e^{-dt}, \Theta_i)$ by (2.1) and areas $A_i e^{-dt}$ so that the continuity equation is satisfied. This is done by

associating the ith element with a plane W_i with gradient

$$\frac{\partial \rho}{\partial x} = \nu + \chi = M_i \qquad (2.18)$$

$$\frac{\partial \rho}{\partial z} = \theta = \theta_i$$

and seeking to construct the surface Y in Fig. 3 such that the area of the face of Y formed by the plane W_i is A_i. Since the angle of each plane W_i with the S axis is fixed by (2.13), the areas can only be changed by moving the W_i up and down. The remainder of the proof shows how this is done.

An arrangement of the planes W_i can be uniquely defined by specifying their s coordinates at x = z = o (Fig. 4).

f. Lemma

Any increase in S_k with S_i $(i \neq k)$ and M_i, θ_i all kept constant will result in an increase in the area A_K balanced by a decrease in the areas of each of the neighbouring faces.

Proof

This follows immediately from the convexity of the surface Y. The effect of a perturbation in S_1 is shown in Fig. 5.

The area A_1 changes monotonically with the same sign.

g. Uniqueness

Given the values of M_i, θ_i and A_i, then if a solution characterised by the coordinates S_i exists, it is unique within a uniform change in all the S_i.

Remark

The fact that Ω is convex is essential for this result.

Proof

Suppose two solutions, represented by surfaces Y_1 and Y_2, exist. Suppose that the associated coordinates are S_i^1 and S_i^2. Calculate the differences $(S_i^2 - S_i^1)$. Divide the planes W_i into two sets, E_A and E_B. E_A contains the planes for which $S_i^2 - S_i^1$ attains its maximum value, say δ, E_B contains

the remainder. Because there are a finite number of planes, E_A contains at least one member. If the theorem is true, then E_B must be empty so that S_i^2 can be obtained from S_i^1 by the simple translation

$$S_i^2 = S_i^1 + \delta \qquad (2.19)$$

Suppose, conversely, that E_B is not empty. Then, after carrying out the translation (2.19), all the coordinates of members of E_A in the two solutions agree, and $S_i^2 < S_i^1$ for all members of E_B. Conversion of Y_1 to Y_2 now requires negative changes to all the S_i associated with E_B, while preserving the areas of all the faces. Since the area of a face changes monotonically with its S_i at the expense of its neighbours, and since at least some members of E_B must adjoin members of E_A; the required changes in S_i must affect the areas of these members of E_A. This contradicts the requirement that the areas are the same in the two solutions. Therefore E_B must be empty and the result proven.

h. Existence

Given bounded M_i, Θ_i and A_i satisfying (2.17), a solution exists.

Proof

We construct a first guess by assigning arbitrary coordinates to planes W_i with the correct slope and constructing the surface Y. The areas of the intersections of W_i with Y need not be correct and some planes may not intersect Y at all. We then adjust the coordinates S_i iteratively to obtain the correct areas. At some stage of the iteration, assume that we have a set of coordinates $\underset{\sim}{S} = \{S_i\}$ yielding areas $\underset{\sim}{A'} = \{A_i'\}$. Some of the A_i' may be zero. Define the

error norm

$$N_2(\underset{\sim}{S}) = \sum_{i=1}^{\hat{}} |A_i - A_i'|^2 \qquad (2.20)$$

Because M_i and Θ_i are bounded, the planes W_i cannot be perpendicular to the plane S = 0. Since there are a finite number of planes, they must all intersect each other and the surface V at a finite angles. (If two planes are parallel then $M_i = M_j$, $\Theta_i = \Theta_j$ and their associated areas can be combined). Therefore all possible intersections of the planes W_i with V can be obtained within a finite range of $\underline{S}$, as can be seen from Fig. 4.

Therefore, for some $\underset{\sim}{S} = \underset{\sim}{S}_0$, $N_2(\underset{\sim}{S})$ attains its minimum. Write the associated values of $(A_i' - A_i)$ as ε_i.

Divide the elements into two sets; E_a and E_b. E_a contains elements for which ε is the maximum, ε_a ; E_b contains the remainder. Then since $N_2(\underset{\sim}{S})$ has been minimised, if E_b is not empty, at least one member, k, of E_a adjoins members of E_b. Element k may also adjoin other elements of E_a. By reducing S_k we can reduce ε_k by an arbitrary amount: say (a + b). By doing so we increase the combined areas of other elements by (a + b).

Suppose that the areas of elements in E_a are increased by a total amount a and in E_b by a total amount b. The resulting change ΔN_2 in N_2 comprises three parts:

(i) $\Delta N_2^{(a)}$ due to changes in areas of elements in E_a excluding k.

(ii) $\Delta N_2^{(k)}$ due to the change in the area of element k

(iii) $\Delta N_2^{(b)}$ due to changes in members of E_b.

Let ξ_b be the maximum error for members of E_b; then, by definition, $\xi_b < \xi_a$. Use the simple inequality, for positive a_i:

$$\left(\sum(A_i + a_i)^2 - \sum A_i^2\right) \leq \left(\text{Max}_i A_i + \sum a_i\right)^2 - \left(\text{Max}_i A_i\right)^2 \quad (2.22)$$

Then it follows that

$$\Delta N_2^{(a)} \leq 2\xi_a a + a^2$$

$$\Delta N_2^{(b)} \leq 2\xi_b b + b^2$$

$$\Delta N_2^{(k)} = -2\xi_a(a+b) + a^2 + b^2 + 2ab$$

hence

$$\Delta N_2 \leq -2(\xi_a - \xi_b)b + 2(a^2 + b^2 + ab) \quad (2.23)$$

Since the rate at which the area of an element changes with respect to its S coordinate is bounded; it is always possible to find a change in S and hence a and b, which is sufficiently small such that

$$(\xi_a - \xi_b)b > (a^2 + b^2 + ab) \quad (2.24)$$

and so $\Delta N_2 < 0$

E_b must therefore be a null set and so the errors ξ_i in the areas of all the segments are equal. Since they sum to zero, they must all vanish. Therefore the desired solution of the problem has been obtained.

h. Continuity

The solution, characterised by the set of coordinates $\{S_i\}$, depends continuously on M_i θ_i and A_i.

Proof

Only an intuitive proof is given here, since a rigorous proof requires considerable care. It is clear that, since M_i and Θ_i are represented in (2.18) by the components of the gradient of the solution $P(x,z)$; we can consider perturbations in M_i and Θ_i together, and treat perturbations in A_i separately.

Given an arbitrarily small quantity ϵ , which measures a change in the coordinates $\{S_i\}$ of the planes making up $P(x,z)$, we seek to find a δ such that for any two sets of data satisfying

$$\sum_{i=1}^{n} (M_i^1 - M_i^2)^2 + (\Theta_i^1 - \Theta_i^2)^2 < \delta$$
$$A_i^1 = A_i^2 \qquad (2.25)$$

we have

$$\int (P^1(x,z) - P^2(x,z))^2 < \epsilon \qquad (2.26)$$

The proof first of all assumes that, for any i and j;

$$(M_i^1 - M_j^2)^2 + (\Theta_i^1 - \Theta_j^2)^2 > \delta \qquad (2.27)$$

If (2.27) is not satisfied for all element pairs, then elements with similar M_i and Θ_i are combined into larger elements. By convexity, all such elements with almost equal slopes must be contiguous on the solution surface. Once (2.27) is satisfied, then a perturbation to the slope of any one element will not change its position on the solution surface; (Fig. 6). Thus the effect on $P(x,z)$ is localised and can be estimated in terms of the size of the perturbation δ , and the maximum linear dimension of the region. This is because $P^1(x,z)$ and $P^2(x,z)$ have gradients everywhere which either differ

by less than δ or differ by a larger amount α, representing the difference in slope between adjoining elements, on a region of size δ/α, where the inter-element boundaries have moved in (x, z) space (Fig. 7). A similar argument proves that combining elements with nearly equal slopes so that (2.27) is satisfied has a small effect on $P(x,z)$.

i. General initial data

The theorem that has been proved establishes the existence of a bounded solution to (2.1 - 2.3) for general piecewise constant initial data. If the initial data can be approximated to arbitrarily high accuracy by a piecewise constant field, then a corresponding sequence of approximate solutions can be generated. Provided that these can be bounded independently of the degree of approximation, they will converge to a limit which is the solution for general initial data. While it seems clear that any meteorologically relevant data satisfy this condition, provided turbulent regions are not resolved; the correct mathematical conditions are not obvious.

3. Three dimensional semi-geostrophic theory.

a. Basic equations

The three-dimensional equations are the natural generalisation of the two dimensional set studied in the previous section. The potential momentum now has two components M and N, closely associated with 'geostrophic coordinates', X and Y, which change non-trivially in time. A detailed study of the equations is given in Hoskins and Draghici (1977). The essential dimensionless problem, where the f plane and Boussinesq approximations have been made, is

$$\frac{Du_g}{Dt} - v_{ag} = 0$$

$$\frac{Dv_g}{Dt} + u_{ag} = 0 \qquad (3.1)$$

$$\frac{D\theta}{Dt} = 0$$

with

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{3.2}$$

$$v_g = \frac{\partial \phi}{\partial x} \; ; \quad -u_g = \frac{\partial \phi}{\partial y} \; ; \quad \theta = \frac{\partial \phi}{\partial z}$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + (u_g + u_{ag})\frac{\partial}{\partial x} + (v_g + v_{ag})\frac{\partial}{\partial y} + w\frac{\partial}{\partial z} \tag{3.3}$$

$$u = u_g + u_{ag} \; ; \quad v = v_g + v_{ag}$$

Defining:

$$M = v_g + x \; , \qquad N = -u_g + y \tag{3.4}$$

then

$$\frac{DM}{Dt} = u_g \; , \qquad \frac{DN}{Dt} = v_g \tag{3.5}$$

The set (3.5) can be rewritten in the form

$$\frac{\partial M}{\partial t} + \underset{\sim}{u}_g \cdot \nabla M + \underset{\sim}{u}_{ag} \cdot \nabla M = u_g$$

$$\frac{\partial N}{\partial t} + \underset{\sim}{u}_g \cdot \nabla N + \underset{\sim}{u}_{ag} \cdot \nabla N = v_g \tag{3.6}$$

$$\frac{\partial \theta}{\partial t} + \underset{\sim}{u}_g \cdot \nabla \theta + \underset{\sim}{u}_{ag} \cdot \nabla \theta = 0$$

where

$$\underset{\sim}{u}_g = (u_g \, , v_g \, , 0)$$

$$\underset{\sim}{u}_{ag} = (u_{ag} \, , v_{ag} \, , w) \tag{3.7}$$

and we have

$$\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} = 0 \qquad (3.8)$$

$$\frac{\partial u_{ag}}{\partial x} + \frac{\partial v_{ag}}{\partial y} + \frac{\partial w}{\partial z} = 0$$

These equations can be solved by a splitting technique, as is often used in primitive equation models (e.g. Gadd (1978)). Suppose that at any given time, the value of ϕ is known everywhere. Then Θ, M and N are determined as the gradients of ϕ by the geostrophic and hydrostatic relations (3.2) and the definitions (3.4). The ageostrophic wind is implicity determined by the equations of motion (3.6) and the requirement that $\frac{\partial M}{\partial t}$, $\frac{\partial N}{\partial t}$ and $\frac{\partial \Theta}{\partial t}$ must be related by the geostrophic and hydrostatic relations. This can be achieved by regarding the fluid at the given time as made up of elements, each with a definite value of M, N and Θ. In the first step of the solution the values of M and N are updated by $u_g \Delta t$ and $v_g \Delta t$, where Δt is the timestep; and the elements are then advected into new positions by the geostrophic wind. The behaviour of this part of the equations can be analysed by considering the associated ordinary differential equations following fluid elements:

$$\frac{dM}{dt} = y - N$$

$$\frac{dN}{dt} = M - x \qquad (3.9)$$

On a doubly periodic domain, x and y are bounded; and the part of the solution of (3.9) which depends on M and N is just an oscillation. Therefore there is unlikely to be any difficulty in this part of the solution to (3.6).

The remainder of (3.6) describes advection by an unknown wind $\underset{\sim}{u}_{ag}$ satisfying the incompressibility condition (3.8), so that fluid elements are moved without change of volume or of their values

of M, N and Θ to new positions such that the geostrophic and hydrostatic relations are satisfied. We now prove the existence of such an incompressible ageostrophic wind field.

b. Existence of implied ageostrophic circulation

The natural definition of potential vorticity, generalizing (2.4), is

$$q = \frac{\partial(N, -M, \Theta)}{\partial(x, y, z)} \tag{3.10}$$

Define

$$P = \phi + \tfrac{1}{2}(x^2 + y^2) \tag{3.11}$$

then

$$M = \frac{\partial P}{\partial x}, \quad N = \frac{\partial P}{\partial y}, \quad \Theta = \frac{\partial P}{\partial z} \tag{3.12}$$

Hence

$$q = \det(\underset{\sim}{Q})$$

where

$$\underset{\sim}{Q} = \begin{bmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{bmatrix} \tag{3.13}$$

For differentiable solutions, the condition for dynamical stability is, as in the two dimensional case, that $\underset{\sim}{Q}$ nowhere has a negative eigenvalue. For a general (not necessarily smooth) solution $P = P(x, y, z)$, the condition for stability may be formally stated.

Definition

A solution of the three dimensional semi-geostrophic equations is called dynamically stable within the (x, y, z) domain D; if the hypersurface in (x, y, z, s) space defined by $s \geq P(x, y, z)$ is convex.

Theorem 2

Inside a convex three dimensional domain D, there is a unique arrangement of a given finite collection of elements of the fluid, each with uniform prescribed values M_i, N_i and Θ_i and volumes τ_i, that is dynamically stable provided that

$$\sum_i \tau_i = \tau_D$$

where τ_D is the volume of D.

Proof

This is identical to the proofs of existence and uniqueness for the two dimensional problem in section 2; since convexity is an n-dimensional property.

Remarks

The proofs of continuous dependence on the data and extensions to general initial data are also identical to the two dimensional case.

4. Further results

The theorems proved in the previous two sections show that a great deal can be inferred about the solution of the Lagrangian conservation law form for the semi-geostrophic equations from geometrical arguments, in particular the identification of dynamically stable solutions with convexity of the solution surface. These arguments can be used to establish a number of additional results, some of which are described in this section. They will be stated for the two dimensional problem so that the arguments behind them can

be easily illustrated by accompanying diagrams. Similar results hold in three dimensions.

a. Non-existence of internal fronts

Theorem 3

If the potential vorticity is bounded, solutions to the semi-geostrophic equations cannot possess localized internal discontinuities in M and Θ. Any discontinuities must intersect the boundary of the domain.

Proof

The preceding results show that there is a unique convex solution surface $P(x,z)$. The gradients of this surface are M and Θ. Suppose that its gradient changes discontinuously at some isolated point. By changing coordinates, this can be treated as the origin. Thus suppose that at (0,0) ∇P changes from $(-g, 0)$ to $(g, 0)$ and that $P(0,0) = 0$. As x decreases from zero, the convexity of P means that $\frac{\partial P}{\partial x}$ must decrease from $-g$ and so P must satisfy

$$\begin{aligned} P(x,z) &\geq -gx \qquad x \leq 0 \\ P(x,z) &\geq gx \qquad x \geq 0 \end{aligned} \tag{4.1}$$

The magnitude of ∇P must increase as $\underset{\sim}{x}$ moves away from the origin and so

$$\nabla P(\underset{\sim}{x}) \cdot \underset{\sim}{x} \geq P(\underset{\sim}{x}) \quad \text{for all } \underset{\sim}{x} = (x,z) \tag{4.2}$$

Consider the rectangular circuit ABCD shown in Fig. 8(a), whose dimension l is chosen such that on segment BC

$$P(x) \geq g\delta \tag{4.3}$$

It must be possible to satisfy this condition because the discontinuity at the origin is isolated. At each point on ABCD, inequality (4.2) defines a set of values which cannot be taken by the gradients of P. This is shown as a region in gradient space in Fig. 8(b). Therefore, the potential circulation

$$C(\delta, l) \equiv \int_{ABCD} q \, dx\, dz$$

around ABCD is greater than that defined by the region in Fig. 8, ie.

$$C(\delta, l) \geq g^2\delta/l \tag{4.4}$$

However, since the potential vorticity everywhere is bounded; say by q_{max}, and the area of circuit ABCD is $2l\delta$; then

$$q \leq q_{max}$$

and

$$l^2 \geq g^2/q_{max} \tag{4.5}$$

Conversely, for any l , with $|l| < g\,q_{max}^{-1/2}$

$q \leq q_{max}$ implies that

$$P(x,z) \leq g\delta \tag{4.6}$$

somewhere on BC for all δ . Hence

$$P(0,l) = 0 \tag{4.7}$$
$$\text{if} \quad |l| < g\,q_{max}^{-1/2}$$

(4.1) and (4.7) together imply that the discontinuity in gradient must extend from the origin to $\pm g\,q_{max}^{-1/2}$ along the z axis . This argument can be repeated at the ends of this line segment to continue it as far as the boundary, thus completing the proof.

Analogous constructions may be made in the three-dimensional equations of Hoskins and Draghici (1977), initially by replacing circuit ABCD by the surface of a disc formed by rotating ABCD about AD. The corresponding excluded region of gradient space is the double-cone formed by rotating triangle $\alpha\beta\gamma$ about $\alpha\gamma$.

This theorem is a stronger statement than that in Hoskins and Bretherton (1972, p.16) as it prohibits fronts even at a finite

discontinuity of potential vorticity.

b. Behaviour of fluid elements on the boundary

Consider the case used to prove the theorem in section 2, where the data are piecewise constant and defined by values M_i and Θ_i on elements of area A_i.

Definition

The 'convex hull' of a set is the smallest convex set containing it. The convex hull of the set of points in gradient space with coordinates (M_i, Θ_i) is shown in Fig. 9, denote its boundary by G.

Theorem 4

Elements of the fluid with gradients on G remain in contact with the boundary of the convex physical domain Ω .

Proof

Suppose that there is a fluid element with associated values $\underset{\sim}{g}_* = (M_*, \Theta_*)$ such that $\underset{\sim}{g}_*$ lies on G, (Fig.10). If $\hat{\underset{\sim}{n}}$ is an outward pointing vector from G at $\underset{\sim}{g}_*$, then it is clear from Fig.10 that no fluid element can have gradients

$$\underset{\sim}{g}_* + \alpha \hat{\underset{\sim}{n}} \tag{4.8}$$

for any positive α . Suppose that this fluid element is not in contact with the boundary of the physical domain. Let $\underset{\sim}{x}_*$ be a point in the element. Then, the convexity of the solution surface means that

$$\frac{\nabla\underset{\sim}{p}(\underset{\sim}{x}) \cdot (\underset{\sim}{x} - \underset{\sim}{x}_*)}{| \underset{\sim}{x} - \underset{\sim}{x}_* |}$$

increases as $| \underset{\sim}{x} - \underset{\sim}{x}_* |$ increases. Thus if $(x - x_*)$ is chosen to be parallel to $\hat{\underset{\sim}{n}}$, $\nabla\underset{\sim}{p}(\underset{\sim}{x})$ must take on values of the form (4.8) with $\alpha > 0$ (Fig.11). This contradicts the hypothesis, and proves that the element must be in contact with the boundary.

Note that this does not mean that elements of the fluid originally on the boundary of the physical domain Ω remain there.

c. Characterization of data leading to frontogenesis

The process of frontogenesis studied here involves the pinching together of boundary points of the fluid, usually followed by the intrusion of the resulting contact discontinuity into the interior of the domain. When the initial potential vorticity is bounded, the possible locations of frontogenesis are restricted by the following theorem:

Theorem 5

Given initial values of M and Θ which are continuous functions of (x, z) and bounded potential vorticity q. The values of (M, Θ) for all points in Ω define a region $\mathcal{H}$ in (M, Θ) space with boundary Γ (Fig. 12). Then frontogenesis can only occur at points (x, z) whose associated values of (M, Θ) are on a concave portion of Γ. If Γ is convex, then M and Θ remain continuous functions of (x, z).

Proof

Because of the convexity of the solution surface $P(x,z)$ and the continuity of the initial data, points not on the boundary of Ω have gradients strictly inside the region $\mathcal{H}$ in (M, Θ) space. This follows immediately from Theorem 4. It a discontinuity forms at a later time, this corresponds to a fold in the solution surface where the gradients necessary to smooth out the fold are not in $\mathcal{H}$. Suppose that the points either side of the intersection of the front with the boundary of Ω are g_1 and g_2; these must both be on Γ in gradient space (Fig. 13). Then the missing gradients lie on the straight line connecting g_1 and g_2; since they are not in $\mathcal{H}$, they must correspond to a concave portion of Γ. If Γ is everywhere convex, no discontinuity can form and $P(x,z)$ stays smooth for all time.

This theorem generalises the observation of Hoskins (1971, p.143) that, in studies using the $\tan^{-1}(x)$ profile of potential temperature, the surface front forms in the warmer half of the fluid and the upper front forms in the colder half of the fluid, as seen schematically in Fig. 14.

d. Conservation of potential circulation

An important and rather paradoxical feature of these frontogenesis models concerns the total potential circulation, C, associated with a material cross-section. Equations (2.1) suggest that C should obey

the conservation law

$$\frac{D}{Dt}\left(C(t)e^{dt}\right) = 0 \qquad (4.9)$$

where the circulation round any circuit can only be changed by the basic deformation field. This would certainty be the case if the circuit in (x, z) space associated with C were purely advected. However, if C is associated with a circuit at the domain boundary, (4.9) is no longer obeyed after a front forms because the injection of a new region of (M, Θ) space, (the shaded area of Fig. 14(b)) into the interior of the fluid acts like an impulsive line source of potential vorticity. Thus, while the potential vorticity is conserved following each fluid parcel, the Eulerian mean potential vorticity may change.

e. Non-convex physical domains

The uniqueness of the solution to (2.1), (2.2) was proved only for convex physical domains Ω. Since the z coordinate is a function of pressure, this corresponds to a convex domain in (x, p). If the upper boundary is taken as $p=0$ there is no problem there. The lower boundary will normally have concave portions where surface pressure is locally a maximum, e.g. in valleys. In these regions distinct fluid elements can be trapped and not interact with the rest of the solution. These correspond to trapped stable layers near the ground in mountain valleys, (Fig. 15). Any of these can be interchanged without affecting the rest of the solution.

The non-uniqueness only affects the rearrangement in Theorem 2. Uniqueness of the solution to the true physical problem is obtained by assuming that there is no sudden rearrangement of stable layers.

f. Dynamically unstable data

The theorems proved in sections 2 and 3 give unique dynamically stable solutions. These can formally be found even from an initially given unstable configuration of the fluid. However, under these conditions the assumptions implicit in the semi-geostrophic approximation are no longer valid, since large ageostrophic accelerations will be generated. The consequent rearrangement of fluid elements is thus unlikely to proceed realistically. Thus the theorems should not be used as a substitute for convective parametrizations.

5. Finite difference solutions of the slightly viscous semi-geostrophic equations.

a. Basic requirement

The physical validity of the inviscid semi-geostrophic equations first breaks down because the Richardson number Ri falls below $\frac{1}{4}$ and turbulent mixing with associated large accelerations sets in (Hoskins and Bretherton (1972), Hoskins (1982)). If the effects of turbulence are, as usual in numerical models, represented by an eddy viscosity; then (2.1) becomes

$$\frac{DM}{Dt} + \alpha M = \epsilon \nabla^2 M \qquad (5.1)$$

$$\frac{D\theta}{Dt} = \epsilon \nabla^2 \theta$$

$$\frac{DA}{Dt} + \alpha A = 0$$

If ϵ is chosen so that Ri $> \frac{1}{4}$ for all time, then scale analysis implies that the vorticity will remain less than 10f, where f is the Coriolis parameter, and the neglect of the ageostrophic accelerations will remain valid for the 'viscous' problem. In reality the neglected terms will be crucial in the detailed dynamics of the turbulent mixing.

It is then hoped that the solution of (5.1), which will contain strong but smooth shear layers, will be a good 'broad-brushed' approximation to the true solution of the primitive equations, which contain turbulent shear layers. Such approximations are routine in other areas of computational fluid dynamics, for instance shock modelling.

It seems highly likely that, since the existence of solutions to the inviscid problem (2.1) has been proved, that solutions to (5.1) will also exist and be slightly smoothed versions of the inviscid solution. However, a proof of this is not yet available. A finite difference solution of (2.1) on a fixed grid could only be obtained by solving (5.1), with ϵ regarded as an artificial viscosity. Thus it is interesting to study finite difference solutions of (5.1) to check their convergence to the solutions derived in C by explicit construction; and the solutions obtained in C by finite difference approximations to the two dimensional primitive equations with artificial viscosity.

b. Solutions for the cross-front circulation

Consider first how (2.1) could be solved by finite difference methods. Write these in the standard dimensionless Eulerian form:

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + dv + w\frac{\partial v}{\partial z} + dx + u = 0$$

$$\frac{\partial \theta}{\partial t} + u\frac{\partial \theta}{\partial x} + w\frac{\partial \theta}{\partial z} = 0 \qquad (5.2)$$

$$\frac{\partial u}{\partial x} + d + \frac{\partial w}{\partial z} = 0 \qquad (5.3)$$

$$\frac{\partial \phi}{\partial x} = v$$

$$\frac{\partial \phi}{\partial z} = \theta$$

$$M = v + x$$

$$w = 0 \quad \text{at} \quad z = 0, 1$$

This set contains explicit predictions for v and θ, but u and w are determined implicity by the diagnostic equations (5.3). The continuity equation implies that u and w can be represented in terms of a streamfunction ψ by

$$u = -\alpha x - \partial\psi/\partial z$$
$$w = \partial\psi/\partial x \qquad (5.4)$$

Then (5.2) and (5.3) determine ψ implicitly by the equation

$$\frac{\partial}{\partial z}\left[\left(-\alpha x - \partial\psi/\partial z\right)\partial M/\partial x + \partial\psi/\partial x\, \partial M/\partial z + \alpha M\right]$$
$$= \frac{\partial}{\partial x}\left[\left(-\alpha x - \partial\psi/\partial z\right)\partial\theta/\partial x + \partial\psi/\partial x\, \partial\theta/\partial z\right] \qquad (5.5)$$

It is well known that this equation changes type according to the sign of the potential vorticity q defined by (2.4), being elliptic, parabolic or hyperbolic according to whether q is positive, zero, or negative. The discriminant of (5.5) is, in fact,

$$\left(\frac{\partial\theta}{\partial x} + \frac{\partial M}{\partial z}\right)^2 - 4\frac{\partial M}{\partial x}\frac{\partial\theta}{\partial z} \qquad (5.6)$$

When a discontinuity forms; $\frac{\partial M}{\partial x}$, $\frac{\partial M}{\partial z}$, $\frac{\partial\theta}{\partial x}$ and $\frac{\partial\theta}{\partial z}$ all tend to infinity in the fixed ratio

$$1 : S : S : S^2$$

where S is the slope of the discontinuity, $[\theta]/[M]$. If these ratios are taken for finite values of the gradients, (5.6) vanishes. This suggests that (5.5) will exhibit parabolic behaviour near a discontinuity, and may thus be difficult to solve.

Now consider the problem of finding ψ after a discontinuity has formed and the solution is similar to those shown in C. The velocity normal to the frontal surface must remain continuous, but

the tangential velocity may be discontinuous. In the special case solved in C, where lines of constant M and θ coincide, it is possible to construct an equation for ψ along the isotherms. The geostrophic and hydrostatic relations (5.3) then show that the slope of an isotherm is $\frac{d\theta}{dM}$ where the solution is smooth, and $[\theta]/[M]$ at a frontal surface. Since values of θ are conserved and M is proportional to $e^{-\alpha t}$ following fluid particles, the slopes of isotherms are proportional to $e^{\alpha t}$. Thus the streamfunction ψ obeys the equation

$$\frac{\partial^2 \psi}{\partial l^2} = \alpha s \qquad (5.7)$$

where l is a coordinate along an isotherm with slope s.

The difficulty of solving (5.7) is shown in Fig. 16. If an isotherm intersects both boundaries, (5.7) can be solved with the boundary conditions $\psi = 0$ at $z = 0, 1$. Otherwise, the problem can only be solved by establishing the geometry of intersections of isotherms and requiring ψ to be continuous at them. The well-posedness of this is guaranteed by Theorem 1. However, it seems impossible to solve such a problem directly where the discontinuities are smeared over grid-lengths. The viscous problem (5.1) would become similarly difficult to solve as ϵ is reduced. It therefore seems necessary to use an iterative method to calculate ψ.

This behaviour in the inviscid case suggests that the solution is not locally determined, even though the ageostrophic accelerations are small. The method of solution must therefore allow information to propagate a large distance at each timestep.

c. Iterative method of solution

Solve (5.2), (5.3) by splitting as follows:

Step 1:
$$\frac{\partial}{\partial t}(v+x) - \alpha x \frac{\partial}{\partial x}(v+x) + \alpha(v+x) = 0$$
$$\frac{\partial \theta}{\partial t} - \alpha x \frac{\partial \theta}{\partial x} = 0 \qquad (5.8)$$

At the end of this step v and θ do not satisfy (5.3).

Step 2: Introduce an imaginary iteration time τ and solve

$$\frac{\partial (v+x)}{\partial \tau} + u' \frac{\partial (v+x)}{\partial x} + w' \frac{\partial (v+x)}{\partial z} = \epsilon' \nabla^2 (v+x)$$
$$\frac{\partial \theta}{\partial \tau} + u' \frac{\partial \theta}{\partial x} + w' \frac{\partial \theta}{\partial z} = \epsilon' \nabla^2 \theta \qquad (5.9)$$

where u' and w' evolve in τ according to

$$\frac{\partial u'}{\partial \tau} + \frac{\partial \phi}{\partial x} - v = K \nabla^2 u'$$
$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0 \quad , \quad w' = 0 \text{ at } z = 0, 1 \qquad (5.10)$$
$$\frac{\partial \phi}{\partial z} = \theta$$

If we assume that (5.9), (5.10) have to be integrated for a time $\tau = \eta t$ to obtain convergence, then the artifical viscosity ϵ' must be set to $\eta^{-1}\epsilon$.

The viscous term in (5.10) ensures that the magnitudes of u' and hence w' decrease with τ . Since $(v+x)$ and θ are conserved following the incompressible motion (u, w) under (5.9); this iteration converges to the unique rearrangement proved possible in Theorem 1; where areas and values of $(v+x)$ and θ are preserved and the conditions (5.3) satisfied.

This iteration differs from a solution of the primitive equations for this problem in that (5.9, (5.10) have to be integrated in τ for longer than real time to obtain convergence. This is

because the internal gravity waves carrying out the adjustment only have a phase speed of greater than 10 ms^{-1} for the first five internal modes. The solution of the semi-geostrophic equations requires information on detailed vertical structure to be transmitted equally fast for all internal modes. If the semi-geostrophic solution is to be physically valid, a fast mechanism must exist for transferring the information in the atmosphere.

A similar method can be used for the three dimensional semi-geostrophic equations. Though the direct proof of Theorem 2 only holds on an f plane, the iteration can be written for variable f as follows, where a dimensional form is used.

Step 1:

$$\begin{aligned} \frac{\partial u_g}{\partial t} + \underset{\sim}{u}_g \cdot \nabla u_g &= 0 \\ \frac{\partial v_g}{\partial t} + \underset{\sim}{u}_g \cdot \nabla v_g &= 0 \\ \frac{\partial \theta}{\partial t} + \underset{\sim}{u}_g \cdot \nabla \theta &= 0 \\ \underset{\sim}{u}_g &= \left(-f^{-1} \frac{\partial \phi}{\partial y}, f^{-1} \frac{\partial \phi}{\partial x}, \right) \end{aligned} \qquad (5.11)$$

such that

$$\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} + \frac{\partial w_g}{\partial z} = 0$$

Step 2:

$$\begin{aligned} \frac{\partial u_g}{\partial \tau} + \underset{\sim}{u}' \cdot \nabla u_g - f v' &= \epsilon' \nabla^2 u_g \\ \frac{\partial v_g}{\partial \tau} + \underset{\sim}{u}' \cdot \nabla v_g + f u' &= \epsilon' \nabla^2 v_g \\ \frac{\partial \theta}{\partial \tau} + \underset{\sim}{u}' \cdot \nabla \theta &= 0 \end{aligned} \qquad (5.12)$$

where

$$\frac{\partial u'}{\partial \tau} + \frac{\partial \phi}{\partial x} - f v_g = k \nabla^2 u'$$

$$\frac{\partial v'}{\partial \tau} + \frac{\partial \phi}{\partial y} + f u_g = k \nabla^2 v' \qquad (5.13)$$

$$\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} = 0$$

$$\frac{\partial \phi}{\partial z} = g\theta/\theta_0$$

On an f plane, Theorem 2 shows that this scheme will converge to a solution where u_g, v_g satisfy the definitions

$$\frac{\partial \phi}{\partial x} = f v_g \quad , \quad \frac{\partial \phi}{\partial y} = - f u_g$$

If f is variable, convergence can only be established by experiment.

d. Lower boundary condition

The boundary condition $w = 0$ at $z = 0$ and 1 corresponds to assuming rigid boundaries at $p = 0$ and p_0, where p_0 is a fixed surface pressure. Heckley and Hoskins (1982) used more realistic boundary conditions. It is also desirable to be able to use (5.11) to (5.13) in sigma coordinates, to avoid the use of the Boussinesq approximation as in (5.1). The correct lower boundary condition at $z = 0$ is

$$\frac{\partial p_*}{\partial t} + u \frac{\partial p_*}{\partial x} = - \rho_* w \qquad (5.14)$$

where

$$v = \frac{\partial \phi_*}{\partial x} \qquad (5.15)$$

and p_*, ϕ_* are values of p and ϕ at $z = 0$.

The condition (5.14) can be approximated by

$$\frac{\partial \phi_*}{\partial t} + u \frac{\partial \phi_*}{\partial x} = -w \tag{5.16}$$

and by the continuity equation

$$-w = -\alpha - \int_0^1 \frac{\partial u}{\partial x}\, dz \tag{5.17}$$

The effect of using this condition instead of the rigid wall condition in Theorem 1 is that the areas of segments of constant M and θ are not exactly conserved, since the geopotential at $z = 0$ must adjust to satisfy (5.15). Since p_* only has to vary by about $\pm$ 5% of its mean value under normal conditions, this represents only a small relative change in areas of segments. It should therefore be possible to extend the explicit construction used in the proof of Theorem 1 iteratively. This involves the following procedure:

(i) Solve with fixed areas, and obtain v at z = 0

(ii) Calculate ϕ_* and hence p_* from (5.15)

(iii) Correct areas of segments in contact with lower boundary to allow for transfer of mass across z = 0.

(iv) Solve again with new areas, and repeat to convergence.

The result of implementing this procedure in the case studied in C will be shown later. Convergence of this iteration requires the property from Theorem 1, that the solution depends continuously on the areas of the elements.

The condition (5.14) can be implemented in the iterative procedure (5.9), (5.10) by modifying (5.10) as follows:

$$\frac{\partial u'}{\partial \tau} + \frac{\partial \phi}{\partial x} - v = k \nabla^2 u' \tag{5.18}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0$$

$$w' = 0 \quad \text{at} \quad z = 1, \quad \phi = \phi_* \quad \text{at} \quad z = 0$$

$$w' = -\left(\frac{\partial \phi_*}{\partial t} + u \frac{\partial \phi_*}{\partial x} \right) \quad \text{at} \quad z = 0$$

$$\frac{\partial \phi}{\partial z} = \theta$$

e. Finite difference approximations, and horizontal boundary condition

The iterative methods (5.8) to (5.10) for the two dimensional deformation problem, (5.11) to (5.13) for the three dimensional semi-geostrophic model, and (5.18) with the correct lower boundary condition can all be solved by standard finite difference procedures. For the data used in C, it was found sufficient to use a forward timestep in (5.8) and a backward-implicit scheme in (5.9) and (5.10).

The horizontal boundary conditions must be precisely specified for the finite difference calculations. These have to be consistent with (2.3); and so the required conditions are

$$\left.\begin{aligned} u' &= 0 \\ \frac{\partial \theta}{\partial t} &= \frac{\partial (v_+ x)}{\partial t} = 0 \end{aligned}\right\} \quad \text{at} \quad x = \pm e^{-\alpha t} \qquad (5.19)$$

The second condition ensures that spurious values of θ and $(v_+ x)$ are not generated at the boundary.

It is important to note the difference in the way in which artificial viscosity is used in (5.8) to (5.10). A large viscosity is used in (5.10) to accelerate convergence, it has no effect on the history carrying variables since the "velocities" u', w' are not

physically realistic, but just iteration parameters. The artificial viscosity in (5.9) is used to capture the fronts, and since Theorem 1 proves the existence of a solution to the inviscid problem, this viscosity can be reduced as the mesh is refined, so ϵ will be proportional to Δx^2. In a 3 dimensional PE integration there is a single set of velocity components and there is no way of applying different diffusion coefficients to geostrophic and ageostrophic winds; the best that can be done is divergence damping (Dey (1978)). In two dimensions, as here, it is possible to use different diffusion coefficients for u and v.

6. Results

a. Zero potential vorticity

The first case for which results are presented uses the same data as in C. Plots of this data are given in Figs. 17 and 18. Though zero potential vorticity is not the most physically realistic case, it has not so far proved possible to implement the exact geometrical construction used to prove Theorem 1 as a practical computational tool in any other case. Therefore this data was used to allow a comparison of finite difference solutions with the exact construction. Using this data yields a discontinuity at the boundaries when $\alpha t = 1.5$. The initial data is derived from

$$u = -\alpha x$$
$$v = \underline{X}(x,z) - x \qquad (6.1)$$
$$\theta = 1 + \tanh^{-1}(S\,\underline{X}(x,z))$$

where $\underline{X}(x,z)$ is defined by

$$\frac{x - \underline{X}(x,z)}{z - \tfrac{1}{2}} = \frac{-S}{(1+(S\,\underline{X}(x,z))^2)} \qquad (6.2)$$

and boundary conditions (5.14) at $z = 0$, $w = 0$ at $z = 1$ and (2.3) at $x = \pm e^{-\alpha t}$. The models compared are

(i) Explicit construction as in Theorem 1 with iterative modification to treat the lower boundary condition (referred to as EC) This model will exactly reproduce the Hoskins and Bretherton solution up to when the front forms.

(ii) Finite difference solution to the primitive two-dimensional equations with artificial viscosity added (referred to as PE) Second order central differencing is used throughout, as in C.

(iii) Finite difference semi-geostrophic solution of (5.8) and (5.18) (referred to as SG). The same finite difference approximations are used as in (ii), with a 200 x 20 grid.

The results from EC are given in Figs. 19 and 20, which show the potential temperature and long-front component of velocity. The comparison between PE and SG is complicated by the fact that the proved possible to run SG with a much lower value for the artifical viscosity (190 $m^2 s^{-1}$) than PE (2.5×10^4 $m^2 s^{-1}$), if the same coefficient is used for all variables in PE. Since the frontal structure is sensitive to this coefficient e.g. Williams (1974), two cases are presented.

In Figs. 21 to 23 the potential temperature, long-front velocity and vertical velocity are shown from an integration of PE using a coefficient of 190 $m^2 s^{-1}$ for v and θ but an increased value for u. These can be compared directly with Figs 24 to 26, obtained using SG. No vertical diffusion was used, so that there is no attempt to model atmospheric boundary layer effects in this work.

The results from PE look rather different to those presented in C because of the different diffusion coefficients used here. The overall width of the frontal zone is greater in PE than in EC and SG. The potential temperature comparison (Figs. 19, 21 and 24) shows that above and below the frontal surface the slopes of the isotherms in PE disagree with those in SG and EC. Near $z = \frac{1}{2}$, the slopes are much shallower in EC than in either PE or SG, which are similar. In the case of PE, this

difference remains as the resolution is further increased, in SG the slopes show slow convergence towards the EC results. This suggests that the errors at this level are caused by lack of resolution. The problem only becomes serious after the initial formation of the front. The dimensionless slope given by thermal wind balance is given by $\partial M/\partial\theta$ which for the data (6.1) is proportional to $\sec^2(X)$. This has a sharp minimum near the $\theta = 0^\circ$ isotherm. Smearing of the profiles by truncation error could result in this minimum not being well resolved after a long integration period.

The differences in the long-front velocity (Figs. 20, 22 and 25) are less obvious. The magnitude of v is consistently greater in PE than in SG and EC. The jump in v extending away from the frontal zone in EC is because of the use of data piecewise constant in M and the change in alignment of the segments because of the presence of the front. This indicates a lack of realism in the choice of initial data. The effect is lost in the finite difference integrations by the numerical smoothing. The maximum vorticity in SG and PE is about 12f, roughly the maximum amount allowed by the validity of the SG equations discussed by Hoskins and Bretherton.

The vertical velocities are considerably different in character in PE and SG. In PE there are regions of large values on either side of the front. In Fig. 23 they fill much of the space between the vertical lines defining the overall width of the frontal zone. In SG (Fig. 26), the large values are confined much closer to the frontal zone and to separate regions extenting almost vertically from the points where the frontal zone intersects the boundaries. There are very large values near the boundaries. It has not yet been possible to calculate vertical velocities from EC. However, in order to produce the effect of propagating the front into the interior of the fluid large values will be necessary concentrated near the boundaries.

The experiments were repeated using the same diffusion coefficient for all variables in PE. The value required to allow a stable integration was 2.5×10^4 $m^2 sec^{-1}$. The SG integration was then repeated using this value. The results for the θ fields are shown in Figs. 27 and 28. The differences are less clear-cut than with the lower value of diffusion. The vertical velocities (not shown) from both integrations are much smoother. Those from SG have larger values closer to the boundary than from PE and are less smooth. The reason why the PE integration fails with less diffusion appears to be unrelated to the roughness of any of the fields but concerned with the development of a singularity at the boundary.

In Fig. 29 we show a comparison between $\Delta\theta$ across the front as given by EC and the maximum value of $\partial\theta/\partial x$ as given by SG and PE. The effect of the diffusion coefficient is very marked. Both SG and PE with low diffusion underestimate the gradients in the mid-troposphere. SG gives consistently 10 to 20 per cent greater values except near the boundary. The value of $\Delta\theta$ across the jump decreases very rapidly away from the boundary even in EC, in common with observations of frontal strength.

Though the use of data with q = 0 was necessary because of the inability at present to implement the EC construction in any other case, typical atmospheric data has positive q. There is a danger that the case q = 0 may present special problems in a numerical model. The data is only neutrally stable to symmetric instability, and numerical truncation error may create regions of negative q. We therefore present a comparison between the θ fields given by SG and PE for positive q in Figs. 30 and 31. The same qualitative differences remain as in the q = 0 case.

The final experiment shown is to solve SG using data with a discontinuity in q, representing a tropopause. The data used has zero potential vorticity for $0 \leq z \leq \frac{3}{4}$, and large positive potential vorticity for $\frac{3}{4} < z < 1$. A rigid wall condition is imposed at z = 1,

and θ is not constant on $z = 1$. This is intended as a comparison with the results of Hoskins and Bretherton on upper tropospheric fronts. In agreement with them, and with Theorem 3, no discontinuity is formed at the tropopause. The model produces a statically unstable profile just below the tropopause, (Fig. 32). This still occurs if the data is modified to have $q \neq 0$ in the troposphere, and appears to be caused by the smearing of the sharp discontinuity in temperature gradient by the finite difference scheme. The long-front velocity field (Fig. 33) shows a well developed jet just below the tropopause with a secondary maximum near $z = 1$. There is no tendency to form a discontinuity in this field near the jets.

7. Discussion

In this paper the extension to the model of Hoskins and Bretherton (1972) introduced in C is further developed. By extending it beyond the initial formation of the front, a solution can be obtained which reflects a balance between frontogenetic and frontolytic forces. It is more reasonable to compare this solution with observations of a front than to use the solution before the front forms, and it would be interesting to use it to repeat the detailed comparisons of Blumen (1980) and Ogura and Portis (1982). The results shown for the vertical velocity suggest that closer agreement with observation may be obtained.

The solution obtained by Lagrangian conservation law arguments has been shown to agree reasonably with that obtained by finite difference solution of the semi-geostrophic equations. The latter model can easily be extended to include boundary layer and moist processes, but it would also be interesting to attempt this with the Lagrangian model. In view of the powerful theorems that can be proved using the Lagrangian model, it is important to study its physical validity in much more detail; a clear starting point is the scaling assumption that turbulent mixing sets in before the ageostrophic accelerations become large. It is to be expected that this model would only be in agreement with observations when this assumption is correct. The predicted separation effect at the boundary may provide a good test.

The differences between the SG and PE finite difference solutions are disturbing. Some of the difficulty must be caused by the use of two-dimensional models. The real turbulent breakdown, which should be predicted by PE, is three dimensional. It would therefore be interesting to extend the comparison into three dimensions. However, since the PE can describe a scale collapse in three dimensions; the success of a PE model depends on a correct parametrization of sub grid-scale turbulence, since the viscous dissipation scale cannot be resolved. It is therefore not surprising that it is difficult to model a solution where the small scales are generated in as organised a way as near a front.

Acknowledgment

The authors wish to thank two referees for constructive comments, and Mr C A Parrett for carrying out some of the computer runs.

REFERENCES

Blumen, W., 1980: A comparison between the Hoskins-Bretherton model of frontogenesis and the analysis of an intense surface frontal zone. J. Atmos. Sci. 37, 64-77.

Cullen, M.J.P., 1983: Solutions to a model of a front forced by deformation. Quart. J. Roy. Meteor. Soc., to appear.

Dey, C.H., 1978: Noise suppression in a primitive equation model. Mon. Weath. Rev., 106, 159-173.

Heckley, W.A. and Hoskins, B.J. 1982: Baroclinic waves and frontogenesis in a non-uniform potential vorticity semi-geostrophic model. J. Atmos.Sci. 39, 1999-2016.

Hoskins, B.J., 1971: Atmospheric frontogenesis: some solutions. Quart. J. Roy. Meteor. Soc., 98, 532-541.

" " 1982: The mathematical theory of frontogenesis. Annual Reviews of Fluid Mechanics, Vol. 14, Annual Reviews Inc., 131-151.

Hoskins, B,J. and Bretherton, F.P. 1972: Atmospheric frontogenesis models: Mathematical formulation and solution. J. Atmos. Sci., 29, 11-37.

Hoskins, B.J. and Draghici, I. 1977: The forcing of ageostrophic motion according to the semi-geostrophic equations and in an isentropic coordinate model. J. Atmos. Sci. 34, 1859-1867.

Lorenz, E.N., 1969: The predictability of a flow which possesses many scales of motion. Tellus 21, 289-307.

Ogura, Y. and Portis, D. 1982: Structure of the cold front observed in SESAME-AVE III and its comparison with the Hoskins-Bretherton frontogenesis model. J. Atmos. Sci. 39, 2773-2792

Orlanski, I and Ross, B. B. 1977: The circulation associated with a cold front. Part I: Dry case. J. Atmos. Sci., 34, 1619-1633.

Ross, B. B. and Orlanski, I. 1982: The evolution of an observed cold front. Part I: numerical simulation. J. Atmos. Sci., 39, 296-327.

Rockafellar, R.T. 1970: Convex analysis. Princeton University Press.

Sanders, F. 1955: An investigation of the structure and dynamics of an intense surface frontal zone. J. Meteor., 12, 542-552.

Williams, R. T. 1974: Numerical simulation of steady-state fronts. J. Atmos. Sci., 31, 1286-1296.

LIST OF FIGURES

Figure 14 (a) Schematic illustration of a stage in the growth of the frontal discontinuties for uniform potential vorticity. A $\tan^{-1}(x)$ temperature profile is used in (x, z) space.

(b) The same stage in gradient space. Points a',b',c',d',e' correspond to points a,b,c,d,e in real space (Fig. 14(a)). The shaded regions correspond to the frontal discontinuties.

Figure 15 A particular example of the non-uniqueness of solutions when the domain is not convex. M is uniform. θ takes either of two values θ_1 (shaded) or θ_2 ($> \theta_1$).

Figure 16 Solution of equation (5.7) assuming zero potential vorticity, with potential isotherms shown. ψ is determined along AB by continuity of ψ at A, B.

Figure 17 Initial potential temperature data given by equation (6.1).

Figure 18 Initial y-velocity field given by equation (6.1).

Figure 19 Potential temperature after 13.9 hr for data (6.1), EC.

Figure 20 y-velocity after 13.9 hr for data (6.1), EC.

Figure 21 Potential temperature after 13.9 hr, PE, 200 x 20 grid, diffusion coefficient 190 $m^2 s^{-1}$ for v and θ, 3.75×10^4 $m^2 s^{-1}$ for u.

Figure 22 y-velocity after 13.9 hr, model as Fig. 21.

Figure 23 Vertical velocity after 13.9 hr, model as Fig. 21.

Figure 24 Potential temperature after 13.9 hr., SG, 200 x 20 grid, diffusion coefficient 190 $m^2 s^{-1}$.

Figure 25 y-velocity after 13.9 hr, model as Fig. 24.

Figure 26 Vertical velocity after 13.9 hr, model as Fig. 24.

Figure 27 Potential temperature after 13.9 hr, PE, model as Fig. 21 with diffusion coefficient 2.5×10^4 $m^2 s^{-1}$ for all variables.

Figure 28 Potential temperature after 13.9hr, SG, model as Fig. 24 with diffusion coefficient 2.5×10^4 $m^2 s^{-1}$.

Figure 29 (A) Graph of temperature difference across front, EC.

(B) Graph of the maximum $\partial\theta/\partial x$ against z for SG and PE with diffusion coefficients 190 $m^2 s^{-1}$ for θ.

(C) As (B), diffusion coefficients 2.5×10^{4} $m^2 s^{-1}$.

Figure 30 Potential temperature after 13.9 hr, model as Fig. 21, data with q = constant.

Figure 31 Potential temperature after 13.9 hr., model as Fig. 24, data as Fig.30.

Figure 32 Potential temperature after 13.9 hr., model as Fig. 24, data set with discontinuity in potential vorticity.

Figure 33 y-velocity after 13.9 hr., model and data as Fig. 32.

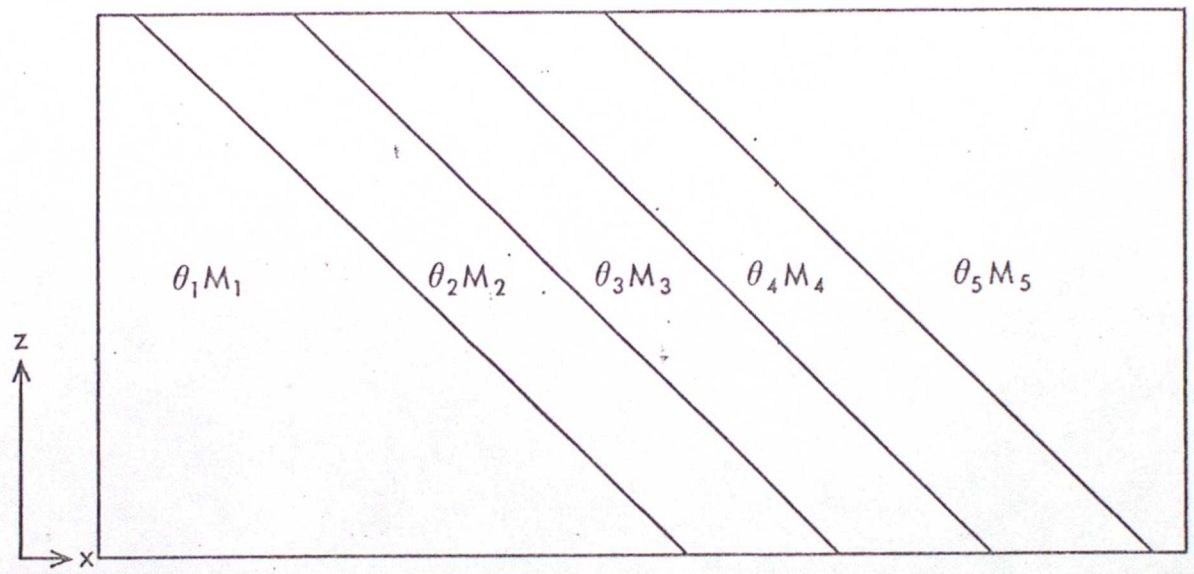


FIG 1

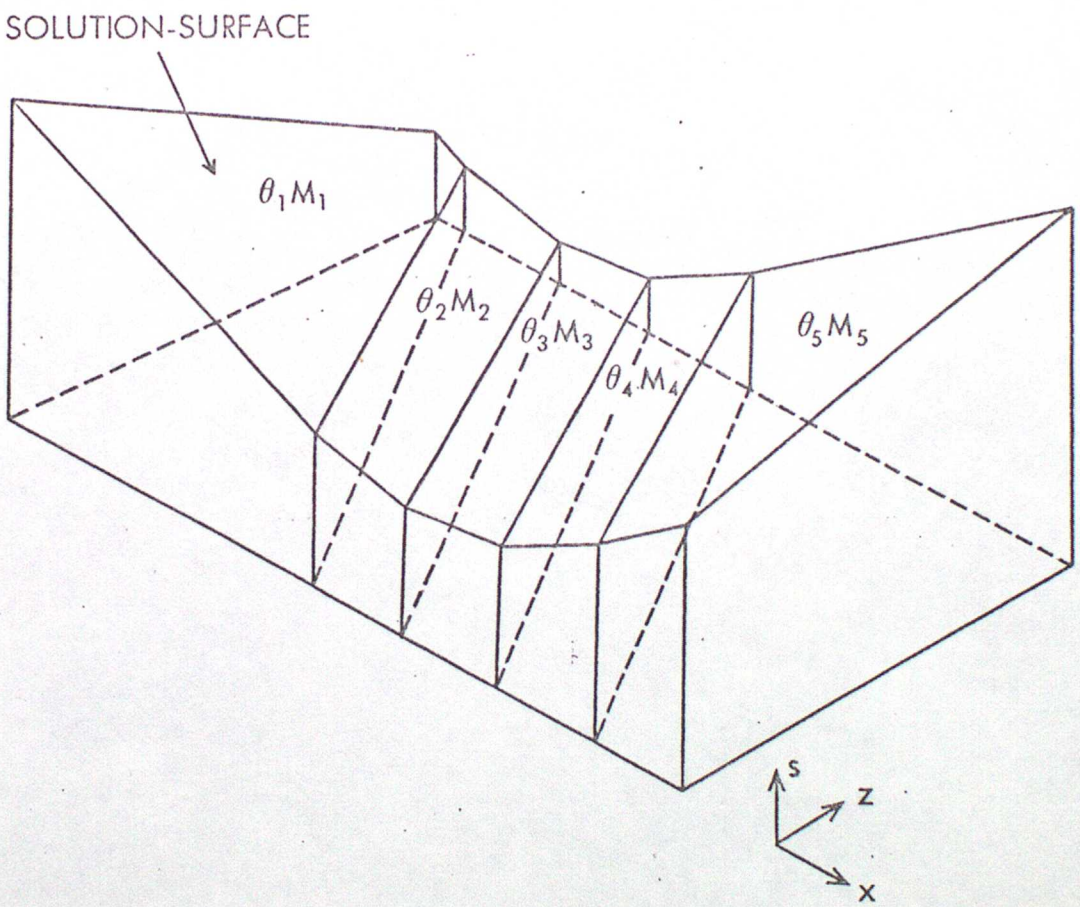
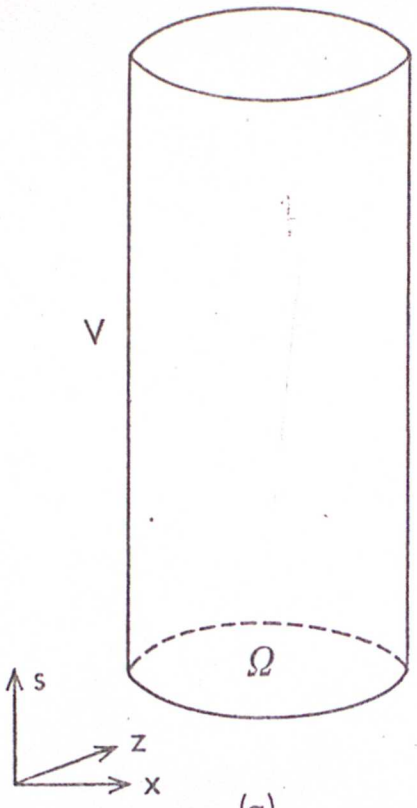
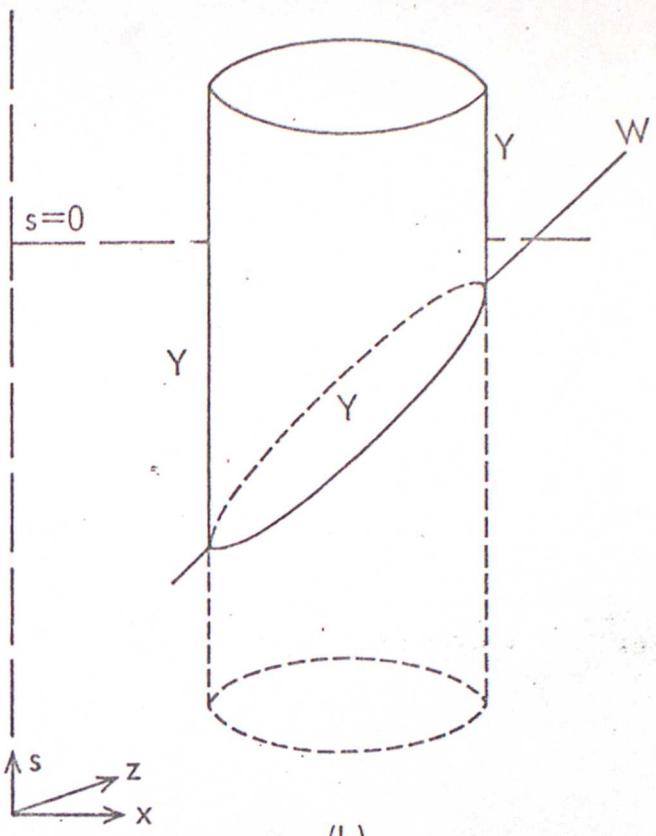

SOLUTION-SURFACE
$\theta_1 M_1$
$\theta_2 M_2$
$\theta_3 M_3$
$\theta_4 M_4$
$\theta_5 M_5$
S
Z
X

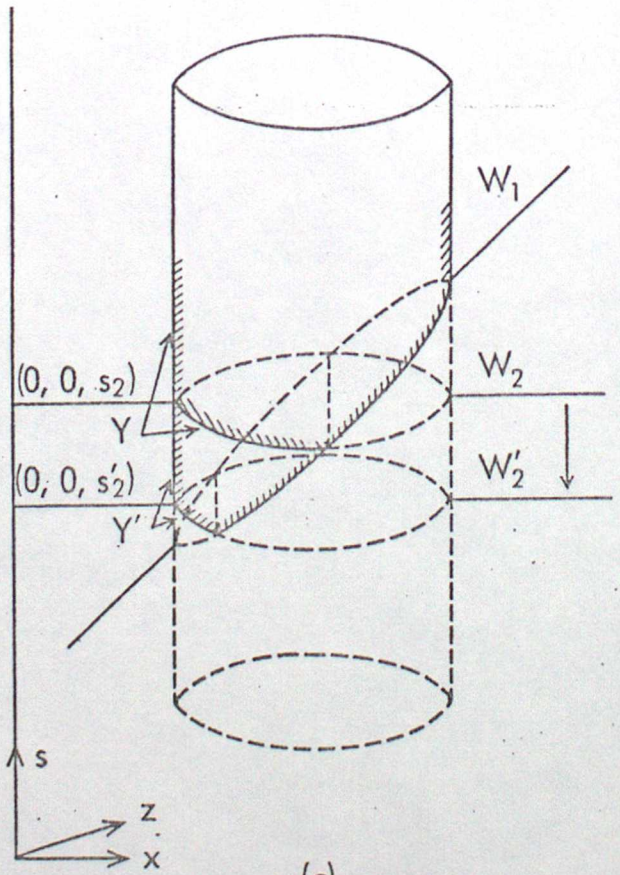
FIG 2



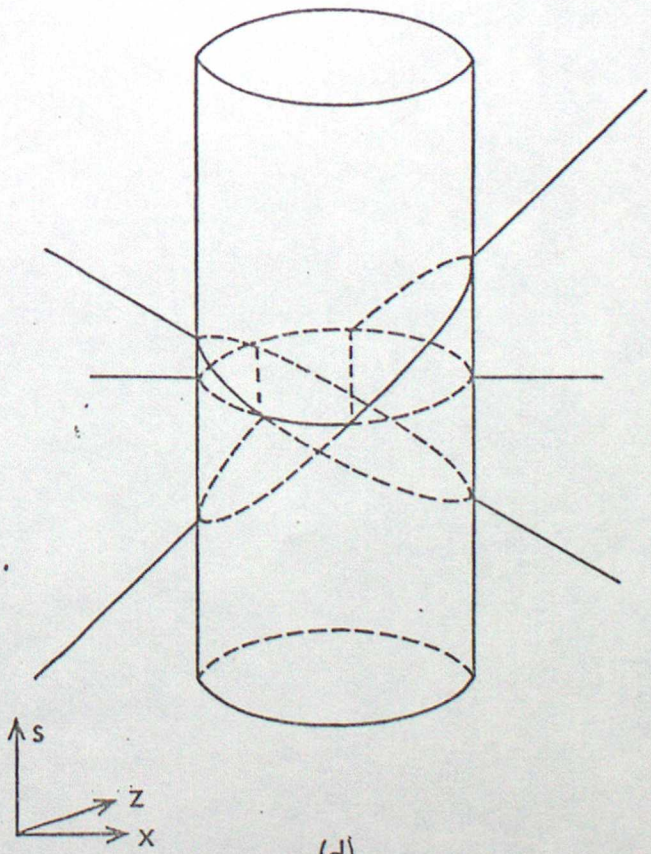
(a)



(b)



(c)



(d)

FIG 3

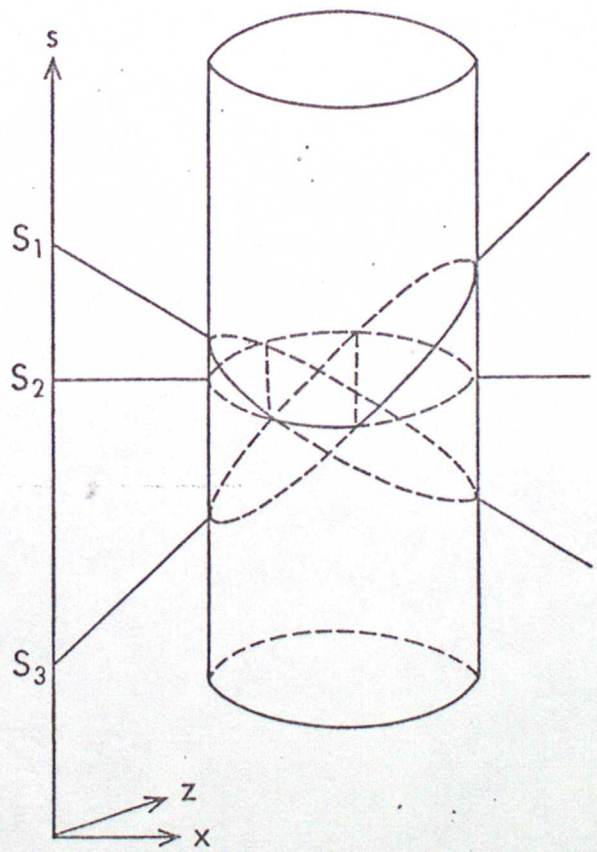


FIG 4

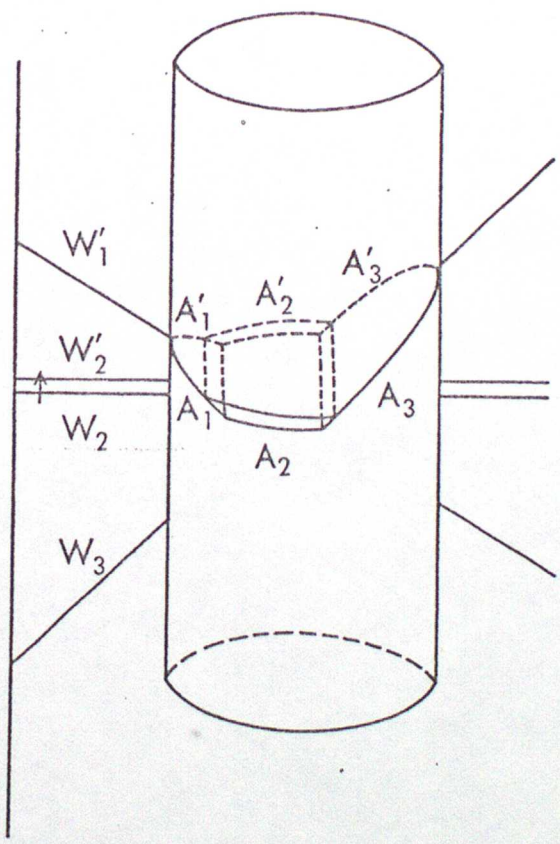


FIG 5

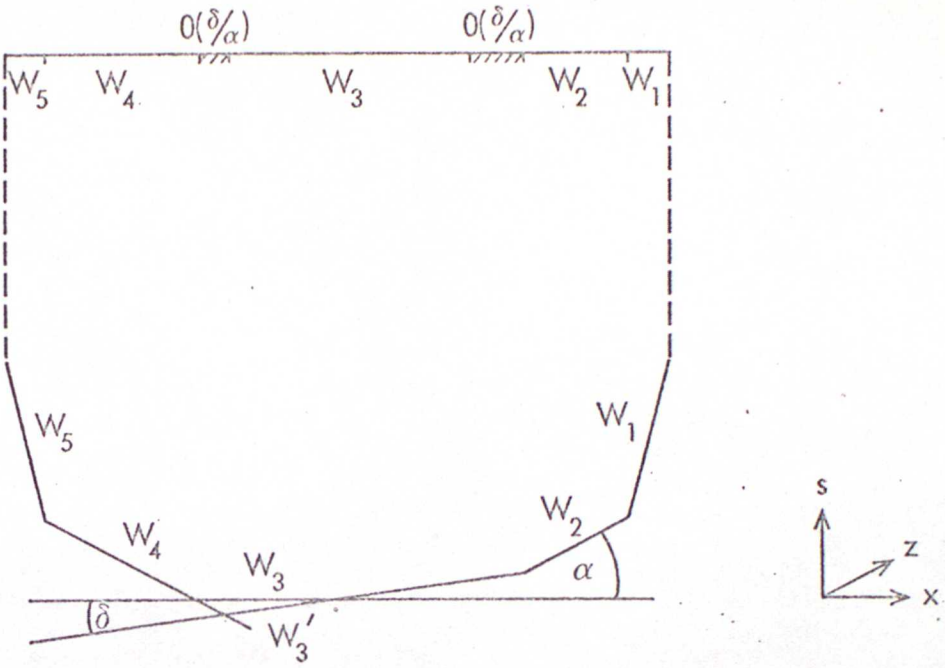

$O(\delta/\alpha)$
$O(\delta/\alpha)$
W_5
W_4
W_3
W_2
W_1
W_5
W_1
W_4
W_2
W_3
α
δ
W_3'
S
Z
X

FIG 6

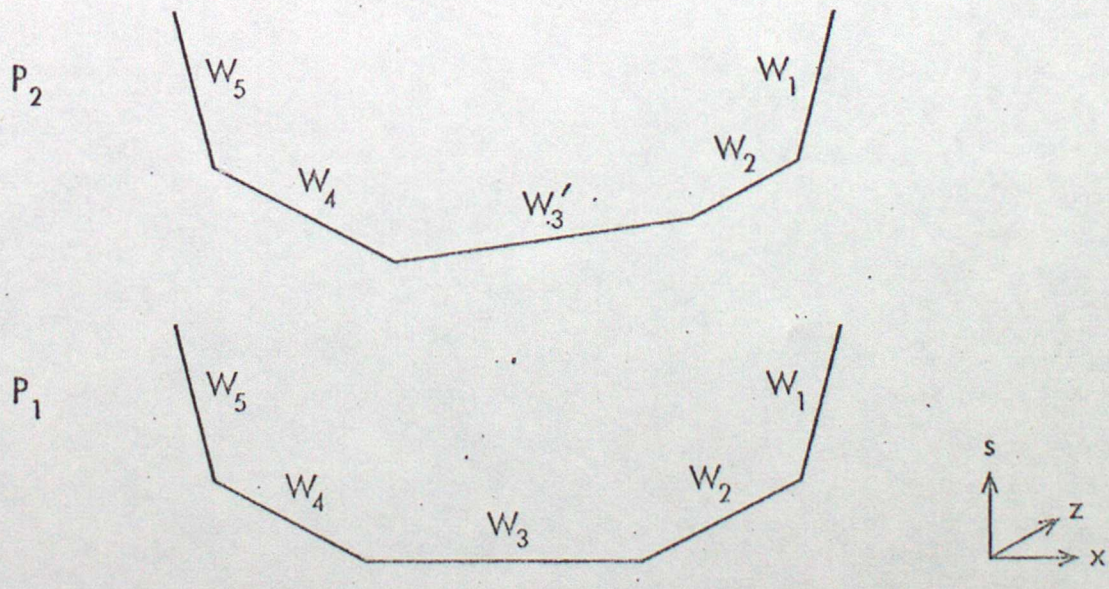
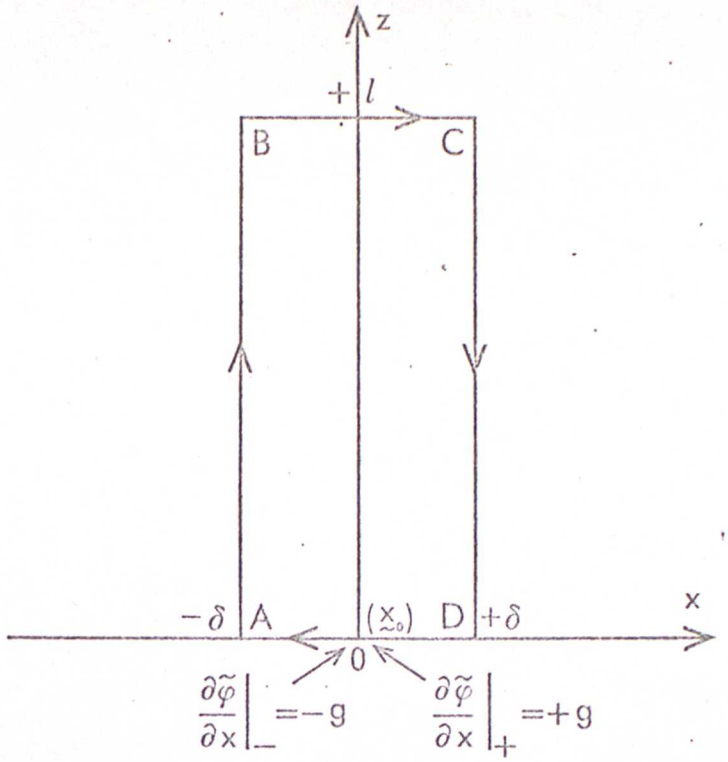
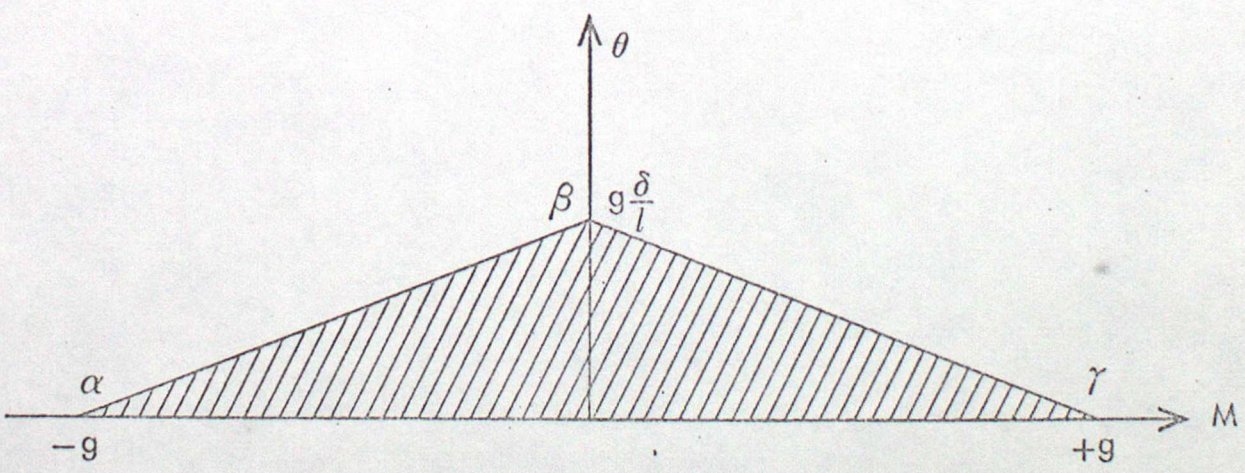

P_2
W_5
W_1
W_2
W_4
W_3'
P_1
W_5
W_1
W_4
W_2
W_3
S
Z
X

FIG 7



(a)



(b)

Fig 8

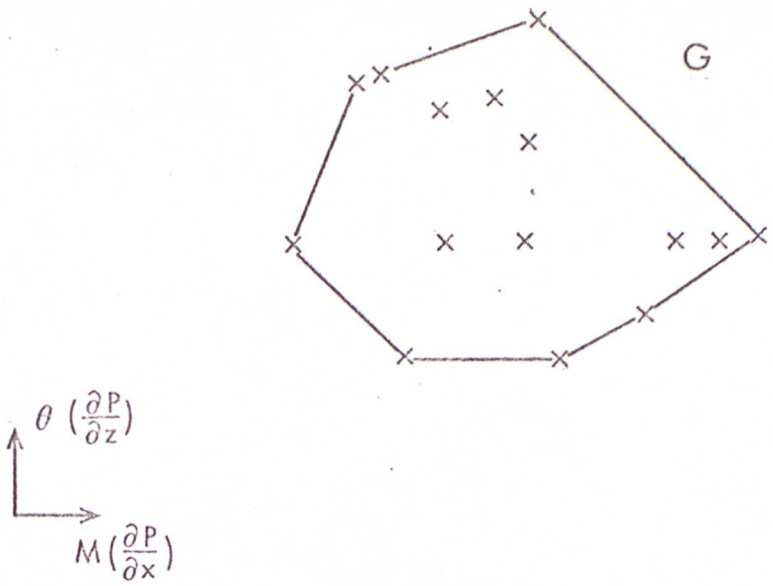

G
$\theta\left(\frac{\partial P}{\partial z}\right)$
$M\left(\frac{\partial P}{\partial x}\right)$

FIG 9

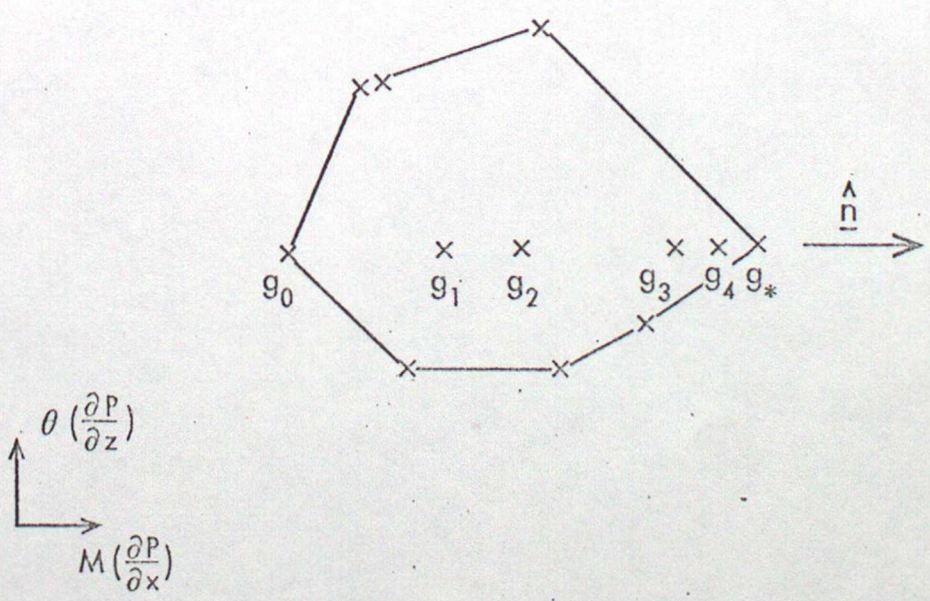

g_0
g_1
g_2
g_3
g_4
g_*
$\hat{n}$
$\theta\left(\frac{\partial P}{\partial z}\right)$
$M\left(\frac{\partial P}{\partial x}\right)$

FIG 10

FIG 11

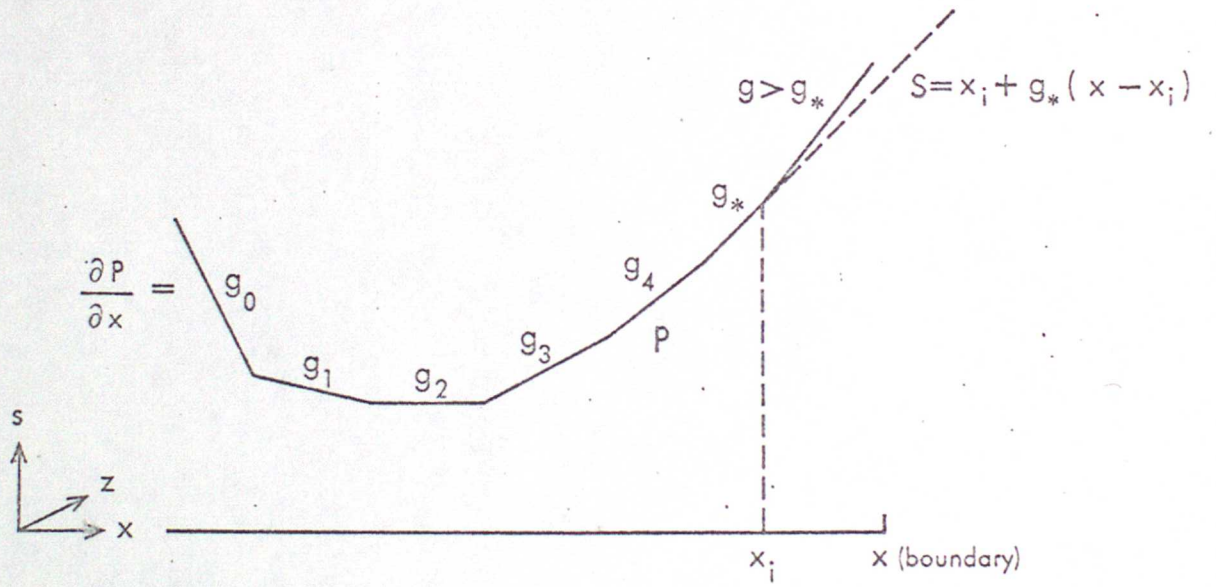


FIG 2

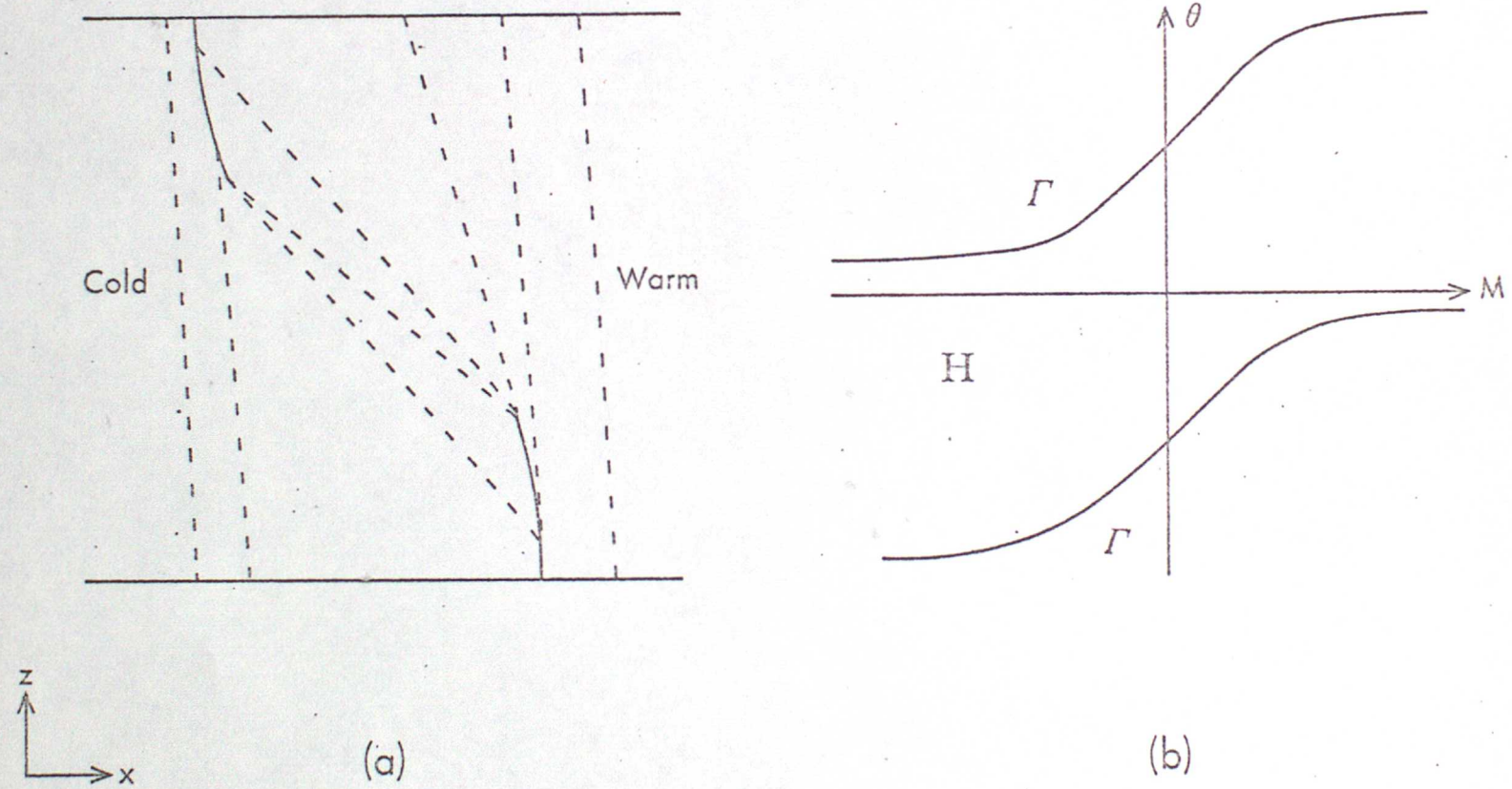
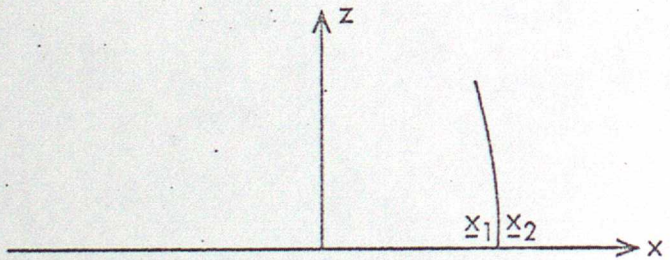
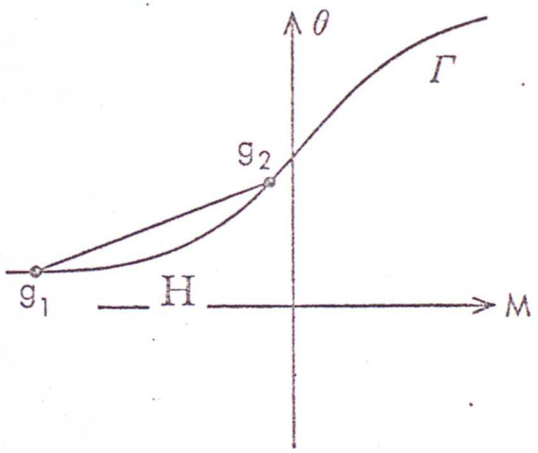
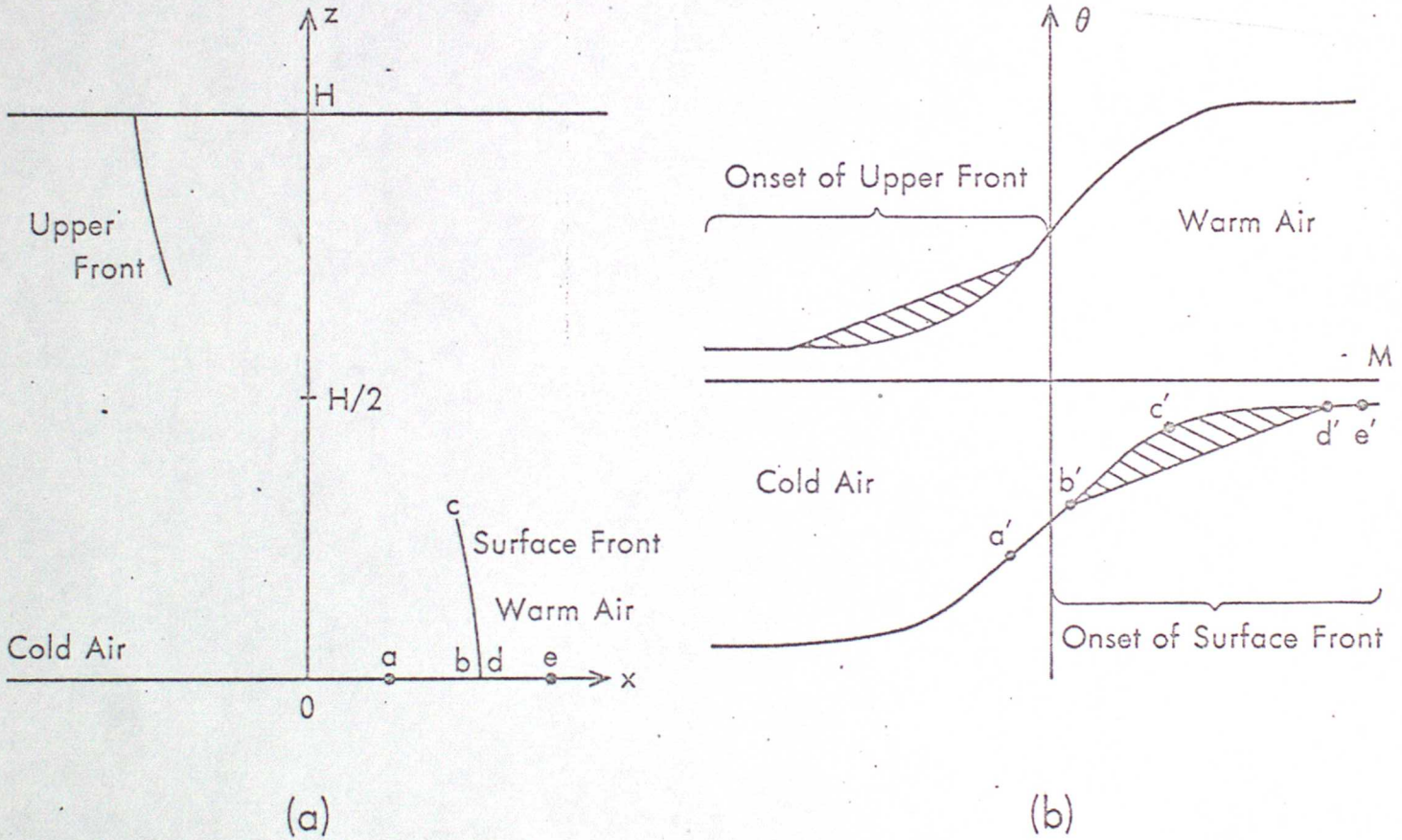
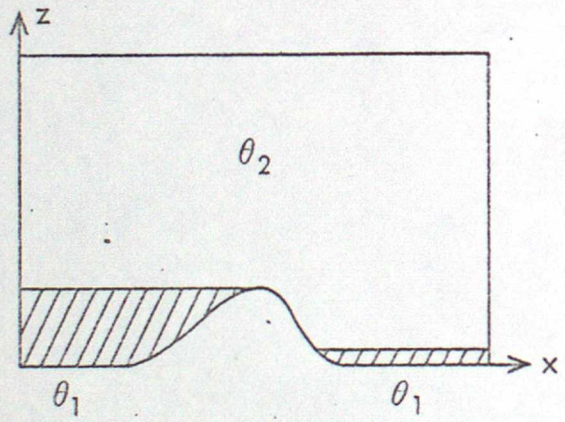


Fig 13

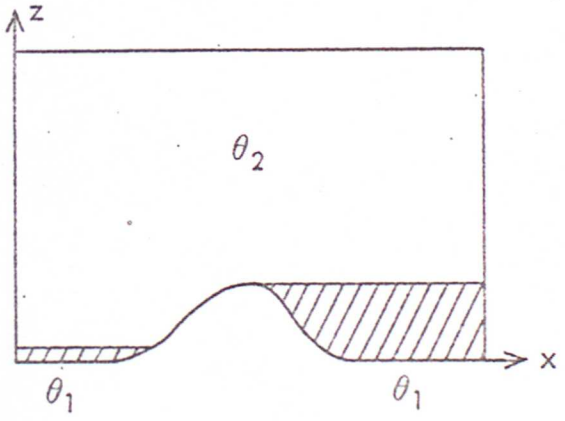





z
H
Upper
Front
H/2
c
Surface Front
Warm Air
Cold Air
a
b
d
e
x
0
(a)
θ
Onset of Upper Front
Warm Air
M
c′
d′
e′
Cold Air
b′
a′
Onset of Surface Front
(b)



(a)



(b)

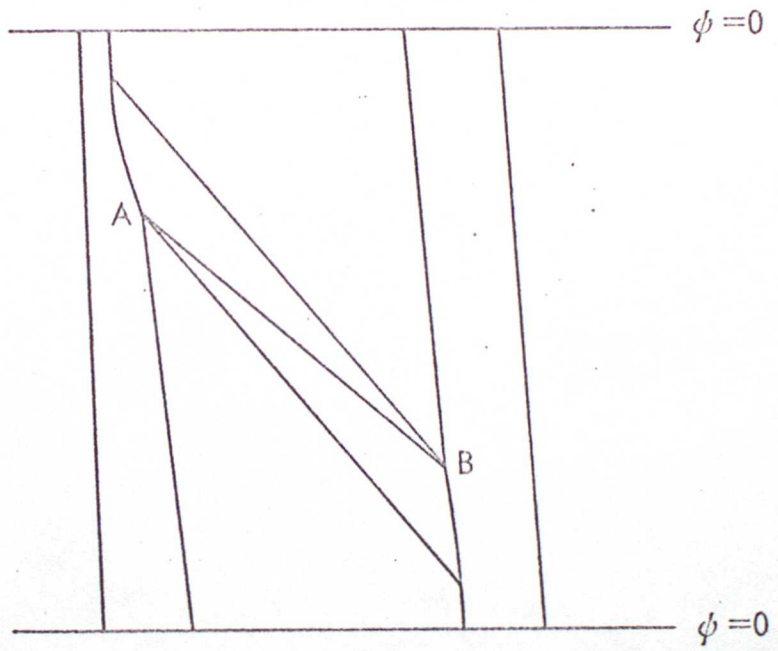


FIG 16

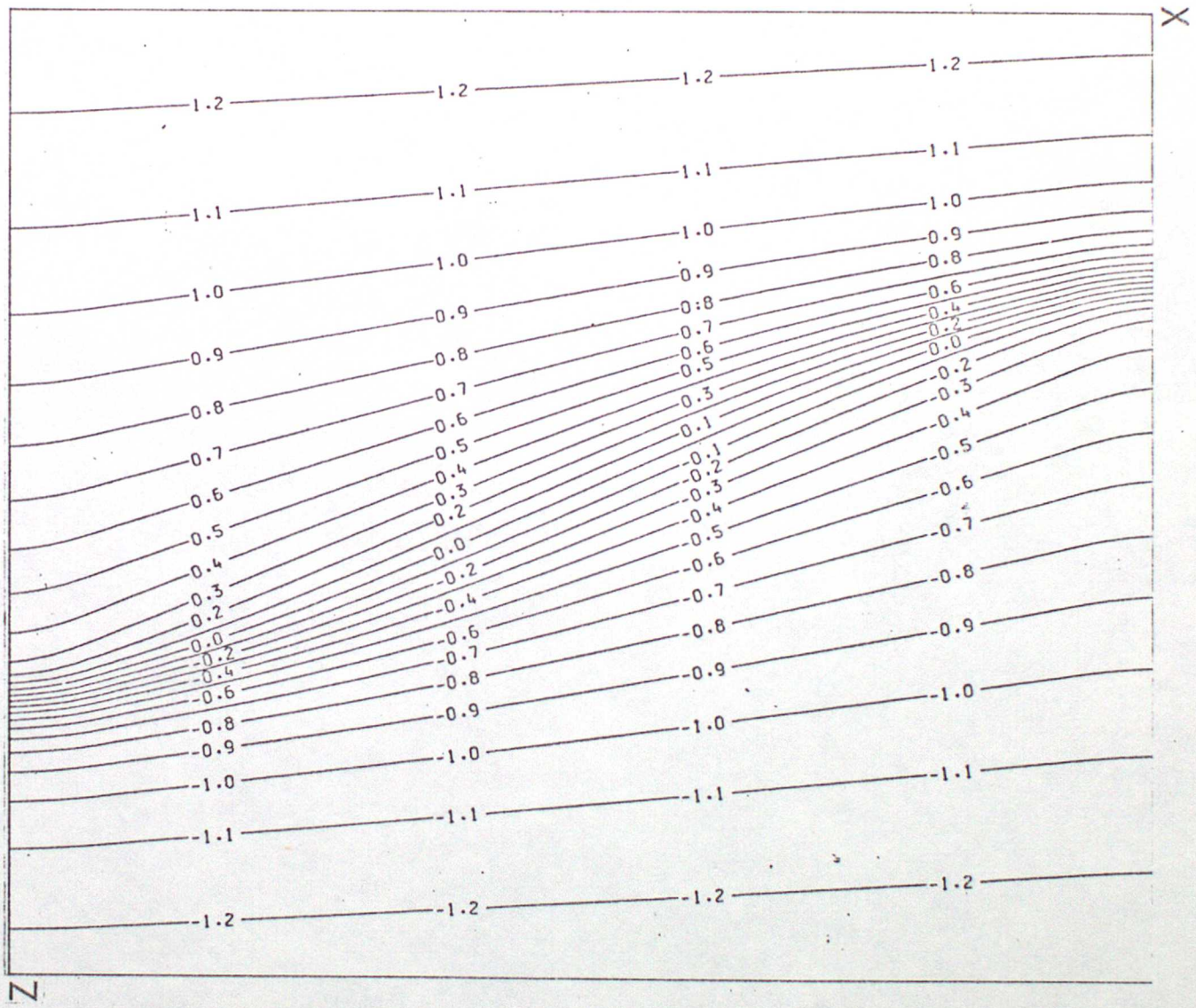

X
N
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1.0
-1.1
-1.2

Fig 17

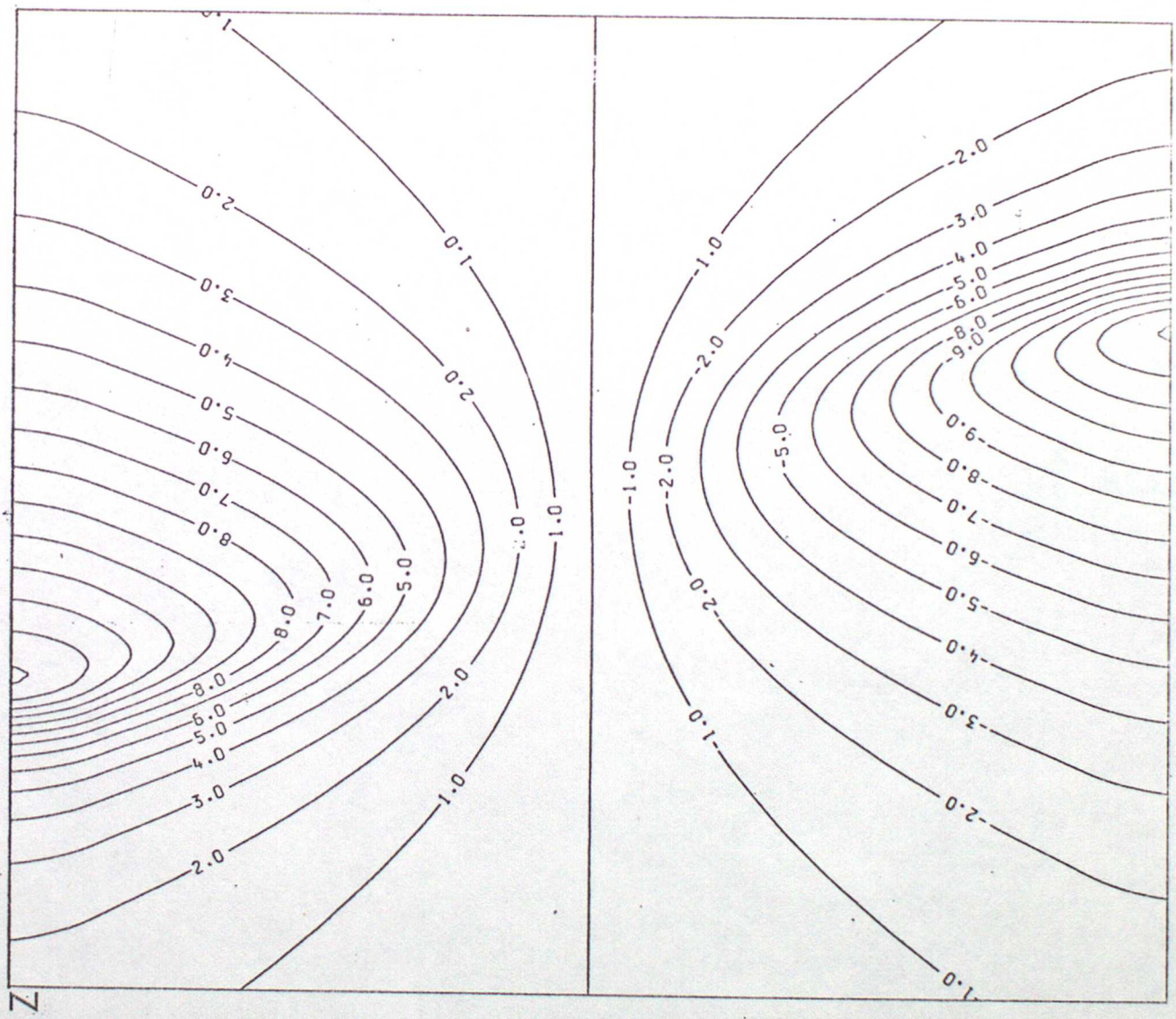


FIG 18

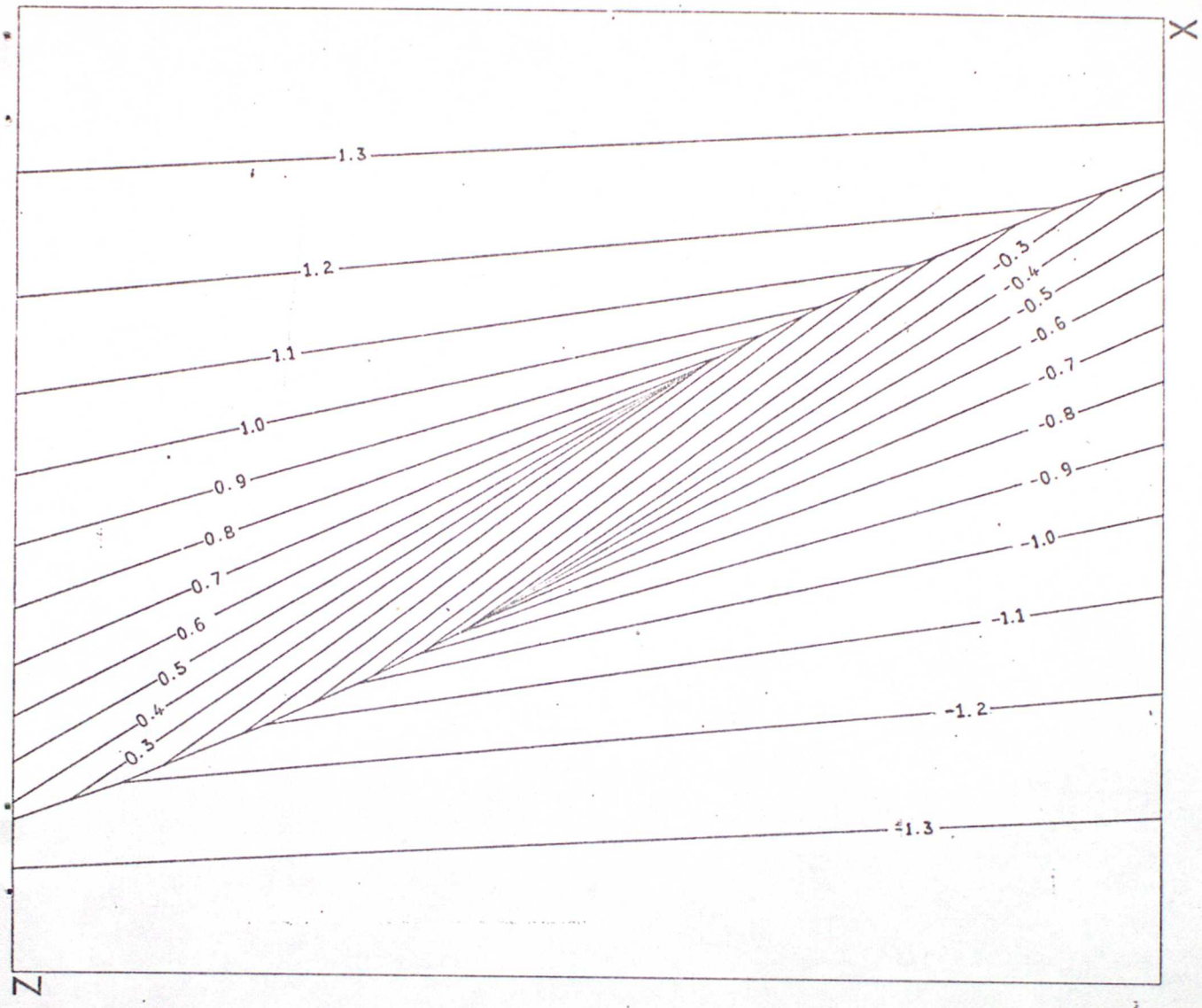


FIG 19

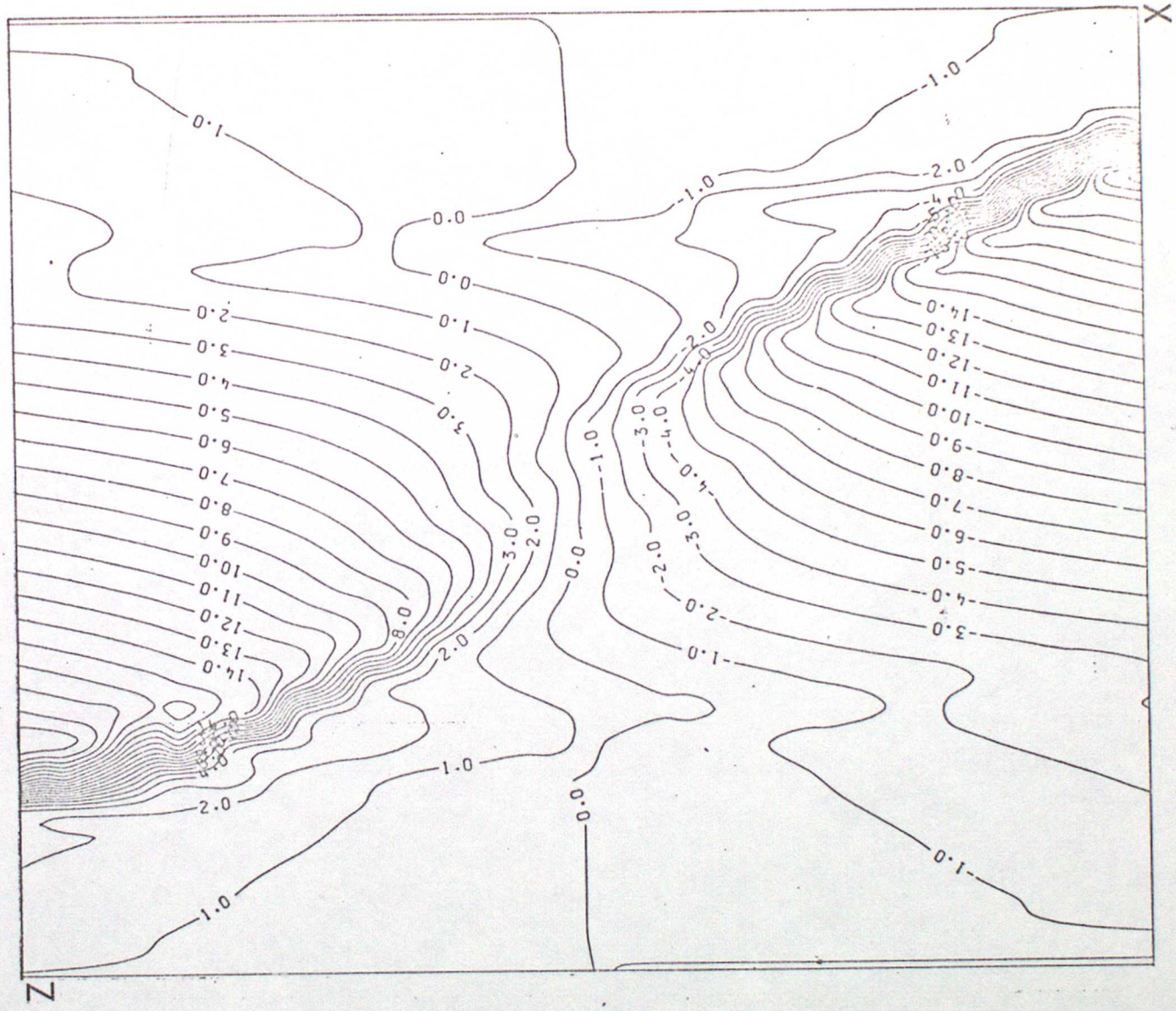

X
Z
1.0
0.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
12.0
13.0
14.0
-1.0
-2.0
-3.0
-4.0
-5.0
-6.0
-7.0
-8.0
-9.0
-10.0
-11.0
-12.0
-13.0
-14.0

FIG 20

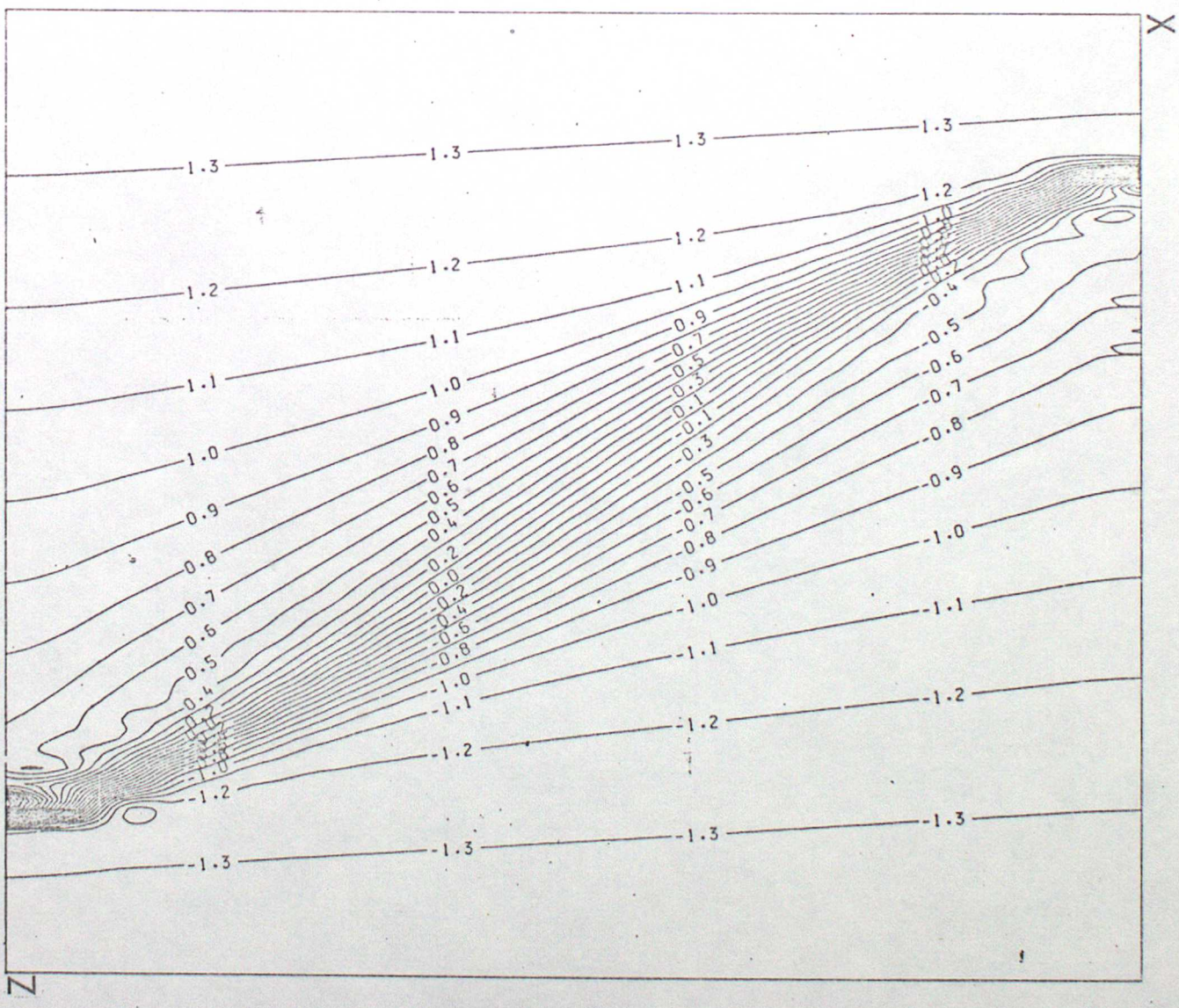


FIG 21

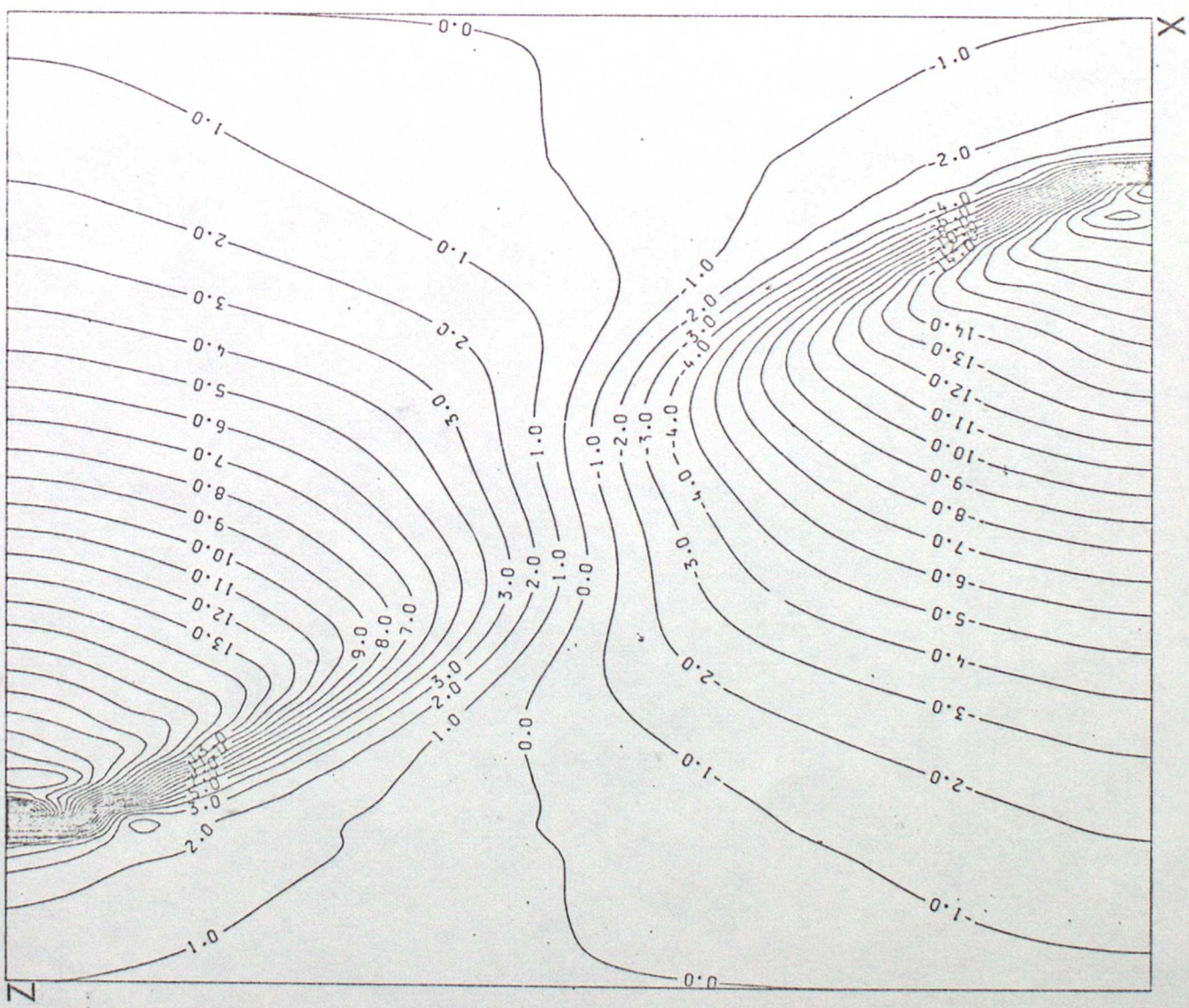


FIG 22

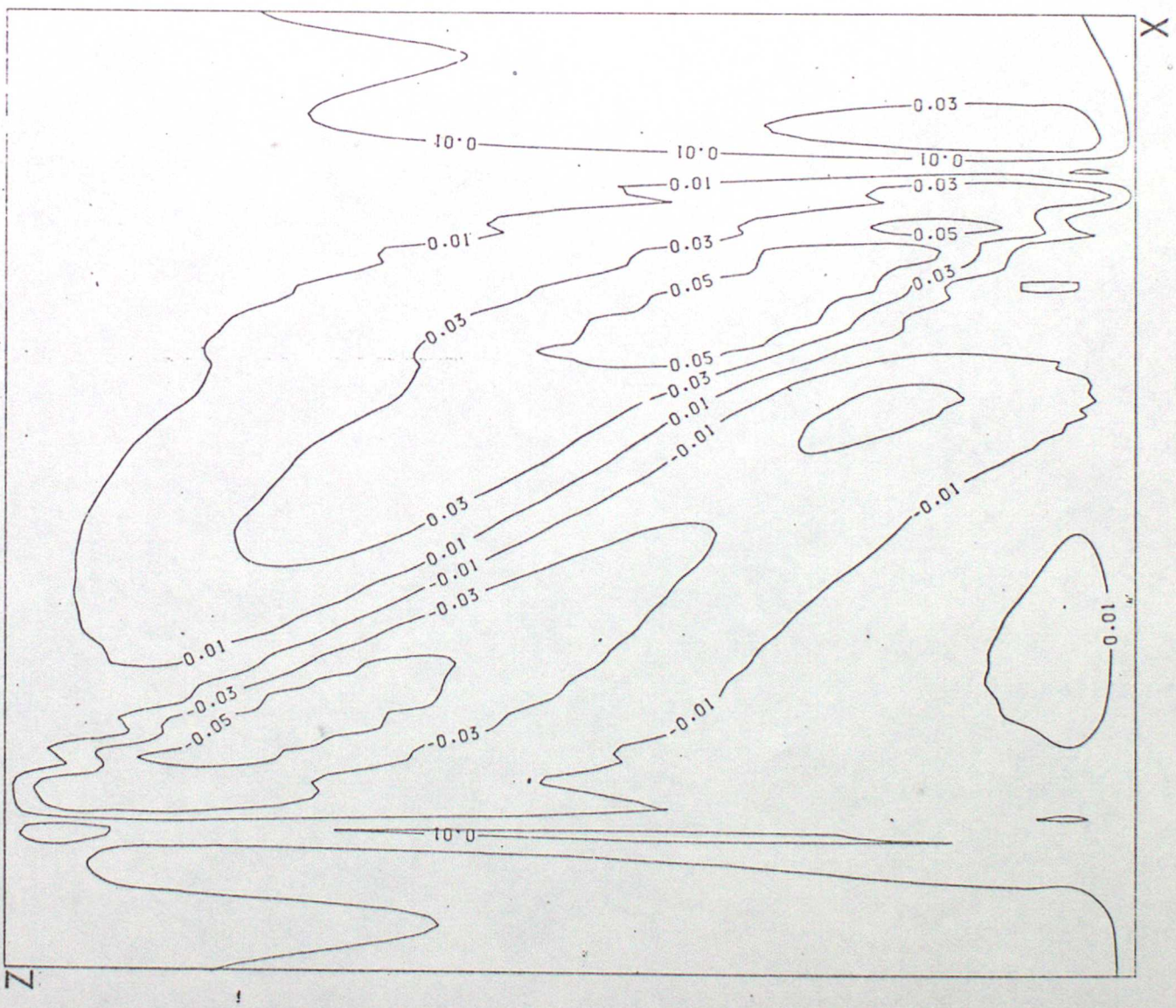

X
Z
0.03
10·0
10·0
10·0
0.01
0.03
0.01
0.03
0.05
0.05
0.03
0.03
0.05
0.03
0.01
0.01
0.03
0.01
0.01
0.03
0.01
-0.01
-0.01
-0.03
-0.03
-0.05
-0.03
-0.01
10·0

FIG 23

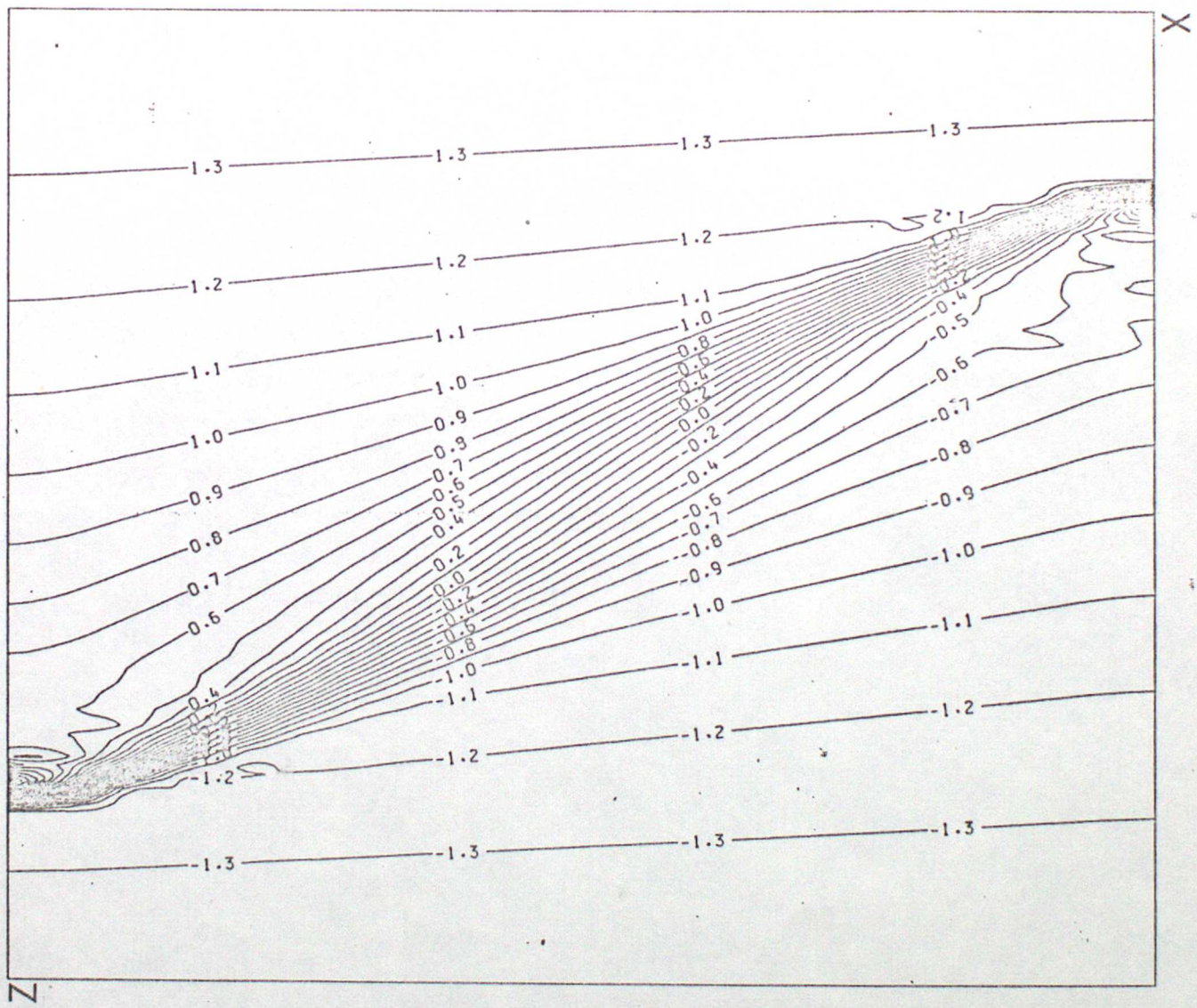


Fig 24

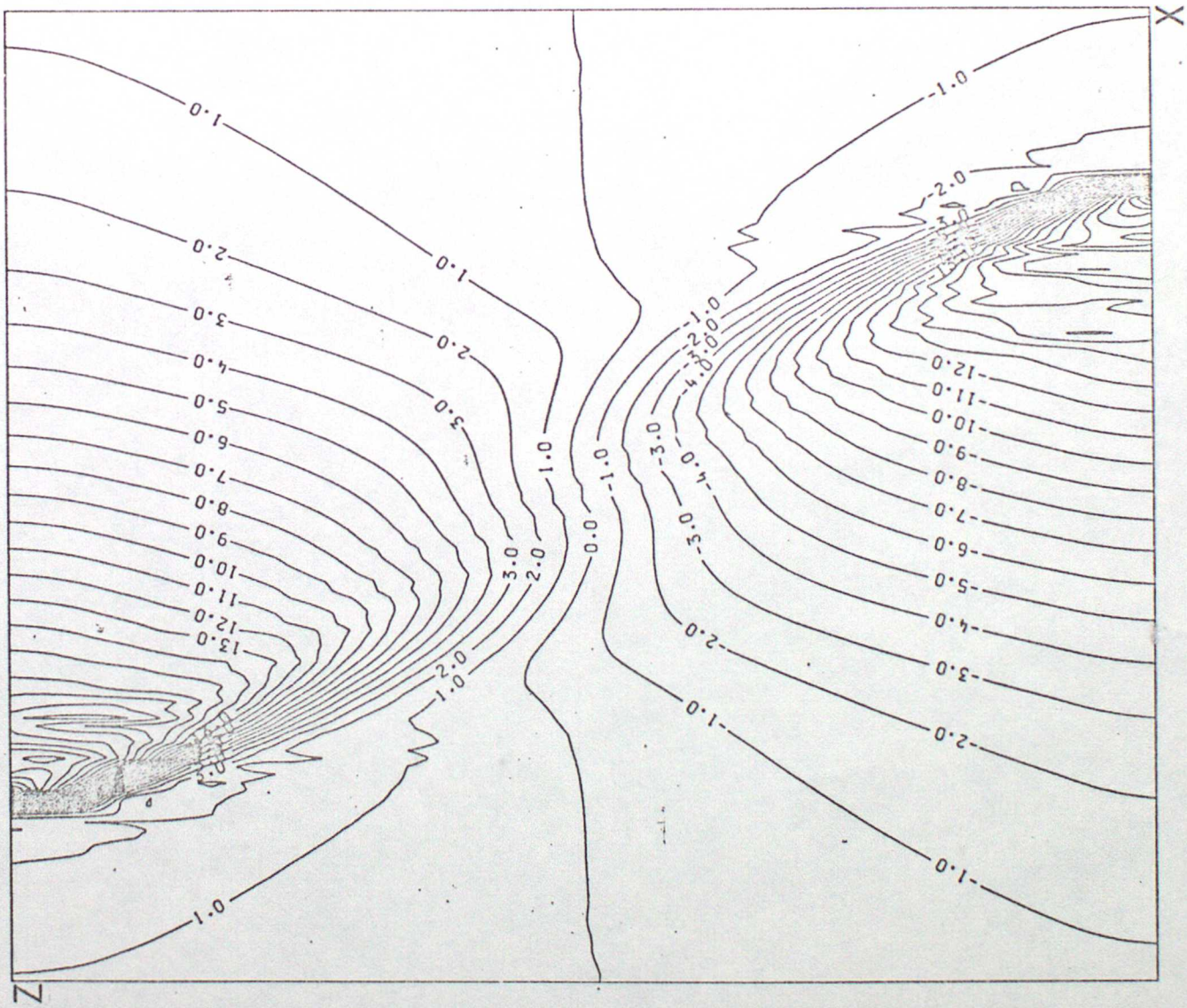


FIG 25

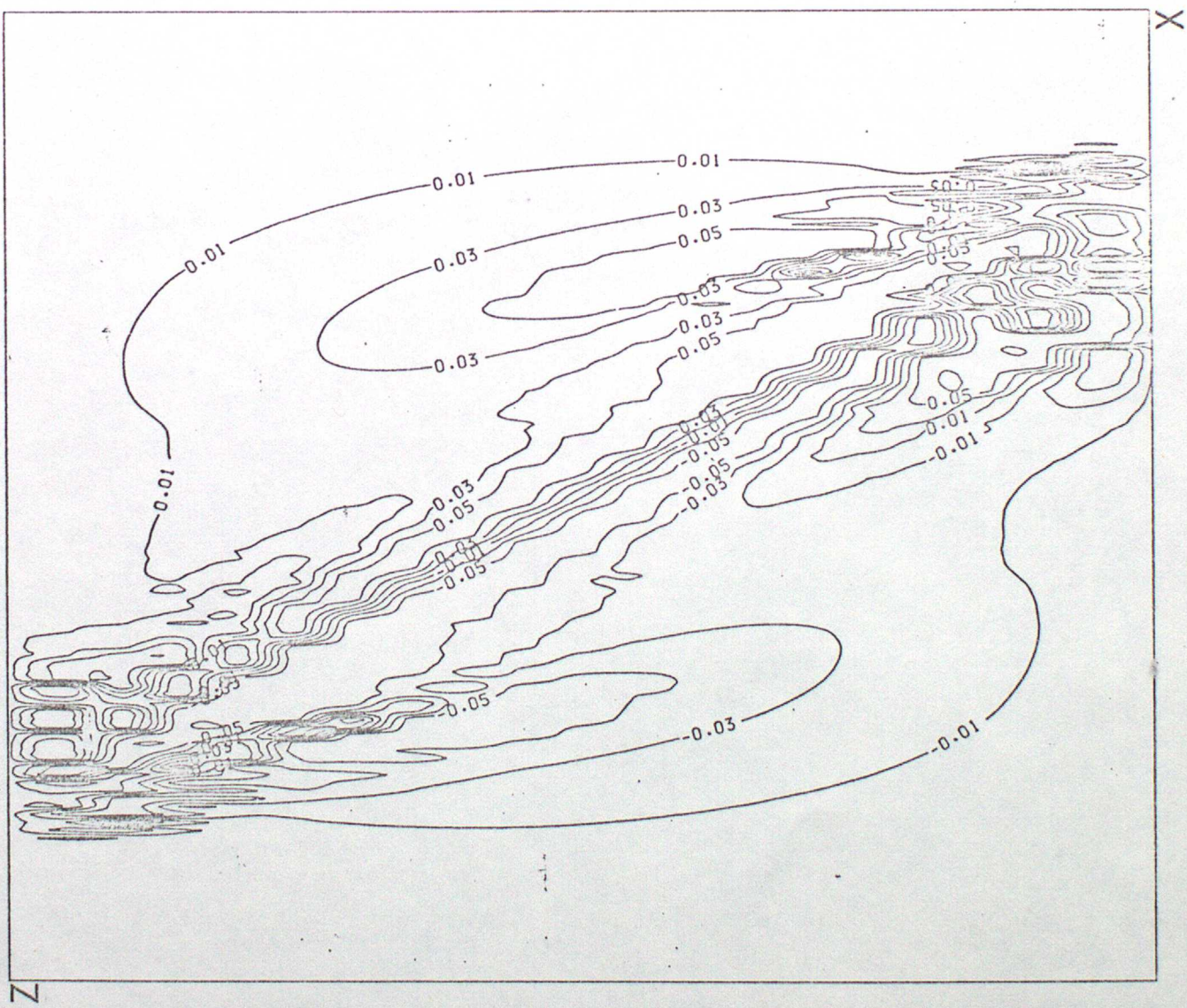


FIG 26

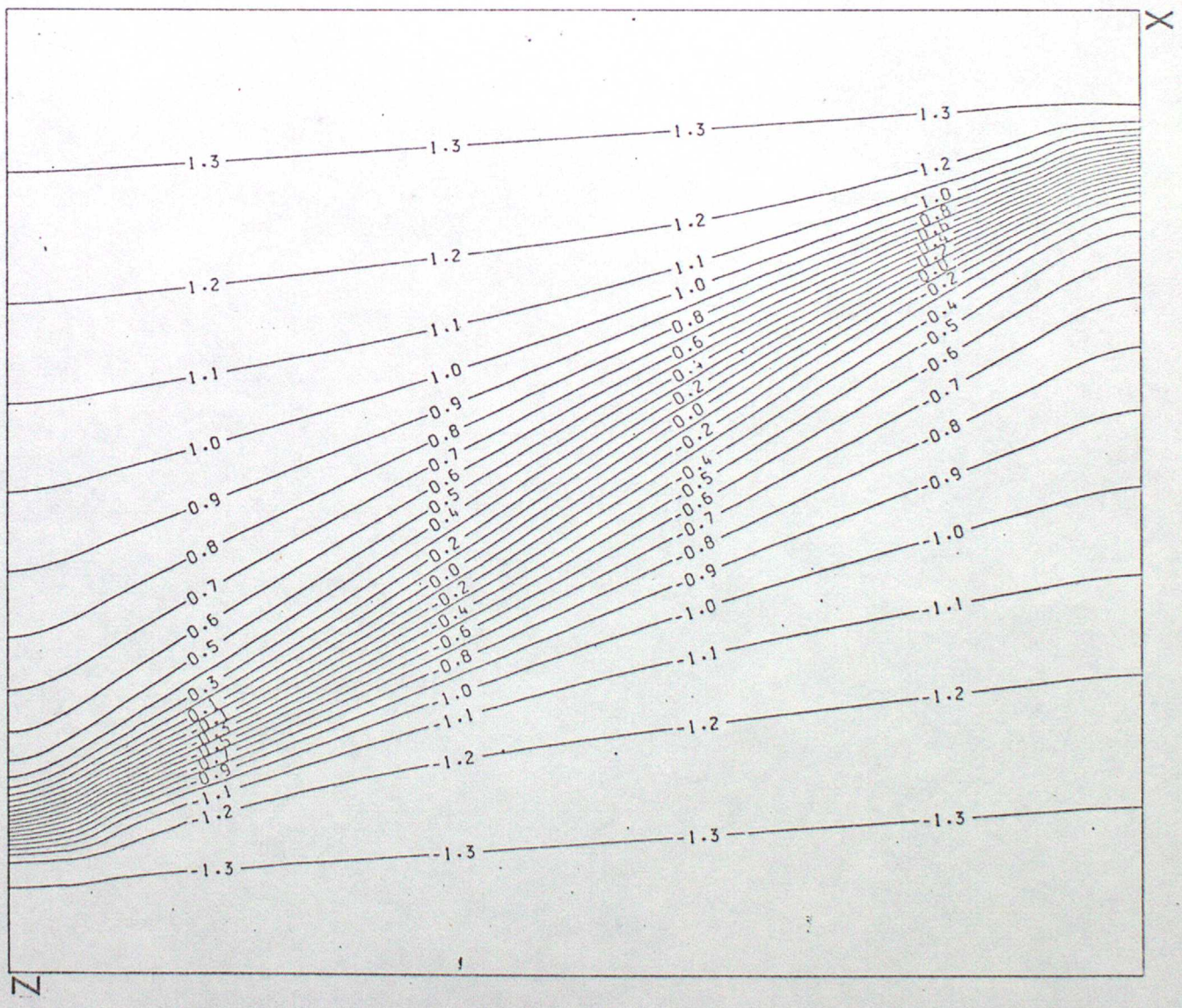
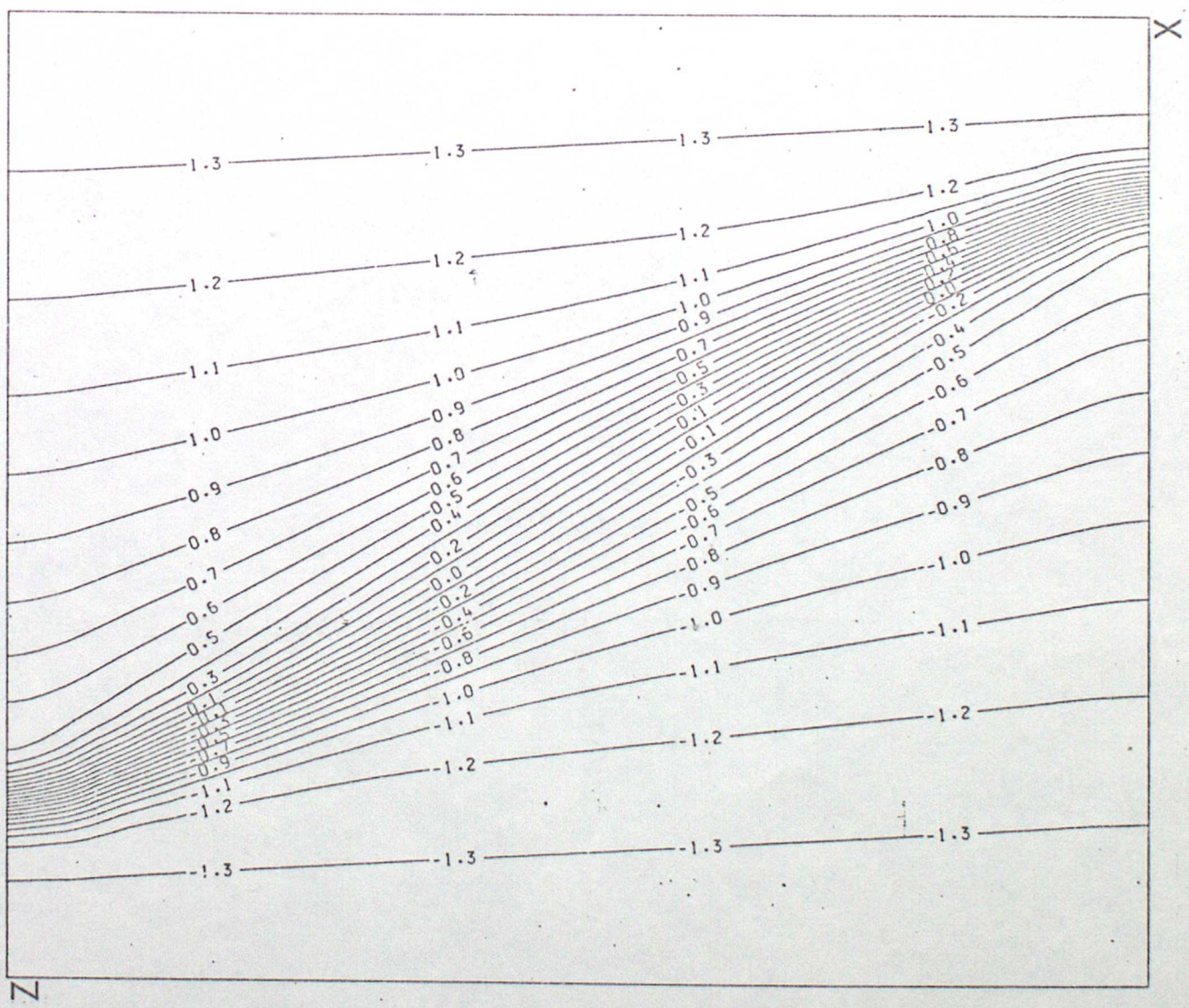
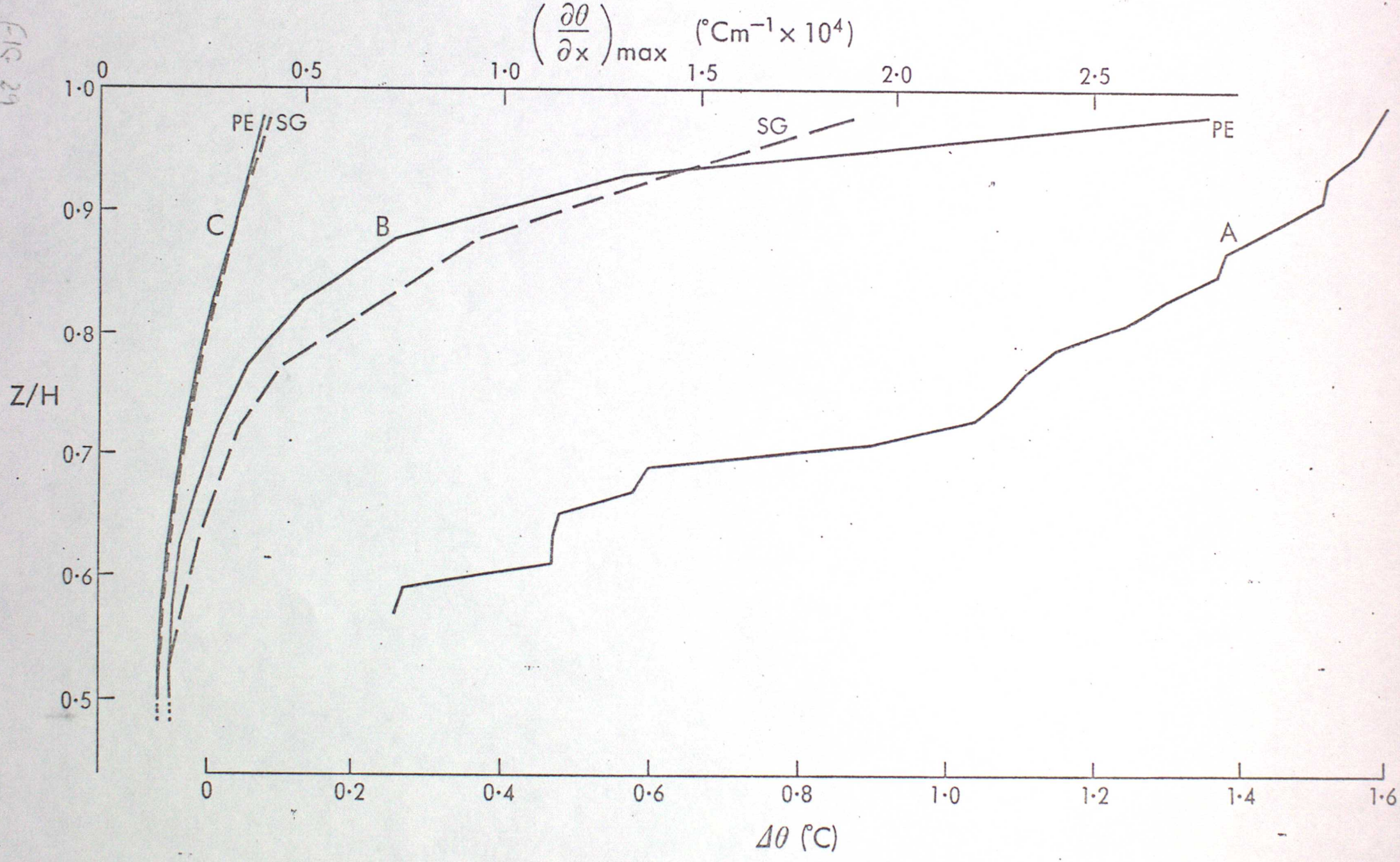


FIG 27


X
N
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1.0
-1.1
-1.2
-1.3

FIF 28


$\left(\frac{\partial\theta}{\partial x}\right)_{max}$ (°Cm^{-1} × 10^4)
0
0·5
1·0
1·5
2·0
2·5
PE
SG
SG
PE
C
B
A
1·0
0·9
0·8
0·7
0·6
0·5
Z/H
0
0·2
0·4
0·6
0·8
1·0
1·2
1·4
1·6
Δθ (°C)

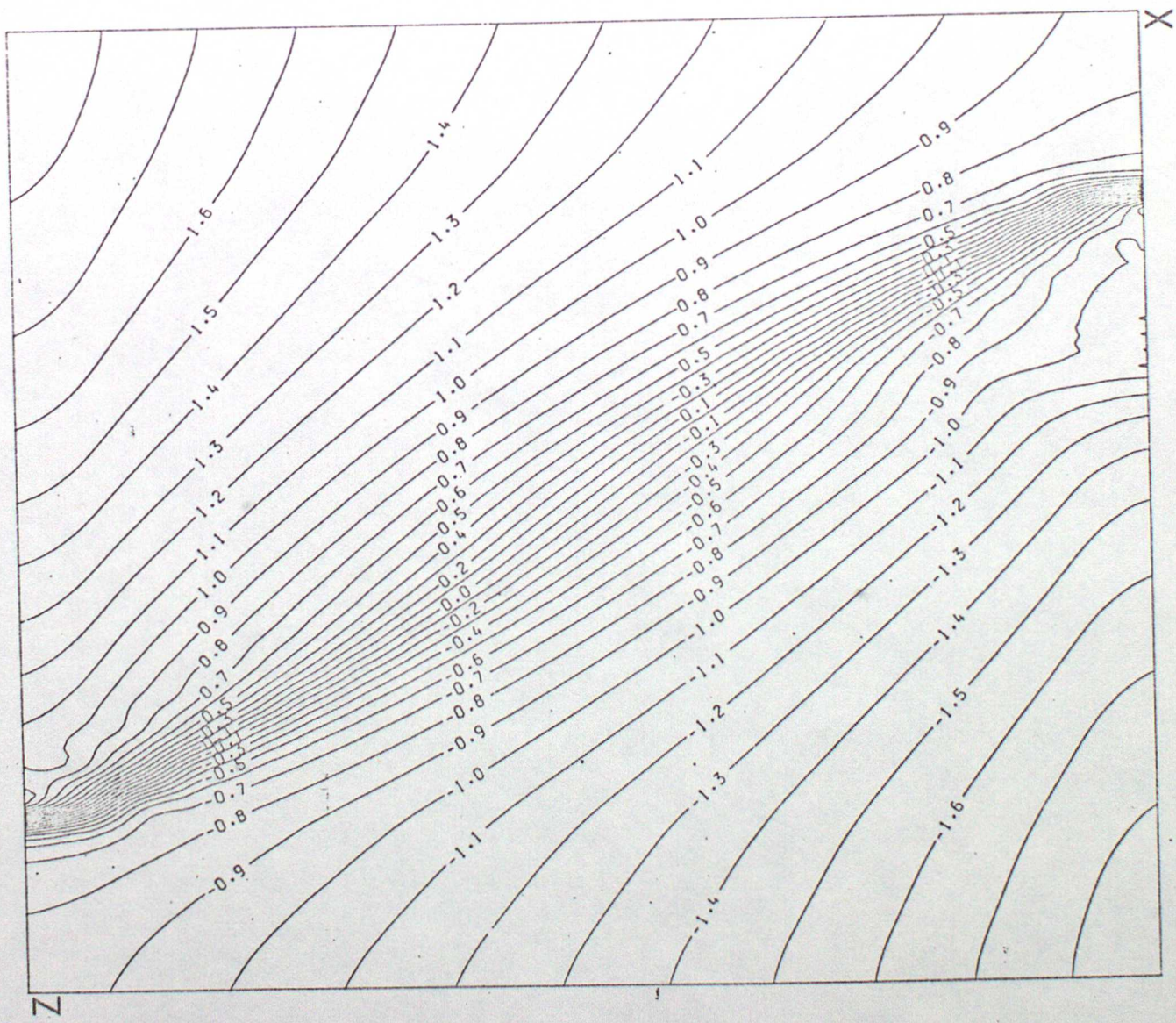


FIG 30

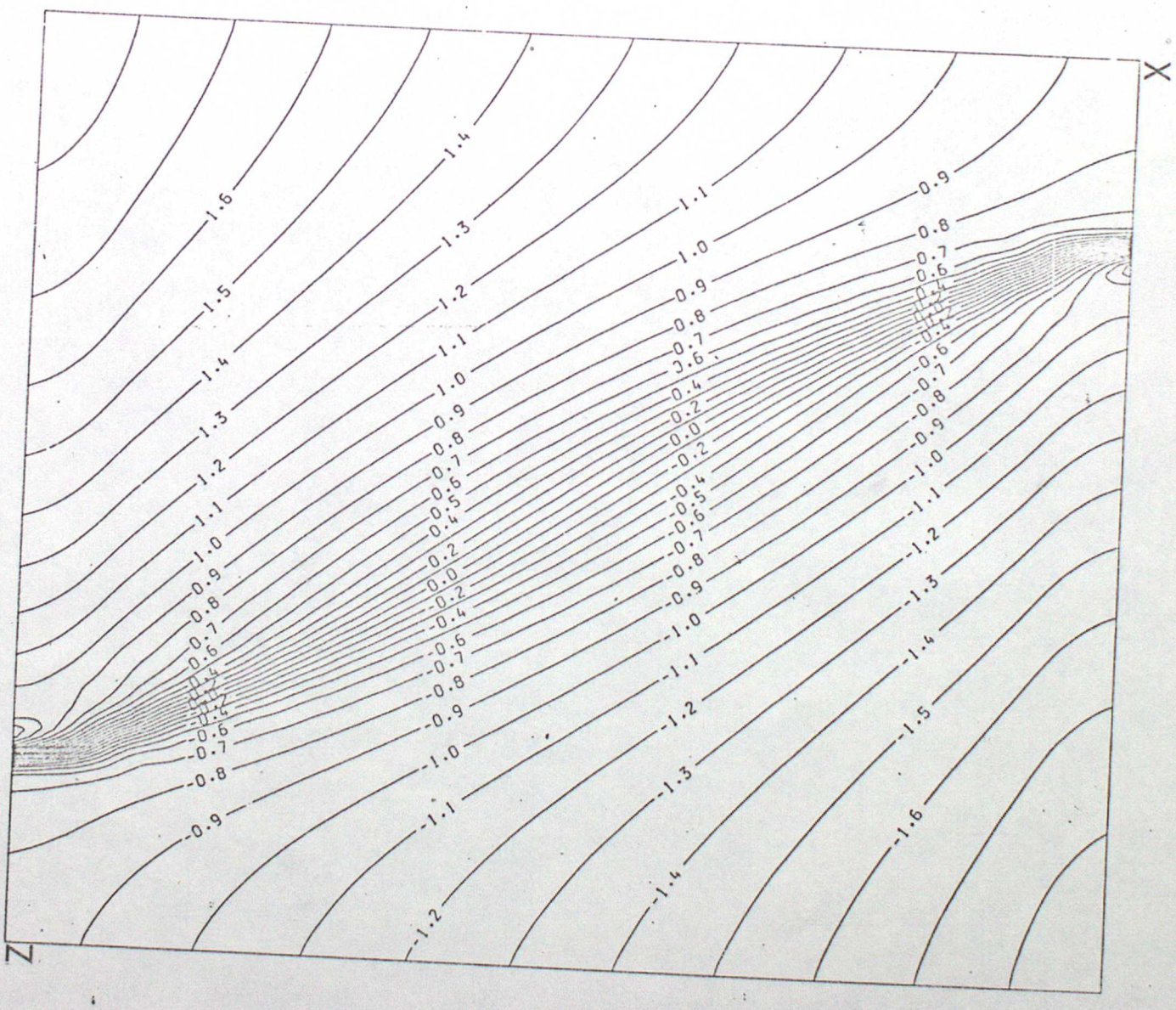


FIG 31

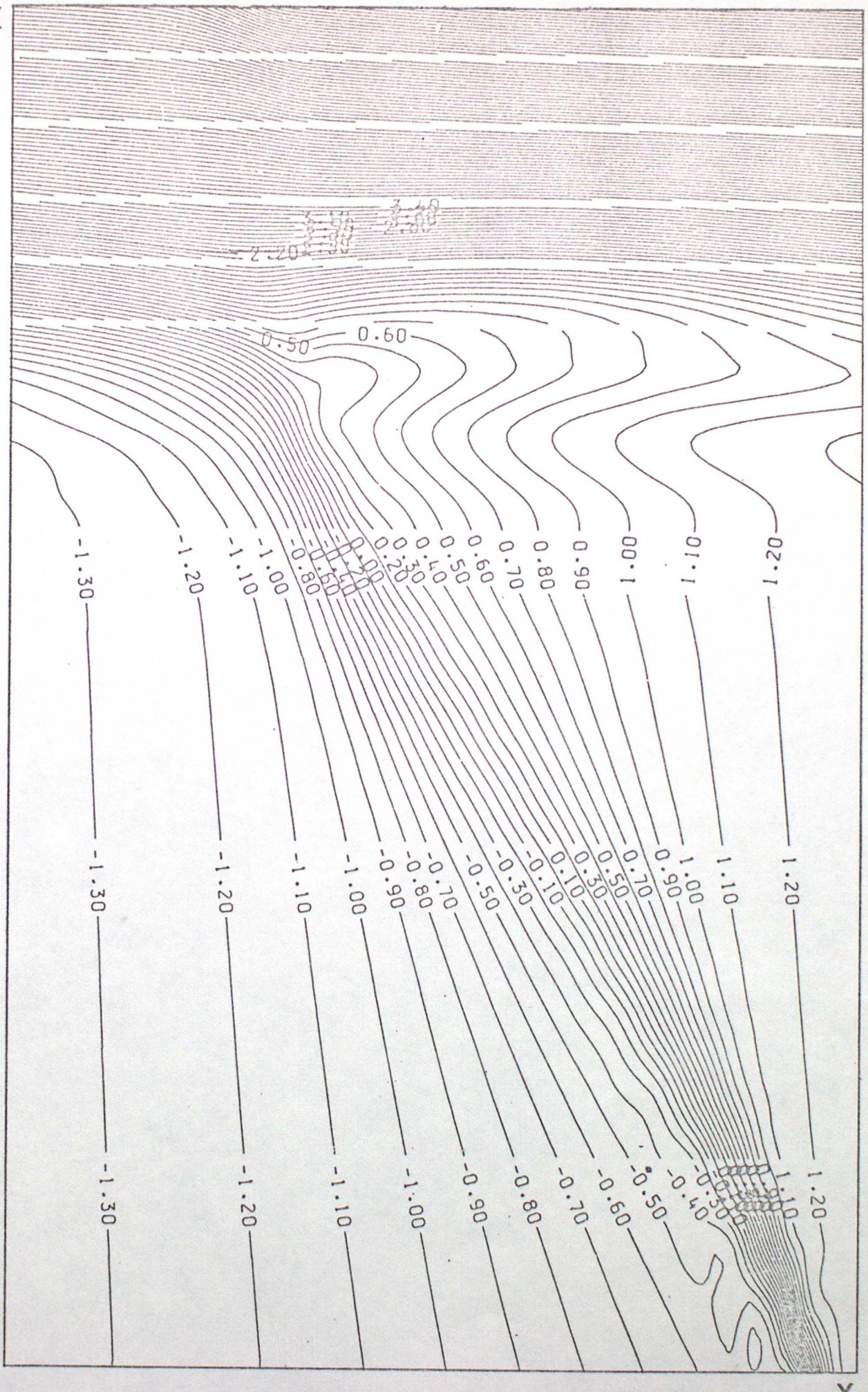

Z
X
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20
0.40
0.30
0.20
0.10
-0.10
-0.30
-0.50
-0.70
-0.80
-0.90
-1.00
-1.10
-1.20
-1.30
-0.40
-0.60

FIG 32

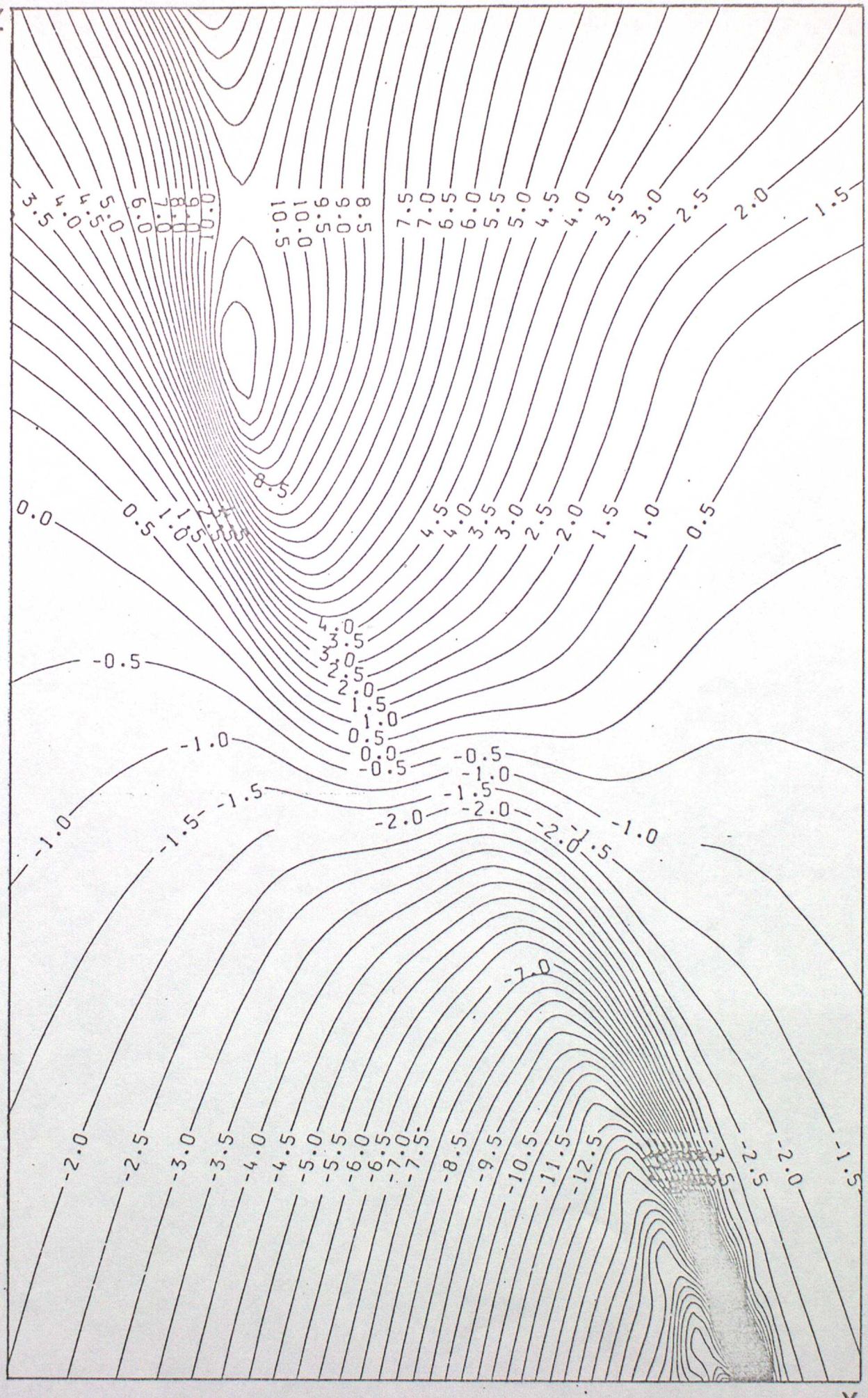


FIG 33