J. Fluid Mech. (1993), vol. 253, pp. 341–358

A mathematical model of turbulent heat and mass transfer in stably stratified shear flow

By G. I. BARENBLATT[1], M. BERTSCH[2], R. DAL PASSO[3], V. M. PROSTOKISHIN[4] AND M. UGHI[5]

[1] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK.

[2] Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 00133 Roma, Italy.

[3] Instituto per le Applicazioni del Calcolo Mauro Picone, Viale del Policlinico 137, 00161 Roma, Italy.

[4] P. P. Shirshov Institute of Oceanology, Russian Academy of Sciences, 23 ul. Krasikova, Moscow 117218, Russia.

[5] Instituto delle Scienze delle Costruzioni, Università di Trieste, Piazzale Europa 1, 34127 Trieste, Italy.

(Received 6 August 1992 and in revised form 19 February 1993)

It is commonly assumed that heat flux and temperature diffusivity coefficients obtained in steady-state measurements can be used in the derivation of the heat conduction equation for fluid flows. Meanwhile it is also known that the steady-state heat flux as a function of temperature gradient in stably stratified turbulent shear flow is not monotone: at small values of temperature gradient the flux is increasing, whereas it is decreasing after a certain critical value of the temperature gradient. Therefore the problem of heat conduction for large values of temperature gradient becomes mathematically ill-posed, so that its solution (if it exists) is unstable.

In the present paper it is shown that a well-posed mathematical model is obtained if the finiteness of the adjustment time of the turbulence field to the variations of temperature gradient is taken into account. An evolution-type equation is obtained for the temperature distribution (a similar equation can be derived for the concentration if the stratification is due to salinity or suspended particles). The characteristic property which is obtained from a rigorous mathematical investigation is the formation of step-wise distributions of temperature and/or concentration from continuous initial distributions.

1. Introduction

The problem of turbulent heat and mass transfer in stably stratified shear flows has been considered for a long time as one of geophysical importance (Monin & Yaglom 1971) and, more recently, as a problem of technical importance (Petukhov & Polyakov 1989). Stable stratification (i.e. a density decrease in the vertical direction in the presence of a strong gravity field) suppresses turbulence, and therefore reduces the heat and mass exchange efficiency. A challenge for theoreticians remains the formation of step-wise temperature and density distributions in isolated water masses in the ocean, induced by stratification and turbulence.

Phillips (1972) suggested that the formation of steps can be explained by the instability of the density field in strongly stably stratified turbulent shear flows. This instability is due to the strong decrease of turbulent heat exchange coefficients with growing temperature gradients. Due to this strong decrease, the increment of heat flux due to the increasing temperature gradient is, at large values of the gradient, less than its decrease due to the corresponding suppression of turbulence, and consequently, of the heat exchange coefficient. Such non-monotone behaviour of the heat flux was observed long ago, and the papers by Rossby & Montgomery (1935), Munk & Anderson (1948), Mamayev (1958) and Felsenbaum & Boguslavsky (1977) should be mentioned in this context.

Posmentier (1977) proposed, independently of Philips, a similar idea in connection with salinity flux. His paper, concerning the non-monotonic dependence of mass flux on the salinity gradient, is of particular interest because it contains the results of numerical calculations which show the formation of steps. The paper by Linden (1979) contains a comprehensive review and an extension of the stability considerations mentioned before. The more recent paper of Ruddick, McDougall & Turner (1989) contains a review of the basic contributions to the problem and also the results of laboratory experiments with stirred sugar and salt solutions demonstrating the formation of density steps. The compilation of Ivey & Imberger (1991) regarding the inhibition of flux due to stratification should also be mentioned.

All these papers however had two important points in common which remained unclear. First of all, it was always assumed in the derivation of the equation for temperature and salinity that the local relation between heat and/or mass flux and the corresponding gradients which is valid for steady equilibrium conditions, can be substituted into the non-steady equation of energy or mass balance, exactly as we do usually in deriving, for instance, Fourier or Fick equations. A common argument is that the flux relaxation times are negligibly small, so the limiting case can be considered when the relaxation time vanishes. However, the equations obtained in this way reduce, for large values of the gradients, to diffusion or heat conduction equations with negative coefficients which lead to mathematically incorrect initial value problems – that was the second point which remained unclear: the possible ill-posedness of the evolution equations obtained.

In the present paper it is shown that taking into account the finiteness of the adjustment time of the turbulence field (and, consequently, of the turbulent exchange coefficients) to the variations of the temperature and/or concentration field leads to a correct mathematical model of the phenomenon for large values of the gradients also. In particular the following evolution equation is obtained for temperature and/or concentration:

$$\theta_t = \phi(\theta_z)_z + \tau\psi(\theta_z)_{zt}. \tag{1}$$

Here $\theta(z,t)$ is the temperature distribution or concentration as a function of the vertical coordinate z and time t, $\phi(\theta_z)$ is the absolute value of the 'temperature flux' (heat flux divided by specific heat) as the function of temperature gradient, and subscripts denote partial derivatives, for example

$$\phi(\theta_z)_z = \frac{\partial}{\partial z}\left(\phi\left(\frac{\partial\theta}{\partial z}\right)\right).$$

The small parameter τ represents the positive relaxation (adjustment) time. The smooth function $\phi(p)$ is *non-monotone*, satisfies the conditions

$$\phi(0) = 0; \quad \phi(+\infty) = 0; \quad \phi(p) > 0 \quad \text{for} \quad p > 0;$$

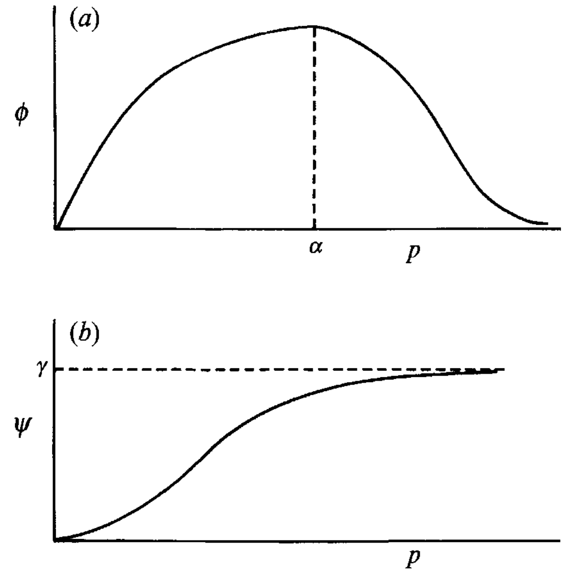

FIGURE 1. The functions ϕ, (a), and ψ, (b).

and has its maximum at a certain value $p = \alpha$. A typical choice of ϕ suggested by experimental data (see figure 1a) is

$$\phi(p) = \frac{Ap}{1+Bp^2}.$$

The function $\psi(p)$ (see figure 1b) is determined by the temperature flux (function ϕ) in the following way:

$$\psi'(p) = -\phi'(p) + \frac{\phi(p)}{p} \quad \text{for} \quad p > 0. \tag{2}$$

The function $\psi(p)$ is smooth and strictly increasing; it satisfies the relations

$$\psi(0) = 0, \quad \psi(+\infty) = \gamma < +\infty,$$

In particular, from (2) follows the inequality

$$\phi'(p) \geqslant -\psi'(p) \quad \text{for} \quad p > 0, \tag{3}$$

which will be of importance below and which implies that the last term of the equation is strong enough to control the possibly negative diffusion coefficient $\phi'(\theta_z)$.

It is plausible that if $\tau = 0$, equation (1) leads to ill-posed initial-boundary-value problems since ϕ is not monotone. An indication in this direction is supplied by the ill-posedness of the backwards heat equation and also by a non-uniqueness result by Höllig (1983), who constructed infinitely many solutions of the equation $\theta_t = \phi(\theta_z)_z$ which have the same initial function in the special case in which ϕ is piecewise linear, decreasing in an interval and increasing elsewhere.

If $\tau > 0$, the evolution equation (1) is of degenerate pseudoparabolic type. Here the word degenerate indicates the fact that, due to relation (2), $\psi'(p)$ is not uniformly bound away from zero since $\psi'(p) \to 0$ as $p \to \infty$.

Degenerate and non-degenerate pseudoparabolic equations arise in many applications and they have been studied by many authors (for references to the literature see Barenblatt *et al.* 1993). Of special interest for us is the work by Padron (1990) in which (1) (or rather its integrated version for the gradient $v = \theta_z : v_t = \phi(v)_{zz} + \tau\psi(v)_{zzt}$) is studied in the non-degenerate case in which $\psi(p) = p$. The behaviour of the function ϕ is the same as in our case. This work contains an existence theorem, a result which indicates the stability of delta-function-type solutions (which correspond to our step-wise solutions), and some numerical calculations. The mathematical techniques which are used are quite different from the ones in the present paper. In addition we shall see that, namely, the degeneracy of the function ψ leads to the generation of the discontinuous solutions.

After the derivation of the physical model in §2, we list in §3 the precise hypotheses on the data and we present and discuss our main mathematical results concerning the initial-boundary-value problem

$$\left.\begin{aligned} &\theta_t = \phi(\theta_z)_z + \tau\psi(\theta_z)_{zt} && \text{for} \quad 0 < z < L, t > 0, \\ &\theta_z(0,t) = \theta_z(L,t) = 0 && \text{for} \quad t > 0, \\ &\theta(z,0) = \theta_0(z) && \text{for} \quad 0 < z < L, \end{aligned}\right\} \tag{4}$$

to which the problem under consideration was reduced. In particular we shall explain what we mean by a solution of this problem: this is non-trivial because the solution is a generalized one. The top and bottom boundaries in problem (4) correspond to no-flux conditions, but the mathematical analysis does not depend strongly on the choice of the boundary conditions; in particular a similar analysis is possible for the corresponding Cauchy problem in which $-\infty < z < +\infty$ and no boundary conditions are imposed. Results remain practically the same.

Our mathematical results presented below justify the well-posedness of problem (4), and give a rather complete qualitative picture of the transient and large-time behaviour of the solutions. This behaviour strongly depends on the properties of the function ψ and, in particular, on the threshold value α for the temperature gradient. We shall show that if the initial temperature gradient $\theta_z(z,0)$ is smaller than α for all z, then the temperature gradient $\theta_z(z,t)$ remains subcritical for all z and t (in this case the problem without third-order term is well-posed). On the other hand, if $\theta_z(z,0)$ is sufficiently large at some points or in an interval, then $\theta(z,t)$ becomes discontinuous after some finite time t_0; in addition, $\theta(z,t)$ remains discontinuous for all later times $t > t_0$, and stabilizes (we shall explain exactly what this means) to a step-wise (i.e. piecewise constant) temperature distribution as $t \to +\infty$.

In the last section we shall present some numerical computations, which strongly suggest that if the temperature gradient is supercritical in some interval, then the asymptotic temperature distribution is discontinuous for sufficiently small values of τ, and the number of layers increases with decreasing τ.

2. Basic physical hypotheses and the model

For simplicity we shall only consider the case of a thermally stratified fluid; the model for the case of salinity stratification or stratification by suspended particles can be obtained in a completely analogous way.

The mean potential temperature $\theta(z,t)$ in a statistically horizontally homogeneous layer satisfies the energy balance equation

$$\theta_t = (k\theta_z)_z. \tag{5}$$

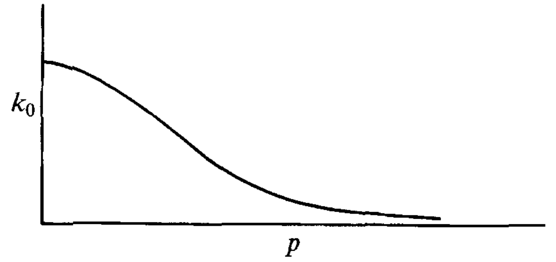

FIGURE 2. The steady heat diffusivity $k_0(p)$.

Here z is the vertical coordinate, t is the time and $k(z,t)$ is the turbulent temperature diffusivity, defined by the relation

$$k = -\frac{\Phi}{\rho c_p \theta_z}, \tag{6}$$

where $\Phi(z,t)$ is the turbulent heat flux, and ρ and c_p are, respectively, the reference fluid density and fluid specific heat per unit mass under constant pressure, so that ρc_p is the specific heat per unit volume. We stress that under conditions of horizontal statistical homogeneity relation (6) does not contain any additional assumption; it is nothing other than the definition of turbulent temperature diffusivity.

Under fixed external hydrodynamic conditions (e.g. applied pressure gradient or stress) and a fixed value of the temperature gradient, the turbulent temperature diffusivity tends to a certain limiting value. For fixed hydrodynamic conditions this limiting value should be a function of the temperature gradient only:

$$k = k_0(\theta_z). \tag{7}$$

It is well known (see e.g. Monin & Yaglom 1981; Ivey & Imberger 1991) that positive temperature gradient inhibits the turbulence, and therefore the function $k_0(p)$ is decreasing:

$$k_0'(p) \leqslant 0 \quad \text{for} \quad p \geqslant 0$$

(k_0' possibly vanishes at $p = 0$ and definitely as $p \to \infty$).

At large values of p the function $k_0(p)$ decreases rapidly, so that the heat flux tends to zero for large temperature gradients θ_z (we neglect the molecular temperature diffusivity even at large temperature gradients), i.e. defining the function ϕ as the absolute value of 'temperature flux'

$$\phi(\theta_z) = k_0(\theta_z)\,\theta_z, \tag{8}$$

we obtain that the graph of $\phi(p)$ has the shape which we have indicated in figure 1(*a*): $\phi(p)$ is increasing for $0 < p < \alpha$ and decreasing for $p > \alpha$. Qualitatively the graph of ϕ in figure 1(*a*) has the same form as the one proposed by Rossby & Montgomery (1935), Munk & Anderson (1948), Posmentier (1977), Felsenbaum & Boguslavsky (1977), and Ruddick *et al.* (1989); a reasonable choice for k_0 based on experimental data is (see figure 2)

$$k_0(p) = \frac{A}{1+Bp^2}. \tag{9}$$

We repeat that $k_0(\theta_z)$ and $\phi(\theta_z)$ correspond to limiting values which are obtained for large times under fixed external conditions, including temperature gradient.

At first sight it seems natural to use the steady temperature diffusivity k_0 to close the energy balance equation (5). Indeed, the relaxation time of the turbulent field to a certain given value of the temperature gradient seems to be small in comparison with the characteristic time of the temperature field redistribution. This approach, similar to the traditional derivation of the heat conduction and diffusion equations, leads to the following equation for the potential temperature:

$$\theta_t = \phi(\theta_z)_z = \phi'(\theta_z)\,\theta_{zz}. \tag{10}$$

Such an approach was used in particular by Posmentier (1977). If the temperature gradient is less than the critical value α the coefficient $\phi'(\theta_z)$ of the second derivative in (10) is positive and (10) leads to correct initial-boundary-value problems. However, if somewhere the temperature gradient is larger than α, this coefficient becomes negative and (10) is the backward heat equation. This suggests that the mathematical model should be modified.

To overcome this difficulty we have to take into account that (5) is only one equation of a system, governing the whole field, which should also contain the equations for a model of turbulent shear flow. At this moment a generally accepted system of equations to describe turbulent shear flow is not available, but all such models should contain the equation of turbulent energy balance which can be written in the form (see e.g. Monin & Yaglom 1971)

$$\bar{b}_t = -\overline{u'w'}\bar{u}_z - (\overline{(p'/\rho + b)\,w'})_z - \epsilon + \beta g \frac{\Phi}{\rho c_p}. \tag{11}$$

Here $u(z,t)$, $v(z,t)$ and $w(z,t)$ denote the components of the velocity along, respectively, the horizontal x-axis, the horizontal y-axis, and the vertical z-axis (we recall that the mean velocity is directed, by definition, along the x-axis), p is the pressure, β the volume thermal extension coefficient of the fluid, g is the gravitational acceleration, primes denote fluctuations, bars indicate ensemble means values, b is the specific turbulent energy per unit mass:

$$b = \tfrac{1}{2}(u'^2 + v'^2 + w'^2), \tag{12}$$

ϵ the viscous dissipation rate of turbulent energy per unit mass, and $\Phi(z,t)$ is the heat flux which in the heat balance equation (5) was replaced by $-\rho c_p k \theta_z$ according to (6). The first term on the right-hand side of (11) represents the inflow rate of turbulent energy due to the work of the Reynolds stresses on the mean velocity gradient, the second term is the divergence of the mean turbulent flux of turbulent energy, the third term $(-\epsilon)$ represents the viscous dissipation rate of turbulent energy, and the last term is the decay rate of turbulent energy due to the work against the buoyancy force. In view of the positivity of the temperature gradient θ_z, the last term in (11) is negative; it represents the basic sink of turbulent energy in a strongly stratified flow and also the inhibition of turbulence by flow stratification.

Now to the point most essential for the present work. The turbulent temperature diffusivity k is governed by instantaneous turbulence properties at the moment t. Generally speaking it cannot be replaced by the function $k_0(z,t)$, i.e. by the limiting value of the turbulent temperature diffusivity, corresponding to the instantaneous value of the temperature gradient. Indeed according to every model of turbulence which contains the turbulence energy balance equation (11) turbulence energy needs some time τ to assimilate the current value of the temperature gradient. Therefore, and due to monotone dependence of k_0 on the temperature gradient at the moment t, turbulence temperature diffusivity can be taken as corresponding to the equilibrium value k_0 related to the temperature gradient at a certain delayed moment $t-\tau$, where τ is the time of delay governed by the turbulence itself:

$$\tau \sim l/u_*. \tag{13}$$

Here l is the mean lengthscale of the vortex system, which is proportional to the integral lengthscale of the velocity field and $u_* = (\sigma/\rho)^{\frac{1}{2}}$ is the flow friction velocity, the characteristic velocity scale determined by the tangential stress σ and fluid density ρ, which is proportional to mean velocity fluctuation. We assume that the delay time τ is constant, small in comparison with the characteristic timescale of temperature stabilization, so we deal with an average over the whole field quantity. Our model can incorporate the temperature-gradient dependence of τ without any complications. We do not do it here for the following two reasons. Firstly, we have at this time no reliable information concerning temperature-gradient dependence of τ. Secondly, this dependence (it is plausible that τ should decrease with growing temperature gradient) will not lead to qualitative differences, only some quantitative ones.

Thus, our basic hypothesis is that the current turbulent temperature diffusivity corresponds to the equilibrium one for the temperature gradient at the moment $t-\tau$:

$$k(z,t) = k_0(\theta_z(z,t-\tau)). \tag{14}$$

Bearing in mind that the delay time τ is small in comparison with the characteristic timescale of the temperature field we obtain, developing (14) in a linear expansion with respect to τ,

$$\theta_z(z,t-\tau) \approx \theta_z(z,t) - \tau\theta_{zt}(z,t)$$

and

$$k \approx k_0(\theta_z - \tau\theta_{zt}) \approx k_0(\theta_z) - \tau k_0'(\theta_z)\,\theta_{zt}. \tag{15}$$

Combining (5) and (15), we obtain the following equation for the temperature:

$$\theta_t = \phi(\theta_z)_z + \tau\psi(\theta_z)_{zt}, \tag{16}$$

where the functions ϕ and ψ are defined by

$$\phi(p) = pk_0(p) \quad \text{for} \quad p \geqslant 0 \tag{17}$$

and

$$\psi(p) = -\int_0^p sk_0'(s)\,\mathrm{d}s = -\phi(p) + \int_0^p k_0(s)\,\mathrm{d}s \quad \text{for} \quad p \geqslant 0. \tag{18}$$

Since k_0 is strictly decreasing, ψ is strictly increasing for $t > 0$. In addition we shall assume that

$$-\int_0^{+\infty} sk_0'(s)\,\mathrm{d}s < +\infty,$$

i.e. that k_0 tends to zero sufficiently fast as $p \to +\infty$, which implies that

$$\psi(+\infty) = \gamma < +\infty.$$

The graph of ψ is represented in figure 1 (*b*).

Differentiating (16) with respect to z and denoting θ_z by v, we obtain another form of the basic equation:

$$v_t = \phi(v)_{zz} + \tau\psi(v)_{zzt}. \tag{19}$$

It is a natural question to ask why the third-order terms in (1) and (19) may be omitted if $\theta_z < \alpha$ everywhere in the field, and at the same time they should be necessarily retained if somewhere ϕ_z is larger than α. A simple answer to this question is supplied if we consider equation (1) with 'frozen' coefficients

$$\theta_t = A\theta_{zz} + B\theta_{zzt}. \tag{20}$$

Expanding the initial conditions in a Fourier series, we may, as usual, consider solutions of the type

$$\theta = c\mathrm{e}^{\mathrm{i}(kz-\omega t)}. \tag{21}$$

From (20) and (21) we obtain

$$-\mathrm{i}\omega = -Ak^2/(1+Bk^2), \tag{22}$$

and therefore

$$\theta = c\exp\left[\mathrm{i}kz - Ak^2t/(1+Bk^2)\right]. \tag{23}$$

If $B = 0$ (i.e. we neglect third-order terms), then for $A > 0$ (subcritical values of the temperature gradient) all perturbations decay, and for $A < 0$ they grow exponentially and the growth rate of high-spatial-frequency solutions (large k) increases with frequency.

If $B > 0$ (the case considered in this paper), all frequencies for $A < 0$ grow uniformly, and incorrectness is prevented. For $A > 0$ the third-order term does not play a significant role because perturbations decay exponentially for all k.

We observe that it is not surprising that steps of the temperature arise in numerical computations if we omit the third-order term (Posmentier 1977; Djumagazieva 1983). Numerically difference (rather than differential) equations are always solved, and expanding first- and second-order differences into series we obtain 'numerical' higher-order terms which influence the computations and which play a stabilizing role. These terms, however, are uncontrolled and have no direct physical sense. Therefore, analytically we cannot omit the third-order term: it is of crucial importance if somewhere in the region the temperature gradient is supercritical.

Equation (20) can also be represented in the form

$$\theta_t + Q_z = 0, \quad Q = -A\theta_z - B(\theta_z)_t.$$

Therefore in the case $A < 0$, the impression can arise that for the steady state the 'temperature flux' $Q = -A\theta_z$ corresponds to the heat flow along the temperature gradient, which contradicts the second law of thermodynamics. In fact, for every unsteady temperature distribution the second term of the temperature flux $B(\theta_z)_t$ overweighs the first terms and preserves the right direction of the heat flow.

3. Mathematical analysis

First we specify and discuss several hypotheses on the functions entering the equation and initial condition which we shall use for the mathematical analysis of the basic problem (4):

H1. the function ψ is a smooth function defined on the interval $0 \leqslant p < +\infty$, $\psi(0) = 0, \psi'(p) > 0$ for $p \geqslant 0$, and $\psi(+\infty) = \gamma$, where $0 < \gamma < +\infty$;

H2. the function ϕ is a smooth function defined on the same interval, $\phi(0) = 0$, $\phi(+\infty) = 0, 0 < \phi(p) \leqslant \phi(\alpha)$ for $0 < p < +\infty$ for some $\alpha > 0$, and there exists constants, k_1 and k_2 such that

$$|\phi'(p)| \leqslant k_1 \psi'(p) \quad \text{and} \quad \left|\left(\frac{\phi'(p)}{\psi'(p)}\right)' \frac{1}{\psi'(p)}\right| \leqslant k_2 \quad \text{for} \quad p \geqslant 0; \tag{24}$$

H3. the function $\theta_0(z)$ (giving the initial temperature distribution) is a smooth and non-decreasing function defined on the interval $0 \leqslant z \leqslant L$, and $\theta_0'(0) = \theta_0'(L) = 0$.

Hypothesis H1 implies that

$$\psi'(0) > 0; \tag{25}$$

on the other hand it follows from (18) that $\psi'(p) = -pk_0'(p)$, whence $\psi'(0) = 0$ if $k_0'(p)$ remains bounded as $p \to 0$. Condition (25) is mathematically convenient since it avoids additional technical difficulties at points where the temperature gradient vanishes; we remind the reader that the interesting physical phenomena occur for temperature gradients which are larger than the threshold value α, and in this paper we focus our attention on the mathematical description of what happens for large gradients. In addition we observe that, for small values of the temperature gradient, it is impossible to determine the steady diffusivity $k_0(\theta_z)$ accurately from experimental data, since the quantity measured is the flux, which should be divided by the temperature gradient to obtain the diffusivity. Therefore we can impose at small values of temperature gradient mathematically convenient condition (25).

A more serious mathematical restriction is condition (24), which replaces the natural condition (3). Actually condition (3) is sufficient to prove the existence of a solution (see in Barenblatt *et al.* 1993, the remark after the proof of the lemma 5.2) and the condition (24) is used to prove its uniqueness. In terms of k_0 condition (24) can be rewritten as

$$\left|\frac{k_0(p)}{pk_0'(p)}\right| \leqslant k_1 + 1 \quad \text{and} \quad \left|\left(\frac{k_0(p)}{pk_0'(p)}\right)' \frac{1}{pk_0'(p)}\right| \leqslant k_2 \quad \text{for} \quad p \geqslant 0,$$

and it is essentially a condition of the behaviour of $k_0(p)$ for large values of p. We observe that (24) is satisfied for large values of p if $k_0(p)$ is defined by

$$k_0(p) = \frac{A}{B + p^\nu} \quad \text{if} \quad \nu > 0,$$

or

$$k_0(p) = \frac{A}{(B+p)^\nu} \quad \text{if} \quad 0 < \nu \leqslant 2.$$

Note that we did not assume in hypothesis H2 that $\phi(p)$ is increasing for $0 < p < \alpha$ and decreasing for $p > \alpha$; we only assumed that ϕ has a maximum at $p = \alpha$, and that $\phi \leqslant \phi(\alpha)$ everywhere. Therefore, generally speaking, there may be several intervals in which the function ϕ decreases or increases.

As we have announced in the Introduction, the solution to problem (4) may become discontinuous within finite time. Since such a solution is not a classical one, we have to define what we mean by a (generalized) solution of problem (4). In particular we have to define what we mean by $\psi(\theta_z)$ and $\phi(\theta_z)$ at points at which θ is discontinuous. The main idea of the definition which we give below is to replace the function $\psi(\theta_z)$ in the equation by a continuous function $w(z, t)$ which coincides with $\psi(\theta_z)$ at the points where θ is smooth, and which is equal to the value of ψ at infinity, γ, at the points where θ is discontinuous; more precisely, we replace

$$\psi(\theta_z(z,t)) = \psi\left(\lim_{h\to 0} \frac{\theta(z+h,t) - \theta(z,t)}{h}\right)$$

by

$$w(z,t) = \lim_{h\to 0} \psi\left(\frac{\theta(z+h,t) - \theta(z^\pm,t)}{h}\right),$$

where $\theta(z^+, t)$ and $\theta(z^-, t)$ denote respectively the left and right limit of θ.

Throughout this paper we shall use the notation

$$Q = \{(z,t): 0 < z < L, t > 0\} \quad \text{and} \quad Q_T = \{(z,t): 0 < z < L, 0 < t \leqslant T\},$$

where $T > 0$.

Definition. A bounded function θ defined on Q which is non-decreasing with respect to z is a solution of the problem (4) if for any $T > 0$:

(i) there exists a continuous function w defined on Q such that $0 \leqslant w \leqslant \gamma$ in Q, and

$$\begin{aligned} w(z,t) &= \lim_{h\to 0} \psi\left(\frac{\theta(z+h,t)-\theta(z^+,t)}{h}\right) \\ &= \lim_{h\to 0} \psi\left(\frac{\theta(z+h,t)-\theta(z^-,t)}{h}\right) \quad \text{for} \quad 0 < z < L \quad \text{and} \quad t > 0; \end{aligned} \tag{26}$$

(ii) θ_t, w_t and w_{zt} are squared integrable in Q_T, there exists a constant C such that

$$\int_0^L w^2(z,t)\,\mathrm{d}z < C \quad \text{for} \quad 0 \leqslant t \leqslant T,$$

θ and s satisfy the condition that the relation (ψ^{-1} is the inverse of ψ)

$$\theta_t = \phi(\psi^{-1}(w))_z + \tau w_{zt} \tag{27}$$

holds almost everywhere in Q_T, i.e. except a set of zero measure, and

$$\theta(z,0) = \theta_0(z) \quad \text{for} \quad 0 < z < L.$$

It turns out that problem (4) has exactly one solution in this class of solutions.

THEOREM 1 (*Existence and uniqueness*). *Let hypotheses H*1, *H*2 *and H*3 *be satisfied. Then problem* (4) *possesses a unique solution* θ.

For the proof we refer to Barenblatt *et al.* (1993).

Now we are ready to state the main results about the qualitative behaviour of the solutions, which are of primary importance here. The first result says that if the temperature gradient in a point z_0 at a certain time t_0 is smaller than the threshold value α, then it remains smaller than α at z_0 for all later times $t > t_0$:

THEOREM 2 (*Gradient estimate*). *Let hypotheses H*1, *H*2 *and H*3 *be satisfied, and let* $\theta(z,t)$ *be the solution of problem* (4). *If* $\theta_z(z_0, t_0) \leqslant \alpha$ *for some* $z_0 \in (0, L)$ *and* $t_0 \geqslant 0$, *then* $\theta_z(z_0, t) \leqslant \alpha$ *for all* $t > t_0$.

It follows from this result that if the initial temperature gradient is smaller than α at every point, then it remains always smaller than α. There exist however initial temperature distributions for which the gradient becomes infinite in finite time:

THEOREM 3 (*Formation of discontinuities*). *Let hypotheses H*1 *and H*2 *be satisfied. Then there exists smooth initial functions* $\theta_0(z)$ *which satisfy hypothesis H*3, *such that the corresponding solutions of problem* (4) *are not continuous in* Q.

Once a solution is discontinuous at a point z_0, the solution remains discontinuous at that point and the temperature jump $\theta(z_0^+, t) - \theta(z_0^-, t)$ is non-decreasing in time:

THEOREM 4 (*Persistence of discontinuities*). *Let hypotheses H1, H2 and H3 be satisfied. Let θ be the solution of problem* (4) *in the sense determined above, and let w be defined by* (26). *If for some* $0 < z_0 < L$ *and* $t_0 > 0$

$$w(z_0, t_0) = \gamma,$$

then

$$w(z_0, t) = \gamma \quad \text{for} \quad t > t_0,$$

and

$$\left.\begin{array}{l}\textit{the function } \theta(z_0^-, t) \textit{ is non-increasing in } (t_0, +\infty),\\ \textit{the function } \theta(z_0^+, t) \textit{ is non-decreasing in } (t_0, +\infty).\end{array}\right\} \tag{28}$$

Finally we consider the asymptotic behaviour of the temperature distributions; if a solution becomes discontinuous, it converges to a step-wise temperature distribution as $t \to +\infty$:

THEOREM 5 (*Convergence to step-wise solutions*). *Let hypotheses H1, H2 and H3 be satisfied and let θ be the solution of problem* (4). *Then there exists a non-decreasing function q defined on the interval* $0 \leqslant z \leqslant L$ *which satisfies*

$$q'(z) = 0 \textit{ for almost every } z \in (0, L)$$

such that

$$\theta(z, t) \to q(z) \textit{ as } t \to \infty \textit{ for almost every } 0 < z < L.$$

If θ is not everywhere continuous in Q, then q is non-constant in the interval $0 < z < L$.

In the following section we shall indicate the formal proofs of some of these results. To keep the proofs as transparent as possible, we shall apply techniques to the generalized solutions which can only be justified when applied to smooth solutions. For the rigorous proofs and for the proof of Theorem 3, which, because of its length and its non-constructive nature, will be omitted in the present work, we refer to Barenblatt *et al.* (1993).

4. Outline of the proofs

All proofs are based on the following (trivial) version of the maximum principle: if the function $a(z)$ is non-negative in an interval $a < z < b$ and if $u(z)$ is a smooth function in the interval $a \leqslant z \leqslant b$, which satisfies the conditions

$$\left.\begin{array}{l} M_1 \leqslant u(z) - a(z)\,u''(z) \leqslant M_2 \\ M_1 \leqslant u(a) \leqslant M_2, M_1 \leqslant u(b) \leqslant M_2 \end{array}\right\} \quad \text{for} \quad a < z < b,$$

then

$$M_1 \leqslant u(z) \leqslant M_2 \quad \text{for} \quad a \leqslant z \leqslant b.$$

We begin with the proof of the gradient estimate in Theorem 2, and define the function

$$v(z, t) = \int_0^z \theta(y, t)\,\mathrm{d}y \quad \text{for} \quad 0 \leqslant z \leqslant L, t \geqslant 0. \tag{29}$$

Then

$$v_z(z, t) = \theta(z, t), \tag{30}$$

and, by (1),

$$v_t(z, t) = \int_0^z \theta_t(y, t)\,\mathrm{d}y = \tau\psi(\theta_z(z, t))_t + \phi(\theta_z(z, t)) - \tau\psi(\theta_z(0, t))_t - \phi(\theta_z(0, t)).$$

In view of the boundary condition at $z = 0$, this relation assumes the form

$$v_t = \tau\psi(\theta_z)_t + \phi(\theta_z), \tag{31}$$

and we obtain from (30) and the boundary condition at $z = L$ that $v_t(z, t)$ satisfies, for any $t > 0$,

$$\left.\begin{aligned} &v_t - \tau\psi'(\theta_z)\, v_{tzz} = \phi(\theta_z) \\ &v_t(0, t) = v_t(L, t) = 0. \end{aligned}\right\} \quad \text{for} \quad 0 < z < L \tag{32}$$

Since $0 \leqslant \phi(\theta_z(z, t)) \leqslant \phi(\alpha)$, we may apply, for fixed but arbitrary $t > 0$, the maximum principle and we find that

$$0 \leqslant v_t(z, t) \leqslant \phi(\alpha) \quad \text{for} \quad 0 \leqslant z \leqslant L, t > 0. \tag{32}$$

Combining the second inequality in (33) with (31), we have that

$$\tau\psi'(\theta_z)\,\theta_{zt} \leqslant \phi(\alpha) - \phi(\theta_z) \quad \text{for} \quad 0 \leqslant z \leqslant L, t > 0. \tag{34}$$

If we fix a value $z = z_0$, (34) can be considered as an ordinary differential inequality for the gradient $\theta_z(z_0, t)$, and since the right-hand side of (34) vanishes if the gradient attains the value α, it follows at once that

$$\theta_z(z_0, t_0) \leqslant \alpha \quad \text{implies that} \quad \theta_z(z_0, t) \leqslant \alpha \quad \text{for all} \quad t > t_0,$$

and we have obtained the gradient estimate of Theorem 2.

Next we consider the proof of Theorem 4: assuming that the gradient

$$\theta_z(z_0, t_0) = +\infty \tag{35}$$

at some point (z_0, t_0), we have to show that

$$\theta_z(z_0, t) = +\infty \quad \text{for all} \quad t > t_0 \tag{36}$$

and

$$\left.\begin{aligned} &\theta(z_0^-, t) \text{ is non-increasing with respect to } t > t_0, \\ &\theta(z_0^+, t) \text{ is non-decreasing with respect to } t > t_0 \end{aligned}\right\} \tag{37}$$

We observe that, intuitively, since $\psi'(+\infty) = \phi'(+\infty) = 0$ relation (36) means that, for $t > t_0$ there is no interaction between the temperature fields in the regions where $z < z_0$ and $z > z_0$. Below we use this decoupling (in a rigorous way), constructing the solution in the two regions independently.

First we consider the problem for $z > z_0$ and $t > t_0$, requiring that the gradient is infinite at z_0:

$$\left.\begin{aligned} &\bar\theta_t = \phi(\bar\theta_z)_z + \tau\psi(\bar\theta_z)_{zt} && \text{for} \quad z_0 < z < L, t > t_0, \\ &\bar\theta_z(z_0, t) = +\infty && \text{for} \quad t > t_0, \\ &\bar\theta_z(L, t) = 0 && \text{for} \quad t > t_0, \\ &\bar\theta(z, t_0) = \theta(z, t_0) && \text{for} \quad z_0 < z < L. \end{aligned}\right\} \tag{38}$$

It can be proved that (38) possesses a unique (generalized) solution $\bar\theta(z, t)$, which, as we shall show below, satisfies the condition

$$\bar\theta(z_0^+, t) \text{ is non-decreasing with respect to } t > t_0. \tag{39}$$

the temperature θ at some point and at some time implies a discontinuity of the limiting temperature q at the same point. We emphasize that this discontinuity of the temperature is possible because the flux is equal to zero at both zero and infinite temperature gradient (see figure 1 *a*).

5. Numerical results

The results of several numerical calculations will be presented below to illustrate step formation from smooth temperature distributions. Assuming that the equilibrium temperature diffusivity coefficient is given by the relation (9), we can rewrite (19) for the temperature gradient in the form

$$u_\vartheta = \left(\frac{u}{1+u^2}\right)_{\xi\xi} + \psi\left(\frac{u^2 u_\vartheta}{(1+u^2)^2}\right)_{\xi\xi}, \tag{41}$$

where the following dimensionless variables have been introduced:

$$\vartheta = \frac{At}{L^2}, \quad \xi = \frac{z}{L}, \quad u = B^{\frac{1}{2}}\theta_z, \quad \sigma = \frac{2A\tau}{L^2}. \tag{42}$$

We observe that the maximum of the heat flux corresponds to the value $u = 1$ of the dimensionless temperature gradient.

We consider the initial-boundary value problem for (41) in the region $0 \leqslant \xi \leqslant 1$, $\vartheta \geqslant 0$, under the following boundary and initial conditions:

$$u(0,\vartheta) = 0, \quad u(1,\vartheta) = 0, \tag{43}$$

$$u(\xi, 0) = u_0(\xi), \tag{44}$$

where $u_0(\xi)$ is given smooth and positive function.

The numerical solution is simplified because the problem is formulated in terms of the temperature gradients only. Our main purpose here is to demonstrate numerically the formation of steps, and to investigate how the number of steps depends on the dimensionless relaxation time σ. It is clear that for step formation from a smooth initial temperature distribution a supercritical part of the initial distribution is needed, i.e. the existence of a part of the interval $0 \leqslant \xi \leqslant 1$ where $u_0(\xi) > 1$.

We performed our computations for the following initial condition:

$$u_0(\xi) = 4U_0(\tfrac{4}{3}\xi)^{\frac{7}{3}}(1-\xi), \tag{45}$$

where U_0 is a positive constant (the initial dimensionless temperature is determined by its gradient, given by (45), and the condition that it vanishes at $\xi = 0$).

The algorithm for finding the numerical solution of problem (41), (43), (45) is rather simple; we have only to note that an implicit approximation, and therefore an iteration procedure, is needed for the right-hand side of (41) due to its nonlinearity. We use the difference net in the region $0 \leqslant \xi \leqslant 1$ defined by $\xi_i = ih, i = 0, 1, \ldots, N$, where$h = 1/N$ and N is the number of nodes of the difference net. In all calculations the value $N =$ 200 was used. The time discretization of (41) was performed in the following way:

$$\frac{\hat{u}-\check{u}}{\Delta\vartheta} = \left(\frac{\hat{u}+\check{u}}{2(1+\bar{u}^2)}\right)_{\xi\xi} + \sigma\left(\frac{\bar{u}^2}{(1+\bar{u}^2)^2}\frac{\hat{u}-\check{u}}{\Delta\vartheta}\right)_{\xi\xi}. \tag{46}$$

here $\Delta\vartheta = \vartheta_{n+1} - \vartheta_n$ is the increment of the dimensionless time, $\check{u} = u(\vartheta_n), \hat{u} = u(\vartheta_{n+1})$, $\bar{u} = \eta\hat{u} + (1-\eta)\check{u}, 0 \leqslant \eta \leqslant 1$.

Indeed, defining

$$\bar{v}(z,t) = \int_{z_0}^{z} \bar{\theta}(y,t)\,\mathrm{d}y,$$

and arguing as in the proof of the gradient estimate, we obtain that for any $t > t_0$

$$\left.\begin{aligned} &\bar{v}_t - \tau\psi'(\bar{\theta}_z)\,\bar{v}_{tzz} = \phi(\bar{\theta}_z) \geqslant 0 \\ &\bar{v}_t(z_0,t) = \bar{v}_t(L,t) = 0. \end{aligned}\right\} \quad \text{for} \quad z_0 < z < L.$$

It follows from the maximum principle that $\bar{v}_t(z,t) \geqslant 0$, i.e. that for any $z_0 < z < L$

$$\bar{v}(z,t) \text{ is non-decreasing with respect to } t > t_0.$$

Since $\bar{v}(z_0, t) = 0$ for all $t > t_0$, this implies that

$$\bar{v}_z(z_0,t) \text{ is non-decreasing with respect to } t > t_0$$

and (39) follows from the relation $\bar{\theta}(z,t) = \bar{v}_z(z,t)$.

In a completely analogous manner we construct a solution $\underline{\theta}(z,t)$ for $0 \leqslant z > z_0$ and $t \geqslant t_0$ with boundary conditions $\underline{\theta}_z(0,t)$ and $\underline{\theta}_z(z_0,t) = +\infty$, and the property

$$\underline{\theta}(z_0^-, t) \text{ is non-increasing with respect to } t > t_0. \tag{40}$$

In view of (35), (39) and (40), a straightforward calculations shows that the function

$$\hat{\theta}(z,t) = \begin{cases} \underline{\theta}(z,t) & \text{if} \quad 0 \leqslant z \leqslant z_0, t \geqslant t_0 \\ \bar{\theta}(z,t) & \text{if} \quad z_0 < z \leqslant L, t \geqslant t_0 \end{cases}$$

is a (generalized) solution of the original problem (4) for $t \geqslant t_0$. Since $\hat{\theta}(z,t_0) = \theta(z,t_0)$ for all z, it follows from the uniqueness of solutions of (4) that $\hat{\theta}(z,t)$ and $\theta(z,t)$ coincide for all $t > t_0$. Finally it follows from the construction of $\hat{\theta}$ and from (39) and (40) that the desired properties of (36) and (37) are satisfied.

We conclude this section with a discussion of the large-time behaviour of the solutions. First we indicate the construction of the limiting temperature profile $q(z)$.

Let $v(z,t)$ be defined by (29). It follows from the first inequality in (33) that

$$v(z,t) \text{ is non-decreasing with respect to } t.$$

By (30), $v_{zz} = \theta_z \geqslant 0$, i.e. $v(z,t)$ is convex with respect to z and, being constant at the boundaries $z = 0$ and $z = L$, v is uniformly bounded. Hence the pointwise limit

$$\bar{v}(z) = \lim_{t\to+\infty} v(z,t) \quad \text{for} \quad 0 \leqslant z \leqslant L$$

exists, and $\bar{v}$ is a convex function. It is well-known that convex functions are differentiable at all points, and we denote the derivative at the points where $\bar{v}$ is differentiable by q:

$$q(z) = \bar{v}'(z).$$

The fact that q is a piecewise-constant function follows essentially from the fact that q is a steady-state solution, i.e. it satisfies the condition

$$\phi(q'(z)) = 0 \quad \text{for} \quad 0 < z < L.$$

Hence $q'(z)$ vanishes at all points at which the gradient is finite, but since $\phi(+\infty) = 0$ we cannot exclude discontinuities of $q(z)$. Indeed it follows from Theorem 4 that if $\theta(z_0^+, t_0) > \theta(z_0^-, t_0)$, then $q(z_0^+) \geqslant \theta(z_0^+, t_0) > \theta(z_0^-, t_0) \geqslant q(z_0^-)$, i.e. a discontinuity of

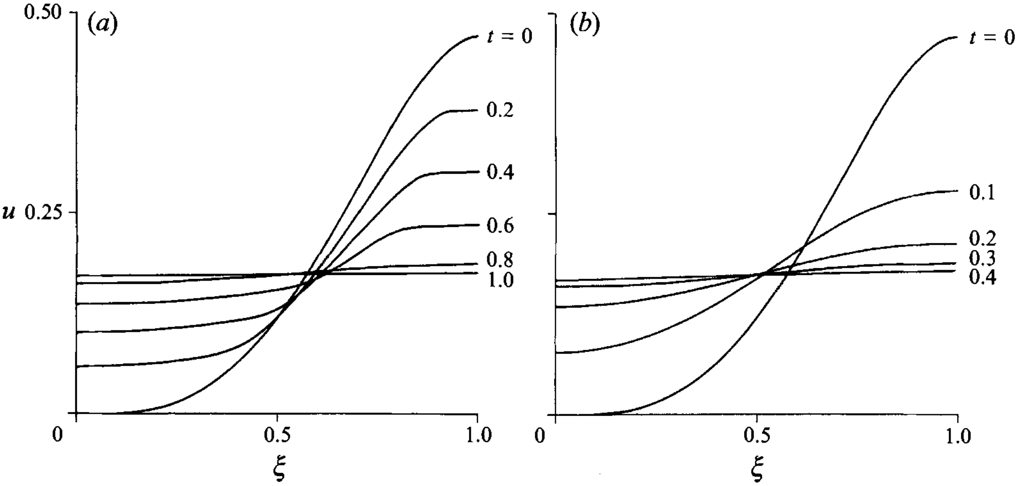

FIGURE 3. Evolution of subcritical ($U_0 = 0.9$) initial temperature distribution. No steps are formed for $\sigma = 1$, (*a*), and for $\sigma = 10^{-3}$, (*b*).

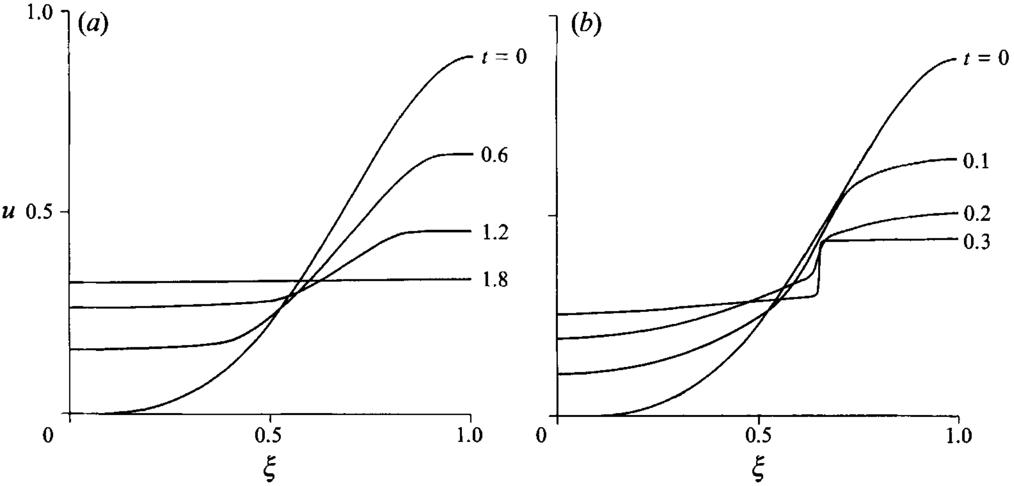

FIGURE 4. Evolution of supercritical ($U_0 = 1.7$) initial temperature distribution. No steps are formed at large ($\sigma = 1$) relaxation time, (*a*), one step is formed at medium ($\sigma = 10^{-2}$) relaxation time, (*b*).

For the approximation of the second space derivative on the difference net we used the formula

$$f_{\xi\xi} \approx \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2} \tag{47}$$

with second-order accuracy. Due to the implicit time approximation we could perform the computation with $\Delta\vartheta = O(h)$. Because of the nonlinearity of the resulting system of algebraic equations, an iteration procedure was necessary to compute the distribution of $\hat{u}$ over the next time layer. Ten iterations were usually required on each time layer to obtain convergence with accuracy 10^{-3}, except for time intervals where the fast evolution of the solution took place; in the latter case several tens of iterations were necessary. The best convergence results were obtained for the completely implicit

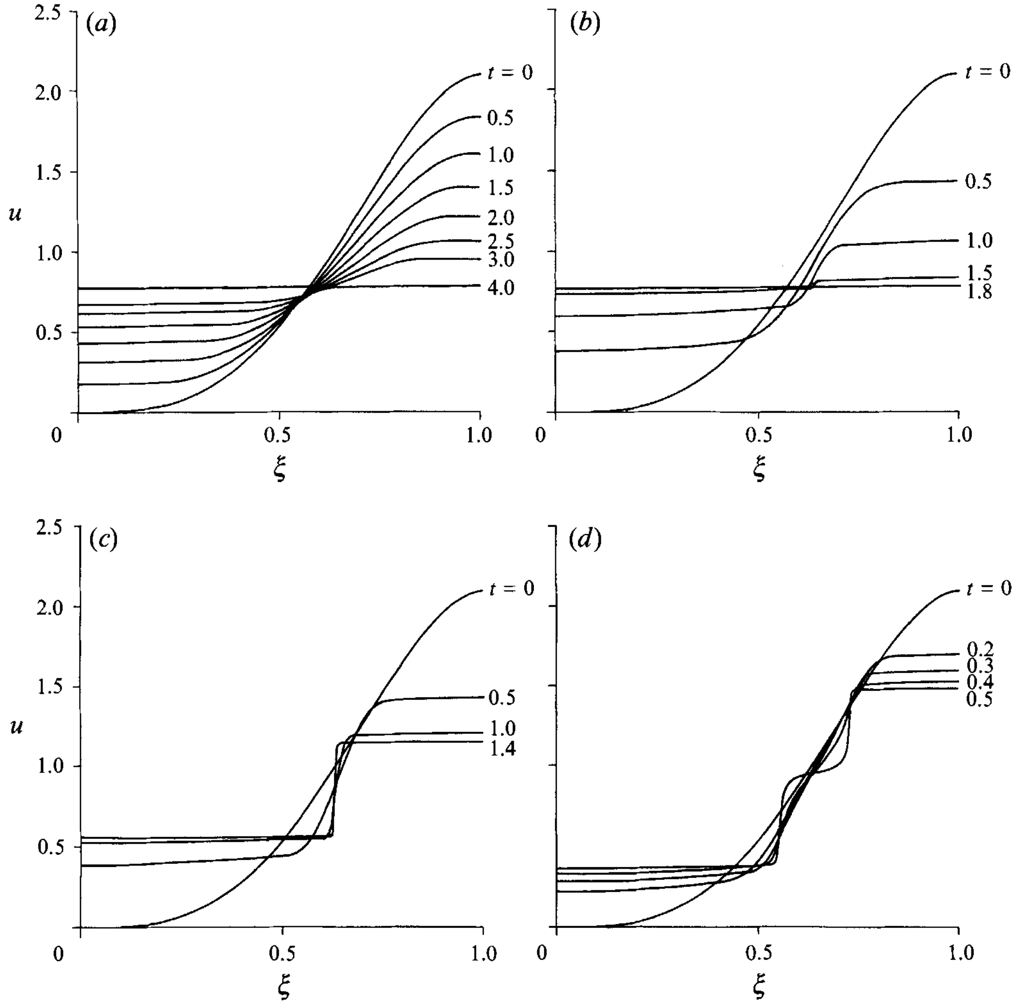

FIGURE 5. Evolution of strongly supercritical ($U_0 = 4$) initial temperature distribution. (a) $\sigma = 1$: no steps are formed at large relaxation time. (b) $\sigma = 10^{-1}$: tendency to the formation of one step is observed. (c) $\sigma = 5\times10^{-2}$: one step is formed at medium relaxation time. (d) $\sigma = 2.5\times10^{-2}$: two steps are formed at medium relaxation time.

scheme, i.e. $\eta = 1$ (the first iteration on each time layer was always performed with $\eta = 0$).

After obtaining the gradient on the new time layer, the temperature was calculated by simple integration, and the integration constant was determined from the condition

$$\int_0^1 \theta(\xi,\vartheta)\,\mathrm{d}\xi = \int_0^1 \theta(\xi,0)\,\mathrm{d}\xi.$$

In figures 3–6 the results of some calculations are represented for three different values of the constant U_0 (corresponding to an initial temperature with, respectively, subcritical, supercritical and strongly supercritical gradient) and for decreasing values of the dimensionless relaxation time σ.

We may conclude that the numerical calculations confirm the analytic results presented above, and they indicate that if the gradient is supercritical, more and more steps are formed with decreasing σ.

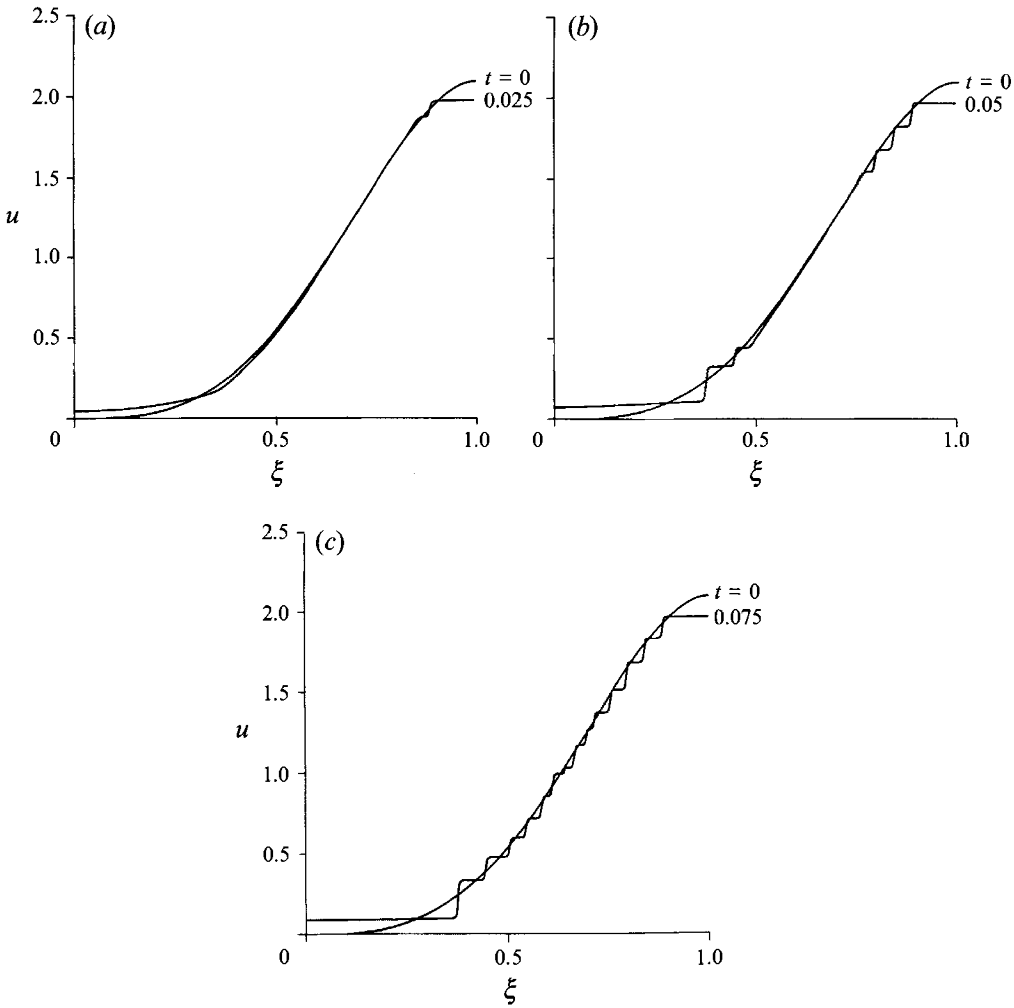

FIGURE 6. Evolution of strongly supercritical ($U_0 = 4$) initial temperature distribution. Multiple steps are formed at small ($\sigma = 10^{-3}$) relaxation time at the dimensionless time moment $\vartheta = 0.025$, (*a*) $\vartheta = 0.05$, (*b*), and $\vartheta = 0.075$, (*c*).

The authors wish to express their sincere thanks to the Institute of Mathematics and its Applications, University of Minnesota, USA, where this work was started, and which has supported this research in part with funds provided by the US National Science Foundation. The coordinating role of Professor Shoshana Kamin is warmly appreciated.

We thank Professor G. K. Batchelor and Dr P. F. Linden (DAMTP, University of Cambridge) for valuable advice and for supplying us with important references.

REFERENCES

BARENBLATT, G. I., BERTSCH, M., DAL PASSO, R. & UGHI, M. 1993 A degenerate pseudoparabolic regularisation of a nonlinear forward backward heat equation arising in the theory of heat and mass transfer in stably stratified turbulent shear flow. *SIAM J. Nonlin. Anal.* (to appear).

DJUMAGAZIEVA, S. KH. 1983 Numerical integration of a certain partial differential equation. *J. Numer. Math. Phys.* **23**, 839–847.

FELSENBAUM, A. I. & BOGUSLAVSKY, S. G. 1977 On a relation between heat flux and temperature gradient in the sea. *Dokl. (C. R.) USSR Acad. Sci.* **235-3**, 557–559.

HÖLLIG, K. 1983 Existence of infinitely many solutions for a forward backward heat equation. *Trans. Am. Math. Soc.* **278**, 299–316.

IVEY, G. N. & IMBERGER, G. 1991 On the nature of turbulence in stratified fluid. Part 1. The energetics of mixing. *J. Phys. Oceanogr.* **21**, 650–658.

LINDEN, P. F. 1979 Mixing in stratified fluids. *Geophys. Astrophys. Fluid Dyn* **13**, 3–23.

MAMAYEV, O. I. 1958 The influence of stratification on vertical turbulent mixing in the sea. *Izv. (Bull.) USSR Acad. Sci. Ser. Geophys.* **4**, 870–875.

MONIN, A. S. & YAGLOM, A. M. 1971 *Statistical Fluid Mechanics: Mechanics of Turbulence*, vol. 1. MIT Press.

MUNK, W. H. & ANDERSON, F. R. 1948 Notes on a theory of the thermocline. *J. Mar. Res.* **7**, 276–295.

PADRON, V. 1990 Sobolev regularization of some nonlinear ill-posed problems. Thesis, Department of Mathematics, University of Minnesota.

PETUKHOV, B. S. & POLYAKOV, A. F. 1989 *Heat Transfer in Turbulent Mixed Convection*. Hemisphere.

PHILLIPS, O. M. 1972 Turbulence in a strongly stratified fluid is unstable? *Deep-Sea Res.* **19**, 79–81.

POSMENTIER, E. S. 1977 The generation of salinity finestructures by vertical diffusion. *J. Phys. Oceanogr.* **7**, 298–300.

ROSSBY, C. G. & MONTGOMERY, R. B. 1935 The layer of frictional influence in wind and ocean currents. *Pap. Phys. Oceanogr. Met.*, vol. 3.

RUDDICK, B. R., MCDOUGALL, T. J. & TURNER, J. S. 1989 The formation of layers in a uniformly stirred density gradient. *Deep-Sea Res.* **36**, 597–609.