

The effect of temperature-dependent solubility on the onset of thermosolutal convection in a horizontal porous layer

By DAVID PRITCHARD¹
AND CHRIS N. RICHARDSON²

¹ Department of Mathematics, University of Strathclyde, 26 Richmond St, Glasgow G1 1XH, Scotland. E-mail: dtp@maths.strath.ac.uk.

² B.P. Institute for Multiphase Flow, Department of Earth Sciences, University of Cambridge, Madingley Rise, Madingley Road, Cambridge CB3 0EZ, England.

(Received 20 June 2006)

We consider the onset of thermosolutal (double-diffusive) convection of a binary fluid in a horizontal porous layer subject to fixed temperatures and chemical equilibrium on the bounding surfaces, when the solubility of the dissolved component depends on temperature. We use a linear stability analysis to investigate how the dissolution or precipitation of this component affects the onset of convection and the selection of an unstable wavenumber; we extend this analysis using a Galerkin method to predict the structure of the initial bifurcation, and we compare our analytical results with numerical integration of the full nonlinear equations. We find that the reactive term may be stabilising or destabilising, with subtle effects particularly when the thermal gradient is destabilising but the solutal gradient is stabilising. The preferred spatial wavelength of convective cells at onset may also be substantially increased or reduced, and strongly reactive systems tend to prefer direct to subcritical bifurcations. These results have implications for geothermal reservoir management and ore prospecting.

1. Introduction

There are many geological systems in which buoyancy forces cause a fluid carrying dissolved minerals to circulate vertically through a porous medium. In many of these situations, the amount of solute in the fluid is not conserved, because it may be dissolved from or precipitated onto the porous matrix as its solubility varies with temperature, pressure and the local rock chemistry. The effect which this gain or loss of dissolved minerals may have on the buoyancy of the fluid, and thus on its convective motion, is largely unknown; the aim of this study is to develop understanding of this effect by examining the effect of reactions on the onset of convective motion.

1.1. *Thermal and thermosolutal convection*

Natural convection in a fluid layer, first investigated by Bénard and by Rayleigh in the early 20th century, remains a subject of active research and provides an important paradigm for instability, for extended pattern formation and for the transition of a system to chaos or turbulence (see e.g. Bodenschatz, Pesch & Ahlers 2000). The problem of thermal convection in a porous medium dates back to the stability analyses by Horton & Rogers (1945) and Lapwood (1948), and shares many features with the pure-fluid problem, although in general the mathematical formulation is slightly simpler.

Thermosolutal convection (also known as thermochemical convection) occurs when both temperature and some dissolved substance contribute to the buoyancy of the fluid. In a fluid layer, the crucial difference between the components is that they diffuse at different rates, and the process is often referred to as double-diffusive convection (DDC). In a porous medium, the system is also double-advective: because heat must be shared between the fluid and the porous medium while solute is confined to the pore space, thermal signals are advected more slowly than solutal signals, and this difference in advection rates is crucial to the stability properties of the system (Phillips 1991, §5.2). The first linear stability analysis of a double-diffusive system in a porous layer was carried out by Nield (1968), who considered the limiting case in which transport is isoadvective, and there has been much subsequent development of this work (Rudraiah, Siddheshwar & Masuoka 2003).

The most interesting dynamics in DDC occur when the (rapidly-diffusing) thermal and (slowly-diffusing) solute fields contribute in opposite senses to the buoyancy gradient. What happens depends on which of these components is destabilising, and in particular on whether the solutal gradient is stabilising or destabilising. In the solutally unstable (SU) regime, as the solutal or thermal Rayleigh number is increased the motionless conductive state loses stability through a supercritical bifurcation to a state of steady convective motion with finite amplitude. The initial instability is through exponentially growing perturbations: we will refer to this below as monotonic instability. The resulting flow pattern is known as ‘fingering’ (Turner 1985); in some studies, the fingering regime is implicitly identified with the SU regime, but since the nonlinear development of the flow lies beyond the current study we will avoid this terminology here.

The solutally stable (SS) regime of convection, in which the solute is stabilising and the temperature destabilising, is more complicated, and there are two important features of the onset of convection in this regime. One is that the principle of exchange of stabilities does not always hold, so the conductive base state may become unstable through a growing oscillatory (‘overdamped’) mode. The other is that stable steady convective rolls may occur even when the thermal Rayleigh number Ra_T is below its critical value, so the flow observed in practice may not be well predicted by a linear analysis taking the conductive state as its basis (Murray & Chen 1989). The onset of convection in this case is through a subcritical bifurcation, with hysteresis possible as Ra_T is varied. The numerical investigation of Mamou & Vasseur (1999) has indicated that additional unsteady solutions may exist, including oscillatory nonlinear convection and travelling waves. Ultimately, the system may become layered, with a series of vertically stacked convective cells, or penetrative, where cells occupy the entire porous layer and may extend into an overlying layer (see the discussion by Oldenburg & Pruess 1998). At higher Rayleigh numbers still, fully chaotic motions are possible (Schoofs & Spera 2003).

1.2. *Geological applications and reactive effects*

Solutal and thermosolutal convection in porous media may be important in a range of geological processes, including the dolomitisation of carbonate platforms (Kaufman 1994), soil salinisation (Gilman & Bear 1994; Wooding, Tyler & White 1997), and heat transfer in geothermal reservoirs (Oldenburg & Pruess 1998). Large-scale convective circulations of groundwater containing dissolved minerals in sandstone aquifers have also been proposed to explain the location of uranium ore deposits (Raffensperger & Garven 1995*a,b*). In many of these situations, the exchange of dissolved species with the porous medium is directly significant, or may be significant through its effect on convection, but these effects have received relatively little attention.

The earliest analysis of reactive effects on convection in a fluid layer was carried out

by Wollkind & Frisch (1971*a,b*), who considered convection combined with a dissociation reaction; complementary work (e.g. Bdzil & Frisch 1971) investigated situations where this was catalysed at the lower boundary of the layer. This analysis was revisited and developed by Bdzil & Frisch (1980), while similar work was carried out simultaneously by Gutkowitz-Krusin & Ross (1980). In these pure-fluid systems, dissolved species cannot be lost from the solution, but their reaction still affects the bulk density through changes to the solutal component and (often more significantly) to the temperature.

The first study of reactive convection in a porous medium was due to Steinberg & Brand (1983, 1984). The model which they considered was rather similar to that which we will explore below, but they restricted their analysis to the regime where the reaction was sufficiently fast that solutal diffusion could be neglected: we will discuss the effect of this neglect in §3.6. Subsequent studies were carried out by Gatica, Viljoen & Hlavacek (1989) and Viljoen, Gatica & Hlavacek (1990), who considered a more complicated exothermic reaction term, and made some progress in investigating the stability boundaries of the system; we note that they concentrated on the limiting case where the Lewis number $Le = 1$, so that the thermal and solutal diffusion rates are identical and overdamped behaviour is impossible.

1.3. *Aims and structure of the current study*

The aim of this study is to elucidate some fundamental mechanisms of reactive-convective instability. We will consider convection in a horizontal porous layer, bounded above and below by layers of constant temperature and solute concentration. This geometry is particularly relevant to geothermal reservoirs, where the vertical flux of heat through the layer may be substantially increased if convective motion occurs. It also resembles the problem considered by Raffensperger & Garven (1995*a,b*), although our model is considerably idealised. We will neglect the thermal contribution of the reaction, which is reasonable when the fluid motion is driven by large geothermal temperature gradients, and we will focus instead on the effect on convection of the solute exchange between the fluid and the porous matrix as the temperature varies. We will not confine ourselves to particular regimes of the Lewis number or reaction rate, and we will investigate more thoroughly than in previous studies how the stability diagram changes with the reaction rate.

Specifically, we will investigate the effect of temperature-dependent solubility on the following questions. (i) Under what circumstances is purely diffusive heat and mass transfer stable? (ii) What spatial structures (in particular, what horizontal wavelength) are favoured at the onset of instability? (iii) What is the bifurcation structure at the onset of instability? (iv) Under what circumstances does subcritical steady convection occur? In a future study we intend to investigate in more detail the nonlinear behaviour possible and to quantify the heat and mass transfer across the porous layer.

In §2 we derive a non-dimensional model of double-diffusive convection with thermally controlled equilibrium solubility. In §3 we carry out a linear stability analysis of the conductive basic state, with particular emphasis on the effect of precipitation or dissolution on the stability. In §4 we extend the linear results into the weakly non-linear regime using a Galerkin model, in order to investigate the bifurcation structure and to obtain some predictions for the occurrence of subcritical modes. In §5 we test our analytical results further by integrating the full system numerically. Finally, in §6 we summarise and discuss the physical implications of our results.

2. Description of a double-diffusive reaction–convection model

We consider a homogeneous, isotropic porous layer of depth \hat{H} with temperatures \hat{T}_0 and \hat{T}_1 and solutal mass concentrations \hat{C}_0 and \hat{C}_1 imposed at the bottom and top respectively. The bottom and top layers are presumed to be impermeable, and we also assume chemical equilibrium at the boundaries. Boundary conditions which represent more complicated geological structures and geochemical conditions could readily be considered (cf. Mamou & Vasseur 1999), but are beyond the scope of the current analysis.

We take a horizontal co-ordinate \hat{x} and a vertical co-ordinate \hat{z} which increases upwards. Making a Boussinesq approximation, the flow is governed by

$$\hat{\nabla} \cdot \hat{\mathbf{u}} = 0 \quad \text{and} \quad \hat{\mathbf{u}} = -\frac{\hat{K}}{\hat{\mu}} \hat{\nabla} \hat{p} - \frac{\hat{K}}{\hat{\mu}} \hat{\rho}(\hat{C}, \hat{T}) \hat{g} \mathbf{e}_z, \quad (2.1)$$

where \hat{K} is the matrix permeability, $\hat{\mu}$ is the fluid viscosity and $\hat{\rho}$ is the fluid density, and the (two-dimensional) transport velocity is $\hat{\mathbf{u}} = (\hat{u}, \hat{w})$.

The transport of heat and solute is described by the advection-diffusion equations (Phillips 1991, §§2.8 and 2.9)

$$(\hat{\rho}\hat{c})_m \frac{\partial \hat{T}}{\partial \hat{t}} + (\hat{\rho}\hat{c})_f (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{T} = (\hat{\rho}\hat{c})_m \hat{\kappa}_T \hat{\nabla}^2 \hat{T} \quad \text{and} \quad (2.2)$$

$$\phi \frac{\partial \hat{C}}{\partial \hat{t}} + (\hat{\mathbf{u}} \cdot \hat{\nabla}) \hat{C} = \phi \hat{\kappa}_C \hat{\nabla}^2 \hat{C} + \hat{k} (\hat{C}_{\text{eq}}(\hat{T}) - \hat{C}). \quad (2.3)$$

Here \hat{T} represents the temperature (we assume local thermal equilibrium between the fluid and the matrix), \hat{C} the mass concentration of solute in the fluid and $\hat{C}_{\text{eq}}(\hat{T})$ the equilibrium concentration of solute at a given temperature. $\hat{\kappa}_C$ is the molecular diffusivity of the solute through the fluid, while $\hat{\kappa}_T$ is the effective diffusivity of heat through the saturated medium; ϕ is the porosity of the matrix and $\hat{k} > 0$ is a lumped effective reaction rate. The volumetric heat capacity of the fluid is denoted by $(\hat{\rho}\hat{c})_f$, and that of the saturated medium as a whole by $(\hat{\rho}\hat{c})_m = \phi(\hat{\rho}\hat{c})_f + (1 - \phi)(\hat{\rho}\hat{c})_s$, where a subscript s denotes the properties of the solid matrix. In the spirit of the Boussinesq approximation, we take all material properties except $\hat{\rho}$ to be constant. In writing (2.2) and (2.3), we have also assumed that the permeability and porosity of the matrix are constant in time and space, and in particular that they do not evolve due to the reaction over the time considered: this is equivalent (Phillips 1991, §4.7) to assuming that we need only consider timescales of the order of $\hat{k}^{-1} \hat{\rho}_s / \hat{C}_{\text{typ}}$, where $\hat{\rho}_s$ is the solid density of the mineral, and \hat{C}_{typ} is a typical volumetric mass concentration of the dissolved mineral. In general, we may expect $\hat{\rho}_s / \hat{C}_{\text{typ}}$ to be quite large, and so our assumption is appropriate for a study of the onset of convection; however, when considering the further evolution of the system we would expect this interaction between the flow and the matrix properties to become important.

We take the density to be given by

$$\hat{\rho}(\hat{C}, \hat{T}) = \hat{\rho}_0 + \hat{\alpha}_C (\hat{C} - \hat{C}_0) + \hat{\alpha}_T (\hat{T} - \hat{T}_0), \quad (2.4)$$

and, following e.g. Jupp & Woods (2003), we will take the equilibrium solute concentration to vary linearly in \hat{T} . We write $\hat{C}_{\text{eq}}(\hat{T}) = \hat{C}_0 + \hat{\gamma}(\hat{T} - \hat{T}_0)$, and if we assume chemical equilibrium at the boundaries we obtain $\hat{\gamma} = (\hat{C}_1 - \hat{C}_0) / (\hat{T}_1 - \hat{T}_0)$. The coefficient $\hat{\alpha}_T$ may be expected to be negative, while $\hat{\alpha}_C$ is positive. The static stability of the system is then determined by the signs of $\hat{\alpha}_T(\hat{T}_1 - \hat{T}_0)$ and $\hat{\alpha}_C(\hat{C}_1 - \hat{C}_0)$: if either of these is positive, it represents a destabilising contribution, while if it is negative, it represents a stabilising

contribution. Meanwhile, the coefficient $\hat{\gamma}$ may be positive or negative: if $\hat{\gamma} > 0$, the solubility increases with temperature (a *prograde* mineral such as silica), while if $\hat{\gamma} < 0$, the solubility decreases with temperature (a *retrograde* mineral such as anhydrite: see Jupp & Woods (2003) and references therein). Because we assume chemical equilibrium at the boundaries, we cannot choose the stability properties independently from the pro- or retrograde nature of the mineral: for a prograde mineral, the gradients of concentration and temperature must be the same sign (so they contribute in opposite senses to stability), while for a retrograde mineral, the gradients of concentration and temperature must be of opposite sign, and so be either both stabilising or both destabilising.

We seek an initial steady state in which $\hat{\mathbf{u}} = \mathbf{0}$ and there is no lateral variation. We then find a linear distribution of temperature and thus of the equilibrium solute concentration,

$$\hat{T}_b(\hat{z}) = \hat{T}_0 + (\hat{T}_1 - \hat{T}_0) \frac{\hat{z}}{\hat{H}} \quad \text{and} \quad \hat{C}_b(\hat{z}) = \hat{C}_0 + (\hat{C}_1 - \hat{C}_0) \frac{\hat{z}}{\hat{H}}. \quad (2.5)$$

The initial distribution of solute is simply $\hat{C}_b = \hat{C}_{\text{eq}}(\hat{T}_b)$. We note that by taking \hat{C}_{eq} to be linear in \hat{T} , we permit the existence of a steady basic state in which the solute is everywhere in chemical equilibrium with the solid matrix, and so the vertical flux of solute is constant in space. Under a non-linear model for \hat{C}_{eq} , we would expect chemical disequilibrium to cause a non-uniform solutal flux, as solute was exchanged with the matrix. We will not consider here how this might affect the stability of the system, but we note it as a topic for further investigation, which might exhibit new dynamical effects (Bdzil & Frisch 1980).

2.1. Streamfunction formulation and non-dimensionalisation

It is helpful to write $\hat{\mathbf{u}}(\hat{x}, \hat{z}, \hat{t})$ in terms of a streamfunction $\hat{\psi}(\hat{x}, \hat{z}, \hat{t})$, so $\hat{u} = -\partial\hat{\psi}/\partial\hat{z}$ and $\hat{w} = \partial\hat{\psi}/\partial\hat{x}$. We then define dimensionless variables by

$$\hat{\mathbf{x}} = \hat{H}\mathbf{x}, \quad \hat{t} = \frac{\hat{H}^2}{\hat{\kappa}_T}t, \quad \hat{\psi} = \phi\hat{\kappa}_T\psi, \\ \hat{T} = \hat{T}_b(\hat{z}) + (\hat{T}_1 - \hat{T}_0)T' \quad \text{and} \quad \hat{C} = \hat{C}_b(\hat{z}) + (\hat{C}_1 - \hat{C}_0)C'; \quad (2.6)$$

we can then eliminate \hat{p} to obtain

$$\nabla^2\psi = -\mathcal{R}_C \frac{\partial C'}{\partial x} - \mathcal{R}_T \frac{\partial T'}{\partial x}, \quad (2.7)$$

$$\frac{\partial T'}{\partial t} + \lambda \frac{\partial \psi}{\partial x} + \lambda \left(\frac{\partial \psi}{\partial x} \frac{\partial T'}{\partial z} - \frac{\partial \psi}{\partial z} \frac{\partial T'}{\partial x} \right) = \nabla^2 T', \quad (2.8)$$

$$\frac{\partial C'}{\partial t} + \frac{\partial \psi}{\partial x} + \left(\frac{\partial \psi}{\partial x} \frac{\partial C'}{\partial z} - \frac{\partial \psi}{\partial z} \frac{\partial C'}{\partial x} \right) = \frac{1}{Le} \nabla^2 C' + k(T' - C'), \quad (2.9)$$

where the remaining parameters are

$$\mathcal{R}_C = \frac{\hat{\alpha}_C(\hat{C}_1 - \hat{C}_0)\hat{g}\hat{H}\hat{K}}{\phi\hat{\mu}\hat{\kappa}_T}, \quad \mathcal{R}_T = \frac{\hat{\alpha}_T(\hat{T}_1 - \hat{T}_0)\hat{g}\hat{H}\hat{K}}{\phi\hat{\mu}\hat{\kappa}_T}, \\ \lambda = \frac{\phi(\hat{\rho}\hat{c})_f}{\phi(\hat{\rho}\hat{c})_f + (1-\phi)(\hat{\rho}\hat{c})_s}, \quad Le = \frac{\hat{\kappa}_T}{\hat{\kappa}_C} \quad \text{and} \quad k = \frac{\hat{k}\hat{H}^2}{\phi\hat{\kappa}_T}. \quad (2.10)$$

The parameters \mathcal{R}_T and \mathcal{R}_C are related to the Rayleigh numbers which emerge naturally from the pure thermal and pure solutal problems. We define

$$Ra_T \equiv \frac{\hat{\alpha}_T(\hat{T}_1 - \hat{T}_0)(\hat{\rho}\hat{c})_f \hat{g} \hat{H} \hat{K}}{\hat{\mu}(\hat{\rho}\hat{c})_m \hat{\kappa}_T} = \lambda \mathcal{R}_T \quad \text{and} \quad Ra_C \equiv \frac{\hat{\alpha}_C(\hat{C}_1 - \hat{C}_0) \hat{g} \hat{H} \hat{K}}{\phi \hat{\mu} \hat{\kappa}_C} = Le \mathcal{R}_C. \quad (2.11)$$

We define \mathcal{R}_C in this way so that we may later take the limit $1/Le \rightarrow 0$ while keeping \mathcal{R}_C finite. We recall that either of \mathcal{R}_T and \mathcal{R}_C may be positive or negative, with a positive value encouraging instability. Meanwhile, the differential heat transport rate λ clearly satisfies $0 < \lambda \leq 1$; the Lewis number Le is known to be greater than unity (frequently much greater than unity); and the dimensionless reaction rate (Damköhler number) $k \geq 0$.

Finally, the non-dimensional boundary conditions are

$$\frac{\partial \psi}{\partial x} = 0, \quad T' = 0 \quad \text{and} \quad C' = 0 \quad \text{at} \quad z = 0 \quad \text{and at} \quad z = 1. \quad (2.12)$$

The governing equations derived above differ from those of Steinberg & Brand (1983, 1984) by including solutal diffusivity and by omitting any thermal contributions from the reaction. As explained above, we expect the latter to be small compared to the geothermal gradient, while we will examine the importance of solutal diffusion in §3.6.

3. Linear stability analysis

3.1. The linearised perturbation model

When we assume that the magnitudes of the perturbations to the base state are small, we may simplify the equations (2.7)–(2.9) further to

$$\nabla^2 \psi = -\mathcal{R}_C \frac{\partial C'}{\partial x} - \mathcal{R}_T \frac{\partial T'}{\partial x}, \quad \frac{\partial T'}{\partial t} + \lambda \frac{\partial \psi}{\partial x} = \nabla^2 T' \quad (3.1)$$

$$\text{and} \quad \frac{\partial C'}{\partial t} + \frac{\partial \psi}{\partial x} = \frac{1}{Le} \nabla^2 C' + k(T' - C'). \quad (3.2)$$

The boundary conditions are still given by (2.12).

We seek Fourier-mode solutions of the form

$$\psi = \Psi_0 e^{\sigma t} e^{imx} \sin(\pi n z), \quad T' = \Theta_0 e^{\sigma t} e^{imx} \sin(\pi n z), \quad C' = \chi_0 e^{\sigma t} e^{imx} \sin(\pi n z), \quad (3.3)$$

where the real parts are assumed, where m is a positive real number, where n is a positive integer, and where Ψ_0 , Θ_0 , χ_0 and σ are in general complex. Substituting these in to (3.1) and (3.2), we obtain the equations

$$-(m^2 + \pi^2 n^2) \Psi_0 = -im \mathcal{R}_C \chi_0 - im \mathcal{R}_T \Theta_0, \quad (3.4)$$

$$\sigma \Theta_0 + im \lambda \Psi_0 = -(m^2 + \pi^2 n^2) \Theta_0 \quad \text{and} \quad (3.5)$$

$$\sigma \chi_0 + im \Psi_0 = -\frac{m^2 + \pi^2 n^2}{Le} \chi_0 + k \Theta_0 - k \chi_0. \quad (3.6)$$

The solvability condition for this set of equations can be written as a quadratic, $a\sigma^2 + b\sigma + c = 0$, and thus

$$\sigma = \sigma_{\pm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad (3.7)$$

where, defining $M = m^2 > 0$ and $N = \pi^2 n^2 \geq \pi^2$ for convenience,

$$\begin{aligned} a &= M + N, & b &= \left[1 + \frac{1}{Le}\right] (M + N)^2 + k(M + N) - M(\mathcal{R}_C + \lambda\mathcal{R}_T) \quad \text{and} \\ c &= k(M + N)^2 - kM\lambda(\mathcal{R}_T + \mathcal{R}_C) - M(M + N) \left(\mathcal{R}_C + \frac{\lambda\mathcal{R}_T}{Le}\right) + \frac{1}{Le}(M + N)^3. \end{aligned} \quad (3.8)$$

For very large M or N , we find that $\sigma_+ \sim -(M, N)/Le$; hence, for given parameter values, $\Re(\sigma_+)$ will be maximised for some finite values of M and N .

The stability boundaries of the system correspond to the condition $\max_{M,N} \Re(\sigma_+) = 0$: much of this section will be concerned with identifying these boundaries, and in particular with drawing them in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane for various values of the other parameters. Before doing this, however, it is useful to recall the classic results for non-reactive convection.

3.2. Pure double-diffusive convection, $k = 0$

When we take $k = 0$, so there is no reaction, we recover classic double-diffusive convection (see e.g. Phillips 1991, §5.3). The growth rate σ_+ is given by equations (3.7) and (3.8) with $k = 0$; instability will occur whenever we can find $M > 0$ and $N/\pi^2 \in \mathbb{N}$ such that $\Re(\sigma_+) > 0$. This, in turn, will occur whenever either (i) $b < 0$ or (ii) $b > 0$ and $c < 0$ (so that $\sqrt{b^2 - 4ac} > b$). We will consider these in turn.

(i) The condition $b < 0$ can be written as

$$\mathcal{R}_T > \left[\frac{1 + Le}{\lambda Le}\right] \frac{(M + N)^2}{M} - \frac{\mathcal{R}_C}{\lambda}. \quad (3.9)$$

For a given set of parameters $(\lambda, Le, \mathcal{R}_C)$, the lowest value of \mathcal{R}_T for which this condition is satisfied occurs when $N = \pi^2$ and $M = \pi^2$; the condition then becomes

$$\mathcal{R}_T > \mathcal{R}_T^{(b)}(\mathcal{R}_C) = \frac{4\pi^2}{\lambda} \left[1 + \frac{1}{Le}\right] - \frac{\mathcal{R}_C}{\lambda}. \quad (3.10)$$

(ii) The condition $c < 0$ can be written as

$$\mathcal{R}_T > \frac{1}{\lambda} \frac{(M + N)^2}{M} - \frac{Le\mathcal{R}_C}{\lambda}. \quad (3.11)$$

Again, for a given $(\lambda, Le, \mathcal{R}_C)$, the lowest value of \mathcal{R}_T for which this condition is satisfied occurs when $N = \pi^2$ and $M = \pi^2$, and the condition then becomes

$$\mathcal{R}_T > \mathcal{R}_T^{(c)}(\mathcal{R}_C) = \frac{4\pi^2}{\lambda} - \frac{Le\mathcal{R}_C}{\lambda}. \quad (3.12)$$

with the additional requirement that the condition (3.10) is not satisfied.

In summary, then, instability will occur if either (3.10) or (3.12) is satisfied, and the system will be stable if neither is satisfied. Note that in general, the principle of exchange of stabilities holds when crossing the stability boundary $\mathcal{R}_T = \mathcal{R}_T^{(c)}$, so the loss of stability is through an exponentially growing perturbation (monotonic instability); crossing the boundary $\mathcal{R}_T = \mathcal{R}_T^{(b)}$, on the other hand, in general $b^2 - 4ac < 0$, and so the instability is oscillatory (cf. Turner 1979, §8.1).

The intersection of $\mathcal{R}_T^{(c)}$ with the \mathcal{R}_T -axis occurs at $\mathcal{R}_T = 4\pi^2/\lambda < \mathcal{R}_T^{(b)}(0)$, and so it is $\mathcal{R}_T^{(c)}$ which determines the well-known stability criterion $Ra_T < 4\pi^2$ for pure thermal convection (Lapwood 1948). Similarly, the stability criterion for pure solutal convection

is given by $\mathcal{R}_C < 4\pi^2/Le$, or $Ra_C < 4\pi^2$. The lines $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ and $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ intersect at

$$\mathcal{R}_C^{\text{crit}} = -\frac{4\pi^2}{Le(Le-1)}; \quad (3.13)$$

for $\mathcal{R}_C < \mathcal{R}_C^{\text{crit}}$, the criterion $\mathcal{R}_T < \mathcal{R}_T^{(b)}$ gives stability, while for $\mathcal{R}_C > \mathcal{R}_C^{\text{crit}}$, the criterion $\mathcal{R}_T < \mathcal{R}_T^{(c)}$ gives stability. An illustration of the stability boundaries will be given in figures 2 a and b.

3.3. Reactive cases, $k > 0$: general features of the stability diagram

We now consider the full double-diffusive, double-advective, reactive problem. The presence of the extra parameter k makes this somewhat more complicated than pure double-diffusive convection; our aim will be to determine as broadly as possible how the results for pure DDC alter as k increases. As in the case of pure double-diffusive convection, we will have to determine two boundaries in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane: the boundary $\mathcal{R}_T = \mathcal{R}_T^{(b)}(\mathcal{R}_C)$ above which there exist (M, N) such that $b < 0$, and the boundary $\mathcal{R}_T = \mathcal{R}_T^{(c)}(\mathcal{R}_C)$ above which there exist (M, N) such that $c < 0$. The unstable region is then the union of the regions $\mathcal{R}_T > \mathcal{R}_T^{(b)}$ and $\mathcal{R}_T > \mathcal{R}_T^{(c)}$.

(i) The condition $b < 0$ corresponds to

$$\mathcal{R}_T > f(M, N) = \frac{1}{\lambda} \left(\left[1 + \frac{1}{Le} \right] \frac{(M+N)^2}{M} + k \frac{(M+N)}{M} - \mathcal{R}_C \right). \quad (3.14)$$

This is clearly least restrictive when N takes its minimum value of $N = \pi^2$, and given this it will be least restrictive when we choose M such that $f(M, \pi^2)$ is minimised over M . The condition $\partial f / \partial M = 0$ has the unique solution for positive M

$$M = M_b = \pi \left[\frac{kLe + \pi^2(1+Le)}{1+Le} \right]^{1/2}, \quad (3.15)$$

and we obtain the instability condition

$$\mathcal{R}_T > \mathcal{R}_T^{(b)}(\mathcal{R}_C) = \frac{\pi^2}{\lambda} \left(1 + \frac{1}{Le} \right) \left(1 + \sqrt{1 + \frac{kLe}{\pi^2(1+Le)}} \right)^2 - \frac{\mathcal{R}_C}{\lambda}. \quad (3.16)$$

We note that M_b and $\mathcal{R}_T^{(b)}$ are increasing functions of k , and so the effect of reaction is to increase the threshold of instability given by this condition, and to increase the wavenumber at the threshold of instability. To interpret this variation physically, we recall that $\mathcal{R}_T^{(b)}$ represents a boundary across which the rest state loses stability through an oscillatory instability, $\Im(\sigma_+) \neq 0$. Oscillatory instabilities require hysteresis between the temperature and the concentration fields: this hysteresis is reduced by the coupling between T' and C' introduced by the reaction, inhibiting the onset of instability. However, this interpretation does not explain the increase in M_b with k , for which we are unable to identify a simple physical mechanism.

(ii) The other boundary is harder to locate. The condition $c < 0$ corresponds to

$$\mathcal{R}_T > \mathcal{R}_T^{(c)}(\mathcal{R}_C) = \min_{M,N} R(M, N), \quad \text{where}$$

$$R(M, N) \equiv \frac{\frac{(M+N)^2}{M\lambda} + \frac{1}{Le} \frac{(M+N)^3}{kM\lambda} - \mathcal{R}_C \left[1 + \frac{(M+N)}{k\lambda} \right]}{1 + \frac{(M+N)}{kLe}}. \quad (3.17)$$

(We will write $\mathcal{R}_T = R(M, N)$ to indicate the boundary $c = 0$ for a given mode, while $\mathcal{R}_T^{(c)}(\mathcal{R}_C)$ will represent the minimum over (M, N) of R , i.e. the stability boundary if M and N are not imposed.)

In appendix A.1, we show that the global minimum of $R(M, N)$ over M and N always occurs on the boundary $N = \pi^2$, and corresponds to a local minimum over M : for the rest of this section we will use the notation $M = M_c$ to denote this minimum and take $N = \pi^2$ throughout. The condition $\partial R / \partial M = 0$ then leads to a quartic equation for M_c ,

$$\frac{M_c^4}{Le^2} + \left(\frac{2k}{Le} + \frac{2\pi^2}{Le^2} \right) M_c^3 + \left(k^2 - \mathcal{R}_C k + \frac{2k\pi^2 + \mathcal{R}_C k\lambda}{Le} \right) M_c^2 + \left(-\frac{2k\pi^4}{Le} - \frac{2\pi^6}{Le^2} \right) M_c - k^2\pi^4 - \frac{2k\pi^6}{Le} - \frac{\pi^8}{Le^2} = 0, \quad (3.18)$$

and we will have $\mathcal{R}_T^{(c)}(\mathcal{R}_C) = R(M_c, \pi^2)$.

We can in principle obtain an analytical expression for the positive real root of (3.18), and thus for the stability boundary in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane. However, these expressions are not informative, so before we carry out these calculations we will obtain some more results to gain insight into the general shape of the stability boundaries. In particular, to draw these boundaries in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane, we need to locate the axis-intercepts and the limiting asymptotic behaviour of $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(c)}$.

3.3.1. Axis-intercepts

We first consider the intercepts at $\mathcal{R}_C = 0$: here there is no direct solutal contribution to buoyancy, so we have pure thermal convection. When $\mathcal{R}_C = 0$, equation (3.18) has the unique positive solution $M_c = \pi^2$. Making this substitution in $\mathcal{R}_T^{(c)}(M_c, \pi^2)$ from equation (3.17), we deduce that, for all cases no matter what the relative magnitudes of the parameters, instability will occur for $\mathcal{R}_C = 0$ if $\mathcal{R}_T > \mathcal{R}_T^{(c)}(0) = 4\pi^2/\lambda$. (This result can be obtained immediately by arguing that if $\mathcal{R}_C = 0$ then the quantities k and Le which affect only the solutal component must be irrelevant, and the result must agree with that for pure thermal convection.) We note from (3.16) that the axis-intercept of $\mathcal{R}_T^{(b)}$ is given by

$$\mathcal{R}_T^{(b)}(0) = \frac{\pi^2}{\lambda} \frac{(1+Le)}{Le} \left(1 + \sqrt{1 + \frac{kLe}{\pi^2(1+Le)}} \right)^2 > \mathcal{R}_T^{(c)}(0), \quad (3.19)$$

so that the boundary $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ always lies below the boundary $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ at $\mathcal{R}_C = 0$.

The other special case which we can consider occurs when the direct thermal contribution to buoyancy is zero, $\mathcal{R}_T = 0$, but the reactively-affected solute concentration does affect the buoyancy. As before, instability will occur if there exists some $M > 0$ such that either (i) $b < 0$ or (ii) $b > 0$ and $c < 0$.

(i) We already have the full solution (3.16) for the boundary $\mathcal{R}_T = \mathcal{R}_T^{(b)}$. Requiring

that $\mathcal{R}_T^{(b)} = 0$ yields

$$\mathcal{R}_C = \mathcal{R}_C^{(b)} = \pi^2 \left(1 + \frac{1}{Le}\right) \left(1 + \sqrt{1 + \frac{kLe}{\pi^2(1+Le)}}\right)^2. \quad (3.20)$$

(ii) The condition $c < 0$ for $N = \pi$ and general M becomes

$$\mathcal{R}_C > S(M) \equiv \frac{k(M + \pi^2)^2 + \frac{1}{Le}(M + \pi^2)^3}{M(k\lambda + M + \pi^2)}. \quad (3.21)$$

We have been unable to obtain a simple expression for the positive real root of $dS/dM = 0$; however, we note that

$$S(M) \leq \frac{k(M + \pi^2)}{M} + \frac{1}{Le} \frac{(M + \pi^2)^2}{M} < S_1(M) \equiv \left[1 + \frac{1}{Le}\right] \frac{(M + \pi^2)^2}{M} + \frac{k(M + \pi^2)}{M}, \quad (3.22)$$

and the quantity $S_1(M)$ represents (from equation (3.14)) the value of \mathcal{R}_C for which $c = 0$ when $\mathcal{R}_T = 0$ and $N = \pi$. This in turn means that $\min_M S(M) < \min_M S_1(M)$, so the curve $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ always intercepts the \mathcal{R}_C -axis to the left of $\mathcal{R}_T = \mathcal{R}_T^{(b)}$. Together with the result that $\mathcal{R}_T^{(c)}(0) < \mathcal{R}_T^{(b)}(0)$, this suggests (although it does not prove) that for $\mathcal{R}_C > 0$ the stability boundary when $k > 0$ is still given by the criterion $c = 0$ rather than by $b = 0$. (We will see that this is confirmed by direct evaluation of the stability boundaries in §3.5.)

3.3.2. Asymptotic behaviour and regimes of relevance of the stability boundaries

For pure DDC, as we have seen, $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ provides the stability boundary for sufficiently large negative values of \mathcal{R}_C , while $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ provides the stability boundary for small negative and for positive values of \mathcal{R}_C . To investigate how this picture changes as k increases, we will now examine the limits $\mathcal{R}_C \rightarrow \pm\infty$.

From the explicit solution (3.16), we obtain $\mathcal{R}_T^{(b)} \sim -\mathcal{R}_C/\lambda$ as $|\mathcal{R}_C| \rightarrow \infty$. To establish the behaviour of $\mathcal{R}_T^{(c)}$ we start from (3.18). We assume that, whatever the values of k and $1/Le$, we can take $|\mathcal{R}_C|$ sufficiently large that only the magnitude of $|\mathcal{R}_C|$ is relevant; we then need to look for the possible asymptotic balances in (3.18).

As $|\mathcal{R}_C|$ becomes large, the leading-order balance in (3.18) must be between the M_c^2 term and either the M_c^4 term or the $\mathcal{O}(1)$ term. The former case gives

$$M_c = \sqrt{\mathcal{R}_C k Le} \left(1 - \frac{\lambda}{Le}\right)^{1/2} - (kLe + \pi^2) + \mathcal{O}(\mathcal{R}_C^{-1/2}), \quad (3.23)$$

and clearly can be valid only for $\mathcal{R}_C > 0$; we then have

$$\mathcal{R}_T^{(c)} = -\frac{Le}{\lambda} \mathcal{R}_C + \frac{2\sqrt{kLe(Le - \lambda)}}{\lambda} \sqrt{\mathcal{R}_C} + \frac{\pi^2 - kLe}{\lambda} + \mathcal{O}(\mathcal{R}_C^{-1/2}). \quad (3.24)$$

We note that the wavenumber only affects the result at $\mathcal{O}(1)$, so even though the value of k controls at leading order the selected wavenumber M_c (cf. §3.2), it does not affect the leading-order term in the position of the stability boundary.

The other possible balance gives us

$$M_c = \frac{1}{\sqrt{-\mathcal{R}_C}} \left[\frac{\pi^2(\pi^2 + Lek)}{\sqrt{kLe^2 - k\lambda Le}} + \frac{\pi^4(\pi^2 + Lek)}{Le k(Le - \lambda)} \frac{1}{\sqrt{-\mathcal{R}_C}} + \mathcal{O}\left(\frac{1}{|\mathcal{R}_C|}\right) \right], \quad (3.25)$$

which is clearly applicable only for $\mathcal{R}_C < 0$. Expanding $\mathcal{R}_T^{(c)}$ for $-\mathcal{R}_C \gg 1$, we find

$$\mathcal{R}_T^{(c)} = -\frac{\lambda k + \pi^2}{Le k + \pi^2} \frac{Le}{\lambda} \mathcal{R}_C + \frac{2\pi^2 Le k (Le - \lambda) \sqrt{-\mathcal{R}_C}}{\lambda(\pi^2 + Le k) \sqrt{k Le (Le - \lambda)}} + \frac{\pi^2(\pi^2 + 2k Le)}{\lambda(\pi^2 + k Le)} + \mathcal{O}\left(\frac{1}{\sqrt{-\mathcal{R}_C}}\right). \quad (3.26)$$

This is an obvious modification of the pure DDC result, with the stability boundary lowering systematically as k increases from zero.

Putting these results together, we see that as $\mathcal{R}_C \rightarrow \infty$, the curve $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ always lies below the curve $\mathcal{R}_T = \mathcal{R}_T^{(b)}$, and so always provides the stability boundary. In the other limit $\mathcal{R}_C \rightarrow -\infty$, for sufficiently small k the lower boundary is given by $\mathcal{R}_T = \mathcal{R}_T^{(b)}$, but as k increases the curve $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ drops below $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ and so becomes the stability boundary. The condition that $\mathcal{R}_T^{(c)} < \mathcal{R}_T^{(b)}$ for large enough $-\mathcal{R}_C$ is equivalent to

$$k > k_{\text{crit}} = \frac{\pi^2}{1 - \lambda} \left(1 - \frac{1}{Le}\right). \quad (3.27)$$

We note that although the condition $k > k_{\text{crit}}$ does not itself guarantee that $\mathcal{R}_T^{(c)}$ lies below $\mathcal{R}_T^{(b)}$ for all \mathcal{R}_C , numerical investigation suggests that $\mathcal{R}_T^{(c)} < \mathcal{R}_T^{(b)}$ for all \mathcal{R}_C once k is even very slightly above this boundary, so we may take $k \approx k_{\text{crit}}$ as marking approximately the reaction rate above which $\mathcal{R}_T^{(b)}$ becomes irrelevant.

3.4. Asymptotics for slow or fast reactions

We have established by looking at the asymptotic behaviour of the stability boundaries that varying k may change the shape of the stability diagram, in particular for $\mathcal{R}_C < 0$ (the solutally stabilised ‘SS’ regime). To gain some more insight into the effect of varying k , we now consider how either a very small or a very large reaction term affects the picture for pure DDC which was sketched in §3.2.

3.4.1. Asymptotics for slow reactions, $0 < k \ll 1$

We will first consider slow reactions. Formally, we will consider the asymptotic regime $0 < k \ll 1$ and $Le = \mathcal{O}(1)$, so reaction is a weaker influence on the solute concentration than diffusion is. As before, we consider the two criteria $b < 0$ and $c < 0$ separately. We have already dealt with the case $b < 0$ (equation (3.16)), and it is evident that the effect of a weak reaction will be to raise this boundary slightly. We therefore concentrate on the case $c < 0$.

As before, we start from equation (3.18). (For sufficiently small k , this will be the relevant boundary for positive or sufficiently small negative values of \mathcal{R}_C , as in §3.2.) Seeking solutions of the form $M_c = M_0 + M_1 k + \mathcal{O}(k^2)$, we find

$$M_c = \pi^2 + \frac{\mathcal{R}_C Le (Le - \lambda)}{8\pi^2} k + \mathcal{O}(k^2), \quad (3.28)$$

and so

$$\mathcal{R}_T^{(c)} = \frac{4\pi^2}{\lambda} - \frac{\mathcal{R}_C Le}{\lambda} + \frac{Le (Le - \lambda) \mathcal{R}_C}{2\pi^2 \lambda} k + \mathcal{O}(k^2). \quad (3.29)$$

The effect of a small reaction term is to stabilise the system against monotonic instabilities for $\mathcal{R}_C > 0$, as destabilising solute is precipitated from the fluid, and to destabilise it against monotonic instabilities for $\mathcal{R}_C < 0$, as stabilising solute is lost.

3.4.2. Asymptotics for fast reactions, $k \gg 1$

It is also natural to consider the limit in which the reaction is very fast compared to the other terms in the system. Formally, we consider the regime $k \gg 1$ with all other parameters of order unity (compare the ‘fast reaction’ analysis of Steinberg & Brand (1983), which assumes that $\hat{k}\hat{H}^2/\hat{\kappa}_C = kLe \gg 1$).

It is simple to expand the expression (3.16) to obtain $\mathcal{R}_T^{(b)} \sim k/\lambda$. To obtain asymptotic results for M_c and $\mathcal{R}_T^{(c)}$ we consider equation (3.18) in the limit $k \rightarrow \infty$. An asymptotic analysis indicates that the only positive root for M_c is of order unity, and has the expansion

$$M_c \sim \pi^2 + \frac{\pi^2(Le - \lambda)\mathcal{R}_C}{2Le} \frac{1}{k} + \mathcal{O}\left(\frac{1}{k^2}\right), \quad (3.30)$$

and thus

$$\mathcal{R}_T^{(c)} \sim \frac{4\pi^2}{\lambda} - \mathcal{R}_C - \frac{2\pi^2(Le - \lambda)\mathcal{R}_C}{\lambda Le} \frac{1}{k} + \mathcal{O}\left(\frac{1}{k^2}\right). \quad (3.31)$$

It is evident from these results that in the regime $k \gg 1$, the boundary $\mathcal{R}_T^{(c)}$ must lie below the boundary $\mathcal{R}_T^{(b)}$, so as the reaction rate is increased with other conditions constant, the stability boundary always comes to be given by $\mathcal{R}_T = \mathcal{R}_T^{(c)}$.

Two points are worth noting. The first is that in this limit, the stability boundary is independent at leading order of both k and Le . Essentially, the solute concentration equilibrates instantaneously with the temperature field, so the condition for instability is simply that the aggregate contribution to buoyancy is sufficient to overcome stabilisation by thermal diffusion, $\mathcal{R}_T + \mathcal{R}_C > 4\pi^2/\lambda$. The next order correction depends on solutal diffusion, and its sign is controlled by \mathcal{R}_C . For $\mathcal{R}_C > 0$, the effect of making the reaction rate finite rather than infinite is to lower the boundary, i.e. to destabilise the system, as destabilising solute is allowed to persist a little longer and so to drive fingering. Conversely, for $\mathcal{R}_C < 0$, a finite reaction rate tends to raise the boundary (stabilise the system).

We also note that as $k \rightarrow \infty$, M_c approaches its limiting value of π^2 , with a correction term which is controlled by \mathcal{R}_C . We may expect that the stabilising or destabilising effect of the ‘extra’ solute which a finite reaction rate permits will be felt most strongly at higher wavenumber, as in other fingering instabilities when diffusive effects are neglected (cf., e.g. Drazin 2002, §3.8). If the effect is stabilising, then, it will tend to lower the value of $\sigma_+(M)$ for higher M , and so shift the maximum to lower values of M ; conversely, if the effect is destabilising, it will tend to shift the maximum to higher values of M .

3.5. Examples of stability diagrams and dispersion relations

We are now in a position to examine in some detail how the stability boundaries vary with the governing parameters, and in particular how the reaction rate k controls them. We will first show some examples of the stability boundaries, then discuss how they fit into the more general variation of the maximal growth rate σ_+ .

3.5.1. Variation of stability boundaries with k

Figures 1 a and b illustrate how the positions of the stability boundaries $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(c)}$ vary with k and with \mathcal{R}_C , while figures 1 c and d show the corresponding variation of the critical wavenumber M at the onset of instability.

We first consider how the stability boundaries change with k and \mathcal{R}_C . As figure 1 a illustrates, both $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(c)}$ increase monotonically as \mathcal{R}_C decreases: a stronger thermal instability is required to compensate for a weaker solutal instability. The boundary $\mathcal{R}_T^{(b)}$

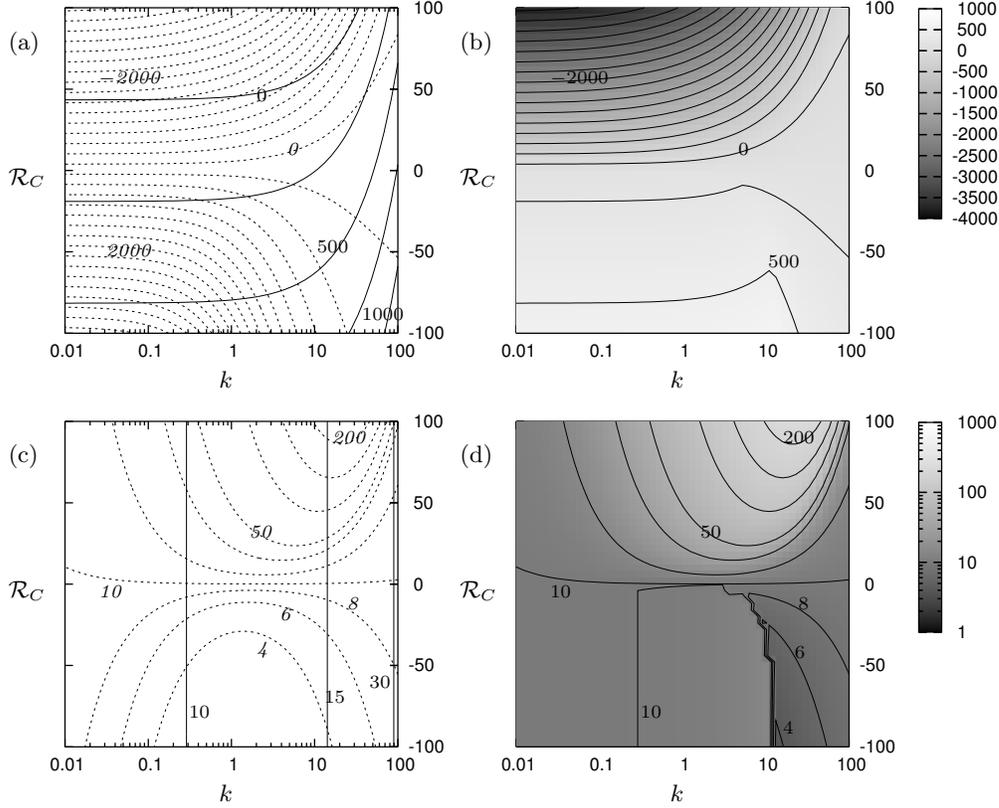


FIGURE 1. Contour plots showing stability boundaries and wavenumbers of marginally unstable perturbations, plotted as a function of k and \mathcal{R}_C , for $\lambda = 0.25$ and $Le = 10$. (a) Stability boundaries $\mathcal{R}_T^{(b)}$ (solid; roman labels) and $\mathcal{R}_T^{(c)}$ (dashed; italic labels): contours of $\mathcal{R}_T^{(b)}$ are at 0 to 1000 by 250, while contours of $\mathcal{R}_T^{(c)}$ are from -3750 to 4000 by 250. (b) Position of the stability boundary $\mathcal{R}_T^{(\min)} = \min(\mathcal{R}_T^{(b)}, \mathcal{R}_T^{(c)})$: contours are at intervals of 250 with the highest at 500. (c) Wavenumbers M_b (solid; roman labels) and M_c (dashed; italic labels) corresponding to the conditions $b = 0$ and $c = 0$ respectively: contours of M_b are at 10, 15 and 30, while contours of M_c are at 4, 6, 8, 10, 15, 30, 50, 100, 150 and 200. (d) Wavenumber M_{\min} of the most unstable mode: contour values as in (c). The blocky effect on contours near the discontinuity in M_{\min} is an artefact of the contour interpolation routine.

also increases with k , whereas the boundary $\mathcal{R}_T^{(c)}$ merely varies less strongly with \mathcal{R}_C for higher values of k : this shows the stabilising effect of the reaction in the SU regime $\mathcal{R}_C > 0$ and a destabilising effect in the SS regime $\mathcal{R}_C < 0$. Combining the results for $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(c)}$ to obtain the actual stability boundary $\mathcal{R}_T^{(\min)}$, we obtain the more complicated variation shown in figure 1 b. In the SU regime $\mathcal{R}_C > 0$, the stability is entirely controlled by $\mathcal{R}_T^{(c)}$ and the variation of $\mathcal{R}_T^{(\min)}$ with k remains monotonic. In the SS regime $\mathcal{R}_C < 0$, for low values of k the boundary is given by $\mathcal{R}_T^{(b)}$, while only for higher values is it given by $\mathcal{R}_T^{(c)}$. The result is that the boundary $\mathcal{R}_T^{(\min)}$ varies non-monotonically with k , first increasing (so, as we have seen in §3.4.1, a weak reaction stabilises the system) and then decreasing, so a sufficiently strong reaction removes stabilising solute from the fluid and destabilises the system.

The variation of the critical wavenumber at the onset of instability is more complicated.

We first consider figure 1 c. The wavenumber M_b which minimises $\mathcal{R}_T^{(b)}$ is independent of \mathcal{R}_C , and simply increases monotonically with k (see equation (3.15)). The wavenumber M_c which minimises $\mathcal{R}_T^{(c)}$, on the other hand, increases with small but increasing k in the SU regime and decreases with small but increasing k in the SS regime (§3.4.1); for large k (§3.4.2) it approaches π^2 again, and so it must have a local extremum over k for any given \mathcal{R}_C . Combining these results yields the variation of M_{\min} shown in figure 1 d. In the SU regime $\mathcal{R}_C > 0$, the maximum value of M_{\min} occurs for higher k as \mathcal{R}_C increases. Rather large values of M_{\min} can be attained for moderate k and \mathcal{R}_C , indicating that the shape of the cells deviates strongly from the square rolls characteristic of pure DDC, and throughout this regime the wavenumber is greater than π^2 . The SS regime $\mathcal{R}_C < 0$ may be divided into two parts as before. For small k , the boundary is given by $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ and the wavenumber remains close to π^2 . The line where $\mathcal{R}_T^{(b)} = \mathcal{R}_T^{(c)}$ is marked by a discontinuity in the value of M_{\min} (which corresponds to the ‘blocky’ line in the interpolated contour plots). (It is not surprising that M_{\min} is discontinuous across this line, since the mode of instability changes as we cross it: this will be discussed further in §4.) For higher values of k , where $\mathcal{R}_T^{(\min)} = \mathcal{R}_T^{(c)}$ and reactions are stabilising, the critical wavenumber is less than π^2 , though it increases towards π^2 with increasing k (see §3.4.2).

3.5.2. Dispersion relations

The results which we have obtained for the marginal stability curves in the $(\mathcal{R}_C, \mathcal{R}_T)$ plane only tell part of the story: to get a fuller picture of what the dispersion relation (3.7) means, we must evaluate some cases directly. For the moment, we will consider only cases for which $N = \pi^2$, although it is only in the region of the stability boundary that we can assume that this represents the fastest-growing modes. Figure 2 illustrates how the stability boundaries and the growth rate σ_+ and wavenumber M_{\min} of the most unstable mode vary as k is increased from zero: the results presented here were obtained numerically using Maple 9.

In figures 2 a, c and e, we see the evolution of the stability boundaries in the $(\mathcal{R}_C, \mathcal{R}_T)$ plane. As we have seen, as k increases, the unstable region first decreases in size, with $\mathcal{R}_T^{(b)}$ moving up everywhere and $\mathcal{R}_T^{(c)}$ moving upwards and rightwards in the region $\mathcal{R}_C > 0$, though downwards in the region $\mathcal{R}_C < 0$ (see figure 1 a). The result is that the intersection between $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(c)}$ moves to the left, so that $\Im(\sigma_+) = 0$ along a greater part of the stability boundary. As k is increased further, the region $\mathcal{R}_C > 0$ is further stabilised, and as k passes k_{crit} , the descending boundary $\mathcal{R}_T^{(c)}$ for $\mathcal{R}_C < 0$ passes below $\mathcal{R}_T^{(b)}$ entirely: the transition to instability is now always through monotonic rather than oscillatory perturbations, and the unstable region in $\mathcal{R}_C < 0$ grows everywhere as k is increased.

Figure 2 b shows the variation of the growth rate σ_+ and the most unstable wavenumber for pure double-diffusive convection. The variation of $\Re(\sigma_+)$ is simple, increasing with increasing \mathcal{R}_T and \mathcal{R}_C . There is a region in $\mathcal{R}_C < 0$ in which the dominant mode has a non-zero imaginary part: this region actually comes into existence for \mathcal{R}_T slightly smaller than $\mathcal{R}_T^{(b)}$, and for any given \mathcal{R}_C there is an upper limit on the values of \mathcal{R}_T for which $\Im(\sigma_+) \neq 0$. Along the stability boundary, the wavenumber of the marginally stable mode is constant at $M_{\min} = \pi^2$, but it rapidly increases, particularly in the unstable region for which $\mathcal{R}_C > 0$ and $\mathcal{R}_T < 0$.

For $0 < k < k_{\text{crit}}$ (figures 2 c and d), this picture does not change significantly. Larger negative values of \mathcal{R}_C are now required to find significant values of $\Im(\sigma_+)$; numerical experimentation suggests that the imaginary region does shift to the left and diminish in size as k increases, but that it persists while $k < k_{\text{crit}}$. (Overall, the criteria which

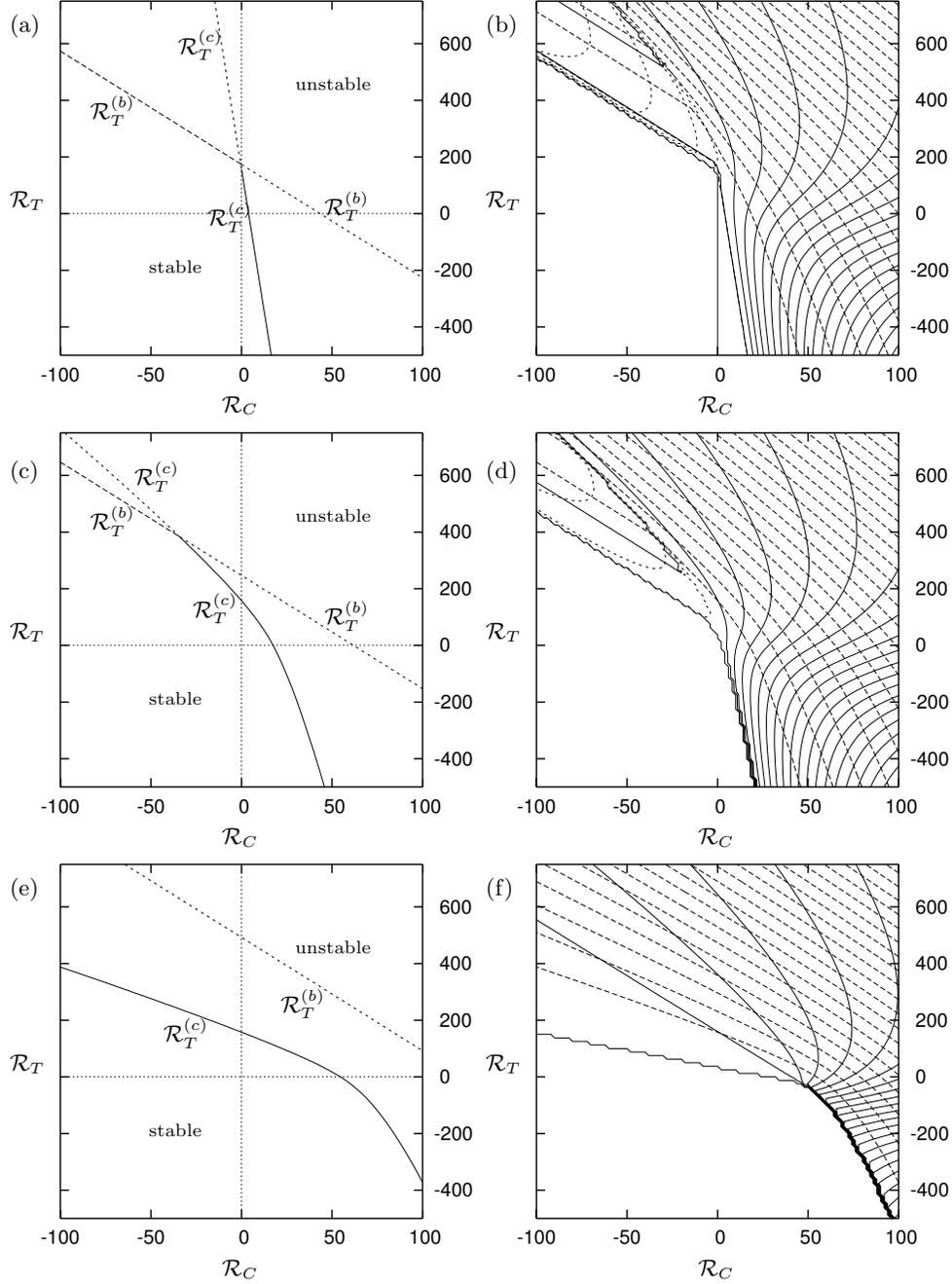


FIGURE 2. (a, c, e) Linear stability boundaries and (b, d, f) growth rates σ_+ in the $(\mathcal{R}_C, \mathcal{R}_T)$ plane for $\lambda = 0.25$, $Le = 10$ and (a,b) $k = 0$; (c, d) $k = 10$; (e, f) $k = 50$ (note that $k_{\text{crit}} \approx 11.84$). In (a), (c) and (e) the solid lines represent stability boundaries on which $\sigma_+ = 0$, the heavy dashed lines represent boundaries on which $\Re(\sigma_+) = 0$ but $\Im(\sigma_+) \neq 0$, and the light dashed lines represent the continuation of $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(c)}$ above the boundary. Plots (b), (d) and (f) show contours of $\Re(\sigma_+)$ (heavy dashed), $\Im(\sigma_+)$ (light dashed) and M_{\min} (solid): the contours in each case are at intervals of 10 with the lowest being at 10^{-5} . The blocky effect on some contours is an artefact of the contour interpolation routine.

describe the boundaries do not give much insight into the behaviour in the rest of the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane.) The main difference from the case $k = 0$ is that, as we have seen in figure 1 d, M_{\min} now varies along the stability boundary, increasing with increasing \mathcal{R}_C .

For still larger values of k , more significant changes occur. In figures 2 e and f, there is now no evidence at all of a region in which $\Im(\sigma_+) \neq 0$. The behaviour of the fastest-growing wavenumber M_{\min} has also changed, and this change appears to happen at around $k = k_{\text{crit}}$: instead of continuous variation of M_{\min} with \mathcal{R}_C and \mathcal{R}_T , there is now a sharp discontinuity across which M_{\min} jumps from zero to a value which increases rapidly with \mathcal{R}_C . Along the stability boundary, M_{\min} now increases rapidly with \mathcal{R}_C .

3.6. Asymptotics for weak solutal diffusion, $1/Le \ll 1$

For plotting purposes, and to gain insight into the effect of a reaction term, we have so far assumed implicitly that the Lewis number $Le = \mathcal{O}(1)$, so the solutal diffusivity is not asymptotically small. In geological applications, however, the ratio between the molecular diffusivities of many minerals and the thermal diffusivity in the rock-water system may be of the order of 100 (see e.g. Bear 1972, p. 651). This makes it useful to investigate the asymptotic regime $Le \gg 1$ (we again recall that the analysis of Steinberg & Brand (1983) assumes that $kLe \gg 1$, and on this basis disregards the solutal diffusion term altogether).

We note that in the limit $1/Le = 0$, the results for pure DDC break down in the fingering regime $\mathcal{R}_C > 0$, because there is no longer any diffusion to stabilise the system. We may therefore expect this limit to be singular: knowing that in this regime the reaction term tends to stabilise the system by removing solute from solution, it is natural to ask in particular whether the reaction term on its own can ever provide stability, or whether diffusion remains essential.

We note first that in the absence of reaction ($k = 0$), in the limit $1/Le \rightarrow 0$ the criterion for instability derived in section 3.2 becomes simply $\mathcal{R}_C > 0$ or $\lambda\mathcal{R}_T + \mathcal{R}_C > 4\pi^2$: the SU regime is completely unstable while in the SS regime the system can be stabilised by the non-diffusing solutal buoyancy gradient. In the SS regime, the marginally stable perturbations at the stability boundary have wavenumber $M_{\min} = \pi^2$. We will now consider in more detail the asymptotics of the stability boundaries when k is of order unity; at the end we will make some brief comments on the case when k is also asymptotically small.

As before, we consider separately the conditions (i) $b < 0$ and (ii) $c < 0$. (We recall from §3.3.2 that we expect the boundary to be given by $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ for positive or small negative values of \mathcal{R}_C whatever the value of the other parameters, while for sufficiently large values of k it may be given by $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ for all values of \mathcal{R}_C .)

(i) Starting from the expression for $\mathcal{R}_T^{(b)}$ given by equation (3.16), we expand in powers of $1/Le$ to obtain

$$\mathcal{R}_T^{(b)} = -\frac{\mathcal{R}_C}{\lambda} + \frac{\pi^2}{\lambda} \left(1 + \sqrt{1 + \frac{k}{\pi^2}}\right)^2 \left[1 + \left(1 + \frac{k}{\pi^2}\right)^{-1/2} \frac{1}{Le}\right] + \mathcal{O}\left(\frac{1}{Le^2}\right). \quad (3.32)$$

We can also expand the most unstable wavenumber $M = M_b$ given by equation (3.15), obtaining

$$M_b = \pi \sqrt{\pi^2 + k} - \frac{\pi k}{2\sqrt{\pi^2 + k}} \frac{1}{Le} + \mathcal{O}(Le^{-2}). \quad (3.33)$$

Including diffusion is then a regular perturbation to this stability curve, slightly increasing the threshold for instability and reducing the wavenumber of the most unstable mode.

This is not unexpected, since diffusion will tend to preferentially inhibit the growth of the higher-wavenumber modes.

(ii) The criterion $c < 0$ is more problematic. As we have seen, the condition that $\partial R/\partial M = 0$ at $M = M_c$ provides a quartic in M_c , equation (3.18). When we take $1/Le = 0$, we lose the first two terms in M_c , indicating that this limit is singular. In fact, three different asymptotic balances are possible in this equation for $1/Le \gg 1$. The simplest balance occurs when $M_c = \mathcal{O}(1)$, and (3.18) reduces at leading order to a quadratic in M_c . Another balance occurs when we seek solutions which preserve the higher-order terms in (3.18): this corresponds to very short-wave perturbations damped by lateral diffusion, as the wavenumber M_c ‘comes in’ from infinity with decreasing Le . A final balance occurs when \mathcal{R}_C is very close to k , so the coefficient of M_c^2 in (3.18) is $\mathcal{O}(Le^{-1})$. It can be shown (appendix A.2.1) that each of these balances corresponds to a particular range of \mathcal{R}_C .

When $\mathcal{R}_C < k$ and $|\mathcal{R}_C - k| = \mathcal{O}(1)$ or larger, we find that

$$M_c \sim \pi^2 \left(\frac{k}{k - \mathcal{R}_C} \right)^{1/2} - \frac{1}{Le} \frac{\pi^2 \mathcal{R}_C}{2} \frac{\left(2\pi^2 \sqrt{\frac{k}{k - \mathcal{R}_C}} + 2\pi^2 + k\lambda \right)}{\sqrt{k}(k - \mathcal{R}_C)^{3/2}} + \mathcal{O}\left(\frac{1}{Le^2}\right) \quad (3.34)$$

and

$$\mathcal{R}_T^{(c)} \sim \frac{\pi^2}{\lambda} \left(1 + \sqrt{\frac{k - \mathcal{R}_C}{k}} \right)^2 - \mathcal{R}_C + \frac{\mathcal{R}_T^{(c1)}}{Le} + \mathcal{O}(Le^{-2}), \quad (3.35)$$

where $\mathcal{R}_T^{(c1)}$ is given by equation (A 10). In this regime, then, the system can be stabilised by the combined effects of solutal buoyancy and the reaction term; the marginally stable wavenumber is controlled by the reaction rate. It is particularly noteworthy that in the absence of solutal diffusion the reaction alone can stabilise the system even for some positive values of \mathcal{R}_C .

When $\mathcal{R}_C > k$ and $|\mathcal{R}_C - k| = \mathcal{O}(1)$ or larger, we find that

$$M_c \sim \sqrt{k}(\sqrt{\mathcal{R}_C} - \sqrt{k})Le \quad \text{and} \quad \mathcal{R}_T^{(c)} \sim -\frac{(\sqrt{\mathcal{R}_C} - \sqrt{k})^2 Le}{\lambda}. \quad (3.36)$$

In this regime, stabilisation depends crucially on the presence of solutal diffusion: in the limit $Le \rightarrow \infty$, the critical wavenumber and the (negative) value of \mathcal{R}_T required for stability both become unbounded. In other words, if the solutal buoyancy gradient is sufficiently destabilising then the reaction cannot stabilise the system on its own.

Finally, when $|\mathcal{R}_C - k| = \mathcal{O}(1/Le)$, we find that

$$M_c = \left(\frac{Le k \pi^4}{2} \right)^{1/3} - \frac{\mathcal{R} + k\lambda + 2\pi^2}{3} + \mathcal{O}(Le^{-1/3}) \quad \text{where} \quad \mathcal{R} = Le(k - \mathcal{R}_C), \quad (3.37)$$

and thus

$$\mathcal{R}_T^{(c)} = \frac{\pi^2}{\lambda} - k + \frac{3}{\lambda} \left(\frac{\pi^8}{4k} \right)^{1/3} \frac{1}{Le^{1/3}} + \mathcal{O}(Le^{-2/3}). \quad (3.38)$$

In this regime the system can in principle be stabilised by reaction in the absence of diffusion, but asymptotically high wavenumbers are favoured since $M_c \sim Le^{1/3}$.

Finally, we comment briefly on the regime in which $k = \mathcal{O}(1/Le) \ll 1$, so the reaction is slow while solutal diffusion remains weak. (This may be relevant in geological situations.) The details of the analysis for this case are given in appendix A.2.2; the principal conclusions are that in the SU regime $\mathcal{R}_C > 0$ the weak reaction term is unable to sta-

bilise a solutally unstable system, while in the SS regime $\mathcal{R}_C < 0$ the boundary, given by $\mathcal{R}_T^{(b)}(\mathcal{R}_C)$, is identical to leading order with that for $k = 0$.

Our overall conclusion from this brief examination of the limit $Le \rightarrow \infty$ is that while a reasonably strong reaction term may be able to stabilise a somewhat solutally unstable system, and to control the wavenumber of instabilities in the presence of weak diffusion, it cannot on its own stabilise a system in which the solutal buoyancy gradient is sufficiently strongly unstable.

4. Nonlinear development: a Galerkin model

To explore how a reaction term affects the nonlinear development of double-diffusive convection it is necessary to integrate the full nonlinear system (2.7)–(2.9) numerically: this will be the subject of §5, and will be developed in a future study. Before doing so, however, we will briefly consider the early stages of nonlinear convection, when the basic structure of the convective rolls is still determined by the behaviour of the linearised solution.

In the immediate vicinity of the stability boundary, we could develop a weakly-nonlinear analysis in which the small parameter is no longer the perturbation amplitude but the deviation of one or another Rayleigh number from its critical value (e.g. Palm, Weber & Kvernfold 1972). Such analyses tend to become mathematically complicated, so to avoid these complexities, we will instead employ a simpler but more heuristic Galerkin method. In this method (e.g. Rudraiah, Srimani & Friedrich 1982), we choose a suitable ansatz for the form of the solution and develop a series of ordinary differential equations for the evolution of this structure. Mamou & Vasseur (1999) have employed it to investigate the bifurcation structure at the onset of thermosolutal convection, and demonstrated reasonable agreement between the Galerkin results and their numerical integrations of the full nonlinear equations.

4.1. Galerkin ansatz

Following Mamou & Vasseur (1999), we will assume that close to the threshold of convection, the basic circulation remains undistorted, but the temperature and concentration fields are distorted by the addition of a second harmonic with no x -dependence, so

$$\begin{aligned} \psi(x, z, t) &= \Psi_1(t) \sin(mx) \sin(\pi z), \\ T'(x, z, t) &= \Theta_1(t) \cos(mx) \sin(\pi z) + \Theta_2(t) \sin(2\pi z) \\ \text{and } C'(x, z, t) &= \chi_1(t) \cos(mx) \sin(\pi z) + \chi_2(t) \sin(2\pi z), \end{aligned} \quad (4.1)$$

where the amplitudes Ψ_1 , $\Theta_{1,2}$ and $\chi_{1,2}$ are real. We assume that the horizontal wavenumber m is known, having been selected, for example, by the dynamics of the linear regime. (We will consider briefly below how the region of subcritical convection may be maximised by choosing an appropriate value of m .)

To set up the Galerkin model, we substitute the ansatz (4.1) into (2.7)–(2.9) and equate the Fourier components up to $2\pi z$ inclusive, ignoring the higher-frequency terms.

We obtain the five equations

$$-(m^2 + \pi^2)\Psi_1 - m\mathcal{R}_C\chi_1 - m\mathcal{R}_T\Theta_1 = 0, \quad (4.2)$$

$$\frac{d\Theta_1}{dt} - \lambda\pi m\Psi_1\Theta_2 + \lambda m\Psi_1 + (m^2 + \pi^2)\Theta_1 = 0, \quad (4.3)$$

$$\frac{d\Theta_2}{dt} + 4\pi^2\Theta_2 + \frac{1}{2}\lambda\pi m\Psi_1\Theta_1 = 0, \quad (4.4)$$

$$\frac{d\chi_1}{dt} - \pi m\Psi_1\chi_2 + m\Psi_1 + \frac{m^2 + \pi^2}{Le}\chi_1 + k(\chi_1 - \Theta_1) = 0, \quad (4.5)$$

$$\frac{d\chi_2}{dt} + \frac{4\pi^2}{Le}\chi_2 + \frac{1}{2}\pi m\Psi_1\chi_1 + k(\chi_2 - \Theta_2) = 0. \quad (4.6)$$

4.2. Galerkin analysis for the non-reacting problem

We will first briefly consider the Galerkin analysis for the non-reacting problem (cf. Mamou & Vasseur 1999): in other words, we take $k = 0$ in equations (4.2)–(4.6).

4.2.1. Steady state solutions

The first step towards understanding the system (4.2)–(4.6) with $k = 0$ is to seek steady-state solutions. Setting the time derivative terms equal to zero and eliminating $\Theta_{1,2}$ and $\chi_{1,2}$, we obtain the equation $\Psi_1(A\Psi_1^4 + B\Psi_1^2 + C) = 0$, where

$$\begin{aligned} A &= -\frac{1}{64}Le\lambda^2(m^2 + \pi^2)m^4, \\ B &= -(m^2 + \pi^2)\left[\frac{\lambda^2(m^2 + \pi^2)m^2}{8Le} + \frac{Le(m^2 + \pi^2)m^2}{8}\right] + \frac{m^4(\lambda^2\mathcal{R}_C + \lambda Le\mathcal{R}_T)}{8}, \\ C &= -\frac{(m^2 + \pi^2)^3}{Le} + (m^2 + \pi^2)m^2\mathcal{R}_C + \frac{\lambda(m^2 + \pi^2)m^2\mathcal{R}_T}{Le}. \end{aligned} \quad (4.7)$$

Since $A < 0$, the non-trivial solutions can be written as $\Psi_1^2 = (B \pm \sqrt{B^2 + 4|A|C})/2|A|$. We therefore conclude that: (i) if $B^2 - 4AC < 0$ then there are no non-trivial solutions; (ii) if $B^2 - 4AC > 0$ and $C > 0$ then there is exactly one positive solution for Ψ_1^2 ; (iii) if $B^2 - 4AC > 0$ and $C < 0$ then the roots both have the same sign as B , so (a) if $B > 0$ there are two positive solutions for Ψ_1^2 , and (b) if $B < 0$ there are no positive solutions. We will now write each of the marginal conditions $B = 0$, $C = 0$ and $B^2 - 4AC = 0$ as a curve in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane.

The condition $B = 0$ gives

$$\mathcal{R}_T = \mathcal{R}_T^{(B)}(\mathcal{R}_C; m) \equiv \frac{(m^2 + \pi^2)^2}{m^2} \left[\frac{1}{\lambda} + \frac{\lambda}{Le^2} \right] - \frac{\lambda}{Le}\mathcal{R}_C. \quad (4.8)$$

The condition $C = 0$ gives

$$\mathcal{R}_T = \mathcal{R}_T^{(C)}(\mathcal{R}_C; m) \equiv \frac{(m^2 + \pi^2)^2}{m^2}\lambda - \frac{Le}{\lambda}\mathcal{R}_C, \quad (4.9)$$

which is identical to the $c = 0$ condition $\mathcal{R}_T = R(M, \pi^2)$ from the linear analysis for non-reactive DDC, equation (3.17).

The condition $B^2 - 4AC = 0$ gives a bivariate quadratic equation in \mathcal{R}_C and \mathcal{R}_T , which can be written after a little manipulation as the parabola

$$(X + Y)^2 + 2\delta(X - Y) + \delta^2 = 0, \quad (4.10)$$

where $X = \lambda \mathcal{R}_C$ and $Y = Le \mathcal{R}_T$, and where the single remaining parameter is defined as

$$\delta = \frac{(m^2 + \pi^2)^2}{m^2} \left(\frac{Le}{\lambda} - \frac{\lambda}{Le} \right) > 0. \quad (4.11)$$

The parabola intercepts the lines $\mathcal{R}_T = \mathcal{R}_T^{(B)}(\mathcal{R}_C)$ and $\mathcal{R}_T = \mathcal{R}_T^{(C)}(\mathcal{R}_C)$ at the point

$$(\mathcal{R}_C^{(BC)}, \mathcal{R}_T^{(BC)}) = \frac{(m^2 + \pi^2)^2}{m^2} \frac{1}{Le^2 - \lambda^2} \left(-\frac{\lambda^2}{Le}, \frac{Le^2}{\lambda} \right); \quad (4.12)$$

note that $\mathcal{R}_T^{(B)}$ is parallel to the axis of the parabola while $\mathcal{R}_T^{(C)}$ is tangent to the parabola. To the left of $\mathcal{R}_C^{(BC)}$, steady subcritical convection is possible above the upper boundary of the parabola, while to the right of it no steady subcritical convection can occur: we will denote the portion of the parabola to the left of $\mathcal{R}_C^{(BC)}$ by $\mathcal{R}_T = \mathcal{R}_T^{(\text{sub})}$. Finally, combining the conditions $B^2 - 4AC = 0$ and $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ yields a quadratic for \mathcal{R}_C which has discriminant $(1 - \lambda^2)(\lambda^2 - Le^2) < 0$, and so there are no real intersections between the parabola and $\mathcal{R}_T = \mathcal{R}_T^{(b)}$.

The analysis so far has been carried out assuming that m is given. We recall that the lowest positions of the boundaries $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(C)}$ occur for $m = \pi$, and it is easy to see from (4.10) that the choice $m = \pi$ also minimises the value of \mathcal{R}_T for which subcritical nonlinear solutions are available: in other words, not just the linear perturbations to the base state, but also the subcritical non-linear rolls, occur most readily with equal height and width. (We recall, though, that the Galerkin analysis is unlikely to predict accurately the solution some distance from the boundary of linear stability, or the structure of nonlinear modes even near the boundary, since there is no guarantee that the higher Fourier modes will be insignificant in these cases.) Since it appears that this wavenumber is favoured at all transitions of the system, we are justified in restricting our investigation to $m = \pi$ when considering $k = 0$. This will not be the case when we come to consider $k > 0$.

4.2.2. Stability diagram and bifurcations

Combining the results from the linear stability analysis and the steady-state Galerkin analysis, we have four ‘boundary’ curves to consider: the three straight lines $\mathcal{R}_T^{(b)}$, $\mathcal{R}_T^{(B)}$ and $\mathcal{R}_T^{(C)}$, and the parabolic arc $\mathcal{R}_T^{(\text{sub})}$. The line $\mathcal{R}_T^{(b)}$ is a stability boundary, across which the zero solution $\Psi_1 = 0$ becomes unstable through an oscillatory instability when $\mathcal{R}_C < \mathcal{R}_C^{\text{crit}}$. The line $\mathcal{R}_T^{(C)}$ is a bifurcation boundary everywhere and a stability boundary when $\mathcal{R}_C > \mathcal{R}_C^{\text{crit}}$. For $\mathcal{R}_C < \mathcal{R}_C^{\text{crit}}$ it represents a subcritical bifurcation in which two nonlinear solutions merge with the zero solution as \mathcal{R}_T increases, while for $\mathcal{R}_C > \mathcal{R}_C^{\text{crit}}$ it represents a supercritical pitchfork bifurcation, in which the zero solution loses stability, through monotonically growing perturbations, to two nonlinear steady solutions which branch out from the zero solution. The line $\mathcal{R}_T^{(B)}$ has no direct significance, but determines the position of the parabolic arc $\mathcal{R}_T^{(\text{sub})}$; this arc marks a double saddle-node bifurcation at which two pairs of finite-amplitude modes appear.

Figure 3 a shows a typical stability diagram in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane, combining the results of the linear stability analysis with the (less rigorous but suggestive) results of the Galerkin analysis and of some numerical integrations of the Galerkin system. (We note that $\mathcal{R}_T^{(C)}(0) < \mathcal{R}_T^{(B)}(0) < \mathcal{R}_T^{(b)}(0)$; meanwhile $|\text{d}\mathcal{R}_T^{(B)}/\text{d}\mathcal{R}_C| < |\text{d}\mathcal{R}_T^{(b)}/\text{d}\mathcal{R}_C| < |\text{d}\mathcal{R}_T^{(C)}/\text{d}\mathcal{R}_C|$; these inequalities ensure that the relative positions of the boundary curves

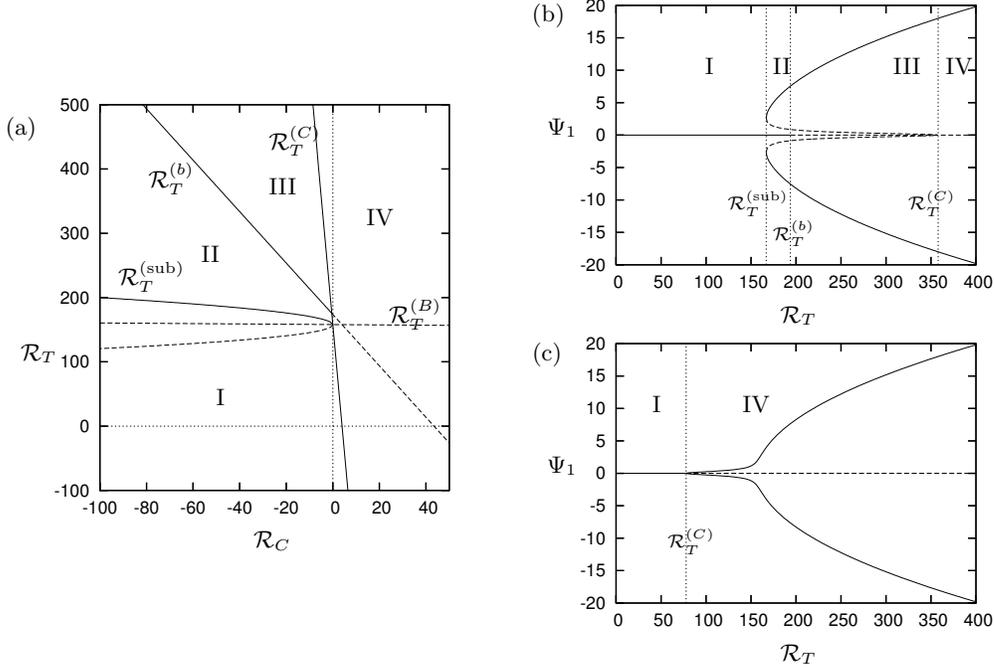


FIGURE 3. Results of Galerkin analysis with $k = 0$, for $\lambda = 0.25$, $Le = 10$ and $m = \pi$ (note that $\mathcal{R}_C^{\text{crit}} \approx -0.44$ and $\mathcal{R}_C^{(BC)} \approx -0.0025$). (a) Boundaries of stability regions in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane. Solid lines represent regime boundaries; dashed lines represent extensions of these lines or other boundaries between cases considered. (See text for explanation of regions I to IV.) (b,c) Amplitude of the steady solution for Ψ_1 against \mathcal{R}_T for $\lambda = 0.25$, $Le = 10$, $m = \pi$ and (b) $\mathcal{R}_C = -5$; (c) $\mathcal{R}_C = 2$. Solid lines represent solutions found (numerically or analytically) to be stable; dashed lines represent unstable solutions; dotted lines represent regime boundaries (regimes labelled I through IV).

are always as shown in figure 3 a.) Meanwhile, figures 3 b and c show the bifurcations of the steady solutions as \mathcal{R}_T varies, for two values of \mathcal{R}_C .

There are four distinct stability regions. In region I ($\mathcal{R}_C < \mathcal{R}_C^{(BC)}$ and $\mathcal{R}_T < \mathcal{R}_T^{(sub)}$), the only steady solution available is the unperturbed solution, $\Psi_1 = 0$, and this is linearly stable. In region II ($\mathcal{R}_C < \mathcal{R}_C^{(BC)}$ and $\mathcal{R}_T^{(sub)} < \mathcal{R}_T < \min(\mathcal{R}_T^{(b)}, \mathcal{R}_T^{(C)})$), $\Psi_1 = 0$ remains linearly stable but there are two pairs of steady nonlinear solutions (the positive and negative solutions for Ψ_1 correspond to different senses of rotation of the cells). Numerical integration of the Galerkin system (4.2)–(4.6) with $k = 0$ indicates that the outer branches (the larger magnitudes of Ψ_1) are stable, while the inner branches are unstable. A perturbation of sufficiently small initial magnitude tends to collapse onto $\Psi_1 = 0$ (generally after some oscillation), while a perturbation of rather larger magnitude collapses after some oscillation onto the outer steady solution. In region III ($\mathcal{R}_C < \mathcal{R}_C^{\text{crit}}$ and $\mathcal{R}_T^{(b)} < \mathcal{R}_T < \mathcal{R}_T^{(C)}$), $\Psi_1 = 0$ is linearly unstable, while the four non-zero steady solutions persist. Finally, in region IV ($\mathcal{R}_T > \mathcal{R}_T^{(C)}$), $\Psi_1 = 0$ is linearly unstable and there are only two non-zero solutions, which are found numerically to be attracting. (Numerically, there is no observable difference between the behaviour of the system in regions III and IV, because they differ only in the number of unstable solutions.)

If we fix \mathcal{R}_C and treat \mathcal{R}_T as the bifurcation parameter, then, three patterns of transition are possible. For $\mathcal{R}_C < \mathcal{R}_C^{\text{crit}}$, the system passes through all four regions as \mathcal{R}_T increases (figure 3 b): from region I it passes through a double-saddle-node bifurcation at $\mathcal{R}_T^{(\text{sub})}$ into region II; it then crosses the stability boundary at $\mathcal{R}_T^{(b)}$ to enter region III, and finally passes through the subcritical pitchfork bifurcation at $\mathcal{R}_T^{(C)}$ to enter region IV. For $\mathcal{R}_C^{\text{crit}} < \mathcal{R}_C < \mathcal{R}_C^{(BC)}$, the transition is from region I to region II as before; the loss of stability of the zero state then coincides with the subcritical pitchfork bifurcation at $\mathcal{R}_T^{(C)}$, and the system passes into region IV. Finally, for $\mathcal{R}_C > \mathcal{R}_C^{(BC)}$ (figure 3 c), the system passes directly from region I to region IV as \mathcal{R}_T is increased: the transition corresponds to the supercritical pitchfork bifurcation at $\mathcal{R}_T^{(C)}$.

4.3. Galerkin analysis for the reacting problem

To investigate the full system (4.2)–(4.6), we start as above by seeking steady solutions to this set of equations. We will first consider the problem assuming that m has been set (for example, by the dynamics associated with the loss of stability of the base state); we will consider afterwards how the variation of m affects the results. Solving for $\Theta_{1,2}$ and $\chi_{1,2}$, we obtain in particular

$$\Theta_2 = \frac{\lambda^2 m^2 \Psi_1^2}{\pi[8(m^2 + \pi^2) + \lambda^2 m^2 \Psi_1^2]} \quad \text{and} \quad (4.13)$$

$$\chi_2 = \Theta_2 \frac{2\kappa^2 + (8\pi^2 \frac{Le}{\lambda} + 2(m^2 + \pi^2))\kappa + 8\pi^2 \frac{Le^2}{\lambda^2} (m^2 + \pi^2) + Le^2 \pi^2 m^2 \Psi_1^2}{2\kappa^2 + (10\pi^2 + 2m^2)\kappa + 8\pi^2 (m^2 + \pi^2) + Le^2 \pi^2 m^2 \Psi_1^2} \quad (4.14)$$

where $\kappa = Lek$. Changing the reaction rate k , then, affects both the amplitude of the convective motion and the relative magnitudes of the thermal and solutal perturbations.

Continuing as before, we obtain the equation $\Psi_1(A\Psi_1^4 + B\Psi_1^2 + C) = 0$, where

$$\begin{aligned} A &= -\pi^2 m^4 Le^2 \lambda^2 (m^2 + \pi^2), \\ B &= -2\lambda^2 Le^2 k^2 m^2 (m^2 + \pi^2) - 2\lambda^2 Lek m^2 (m^2 + \pi^2)(m^2 + 5\pi^2) \\ &\quad + 8\pi^2 m^4 \lambda Le (\lambda \mathcal{R}_C + Le \mathcal{R}_T) - 8\pi^2 m^2 (m^2 + \pi^2)^2 (Le^2 + \lambda^2), \\ C &= 16Le^2 k^2 [\lambda m^2 (\mathcal{R}_C + \mathcal{R}_T) - (m^2 + \pi^2)^2] + 16Lek (m^4 (\lambda \mathcal{R}_T + Le \mathcal{R}_C) \\ &\quad + \pi^2 m^2 (5\lambda \mathcal{R}_T + 4\lambda \mathcal{R}_C + Le \mathcal{R}_C) - (m^2 + \pi^2)^2 (m^2 + 5\pi^2)) \\ &\quad + 64 [m^2 \pi^2 (m^2 + \pi^2) (Le \mathcal{R}_C + \lambda \mathcal{R}_T) - \pi^2 (m^2 + \pi^2)^3]. \end{aligned} \quad (4.15)$$

As before, we need to locate the boundaries in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane which correspond to the conditions $C = 0$, $B = 0$ and $B^2 - 4AC = 0$.

The condition $B = 0$ becomes $\mathcal{R}_T = \mathcal{R}_T^{(B)}(\mathcal{R}_C; m)$, where

$$\mathcal{R}_T^{(B)} = \frac{\lambda (m^2 + \pi^2)}{4} \frac{k^2}{m^2 \pi^2} + \frac{\lambda (m^2 + \pi^2)(m^2 + 5\pi^2)}{4} \frac{k}{\pi^2 m^2 Le} + \frac{(m^2 + \pi^2)^2}{m^2} \left[\frac{1}{\lambda} + \frac{\lambda}{Le^2} \right] - \frac{\lambda}{Le} \mathcal{R}_C. \quad (4.16)$$

The condition $C = 0$ becomes

$$\mathcal{R}_T = \mathcal{R}_T^{(C)}(\mathcal{R}_C; m) = \frac{Lek(m^2 + \pi^2)^2 + (m^2 + \pi^2)^3}{\lambda(Lek + m^2 + \pi^2)m^2} - \frac{\lambda Lek + Le(m^2 + \pi^2)}{\lambda(Lek + m^2 + \pi^2)} \mathcal{R}_C. \quad (4.17)$$

As for the non-reacting case, this is identical to the $c = 0$ boundary $\mathcal{R}_T = R(M, \pi^2)$ given by equation (3.17), and this suggests that the presence of reactions does not alter the structure of the bifurcation across this boundary.

The condition $B^2 - 4AC = 0$ also defines a parabola in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane. The details

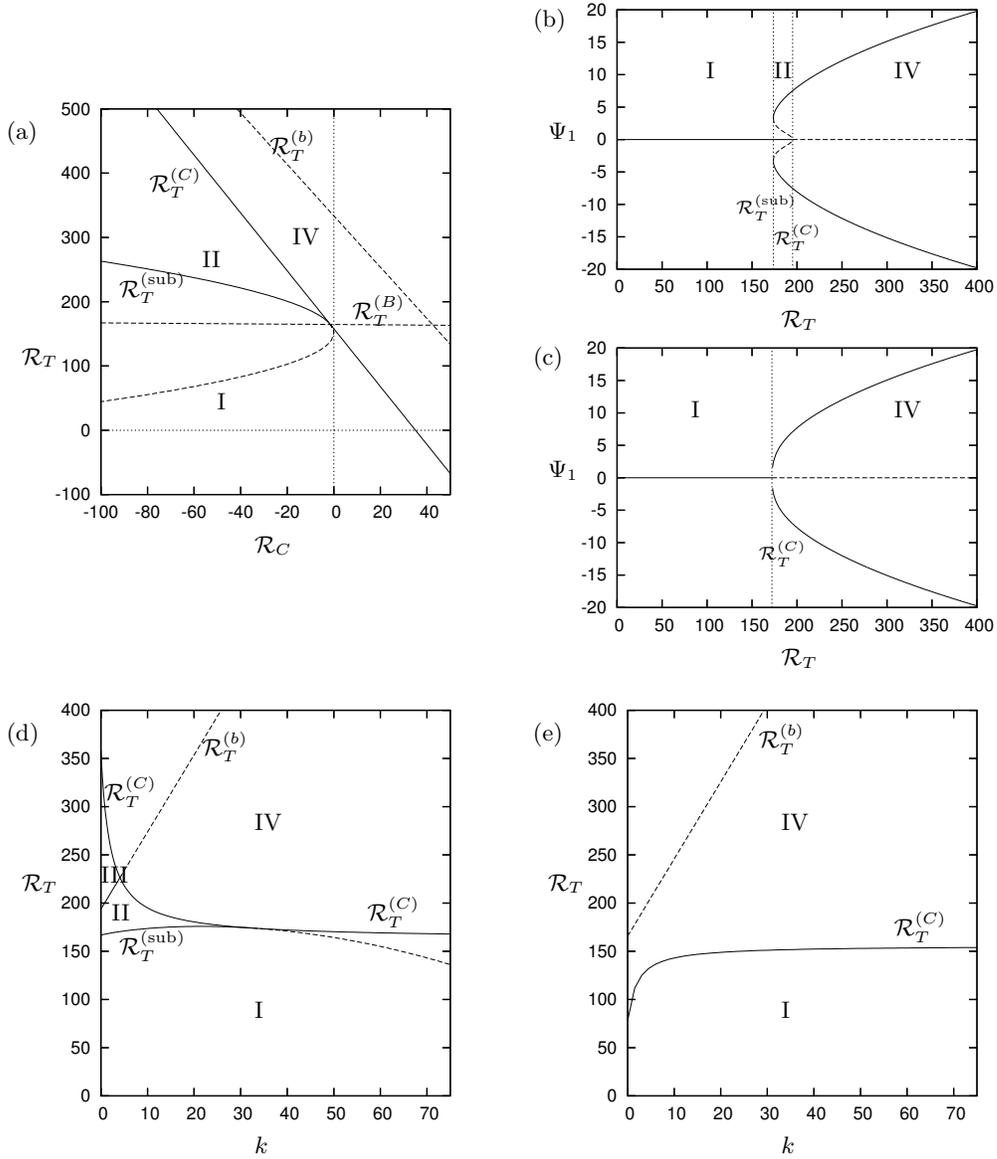


FIGURE 4. Results of Galerkin analysis with $k > 0$, for $\lambda = 0.25$, $Le = 10$ and $m = \pi$. (a) Boundaries of stability regions in the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane, for $k = 20$; legend as for figure 3 a. (b,c) Amplitude of the steady solution for Ψ_1 against \mathcal{R}_T for $\lambda = 0.25$, $Le = 10$, $m = \pi$, $\mathcal{R}_C = -5$ and (b) $k = 10$; (c) $k = 40$; legend as for figures 3 b and c. (d,e) Stability and bifurcation boundaries in the (k, \mathcal{R}_T) -plane, for (d) $\mathcal{R}_C = -5$, (e) $\mathcal{R}_C = 2$. Legend is the same as in (a).

are given in appendix A.3: as for $k = 0$, the parabola is centred on the line $\mathcal{R}_T = \mathcal{R}_T^{(B)}$, and the line $\mathcal{R}_T = \mathcal{R}_T^{(C)}$ is tangent to the parabola.

Figure 4 illustrates how the linear and nonlinear stability properties of the system change as k is varied, for fixed m . We first consider figure 4 a, comparing it with figure 3 a for $k = 0$. As k increases, the intersection $(\mathcal{R}_C^{(BC)}, \mathcal{R}_T^{(BC)})$ between $\mathcal{R}_T^{(B)}$ and $\mathcal{R}_T^{(C)}$

moves up and to the left. The parabola $B^2 - 4AC = 0$ continues to be tangent to the \mathcal{R}_T -axis, and it becomes correspondingly wider as $\mathcal{R}_C^{(BC)}$ moves away from this axis. Meanwhile the gradient of $\mathcal{R}_T^{(C)}$ reduces more rapidly than that of $\mathcal{R}_T^{(b)}$, with the effect that region IV occupies an increasing region (this is the contribution of reaction to instability). For sufficiently large k , the slope $|\mathrm{d}\mathcal{R}_T^{(C)}/\mathrm{d}\mathcal{R}_C|$ becomes less than the slope $|\mathrm{d}\mathcal{R}_T^{(b)}/\mathrm{d}\mathcal{R}_C| = 1/\lambda$, and the shape of the stability diagram changes. This transition occurs when $k = (1 - 1/Le)(m^2 + \pi^2)/(1 - \lambda) > k_{\text{crit}}$; we will not consider it further, because when $k \gtrsim k_{\text{crit}}$ we know that $\mathcal{R}_T^{(c)}$ provides the stability boundary everywhere, and the position of $\mathcal{R}_T^{(b)}$ is no longer physically relevant.

Figures 4 b–e show how the stability picture alters as k is increased for a fixed value of \mathcal{R}_C , both in the SS regime $\mathcal{R}_C < 0$ (figures 4 b–d) and the SU regime $\mathcal{R}_C > 0$ (figure 4 e). In the SU regime, the transition is always the pitchfork bifurcation from I to IV at $\mathcal{R}_T = \mathcal{R}_T^{(C)}$. In the SS regime, at low k , the picture is essentially identical to that for $k = 0$ (see figure 3 b). As k increases, region III disappears (figure 4 b): the loss of stability of the zero solution coincides with the disappearance of the unstable finite-amplitude solution branches, so $\mathcal{R}_T^{(C)}$ marks a subcritical pitchfork bifurcation. Finally, for sufficiently large k , our fixed value of \mathcal{R}_C lies to the right of the point where $\mathcal{R}_T^{(C)}$ glances against the parabola, and we again have a supercritical pitchfork bifurcation from region I to region IV (figure 4 c).

4.3.1. Interpretation of results when m is not externally imposed

It is necessary to consider how we should interpret these results for $k > 0$, given that the preferred wavenumber at the onset of instability may now vary strongly with the other parameters. The interpretation is easiest for values of \mathcal{R}_C such that $\mathcal{R}_T^{(c)} < \mathcal{R}_T^{(b)}$, so the stability boundary is given by $\mathcal{R}_T^{(c)}$. In this case, the stability boundary for any given m is identical to the line $\mathcal{R}_T^{(C)}$ for the same value of m ; consequently, whichever mode is selected by the linear instability, the mechanism will be the same, and the loss of stability is through a pitchfork bifurcation.

When $\mathcal{R}_T^{(b)} < \mathcal{R}_T^{(c)}$, the stability boundary is given by $\mathcal{R}_T^{(b)}$, and we know that the onset will be through an oscillatory instability which, when \mathcal{R}_T is close to the stability boundary, selects a mode of wavenumber $m^2 = M_b$. For the cases plotted in figure 4, the stability boundary is always above the bifurcation boundary $\mathcal{R}_T^{(\text{sub})}$ at which the nonlinear steady solutions appear, and so it is always possible for the growing perturbation ultimately to ‘lock on’ to one of these solutions. It is necessary to show that these solutions are always available (otherwise the wavenumber of the convective rolls would have to change for a steady solution to emerge, and it is conceivable that no steady behaviour would be available). This will occur if the bifurcation boundary $\mathcal{R}_T^{(\text{sub})}$ for the mode $m^2 = M_b$ lies below the boundary $\mathcal{R}_T^{(b)}$ for all values of \mathcal{R}_C , λ , Le and k . This in turn will occur if there are no intersections between the line $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ and the parabola $B^2 - 4AC = 0$, where A , B and C are evaluated at $m^2 = M_b$. Seeking such an intersection leads to a quadratic equation for \mathcal{R}_C , and while we have not been able to prove that the discriminant of this equation is always negative, extensive numerical experimentation indicates that this is the case across the whole range of parameters ($0 < \lambda < 1$, $1 < Le \leq 100$, $0 \leq k \leq 50$) investigated. We may therefore be fairly confident that the bifurcation structure suggested by figure 4 does represent the means by which the onset of convection occurs.

Finally, it is also interesting to consider how the position of the boundary $\mathcal{R}_T = \mathcal{R}_T^{(\text{sub})}$

between regions I and II varies with m . In general, there does not appear to be a closed-form expression for the value of m that minimises $\mathcal{R}_T^{(\text{sub})}$. However, it is straightforward to carry out the minimisation numerically for given parameter values and thus to obtain an estimate of the maximal region in which the Galerkin analysis predicts that steady subcritical convection is possible. An example of this estimate will be shown below in figure 6.

5. Numerical tests of the linear stability analysis

In general, linear stability analysis and the Galerkin method described in §4 have proved reliable methods of investigating convective systems. However, it is useful to test their predictions by comparing them with a full numerical integration of the system (2.1)–(2.4), for several reasons. Firstly, the complexity of the analysis makes an independent check desirable. Secondly, numerical methods can give information about the fully non-linear behaviour of the system, including the presence or absence of subcritical modes, which the methods employed above can only suggest rather than establish. Thirdly, a numerical approach is not restricted to considering initial conditions which differ only slightly from the conductive state, and it can therefore test not only whether the conductive state is stable, but whether it is more generally attracting, thus strengthening the conclusions which can be drawn from the linear analysis. (This is particularly valuable in regions where the Galerkin analysis suggests that multiple steady solutions are available.) Finally, the inevitable restrictions of numerical methods (for example, to a horizontally bounded or periodic domain) provide a test of how applicable the conclusions drawn from the idealised analysis are to less idealised geometries. We have therefore carried out a series of numerical experiments to test the predictions of §3 and to examine briefly the possibility of subcritical instability.

5.1. Numerical approach

5.1.1. Numerical integration scheme

The equations (2.1)–(2.4), nondimensionalised according to equation (2.6), were integrated numerically using a standard Cholesky factorisation solver for the pressure field, combined with the semi-Lagrangian Crank–Nicolson (SLCN) scheme described by Spiegelman & Katz (2006) for the transport–reaction equations. This scheme, which was developed for geological transport–reaction problems, is computationally efficient and is stable even for large timesteps, allowing large series of numerical experiments to be run in a relatively short time.

In the SLCN scheme, the transport–reaction equations are discretised in time over the interval $t_n \leq t \leq t_{n+1}$ using the values of T and C on the spatial grid points at $t = t_{n+1}$, while the values of T and C at $t = t_n$ are calculated by interpolation at spatial points which are ‘back-projected’ from the grid points using a velocity field interpolated at the half-way time $t_{n+1/2} = (t_n + t_{n+1})/2$. The purely advective part of the transport equation is thus accounted for automatically in the time-discretisation; meanwhile the diffusion and reaction terms are evaluated using interpolated values of T and C on the spatial mid-way points at $t_{n+1/2}$. This gives an implicit scheme which, given the variables at $t = t_n$ together with some estimate of the velocity at $t_{n+1/2}$, may readily be solved for T and C at $t = t_{n+1}$ by standard methods.

The integration procedure at each timestep $t = t_n$ is iterative. First the pressure is calculated from the existing T and C fields, and the velocity at t_n is calculated from this. A ‘half-way’ velocity field at $t_{n+1/2}$ is now obtained by interpolation, and this is used in the

SLCN scheme to obtain the T and C fields at $t = t_{n+1}$. (On the first iteration, the estimate $\mathbf{u}(\mathbf{x}, t_{n+1/2}) = \mathbf{u}(\mathbf{x}, t_n)$ is used; thereafter $\mathbf{u}(\mathbf{x}, t_{n+1/2}) = [\mathbf{u}(\mathbf{x}, t_n) + \mathbf{u}(\mathbf{x}, t_{n+1})]/2$.) An improved estimate of the pressure and velocity fields at t_{n+1} is now calculated by solving for p with the new estimates of T and C at this time; the procedure can then be iterated until it has converged. (In the computations shown here, all fields were calculated to an accuracy of 10^{-9} .) To reduce diffusive errors, the time steps in all calculations were chosen adaptively to be close to the Courant condition.

5.1.2. Initial and boundary conditions, numerical parameters and validation

For all the numerical experiments described here, the initial condition was taken to be

$$C = T = \frac{1}{2} - \frac{1}{2} \tanh \left(10 \left[z - \frac{1}{2} + \delta \sin \left(\frac{2\pi j x}{5} \right) \right] \right) \quad \text{for } j \in \mathbb{N}. \quad (5.1)$$

The boundary conditions (2.12) were applied on the horizontal boundaries, while periodicity was imposed at the vertical boundaries.

Equation (5.1) provides a smooth initial condition in which the initial vertical gradients occur away from the boundaries: by varying the parameter δ the integration can be ‘seeded’ more strongly with a periodic disturbance having wavenumber $m = 2\pi j/5$ (so there are j periods within the computational domain of width 5 used in our calculations). Values $\delta = 1$ (a ‘strongly periodic’ initial condition) and $\delta = 10^{-5}$ (a ‘weakly periodic’ initial perturbation) were used in the numerical experiments described below: j was taken to be 1 unless otherwise stated. This initial condition clearly does not represent a small perturbation to the conductive base state, as investigated in §3. In comparing our numerical results with the linear analysis, then, we do not investigate directly the growth or decay of perturbations to this base state; rather, we use the initial condition (5.1) to investigate the long-term state of the system, thus testing whether the linear and Galerkin analyses, despite their specialised nature, capture the overall behaviour of the system. A further advantage of starting far from the conductive base state is that it is easier to access the nonlinear steady states available in region II of the stability plots (figures 3 and 4). Within this region, we expect the initial condition to be important; outwith it, we do not expect the initial condition to matter, and a series of preliminary experiments, varying j and δ in equation (5.1) as well as considering initial conditions comprising a conductive profile plus some small-amplitude noise, supported this expectation.

The equations were solved in a rectangular domain of 30×150 square grid cells, i.e. $0 < x < 5$ and $0 < z < 1$. This size was chosen to allow a large number of calculations in a short time, while providing an aspect ratio of 5 (which is sufficient for multiple cells to develop). The stability calculations for $k = 50$ and for selected values of $(\mathcal{R}_C, \mathcal{R}_T)$ with $k = 10$, and the wavenumber calculations in §5.3, were repeated at double this resolution with practically identical results, suggesting that the resolution was adequate to resolve the convection. The only region in which there were systematic differences was that of very high \mathcal{R}_C and k , where the number of convective cells becomes large and could not be adequately resolved at the lower resolution. (The results shown in §5.3 were obtained at the higher resolution.) In addition to these repeated experiments, the method was validated for pure DDC by testing it against simple analytical results for the advection and diffusion of a point source of heat or solute, and by testing its predictions of heat transfer across the layer against the results tabulated by Schoofs (1999) (§3.4.1): agreement to within 1% was obtained in the latter case. Finally, its ability to reproduce the stability diagram for $k = 0$ (figure 2 b) was tested and found to be satisfactory.

5.2. Stability boundaries and subcritical convection

Figures 5 and 6 summarise the results of a large number of numerical experiments carried out to validate the linear stability analysis against the numerics. In each experiment, the integration was started from the initial condition (5.1), and the numerical experiments were continued for 1000 iterations, corresponding to dimensionless times t between around 40 and 2500; the criterion used to distinguish between ‘conductive’ and ‘convective’ cases was whether by the end of the experiment the maximum value of $|\mathbf{u}|$ in the domain was greater or less than 10^{-5} . (In a real reactive fluid-rock system, the reactions might conceivably cause significant changes to the porosity or permeability of the rock over these very long times (cf. §2): our object here is not to consider the effect of such feedback, but we note it as an important topic which we hope to address in future work.)

Figure 5 a shows the results for $k = 50$, when the stability boundary is given everywhere by $\mathcal{R}_T = \mathcal{R}_T^{(c)}$. The linear analysis agrees very well with the numerical results, with only a few anomalies very near the boundary. Possible explanations for these include the fact that growth or decay rates in this region are small (making it hard to distinguish between stable and unstable cases); the restrictions on m imposed by the horizontally periodic domain; and the limits of the numerical resolution (see the discussion of figure 7 f, below.) However, the results overall are very encouraging and confirm the strong effect of the reaction term on the stability of the system.

The picture for weaker reactions is more complicated. In the SU regime $\mathcal{R}_C > 0$, the linear analysis again predicts the boundary between conductive and convective behaviour very well (figure 5 b). In the SS regime $\mathcal{R}_C < 0$, there is clear evidence of a region of subcritical convection. This is explored further in figure 6, which summarises the results for six series of experiments carried out in the SS regime, with $\delta = 1$ and 10^{-5} and $j = 1, 3$ and 10 .

The most striking feature of the results for a strongly periodic initial condition, $\delta = 1$ (figure 6 a) is the strong dependence of the eventual state on the initially imposed wavenumber j . Of the values tried, $j = 3$ was the one from which steady nonlinear convection was most readily accessible, and thus corresponds to the lowest pair of lines in figure 6 a. With $j = 1$, the subcritical convective solutions were slightly less accessible, and for $j = 10$ there is no evidence that subcritical convective solutions were accessible at all: the highest conductive and lowest convective cases neatly straddle the linear stability boundary. (In all cases in which a subcritical convective state was established, it consisted of three complete convective cells within the computational domain, i.e. $m = 6\pi/5$; it is therefore not surprising that these states were most readily accessed from the initial condition with $j = 3$.) Bearing in mind that even for $\delta = 0$, the initial condition (5.1) is somewhat different from the conductive base state, it appears that the key to accessing subcritical convection is the presence of a marked horizontal variation with an appropriate scale, which favours horizontally non-uniform convective rather than horizontally uniform conductive behaviour.

In contrast, the results with small δ (figure 6 b) all follow the linear stability boundary reasonably closely, and there is very little difference between the results for various values of the imposed wavenumber j . (A close examination reveals that the results for $j = 3$ track the linear stability boundary almost exactly, while those for $j = 1$ and $j = 10$ are slightly more stable than the linear analysis predicts: this may be an effect of the very slow growth rates near the stability boundary, which mean that perturbations take an extremely long time to emerge unless the initial state is ‘seeded’ with some disturbance close to the maximally unstable wavenumber.)

The difference between the small- and large- δ results makes it clear that in the region

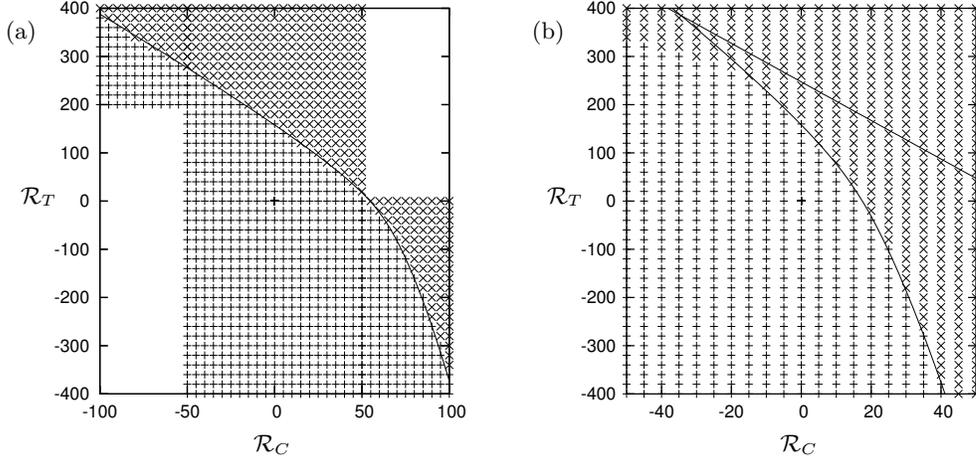


FIGURE 5. Numerical tests of the stability boundaries for $\lambda = 0.25$, $Le = 10$ and (a) $k = 50$; (b) $k = 10$, with a strongly periodic initial perturbation ($\delta = 1$ and $j = 1$). Solid lines represent $\mathcal{R}_T = \mathcal{R}_T^{(c)}$ and $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ (cf. figures 2 c and e). In all plots, points represent conductive (+) and convective (\times) cases.

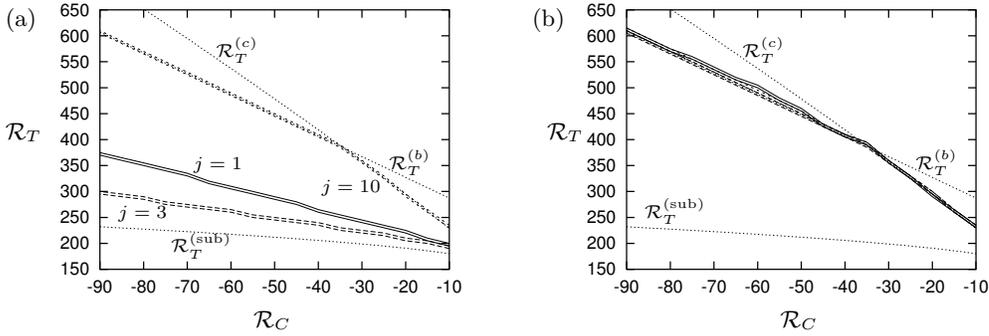


FIGURE 6. Effect of initial condition on development of the system in the region of subcritical convection: plots summarise several series of numerical experiments for $\lambda = 0.25$, $Le = 10$, $k = 10$ and (a) ‘strongly periodic’ initial condition with $\delta = 1$; (b) ‘weakly periodic’ initial condition $\delta = 10^{-5}$. For each value of j , two lines are shown, marking the highest conductive and lowest convective values of \mathcal{R}_T found. Values of j considered were $j = 1$ (solid); $j = 3$ (heavy dashed); $j = 10$ (light dashed). Dotted lines mark the stability boundaries $\mathcal{R}_T^{(b)}$ and $\mathcal{R}_T^{(c)}$ (cf. figure 2 c), along with the lower boundary $\mathcal{R}_T^{(\text{sub})}$ of the region where the Galerkin analysis predicts steady subcritical convection.

immediately below the stability boundary for $\mathcal{R}_C < 0$, nonlinear convective solutions are available even though the conductive base state is stable, and that the accessibility of these solutions may depend crucially on the initial condition of the system. It is interesting that the Galerkin analysis predicts that subcritical convection can occur in a slightly wider region than that in which we have found it: this may reflect the inaccuracy of the Galerkin ansatz far from the linear stability boundary, or simply the restrictions of the particular initial conditions we have employed. We intend to investigate this region further in a future study.

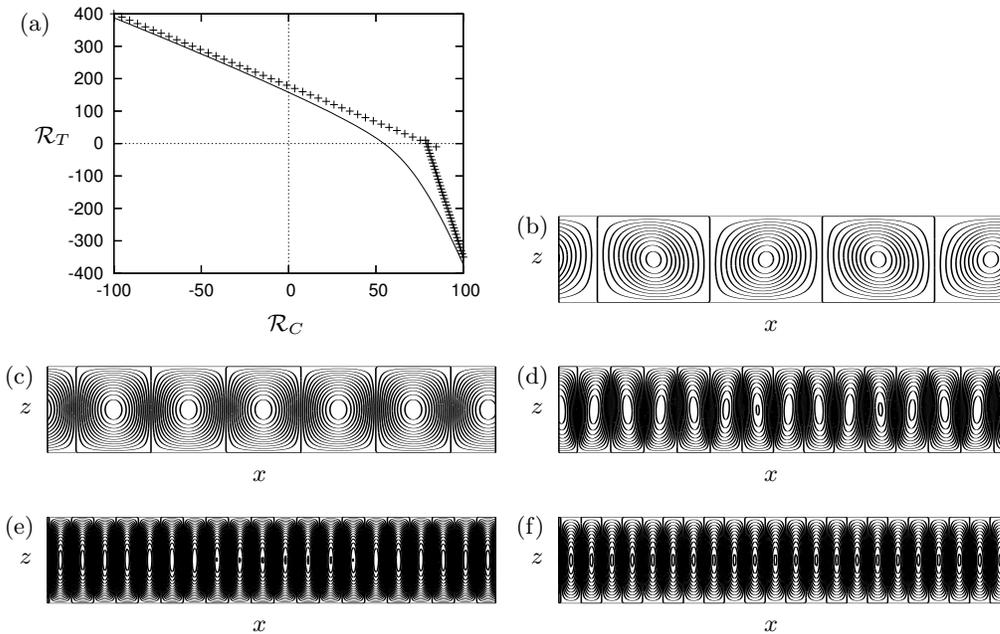


FIGURE 7. Summary of numerical experiments to test wavenumber increase, with $\lambda = 0.25$, $Le = 10$ and $k = 50$. (a) Locations of points $(\mathcal{R}_C, \mathcal{R}_T)$ for which experiments were carried out. (b)–(f) Streamlines of the final steady convective circulation, for (b) $\mathcal{R}_C = -100$, $\mathcal{R}_T = 400$ ($|\mathbf{u}|_{\max} \approx 1.50$ at $t = 410$); (c) $\mathcal{R}_C = -10$, $\mathcal{R}_T = 200$ ($|\mathbf{u}|_{\max} \approx 0.92$ at $t = 272$); (d) $\mathcal{R}_C = 80$, $\mathcal{R}_T = 0$ ($|\mathbf{u}|_{\max} \approx 0.88$ at $t = 47.7$); (e) $\mathcal{R}_C = 91$, $\mathcal{R}_T = -200$ ($|\mathbf{u}|_{\max} \approx 0.37$ at $t = 142$); (f) $\mathcal{R}_C = 100$, $\mathcal{R}_T = -350$ ($|\mathbf{u}|_{\max} \approx 6.9 \times 10^{-3}$ at $t = 401$).

5.3. Wavenumbers

The other key prediction of the linear stability analysis is that for $k > 0$ the favoured wavenumber of the convective motion increases strongly with increasing \mathcal{R}_C . To test this prediction, we carried out a series of experiments, taking transects across the $(\mathcal{R}_C, \mathcal{R}_T)$ -plane just above the stability boundary. In each case, the perturbation was initiated and allowed to reach a steady or near-steady state as in §5.2. The number n_c of cells in the periodic domain $0 \leq x \leq 5$ was counted for each steady state, providing the wavenumber $M = (2\pi n_c/5)^2$.

A summary of these experiments and some specimen plots of the steady convective motion are shown in figure 7. These illustrate well the changing pattern of convection which occurs along the stability boundary as \mathcal{R}_C is increased: for large negative \mathcal{R}_C (figure 7 b) the convective rolls are slightly wider than they are square; as \mathcal{R}_C increases the cells become increasingly narrow, until in figure 7 f they have reached an aspect ratio of more than 4:1. Note also that there is a slight suggestion of irregularity in figures 7 d and e, which suggests that the convection may not have completely stabilised by the end of these experiments. (Typically in our experiments, the dominant wavenumber became established rapidly, but there was some long-wave variation inherited from the initial condition which took longer to die away.)

This variation is described quantitatively in figure 8. The results cover a range of predicted wavenumbers from $M = 6.15$ (at $\mathcal{R}_C = -100$, $\mathcal{R}_T = 400$) to $M = 184.2$ (at $\mathcal{R}_C = 100$, $\mathcal{R}_T = 350$), capturing the strong increase of M along the stability boundary

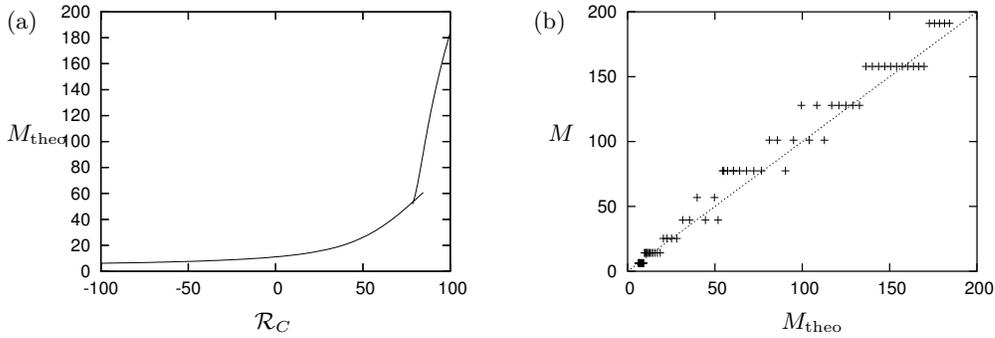


FIGURE 8. Results from the series of numerical experiments to test predictions of favoured wavenumber. Parameters used were $\lambda = 0.25$, $Le = 10$, $k = 50$ and values of $(\mathcal{R}_C, \mathcal{R}_T)$ along the transects plotted in figure 7 a. (a) Predictions for M from the linear stability analysis at the various points along the two transects. (b) Numerically obtained wavenumbers M compared to predictions from the linear stability analysis. The dotted line marks $M = M_{\text{theo}}$.

which is apparent in figures 1 d and 2 d. Allowing for the quantisation of M which is imposed by the horizontal periodicity of the computational domain, the agreement between the numerical results and the linear stability analysis is good, indicating that the eventual convection pattern is indeed determined by the favoured linear mode, and that a strong change in the convection pattern can be caused by the reaction. There are occasional anomalies, where the number of cells is one more or fewer than might be expected. This indicates that the growth rate of the favoured wavenumber is sometimes not very much greater than those of the adjacent wavenumbers, and so the eventual steady state may be affected by the choice of initial perturbation or by numerical error. However, the overall trend in M survives these minor anomalies.

The results of these numerical tests suggest that the linear stability analysis successfully captures the processes around the onset of instability, and in particular that imposing horizontal periodicity does not affect the character of these processes. They also indicate the need to investigate more thoroughly the nonlinear dynamics in the SS regime and possibly also further from the linear stability boundary. However, these investigations lie beyond the scope of the current study.

6. Discussion and conclusions

We have used a combination of stability analysis and numerical integration to build up a picture of how temperature-dependent mineral solubility affects the onset of thermosolutal convection in a porous medium. In contrast to most previous studies, we have focussed on cases in which the dominant contribution of the reaction to buoyancy comes from the change in solute concentration rather than from changes in temperature. As expected, we have found that the precipitation and dissolution of the mineral may have a significant effect on the stability of the conductive base state and on the patterns of convection which occur.

One important finding is that in general, the presence of reactions affects the critical wavenumber at which the onset of instability occurs: similarly strong, though qualitatively different, dependences of critical wavenumber on reaction rate were found in the problems considered by Gutkiewicz-Krusin & Ross (1980) and by Steinberg & Brand (1983). In the solutally unstable (SU) regime $\mathcal{R}_C > 0$, reactions may strongly increase

this wavenumber, so the convective cells which result are narrower than a model of pure double-diffusive convection would predict. In the solutally stabilised (SS) regime $\mathcal{R}_C < 0$, it may increase or decrease the wavenumber, depending whether the conductive state is subject to a monotonic or an oscillatory instability. This change in the spatial pattern of convection has obvious applications to ore deposition problems such as that considered by Raffensperger & Garven (1995*a,b*).

Precipitation and dissolution also change the positions of the stability boundaries from those in pure DDC. In the SU regime, they tend to stabilise the system, as destabilising solute is removed from solution. For weakly destabilised cases ($0 < \mathcal{R}_C < k$), the stabilising effect of the reaction term can stabilise the system even in the absence of diffusion, while for higher values of \mathcal{R}_C , diffusion is essential to the stability. This demonstrates the danger of disregarding the solutal diffusion term even for very fast reactions (cf. Steinberg & Brand 1983), and indicates when particular care may be necessary in the handling of solutal diffusion in more detailed numerical simulations of convecting geothermal systems. In the SU regime, the transition to instability is always through a supercritical (pitchfork) bifurcation, as in pure DDC.

In the SS regime, the effect of precipitation and dissolution is more complicated. In pure DDC, instability may occur, as \mathcal{R}_T is increased, through either an oscillatory or a monotonically growing perturbation; the bifurcation structure near the stability boundary may be sub- or supercritical. For sufficiently solutally stable systems ($\mathcal{R}_C < \mathcal{R}_C^{\text{crit}}$), the instability is oscillatory and stability is lost to a pair of finite-amplitude steady solutions. For very weakly solutally stable systems ($\mathcal{R}_C^{(BC)} < \mathcal{R}_C < 0$), the instability is monotonic and the bifurcation is a supercritical pitchfork (as in the SU regime). For intermediate cases ($\mathcal{R}_C^{\text{crit}} < \mathcal{R}_C < \mathcal{R}_C^{(BC)}$), the instability is still monotonic, but stability is again lost to a pair of finite-amplitude steady solutions.

One effect of the reaction is to inhibit oscillatory instabilities, increasing the critical thermal Rayleigh number $\mathcal{R}_T^{(b)}$ which corresponds to oscillatory onset. This may occur because the reaction term couples together the temperature and concentration fields and so inhibits double-diffusive and double-advective effects. For $\mathcal{R}_C < 0$, the effect of the reaction on the monotonic mode of instability is destabilising: the corresponding boundary $\mathcal{R}_T^{(c)}$ is reduced as k increases and stabilising solute is removed from the system. The consequence of this is that, if we hold $\mathcal{R}_C < 0$ fixed and increase k , we may see the stability threshold first rise and then fall with increasing k . For $\mathcal{R}_C > \mathcal{R}_C^{\text{crit}}$, the instability is always monotonic, and the system becomes more unstable as k increases. For $\mathcal{R}_C < \mathcal{R}_C^{\text{crit}}$, initially the preferred mode of instability is oscillatory and the reaction is stabilising; as k is increased further, the preferred mode of instability becomes monotonic, and the reaction becomes destabilising. As k increases, the region of supercritical bifurcations increases at the expense of the region of subcritical bifurcations, and ultimately the subcritical region disappears altogether for sufficiently high k . These predictions are all supported by the numerical experiments described in §5.

While the results presented here were obtained for a rather idealised geometry, we expect our qualitative findings to be more widely applicable. From a practical point of view, the key implication of our results is that when simulating convective geochemical systems, the treatment of reaction rates may be crucially important. Reaction rates in geochemistry are often poorly constrained to within orders of magnitude, and as we have show here, this uncertainty might lead to significant errors in the prediction of fluid motion. This may be of particular interest in studies of geothermal systems, where the onset of convection may substantially increase the heat flux in the system: our

results suggest that the geochemical controls on this process should be more thoroughly investigated.

There are many opportunities for further theoretical work to explore the thermosolutal reaction–convection system described here. Natural extensions of the stability analysis include considering other boundary conditions at the upper and lower surfaces (for example, heat flux conditions or chemical disequilibrium), or studying the role of effects such as anisotropic permeability or nonlinear equilibrium solubility curves. In a future study, we intend to investigate further the nonlinear dynamics of the system, with a particular focus on behaviour in the subcritical SS regime (which Mamou & Vasseur (1999) found to be particularly rich in the case of pure DDC), on quantifying heat and mass transfer across the layer, and on the long-term evolution of the porous matrix as precipitation and dissolution affects its porosity and permeability. It is likely that in this paper we have done no more than to scratch the surface of an interesting and complicated problem.

We are grateful to Professor Andy Woods for helpful comments and encouragement. We would also like to thank three anonymous referees for their comments, which have helped to clarify and strengthen this paper substantially. DP acknowledges financial support at the BP Institute through a postdoctoral fellowship under the NERC/EPSRC EMS scheme (ref. NE/B50188X/1), as well as the hospitality of the Institute during a visit in April 2006.

Appendix A. Technical details of various results

This appendix includes various mathematical details which have been omitted from the main text for brevity.

A.1. Minimising $R(M, N)$ over M and N

We start from the condition (3.17). We wish to minimise $R(M, N)$ over M and N in the semi-open ranges $M \geq 0$ and $N \geq \pi^2$, with the additional condition that $N/\pi^2 \in \mathbb{N}$. For the moment we will ignore the last condition and address the problem assuming that N varies continuously; we will consider the consistency of this later.

The global minimum may occur within the (M, N) -region of interest, in which case it must also be a local minimum over M and N simultaneously, or it may occur at the boundary of the region, or in one of the limits $N \rightarrow \infty$ or $M \rightarrow \infty$. If it occurs at the boundary $N = \pi^2$ but not in one of the limits of M , then it must correspond to a local minimum over M ; similarly if it occurs at the boundary $M = 0$ but not in one of the limits of N , it must correspond to a local minimum over N .

We will first dispose of the limits of large M or N . As $M \rightarrow \infty$, $R \sim M/\lambda$, while as $N \rightarrow \infty$ or $M \rightarrow 0$, $R \sim N^2/(M\lambda)$. The global minimum therefore cannot occur in any of these limits. We are left with two possibilities: either the global minimum occurs internally or on the boundary $N = \pi^2$. We will attempt to locate an internal local minimum in the region $N > 0$, $M > 0$: if no such minimum exists, then the global minimum must occur on $N = \pi^2$, while if such a minimum does exist then we will have the additional task of determining whether it occurs for $N \geq \pi^2$ and whether it is a global minimum.

We now consider how many local minima we may expect over M and over N .

For general values of M , we find that $\partial R/\partial N = 0$ if and only if $f(N) = 0$, where

$$f(N) = N^3 + (2kLe + 3M)N^2 + (k^2Le^2 + 4kLeM + 3M^2)N + M(M + kLe)^2 + \frac{1}{2}kLeM\mathcal{R}_C(\lambda - Le). \quad (\text{A } 1)$$

We note that $df/dN > 0$ for all $N > 0$, and so $f(N)$ has a root for $N > 0$ if and only if $f(0) < 0$; i.e. $R(M, N)$ has a local minimum over N if and only if

$$M < M_f = \left[\frac{1}{2}kLe\mathcal{R}_C(Le - \lambda) \right]^{1/2} - kLe. \quad (\text{A } 2)$$

In other words, any local minimum over N can be a local minimum over both M and N only if $M < M_f$, and in our search for a local minimum over (M, N) we can confine ourselves to the range $0 < M < M_f$. This immediately tells us that for $\mathcal{R}_C < 0$, the global minimum over N must occur at $N = \pi^2$. It also means that if $M_f < 0$, then the global minimum over N must occur at the boundary $N = \pi^2$ for any physical value of M ; we need therefore worry about a local minimum over N if and only if $M_f > 0$, i.e. if and only if

$$k < k_f = \frac{(Le - \lambda)}{2Le}\mathcal{R}_C. \quad (\text{A } 3)$$

Keeping N general, we now consider the variation of $R(M, N)$ with M . The condition $\partial R/\partial M = 0$ at $M = M_c$ leads to the quartic equation (3.18), so there can be at most four local extrema of $R(M, N)$. Looking at the limiting behaviour of $R(M, N)$ for small and large M , we find that there must be at least one local minimum over M in the (physical) range $M > 0$, and one local maximum over M in the (unphysical) range $M < 0$. With a little further effort, we can show that

$$\frac{\partial^2 R}{\partial M^2} = \frac{2N^2}{\lambda M^3} + \frac{2kLe(Le - \lambda)}{\lambda(kLe + M + N)^3}\mathcal{R}_C, \quad (\text{A } 4)$$

and so, at least for $\mathcal{R}_C > 0$, we cannot have more than a single local minimum in the range $M > 0$. Since our search for a global minimum has already restricted us to the range $\mathcal{R}_C > 0$, we know that there is exactly one minimum over M to be concerned about.

We now return to the question of optimising over N . If we can prove that the (unique) value of M which minimises R must always be greater than M_f (regardless of the value of N), then we will have shown that the minimum must correspond to $N = \pi^2$. This in turn will follow if we can prove that $\partial R/\partial M(M_f, N) < 0 \forall N > 0$.

The sign of $\partial R/\partial M$ is the same as the sign of $g(M, N)$, where $g(M, N)$ is the LHS of (3.18), multiplied by a factor of Le^2 . Substituting in $M = M_f$ and rearranging, we find that

$$g(M_f, N) = -P^4 - P^3 - \beta \left(1 - \frac{\beta}{4} \right) P^2 + \frac{1}{4}(1 - \beta)^2 P - \frac{1}{16}(1 - \beta)^2, \quad (\text{A } 5)$$

where

$$P = \frac{N}{\sqrt{2kLe(Le - \lambda)\mathcal{R}_C}} \quad \text{and} \quad \beta = \sqrt{\frac{2kLe}{\mathcal{R}_C(Le - \lambda)}}. \quad (\text{A } 6)$$

We note that $0 < \beta < 1$ for $k < k_f$, and it is simple to show by plotting $g(P; \beta)$ that there are no points in the range $P > 0$, $0 < \beta < 1$ at which $g \geq 0$.

We have now established that there is no local minimum over (M, N) within the range

$M > 0$, $N > 0$, and this, in conjunction with our earlier results, means that the global minimum over (M, N) must occur on the boundary $N = \pi^2$. We note, finally, that this is consistent with the condition that $N/\pi^2 \in \mathbb{N}$.

A.2. Asymptotics for $\mathcal{R}_T^{(c)}$ and M_c when $1/Le \ll 1$

A.2.1. Fast or moderately fast reactions, $k = \mathcal{O}(1)$

We will first consider the case where $1/Le \ll 1$ but the reaction rate k is of order unity. As we have seen, the limit $Le \rightarrow \infty$ is a singular limit of the equation (3.18), and it is possible to find three different leading-order balances. We will consider each of them in turn.

(a) $M = \mathcal{O}(1)$. We seek solutions of the form $M = M_0 + M_1/Le + \mathcal{O}(Le^{-2})$, where $M_0 > 0$. At leading order, (3.18) becomes

$$(k^2 - \mathcal{R}_C k) M_0^2 - k^2 \pi^4 = 0. \quad (\text{A } 7)$$

When $\mathcal{R}_C < k$, this has the unique positive solution $M_0 = \pi^2 \sqrt{k/(k - \mathcal{R}_C)}$. When $\mathcal{R}_C \geq k$, on the other hand, there is no positive solution for M_0 . This indicates that in the non-diffusive limit $1/Le = 0$, reaction can only stabilise the system at finite wavenumbers for $\mathcal{R}_C < k$, and for higher values of \mathcal{R}_C some diffusion must be necessary.

Considering the next-order terms in $1/Le$, we find that

$$M_1 = -\frac{\pi^2 \mathcal{R}_C}{2} \frac{\left(2\pi^2 \sqrt{\frac{k}{k - \mathcal{R}_C}} + 2\pi^2 + k\lambda\right)}{\sqrt{k}(k - \mathcal{R}_C)^{3/2}}, \quad (\text{A } 8)$$

and substituting the expansion for M into $\mathcal{R}_T^{(c)}$ and expanding in $1/Le$, we obtain

$$\mathcal{R}_T^{(c)} \sim \frac{\pi^2}{\lambda} \left(1 + \sqrt{\frac{k - \mathcal{R}_C}{k}}\right)^2 - \mathcal{R}_C + \frac{\mathcal{R}_T^{(c1)}}{Le} + \mathcal{O}(Le^{-2}), \quad (\text{A } 9)$$

where

$$\mathcal{R}_T^{(c1)} = -\left(\pi^2 + \frac{\pi^4}{\lambda k}\right) \frac{k - \mathcal{R}_C}{k} + \left(\pi^2 + \frac{2\pi^4}{\lambda k}\right) \frac{\mathcal{R}_C}{\sqrt{k}(k - \mathcal{R}_C)} + \pi^2 + \frac{\pi^4}{(k - \mathcal{R}_C)\lambda}. \quad (\text{A } 10)$$

We note that in the region of $\mathcal{R}_C = k$, this expansion breaks down since $\mathcal{R}_T^{(c1)}/Le$ is no longer small compared to the zero-order term.

(b) To investigate the irregular root, we have to take a little more care. We now seek solutions to (3.18) of the form $M = \mu Le^\alpha$ where $\mu = \mathcal{O}(1)$; substituting this in, we find that the only consistent dominant balance which includes the leading terms in M occurs when $\alpha = 1$, and at leading order we have

$$Le^2 \mu^4 + 2kLe^2 \mu^3 + (k^2 - \mathcal{R}_C k) Le^2 \mu^2 + \mathcal{O}(Le) = 0. \quad (\text{A } 11)$$

The positive solution is easy to obtain: we find

$$\mu \sim -k + \sqrt{\mathcal{R}_C k}, \quad \text{i.e.} \quad M \sim \sqrt{k}(\sqrt{\mathcal{R}_C} - \sqrt{k})Le, \quad (\text{A } 12)$$

with the condition that $\mathcal{R}_C > k$ in order that $M > 0$. (This result is independent of the vertical wavenumber π^2 , because it represents a short-wavelength mode in which lateral rather than vertical diffusion is dominant.) Seeking an expansion in powers of Le^{-1} with

this leading-order term, we find

$$M = Le \left(-k + \sqrt{\mathcal{R}_C k} - \left[\pi^2 + \frac{\lambda \sqrt{\mathcal{R}_C k}}{2} \right] \frac{1}{Le} + \mathcal{O}(Le^{-2}) \right), \quad (\text{A } 13)$$

and substituting this expansion into $\mathcal{R}_T^{(c)}$ yields

$$\mathcal{R}_T^{(c)} \sim -\frac{(\sqrt{\mathcal{R}_C} - \sqrt{k})^2 Le}{\lambda} + \frac{\pi^2}{\lambda} - \sqrt{\mathcal{R}_C k} + \mathcal{O}(Le^{-1}). \quad (\text{A } 14)$$

(c) The final case occurs when $k - \mathcal{R}_C = \mathcal{R}/Le$, for some $\mathcal{R} = \mathcal{O}(1)$. Substituting this into (3.18), and again seeking solutions of the form $M = Le^\alpha \mu$, we find that the only consistent balance for $\mu > 0$ occurs when $\alpha = 1/3$ and

$$2k\mu^3 - k^2\pi^4 + \mathcal{O}(Le^{-1/3}) = 0. \quad (\text{A } 15)$$

The leading term in the solution is then given by $M = (Le k \pi^4 / 2)^{1/3}$. To find the next-order term, we substitute into (3.18) the expansions $\mathcal{R}_C = k - \mathcal{R}/Le$ and $M = Le^{1/3}(M_0 + Le^{-1/3}M_1 + \mathcal{O}(Le^{-2/3}))$, obtaining

$$M = \left(\frac{Le k \pi^4}{2} \right)^{1/3} - \frac{\mathcal{R} + k\lambda + 2\pi^2}{3} + \mathcal{O}(Le^{-1/3}) \quad (\text{A } 16)$$

and thus

$$\mathcal{R}_T^{(c)} = \frac{\pi^2}{\lambda} - k + \frac{3}{\lambda} \left(\frac{\pi^8}{4k} \right)^{1/3} \frac{1}{Le^{1/3}} + \mathcal{O}(Le^{-2/3}). \quad (\text{A } 17)$$

A.2.2. Slow reactions $k = \mathcal{O}(1/Le)$

It is also natural to investigate the limit of a large Lewis number and a slow reaction rate. We may express this formally as $0 < k \sim 1/Le \ll 1$, so we set $k = K/Le$ where $K = \mathcal{O}(1)$, and consider the regime $1/Le \ll 1$. As before, we consider the two criteria $b < 0$ and $c < 0$.

(i) The condition $b < 0$ is given as before by equation (3.16). Setting $k = K/Le$ and expanding for $1/Le \ll 1$, we find

$$\mathcal{R}_T^{(b)} = \frac{4\pi^2}{\lambda} - \frac{\mathcal{R}_C}{\lambda} + \frac{2(2\pi^2 + K)}{\lambda} \frac{1}{Le} - \frac{K^2}{4\pi^2\lambda} \frac{1}{Le^2} + \mathcal{O}(Le^{-3}), \quad (\text{A } 18)$$

so the two stabilising terms combine to raise slightly the threshold of instability.

(ii) To tackle the condition $c < 0$, we investigate equation (3.18). Substituting in $k = K/Le$, we find that there are two possible consistent scalings for M_c , depending on the sign of \mathcal{R}_C .

(a) If $\mathcal{R}_C > 0$, the only consistent scaling is $M_c \sim Le^{1/2}$, and we then obtain

$$M_c = Le^{1/2} \left[\sqrt{\mathcal{R}_C K} - \frac{K + \pi^2}{\sqrt{Le}} + \frac{\pi^4 - \mathcal{R}_C K \lambda}{2\sqrt{\mathcal{R}_C K}} \frac{1}{Le} + \mathcal{O}(Le^{-3/2}) \right]. \quad (\text{A } 19)$$

It follows that

$$\mathcal{R}_T^{(c)} = -\frac{\mathcal{R}_C Le}{\lambda} + \frac{2\sqrt{\mathcal{R}_C K} \sqrt{Le}}{\lambda} + \frac{\pi^2 - K}{\lambda} + \mathcal{O}(Le^{-1/2}). \quad (\text{A } 20)$$

The weak reaction term now does not alter the leading-order result from pure DDC, and cannot stabilise a solutally unstable system.

(b) If $\mathcal{R}_C < 0$, the only consistent scaling is $M_c \sim Le^{-1/2}$ (very long waves), and the corresponding value of $\mathcal{R}_T^{(c)}$ scales as $\mathcal{R}_T^{(c)} \sim -Le\mathcal{R}_C$. This boundary lies above the boundary $\mathcal{R}_T = \mathcal{R}_T^{(b)}$ and so is physically irrelevant.

(c) The results above suggest that we should consider separately the regime in which $\mathcal{R}_C \sim 1/Le$. However, this regime does not offer a significant simplification of equation (3.18), so we do not pursue this analysis here.

A.3. Properties of the boundary $B^2 - 4AC = 0$ when $k > 0$

In the reactive problem, $k > 0$, the condition $B^2 - 4AC = 0$ becomes

$$(X+Y)^2 + (\delta_{X2}k^2 + \delta_{X1}k + 2\delta)X + (\delta_{Y2}k^2 + \delta_{Y1}k - 2\delta)Y + (\delta_4k^4 + \delta_3k^3 + \delta_2k^2 + \delta_1k + \delta^2) = 0, \quad (\text{A 21})$$

where we have defined $X = \lambda\mathcal{R}_C$ and $Y = Le\mathcal{R}_T$ and δ as before, and where

$$\begin{aligned} \delta_{X2} &= (2Le - \lambda) Le \frac{(m^2 + \pi^2)}{2\pi^2 m^2} > 0, \\ \delta_{X1} &= (2Le [4\pi^2 \lambda + Le(m^2 + \pi^2)] - \lambda^2(m^2 + 5\pi^2)) \frac{(m^2 + \pi^2)}{2\pi^2 m^2 \lambda} > 0, \\ \delta_{Y2} &= \lambda Le \frac{(m^2 + \pi^2)}{2\pi^2 m^2} > 0, \quad \delta_{Y1} = (m^2 + 5\pi^2) \lambda \frac{(m^2 + \pi^2)}{2\pi^2 m^2} > 0, \\ \delta_4 &= Le^2 \lambda^2 \frac{(m^2 + \pi^2)^2}{16\pi^4 m^4} > 0, \quad \delta_3 = (m^2 + 5\pi^2) \lambda^2 Le \frac{(m^2 + \pi^2)^2}{8\pi^4 m^4} > 0, \\ \delta_2 &= (\lambda^2(m^4 + 18\pi^2 m^2 + 33\pi^4) - 8\pi^2 Le^2(m^2 + \pi^2)) \frac{(m^2 + \pi^2)^2}{16\pi^4 m^4} \geq 0, \\ \text{and} \quad \delta_1 &= (\lambda^2 - Le^2)(m^2 + 5\pi^2) \frac{(m^2 + \pi^2)^3}{2\pi^2 m^4 Le} < 0. \end{aligned} \quad (\text{A 22})$$

The form of the quadratic terms indicates that this curve always represents either a parabola or a pair of parallel lines. Numerical experimentation further indicates that the boundary remains a non-degenerate parabola for $0 < \lambda < 1$ and $Le > 1$, regardless of the value of k .

It can be shown (we omit the algebra here for simplicity) that when $X = 0$, equation (A 21) has a single solution for \mathcal{R}_T : in other words, the parabola is tangent to the \mathcal{R}_T -axis (as it was in the case $k = 0$). However, the parabola is no longer tangent to the \mathcal{R}_C -axis. We can also show that the intercept of $\mathcal{R}_T^{(B)}$ and $\mathcal{R}_T^{(C)}$ always occurs for some value of $\mathcal{R}_C < 0$ regardless of the value of k . As before, the parabola $B^2 - 4AC$ must pass through this point, and it is the only point at which it can intersect either line. There are only two ways in which a parabola can have a single intersection with a straight line: either the line is parallel to the symmetry axis of the parabola, or the parabola is tangent to the line at the point of intersection. Since the gradient of the tangent is unique, and the gradient of the symmetry axis is unique, we conclude that one of $(\mathcal{R}_T^{(B)}, \mathcal{R}_T^{(C)})$ must be the symmetry axis of the parabola and the other must be tangent to it; we already know which is which for $k = 0$, and assuming smooth variation we conclude that as before, $\mathcal{R}_T^{(B)}$ is parallel to the symmetry axis and $\mathcal{R}_T^{(C)}$ is tangent to the parabola. Numerical experimentation confirms this deduction.

REFERENCES

- BDZIL, J. & FRISCH, H. L. 1971 Chemical instabilities. II. Chemical surface reactions and hydrodynamic instability. *Phys. Fluids* **14** (3), 475–482.
 BDZIL, J. B. & FRISCH, H. L. 1980 Chemically driven convection. *J. Chem. Phys.* **72** (3), 1875–1886.

- BEAR, J. 1972 *Dynamics of Fluids in Porous Media*. Dover Publications.
- BODENSCHATZ, E., PESCH, W. & AHLERS, G. 2000 Recent developments in Rayleigh–Bénard convection. *Annu. Rev. Fluid Mech.* **32**, 709–778.
- DRAZIN, P. G. 2002 *Introduction to Hydrodynamic Stability*. Cambridge University Press.
- GATICA, J. E., VILJOEN, H. J. & HLAVACEK, V. 1989 Interaction between chemical reaction and natural convection in porous media. *Chem. Engng. Sci.* **44** (9), 1853–1870.
- GILMAN, A. & BEAR, J. 1994 The influence of free convection on soil salinization in arid regions. *Transport Porous Media* **23**, 275–301.
- GUTKOWICZ-KRUSIN, D. & ROSS, J. 1980 Rayleigh–Bénard instability in reactive binary fluids. *J. Chem. Phys.* **72** (6), 3577–3587.
- HORTON, C. W. & ROGERS, F. T. 1945 Convection currents in a porous medium. *J. Appl. Phys.* **16**, 367–370.
- JUPP, T. E. & WOODS, A. W. 2003 Thermally-driven reaction fronts in porous media. *J. Fluid Mech.* **484**, 329–346.
- KAUFMAN, J. 1994 Numerical models of fluid flow in carbonate platforms: implications for dolomitization. *J. Sed. Res. A* **64**, 128–139.
- LAPWOOD, E. R. 1948 Convection of a fluid in a porous medium. *Proc. Camb. Phil. Soc.* **44**, 508–521.
- MAMOU, M. & VASSEUR, P. 1999 Thermosolutal bifurcation phenomena in porous enclosures subject to vertical temperature and concentration gradients. *J. Fluid Mech.* **395**, 61–87.
- MURRAY, B. T. & CHEN, C. F. 1989 Double-diffusive convection in a porous medium. *J. Fluid Mech.* **201**, 147–166.
- NIELD, D. A. 1968 Onset of thermohaline convection in a porous medium. *Water Resources Res.* **4** (3), 553–560.
- OLDENBURG, C. M. & PRUESS, K. 1998 Layered thermohaline convection in hypersaline geothermal systems. *Transport Porous Media* **33**, 29–63.
- PALM, E., WEBER, J. E. & KVERNOLD, O. 1972 On steady convection in a porous medium. *J. Fluid Mech.* **54** (1), 153–161.
- PHILLIPS, O. M. 1991 *Flow and Reactions in Permeable Rocks*. Cambridge University Press.
- RAFFENSPERGER, J. P. & GARVEN, G. 1995a The formation of unconformity-type uranium ore deposits. 1. Coupled groundwater flow and heat transport modelling. *Am. J. Sci.* **295** (5), 581–636.
- RAFFENSPERGER, J. P. & GARVEN, G. 1995b The formation of unconformity-type uranium ore deposits. 2. Coupled hydrochemical modelling. *Am. J. Sci.* **295** (6), 639–696.
- RUDRAIAH, N., SIDDHESHWAR, P. G. & MASUOKA, T. 2003 Nonlinear convection in porous media: a review. *J. Porous Med.* **6** (1), 1–32.
- RUDRAIAH, N., SRIMANI, P. K. & FRIEDRICH, R. 1982 Finite amplitude convection in a two-component fluid saturated porous layer. *Int. J. Heat Mass Transfer* **25** (5), 715–722.
- SCHOofs, S. 1999 Thermochemical convection in porous media: an application to hydrothermal systems and magmatic intrusions. PhD thesis, Universiteit Utrecht.
- SCHOofs, S. & SPERA, F. J. 2003 Transition to chaos and flow dynamics of thermochemical porous medium convection. *Transport Porous Media* **50**, 179–195.
- SPIEGELMAN, M. & KATZ, R. F. 2006 A semi-Lagrangian Crank-Nicolson algorithm for the numerical solution of advection-diffusion problems. *Geochem. Geophys. Geosystems* **7**, Q04014, doi: 10.1029/2005GC001073.
- STEINBERG, V. & BRAND, H. 1983 Convective instabilities of binary mixtures with fast chemical reaction in a porous medium. *J. Chem. Phys.* **78** (5), 2655–2660.
- STEINBERG, V. & BRAND, H. R. 1984 Amplitude equations for the onset of convection in a reactive mixture in a porous medium. *J. Chem. Phys.* **80** (1), 431–435.
- TURNER, J. S. 1979 *Buoyancy Effects in Fluids*. Cambridge University Press.
- TURNER, J. S. 1985 Multicomponent convection. *Annu. Rev. Fluid Mech.* **17**, 11–44.
- VILJOEN, H. J., GATICA, J. E. & HLAVACEK, V. 1990 Bifurcation analysis of chemically driven convection. *Chem. Engng. Sci.* **45** (2), 503–517.
- WOLLKIND, D. J. & FRISCH, H.L. 1971a Chemical instabilities: I. A heated horizontal layer of dissociating fluid. *Phys. Fluids* **14** (1), 13–18.

- WOLLKIND, D. J. & FRISCH, H. L. 1971*b* Chemical instabilities. III. Nonlinear stability analysis of a heated horizontal layer of dissociating fluid. *Phys. Fluids* **14** (3), 482–487.
- WOODING, R. A., TYLER, S. W. & WHITE, I. 1997 Convection in groundwater below an evaporating salt lake: 1. Onset of instability. *Water Resources Res.* **33** (6), 1199–1217.