The relationship between double-diffusive intrusions and staircases in the Arctic Ocean

YANA BEBIeva ∗ AND MARY-LOUISE TIMMERMANS †

Department of Geology and Geophysics, Yale University, 210 Whitney Ave, New Haven CT 06511, USA.

ABSTRACT

The origin of double-diffusive staircases in the Arctic Ocean is investigated for the particular background setting in which both temperature and salinity increase with depth. Motivated by observations that show the co-existence of thermohaline intrusions and double-diffusive staircases, a linear stability analysis is performed on the governing equations to determine the conditions under which staircases form. It is shown that a double-diffusive staircase can result from interleaving motions if the observed bulk vertical density ratio is below a critical vertical density ratio estimated for particular lateral and vertical background temperature and salinity gradients. Vertical temperature and salinity gradients dominate over horizontal gradients in determining whether staircases form. Examination of Arctic Ocean temperature and salinity measurements indicates that observations are consistent with the theory for reasonable choices of eddy diffusivity and viscosity.

1. Introduction

The Arctic Ocean has a strong halocline and deeper water layers that are warmer than those at the surface in contact with sea-ice cover (e.g., Aagaard et al. 1981). Understanding the mechanisms and magnitude of upward fluxes of deep ocean heat is essential to predictions of Arctic sea ice and climate (e.g., Maykut and Untersteiner 1971; Wettlaufer 1991; Perovich et al. 2008; Carmack et al. 2015; Timmermans 2015). Relatively cold and fresh surface waters, originating from net precipitation, river runoff, inflows from the Pacific Ocean and seasonal sea-ice melt, occupy the upper ~150–200 m water column (e.g., Steele et al. 2008; Timmermans et al. 2014). Below these upper layers lie relatively warm and salty waters associated with Atlantic Water (AW) inflows, centered around 400 m depth in the Arctic Ocean’s Canada Basin, with lateral temperature gradients indicating a general cooling moving east from the warm core of the AW layer on the western side of the basin (Figure 1a).

The basic vertical stratification, in which temperature and salinity both increase with depth, provides conditions amenable to double-diffusive instability, believed to be a key physical process generating thermohaline intrusions and staircases in the Arctic Ocean. Throughout much of the central Arctic Ocean basins, heat transfer from the AW layer is via double-diffusive convection (e.g., Melling et al. 1984; Padman and Dillon 1987, 1988; Timmermans et al. 2008; Polyakov et al. 2012; Sirevaag and Fer 2012; Guthrie et al. 2015). Two types of double-diffusive convection can arise in a stably-stratified ocean: the case when both temperature \( T \) and salinity \( S \) increase with depth is referred to as diffusive convection (the DC regime, an overview is given by Kelley et al. 2003), while the case when both temperature and salinity decrease with depth is referred to as the salt-finger (SF) regime (for overviews see, e.g., Kunze 2003; Schmitt 2003). These stratifications may be characterized by a density ratio, which we define here as 
\[
\frac{\bar{\rho}}{\rho_0} = \frac{\partial \rho}{\partial S} / \alpha_T,
\]
where the subscript \( \bar{z} \) denotes vertical gradient, and \( \beta = (1/\rho_0) (\partial \rho / \partial S)_{T, p} \) and \( \alpha = -(1/\rho_0) (\partial \rho / \partial T)_{S, p} \) (where \( \rho_0 \) is a reference density) are the saline contraction and thermal expansion coefficients, respectively. The overlines indicate a bulk gradient, taken to be linear over some specified depth range.

At the top boundary of the AW layer in the Canada Basin, a prevalent DC staircase is characterized by a sequence of mixed layers of thickness on the order of several meters separated by sharp gradients in temperature and salinity (Neal et al. 1969; Padman and Dillon 1987; Timmermans et al. 2008) (Figure 1b,c). Thermohaline intrusions, characterized by DC and SF regions alternating in depth, are often found underlying the staircase (e.g., Carmack et al. 1998; Merryfield 2002; Woodgate et al. 2007). Intrusions are believed to be associated with lateral (in addition to vertical) gradients in temperature and salinity, and are driven partly by vertical buoyancy flux di-
vergences (overviews of thermohaline intrusions are given by Ruddick and Kerr 2003; Ruddick and Richards 2003). Understanding and predicting the observed vertical temperature-salinity structure (whether staircases or intrusions), and associated vertical and lateral heat fluxes from the AW layer, requires knowledge of the origins of these features and their relationship to each other (see Kelley 2001). Extensive efforts have been made to explain staircase and intrusion origins and evolution with respect to the SF configuration of double diffusion. There exist around six theories for the origin of SF staircases, as reviewed by Radko (2013). One of these theories is that interleaving motions can develop into a staircase (Merryfield 2000). The idea relies on the presence of lateral temperature and salinity gradients and builds on previous studies that invoke a standard parametric flux model (Walsh and Ruddick 1995, 2000). This model was first introduced by Stern (1967), who delineated three separate scales of motion: “small” to describe double-diffusive processes on centimeter to meter scales, “medium” to describe the scales of interleaving (order tens of meters vertically and 1 km laterally), and “large” to characterize the background state (the full thickness of the double-diffusive thermocline and basin scales laterally). The main assumption here is that medium-scale dynamics are qualitatively similar to small-scale dynamics and that the effects of double diffusion on medium-scale vertical fluxes can be parameterized in terms of effective diffusivities and large-scale vertical gradients. Merryfield (2000) applied this formalism in his calculations showing that interleaving motions evolve into SF staircases when the density ratio of the background (SF-stratified) state is below a certain value (i.e., there exists a critical density ratio that delineates the boundary between staircase and intrusion formation).

Very little has been done with respect to analysis of staircase origins in the DC-stratified setting. The purpose of this study is to address the relationships between and origins of DC staircases and intrusions, with application to the Arctic Ocean. Motivated by observations which show staircases and intrusions co-existing and evolving, we assess the main factors that determine whether staircases or intrusions will be observed. Perhaps the major limitation on developing a mechanism for staircase formation in the Arctic Ocean has been considered to relate to the fact that the magnitude of the observed vertical density ratio ($\frac{\rho_F}{\rho_0}$) falls outside the DC-unstable range ($1 < \frac{\rho_F}{\rho_0} < 1.1$) derived by linear stability analysis of the Boussinesq equations (Vernon 1965). However, this derivation considers only small-scale dynamics (i.e., molecular diffusivities are used in the computation of fluxes), and no horizontal background gradients in temperature and salinity. Here, we invoke the parametric flux model (as used by Merryfield (2000) for the SF case) to determine a critical vertical density ratio $\frac{\rho_F}{\rho_0}$ for the DC stratification at which a transition between staircases and intrusions occurs for the observed vertical and lateral temperature gradients in the Arctic Ocean. In this formalism, instability does not require $\frac{\rho_F}{\rho_0}$ to be bounded by 1.1.

The paper is organized as follows. In the next section, we formulate the governing equations (section 2a) and then determine a constraint on the growth rate for growing perturbations to evolve towards a DC staircase (section 2b). Next, in section 2c, we perform a linear stability analysis on the governing Boussinesq equations to determine the fastest growing mode as a function of vertical and lateral temperature and salinity gradients. Together with the result from section 2b, this allows us to define a critical vertical density ratio $\frac{\rho_F}{\rho_0}$ above which intrusions form, and below which staircases form. In section 3, we apply the theory to Ice-Tethered Profiler (ITP; Krishfield et al. (2008); Toole et al. (2011)) measurements of temperature and salinity through the Canada Basin double-diffusive thermocline, which features both a staircase and intrusions. We examine ITP temperature and salinity profiles to show that the observations are consistent with the theory for reasonable choices of appropriate parameters. Findings are summarized and discussed in section 4.

2. Theory

a. Formulation of the governing equations

We begin by formulating the governing equations accounting for three scales of motion as described in section 1. The governing set of Boussinesq equations (2D: horizontal and vertical dimensions) is as follows

$$
\begin{align*}
\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z} &= -\frac{\partial p}{\partial x} + \nu \frac{\partial^2 u}{\partial x^2} + \alpha \frac{\partial^2 u}{\partial z^2} \\
\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + w \frac{\partial w}{\partial z} &= -\frac{\partial p}{\partial z} + g \frac{\partial \rho}{\partial z} + \nu \frac{\partial^2 w}{\partial z^2} + \frac{\partial T}{\partial z} \\
u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z} &= 0 \\
T_t + u T_x + w T_z &= \kappa_T \nabla^2 T \\
S_t + u S_x + w S_z &= \kappa_S \nabla^2 S \\
\rho &= \rho_0 (1 - \alpha (T - T_0) + \beta (S - S_0)),
\end{align*}
$$

where $u$ and $w$ are horizontal and vertical components of velocity, $p$ is pressure, $\kappa_T$ and $\kappa_S$ are the molecular coefficients of heat and salt diffusion respectively, $\nu$ is the molecular kinematic viscosity, and $T_0$ and $S_0$ are reference values of temperature and salinity (corresponding to density $\rho_0$). We neglect rotation because we are primarily interested in the fastest growth rate; with respect to the SF case, Kerr and Holyer (1986) have shown that the structure of the fastest-growing mode of perturbation (i.e., vertical and horizontal wave numbers along with the growth rate) is unaffected by rotation. This set of equations consists of two momentum equations, the continuity equation, conservation equations for $T$ and $S$ and the equation of state.

We proceed by decomposing variables into mean large scale values (basin-wide scales, denoted with overbars),...
medium scales (intrusions, denoted with tildes) and small scales (double-diffusive mixing, denoted by primes), for example, $u = \bar{u} + \tilde{u} + \dot{u}$, $T = \bar{T} + \tilde{T} + T'$ etc.; this is similar to the separation of scales under Reynolds averaging (Reynolds 1894). All equations are next averaged over small scales associated with double-diffusive mixing on a time period that is sufficiently long to smooth transient fluctuations, yet short enough to retain the slow evolution of the interleaving motion (operating on medium scales). Further, we assume that the background state is motionless (i.e. $\bar{u} = 0$ and $\bar{w} = 0$). Note also that the background temperature and salinity structure does not change on the timescales of medium motion, so that $\partial \bar{T} / \partial t = 0$ and $\partial \bar{S} / \partial t = 0$; in section 3b, we show this to be an appropriate approximation for the setting considered here. After time averaging (denoted by angle brackets, and note that time averaging of any medium or large scale variable is equal to that variable: e.g., $\langle \tilde{u} \rangle = \bar{u}$) the governing equations become

\[
\begin{align*}
\bar{u}_t + \bar{u} \bar{u}_x + \bar{w} \bar{w}_z &= -\bar{p}_x / \rho_0 + \nu \nabla^2 \bar{u} \\
\bar{w}_t + \bar{u} \bar{w}_x + \bar{w} \bar{w}_z &= -\bar{p}_z / \rho_0 - g \bar{\rho} / \rho_0 + \nu \nabla^2 \bar{w} \\
\bar{u}_x + \bar{w}_z &= 0
\end{align*}
\]

The last two terms (Reynolds stresses) on the right hand side of the two momentum, heat and salt equations represent the effects of fluctuations (associated with double-diffusive mixing) on interleaving motions. Horizontal divergences can be neglected with respect to vertical divergences since interleaving motions have small aspect ratio.

Analogous to Reynolds averaging, we combine Reynolds stresses and frictional terms to define vertical eddy viscosity ($A$) and vertical eddy diffusivities for heat ($K_T$) and salt ($K_S$) as follows

\[
\begin{align*}
\nu \bar{u}_z - \langle \tilde{u}' \dot{w}' \rangle &= A \bar{u}_z \\
\nu \bar{w}_z - \langle \dot{w}' \dot{w}' \rangle &= A \bar{w}_z \\
\kappa_T (\bar{T} + \tilde{T})_z - \langle \dot{w}' T' \rangle &= K_T \bar{T}_z \\
\kappa_S (\bar{S} + \tilde{S})_z - \langle \dot{w}' S' \rangle &= K_S \bar{S}_z.
\end{align*}
\]

The resulting linearized system of equations describes motions on the medium scale:

\[
\begin{align*}
\bar{u}_t + \bar{p}_x / \rho_0 + F'_u &= 0 \\
\bar{w}_t + \bar{p}_z / \rho_0 + g \bar{\rho} / \rho_0 + F'_w &= 0 \\
\bar{u}_x + \bar{w}_z &= 0 \\
\bar{T}_t + \bar{u} \bar{T}_x + \bar{w} \bar{T}_z + F'_T &= 0 \\
\bar{S}_t + \bar{u} \bar{S}_x + \bar{w} \bar{S}_z + F'_S &= 0,
\end{align*}
\]

where vertical fluxes of horizontal ($F'_u$) and vertical ($F'_w$) momentum as well as vertical fluxes of heat ($F'_T$) and salt ($F'_S$) are defined as

\[
\begin{align*}
F'_u &= -A \bar{u}_z, \\
F'_w &= -A \bar{w}_z, \\
F'_T &= -K_T \bar{T}_z, \\
F'_S &= -K_S \bar{S}_z.
\end{align*}
\]

The salt flux may be expressed via $F'_T$ and the ratio of the density flux of salt to the density flux of heat (i.e., the flux ratio $R_F = \beta F'_S / (\alpha F'_T)$) as

\[
F'_S = R_F \frac{\alpha}{\beta} F'_T.
\]

For the DC-type of double-diffusive convection it has been shown experimentally that $R_F \approx 0.15$ when $\overline{K_T} \gtrsim 2$ (Turner 1965). Here we take $R_F = 0.15$ for the medium scale motions under the assumption that fluxes on medium-scales can be parameterized as double-diffusive mixing on medium scales (see Stern 1967) and that the DC fluxes dominate (cf. Toole and Georgi 1981 and Walsh and Ruddick (1995, 2000) who followed the same reasoning for the SF case).

1) ESTIMATES OF $K_T$ AND $A$

Eddy diffusivity $K_T$ parameterizes the effects of double-diffusive mixing (equation 5) on medium scales. In the staircase region $K_T$ may be estimated using a double-diffusive heat flux parameterization based on a 4/3-flux law (Kelley 1990) and the typical background vertical temperature gradient; Guthrie et al. (2013) show this 4/3-flux law to be a reasonable representation of the fluxes. This yields $K_T = O(10^{-6}) \text{m}^2\text{s}^{-1}$ for typical double-diffusive fluxes in the Canada Basin around $0.1 - 0.2 \text{W/m}^2$ (Padman and Dillon 1987; Timmermans et al. 2008). This estimate for $K_T$ is consistent with estimates by Guthrie et al. (2013) in the Canada Basin in the region of the staircase, where they find $K_T = 1 \times 10^{-6} - 5 \times 10^{-6} \text{m}^2\text{s}^{-1}$.

Eddy viscosity must be formulated to represent momentum transfer on medium scales due to double-diffusive processes. In the theoretical framework, we consider growing perturbations in the absence of large-scale background shear. As perturbations grow, however, small velocities arise associated with interleaving. We assume that
interleaving motions achieve a buoyancy-friction balance (McDougall 1985) some time after the onset of perturbations in the absence of large-scale shear and pressure work (Tennekes and Lumley 1972). An effective viscosity can then be estimated using the kinetic energy balance (assuming that horizontal shear is the dominant term in the rate of viscous dissipation of turbulent kinetic energy) as follows

$$\text{Au}_{zz} = g \alpha T_e (1 - R_F) \Rightarrow A = \frac{g \alpha K_T T_e (1 - R_F) H^2}{2U_0},$$

where $u_{zz}$ is scaled as $2U_0/H^2$, and typical interleaving velocity along a layer (of characteristic height $H = O(1)$ m) is $U_0 = O(1)$ mm/s (Ruddick and Hebert 1988; Walsh and Carmack 2003). Our estimated value for $K_T$ together with the typical mean vertical temperature gradient in our observational setting ($T_e \approx 0.01^\circ$C/m) yields $A = O(10^{-8})$ m$^{-2}$s$^{-1}$. This estimate of eddy viscosity is smaller than molecular viscosity, which may not be surprising in the context of pure double-diffusive mixing. For example, for the SF case, Ruddick (1985) estimated eddy viscosities in the range $10^{-9} - 10^{-14}$ m$^{-2}$s$^{-1}$, where he attributed these small values to the small lateral velocities between mixed layers in an SF staircase. Here we consider shear generated only due to interleaving motions originating from a motionless background state (due to convergences and divergences of double-diffusive fluxes). This is principally different from the case where momentum transport is considered through thermohaline steps in the presence of large-scale background shear (e.g., Padman 1994). Finally, the estimates for $A$ are consistent with those found in a laboratory study on double-diffusive interleaving: Krishnamurti (2006) showed that the ratio of Reynolds stress momentum fluxes to viscous stress momentum fluxes (i.e., $\langle u'w' \rangle / \nu \tilde{u}$, in our notation) can vary from 0.1 to 1 within the field of interleaving motion. Following (3), this yields a broad range for $A$ with variations by an order of magnitude ($10^{-8}$ to $10^{-7}$ m$^{-2}$s$^{-1}$).

\[ \text{b. Constraint leading to a staircase} \]

In order to derive a constraint on the growth rate that must be satisfied for a perturbation to evolve to a staircase, we consider wave-like solutions of the system (7)–(11). We write temperature and salinity fields as the sum of linear background gradients plus deviations (resulting from interleaving motions) in the form of a plane-wave as follows:

$$T(x, z, t) = T_{zz} + T_{x}x + T = T_{zz} + T_{x}x + T e^{ikx + imz + \lambda t}$$

$$S(x, z, t) = S_{zz} + S_{x}x + S = S_{zz} + S_{x}x + S e^{ikx + imz + \lambda t},$$

where $k$ and $m$ are horizontal and vertical wavenumbers respectively, $\lambda$ is growth rate, and $T$ and $S$ are the amplitudes of the temperature and salinity perturbations. Merryfield (2000) provides an effective explanation for the possible outcomes of a perturbation given initial conditions of an SF stratification; we follow the same reasoning for the DC case. At a particular $x = x_0$ the evolution of temperature and salinity can be illustrated on the $(\alpha T, \beta S)$ plane (Figure 2a). Initially, the amplitudes of the perturbations are small and the $T$ and $S$ fields are represented by a single point on the plane. A growing perturbation is a line segment with slope $\beta S/\alpha T$. If this slope is smaller than the vertical density ratio $R_F$ characterizing the background stratification (i.e., large scale) from which a perturbation begins to grow, then the value of $\beta S/\alpha T$ increases in time on one end of the segment and decreases on the other end (Figure 2a, line denoted $I$ for intrusion). An increase of $\beta S/\alpha T$ leads to a doubly-stable stratification (i.e., the vertical temperature gradient reverses sign to decreasing temperature with depth, while the salinity gradient remains stable), with further increases leading to an SF-unstable stratification when the salinity gradient also reverses sign. The other end of the line segment $I$ represents DC-favorable conditions with decreasing $\beta S/\alpha T$ towards unity, where convective instability takes place and a mixed layer is formed. Therefore, the evolution along line $I$ corresponds to the development of intrusions consisting of alternating mixed layers, DC-gradients, stable regions, and SF-gradients. The second scenario (depicted by the line $S$ for staircase in Figure 2a) arises when the evolution of $T$ and $S$ gradients are on a slope that is larger than $R_F$. In this scenario, as the perturbation grows, $\beta S/\alpha T$ decreases on one end of the segment (towards $\beta S/\alpha T = 1$ and the convectively unstable region) and increases on the other end of the segment in the DC-unstable region of the $(\alpha T, \beta S)$ plane (Figure 2a, $S$). The growing perturbations result in a series of convectively unstable mixed layers and DC gradients, i.e., a DC-type staircase. Whether a perturbation evolves to intrusions (along line $I$) or to a DC-staircase (along line $S$) depends upon the growthrate and wavelengths of the perturbations, a function of the initial background temperature and salinity gradients. The implicit assumption here is that the evolution of temperature and salinity is in the linear phase (i.e., perturbations develop and grow exponentially, and retain the same plane-wave spatial dependence as they grow from infinitesimally small perturbations).

We begin by examining the case in which perturbations to the linear background gradient evolve into a staircase. That is,

$$\frac{\beta S}{\alpha T} > R_F.$$  

(17)
Combining (15)–(17) and introducing a stream function \( \tilde{\psi} = \Psi e^{ikx + imz + \lambda t} \) (where \( \Psi \) is the amplitude) such that
\[
\tilde{\psi}_x = \tilde{\psi}_z = 0,
\]
we obtain the following criterion that the growth rate \( \lambda \) must satisfy
\[
\lambda < \frac{m^2K_T}{R_p - 1} \left[ 1 - R_F - \frac{s}{\Gamma} (R_p - R_F) \right],
\]
in order for a perturbation to evolve to a staircase (from criterion 17). Here, \( \Gamma = \frac{T_s}{T_c} \) and \( s = k/m \) (the slope of the growing perturbations). We have assumed that the aspect ratio of the background state is always larger than the aspect ratio of the perturbations (i.e., \( \Gamma/(\Gamma - s) > 0 \)), which is always satisfied for the observed values of \( T_s \) and \( T_c \) (see section 2d).

We next perform a linear stability analysis (see e.g., Toole and Georgi 1981; Walsh and Ruddick 1995, 2000) on the system (7)–(11) to compute the most unstable mode (i.e., maximal \( \lambda \) and corresponding \( k \) and \( m \)) for a specified \( T_s \), \( T_c \) and \( R_p \). The result is used together with the condition (19) to derive a critical density ratio \( R_p^c \) below which staircases are expected to be the end result of an interleaving perturbation.

c. Critical density ratio

To reduce the number of equations in the system, we combine (7)–(9) to yield an evolution equation for vorticity \( \nabla^2 \tilde{\psi} \)
\[
\nabla^2 \tilde{\psi}_t + g (\beta \tilde{S}_x - \alpha \tilde{T}_x) + F_{xx}^{m} - F_{zz}^{m} = 0.
\]
Horizontal background gradients in temperature and salinity are taken to be density compensating such that \( \alpha \tilde{T}_x = \beta \tilde{S}_x \). For application to the observations, we consider along-isopycnal gradients to avoid the effects of isopycnal heating (induced by mesoscale eddies and other dynamics) that do not lead to water mass transformation. In this way, we are effectively imposing \( \tilde{S}_x = 0 \) and as a consequence \( \alpha \tilde{T}_x = \beta \tilde{S}_x \). Using this fact, and the definition of \( \tilde{R}_p \), we express background salinity gradients in terms of background temperature gradients as
\[
\tilde{S}_x = \frac{\alpha T_x}{\beta},
\]
\[
\tilde{S}_z = \frac{R_p}{\alpha T_x / \beta}.
\]

Multiplying (10) by \( \alpha \) and (11) by \( \beta \), the final system of equations can be expressed as
\[
\nabla^2 \tilde{\psi}_t + g (\beta \tilde{S}_x - \alpha \tilde{T}_x) - A \nabla^2 \tilde{\psi}_{zz} = 0
\]
\[
\alpha \tilde{T}_x - \tilde{\psi}_x \alpha \tilde{T}_x + \tilde{\psi}_t \alpha \tilde{T}_x - \alpha K_T \tilde{T}_x = 0
\]
\[
\beta \tilde{S}_z - \tilde{\psi}_{zz} \alpha \tilde{T}_x + \tilde{\psi}_z \tilde{R}_p \alpha \tilde{T}_x - R_F \alpha K_T \tilde{T}_x = 0.
\]

Again assuming plane-wave solutions for the medium scale motions \( \tilde{\psi}, \tilde{T}, \tilde{S} = (\Psi, T, S) e^{ikx + imz + \lambda t} \), the system (23)–(25) reduces to
\[
\begin{pmatrix}
-(k^2 + m^2)(A + \lambda) & -igk & igk \\
\alpha (k T_s - m T_x) & \lambda + K_T m^2 & 0 \\
\alpha (R_p k T_s - m T_x) & R_F K_T m^2 & \lambda
\end{pmatrix}
\begin{pmatrix}
\Psi \\
\alpha \tilde{T} \\
\beta \tilde{S}
\end{pmatrix} = 0.
\]

A solution to this exists only if the determinant of the coefficient matrix is equal to zero, which gives
\[
\lambda^3 m^2 (s^2 + 1) + \lambda^2 m^4 (A + K_T)(s^2 + 1) + \lambda [AK_T m^6 (s^2 + 1) + \alpha \tilde{T}_x g^2 m^2 (1 - \tilde{R}_p)] + \alpha K_T g m^4 s (1 - R_F) - \tilde{T}_x (R_p - R_F) = 0.
\]

The growth rate \( \lambda_{\text{max}} \) of the fastest-growing mode (and corresponding values \( s_{\text{max}} \) and \( m_{\text{max}} \)) can be determined from (26) by applying a constrained optimization technique using the method of Lagrange multipliers (Bertsekas 2014), for a given \( T_s \) and \( T_c \). We wish to maximize the growth rate \( \lambda \), which may be expressed as a function of three variables \( f(\lambda, m, s) = \lambda \) subject to the constraint given by (26), which we denote as \( G(\lambda, m, s) = 0 \). In the method of Lagrange multipliers, we construct a system of equations that satisfy \( \nabla f(\lambda, m, s) = a \nabla G(\lambda, m, s) \) (where \( a \) is a constant) subject to \( G(\lambda, m, s) = 0 \). This yields four equations and four unknowns. Next, using solutions of the system (i.e., \( \lambda_{\text{max}}, m_{\text{max}}, s_{\text{max}} \)) we can obtain \( R_p^c \) from the criterion (19). That is, we solve the following equation
\[
\frac{\lambda_{\text{max}}}{m_{\text{max}}^2} = \frac{K_T}{R_p - 1} \left[ 1 - R_F - \frac{s}{\Gamma} (R_p - R_F) \right].
\]

It is instructive to explore how \( R_p^c \) is influenced by \( T_s \), \( T_c \) and \( R_p \) for the general ranges observed in the Canada Basin, and \( A \) and \( K_T \) values described in section 2a. The linear stability analysis together with (27) indicates \( R_p^c \) is insensitive to variations in \( T_s \) for some specified \( T_c \) and \( \tilde{S}_x \) (Figure 2b, red contours). This may be explained by the fact that the ratio of horizontal to vertical wavenumbers of the most unstable mode varies proportionally to the aspect ratio of the background temperature gradients. That is, \( s_{\text{max}}/\Gamma = \text{const} \) (\( \approx 0.1 \) for the range of \( T_c \) and \( T_s \) here). Furthermore, \( \lambda_{\text{max}}/m_{\text{max}}^2 \) also remains constant with varying \( T_c \), implying that (27) has the same coefficients over a range of \( T_c \), and as a consequence \( R_p^c \) remains unchanged.

While the growth rate is larger for larger \( T_c \), this does not affect whether a staircase or intrusions are the end result of a perturbation to a linear background stratification. The main influencing parameters in this regard are \( T_s \) and \( \tilde{S}_x \) (and \( R_p \)). When \( \tilde{S}_x \) is held fixed (e.g., \( \tilde{S}_x = -0.002 m^{-1} \)) for smaller magnitude \( T_s, \tilde{T}_x \), \( R_p^c \) increases faster than \( R_p^c \) (Figure 2b). For the weaker \( T_s \), it is more likely that \( R_p^c < R_p \),
implying intrusions are the end state of an interleaving perturbation. In the next section, we explore these relationships with respect to observed staircases and intrusions in the Arctic Ocean.

3. Context with ITP observations

As a consistency check, we aim to examine whether the presence of intrusions or staircases in the Arctic Ocean water column is commensurate with the linear theory described in the previous section. Water-column measurements are from Ice-Tethered Profilers (ITPs, Krishfield et al. 2008; Toole et al. 2011) that drifted in the Canada Basin (Figure 1a). ITPs are automated profiling instruments that provide measurements of temperature, salinity and depth from several meters depth beneath the sea ice, through the Atlantic Water layer to about 750 m depth. The final processed data for two ITP systems are analyzed here: ITP 1, operating between August 2005 and September 2006, and ITP 41, operating between October 2010 and January 2012. These ITPs were deployed in the same general location close to the Northwind Ridge, and drifted along similar trajectories across the Canada Basin (Figure 1a) over the same seasons, separated by five years. Measurements have a vertical resolution of about 25 cm, and a horizontal profile spacing of a few kilometers; full processing procedures are given by Krishfield et al. (2008). ITP data have been analyzed in several past studies of double diffusion in the Canada Basin (e.g., Timmermans et al. 2008; Radko et al. 2014; Bebiva and Timmermans 2016).

We will examine ITP profiles to compare the observed vertical density ratio $\bar{\rho}_p$ in depth regions exhibiting staircases and intrusions with the critical density ratio $\bar{\rho}_{cr}$ computed for the observed vertical and horizontal background temperature gradients and $\bar{\rho}_p$.

a. Quantifying the temperature and salinity gradients

Before applying the theory to the observations, we require some method to assess the bulk temperature and salinity gradients ($T_g$, $S_g$ and $T_s$, $S_s$) applicable for a given ITP profile.

Vertical gradients. To compute $T_g$ and $S_g$ (and therefore $\bar{\rho}_p$), we consider the depth range between a shallow bound of the AW layer (in the range around ~230–300 m depth, Figure 3, 4) to the depth of the AW temperature maximum around ~380–410 m. Vertical profiles of $T_g$ are constructed by first performing a cubic spline with a smoothing parameter chosen to smooth the smallest finestructure scales in the profiles. In this smoothed profile, maxima in the 2nd vertical derivative of temperature ($T_{zz}$) are selected as boundaries that allow for segmentation of the profile into linear gradients $T_s$. If the estimated $T_s$ within a segment is smaller than a specified threshold (taken to be 0.001°C/m), this segment is merged with the adjacent deeper segment, and $T_s$ is recomputed. This process is repeated until either $T_s$ is greater than 0.001°C/m or a depth segment extends to the depth level of the AW temperature maximum. The final distribution of segments over a profile is insensitive to the choice of threshold gradient for thresholds in the range 0.001 – 0.003 °C/m. Given the tight relationship between temperature and salinity, the same depth segments can be used to estimate vertical salinity gradients $S_s$. In computing $\bar{\rho}_p$, $\alpha$ and $\beta$ are averaged over a given depth segment. This segmentation method is optimal for the explicit purpose of returning a linear gradient for each different region of the profile.

Horizontal gradients. Considering the set of profiles sampled by the ITP across the Canada Basin, $T_s$ is estimated along isopycnals lying within selected depth segments (determined for the profile of interest). The lateral temperature gradient $T_x$ is computed using standard linear regression of the mean temperature within the range of isopycnals (in a given depth segment) versus distance from the most western profile near the Northwind Ridge and in close proximity to the warm Atlantic Water boundary current (the general spatial gradient is clear in Figure 1a).

b. Application of the theory

A selection of representative profiles are chosen here that allow us to compare and contrast a variety of scenarios. First, we consider a profile taken by ITP 1 on June 17, 2006 (Figure 3a), and a profile taken by ITP 41 on June 17, 2011 (Figure 3d), where both profiles are in close proximity (black and grey circles, respectively in Figure 1a). For each profile, $\bar{\rho}_p$ and $\bar{\rho}_{cr}$ are computed using estimated $T_g$, $S_g$ and $T_s$ (further taking $K_T = 10^{-6}$ m$^2$s$^{-1}$ and $A = 10^{-8}$ m$^2$s$^{-1}$). The ITP 1 profile shows a robust staircase around 250–300 m that is disrupted (in depth) by two thick apparently “mixed” layers characterized by a weak positive temperature gradient (temperature decreases with depth). This is a characteristic feature of the SF portion of an intrusion; there is evidence for weak salinity gradients also decreasing with depth, however salinity changes over the SF portion (about 0.01%) are much smaller than temperature changes (about 5%). This structure is consistent with $\bar{\rho}_p < \bar{\rho}_{cr}$ except in the depth range ~270–290 m where $\bar{\rho}_p > \bar{\rho}_{cr}$, consistent with the presence of intrusions. In the deeper portions of the profile above the AW temperature maximum (~300–380 m) relatively strong reversed temperature gradients associated with intrusions are observed (Figure 3a inset). This is consistent with $\bar{\rho}_p > \bar{\rho}_{cr}$ in this depth range. The portions showing intrusions are characterized by slightly higher $\bar{\rho}_p$ compared to the staircase regions. However $\bar{\rho}_{cr}$ also varies with depth, indicating that there is not a unique value of $\bar{\rho}_{cr}$ that determines...
whether staircases or intrusions will form, rather $\overline{K}^2 \rho$ depends upon the background vertical and horizontal gradients.

Detailed examination of the ITP 41 temperature profile (Figure 3d) illustrates that the upper part of the AW layer consists only of a staircase, as does the deep portion indicating thicker mixed layers around $\sim 340-410$ m which do not show the temperature gradient reversals characteristic of intrusions. The entire profile showing a staircase structure is consistent with $\overline{K}^2 \rho < \overline{K}^2 \rho$ throughout the depth range $\sim 300-410$ m (Figure 3f). The major differences between the ITP 1 (2006) and ITP 41 (2011) profiles is that the later year was characterized by a weaker vertical background salinity gradient and a stronger vertical background temperature gradient (both of which resulted in smaller $\overline{K}^2 \rho$ in 2011 compared to 2006). Referring to the general relationships derived in the previous section (see Figure 2b) suggests the 2011 profile is more likely to show staircases. $\overline{T}_z$ estimated over the basin in 2010–2011 is about 4 times weaker than $\overline{T}_z$ in 2005–2006 throughout the depth range of interest (Figure 3b,e). However, our linear stability analysis has shown that $\overline{T}_z$ only has a role in setting the aspect ratio of the initial interleaving motion (probably not representative of the scales of the final step structure), rather than the final response of a system to a perturbation.

Although the profiles considered in Figure 3 are generally representative of the Canada Basin during each of 2005-06 and 2010-11, we find that in the western survey region in the former period only a staircase is observed in some profiles, while in the eastern survey region in the latter period, both staircases and intrusions are observed in some profiles. We next compare and contrast these cases in each year.

A profile from the western part of the ITP 1 drift track (taken on August 31, 2005) exhibits predominantly staircases (Figure 4a), by contrast with the ITP 1 profile to the east (Figure 3a). This is likely because $\overline{T}_z$ in this region is larger compared to $\overline{T}_z$ in the east in that same year, while $\overline{S}_z$ is does not change between the two regions. Thus, $\overline{K}^2 \rho$ is lower for the western profile, and the observed staircase here is consistent with the smaller $\overline{K}^2 \rho$. This lateral distribution of the vertical gradients is consistent with a pulse of anomalously warm AW into the basin in the early to mid 2000s (e.g., see Figure 2 in Polyakov et al. 2011). That is, the warmest waters are located close to the Northwind Ridge providing a strong vertical temperature gradient (at the top boundary of the AW layer) in the western part of the basin.

A profile from the eastern part of the ITP 41 drift track (taken on December 31, 2011) exhibits weak intrusions in the deeper regions (Figure 4d), by contrast with the ITP 41 profile to the west that showed only a staircase (Figure 3d). At depths where intrusions are observed in the eastern profile, the magnitude of $\overline{T}_z$ is lower by a factor of 3, while $\overline{S}_z$ is lower by a factor of 1.5 compared to the values for the eastern profile to the west. The net effect of these salinity and temperature changes is an overall larger $\overline{K}^2 \rho$ in the deep regions, where $\overline{K}^2 \rho \sim 4$ in the west in 2011 and $\overline{K}^2 \rho \sim 9.5$ in the east in 2011. $\overline{K}^2 \rho$ is also larger for the eastern profile but remains below $\overline{K}^2 \rho$ consistent with the observed intrusions. The weaker $\overline{T}_z$ and $\overline{S}_z$ to the east is consistent with a general cooling and salinity diffusion of the AW core boundary current during its cyclonic transit around the basin.

The horizontal and vertical wavelengths of the most unstable modes are of the order of 1000–10,000 km and 15–50 m respectively in all regions surveyed and in both time periods (i.e., 2005-06 and 2010-11). Similar scales of the most unstable modes leading to either a staircase or intrusions suggest that these two end states are of the same nature. Of course the linear theory cannot predict the transient evolution of the linear profile (e.g., how interleaving motions evolve, and the associated scale adjustments such as layer splitting/merging (Radko et al. 2014)); this requires a separate analysis (as performed, for example by Walsh and Ruddick 1998; Li and McDougall 2015, for the SF case). The time scale of the instability (determined from the fastest growing mode) ranges from about one year for intrusions to several years for a staircase. A smaller magnitude of the horizontal temperature gradient shallower in the water column (e.g., Figure 4b) leads to an approximately order of magnitude decrease in $\lambda$ (i.e., an order of magnitude increase in time for a linear profile to develop staircases). Furthermore, at these shallower depths, it may be that other mechanisms (possibly more one-dimensional in nature) dominate staircase formation and evolution.

In the absence of a continuous time series of temperature and salinity at a given location, the evolution of a profile towards a staircase or intrusions in response to a changing background state is difficult to infer. (As we have shown here, detailed knowledge of the background horizontal gradients is less essential.) Nevertheless, the above analysis serves as a consistency argument between the predictions of linear theory and observations. With this analysis is the implicit assumption that the development of intrusions or a staircase does not affect the magnitudes of the background gradient (i.e., that the background gradients do not evolve on timescales faster than the fastest growing mode). An order-of-magnitude estimate for the timescale on which double-diffusive fluxes modify the background temperature gradients can be estimated as $L^2 / \overline{K} T$, where $L$ is a characteristic length scale for the background gradient. Taking $L \sim 30$ m for a vertical scale (a reasonable estimate over which linear gradients have been fit to the typical profiles), and effective diffusivity $\overline{K} T = 10^{-6}$ m$^2$s$^{-1}$, yields timescales of decades for modification of the vertical background temperature
gradient. Similarly, considering an effective vertical diffusivity for salinity of around $10^{-7}$ m$^2$s$^{-1}$ (e.g., Bebieva and Timmermans 2016), we find timescales for modification of the vertical background salinity gradient to be an order of magnitude longer than this. Following the same reasoning for the evolution of the lateral gradients (over O(100) km horizontal scales), and taking isopycnal diffusivities in the range 5 - 50 m$^2$s$^{-1}$ (Hebert et al. 1990; Walsh and Carmack 2003), also yields timescales of decades for the lateral gradients to evolve. Comparison of these timescale estimates to the maximal growth rates found here (i.e., one to several years) suggests background gradients do not evolve on timescales shorter than those associated with the development of a perturbation towards either a staircase or intrusions; thus in this analysis, the background gradients may be approximated as independent of time.

4. Summary and discussion

We have examined a scenario for the origin of double-diffusive staircases and intrusions that are observed to co-exist in the Arctic Ocean’s AW. A linear stability analysis of the governing equations was performed to determine the most unstable mode for a given horizontal and vertical linear temperature and salinity stratification that would lead towards either staircases or intrusions. Staircases are the end result of a perturbation if the observed vertical density ratio is below a critical vertical density ratio ($R_p < R_p^{cr}$), and intrusions are expected to form otherwise ($R_p > R_p^{cr}$). The poor constraints on estimates of $K_T$ and $A$ preclude any predictive capacity for our formalism; the theoretical results cannot be used to make viable predictions based on the relative values of $R_p$ and $R_p^{cr}$. Within the uncertainties on $A$ and $K_T$, a wide range of $R_p^{cr}$ is possible. For example, over possible ranges for $A$ ($10^{-8}$ to $10^{-7}$ m$^2$s$^{-1}$) and $K_T$ ($10^{-6}$ to $10^{-5}$ m$^2$s$^{-1}$), $R_p^{cr}$ changes by a factor of 3. Nevertheless, application of our theoretical formalism to observations serves as a consistency assertion, and we have shown that the linear theory is consistent with the observations for some reasonable choice of $A$ and $K_T$. We do not implement a thorough statistical analysis of all ITP measurements here, but rather provide observational examples to demonstrate the possibility of staircase formation in a DC background stratification subject to interleaving perturbations. Future analyses are required to explore the parameter space further, either numerically or in a laboratory setting.

We have shown that the dominant factors that determine the presence of either a staircase or intrusions are $T_C$, $S_C^z$ and $R_p$, with lateral temperature gradients having little influence. In general, we expect staircases in regions of relatively strong $T_C$ (and small $R_p$) and intrusions where $T_C$ is weaker. Consider, for example, an influx of warm AW in the water column. This would give rise to a stronger $T_C$ at the top boundary of the AW, and a weaker $T_C$ above and close to the AW temperature maximum (i.e., within an approximately homogeneous core). Such a modification of the background state would result in the formation of staircases at the AW top boundary and intrusions in the underlying portions. This stratification allows for vertical as well as lateral heat transfer via the intrusions. If over time the background temperature gradient in the deeper portion increases, this new $T_C$ may be susceptible to a staircase end state. One potential mechanism for increased $T_C$ could be that intrusions distribute heat more effectively downward via SF fluxes, rather than upward via DC fluxes (e.g., see heat flux estimates in Bebieva and Timmermans 2016), although the overall effect of intrusive fluxes on the background gradient is unclear. Formation of a staircase in place of intrusions would lead to reduced lateral transport of heat, and vertical fluxes would dominate heat transfer. In this respect, the lateral and vertical transfer of AW heat in the Arctic Ocean is dictated by the interplay between timescales for intrusive fluxes to modify the background gradients and the lateral supply of AW heat. Further investigation is required to quantify these timescales.

For the SF configuration, Merryfield (2000) showed that staircases are the end result of a perturbation when the vertical density ratio (defined in the conventional way for the SF configuration, inverse to the definition given here) is small, while intrusions are the end state when that density ratio takes larger values. We have used the same formalism here to demonstrate that an analogous result is applicable to the DC case. That is, a staircase is a possible end state of an interleaving perturbation. Merryfield (2000) showed that a staircase is more likely for smaller values of the effective diffusivity for heat. By contrast, we have shown that (for the DC configuration) a staircase is more likely for larger values of effective eddy diffusivity for salt. While the linear stability analysis and separation of scales framework described here is instructive for understanding the general structure of a water-column profile (i.e., staircases or intrusions), the observations often demonstrate a detailed finestructure that is somewhat more complicated. This analysis cannot describe layers in the observations that are observed between and within the main staircase mixed layer and intrusive structures (e.g., the smallest-scale temperature finestructure in the insets of Figure 4a). These features, however, may be key to the transition between interleaving structures and staircases – a conjecture that will require further analysis.

Acknowledgments. The Ice-Tethered Profiler data were collected and made available by the Ice-Tethered Profiler Program based at the Woods Hole Oceanographic Institution (Krishfield et al. 2008; Toole et al. 2011); data are available at http://www.whoi.edu/itp/data. Funding was
References


**Fig. 1:** a) Map showing locations of ITP 1 and ITP 41 profiles over the course of their drift; colors indicate the AW potential temperature maximum (°C). b) Potential temperature (°C, referenced to the surface) and c) salinity profiles measured by ITP 1 (black lines, 17 June 2006) and ITP 41 (grey lines, 17 June 2011); locations where both profiles were sampled are marked by circles with corresponding colors on the map in a). The expanded scales show the double-diffusive staircase and intrusions for the ITP 1 profile. Diamonds mark the locations of two other profiles examined here, Figure 4.
Fig. 2: a) Schematic showing evolution of temperature and salinity according to (15) and (16) (following Merryfield 2002). Dashed lines are constant $R_\rho$ contours, bounded by two solid lines where $R_\rho = 1$ and $R_\rho \to \infty$. Blue lines show evolution of the initial perturbation from a background state (red dot, characterized by $R_\rho$) towards either a staircase (S) or intrusions (I). There are four regimes shown on the diagram depending upon values of $(\alpha T_z, \beta S_z)$: DC-unstable, convectively unstable, SF-unstable and stable stratification. b) Solutions to (27) for $\lambda_{\text{max}}$ (black curves) and $R_\rho^{cr}$ (red curves) as a function of $T_x$ and $T_z$, where the range of temperature gradients approximately correspond to observed values. Values $A$ and $K_T$ are as given in the text and the background salinity gradient is held fixed for the calculations ($S_z = -0.002 \text{m}^{-1}$). The approximate delineation between when staircases are expected and when intrusions are expected (based on values of $R_\rho$ and $R_\rho^{cr}$) is shown by a wavy line.
Fig. 3: a) Potential temperature ($^\circ$C, referenced to the surface) profile measured by ITP 1 in the Canada Basin at the location of the black circle in Figure 1a. Red lines overlain indicate a piece-wise linear approximation to the profile. Zoomed-in potential temperature profiles are shown in the insets, with salinity shown by a blue line (bottom inset) with the scale above. b) Lateral background temperature gradient ($10^{-7}$ $^\circ$C/m) for each depth segment shown in a). c) Corresponding profiles of $\overline{\rho c_T}$ and $\overline{\rho}$ (see text). d), e) and f) as for a), b) and c) respectively but for a profile taken by ITP 41 (grey circle in Figure 1a).
Fig. 4: a) Potential temperature (°C, referenced to the surface) profile measured by ITP 1 in the Canada Basin at the location of the black diamond in Figure 1a. Red lines overlain indicate a piece-wise linear approximation to the profile. b) Lateral background temperature gradient (10^{-7} °C/m) for each depth segment shown in a). c) Corresponding profiles of \( \overline{R_p} \) and \( \overline{R_\rho} \). d), e) and f) as for a), b) and c) respectively but for a profile taken by ITP 41 (grey diamond in Figure 1a).