Point Vortex Dynamics For Coupled Surface/interior Qg And Propagating Heton Clusters In Models For Ocean Convection

Article *in* Geophysical & Astrophysical Fluid Dynamics · May 2000

DOI: 10.1080/03091920108203407 · Source: CiteSeer

Chjan Lim
Rensselaer Polytechnic Institute

POINT VORTEX DYNAMICS FOR COUPLED SURFACE/INTERIOR QG AND PROPAGATING HETON CLUSTERS IN MODELS FOR OCEAN CONVECTION

by

Chjan C. Lim
Mathematical Sciences Dept.
Rensselaer Polytechnic Inst.
Troy, NY 12180

Andrew J. Majda
Courant Inst. Mathematical Sciences
New York University
New York, NY 10012

Abstract

The dynamic behavior of baroclinic point vortices in two-layer quasi-geostrophic flow provides a compact model for studying the transport of heat in a variety of geophysical flows including recent heton models for open ocean convection as a response to spatially localized intense surface cooling. In such heton models, the exchange of heat with the region external to the compact cooling region reaches a statistical equilibrium through the propagation of tilted heton clusters. Such tilted heton clusters are aggregates of cyclonic vortices in the upper layer and anti-cyclonic vortices in the lower layer which collectively propagate almost as an elementary tilted heton pair even though the individual vortices undergo shifts in their relative locations. One main result in this paper is a mathematical theorem demonstrating the existence of large families of long-lived propagating heton clusters for the two-layer model in a fashion compatible to a remarkable degree with the earlier numerical simulations. Two-layer quasi-geostrophic flow is an idealization of coupled surface/interior quasi-geostrophic flow. The second family of results in this paper involves the systematic development of Hamiltonian point vortex dynamics for coupled surface/interior QG with an emphasis on propagating solutions that transport heat. These are novel vortex systems of mixed species where surface heat particles interact with quasi-geostrophic point vortices. The variety of elementary two-vortex exact solutions that transport heat include two surface heat particles of opposite strength, tilted pairs of a surface heat particle coupled to an interior vortex of opposite strength and two interior tilted vortices of opposite strength at different depths. The propagation speeds of the tilted elementary hetons in the coupled surface/interior QG model are compared and contrasted with those in the simpler two-layer heton models. Finally, mathematical theorems are presented for the existence of large families of propagating long-lived tilted heton clusters for point vortex solutions in coupled surface/interior QG flow.

Key words: Ocean convection, baroclinic vortices, Hamiltonian structure

1. Introduction

Beginning with the seminal work of Hogg and Stommel ([10], [11]), the dynamic behavior of systems of point vortices in two-layer quasi-geostrophic models have become very attractive compact models for understanding the transport of heat in a variety of prototype geophysical flows in the ocean ([13], [27], [14], [15]). In particular, Legg, Marshall, and their collaborators ([13], [14], [15]) have introduced and analyzed two-layer heton models for studying the spreading phase of penetrative open ocean convection. Deep open-ocean convection, which occurs in the Labrador Sea, the Greenland Sea, and the Mediterranean Sea in the current world climate, is an important phenomenon that strongly influences the thermohaline circulation governing the poleward transport of heat in the ocean. For a recent comprehensive survey, see the review by Marshall and Schott [24].

Hogg and Stommel, ([10], [11]), introduced cold hetons as steady purely baroclinic point vortices with cyclonic flow in the upper layer and anti-cyclonic flow in the lower layer while hot hetons have the flows in the upper and lower layers exchanged. They also observed that vertically tilted hetons are elementary exact solutions which propagate at a constant velocity depending on the separation distance and transport heat. In the two-layer heton models for the spreading phase of convection from a localized region of compact intense surface cooling (see [13], [14], [15]), the surface cooling in the rapidly rotating fluid leads to localized convective overturning which is modelled by a distribution of elementary cold hetons spread over the cooling region. Direct numerical simulations of the resulting baroclinic point vortex system, ([13], [14]), show that the exchange of heat with the region external to the compact cooling region reaches a statistical equilibrium through the propagation of tilted heton clusters. Such tilted heton clusters are aggregates of cyclonic vortices in the upper layer and anti-cyclonic vortices in the lower layer which collectively propagate almost as an elementary tilted heton even though the individual vortices undergo small shifts in their relative locations. Unambiguous graphical evidence for the role of such tilted heton clusters is given in Figures 11 C), D), and 12 C), D) from ref. [13]; furthermore, we note for comparison with the theory presented here that such tilted heton clusters have a mismatch of one typically and occasionally two vortices between the number of vortices in the upper and lower layers. Legg and Marshall, [13], also emphasize that the two-layer heton models for ocean convection are idealized from a more complete physical model involving coupled surface/interior quasi-geostrophic flow (see [26], Chapter 6).

We will report here a detailed count of the number of vortices in the propagating heton clusters in figures 11 C), D) and 12 D); the figures 11 A), B) and 12 A), B), C) are not relevant to this vortex count because they represent the early development of the initial cluster of hetons, before the ejection of tilted heton clusters from the rim current. To fix some parameters for the following discussion, we note that in both sets of figures, the initial cluster of hetons has radius 5λ, and the furthest tilted heton cluster in 11 D)and 12 D) respectively is at distance 20λ and 25λ from the center of their frames. We will focus on well-defined heton clusters that are further than 10λ from the center. For figure 11 C), there are four such clusters: cluster 1 with coordinates (8,10) has 5 reds and 8 greens, cluster 2 at (5, -12) has 3 reds and 4 greens, cluster 3 at (-15, 2) has 5 reds and 6 greens, and cluster 4 at (-10, 6) has 5 reds and 3 greens. There are again 4 well-defined clusters in 11 D): cluster 1 at (-9, 7) has 5 reds and 3 greens, cluster 2 at (-10, 2) has 3 reds and 3 greens, cluster 3 at (-20, -5) has 5 reds and 6 greens, and cluster 4 at (10, -13) has 3 reds and 4 greens. For figure 12 D), there are 5 such clusters: cluster 1 at (-15, 4) has 12 reds and 14 greens, cluster 2 at (-10, -10) has 13 reds and 8 greens, cluster 3 at (2, -10) has 2 reds and 4 greens, cluster 4 at (25, -5) has 8 reds and 8 greens, and cluster 5 at (8, 12) has 7 reds and 6 greens. Figure 12 C) has only one such well-defined cluster and it has 8 reds and 8 greens. Although the detailed initial conditions underlying the numerical experiments in Legg and Marshall [13] differ somewhat from the theoretical conditions of our KAM analysis (for example, the cone condition in section 6.2 is not satisfied), the above vortex count based on the later stages of the development of propagating heton clusters suggests that there is a mismatch in the vortex number between the upper and lower layers. Initial configurations such as those in figures 11 and 12 which clearly do not satisfy the cone condition, can evolve later into separate heton clusters which more nearly satisfy this condition, because in the higher dimensional phase space of the heton model, the regions near the KAM tori (if they exists) are not dynamically invariant.

The primary goals of this paper are twofold: 1) to develop the point vortex dynamics of coupled surface/interior quasi-geostrophic flow including the variety of elementary solutions which transport heat beyond those in the heton models; 2) to develop a mathematical framework to demonstrate the existence of large families of long-lived propagating heton clusters for both the two-layer heton models and the more general point vortex dynamics of coupled surface/interior quasi-geostrophic flow. In particular, as regards the second goal of this paper, there is a remarkable serendipity between the numerical results of Legg and Marshall from

[13] which we summarized briefly in the preceding paragraph and the following Theorem proved in section 6.2 below:

Theorem: The Heton point vortex model for two-layer quasi-geostrophic flows supports large families of long-lived propagating heton clusters where the upper layer cluster consists of k positive potential vortices and the lower-layer cluster consists of $k \pm \delta$ negative potential vortices of the same strength such that k satisfies $k \geq 3$ with $\delta = 1$ or $k \geq 4$ with $\delta = 2$. More detailed structure of these solutions is presented in sections 6.2 and 6.3 below.

Next we summarize the remaining contents of the paper. In section 2, we introduce the continuum equations for coupled surface/interior quasi-geostrophic flow. In section 3, we establish that a natural discretization of these equations as a particle method leads to Hamiltonian point vortex dynamics involving surface heat particles coupled to quasi-geostrophic interior point vortices. The conserved quantities which arise from symmetries of this Hamiltonian as well as some novel features of surface/interior coupling are also developed in section 3. In section 4 we summarize several features and elementary solutions of the two-layer heton models which are useful for subsequent developments in the paper. Elementary exact solutions for the point vortex dynamics of coupled surface/interior quasi-geostrophic flow are developed in section 5 with an emphasis on the wide variety of exact solutions which transport heat and direct comparison with both elementary solutions for the two-layer heton models and also standard barotropic point vortices (see [2], [22], [23] and references therein). Section 6 contains the mathematical machinery needed to establish the existence of plentiful families of propagating tilted heton clusters as in the Theorem above for both the two-layer heton models and the point vortex models of coupled surface/interior quasi-geostrophic flow. The main mathematical tools involve a combinatorial version of the KAM theorem for N-body Hamiltonians developed by the first author in a series of papers ([16] [17] [18] [19]), with several novel aspects that arise due to the mixed species present in the point vortex problems considered here. Section 6.1 is an elementary introduction and demonstration of these techniques with the geophysical applications presented in section 6.2. Any reader can skip the mathematical details yet understand the main results by reading the introductions to section 6 and 6.2 and then section 6.3. On the other hand, one can read section 6.1 to get the spirit of the mathematical arguments for an idealized simpler problem.

We end the introduction by mentioning that there is recent work by DiBattista and the second author ([3], [4], [5]) on utilizing suitable equilibrium statistical mechanics for heton models to predict the spreading phase of a basin-wide cooling event (as opposed to the localized cooling from [13]) without detailed resolution of the dynamics.

2. Equations of Motion for Coupled Surface/Interior QG

A basic reference for this section is chapter 6 of [26]. The fluid domain is

$$D = R^2 \times [-H, 0].$$

Eventually, we will take H to be infinite in our Green function calculations below. Let the potential vorticity be

$$q = \Delta_H \Psi + k^2 \frac{\partial^2 \Psi}{\partial z^2}, \tag{2.1}$$

where Ψ is the stream function in the problem and k is a parameter, essentially the Rossby deformation radius, that is determined by the stratification profile, the Coriolis parameter f_o and the Brunt-Vaisala frequency N. By selecting the vertical unit of length to be k^{-1}, without loss of generality, we set $k \equiv 1$ in the discussion below. The horizontal velocity is given by

$$\vec{v}_H = \nabla_H^{\perp} \Psi = \begin{pmatrix} -\frac{\partial \Psi}{\partial y} \\ \frac{\partial \Psi}{\partial x} \end{pmatrix}.$$

The potential temperature θ is given by the hydrostatic approximation

$$\theta = \frac{\partial \Psi}{\partial z}, \tag{2.2}$$

and the stream function Ψ is identified with the pressure P (cf. [26] chapter 6). The quasi-geostrophic potential vorticity equation for q is given by the first of the following pair of equations

$$\frac{\partial q}{\partial t} + J_H(\Psi, q) = 0, \text{ for } (x, y, z) \in D, \tag{2.3}$$

$$\frac{\partial \theta}{\partial t} + J_H(\Psi, \theta) = 0, \text{ at } z = 0, \tag{2.4}$$

where

$$J_H(\Psi, \circ) = \nabla_H^{\perp} \Psi \cdot \nabla \circ .$$

The second equation is a consequence of the thermodynamic equation

$$\frac{D_H}{Dt}\theta + w = 0,$$

and the boundary condition at the upper boundary of the fluid $z = 0$ for zero viscosity (no wind stress), and no internal heating, i.e.,

$$w = -\frac{D_H}{Dt}\left(\frac{\partial\Psi}{\partial z}\right)|_{z=0} \equiv -\frac{\partial}{\partial t}\left(\frac{\partial\Psi}{\partial z}\right)|_{z=0} - J_H(\Psi, \frac{\partial\Psi}{\partial z}|_{z=0}) = 0, \tag{2.5}$$

where w is the vertical component of the velocity. In the quasi-geostrophic approximation the kinematic boundary condition at $z = 0$ is essentially that for a rigid upper lid [26]. On the other hand, the isopycnal surfaces in the ocean deform more than the free surface, and are responsible for the change in potential vorticity due to vortex stretching.

The boundary analysis at $z = -H$ for trivial topography, no heating and zero viscosity [26], i.e.,

$$\frac{\partial}{\partial t}\left(\frac{\partial\Psi}{\partial z}\right)|_{z=-H} + J_H(\Psi, \frac{\partial\Psi}{\partial z}|_{z=-H}) = 0$$

implies the boundary condition for our model,

$$\frac{\partial\Psi}{\partial z}|_{z=-H} = 0. \tag{2.6}$$

Alternatively, the pair of equations (2.3), (2.4) and boundary condition (2.6) can be viewed as the equations of motion for a coupled interior-surface quasi-geostrophic model. The first equation governs the evolution of the potential vorticity $q = \Delta\Psi$, and the second equation governs the evolution of the surface potential temperature θ. These two quantities are in turn related by $\theta = \frac{\partial\Psi}{\partial z}|_{z=0}$.

The above equations form an infinite-dimensional Hamiltonian system [29], [30] with Hamiltonian function

$$H_\infty = \frac{1}{2}\int_D \left(|\nabla_H\Psi|^2 + \left|\frac{\partial\Psi}{\partial z}\right|^2\right) dV.$$

The first term in the integrand is the kinetic energy of horizontal motion, and the second term is the potential energy. Since there is no viscosity in our model, there is no dissipative loss, through the Ekman layer for instance. Because the vertical velocity $w = 0$ at the boundary $z = 0$ in (2.5), no work is done by the fluid pressure at the upper surface. Therefore, H_∞ is a conserved energy functional for this model [26]. We split the potential vorticity stream function into two parts

$$\Psi = \Psi_I + \Psi_B, \tag{2.7}$$

where Ψ_I and Ψ_B are chosen so that

$$\begin{aligned} \Delta\Psi_I &= q, \\ \Delta\Psi_B &= 0, \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} \frac{\partial \Psi_I}{\partial z}|_{z=0} &= 0, \ \frac{\partial \Psi_I}{\partial z}|_{z=-H} = 0, \\ \frac{\partial \Psi_B}{\partial z}|_{z=0} &= \theta(z=0), \ \frac{\partial \Psi_B}{\partial z}|_{z=-H} = 0. \end{aligned} \tag{2.9}$$

Integrating by parts we obtain

$$\begin{aligned} H_\infty &= -\frac{1}{2}\int_D \Psi\Delta\Psi \ dV \ + \ \frac{1}{2}\int_{R^2} \Psi\frac{\partial\Psi}{\partial z}|_{z=0} \ dA - \frac{1}{2}\int_{R^2} \Psi\frac{\partial\Psi}{\partial z}|_{z=-H} \ dA \\ &= -\frac{1}{2}\int_D \Psi_I \ q \ dV \ + \ \frac{1}{2}\int_{R^2} \Psi_B|_{z=0} \ \theta(0) \ dA \\ &\quad +\frac{1}{2}\int_D -\Psi_B \ q \ dV \ +\frac{1}{2}\int_{R^2} \Psi_I(0) \ \theta(0) \ dA. \end{aligned}$$

Letting

$$\begin{aligned} H_I &= -\frac{1}{2}\int_D \Psi_I \ q \ dV, \\ H_B &= \frac{1}{2}\int_{R^2} \Psi_B|_{z=0} \ \theta(0) \ dA, \\ H_{IB} &= \frac{1}{2}\int_D -\Psi_B \ q \ dV \ +\frac{1}{2}\int_{R^2} \Psi_I|_{z=0} \ \theta(0) \ dA, \end{aligned} \tag{2.10}$$

we obtain a useful splitting of the Hamiltonian into an interior term H_I, a surface term H_B and an interaction term H_{IB},

$$H_\infty = H_I + H_B + H_{IB.}$$

3. Point vortex equations for coupled surface - interior quasi-geostrophic flows

We take the approach here of a particle method discretization of the coupled quasi-geostrophic equations in (2.3),(2.4) with the lower boundary H at infinity for explicit calculational simplicity. The derivation of the finite dimensional Hamiltonian model for the coupled surface temperature interior potential vorticity quasi-geostrophic model (the coupled QG model in short) is based on standard methods related to the vortex method in planar fluid mechanics [2]. In later subsections we will discuss several interesting special cases of this model.

For the surface Hamiltonian we write

$$\theta(\vec{\alpha}, t) \simeq \sum_{i=1}^{n} \theta_i(t)\delta(\vec{\alpha} - \vec{\alpha}_i(t)) \tag{3.1}$$

where $\vec{\alpha}_i(t) = (\alpha_i(t), \beta_i(t))$ and $\theta_i(t)$ are the position and temperature charge of the surface heat particle at time t. From the relation between Ψ_B and θ in (2.9), we obtain the equation

$$-(-\Delta)^{1/2}\Psi_B = \theta$$

which has the solution

$$\Psi_B(\vec{x}, z, t) = -\int_{R^2} \frac{1}{|\vec{x} - \vec{x}'|}\theta(\vec{x}', t)\; dA'.$$

This yields the potential density Hamiltonian in the continuous problem

$$H_B = -\int_{R^2\times R^2} \frac{\theta(\vec{x}')\theta(\vec{x})}{|\vec{x} - \vec{x}'|} dA dA'.$$

Using (3.1) we get the discrete surface Hamiltonian

$$H_B = -\frac{1}{2\pi}\sum_{i\neq j}^{n} \frac{\theta_i\theta_j}{|\vec{\alpha}_i - \vec{\alpha}_j|}.$$

It turns out that θ_i is constant in time.

Next we discretize the interior Hamiltonian [8]. The continuous potential vorticity is approximated by a sum over m point vortices

$$q(\vec{x}, z, t) \simeq \sum_{i=1}^{M} \lambda_i(t)\delta(x - x_i(t), y - y_i(t), z - z_i(t)) \tag{3.2}$$

with potential vorticity charge $\lambda_i(t)$ (which turns out to be independent of time t), and position $(\vec{x}_i, z_i) = (x_i, y_{i,} z_i)$. From the relation between q and Ψ_I, i.e.,

$$\Delta \Psi_I = q$$

we obtain the following Green's function in the semi-infinite domain where H is taken for simplicity to be infinite:

$$G(x - x_i, y - y_i, z, z_i) = \left(\begin{array}{c} \frac{1}{[(x-x_i)^2+(y-y_i)^2+(z-z_i)^2]^{1/2}} \\ +\frac{1}{[(x-x_i)^2+(y-y_i)^2+(z+z_i)^2]^{1/2}} \end{array} \right).$$

Using the finite sum (3.2), we obtain

$$\Psi_I(x, y, z, t) = \frac{k}{4\pi} \sum_{i=1}^{m} \lambda_i(t) G(x - x_i(t), y - y_i(t), z, z_i(t)).$$

The discrete Hamiltonian for the interior is then given by

$$H_I = -\frac{1}{4\pi} \sum_{i \neq j}^{m} \lambda_i \lambda_j \left(\begin{array}{c} \frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{1/2}} \\ +\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2]^{1/2}} \end{array} \right).$$

It is easy to show that the vertical coordinates z_i of the vortices are constants under the flow of the Hamiltonian H_I.

The discrete version of the interaction Hamiltonian $H_{IB} = \frac{1}{2} \int_D -\Psi_B \; q \; dV$ $+\frac{1}{2} \int_{R^2} \Psi_I|_{z=0} \; \theta \; dA$ is given by

$$\frac{1}{2\pi} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\lambda_i \theta_j}{|\vec{x}_i - \vec{\alpha}_j|} + \frac{1}{2\pi} \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i \theta_j \left(\frac{1}{[(\alpha_j - x_i)^2 + (\beta_j - y_i)^2 + (z_i)^2]^{1/2}} \right).$$

Summarizing we have the finite dimensional Hamiltonian function:

$$\begin{aligned} H \quad = \quad & -\frac{1}{4\pi} \sum_{i \neq j}^{M} \lambda_i \lambda_j \left(\begin{array}{c} \frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{1/2}} \\ +\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2]^{1/2}} \end{array} \right) \\ & -\frac{1}{2\pi} \sum_{i \neq j}^{n} \frac{\theta_i \theta_j}{|\vec{\alpha}_i - \vec{\alpha}_j|} \\ & +\frac{1}{2\pi} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\lambda_i \theta_j}{|\vec{x}_i - \vec{\alpha}_j|} \\ & +\frac{1}{2\pi} \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i \theta_j \left(\frac{1}{[(\alpha_j - x_i)^2 + (\beta_j - y_i)^2 + (z_i)^2]^{1/2}} \right). \end{aligned} \tag{3.3}$$

The equations of motion for the variables $\vec{x}_i = (x_i, y_i)$ and $\vec{\alpha}_i = (\alpha_i, \beta_i)$ are:

$$\begin{aligned}
\lambda_i \frac{d}{dt} x_i &= -\sum_{j \neq i}^{m} \frac{\lambda_i \lambda_j}{4\pi} (y_i - y_j) \left(\begin{array}{c} \frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{3/2}} \\ +\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2]^{3/2}} \end{array} \right) \\
+\frac{\lambda_i \theta_j}{2\pi} \sum_{j=1}^{n} & \quad \frac{(y_i - \beta_j)}{|\vec{x}_i - \vec{\alpha}_j|^3} + \frac{(y_i - \beta_j)}{[(\alpha_j - x_i)^2 + (\beta_j - y_i)^2 + (z_i)^2]^{3/2}}, \\
\lambda_i \frac{d}{dt} y_i &= \sum_{j \neq i}^{m} \frac{\lambda_i \lambda_j}{4\pi} (x_i - x_j) \left(\begin{array}{c} \frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{3/2}} \\ +\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2]^{3/2}} \end{array} \right) \\
-\frac{\lambda_i \theta_j}{2\pi} \sum_{j=1}^{n} & \quad \frac{(x_i - \alpha_j)}{|\vec{x}_i - \vec{\alpha}_j|^3} + \frac{(x_i - \alpha_j)}{[(\alpha_j - x_i)^2 + (\beta_j - y_i)^2 + (z_i)^2]^{3/2}}; \\
\theta_i \frac{d}{dt} \alpha_i &= -\frac{1}{2\pi} \sum_{i \neq j}^{n} \frac{\theta_i \theta_j (\beta_i - \beta_j)}{|\vec{\alpha}_i - \vec{\alpha}_j|^3} \\
-\frac{\lambda_j \theta_i}{2\pi} \sum_{j=1}^{m} & \quad \frac{(y_j - \beta_i)}{|\vec{x}_j - \vec{\alpha}_i|^3} + \frac{(y_j - \beta_i)}{[(\alpha_i - x_j)^2 + (\beta_i - y_j)^2 + (z_j)^2]^{3/2}}, \\
\theta_i \frac{d}{dt} \beta_i &= \frac{1}{2\pi} \sum_{i \neq j}^{n} \frac{\theta_i \theta_j (\alpha_i - \alpha_j)}{|\vec{\alpha}_i - \vec{\alpha}_j|^3} \\
+\frac{\lambda_j \theta_i}{2\pi} \sum_{j=1}^{m} & \quad \frac{(x_j - \alpha_i)}{|\vec{x}_j - \vec{\alpha}_i|^3} + \frac{(x_j - \alpha_i)}{[(\alpha_i - x_j)^2 + (\beta_i - y_j)^2 + (z_j)^2]^{3/2}}.
\end{aligned} \tag{3.4}$$

These equations are Hamilton's equations

$$\begin{aligned}
\frac{d}{dt} q_i &= \frac{d}{dt} x_i = -\frac{\partial H}{\partial(\lambda_i y_i)} = -\frac{\partial H}{\partial p_i}, \\
\frac{d}{dt} p_i &= \lambda_i \frac{d}{dt} y_i = \frac{\partial H}{\partial x_i} = \frac{\partial H}{\partial q_i}, \ i = 1, ..., m; \\
\frac{d}{dt} q_j &= \frac{d}{dt} \alpha_j = -\frac{\partial H}{\partial(\theta_j \beta_j)} = -\frac{\partial H}{\partial p_j}, \\
\frac{d}{dt} p_j &= \theta_j \frac{d}{dt} \beta_j = \frac{\partial H}{\partial \alpha_j} = \frac{\partial H}{\partial q_j}, \ j = m+1, ..., m+n,
\end{aligned}$$

for the symplectic variables:

$$\begin{aligned}
q_i &= x_i, \ p_i = \lambda_i y_i \ , \ i = 1, ..., m; \\
q_j &= \alpha_j, \ p_j = \theta_j \beta_j, \ j = m+1, ..., m+n.
\end{aligned} \tag{3.5}$$

The phase space of the system of $N = m+n$ particles is fixed once the number L of levels z_j and the number N_j of particles in each level are chosen. Here L is the number of levels of the interior vortices plus one for the surface hetons. The actual values of the levels z_j and the numbers N_j depend on the vertical structure of the infinite dimensional problem.

3.1. Symmetries and Invariants

Besides the Hamiltonian function in (3.3), we have the standard invariants associated with linear and angular momenta. The first invariant arises from the horizontal translational symmetry of the Hamiltonian (3.3):

$$G = \sum_{i=1}^{M} \lambda_i \vec{x}_i + \sum_{j=1}^{N} \theta_j \vec{\alpha}_j. \tag{3.6}$$

The second invariant comes from the $SO(2)$ (rotational) symmetry of the Hamiltonian (3.3):

$$K = \sum_{i=1}^{M} \lambda_i |\vec{x}_i|^2 + \sum_{j=1}^{N} \theta_j |\vec{\alpha}_j|^2. \tag{3.7}$$

These invariants correspond to conserved quantities in the continuous problem. For a rather complete discussion of the role of symmetries in vortex dynamics in another context, we refer the reader to Lim, Montaldi and Roberts [20]. Furthermore the equations are time-reversible [7], [22] with respect to the following order two operations:

$$\begin{aligned} R_1 &: (\vec{q}, \vec{p}) \to (\vec{q}, -\vec{p}), \\ R_2 &: (\vec{q}, \vec{p}) \to (-\vec{q}, \vec{p}). \end{aligned}$$

However, these reversible symmetries do not lead to additional first integrals in this problem.

3.2. Special case: (a) pure surface heat particles

Here we look at the special case where there are only surface heat particles and no interior vortices. The existence of surface heat particles in the Coupled QG model allows us to go beyond the special case in Gryanik [8] of purely interior PV and also the two-layer Heton model ([10], [11],[13]), in treating the heat transfer at the air-ocean interface. The discrete Hamiltonian from (3.3) in the special case with only surface particles is given by

$$
\begin{aligned}
H_B(\vec{\alpha}_i) &= -\frac{1}{2\pi}\sum_{i\neq j}^{n}\frac{\theta_i\theta_j}{|\vec{\alpha}_i-\vec{\alpha}_j|}\\
&= -\frac{1}{2\pi}\sum_{i\neq j}^{n}\frac{\theta_i\theta_j}{\left((q_j-q_i)^2+\left(\frac{p_j}{\theta_j}-\frac{p_i}{\theta i}\right)^2\right)^{1/2}},
\end{aligned}
$$

where the potential temperature of particle j is given by θ_j and its position is given by $\vec{\alpha}_j=(\alpha_j,\,\beta_j)\in R^2$. In terms of the symplectic variables

$$
q_j=\alpha_j,\;\; p_j=\theta_j\beta_j,
$$

the equations of motion in Hamiltonian form are

$$
\begin{aligned}
\frac{d}{dt}q_j &= -\frac{\partial H_S}{\partial p_j}=\frac{1}{2\pi}\sum_{i\neq j}^{n}\frac{\theta_i(\frac{p_j}{\theta_j}-\frac{p_i}{\theta_i})}{\left((q_j-q_i)^2+\left(\frac{p_j}{\theta_j}-\frac{p_i}{\theta i}\right)^2\right)^{3/2}},\\
\frac{d}{dt}p_j &= \frac{\partial H_S}{\partial q_j}=-\frac{1}{2\pi}\sum_{i\neq j}^{n}\frac{\theta_j\theta_i(q_j-q_i)}{\left((q_j-q_i)^2+\left(\frac{p_j}{\theta_j}-\frac{p_i}{\theta i}\right)^2\right)^{3/2}}.
\end{aligned}
\tag{3.8}
$$

We note that the interaction in the surface Hamiltonian H_B has a $O(\frac{1}{r})$ singularity at the origin and a $O(\frac{1}{r})$ decay at infinity in the distance r between two surface particles. Thus there is a stronger singularity at the origin but a weaker interaction at large distances compared with planar point vortices. Elementary exact solutions in this special case are given in (5.1).

3.3. Special case: (b) pure interior vortices

In the case where there are no surface heat particles, the discrete model in (3.3) reduces to the following system

$$
H_{PV}=-\frac{1}{4\pi}\sum_{i\neq j}^{M}\lambda_i\lambda_j\left(\begin{array}{c}\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{1/2}}\\ +\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2]^{1/2}}\end{array}\right),
$$

where z_j is the depth of the $j-th$ interior vortex. Hamilton's equations of motion in this case are

$$\begin{aligned}
\lambda_i \frac{d}{dt} x_i &= -\sum_{j \neq i}^{m} \frac{\lambda_i \lambda_j}{4\pi} (y_i - y_j) \left(\begin{array}{c} \frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{3/2}} \\ +\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2]^{3/2}} \end{array} \right) \qquad (3.9) \\
\lambda_i \frac{d}{dt} y_i &= \sum_{j \neq i}^{m} \frac{\lambda_i \lambda_j}{4\pi} (x_i - x_j) \left(\begin{array}{c} \frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{3/2}} \\ +\frac{1}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2]^{3/2}} \end{array} \right).
\end{aligned}$$

The elementary dynamics of this special case have been studied by Gryanik [8] and later we will refer to his results. If we let

$$d_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2]^{1/2}$$

then for

$$d_{ij} \gg |zi - zj|,$$

the dynamics of the interior point vortices is essentially barotropic, that is, independent of the vertical coordinate z. When the horizontal and vertical distances between vortices are comparable or the vertical separation exceeds the horizontal separation, baroclinic contributions to the dynamics become important [8].

In comparisons with the Heton model discussed in section 4 below, it is this special case of the Coupled QG model of purely interior PV that corresponds most directly to the point vortices of the two layer model. This special case effectively generalizes the two layer model to an arbitrary number of layers where the dynamics is purely advective, but interacts with the PV in other layers. Another way to look at this special case, is to consider an interacting multiple species bath of interior PV, where the particle numbers of the species are conserved. The above derivation of the full Coupled QG model in terms of the Green functions of three dimensional Laplace-Poisson problems, provides a rational basis for this generalization of the two layer model.

4. Heton model

For later comparisons with the Coupled QG model we summarize the basic properties of the discrete Hamiltonian system that governs the dynamics of the Heton model introduced by Hogg and Stommel [10] and used by Legg and Marshall in their model for open ocean convection [13]. We also need these formulas for the

Heton model for our discussion in section 6. Hetons arise in the standard two-layer quasi-geostrophic equations from a crude vertical discretization of the QG equations with equal equivalent depths. See the last sections of Chapter 6 of [26]. For such two-layer models, the potential vorticity of an aligned heton (pair of point vortices stacked one on top of the other) consists of a point vortex in each layer with potential vorticities given by

$$\begin{aligned} q_1 &= \lambda_i \delta(\vec{x} - \vec{x}_i) \\ q_2 &= -\lambda_i \ \delta(\vec{x} - \vec{x}_i), \end{aligned}$$

where $\Delta Q = 2\lambda_i$ is the constant amplitude of the PV which is determined by the strength of the surface cooling, and by the circulation in each layer, and $\vec{x}_i$ is the location of the point vortex in the plane. The stream function Ψ at $\vec{x}$ due to a vortex of strength λ_i located at $\vec{x}_i$ in the same layer is

$$\Psi = \frac{\lambda_i}{2} \left[\ln |\vec{x} - \vec{x}_i| - K_0 \left(\frac{|\vec{x} - \vec{x}_i|}{L_R(\pi/\sqrt{8})} \right) \right], \tag{4.1}$$

and that due to a vortex of strength λ_i located at $\vec{x}_i$ in the other layer is

$$\Psi = \frac{\lambda_i}{2} \left[\ln |\vec{x} - \vec{x}_i| + K_0 \left(\frac{|\vec{x} - \vec{x}_i|}{L_R(\pi/\sqrt{8})} \right) \right], \tag{4.2}$$

where L_R is the Rossby radius of deformation based on the total depth D of the two layers, and K_0 is the modified Bessel function of zeroth order. The Hamiltonian function for m point vortices (with n particles in the top layer) is given by

$$\begin{aligned} H &= \frac{1}{4} \sum_{i \neq j = 1}^{n} \lambda_i^1 \lambda_j^1 \left[\ln |\vec{x}_j^1 - \vec{x}_i^1| - K_0 \left(\frac{|\vec{x}_j^1 - \vec{x}_i^1|}{L_R(\pi/\sqrt{8})} \right) \right] \\ &+ \frac{1}{4} \sum_{i \neq j = 1}^{m-n} \lambda_i^2 \lambda_j^2 \left[\ln |\vec{x}_j^2 - \vec{x}_i^2| - K_0 \left(\frac{|\vec{x}_j^2 - \vec{x}_i^2|}{L_R(\pi/\sqrt{8})} \right) \right] \\ &+ \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{m-n} \lambda_i^1 \lambda_j^2 \left[\ln |\vec{x}_j^2 - \vec{x}_i^1| + K_0 \left(\frac{|\vec{x}_j^2 - \vec{x}_i^1|}{L_R(\pi/\sqrt{8})} \right) \right]. \end{aligned} \tag{4.3}$$

The equations of motion of these m vortices located at $\vec{x}_j^\alpha = (x_j^\alpha, \ y_j^\alpha)$, where $\alpha = 1, 2$ are therefore given by

$$\lambda_j^1 \frac{d}{dt}(x_j^1, \ y_j^1) = \left(\frac{\partial H}{\partial y_j^1}, \ -\frac{\partial H}{\partial x_j^1} \right) \tag{4.4}$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{i\neq j}^{n} \lambda_i^1\lambda_j^1 \left[\frac{1}{|\vec{x}_j^1-\vec{x}_i^1|} + \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{|\vec{x}_j^1-\vec{x}_i^1|}{L_R(\pi/\sqrt{8})}\right)\right]\left(x_i^1-x_j^1,\ y_i^1-y_j^1\right)^{\perp} \\
+\frac{1}{2}\sum_{i=1}^{m-n} \quad & \lambda_i^2\lambda_j^1 \left[\frac{1}{|\vec{x}_j^1-\vec{x}_i^2|} - \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{|\vec{x}_j^1-\vec{x}_i^2|}{L_R(\pi/\sqrt{8})}\right)\right]\left(x_i^2-x_j^1,\ y_i^2-y_j^1\right)^{\perp}
\end{aligned}
$$

$$
\begin{aligned}
\lambda_j^2\frac{d}{dt}(x_j^2,\ y_j^2) &= \left(\frac{\partial H}{\partial y_j^2},\ \ -\frac{\partial H}{\partial x_j^2}\right) \\
&= \frac{1}{2}\sum_{i\neq j}^{m-n} \lambda_i^2\lambda_j^2 \left[\frac{1}{|\vec{x}_j^2-\vec{x}_i^2|} + \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{|\vec{x}_j^2-\vec{x}_i^2|}{L_R(\pi/\sqrt{8})}\right)\right]\left(x_i^2-x_j^2,\ y_i^2-y_j^2\right)^{\perp} \\
+\frac{1}{2}\sum_{i=1}^{m-n} \quad & \lambda_i^1\lambda_j^2 \left[\frac{1}{|\vec{x}_j^2-\vec{x}_i^1|} - \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{|\vec{x}_j^2-\vec{x}_i^1|}{L_R(\pi/\sqrt{8})}\right)\right]\left(x_i^1-x_j^2,\ y_i^1-y_j^2\right)^{\perp}.
\end{aligned}
\tag{4.5}
$$

These equations of motion are based on two tangential velocity profiles around a point vortex, namely $\frac{1}{r} \pm \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{r}{L_R(\pi/\sqrt{8})}\right)$; it is useful to compare them with the corresponding profile for a planar point vortex, i.e., $\frac{1}{r}$. The graph of $\frac{1}{r} + \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{r}{L_R(\pi/\sqrt{8})}\right)$ is similar to but dominates that for $\frac{1}{r}$. Thus, we have a $\frac{1}{r}$ type singularity at the origin and for $r \gg L_R(\pi/\sqrt{8})$,

$$
\frac{1}{r} + \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{r}{L_R(\pi/\sqrt{8})}\right) \to \frac{1}{r}
$$

from above. On the other hand, the graph of $\frac{1}{r} - \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{r}{L_R(\pi/\sqrt{8})}\right)$ is dominated by both $\frac{1}{r}$ and $\frac{1}{r} + \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{r}{L_R(\pi/\sqrt{8})}\right)$; it tends to 0 as r tends to zero, and for $r \gg L_R(\pi/\sqrt{8})$,

$$
\frac{1}{r} - \frac{1}{L_R(\pi/\sqrt{8})}K_1\left(\frac{r}{L_R(\pi/\sqrt{8})}\right) \to \frac{1}{r}
$$

from below (cf. figure 1(a) in [10]).

4.1. Hot and cold hetons and anti-hetons

A heton is a pair of point vortices of opposite signs, one in each layer. From (4.4) and (4.5), it is easy to see that an aligned heton is stationary, while a tilted heton with horizontal separation d between the vortices, translates at the speed

$$\lambda \left[\frac{1}{d} - \frac{1}{L_R(\pi/\sqrt{8})} K_1 \left(\frac{d}{L_R(\pi/\sqrt{8})} \right) \right]. \tag{4.6}$$

If the upper vortex is anticyclonic (negative vorticity) and the lower is cyclonic, then the tilted pair is called a hot heton, and will transport heat in a direction perpendicular to the direction of tilt at the speed in (4.6). If the signatures of the point vortices in the heton are reversed, then we have a cold heton, which will transport heat in a direction opposite to its propagation, again at speed (4.6). The rationale behind these labels is the following. Anticyclonic PV in the upper layer and cyclonic PV in the lower layer will both depress the interface between the layers yielding a warmer average temperature throughout the column; cyclonic PV in the upper layer and anticyclonic PV in the lower layer will both raise this interface yielding a colder average temperature throughout the column. Clearly there are other configurations for a pair of point vortices such as the anti-heton which consists of two like-signed vortices in different layers, but none of them can transport heat [10]. For example, equal like-signed vortices yield purely barotropic flow. Heton models have been used by DiBattista and the second author as the basis for equilibrium statistical theories to predict the spreading phase of open ocean convection [3].

We will compare formula (4.6) with the corresponding formula from the Coupled QG model in section 5.2. For this comparison, it must be kept in mind that the Heton model is based on two dimensional considerations, i.e., the stream functions in (4.1) and (4.2) are derived from the Green functions for a pair of two dimensional Lapace-Poisson problems. The Coupled QG model, on the other hand, is based on the Green functions of the three-dimensional Laplace-Poisson problems in (2.8) and (2.9). They have $\frac{1}{r}$ type singularities at the origin, unlike the logarithmic singularities in the barotropic component $\ln r$ and baroclinic component $K_0(\frac{r}{L_R(\pi/\sqrt{8})})$ of the Heton model.

We should keep in mind that the special case of the Coupled QG model involving purely interior PV and studied by Gryanik [8], is most closely related to the Heton model. There is really no counterpart of the surface heat particles in the Heton model. Indeed, in [13], Legg and Marshall substitute an isopycnal sheet of PV just below the ocean surface in place of the surface potential density anomaly. They choose this upper sheet of PV to be cyclonic, and the lower PVs to be anticyclonic to model the stratification reduction in the cold chimney zone of open ocean convection. We introduce the Coupled QG model to avoid this ad-hoc substitution and return the surface potential temperature anomaly to its proper place in our model. In the event of a cold air mass over warmer water, the surface potential temperature anomaly θ in the Coupled QG model should be set negative to represent an increase in density, and a concurrent loss of buoyancy in the surface mixed layer.

5. Elementary solutions of the Coupled QG model and comparisons with the Heton model.

We will show that unlike the Heton model where only tilted hetons can transport heat, two surface heat particles, a tilted surface heat particle interior PV pair, and a pair of interior PVs of opposite vorticities at different levels can transport heat in the Coupled QG model. As discussed earlier, the latter corresponds most closely to the heton in the two layer model, and like hetons, they transport heat horizontally away from a deep convection zone for instance. The mixed pair turns out to model the coupled transport of heat from the air-water interface and the interior of the fluid.

5.1. Two surface particles

Starting from the Hamilton equations of motion for the surface particles (3.8), we obtain the following expression for the rate of *change* of the distance between two particles of strengths θ_1 and θ_2

$$
\begin{aligned}
&\frac{d}{dt}\left[(q_1-q_2)^2+(\frac{p_1}{\theta_1}-\frac{p_2}{\theta_2})^2\right]\\
= \;& 2\left[(q_1-q_2)\left(\frac{d}{dt}q_1-\frac{d}{dt}q_2\right)+(\frac{p_1}{\theta_1}-\frac{p_2}{\theta_2})\left(\frac{1}{\theta_1}\frac{d}{dt}p_1-\frac{1}{\theta_2}\frac{d}{dt}p_2\right)\right]\\
= \;& 2\left[\begin{array}{c}(q_1-q_2)\left(\frac{\theta_2(\frac{p_1}{\theta_1}-\frac{p_2}{\theta_2})+\theta_1(\frac{p_1}{\theta_1}-\frac{p_2}{\theta_2})}{\left((q_1-q_2)^2+(\frac{p_1}{\theta_1}-\frac{p_2}{\theta_2})^2\right)^{3/2}}\right)\\ -(\frac{p_1}{\theta_1}-\frac{p_2}{\theta_2})\left(\frac{\theta_2(q_1-q_2)+\theta_1(q_1-q_2)}{\left((q_1-q_2)^2+(\frac{p_1}{\theta_1}-\frac{p_2}{\theta_2})^2\right)^{3/2}}\right)\end{array}\right]\\
= \;& 0.
\end{aligned}
$$

Thus, no matter what the strengths of the heat particles they move in such a way as to keep the same distance apart. In other words it is enough to compute the change in the relative angle between the particles to completely determine the motion of a pair of surface particles. To summarize, the separation d between two heat particles is a conserved quantity which means that two such particles of any strengths are a completely integrable system, and they form a relative equilibrium.

A further calculation gives the rate of change of the vector separation between two surface particles $\Big((q_1 - q_2),\ (\frac{p_1}{\theta_1} - \frac{p_2}{\theta_2})\Big)$ to be

$$
\begin{aligned}
\frac{d}{dt}\Big((q_1 - q_2), (\frac{p_1}{\theta_1} - \frac{p_2}{\theta_2})\Big) &= \begin{pmatrix} \frac{\theta_2(\frac{p_1}{\theta_1} - \frac{p_2}{\theta_2}) + \theta_1(\frac{p_1}{\theta_1} - \frac{p_2}{\theta_2})}{\left((q_1-q_2)^2 + (\frac{p_1}{\theta_1} - \frac{p_1}{\theta_2})^2\right)^{3/2}}, \\ \frac{-\theta_2(q_1-q_2) - \theta_1(q_1-q_2)}{\left((q_1-q_2)^2 + (\frac{p_1}{\theta_1} - \frac{p_1}{\theta_2})^2\right)^{3/2}} \end{pmatrix} \qquad (5.1) \\
&= \frac{(\theta_1 + \theta_2)}{\left((q_1 - q_2)^2 + (\frac{p_1}{\theta_1} - \frac{p_1}{\theta_2})^2\right)^{3/2}} \left(\frac{p_1}{\theta_1} - \frac{p_2}{\theta_2},\ \ -(q_1 - q_2)\right) \\
&= \frac{(\theta_1 + \theta_2)}{\left((q_1 - q_2)^2 + (\frac{p_1}{\theta_1} - \frac{p_1}{\theta_2})^2\right)^{3/2}} \left((q_1 - q_2), \frac{p_1}{\theta_1} - \frac{p_2}{\theta_2}\right)^{\perp}.
\end{aligned}
$$

Thus, we may conclude that a pair of equal surface particles of strength θ and initial separation d rotates about their common center at angular rate equal to

$$\frac{2\theta}{d^2}. \qquad (5.2)$$

In the cases where the surface particles have unequal temperatures, we find a variety of simple dynamical motions. For instance, when $\theta_1 = -\theta_2$, the above equation (5.1) reduces to

$$\frac{d}{dt}\Big((q_1 - q_2), (\frac{p_1}{\theta_1} - \frac{p_1}{\theta_2})\Big) = (0, 0).$$

This implies that two oppositely signed surface particles travel in a straight line perpendicular to the line that joins them, at speed equal to

$$v = \frac{\theta_2}{d^2}. \qquad (5.3)$$

Intermediate values of the two heat strengths θ_1 and θ_2 yield a family of curvilinear trajectories in which the separation d between the particles is held constant.

A comparison of the speeds (5.2) and (5.3) with the corresponding speeds in the planar point vortex model show that the singularity at the origin is stronger than the $O(\frac{1}{r})$ singularity in the latter model.

5.2. One surface and one interior particle

We now study the dynamics of a tilted pair which consists of a heat particle and a interior vortex located at depth $z = -h$. The case of an aligned pair consisting of a heat particle located vertically over a interior vortex will be mentioned in section 5.4. Our calculations show that such an aligned pair is not an equilibrium but rather a singular point of the vector field in the sense that the norm of the vector field blows up at such points in phase space.

We set the value of the surface potential temperature to be θ and the PV to be λ. Let $\vec{x} = (x, y)$ denote the position of the interior vortex and $\vec{\alpha} = (\alpha, \beta)$ denote the position of the heat particle. The equations of motion for this pair are given by the Hamiltonian system

$$\begin{aligned}
\frac{d}{dt}x &= \frac{\theta}{2\pi}\left\{\frac{y-\beta}{|\vec{x}-\vec{\alpha}|^3} + \frac{y-\beta}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{3/2}}\right\}, \\
\frac{d}{dt}y &= \frac{-\theta}{2\pi}\left\{\frac{x-\alpha}{|\vec{x}-\vec{\alpha}|^3} + \frac{x-\alpha}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{3/2}}\right\}, \\
\frac{d}{dt}\alpha &= -\frac{\lambda}{2\pi}\left\{\frac{y-\beta}{|\vec{x}-\vec{\alpha}|^3} + \frac{y-\beta}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{3/2}}\right\}, \\
\frac{d}{dt}\beta &= \frac{\lambda}{2\pi}\left\{\frac{x-\alpha}{|\vec{x}-\vec{\alpha}|^3} + \frac{x-\alpha}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{3/2}}\right\},
\end{aligned} \tag{5.4}$$

with the Hamiltonian function

$$H(\vec{x}, \vec{\alpha}) = -\frac{\theta\lambda}{2\pi}\left\{\frac{1}{|\vec{x}-\vec{\alpha}|} + \frac{1}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{1/2}}\right\}.$$

In order to calculate the effect of having a tilted pair with $\theta \neq -\lambda$, we will compute the following expression from (5.4),

$$\begin{aligned}
&\frac{d}{dt}(x-\alpha,\ y-\beta) \\
&= (\theta+\lambda)\begin{pmatrix} \left\{\frac{y-\beta}{|\vec{x}-\vec{\alpha}|^3} + \frac{y-\beta}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{3/2}}\right\}, \\ -\left\{\frac{x-\alpha}{|\vec{x}-\vec{\alpha}|^3} + \frac{x-\alpha}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{3/2}}\right\} \end{pmatrix} \\
&= (\theta+\lambda)\left[\frac{1}{|\vec{x}-\vec{\alpha}|^3} + \frac{1}{[(\alpha-x)^2+(\beta-y)^2+h^2]^{3/2}}\right](x-\alpha,\ y-\beta)^{\perp},
\end{aligned} \tag{5.5}$$

which gives the rate of change of the horizontal components of the vector joining the heat particle to the vortex. From (5.5) we conclude that the length $|\vec{x}-\vec{\alpha}|$

of the relative co-ordinate is conserved and $\vec{x} - \vec{\alpha}$ rotates with a constant angular velocity. Also the linear momentum $\lambda\vec{x} + \theta\vec{\alpha}$ is conserved as a consequence of (3.6) so that the equations in (5.4) are completely integrable.

In the case of $\theta = -\lambda$ we compute the speed of rigid translation in a straight line to be

$$\begin{aligned} v &= \frac{\lambda d}{2\pi}\left\{\frac{1}{d^3} + \frac{1}{(d^2+h^2)^{3/2}}\right\}, \\ d &= |\vec{x} - \vec{\alpha}|. \end{aligned} \tag{5.6}$$

Thus, for example, a pair in which the heat particle is directly south of the vortex will propagate westward, and so on. In these examples, we also note that the speed of propagation is independent of the orientation of the pair but depends only on the horizontal separation d and the vertical separation h. We note also that for a fixed nonzero value of the horizontal separation d, the velocity v in (5.6) tends to $\frac{\lambda}{2\pi d^2}$ as h tends to ∞ and also v tends to $\frac{\lambda}{\pi d^2}$ as h tends to 0 which is compatible with (5.2) for two surface heat particles. This case, with $\theta = -\lambda$, is the analogue for a surface heat particle coupled with an interior vortex of the hot and cold propagating hetons for the two-layer model discussed in section 4.1 above. The above exact solutions reveal significant differences in propagation velocity by comparing (4.6) and (5.6).

We note that equation (5.5) implies that in the cases $\theta \neq -\lambda$, the tilted pair does not propagate in a straight line but instead in a horizontally curvilinear trajectory while its horizontal relative vector rotates at angular velocity

$$v_\phi = (\theta + \lambda)\left[\frac{1}{d^2} + \frac{d}{[d^2+h^2]^{3/2}}\right]. \tag{5.7}$$

It is noteworthy that this expression for the angular velocity depends on the strengths of the particles only in the factor $(\theta+\lambda)$. These interaction velocities have a $O(d^{-2})$ singularity at the origin and a $O(d^{-2})$ decay at infinity in the horizontal separation d. Comparing (5.6) and (5.7) with the corresponding formula (4.6) for the Heton model, we note the significant difference in the $O(d)$ behaviour at the origin and the $O(d^{-1})$ decay at infinity in (4.6). Thus, the interaction between a surface heat particle and a interior PV is stronger at close range than both the interactions of the Heton model and the planar point vortex model for the 2-D Euler equation but decays faster at long range than both these models. Just like the Heton model, the speeds (5.6) and (5.7) of horizontal propagation of the tilted heat particle-interior vortex pair give us a measure of the rate of heat transfer in the Coupled QG model. The same remarks apply for the dipole pair in (5.6).

5.3. Two interior point vortices

Gryanik [8] has reported on the dynamics of a bath of interior point vortices in the quasi-geostrophic approximation. Our calculations in this subsection follow his work and are given here for completeness in order to make comparisons with both the surface/interior QG hetons in section 5.2 and the two-layer hetons from section 4.1. In the case of two interior point vortices there are obviously two interesting subcases, namely (a) both vortices are at the same depth h, and (b) the vortices are at different depths z_1 and z_2, and moreover, they are not vertically aligned. The vertically aligned case is an equilibrium and will be briefly mentioned in the next subsection. From (3.9), we deduce the equations of motion for case (a):

$$\begin{aligned} \frac{d}{dt}x_1 &= -\frac{\lambda_2 D}{4\pi}(y_1 - y_2) \\ \frac{d}{dt}y_1 &= \frac{\lambda_2 D}{4\pi}(x_1 - x_2) \\ \frac{d}{dt}x_2 &= -\frac{\lambda_1 D}{4\pi}(y_2 - y_1) \\ \frac{d}{dt}y_2 &= \frac{\lambda_1 D}{4\pi}(x_2 - x_1), \end{aligned} \tag{5.8}$$

where

$$\begin{aligned} D &= \left(\frac{1}{[(x_1 - x_2)^2 + (y_1 - y_2)^2]^{3/2}} + \frac{1}{[(x_1 - x_2)^2 + (y_1 - y_2)^2 + 4z^2]^{3/2}}\right) \\ &= \left(\frac{1}{d^3} + \frac{1}{(d^2 + 4z^2)^{3/2}}\right). \end{aligned} \tag{5.9}$$

A further calculation yields

$$\begin{aligned} \frac{d}{dt}(x_1 - x_2,\ y_1 - y_2) &= \frac{(\lambda_1 + \lambda_2)D}{4\pi}\left(-(y_1 - y_2),\ (x_1 - x_2)\right) \\ &= \frac{(\lambda_1 + \lambda_2)D}{4\pi}(x_1 - x_2,\ y_1 - y_2)^{\perp}. \end{aligned}$$

This implies that two vortices on the same level behave much like two planar point vortices except for the range and the singular strength of the interaction. Here the interaction velocities have a $O(d^{-2})$ singularity at the origin and a $O(d^{-2})$ decay at infinity. The Heton model predicts that for point vortices in the same layer, the interaction has a $O(d^{-1})$ singularity at the origin and a $O(d^{-1})$ decay at infinity. Thus, the interior PVs have a weaker long range interaction than the Heton model and planar point vortex model, but it has a stronger singularity at the origin. For example two vortices of opposite sign will travel in a straight line perpendicular to the line joining them at speed equal to

$$v = \frac{D}{4\pi} d. \tag{5.10}$$

Two particles of the same strengths λ will rotate about their center at angular velocity

$$v_a = \frac{\lambda D}{2\pi} d. \tag{5.11}$$

For arbitrary strengths the two vortices travel rigidly as a relative equilibria on a curvilinear trajectory. Remarkably, the angular rate of rotation of the vector joining two vortices of any strengths has the same form as (5.11) except for the vortex strengths, i.e.,

$$v_a = \frac{(\lambda_1 + \lambda_2) D}{4\pi} d. \tag{5.12}$$

For case (b) the equations are given by (5.8) but the expression for the factor D takes the form

$$\begin{aligned} D &= \left(\begin{array}{c} \frac{1}{[(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2]^{3/2}} \\ +\frac{1}{[(x_1-x_2)^2+(y_1-y_2)^2+(z_1+z_2)^2]^{3/2}} \end{array} \right) \\ &= \left(\frac{1}{(d^2+(z_1-z_2)^2)^{3/2}} + \frac{1}{(d^2+(z_1+z_2)^2)^{3/2}} \right). \end{aligned} \tag{5.13}$$

Thus, the analyses for case (a) carries over to case (b). In particular, two vortices of arbitrary strengths on different levels move horizontally on a curvilinear trajectory as a relative equilibria, keeping their horizontal separation fixed. The only difference is in the rate of translation and angular velocities which retain the forms in (5.10) and (5.12), except that D now takes the form in (5.13). This means that two vortices which are not vertically aligned on two different levels have somewhat weaker interactions than their counterparts on the same level; they differ by the extra factor of $(z_1 - z_2)^2$ in the first term of D. This crucial difference implies that the interaction in this case is regular in the sense that as the horizontal separation d tends to 0, the speeds (5.10) and (5.12) tends to 0.

Thus two interior PV vortices at different levels interact much like two point vortices in different layers in the Heton model. The case of opposite vorticities corresponds to the heton while the case of like-signed vorticities correspond to the anti-heton case. Much like (4.6) for the hetons and anti-hetons, the speeds of propagation (5.10) and (5.12) in the case of two vertically separated interior PV vortices, tends to zero as the horizontal separation d tends to zero and ∞. However, the speed here is $O(d)$ as $d \to 0$ and $O(d^{-2})$ as $d \to \infty$, unlike the $O(d^{-1})$ decay at infinity in (4.6). The interaction between PV at different levels is weaker in the long range than in the Heton model or the planar point vortex model. Hogg and Stommel [10], [11] numerically computed the behaviour of two hetons, and showed a variety of dynamics, some of which led to an exchange of particles between the hetons.

5.4. Equilibria and singularities of the equations of motion

The following equilibria of (3.4) can be easily verified from the equation of motion and the details are left to the reader: any number of vertically aligned interior vortices is an equilibrium. On the other hand, we observe that there are also the following singular points: any number of interior vortices and a heat particle aligned vertically. A noteworthy special case of the latter is a vertically aligned heat particle over a interior PV. There are two distinguished subcases, namely (a) the heat particle and PV have opposite signature and (b) they have the same signature. Since these configurations are associated with singular points of the Hamiltonian vector field (and not equilibria), they are highly unstable to perturbations. In case (a), any slight misalignment of the pair will send it moving rapidly in a horizontal direction perpendicular to the line of tilt. As for case (b), a slight misalignment will result in a rapidly spinning pair. The vertically aligned pairs of type (b) are clearly more stationary than the pairs of type (a).

6. Propagating tilted Clusters for the Coupled QG and Heton models

In this section we will set up the KAM machinery [1] to show that propagating tilted clusters are long-lived objects that are supported by both the coupled QG and Heton models. In [10], Hogg and Stommel proposed the heton as a candidate for heat transport, and showed that they propagate at the speeds in (4.6). Their numerical results showed hetons dispersing from a central bath, and also indicated that pairs of hetons can exchange particles in near collision interactions. More recently Legg and Marshall [13] used the Heton model to study open ocean deep convection. The main motivation for our work in this section is the numerical results of Legg and Marshall [13] on the Heton model which provided strong evidence that propagating tilted clusters are a characteristic phenomenon in the spreading phase of a localized cooling event. By introducing point vortices of opposite signatures in two layers in a fixed localized region at a fixed cooling rate, they showed that initially these point vortices are constrained to move in the convection zone by a rim current, and later after the rim current had decayed by baroclinic instability, clusters of point vortices form in each layer, which then self-organize into propagating tilted cluster pairs. While individual vortices in the cluster interact and change their phases, these overall cluster pairs move at overall speeds close to (4.6), and serve to transport cold water laterally away from the convection zone. Here we establish the existence of large families of solutions with precisely this structure.

The Coupled QG model that we introduced in this paper has, as was shown earlier in section 5, two distinct elementary modes of heat transport, namely (a) a tilted hetonic pair of opposite signature interior PVs (see section 5.3), and (b) a tilted surface heat particle - interior PV pair of opposite signature (see section 5.2 and (5.6)). Mode (a) is nearest to the heton in the Heton model, and can be used to give additional vertical resolution in the modelling of heat transport in the spreading phase. Both modes can be generalized to increase the vertical resolution of the model. Instead of the pair of particles in mode (b), we could consider a tilted pair of clusters, where the upper cluster consists of heat particles, and the lower cluster consists of PVs of opposite signature at possibly different depths. Instead of a hetonic pair of PVs in mode (a), we could consider a tilted pair of clusters of PVs, where each cluster consists of PVs of like sign at possibly different depths. Here for the coupled QG model, we concentrate on establishing the existence of tilted clusters in situation b). The physically significant consequence of our KAM analyses is the fact that these tilted pairs of clusters are long-lived objects that propagate at speeds close to the respective horizontal speeds of the pure modes (a) and (b). Thus, these tilted cluster pairs transport heat just like their respective pure mode counterparts, but the clusters now provide valuable additional physical resolution in the modelling process.

The main results in this paper are now stated in the form of two theorems, for the Heton model and Coupled QG model respectively.

Theorem 1: The Heton point vortex model for two-layer quasi-geostrophic flows supports long-lived propagating tilted pairs of clusters where the upper cluster consists of k positive potential vortices and the

lower cluster consists of $k \pm \delta$ negative potential vortices of the same strength, such that the integer δ has values 1 or 2.

Theorem 2: The Hamiltonian system $H_\infty = H_I + H_B + H_{IB}$ for coupled surface-interior quasi-geostrophic flows supports long-lived propagating tilted pairs of clusters where the surface cluster consists of m temperature particles, and the interior cluster consists of $m \pm \delta$ potential vortices of opposite strength, such that the integer δ has value 1.

The necessary background for the KAM machinery, such as the combinatorial symplectic transformations that will be used in both models are included in a self-contained appendix for the sake of completeness. Basic concepts from combinatorics that are relevant to the discussion in this section and in the appendix will be discussed next.

6.1. Combinatorics of Jacobi variables and the KAM theory

The application of combinatorial concepts in $N-$*body problems* is based on three key ideas: (i) Hamiltonian functions of $N-$*body* problems can be conveniently viewed as functions on certain graphs, (ii) a special set of canonical transformations (known as Jacobi transformations) is generated by a combinatorial algorithm defined on the class of binary trees, and (iii) the new Hamiltonians (after the Jacobi transformations) which are now functions on binary trees, can be written in the perturbation ansatz $H = H_0 + H_1$, where the completely-integrable term H_0 is based on certain subgraphs of the binary trees. We will now go into the details of each of these key ideas.

The first idea is based on the fact that the Hamiltonian functions we are analyzing, are essentially sums over all distinct pairs of particles, provided that the interaction in the problem is of two-body type. Now such Hamiltonians H can be represented as being defined on a complete graph K_N which is a graph with N vertices and $\frac{N(N-1)}{2}$ edges. For every pair of vertices x, y in K_N, there is an associated edge $e = (x, y)$; hence the name complete graph. Each vertex x_i in K_N represents a particle $\{\lambda_i,\, z_i\}$ in the N-body problem, where λ_i is the weight associated with particle i and z_i is its position in the physical domain D which is a subset of either the plane, sphere or R^3. Another way to see this is to treat the vertices x_i in K_N as the original canonical variables $\{\lambda_i,\, z_i\}$ in the problem. Each edge e in K_N represents a term in the sum that is the Hamiltonian H. The special nature of the two-body interaction in any particular problem is represented then by the form of the Green function F in the Hamiltonian

$$H = \sum_{i \neq j} \lambda_i \lambda_j F(z_i, z_j). \tag{6.1}$$

These weights are real numbers or even real vectors representing for instance the vorticity, potential vorticity or the potential temperature and density of the Lagrangian particle. It is usual to study Hamiltonians which have at least translational symmetry, in which case,

$$F(z_i, z_j) = F(z_i - z_j).$$

If moreover, H has rotational symmetry as well, then

$$F(z_i, z_j) = F(|z_i - z_j|).$$

These symmetries imply that it is advantageous to analyse the Hamiltonian system H using relative coordinates rather than the original canonical variables which are absolute coordinates.

The second key idea in $N - body$ problems is to be found in the area of the ubiquitous canonical or symplectic transformations. A special class of linear canonical transformations to essentially relative coordinates in $N - body$ problems have been used in Celestial Mechanics since the days of Jacobi [28], and has recently been generalized and adapted for other $N - body$ problems like those of first order type in vortex dynamics [16], [17], [19], [12]. It turns out that this special class of transformations can be represented by the class of binary trees with N terminal vertices or leaves. A tree is a graph which is connected (has no isolated pieces) and has no circuits where a circuit is a simple path in a graph, consisting of consecutive adjacent vertices that return to the starting vertex. Associated with a given tree T and a distinguished vertex called the root r, is a partial order relation $<$ between the vertices in $T : x < y$ if the path from x to the root r passes through y. In this partial order the root r of the tree T is highest in rank, and the terminal vertices or leaves of T are the lowest. A vertex in a tree T which is not a leaf is called an internal vertex. A binary tree is a tree where each internal vertex v has exactly two vertices v_l and v_r directly below it in this partial order; they are called the left child v_l and right child v_r of v. A left (right) descendent of a given internal vertex v in a binary tree T is a vertex $w < v$ such that the last vertex before v in the path from w to v is a left (resp. right) child v_l (resp. v_r) of v.

Now the Jacobi transformations are generated by an algorithm (given in detail in the appendix) which takes the original canonical variables in H, namely the pairs $\{\lambda_i,\, z_i\}$ associated with vertices x_i in K_N, and sends them to new canonical variables $\{\Gamma_j, Z_j\}_{j=1}^{N-1}$ which are now associated with the internal or non-terminal vertices v_j, $j = 1, .., N-1$ of a binary tree $T(N)$ whose N leaves are just the N vertices of K_N. For any given choice of binary tree $T(N)$ with N leaves and $N-1$ internal vertices, there belongs a unique Jacobi canonical transformation. Since there are the Catalan numbers $C(N)$ distinct unlabelled binary trees with N leaves, there are in principle $C(N)$ Jacobi transforms for each $N-$*body* Hamiltonian of the type in (6.1). It will be clear from the warm-up example in the next subsection that these new canonical variables known as Jacobi variables are generalized relative coordinates. Let us consider a pair of vertices $x_i = \{\lambda_i,\, z_i\}$ and $x_j = \{\lambda_j,\, z_j\}$ in K_N which are also two leaves v_l and v_r in a specific tree $T(N)$, such that x_i is the left child v_l and x_j is the right child v_r of an internal vertex v in $T(N)$. The Jacobi algorithm uses the binary structure of $T(N)$ in the following way. Construct the new canonical variable $\{\Gamma, Z\}$ for internal vertex v to be the pair

$$\{\Gamma, Z\} = \left\{ \frac{\lambda_i \lambda_j}{\lambda_i + \lambda_j}, \frac{\lambda_j z_j - \lambda_i z_i}{\lambda_i + \lambda_j} \right\}. \tag{6.2}$$

The second component $Z = \frac{\lambda_j z_j - \lambda_i z_i}{\lambda_i + \lambda_j}$ is clearly a relative variable, and has the dimension of length. The first component $\Gamma = \frac{\lambda_i \lambda_j}{\lambda_i + \lambda_j}$ is the generalized weight of the new canonical variable, and in physical terms, it is the weight assigned to the "center of mass" of the pair of particles $\{\lambda_i,\, z_i\}$ and $\{\lambda_j,\, z_j\}$. It is already abundantly clear from (6.2) that the Jacobi construction fails if $\lambda_i + \lambda_j = 0$. In fact, there are $N-1$ such admissibility conditions (A.3) on the original weights λ_j, $j = 1, ..., N$ of the problem, which are necessary for the Jacobi transformation to be well-defined; they are naturally associated with the $N-1$ internal vertices of $T(N)$ (cf. theorem 1 and the second remark in the appendix.) These admissibility conditions can be interpreted in terms of mismatches in the number of vortices between the two layers in the heton clusters, if point vortices of equal numerical strength are used in the model.

6.1.1. KAM framework

Arnold's form of the KAM theorem [1] will be used to establish the existence of invariant tori in phase space which are continuations of the tori $S^1 \times .... \times S^1$ of a completely integrable Hamiltonian $H_o(T(n,n))$. The full Hamiltonian in the problem is to be viewed as a small perturbation of $H_o(T(n,n))$ in special regions M of phase space. Regions M in phase space correspond to tilted cluster configurations where the particles in the clusters are much closer to each other than the horizontal distance d between the clusters.

We will use the special set of coordinates discussed in the appendix to rigorously prove that there are positive measure sets in phase space which are invariant, and where each phase point corresponds to a tilted cluster configuration with separated clusters of nearly equal numbers of vortices. While it is not true that all arrangements of the particles in each cluster will be maintained in the tilted cluster configuration forever, a certain proportion of these arrangements will be and others that are not maintained forever will nonetheless be maintained for exponentially long times by Nekhoroshev's estimates [1].

For simplicity, we will focus on the case of $m = 2n \pm 1$ particles, where the top cluster has n particles and the bottom cluster has $n \pm 1$ particles. We will show that a fundamental and simple result from the theory of canonical transformations of the Jacobi type [17] has the surprising consequence that the two clusters in the problem which correspond to a KAM torus in the context of the combinatorial perturbation method (CPM), must have a mismatch in particle numbers. For these configurations, the optimal Jacobi variables are derived from the binary tree $T = T(n, n \pm 1)$. The root r in this binary tree has n right descendents and $n \pm 1$ left descendents that are leaves.

The third key idea in our application of combinatorics to KAM dynamics is based on the observation that the combinatorial Jacobi variables introduced above via the binary tree T (and given in full in the appendix) can be used to write the original Hamiltonian H in perturbation form,

$$H = H_o(T) + H_1(T).$$

With this ansatz in mind, we will construct a completely integrable Hamiltonian function $H_o(T)$ on the basis of the binary tree $T = T(n, n \pm 1)$, and show that the difference $H_1(T) = H - H_o(T)$ is small for clustered configurations. Both H_0 and H_1 depend on the combinatorial structure of the binary tree T, in particular the partial order relation $<$ induced by T, through the set of Jacobi canonical variables. In fact, once the tree T is fixed, the forms of H_0 and H_1 are completely determined by parameters which are functions only of the original weights λ_j of the particles. Clearly the independence of these parameters in H_0 and H_1 from the values of the canonical variables is a necessary condition for these perturbation Hamiltonians to be well-defined (cf. the first remark in the appendix).

The well-known KAM nondegeneracy conditions [1] turn out to have simple forms in terms of the combinatorial structure of T. Moreover, the set of $m-1$ necessary and sufficient conditions (A.3) on the original weights λ_j of the particles in order for the Jacobi variables based on T to be well-defined turns out to be exactly the conditions required by KAM nondegeneracy or twist. This is a subtle but important point in the CPM that we are proposing here. It means that the Jacobi variables based on binary trees T are in some sense optimal for the existence proof of long-lived tilted clusters because the KAM-twist conditions do not impose additional constraints on the weights of the particles, beyond those already in place from the Jacobi transformation. Moreover, it is surprising and interesting that in the context of the CPM, the KAM twist conditions should have a physical consequence in the mismatch of vortex numbers between the two layers in the heton model.

Before we work out the details of the CPM for the Heton model, we will present a warm-up example involving two particles of one species and three particles of the other species; it is based on the simple logarithmic Hamiltonian for planar point vortex dynamics.

6.1.2. Simple 5 particles example

Let the Hamiltonian function for two species of point vortices with vorticities λ and $-\lambda$ be given by

$$H = \lambda^2 \sum_{i=1}^{3} \sum_{j\neq i=1}^{3} \ln|\vec{x}_i - \vec{x}_j| + \lambda^2 \sum_{i=4}^{5} \sum_{j\neq i=4}^{5} \ln|\vec{x}_i - \vec{x}_j| - \lambda^2 \sum_{i=1}^{3} \sum_{j=4}^{5} \ln|\vec{x}_i - \vec{x}_j|. \quad (6.3)$$

The symplectic variables are

$$(q_i = x_i,\, p_i = \sqrt{\lambda} y_i$$

for $i = 1, 2, 3$; and

$$(q_i = x_i,\, p_i = -\sqrt{|\lambda|} y_i$$

for $i = 4, 5$, where $\vec{x}_i = (x_i,\, y_i) \in R^2$. Vortices 4 and 5 have strengths $-\lambda$. The Hamiltonian function (6.3) can be viewed as a function which is defined on the complete graph K_5 on five vertices, where the vertices correspond to the symplectic variables (q_i, p_i) and the edges correspond to the terms in the three sums of (6.3).

The special symplectic transformation described in the appendix is based on a suitable binary tree. We will use the tree $T(3, 2)$ depicted in figure 1, which is the most balanced tree with two leaves (terminal vertices) labeled 4 and 5 on the left branch and three leaves labeled $1, 2, 3$ on the right branch at the root which is labeled $4'$. Internal (non-leaf) vertices are labeled $1', 2', 3'$ and $4'$. By inspection, it is clear that there is a partial order relation induced by $T(3,2)$ on the set of internal vertices. In this case, we have

$$1' < 2',\ 2' < 4',\ 3' < 4'.$$

Following the expressions (A.1) and (A.2) in the appendix, the set of Jacobi variables for this specific example is completely defined by the structure of $T(3,2)$ and the weights $\{\lambda, \lambda, \lambda, -\lambda, -\lambda\}$, and given by

$$\begin{aligned} Q_1 &= q_1 - q_2, \\ Q_2 &= q_3 - \frac{1}{2}(q_1 + q_2), \\ Q_3 &= q_4 - q_5, \\ Q_4 &= \frac{1}{3}(q_1 + q_2 + q_3) - \frac{1}{2}(q_4 + q_5); \end{aligned} \tag{6.4}$$

and

$$\begin{aligned} P_1 &= \frac{1}{2}(p_1 - p_2), \\ P_2 &= \frac{2}{3}p_3 - \frac{1}{3}(p_1 + p_2), \\ P_3 &= \frac{1}{2}(p_4 - p_5), \\ P_4 &= -2(p_1 + p_2 + p_3) - 3(p_4 + p_5). \end{aligned}$$

As discussed above, this canonical transformation from the original variables (q_i, p_i) to the new variables (Q_j, P_j) can be viewed as a transformation from the complete graph K_5 to the binary tree $T(3,2)$, because the internal vertices of the tree correspond to the new canonical variables while the original variables are associated with the leaves. In terms of the new variables (6.4), and after a further transformation to action-angle variables $(J_k,\ \Psi_k)$ where

$$(Q_k, P_k) = J_k \exp(i\Psi_k),$$

the Hamiltonian H takes the form H' given below

$$H' = \lambda^2 \sum_{i=1}^{3} \sum_{j \neq i = 1}^{3} \ln |P(i,j)| + \lambda^2 \sum_{i=4}^{5} \sum_{\substack{j \neq i \\ j=4}}^{5} \ln |P(i,j)| - \lambda^2 \sum_{i=1}^{3} \sum_{j=4}^{5} \ln |P(i,j)|,$$

where $P(i,j)$, the geometrical path between leaf i and j, is given by the following linear combination

$$\vec{x}_i - \vec{x}_j = P(i,j) = \sum_{k\in(i,j)} b_k(i,j) J_k \exp(i\Psi_k) \tag{6.5}$$

in terms of the new canonical variables $(J_k,\ \Psi_k)$ corresponding to each internal vertex k in the combinatorial path (i,j) between the same pair of leaves. For example the geometrical path $P(1,4)$ between leaf 1 and 4 is given by

$$\begin{aligned} P(1,4) &= b_1(1,4)J_1\exp(i\Psi_1) + b_2(1,4)J_2\exp(i\Psi_2) \\ &\quad + b_3(1,4)J_3\exp(i\Psi_3) + b_4(1,4)J_4\exp(i\Psi_4). \end{aligned} \tag{6.6}$$

This new Hamiltonian function H' can therefore be viewed as defined on the tree $T(3,2)$ where the internal vertex correspond to the canonical variables $(J_k,\ \Psi_k)$, $j = 1,2,3,4$, and the set $P(T)$ of leaf-to-leaf paths (i,j) corresponds to the terms in H'.

To proceed with the combinatorial perturbation method (CPM), we will define the following subsets of $P(T)$: for each $s = 1,2,3,4$, let $P(s)$ be the set of leaf-to-leaf paths (i,j), whose highest node (according to the partial relation $<$) $s(i,j)$ is internal vertex s. These sets are given explicitly for $T(3,2)$:

$$\begin{aligned} P(1) &= \{(1,2)\},\ P(2) = \{(1,3),(2,3)\},\ P(3) = \{(4,5)\}, \\ P(4) &= \{(1,4),(1,5),(2,4),(2,5),(3,4),(3,5)\}, \end{aligned}$$

where the set $P(4)$ consists of the leaf-to-leaf paths that passes through the root vertex 4. We resume the terms in the Hamiltonian H' to construct H_0 and H_1 as follows: for each $s = 1,2,3,4$, collect the terms in H' according to the sets $P(s)$, and call them

$$H'_s = \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \ln|P(i,j)|.$$

They are given explicitly for $T(3,2)$ by:

$$\begin{aligned} H_1^{'} &= \lambda^2 \ln|P(1,2)|, \\ H_2^{'} &= \lambda^2(\ln|P(1,3)| + \ln|P(2,3)|), \\ H_3^{'} &= \lambda^2 \ln|P(4,5)|, \\ H_4^{'} &= -\lambda^2 \sum_{i=1}^{3}\sum_{j=4}^{5} \ln|P(i,j)|, \end{aligned}$$

where $P(i,j)$ denotes the geometrical path associated with the combinatorial path (i,j) between leaf i and leaf j, and $|P(i,j)|$ denotes the Euclidean length of $P(i,j)$. The resumed Hamiltonian function in this example is now given by

$$H' = H_1^{'} + H_2^{'} + H_3^{'} + H_4^{'},$$

and in general by

$$H' = \sum_{s=1}^{m-1} H_s^{'},$$

where the sum is taken over the set of internal vertex including the root node $m-1$ for a tree $T(m)$ with m leaves.

In the next step of the CPM, we will impose the cone conditions on the ratios of the actions J_s, i.e.,

$$\frac{J_k}{J_l} = \varepsilon \ll 1 \tag{6.7}$$

for all $k < l$ (according to the partial order $<$ induced by $T(3,2)$), to rewrite the paths $P(i,j)$ in (6.5) in the form

$$\begin{aligned} P(i,j) &= \sum_{k\in(i,j)} b_k(i,j) J_k \exp(i\Psi_k) \\ &= b_s(i,j) J_s \exp(i\Psi_s) \left[1 + \frac{\sum_{\substack{k\in(i,j)\\ k<s}} b_k(i,j) J_k \exp(i\Psi_k)}{b_s(i,j) J_s \exp(i\Psi_s)} \right]. \end{aligned}$$

From (6.6), we get the specific example

$$\begin{aligned} P(1,4) &= b_4(1,4) J_4 \exp(i\Psi_4) \\ &\times \left[1 + \frac{1}{b_4(1,4) J_4 \exp(i\Psi_4)} \left(\begin{array}{c} b_1(1,4) J_1 \exp(i\Psi_1) + b_2(1,4) J_2 \exp(i\Psi_2) \\ + b_3(1,4) J_3 \exp(i\Psi_3) \end{array} \right) \right]. \end{aligned}$$

Then

$$|P(i,j)| = |b_s(i,j)| J_s \left| 1 + \frac{\sum_{\substack{k\in(i,j)\\ k<s}} b_k(i,j) J_k \exp(i\Psi_k)}{b_s(i,j) J_s \exp(i\Psi_s)} \right|,$$

and by (6.7) and the definition of $P(s)$, we have for each $s = 1,2,3,4$,

$$H_s^{'} = \sum_{(i,j)\in P(s)} \lambda_i \lambda_j \ln |P(i,j)|$$

$$
\begin{aligned}
&= \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \ln\left(|b_s(i,j)|J_s\left|1+\frac{\sum_{\substack{k\in(i,j)\\k<s}} b_k(i,j)J_k\exp(i\Psi_k)}{b_s(i,j)J_s\exp(i\Psi_s)}\right|\right)\\
&= \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \ln J_s + \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \ln\left(\left|1+\frac{\sum_{\substack{k\in(i,j)\\k<s}} b_k(i,j)J_k\exp(i\Psi_k)}{b_s(i,j)J_s\exp(i\Psi_s)}\right|\right)\\
&= \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \ln J_s + \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \sum_{\substack{k\in(i,j)\\k<s}} \frac{|b_k(i,j)|J_k}{|b_s(i,j)|J_s} + h.o.t.\\
&= H_s^0 + H_s^1.
\end{aligned}
$$

Collecting all the terms H_s^0 we get the completely decoupled nonlinear oscillator term

$$
\begin{aligned}
H_0 &= \sum_{s=1}^4 H_s^0 = \sum_{s=1}^4 \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \ln J_s\\
&= \lambda^2 \sum_{s=1}^3 \left(\sum_{(i,j)\in P(s)} \ln J_s\right) - \lambda^2 \sum_{(i,j)\in P(4)} \ln J_4\\
&= \lambda^2 \sum_{s=1}^3 |P(s)|J_s - \lambda^2 |P(4)| \ln J_4\\
&= \lambda^2 (\ln J_1 + 2\ln J_2 + \ln J_3) - 6\lambda^2 \ln J_4,
\end{aligned}
$$

and the remainder H_1 is then given by

$$
H_1 = \sum_{s=1}^4 H_s^1 = \sum_{(i,j)\in P(s)} \lambda_i\lambda_j \sum_{\substack{k\in(i,j)\\k<s}} \frac{|b_k(i,j)|J_k}{|b_s(i,j)|J_s} + h.o.t.
$$

H_1 is of order ε on the cone sets in which (6.7) holds because by explicit calculation in this simple example (and by the first remark in the appendix in the general case), we know that the Jacobi coefficients $|b_k(i,j)|$ depend only on the weights $\{\lambda,\lambda,\lambda,-\lambda,-\lambda\}$ once the tree is fixed. In other words, when we change the ratios ε in (6.7), these coefficients remain fixed.

Finally it is easy to check that the KAM nondegeneracy condition on the Hessian of H_0 is satisfied since the decoupled form of H_0 implies that the Hessian matrix for H_0 is diagonal.

6.2. Proof of Theorem 1 for the Heton model

The detail vortex count in the introduction suggests that the upper and lower clouds in propagating heton clusters differ in number by one or more vortices. We will show below that our method of proof, the CPM, requires such a mismatch of numbers between the upper and lower cluster. It is possible however, that coordinates besides the above Jacobi variables could be found and used in a different KAM analysis of the heton model to prove the existence of long-lived propagating heton clusters which do not require this mismatch.

We will consider the following setup:- a tightly clustered cloud of n equal vortices in the upper layer is offset by an almost equal number of tightly clustered vortices of opposite circulation in the lower layer. At least initially the center of vorticity of the upper cloud is horizontally offset from the center of vorticity of the lower cloud. Thus, the initial position of these two centers of vorticity is that of a tilted heton with circulation $n\lambda$, where λ is the circulation of each point vortex in the upper cloud. Since an isolated tilted heton is a relative equilibria of the equations of motion which translates rigidly at speed (4.6), it is an easy consequence of the cone condition that these tightly clustered clouds of opposite circulations will propagate macroscopically as a tilted heton at speeds near (4.6). The KAM dynamics predicts that the individual vortices in each cloud are expected to mill around their respective centers of vorticity forever.

Remark 1. *In the Tilted Heton problem, the weights are λ and $-\lambda$, and one of the full binary trees that will be used is the balanced tree $T(n,n)$ for $m = 2n$ leaves. At this juncture, we have two alternatives: (a) continue to use the simplest tree $T(n,n)$ and change the weight of one of the point vortices to $\lambda + \epsilon$ where $\epsilon \ll \lambda$, so that the conditions (A.3) hold for $s = m - 1$, and (b) keep the weights unchanged but introduce a mismatch of* exactly *one so that the upper cloud has n vortices of circulation λ and the lower cloud has $m - n = n \pm 1$ vortices with circulation $-\lambda$, and use a unbalanced full binary tree $T(n, n \pm 1)$.*

Remark 2. *The first choice in the above remark will work because changing the weight of one out of $2n$ vortices by a very small amount $\epsilon \ll \lambda$ will only change the dynamics of the external variable in the generalized tilted pair slightly. Instead of rigid translation of the dumb-bell configuration in a straight line, we should now see a slightly curved trajectory for the dumb-bell with a very large radius of curvature. The second choice coincides with the practical conditions of the Legg and Marshall numerical experiment, and will be adopted in the rest* *of this section. The facts that numerical simulations support a mismatch in vortex numbers between the two layers in the Heton model when the vorticities equal $\pm\lambda$, and that such mismatches are required by the KAM twist condition for proving the existence of geophysically relevant KAM tori within the combinatorial perturbation framework, are compelling evidence that this method is indeed the correct one for the propagating heton clusters problem.*

6.2.1. Combinatorial canonical transformations

First, we apply the Jacobi transformation given by (A.4), (A.5), (A.6), and (A.7) in the appendix, in terms of the full binary tree $T(n, n\pm 1)$ to the discrete Hamiltonian function on $m = 2n \pm 1$ particles,

$$
\begin{aligned}
H &= \frac{1}{4}\sum_{i\neq j=1}^{n} \lambda_i^1\lambda_j^1 \left[\ln|\vec{x}_j^1 - \vec{x}_i^1| - K_0\left(\frac{|\vec{x}_j^1 - \vec{x}_i^1|}{L_R(\pi/\sqrt{8})}\right)\right] \\
&+\frac{1}{4}\sum_{i\neq j=1}^{m-n} \lambda_i^2\lambda_j^2 \left[\ln|\vec{x}_j^2 - \vec{x}_i^2| - K_0\left(\frac{|\vec{x}_j^2 - \vec{x}_i^2|}{L_R(\pi/\sqrt{8})}\right)\right] \\
&+\frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{m-n} \lambda_i^1\lambda_j^2 \left[\ln|\vec{x}_j^2 - \vec{x}_i^1| + K_0\left(\frac{|\vec{x}_j^2 - \vec{x}_i^1|}{L_R(\pi/\sqrt{8})}\right)\right].
\end{aligned}
$$

with weights as outlined above. In terms of the symplectic variables for $\alpha = 1, 2$,

$$q_i^\alpha = x_i^\alpha,\ p_i^\alpha = \lambda_i^\alpha y_i^\alpha,$$

where the location of a vortex is given by

$$\vec{x}_i^\alpha = (x_i^\alpha, y_i^\alpha).$$

¿From here on, we will fix $\lambda_i^1 = \lambda$ and $\lambda_i^2 = -\lambda$. Then we have

$$
\begin{aligned}
q_i^1 &= x_i^1,\ p_i^1 = \lambda y_i^1, \\
q_i^2 &= x_i^2,\ p_i^2 = -\lambda y_i^2,
\end{aligned}
$$

or

$$|\vec{x}_j^1 - \vec{x}_i^1| = \sqrt{(q_j^1 - q_i^1)^2 + \left(\frac{p_j^1}{\lambda} - \frac{p_i^1}{\lambda}\right)^2},$$

$$
\begin{aligned}
|\vec{x}_j^2 - \vec{x}_i^2| &= \sqrt{(q_j^2 - q_i^2)^2 + \left(\frac{p_j^2}{\lambda} - \frac{p_i^2}{\lambda}\right)^2}, \\
|\vec{x}_j^2 - \vec{x}_i^1| &= \sqrt{(q_j^2 - q_i^1)^2 + \left(\frac{p_j^2}{\lambda} + \frac{p_i^1}{\lambda}\right)^2},
\end{aligned}
$$

and the Hamiltonian function takes the form

$$
\begin{aligned}
H &= \frac{\lambda^2}{4}\sum_{i\neq j=1}^{n}\left[\begin{array}{l} \ln\sqrt{(q_j^1 - q_i^1)^2 + \left(\frac{p_j^1}{\lambda} - \frac{p_i^1}{\lambda}\right)^2} \\ -K_0\left(\frac{\sqrt{(q_j^1 - q_i^1)^2 + \left(\frac{p_j^1}{\lambda} - \frac{p_i^1}{\lambda}\right)^2}}{L_R(\pi/\sqrt{8})}\right)\end{array}\right] \qquad (6.8) \\
&+\frac{\lambda^2}{4}\sum_{i\neq j=1}^{m-n}\left[\begin{array}{l} \ln\sqrt{(q_j^2 - q_i^2)^2 + \left(\frac{p_j^2}{\lambda} - \frac{p_i^2}{\lambda}\right)^2} \\ -K_0\left(\frac{\sqrt{(q_j^2 - q_i^2)^2 + \left(\frac{p_j^2}{\lambda} - \frac{p_i^2}{\lambda}\right)^2}}{L_R(\pi/\sqrt{8})}\right)\end{array}\right] \\
&-\frac{\lambda^2}{4}\sum_{i=1}^{n}\sum_{j=1}^{m-n}\left[\begin{array}{l} \ln\sqrt{(q_j^2 - q_i^1)^2 + \left(\frac{p_j^2}{\lambda} + \frac{p_i^1}{\lambda}\right)^2} \\ +K_0\left(\frac{\sqrt{(q_j^2 - q_i^1)^2 + \left(\frac{p_j^2}{\lambda} + \frac{p_i^1}{\lambda}\right)^2}}{L_R(\pi/\sqrt{8})}\right)\end{array}\right].
\end{aligned}
$$

From (A.7), we deduce that the quantities $(q_i^\alpha - q_j^\alpha)$, $(q_j^2 - q_i^1)$, $(p_i^\alpha - p_j^\alpha)$, and $(p_j^2 + p_i^1)$ are given by the following expressions on leaf i to leaf j paths, $P(i,j,\alpha,\beta)$:

$$
\begin{aligned}
(q_i^\alpha - q_j^\alpha) &= \sum_{k=1}^{s(i,j,\alpha)} c_k(i,j,\alpha)Q_k, \qquad (6.9) \\
(p_i^\alpha - p_j^\alpha) &= \sum_{k=1}^{s(i,j,\alpha)} d_k(i,j,\alpha)P_k, \\
(q_j^2 - q_i^1) &= \sum_{k=1}^{s(i,j)} c_k^{'}(i,j)Q_k,
\end{aligned}
$$

$$(p_i^1 + p_j^2) \;=\; \sum_{k=1}^{s(i,j)} d_k^{'}(i,j)P_k,$$

where $s(i,j,\alpha)$, $\alpha = 1,2$, is the internal node nearest to the root $m-1$ of $T(n, n\pm 1)$,in the path $P(i,j,\alpha)$ (resp. $s(i,j)$ in the path $P(i,j)$ for the case where vortex j is in layer 2 and vortex i is in layer 1). The sum $\sum_{k=1}^{s(i,j,\alpha)}$ is over all the internal nodes in the path $P(i,j,\alpha)$ (resp. $\sum_{k=1}^{s(i,j)}$ in the path $P(i,j)$ for the case where vortex j is in layer 2 and vortex i is in layer 1), and the coefficients $c_k(i,j,\alpha)$, $c_k^{'}(i,j)$, $d_k(i,j,\alpha)$ and $d_k^{'}(i,j)$ are functions of the weights (in this case just λ) which are obtained from those in the inversion (A.7).

Remark 3. *The quantities such as* $\left[(q_i^\alpha - q_j^\alpha)^2 + (\frac{p_i^\alpha}{\lambda} - \frac{p_j^\alpha}{\lambda})^2\right]^{1/2}$ *and* $\left[(q_j^2 - q_i^1)^2 + (\frac{p_i^1}{\lambda} + \frac{p_j^2}{\lambda})^2\right]^{1/2}$ *in the arguments of the logarithms and K_o in (6.8) are actually horizontal separations between particle i and particle j.*

This yields a new Hamiltonian H' $(Q_j,\ P_j,\ \lambda, T(n, n\pm 1))$ where the new symplectic variables Q_j, P_j, $j = 1, ..., m-1$ are associated with the internal nodes of $T(n, n\pm 1)$:

$$\begin{aligned}
H \;=\; & \frac{\lambda^2}{4}\sum_{i\neq j=1}^{n}\left[\begin{array}{c} \ln\left[(\sum_{k=1}^{s(i,j,1)} c_k(i,j,1)Q_k)^2 + \frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j,1)} d_k(i,j,1)P_k)^2\right]^{1/2} \\ -K_0\left(\frac{\left[(\sum_{k=1}^{s(i,j,1)} c_k(i,j,1)Q_k)^2+\frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j,1)} d_k(i,j,1)P_k)^2\right]^{1/2}}{L_R(\pi/\sqrt{8})}\right)\end{array}\right] \\
+\frac{\lambda^2}{4}\sum_{i\neq j=1}^{m-n} & \left[\begin{array}{c} \ln\left[(\sum_{k=1}^{s(i,j,2)} c_k(i,j,2)Q_k)^2 + \frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j,2)} d_k(i,j,2)P_k)^2\right]^{1/2} \\ -K_0\left(\frac{\left[(\sum_{k=1}^{s(i,j,2)} c_k(i,j,2)Q_k)^2+\frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j,2)} d_k(i,j,2)P_k)^2\right]^{1/2}}{L_R(\pi/\sqrt{8})}\right)\end{array}\right] \\
-\frac{\lambda^2}{4}\sum_{i=1}^{n}\sum_{j=1}^{m-n} & \left[\begin{array}{c} \ln\left[(\sum_{k=1}^{s(i,j)} c_k^{'}(i,j)Q_k)^2 + \frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j)} d_k^{'}(i,j)P_k)^2\right]^{1/2} \\ +K_0\left(\frac{\left[(\sum_{k=1}^{s(i,j)} c_k^{'}(i,j)Q_k)^2+\frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j)} d_k^{'}(i,j)P_k)^2\right]^{1/2}}{L_R(\pi/\sqrt{8})}\right)\end{array}\right].
\end{aligned} \tag{6.10}$$

Remark 4. *The quantities*

$$\left[(\sum_{k=1}^{s(i,j,\alpha)} c_k(i,j,\alpha)Q_k)^2 + \frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j,\alpha)} d_k(i,j,\alpha)P_k)^2\right]^{1/2}$$

are horizontal inter-particle separations in layer α and

$$\left[(\sum_{k=1}^{s(i,j)} c_k^{'}(i,j)Q_k)^2 + \frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j)} d_k^{'}(i,j)P_k)^2 \right]^{1/2}$$

are horizontal separations between vortex j in layer 2 and vortex i in layer 1.

6.2.2. Combinatorial perturbation method

The second step in our combinatorial approach is to put H' $(Q_j,\ P_j,\ \lambda, T(n,n))$ in the form

$$H' = H_o + H_1,$$

using explicitly the partial order relation induced by the structure of $T(n, n\pm 1)$. To begin we note that each of the first two sums in (6.10) has the same basic form, namely

$$\sum_{i \neq j=1}^{M} \left[\begin{array}{c} \ln \left[(\sum_{k=1}^{s(i,j,\alpha)} c_k(i,j,\alpha)Q_k)^2 + \frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j,\alpha)} d_k(i,j,\alpha)P_k)^2 \right]^{1/2} \\ -K_0 \left(\frac{\left[(\sum_{k=1}^{s(i,j,\alpha)} c_k(i,j,1)Q_k)^2 + \frac{1}{\lambda^2}(\sum_{k=1}^{s(i,j,\alpha)} d_k(i,j,\alpha)P_k)^2 \right]^{1/2}}{L_R(\pi/\sqrt{8})} \right) \end{array} \right] \qquad (6.11)$$

where for fixed $i \neq j$, the sum $\sum_{k=1}^{s(i,j,\alpha)}$ is in fact the sum over all internal nodes $k \leq s(i,j,\alpha)$ in the path $P(i,j,\alpha)$ from leaf (particle) i to leaf (particle) j in layer α, where $<$ is the partial order on the internal nodes of $T(n, n\pm 1)$ and $s = s(i,j,\alpha)$ is the internal node in $P(i,j,\alpha)$ which is nearest (in the sense of $<$) to the root node $m-1$ of the tree. The third sum has a similar form with $+K_o$ replacing $-K_o$.

In terms of the action-angle variables

$$(Q_s, P_s) = J_s \exp(i\varphi_s) \text{ for } s = 1, ..., m-1, \qquad (6.12)$$

and the coefficients $b_k(i,j,\alpha)$ which are functions only of λ just as c_k, $c_k^{'}$, d_k and $d_k^{'}$ are by the first remark in the appendix, one can write (6.11) as follows:

$$\sum_{i \neq j=1}^{M} \left[\ln \left| \sum_{k=1}^{s(i,j,\alpha)} b_k(i,j,\alpha) J_k \exp(i\varphi_k) \right| \pm K_0 \left(\frac{\left| \sum_{k=1}^{s(i,j,\alpha)} b_k(i,j,\alpha) J_k \exp(i\varphi_k) \right|}{L_R(\pi/\sqrt{8})} \right) \right],$$

where $|\cdot|$ denotes the modulus.

By writing

$$\left|\sum_{k=1}^{s(i,j,\alpha)} b_k(i,j) J_k \exp(i\varphi_k)\right| = |b_s(i,j,\alpha) J_s \exp(i\varphi_s)| \left|\exp(i0) + \sum_{k<s} \frac{b_k(i,j,\alpha) J_k \exp(i\varphi_k)}{b_s(i,j,\alpha) J_s \exp(i\varphi_s)}\right|,$$

we get

$$\begin{aligned}
K &= \sum_{i\neq j=1}^{M} \left[\ln\left|\sum_{k=1}^{s(i,j,\alpha)} b_k(i,j,\alpha) J_k \exp(i\varphi_k)\right| \pm K_0\left(\frac{\left|\sum_{k=1}^{s(i,j,\alpha)} b_k(i,j,\alpha) J_k \exp(i\varphi_k)\right|}{L_R(\pi/\sqrt{8})}\right)\right] \\
&= \sum_{i\neq j=1}^{M} \left[\begin{array}{c} \ln|b_s(i,j,\alpha) J_s \exp(i\varphi_s)| + \ln\left|1 + \sum_{k<s} \frac{b_k(i,j,\alpha) J_k \exp(i\varphi_k)}{b_s(i,j,\alpha) J_s \exp(i\varphi_s)}\right| \\ \pm K_0\left(\frac{|b_s(i,j,\alpha) J_s \exp(i\varphi_s)|\left|1+\sum_{k<s} \frac{b_k(i,j,\alpha) J_k \exp(i\varphi_k)}{b_s(i,j,\alpha) J_s \exp(i\varphi_s)}\right|}{L_R(\pi/\sqrt{8})}\right) \end{array}\right] \\
&= \sum_{i\neq j=1}^{M} \left[\begin{array}{c} \ln|b_s(i,j,\alpha) J_s \exp(i\varphi_s)| + \ln\left|1 + \sum_{k<s} \frac{b_k(i,j,\alpha) J_k \exp(i\varphi_k)}{b_s(i,j,\alpha) J_s \exp(i\varphi_s)}\right| \\ \pm K_0\left(\frac{|b_s(i,j,\alpha) J_s \exp(i\varphi_s)|\left|1+\sum_{k<s} \frac{b_k(i,j,\alpha) J_k \exp(i\varphi_k)}{b_s(i,j,\alpha) J_s \exp(i\varphi_s)}\right|}{L_R(\pi/\sqrt{8})}\right) \end{array}\right] \\
&= \sum_{i\neq j=1}^{M} \ln J_s \pm K_0\left(\frac{|b_s(i,j,\alpha)||J_s|}{L_R(\pi/\sqrt{8})}\right) + \qquad (6.13)
\end{aligned}$$

$$\sum_{i\neq j=1}^{M} \quad \sum_{k<s(i,j,\alpha)} \left|\frac{b_k(i,j,\alpha) J_k \exp(i\varphi_k)}{b_s(i,j,\alpha) J_s \exp(i\varphi_s)}\right| + h.o.t.$$

In the next step, we perform a re-summing of the first sum in (6.13) as follows: for each internal node $s = 1, ..., W$, where W depends on which of the three sums in (6.10) we are concerned with at the moment, collect all pairs (i,j), $i \neq j$ whose highest node $s(i,j,\alpha) = s$ (resp. $s(i,j) = s$ in the case of the third sum in the Hamiltonian) and call this set of leaf-to-leaf paths $P(s)$; then the sum becomes

$$G_o = \sum_{s=1}^{W}\left(\sum_{(i,j)\in P(s)}\left[\ln J_s \pm K_0\left(\frac{|b_s(i,j,\alpha)|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right]\right).$$

Collecting the higher order terms into G_1 we get

$$\begin{aligned}
G &= G_o + G_1 \\
&= \sum_{s=1}^{W}\left(\sum_{(i,j)\in P(s)}\left[\ln J_s \pm K_0\left(\frac{|b_s(i,j,\alpha)|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right]\right) \\
&+ \sum_{i\neq j=1}^{M} \quad \sum_{k<s(i,j,\alpha)} \left|\frac{b_k(i,j,\alpha) J_k \exp(i\varphi_k)}{b_s(i,j,\alpha) J_s \exp(i\varphi_s)}\right| + h.o.t.
\end{aligned}$$

The completely integrable term G_o after the re-summing step can be pictured as a change from summing over all leaf-to-leaf paths $P(i,j)$ in $T(n, n \pm 1)$ to first summing over subsets $P(s)$ of leaf-to-leaf paths, and then summing over the W internal nodes $s = 1, ..., W$ which are in each case, the common highest node in $T(n, n \pm 1)$ of all the members of the set $P(s)$. In this way, we put the full Hamiltonian in the form $H' = H_o + H_1$, where using superscript to denote the association with the three sums in (6.10), and the fact that $m = 2n \pm 1$, and a labelling scheme of the internal nodes of $T(n, n \pm 1)$ whereby each of its internal nodes whose leaf descendents are vortices in layer 1 are named $s = 1, ..., n-1$, and those whose leaf descendents are vortices in layer 2 are named $s = n, ..., m-2$, and finally the root is $s = m - 1$, we get

$$\begin{aligned} H_o &= G_o^{(1)} + G_o^{(2)} + G_o^{(3)} \\ &= \frac{\lambda^2}{4}\sum_{s=1}^{n-1}\left(\sum_{(i,j)\in P(s)}\left[\ln J_s - K_0\left(\frac{|b_s(i,j,1)|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right]\right) \\ +\frac{\lambda^2}{4}\sum_{s=n}^{m-2} &\left(\sum_{(i,j)\in P(s)}\left[\ln J_s - K_0\left(\frac{|b_s(i,j,2)|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right]\right) \\ -\frac{\lambda^2}{4}\sum_{i=1}^{n} &\sum_{j=1}^{m-n}\left[\ln J_{m-1} + K_0\left(\frac{|b_{m-1}(i,j)|\ J_{m-1}}{L_R(\pi/\sqrt{8})}\right)\right]. \end{aligned} \tag{6.14}$$

It can be easily shown that:

Lemma 1. *If $T(n, n \pm 1)$ is the most balanced binary tree with a mismatch of 1 at the level of the root, then (a) for each $s = 1, ..., n-1$ or $s = n, ..., m-2$, the weights $b_s(i,j,\alpha)$, $\alpha = 1, 2$ are the same for all pairs of leaves (i,j) in the set $P(s)$ and (b) $b_{m-1}(i,j)$ are the same for all pairs (i,j) such that $i = 1, ..., n$, and $j = n+1, ..., m$.*

An application of this lemma to (6.14) gives

$$\begin{aligned} H_o &= G_o^{(1)} + G_o^{(2)} + G_o^{(3)} \\ &= \frac{\lambda^2}{4}\sum_{s=1}^{n-1}\left(\sum_{(i,j)\in P(s)}\left[\ln J_s - K_0\left(\frac{|b_s|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right]\right) \\ +\frac{\lambda^2}{4}\sum_{s=n}^{m-2} &\left(\sum_{(i,j)\in P(s)}\left[\ln J_s - K_0\left(\frac{|b_s|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right]\right) \end{aligned} \tag{6.15}$$

$$-\frac{\lambda^2}{4}\sum_{i=1}^{n} \qquad \sum_{j=1}^{m-n}\left[\ln J_{m-1}+K_0\left(\frac{|b_{m-1}|\ J_{m-1}}{L_R(\pi/\sqrt{8})}\right)\right].$$

Summing over the elements of $P(s)$ in the first two sums, and using the fact that there are $n(n\pm 1)$ terms in the last sum, we get

$$\begin{aligned} H_o &= \frac{\lambda^2}{4}\sum_{s=1}^{n-1}|P\,(s)\,|\left[\ln J_s - K_0\left(\frac{|b_s|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right] \\ +\frac{\lambda^2}{4}\sum_{s=n}^{m-2} & \quad |P\,(s)\,|\left[\ln J_s - K_0\left(\frac{|b_s|\ J_s}{L_R(\pi/\sqrt{8})}\right)\right] \\ & -\frac{2\lambda^2|P(m-1)|}{4}\left[\ln J_{m-1}+K_0\left(\frac{|b_{m-1}|\ J_{m-1}}{L_R(\pi/\sqrt{8})}\right)\right], \end{aligned} \tag{6.16}$$

where the number $|P\,(s)\,|$ of elements in the sets $P(s)$ of leaf-to-leaf paths whose highest node is s, is determined by the most balanced full binary tree $T(n,n\pm 1)$. For example when $s=m-1$, $P(m-1)$ is the set of such paths that passes through the root node $s=m-1$; thus, $|P(m-1)|=n(n\pm 1)/2$. It is clear now that H_o is a completely integrable Hamiltonian— in fact it is totally decoupled in its dependence on the actions J_s, $s=1,...,m-1$.

The perturbation H_1 which consists of sums of the form

$$\begin{aligned} G_1 &= \sum_{i\neq j}^{n}\sum_{k<s(i,j)}\left|\frac{b_k(i,j)J_k\exp(i\varphi_k)}{b_s(i,j)J_s\exp(i\varphi_s)}\right| + h.o.t. \\ &= \sum_{i\neq j}^{n}\sum_{k<s(i,j)}\left|\frac{b_k(i,j)}{b_s(i,j)}\right|\frac{J_k}{J_s}, \end{aligned}$$

is not small at all points of the $2(m-1)$ dimensional phase-space M. But they can be controlled in special regions where the following ratios are small, i.e.,

$$\frac{J_k}{J_s} < \epsilon \ \ll 1,\ \ k<s, \tag{6.17}$$

where internal nodes $k<s$ means that k is lower in the tree than s (according to the partial order relation of $T(n,n\pm 1)$). We shall call the regions where (6.17) hold, cones:

$$C(J_s,\varphi_s,T(n,n),\epsilon)=\{(J_1,\varphi_1,...,J_{m-1},\varphi_{m-1})\in M \mid \frac{J_k}{J_s}<\epsilon \ \ll 1,\ \ k<s\}. \tag{6.18}$$

It is clear that the cones are predetermined by the tree $T(n, n\pm1)$ and the number $\epsilon \ll 1$. By construction the following quantities in H_1 are small in the cones:

$$\left|\frac{b_k(i,j)}{b_s(i,j)}\right| \frac{J_k}{J_s}.$$

6.2.3. KAM twist condition

Let $X_s = \frac{|b_s|\ J_s}{L_R(\pi/\sqrt{8})}$, and for $s = 1, ...$

$$Z_s = \left(\frac{1}{J_s^2} - \frac{\partial^2}{\partial J_s^2} K_0(X_s)\right).$$

Since H_o is in fact decoupled, and b_s is independent of J_s by the first remark in the appendix, the Hamiltonian H_o (6.16) has Hessian

$$\begin{aligned}
&\det \begin{bmatrix} \partial^2_{J_1J_1}H_0 & \partial^2_{J_1J_2}H_0 & ... & \partial^2_{J_1J_{m-1}}H_0 \\ \partial^2_{J_2J_1}H_0 & \partial^2_{J_2J_2}H_0 & .. & \partial^2_{J_2J_{m-1}}H_0 \\ .. & .. & .. & .. \\ \partial^2_{J_{m-1}J_1}H_0 & \partial^2_{J_{m-1}J_2}H_0 & .. & \partial^2_{J_{m-1}J_{m-1}}H_0 \end{bmatrix} \\
= \ &\det\left(-\frac{\lambda^2}{4}\begin{bmatrix} |P(1)|Z_1 & 0 & ... & 0 \\ 0 & |P(2)|Z_2 & 0 & .. \\ .. & 0 & ... & 0 \\ 0 & .. & 0 & 2|P(m-1)|\left(\frac{1}{J_{m-1}^2} + \frac{\partial^2}{\partial J_{m-1}^2}K_0(X_{m-1})\right) \end{bmatrix}\right) \\
= \ &(-1)^{m-1} \times 2\left(\frac{\lambda^2}{4}\right)^{m-1} |P(m-1)| \left(\frac{1}{J_{m-1}^2} + \frac{\partial^2}{\partial J_{m-1}^2}K_0(X_{m-1})\right) \prod_{s=1}^{m-2} |P(s)|Z_s.
\end{aligned}$$

This Hessian is nonzero on the cones (6.18) because

$$\begin{aligned}
\frac{1}{J_s^2} - \frac{\partial^2}{\partial J_s^2}K_0(X_s) &= \frac{1}{J_s^2} - \left(\frac{|b_s|}{L_R(\pi/\sqrt{8})}\right)^2 \left(\frac{\partial^2}{\partial x^2}K_0(x)\right)\left(\frac{|b_s|\ J_s}{L_R(\pi/\sqrt{8})}\right) \\
&\simeq O(\frac{1}{J_s^2}) \neq 0,
\end{aligned}$$

and the term associated with the root node $s = m-1$,

$$\frac{1}{J_{m-1}^2} + \frac{\partial^2}{\partial J_{m-1}^2}K_0(X_{m-1}) \neq 0$$

for finite values of J_{m-1}, which follows from the properties of K_0. This completes the rigorous justification for the existence of positive measure sets of KAM tori in the reduced phase space of the heton model. These tori correspond to permanent, tightly clustered, tilted heton configurations that consists of a cloud of hetons in each layer because the parameter ϵ in the cone condition gives a measure of the tightness of each cloud of hetons relative to the separation between them.

In order to deduce that these tori actually correspond to heton clusters that move on slightly curved trajectories in physical space, we recall that the eliminated degree of freedom in the CPM is associated with the absolute motion of the center of the heton clusters. It is easy to show that to leading order, the absolute motion of heton clusters is given by putting all the hetons in each cloud at the center of the cloud. We conclude that if we replace the simple tilted heton by clouds of hetons, then the tilted clouds will behave to $O(\epsilon)$ like the elementary solution of the simple tilted heton, that is, according to equation (4.6) with λ replaced by $n\lambda$. This completes the proof of theorem 1.

6.3. A special consideration for the Heton model

In the above KAM analysis, we made good use of the cone condition (6.17) which is sufficient to imply the physically crucial fact that the velocities of all distinct hetonic pairs of point vortices in the cluster are very well approximated by the velocity of a canonical hetonic pair whose point vortices are in fact located at the centers of mass of the upper and lower clusters. Our analysis of the long-lived propagating heton cluster can be improved by noting that the velocity of a single hetonic pair given by (4.6) has a very flat bell-shaped dependence on the scaled horizontal separation $r = \frac{d}{L_R(\pi/\sqrt{8})}$ between the point vortices in the two layers. This property of the graph of (4.6) implies that there is a unique value of $r = r^*$ at which the velocity (4.6) is a maximum and moreover, its derivative is almost zero for a substantial range of $r > r^*$. Rewriting (4.6) in terms of r, we get the formula

$$\frac{\lambda}{L_R(\pi/\sqrt{8})}\left[\frac{1}{r} - K_1(r)\right], \tag{6.19}$$

which has a maximum at $r^* \simeq 1$. This implies that the velocity of a hetonic pair is maximum if the separation between the vortices are near the radius of deformation $L_R(\pi/\sqrt{8})$. Since the horizontal separation d between the centers of mass of the two clusters is represented by the action J_{m-1} associated with the root (or node $m-1$) of the binary tree $T(m)$, it will be useful to ask, what if any benefit, can be derived by setting J_{m-1} near $L_R(\pi/\sqrt{8})$.

One of the consequence of the flatness of the graph of (6.19) for $r > 1$, is that the velocities of all distinct hetonic pairs of point vortices in the clusters are very well approximated by the velocity of the canonical hetonic pair whose horizontal separation $d \simeq L_R(\pi/\sqrt{8})$, even if the spread of the values of horizontal separations of these hetonic pairs is comparable with the radius of deformation. This means that the Hessian of H_0 is small, i.e, the twist is small for these configurations. Since the KAM estimate for the upper bound ϵ_0 on $|H_1|$ depends inversely on the twist, this bound ϵ_0 is relatively large. This means that if we selected $J_{m-1} \simeq L_R(\pi/\sqrt{8})$, then the parameter ϵ in the cone condition (6.17) can be relatively large, i.e., the ratios J_{s-1}/J_s do no have to be vanishingly small. Thus, we deduce, albeit heuristically, that there exists long-lived propagating hetonic clusters where the horizontal separation d between their respective centers of mass, is of the same order of magnitude as the largest horizontal separations within each cluster, provided $d \simeq L_R(\pi/\sqrt{8})$. This reasoning can be made rigorous along the lines of the above KAM analysis, and suggests that the existence of KAM tori in the Heton model is not merely formal, i.e, the size of the allowed perturbations do not have to be vanishingly small. It is moreover compatible with the common observation in computer simulations of KAM dynamics that the size of the perturbation given by the ratio J_{s-1}/J_s in (6.17) is often much larger than theoretical KAM estimates.

The dynamical significance of the radius of deformation has been discussed in [10] and [13]. Indeed, the propagating hetonic clusters in [13] not only display the remarkable particle numbers mismatch discussed earlier and verified by our combinatorial KAM analysis, but also have the very cluster characteristics just analysed.

7. Concluding Discussion

In this paper, Hamiltonian point vortex dynamics for coupled surface/interior QG has been developed systematically. These are novel vortex systems of mixed species where surface heat particles interact with interior quasi-geostrophic point vortices. As discussed in section 5, there is a large variety of elementary two-vortex exact solutions that transport heat including two-surface heat particles of opposite strength (see section 5.1), the horizontally tilted pairs consisting of a surface heat particle coupled to an interior vortex of opposite strength described in section 5.2, and the horizontally tilted interior vortices of opposite sign discussed in section 5.3.

Through comparison of explicit formulas, we established in section 5.2 that the interaction between a surface heat particle and an interior vortex is stronger at close range than both the interaction between a tilted heton pair in the two-layer model and the planar barotropic point vortex model for a fixed separation distance but is weaker at large separation distances; furthermore, as the vertical separation distance decreases, the behavior of the horizontally tilted surface/interior vortex pair mimics that for two surface particles studied in section 5.1. The propagating pairs of interior quasi-geostrophic point vortices with opposite strength at different vertical levels studied in section 5.3 have a behavior which qualitatively closely parallels the structure for hetons in the two-layer model; however, the interaction at large horizontal distances is weaker for the 3-D quasi-geostrophic model when compared with the two-layer model. In section 6, we have established the existence of large families of long-lived tilted heton clusters for both the two-layer heton models and the more general surface/interior coupled point vortex system. As we have discussed extensively in section 6, the structure of these tilted heton clusters is in remarkable detailed agreement with those observed by Legg and Marshall [13] in direct numerical simulation of the baroclinic point vortex system. The authors hope that the results presented here, stimulate the development of improved qualitative models for open ocean convection utilizing coupled surface/interior point vortices. For numerical purposes, one could utilize appropriate modified smoothed core vortices as for barotropic flow, ([2], [23]). Also it would be interesting to generalize the results of this paper to include both the large scale β-effect and topography. Equilibrium statistical mechanics for the two-layer heton models have been developed recently [6]. It would be interesting to extend that work to develop the equilibrium statistical mechanics of point vortices in coupled surface/interior QG.

A. Appendix

A.1. Jacobi variables

In previous work [17], we extended the Jacobi coordinates first to Newtonian $n-$body problems for arbitrarily large n, and then to planar point vortex problems [16] that involve two species of particles, that is, the vorticities have different signs (cf. also [12]). In [19], we extended the Jacobi variables to point vortex problems on the sphere. Here we are interested in a set of Jacobi coordinates for the Tilted clusters problem in both the Coupled QG and Heton models, where one species of particles have weight λ and the others have weight $-\lambda$. For economy of notation we have elected at this point not to distinguish the two species of particles and to use λ_j for the weight of the $j-th$ particle, in the following discussion.

A.1.1. Binary Trees and Canonical Transformations

The combinatorial algorithm that we will now describe, gives symplectic matrices of the form

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$$

where the m by m matrices $\mathbf{A}$ and $\mathbf{D}$ have the same zero pattern as $\mathbf{H}'$ (that is, modulo signs they have the same pattern of zeroes and nonzeroes), and satisfy the relation

$$\mathbf{A}^t\mathbf{D} = \mathbf{I}_m.$$

The last relation is equivalent to the usual condition for symplectic matrices,

$$\mathbf{M}^t\mathbf{J}\mathbf{M} = \mathbf{J}.$$

The first step in this algorithm is the introduction of a hybrid incidence type matrix $\mathbf{H}(\ T(m))$ (with $0, \pm 1$ entries) for the class of full binary trees $T(m)$ with m leaves. A full binary tree is a rooted tree where each (internal) node has exactly two descendents, except the terminal nodes which are called leaves. This matrix was introduced in [17], and has remarkable combinatorial properties [18]. The tree $T(m)$ has $m-1$ internal nodes which are partially ordered by the relation that node s is below node t if the path from node s to the root r contains node t. There are $C(m)$ distinct full binary trees with m leaves where the Catalan numbers $C(m) = \frac{1}{2m-1}\begin{pmatrix} 2m-1 \\ m \end{pmatrix}$. For each of the $m-1$ internal nodes $\mathbf{s}$ of a given $T(m)$, we define $\mathbf{\Lambda}^{\pm}(s)$ to be the set of leaf $\mathbf{j}$ below node $\mathbf{s}$ and connected to $\mathbf{s}$ via the right (left resp.) branch, and

$$\mathbf{\Lambda}(s) = \mathbf{\Lambda}^+(s) \cup \mathbf{\Lambda}^-(s), \qquad\qquad s = 1, ..., m-1.$$

The incidence type matrix $\mathbf{H}(\ T(m))$ is defined as follows

$$\begin{aligned} \mathbf{H}_{sj}(T(m)) &= \begin{pmatrix} 1 \text{ if } j \in \mathbf{\Lambda}^+(s) \\ -1 \text{ if } j \in \mathbf{\Lambda}^-(s) \\ 0 \text{ if } j \notin \mathbf{\Lambda}(s) \end{pmatrix} \text{ for } s = 1, ..., m-1; \\ \mathbf{H}_{mj}(T(m)) &= 1 \text{ for } j = 1, ..., m; \end{aligned}$$

the columns of $\mathbf{H}(\ T(m))$ are indexed by the leaf labels, and the first $m-1$ rows are indexed by the labels of the $m-1$ internal nodes in $T(m)$.

The second step in this combinatorial algorithm constructs two matrices $\mathbf{A}$ and $\mathbf{D}$ which have the same sign pattern as $\mathbf{H}(\ T(m))$ in the case where all the weights $\lambda_j,\ j=1,...,m$ have the same signature, and the same zero pattern $\mathbf{H}'(\ T(m))$ but not the same sign pattern in the case where the λ_j have mixed signs. Note that the number m of leaves equals the number of particles in the problem. They are given by

$$\begin{aligned}
\mathbf{A}_{sk} &= \left(\begin{array}{c} \dfrac{\lambda_k}{\sum\limits_{j\in\mathbf{\Lambda^+(s)}} \lambda_j} \text{ if } k \in \mathbf{\Lambda}^+(s) \\ \dfrac{-\lambda_k}{\sum\limits_{j\in\mathbf{\Lambda^-(s)}} \lambda_j} \text{ if } k \in \mathbf{\Lambda}^-(s) \\ 0 \text{ if } k \notin \mathbf{\Lambda}(s) \end{array} \right) \text{ for } s=1,....,m-1, \\
\mathbf{A}_{mk} &= \frac{\lambda_k}{\sum\limits_{j=1}^{m} \lambda_j} \text{ for } k=1,...,m;
\end{aligned} \tag{A.1}$$

and

$$\begin{aligned}
\mathbf{D}_{sk} &= \left(\begin{array}{c} \dfrac{g_s}{\sum\limits_{j\in\mathbf{\Lambda^+(s)}} \lambda_j} \text{ if } k \in \mathbf{\Lambda}^+(s) \\ \dfrac{-g_s}{\sum\limits_{j\in\mathbf{\Lambda^-(s)}} \lambda_j} \text{ if } k \in \mathbf{\Lambda}^-(s) \\ 0 \text{ if } k \notin \mathbf{\Lambda}(s) \end{array} \right) \text{ for } s=1,....,m-1, \\
\mathbf{D}_{mk} &= 1 \text{ for } k=1,...,m,
\end{aligned} \tag{A.2}$$

where the so-called Jacobi masses are

$$g_s = \frac{(\sum\limits_{j\in\mathbf{\Lambda^+(s)}} \lambda_j)(\sum\limits_{j\in\mathbf{\Lambda^-(s)}} \lambda_j)}{\sum\limits_{j\in\mathbf{\Lambda(s)}} \lambda_j} \text{ for } s=1,....,m-1.$$

Theorem 1. *If the following conditions hold*

$$\sum_{j\in\mathbf{\Lambda^+(s)}} \lambda_j \neq 0, \sum_{j\in\mathbf{\Lambda^-(s)}} \lambda_j \neq 0, \sum_{j\in\mathbf{\Lambda(s)}} \lambda_j \neq 0 \textit{ for } s=1,....,m-1, \tag{A.3}$$

then the matrices $\mathbf{A}$ *and* $\mathbf{D}$ *(A.1), (A.2) satisfy the relation* $\mathbf{A}^t\mathbf{D}=\mathbf{I}_m$.

Proof: The proof is by direct verification and can be found in Lim [17]. QED
The Jacobi transformation is now given by

$$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \tag{A.4}$$

where $\mathbf{q} = \{q_1, ..., q_m\}$, $\mathbf{p} = \{p_1, ..., p_m\}$, $\mathbf{Q} = \{Q_1, ..., Q_m\}$, and $\mathbf{P} = \{P_1, ..., P_m\}$. We note especially that

$$Q_m = \frac{\sum_{k=1}^{m} \lambda_k q_k}{\sum_{j=1}^{m} \lambda_j} \tag{A.5}$$

is the average *zonal coordinate* of the m particles on the f plane, and

$$P_m = \sum_{k=1}^{m} p_k \tag{A.6}$$

is the weighted sum of the m meridional coordinates. The two new variables Q_m and P_m are external variables in the sense that together they determine the absolute location of the configuration of m particles on the f plane. The remaining variables, $Q_1, ..., Q_{m-1}; P_1, ..., P_{m-1}$, are internal or shape variables, that is, they give the relative positions of the m particles in the configuration. They correspond to the $m-1$ internal nodes of the full binary tree $T(m)$, that was used to generate the Jacobi transformation (A.4).

Inverting (A.4), one obtains

$$\begin{aligned} q_j &= Q_m + \sum_{s=1}^{m-1} c_s^j(\boldsymbol{\lambda}) Q_s, \qquad \text{(A.7)} \\ p_j &= P_m + \sum_{s=1}^{m-1} d_s^j(\boldsymbol{\lambda}) P_s, \end{aligned}$$

for $j = 1, ..., m$. The coefficients c_s^j and d_s^j are zero if the internal node s is not in the path from leaf j to the root $r = m - 1$ (inclusive of r) of the tree $T(m)$. Thus, the inverse Jacobi transformation (A.7) based on a particular tree $T(m)$, corresponds to the set of paths from leaf j to the root r.

Remark 5. *Once the tree $T(m)$ is fixed, the coefficients $c_s^j(\boldsymbol{\lambda})$ and $d_s^j(\boldsymbol{\lambda})$ in (A.7) are functions only of $\boldsymbol{\lambda}$ by construction. This fact will be important for the* *applications of the Jacobi coordinates and the combinatorial perturbation method (CPM) described in this paper to the Coupled QG and Heton models. It enters into the evaluation of second order derivatives in the Hessian of the integrable H_0. It also figures in the way that we control the perturbation term H_1 by scaling the ratios of the actions J_s (6.12). More importantly, the very construction of H_0 and H_1 as functions which are generated by the tree $T(m)$, requires that these coefficients depend only on the weights $\boldsymbol{\lambda}$ of the particles once the tree $T(m)$ is fixed. In other words, the functions H_0 and H_1 in the CPM contain these coefficients as parameters, and if they depend on the symplectic variables, then H_0 and H_1 are clearly not well-defined.*

Remark 6. *If any of the quantities $\sum_{j\in\mathbf{\Lambda}^+(\mathbf{s})} \lambda_j$, $\sum_{j\in\mathbf{\Lambda}^-(\mathbf{s})} \lambda_j$, $\sum_{j\in\mathbf{\Lambda}(\mathbf{s})} \lambda_j$ for $s = 1, ..., m-1$ vanishes then the above matrices $\mathbf{A}$ and $\mathbf{D}$ are not defined. For the same set of weights, the above condition may hold for some tree $T(m)$ but not necessarily for another. In the case where all the weights λ_j have the same sign, $\mathbf{A}$ and $\mathbf{D}$ are always well-defined.*

Remark 7. *The full binary trees that will be used in this paper are the most balanced trees $T(n, n\pm1)$ where there is a mismatch of exactly one in the number of leaves in each descending branch of the root.*

Acknowledgments: The authors thank Mark DiBattista for his interest and comments on this work. They also wish to acknowledge the careful reading and thoughtful critique of an earlier version of this paper by two anonymous referees. The research of Andrew J. Majda is partially supported by grants NSF DMS-9596102-001, NSF DMS-9625795, ONR N00014-96-0043, and ARO DAAG55-98-1-0129.

References

[1] V. Arnold, *Mathematical methods in classical mechanics,* Springer-Verlag, New York, (1982)

[2] A. J. Chorin, *Vorticity and Turbulence,* Springer-Verlag, New York, (1994).

[3] M. DiBattista and A.J. Majda, *"An equilibrium statistical theory for large-scale features of open-ocean convection"* accepted and to appear in J. Phys. Oceanography.

[4] M. DiBattista and A. Majda, *An equilibrium statistical model for the spreading phase of open-ocean convection*, Proc. Natl. Acad. Sci., 96, 6009-6013 (1999)

[5] M. DiBattista, A. Majda, and J. Marshall, *The effect of preconditioning on statistical predictions for the spreading phase of open ocean convection*, submitted to J. Phys. Oceanography, September 1999.

[6] M. DiBattista and A. Majda, *"Equilibrium statistical mechanics for two-layer heton models"* (in preparation).

[7] M. Golubitsky, M. Krupa and C. Lim, *"Time-reversibility and particle sedimentation"*, SIAM J. Appl. Math., 51(1), 49-72, (1990).

[8] V.M. Gryanik, *"Dynamics of Localized Vortex Perturbations — "Vortex Charges" in a Baroclinic Fluid",* Isvestiya, Atmospheric and Oceanic Physics, Vol. 19(5), 347-352, (1983).

[9] I.M. Held, R.T. Pierrehumbert, S.T. Garner and K.L. Swanson, *"Surface quasi-geostrophic dynamics",* J. Fluid Mech. 282, 1-20, (1995).

[10] N.G. Hogg and H.M. Stommel, *"The Heton: an elementary interaction between discrete baroclinic vortices, and its implication concerning eddy heat flow",* Proc. Roy. Soc. Lond. A397, 1-20, (1985).

[11] N.G. Hogg and H.M. Stommel, *"Hetonic explosions: The breakup and spread of warm pools as explained by baroclinic point vortices",* J. Atmos. Sci., 42, 1465-1476, (1985).

[12] K. M. Khanin, *"Quasiperiodic motions of vortex systems"*, Physica D, 2, 261-269, 1982.

[13] S. Legg and J. Marshall, *"A Heton model of the spreading phase of Open Ocean Deep Convection",* J. Phys. Oceanography, 23(6), 1040-1056, (1993).

[14] S. Legg, H. Jones, and M. Visbeck, *A heton perspective of baroclinic eddy transfer in localized ocean deep convection*, J. Phys. Oceanogr.,26, 2251-2266 (1996).

[15] S. Legg and J. Marshall, *"The influence of the ambient flow on the spreading of convective water masses"*, J. Mar. Res., 56, 107-139 (1998)

[16] C. C. Lim, *"A combinatorial perturbation method and whiskered tori in vortex dynamics"*, Physica 64D, 163-184, (1993).

[17] C.C. Lim, *"Binary trees, symplectic matrices, and the Jacobi coordinates of celestial mechanics"*, Arch. rat. Mech Anal., 115, 153-165, (1991).

[18] C.C. Lim, *"Nonsingular sign patterns and the orthogonal group,"* Linear Algebra and Applic., 184, 1-12, (1993).

[19] C.C. Lim, , *"Relative equilibria of symmetric n-body problems on the sphere: Inverse and Direct results,"* Comm. Pure and Applied Math. Vol. LI, 341-371 (1998).

[20] C.C. Lim, J. Montaldi and M. Roberts , *"Relative equilibria of point vortices on the sphere"*, submitted to Physica D 1999.

[21] C.C. Lim and I-H. McComb, *"Resonant Normal Modes in Time-reversible equivariant vectorfields,"* J. Dynamics

[22] P.L. Lion and A.J. Majda, *"Equilibrium statistical theory for nearly parallel vortex filaments"*, to appear in CPAM (1999).

[23] A. Majda and A. Bertozzi, *"Vorticity and Incompressible Flow"*, Cambridge Univ. Press, (to appear in 2000)

[24] J. Marshall and F. Schott, *"Open-ocean convection: Observations, theory and models"*Rev. Geophysics, 37, 1-64 (1999).

[25] J. McWilliams, *"The emergence of isolated vortices in turbulent flows,"*J. Fluid Mech., 146, 21-, (1984).

[26] J. Pedlosky, *Geophysical Fluid Dynamics,* Springer-Verlag, 2nd Ed.

[27] J. Pedlosky, *"The instability of continuous heton clouds,"* J. Atmos. Sci., 42, 1477-1486, (1985).

[28] H. Pollard, *Celestial Mechanics,* Carus Math Pub. 1978.

[29] R. Salmon, *Lectures in Geophysical Fluid Dynamics,* Oxford U. Press, 1998.

[30] T.G. Shephard, *"Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics"*, Adv. Geophysics, 32, 287-338 (1990)