Exploring bistability in rotating Bose–Einstein condensates by a quotient transformation invariant continuation method

Yueh-Cheng Kuo a, Wen-Wei Lin b, Shih-Feng Shieh c, Weichung Wang d,∗

a Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 811, Taiwan
b Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan
c Department of Mathematics, National Taiwan Normal University, Taipei 116, Taiwan
d Department of Mathematics, National Taiwan University, Taipei 106, Taiwan

1. Introduction

Bose–Einstein condensates (BEC) have been produced and extensively studied in laboratories in the last 15 years. Studies include the exploration of quantized vortex states and their connections with superfluidity [1,2]. In many circumstances, a laser beam rotating with an angular velocity ω is applied to the magnetic trap holding the atoms to create a harmonic and anisotropic potential. Vortex nucleation [2] and vortex arrays [3] have also been observed.

In the rotating frame, a BEC with an external driven field at temperatures T much smaller than the critical temperature Tc [4] can be well described by time-independent nonlinear Schrödinger (NLS) equations [5–7]. In particular,

\[ \frac{1}{2} \nabla^2 \phi(x) + V(x)\phi(x) + \alpha |\phi|^2 \phi(x) + \omega \alpha \phi \nabla \phi(x) = \lambda \phi(x) \]  

(1.1a)

for x ∈ Ω ⊆ \mathbb{R}^2

\[ \int_{\Omega} |\phi(x)|^2 dx = 1, \]  

(1.1b)

where Ω is a bounded, smooth domain and \( \lambda = \sqrt{-1} \). In (1.1a), V(x) ≥ 0 is the magnetic trapping potential, \( \alpha \) = x\hat{\alpha}−y\hat{\beta} is the z-component of the angular momentum, \( \omega \) is the angular velocity of the rotating laser beam, \( \alpha \) denotes the intra-component scattering length and \( \lambda \) is the chemical potential or eigenvalue. The ground state solution of a BEC can be found by minimizing the energy functional E(\phi) under the conditions given by (1.1b) [5,6]. That is, Minimize E(\phi) subject to \( \int_{\Omega} |\phi(x)|^2 dx = 1 \) and \( \phi(x) = 0 \) for \( x \in \partial \Omega \),

\[ E(\phi) = \int_{\Omega} \left( \frac{1}{2} |\nabla \phi|^2 + V(x)|\phi|^2 + \frac{\omega}{2} |\phi|^4 + \alpha \phi^* \nabla \phi \right) \]  

(1.2a)

where \( \phi^* \) denotes the complex conjugate of \( \phi \). In (1.2b), the value of \( \int_{\Omega} \phi^* \nabla \phi \) is purely imaginary. The NLS equations (1.1) are also regarded as the Euler–Lagrange equation of the optimization problem given by (1.2).

To numerically solve a rotating BEC, a continuous normalized gradient flow (CNGF) method with a backward Euler finite difference discretization has been proposed in [8] for computing the ground, symmetric and central vortex states. The paper analyzes the existence/nonexistence problem for ground states...
We propose a quotient transformation invariant continuation method for solving the NLS equations numerically. In Section 3 we show bifurcation for a rotating BEC. In Section 4 we present the numerical results for stable and bound state solutions of a rotating BEC by using the QTICM. Finally, a conclusion is given in Section 5.

1.1. Discretization of the NLS equations

To investigate the NLS equations (1.1) and the associated optimization problem, we can discretize the differential equations, for example, by a finite difference method. In particular, we suppose that the operator $-\frac{1}{2}\nabla^2 + V(x)$ and the angular momentum $\partial_{\theta}$ in (1.1a) are discretized by the central difference approximation with a grid size $h$. Let $A \in \mathbb{R}^{N \times N}$ be the discretization matrix of the operator $-\frac{1}{2}\nabla^2 + V(x)$ combined with the Dirichlet boundary condition. We can see that $A$ is an irreducible and symmetric positive definite matrix with non-negative off-diagonal entries (i.e., an irreducible $M$-matrix). Let $S \in \mathbb{R}^{N \times N}$ be the discretization matrix of $\partial_{\theta}$ where $S$ is a skew-symmetric matrix.

We denote the approximation of the wave function $\phi$ in (1.1a) as $u_\mathbf{r} = u_1 + i u_2$, where $u_1, u_2 \in \mathbb{R}^N$. Then, the discrete nonlinear Schrödinger (DNLS) equations corresponding to (1.1) can be formulated as

$$Au_\mathbf{r} + \lambda u_\mathbf{r} \otimes u_\mathbf{r} + \omega Su_\mathbf{r} = \lambda u_\mathbf{r},$$

$$u_1^2 - u_2^2 = \frac{1}{h^2},$$

(1.3a

(1.3b)

Here $u_\mathbf{r}^* = u_\mathbf{r} \circ \mathbf{r}$ and $\circ$ denotes the Hadamard product. Similarly, the optimization problem (1.2a) becomes the finite-dimensional optimization problem:

$$\text{Minimize} \ E(u_1, u_2)$$

subject to $u_1^* u_1 + u_2^* u_2 = 1/h^2$, with the energy functional

$$E(u_1, u_2) = \frac{1}{2} \| u_1 \|^2 + \frac{1}{2} \| u_2 \|^2 + \omega \| Su_\mathbf{r} \|^2$$

$$= \frac{1}{2} \sum_{i=1}^{2} u_i^* A u_i + \frac{\alpha}{2} (u_1^* + u_2^*)^2$$

$$\times (u_1^2 + u_2^2) + 2 \omega \| Su_\mathbf{r} \|.$$ (1.4a)

Throughout this paper, we use bold face letters or symbols to denote a matrix or a vector. For $\mathbf{u} = (u_1, \ldots, u_n)^T$, $\mathbf{v} = (v_1, \ldots, v_n)^T \in \mathbb{C}^N$, $\mathbf{u} \circ \mathbf{v} = (u_1 v_1, \ldots, u_n v_n)^T$ denotes the Hadamard product of $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{u}^\Box = \mathbf{u} \circ \cdots \circ \mathbf{u}$ denotes the $r$-time Hadamard product of $\mathbf{u}$, and $\| \mathbf{u} \| := \text{diag}(\mathbf{u})$ denotes the diagonal matrix of $\mathbf{u}$. For $A \in \mathbb{R}^{N \times N}$, $\mathbb{N}(A)$ and $\mathbb{R}(A)$ denote the null space and column space of $A$, respectively. For $A \in \mathbb{R}^{N \times N}$, $A > 0$ (\geq 0) denotes a symmetric positive (semi-positive) definite matrix.

2. Quotient transformation invariant continuation method

In this section, we develop a quotient transformation invariant continuation method (QTICM) to solve the DNLS equations (1.3). First, we derive the parameter dependent polynomial systems that are solved at each iteration of the continuation method in Section 2.1. A strategy for circumventing the difficulties associated with transformation invariant solutions is presented in Section 2.2. This strategy is discussed in detail in Section 2.3. Methods for detecting the bifurcation points and stability of the solutions are proposed in Sections 2.4 and 2.5, respectively.

2.1. A parameter dependent polynomial system

A continuation method usually follows solution curves by solving sequences of parametrized polynomial systems. The polynomial systems considered here are described as follows.

Letting $u_1$ and $u_2$ be the real and imaginary parts of $u_\mathbf{r}$, respectively, we can rewrite $u_\mathbf{r} = u_1 + i u_2 \in \mathbb{C}^N$ and the DNLS equations (1.3) as:

$$Au_\mathbf{r} + \alpha (\| u_1 \|^2 + \| u_2 \|^2) u_1 - \omega Su_\mathbf{r} = \lambda u_\mathbf{r}$$

(2.1a)

$$Au_\mathbf{r} + \alpha (\| u_1 \|^2 + \| u_2 \|^2) u_2 + \omega Su_\mathbf{r} = \lambda u_\mathbf{r}$$

(2.1b)

$$u_1^2 - u_2^2 = \frac{1}{h^2}.$$ (2.1c)

Now, we denote $p$ as the positive continuation parameter, such that $p$ is incorporated into the system (2.3) by letting

$$\alpha = \alpha_0 + \mu p$$

$$\omega = \omega_0 + \nu p$$

(2.2)

for the known constants $\alpha_0$, $\mu_0$, $\omega_0$, and $\nu_0$. Furthermore, by letting $\mathbf{u} = (u_1, u_2)^T \in \mathbb{R}^{2N}$ and $\mathbf{z} = (u_1^*, \lambda)^T \in \mathbb{R}^{2N+1}$, we can rewrite the DNLS equations (2.1) as a parameter dependent polynomial system:

$$G(\mathbf{z}, p) = 0$$

(2.3)
where \( G \equiv (G_1, G_2, g) : \mathbb{R}^{2N+1} \times \mathbb{R} \rightarrow \mathbb{R}^{2N+1} \) is given by
\[
G_1(z, p) = A_1^u + \alpha((u_1^+)^T + (u_1^-)^T)u_1 - \omega Su_2 - \lambda u_1, \quad (2.4a)
\]
\[
G_2(z, p) = A_2^u + \alpha((u_2^+)^T + (u_2^-)^T)u_2 + \omega Su_1 - \lambda u_2, \quad (2.4b)
\]
\[
g'(z) = \frac{1}{2} \left( u_1^+ u_1^- + u_2^+ u_2^- - \frac{1}{\tau^2} \right). \quad (2.4c)
\]

Consequently, the Jacobian of \( G \) is given by
\[
\mathcal{D}G = [G_z, G_p] \in \mathbb{R}^{(2N+1) \times (2N+2)}. \quad (2.5)
\]

Thus,
\[
G_2(z, p) = [G_u|G_s] = \begin{bmatrix} H(z, p) & -u \end{bmatrix} \in \mathbb{R}^{(2N+1) \times (2N+1)}, \quad (2.6a)
\]
\[
G_p(z, p) = \mu \begin{bmatrix} -Su_2 & Su_1 & 0 \end{bmatrix} + \nu \begin{bmatrix} (u_1^+)^T + (u_1^-)^T & 0 \end{bmatrix} \in \mathbb{R}^{2N+1}, \quad (2.6b)
\]
\[
H(z, p) = \begin{bmatrix} \tilde{A} + 2\alpha([u_1^+]^T + [u_1^-]^T)u_1 - \omega S & 2\alpha([u_2^+]^T + [u_2^-]^T)u_2 \end{bmatrix} \quad (2.7)
\]
is a \( 2N \times 2N \) symmetric matrix, and
\[
\tilde{A} = A + \alpha([u_1^+]^T + [u_1^-]^T) - \lambda I_N. \quad (2.8)
\]

2.2. The transformation invariant solutions and quotient solutions

Now, we discuss why a standard continuation method is inappropriate for our problem and how we may circumvent the difficulty. Let the solution set \( \mathcal{M} \) of (2.3) be defined as follows:
\[
\mathcal{M} = \{ \mathbf{z} \in \mathbb{R}^N \mid G(z, p) = 0 \}. \quad (2.9)
\]
We define a function \( u(\theta) : [0, 2\pi] \rightarrow \mathbb{R}^N \) by
\[
u(\theta) = \begin{bmatrix} \cos \theta u_1 + \sin \theta u_2 \\ -\sin \theta u_1 + \cos \theta u_2 \end{bmatrix}. \quad (2.10)
\]
It is clear that if \((u_1^+)^T, (u_1^-)^T, \ldots, (u_1^-)^T, \ldots, (u_1^-)^T, \ldots, (u_1^-)^T)^T \in \mathcal{M} \) and both \( \lambda \) and \( p \) are fixed, then
\[
G(u(\theta), \lambda, p) = 0 \quad (2.11)
\]
for all \( \theta \in [0, 2\pi] \). That is, the solution set \( \mathcal{M} \) is actually a two-dimensional manifold on \( \mathbb{R}^{2N+2} \) containing the transformation invariant solutions \( u(\theta) \). In other words, for a particular solution \((u(0), \lambda, p)^T \in \mathcal{M} \), all of the solutions in the form \((u(\theta)^T, \lambda, p)^T \) have the same energy and density functions with the same shape. Namely, \( E(u(0)) = E(u(\theta)) \) and \([u(0)]^T = [u(\theta)]^T \) for all \( \theta \in [0, 2\pi] \). This two-dimensional solution manifold thus may prevent a standard continuation method from successfully following a desired solution curve by varying the parameter \( p \).

This observation can be understood conceptually with Fig. 2.1. We image the shape of the solution is a pipeline in \( \mathbb{R}^{2N+2} \); the solution curve \((u(\theta)^T, \lambda, p)^T \) for a certain \( \lambda \) and \( p \) is given by the black line in the middle of the pipeline of Fig. 2.1(b). A standard continuation method may irregularly circle around the pipeline solution set as shown in Fig. 2.1(a).

Now, let
\[
G(y(\theta)) = 0 \quad (2.12)
\]
be the arc length parameter, where \( \theta \in \mathbb{R} \). By differentiating Eq. (2.12) with respect to \( \theta \), we obtain:
\[
\mathcal{D}G(y(\theta))y'(\theta) = 0. \quad (2.13)
\]
where \( \mathcal{D}G(y(s)) \) is defined in (2.5) and \( y(s) = (z(s)^T, p(s)^T) \) denotes a tangent vector to \( \mathcal{M} \) at \( y(\theta) \). Eq. (2.13) suggests that the tangent vector \( y(\theta) \) to \( \mathcal{M} \) at \( y(\theta) \) is the nontrivial solution of the \((2N+1) \times (2N+2)\) homogeneous system \( \mathcal{D}G(y(s))w = 0 \), when \( \mathcal{D}G(y(s)) \) has a full row rank. However, the Jacobian matrix \( \mathcal{D}G(y(s)) \) considered here does not have a full row rank, as the dimension of the solution set \( \mathcal{M} \) is 2. In this case, standard continuous methods may not follow the solution set successfully.

To avoid circling around the solution manifold, our approach is to consider a quotient solution set by applying additional hyperplane constraints on the system. First, we parametrize the solution curve (2.3) by introducing arc length parameter \( s \), so that the equivalent solutions \( u(\theta) \) with \( \theta \in [0, 2\pi] \) are quotient. Furthermore, the quotient solution curve can be parametrized as:
\[
\mathcal{C} = \{ \mathbf{y}(s) = (z(s)^T, p(s)^T) \mid G(y(s)) = 0, s \in \mathbb{R} \}. \quad (2.14)
\]

Now, we shall discuss how we can trace the quotient solution curve \( \mathcal{C} \) by using a continuation method with additional hyperplane constraints. To do so, we can compute the tangent vector of \( u(\theta) \) for a certain \( \theta \). Without losing generality, we may choose \( \theta = 0 \) to obtain
\[
\frac{\partial u}{\partial \theta}(0) = (u_1^+, -u_1^-)^T, \quad (2.15)
\]
and then develop a quotient transformation invariant continuation method by incorporating the following prediction and correction schemes.

1. In the prediction step, a prediction direction \( \mathbf{y}_p = (\tilde{z}, \tilde{p})^T \) should satisfy (2.13) and the vector \( \tilde{z} \) should be orthogonal to the tangent vector \((u_1^+, -u_1^-)^T \).\( \uparrow \)

2. The correction step, we compute the correction vector that satisfies an additional hyperplane constraint whose normal vector is \((u_1^+, -u_1^-)^T, 0\).\( \uparrow \)

Fig. 2.1(b) demonstrates this idea visually.

We have introduced the main ideas of the quotient transformation invariant continuation method. In the next subsection, we will discuss in detail how this QTCM follows the parametrized quotient solution curve with modified prediction and correction steps.

2.3. Prediction and correction

Let \( \mathbf{y}_p = (\tilde{z}, \tilde{p})^T \in \mathbb{R}^{2N+2} \) be a point that has been accepted as an approximate solution on \( \mathcal{C} \). A “good” prediction direction \( \mathbf{y}_p = (\tilde{z}, \tilde{p})^T \) should satisfy (2.13) and the vector \( \tilde{z} \) should be orthogonal to \((u_1^+, -u_1^-)^T \).

Consequently, we apply an additional hyperplane constraint to Eq. (2.13):
\[
\tilde{a}_3^T \tilde{z} = 0, \quad (2.16)
\]
where \( \tilde{a}_3 = (\frac{\partial G_z}{\partial \theta}(0)^T, 0)^T = (u_1^+, -u_1^-)^T \). It follows that the prediction direction \( \mathbf{y}_p = \hat{\mathbf{y}} \in \mathbb{R}^{2N+2} \) should satisfy the bordered linear system
\[
\begin{bmatrix} G_z & G_p \end{bmatrix} \begin{bmatrix} \mathbf{a}_3^T \\ \mathbf{c}_3^T \end{bmatrix} \hat{\mathbf{y}} = \begin{bmatrix} \mathbf{0}^T \\ \mathbf{1} \end{bmatrix}, \quad (2.17)
\]
where \((\mathbf{c}_3^T, \mathbf{c}_3^T)^T \in \mathbb{R}^{2N+2} \) is a suitable constant vector. Note that Eq. (2.17) can be geometrically interpreted as follows. We first use the Euler predictor
\[
y_{i+1,1} = y_i + h(\mathbf{y}_i^+ / \|\mathbf{y}_i^+\|) \quad (2.13)
\]
to predict a new point $y_{i+1}$, where $h_i > 0$ is a suitable step length and $\dot{y}_i$ is the tangent vector at $y_i$ that is obtained by solving the bordered linear system (2.21).

In the correction step, the next acceptable solution $y_{i+1}$ must pass through the hyperplane whose normal vector is $(a_i, 0)^T$ with the Euler predictor $y_{i+1} = (z_{i+1}, \tilde{p}_{i+1})$. In other words, the solution curve $C$ is now determined by the following two remarks.

1. The sequence $(y_{i+1})$ converges numerically for $l = 1, 2, \ldots$, so we accept $(y_{i+1})$ as a new approximation to the solution curve $C$.

2. Both linear systems (2.17) and (2.18) are overdetermined and have unique solutions, an efficient way to solve these two systems is to rewrite them in the form:

$$
G(z, p) = 0,
$$

$$
a_i^T z = \tilde{z}_{i+1},
$$

$$
\tilde{z}^T z + \tilde{p}^T p = \tilde{y}^T y_{i+1}.
$$

Starting from the predictor $y_{i+1}$. Newton’s method is typically used to compute the correctors. By setting $y_{i+1} = y_{i+1} + \delta_i$ for $l = 1, 2, \ldots$, we solve the bordered linear system using an iterative method:

$$
\begin{align*}
G(y_{i+1}) &= [a_i^T, \tilde{z}_{i+1}^T, 0^T],
G_0(y_{i+1}) &= [0^T, \tilde{p}_{i+1}^T, -\rho_{i+1}^T, \rho_{i+1}^T],
\delta_i &= [-G(y_{i+1})],
\end{align*}
$$

(2.18)

with $\rho_{i+1} = a_i^T (z_{i+1} - z_{i+1})$ and $\rho_{i+1} = \tilde{y}_{i+1} (y_{i+1} - y_{i+1})$. If the sequence $(y_{i+1})$ converges numerically for $l = 1, 2, \ldots$, we accept $y_{i+1}$ as a new approximation to the solution curve $C$.

Let $C$ be the solution curve defined in (2.14), where $y(s) \in C$ and

$$
J(s) = \begin{bmatrix}
G(y(s)) \\
G_0(y(s))
\end{bmatrix} \in \mathbb{R}^{(2N+2) \times (2N+2)},
$$

(2.21a)

$$
\tilde{J}(s) = \begin{bmatrix}
G(y(s)) \\
G_0(y(s))
\end{bmatrix} \in \mathbb{R}^{(2N+2) \times (2N+1)}.
$$

(2.21b)

As described in [17, 18, 16], a point $y(s) \in C$ is said to be a regular point if rank $(J(s)) = 2N + 1$ (i.e., dim $\mathcal{N}(J(s)) = 1$) and a singular point if rank $(J(s)) \leq 2N$ (i.e., dim $\mathcal{N}(J(s)) \geq 2$). For a regular point $y(s)$, the tangent vector $\dot{y}(s)$ is uniquely determined by the linear system (2.17).

We propose to use the inverse power method to detect the singularity of the symmetric matrix $B(s)$ described in (2.20) at the point $y(s) \in C$. Furthermore, the tangent vectors at the singular point $y(s) \in C$ can also be found using the methods reported in [16, pp. 88–99] and [19]. This proposed scheme can be justified by Theorem 2.1 for the following mild assumption. Suppose we consider only the case

$$
G(y(s)) \in \mathcal{N}(\dot{J}(s)) \text{ for each singular point } y(s) \in C,
$$

(2.22)

i.e., dim $\mathcal{N}(\dot{J}(s)) = \dim \mathcal{N}(J(s)) - 1$ and dim $\mathcal{N}(J(s)) \geq 2$, when detecting the bifurcation points. For this situation, the tangent vector at a singular point has a nonzero component at $p(s)$ and is expected to appear in the solution curve $C$ of (2.12). Furthermore, as shown in Theorem 2.1, detecting the singular points of the solution curve $C$ is equivalent to detecting the singularity of the symmetric matrix $B$ in (2.20).

**Theorem 2.1.** If the condition (2.22) holds, then the following statements are equivalent: (i) rank $(J(s)) \leq 2N + 1$, (ii) $\mathcal{N}(J(s)) \neq \{0\}$, and (iii) $B(s)$ is singular.

**Proof.** See Appendix. □

2.5. Testing for stability

Along the solution curves determined by the QTCM, several types of solutions can be computed. Except for the ground and excited state solutions, we can define the so-called stable state solutions and then explain how to detect whether or not the computed solution $u^*$ of the DNLS equations (2.3) is stable.

It is well-known that the ground state solution of the set of DNLS equations (2.3) is found by minimizing the energy functional $E(u_1, u_2)$ in (1.4b) on the sphere:

$$
\delta = \begin{bmatrix}
u^T_1 u^T_2
\end{bmatrix} \in \mathbb{R}^{2N+1}, u_1^T u_1 + u_2^T u_2 = \frac{1}{l^2}.
$$

(2.23)

On the other hand, an excited state solution is a saddle point of the energy functional $E(u_1, u_2)$ on the sphere $\delta$. We can further define a stable state solution as follows: note that the definition suggests that the ground state solution is also a stable state solution.
**Definition 2.1.** The solution \( u = (u^T, u^3)^T \in \delta \) of the DNLSE (2.3) is a stable state solution if it is a local minimum of the energy functional \( E(u) \) on the sphere \( \delta \).

Verifying local minimum of \( E(u) \) on \( \delta \) can be achieved by applying standard optimization results to the optimization problem (1.4a). We further define the Lagrangian function of the optimization problem (1.4a):

\[
L(u, \lambda) = E(u) - \lambda g(u),
\]

where \( g(u) = (u^T u_1 + u^3 u_2 - 1) \) is given by (2.1c). By combining the following two theorems, we find a practical way to identify stable state solutions.

**Theorem 2.2 ([20]).** Suppose that the Karush–Kuhn–Tucker (KKT) conditions

\[
\nabla_u L(u^*, \lambda^*) = 0
\]

and

\[
u^* \in \delta
\]

are satisfied for a certain \( u^* \in \mathbb{R}^{2N} \) and \( \lambda^* \). Suppose also that

\[
w^T \nabla_u^2 L(u^*, \lambda^*) w > 0, \quad \text{for all } w^T u^* = 0, \quad w \neq 0.
\]

Then \( u^* \) is a strict local minimum solution of (1.4a).

**Theorem 2.3.** (i) The KKT condition (2.25) is equivalent to Eq. (2.3). (ii) The Hessian condition (2.26) is equivalent to

\[
H_u(u) := Z^T(u) H(y) Z(u) > 0,
\]

where \( Z(u) \) is a matrix whose columns form a basis of \( u^\perp = \{ x \in \mathbb{R}^{2N} \mid x^T u = 0 \} \) and \( H(y) = H(u, \lambda, p) = \nabla^2 u^2 L(u, \lambda) \) is defined by (2.7) and (2.8).

**Proof.** A proof by definition is straightforward for this theorem.

Since Theorem 2.3 shows that the conditions (2.25) and (2.26) are equivalent to the Eqs. (2.3) and (2.27), respectively, we can make the following observation: Let \( y^* = (u^*, \lambda^*, p^*) \in \mathbb{R}^{2N+2} \) be a solution of DNLSE equations (2.3); then the detection of the stability of \( E(u) \) in (1.4b) over \( \delta \) at \( y^* \) is equivalent to the detection of the positivity of the projected Hessian matrix \( H_u(u^*) \). In the following theorem, we give a sufficient condition for determining if the projected Hessian matrix \( H_u(u^*) \) is positive definite.

**Theorem 2.4.** Let \( y^* = (u^*, \lambda^*, p^*) \in \mathbb{R}^{2N+2} \) be a solution of DNLSE equations (2.3). If \( a = u^* + \lambda^* p^* \neq 0 \), then

\[
i \left( \begin{array}{c}
\nabla y^*(u^*)^T \nabla\nabla^2 y^*(u^*)^T
\end{array} \right) = i \left( \begin{array}{c}
H_u(u^*) \nabla\nabla^2 y^*(u^*)^T
\end{array} \right) + (1, 1, 0),
\]

where \( i(\cdot) = (i_0(\cdot), i_{-1}(\cdot), i_0(\cdot)) \) and \( i_{-1}(\cdot), i_0(\cdot) \) and \( i_0(\cdot) \) are the number of positive, negative, and zero eigenvalues of a symmetric matrix, respectively.

**Proof.** See Appendix.

In short, by checking whether or not the projected Hessian matrix \( H_u(u^*) \) is positive definite, we may determine the stability of the corresponding solution \( u^* \).

**2.6. A short summary**

To solve the DNLSE equations (1.3), we have presented the main ideas of the QTICM from analytical and geometrical viewpoints. We have also discussed how the predictors, correctors, and bifurcation points can be computed. In addition, we have shown how the stability of the solutions can be verified.

We conclude the development of the QTICM by a remark regarding the choice of discretization schemes. While the central finite differences is used to demonstrate the idea of QTICM, we can apply other spatial discretizations in the QTICM to gain efficiency or accuracy improvements. Some possible discretizations include (higher order) finite differences, finite volume, finite element, and pseudospectral methods. After choosing a discretization method, we need to discretize the NLS equations (1.1) to obtain the corresponding DNLSE equations that are similar to (1.3). We can separate the real and imaginary parts of the DNLSE equations in a similar way as shown in (2.4). The additional hyperplane constraint can then be incorporated by taking (2.15) as the normal vector of the hyperplane. As long as the computed predictors and correctors along the solution curves are accurate enough, the QTICM with the particular discretization can trace the solution curves as intended.

**3. Bifurcation of a rotating BEC with isotropic potential traps**

We proposed the QTICM for (1.3) in Section 2. Hereafter, we focus on the DNLSE equations with a harmonic trap potential in the two dimensions. In particular, we assume that:

\[
V(x, y) = \gamma_2 x^2 + \gamma_3 y^2,
\]

where \( \gamma_2 \) and \( \gamma_3 \) are the trap frequencies.

In this section, we study the bifurcation of a rotating BEC with isotropic traps in a discoidal domain, i.e. the trap frequencies are equal (\( \gamma_2 = \gamma_3 \)) and \( \Omega = \{ x \in \mathbb{R}^2 \mid \| x \|_2 \leq 1 \} \). Now, we consider the quotient solution curve \( C \) of (2.3) by letting \( p = \omega \) and fixing \( \alpha = \alpha_0 \). For any \( \alpha > 0 \), the ground state solution \( \phi_\alpha(x) \) is unique [21] and is real radial-symmetric for \( \omega = 0 \), i.e. \( \phi_\alpha(x) = 0 \). In fact, the NLS equations (1.1) always have a symmetric solution \( \phi_\alpha(x) \) that is independent of the angular velocity. It is well-known [8,22,23] that when the angular velocity increases to a critical value, the bifurcations occur in the ground state.

Similar problems also arise in the discretized system when Eq. (1.3) is considered for a radially symmetric trap in the discoidal domain and \( \omega = 0 \). The ground state solution \( u^\delta \) of the optimization problem

\[
\text{Minimize } u^T A u + \frac{\alpha}{2} \frac{u^T u^2}{h^2},
\]

subject to \( u^T u = 1 \), is radial-symmetric. We make the following assumption:

\[(\mathcal{H}) \quad \text{The ground state solution of (3.1) } u^\delta \in \mathcal{N}(S).
\]

Note that the matrix \( S \) gives the discretization of the operator \( \partial_\alpha \) and the assumption \( (\mathcal{H}) \) is true for the discretization scheme using polar coordinates (see e.g. [24]). Using the assumption \( (\mathcal{H}) \) and letting \( \lambda^\delta \) be the eigenvalue of the DNLSE equations (2.1) corresponding to the eigenvector (ground state) \( (u^\delta / \sqrt{2}, u^\delta / \sqrt{2}) \) for \( \omega = 0 \), it can be seen that the primal stall of the solution curve of the DNLSE equations (2.3) is

\[
C_\delta = \left\{ \left( \frac{1}{\sqrt{2}} u^\delta, \frac{1}{\sqrt{2}} u^\delta, \lambda^\delta, \omega \right)^T \mid 0 \leq \omega \leq \infty \right\} \subset \mathcal{C}.
\]
In this section, we prove that the curve $C$ defined in (2.14) will have a bifurcation point at a finite $\omega = \omega^*$. To this end, we first define the matrices $H_1(z)$ and $H_2 \in \mathbb{R}^{2N \times 2N}$ by
\[
H_1(z) + \alpha H_2 = \begin{bmatrix} A + 2\alpha [u^T] & 2\alpha [u_1 \cdots u_N] \\ 2\alpha [u_1 \cdots u_N] & A + 2\alpha [u^T] \end{bmatrix} + \alpha \begin{bmatrix} 0 & -S \\ S & 0 \end{bmatrix}
\]
\[
= \begin{bmatrix} \hat{A} + 2\alpha [u^T] & 2\alpha [u_1 \cdots u_N] - S \\ 2\alpha [u_1 \cdots u_N] + \alpha S & \hat{A} + 2\alpha [u^T] \end{bmatrix},
\]
where $\hat{A}$ and $H(y)$ are defined in (2.8) and (2.7), respectively.

**Theorem 3.1.** Suppose that $0 < \alpha < \infty$ and $p = \omega$ in (2.2), and the assumption (H) holds. Then, the primal stalk described by (3.2) has at least $n = \text{rank} (S) - 3$ bifurcation points at a finite valued $\omega = \omega^*_i, i = 1, \ldots, n$. That is, the matrix $B(z^\omega, \omega^*_i) = \begin{bmatrix} G_{i} \alpha \omega \end{bmatrix}$ in (2.20) is singular on $C^*_{\omega} = \omega = \omega^*_i$, where $z^\omega = \left( \frac{1}{\sqrt{2}} u^T, \frac{1}{\sqrt{2}} u^T, \lambda \right)^T$.

**Proof.** From (2.6a), (2.15) and (3.3), we see that
\[
B(z^\omega, \omega) \equiv \begin{bmatrix} H & G_{i} \omega \\ G^T \alpha & 0 \end{bmatrix} \in \mathbb{R}^{(2N+2) \times (2N+2)},
\]
where $G_{i} \omega = \left( \frac{1}{\sqrt{2}} u^T, \frac{1}{\sqrt{2}} u^T \right)^T$ and $a_0 = \left( \frac{1}{\sqrt{2}} u^T, \frac{1}{\sqrt{2}} u^T \right)^T$.

For $\omega = 0$, we have $H_0 \equiv H(z^\omega, 0) = H_1(z^\omega) = \begin{bmatrix} \hat{A} + 2\alpha [u^T]^2 & \alpha [u^T]^2 \\ \alpha [u^T]^2 & \hat{A} + 2\alpha [u^T]^2 \end{bmatrix} \in \mathbb{R}^{2N \times 2N}$, with
\[
\hat{A} = A + \alpha [u^T]^2 - \lambda_0 I_N.
\]

Using the fact that $u^T \in \mathbb{R}^N$ is the ground state of (3.1), $\lambda_0$ is the smallest eigenvalue of the matrix $A + \alpha [u^T]^2$. On the other hand, $A + \alpha [u^T]^2$ is an irreducible $M$-matrix and $u^T > 0$ is the unique eigenvector corresponding to the smallest eigenvalue $\lambda_0$. Hence, $\hat{A}$ in (3.5b) is positive semi-definite with a simple zero eigenvalue. From (3.5a), it follows that $\lambda(H_0) = \lambda(\hat{A} + 2\alpha [u^T]^2) \cup \lambda(\hat{A})$. Here, $\lambda(\cdot)$ denotes the set of eigenvalues of a matrix. However, $\hat{A} + 2\alpha [u^T]^2$ is positive definite since the ground state $u^T$ is a positive vector, implying that $H_0$ is positive semi-definite with one-dimensional null space. From (3.4) and since $\min_{|x|=1, x \in \mathbb{R}^N} \langle x, 0, 0 \rangle B_1 \frac{z^\omega}{(x', 0, 0)} \frac{z^\omega}{(x', 0, 0)} \frac{z^\omega}{(x', 0, 0)} = 0$, using Courant-Fischer Theorem [25, p. 179] we find that the smallest third eigenvalue of $B(z^\omega, 0)$ is non-negative. Hence, $B(z^\omega, 0)$ has 3 negative eigenvalues at most.

Next, we claim that the smallest rank (S) eigenvalues of $B(z^\omega, \omega)$ will become negative when $\omega$ is sufficiently large. Since $S$ is a real $N \times N$ skew-symmetric matrix, the eigenvalues of $S$ can be written as $\pm \mu_1, \ldots, \pm \mu_m, 0, \ldots, 0$, where $\mu_i > 0$. Note that $m = \text{rank} (S)/2$. Therefore, the eigenvalues of $H_2 = \begin{bmatrix} 0 & -\alpha \omega \end{bmatrix}$ are $\mp \alpha \omega \mu_1, \ldots, \mp \alpha \omega \mu_m, 0, \ldots, 0$. From (3.3) and (3.4), it follows that
\[
\lim_{\omega \to +\infty} -B(z^\omega, \omega) = \begin{bmatrix} H_2 & 0 \\ 0 & 0 \end{bmatrix}.
\]

This implies that the smallest $2m = \text{rank} (S)$ eigenvalues of the matrix $B(z^\omega, \omega)$ are negative for values of $\omega$ that are sufficiently large. Also, using the fact that there are no more than 3 non-positive eigenvalues of $B(z^\omega, 0)$, it follows that there exist at least $n = 2m - 3$ bifurcation points, $\omega^*_i \in (0, \infty), i = 1, \ldots, 2m - 3$ such that the matrix $B(z^\omega, \omega^*_i)$ is singular. The proof is complete.

**4. Bistability of a rotating BEC**

In Section 3, we analyzed bifurcation phenomena for the DNLS equations with isotropic traps $(\gamma_\alpha = \gamma_\beta)$. We found that the curve of the ground state solution has a bifurcation point. However, it remains unclear whether bistable states exist for certain $\omega$ values. In this section we perform numerical studies by considering various settings for $\gamma_\alpha$ and $\gamma_\beta$ using our MATLAB version of the QTICM. Finding an S-shape-like bistability in the DNLS equations is not only a novel result to the best of the authors’ knowledge, but it would also provide a clue as to why both 0- and 1-vortex configurations can be found experimentally for a certain angular velocity, as reported in [2].

To conduct the numerical experiments, we assume that $\omega$ is the continuation parameter $\beta$. That is, we set $\mu_0 = 0, \omega_0 = 0,$ and $\nu_0 = 1$ in (2.2). The mesh size of the grid domain $\Delta \kappa$ is chosen to be $\kappa = 0.2$. The tolerance of each step in Newton’s correction is chosen to be $\epsilon_\gamma = 10^{-8}$. To compute the initial solution of the continuation method, we set $\omega = 0$ and use the method discussed in [12].

Our computational results are summarized in the following three subsections. In Section 4.1, we assert the existence of a bistable region by examining the bifurcation diagram, solution profiles, and the corresponding solution stability. In particular, we find an interval $[\omega_1, \omega_2]$ where the DNLS equations (2.3) have two types of stable state solutions containing 0- and 1-vortex solutions. Since the trap potential and intra-component scattering length are also tunable factors in BEC experiments, we study how the trap potentials and intra-component scattering length affect the bistable region and the corresponding energy curves in Sections 4.2 and 4.3, respectively.

**4.1. Bistability and bifurcation diagram**

In this subsection, we assume that $\alpha = 100, V(x) = 1.5x^2 + y^2$, and $\omega = [-5, 5] \times [-6, 6]$. Fig. 4.2 shows the bifurcation diagram of the DNLS equations (2.3) for $\omega \in (0, 0.8)$. The stable and unstable state solution curves are plotted with blue lines and red lines, respectively, and the corresponding nodal domains of density $u_1 + u_2$ are attached next to the solution curves. Note that the solution curves are defined in (2.14), the solution stability is defined in Definition 2.1, and the stability criteria are described in Theorem 2.4.

Fig. 4.2(a) shows how the solution curves bifurcate along the continuation parameter $\omega \in (0, 0.8)$. Starting at $\omega = 0$, we trace the stable solution curve $\frac{c_0}{\epsilon_0}$ with a 0-vortex until the first bifurcation point, which is located at $\omega = 0.607$. The corresponding (excited) solutions of the bifurcation branch $\frac{c_0}{\epsilon_0}$ and $\frac{c_0}{\epsilon_1}$ contain vortices entering the domain from the right and left hand side, respectively. The solution curves bifurcate again at $\omega = 0.397$, where a single vortex is found at the center of the domain. We can continuously increase $\omega$ to obtain a stable solution curve $\frac{c_0}{\epsilon_0}$ with a 1-vortex until the next bifurcation point occurs. Note that two turning points at $\omega = 0.520$ and 0.728 are also marked.

Fig. 4.2(a) also shows that there is bistability in the DNLS equations. We denote the bistable region in terms of $\omega$ as $[\omega_1, \omega_2] = [0.397, 0.607]$. In this region, the set of DNLS equations contains two stable solutions. We can denote the 0-, 0-1- and 1-vortex regions as $[0, 0.397], [0.397, 0.607], [0.607, 0.728]$, respectively. These numerical results agree with the experiments.
we conclude that the trap frequency marginally affects the width of the bistable regions, while the bistable region moves to the right as $\gamma_x$ increases.

4.3. Effect of the Intra-component scattering length

Finally, we study how the intra-component scattering length affects the bistable region. We assume the intra-component scattering length $\alpha \in [10, 200]$, the potential trap $V(x) = 1.5x^2 + y^2$ and computational domain $\Omega = [-5, 5] \times [-6, 6]$. We use $[\omega_1(\gamma_x), \omega_2(\gamma_x)]$ to denote the bistable region, which is a function of $\gamma_x$.

Fig. 4.3(a) demonstrates the values of $\omega_1(\gamma_x)$ and $\omega_2(\gamma_x)$ versus $\gamma_x \in [1, 2]$. The figure also plots $\omega_x$ to indicate the frequency at which the energy curves corresponding to the 0- and 1-vortices cross. The figure shows that both $\omega_1(\gamma_x)$ and $\omega_2(\gamma_x)$ increase almost linearly as $\gamma_x$ increases. In addition, the widths of the bistable regions, i.e., $[\omega_2(\gamma_x) - \omega_1(\gamma_x)]$, only slightly vary as $\gamma_x$ varies. In Fig. 4.3(b), the energy curves for the stable state solutions are plotted for $\gamma_x = 1, 0.5, 1.2, 1.4, 1.6, 1.8,$ and $2$. The energy curves gradually change from linear to quadratic-like functions.

We use $\omega_1(\alpha)$, $\omega_2(\alpha)$ to denote the bistable region, which is a function of $\alpha$.

Fig. 4.4(a) demonstrates the values of $\omega_1(\alpha)$ and $\omega_2(\alpha)$ versus $\alpha \in [10, 200]$. The figure also plots $\omega_x(\alpha)$ to indicate the frequency at which the energy curves corresponding to 0- and 1-vortices intersect. The curves of $\omega_1(\alpha)$ and $\omega_2(\alpha)$ decrease exponentially. Unlike the behavior reported in Section 4.2, the widths of the bistable regions grow larger, but gradually approach a constant value as $\alpha$ increases. In Fig. 4.4(b), the energy curves for the stable state solutions are plotted for $\alpha = 10, 40, 80, 120, 160,$ and $200$.

Fig. 4.4 shows that the intra-component scattering length significantly affects the width of the bistable regions, while the bistable region moves to the left as $\alpha$ increases.

5. Conclusions

We have developed a quotient transformation invariant continuation method for the numerical computation of the ground and excited states of a rotating BEC. Our method circumvents the difficulties associated with transformation invariant solutions. We analytically studied the bifurcation properties of the primal stalk solution curve corresponding to the DNLS equations for an isotropic trap. In addition, numerical experiments have shown that our method is reliable and finds the bistable solutions corresponding to stable states with a 0- or 1-vortex. This result agrees with experimental observations and explains these observations numerically. We have also reported how the trap frequency and the intra-component scattering length affect the bistability.

Behaviors like vortex patterns and stability of rotating Bose–Einstein condensates in rapid rotating region have not been understood completely. As reported in [2], multiple vortices are observed as the value of $\omega$ is increased, which complicates the dynamics of the system. We use the QTICM to trace some solution curves as $\omega$ approaches 1. Fig. 5.5 shows some of the connected numerical experiment results based on the method also lead to the discovery of the bistable region.

4.2. Effect of trap potential

Now, we study how the trap potential affects the bistable region. In particular, we consider the trap $V(x) = \gamma_x x^2 + \gamma_y y^2$ for the trap frequencies $\gamma_x \in [1, 2]$ and $\gamma_y = 1$. In addition, we choose the intra-component scattering length to be $\alpha = 100$ and the computational domain to be $\Omega = [-6, 6] \times [-6, 6]$. We use $[\omega_1(\gamma_x), \omega_2(\gamma_x)]$ to denote the bistable region, which is a function of $\gamma_x$.

Fig. 4.3(a) demonstrates the values of $\omega_1(\gamma_x)$ and $\omega_2(\gamma_x)$ versus $\gamma_x \in [1, 2]$. The figure also plots $\omega_x$ to indicate the frequency at which the energy curves corresponding to the 0- and 1-vortices cross. The figure shows that both $\omega_1(\gamma_x)$ and $\omega_2(\gamma_x)$ increase almost linearly as $\gamma_x$ increases. In addition, the widths of the bistable regions, i.e., $[\omega_2(\gamma_x) - \omega_1(\gamma_x)]$, only slightly vary as $\gamma_x$ varies. In Fig. 4.3(b), the energy curves for the stable state solutions are plotted for $\gamma_x = 1, 0.5, 1.2, 1.4, 1.6, 1.8,$ and $2$. The energy curves gradually change from linear to quadratic-like functions.

We conclude that the trap frequency marginally affects the width of the bistable regions, while the bistable region moves to the right as $\gamma_x$ increases.

4.3. Effect of the Intra-component scattering length

Finally, we study how the intra-component scattering length affects the bistable region. We assume the intra-component scattering length $\alpha \in [10, 200]$, the potential trap $V(x) = 1.5x^2 + y^2$ and computational domain $\Omega = [-5, 5] \times [-6, 6]$. We use $[\omega_1(\alpha), \omega_2(\alpha)]$ to denote the bistable region, which is a function of $\alpha$.

Fig. 4.4(a) demonstrates the values of $\omega_1(\alpha)$ and $\omega_2(\alpha)$ versus $\alpha \in [10, 200]$. The figure also plots $\omega_x(\alpha)$ to indicate the frequency at which the energy curves corresponding to 0- and 1-vortices intersect. The curves of $\omega_1(\alpha)$ and $\omega_2(\alpha)$ decrease exponentially. Unlike the behavior reported in Section 4.2, the widths of the bistable regions grow larger, but gradually approach a constant value as $\alpha$ increases. In Fig. 4.4(b), the energy curves for the stable state solutions are plotted for $\alpha = 10, 40, 80, 120, 160,$ and $200$.

Fig. 4.4 shows that the intra-component scattering length significantly affects the width of the bistable regions, while the bistable region moves to the left as $\alpha$ increases.
Table 4.1
The points $ω_1$ and $ω_2$ corresponding to the bistable region and the energy for $ω = 0$ are computed by using different mesh size $h_i = \frac{1}{2^{i+1}}$ for $i = 0, 1, 2, 3$. The convergence rates are presented in the columns labeled by CR.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$ω_1$</th>
<th>CR</th>
<th>$ω_2$</th>
<th>CR</th>
<th>Energy ($ω = 0)$</th>
<th>CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_0$</td>
<td>0.3967274381981</td>
<td>−</td>
<td>0.6065765780083</td>
<td>−</td>
<td>4.369505165262</td>
<td>−</td>
</tr>
<tr>
<td>$h_1$</td>
<td>0.3978574022653</td>
<td>−</td>
<td>0.6059499956020</td>
<td>−</td>
<td>4.369781584647</td>
<td>−</td>
</tr>
<tr>
<td>$h_2$</td>
<td>0.3984879007864</td>
<td>1.62</td>
<td>0.6057003973605</td>
<td>2.27</td>
<td>4.369781584646</td>
<td>2.15</td>
</tr>
<tr>
<td>$h_3$</td>
<td>0.3987605482082</td>
<td>1.95</td>
<td>0.6055944117548</td>
<td>2.11</td>
<td>4.369950507681</td>
<td>1.92</td>
</tr>
</tbody>
</table>

Fig. 4.3. (a) The change to the bistable region [$ω_1(γ_x), ω_2(γ_x)$] in terms of $γ_x \in [1, 2]$. Here, we assume that the potential is given by $V(x) = γ_x x^2 + y^2$ and $α = 100$. The frequency corresponding to the intersection of 0- and 1-vortex energy curves is denoted by $ω_c(γ_x)$. (b) Energy curves containing 0- or 1-vortex for $γ_x = 1.0, 1.2, 1.4, 1.6, 1.8$, and 2.

solution curves that $ω$ ranges from 0.500 to 0.986. The figure suggests that more and more bifurcation branches can be found by the QTICM as $ω$ increases. However, it is not clear such a path following scheme may find all the solution curves of the DNLS equations and then illustrate the complete solution behaviors. In other words, the QTICM may act as a tool to find some of the connected solutions of the DNLS equations, it remains a challenge to completely understand the dynamic system as $ω$ increases to 1.

Acknowledgements
The authors are grateful to the anonymous referees for their valuable suggestions and comments. This work is partially supported by the National Science Council of Taiwan, the National Center for Theoretical Sciences, and the Taida Institute of Mathematical Sciences.

Appendix

A.1. Proof of Theorem 2.1

It is straightforward to see that statements (i) and (ii) are equivalent. Thus, we show that $N(\mathbf{j}(s)) \neq \{0\}$ is equivalent to the fact that $\mathbf{B}(s)$ is singular.

Necessity condition. If $N(\mathbf{j}(s)) \neq \{0\}$ then there exists a nonzero vector $\mathbf{x} \in \mathbb{R}^N$ such that $\mathbf{j}(s)x = 0$. It is easy to verify that $\mathbf{B}(s)(\mathbf{x}^T, 0)^T = 0$ and, therefore, $\mathbf{B}(s)$ is singular.
Fig. 4.4. (a) The change to the bistable region $[\omega_1(\alpha), \omega_2(\alpha)]$ in terms of $\alpha \in [10, 200]$. Here, we assume that the potential is given by $V(x) = 1.5x^2 + y^2$. The frequency corresponding to the intersection of 0- and 1-vortex energy curves is denoted by $\omega_c(\alpha)$. (b) The energy curves containing a 0- or 1-vortex for $\alpha = 10, 40, 80, 120, 160$, and 200.

Fig. 5.5. A snapshot of bifurcation diagram for $\omega$ ranges from 0.500 to 0.986. We set $V(x) = 1.5x^2 + y^2$ and $\alpha = 100$. 
Suppose $B(s)$ is singular, then there exists a nonzero vector $(x^T, \xi)^T \in \mathbb{R}^{2N+2}$ such that
\begin{equation}
B(s) \begin{bmatrix} x \\ \xi \end{bmatrix} = 0, \tag{A.1}
\end{equation}
which implies
\begin{equation}
G_\vartheta x + \xi a_\vartheta = 0. \tag{A.2}
\end{equation}
Differentiating (2.11) with respect to $\vartheta$, we have
\begin{equation}
G_\vartheta a = 0. \tag{A.3}
\end{equation}
Multiplying $a_\vartheta^T$ from the left to (A.2), we have
\begin{equation}
a_\vartheta^T G_\vartheta x + (a_\vartheta^T a_\vartheta) \xi = 0. \tag{A.4}
\end{equation}
Since $G_\vartheta$ is symmetric, from (A.4), we see that $a_\vartheta^T G_\vartheta x = 0$. Hence $\xi = 0$. From (A.1), it can be seen that $f(s)x = 0$. Namely, statement (ii) holds.

### A.2. Proof of Theorem 2.4

In order to prove the Theorem 2.4, we first show some preliminary propositions and theorems.

**Definition A.1.** $A$ is congruent to $B$ if there is a non-singular matrix $U$ such that $A = U^T B U$.

**Proposition A.1.** If $A$ is an $n \times n$ symmetric matrix, $\|u\| = 1$ and $a = u^T A u \neq 0$, then $\begin{bmatrix} A & u \\ u^T & 0 \end{bmatrix}$ is congruent to $\begin{bmatrix} P^T \bar{\varphi} P - \frac{1}{a} u u^T & 0 \\ 0 & a \end{bmatrix}$, where $P = I_n - uu^T$, $a = u^T A u$ and $\bar{\varphi} = P^T A u + u$.

**Proof.** Let $U_1 = \begin{bmatrix} P & u \\ u^T & 0 \end{bmatrix}$, where $P = I_n - uu^T$. Since $\|u\| = 1$, then the matrix $P$ is an orthogonal projection of $\mathbb{R}^n$ on $u^\perp$, i.e., $P$ is a symmetric matrix with $P = P^T$. It is easy to check that $U_1$ is a symmetric orthogonal matrix. Then we compute
\begin{equation}
U_1^T \begin{bmatrix} A & u \\ u^T & 0 \end{bmatrix} U_1 = \begin{bmatrix} P^T \bar{\varphi} P & P^T A u + u \\ u^T P^T A u + u & uu^T \end{bmatrix}. \tag{A.5}
\end{equation}
Let $U_2 = \begin{bmatrix} I_{n-1} & 0 \\ -1/a & 1 \end{bmatrix}$ where $\bar{\varphi} = P^T A u + u$. Since $a = u^T A u \neq 0$, from (A.5) we have
\begin{equation}
U_1^T U_2^T \begin{bmatrix} A & u \\ u^T & 0 \end{bmatrix} U_1 U_2 = \begin{bmatrix} P^T \bar{\varphi} P - \frac{1}{a} uu^T & 0 \\ 0 & a \end{bmatrix}. \tag{A.6}
\end{equation}
The proof is completed. □

Since the eigenvalues of the symmetric matrix $A$ are real, we adopt the convention that they are labeled according to non-decreasing size:
\begin{equation}
\lambda_{min} = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n-1} \leq \lambda_n = \lambda_{max}.
\end{equation}

**Definition A.2.** Let $A$ be an $n \times n$ real symmetric matrix. The inertia of $A$ is the ordered triple
\begin{equation}
i(A) = (i_+(A), i_-(A), i_0(A))
\end{equation}
where $i_+(A)$ is the number of positive eigenvalues of $A$, $i_-(A)$ is the number of negative eigenvalues of $A$ and $i_0(A)$ is the number of zero eigenvalues of $A$ (multiplicity is counted for each).

The following propositions can then be verified.

**Proposition A.2.** Let $A$, $B$ be $n \times n$ real symmetric matrices. Then $A$ is congruent to $B$ if and only if $i(A) = i(B)$.

**Proposition A.3.** Suppose $A$ is real symmetric. Let $\bar{u} = P^T A u + u$ and let $P = I_n - uu^T$ be an orthogonal projection of $\mathbb{R}^n$ on $u^\perp$. Then
\begin{equation}
\begin{array}{ll}
ia. & i(P^T A P - \tau uu^T) = i(Z^T AZ) + (0, 1, 0) \quad \text{if } \tau > 0, \\
b. & i(P^T A P - \tau uu^T) = i(Z^T AZ) + (1, 0, 0) \quad \text{if } \tau < 0,
\end{array}
\end{equation}
where $Z$ is a matrix whose columns form an orthonormal basis of $u^\perp$.

**Proof.** Let $U_1 = [u, u_2, \ldots, u_n] = [u, Z]$ be an orthogonal matrix with first column is $u$, then
\begin{equation}
PU_1 = (I - uu^T)U_1 = [0, Z], \tag{A.6}
\end{equation}
and since $u^T P A u = 0$, then we have
\begin{equation}
U_1^T \bar{u} = U_1^T (u + PAu) = e_1 + [0, w^T]^T, \tag{A.7}
\end{equation}
where $w = Z^T PAu \in \mathbb{R}^{n-1}$. From (A.6) and (A.7), we compute
\begin{equation}
U_1^T (P^T A P - \tau uu^T)U_1 = U_1^T P^T A P U_1 - \tau U_1 u u^T U_1 = \begin{bmatrix} 0^T \\ Z \end{bmatrix} A[0, Z] - \tau \begin{bmatrix} 1 \\ w \end{bmatrix} [1, w^T]. \tag{A.8}
\end{equation}
Let $U_2 = \begin{bmatrix} 1 & -w \\ 0 & 1 \end{bmatrix}$ and from (A.8) we have
\begin{equation}
U_1^T U_2^T (P^T A P - \tau uu^T)U_1 U_2 = \begin{bmatrix} 0 & 0^T \\ 0 & Z^T AZ \end{bmatrix}. \tag{A.9}
\end{equation}

From (A.9), (A.10) and Proposition A.3, it is easy to verify
\begin{equation}
i(P^T A P - \tau uu^T) = i(Z^T AZ) + (0, 1, 0), \quad \text{if } \tau > 0,
\end{equation}
\begin{equation}
i(P^T A P - \tau uu^T) = i(Z^T AZ) + (1, 0, 0), \quad \text{if } \tau < 0. \quad \Box
\end{equation}

**Proof of Theorem 2.4.** Let $v = u'/\|u\|_2$, then it is easily seen that
\begin{equation}
i \begin{bmatrix} H(y') \\ u'^T \end{bmatrix} = i \begin{bmatrix} H(y') \\ v'^T \end{bmatrix}. \tag{A.11}
\end{equation}
First, we consider the case $a > 0$. From Propositions A.1-A.3, we have
\begin{equation}
i \begin{bmatrix} H(y') \\ v'^T \end{bmatrix} = i \begin{bmatrix} P^T H(y')P - \|u\|^2_a v v^T \\ 0^T \end{bmatrix} \tag{A.12}
\end{equation}
and
\begin{equation}
i \begin{bmatrix} H(y') \\ u'^T \end{bmatrix} = i(H_{u+}(y')) + (1, 1, 0). \tag{A.13}
\end{equation}
where $P = I_{2N} - vv^T$ and $v = P^T H(y')v + v$. From (A.11) and (A.12), we obtain
\begin{equation}
i \begin{bmatrix} H(y') \\ v'^T \end{bmatrix} = i \begin{bmatrix} P^T H(y')P - \|u\|^2_a v v^T \\ 0^T \end{bmatrix} + (1, 0, 0) \tag{A.14}
\end{equation}
and
\begin{equation}
i \begin{bmatrix} H(y') \\ u'^T \end{bmatrix} = i(H_{u+}(y')) + (1, 1, 0). \tag{A.15}
\end{equation}
Similarity, in the case $a < 0$, we also have
\begin{equation}
i \begin{bmatrix} H(y') \\ u'^T \end{bmatrix} = i(H_{u+}(y')) + (1, 1, 0). \tag{A.16}
\end{equation}
The proof is complete. □
References


