# Computation of invariant tori by Newton-Krylov methods in large-scale dissipative systems.

J. Sánchez[a,*], M. Net[a], C. Simó[b]

[a] *Dept. de Física Aplicada, Universitat Politècnica de Catalunya, Jordi Girona Salgado s/n. Campus Nord. Mòdul B4, 08034 Barcelona, Spain*
[b] *Dept. de Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, Gran Via de les Corts Catalanes, 585, 08071 Barcelona, Spain*

**Abstract**

A method to compute invariant tori in high-dimensional systems, obtained as discretizations of PDEs, by continuation and Newton-Krylov methods is described. Invariant tori are found as fixed points of a generalized Poincaré map without increasing the dimension of the original system. Due to the dissipative nature of the systems considered, the convergence of the linear solvers is extremely fast. The computation of periodic orbits inside the Arnold's tongues is also considered. Thermal convection of a binary mixture of fluids, in a rectangular cavity, has been used to test the method.

*Key words:* Continuation methods, invariant tori, generalized Poincaré map, variational equations, Newton-Krylov methods, resonances.

## 1. Introduction

Invariant tori arise in general continuous or discrete dynamical systems, usually, when a branch of periodic orbits loses stability at a Neimark-Sacker bifurcation. The case of Hamiltonian systems is more subtle: invariant tori appear typically around a (totally or partially) elliptic fixed point, periodic orbit or lower dimensional torus (see, e.g. [1],[2]). In a general dissipative system attracting 2D invariant tori can be created by a Neimark-Sacker bifurcation of a periodic orbit. Later on, as illustrated in Section 6.1, the invariant tori can have a cascade of period-doubling bifurcations leading to chaos. Typically only a finite number of period doublings is found before the chaotic range is reached (see, e.g., [3] in the context of fluid mechanics). Another common possibility is the occurrence of a second Neimark-Sacker bifurcation and 3D invariant tori are created by bifurcation of a 2D torus which becomes unstable. When the dynamics on an attracting 3D torus comes close enough to a double resonance one can expect the occurrence of strange attractors. This is the well-known Ruelle-Takens route to turbulence. A description of the mechanisms leading to these attractors can be found on [4] for a family of 3D diffeomorphisms. Hence, invariant tori play a relevant role in several routes to chaos for dissipative systems.

*Corresponding author's e-mail: sanchez@fa.upc.edu

For low-dimensional conservative or dissipative systems, several algorithms have been developed in the past. In a local case, close to a fixed point, Lindstedt–Poincaré methods can be used, see [5, 6], even to compute the invariant manifolds for partially normally hyperbolic tori [7]. Far away from fixed points different methods have been proposed, see for instance [8, 9, 10, 2, 11], including finding a point on an invariant curve as a fixed point of a generalized Poincaré map, or expanding the tori or the invariant curves in Fourier series. We address the interested reader to the historical introduction in [12] which includes some examples. In the case of high-dimensional systems, tori are usually computed by time evolution of the equations. Only attracting tori are obtained by this method. Therefore it is very expensive to study their dependence on parameters, specially near bifurcations, where due to the weak attraction the transients are very long.

We report here a method to compute stable or slightly unstable invariant tori which, with minor differences, has been used in the past in low-dimensional dissipative [9] and the conservative [10, 2] cases. Up to now, these methods have never been used for problems coming from discretizations of partial differential equations, because the matrices of the linear algebra involved are full, no matter the spatial discretization employed. For instance, suppose a system of differential equations has the form $\dot{x}_i = f_i(x_{i-1}, x_i, x_{i+1})$, $i = 1, \cdots, n$, with periodic conditions $x_0 = x_n$ and $x_{n+1} = x_1$, and an initial condition $x_i^0$. The Jacobian of the system is always tridiagonal. If it is integrated from $t = 0$ with an explicit Euler method with step $\Delta t$, after the first step $x_i(\Delta t)$ depends on $x_{i-1}^0$, $x_i^0$, and $x_{i+1}^0$, after the second it will depend on five of the initial conditions, and after $(n-1)/2$ steps, on all initial conditions. Then $\partial_{x_j^0} x_i(t) \neq 0$ for all $i$ and $j$, if $t > (n-1)\Delta t/2$. If an implicit Euler method is used, this holds from the first step. This gives a full Jacobian for the time evolution operator, and for any derivated operator as, for instance, the Poincaré map.

Instead of using direct methods to solve the linear systems we use iterative Krylov methods. Due to the dissipative nature of the problems the method is addressed to, the number of iterations required is very small, and the full version of GMRES can be used, which is optimal among the family of projection Krylov methods [13]. The technique generalizes that described previously in [14] for the computation of fixed points and periodic orbits.

In order to perform the numerical experiments, we have applied the method to the thermal convection problem of a binary mixture, filling a two-dimensional rectangular domain. This problem has been studied widely by several authors, mainly in large aspect ratio domains. From weakly nonlinear theory, laboratory experiments and direct numerical simulations (see [15, 16, 17, 18, 19] among many others) it is well known that the onset of convection is subcritical and oscillatory below a critical negative value of the separation ratio, and that, depending on $\Gamma$, it can give rise to a very complex dynamics. Although the analysis of the physical problem is out of the scope of the paper, it is worthwhile to mention that this is the first time the unstable branch of periodic orbits bifurcated from the steady basic state, and branches of invariant tori have been computed, and their stability analyzed. Therefore the continuation techniques shown in this paper, applied to the computation of complete bifurcation diagrams, can help to confirm the validity of the existent theories about the origin of the weakly non-linear, but complex time-dependent states, found when the

convection of a binary mixture sets in.

In general, there is a broad range of areas in which multi-diffusive convection has to be taken into account to understand the nonlinear behaviour, not only in nature (Oceanography, Astrophysics, Geophysics, Geology, etc.), but also in Engineering (Metallurgy, Materials Science, etc.) (see[20] for a detailed review). Moreover, as it was said before, the continuation techniques described below can be applied to any dissipative process, which admits a mathematical model, for instance, those that occur in many Biological and Medical or Ecological systems.

Section 2 sums up the generic continuation method used for fixed points of maps. Section 3 explains the computation of Poincaré maps, and the action by their differentials. In Section 4 the generalized Poincaré map is introduced, as well as the way of computing the action by its differentials. The problem used as test for the method is presented in Section 5. The results of the computation of invariant tori in the test problem are summarized in Section 6. Some data on the efficiency of the problem are given in Section 7, and, finally, we end with the conclusions in Section 8.

## 2. Continuation of fixed points of maps by Newton-Krylov methods

Consider a dissipative map

$$x \to G(x,\lambda), \quad (x,\lambda) \in \mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}, \tag{1}$$

with $n \gg 1$. We are interested in its fixed points and their dependence on $\lambda$. Predictor-corrector parameter or pseudo-arclength-like continuation methods can be used to trace the curves $(x(s), \lambda(s))$ of fixed points parameterized by the arclength [21]. They admit an unified formulation by adding the equation of the hyperplane

$$m(x,\lambda) \equiv \theta < v_x, x - x_0 > +(1-\theta)v_\lambda(\lambda - \lambda_0) = 0, \tag{2}$$

$(x_0, \lambda_0)$ and $(v_x, v_\lambda)$ being the predicted point and the tangent to the curve of solutions, obtained from previously computed points by extrapolation (see Fig. 1), and $\theta \in [0,1]$ . If $\theta = 0$ the equation fixes $\lambda$ to the predicted value $\lambda_0$, and if $\theta = 1/2$ the hyperplane $m(x,\lambda) = 0$ is approximately normal to the curve. The hyperplane will cut the curve of solutions if the prediction is not too far away from the last computed point. Then, the system that determines a unique pair, $(x,\lambda)$, is

$$F(x,\lambda) = \begin{pmatrix} x - G(x,\lambda) \\ m(x,\lambda) \end{pmatrix} = 0 \in \mathbb{R}^n \times \mathbb{R}\,. \tag{3}$$

The existence of a unique local solution follows from the Implicit Function Theorem if $DF$ is regular. System (3) can be solved by Newton's method, $(x^{i+1}, \lambda^{i+1}) = (x^i, \lambda^i) + (\Delta x^i, \Delta\lambda^i)$. At each iteration, the linear system

$$\begin{pmatrix} I - D_x G(x^i,\lambda^i) & -D_\lambda G(x^i,\lambda^i) \\ \theta v_x^\top & (1-\theta)v_\lambda \end{pmatrix} \begin{pmatrix} \Delta x^i \\ \Delta\lambda^i \end{pmatrix} = \begin{pmatrix} -x^i + G(x^i,\lambda^i) \\ -m(x^i,\lambda^i) \end{pmatrix}, \tag{4}$$

is solved iteratively by matrix-free methods (GMRES, GMRES(M), BICGSTAB, etc.), which only require the computation of matrix products, i.e., products of the form

$$D_x G(x^i,\lambda^i)\Delta x^i + D_\lambda G(x^i,\lambda^i)\Delta\lambda^i.$$

The key point to solve these linear systems, is that they need no preconditioning if the Jacobian $D_x G(x^i, \lambda^i)$ has most of its spectrum clustered around the origin. This is what happens for the fixed points of dissipative problems we want to compute. Dissipative systems contract volumes in phase space, namely, trajectories near any invariant object (fixed points, periodic orbits, invariant tori, etc.) tend to the object, except along its unstable manifold which we assume will be of very small dimension compared with the total dimension of the system. This makes the evolution operator be a contraction (except along the low dimensional unstable manifold). The longer the time integration, the more contractive along the stable manifold. This is also related to the fact that the spectrum of a linearization of the field defining a system of dissipative differential equations is located on the left of the imaginary axis, except by a few elements with non-negative real part.

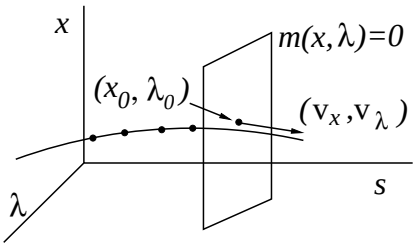


Figure 1: Predictor-corrector continuation method.

## 3. Computation of Poincaré maps

Consider a system of ordinary differential equations (ODEs)

$$\dot{x} = f(x, \lambda), \quad (x, \lambda) \in \mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}, \tag{5}$$

obtained after the spatial discretization of a set of parabolic partial differential equations (PDEs), such as those governing reaction-diffusion or fluid mechanic problems. We define the Poincaré map, $P : \mathcal{V} \subset \Sigma_1 \to \Sigma_1$, for simplicity, on a hyperplane

$$\Sigma_1 = \{\, x \in \mathbb{R}^n \,/\, < v_1, x - x_1 >= 0\}$$

by $P(x, \lambda) = \phi(t(x), x)$, where $\phi(t, x)$ is the solution of (5) with initial condition $x$, and $t(x) > 0$ is the first time at which the orbit starting at $x$ intersects $\Sigma_1$ again in the sense defined by the vector $f(x, \lambda)$. In some cases, depending on the geometry of the orbits, instead of the first intersection one should consider another one, or to look for intersections on a suitable target domain and, furthermore, it can be convenient to change the Poincaré section when some parameter in the family changes. But this will not be considered in what follows. When $x$ is on a periodic orbit, $t(x)$ is its period. If the orbit doubles the period $t(x)$ is the second time it intersects $\Sigma_1$ and so on.

To avoid the construction of an orthonormal basis on $\Sigma_1$, $P$ can be parameterized as sketched in Fig. 2a. If $k$ is the index of the largest component of $v_1$, and $R_k$ is the orthogonal projection from $\Sigma_1$ onto the hyperplane $x_k = 0$, then we define $\bar{P} : \bar{\mathcal{V}} \subset \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ as

$$\bar{P}(\bar{x}, \lambda) = R_k(P(R_k^{-1}(\bar{x}), \lambda)). \tag{6}$$

Details on the explicit computation of the map $R_k$, and its inverse are given in [14]. There and in [22] periodic orbits of (5) were obtained by computing the fixed points of $\bar{P}$.

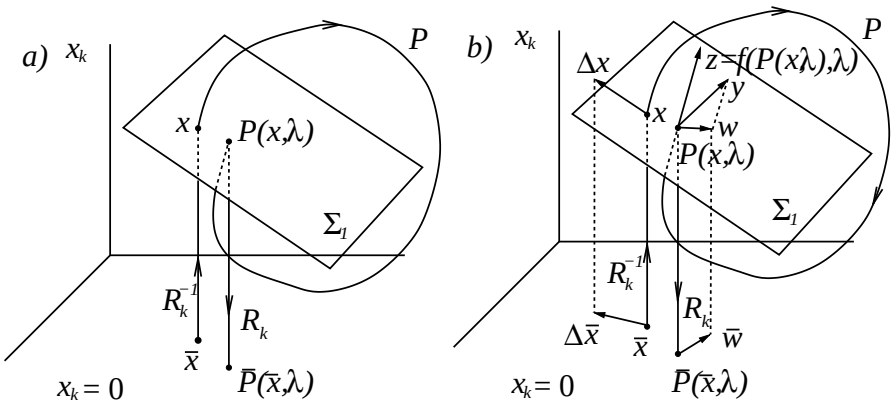


Figure 2: Parameterized Poincaré map, and the action by its differential. a) The scheme shows the process required to compute the parameterized Poincaré map. Starting from $\bar{x}$ belonging to the hyperplane $x_k = 0$, $x$ is computed as the point in $\Sigma_1$ whose orthogonal projection onto $x_k = 0$ is $\bar{x}$. The Poincaré map is applied next to obtain $P(x, \lambda)$, which is finally projected onto $x_k = 0$, providing $\bar{P}(\bar{x}, \lambda)$. b) The analogous process required to obtain the action of the Jacobian of $\bar{P}$ on a vector $\Delta\bar{x}$ (see eq. (7)).

By applying the chain rule to (6), a matrix product of the Jacobian of $\bar{P}$ times $(\Delta\bar{x}, \Delta\lambda)$, i.e., $\bar{w} = D_{\bar{x}}\bar{P}(\bar{x}, \lambda)\Delta\bar{x} + D_\lambda\bar{P}(\bar{x}, \lambda)\Delta\lambda$, requires the computation of

$$w = D_x P(x, \lambda)\Delta x + D_\lambda P(x, \lambda)\Delta\lambda = y - \frac{< v_1, y >}{< v_1, z >} z, \tag{7}$$

where $x = R_k^{-1}(\bar{x})$, $\Delta x = DR_k^{-1}(\bar{x})\Delta\bar{x}$, $z = f(P(x, \lambda), \lambda)$, and $y$ is the solution of the first variational equation

$$\begin{aligned} \dot{x} &= f(x, \lambda), \\ \dot{y} &= D_x f(x, \lambda)y + D_\lambda f(x, \lambda)\Delta\lambda, \end{aligned}$$

with initial conditions

$$\begin{aligned} x(0) &= x, \\ y(0) &= \Delta x = DR_k^{-1}(\bar{x})\Delta\bar{x} \end{aligned}$$

at the time defined by the Poincaré map (see Fig. 2b). The second equality in (7) can be obtained by an application of the implicit function theorem (see [6] for the details, or [23], where, in addition, the derivatives of the Poincaré map up to, and including third order, are given). Therefore, each matrix product requires the time integration of a system of $2n$ equations.

## 4. Continuation of invariant 2-tori of ODEs (invariant curves of maps)

A map $G$ is defined now, such that its fixed points are approximations to points on invariant tori. We assume that the intersection of the 2D-tori with

the selected Poincaré section is a smooth invariant curve. The map we are going to define can be seen as a *synthesized return map* of the curve to the vicinity of an initial point.

Let $P : \mathcal{V} \subset \Sigma_1 \to \Sigma_1$ be the Poincaré map defined, as before, on a hyperplane $\Sigma_1$, and $\Sigma_2$ another hyperplane given by $< v_2, x - x_2 >= 0$. Let us suppose it is transversal to $\Sigma_1$ and to the invariant 2-tori we are interested to find (see Fig. 3). We want to obtain the intersection of the invariant curve, and $\Sigma_2$. Suppose $x$ is an initial point on the intersection $\Sigma_1 \cap \Sigma_2$, and consider the ball of radius $\varepsilon$ centered at $x$. A time integration with initial condition $x$ is started to find the first $q+1$ powers of the Poincaré map, $P^{k_j}(x, \lambda)$ with $j = 1, \cdots, q+1$ and, say, $k_1 < k_2 < \ldots k_{q+1}$, which fall inside the ball ($q = 3$ in Fig. 3). Then we consider the polynomial which interpolates these points, and its intersection with $\Sigma_2$. This process defines a map $G$ from $\mathcal{U} \subset \Sigma_1 \cap \Sigma_2$ onto itself. Its fixed points are the approximations we are looking for. As in the case of the Poincaré map $P$, the map $G(x, \lambda) : \mathcal{U} \subset \mathbb{R}^{n-2} \times \mathbb{R} \to \mathbb{R}^{n-2}$ can be parameterized as follows. Suppose that $j_i$ is the index of the largest component of $v_i$ for $i = 1, 2$, and that $R : \Sigma_1 \cap \Sigma_2 \to \mathbb{R}^{n-2}$ is the orthogonal projection from $\Sigma_1 \cap \Sigma_2$ onto the subspace $\{x_{j_1} = 0, x_{j_2} = 0\}$. In practice, the map $R$ drops the components $j_1$ and $j_2$ from a point in $\Sigma_1 \cap \Sigma_2$, and its inverse fills these components such that the resulting point is in $\Sigma_1 \cap \Sigma_2$. Therefore $G$ is defined as

$$G(\bar{x}, \lambda) = R \sum_{j=1}^{q+1} l_j(0) P^{k_j}(R^{-1}(\bar{x}), \lambda), \tag{8}$$

where the $l_j(\mu)$ are the Lagrange polynomials of degree $q$ at the points

$$\mu_j =< v_2, P^{k_j}(R^{-1}(\bar{x}), \lambda) - x_2 >, \quad j = 1, \cdots, q+1, \tag{9}$$

$\mu$ being the projection of the points inside the ball onto a line normal to $\Sigma_2$ (see Fig. 3). The action of the Jacobian of $G$ on a vector $(\bar{v}, \mu)$ reduces to the case of the differential of the Poincaré map, which we have already seen. By differentiating (8) it can be seen that

$$DG(\bar{x}, \lambda)(\bar{v}, \mu) = R \sum_{i=1}^{q+1} \left[ l_i(0) DP^{k_i}(x, \lambda) + P^{k_i}(x, \lambda) \sum_{j=1}^{q+1} \partial_{\mu_{k_j}} l_i(0) v_2^T DP^{k_j}(x, \lambda) \right] (D_{\bar{x}} R^{-1}(\bar{x}) \bar{v}, \mu), \tag{10}$$

with $x = R^{-1}(\bar{x})$.

The basic idea of the construction presented here is that if the invariant curve of the Poincaré map is analytical and reducible (that is, the linearized map can be reduced to constant coefficients with a suitable change of variables), then the interpolating functions in (8) give better and better estimates of the synthesized return map when $\varepsilon$ decreases, and (10) gives a correct approximation of the differential of the return map. Of course, in practice, these theoretical improved estimates can be affected by increasing round-off errors.

The reducibility assumption is essential in the sense that, otherwise, the differential of the return map is not well defined in general. Anyway, if the

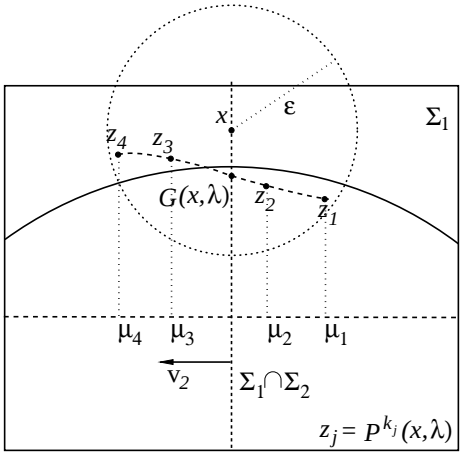


Figure 3: Map defining an approximation of a point on the torus. The solid line represents the invariant curve in $\Sigma_1$, and its intersection with $\Sigma_1 \cap \Sigma_2$ is the point we want to approximate. Note that the points $\mu_j$ on the direction of $v_2$ need not to appear ordered as illustrated in the figure.

"lack of reducibility" is small and the rotation number of the invariant curve satisfies some Diophantine Condition, still a formal process aiming at obtaining reducibility can be used. The process will not be convergent, but will produce good approximations to a reducible system. The interested reader can check theoretical results and applications in [24] and [25]. Far away from reducibility one can not even be sure of the existence of the invariant torus. The lack of reducibility can also produce "non-typical" behaviour, as can be shown by the evolution of Lyapunov exponents (see, e.g. [3]).

A fixed point of the map $G(\bar{x}, \lambda)$ corresponds, via the map $x = R^{-1}(\bar{x})$, to an approximation to a point of the invariant 2-tori in $\Sigma_1 \cap \Sigma_2$. The map $G$ depends on the radius $\varepsilon$ and the powers $k_j$. Several strategies can be adopted to ensure that the interpolation has the same number of points at both sides of $\Sigma_2$, or to discard a point $P^{k_j}(R^{-1}(\bar{x}))$ inside the ball if it is too close to a previous one. To make the process adaptive, $\varepsilon$ must be changed during the continuation. The method we have employed in the example shown below uses an initial $\varepsilon$, and, once a torus has been found, it is set to a fraction of the diameter of the invariant curve passing through the fixed point of $G$. The continuation process will always start from an initial known condition. The initial diameter, and $\varepsilon$ can be estimated, from a time integration, by computing enough Poincaré sections to complete a turn on the invariant curve. If the continuation starts close to a Neimark-Sacker bifurcation, the diameter grows proportionally to $\sqrt{\lambda - \lambda_c}$, $\lambda_c$ being the value of the parameter at the bifurcation. During the computation of the tori the diameter can be estimated by calculating the distances from the initial $x$ to the powers of the Poincaré map obtained during the computation of $G$.

There are two reasons that prevent this process to be considered as a "typical" continuation. The first is that the function $G$ varies during the continuation process because the rotation number of the invariant curve changes, and so do the powers which define $G$. So we do not follow the fixed points of a single map,

but those of a family of maps which change with the parameter $\lambda$. The second is that the objects found during the continuation will have different nature. The invariant curve will correspond to a quasi-periodic regime if the rotation number satisfies some Diophantine conditions. This occurs for a Cantor subset of the parameter interval. Otherwise the invariant curve can be the union of the unstable invariant manifolds of the saddle periodic points of the Poincaré map restricted to the invariant curve in case the rotation number is rational. These manifold meet at the periodic points of node type. For non-rational and non-Diophantine numbers a remnant of the invariant curve can subsist as a Cantor set. As mechanisms describing the destruction of the invariant curve in the rational case, one can mention the fact that the manifolds of the saddle points do not enter the nodal ones (and then they typically give rise to strange attractors) and the fact that the nodal points become focus and, later on, they have a period doubling. We can refer to [26] for theoretical study and numerical examples of these and other possibilities.

## 5. Thermal convection in binary fluid mixtures

The thermal convection of a binary mixture, filling a two-dimensional rectangular domain $\Omega$ heated from below is studied. The problem is governed by the mass, momentum, energy, and one of the concentrations (the denser in what follows) equations [27]. To write them in non-dimensional form the height of the domain $h$, the thermal diffusion time $h^2/\kappa$, $\kappa$ being the thermal diffusivity, the temperature difference between the top and bottom sides $\Delta T$, and $\bar{C}(\bar{C}-1)D'\Delta T/D$, $\bar{C}$ being the volume-average concentration, $D>0$ the mass diffusion coefficient, and $D'$ the thermal diffusion coefficient, are taken to scale the problem. In non-dimensional units $\Omega=[0,\Gamma]\times[0,1]$, $\Gamma$ being the width $l$ to the height ratio, and $x$ and $y$ are the horizontal and vertical coordinates, respectively.

The basic conductive and linearly stratified state, which is a solution of the equations for any value of the parameters, is given by zero velocity $\mathbf{v}_b=0$, and non-dimensional linear profiles for the temperature $T_b=T_b(0)-y$, and the concentration $C_b=C_b(0)-y$. The values $T_b(0)$ and $C_b(0)$ are related constants because of the boundary conditions defined below.

The Boussinesq approximation of the equations for the perturbation $(\mathbf{v},\Theta,\Sigma)$, of the basic state $(\mathbf{v}_b,T_b,C_b)$, are

$$
\begin{aligned}
&\partial_t\mathbf{v}+(\mathbf{v}\cdot\nabla)\mathbf{v}=-\nabla\pi+\sigma\nabla^2\mathbf{v}+\sigma Ra(\Theta+S\Sigma)\hat{e}_y,\\
&\partial_t\Theta+(\mathbf{v}\cdot\nabla)\Theta=\nabla^2\Theta+v_y,\\
&\partial_t\Sigma+(\mathbf{v}\cdot\nabla)\Sigma=L(\nabla^2\Sigma-\nabla^2\Theta)+v_y,\\
&\nabla\cdot\mathbf{v}=0,
\end{aligned}
$$

where $\mathbf{v}=(v_x,v_y)$. The term $-L\nabla^2\Theta$, which appears in the third equation is the Soret term, and is responsible for the generation of flows of matter caused by temperature gradients. The reciprocal term (Dufour term) in the $\Theta$ equation which would give rise to heat flows caused by concentration gradients has been neglected because for liquids it is much less important than for gases.

The problem depends on the aspect ratio, the non-dimensional Rayleigh, Prandtl and Lewis numbers, and the separation ratio, defined as

$$\Gamma = \frac{l}{h}, \quad Ra = \frac{\gamma\alpha\Delta T h^3}{\kappa\nu}, \quad \sigma = \frac{\nu}{\kappa}, \quad L = \frac{D}{\kappa}, \quad S = \frac{\bar{C}(1-\bar{C})\beta D'}{\alpha D}$$

respectively. In the definitions of the parameters $\nu$ means the kinematic viscosity, and $\alpha$ and $\beta$ (taken positive) are the thermal and solutal expansion coefficients respectively. The Prandtl and Lewis numbers represent ratios of time scales involving the viscous, thermal and solutal diffusions, but the separation ratio is also related with the temperature and concentration gradients of the flow. As usual, the Rayleigh number represents the ratio between the gravitational potential energy released when the light fluid rises (and the heavier falls), and the viscous and thermal dissipation of energy. The physical parameters must be evaluated at $\bar{C}$, $\bar{T}$, and $\bar{\rho}$.

In the continuation experiments we fix $\Gamma = 4$, $\sigma = 0.6$, $L = 0.03$ and $S = -0.1$. The last three values correspond to a mixture of two isotopes of Helium in liquid state. According with the definition of $S$, $D' < 0$, and initially the concentration gradient is stabilizing in opposition to the destabilizing temperature gradient. Then, if $S$ is below a negative critical value, as it is in the test problem, the primary bifurcation from the basic state is a Hopf bifurcation, and, when convection sets in, the denser component tends to migrate towards the hotter region. The Rayleigh number, which is proportional to the difference between the bottom and top temperatures, will be the control parameter.

The boundary conditions taken are non-slip for the velocity field ($\mathbf{v} = 0$ on $\partial\Omega$), constant temperatures at the top and bottom sides, insulating lateral sides, and non-porous boundaries.

The above equations are rewritten in terms of a stream-function, $\psi$, i.e., $\mathbf{v} = (-\partial_y\psi, \partial_x\psi)$, and an auxiliary function $\eta = \Sigma - \Theta$. They are

$$\begin{aligned}
&\partial_t\nabla^2\psi + J(\psi, \nabla^2\psi) = \sigma\nabla^4\psi + \sigma Ra\left[(S+1)\partial_x\Theta + S\partial_x\eta\right], \\
&\partial_t\Theta + J(\psi, \Theta) = \nabla^2\Theta + \partial_x\psi, \\
&\partial_t\eta + J(\psi, \eta) = L\nabla^2\eta - \nabla^2\Theta,
\end{aligned}$$

with $J(f, g) = \partial_x f\partial_y g - \partial_y f\partial_x g$, and the boundary conditions become

$$\begin{aligned}
\psi = \partial_n\psi = \partial_n\eta = 0 \quad &\text{on} \quad \partial\Omega, \\
\Theta = 0 \quad &\text{on} \quad y = 0, 1, \\
\partial_x\Theta = 0 \quad &\text{on} \quad x = 0, \Gamma.
\end{aligned}$$

In this way the incompressibility condition is identically fulfilled, the boundary conditions for $\Theta$ and $\Sigma$ decouple, and the number of unknowns is reduced.

The group of symmetries of this system is $\mathbb{Z}_2 \times \mathbb{Z}_2$ generated by the reflections $R_x$ and $R_y$, with respect to the vertical, and horizontal mid-planes, i.e., changing $x$ by $\Gamma - x$ and the sign of $\psi$, or changing $y$ by $1 - y$ and the sign of all three functions, leaves the system invariant. These symmetries give rise to pitchfork bifurcations of fixed points, periodic orbits, and also of invariant tori, at which any of the two symmetries can be broken.

To obtain the numerical solutions, the functions $\psi$, $\Theta$, and $\eta$ are approximated by a pseudo-spectral method. Collocation on a mesh of $n_x \times n_y = 64 \times 16$

Gauss-Lobatto points has been used in all the calculations shown. This gives a total dimension $n = 3072$. This mesh is enough to have a good accuracy in the interval of $Ra$ considered because the solutions are quite smooth. Finer resolutions with $n_x \times n_y = 96 \times 24$ ($n = 6912$), and $n_x \times n_y = 128 \times 32$ ($n = 12288$) have also been used to check some of the results. Details on the accuracy of the calculations are given in Section 7. The stiff system of ODEs obtained after the spatial discretization can be written as $B\dot{u} = Lu + N(u)$, where the vector $u = (\psi_{ij}, \Theta_{ij}, \eta_{ij})$ contains the values of $\psi$, $\Theta$ and $\eta$ at the mesh of collocation points. The operators $L$ and $N$ represent the linear and non-linear terms in the equations. They are integrated by using fixed-time-step sixth-order BDF-extrapolation formulas

$$\frac{1}{\Delta t} B \left( \gamma_0 u^{n+1} - \sum_{i=0}^{k-1} \alpha_i u^{n-i} \right) = \sum_{i=0}^{k-1} \beta_i N(u^{n-i}) + L u^{n+1},$$

where the superscripts indicate the time step. The coefficients, up to sixth order, are given in [14]. The initial points required to start the time integration are obtained by a fully implicit BDF method. The subroutine DLSODPK of the ODEPACK package [28] has been used.

The size of the fixed time-step is limited by stability reasons. It has been selected as large as possible. For the mesh $n_x \times n_y = 64 \times 16$, $\Delta t = 10^{-3}$ has been used. Finer meshes up to $n_x \times n_y = 128 \times 32$ require halving it. In the interval of $Ra$ considered, the period of the periodic orbits, on the branch from which the tori bifurcate, ranges from 1.4 to 4 non-dimensional time units, approximately. Therefore, it takes $\mathcal{O}(1000)$ time steps to complete a period.

To compute the Poincaré map, the function defining the hyperplane $\Sigma_1$, $< v_1, x - x_1 >$, is evaluated during the integration. When there is a change of sign in the right direction, the intersection is computed by interpolating the curve using the available previous points required by the multistep time integrator, and substituting the interpolation polynomial $x_{interp}(t)$ into $< v_1, x_{interp}(t) - x_1 >= 0$. This gives a scalar polynomial equation, which is solved by a secant method. This determines the arrival time (the period in the case of a periodic orbit), and the intersection.

## 6. Invariant tori

Fig. 4(a) shows a part of the bifurcation diagram for the thermal convection problem, which contains the main branch of periodic orbits, and the bifurcated branches of tori. The Euclidean norm of the solution is plotted versus $Ra$. The points shown correspond to the intersection of the solution with the hyperplane $\Sigma_1$ if it is a periodic orbit, and with $\Sigma_1 \cap \Sigma_2$ if it is a torus. Solid and dashed lines mean stable and unstable solutions, respectively.

Before describing the diagram, it is important to notice that the sequence of bifurcations leading to invariant tori takes place in a short interval of the parameter $Ra$ compared with the classical Rayleigh-Bénard problem. Moreover, due to the narrow region in which the periodic orbits are stable, the first regime which would be observed in a laboratory experiment in a short box, by slowly increasing Ra, would probably be quasi-periodic.

The horizontal line in Fig. 4(a) corresponds to the basic state which loses stability at a subcritical Hopf bifurcation, where a branch of periodic orbits

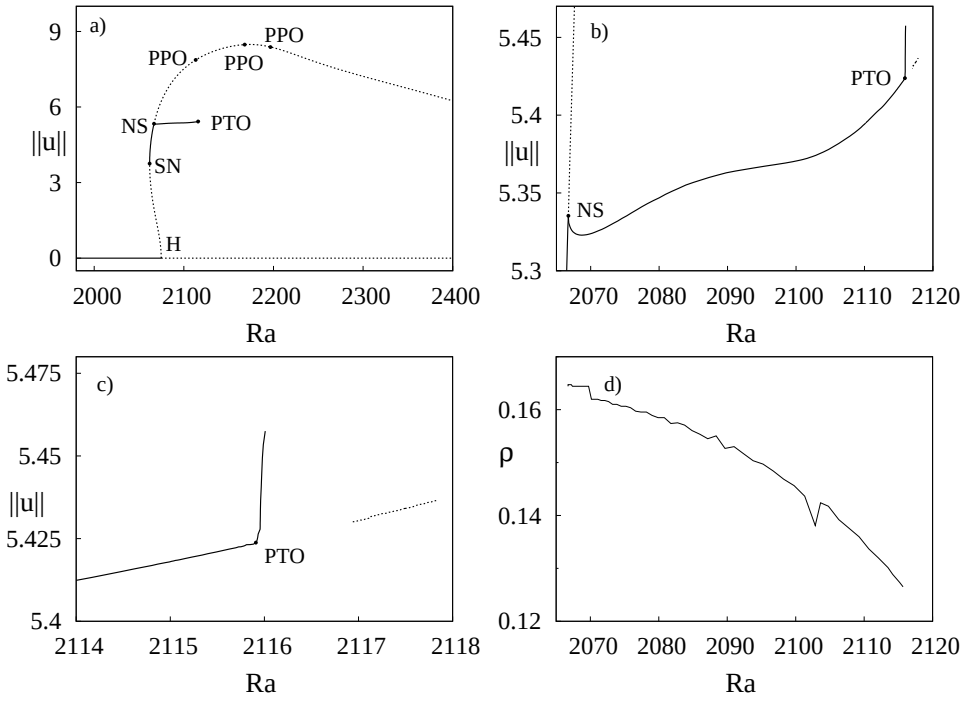


Figure 4: a) Bifurcation diagram. The horizontal line corresponds to the basic state. It loses stability at a Hopf bifurcation (H). The periodic orbits become stable at a saddle-node (SN) and again unstable at a Neimark-Sacker (NS) bifurcations. Three pitchfork bifurcations of periodic orbits (PPO) are found along the branch of periodic orbits, and a pitchfork bifurcations of tori (PTO) at almost the end of the branch of tori. b) Detail of the branch of invariant tori. c) Detail of the pitchfork bifurcation of invariant tori. Only one of the two stable branches is shown. d) Estimation of the rotation number along the main branch of tori.

emerges. The periodicity of the solutions consists in reversals of the direction of rotation of the vortices which fill the box, each half period. The latter becomes stable after a saddle-node, and, very near, there is a Neimark-Sacker bifurcation giving rise to the branch of invariant tori we have computed as test. The tangency of the curves of periodic orbits and invariant tori at the Neimark-Sacker bifurcation in Fig. 4(b) is just an effect of the scalings of the figure, which disappears as the figure is zoomed. After the Neimark-Sacker, the branch of periodic orbits has secondary pitchfork bifurcations, marked in Fig. 4(a) with three more dots. A detail of the branch of tori is given in Fig. 4(b). It starts at $Ra \approx 2066.74$, and it is stable up to the pitchfork bifurcation at $Ra \approx 2115.92$, which can be seen in more detail in Fig. 4(c). Only one of the two stable branches of tori after the pitchfork has been computed and shown in the diagrams. This stable branch has been continued until $Ra \approx 2116.02$. At that point the continuation method could not find additional tori, for a narrow range until $Ra \approx 2116.15$. At that value the arclength step becomes too small for the method to progress efficiently due to the closeness to a 1/8 resonance, which is also very close to the previous pitchfork bifurcation. At $Ra \approx 2116.15$ a torus was found by forward integration and a subsequent passage through the 1/8 resonance at $2116.18 < Ra < 2116.20$ could be studied using continuation

again. The results concerning that passage will be shown in Section 6.3

We have been able to compute a small portion of the unstable branch after the pitchfork bifurcation at $Ra \approx 2115.92$ (see Fig. 4(c)). It was started with an initial condition obtained by a continuation of the stable branch with a large arclength step, which allowed to cross to the other side of the pitchfork. The gap in the unstable branch is also due to the difficulty in computing the tori, with a reasonable arclength step, so close to the pitchfork bifurcation, and to the 1/8 resonance, which is also present on the unstable branch. Fig. 7 shows the transition, obtained by time evolution, from a torus of the unstable branch at $Ra = 2117.4954$ to another of the stable branch, symmetric to that shown in Fig. 5. From now on, the values of $Ra$ given with four decimal figures correspond to solutions obtained with the pseudo-arclength continuation code.

The rotation number $\rho$ along the main branch of tori, shown in Fig. 4(d), decreases by increasing $Ra$, starting near but below 1/6. The estimations of $\rho$ are computed by a method which provides lower and upper bounds. We refer to the Appendix for details on the method and error estimates.

During the continuation, to compute the rotation number we use only the points corresponding to the powers required to construct the map (8), which are selected by the algorithm. This is the reason why the curve looks irregular. If more precision is required, as is the case in Section 6.3, longer time integrations are needed. In Fig. 4(d) the mean of the two bounds is plotted. There is a 1/7-resonance interval at approximately $2102.79 < Ra < 2102.80$. It can be recognized in Fig. 4(d) by a peak in the curve. Close to this resonance the powers used in (8) are 7, 14 and 21, and, as they are small, the estimation is worse than for nearby points at which higher powers are selected.

*6.1. A period-doubling route to chaos*

After the 1/8-phase-locking interval and again on the stable branches which appear at the pitchfork, there are two period doubling bifurcations at $Ra \approx 2118.40$ and $Ra \approx 2118.55$, and, finally, a breakdown of the tori at $Ra \approx 2118.60$. The corresponding orbits have been found by forward integration. This sequence can be seen in Fig. 5, on one of the stable branches. We have checked that the same sequence of events is obtained along the conjugated branch. In all the figures representing an invariant curve or a strange attractor, $\eta$ at $(x, y) = (3.89, 0.45)$ versus $\eta$ at $(x, y) = (3.14, 0.45)$ is shown. The rotation numbers of the tori of Fig. 5 indicate that near $Ra = 2117.5$ there is a 1/9-resonance interval. For $Ra = 2118$ $\rho = 0.104164$, very close but below the 5/48-resonance. We have checked that if the time integration is prolonged, the iterates tend to fill the curve. Furthermore, for $2118.40 < Ra < 2118.55$, $0.1007 > \rho > 0.0998$, i.e., $\rho \approx 1/10$, predicting the presence of a new resonance.

For the values $Ra = 2118.6, Ra = 2118.7$ the rotation number is not defined, in contrast with the case of existence of an invariant curve. The upper and lower bounds provided by the method differ by a significant amount. But one can estimate a kind of "average rotation number" $\bar{\rho}$. The idea is to consider some of the variables of the problem at a given $(x, y)$ point, and look for the number of relative maxima (or minima) after a large number of Poincaré iterates. The ratio of maxima to iterates provides an estimate of $\bar{\rho}$. In this way for $Ra = 2118.6$ one obtains $\bar{\rho} \approx 0.09936$ and for $Ra = 2118.7$ the value $\bar{\rho} \approx 0.09819$. These values have little changes when the function of the point is changed. This suggests that the strange attractors observed in the simulations are strongly related to the

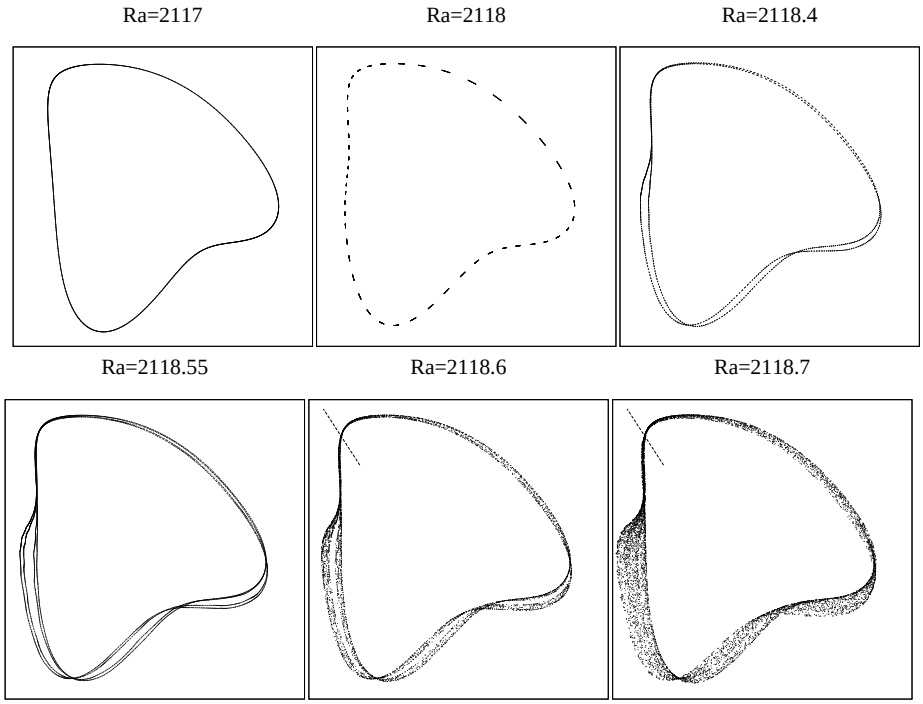


Figure 5: Breakdown of the tori on one of the stable branches after the pitchfork bifurcation of tori. The small segments in the plots for $Ra = 2118.6$ and $Ra = 2118.7$ indicate the common section used to show the transversal structure of the strange attractors in Fig 6. In all the plots the domain shown is $[-0.28, -0.06] \times [-0.08, 0.14]$, corresponding to the values of the function $\eta$ at the points $(3.89, 0.45)$ (horizontal axis), and $(3.14, 0.45)$ (vertical axis) of $\Omega$.

manifolds of unstable invariant curves having rotation numbers close to these values.

Fig. 6 shows the transversal structure of the strange attractors for $Ra = 2118.6$ and $Ra = 2118.7$. The points in a thin angular sector of half a degree of width, with vertex at the barycenter of the set of all the computed points, are selected. The position of the section is shown in Fig. 5. It was chosen to contain as much points as possible and to clearly display the shape of the attractor. The points selected are plotted using two new variables, which correspond to the values of the stream-function, $\psi$, at $(x, y) = (3.62, 0.45)$ and $(x, y) = (3.14, 0.45)$. Although the number of points in the figure is not very large, the structures are reminiscent of a quasi-periodic Hénon-like attractor [29]. Notice that for $Ra = 2118.6$ it has two separated pieces, which have joined for $Ra = 2118.7$. See [30] for the geometric explanation of this "fusion" of attractors. In the same way that the Hénon attractor sits on the closure of the unstable manifold of a hyperbolic fixed point, it seems reasonable to expect that the present attractors sit on the closure of the unstable manifolds of unstable invariant curves, as pointed out before.

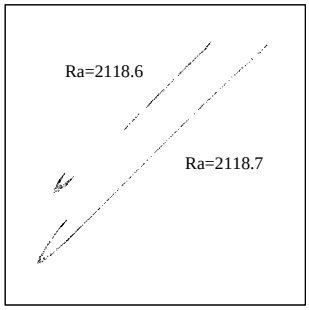


Figure 6: Detail of the transversal structure of the chaotic attractors at $Ra = 2118.6$, and $Ra = 2118.7$. The dots corresponding to $Ra = 2118.6$ have been displaced vertically 0.003 units, to make the figure clearer. The domain shown is $[-0.153, -0.143] \times [-0.442, -0.424]$, corresponding to the values of the function $\psi$ at the points $(3.62, 0.45)$ (horizontal axis), and $(3.14, 0.45)$ (vertical axis) of $\Omega$.

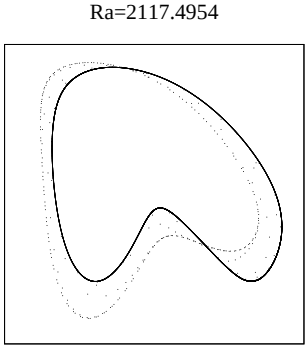


Figure 7: Evolution from an unstable to a stable torus at 2117.4954. The domain shown is $[-0.28, -0.025] \times [-0.08, 0.14]$, corresponding to the values of the function $\eta$ at the points $(3.89, 0.45)$ (horizontal axis), and $(2.61, 0.45)$ (vertical axis) of $\Omega$.

### *6.2. Symmetries of the tori*

To check which are the symmetries of the tori, and which are broken at the pitchfork bifurcation we proceed as in [31]. The distance function $d(t) = ||\mathcal{T}u(0) - u(t)||$ has been computed for the transformations $\mathcal{T} = I$, $R_x$, $R_y$ and $R_xR_y$ indicated respectively with the symbols $+$, $\circ$, $\times$ and $\square$ in Fig. 8. A symbol is plotted only when $d(t)$ becomes less than 0.04. The norms, $||u(0)||$, of the four initial conditions, are 5.40, 5.42, 5.43, and 5.74, respectively. The first plot at $Ra = 2110.5948$ indicates that, before the pitchfork bifurcation, the tori are invariant, as sets, with respect to all the symmetries of the group. Notice that all the symbols peak near zero for some long enough time. The second plot at $Ra = 2115.9241$ seems to be still before the bifurcation, because all the transformations peak at the same level. However, the third plot, at $Ra = 2115.9596$ shows that the symmetries $R_x$ and $R_xR_y$ have been broken. The symbols corresponding to these symmetries can still be seen in the upper part of the plot, because the torus is very near the bifurcation. These two figures give the best determination we have of the location of the bifurcation

point. At $Ra = 2116$ the torus is away enough from the bifurcation, so that only the symbols corresponding to $I$ and $R_y$ can be seen.

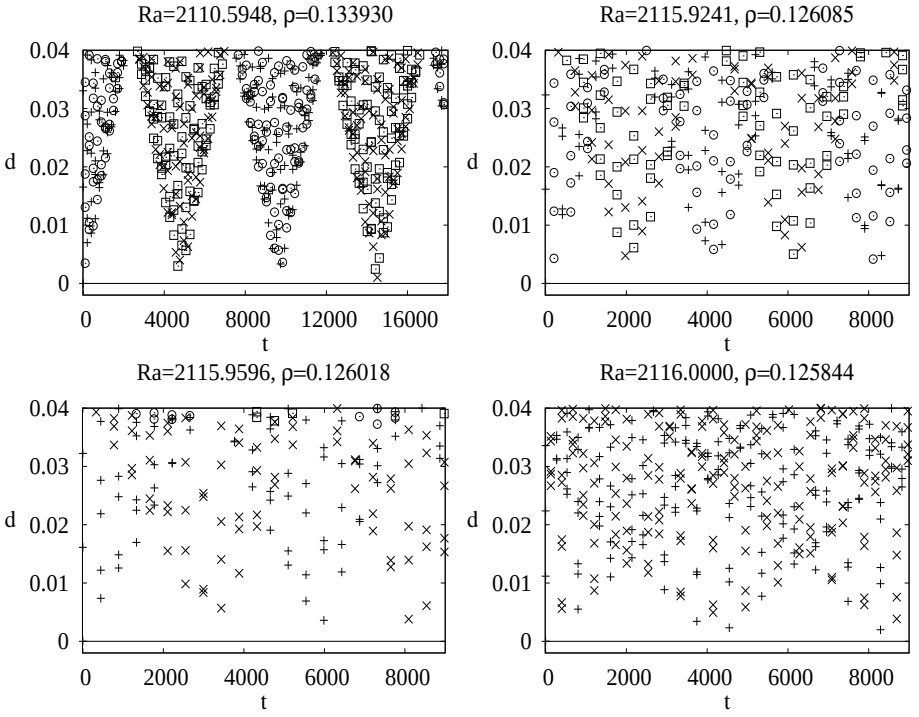


Figure 8: Distances of the transformation, $\mathcal{T}$, of an initial condition on a torus to the trajectory with that initial condition. The symbols $+$, $\circ$ $\times$ and $\square$ correspond, respectively, to $\mathcal{T} = I$, $R_x$, $R_y$ and $R_xR_y$. The average return times (the times between two consecutive sections of the Poincaré map) is between 1.77 and 1.80 for the four cases in this figure.

The position of the symbols in Fig. 8 can be explained as follows. Starting from u(0), the trajectory passes close to u(0) after a certain time $T'$ which, for instance, is near 10000 for $Ra = 2110.5948$. The time it takes to pass close to $R_yu(0)$ is approximately $T'/2$, i.e., $R_yu(0) \approx u(T'/2)$. This is inherited from the periodic orbits from which the tori bifurcate, which are symmetric periodic orbits, i.e., $R_yu(0) = u(T/2)$ , $T$ being the period ( $T$ is close to 1.5 at the Neimark-Sacker bifurcation). The same relation holds between $R_xu(0)$ and $R_yR_xu(T'/2)$ (bear in mind that $R_y$ and $R_x$ commute). Moreover the distance in time between $u(0)$ and $R_xu(0)$ is small compared to $T'$. The change in the time it takes to pass again close to the initial condition $u(0)$, between the upper two plots in Fig. 8, is due to the variation of the rotation number. The first plot has a rotation number which is different from the rest, which are very similar to each other.

To check that the two stable branches, born at the pitchfork bifurcation of tori, are conjugated by the symmetry broken at bifurcation, $d'(t) = ||\mathcal{T}u_1(0) - u_2(t)||$ has been computed as a function of time, $u_1$ and $u_2$ being two trajectories one on each branch, and both at $Ra = 2118.2$ (see again [31]). The transformations and the symbols used to designate them are the same as before. Fig. 9

shows the two invariant curves, and Fig. 10 the distance $d'$ as a function of time. It indicates that the tori are conjugated by the transformations $R_x$ and $R_xR_y$. As the tori are invariant under $R_y$, if they are conjugated by $R_x$ they also are by $R_xR_y$.

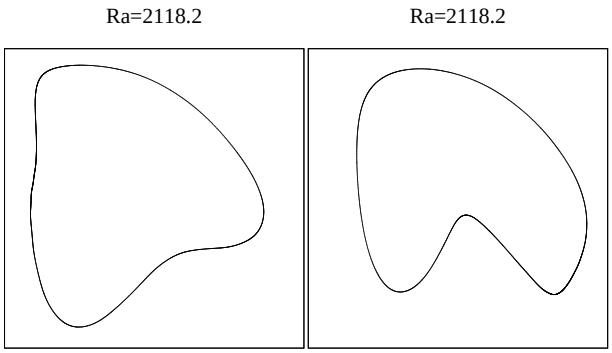


Figure 9: Two invariant tori conjugated by $R_x$ and $R_xR_y$. In the plots the domain shown is $[-0.28, -0.06] \times [-0.08, 0.14]$ (left), and $[-0.28, -0.025] \times [-0.08, 0.14]$ (right), corresponding to the values of the function $\eta$ at the points $(3.89, 0.45)$ (horizontal axis), and $(2.61, 0.45)$ (vertical axis) of $\Omega$.

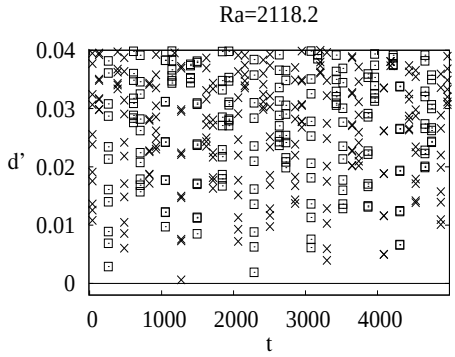


Figure 10: Distances of the transformation, $\mathcal{T}$, of an initial condition on a torus to a trajectory on another torus on the conjugated branch. The symbols $+$, $\circ$ $\times$ and $\square$ correspond, respectively, to $\mathcal{T} = I$, $R_x$, $R_y$ and $R_xR_y$.

### *6.3. Inside the $\rho = 1/8$ Arnold's tongue*

Figure 11(a) shows the dependence of the rotation number, $\rho$, with the parameter $Ra$ close to the $1/8$ resonance, with the characteristic parabolic profile close to the boundaries of the resonance interval. The dots indicate the computed $\rho$.

Figure 11(b) shows the continuation of the periodic orbits inside the Arnold's tongue. With $n_x \times n_y = 64 \times 16$ the interval of resonance is $2116.1760 \leq Ra \leq 2116.1981$ (see section 7). This diagram was computed with the continuation code of periodic orbits, with a single initial condition. Sixteen saddle-node bifurcations of periodic orbits were found at each limit of the interval, before

the continuation curve closes. Therefore, for each value of $Ra$ in the resonance region, the Poincaré map has 32 points of period 8, corresponding to the intersection of 4 periodic orbits with the Poincaré section. This non-generic situation is due to the remaining symmetry, $R_y$, of the tori. Fig. 11(c) shows the 32 intersections for $Ra = 2116.1872$. The symbols $+$, and $\times$ correspond to stable periodic points, and $*$ and $\square$ to the unstable.

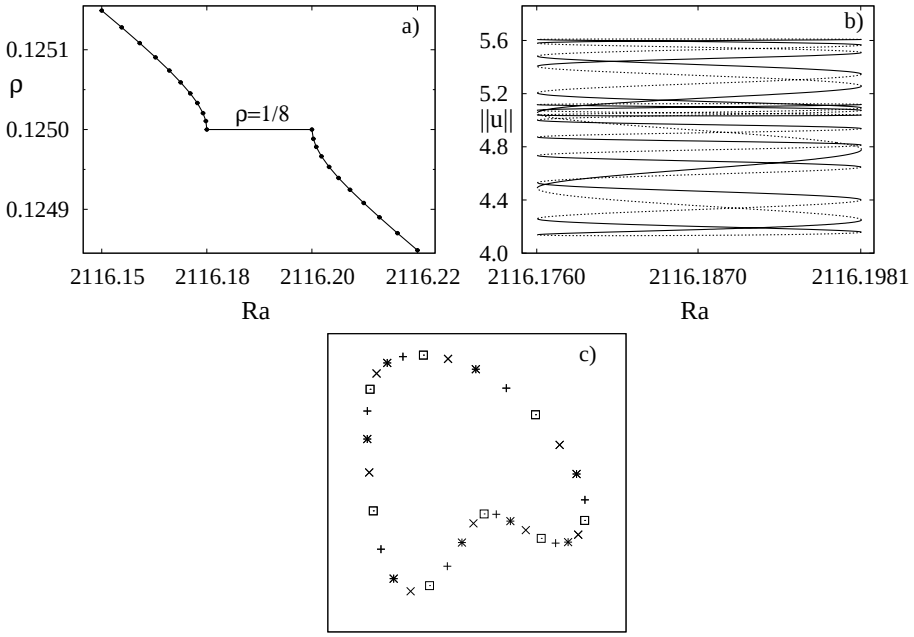


Figure 11: a) Rotation number as a function of $Ra$ close to the resonance interval of $\rho = 1/8$. b) Branches of periodic orbits inside this resonance interval. c) Periodic points of the Poincaré map for $Ra = 2116.1872$ in the middle of the resonance interval. The domain shown is $[-0.28, -0.025] \times [-0.08, 0.14]$ corresponding to the value of the function $\eta$ at the points $(3.89, 0.45)$ (horizontal axis) and $(2.61, 0.45)$ (vertical axis) of $\Omega$. The symbols $+$ and $\times$ correspond to the two stable periodic orbits, and $*$ and $\square$ to the two unstable.

## 7. Efficiency and accuracy

All the computations shown were performed on a Intel Core 2 Quad machine at 2.4 GHz, which has four cores (CPUs), although no parallelism was implemented. The time integrations were performed by using BDF-extrapolation formulas of order six. The Newton's iterations to solve (3) were stopped when the Euclidean norm of the difference between two iterates was below $tol_1$, and the Euclidean norm of the function on the last iterate was below $tol_2$. We took $tol_1$ and $tol_2$ in the interval $[10^{-9}, 10^{-7}]$, depending on the branch being computed. Most of the main stable branch could be computed with $tol_1 = tol_2 = 10^{-9}$. The computation of the unstable branch required to use tolerances in the upper limit. The reason is that, when the tori are unstable, any perturbation of the initial condition grows exponentially with time. This makes less accurate the computation of the map (8), and therefore the tolerances must be increased.

The linear systems (4) were solved using GMRES, which was stopped when the residual was below $tol_2/10$. The relation between this residual, and $tol_1$ or $tol_2$ must be obtained by experimentation to keep the quadratic convergence of Newton's method. Our particular selection proved to be enough for this purpose. Reducing the residual of the linear systems to these values took no more than eight iterations of GMRES. This fast convergence is easy to explain. GMRES converges very quickly if the spectrum of the matrix is highly clustered around a single point in the complex plane. In our case the linear operator $D_xG$ is strongly contractive, because $G$ is a linear combination of high powers of the Poincaré map (see eq. (8)). Then $I-D_xG$ has all the spectrum clustered around $+1$ if the torus is stable. In our computations we used second or third degree interpolation in the definition of $G$ ($q = 2$ or 3), and a value of the diameter of the ball to select the interpolated points of $0.05D \leq \varepsilon \leq 0.1D$, $D$ being the estimation of the diameter of the torus (see Fig. (3)). Then, the powers $k_j$ in (8) were at least 7. If the torus is unstable, with an unstable manifold of dimension three, as in the unstable case we have computed, $D_xG$ has only a one-dimensional unstable direction which is quickly resolved by GMRES (see the convergence analysis in the original paper on GMRES [13], or for the case of periodic orbits in [14]).

To construct the map $G$, $q+1$ points are required in the ball of radius $\varepsilon$. Depending on the rotation number, the degree of interpolation $q$, and the size of $\varepsilon$, it can take more or less powers to obtain them. No limit is put, in principle, to the power. If they become too large, it indicates that $\varepsilon$ is very small or that $\rho$ is close to some rational $p/q$, with $q$ not too large, but not so close that just powers of the form $P^q$ could be used for the algorithm. Then $\varepsilon$ must be increased. The maximal power arrived to 200 in some computations with $q = 3$, a small $\varepsilon$, low tolerances and close to a resonance. It must be stressed that, if $\varepsilon$ is large and $q$ small, there is no sense in asking low tolerances in Newton's method. The solution obtained will still be a poor approximation to a point on the torus.

| Mesh | Hopf bif. | S-N bif. | N-S bif. | Mult. at the NS bif. |
|---|---|---|---|---|
| 64x16 | 2074.7564 | 2061.8518 | 2066.7371 | 0.520158±0.854069i |
| 96x24 | 2074.7557 | 2061.8567 | 2066.7494 | 0.520133±0.854085i |

Table 1: Comparison between the first two meshes, of the position of the Hopf, Saddle-Node, and Neimark-Sacker bifurcations along the curve of periodic orbits, and of the multiplier which crosses the unit circle at the Neimark-Sacker bifurcation.

| Mesh | Lower bound | Upper bound |
|---|---|---|
| 64x16 | 0.12500000 | 0.12500000 |
| 96x24 | 0.12508352 | 0.12508354 |
| 128x32 | 0.12508475 | 0.12508498 |

Table 2: Comparison, among the three meshes, of the bounds for the rotation number at $Ra = 2116.1870$.

Several tests were performed to establish the accuracy of the calculations, by using meshes with $n_x \times n_y = 96 \times 24$, and $n_x \times n_y = 128 \times 32$. They required computing some continuations of periodic orbits close to the bifurcation points,

| Mesh | Lower limit | Upper limit | Width |
|---|---|---|---|
| 64x16 | 2116.1760 | 2116.1981 | 0.0221 |
| 96x24 | 2116.1965 | 2116.2189 | 0.0224 |

Table 3: Comparison, between the first two meshes, of the limits of the 1/8-resonance region, and its width.

and in the resonance region of Fig. 11. In particular, the computation shown in Fig. 11b was repeated with $n_x \times n_y = 96 \times 24$.

Table 1 shows the position of the Hopf, Saddle-Node, and Neimark-Sacker bifurcations along the curve of periodic orbits, and the multiplier which crosses the unit circle at the Neimark-Sacker bifurcation for $n_x \times n_y = 64 \times 16$, and $n_x \times n_y = 96 \times 24$. The maximal relative error in the bifurcation points is $6 \times 10^{-6}$, and four exact significant digits are obtained for the multiplier.

To determine the sensitivity of the position of the interval in the Arnold's tongue with the resolution, we have compute the rotation number of a point ($Ra = 2116.1870$) belonging to the 1/8-resonance interval with the coarser mesh, and the bounds of the tongue with $n_x \times n_y = 96 \times 24$. Table 2 shows the dependence of $\rho$ with the size of the mesh. With the resolution employed ($n_x \times n_y = 64 \times 16$) a relative error of $7 \times 10^{-4}$ is obtained in the determination of $\rho$. Furthermore, as can be seen in table 3, this small difference shifts the position of the interval by at least $\Delta Ra = 0.0205$, and its width is predicted with a relative error of 0.013. So, a higher spatial resolution is required to stabilize the value of its limits. In any case, the qualitative behaviour of the system does not change, and, for the purpose of this paper, the estimated errors for the coarser mesh used are sufficiently small as to allow to compute with $n_x \times n_y = 64 \times 16$ points.

Notice that the computations involve solving initial value problems (IVP), and systems of equations by Newton-Krylov methods. The time integrator should solve properly the IVP. If, say, a second order multistep method is used for the time integrator, the initialization process must be accurate enough to avoid adding first order perturbations that will ruin the rest of the integration. As the integrations are carried out on long time intervals in the computation of the map $G$, and also for the computation of rotation numbers, it is convenient to use high order time-steppers to reduce the accumulated errors. On the other hand, if one is interested in fine details, like the position of resonance regions, it is clear from the above comparisons, that a high spatial resolution is needed.

The computation of the main branch of invariant tori, from the Neimark-Sacker to the pitchfork bifurcation, took approximately six days of CPU. One hundred points on the curve were obtained, giving an average of 1.4 hours per computed tori. The solution of the linear systems and the calculation of the dominant multipliers, to study the stability of the solutions, involve only basic algebraic operations such as orthogonalization of basis of vectors. The expensive part is the evaluation of matrix products, which involve time integrations. More than 98% of the computing time was spent in the time integrations of the initial system, and the variational equations. Therefore, any attempt to accelerate the calculation should focus on this part of the process.

## 8. Conclusions

We have shown that it is possible to perform numerical continuations of stable or weakly unstable invariant tori in large-scale systems. If they are strongly unstable the map $G$ cannot be computed before the time evolution has escaped away from the tori.

The method we have used has the advantage that the size of the linear problems to be solved is the original size of the system. Moreover, the Jacobian of the map considered is very well conditioned, so that Krylov-based iterative solvers like GMRES converge very quickly. Other methods, for instance those based in expanding the invariant curve in Fourier series, increase the size prohibitively for high-dimensional systems, unless the tori are smooth enough so that the number of terms in the expansion is small. In addition, in this method the spectra of the linear operators spread in circles [11] complicating the solution by iterative methods, as happens when multiple shooting is used to compute periodic orbits [32]. A possible cure to this latter problem, for smooth tori, could be to use a preconditioner for the linear systems, based on the knowledge of the invariant subspaces corresponding to the leading eigenvalues. This technique has been used in [32], where optimal speedups were obtained when using parallelism.

On the other hand, the method we have used can be computationally expensive if high powers of the Poincaré maps are involved. For instance for very small rotation numbers, or for rotation numbers rather close to a rational with small denominator. In the first case, averaging techniques could be employed to eliminate the fast frequency, and, in the latter, one could switch to the continuation of periodic orbits inside the resonance regions as has been done in Section 6.3. It will be also computationally expensive when looking for higher-dimensional tori, but this seems to be an intrinsic problem, independent of the method.

## 9. Acknowledgments

The research of J.S. and M.N. has been supported by Spain MEC-DGI and Catalonia GENCAT under projects FIS2007-64993 and 2009-SGR-67, respectively. The research of C.S. has been supported by projects MTM2006-05849/Consolider and 2009-SGR-67.

## Appendix: On the computation of $\rho$ and error estimates

To estimate $\rho$ for an invariant curve of a discrete map we use a method which provides lower and upper bounds. If it is known that the curve really exists the method can be simplified but, as it will be presented here, it allows to detect if the curve is no longer existing. For a preliminary presentation see, e.g., [33]. The method is purely topological: it only depends on the relative position of successive iterates along the invariant curve. This is relevant when the density of iterates along the curve changes in a significative way.

We assume that there exists some projection of the invariant curve on a 2D-plane such that the domain bounded by the projected curve is starred with respect to some interior point $O$. This allows to define an angle associated to each iterate. Otherwise to introduce a "natural" angle requires more work. Let $\alpha_i \in [0, 2\pi)$, $i = 1, \cdots, M$, where $M$ is the number of available points on the

curve, be the angles between a reference line $\ell$ passing through $O$ and the initial point $p_1$ and that joining each point $p_i$ on the invariant curve with $O$. We take an orientation such that the angles between successive iterates are increasing on the lift.

Let $n(i)$ be the number of turns along the invariant curve given to arrive to the point $p_i$, computed according to the rule: $\alpha_1 = 0$, and if $\alpha_{i+1} > \alpha_i$ then $n(i+1) = n(i)$, otherwise $n(i+1) = n(i)+1$, for $i = 1, \cdots, M-1$. The angles are then sorted so that $\alpha_{i_1} < \alpha_{i_2} < \alpha_{i_3} < \cdots$. Then the two bounds are initialized as $\rho_{min} = 0$ and $\rho_{max} = 1$, and updated following the rule: let $q_j = (n(i_{j+1}) - n(i_j))/(i_{j+1} - i_j)$, then, if $i_{j+1} > i_j$, $\rho_{min} = \max\{\rho_{min}, q_j\}$, and otherwise $\rho_{max} = \min\{\rho_{max}, q_j\}$, for $j = 1, \cdots, M-1$. If during the computation it happens $\rho_{max} < \rho_{min}$ this gives evidence that the invariant curve does not exist. At the end lower and upper bounds of $\rho$ are obtained, and its difference gives an indication on the "quality" of the determination. Of course $\rho \in [0, 1)$ and, by changing orientation if necessary, one can assume in fact that $\rho \in [0, 1/2]$.

To produce estimates of the error we rely on some elementary concepts. We recall that a real number $\alpha$, that we assume in $[0, 1)$, satisfies a $DC(c, \tau)$ Diophantine Condition if $|\alpha - p/q| \geq cq^{-\tau}$ with $c > 0, \tau \geq 2$ for all $p/q \in \mathbb{Q}$, $p/q \geq 0$. If $\alpha$ satisfies $DC(c, \tau_1)$ it also satisfies $DC(c, \tau_2)$ for $\tau_2 > \tau_1$ and, eventually, one can increase $c$. As it is well known, the measure of the set of the $DC(c, \tau)$ numbers in $[0, 1]$ is zero if $\tau = 2$ and of the form $1 - \mathcal{O}(c)$ for $\tau > 2$ when $c \to 0$. We shall also assume that the constant $c$ is optimal in the sense that can not be replaced by any smaller quantity.

Given $\alpha$ one can compute the quotients of the continued fraction expansion $\alpha : [q_1, q_2, q_3, \ldots]$ and recover the approximants of the form $0/1, 1/q_1, \ldots, N_{k-1}/D_{k-1}, N_k/D_k, \ldots$. We recall the elementary formulas for next approximant $N_{k+1}/D_{k+1}$ given by $N_{k+1} = N_k q_{k+1} + N_{k-1}, D_{k+1} = D_k q_{k+1} + D_{k-1} = D_k(q_{k+1} + \sigma)$ for some $\sigma \in (0, 1)$ and the relation $|N_k D_{k-1} - N_{k-1} D_k| = 1$. Furthermore, if we introduce $\hat{q}_{k+1} = q_{k+1} + \dfrac{1}{q_{k+2}+} \dfrac{1}{q_{k+3}+} \ldots$ which obviously satisfies $1 \leq q_{k+1} < \hat{q}_{k+1} < q_{k+1} + 1$, then

$$\alpha = \frac{N_k \hat{q}_{k+1} + N_{k-1}}{D_k \hat{q}_{k+1} + D_{k-1}},$$

from which it follows the following expression for the error $|\alpha - N_k/D_k| = D_k^{-2}/(q_{k+1} + \theta)$ for some $\theta \in (0, 2)$. Let us denote the error as $\varepsilon$. Hence

$$\varepsilon = \frac{(q_{k+1} + \sigma)^2}{D_{k+1}^2 (q_{k+1} + \theta)} \approx \frac{q_{k+1}}{D_{k+1}^2}$$

if $q_{k+1}$ is large.

Now we want to estimate which is the last approximant $N_k/D_k$ that one can obtain after doing the iterate number $M$. We should have $D_k \leq M - 1 < D_{k+1}$. Hence the problem is how big can be $D_{k+1}$, compared to $D_k$, for a $DC(c, \tau)$ number or, in other words, to bound $q_{k+1}$. Combining the first equality for $\varepsilon$ above with the $DC(c, \tau)$ condition we have $q_{k+1} + \theta \leq \left(D_{k+1}/(q_{k+1} + \sigma)\right)^{\tau-2}/c$. Neglecting again the small values $\theta, \sigma$ in front of $q_{k+1}$ in the worst cases, we have $q_{k+1}^{\tau-1} \leq D_{k+1}^{\tau-2}/c$. Therefore $\varepsilon < c^{-1/(\tau-1)} D_{k+1}^{(\tau-2)/(\tau-1)} D_{k+1}^{-2}$. Summarizing we have proved the following

**Proposition 1.** *If the rotation number $\rho$ satisfies a $DC(c,\tau)$ Diophantine Condition, then the algorithm presented above to compute estimates of $\rho$ has an error $\varepsilon$ satisfying $\varepsilon < c^{-1/(\tau-1)} M^{-\tau/(\tau-1)}$.*

**Remark 1.** *We can consider what happens in some cases.*

- *For quadratic irrational number or, more generally, for constant type numbers, the one which have all quotients $q_k$ bounded, one has $\tau = 2$ and, hence, $\varepsilon = \mathcal{O}(M^{-2})$.*
- *In a case with $\tau = 5/2$ then $\varepsilon = \mathcal{O}(M^{-5/3})$. The role of $c$ appears in the $\mathcal{O}$ symbol.*
- *For given $c, \tau$ we can look for the measure of the numbers in $(0,1)$ satisfying a $DC(c,\tau)$ condition. A finite computation, taking into account the overlap of intervals to be suppressed from $(0,1)$, estimates on the remainder and suitable extrapolation provide the following examples:*

$$\mu(DC(0.01, 2.1)) = 0.870\ldots, \quad \mu(DC(0.10, 2.5)) = 0.63114\ldots,$$

$$\mu(DC(0.05, 3.0)) = 0.8635141\ldots,$$

*where $\mu$ denotes the Lebesgue measure. Note that in the last two cases the intervals suppressed around 0 and 1 are responsible for most of the measure lost with respect to 1.*

## References


[1] À. Jorba, J. Villanueva, On the Normal Behaviour of partially Elliptic Lower Dimensional Tori, Nonlinearity 10 (1997) 783–822.

[2] C. Simó, Effective computations in celestial mechanics and astrodynamics, in: V. Rumyantsev, A. Karapetyan (Eds.), Modern Methods of Analytical Mechanics and their Applications (CISM Courses and Lectures 387), Springer, 1998, pp. 55–102.

[3] L. van Veen, The quasi-periodic doubling cascade in the transition to weak turbulence, Physica D 210 (2005) 249–261.

[4] H. Broer, C. Simó, R. Vitolo, The Hopf-saddle-node bifurcation for fixed points of 3D-diffeomorphisms: the Arnol'd resonance web, Bull. Belg. Math. Soc. Simon Stevin 15 (2008) 769–787.

[5] G. Gómez, J. Llibre, R. Martínez, C. Simó, Station keeping of libration point orbits, chapter 7, Tech. rep., ESA Final Technical Report, 689 +xii pp., reprinted as Dynamics and Mission Design Near Libration Points. Volume 1: Fundamentals: The Case of Collinear Libration Points, World Sci. Pub., Monograph Ser. Math. Vol. 2, Singapore, xx+442 pp., ISBN 981–02–4285–9, 2001. (1985).

[6] C. Simó, Analytical and numerical computation of invariant manifolds, in: C. Benest, C. Froeschlé (Eds.), Modern methods in celestial mechanics, Editions Frontières, 1990, pp. 285–330,
Also at `http://www.maia.ub.es/dsg/2004/`.

[7] C. Simó, Averaging under fast quasiperiodic forcing, in: I. Seimenis (Ed.), Integrability and chaos in Hamiltonian Systems, Plenum Pub. Co., 1994, pp. 13–34.

[8] L. O. Chua, A. Ushida, Algorithms for computing almost periodic steady-state response of nonlinear systems to multiple input frequencies, IEEE Tran. Circuits Sys. 28 (10) (1981) 953–971.

[9] C. Kaas-Petersen, Computation, continuation, and bifurcation of torus solutions for dissipative maps and ordinary differential equations, Physica D 25 (1–3) (1987) 288–306.

[10] C. Simó, Effective Computations in Hamiltonian Dynamics, in: Cent ans après les Méthodes Nouvelles de H. Poincaré, Société Mathématique de France, 1996, pp. 1–23.

[11] À. Jorba, Numerical computation of the normal behaviour of invariant curves on $n$-dimensional maps, Nonlinearity 14 (2001) 943–976.

[12] F. Schilder, H. M. Osinga, W. Vogt, Continuation of Quasi-periodic Invariant Tori, SIAM J. Appl. Dinam. Systems 4 (3) (2005) 459–488. doi:10.1137/040611240.
URL `http://link.aip.org/link/?SJA/4/459/1`

[13] Y. Saad, M. H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comput. 7 (1986) 865–869.

[14] J. Sánchez, M. Net, B. García-Archilla, C. Simó, Newton-Krylov continuation of periodic orbits for Navier-Stokes flows, J. Comput. Phys. 201 (1) (2004) 13–33.

[15] M. C. Cross, Structure of nonlinear travelling waves in finite geometries, Phys. Rev. A 38 (7) (1988) 3593–3600.

[16] C. Surko, P. Kolodner, Oscillatory travelling-wave convection in a finite container, Phys. Rev. Lett. 58 (20) (1987) 2055–2058.

[17] P. Kolodner, Repeated transients of weakly nonlinear traveling-wave convection, Phys. Rev. E 47 (2) (1993) 1038–1048.

[18] L. Ning, Y. Harada, H. Yahata, Formation process of the travelling-wave state with a defect in binary fluid convection, Progress of Theoretical Physics 98 (3) (1997) 551–566.

[19] O. Batiste, E. Knobloch, I. Mercader, M. Net, Simulations of oscillatory binary fluid convection in large aspect ratio containers, Phys. Rev. E 65 (2002) 016303–1–016303–19.

[20] J. Turner, Multicomponent convection, Ann. Rev. Fluid Mech 17 (1985) 11–40.

[21] H. B. Keller, Numerical solution of bifurcation and nonlinear eigenvalue problems, in: P. H. Rabinowitz (Ed.), Applications of Bifurcation Theory, Academic Press, 1977, pp. 359–384.

[22] J. Sánchez, M. Net, J. Vega, Amplitude equations close to a triple-(+1) bifurcation point of $D_4$-symmetric periodic orbits in $O(2)$-equivariant systems, Discrete and Continuous Dynamical Systems-Series B 6 (6) (2006) 1357–1380.

[23] Y. A. Kuznetsov, H. G. E. Meijer, L. van Veen, The fold-flip bifurcation, Int. J. Bifurcation and Chaos 14 (7) (2004) 2253–2282.

[24] À. Jorba, C. Simó, On the Reducibility of Linear Differential Equations with Quasiperiodic Coefficients, J. Diff. Eq. 98 (1992) 111–124.

[25] À. Jorba, R. Ramírez-Ros, J. Villanueva, Effective Reducibility of Quasiperiodic Linear Equations close to Constant Coefficients, SIAM J. Math. Anal. 28 (1997) 178–188.

[26] H. Broer, C. Simó, J. Tatjer, Towards global models near homoclinic tangencies of dissipative diffeomorphisms, Nonlinearity 11 (1998) 667–770.

[27] S. R. de Groot, P. Mazur, Non-equilibrium thermodynamics, Dover Publications, 1962.

[28] A. C. Hindmarsh, ODEPACK, a systematized collection of ODE solvers, in: R. S. S. et al. (Ed.), Scientific Computing, North-Holland, Amsterdam, 1983, pp. 55–364.

[29] H. Broer, C. Simó, R. Vitolo, Quasi-periodic Hénon-like attractors in the Lorenz-84 climate model with seasonal forcing, in: F. Dumortier, H. Broer, J. Mawhin, A. Vanderbauwhede, S. V. Lunel (Eds.), Proceedings of the Equadiff-2003 conference, World Scientific, 2005, pp. 601–606.

[30] C. Simó, On the Hénon-Pomeau attractor, J. of Stat. Physics 21 (4) (1979) 465–494.

[31] M. Net, J. Sánchez, Symmetric periodic orbits and global dynamics of quasi-periodic flows in an O(2) equivariant system: Two-dimensional thermal convection, Int. J. Bifurcation and Chaos 15 (12) (2005) 3953–3972.

[32] J. Sánchez, M. Net, On the multiple shooting continuation of periodic orbits by Newton-Krylov methods, Int. J. Bifurcation and Chaos 20 (1) (2010) to appear.

[33] C. Simó, Global dynamics and fast indicators, in: H. W. Broer, B. Krauskopf, G. Vegter (Eds.), Global Analysis of Dynamical Systems, IOP Publishing, Bristol, 2001, pp. 373–390.