Deformation of invariant tori under perturbation

OPEN ACCESS
IOP Publishing
Plasma Physics and Controlled Fusion
Plasma Phys. Control. Fusion **67** (2025) 075006 (21pp)
https://doi.org/10.1088/1361-6587/ade007

Deformation of invariant tori under perturbation

Wenyin Wei[1,2,3], **Jiankun Hua**[3,4], **Alexander Knieps**[3], **Shaocheng Liu**[5,*] and **Yunfeng Liang**[1,3,*]

[1] Institute of Plasma Physics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, People's Republic of China
[2] University of Science and Technology of China, Hefei 230026, People's Republic of China
[3] Forschungszentrum Jülich GmbH, Institute of Fusion Energy and Nuclear Waste Management—Plasma Physics, 52425 Jülich, Germany
[4] International Joint Research Laboratory of Magnetic Confinement Fusion and Plasma Physics, State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China
[5] College of Physics, Donghua University, Shanghai 201620, People's Republic of China

E-mail: scliu@dhu.edu.cn and y.liang@fz-juelich.de

Received 29 January 2025, revised 19 March 2025
Accepted for publication 3 June 2025
Published 23 June 2025

Abstract

Invariant tori (flux surfaces) reside in the ordered regions of phase space of a dynamical system and represent the well-confined region of a magnetic confinement fusion (MCF) device, which may be threatened in non-axisymmetric cases such as tokamaks under resonant magnetic perturbation and stellarators. In MCF devices, the structure of nested closed flux surfaces governs radial transport and thus plays a critical role in confinement performance. Using the method of variation as a mathematical foundation (the vector field itself as a spatial function is considered as an argument of the geometry of these tori), this paper derives the formulae that describe

$$\text{how invariant tori (flux surfaces) } \mathcal{T} \text{ deform under perturbation } \Delta\mathcal{B},$$

which can calculate the deformation in tens of seconds for stellarator configuration optimisation by Julia programming language without delicate hardware acceleration and almost instantly for tokamaks.

Keywords: deformation of flux surfaces, invariant torus, functional perturbation theory, integrable system

* Authors to whom any correspondence should be addressed.

1. Introduction

Invariant tori are ordered structures within dynamical systems, in contrast to chaotic regions where long-term behaviour is unpredictable. Systems are commonly classified into two broad classes: *integrable* and *non-integrable*. For conservative systems like a time slice of magnetic field, integrability is equivalent to the following well-known condition:

> *The entire phase space is exactly the union of all invariant tori.*

To accurately measure the extent of integrability in a conservative system, one can assess the ratio of the phase space occupied by invariant tori to the total phase space. (In contrast, non-conservative systems, where the divergence is non-zero, may exhibit components that are not invariant tori but still possess integrable characteristics. For example, the map $\mathcal{P}(x) = x/e$ simply shows a converging tendency toward zero.)

In magnetic confinement fusion (MCF) devices, the nested closed flux surfaces, which act as invariant tori [1], are crucial for achieving optimal performance. These surfaces dictate the radial transport of charged particles but can be vulnerable to disturbances like the deliberately imposed resonant magnetic perturbation [2–4] (RMP), the plasma instability modes, the plasma beta β effect (as revealed by simulations of EMC3-EIRENE [5], HINT [6] and SPEC [7]), the complex current redistribution due to the plasma-wall interaction (PWI), and the plasma disruptions [8] due to (or leading to) the collapse of these surfaces.

The stellarator community has been continuously concerned about three-dimensional (3D) magnetic topology due to the direct realistic needs from the infinite-dimensional degrees of freedom space of design parameters and the intrinsic difficulty [9] in handling chaos. The shaping capability [10] of a stellarator determines its operational limit, which is a strong motivation for this paper to develop the functional perturbation theory (FPT) to address flux surface deformation under perturbation by considering the whole magnetic field as an argument of the geometry representation of flux surfaces.

> *Shaping is the primary means of control of a toroidal fusion plasma.*
>
> —Allen H. Boozer [11]

Functional is a familiar concept to stellarator designers, because the 3D shape of a stellarator coil is essentially a curve function, while the optimisation objective function dependent on coil shapes are naturally a function of functions [12, 13] (usually the coil shapes are discretised to facilitate optimisation). Robustness of a magnetic configuration [14] to the variations of pressure and bootstrap current can now be evaluated by FPT by computing the first-order flux surface deformation $\delta\boldsymbol{\chi}(\theta, r, \phi)$.

Sufficient numerical and experimental evidence emphasises the importance of investigating the behaviour of flux surfaces under perturbation, e.g. whether they are to become island chains, break down into debris or deform [15]. The extended-magnetohydrodynamic code M3D-C1 has been adapted to accommodate a non-axisymmetric computation domain and its simulation [16] of a reactor-scale quasisymmetric stellarator equilibrium shows a flux surface destruction tendency due to interchange-like pressure-driven instability near the plasma edge. The experimental evidence [17] [from Wendelstein 7-AS [18] and large helical device [19]] has shown that the chaotisation of the magnetic field structure can be a major subject in plasmas with higher-beta values, which reminds people of interests in stellarator optimisation: eliminate islands [20, 21], protect good flux surfaces [22, 23], and synergistic optimisation of plasma physics and coil engineering objectives [24, 25] and reserve a margin of coil adjustability to counteract detrimental effects from error fields [26] or plasma spontaneous chaotisation in 3D configurations [27].

The FPT developed in this paper provides a universal theory framework to tackle the flux surface deformation problem, without the need to derive formulae for a specific configuration such as negative triangularity [28]. The negative triangularity configuration is a possible scenario of tokamak reactor [29], which has a larger power handling area for larger major radius at the low-field side, and reported to have no or very weak ELM [30] and H-mode grade confinement integrated with L-mode edge [31, 32]. Experiments on machines such as ASDEX-U [33], DIII-D [34, 35], TCV [36, 37], *etc* have shown advantages of negative triangularity configuration. TCV is renowned for its strong shaping capabilities and a wide feasible scenario range from conventional to advanced scenarios [38, 39], but other machines can also exploit their operation limit of feasible configurations with the aid of FPT.

Preserving the confined volume enclosed by the last closed flux surface (LCFS) is important for economic reasons in fusion reactors, given the high costs associated with the vacuum vessel's volume. Additionally, the 'stickiness' [40, 41] of the LCFS may influence the flux of particles crossing it into the scrape-off layer, the 'width' (or more accurately, the scope) of which can be largely dependent on the degree of chaoticity [42, 43] when chaos arises, instead of being determined by the plasma diffusion according to Eich scaling [44].

Understanding *the deformation of invariant tori under perturbation* is crucial for predicting system behaviour and optimising design, thereby reducing the need for expensive or impractical real-world tests. Such understanding elucidates the strong correlation between changes in magnetic topology (global structure) and variations in the local field values, which is revealed by the formulae of the FPT presented in this paper. Notably, FPT does not require:

- The number of dimensions of the system to be three or
- The divergence of the field to vanish.

These two conditions are found to be unnecessary during derivation, which enables FPT to harvest a broad range of applications. Systems with higher dimensionality, such as the N-body interaction system with $6N$ dimensions in 3D space,

can also be investigated by this theory so astrophysicists and statistical physicists do not need to write and read another paper reminiscent of this one.

It is not yet well understood how frequently and how long two oppositely charged particles can entangle with each other. This is important because such a neutral particle cluster would, for a transient period, fail to be effectively confined by the magnetic field. Compared to a brief, isolated collision, such entanglement has a much prolonged duration if the two particles' velocities are close, and the trajectory in the $6N$-phase space (expressed in the centre-of-mass frame) behaves like a high-dimensional invariant torus. Such entanglement is a central topic of astrophysics because gravity is purely attractive. It is common for small-mass celestial objects to be captured by large ones, such as the moon by the earth. Electromagnetic interaction between charged particles is subject to the ambient electric and magnetic fields, which brings more degrees of freedom and challenges to understanding the entanglement between charged particles.

The destruction of invariant tori is a key subject in the domain of Kolmogorov–Arnold–Moser (KAM) theorem and that of MCF [45], closely tied to their deformation. As invariant tori break down [46], complex structures like island-around-island hierarchies [47, 48] and cantori [49, 50] with infinite gaps can emerge . This paper lays the groundwork that can be further utilised to explore these phenomena.

2. Basic definitions and notations

2.1. Invariant torus

This paper adopts the standard definition of an invariant torus. For a map $\mathcal{P}:\mathbb{R}^N\to\mathbb{R}^N$, if there exists a diffeomorphism $\boldsymbol{\varphi}:\mathbb{R}^N\supset\mathcal{T}^d\to\mathbb{T}^d$ (where $\mathbb{T}^d$ is the standard d-torus) such that the motion on $\mathbb{T}^d$ is uniformly linear but non-static—i.e.

$$\boldsymbol{\varphi}(\mathcal{P}(\boldsymbol{x}))-\boldsymbol{\varphi}(\boldsymbol{x})=\Delta\boldsymbol{\theta}\in\mathbb{R}^d\setminus\{\mathbf{0}\}\text{ remains constant}$$

—then $\mathcal{T}^d$ is termed a d-dimensional invariant torus (or invariant d-torus). For a continuous-time dynamical system, i.e. a flow, such an invariant torus can be defined similarly. It is merely the requirement on $\boldsymbol{\varphi}$ becomes that the conjugate motion on $\mathcal{T}^d$ has a constant non-vanishing angular velocity

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\varphi}(\boldsymbol{X}(\boldsymbol{x}_0,t))=\boldsymbol{\omega}\in\mathbb{R}^d\setminus\{\mathbf{0}\}.$$

Here, $\boldsymbol{\omega}$ and $\Delta\boldsymbol{\theta}$ are referred to as the *frequency vector* of $\mathcal{T}^d$, while the corresponding *period vector* is

$$\boldsymbol{m}:=[2\pi/\Delta\theta_1,\ldots,2\pi/\Delta\theta_d].$$

This convention aligns with the MCF community's practice, where field lines on a flux surface with a rotation transform $\iota=n/m=1/q$ (n and m being coprime) complete m toroidal turns before returning. For Poincaré mapping defined for one toroidal turn, $\Delta\theta$ is simply a scalar and $\Delta\theta\equiv 2\pi n/m \mod 2\pi$. Note that the value of $\Delta\theta$ is not unique, as any integer multiple of 2π can be added to it. For simplicity the $\Delta\theta\in[0,2\pi)$ is adopted in practical calculations.

An invariant torus with a commensurable rotation vector is termed $\boldsymbol{\omega}$-commensurable, otherwise $\boldsymbol{\omega}$-incommensurable (the notation $\boldsymbol{\omega}$ can be omitted for brevity if clear). An $\boldsymbol{\omega}$-commensurable torus can be further decomposed into a bunch of invariant tori of lower dimensionality until the $\boldsymbol{\omega}$ of the subtori are incommensurable, which implies that $\boldsymbol{\omega}$-commensurable tori are more fragile than those incommensurable ones under perturbation because the subtori individually deform against perturbation rather than together as an incommensurable one. The well-known Diophantine condition is not introduced because the generalised exponent of mapping $\mathcal{P}^{\boldsymbol{m}}$ (from an integer m to a real number vector $\boldsymbol{m}$) helps to work-around the small denominator problem.

There exists a widespread misunderstanding of KAM theorem: a commensurable torus is fragile, i.e. it would be destroyed by any finite perturbation. This statement is false because the commensurable torus can be preserved if the perturbation meets some condition or is dedicatedly designed. The incommensurable condition of a torus in KAM theorem is a sufficient condition to ensure the torus survives, not a sufficient and necessary condition (if and only if). A very easy-to-understand example comes from tokamaks: in an axisymmetric magnetic field, rational and irrational flux surfaces all can survive if the perturbation is also axisymmetric, which is a natural consequence of Poincaré-Bendixson theorem because field line tracing along an axisymmetric field degrades to a 2D system.

2.2. Parameterisation of invariant tori

Let $\boldsymbol{\chi}(\theta_1,\ldots,\theta_d)$ be a parameterisation of $\mathcal{T}^d$ such that mapping once corresponds to a constant increment of the angles θ_i on the torus. The choice of this notation $\boldsymbol{\chi}$ is to resemble $\boldsymbol{X}$, both representing the state vector, yet $\boldsymbol{X}$ has been commonly employed in the literature of dynamical systems to denote a trajectory, hence $\boldsymbol{\chi}$ serves as a suitable alternative, suggesting a 'curved' version of $\boldsymbol{X}$ to represent a torus. By definition, mapping once on a point on the torus shifts its angles by a constant vector increment $\Delta\boldsymbol{\theta}$, as expressed by

$$\mathcal{P}(\boldsymbol{\chi}(\theta_1,\ldots,\theta_d))=\boldsymbol{\chi}(\theta_1+\Delta\theta_1,\ldots,\theta_d+\Delta\theta_d).$$

One can define a vector $\boldsymbol{k}$ as the exponent of the map $\mathcal{P}$, firstly continuing $\mathcal{P}^k$ from $k\in\mathbb{Z}$ to $\mathbb{R}$, and then generalising to $\boldsymbol{k}\in\mathbb{R}^d$:

$$\mathcal{P}^{\boldsymbol{k}}(\boldsymbol{\chi}(\boldsymbol{\theta})):=\boldsymbol{\chi}(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}),\tag{1}$$

where $*$ denotes element-wise multiplication. Thus, $\mathcal{P}^{\boldsymbol{m}}$ is, by definition, a returning map on $\mathcal{T}^d$. When $k\in\mathbb{R}$ instead of $\mathbb{R}^d$, $\boldsymbol{k}*\Delta\boldsymbol{\theta}$ reduces to $k\Delta\boldsymbol{\theta}$. The motivation for generalising the power of mapping is to allow the returning map to be expressed by a power of mapping so that the returning map can be well-defined even on an $\boldsymbol{\omega}$-incommensurable

torus and expressed concisely as $\mathcal{P}^{\boldsymbol{m}}$. A concise notation for the returning map is essential since the full-period Jacobian of mapping can be then easily expressed as $\mathcal{D}\mathcal{P}^{\boldsymbol{m}}$, which is a matrix of significant importance for studying the nearby motion pattern around an invariant torus, as already shown by the well-known Greene's residue criterion for the planar case $N=2$. $\mathcal{D}$ denotes the Jacobian matrix, which collects all the partial derivative components in the initial condition $\boldsymbol{x}_0$.

For the case $d=N-1$, a sequence of nested invariant $(N-1)$-tori can be parameterised as $\boldsymbol{\chi}(\boldsymbol{\theta},r)$, where r serves as the torus label, often interpreted as the radial direction. For the case of an N-degree-of-freedom Hamiltonian system with a $2N$-dimensional phase space and up to N integral invariants (*a.k.a.* actions), the torus dimensionality $d\leqslant N$ and (at most) N actions can serve as torus labels (often denoted by $\{I_i\}$) to distinguish different tori. In this paper, the denotations are kept simple by writing the index/label as a scalar r (short for ***r***adial) to make it easier for readers to understand, yet it does not essentially restrict readers from employing the equations to the cases where two or more scalars are needed to label a torus.

Given the parameterisation $\boldsymbol{\chi}(\boldsymbol{\theta},r)$ for all existing tori, its inverse functions are naturally the angle and index distributions,

$$\boldsymbol{\theta}(\boldsymbol{x}):F\to\mathbb{T}^d \text{ and } r(\boldsymbol{x}):F\to\mathbb{R}.$$

2.3. The union of invariant tori can be fragmented

Due to the complicated nature of the long-term behaviour of a system involving chaos, the union of all invariant tori $\cup_r\mathcal{T}_r$ is usually not the full $\mathbb{R}^N$ space in general systems that are not guaranteed to be completely integrable. Similarly, the domain of r is not guaranteed to be the full $\mathbb{R}$ because an invariant torus corresponding to a specific r may not exist. Hence, one can collect the r labels of all existing tori to be a subset of $\mathbb{R}$, which may be ***f***ragmented or not, thereby denoted by F_r. The first letter of '***f***ragmented' is taken and uppercased to indicate the notation F_r represents a possibly fragmented set and the subscript r shows it is the domain of r.

The adjective fragmented is not a strict mathematical definition here (For a rigorous definition, readers are directed to appendix B), yet it can vividly convey the meaning that such a set is so irregular that it may be neither open nor closed, neither everywhere-Cantor-like nor composed by finite length intervals. The union of all the existing invariant tori $\cup_r\mathcal{T}_r$ is also not guaranteed to be regular enough that the usual derivative can be done, therefore this union is similarly denoted by $F\subseteq\mathbb{R}^N$. For MCF experiment shots, one can evaluate how good the confinement is by the ratio of the volume occupied by the primary flux surfaces (not those in island chains) to that enclosed by the vacuum vessel as shown below (FS subscripts are short for flux surfaces),

$$V_{\mathrm{FS}}=\text{measure}\left(\text{the union of all primary flux surfaces}\right)$$

$$Q_{\mathrm{FS}}=\frac{V_{\mathrm{FS}}}{\text{measure(the volume enclosed by vacuum vessel)}}.$$

2.4. Fragmentation requires a weak derivative

Much care should be put into the fragmented characteristic of the defining domain, which brings a challenge to differential calculus. The fragmented nature of the defining domains of $\boldsymbol{\chi}(\boldsymbol{\theta},r)$, $\boldsymbol{\theta}(\boldsymbol{x})$ and $r(\boldsymbol{x})$ for general systems that are not completely integrable leads us to define a weaker form of derivative. For instance, the partial derivative in r is redefined in a weaker sense to be

$$\frac{\partial}{\partial r}\boldsymbol{\chi}(\boldsymbol{\theta},r):=\lim_{\substack{r'\to r\\ r'\in F_r}}\frac{\boldsymbol{\chi}(\boldsymbol{\theta},r')-\boldsymbol{\chi}(\boldsymbol{\theta},r)}{r'-r},$$

in which the lim operation only needs to take r' in the defining domain $F_r\subseteq\mathbb{R}$. In other words, it only requires the finite difference ratio

$$\frac{\boldsymbol{\chi}(\boldsymbol{\theta},r')-\boldsymbol{\chi}(\boldsymbol{\theta},r)}{r'-r},$$

limits to an identical value for all sequences $(r_i')\subseteq F_r$ that approximate r, but not necessarily does so for all such sequences (r_i') in a neighbourhood of r. Such relaxation is common in KAM-relevant studies due to the fragmented nature of the union of invariant tori, yet a classical one of these relaxations given by Pöschel [51] may not be favoured by physicists because it requires a sequence of functions defined on broader domains to converge uniformly to the essential one defined here, which are probably deemed abundant by physicists.

The usual derivative that requires the function to be well-defined on a neighbourhood of r, such as an open interval $(r-\epsilon,r+\epsilon)$ in 1D cases, i.e.

$$\begin{aligned}\frac{\partial}{\partial r}\boldsymbol{\chi}(\boldsymbol{\theta},r)&:=\lim_{\substack{r'\to r\\ r'\in(r-\epsilon,r+\epsilon)}}\frac{\boldsymbol{\chi}(\boldsymbol{\theta},r')-\boldsymbol{\chi}(\boldsymbol{\theta},r)}{r'-r}\\ \left(\begin{smallmatrix}\text{or more familiarly}\\ \text{defined to be}\end{smallmatrix}\right)&:=\lim_{h\to0}\frac{\boldsymbol{\chi}(\boldsymbol{\theta},r+h)-\boldsymbol{\chi}(\boldsymbol{\theta},r)}{h}\end{aligned}$$

simply fails to work when the defining domain is fragmented. There may exist infinite gaps between r and another $r'\in F_r$ due to the destruction of invariant tori in between. When one torus is isolated from other tori for a finite distance, even the weak derivative has to be replaced by a finite difference with a sacrifice of exact equality. Such a situation is common for non-conservative systems but can also happen in conservative ones if a torus survives under perturbation while its neighbours all break down.

The definition of gradient operator ∇ is similarly relaxed when applied on $\boldsymbol{\theta}(\boldsymbol{x})$ and $r(\boldsymbol{x})$ so that they are still meaningful in a non-integrable system. Note that $\boldsymbol{\theta}(\boldsymbol{x})$ can be considered as $\cup_r \boldsymbol{\varphi}_r(\boldsymbol{x})$, the union of all the conjugate diffeomorphisms $\boldsymbol{\varphi}_r$ of all existing invariant tori. This is a union of functions.

The partial derivative $\partial_k \mathcal{P}^k$ can then be easily computed with the aid of the $\boldsymbol{\theta}$ grid because the derivative in the mapping times (i.e. its exponent) can be converted to partial derivatives of the parameterisation of the torus *w.r.t.* its angles:

$$\partial_k \mathcal{P}^k = \left.\frac{\partial \boldsymbol{\chi}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}+k\Delta\boldsymbol{\theta}} \cdot \Delta\boldsymbol{\theta} = \sum_{i=1}^{d} \Delta\theta_i \left.\frac{\partial \boldsymbol{\chi}}{\partial \theta_i}\right|_{\boldsymbol{\theta}+k\Delta\boldsymbol{\theta}}, \tag{2}$$

and for a vectorised $\boldsymbol{k}$,

$$\begin{aligned}\partial_{\boldsymbol{k}} \mathcal{P}^{\boldsymbol{k}} &= \left.\frac{\partial \boldsymbol{\chi}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}} \cdot \overbrace{\frac{\partial(\boldsymbol{k}*\Delta\boldsymbol{\theta})}{\partial \boldsymbol{k}}}^{=\mathrm{diag}(\Delta\boldsymbol{\theta})_{d\times d}} \\ &= \left.\begin{bmatrix} | & & | \\ \Delta\theta_1 \partial_{\theta_1}\boldsymbol{\chi} & \cdots & \Delta\theta_d \partial_{\theta_d}\boldsymbol{\chi} \\ | & & | \end{bmatrix}\right|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}}\end{aligned} \tag{3}$$

where the subscripts $\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}$ denote where to evaluate the functions and the subscript $d\times d$ is used to help clarify the dimension of the matrix. The vertical lines in the matrix are to help clarify they are column vectors and will be used throughout this paper along with horizontal lines for row vectors when column/row needs to be distinguished. The subscripts denoting where the term is evaluated or the shape of the matrix will also be used when necessary.

2.5. Functional derivative

As powerful mathematical tools introduced from functional analysis, partial and total functional derivatives [52] are denoted by

$$\delta/\delta\mathcal{B} \quad \text{and} \quad \mathrm{d}/\mathrm{d}\mathcal{B},$$

which become directional derivatives

$$\Delta\mathcal{B}\cdot\delta/\delta\mathcal{B} \quad \text{and} \quad \Delta\mathcal{B}\cdot\mathrm{d}/\mathrm{d}\mathcal{B}$$

when accompanied by a given perturbation. For brevity, the former one can be simply denoted δ if the system to be perturbed and the perturbation are clear. For example, $\delta\mathcal{P}$ and $\delta\boldsymbol{X}_{\mathrm{pol}}$ are short for the first variation of Poincaré map and the poloidal shift of a trajectory expressed in standard cylindrical coordinates, that is $(\Delta\mathcal{B}\cdot\delta/\delta\mathcal{B})\mathcal{P}$ and $(\Delta\mathcal{B}\cdot\delta/\delta\mathcal{B})\boldsymbol{X}_{\mathrm{pol}}$. The latter total directional derivative may not be abbreviated to just one letter 'd' because much confusion can arise, e.g. one may be puzzled by whether $\mathrm{d}\mathcal{P}$ is an infinitesimal differential or a total functional derivative.

The aforementioned $\boldsymbol{X}_{\mathrm{pol}}$ is short for $\boldsymbol{X}_{\mathrm{pol}}[\mathcal{B}](\boldsymbol{x}_{0,\mathrm{pol}},\phi_s,\phi_e)$, the representation for a trajectory that ***s***tarts at a point with (R,Z,ϕ)-coordinates $(\boldsymbol{x}_{0,\mathrm{pol}},\phi_s)$ and ***e***nds at ϕ_e azimuthal angle. Then the Poincaré map usually corresponds to one toroidal turn and $\delta\mathcal{P}$ is simply $\delta\boldsymbol{X}_{\mathrm{pol}}(\boldsymbol{x}_{0,\mathrm{pol}},\phi_s,\phi_e=\phi_s+2\pi)$ by definition. In the MCF community, if for each toroidal turn (2π change in ϕ) the poloidal angle increment $\Delta\theta$ is constant, that means the poloidal angle θ is that of PEST coordinates. One can also trace the trajectory in Cartesian coordinates, let $\delta\boldsymbol{X}(\boldsymbol{x}_0,t)$ progress along the trajectory. The formula describing the progression of the first variation, $\delta\boldsymbol{X}$, along a trajectory can be easily derived by the chain rule. With $\Delta\mathcal{B}\cdot\mathrm{d}/\mathrm{d}\mathcal{B}$ applied on both sides, the field line tracing equation for an autonomous flow

$$\frac{\partial}{\partial t}\boldsymbol{X}[\mathcal{B}](\boldsymbol{x}_0,t) = \boldsymbol{B}[\mathcal{B}](\boldsymbol{X}[\mathcal{B}](\boldsymbol{x}_0,t)) \tag{4}$$

converts to

$$\frac{\partial}{\partial t}\delta\boldsymbol{X}[\mathcal{B};\Delta\mathcal{B}](\boldsymbol{x}_0,t) = \nabla\boldsymbol{B}\cdot\delta\boldsymbol{X} + \delta\boldsymbol{B}, \tag{5a}$$

which is applicable for *arbitrary finite-dimensional* flows. The other similar formula describing progression of the poloidal variation, $\delta\boldsymbol{X}_{\mathrm{pol}}$, in standard 3D *cylindrical coordinates* is shown below

$$\begin{aligned}&\frac{\partial}{\partial\phi_e}\delta\boldsymbol{X}_{\mathrm{pol}}[\mathcal{B};\Delta\mathcal{B}](\boldsymbol{x}_{0,\mathrm{pol}},\phi_s,\phi_e) \\ &\quad= \frac{\partial(R\boldsymbol{B}_{\mathrm{pol}}/B_\phi)}{\partial(R,Z)}\cdot\delta\boldsymbol{X}_{\mathrm{pol}} + \delta\frac{R\boldsymbol{B}_{\mathrm{pol}}}{B_\phi},\end{aligned} \tag{5b}$$

where $\delta(R\boldsymbol{B}_{\mathrm{pol}}/B_\phi)$ is short for $(\Delta\mathcal{B}\cdot\delta/\delta\mathcal{B})(R\boldsymbol{B}_{\mathrm{pol}}/B_\phi)$, equal

$$\frac{R\delta\boldsymbol{B}_{\mathrm{pol}}}{B_\phi} - \frac{R\boldsymbol{B}_{\mathrm{pol}}}{B_\phi^2}\,\delta B_\phi,$$

by the product rule of differentiation.

3. Deduction and demonstration

Based on the above calculation of $\delta\mathcal{P}^m$, one can derive the shift of a hyperbolic cycle via the simple geometry analysis below:

$$\delta\boldsymbol{x}_{\mathrm{cyc}} = \delta\mathcal{P}^{\boldsymbol{m}} + \mathcal{D}\mathcal{P}^{\boldsymbol{m}}\cdot\delta\boldsymbol{x}_{\mathrm{cyc}}$$

of which LHS is the cycle starting point shift and RHS is the cycle ending point shift: the first term $\mathcal{D}\mathcal{P}^{\boldsymbol{m}}\cdot\delta\boldsymbol{x}_{\mathrm{cyc}}$ is propagated from the starting shift and the second term $\delta\mathcal{P}$ is the ending point shift due to perturbation. Thereby, $\delta\boldsymbol{x}_{\mathrm{cyc}}$ can now be solved for by

$$\delta\boldsymbol{x}_{\mathrm{cyc}} = -\left[\mathcal{D}\mathcal{P}^{\boldsymbol{m}} - \mathbf{I}\right]^{-1}\cdot\delta\mathcal{P}^{\boldsymbol{m}}. \tag{6}$$

However, this cannot be directly extended to the case of a periodic orbit on an invariant d-torus $\mathcal{T}^d$, which is the focus of this paper. The challenge lies in the fact that these cycles can be parabolic and then $\mathcal{D}\mathcal{P}^{\boldsymbol{m}} - \mathbf{I}$ possess eigenvalue(s) equal to zero, rendering it non-invertible. The underlying reason is that the tangent component of $\delta\boldsymbol{x}_{\mathrm{cyc}}$ along $\mathcal{T}^d$ can take arbitrary values (If a point moves in the tangent direction for an infinitesimal distance, it remains in the torus $\mathcal{T}^d$). To resolve this

indeterminacy, it is necessary to revisit the initial geometric analysis:

$$\delta \boldsymbol{x}_{\rm cyc} = \delta\mathcal{P}^{\boldsymbol{m}}\left(\boldsymbol{x}_{\rm cyc}\right) + \mathcal{D}\mathcal{P}^{\boldsymbol{m}}\left(\boldsymbol{x}_{\rm cyc}\right)\cdot\delta\boldsymbol{x}_{\rm cyc}, \tag{7}$$

and focus solely on the normal component perpendicular to $\mathcal{T}^d$ by replacing $\delta\boldsymbol{x}_{\rm cyc}$ with $\delta_\perp := \sum_{i=1}^{N-d}\hat{\boldsymbol{n}}_i\hat{\boldsymbol{n}}_i^{\rm T}\cdot\delta$, where $\{\hat{\boldsymbol{n}}_i\}_{i=1}^{N-d}$ is an orthonormal basis of the local normal space $N_p\mathcal{T}^d$. Thus, the equation becomes

$$\begin{aligned}\delta_\perp\boldsymbol{x}_{\rm cyc} &= \delta\mathcal{P}^{\boldsymbol{m}}\left(\boldsymbol{x}_{\rm cyc}\right) + \mathcal{D}\mathcal{P}^{\boldsymbol{m}}\left(\boldsymbol{x}_{\rm cyc}\right)\cdot\delta_\perp\boldsymbol{x}_{\rm cyc},\\ \left(\mathcal{D}\mathcal{P}^{\boldsymbol{m}}-\mathbf{I}\right)\cdot\delta_\perp\boldsymbol{x}_{\rm cyc} &= -\delta\mathcal{P}^{\boldsymbol{m}}.\end{aligned} \tag{8}$$

However, another issue arises: $\left(\mathcal{D}\mathcal{P}^{\boldsymbol{m}}-\mathbf{I}\right)\cdot\delta_\perp\boldsymbol{x}_{\rm cyc}$ can only have a tangent component to $\mathcal{T}^d$ if all the eigenvalues of $\mathcal{D}\mathcal{P}^{\boldsymbol{m}}$ are equal to one, which is typically the case for conservative systems like flux-preserving maps. Yet, the RHS of equation (8), $-\delta\mathcal{P}^{\boldsymbol{m}}$, depends on the perturbation and may include a normal component. There can be three possibilities:

1. $\delta\mathcal{P}^{\boldsymbol{m}}(\boldsymbol{x}_{\rm cyc})$ has only a tangent component to $\mathcal{T}^d$.
2. The eigenvalue(s) of $\mathcal{D}\mathcal{P}^{\boldsymbol{m}}$ corresponding to the eigenvectors not tangent to $\mathcal{T}^d$ are allowed to deviate from one. This is the case for non-conservative systems where the divergence is non-zero.
3. The invariant torus is destroyed, rendering $\delta_\perp\boldsymbol{x}_{\rm cyc}$ meaningless. In this case, the normal component of $-\delta\mathcal{P}^{\boldsymbol{m}}$ indicates to what extent the torus will be destroyed.

For conservative systems, it is the first case that allows the equation (8) to hold. The tangency of $\delta\mathcal{P}^{\boldsymbol{m}}(\boldsymbol{x}_{\rm cyc})$ can be immediately acquired by imposing the functional total derivative $\Delta\mathcal{P}\cdot{\rm d}/{\rm d}\mathcal{P}$ on both sides of the equation defining the returning map $\mathcal{P}^{\boldsymbol{m}}$,

$$\mathcal{P}^{\boldsymbol{m}[\mathcal{P}](\boldsymbol{x})}\left[\mathcal{P}\right](\boldsymbol{x}) := \boldsymbol{x}, \tag{9}$$

$$\underbrace{\partial_{\boldsymbol{m}}\mathcal{P}^{\boldsymbol{m}}\cdot\delta\boldsymbol{m}}_{=\sum_{i=1}^{d}\partial_{\theta_i}\boldsymbol{\chi}\,\Delta\theta_i\,\delta m_i} + \delta\mathcal{P}^{\boldsymbol{m}} = \mathbf{0}, \tag{10}$$

which clearly shows that $\delta\mathcal{P}^{\boldsymbol{m}}$ must be tangent to $\mathcal{T}^d$ if the returning map $\mathcal{P}^{\boldsymbol{m}}$ remains well-defined under perturbation. This is consistent with the classical magnetic spectrum analysis technique of RMP. The resonant radial component breaks the torus, while the non-resonant radial components deform the torus. A tangent perturbation does not change the torus shape but other properties like rotation transform.

For non-conservative systems, $\delta\boldsymbol{m}[\mathcal{P};\Delta\mathcal{P}](\boldsymbol{x})$ is not well-defined because each invariant torus must be separate from others for a finite distance, implying that $\boldsymbol{m}(\boldsymbol{x})$ is defined on a discrete subset of $\mathbb{R}^N$ (every torus is isolated and behaves as a sink or source in the normal direction).

Although the non-invertibility of $\mathcal{D}\mathcal{P}^{\boldsymbol{m}}$ is an issue for acquiring the solution of $\delta_\perp\boldsymbol{x}_{\rm cyc}$ directly, it can still be solved for by excluding the tangent component. One simply needs to solve for all normal components $\hat{\boldsymbol{n}}_i^{\rm T}\cdot\delta_\perp\boldsymbol{x}_{\rm cyc}$ (underlined below) by

$$\sum_{i=1}^{N-d}\left(\mathcal{D}\mathcal{P}^{\boldsymbol{m}}-\mathbf{I}\right)\cdot\hat{\boldsymbol{n}}_i\underline{\hat{\boldsymbol{n}}_i^{\rm T}\cdot\delta_\perp\boldsymbol{x}_{\rm cyc}} = -\delta\mathcal{P}^{\boldsymbol{m}}.$$

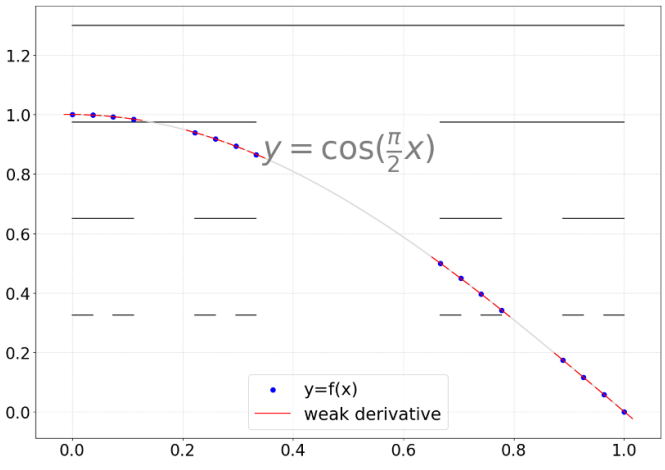

Figure 1. Illustration of weak differentiability: A function $y=\cos(\frac{\pi}{2}x)$ restricted to the classical Cantor set, whose usual derivative does not exist due to the totally disconnected nature of this defining domain, still has its weak derivative equal $f'(x) = -\frac{\pi}{2}\sin(\frac{\pi}{2}x)$.

Nonetheless, this is merely a single-cycle analysis. There are many cycles on an invariant torus to analyse. To consider an invariant torus as a whole, the $(\boldsymbol{\theta},r)$-based representation of $\boldsymbol{\chi}(\boldsymbol{\theta},r)$ will be analysed later.

In this paper, the standard map (see figure 2) is used as an example to demonstrate the formulae, with its parameter set to $p=0.975$ (the same as in figure 1 of [49]). Additionally, the magnetic field from the standard configuration of Wendelstein 7-X is also considered for demonstration. To determine the exponent m such that $\mathcal{P}^m$ is a returning map for points on an irrational invariant torus, map the initial point $\boldsymbol{x}_0$ for numerous times. Define an angle difference $\Delta\theta\in[0,2\pi)$ for once mapping, then $\boldsymbol{x}_i$ compared to $\boldsymbol{x}_0$ has an angle difference $i\Delta\theta$. Denote by n_i how many times the orbit crosses $\boldsymbol{x}_0$ counterclockwise until the i-th point, then one knows the angle increment $i\Delta\theta$ from $\boldsymbol{x}_0$ to $\boldsymbol{x}_i$ is between $2\pi n_i \leqslant i\Delta\theta \leqslant 2\pi(n_i+1)$. Notice $m = 2\pi/\Delta\theta$, therefore

$$\frac{i}{n_i+1} \leqslant m \leqslant \frac{i}{n_i},\quad \forall i\in\mathbb{N}. \tag{11}$$

The two variable needed to solve for $\delta_\perp\boldsymbol{x}_{\rm cyc}$ by equation (8) are $\mathcal{D}\mathcal{P}^{\boldsymbol{m}}$ and $\delta\mathcal{P}^{\boldsymbol{m}}$. An expression of $\mathcal{D}\mathcal{P}^{\boldsymbol{k}}$ in terms of $\partial_r\boldsymbol{\chi}$ and $\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}$ is acquired by exerting total derivatives in r and $\boldsymbol{\theta}$ *resp.* on the equation (1) defining $\mathcal{P}^{\boldsymbol{k}}$,

$$\begin{aligned}\mathcal{D}\mathcal{P}^{\boldsymbol{k}}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right)\right)\cdot\partial_r\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right) &= \partial_r\boldsymbol{\chi}\left(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r\right)\\ &+\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\left(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r\right)\cdot\left(\boldsymbol{k}*\frac{{\rm d}\Delta\boldsymbol{\theta}}{{\rm d}r}\right),\end{aligned} \tag{12}$$

$$\mathcal{D}\mathcal{P}^{\boldsymbol{k}}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right)\right)\cdot\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right) = \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\left(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r\right), \tag{13}$$

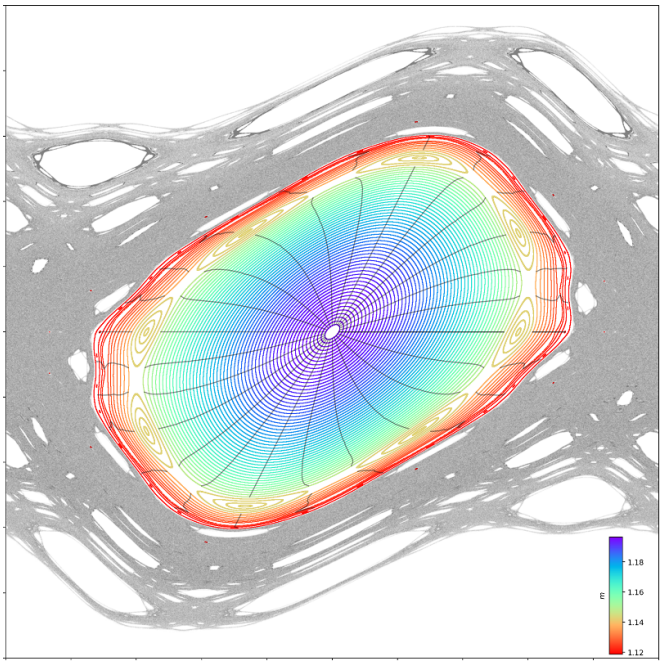

Figure 2. Distribution of m for the Chirikov standard map at $k = 0.975$ for $\boldsymbol{x} \in [-0.5, 0.5] \times [-0.5, 0.5]$. Iso-$\theta$ contours are plotted with spacing $\pi/12$.

$$\begin{aligned}\mathcal{DP}^{\boldsymbol{k}}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right)\right) = &\left[\begin{array}{cc} \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\cdot\left(\boldsymbol{k}*\frac{\mathrm{d}\Delta\boldsymbol{\theta}}{\mathrm{d}r}\right) & \partial_{\boldsymbol{\theta}}\boldsymbol{\chi} \\ +\partial_r\boldsymbol{\chi} & \end{array}\right]\Bigg|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}} \\ &\cdot\left[\partial_r\boldsymbol{\chi} \quad \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\right]\Big|_{\boldsymbol{\theta}}^{-1} \\ = &\left[\begin{array}{cc} | & |\,|\,| \\ \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\cdot\left(\boldsymbol{k}*\frac{\mathrm{d}\Delta\boldsymbol{\theta}}{\mathrm{d}r}\right) & \boldsymbol{0}_{N\times(N-1)} \\ | & |\,|\,| \end{array}\right]\Bigg|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}} \\ &\cdot\left[\partial_r\boldsymbol{\chi} \quad \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\right]\Big|_{\boldsymbol{\theta}}^{-1} \\ &+\left[\begin{array}{cc} | & |\,|\,| \\ \partial_r\boldsymbol{\chi} & \partial_{\boldsymbol{\theta}}\boldsymbol{\chi} \\ | & |\,|\,| \end{array}\right]\Bigg|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}} \cdot\left[\partial_r\boldsymbol{\chi} \quad \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\right]\Big|_{\boldsymbol{\theta}}^{-1},\end{aligned}$$

where the vertical and horizontal lines could be helpful to tell how the vectors fill the matrix, whether as a column or a row. Note the inverse of $\left[\begin{array}{cc} | & |\,|\,| \\ \partial_r\boldsymbol{\chi} & \partial_{\boldsymbol{\theta}}\boldsymbol{\chi} \\ | & |\,|\,| \end{array}\right]$ is $\left[\begin{array}{c} -\nabla r- \\ \equiv\nabla\boldsymbol{\theta}\equiv \end{array}\right]$, so the above equation can be further reduced to an expression of $\mathcal{DP}^{\boldsymbol{k}}$ in terms of the $\boldsymbol{\theta}$ and r grid,

$$\begin{aligned}\mathcal{DP}^{\boldsymbol{k}}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right)\right) = &\left[\begin{array}{c} | \\ \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\cdot\left(\boldsymbol{k}*\frac{\mathrm{d}\Delta\boldsymbol{\theta}}{\mathrm{d}r}\right) \\ | \end{array}\right]\Bigg|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}} \cdot\left[-\nabla r-\right]\Big|_{\boldsymbol{\theta}} \\ &+\left[\begin{array}{cc} | & |\,|\,| \\ \partial_r\boldsymbol{\chi} & \partial_{\boldsymbol{\theta}}\boldsymbol{\chi} \\ | & |\,|\,| \end{array}\right]\Bigg|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}} \cdot\left[\begin{array}{c} -\nabla r- \\ \equiv\nabla\boldsymbol{\theta}\equiv \end{array}\right]\Big|_{\boldsymbol{\theta}}, \quad (14)\end{aligned}$$

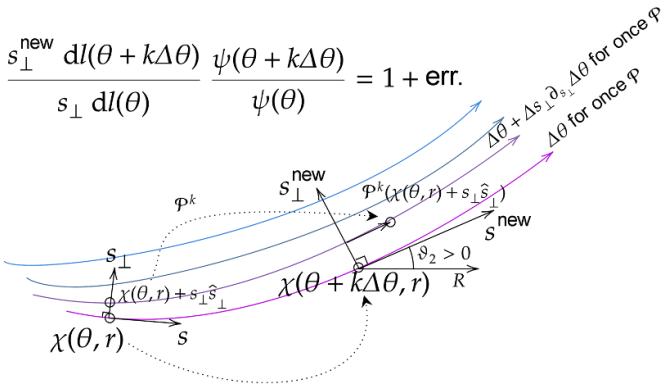

Figure 3. Cartoon to show the $\mathcal{DP}^k$ and the local frames of coordinate systems. The point $\boldsymbol{\chi}(\theta,r)+s_\perp\hat{\boldsymbol{s}}_\perp$, which is a bit shift from $\boldsymbol{\chi}(\boldsymbol{\theta},r)$ in the normal direction $\hat{\boldsymbol{s}}_\perp$ is mapped away from the $s_\perp^{\text{new}}$ axis due to shear.

of which a special case is when $\boldsymbol{k}$ takes the value of $\boldsymbol{m}$ (the latter term on the RHS degrades to an identity matrix $\mathbf{I}$),

$$\mathcal{DP}^{\boldsymbol{m}}\left(\boldsymbol{x}\right) = \left(\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\cdot\left(\boldsymbol{m}*\frac{\mathrm{d}\Delta\boldsymbol{\theta}}{\mathrm{d}r}\right)\right)\nabla r + \mathbf{I}_{N\times N} \quad (15)$$

where all the variables are evaluated at $\boldsymbol{x}$, therefore it is needless to indicate where to evaluate.

The eigenvalues of $\mathcal{DP}^{\boldsymbol{m}}$ are identical at all points of an ($\boldsymbol{\omega}$-incommensurable) invariant torus because the evolution of $\mathcal{DP}^{m}$ does not alter the eigenvalues [42] and evolution on an ω-incommensurable torus can reach almost every point on the torus. Some properties, e.g. this one on eigenvalues invariance, can be transferred between an $\boldsymbol{\omega}$-commensurable torus and an $\boldsymbol{\omega}$-incommensurable torus, because in most cases the former one can be considered as a limit of a sequence of the latter ones nearby and vice versa. However, there exist extreme counterexamples in which the transfer is hindered, e.g. an isolated invariant torus which has no invariant torus nearby. For non-conservative systems, the case of an invariant torus being isolated is not extreme but instead universal because each invariant torus is separate from others for a finite distance.

Be careful that $\mathcal{DP}^{\boldsymbol{k}}(\boldsymbol{x})$ is not $\boldsymbol{m}$-periodic in $\boldsymbol{k}$ as $\mathcal{P}^{\boldsymbol{k}}$ is, which is because points on nearby invariant tori are probably to be mapped gradually away from each other as $\boldsymbol{k}$ increases due to the shear between tori, i.e. neighbouring tori have a bit difference in rotation vectors, as shown in figure 3. $\delta\mathcal{P}^{\boldsymbol{k}}(\boldsymbol{x})$ is also not $\boldsymbol{m}$-periodic but due to a reason other than shear: the impact of perturbation is accumulated all the way.

The distance between two neighbouring invariant tori depends on the point at which the distance is evaluated, as reflected by $\mathcal{DP}^{k}$. Let $(N,d) = (2,1)$ for illustration (see figure 3) and this case is of great importance owing to that the distance variation also reflects the local density of flux surfaces in an MCF machine. Denote a matrix representing φ rad counterclockwise rotation in $\mathbb{R}^2$ by $\mathbf{R}_\varphi$. Construct local

coordinate frames at $\boldsymbol{x}$ and $\mathcal{P}^k(\boldsymbol{x})$ *resp.* with orthonormal bases $\{\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}_\perp\}$ and $\{\hat{\boldsymbol{s}}^{\text{new}},\hat{\boldsymbol{s}}^{\text{new}}_\perp\}$, then

$$\begin{aligned}
\mathrm{d}\boldsymbol{x}_{\text{loc}} &= \mathbf{R}_{-\vartheta_1}\mathrm{d}\boldsymbol{x}\\
\begin{bmatrix}\mathrm{d}s^{\text{new}}\\ \mathrm{d}s^{\text{new}}_\perp\end{bmatrix} = \overbrace{\mathrm{d}\boldsymbol{x}^{\text{new}}_{\text{loc}}} &\equiv \mathrm{d}s\,\hat{\boldsymbol{s}} + \mathrm{d}s_\perp\,\hat{\boldsymbol{s}}_\perp = \begin{bmatrix}\mathrm{d}s\\ \mathrm{d}s_\perp\end{bmatrix}\\
\overbrace{\mathrm{d}\boldsymbol{x}^{\text{new}}_{\text{loc}}} &= \mathcal{D}\mathcal{P}^k_{\text{loc}}\,\overbrace{\mathrm{d}\boldsymbol{x}_{\text{loc}}}\\
\mathrm{d}\boldsymbol{x}^{\text{new}}_{\text{loc}} &= \mathbf{R}_{-\vartheta_2}\mathrm{d}\boldsymbol{x}^{\text{new}}\\
\mathrm{d}\boldsymbol{x}^{\text{new}} &= \underbrace{\mathbf{R}_{\vartheta_2}\mathcal{D}\mathcal{P}^k_{\text{loc}}\mathbf{R}_{-\vartheta_1}}_{=\mathcal{D}\mathcal{P}^k}\mathrm{d}\boldsymbol{x}
\end{aligned}$$

where the subscripts $_{\text{loc}}$ mean the variables are viewed in the local frames. $\mathcal{D}\mathcal{P}^k_{\text{loc}}$ can be deduced from $\mathcal{D}\mathcal{P}^k$ or vice versa. Another relation between them is

$$\mathcal{D}\mathcal{P}^k_{\text{loc}} = \begin{bmatrix}\hat{\boldsymbol{s}}^{\text{newT}}\cdot\mathcal{D}\mathcal{P}^k\cdot\hat{\boldsymbol{s}} & , & \hat{\boldsymbol{s}}^{\text{newT}}\cdot\mathcal{D}\mathcal{P}^k\cdot\hat{\boldsymbol{s}}_\perp\\ \hat{\boldsymbol{s}}^{\text{newT}}_\perp\cdot\mathcal{D}\mathcal{P}^k\cdot\hat{\boldsymbol{s}} & , & \hat{\boldsymbol{s}}^{\text{newT}}_\perp\cdot\mathcal{D}\mathcal{P}^k\cdot\hat{\boldsymbol{s}}_\perp\end{bmatrix}. \tag{16}$$

If a map is flux-preserving, i.e.

$$\psi(\boldsymbol{x})\,\mathrm{d}S(\boldsymbol{x}) = \psi(\mathcal{P}(\boldsymbol{x}))\,\mathrm{d}S(\mathcal{P}(\boldsymbol{x})),$$

(where ψ is the flux density function and $\psi\,\mathrm{d}S$ is the flux) then this property can provide a first-order estimation for the distance $s_\perp(\theta)$ between tori by

$$\begin{aligned}
&\psi(\theta)\,s_\perp(\theta)\,\mathrm{d}l(\theta) = \psi(\theta_0)\,s_\perp(\theta_0)\,\mathrm{d}l(\theta_0) + \ldots,\\
&s_\perp(\theta) = 0 + \underbrace{s_\perp(\theta_0)}_{s_{\perp0}:=}\quad\underbrace{\frac{\mathrm{d}l(\theta_0)}{\mathrm{d}l(\theta)}\frac{\psi(\theta_0)}{\psi(\theta)}}_{\text{can be considered as 1st derivative }\frac{\mathrm{d}s_\perp}{\mathrm{d}s_{\perp0}}}\\
&\qquad + \mathcal{O}\left(|s_{\perp0}|^2\right)\\
&\qquad = s_{\perp0}\frac{|\partial_\theta\boldsymbol{\chi}|_{\theta_0}}{|\partial_\theta\boldsymbol{\chi}|_\theta}\frac{\psi(\theta_0)}{\psi(\theta)} + \mathcal{O}\left(|s_{\perp0}|^2\right),
\end{aligned}\tag{17}$$

where r as an argument in (θ,r) is omitted for brevity since this is an expansion of $s_\perp$ near the one invariant torus of concern. For fusion devices, where the Poincaré map $\mathcal{P}$ is defined for one toroidal turn, the flux density $\psi(\boldsymbol{x}) = B_\phi(\boldsymbol{x})$. For a general high-dimensional system, $(N,d) = (N,N-1)$, the flux-preserving property has a general form as below,

$$\det\left[\begin{array}{ccccc}| & \cdots & | & & |\\ \partial_{\theta_1}\boldsymbol{\chi} & \cdots & \partial_{\theta_d}\boldsymbol{\chi} & & s_\perp(\boldsymbol{\theta})\hat{\boldsymbol{s}}_\perp\\ | & \cdots & | & & |\end{array}\right]\Bigg|_{\boldsymbol{\theta}}\psi(\boldsymbol{\theta}) = \text{const.}, \tag{18}$$

(where $\hat{\boldsymbol{s}}_\perp$ is the unit vector normal to the torus) by which one can have a similar first order estimate to $s_\perp(\boldsymbol{\theta})$.

Hereafter, our focus moves from $\mathcal{D}\mathcal{P}^{\boldsymbol{k}}$ to $\delta\mathcal{P}^{\boldsymbol{k}}$ to provide readers with the formulae describing the deformation of invariant tori. By regarding the location $\boldsymbol{x}$ more fundamental than $(\boldsymbol{\theta},r)$ and considering the whole map $\mathcal{P}$ also as an argument of $\boldsymbol{\theta}$, $\Delta\boldsymbol{\theta}$, r and $\boldsymbol{\chi}$, the defining equation (1) for $\mathcal{P}^{\boldsymbol{k}}$ becomes

$$\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{x}) := \boldsymbol{\chi}[\mathcal{P}](\boldsymbol{\theta}[\mathcal{P}](\boldsymbol{x}) + \boldsymbol{k}*\Delta\boldsymbol{\theta}[\mathcal{P}](\boldsymbol{x}), r[\mathcal{P}](\boldsymbol{x})), \tag{19}$$

which after imposed $\Delta\mathcal{P}\cdot\mathrm{d}/\mathrm{d}\mathcal{P}$ converts to,

$$\begin{aligned}
\delta\mathcal{P}^{\boldsymbol{k}}(\boldsymbol{x}) = {}&\delta\boldsymbol{\chi}(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r)\\
&+\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r)\cdot(\delta\boldsymbol{\theta}(\boldsymbol{x})+\boldsymbol{k}*\delta\Delta\boldsymbol{\theta}(\boldsymbol{x}))\\
&+\partial_r\boldsymbol{\chi}(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r)\cdot\delta r(\boldsymbol{x}),
\end{aligned}\tag{20}$$

that is the $\boldsymbol{x}$-based form of *the first-order deformation formula of invariant tori under perturbation*, allowing for any possible choice of coordinates, e.g. where to define the curve on which $\boldsymbol{\theta} = \boldsymbol{0}$. For readers interested in higher-order derivations of this formula, further details are available in appendix A.

When $\boldsymbol{k}$ takes the value of $\boldsymbol{m}$, one can drop $\boldsymbol{k}*\Delta\boldsymbol{\theta}$ because of the periodicity of $\boldsymbol{\chi}$ in $\boldsymbol{\theta}$,

$$\begin{aligned}
\delta\mathcal{P}^{\boldsymbol{m}}(\boldsymbol{x}) = {}&\delta\boldsymbol{\chi}(\boldsymbol{\theta},r)+\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}(\boldsymbol{\theta},r)\cdot(\delta\boldsymbol{\theta}(\boldsymbol{x})+\boldsymbol{m}*\delta\Delta\boldsymbol{\theta}(\boldsymbol{x}))\\
&+\partial_r\boldsymbol{\chi}(\boldsymbol{\theta},r)\cdot\delta r(\boldsymbol{x}),
\end{aligned}\tag{21}$$

There are three common, easy-to-understand criteria to *identify*, during perturbation being exerted, *an invariant torus*, which can simplify the above equation by removing redundant arbitrariness in the choice of coordinates. ① The first one is to anchor it by a fixed point $\boldsymbol{x}$, which implies that $\delta r(\boldsymbol{x})$ at this point always vanishes. In the meanwhile, if $\theta(\boldsymbol{x})$ at this point is endowed with a constant value no matter what perturbation is imposed, $\delta\theta(\boldsymbol{x})$ also vanishes. Then $\boldsymbol{\chi}(\theta(\boldsymbol{x}),r(\boldsymbol{x}))$ at this point is also fixed, i.e. $\delta\boldsymbol{\chi}(\theta,\phi) = 0$. The equation (20) is simplified to

$$\delta\mathcal{P}^{\boldsymbol{k}}(\boldsymbol{x}) = \delta\boldsymbol{\chi}|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}} + \partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\Big|_{\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta}}\cdot(\boldsymbol{k}*\delta\Delta\boldsymbol{\theta}(\boldsymbol{x})), \tag{22}$$

where $\delta\mathcal{P}^{\boldsymbol{k}}(\boldsymbol{x})$ can be computed for $\boldsymbol{k} = k*\boldsymbol{1}$, $k\in\mathbb{Z}$, by the discrete-time version of the first variation progression equation (5*b*),

$$\delta\mathcal{P}^{k+1}(\boldsymbol{x}_0) = \delta\mathcal{P}(\boldsymbol{x})|_{\boldsymbol{x}=\mathcal{P}^k(\boldsymbol{x}_0)} + \delta\mathcal{P}^k(\boldsymbol{x}_0)\cdot\mathcal{D}\mathcal{P}(\boldsymbol{x})|_{\boldsymbol{x}=\mathcal{P}^k(\boldsymbol{x}_0)} \tag{23}$$

while the other unknowns are $\delta\boldsymbol{\chi}$ (2π-periodic in every θ_i) and $\delta\Delta\boldsymbol{\theta}(\boldsymbol{x})\in\mathbb{R}^d$. One can employ least-squares methods (as used in figure 4) or other fitting techniques to estimate the Fourier series coefficients of $\delta\boldsymbol{\chi}$ and the scalar value of $\delta\Delta\boldsymbol{\theta}(\boldsymbol{x})$.

② The second common choice is a class of winding-based radius choices. When $(N,d) = (2,1)$, a winding-based radius can be *resp.* $\Delta\theta$, rotation transform ι, safety factor $q = 1/\iota$ or any function dependent only on these terms. For simplicity, let $r := \Delta\theta(\boldsymbol{x})$ and endow the initiating point $\boldsymbol{x}$ with a fixed angle θ (so $\delta\boldsymbol{\theta}(\boldsymbol{x})$ is cancelled below), then, equation (20) is simplified into

$$\begin{aligned}
\delta\mathcal{P}^k(\boldsymbol{x}) = {}&\delta\boldsymbol{\chi}|_{\theta+k\Delta\theta} + (\partial_\theta\boldsymbol{\chi})|_{(\theta+k\Delta\theta,r)}\cdot\cancel{\delta\boldsymbol{\theta}(\boldsymbol{x})}\\
&+(k\partial_\theta\boldsymbol{\chi}+\partial_r\boldsymbol{\chi})|_{(\theta+k\Delta\theta,r)}\cdot\delta r(\boldsymbol{x}).
\end{aligned}\tag{24}$$

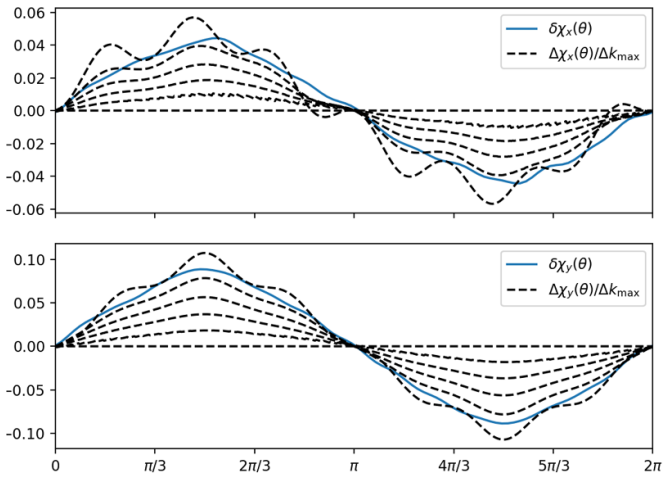

Figure 4. Shifts of $\boldsymbol{\chi}(\theta, r)$ for an invariant torus $\mathcal{T}^1$ anchored by a starting point $\boldsymbol{x}_0 = [.12, .0]$ at which the angle is fixed to be zero. $k_0 = 0.975$ while Δk takes values of $.00, .02, .04, .06, .08, 0.10$. (a) and (b) *resp.* for the x and y components. The first variation $\delta\boldsymbol{\chi}(\theta)$ is computed here by equation (22), while the finite increment $\Delta\boldsymbol{\chi}(\theta)$ is simply the difference of $\boldsymbol{\chi}(\theta)$ between after and before the perturbation is imposed.

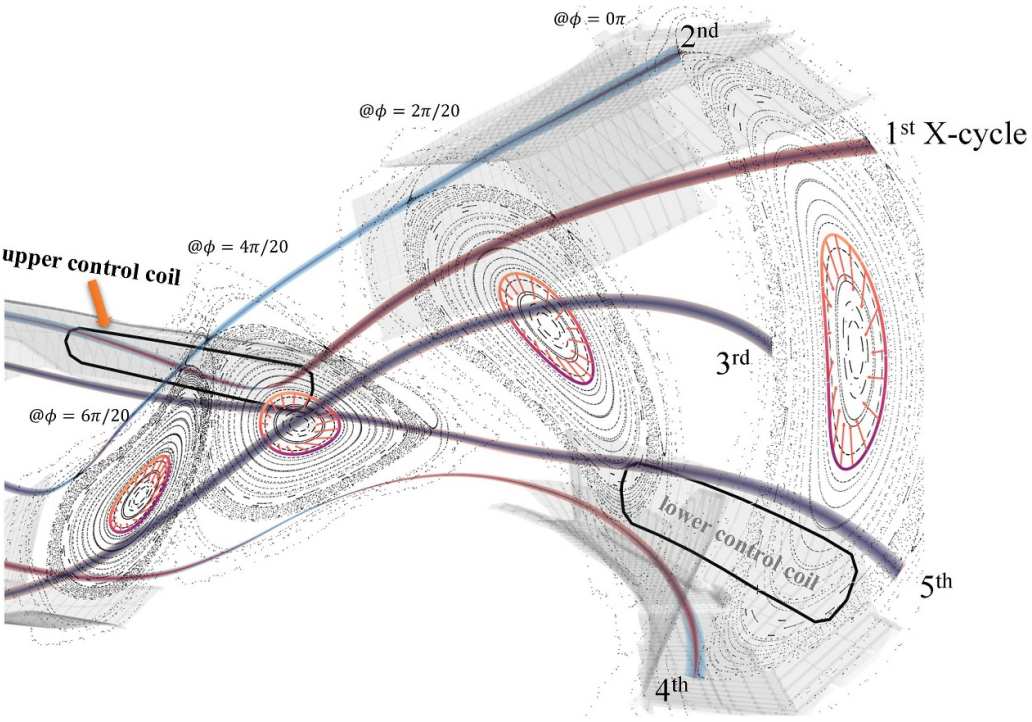

Figure 5. With the perturbation field $\delta\mathcal{B}$ chosen to be the vacuum field generated by the set of **upper control coils** (1 kA per coil), $\delta\boldsymbol{\chi}(\theta, r)$ are drawn as arrows for a flux surface (r is chosen to be $\Delta\theta$, so this torus is identified by its rotation transform during perturbation) in the standard configuration of Wendelstein 7-X (i.e. the unperturbed system and the perturbation are *resp.* $\boldsymbol{B}_{\text{std}}$ and $\boldsymbol{B}_{\text{upCC}}$. CC is short for Control Coil), amplified by a factor of 30 for visibility. The first variation $\delta\boldsymbol{\chi}(\theta, r)$ is computed here by equation (26) on the rightmost cross-section at $\phi = 0$ rad, then evolved to others by equation (5*b*).

Owing to the fact that the perpendicular shift is an intrinsic property (a term from differential geometry) of $\mathcal{T}^d$, $\delta_\perp \boldsymbol{x}_{\text{cyc}}$ computed by equation (8) shall equal the normal part of $\delta\boldsymbol{\chi}(\theta, r)$ computed by equation (24), that is $\sum_{i=1}^{N-d} \delta_\perp \boldsymbol{\chi} = \hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_i^{\mathrm{T}} \cdot \delta\boldsymbol{\chi}$, when the radial label r is chosen to be $\Delta\theta$.

One may find the condition that the initiating point $\boldsymbol{x}$ is bound with a fixed angle inconvenient, e.g. in fusion devices, usually points at the low-field side having the identical Z-coordinates as that of the magnetic axis are considered to have angles $\theta = 0$. To facilitate setting such a condition, consider

$(\boldsymbol{\theta}, r)$ more fundamental than $\boldsymbol{x}$, i.e. let $\boldsymbol{x}$ be a function of $(\boldsymbol{\theta}, r)$, that is $\boldsymbol{\chi}(\boldsymbol{\theta}, r)$. Then, the defining equation (1) for $\mathcal{P}^{\boldsymbol{k}}$ becomes

$$\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{\chi}[\mathcal{P}](\boldsymbol{\theta}, r)) = \boldsymbol{\chi}[\mathcal{P}](\boldsymbol{\theta} + \boldsymbol{k} * \Delta\boldsymbol{\theta}[\mathcal{P}](r), r), \tag{25}$$

which after imposed $\Delta\mathcal{P} \cdot \mathrm{d}/\mathrm{d}\mathcal{P}$ converts to

$$\begin{aligned}&\delta\mathcal{P}^{\boldsymbol{k}}(\boldsymbol{\chi}(\boldsymbol{\theta}, r)) + \mathcal{D}\mathcal{P}^{\boldsymbol{k}}(\boldsymbol{\chi}(\boldsymbol{\theta}, r)) \cdot \delta\boldsymbol{\chi}(\boldsymbol{\theta}, r) \\ &\qquad = \delta\boldsymbol{\chi}(\boldsymbol{\theta} + \boldsymbol{k} * \Delta\boldsymbol{\theta}, r) + (\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}) \Big| \overbrace{\delta(\boldsymbol{k} * \Delta\boldsymbol{\theta})}^{\text{vanishes if } \Delta\boldsymbol{\theta} \text{ is merely dependent on } r, \text{ not on } \mathcal{P}.}_{(\boldsymbol{\theta} + \boldsymbol{k} * \Delta\boldsymbol{\theta}, r)},\end{aligned} \tag{26}$$

that is the $(\boldsymbol{\theta}, r)$-based form of *the first-order deformation formula of invariant tori under perturbation*. The aforementioned condition defined by the horizontal $\theta = 0$ line has a natural expression,

$$\hat{\boldsymbol{e}}_Z \cdot \delta\boldsymbol{\chi}(\theta = 0, r) = \underbrace{\hat{\boldsymbol{e}}_Z \cdot \delta\boldsymbol{\chi}(\theta = 0, r = 0)}_{\text{the magnetic axis shift } Z\text{-component}}. \tag{27}$$

③ The third choice is a class of labels that are positively related to the volume enclosed by the flux surface, e.g. the volume itself V, $R_{\mathrm{LFS}} - R_{\mathrm{ax}}$, the poloidal flux Ψ_{pol}, its normalised square root $s := \sqrt{\Psi_{\mathrm{pol}}}/\sqrt{\Psi_{\mathrm{pol,\,max}}}$, *etc*. Regarding $R_{\mathrm{LFS}} - R_{\mathrm{ax}}$, R_{LFS} is the intersection point R-coordinate of the torus and the horizontal line drawn from the magnetic axis to the low field side (LFS), while R_{ax} is simply that of the magnetic axis. If the radius label r is chosen to be $R_{\mathrm{LFS}} - R_{\mathrm{ax}}$, the condition further simplifies to

$$\delta\boldsymbol{\chi}(\theta = 0, r) = \underbrace{\delta\boldsymbol{\chi}(\theta = 0, r = 0)}_{\text{the magnetic axis shift}}. \tag{28}$$

For tokamaks, $\delta\boldsymbol{\chi}(\theta, r)$ only needs to be solved for one R–Z section because Poincaré mappings on other R-Z sections behave the same. For stellarators, the $\delta\boldsymbol{\chi}(\theta, r)$ calculated for one such section can be evolved to other sections by equation (5*b*), as shown in figure 5. Equation (5*b*) not only serves as a progression formula for $\delta\boldsymbol{X}_{\mathrm{pol}}$ of field line tracing but also an evolution one because the following field line tracing equation

$$\frac{\mathrm{d}}{\mathrm{d}\phi}\boldsymbol{\chi}(\boldsymbol{\theta} + \iota\phi, r, \phi) = \frac{R\boldsymbol{B}_{\mathrm{pol}}}{B_\phi}(\boldsymbol{\chi}(\boldsymbol{\theta}, r, \phi), \phi), \tag{29}$$

with the $\mathcal{B}$ argument explicitly stated and both sides differentiated with $\Delta\mathcal{B} \cdot \mathrm{d}/\mathrm{d}\mathcal{B}$,

$$\frac{\mathrm{d}}{\mathrm{d}\phi}\boldsymbol{\chi}[\mathcal{B}](\boldsymbol{\theta} + \iota(r)\phi, r, \phi) = \frac{R\boldsymbol{B}_{\mathrm{pol}}}{B_\phi}[\mathcal{B}](\boldsymbol{\chi}[\mathcal{B}](\boldsymbol{\theta}, r, \phi), \phi), \tag{30}$$

$$\begin{aligned}&\delta\frac{\mathrm{d}}{\mathrm{d}\phi}\boldsymbol{\chi}[\mathcal{B}](\boldsymbol{\theta} + \iota(r)\phi, r, \phi) \\ &\qquad = \left(\delta\boldsymbol{\chi} \cdot \frac{\partial}{\partial(R, Z)}\right)\frac{R\boldsymbol{B}_{\mathrm{pol}}}{B_\phi} + \delta\left(\frac{R\boldsymbol{B}_{\mathrm{pol}}}{B_\phi}\right),\end{aligned} \tag{31}$$

has the same expression as the progression equation (5*b*). Note the $\mathcal{B}$-dependence of $\iota = \iota[\mathcal{B}](r)$ is removed provided r is a winding-based radius.

4. Conclusion and discussion

In summary, this paper borrows the notions of functional and functional derivative from mathematics to address *invariant tori and their deformation under perturbation*. This approach enables a direct connection between changes in the magnetic field and the resulting flux surface deformations, providing insight into the specific perturbations needed to achieve a desired toroidal magnetic topology (this is an inverse problem of how flux surfaces deform under a given perturbation).

Notably, we do not include plasma response in this analysis but consider a general perturbation from a view of mapping to maintain the theory's broad applicability across finite-dimensional dynamical systems emerging in various domains and to ensure clarity and accessibility for readers. Plasmas with different parameters may behave distinctly, one can use various models of plasma response magnetic field

$$\Delta\boldsymbol{B} = \boldsymbol{B}_{\mathrm{response}}[\mathcal{B}_{\mathrm{external}}] + \boldsymbol{B}_{\mathrm{external}}$$

to calculate the resulting change of Poincaré map $\Delta\mathcal{P}$ or merely its first variation $\delta\mathcal{P}$ by equation (5*b*). Calligraphic font is used for the external magnetic field perturbation as an argument of $\boldsymbol{B}_{\mathrm{response}}[\mathcal{B}_{\mathrm{external}}]$, because it is regarded as a whole rather than evaluated at a specific point. The massless force-balance equation commonly used in magnetohydrodynamics

$$\boldsymbol{J} \times \boldsymbol{B} = \nabla p, \tag{32}$$

which will become the following form if one considers its first-order variation,

$$\delta\boldsymbol{J} \times \boldsymbol{B}_0 + \boldsymbol{J}_0 \times (\delta\boldsymbol{B}_{\mathrm{plasma}} + \delta\boldsymbol{B}_{\mathrm{ext}}) = \delta\nabla p. \tag{33}$$

Ampère's law $\nabla \times \boldsymbol{B} = \mu_0\boldsymbol{J}$ becomes (when applied the variational principle) :

$$\nabla \times \delta\boldsymbol{B}_{\mathrm{plasma}} = \mu_0\delta\boldsymbol{J}. \tag{34}$$

Similarly, the divergence-free condition of the magnetic field $\nabla \cdot \boldsymbol{B} = 0$ converts to:

$$\nabla \cdot \delta\boldsymbol{B}_{\mathrm{plasma}} = 0 \tag{35}$$

$\delta\boldsymbol{B}_{\mathrm{plasma}}$ can be considered generated by $\delta\boldsymbol{J}$ according to Biot–Savart law or Ampère's law if the change of $\boldsymbol{E}$ is comparatively slow. The current distribution $\boldsymbol{J}$ on stellarators has Pfirsch–Schlüter and diamagnetic components [53], which can be determined by the pressure distribution (and magnetic field). Therefore, there are totally three equations (two vector equations and one scalar) that suffice to solve for the three unknowns $\delta\boldsymbol{J}$, $\delta\boldsymbol{B}_{\mathrm{response}}$ and δp (3+3+1 = 7 components).

The notation $\delta\boldsymbol{J}$ can be considered a directional functional derivative $\delta\boldsymbol{J}(x) = \delta\boldsymbol{J}[(\mathcal{J}_0, \mathcal{B}_0, p_0); (\Delta\mathcal{J}, \Delta\mathcal{B}, \Delta p)](x)$. The relevant physical fields are bundled to be a tuple to emphasise the PDE (partial differential equation) nature of the system. If only one physical field is written as an argument in the square bracket, the corresponding functional derivative

would only work for that physical field and fail to do calculus on other physical fields.

Be aware that not necessarily every point on an invariant torus has an accurate solution for $\delta_{\perp}\boldsymbol{x}_{\mathrm{cyc}}$ due to the fact that invariant tori may only be defined on a fragmented domain both in the space of $\mathcal{P}$ and that of $\boldsymbol{x}$. An invariant torus can be destroyed into an island chain or a cantorus that has infinite gaps.

To transfer these formulae from maps to flows, one usually merely needs to replace the symbols as shown below,

$$\mathcal{P}\left(\boldsymbol{x}_0\right) \mapsto \boldsymbol{X}\left(\boldsymbol{x}_0, t\right) \text{abbr. as } \boldsymbol{X}_t\left(\boldsymbol{x}_0\right),$$

where t is fixed, so $\boldsymbol{X}_t$ can be considered as a map,

$$\begin{gathered}
\boldsymbol{\varphi}\left(\mathcal{P}\left(\boldsymbol{x}\right)\right) - \boldsymbol{\varphi}\left(\boldsymbol{x}\right) =: \Delta\boldsymbol{\theta} \mapsto \frac{\mathrm{d}\boldsymbol{\varphi}\left(\boldsymbol{X}\left(\boldsymbol{x}_0, t\right)\right)}{\mathrm{d}t} =: \boldsymbol{\omega}, \\
\boldsymbol{m} := \left[\frac{2\pi}{\Delta\theta_1}, \ldots, \frac{2\pi}{\Delta\theta_d}\right] \mapsto \boldsymbol{T} := \left[\frac{2\pi}{\omega_1}, \ldots, \frac{2\pi}{\omega_d}\right], \\
\mathcal{P}^{\boldsymbol{k}}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta}\right)\right) := \boldsymbol{\chi}\left(\boldsymbol{\theta} + \boldsymbol{k} * \Delta\boldsymbol{\theta}\right) \mapsto \boldsymbol{X}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta}\right), \boldsymbol{t}\right) := \boldsymbol{\chi}\left(\boldsymbol{\theta} + \boldsymbol{\omega} * \boldsymbol{t}\right), \\
\mathcal{P}^{\boldsymbol{m}} \mapsto \boldsymbol{X}_{\boldsymbol{T}}.
\end{gathered}$$

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/WenyinWei/Jynamics.jl.

Acknowledgment

This work was supported by National Magnetic Confined Fusion Energy R&D Program of China (No. 2022YFE03030001) and National Natural Science Foundation of China (Nos. 12275310 and 12175277). Additionally, the author Wenyin Wei would like to express his gratitude to the China Scholarship Council for providing financial support for his doctoral joint cultivation in Europe.

Conflict of interest

The authors have no conflicts to disclose.

Appendix A. High order deformation of invariant tori

With $\boldsymbol{x}$ considered more fundamental than $(\boldsymbol{\theta}, r)$, the defining equation (1) for $\mathcal{P}^{\boldsymbol{k}}$ is complicated into the following form by regarding the system $\mathcal{P}$ as an argument of $(\mathcal{P}^{\boldsymbol{k}}, \boldsymbol{\chi}, \boldsymbol{\theta}, \Delta\boldsymbol{\theta}, r)$,

$$\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{x}) := \boldsymbol{\chi}[\mathcal{P}](\boldsymbol{\theta}[\mathcal{P}](\boldsymbol{x}) + \boldsymbol{k} * \Delta\boldsymbol{\theta}[\mathcal{P}](\boldsymbol{x}), r[\mathcal{P}](\boldsymbol{x})), \tag{19 revisited}$$

which after being exerted $\Delta\mathcal{P}\cdot \mathrm{d}/\mathrm{d}\mathcal{P}$ can yield the first form of the complete formula describing the first-order deformation of invariant tori, repeated as below,

$$\delta\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{x}) = \delta\boldsymbol{\chi} + (\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta}) \cdot \partial_{\boldsymbol{\theta}}\boldsymbol{\chi} + \delta r \cdot \partial_r \boldsymbol{\chi} \tag{20 revisited}$$

The more times we apply $\Delta\mathcal{P}\cdot \mathrm{d}/\mathrm{d}\mathcal{P}$ on both sides, the higher the order of deformation equation we obtain,

$$\begin{aligned}
\delta^2\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{x}) = {} & \delta^2\boldsymbol{\chi} + 2(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta}) \cdot \partial_{\boldsymbol{\theta}}\delta\boldsymbol{\chi} + 2\delta r \cdot \partial_r \delta\boldsymbol{\chi} \\
& + \left(\delta^2\boldsymbol{\theta} + \boldsymbol{k} * \delta^2\Delta\boldsymbol{\theta}\right) \cdot \partial_{\boldsymbol{\theta}}\boldsymbol{\chi} + (\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})^2 \cdot \partial_{\boldsymbol{\theta}}^2\boldsymbol{\chi} \\
& + 2(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})\delta r \cdot \partial_r\partial_{\boldsymbol{\theta}}\boldsymbol{\chi} \\
& + \delta^2 r \cdot \partial_r\boldsymbol{\chi} + (\delta r)^2 \partial_r^2\boldsymbol{\chi}
\end{aligned} \tag{A1}$$

$$\begin{aligned}
\delta^3\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{x}) = {} & \delta^3\boldsymbol{\chi} + 3(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta}) \cdot \partial_{\boldsymbol{\theta}}\delta^2\boldsymbol{\chi} + 3\delta r \cdot \partial_r\delta^2\boldsymbol{\chi} \\
& + 3\left(\delta^2\boldsymbol{\theta} + \boldsymbol{k} * \delta^2\Delta\boldsymbol{\theta}\right) \cdot \partial_{\boldsymbol{\theta}}\delta\boldsymbol{\chi} + 3(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})^2 \cdot \partial_{\boldsymbol{\theta}}^2\delta\boldsymbol{\chi} \\
& + 6(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})\delta r \cdot \partial_{\boldsymbol{\theta}}\partial_r\delta\boldsymbol{\chi} \\
& + 3(\delta r)^2 \cdot \partial_r^2\delta\boldsymbol{\chi} + 3\delta^2 r \cdot \partial_r\delta\boldsymbol{\chi} \\
& + \left(\delta^3\boldsymbol{\theta} + \boldsymbol{k} * \delta^3\Delta\boldsymbol{\theta}\right) \cdot \partial_{\boldsymbol{\theta}}\boldsymbol{\chi} + 3(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})\left(\delta^2\boldsymbol{\theta} + \boldsymbol{k} * \delta^2\Delta\boldsymbol{\theta}\right) \cdot \partial_{\boldsymbol{\theta}}^2\boldsymbol{\chi} \\
& + (\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})^3 \cdot \partial_{\boldsymbol{\theta}}^3\boldsymbol{\chi} \\
& + 3\left(\delta^2\boldsymbol{\theta} + \boldsymbol{k} * \delta^2\Delta\boldsymbol{\theta}\right)\delta r \cdot \partial_{\boldsymbol{\theta}}\partial_r\boldsymbol{\chi} + 3(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})\delta^2 r \cdot \partial_{\boldsymbol{\theta}}\partial_r\boldsymbol{\chi} \\
& + 3(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})(\delta r)^2 \cdot \partial_{\boldsymbol{\theta}}\partial_r^2\boldsymbol{\chi} + 3(\delta\boldsymbol{\theta} + \boldsymbol{k} * \delta\Delta\boldsymbol{\theta})^2\delta r \cdot \partial_{\boldsymbol{\theta}}^2\partial_r\boldsymbol{\chi} \\
& + \delta^3 r \cdot \partial_r\boldsymbol{\chi} + 3\delta^2 r\delta r \cdot \partial_r^2\boldsymbol{\chi}.
\end{aligned} \tag{A2}$$

One can conclude the pattern as follows, that is *the $\boldsymbol{x}$-based form of the high-order deformation formula of invariant torus*,

$$\begin{aligned}
\frac{1}{n!}\delta^n\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{x}) = {} & \sum_{\substack{(n_{\boldsymbol{\theta}i}),(p_{\boldsymbol{\theta}i}),(n_{ri}),(p_{ri}),n_\chi \\ \text{such that} \\ \sum_{i=1}^{d_{\boldsymbol{\theta}}} n_{\boldsymbol{\theta}i}p_{\boldsymbol{\theta}i} + \sum_{i=1}^{d_r} n_{ri}p_{ri} + n_\chi = n}} \binom{p_{\boldsymbol{\theta}}^+}{p_{\boldsymbol{\theta}1}, \ldots, p_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}} \left(\frac{\delta^{n_{\boldsymbol{\theta}1}}\boldsymbol{\theta} + \boldsymbol{k} * \delta^{n_{\boldsymbol{\theta}1}}\Delta\boldsymbol{\theta}}{n_{\boldsymbol{\theta}1}!}\right)^{p_{\boldsymbol{\theta}1}} \cdots \left(\frac{\delta^{n_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}}\boldsymbol{\theta} + \boldsymbol{k} * \delta^{n_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}}\Delta\boldsymbol{\theta}}{n_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}!}\right)^{p_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}} \\
& \times \binom{p_r^+}{p_{r1}, \ldots, p_{rd_r}} \left(\frac{\delta^{n_{r1}} r}{n_{r1}!}\right)^{p_{r1}} \left(\frac{\delta^{n_{rd_r}} r}{n_{rd_r}!}\right)^{p_{rd_r}} \vdots_{(p_{\boldsymbol{\theta}}^+ + p_r^+)} \frac{\partial_{\boldsymbol{\theta}}^{p_{\boldsymbol{\theta}}^+}\partial_r^{p_r^+}\delta^{n_\chi}\chi}{p_{\boldsymbol{\theta}}^+! p_r^+! n_\chi!},
\end{aligned} \tag{A3}$$

where

$$\begin{aligned}
& n_\chi \geqslant 0 \\
& p^+ = p_1 + p_2 + \cdots + p_d, \quad d \text{ is the total number of powers}, \\
& p_i \geqslant 1, \\
& n_1 > n_2 > \cdots > n_d \geqslant 1,
\end{aligned}$$

with subscripts $_{\boldsymbol{\theta}}$ and $_r$ of $(n_{\boldsymbol{\theta}i}, p_{\boldsymbol{\theta}i}, p_{\boldsymbol{\theta}}^+, d_{\boldsymbol{\theta}}, n_{r\,i}, p_{r\,i}, p_r^+, d_r)$ omitted for brevity. The formula can be reduced by removing the redundant arbitrariness in choosing coordinates.

On the other hand, with $(\boldsymbol{\theta}, r)$ considered more fundamental than $\boldsymbol{x}$, the defining equation (1) for $\mathcal{P}^{\boldsymbol{k}}$ is complicated into the following form by regarding the system $\mathcal{P}$ as an argument of $\mathcal{P}^{\boldsymbol{k}}$, $\boldsymbol{\chi}$ and $\Delta\boldsymbol{\theta}$,

$$\mathcal{P}^{\boldsymbol{k}}[\mathcal{P}](\boldsymbol{\chi}[\mathcal{P}](\boldsymbol{\theta}, r)) = \boldsymbol{\chi}[\mathcal{P}](\boldsymbol{\theta} + \boldsymbol{k} * \Delta\boldsymbol{\theta}[\mathcal{P}](r), r), \tag{25 revisited}$$

which after imposed $\Delta\mathcal{P}\cdot \mathrm{d}/\mathrm{d}\mathcal{P}$ converts to

$$\begin{aligned}&\delta\mathcal{P}^{\boldsymbol{k}}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right)\right)+\mathcal{D}\mathcal{P}^{\boldsymbol{k}}\left(\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right)\right)\cdot\delta\boldsymbol{\chi}\left(\boldsymbol{\theta},r\right)\\&\quad=\delta\boldsymbol{\chi}\left(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r\right)+\left(\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}\right)\Big|_{\left(\boldsymbol{\theta}+\boldsymbol{k}*\Delta\boldsymbol{\theta},r\right)}\cdot\delta\left(\boldsymbol{k}*\Delta\boldsymbol{\theta}\right),\end{aligned}\tag{26 \textit{revisited}}$$

with arguments omitted for brevity,

$$\delta\mathcal{P}^{\boldsymbol{k}}+\delta\boldsymbol{\chi}\cdot\mathcal{D}\mathcal{P}^{\boldsymbol{k}}=\delta\boldsymbol{\chi}+\left(\boldsymbol{k}*\delta\Delta\boldsymbol{\theta}\right)\cdot\partial_{\boldsymbol{\theta}}\boldsymbol{\chi}.\tag{A4}$$

The second-order formula is

$$\begin{aligned}&\delta^2\mathcal{P}^{\boldsymbol{k}}+2\left(\delta\boldsymbol{\chi}\cdot\mathcal{D}\right)\delta\mathcal{P}^{\boldsymbol{k}}+\left(\delta^2\boldsymbol{\chi}\cdot\mathcal{D}\right)\mathcal{P}^{\boldsymbol{k}}+\left(\delta\boldsymbol{\chi}\cdot\mathcal{D}\right)^2\mathcal{P}^{\boldsymbol{k}}\\&\quad=\delta^2\boldsymbol{\chi}+2\left(\boldsymbol{k}*\delta\Delta\boldsymbol{\theta}\right)\cdot\partial_{\boldsymbol{\theta}}\delta\boldsymbol{\chi}+\left(\boldsymbol{k}*\delta\Delta\boldsymbol{\theta}\right)^2\cdot\partial_{\boldsymbol{\theta}}^2\boldsymbol{\chi}+\left(\boldsymbol{k}*\delta^2\Delta\boldsymbol{\theta}\right)\cdot\partial_{\boldsymbol{\theta}}\delta\boldsymbol{\chi}.\end{aligned}\tag{A5}$$

Hence, *the* $(\boldsymbol{\theta},r)$*-based form of the high-order deformation formula of invariant torus* is concluded as bellow for the *n*th-order variation with notations defined in a similar manner as those of the first form,

$$\begin{aligned}&\sum_{\substack{(n_{\boldsymbol{\chi}^i}),(p_{\boldsymbol{\chi}^i}),n_{\mathcal{P}}\text{ such that}\\ \sum_{i=1}^{d_{\boldsymbol{\chi}}}n_{\boldsymbol{\chi}^i}p_{\boldsymbol{\chi}^i}+n_{\mathcal{P}}=n}}\binom{p_{\chi}^{+}}{p_{\boldsymbol{\chi}_1},\ldots,p_{\boldsymbol{\chi}d_{\boldsymbol{\chi}}}}\left(\frac{\delta^{n_{\boldsymbol{\chi}^1}}\boldsymbol{\chi}}{n_{\boldsymbol{\chi}^1}!}\right)^{p_{\boldsymbol{\chi}^1}}\cdots\left(\frac{\delta^{n_{\boldsymbol{\chi}^{d_{\boldsymbol{\chi}}}}}\boldsymbol{\chi}}{n_{\boldsymbol{\chi}^{d_{\boldsymbol{\chi}}}}!}\right)^{p_{\boldsymbol{\chi}^{d_{\boldsymbol{\chi}}}}}\vdots_{\left(p_{\chi}^{+}\right)}\frac{\mathcal{D}^{p_{\chi}^{+}}\delta^{n_{\mathcal{P}}}\mathcal{P}^{\boldsymbol{k}}}{p_{\chi}^{+}!n_{\mathcal{P}}!}\\=&\sum_{\substack{(n_{\boldsymbol{\theta}i}),(p_{\boldsymbol{\theta}i}),n_{\boldsymbol{\chi}}\text{ such that}\\ \sum_{i=1}^{d_{\boldsymbol{\theta}}}n_{\boldsymbol{\theta}i}p_{\boldsymbol{\theta}i}+n_{\boldsymbol{\chi}}=n}}\binom{p_{\theta}^{+}}{p_{\boldsymbol{\theta}_1},\ldots,p_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}}\left(\frac{\boldsymbol{k}*\delta^{n_{\boldsymbol{\theta}1}}\Delta\boldsymbol{\theta}}{n_{\boldsymbol{\theta}1}!}\right)^{p_{\boldsymbol{\theta}1}}\cdots\left(\frac{\boldsymbol{k}*\delta^{n_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}}\Delta\boldsymbol{\theta}}{n_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}!}\right)^{p_{\boldsymbol{\theta}d_{\boldsymbol{\theta}}}}\vdots_{\left(p_{\theta}^{+}\right)}\frac{\partial_{\boldsymbol{\theta}}^{p_{\theta}^{+}}\delta^{n_{\boldsymbol{\chi}}}\boldsymbol{\chi}}{p_{\boldsymbol{\theta}}^{+}!n_{\boldsymbol{\chi}}!}.\end{aligned}\tag{A6}$$

Appendix B. Formalising the concept of a fragmented set

In this appendix, the line of thought to define a 'fragmented' set is presented, inspired by the key properties of the classical Cantor set. This will be stated in topology language; therefore, it might not be comfortable for plasma control system designers to read, which is a sacrifice for rigour. The goal is not to directly replicate all structural features of the Cantor set, such as perfectness, total disconnectedness, or lack of interior points. Instead, we seek to identify the essential topological feature that captures the notion of 'fragmentation' for the sets much more complicated than the other regular ones. A formal definition of a fragmented set will be given later in this appendix, while a short one is :

> *A subset F of a topological space* (X,τ) *is called fragmented if there exists at least one accumulation point of the set in the connected component quotient space.*

Functions only defined on a fragmented domain, e.g. the classical Cantor set, can not be done calculus with the usual derivative, which is a strong motivation for proposing a weaker version such as the one proposed at the beginning of this paper.

The Cantor set is often remarkable for its pathological properties: it is closed, perfect (no isolated points), totally disconnected, nowhere dense, and has zero Lebesgue measure. While these properties characterise a highly irregular 'fractal' subset of the real line, we must carefully consider which features are truly essential when generalising to higher dimensions and to sets arising from complicated scenarios in dynamical systems.

In high-dimensional phase spaces, especially those of non-integrable systems, the union of invariant tori may form extremely complicated patterns. Some parts of the phase space may be shattered into a 'dust' or 'debris' (e.g. Cantorus) of invariant objects separated by chaotic regions, creating a labyrinth of intermingled structures. These pieces can be so unstructured that even one of them does not have a uniform dimensionality, which makes it difficult and less worthwhile to define them rigorously.

Requiring that the set $F\subseteq\mathbb{R}^N$ be nowhere dense or perfect in the classical sense can be overly restrictive. Instead, focusing on connected components proves more suitable: each connected piece of F becomes effectively 'collapsed' to a point in the quotient component space, independent of how complicated its dimensionality might be within F. This approach accommodates a wide range of topological complexities, aligning better with realistic dynamical system scenarios.

B.1. Defining a fragmented set through connectivity

B.1.1. Step 1: Connected components and the component space. Let (X,τ) be a topological space and F a subset of X. By introducing an equivalence relation $x\sim y$ if and only if x and y lie in the same connected component, one can define the connected component quotient space:

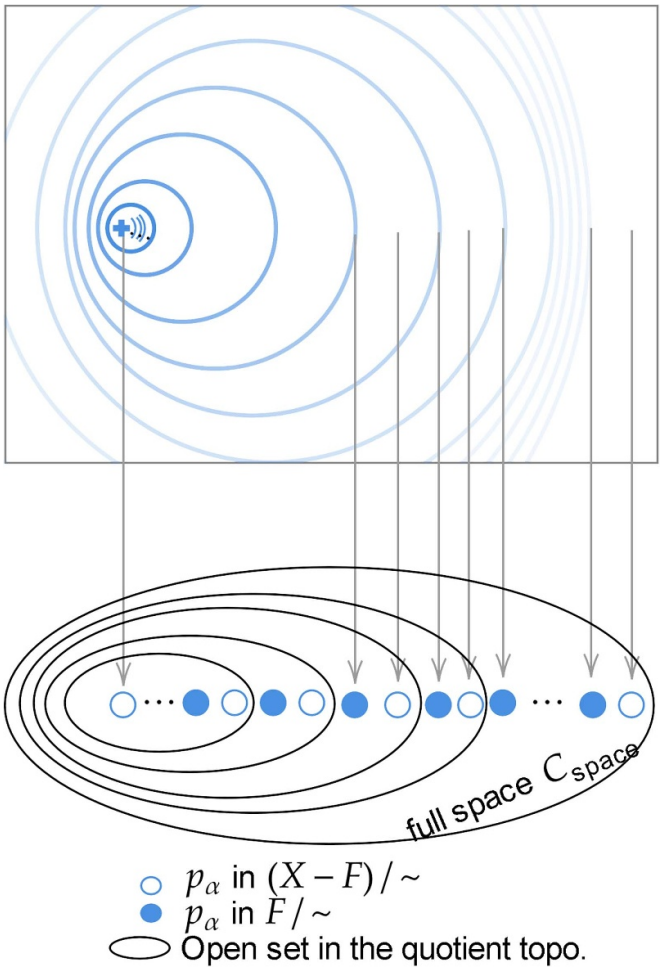

Figure 6. Geometric diagram to show the topology of $\mathbb{R}^2$ partitioned by a fragmented set F that consists of an infinite sequence of circles.

$$\mathcal{C}_{\text{space}} = F/\sim \ \cup \ (X\setminus F)/\sim, \tag{B1}$$

which is to 'collapse' each connected component C_α into a single point p_α. $\mathcal{C}_{\text{space}}$ thus represents the 'skeleton' of X partitioned by F at the level of connectivity. Let $\{C_\alpha\}_{\alpha\in A}$ be the collection of all connected components of F and $X\setminus F$, where α is the index and A is the index set. The notation p_α is used when the component needs to be identified to a point in the topology of $\mathcal{C}_{\text{space}}$.

B.1.2. Step 2: The topology on the component space. The $\mathcal{C}_{\text{space}}$ is not acquired simply by once quotient mapping but plus a union of $F/\sim$ and $(X\setminus F)/\sim$, yet it can also be considered to be acquired by merely once quotient mapping via replacing the equivalence relation by the connectivity of X partitioned by F. In other words, the quotient map could be defined according to the connectivity in F and $X\setminus F$ instead of simply X. Formally, let $q: F\to\mathcal{C}_{\text{space}}$ be the quotient map defined by:

$$q(x) = p_\alpha \quad \text{if} \quad x\in C_\alpha, \text{a connected component of } F \text{ or } X\setminus F.$$

Then, a subset U of $\mathcal{C}_{\text{space}}$ is open in $\mathcal{C}_{\text{space}}$ precisely when $q^{-1}(U)$ is open in X, which is now consistent with the standard definition of the quotient topology. All such open subsets of $\mathcal{C}_{\text{space}}$ are collected to be the quotient topology τ_q equipped to the quotient space

$$\tau_q := \left\{u\subseteq\mathcal{C}_{\text{space}} |\ q^{-1}(U)\in\tau\right\}.$$

The quotient topology τ_q on $\mathcal{C}_{\text{space}}$ reflects the arrangement and clustering of connected components in X partitioned by F. By employing the quotient topology, the connectivity structure of F and $X\setminus F$ is abstracted into $\mathcal{C}_{\text{space}}$ without imposing additional metric constraints.

Example (nested circles in $\mathbb{R}^2$): Consider $F\subseteq\mathbb{R}^2$ composed of infinitely many disjoint circles accumulating to a single point (see figure 6). Here, each circle is a connected component of F. The component space $\mathcal{C}_{\text{space}}$ consists of points corresponding to

each circle. An open neighbourhood around the accumulation point in $\mathcal{C}_{\text{space}}$ includes points corresponding to circles arbitrarily close to the accumulation circle in F. This example illustrates how the quotient topology captures the clustering behaviour of connected components.

B.1.3. Step 3: The essential property—existence of at least one accumulation point. The classical Cantor set is perfect by definition because it is closed and every point in it is an accumulation point. This perfectness requirement is too strict to allow a fragmented set to contain 'regular islands' but still harbour a 'Cantor-like dust' elsewhere. Instead of perfectness, merely one accumulation point in the connected component space suffices to distinguish the set from other regular ones in topological complexity, which leads to a formal definition of a fragmented set:

Definition (fragmented):
Let (X,τ) be a topological space and F a subset of X. The set F is called fragmented if and only if there exists at least one accumulation point of $F/\sim$ in the quotient connected component space

$$\mathcal{C}_{\text{space}} = F/\sim \ \cup \ (X\setminus F)/\sim,$$

where $\sim$ is the equivalence relation defined by connectivity under the topology τ.

Note that $F/\sim$ is a subset of $\mathcal{C}_{\text{space}}$, so this definition is equivalent to whether there exists an accumulation point of F in X in the sense that both of them are lifted to the quotient space. In other words, the condition can also be stated as $F/\sim$ is an accumulation-point-existing subset of $\mathcal{C}_{\text{space}}$, no matter whether the accumulation point is in or not in $F/\sim$ itself.

This single topological requirement captures the essence of a Cantor-like fragmentation: there is at least one location in $F/\sim$ where the connected components cluster endlessly, preventing it from being described as a finite simple union of regular sets. Even if F contains regular subsets that fill some other regions, the presence of at least one such accumulation point of components indicates a fragmented structure lurking beneath.

In this way, a topological criterion for fragmentation is distilled from the Cantor set's properties suitable for generalising to complicated, high-dimensional scenarios. This definition captures the essence of fragmentation without forcing all classical Cantor-like attributes (such as total disconnectedness or lack of interior) onto F. Instead, it generalises the concept in a way that applies naturally to complicated sets arising in dynamical systems, thus providing a robust notion of 'fragmented' beyond 'Cantor-like' in the classical one-dimensional setting.

Examples (common fragmented sets):

- $F=\left\{\dfrac{1}{n}\right\}_{n\in\mathbb{N}^+}$ as a subset of $\mathbb{R}$ with the standard topology.
- $F=\mathbb{Q}$ as a subset of $\mathbb{R}$ with the standard topology.
- $F=\mathbb{R}\setminus\mathbb{Q}$ as a subset of $\mathbb{R}$ with the standard topology.
  Always keep in mind that, even though $\mathbb{R}\setminus\mathbb{Q}$ and $\mathbb{Q}$ are both fragmented, their union $\mathbb{R}$ is not fragmented because the gaps in either $\mathbb{Q}$ and $\mathbb{R}\setminus\mathbb{Q}$ are filled by the union with the other.
- The classical Cantor set as a subset of $\mathbb{R}$ with the standard topology.
- Sierpinski–Menger sponge, a generalisation of Cantor set to 3D space, as a subset of $\mathbb{R}^3$ under the standard topology.
  Note that Sierpinski–Menger sponge is not fragmented because it is merely one connected component.
- $F=\left\{C_r \mid r\in\mathbb{Q}\right\}$, where C_r is the circle centred at the origin with radius r.
  F is fragmented as a subset of $\mathbb{R}^2$ under the standard topology. Note that $\left\{C_r \mid r\in\mathbb{Z}\right\}$ is not fragmented because the connected component containing the origin point is open, thereby impossible to be an accumulation point.
- $F=\left\{\mathcal{T}^2(q) \mid \text{the safety factor } q\in\mathbb{Q}\right\}$, the set of all rational tori in an MCF machine, as a subset of $\mathbb{R}^3$ under the standard topology.

  The set of all irrational tori is also fragmented. When lifted to the quotient space, both sets have the magnetic axis $\mathcal{T}^1_{\text{ax}}/\sim$ as an obvious accumulation point, while each invariant torus $\mathcal{T}^2(q)/\sim$ itself is also an accumulation point. Extreme counter-examples where the magnetic field or its poloidal component vanishes are disregarded due to a lack of practical value.

B.2. Abandoning classical properties not essential to fragmentation

No interior points: the classical Cantor set is nowhere dense and contains no intervals. In high-dimensional scenarios, talking about interior points of F is less pertinent. Our construction bypasses this by working at the level of connected components because the existence of interior points in a regular subset of F does not impede the set F from being fragmented elsewhere. The fragmented property is infectious: *If F has a subset F_{sub} that is fragmented, then F will inherit being fragmented, provided the complement $F - F_{\text{sub}}$ is not large enough to bury all the accumulation points of $F_{\text{sub}}/\sim$ in the component space.* A formal description can be found in the next section of this appendix.

Dimension or smoothness assumptions: a fragmented set defined here is not required to have uniform dimensionality or a differentiable structure. The connected components can be of varying complexity, dimensionality, or smoothness.

Self-similarity: fractals—a well-known term coined by Benoit Mandelbrot in 1975—are often defined or understood through self-similarity (exact or statistical), scaling laws, and non-integer Hausdorff dimensionality. The Cantor set, for instance, is constructed by iterative removal of intervals and exhibits a clear self-similar structure. While fragmented sets may indeed share some of these traits, it is not a requirement. By contrast, fragmentation is a topological notion emphasising the distribution and complexity of connected components.

B.3. Infectivity of the fragmented property

Consider a topological space (X,τ) and subsets $F_{\text{sub}} \subseteq F \subseteq X$. If F_{sub} is fragmented, there is an accumulation point p of $F_{\text{sub}}/\sim$ in the quotient space $\mathcal{C}^{F_{\text{sub}}}_{\text{space}} := (F_{\text{sub}}/\sim) \cup ((X - F_{\text{sub}})/\sim)$. We can approximate p by a sequence (p_i) with each $p_i \in F_{\text{sub}}/\sim$.

Since F_{sub} is included in F, there is a natural map

$$f : F_{\text{sub}}/\sim \to F/\sim$$

sending each connected component of F_{sub} to the (possibly larger) connected component of F containing it. If, for every sequence that approximates p, all but finitely many terms (p_i) land in a single, larger connected component of $F/\sim$, then p ceases to be an accumulation point when viewed in $\mathcal{C}^{F}_{\text{space}} := (F/\sim) \cup ((X - F)/\sim)$. In this case, we say that the accumulation point p of $F_{\text{sub}}/\sim$ has been *buried* upon extending from F_{sub} to F.

If not all approximation sequences are buried, then F inherits the fragmented nature of F_{sub}. We say that F is *infected* with fragmentation or *inherits* fragmentation. This infectivity underlies the ability to produce highly complex fragmented sets by judiciously choosing unions that prevent burying.

On the other hand, if not all approximation sequences are buried this way, then F inherits the fragmented nature from F_{sub}—in other words, F becomes 'infected' and is thus fragmented.

$$\begin{array}{lcl} \mathcal{C}^{F_{\text{sub}}}_{\text{space}} \text{ partitioned by } F_{\text{sub}} & & \mathcal{C}^{F}_{\text{space}} \text{ partitioned by } F \\ \| & & \| \\ F_{\text{sub}}/\sim & \xrightarrow{f} & F/\sim \\ \cup & & \cup \\ (X \setminus F_{\text{sub}})/\sim & \leftarrow & (X \setminus F)/\sim \end{array}$$

It would be helpful to identify a condition on the sets such that their union always inherits the fragmented property without burying each other's accumulation points. The condition is separation.

Proposition (infectivity of fragmentation):
Let (X,τ) be a topological space.

1. (Set union) Suppose $A, B \subset X$ are separated subsets of X. If A is fragmented, then $A \cup B$ is also fragmented.
2. (Set minus) Suppose $C_{\text{sub}} \subset C$ are two subsets of X. If C_{sub} is fragmented while C is non-fragmented, then $C \setminus C_{\text{sub}}$ is fragmented.

Proof. *(Set union)* Since A is fragmented, by definition there exists at least one accumulation point of $A/\sim$ in the corresponding connected component quotient space

$$\mathcal{C}^{A}_{\text{space}} := A/\sim \cup (X \setminus A)/\sim,$$

where $\sim$ is the equivalence relation induced by connectedness under τ. Let p be such an accumulation point.

Consider now

$$\mathcal{C}_{\text{space}}^{A\cup B} := (A\cup B)/\sim \cup (X\setminus(A\cup B))/\sim .$$

Since A and B are separated, there exist disjoint open sets U and V such that $A\subseteq U$ and $B\subseteq V$. This topological separation ensures that no connected component arising from A is forcibly merged or coalesced into a larger connected component due to the union with B, or vice versa, when passing to $(A\cup B)/\sim$. Intuitively, the structure of connected components within A is not 'buried' by the union with B.

Formally, if p is an accumulation point of $A/\sim$ in $\mathcal{C}_{\text{space}}^{A}$, then there exists a sequence (p_i) in $A/\sim$ approximating p. The introduction of B, being separated from A, does not allow the connected components corresponding to (p_i) to collapse into a single larger component that would eliminate p's status as an accumulation point. Thus, p remains an accumulation point in $\mathcal{C}_{\text{space}}^{A\cup B}$.

Therefore, $A\cup B$ is also fragmented.

(Set minus) Only when $C\setminus C_{\text{sub}}$ is fragmented can it merge the connected components approximating the accumulation point in the quotient space partitioned by C_{sub} such that their union is non-fragmented. □

Next, a stronger proposition of infectivity for a family of infinite subsets is presented, but it needs a stronger separation condition than pairwise separation to ensure an accumulation point of one subset will not be buried by the countable or uncountable union. To illustrate it is a stronger condition, consider the family of singletons $\{\{x\} : x\in[0,1]\}$ on $\mathbb{R}$. Each pair of singletons is separated by open sets that isolate the corresponding points. However, the union is simply the interval $[0,1]$, which is connected, and hence corresponds to a single point in the quotient space $\mathcal{C}_{\text{space}}$.

Definition (global separation condition):
A family of subsets $\{A_\alpha\}_{\alpha\in I}$ of a topological space (X,τ) is said to be *globally separated* if for any partition of I into two disjoint subsets I_1 and I_2, there exist disjoint open sets U and V in (X,τ) such that

$$\bigcup_{\alpha\in I_1} A_\alpha \subseteq U \quad \text{and} \quad \bigcup_{\alpha\in I_2} A_\alpha \subseteq V.$$

This global separation is a stronger condition than the pairwise separation for a family of subsets, ensuring that no topological complexity arising from the union of infinitely many subsets collapses into a single connected component.

Proposition (infectivity of fragmentation under global separation):
Let (X,τ) be a topological space and let $\{A_\alpha\}_{\alpha\in I}$ be a globally separated family of subsets of X. If at least one of the A_α is fragmented, then $\bigcup_{\alpha\in I}A_\alpha$ is also fragmented.

Proof. One can begin with the countable case. Assume $\{A_i\}_{i=1}^{\infty}$ is a countably infinite, separated family. Suppose, without loss of generality, that A_1 is fragmented. By the proposition for two sets, since A_1 and A_2 are separated, $A_1\cup A_2$ is fragmented. Now consider $(A_1\cup A_2)\cup A_3$. Since $\{A_1,A_2,A_3\}$ is separated, the fragmentation persists, and thus $(A_1\cup A_2\cup A_3)$ is also fragmented. Proceeding inductively, if at least one A_n is fragmented, the entire union $\bigcup_{i=1}^{\infty}A_i$ remains fragmented.

For an uncountable family $\{A_\alpha\}_{\alpha\in I}$, one can use a transfinite induction. Enabled by Zermelo-Fraenkel set theory with the axiom of choice, one can well-order the index set I and mimic the countable construction at each successor step, and take limits at limit ordinals. The separated structure of the family ensures that no accumulation points become 'buried' as one takes larger and larger unions. Thus, if at least one A_α is fragmented, the entire union $\bigcup_{\alpha\in I}A_\alpha$ is fragmented.

In both the countable and uncountable cases, the key is that the notion of separation extends to the entire family, preventing unwanted mergers of connected components in the quotient space. Hence, the infectivity of fragmentation is not limited to single pairs of sets but holds for arbitrarily large separated families of sets. □

B.4. Fragmented sets exist universally

Let (X,τ) be an infinite topological space. Consider the power set 2^X of X, which has cardinality $2^{|X|}$. By definition, every subset $A\subseteq X$ is either fragmented or non-fragmented. Thus, 2^X is partitioned into two disjoint classes:

$$2^X = \mathcal{F}\cup\mathcal{N},$$

where $\mathcal{F}$ is the family of all fragmented subsets and $\mathcal{N}$ is the family of all non-fragmented subsets. Clearly, $|\mathcal{F}|+|\mathcal{N}| = 2^{|X|}$ in the sense of cardinal arithmetic.

At least one of $\mathcal{F}$ or $\mathcal{N}$ must have cardinality $2^{|X|}$, otherwise, if both had cardinalities strictly less than $2^{|X|}$, their union could not cover all subsets of X. The question then becomes: under what condition can one guarantee that $\mathcal{F}$ attains this maximal cardinality?

If we can find a large family of fragmented sets—at least $|X|$-many distinct ones—this can provide a foothold to argue that $\mathcal{F}$ may achieve cardinality $2^{|X|}$. Based on these $|X|$-many 'blocks', one can combine them to produce $2^{|X|}$ distinct sets. Keep in mind that for infinite X, $2^{|X|}$ is a strictly greater cardinal than $|X|$. The point is to ensure the union of a combination of these fragmented sets is still fragmented, and that these $2^{|X|}$ sets generated are distinct from each other.

Example on $\mathbb{R}$: Consider the real line $\mathbb{R}$ with the standard topology. For each real number $\alpha \in \mathbb{R}$, define:

$$F_\alpha := \left\{ \alpha + \frac{1}{n} \mid n \in \mathbb{N}^+ \right\}.$$

Each set F_α is a countable infinite subset of $\mathbb{R}$ that accumulates at α. Note that $\alpha \notin F_\alpha$, but α is a limit point of F_α. Each F_α is obviously fragmented by definition. Additionally, if $\alpha \neq \beta$, obviously $F_\alpha \neq F_\beta$. Hence, $\{F_\alpha \mid \alpha \in \mathbb{R}\}$ is a family of $|\mathbb{R}|$-many distinct fragmented sets.

The arbitrary union of these fragmented sets

$$\bigcup_{\alpha \in I} A_\alpha, \text{for } \mu\text{-almost all } I \subset \mathbb{R},$$

is still fragmented. μ-almost is to be explained in the proof of the following theorem where μ is a measure on the infinite-dimensional subset space $2^{\mathbb{R}}$ that consists of all subsets of $\mathbb{R}$. Each F_α accumulates at α and thus is fragmented. However, the family $\{F_\alpha\}_{\alpha \in \mathbb{R}}$ is not pairwise separated because they partially overlap with each other. We want to modify each F_α slightly by *removing some points* so that the resulting family $\{G_\alpha\}_{\alpha \in \mathbb{R}}$ is still fragmented and also pairwise separated. Since each F_α is infinite, we can afford to remove infinitely many points and still maintain accumulation at α. This yields a family $\{G_\alpha\}_{\alpha \in \mathbb{R}}$ of cardinality $|\mathbb{R}|$ that is pairwise separated and still entirely composed of fragmented sets. Hence, this construction provides a direct, explicit example suitable for applying directly the following theorem on cardinality dominance of fragmented sets.

Theorem (fragmentation is generic):
Let (X, τ) be an infinite topological space. If X admits a family $\{F_\alpha\}_{\alpha \in I}$ of pairwise-separated fragmented subsets with cardinality

$$|I| \geqslant |X|,$$

then the family $\mathcal{F}$ of all fragmented subsets of X has cardinality

$$|\mathcal{F}| = 2^{|X|}.$$

Proof. The proof is broken down into several conceptual steps, combining intuitive ideas with rigorous argumentation.

Step 1 (Measure-theoretic setup):

Consider the product space

$$\Omega := \{0,1\}^I,$$

equipped with the product σ-algebra generated by cylinder sets and the product measure μ defined as follows: Let μ_α be the measure on $\{0,1\}$ assigning measure $1/2$ to each number. By Kolmogorov's extension theorem, one can obtain a complete product measure μ on Ω:

$$\mu = \bigotimes_{\alpha \in I} \mu_\alpha.$$

Each element $\omega \in \Omega$ corresponds to an index subset $J_\omega \subseteq I$ by

$$J_\omega = \{\alpha \in I \mid \omega(\alpha) = 1\},$$

which is the natural identification bijection $J_\omega : \Omega \ni \omega \mapsto J_\omega \in 2^I$. In this way, 2^I is measured by μ: a measurable set $\mathcal{E} \subseteq 2^I$ is identified with the corresponding set

$$E := \{\omega \in \Omega \mid J_\omega \in \mathcal{E}\}$$

in Ω, and we say $\mathcal{E}$ is a null set (or has measure zero) if $\mu(E) = 0$.

Henceforth, one can speak in a language of real analysis 'for μ-almost all subsets $J \subseteq I$, a property $P(J)$ holds,' which means that the collection of J for which $P(J)$ fails is a μ-null set.

Step 2 (Combinatorial explosion from $|X|$ to $2^{|X|}$):

Consider all subsets $J \subseteq I$, each of which form the union

$$\cup_{\alpha\in J} F_\alpha.$$

The mapping $J \mapsto \cup_{\alpha\in J} F_\alpha$ is an injection from the power set of I (which has cardinality $2^{|X|}$) into the family of all fragmented subsets $\mathcal{F}$.

Consider first a finite collection $\{\alpha_1, \ldots, \alpha_n\}$ of indices. By applying the infectivity of fragmentation, we know that:

$$F_{\alpha_1} \cup F_{\alpha_2} \cup \cdots \cup F_{\alpha_n}$$

remains fragmented. This is a deterministic consequence of pairwise separation and the infectivity proposition.

Step 3 (For μ-almost all index subsets, the union is still fragmented):

Only when it is an infinite union can the union set be threatened to lose fragmentation. The index set $J \subseteq I$ has to be chosen carefully so that their union 'glues together' infinitely many components in a way that buries all accumulation points. However, such glueing requires picking indices in a highly structured and intricate pattern. The measure μ assigns measure zero to these 'rare' patterns because they form a measure-null subset of 2^I. According to the infectivity of fragmentation during set minus operation, any removal of one subset F_α will cause the union to become fragmented again. Removal of finitely and infinitely many subsets is much more possible in the sense of measure theory, which also lets the union cease being fragmented.

To be more explicit, the set of $J \subseteq I$ that fails to preserve fragmentation lies within certain null sets defined by infinitely many measure-zero conditions (akin to requiring infinitely many 'coincidences' or 'matches' of topological conditions). By the completeness and translation-invariance properties of the product measure μ, any such topologically contrived selection pattern has zero measure. Hence, the complement of that collection—the 'generic' subset of 2^I—has full μ-measure.

In other words, for μ-almost all subsets $J \subseteq I$, the union $\bigcup_{\alpha\in J} F_\alpha$ is fragmented. This, combined with the combinatorial explosion, completes the proof. □

There is no need for metrics or additional structures. The infectivity of fragmentation under separation and the combinatorial explosion from $|X|$-many to $2^{|X|}$-many subsets are topological and set-theoretic in nature.

The theorem shows a stark contrast between fragmented and non-fragmented sets: the non-fragmented ones cannot similarly leverage separation and infectivity arguments. Therefore, fragmented sets are usually expected to dominate in terms of cardinality.

If no such initial family $\{F_\alpha\}$ exists, then one cannot conclude that fragmented sets dominate. It could be the case that the family $\mathcal{N}$ of non-fragmented sets achieve the cardinality $2^{|X|}$ or both. However, under a wide range of natural topological spaces (for instance, in spaces resembling $\mathbb{R}$, such as $\mathbb{Q}$ and $\mathbb{R}^N$, where one can easily produce infinitely many distinct fragmented sets), it is quite possible that the family of fragmented sets has a cardinality comparable to the family of all subsets, as stated in the theorem.

For comparison, in a discrete space like $\mathbb{Z}$, every set does not produce accumulation points. Thus, every set is non-fragmented. Such spaces fail to provide the vast collections needed to exhibit fragmentation-richness. Indeed, $\mathbb{Z}$ is too 'coarse' to embed intricate Cantor-like structures.

Remark. The measure-theoretic argument can be viewed as a formalisation of the notion that 'almost no' selection of indices produces exact merges that erase all accumulation points arising from all subsets. In contrast, global separation guarantees every selection is well-behaved, a far stronger condition than the pairwise-separated one, yet (unnecessarily) making it more difficult to construct the subset family. The measure-theoretic perspective highlights that while global separation ensures the union to inherit fragmentation, pairwise separation combined with fragmentation is sufficient to ensure fragmentation for almost all choices of the index subset to do the union, a slightly weaker but still robust conclusion in a measure-theoretic sense.

Example (A $|\mathbb{R}|$-many family of pairwise separated fragmented sets on $\mathbb{R}$):
The construction is shown in the following steps.

Step 1 (Utilising the Cantor set):

The standard Cantor set

$$\mathcal{C} := \left\{ \sum_{n=1}^{\infty} \frac{a_n}{3^n} \mid a_n \in \{0,2\} \ \forall n \right\}$$

has cardinality $|\mathcal{C}| = 2^{\aleph_0}$ and many characteristics that can help construct the family of pairwise separated fragmented sets. Each point of $\mathcal{C}$ is a limit of others, and $\mathcal{C}$ is totally disconnected.

Well-order $\mathcal{C}$ as $\{c_i : i \in I\}$ with a transfinite index set $|I| = 2^{\aleph_0}$. If one can associate to each c_i a fragmented set G_{c_i} accumulating at c_i and ensure that $\{G_{c_i}\}$ is pairwise separated.

Step 2 (Choosing the intervals that will be used to trim fragmented sets so that they are pairwise separated):

To ensure the family is pairwise separated, one can employ disjoint intervals to trim the fragmented sets. Via transfinite recursion, each c_i will be assigned a closed interval I_{c_i} around it or skipped if it is already covered by previously constructed intervals.

Since $\mathcal{C}$ is totally disconnected, one can choose a closed interval I_{c_i} that contains c_i and is disjoint from all previously chosen intervals, provided it is not already contained by such intervals.

To not exclude too many points that have not been enumerated, choose a sufficiently small length of the interval I_{c_i} such as $3^{-(i+5)}$, where i is viewed as an ordinal index enumerating $\mathcal{C}$. As i grows, these intervals become extremely small, especially when i is large enough that can only be considered as a transfinite number. The exact length $3^{-(i+5)}$ is just an example; the key is that the self-similarity of the Cantor set can be utilised to guarantee that each interval I_{c_i} does not include other points at the same level as c_i in the Cantor construction procedure.

Due to the self-similarity, the remaining subset of $\mathcal{C}$ that neither has been enumerated at staged i nor included by I_{c_i} still has cardinality $2^{\aleph_0}$. The cardinality of the points that would be skipped in the enumeration due to being included by I_{c_i} is controlled to be less than or equal to that of the other part, that is the aforementioned remaining subset. The cardinal arithmetic guarantees that at least one of them has to be $2^{\aleph_0}$, so that an overwhelming majority (still of size $2^{\aleph_0}$) of points remain to continue the process.

Step 3 (Trimming the fragmented subsets):

For each c_i, it is natural to choose the following simple set that accumulates at c_i,

$$F_{c_i} := \left\{ c_i + \frac{1}{n} : n \in \mathbb{N}^+ \right\}.$$

It is obvious that for distinct $i \neq j$, $F_{c_i} \cap F_{c_j}$ might not be empty due to accidental coincidence. To prevent intersection, trim these fragmented sets by the intervals constructed in the second step,

$$G_{c_i} := F_{c_i} \cap I_{c_i}.$$

Since I_{c_i} is an interval already constructed and includes c_i, G_{c_i} remains infinite and accumulates at c_i, so G_{c_i} is still fragmented as F_{c_i}. The family $\{G_{c_i}\}_{i \in I}$ is pairwise separated by the deliberate construction. □

ORCID iDs

Wenyin Wei https://orcid.org/0000-0003-3072-4520
Alexander Knieps https://orcid.org/0000-0003-0083-7188
Shaocheng Liu https://orcid.org/0000-0002-7298-0680
Yunfeng Liang https://orcid.org/0000-0002-9483-6911

References

[1] Chandre C, Vittot M, Ciraolo G, Ghendrih P and Lima R 2005 *Nucl. Fusion* **46** 33
[2] Abdullaev S 2014 *Magnetic Stochasticity in Magnetically Confined Fusion Plasmas: Chaos of Field Lines and Charged Particle Dynamics* (*Springer Series on Atomic, Optical and Plasma Physics*) (Springer) (available at: https://books.google.de/books?id = 92RXvgAACAAJ)
[3] Ali H and Punjabi A 2007 *Plasma Phys. Control. Fusion* **49** 1565
[4] Volpe F , Kessler J, Ali H, Evans T E and Punjabi A 2012 *Nucl. Fusion* **52** 054017
[5] Xu S *et al* 2023 *Nucl. Fusion* **63** 066005
[6] Zhou S *et al* 2022 *Nucl. Fusion* **62** 106002
[7] Loizu J, Hudson S R, Nührenberg C, Geiger J and Helander P 2017 *J. Plasma Phys.* **83** 715830601
[8] Abdullaev S S, Finken K H, Wongrach K, Tokar M, Koslowski H R, Willi O and Zeng L 2015 *J. Plasma Phys.* **81** 475810501
[9] Hudson S R and Nakajima N 2010 *Phys. Plasmas* **17** 052511
[10] Paul E J, Landreman M and Antonsen T 2021 *J. Plasma Phys.* **87** 905870214
[11] Boozer A H 2008 *Plasma Phys. Control. Fusion* **50** 124005
[12] Carlton-Jones A, Paul E J and Dorland W 2021 *J. Plasma Phys.* **87** 905870222
[13] McGreivy N , Hudson S R and Zhu C 2021 *Nucl. Fusion* **61** 026020
[14] Strickler D J , Hirshman S P, Spong D A, Cole M J, Lyon J F, Nelson B E, Williamson D E and Ware A S 2004 *Fusion Sci. Technol.* **45** 15
[15] Bachelard R, Chandre C and Leoncini X 2006 *Chaos* **16** 023104
[16] Wright A M and Ferraro N M 2024 *Phys. Plasmas* **31** 082509
[17] Reiman A H 2021 *Plasma Phys. Control. Fusion* **63** 054002
[18] Weller A *et al* 2003 *Plasma Phys. Control. Fusion* **45** A285
[19] Sakakibara S *et al* 2008 *Plasma Phys. Control. Fusion* **50** 124014
[20] Hudson S R, Monticello D A, Reiman A H, Boozer A H, Strickler D J, Hirshman S P and Zarnstorff M C 2002 *Phys. Rev. Lett.* **89** 275003
[21] Hegna C C 2012 *Phys. Plasmas* **19** 056101
[22] Baillod A, Loizu J, Graves J P and Landreman M 2022 *Phys. Plasmas* **29** 042505
[23] Landreman M, Medasani B and Zhu C 2021 *Phys. Plasmas* **28** 092505
[24] Jorge R, Goodman A, Landreman M, Rodrigues J and Wechsung F 2023 *Plasma Phys. Control. Fusion* **65** 074003
[25] Henneberg S A, Hudson S R, Pfefferlé D and Helander P 2021 *J. Plasma Phys.* **87** 905870226
[26] Zhu C , Gates D A, Hudson S R, Liu H, Xu Y, Shimizu A and Okamura S 2019 *Nucl. Fusion* **59** 126007
[27] Knieps A *et al* 2022 *Nucl. Fusion* **62** 026011
[28] Sauter O 2016 *Fusion Eng. Des.* **112** 633
[29] Kikuchi M, Takizuka T and Furukawa M 2014 *Proc. 12th Asia Pacific Physics Conf. (APPC12)* p 015014
[30] Kikuchi M *et al* 2019 *Nucl. Fusion* **59** 056017

[31] Marinoni A *et al* 2019 *Phys. Plasmas* **26** 042515
[32] Marinoni A *et al* 2021 *Nucl. Fusion* **61** 116010
[33] Happel T, Pütterich T, Told D, Dunne M, Fischer R, Hobirk J, McDermott R M, Plank U and the A S D E X U T 2022 *Nucl. Fusion* **63** 016002
[34] Austin M E *et al* 2019 *Phys. Rev. Lett.* **122** 115001
[35] Thome K E *et al* 2024 *Plasma Phys. Control. Fusion* **66** 105018
[36] Camenen Y, Pochelon A, Behn R, Bottino A, Bortolon A, Coda S, Karpushov A, Sauter O and Zhuang G (the TCV Team) 2007 *Nucl. Fusion* **47** 510
[37] Han W , Offeddu N, Golfinopoulos T, Theiler C, Tsui C K, Boedo J A, and Marmar E S (the TCV Team) 2021 *Nucl. Fusion* **61** 034003
[38] Coda S *et al* 2019 *Nucl. Fusion* **59** 112023
[39] Anand H , Coda S, Felici F, Galperti C and Moret J-M 2017 *Nucl. Fusion* **57** 126026
[40] Howard J E and Hohs S M 1984 *Phys. Rev. A* **29** 418
[41] Harsoula M and Contopoulos G 2018 *Phys. Rev. E* **97** 022215
[42] Wei W , Liang Y 2023 *Plasma Sci. Technol.* **25** 095105
[43] Frerichs H , Schmitz O, Bonnin X, Loarte A, Feng Y, Li L, Liu Y, and Reiter D 2020 *Phys. Rev. Lett.* **125** 155001
[44] Eich T , Sieglin B, Scarabosio A, Fundamenski W, Goldston R J and Herrmann A 2011 *Phys. Rev. Lett.* **107** 215001
[45] Nührenberg C, Boozer A H and Hudson S R 2009 *Phys. Rev. Lett.* **102** 235001
[46] MacKay R S, Meiss J D and Percival I C 1984 *Phys. Rev. Lett.* **52** 697
[47] Meiss J D 1986 *Phys. Rev. A* **34** 2375
[48] Alus O, Fishman S and Meiss J D 2014 *Phys. Rev. E* **90** 062923
[49] Meiss J D 2015 *Chaos* **25** 097602
[50] Hudson S R 2006 *Phys. Rev. E* **74** 056203
[51] Pöschel J 1982 *Commun. Pure Appl. Math* **35** 653
[52] Frigyik B A, Srivastava S and Gupta M R 2008 An introduction to functional derivatives *Technical Report* UWEETR-2008-0001 (Department of Electrical Engineering, University of Washington) (available at: https://vannevar.ece.uw.edu/techsite/papers/refer/UWEETR-2008-0001.html)
[53] Wobig H 1999 *Plasma Phys. Control. Fusion* **41** A159