Physica D 467 (2024) 134274

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

A principle of maximum entropy for the Navier–Stokes equations

Gui-Qiang G. Chen [a,*], James Glimm [b,c], Hamid Said [d]

[a] Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK
[b] Stony Brook University, Stony Brook, NY 11794, USA
[c] GlimmAnalytics LLC, USA
[d] Department of Mathematics, College of Science, Kuwait University, Safat 13060, Kuwait

ARTICLE INFO

Communicated by Ye Zhou

MSC:
28D20
76F02
28C20
76D05
49S05
35A15
70G10
35Q30
37A50

Keywords:
Principle of maximum entropy
Energy-enstrophy surface
Navier–Stokes equations
Fully developed turbulence

ABSTRACT

A principle of maximum entropy is proposed in the context of viscous incompressible flow in Eulerian coordinates. The relative entropy functional, defined over the space of L^2 divergence-free velocity fields, is maximized relative to alternate measures supported over the energy–enstrophy surface. Since thermodynamic equilibrium distributions are characterized by maximum entropy, connections are drawn with stationary statistical solutions of the incompressible Navier–Stokes equations. Special emphasis is on the correspondence with the final statistics described by Kolmogorov's theory of fully developed turbulence.

1. Introduction

We are concerned with the reformulation of the principle of maximum entropy, as established in the context of statistical mechanics and information theory, and its applications to mathematical fluid dynamics. More specifically, the Boltzmann–Gibbs entropy under investigation (*cf.* [1,2]):

$$S(f) = -\int f(u)\log f(u)\,\mathrm{d}\eta(u) \tag{1.1}$$

is maximized under appropriate constraints resulting from physical considerations,[1] where η denotes a reference measure defined over the velocity configuration space. From a mathematical standpoint, these considerations impose constraints on the class of admissible velocity fields u and their associated probability distributions $f(u)$.

The principle of maximum entropy, introduced by Jaynes [3] in 1957, revolves around making the most desirable choice, for given prior data. This choice aims to maximize the entropy functional (1.1). In other words, as expressed in [4], "[T]he [principle of maximum entropy] is the probability assignment that is consistent with the available information but is maximally noncommittal with regard to missing information". From the standpoint of statistical mechanics, it is seen as a method to determine the correct probability distribution at the state of equilibrium, depending heavily on the constraints imposed on the system, *i.e.*, our knowledge of prior or given information. As the principle of maximum entropy provides a rule for assigning probabilities to data, it proves particularly useful in "filling gaps" between otherwise scattered data. Therefore, maximum entropy modeling can be applied to a wide range of problems. We refer to [5] for applications to several areas including statistics, biology, and medicine.

The physics underlying our problem is fundamentally rooted in the first and second laws of thermodynamics. The former signifies the conservation of energy E, while the latter imposes a constraint on admissible processes in terms of the production of entropy S. At a

* Corresponding author.
E-mail addresses: chengq@maths.ox.ac.uk (G.-Q.G. Chen), james.glimm@stonybrook.edu (J. Glimm), hamids@sci.kuniv.edu.kw (H. Said).

[1] In case of a discrete system in which probabilities $f_1, f_2, \dots, f_k$ are assigned to a random variable Y taking values $y_1, y_2, \dots, y_k$, respectively, then the entropy is expressed as $S = -\sum_i f_i \log f_i$.

https://doi.org/10.1016/j.physd.2024.134274
Received 21 February 2024; Received in revised form 20 June 2024; Accepted 21 June 2024
Available online 1 July 2024

micro-level, where each micro-state is equally probable, entropy can be expressed in terms of the Boltzmann factor k_B:

$$S = k_B \log(\Omega)$$

where Ω is the number of states present at a given energy E. Eq. (1.1) can be understood as a generalization of the above formula for entropy S when the micro-states of the system may not be equally distributed.

This formulation of the second law in terms of a single entropy can be rewritten in a perhaps more conventional manner. We restrict our analysis to incompressible isothermal fluids without mixing. As such, we introduce two entropies, S_E for the entropy of the energetic degrees of freedom which is related to the viscous and turbulent dissipation of energy, and S_Z for the entropy of the vortical and enstropic degrees of freedom which is related to the viscous and turbulent dissipation of enstrophy. We assume that both the energetic and the enstrophic degrees of freedom are governed by a common temperature, which can be fixed in the context of the second law of thermodynamics. Consequently, the second law takes the form:

$$\mathrm{d}E = T\mathrm{d}S_E + T\mathrm{d}S_Z.$$

The first of the terms on the right-hand side aligns with the familiar expression from standard thermodynamics, a scenario that typically excludes physics models featuring an enstropic degree of freedom. The inclusion of the second term is necessitated by the much richer physical model under consideration here.

In tandem with the modified or reformulated second law, we can express the Legendre transforms and the definition of entropy as the logarithm of the volume of a constant energy surface, now accounting for two energy sources. Consequently, we represent

$$\mathrm{d}E = \mathrm{d}E_K + \mathrm{d}E_Z$$

as a summation of the kinetic and enstrophic energies, resulting in two terms on the left-hand side of the second law. The kinetic energy E_K aligns with the conventional description in thermodynamics, while E_Z is an additional energy source specific to the effect that the entropy considered is for the simplest of the fluid physics theories possible, namely incompressible with no mixtures.

From a mathematical standpoint, the Boltzmann–Gibbs entropy (or its discrete counterpart) is a special case of what is known as the *relative entropy* first introduced by Kullback and Leibler in 1951 [6] (consult [7] for more details). Consider a probability space $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ and two measures η and μ defined on it, where $\mathcal{B}(\mathcal{H})$ denotes the space of Borel probability measures over space $\mathcal{H}$. Then the relative entropy $\mathcal{S}(\mu|\eta)$ (of μ with respect to η) is defined by

$$\mathcal{S}(\mu|\eta) := \begin{cases} -\int_{\mathcal{H}} \frac{\mathrm{d}\mu}{\mathrm{d}\eta} \log(\frac{\mathrm{d}\mu}{\mathrm{d}\eta})\,\mathrm{d}\eta(u) & \text{if } \mu \ll \eta, \\ -\infty & \text{otherwise.} \end{cases} \tag{1.2}$$

In the case where measure μ is absolutely continuous with respect to η, the Radon–Nikodym derivative $\frac{\mathrm{d}\mu}{\mathrm{d}\eta}$ is in $L^1(\mathrm{d}\eta)$. Once η is chosen as the Lebesgue measure (when $\mathcal{H} = \mathbb{R}^n$), then $\frac{\mathrm{d}\mu}{\mathrm{d}\eta}$ can identified with the probability distribution $f(u)$ in Eq. (1.1). The relative entropy turns into a functional that is seen to be maximized over the space of all probability measures, once we fix the *reference* measure η. Therefore, the maximum entropy problem consists of finding a probability measure μ on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$, depending now on the choice of measure η, such that the relative entropy is maximized under some given information. In the probability space, $-\mathcal{S}(\mu|\eta)$ can be perceived as a measure of "*distance*" between the maximum entropy measure μ and the reference measure η, or any two measures for that matter.

In the context of fluid flow, a theory for entropy maximization was first established in [8], in which a well-defined probability measure acting as a linear functional on the configuration space of the Euler and Navier–Stokes equations, reduced to enforce particle interchange symmetries, was proved to maximize the Boltzmann–Gibbs entropy relative to alternate solutions, restricted to the constant energy surface. In this paper, we establish the existence of a probability measure on the configuration space satisfying a maximum entropy admissibility condition relative to alternate solutions of the Navier–Stokes equations, yet with a restriction to the *energy-enstrophy* surface. The significance of this restriction lies in its connection to the pursuit of establishing a rigorous mathematical foundation for the investigation of turbulent phenomena, a link that we briefly outline in the following section.

1.1. Motivation for entropy maximization

The recent nonuniqueness result for the Navier–Stokes equations [9] suggest that weak solutions may fall short from describing the complex nature of fluid flow. On this basis, a probabilistic description of the flow can serve as an alternate framework for studying the Navier–Stokes and Euler equations, as opposed to classical weak solutions and their Euler limits. In fact, any comprehensive exploration of turbulent flow phenomena must contend with some notion of *averaging*; especially as it relates to formulating the Kolmogorov theory of fully developed turbulence, widely considered to be valid by the practitioners of the field. We refer the reader to [10] for a modern comprehensive account on this topic.

In an endeavor to establish a mathematical theory of turbulence, Foias [11] (see also [12]), inspired by earlier work of G. Prodi, initiated a program to give, among other things, a rigorous formulation of the notion of *statistical solutions* for the Navier–Stokes equations and eventually derive the predictions of Kolmogorov for the energy spectrum and the structure functions [13,14]. The Kolmogorov power laws are formulated in relation to the *average* (long time averages or ensemble averages) of certain physically measurable quantities; hence, by constructing a family of time-dependent and spatially homogeneous probability measures μ_t (over an appropriate Hilbert space) that satisfy the Navier–Stokes equations in some averaged sense.[2] Therefore, Foias, later with Manley, Roas, and Temam [11,12], was able to attach an unambiguous meaning to the notion of an ensemble average.[3] In the context of this program, one main assumption embedded in the derivation of the Kolmogorov power laws is that the global energy *identity* for the Navier–Stokes equations (with normalized mass density and viscosity ν):

$$\frac{1}{2}\|u(t)\|^2 + \nu \int_0^t \|\nabla u(s)\|^2\,\mathrm{d}s = \frac{1}{2}\|u_0\|^2 \tag{1.3}$$

is satisfied when integrated against a (spatially) homogeneous statistical solution μ_t for all $t > 0$. Here u_0 is taken to be the (given) initial velocity of the fluid, *i.e.*, $u(x,0) \doteq u_0$, and $\|\cdot\|$ denotes the L^2_x–norm for simplicity. These solutions, which also remain invariant under rescaling, are called *self-similar* homogeneous statistical solutions to the Navier–Stokes equations.

Leray–Hopf solutions are known to satisfy the global energy *inequality* (obtained by replacing "=" with "≤" in relation (1.3)). Whether identity (1.3) holds in three space dimensions remains one of the major open problems for weak solutions of the Navier–Stokes equations. Incidentally, the same is true for statistical solutions: only an averaged version of the global energy inequality is known to hold.

[2] A closely related theory was introduced by Vishik and Furkisov [15], where statistical solutions exist as probability measures on the set of solution trajectories of the Navier–Stokes equations, *e.g.*, $C^{\mathrm{w}}_t L^2_x$. See also the work of Foias, Rosa, and Temam [16] for a connection between these two notions of solutions.

[3] In fact, Foias et al. proved the convergence of long-time averages $\lim_{t\to\infty} \frac{1}{t}\int_0^t \Phi(u(t))\,\mathrm{d}t$ for a suitable class of functions Φ and the weak solutions u of the Navier–Stokes equations. See [12] for the precise statement and more details.

In other words, these claims infer that a probabilistic description consistent with Kolmogorov's spectrum needs to be supported over the set

$$G_e(t) \doteq \left\{ u \;:\; \frac{1}{2}\|u(t)\|^2 + \nu \int_0^t \|\nabla u(s)\|^2 \,\mathrm{d}s = e(t) \right\} \tag{1.4}$$

where the prescribed energy $e(t)$ is taken to be the energy profile for the initial velocity. However, since the existence and uniqueness of such solutions remain in question, energy $e(t)$ can be prescribed *a priori* to ensure that $G_e(t)$ is nonempty. Throughout this paper, we assume that, for each $t \geq 0$, $G_e(t)$ is defined for any prescribed positive energy $e(t)$; such a set $G_e(t)$ is called the *energy-enstrophy surface*.

A similar challenge arises when constructing numerical solutions. It is known that statistical solutions of the incompressible Euler equations and multidimensional systems of conservation laws can be computed. However, due to the lack of the well-posedness of these problems and the lack of convergence of their numerical approximations, it is not known whether these solutions are entropy or entropy maximizing solutions.[4] See [17,18] and references therein for more details.

Next, we demonstrate how the principle of maximum entropy provides a systematic approach for determining probability distributions in the presence of constraints on the mean energy of the system.

1.2. Principle of maximum entropy: classical formulation

As indicated above, the principle of maximum entropy can be formulated in a number of different ways depending on modeling considerations. The formulation in the context of equilibrium statistical mechanics, which is presented here, most resembles the application of the principle for the purposes of this work. Consider the entropy functional (1.1) together with the two constraints:

$$\int g(u)\,\mathrm{d}\xi(u) = 1, \qquad \int g(u)\,\frac{m\,u^2}{2}\,\mathrm{d}\xi(u) = \varepsilon, \tag{1.5}$$

where $\frac{m\,u^2}{2}$ is the kinetic energy of a single particle as measured in a fixed frame, ε is the mean energy per particle which is *a priori* known, and measure ξ is assumed to exist over the configuration space consisting of all velocity fields.[5] Then the principle of maximum entropy asserts that the physical entropy achieves its maximum, denoted by $g^*(u)$, for the constrained system. Thus, we may apply the methods of the calculus of variations to functional (1.1) subject to constraints (1.5) to conclude that there exist two Lagrange multipliers α and β satisfying

$$\int \frac{\delta L}{\delta g}(g(u);\alpha,\beta)\,\phi(u)\,\mathrm{d}\xi(u)\bigg|_{g=g*} = 0$$

for any test function $\phi(u)$, where

$$L(g(u);\alpha,\beta) = -g(u)\log g(u) + \alpha(1-g(u)) + \beta(\varepsilon - g(u)\,\frac{m\,u^2}{2}).$$

Therefore, we obtain

$$-\int \phi(u)\Big(\big(1+\log g^*(u)\big) + \alpha + \beta\frac{m\,u^2}{2}\Big)\,\mathrm{d}\xi(u) = 0. \tag{1.6}$$

If the integrand is assumed to be sufficiently regular, then

$$\big(1+\log g^*(u)\big) + \alpha + \beta\frac{m\,u^2}{2} = 0$$

for all u, so that

$$g^*(u) = c_\alpha \exp(-\beta\frac{m\,u^2}{2}), \tag{1.7}$$

[4] Incidentally, as is mentioned in [17], the lack of rigorous convergence of numerical schemes may be attributed to the emergence of turbulent-like structures at finer scales as the mesh is refined.

[5] In the finite-dimensional setting, ξ is taken to be the Lebesgue measure, and the configuration space consists of all possible velocity components of a single particle: $u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$.

where $c_\alpha = \exp(-(1+\alpha))$. The two constraints (1.5) can be used to find parameters α and β in terms of the known mean energy ε. One can even go a step further and determine the Lagrange multipliers in terms of more fundamental thermodynamics quantities by employing the first law of thermodynamics [19]. In any case, we obtain the well-known Maxwell–Boltzmann distribution describing the thermal equilibrium velocity distribution through the principle of maximum entropy.

In general, entropy maximization methods can also be applied for systems outside their equilibrium states such as in problems of rarefied gas dynamics, where the mean free path is large, compared to the system's characteristic length scales [20]. The fact that the principle of maximum entropy produces a probability distribution under one or multiple constraints implies an overarching approach for finding constitutive relations for general balance-type laws. Indeed, once the maximum entropy probability distribution is determined, explicit expressions of the mechanical and thermodynamic fluxes in terms of the relevant densities, now imposed as constraints and appear as the moments of the probability distribution (*e.g.*, Eq. $(1.5)_2$), can be obtained. The constitutive relations are then substituted into the moments of the Boltzmann equation to close the system. It should also be noted that both the Euler and Navier–Stokes equations can be recovered from the moment equations by presuming the velocity distribution at the equilibrium state or small deviations from the velocity distribution at the equilibrium state, respectively. See [20–22] for more details on non-equilibrium maximum entropy methods and moment methods with applications to rarefied gases.

In terms of the context of this work, such phenomena will not be considered as the Euler and Navier–Stokes-like energy identities are assumed to hold. The maximum entropy probability measure μ_t (whose support is the *energy-enstrophy surface* $G_e(t)$) obtained for the problem of fluid flow should in fact be associated with a universal law of decay in the limit: $t \to \infty$. This is indeed the case for self-similar homogeneous statistical solutions constructed: the average kinetic energy scales like $\frac{M}{t}$ as $t \to \infty$, where constant $M = M(\nu)$ is independent of any initial conditions (*cf.* [12]).

As will be illustrated in Section 2, a version of the energy constraint similar to (1.5) will be presented, but not explicitly: it is included in terms of the support of the maximum entropy probability measure, that is to say over set $G_e(t)$. This choice is motivated by the role that the global (averaged) energy identity for self-similar homogeneous statistical solutions plays in producing the predictions of Kolmogorov's theory of turbulence.

2. Entropy maximization and the physical measure

In this section, we establish the principle of maximum entropy in the context of incompressible flow. We begin first by setting out the notation and formulating the problem.

2.1. Principle of maximum entropy: formulation for the Navier–Stokes equations

The fixed time particle configuration space is the Hilbert space $\mathcal{H}$ of L^2 divergence-free velocity fields defined on a cube $V \subset \mathbb{R}^3$ and satisfying periodic boundary conditions on the boundary of V. For every $t \in [0,T]$ (fixed), define the Banach space X_t to be the space $L_t^\infty L_x^2 \cap L_t^2 H_x^1$ endowed with the norm:

$$\|u\|_t^2 \doteq \frac{1}{2}\|u(t)\|^2 + \nu \int_0^t \|\nabla u(s)\|^2 \mathrm{d}s,$$

where $\|\cdot\|$ denotes the L_x^2-norm. Throughout this section, given a fixed $t \geq 0$, we fix a scalar $e = e(t) \geq 0$ to define the energy-enstrophy surface:

$$G_e(t) = \left\{ u \in X_t \;:\; \|u\|_t^2 = e(t) \right\} \qquad \text{for each } t \in [0,T], \tag{2.1}$$

and make the identification

$$G_e(0) = \{u_0 \in \mathcal{H}(V) \;:\; \|u\|_0 = \frac{1}{2}\|u_0\|^2 = e(0)\}. \tag{2.2}$$

It is clear that $G_e(t)$ and $G_e(0)$ are non-empty. For each $t \geq 0$ and non-negative $e(t)$, a given physical probability measure $\eta_{e,t}$ is supported over $G_e(t)$. For measures $\eta_{e,t}$ and $\mu_{e,t}$ belonging the probability space $\mathcal{M} := (\mathcal{H}, \mathcal{B}(\mathcal{H}))$, we define

$$\mathcal{S}(\mu_{e,t}|\eta_{e,t}) := \begin{cases} -\int_{\mathcal{H}} \frac{\mathrm{d}\mu_{e,t}}{\mathrm{d}\eta_{e,t}} \log(\frac{d\mu_{e,t}}{d\eta_{e,t}})\,\mathrm{d}\eta_{e,t}(u) & \text{if } \mu_{e,t} \ll \eta_{e,t}, \\ -\infty & \text{otherwise}, \end{cases} \tag{2.3}$$

which is known as the *relative entropy* functional of measure $\mu_{e,t}$ with respect to the reference probability measure $\eta_{e,t}$. Since $\log(y)$ is a concave function and $f_{e,t}(u) := \frac{\mathrm{d}\mu_{e,t}(u)}{\mathrm{d}\eta_{e,t}(u)} \in L^1(\eta_{e,t})$, we employ Jensen's inequality to obtain

$$\begin{aligned} \mathcal{S}(\mu_{e,t}|\eta_{e,t}) = \mathcal{S}(f_{e,t}) &= \int_{\mathcal{H}} \log(\frac{1}{f_{e,t}(u)})\,\mathrm{d}\mu_{e,t}(u) \leq \log(\int_{\mathcal{H}} \frac{1}{f_{e,t}(u)}\,\mathrm{d}\mu_{e,t}(u)) \\ &\leq \log(\int_{\mathcal{H}} \mathrm{d}\eta_{e,t}(u)) = \log(\eta_{e,t}(G_e(t))) = 0. \end{aligned}$$

We postpone the discussion on the measure $\eta_{e,t}$ over the energy-enstrophy to the next section.

Our problem consists of maximizing the relative entropy functional (2.3) subject to some appropriate constraints: the solution obtained is a probability distribution

$$\int f_{e,t}(u)\,\mathrm{d}\eta_{e,t}(u) = \int \mathrm{d}\eta_{e,t}(u) = 1, \tag{2.4}$$

and the support of the resulting measure must be supported on $G_e(t)$. For the latter constraint, it will be shown that the fact that $\eta_{e,t}$ is supported on $G_e(t)$ is sufficient to ensure the second constraint is satisfied by the maximum entropy measure. We hence require a solution to the following maximum problem:

$$\sup\{\mathcal{S}(f_{e,t}) \;:\; \mu_{e,t} \in \mathcal{M} \text{ satisfying (2.4)}\} \tag{2.5}$$

for each fixed $t \in [0,T]$ and *a priori* prescribed $e(t)$.

2.2. Main theorem for the principle of maximum entropy

The following basic lemma is key in establishing our main result – Theorem 2.2 – below.

Lemma 2.1. *Fix $t \in [0,T]$ and $e(t) \geq 0$. Then the following statements hold*:

(i) *The relative entropy is non-positive*: $\mathcal{S}(f_{e,t}) \leq 0$ *for all* $\mu_{e,t} \in \mathcal{M}$, *where* $f_{e,t} = \frac{\mathrm{d}\mu_{e,t}}{\mathrm{d}\eta_{e,t}}$.

(ii) *The relative entropy is concave in* $\mu_{e,t}$: *for* $\mu^1_{e,t}, \mu^2_{e,t} \in \mathcal{M}$ *so that* $\mathcal{S}(\mu^1_{e,t}|\eta_{e,t})$ *and* $\mathcal{S}(\mu^2_{e,t}|\eta_{e,t})$ *are finite, then*

$$\mathcal{S}(\mu_{e,t}|\eta_{e,t}) \geq \alpha\mathcal{S}(\mu^1_{e,t}|\eta_{e,t}) + (1-\alpha)\mathcal{S}(\mu^2_{e,t}|\eta_{e,t}),$$

where $\mu_{e,t} = \alpha\mu^1_{e,t} + (1-\alpha)\mu^2_{e,t}$ *and* $\alpha \in [0,1]$.

(iii) *The total variation of* $\mu_{e,t}$, *defined by*

$$|\mu_{e,t}| = \sup_{\Pi} \sum_{i=1}^{m} \int_{A_i} \mathrm{d}\mu_{e,t}(u),$$

satisfies

$$|\mu_{e,t} - \eta_{e,t}|^2 \leq -2\mathcal{S}(\mu_{e,t}|\eta_{e,t}),$$

where the supremum is taken over all finite partitions $\Pi = \{A_1, A_2, \ldots, A_m\}$ *of* $\mathcal{H}$.

(iv) *Functional* $\mathcal{S}(f_{e,t})$ *is upper-semicontinuous: if* $|\mu^n_{e,t} - \mu_{e,t}| \to 0$ *as* $n \to \infty$, *then*

$$\mathcal{S}(\mu_{e,t}|\eta_{e,t}) \geq \limsup_n \mathcal{S}(\mu^n_{e,t}|\eta_{e,t}).$$

Proof. Statement (i) has been proven in Section 2.1. The proofs of (iii)–(iv) can be found in Ch.1 of [4].

A slight modification of (ii) is also stated in [4] but without explicit proof, so we include the proof of (ii) for the sake of completeness. We begin by invoking the inequality:

$$\sum_{i=1}^{m} a_i \log(\frac{a_i}{b_i}) \geq a \log(\frac{a}{b}),$$

where a_i, b_i are any non-negative numbers for all $i = 1, 2, \ldots, m$, with $a = \sum_i a_i$ and $b = \sum_i b_i$. The equality holds if and only if $\frac{a_1}{b_1} = \frac{a_2}{b_2} = \cdots = \frac{a_m}{b_m}$.

Fix $\alpha \in [0,1]$, and let $\mu_{e,t} = \alpha\mu^1_{e,t} + (1-\alpha)\mu^2_{e,t}$. Consider

$$\begin{aligned} &\alpha\frac{d\mu^1_{e,t}}{d\eta_{e,t}} \log(\frac{\mathrm{d}\mu^1_{e,t}}{\mathrm{d}\eta_{e,t}}) + (1-\alpha)\frac{\mathrm{d}\mu^2_{e,t}}{\mathrm{d}\eta_{e,t}} \log(\frac{d\mu^2_{e,t}}{d\eta_{e,t}}) \\ &= \alpha\frac{\mathrm{d}\mu^1_{e,t}}{\mathrm{d}\eta_{e,t}} \log\left(\frac{\alpha\frac{\mathrm{d}\mu^1_{e,t}}{\mathrm{d}\eta_{e,t}}}{\alpha\frac{\mathrm{d}\eta_{e,t}}{\mathrm{d}\eta_{e,t}}}\right) + (1-\alpha)\frac{\mathrm{d}\mu^2_{e,t}}{\mathrm{d}\eta_{e,t}} \log\left(\frac{(1-\alpha)\frac{\mathrm{d}\mu^2_{e,t}}{\mathrm{d}\eta_{e,t}}}{(1-\alpha)\frac{\mathrm{d}\eta_{e,t}}{\mathrm{d}\eta_{e,t}}}\right) \\ &\geq \left(\alpha\frac{\mathrm{d}\mu^1_{e,t}}{\mathrm{d}\eta_{e,t}} + (1-\alpha)\frac{\mathrm{d}\mu^2_{e,t}}{\mathrm{d}\eta_{e,t}}\right) \log\left(\frac{\alpha\frac{\mathrm{d}\mu^1_{e,t}}{\mathrm{d}\eta_{e,t}} + (1-\alpha)\frac{\mathrm{d}\mu^2_{e,t}}{\mathrm{d}\eta_{e,t}}}{\alpha\frac{\mathrm{d}\eta_{e,t}}{\mathrm{d}\eta_{e,t}} + (1-\alpha)\frac{\mathrm{d}\eta_{e,t}}{\mathrm{d}\eta_{e,t}}}\right) \\ &= \frac{\mathrm{d}\mu_{e,t}}{\mathrm{d}\eta_{e,t}} \log(\frac{\mathrm{d}\mu_{e,t}}{\mathrm{d}\eta_{e,t}}). \end{aligned}$$

Integrating with respect to $\eta_{e,t}$ gives the desired conclusion.

We now state and prove the main result of this paper.

Theorem 2.2. *The physical measure $\eta_{e,t}$ is a solution to the maximization problem (2.5). That is, the physical measure maximizes the entropy functional relative to all the comparison probability measures $\mu_{e,t}$, both restricted to the global energy-enstrophy surface $G_e(t)$.*

Proof. We divide the proof into two steps.

1. *Existence.* We follow the argument in Ch.3 of [4]. Fix $t \in [0,T]$ and a non-negative function $e(t)$. Consider the problem

$$I_{\sup} = \sup_{\mu_{e,t} \in \mathcal{A}} \mathcal{S}(f_{e,t}) \leq 0, \tag{2.6}$$

where $\mathcal{A} := \{\mu_{e,t} \in \mathcal{M} \text{ satisfying (2.4)}\}$. Let $\mu^n_{e,t}$ be a maximizing sequence for our problem, *i.e.*, $\mu^n_{e,t} \in \mathcal{A}$, and $\mathcal{S}(f^n_{e,t}) \to I_{\sup}$ as $n \to \infty$. It can directly be checked that

$$\begin{aligned} \mathcal{S}(f^m_{e,t}) + \mathcal{S}(f^n_{e,t}) &= \mathcal{S}(\mu^m_{e,t}|\eta_{e,t}) + \mathcal{S}(\mu^n_{e,t}|\eta_{e,t}) \\ &= 2\mathcal{S}(\mu^{m,n}_{e,t}|\eta_{e,t}) + \mathcal{S}(\mu^m_{e,t}|\mu^{m,n}_{e,t}) + \mathcal{S}(\mu^n_{e,t}|\mu^{m,n}_{e,t}), \end{aligned} \tag{2.7}$$

where $\mu^{m,n}_{e,t} = \frac{1}{2}(\mu^m_{e,t} + \mu^n_{e,t})$. Because of the linear nature of the constraint, we see that $\mu^{m,n}_{e,t} \in \mathcal{A}$. The concavity of the relative entropy, Lemma 2.1, gives

$$2\mathcal{S}(\mu^{m,n}_{e,t}|\eta_{e,t}) \geq \mathcal{S}(\mu^m_{e,t}|\eta_{e,t}) + \mathcal{S}(\mu^n_{e,t}|\eta_{e,t}),$$

which, together with the non-positivity of $\mathcal{S}(\cdot)$, implies that $\lim_{m,n} \mathcal{S}(\mu^{m,n}_{e,t}|\eta_{e,t}) = 2I_{\sup}$ so that $\lim_{m,n} \mathcal{S}(\mu^m_{e,t}|\mu^{m,n}_{e,t}) = 0 = \lim_{m,n} \mathcal{S}(\mu^n_{e,t}|\mu^{m,n}_{e,t})$. By Lemma 2.1(iii),

$$\begin{aligned} |\mu^m_{e,t} - \mu^n_{e,t}| &\leq |\mu^m_{e,t} - \mu^{m,n}_{e,t}| + |\mu^n_{e,t} - \mu^{m,n}_{e,t}| \\ &\leq \sqrt{-2\mathcal{S}(\mu^m_{e,t}|\mu^{m,n}_{e,t})} + \sqrt{-2\mathcal{S}(\mu^n_{e,t}|\mu^{m,n}_{e,t})} \to 0 \qquad \text{as } m,n \to \infty. \end{aligned} \tag{2.8}$$

Then we conclude that there exists $\mu^*_{e,t} \in \mathcal{M}$ satisfying

$$\lim_n |\mu^n_{e,t} - \mu^*_{e,t}| = 0. \tag{2.9}$$

Moreover, due to the linearity of constraint (2.4) again, it readily follows that $\mu^*_{e,t} \in \mathcal{A}$. Now, by the upper-semicontinuity of $\mathcal{S}$, Lemma 2.1(iv), we conclude

$$I_{\sup} = \mathcal{S}(\mu^*_{e,t}|\eta_{e,t}).$$

2. *Euler–Lagrange equations*. Since the constraint is linear in f, then, by the standard methods of the Calculus of Variations, there exists a constant α_t (independent of u) satisfying the Euler–Lagrange equations:

$$-\int \Psi_t(u)\big(1+\log(f^*_{e,t}(u))\big)\,\mathrm{d}\eta_{e,t}(u) - \alpha_t \int \Psi_t(u)\,\mathrm{d}\eta_{e,t}(u) = 0 \tag{2.10}$$

for any test function $\Psi_t(u)$. Then we have

$$-\big(1+\log(f^*_{e,t}(u))\big) - \alpha_t = 0 \tag{2.11}$$

for $\eta_{e,t}$-almost every $u \in G_e(t)$. Thus,

$$\log(f^*_{e,t}(u)) = -(1+\alpha_t), \tag{2.12}$$

or equivalently

$$f^*_{e,t}(u) = \exp\{-(1+\alpha_t)\}. \tag{2.13}$$

That is, $f^*_{e,t}$ is constant in u. Since both $\mu^*_{e,t}$ and $\eta_{e,t}$ are probability measures, constraint (2.4) yields $\alpha_t \equiv -1$, implying that $f^*_{e,t} \equiv 1$ or equivalently $\mu^*_{e,t} = \eta_{e,t}$. In other words, the physical measure $\eta_{e,t}$ achieves the maximum entropy under constraint (2.4).

Remark 2.1. The global energy-enstrophy surface $G_e(t)$ in (2.1) for Theorem 2.2 can be replaced by the local energy-enstrophy surface: For any given $e(x,t) > 0$,

$$\begin{aligned}&\frac{1}{2}\int_V |u(x,t)|^2\psi(x,t)\,\mathrm{d}x + \nu\int_0^t\int_V |\nabla u|^2\psi\,\mathrm{d}x\,\mathrm{d}\tau\\ &-\frac{1}{2}\int_0^t\int_V \big\{|u|^2(\partial_t\psi+\nu\Delta\psi)+(|u|^2+2p)u\cdot\nabla\psi\big\}\,\mathrm{d}x\,\mathrm{d}\tau = \int_V e(x,t)\psi(x,t)\,\mathrm{d}x\end{aligned} \tag{2.14}$$

for any $\psi(x,\tau) \in C^\infty_{\mathrm{c}}(V\times[0,t])$. Then Theorem 2.2 can be stated as follows: the physical measure maximizes the entropy functional relative to all the comparison probability measures $\mu_{e,t}$, both restricted to the local entropy-enstrophy surface (2.14).

Remark 2.2. A connection can be drawn to the classical results highlighted in Section 1.2: at $t = 0$, we set $\eta_{e,0} = g(u)\,\xi$, where $g(u)$ is given by Eq. (1.7); this implies that $\dfrac{\mathrm{d}\mu^*_{e,0}}{\mathrm{d}\xi} = g(u) \in L^1(\mathrm{d}\xi)$ in agreement with the classical principle of maximum entropy. Furthermore, at $t = 0$, the energy condition in (2.2) gives the classical energy constraint $(1.5)_2$, once we identify the mean energy with $\int e(0)g(u)\,\mathrm{d}\xi$.

3. Connections and remarks

In view of the previous work [8] on the principle of maximum entropy applied to the problem of inviscid fluid flow, the current results outline a general approach for studying the case of viscous flow under a common framework. As such, it is clear that the standard energy class associated with the Navier–Stokes equations provides the natural setting for expressing the total energy $\|u\|_t^2$ of the fluid at each time t, and hence defining the energy-enstrophy surface $G_e(t)$.

Equally clear is that having a complete theory of entropy maximization for fluid flow is tied with specifying selection criteria for the reference measure $\eta_{e,t}$. Whether or not such criteria are model dependent is open to question. For instance, in the finite-dimensional setting, a natural candidate for the reference measure is the Lebesgue measure (see [8,23]) as it is the only measure (up to a multiplicative constant) that possesses the property of translation invariance; that is, in the velocity space, it remains invariant under the Galilean transformation. Yet another candidate for the reference measure is given by considering the statistics of fluid flows at a long time; this is related to the *stationary* statistical solutions of the Navier–Stokes equations and Kolmogorov's theory of turbulence (*cf.* [12]).

We first recall the notion of the stationary statistical solutions of the Navier–Stokes equations. Assume that μ is a Borel probability measure over space $\mathcal{H}$ with finite enstrophy, that is,

$$\int_{\mathcal{H}} \|\nabla u\|^2\,\mathrm{d}\mu(u) < \infty.$$

Then the measure μ is said to be a stationary statistical solution of the Navier–Stokes equations if μ satisfies the following (stationary) Liouville-type equation:

$$\int_{\mathcal{H}}\int_V Q(u)\,\Phi'(u)\,\mathrm{d}x\,\mathrm{d}\mu(u) = 0 \tag{3.1}$$

for all suitable test functions Φ, and the energy inequality:

$$\int_{\mathcal{H}}\int_V \big(\nu|\nabla u|^2 - F\cdot u\big)\,\mathrm{d}x\,\mathrm{d}\mu(u) \le 0, \tag{3.2}$$

where $Q(u) = -\nu\Delta u + (u\cdot\nabla)u - F$.

Now from the perspective of thermodynamics, the flow relaxes, after sufficiently long enough time, to its equilibrium state characterized by maximum entropy; so we may assume that the statistics of this (final) state is encoded by the stationary solution $\mu \doteq \mu_\infty$. On the other hand, one of the fundamental postulates of Kolmogorov's theory for fully developed isotropic turbulence is the similarity hypothesis that the statistical behavior of the flow must be determined by only two parameters: the (average) energy dissipation rate $\bar{\epsilon}_\nu(t)$ and viscosity ν (see also [24]). Therefore, the average energy balance, in the absence of body forces, for the case of non-zero energy dissipation rate (*i.e.*, the average of Eq. (1.3)) reads

$$\frac{1}{2}\int_{\mathcal{H}}\|u\|^2\,\mathrm{d}\mu_t(u) + \underbrace{\nu\int_{\mathcal{H}}\int_0^t\|\nabla u(s)\|^2\,\mathrm{d}s\,\mathrm{d}\mu_t(u)}_{\bar{\epsilon}_\nu(t)} = \frac{1}{2}\int_{\mathcal{H}}\|u\|^2\,\mathrm{d}\mu_0(u), \tag{3.3}$$

where the probability measure μ_t is such that Eq. (3.3) is well-defined (*e.g.* a time-dependent statistical solution). However, if we set $\mu_t = \mu_\infty$, then the non-zero energy (finite) dissipation rate can be maintained only if the kinetic energy is infinite, as seen by Eq. (3.2). This non-physical conclusion was pointed out first by Hopf [24] (see also [12]). The resolution of this paradox, according to Foias, Manley, Rosa, and Temam [11,12,16], lies instead in introducing a *family* of probability measures μ_t (parameterized by time $t \in [0,\infty)$),[6] a feature which the physical measure $\eta_{e,t}$ enjoys through the energy-enstrophy surface $G_e(t)$. These solutions, called self-similar homogeneous statistical solutions, were originally postulated for describing the decay of fully developed turbulence according to Kolmogorov [13,14], though to date the question of their existence has not been resolved yet.

Meanwhile, our results suggest replacing the right-hand side of Eq. (3.3) with $\bar{e}(t)$, a quantity that can ostensibly be identified with $\int_{G_e(t)} e(t)\,\mathrm{d}\eta_{e,t}$. Such a choice adds a degree of freedom to the final statistics of the flow, now depending upon the initial prescribed energy $e(t)$. In keeping with the above similarity hypothesis, however, we can in principle prescribe the total energy $e(t)$ at a high Reynolds number in terms of parameters $(\bar{\epsilon}_\nu, \nu)$ and wavenumber κ, removing the ambiguity associated with the term: $\int_{G_e(t)} e(t)\,\mathrm{d}\eta_{e,t}$. Let $\mathcal{E}(\kappa,t)$ be the *energy spectrum* (in the Fourier space with $|\boldsymbol{\kappa}| = \kappa$) associated with the average kinetic energy. Then the average total energy can be expressed in terms of the energy spectrum:

$$\int_0^\infty \mathcal{E}(\kappa,t)\Big(\frac{1}{2}+\nu\kappa^2\Big)\,\mathrm{d}\kappa. \tag{3.4}$$

[6] More precisely, for each homogeneous statistical solution of the Navier–Stokes equations μ_t (with viscosity ν), we define $\mu_t = \mu^{\nu,\bar{\epsilon}_\nu(t)}$ for some $\bar{\epsilon}_\nu(t) > 0$.

Eq. (3.4) provides a natural alternative for our (average) total energy $\bar{e}(t)$. In fact, the amount of total energy contained between wavenumbers κ_1 and κ_2 at a given time t, *i.e.*,

$$\int_{\kappa_1}^{\kappa_2} \mathcal{E}(\kappa, t)\left(\frac{1}{2} + \nu\kappa^2\right) \mathrm{d}\kappa,$$

can now be identified with the change in the total energy $\bar{e}_{\kappa_2}(t) - \bar{e}_{\kappa_1}(t)$.

As $\nu \to 0$, or equivalently at *infinite* Reynolds number, the implications of Kolmogorov's theory give

$$\mathcal{E}(\kappa, t) \sim C(t)\,\bar{\epsilon}(t)^{\frac{2}{3}}\,\kappa^{-\frac{5}{3}} \tag{3.5}$$

for large enough $|\boldsymbol{\kappa}| = \kappa$ but still belonging to the inertial range, where $C(t)$ is a positive dimensionless constant independent of viscosity ν, and $\lim_{\nu\to 0} \bar{\epsilon}_\nu(t) = \bar{\epsilon}(t) > 0$. In both cases, nevertheless, the explicit dependence on $e(t)$ can be removed in the final statistics in favor of the energy dissipation rate and viscosity.

In view of interpreting the total energy $e(t)$ in terms of the energy spectrum, one can conceive of the maximum entropy probability measure, that is the reference probability measure, as assigning a probability for finding the total energy carried by a certain wavenumber κ to be $\bar{e}_\kappa(t)$. When the wavenumber lies in the inertial range and in the limit of infinite Reynolds number, this probability can be calculated via the Kolmogorov energy spectrum (3.4).

CRediT authorship contribution statement

Gui-Qiang G. Chen: Writing – review & editing, Investigation. **James Glimm:** Writing – review & editing, Investigation. **Hamid Said:** Writing – review & editing, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

Gui-Qiang G. Chen's research is partially supported by the UK Engineering and Physical Sciences Research Council Awards EP/L015811/1, EP/V008854/1, and EP/V051121/1.

References

[1] M. Gavrilov, R. Chétrite, J. Bechhoefer, Direct measurement of weakly nonequilibrium system entropy is consistent with Gibbs–Shannon form, Proc. Natl. Acad. Sci. 114 (42) (2017) 11097–11102.
[2] A. Wherl, The many facets of entropy, Rep. Math. Phys. 30 (1) (1991) 119–129.
[3] E.T. Jaynes, Information theory and statistical mechanics, Phys. Rev. 106 (1957) 620.
[4] S. Ihara, Information Theory for Continuous Systems, Vol. 2, World Scientific, 1993.
[5] J.N. Kapur, Maximum-Entropy Models in Science and Engineering, John Wiley & Sons, 1989.
[6] S. Kullback, R.A. Leibler, On information and sufficiency, Ann. Math. Stat. 22 (1) (1951) 79–86.
[7] S. Kullback, Information Theory and Statistics, Courier Corporation, 1997.
[8] J. Glimm, D. Lazarev, G.-Q.G. Chen, Maximum entropy production as a necessary admissibility condition for the fluid Navier–Stokes and Euler equations, SN Appl. Sci. 2 (2020) 1–9.
[9] T. Buckmaster, V. Vicol, Nonuniqueness of weak solutions to the Navier–Stokes equations, Ann. of Math. 189 (2019) 101–144.
[10] U. Frisch, R.J. Donnelly, Turbulence: the legacy of A.N. Kolmogorov, Phys. Today 49 (1996) 82–84.
[11] C. Foias, Statistical study of Navier–Stokes equations, I, Rend. Semin. Mat. Univ. Padova 48 (1972) 219–348.
[12] C. Foias, O. Manley, R. Rosa, R. Temam, Navier–Stokes Equations and Turbulence, Vol. 83, Cambridge University Press, 2001.
[13] A.N. Kolmogorov, Dissipation of energy in the locally isotropic turbulence, Dokl. Akad. Nauk SSSR 32 (1941) 19–21.
[14] A.N. Kolmogorov, The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 434 (1991) 9–13.
[15] M.I. Vishik, A.V. Furkisov, Mathematical Problems of Statistical Hydromechanics, Vol. 9, Springer Science & Business Media, 2012.
[16] C. Foias, R. Rosa, R. Temam, Properties of time-dependent statistical solutions of the three-dimensional Navier–Stokes equations, Ann. Inst. Fourier 63 (2013) 2515–2573.
[17] U.S. Fojordholm, K. Lye, S. Mishra, F. Weber, Statistical solutions of hyperbolic systems of conservation laws: numerical approximation, Math. Models Methods Appl. Sci. 30 (2020) 539–609.
[18] S. Lanthaler, S. Mishra, C. Parés-Pulido, Statistical solutions of the incompressible euler equations, Math. Models Methods Appl. Sci. 31 (2021) 223–292.
[19] F.A. Bais, J.D. Farmer, The physics of information, in: J. van Benthem, P. Adriaan (Eds.), The Handbook on the Philosophy of Information, Elsevier, 2007.
[20] W. Dreyer, Maximisation of the entropy in non-equilibrium, J. Phys. A: Math. Gen. 20 (18) (1987) 6505.
[21] C.D. Levermore, Moment closure hierarchies for kinetic theories, J. Stat. Phys. 83 (1996) 1021–1065.
[22] I. Müller, T. Ruggeri, Extended Thermodynamics, in: Springer Tracts in Nat. Phil, vol. 37, 1993.
[23] J. Dunkel, P. Talkner, P. Hänggi, Relative entropy, Haar measures and relativistic canonical velocity distributions, New J. Phys. 9 (2007) 144.
[24] E. Hopf, Statistical hydromechanics and functional calculus, J. Ration. Mech. Anal. 1 (1952) 87–123.