PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: D'Onofrio R, Ortenzi G, Roulstone I, Rubtsov V. 2023 Solutions and singularities of the semigeostrophic equations via the geometry of Lagrangian submanifolds. *Proc. R. Soc. A* **479**: 20220682. https://doi.org/10.1098/rspa.2022.0682

Received: 18 October 2022
Accepted: 23 January 2023

Subject Areas:
mathematical physics, fluid mechanics

Keywords:
semigeostrophic equationsMonge-Ampère equationsLagrangian submanifoldssingularitiespseudo-Riemannian geometry

Author for correspondence:
Roberto D'Onofrio
e-mail: r.donofrio1@campus.unimib.it

Solutions and singularities of the semigeostrophic equations via the geometry of Lagrangian submanifolds

Roberto D'Onofrio[1,2,3], Giovanni Ortenzi[1,2], Ian Roulstone[3] and Volodya Rubtsov[4,5]

[1]Dipartimento di Matematica e Applicazioni, Università di Milano-Bicocca, Via Roberto Cozzi 55, I-20125 Milano, Italy
[2]INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy
[3]Department of Mathematics, University of Surrey, Guildford GU2 7XH, UK
[4]Département de Mathématiques, University of Angers, CNRS, LAREMA, SFR MATHSTIC, F-49000 Angers, France
[5]IGAP (Institute for Geometry and Physics), Via Beirut 2, 34151 Trieste, Italy

RD'O, 0000-0001-7894-8254

By using Monge–Ampère geometry, we study the singular structure of a class of nonlinear Monge–Ampère equations in three dimensions, arising in geophysical fluid dynamics. We extend seminal earlier work on Monge–Ampère geometry by examining the role of an induced metric on Lagrangian submanifolds of the cotangent bundle. In particular, we show that the signature of the metric serves as a classification of the Monge–Ampère equation, while singularities and elliptic–hyperbolic transitions are revealed by degeneracies of the metric. The theory is illustrated by application to an example solution of the semigeostrophic equations.

1. Introduction

Atmospheric fronts are a salient feature of mid-latitude weather systems. From the viewpoint of mathematical modelling, fronts are understood as material interfaces, advected by the fluid flow, across which the physical

THE ROYAL SOCIETY PUBLISHING

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 479: 20220682

features undergo a jump discontinuity. One of the most successful approaches to mathematical modelling of weather fronts is Hoskins' semigeostrophic (SG) equations [1–3]. In [4], Chynoweth & Sewell recognized the presence of a Legendre duality structure between four different sets of variables in SG theory akin to the classical quartet of dual potentials in thermodynamics. In the same article, Chynoweth & Sewell showed how singularities of the Legendre mapping could be used to model flows containing a weather front. Their approach is reminiscent of the studies of shock waves in stationary gas flows, where it is known under the name of hodograph transformation (e.g. Chapter 12 of [5]).

SG flows are completely described by a single function called the geopotential. Moreover, they conserve a form of Ertel's potential vorticity, which in turn is related to the geopotential by a Monge–Ampère type equation. By denoting the geopotential by $P(x, y, z, t)$ and the potential vorticity by $q_g(x, y, z, t)$, we may write this relation as follows:

$$\begin{vmatrix} \frac{\partial^2 P}{\partial x^2} & \frac{\partial^2 P}{\partial x \partial y} & \frac{\partial^2 P}{\partial x \partial z} \\ \frac{\partial^2 P}{\partial y \partial x} & \frac{\partial^2 P}{\partial y^2} & \frac{\partial^2 P}{\partial y \partial z} \\ \frac{\partial^2 P}{\partial z \partial x} & \frac{\partial^2 P}{\partial z \partial y} & \frac{\partial^2 P}{\partial z^2} \end{vmatrix} = C q_g(x, y, z, t), \tag{1.1}$$

where C is a constant depending on the physical parameters of the model (e.g. [6]). Time dependence is implicit in (1.1) as no time derivatives are involved. Therefore, (1.1) represents a kinematic constraint between the geopotential and the potential vorticity and can be studied by considering the time as a fixed parameter. This approach was also used in [4], where the authors provide several examples of couples (P, q_g) satisfying (the two-dimensional version of) (1.1) and capable of modelling atmospheric fronts frozen in time. Both the kinematic and the dynamic view are taken in [7].

The kinematic approach to singularities can be studied using the geometrical framework of Monge–Ampère equations (MAEs) pioneered by Lychagin and his school (e.g. [8]). Delahaies & Roulstone [9] have explored the implications of Monge–Ampère geometry for the shallow water version of the SG equations, while Roulstone *et al.* [10] and Banos *et al.* [11] have conducted similar studies for the inviscid Navier–Stokes (Euler) equations. This article follows this line of research and investigates the relevance of Monge–Ampère geometry to the study of singularities of the incompressible three-dimensional SG equations. One of the main advantages of the geometric approach to MAEs is a clear and intuitive understanding of the notion of a generalized solution. While a classical solution is a function $P(x, y, z, t)$, a generalized solution is a Lagrangian submanifold L in the cotangent bundle $T^*\mathbb{R}^3$ (the phase space) of the physical space $\mathbb{R}^3$. One thinks of the manifold L as the multi-valued graph of the gradient of the geopotential, understood as a map $\nabla P : \mathbb{R}^3 \to \mathbb{R}^3$, in $T^*\mathbb{R}^3 \simeq \mathbb{R}^3 \times \mathbb{R}^3$. To recover physical information, a generalized solution L must be projected onto the physical space, and singularities can arise in the process. This geometrical perspective on solutions and their singularities was first introduced by Vinogradov & Kupershmidt in their work on Hamilton–Jacobi theory (§8 of [12]). Kossowski [13] has independently proposed the same formalism for studying singularities of symplectic MAEs in two independent variables. A similar viewpoint is adopted by Ishikawa & Machida [14,15] for classifying generic singularities of Hessian type MAEs in two and three independent variables. We refer to [16] for further examples of application of this formalism to more general nonlinear partial differential equations (PDEs).

In this work, we present an alternative approach to singularities based on the pseudo-Riemannian geometry. For classifying symplectic MAEs in three independent variables, Lychagin & Rubtsov [17] introduced a metric tensor g_ω on $T^*\mathbb{R}^3$ (formula (3.1)) and showed that its signature distinguishes the various classes—elliptic, hyperbolic and parabolic (see also [18]). For

the particular case of equation (1.1), this metric has signature (3,3) and gives $T^*\mathbb{R}^3$ the structure of a pseudo-Riemannian manifold.

Every generalized solution $L \subset T^*\mathbb{R}^3$ inherits a metric structure from the ambient space associated with the pull-back metric $h_\omega := g_\omega|_L$, and this constitutes the main focus of the present work. We study the properties of singular solutions to (1.1) through pseudo-Riemannian geometry of Lagrangian submanifolds. Our main goal is to provide an understanding of the pull-back Lychagin–Rubtsov metric h_ω from the viewpoint of the PDE theory. We claim that the h_ω metric plays the same role for MAEs as the coefficient matrix does for linear second-order PDEs. This means that it can be used to define the symbol type of the underlying MAE, and, in hyperbolic regime, to construct the characteristic surfaces. In our general setting, the MAEs under consideration may be of mixed type. We show that h_ω is Riemannian on elliptic branches of L and Lorentzian on hyperbolic ones. Moreover, we prove that elliptic–hyperbolic transitions and kinematic singularities coincide for equation (1.1), implying that h_ω degenerates on the singular locus of L. In this sense, we claim that the pull-back metric is a diagnostic tool for studying singularities.

In §2, we give some background on SG equations and Monge–Ampère geometry. We state and prove our results about the pull-back of the Lychagin–Rubtsov metric and the symbol type of the MAE in §3. Finally, we present an explicit example of a solution to the SG equations, illustrating the aforementioned results, in §4.

2. Background and methods

In this review section, we recall some basics about the SG system and the geometry of MAEs.

(a) Semigeostrophic equations

Hoskins' SG equations [3] are an approximation to the Euler system of fluid dynamics intended to model large-scale motion of the atmosphere. They are usually written as follows:

$$\begin{cases} \dfrac{Du_g}{Dt} - fv + \dfrac{\partial \phi}{\partial x} = 0, \ \dfrac{Dv_g}{Dt} + fu + \dfrac{\partial \phi}{\partial y} = 0, \ \dfrac{g\theta}{\theta_0} = \dfrac{\partial \phi}{\partial z}, & \text{(momentum)} \\ \dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z} = 0, & \text{(mass)} \\ \dfrac{D\theta}{Dt} = 0, & \text{(energy)} \end{cases} \tag{2.1}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z} \tag{2.2}$$

is the material time derivative, and $\{x, y, z\}$ form a Cartesian coordinate system with y directed pole-ward and z directed vertically. The unknowns are the fluid velocity field (u, v, w), the geopotential ϕ and the potential temperature θ (see [4] for a detailed definition of these variables). The positive constants $f \approx 10^{-4}$ Hz and $g \approx 10\,\mathrm{m\ s^{-2}}$ account for the effects of Earth's rotation and gravity, while θ_0 is a reference value for θ. Further, there are two main approximations at work in (2.1). Firstly, hydrostatic balance is assumed, which results in neglecting the vertical acceleration term in the momentum balance. Secondly, the fluid flow is supposed to be close to geostrophic equilibrium (see [3], §3): this is accounted for in (2.1) through Hoskins' 'geostrophic momentum approximation', which replaces the fluid velocity in the horizontal acceleration terms with its geostrophic part,

$$u_g := -\frac{1}{f}\frac{\partial \phi}{\partial y} \quad \text{and} \quad v_g := \frac{1}{f}\frac{\partial \phi}{\partial x}. \tag{2.3}$$

System (2.1) implies an important conservation property, namely, the conservation of Ertel's potential vorticity along fluid trajectories,

$$\frac{Dq_g}{Dt} = 0. \tag{2.4}$$

This quantity represents the projection of the absolute geostrophic vorticity,

$$\zeta_g = \left(-\frac{\partial v_g}{\partial z} + \frac{1}{f}\frac{\partial(u_g, v_g)}{\partial(y,z)}, \frac{\partial u_g}{\partial z} + \frac{1}{f}\frac{\partial(u_g, v_g)}{\partial(z,x)}, f + \frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} + \frac{1}{f}\frac{\partial(u_g, v_g)}{\partial(x,y)} \right), \tag{2.5}$$

along the gradient of the potential temperature,

$$q_g := \zeta_g \cdot \nabla\theta. \tag{2.6}$$

Notice that equations (2.6) and (2.3) and the hydrostatic balance condition (2.1) provide a direct link between the potential vorticity and the geopotential, which comes in the form of a MAE. This statement is made more clear by introducing the modified geopotential:

$$P := \frac{\phi}{f^2} + \frac{x^2}{2} + \frac{y^2}{2}, \tag{2.7}$$

which allows one to write equation (2.6) as follows:

$$q_g = \frac{f^3\theta_0}{g} \det \mathrm{Hess}(P), \tag{2.8}$$

where det Hess(P) denotes the determinant of the Hessian matrix of P with respect to the spatial variables. In this work, we are exclusively interested in kinematic aspects of system (2.1) as encoded in the Monge–Ampère equation (2.8).

The mathematical structure of the SG equations appears even more clearly on introducing dimensionless variables. Following [19], we write the SG system in a dimensionless form as follows:

$$\left.\begin{aligned} &\epsilon\frac{Du_g}{Dt} - v + \frac{\partial\phi}{\partial x} = 0, \quad \epsilon\frac{Dv_g}{Dt} + u + \frac{\partial\phi}{\partial y} = 0 \ \frac{D\theta}{Dt} = 0 \\ \text{and} \quad &\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \quad u_g = -\frac{\partial\phi}{\partial y}, \ v_g = \frac{\partial\phi}{\partial x}, \ \theta = \frac{\partial\phi}{\partial z}. \end{aligned}\right\} \tag{2.9}$$

The dimensionless parameter ϵ is the Rossby number, $\epsilon = U/fL$, which involves the typical horizontal length and velocity scales and represents the ratio between inertial and Coriolis forces. The value of ϵ is typically ≈ 0.1 in SG flows. A consistent dimensionless expression for the modified geopotential is

$$P = \frac{x^2}{2} + \frac{y^2}{2} + \epsilon\phi, \tag{2.10}$$

and the dimensionless version of (2.8) reads accordingly

$$\det \mathrm{Hess}(P) = \epsilon q_g. \tag{2.11}$$

Shutts & Cullen [6] used equation (2.11) to study stability of SG flows with respect to small displacement of fluid parcels. They found that a necessary condition for parcel stability is the (spatial) convexity of the geopotential P, which implies strict positivity of the potential vorticity. Although we place no *a priori* hypothesis on the convexity of P, we shall always assume $q_g > 0$ henceforth.

Much of the current interest in the SG equations is motivated by a change of variable due to Hoskins [3], known as the 'geostrophic momentum transformation', which has drastically improved the general comprehension of SG flows. As later recognized by Chynoweth & Sewell [4], this change of variable may be interpreted as a Legendre type transformation, and we

adopt their perspective for describing it. We start by introducing the (dimensionless) horizontal components of the absolute momentum,

$$M = x + \epsilon v_g = x + \epsilon \frac{\partial \phi}{\partial x} \quad \text{and} \quad N = y - \epsilon u_g = y + \epsilon \frac{\partial \phi}{\partial y}, \tag{2.12}$$

which stand in a special relation with the geopotential,

$$\frac{\partial P}{\partial x} = M \quad \text{and} \quad \frac{\partial P}{\partial y} = N. \tag{2.13}$$

Moreover, the hydrostatic balance condition is written in terms of P as follows:

$$\frac{\partial P}{\partial z} = \epsilon \frac{\partial \phi}{\partial z} = \epsilon \theta. \tag{2.14}$$

Equations (2.13) and (2.14) are the starting point for [4], where a quartet of Legendre transformations of P is identified. In fact, these relations open the way to the geometrization of the kinematic equation (2.11). For the remainder of this section, we will frame the work of Chynoweth and Sewell within the geometrical theory of MAEs (e.g. [8]). For convenience of exposition, we will use the notation,

$$(X, Y, Z) := (M, N, \epsilon\theta) = \nabla P. \tag{2.15}$$

(b) Monge–Ampère structure and Legendre duality

In this work, we only deal with symplectic MAEs in three independent variables. Here, the word *symplectic* means that an equation's coefficients can only depend on the independent variables and the gradient of the dependent variable but not on the value itself of the dependent variable. This is true for equation (2.11) if we understand the potential vorticity q_g as a function of space (and possibly time). As the name suggests, symplectic MAEs are associated with the symplectic geometry of a manifold, the phase space, representing the cotangent bundle to the space of independent variables. In the particular case of equation (2.11), the associated symplectic manifold is $T^*\mathbb{R}^3$, which we endow with coordinates (x, y, z, X, Y, Z) and the canonical symplectic form:

$$\Omega = \mathrm{d}X \wedge \mathrm{d}x + \mathrm{d}Y \wedge \mathrm{d}y + \mathrm{d}Z \wedge \mathrm{d}z. \tag{2.16}$$

A function $P : \mathbb{R}^3 \to \mathbb{R}$ induces a section of the cotangent bundle through its differential, $dP : \mathbb{R}^3 \to T^*\mathbb{R}^3$. Note that the image of dP in $T^*\mathbb{R}^3 \simeq \mathbb{R}^3 \times \mathbb{R}^3$ coincides with the graph of the gradient $\nabla P : \mathbb{R}^3 \to \mathbb{R}^3$. For any given 3-form ω on $T^*\mathbb{R}^3$, we can define a map $\Delta_\omega : \mathrm{C}^\infty(\mathbb{R}^3) \to \Omega^3(\mathbb{R}^3)$ taking functions to 3-forms on $\mathbb{R}^3$, which we call the Monge–Ampère operator associated with ω. It associates a function with the restriction of ω to its graph,

$$\Delta_\omega(P) := (dP)^* \omega, \tag{2.17}$$

where the superscript * denotes the pull-back. The correspondence between Monge–Ampère operators and 3-forms is not 1-to-1 (several forms produce the same operator), but it can be made so by taking a suitable quotient of the space of 3-forms on $T^*\mathbb{R}^3$. Not every 3-form produces a non-zero Monge–Ampère operator and those which do are called *effective*. This induces an equivalence relation on the space of 3-forms on $T^*\mathbb{R}^3$—two forms are equivalent if they differ by a non-effective form. Thus, an equivalence class of 3-forms gives rise to one and the same Monge–Ampère operator. We will exclusively deal with effective forms and shall make no explicit distinction between an effective form and the class it represents. In the symplectic case, effective 3-forms can be characterized as those $\omega \in \Omega^3(T^*\mathbb{R}^3)$, which satisfy $\omega \wedge \Omega = 0$. It is easy to verify that the effective 3-form on $T^*\mathbb{R}^3$ associated with equation (2.11) is

$$\omega = \mathrm{d}X \wedge \mathrm{d}Y \wedge \mathrm{d}Z - \epsilon q_g \, \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z. \tag{2.18}$$

The equation $\Delta_\omega = 0$ is called the MAE associated with ω, and we denote it by E_ω. The MAE corresponding to (2.18) is, in Cartesian coordinates, (2.11), as can be immediately verified by

direct calculation. The next step in the definition of this geometrical framework is the notion of a solution. A *generalized* solution to E_ω is a smooth Lagrangian submanifold[1] $L \in T^*\mathbb{R}^3$ such that $\omega|_L = 0$. A *classical* solution is one that is globally represented as the graph of a function, meaning that L is the range of the differential dP for some twice differentiable function $P(x, y, z)$. Note that in this latter case, the condition $\omega|_L = 0$ reads

$$\Delta_\omega(P) = 0. \tag{2.19}$$

Generalized solutions are precisely those which can be locally (but not globally) represented as the graph of a function near generic points. Stated differently, the mapping $\pi_L := \pi|_L$, where

$$\pi : T^*\mathbb{R}^3 \to \mathbb{R}^3, \quad (x, y, z, X, Y, Z) \mapsto (x, y, z), \tag{2.20}$$

represents the canonical bundle projection, is a diffeomorphism between L and $\mathbb{R}^3$ as long as L is the graph of a classical solution. Otherwise, we must distinguish between *regular* points, where $\mathrm{d}\pi_L$ has maximal rank, and *singular* points, where it has not. In this latter case, the diffeomorphism property is local and only holds near regular points.

Although generalized solutions cannot be globally represented as the graph of a function, they admit an alternative local representation in terms of a single function near any point (either regular or singular). Note that $\Omega = \mathrm{d}\alpha$, where $\alpha := X\,\mathrm{d}x + Y\,\mathrm{d}y + Z\,\mathrm{d}z$ is the tautological 1-form on $T^*\mathbb{R}^3$. If L is a Lagrangian submanifold, then $\alpha|_L$ is closed, and so there exists a function f on L such that, locally,

$$\alpha|_L = df. \tag{2.21}$$

We call f a generating function for the Lagrangian submanifold L. Once a coordinate set on L is selected, (2.21) reduces to an algebraic system of equations whose zero set in $T^*\mathbb{R}^3$ identifies the submanifold L. In the neighbourhood of any given point, coordinates on L can always be chosen as a suitable 3-subset of the cotangent coordinates. Overall, there are $2^3 = 8$ possible choices to pick a 3-subset from $\{x, y, z, X, Y, Z\}$ and, therefore, as many classes of generating functions. The Legendre dual potentials of Chynoweth & Sewell [4],

$$R(X, Y, Z), \quad S(X, Y, z), \quad T(x, y, Z), \tag{2.22}$$

provide some physically relevant examples of generating functions. We explicitly work out the description of a Lagrangian submanifold L in terms of S. Setting

$$f(X, Y, z) = Xx + Yy - S(X, Y, z) \tag{2.23}$$

in equation (2.21) gives,

$$Z\,\mathrm{d}z = x\,\mathrm{d}X + y\,\mathrm{d}Y - \frac{\partial S}{\partial X}\,\mathrm{d}X - \frac{\partial S}{\partial Y}\,\mathrm{d}Y - \frac{\partial S}{\partial z}\,\mathrm{d}z, \tag{2.24}$$

which in turn implies

$$x = \frac{\partial S}{\partial X}(X, Y, z), \quad y = \frac{\partial S}{\partial Y}(X, Y, z), \quad Z = -\frac{\partial S}{\partial z}(X, Y, z). \tag{2.25}$$

The combined zero set of equations (2.25) in $T^*\mathbb{R}^3$ identifies the Lagrangian submanifold L generated by S. Similarly, the choices

$$f(X, Y, Z) = Xx + Yy + Zz - R(X, Y, Z) \tag{2.26}$$

and

$$f(x, y, Z) = Zz + T(x, y, Z) \tag{2.27}$$

lead to a local description of as many classes of Lagrangian submanifolds as the zero set of, respectively,

$$x = \frac{\partial R}{\partial X}(X, Y, Z), \quad y = \frac{\partial R}{\partial N}(X, Y, Z), \quad z = \frac{\partial R}{\partial Z}(X, Y, Z) \tag{2.28}$$

[1]A Lagrangian submanifold L in a symplectic manifold (M^{2n}, Ω) is an isotropic submanifold ($\Omega|_L = 0$) of the maximum possible dimension (dim(L) = n).

and

$$z = -\frac{\partial T}{\partial Z}(x, y, Z), \qquad X = \frac{\partial T}{\partial x}(x, y, Z), \quad Y = \frac{\partial T}{\partial y}(x, y, Z). \tag{2.29}$$

Remark 2.1. In [4], the term 'dual space' is used to refer to the Cartesian space of coordinates (X, Y, z), (x, y, Z) or (X, Y, Z). The geometrical setting brings out the true nature of the dual space as the local coordinate representation of a Lagrangian submanifold.

The condition that L is a solution to the MAE results in a condition on the generating function itself which again takes the form of a MAE. For the Chynoweth–Sewell potentials (2.22), this condition, respectively, reads

$$\det \mathrm{Hess}(R) = \frac{1}{\epsilon q_g}, \tag{2.30}$$

$$\epsilon q_g \left[\frac{\partial^2 S}{\partial X^2}\frac{\partial^2 S}{\partial Y^2} - \left(\frac{\partial^2 S}{\partial X \partial Y} \right)^2 \right] + \frac{\partial^2 S}{\partial z^2} = 0 \tag{2.31}$$

and

$$\frac{\partial^2 T}{\partial x^2}\frac{\partial^2 T}{\partial y^2} - \left(\frac{\partial^2 T}{\partial x \partial y} \right)^2 + \epsilon q_g \frac{\partial^2 T}{\partial Z^2} = 0. \tag{2.32}$$

We close this section by noting that $R(X, Y, Z)$ plays a distinguished role among the Chynoweth–Sewell potentials. Indeed, if (x, y, z) are entirely replaced by (X, Y, Z) in the role of independent variables, the whole SG system (2.1) reduces to just two equations (e.g. [20]) and comprises the MAE (2.30) and a transport equation for the potential vorticity,

$$\left.\begin{aligned} &\epsilon q_g \det \mathrm{Hess}(R) = 1, \\ &\frac{\partial q_g}{\partial t} = \frac{\partial(q_g, \Psi)}{\partial(X, Y)} \\ \text{and} \quad &\Psi := \frac{X^2}{2} + \frac{Y^2}{2} - R. \end{aligned}\right\} \tag{2.33}$$

System (2.33) provides a clear distinction between the model's kinematics, encoded in the MAE, and its dynamics, represented by the transport of vorticity.

3. Pseudo-Riemannian geometry and classification of nonlinear PDEs

To classify[2] symplectic three-dimensional Monge–Ampère operators, Lychagin & Rubtsov [17] introduced a symplectic invariant attached to any given effective 3-form on $T^*\mathbb{R}^3$. It may be defined through the relation (see [11])

$$g_\omega(\xi_1, \xi_2)\frac{\Omega^3}{3!} = \iota_{\xi_1}\omega \wedge \iota_{\xi_2}\omega \wedge \Omega, \tag{3.1}$$

which holds for each pair of vector fields ξ_1, ξ_2 on $T^*\mathbb{R}^3$. Note that g_ω is a symmetric bilinear form on the phase space $T^*\mathbb{R}^3$. Moreover, as shown in [18], g_ω happens to be non-degenerate for certain classes of Monge–Ampère operators. When this holds, g_ω defines a Riemannian (or pseudo-Riemannian) metric on the phase space $T^*\mathbb{R}^3$, which we call the Lychagin–Rubtsov metric. For ω and Ω given by (2.18) and (2.16), (3.1) yields

$$g_\omega = 2\epsilon q_g(\mathrm{d}x\,\mathrm{d}X + \mathrm{d}y\,\mathrm{d}Y + \mathrm{d}z\,\mathrm{d}Z). \tag{3.2}$$

Thus, as long as the MAE (2.11) is concerned, g_ω is a pseudo-Riemannian metric with signature $(3, 3)$ over the phase space $T^*\mathbb{R}^3$. The Lychagin–Rubtsov metric induces a pseudo-metric on every

[2]Two effective 3-forms ω_1 and ω_2 on $T^*\mathbb{R}^3$ are locally equivalent if there is a local symplectomorphism of the phase space pulling ω_2 back to ω_1.

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 479: 20220682

submanifold of $T^*\mathbb{R}^3$, and, in particular, on solutions of the MAE. Let

$$\iota : L \to T^*\mathbb{R}^3, \tag{3.3}$$

be a generalized solution to (2.11), that is, a Lagrangian submanifold such that $\iota^*\omega = 0$. Then, L inherits the (pseudo)-Riemannian structure of ambient space as provided by the pull-back metric

$$h_\omega := \iota^* g_\omega. \tag{3.4}$$

Remark 3.1. The Lychagin–Rubtsov metric (3.2) is pseudo-Riemannian (non-degenerate) as long as $q_g \neq 0$. This is always true in the present work as we assume $q_g > 0$. However, nothing can be said *a priori* about the pull-back metric (3.4), which depends on the solution L and the position on L. In the general case, h_ω can be either Riemannian, pseudo-Riemannian and even degenerate. We call h_ω degenerate at a point $e \in L$ if there exists a tangent vector $\xi_1 \in T_eL$ such that $h_\omega(\xi_1, \xi_2) = 0$ for any $\xi_2 \in T_eL$.

We are now able to describe one of the main results of this article: the characterization of MAEs in terms of the geometry of L. We show that there is a correspondence between the signature of (3.4) and the symbol type (elliptic/parabolic/hyperbolic) of the MAE (2.11). This is made precise in proposition 3.7 and leads to a natural characterization of the equation type in terms of h_ω (definition 3.8). We start with classical solutions, amenable for treatment through linearization. Next, we consider generalized solutions with the aid of generating functions. Then we draw these ideas together in §3c, stating the relationship between elliptic–hyperbolic transitions, projection singularities and the pull-back of the Lychagin–Rubtsov metric. The section ends with an account of characteristic surfaces in terms of the pull-back metric.

(a) Classical solutions

We begin by recalling a classical definition from the theory of PDEs (e.g. [21]):

Definition 3.2 (Type of a linear equation). A second-order linear PDE with principal part

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}, \tag{3.5}$$

is called *elliptic* if the eigenvalues of the symmetric matrix $A = [a_{ij}]$ have the same sign, *hyperbolic* if one eigenvalue has the opposite sign from the others and *parabolic* if there is at least one zero eigenvalue.

The notion of an equation type has been generalized to nonlinear equations by Harvey & Lawson [22] as follows (see also §2 of [23] for a geometrical perspective).

Definition 3.3. The type of a nonlinear equation *at* a given solution is the type of its linearization about the solution.

We are thus led to consider the linearization of equation (2.11) about a fixed solution. Let $P + \delta P$ be a perturbation of some exact solution P to (2.11). Introducing this ansatz into equation (2.11) and using the Jacobi formula for determinants leads, to the first order in δP, to the linear equation satisfied by the perturbation field,

$$\mathrm{Tr}\ [\mathrm{adj}(\mathrm{Hess}(P))\mathrm{Hess}(\delta P)] = 0. \tag{3.6}$$

Here, the coefficient matrix is

$$A = \mathrm{adj}(\mathrm{Hess}(P)), \tag{3.7}$$

where 'adj' denotes the adjugate matrix.

The assumption that $q_g > 0$ implies that (3.6) is elliptic if P is (spatially) convex and hyperbolic if P is saddle shaped. Definition 3.3 allows us to bring this information to the nonlinear equation (2.11) as it stands. Also note that equation (2.11) is nowhere parabolic as long as classical solutions

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 479: 20220682

are considered. In fact, equation (2.11) itself prevents the eigenvalues of Hess(P) (and thus those of A) from vanishing. We are now in a position to prove the following:

Proposition 3.4. *Let P be a classical solution to (2.11) and let $L \subset T^*\mathbb{R}^3$ denote the graph of dP. Then, the pull-back Lychagin–Rubtsov metric on L,*

$$h_\omega := (dP)^* g_\omega, \tag{3.8}$$

has matrix representation

$$h_\omega = 2\mathrm{adj}(A) = 2\epsilon q_g Hess(P), \tag{3.9}$$

where A is the linearization matrix (3.7). Moreover, h_ω has signature $(3,0)$ if equation (2.11) is elliptic at the solution P and signature $(1,2)$ if (2.11) is hyperbolic at P.

Proof. By direct calculation. Recalling that,

$$dP : (x,y,z) \mapsto \left(x, y, z, \frac{\partial P}{\partial x}, \frac{\partial P}{\partial y}, \frac{\partial P}{\partial z}\right), \tag{3.10}$$

equation (3.8) implies

$$h_\omega = (dP)^* g_\omega = 2\epsilon q_g \frac{\partial^2 P}{\partial x^i \partial x^j}\, \mathrm{d}x^i\, \mathrm{d}x^j, \tag{3.11}$$

where we have set $(x^1, x^2, x^3) \equiv (x, y, z)$, and summation on repeated indices is implied. Therefore, h_ω has matrix representation

$$h_\omega = 2\epsilon q_g \mathrm{Hess}(P). \tag{3.12}$$

On the other hand, equation (3.7) plus the algebraic identity

$$\mathrm{adj}(\mathrm{adj}(M)) = \det(M)^{n-2} M, \tag{3.13}$$

holding for any square matrix $M \in \mathbb{R}^{n\times n}$ with $n > 2$, gives

$$\mathrm{adj}(A) = \mathrm{adj}(\mathrm{adj}(\mathrm{Hess}(P))) = \det(\mathrm{Hess}(P))\mathrm{Hess}(P) = \epsilon q_g \mathrm{Hess}(P), \tag{3.14}$$

and thus, equation (3.9) follows. As for the second part, we observe that the determinant,

$$\det(h_\omega) = 8\det(A)^2 = 8\det \mathrm{Hess}(P)^4 = 8(\epsilon q_g)^4, \tag{3.15}$$

is always positive, so the eigenvalues of (3.9) can only be (i) all positive or (ii) one positive and two negative. According to (3.9), (i) occurs when P is convex and case (ii) occurs when P is saddle shaped. Therefore, cases (i) and (ii) correspond to (2.11) being respectively elliptic or hyperbolic. ∎

Remark 3.5. The first equality in (3.9) turns out to be a general property of symplectic three-dimensional MAEs. In particular, it applies to the dual equations (2.30), (2.31) and (2.32).

Definition 3.3 is no longer directly applicable when generalized solutions are allowed, as the whole linearization process is ill defined. However, equation (3.9) suggests a characterization of ellipticity of equation (2.11) based on h_ω, which applies to generalized solutions too. Indeed, the equation type at a generalized solution is directly traceable to the signature of h_ω in a definite way. In the remainder of this section, we will prove the consistency of this characterization by relying on the local description of generalized solutions in terms of generating functions.

(b) Generalized solutions

We recall that any generating function f of a generalized solution L satisfies a MAE which arises from expressing the condition $\omega|_L = 0$ in local coordinates on L. Moreover, as f is a classical solution to this equation, there are no obstructions to linearization. Thus, we can give the following

Definition 3.6. Let L be a generalized solution to (2.11) locally generated by a generating function f. We say that (2.11) is elliptic, parabolic or hyperbolic at some point $e \in L$ if f satisfies a MAE of the same type at the point.

We remark that the symbol type of a differential equation is invariant under a change of variables [21], and this ensures the consistency of definition 3.6. We are thus in a position to prove the

Proposition 3.7. *Let $\iota : L \mapsto T^*\mathbb{R}^3$ be a generalized solution to (2.11). Then, the pull-back metric $h_\omega = \iota^* g_\omega$ has signature $(3,0)$ on elliptic branches of L, $(1,2)$ on hyperbolic branches and degenerates along parabolic branches.*

Proof. This proposition is proved by direct inspection of the linearized MAE satisfied by the generating function f. We explicitly carry out the calculations for the case of $f = Zz + T$ (the remaining cases are addressed similarly and lead to the same conclusions). Thus, let L be some generalized solution locally described by $T(x, y, Z)$ according to (2.29). Introducing a perturbation $T + \delta T$ of an exact solution T to (2.32) leads to a linear equation satisfied by the perturbation field,

$$\frac{\partial^2 T}{\partial x^2}\frac{\partial^2 \delta T}{\partial y^2} + \frac{\partial^2 T}{\partial y^2}\frac{\partial^2 \delta T}{\partial x^2} - 2\frac{\partial^2 T}{\partial x \partial y}\frac{\partial^2 \delta T}{\partial x \partial y} + \epsilon q_g \frac{\partial^2 \delta T}{\partial Z^2} = 0. \tag{3.16}$$

Its 3×3 coefficient matrix is

$$A = \begin{pmatrix} \mathrm{adj}(H) & 0 \\ 0 & \epsilon q_g \end{pmatrix} \quad \text{and} \quad H := \begin{pmatrix} \dfrac{\partial^2 T}{\partial x^2} & \dfrac{\partial^2 T}{\partial x \partial y} \\ \dfrac{\partial^2 T}{\partial x \partial y} & \dfrac{\partial^2 T}{\partial y^2} \end{pmatrix}. \tag{3.17}$$

On the other hand, the Lychagin–Rubtsov metric $h_\omega = g_\omega|_L$ has the local coordinate expression

$$h_\omega = 2\epsilon q_g \left(\frac{\partial^2 T}{\partial x^2}\,\mathrm{d}x^2 + 2\frac{\partial^2 T}{\partial x \partial y}\,\mathrm{d}x\,\mathrm{d}y + \frac{\partial^2 T}{\partial y^2}\,\mathrm{d}y^2 - \frac{\partial^2 T}{\partial Z^2}\,\mathrm{d}Z^2 \right) \tag{3.18}$$

and may be written in the matrix form as follows:

$$h_\omega = 2\begin{pmatrix} \epsilon q_g H & 0 \\ 0 & \det(H) \end{pmatrix}, \tag{3.19}$$

where we have used (2.32). We note in passing that

$$h_\omega = 2\mathrm{adj}(A). \tag{3.20}$$

We see from (3.17) that equation (2.32) is elliptic as long as H is positive definite, which, in light of (3.19), corresponds to h_ω having signature $(3,0)$. Parabolic and hyperbolic cases correspond to $\det(H) = 0$ and $\det(H) < 0$, respectively. Therefore, it follows from equation (3.19) that h_ω is degenerate on parabolic branches and of type $(1,2)$ on hyperbolic ones. ∎

The significance of proposition 3.7 is that h_ω encodes all the essential information about the equation type and may be used to give an invariant definition of the symbol type based on its signature. This may be summarized as follows:

Definition 3.8. Let L be a generalized solution to (2.11) with Lychagin–Rubtsov metric h_ω. We say that (2.11) is elliptic or hyperbolic at a point $e \in L$ if $(h_\omega)_e$ is, respectively, of type $(3,0)$ or $(1,2)$. We say that (2.11) is parabolic at $e \in L$ if $(h_\omega)_e$ is degenerate.

(c) Singularities

In this section, we show that elliptic–hyperbolic transitions in (2.11) can only occur along through some singularity. This result is closely related to the assumption of strictly positive potential vorticity, which, according to (2.11), prevents the eigenvalues of Hess(P) from changing sign as

long as classical solutions are considered. As a result, the pull-back of the Lychagin–Rubtsov metric degenerates on singularities.

We recall that points on L are called *regular* if the tangent map $\mathrm{d}\pi_L$ is surjective, and *singular* otherwise. We denote by ΣL the set of singular points on L.

Proposition 3.9. *Let L be a generalized solution to (2.11). Then the set of parabolic points on L coincides with the singular locus ΣL.*

Proof. Once again, we rely on a local description in coordinates and generating functions to prove our result. Let a solution L be locally generated by a function $f = Zz + T$, that is

$$L = \left\{ (x, y, z, X, Y, Z) \in T^*\mathbb{R}^3 : X = \frac{\partial T}{\partial x}, \quad Y = \frac{\partial T}{\partial y}, \; z = -\frac{\partial T}{\partial Z} \right\}. \tag{3.21}$$

In local coordinates $\{x, y, Z\}$ on L, the projection mapping reads

$$\pi_L(x, y, Z) = \left(x, y, -\frac{\partial T}{\partial Z} \right), \tag{3.22}$$

and so it is singular on points satisfying

$$\det(\mathrm{d}\pi_L) = -\frac{\partial^2 T}{\partial Z^2} = (\epsilon q_g)^{-1} \left(\frac{\partial^2 T}{\partial x^2} \frac{\partial^2 T}{\partial y^2} - \left(\frac{\partial^2 T}{\partial x \partial y} \right)^2 \right) = 0. \tag{3.23}$$

On the other hand, we know from the proof of proposition 3.7 that parabolic points on L satisfy the same equation. To complete the proof, one should examine in turn each of the remaining classes of generating functions. However, calculations are almost identical to those we have already exhibited, and we omit them for conciseness. ∎

Propositions 3.7 and 3.9 are combined to

Corollary 3.10. *Given a generalized solution L to (2.11), the induced Lychagin–Rubtsov metric on L degenerates along the singular locus ΣL.*

Thus, every regular branch of a multi-valued solution $L \subset T^*\mathbb{R}^3$ is of a single type (elliptic or hyperbolic), and transitions are only possible in passing from one branch to another.

Remark 3.11. Chynoweth & Sewell's approach to singularities is by the Legendre transform (see equations (12) of [4]). Once a solution to (2.30), (2.31) or (2.32) is known, the (possibly multi-valued) geopotential is recovered by the inverse Legendre transform,

$$P = Xx + Yy + Zz - R, \quad x = \frac{\partial R}{\partial X}, \; y = \frac{\partial R}{\partial Y}, \; z = \frac{\partial R}{\partial Z}, \tag{3.24}$$

$$P = Xx + Yy - S, \quad x = \frac{\partial S}{\partial X}, \; y = \frac{\partial S}{\partial Y} \tag{3.25}$$

and

$$P = Zz + T, \quad z = -\frac{\partial T}{\partial Z}. \tag{3.26}$$

Singularities are then identified according to their effects on the graph of the multi-valued P. Chynoweth and Sewell's viewpoint is reconciled with the geometric viewpoint as follows. We denote by $J^1(\mathbb{R}^3) \simeq T^*\mathbb{R}^3 \times \mathbb{R}$ the bundle of 1-jets over the physical space $\mathbb{R}^3(x, y, z)$, and we endow it with coordinates

$$(x, y, z, u, X, Y, Z). \tag{3.27}$$

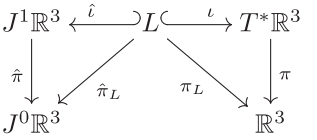
In this extended space, a generalized solution is understood as a Legendrian submanifold[3] $\hat{\imath} : L \to J^1\mathbb{R}^3$ such that $\hat{\imath}^*\omega = 0$ (here, ω is understood as a 3-form on $J^1\mathbb{R}^3$). Then, the projection $\hat{\pi}_L = \hat{\pi} \circ \hat{\imath}$

[3]A Legendrian submanifold of a contact manifold (M^{2n+1}, C) is a n-dimensional integral manifold of the contact distribution C. If $M = J^1\mathbb{R}^3$, C is canonically described as the kernel of the Cartan 1-form $\Theta = du - Xdx - Ydy - Zdz$ (e.g. [8]).

of L to the base space $J^0\mathbb{R}^3$ of the jet bundle, where

$$\hat{\pi} : J^1\mathbb{R}^3 \to J^0\mathbb{R}^3, \quad (x, y, z, u, X, Y, Z) \mapsto (x, y, z, u), \tag{3.28}$$

results in the graph of the (possibly) multi-valued geopotential P, parametrized by the local coordinates on L by equations (3.24), (3.25) or (3.26). The following commutative diagram summarizes the situation.



(d) Characteristic variety

The strong connection between h_ω and the symbol type of (2.11) suggests a link with the characteristic surfaces as well. In this section, we explore the geometry of the light cone of h_ω and use it to introduce a suitable notion of characteristic surfaces in hyperbolic and parabolic regime.

A central role in this subject is played by vectors of null length, which, borrowing terminology from relativity theory, are called *light-like*. The set of light-like vectors based at a point is called the *light cone* or the *characteristic variety*. This notion is made precise by the following definition, which builds on the work of Kossowski [13] on two-dimensional MAEs.

Definition 3.12. Let g_ω be given by (3.2) and let $e \in T^*\mathbb{R}^3$. The cotangent characteristic variety (or simply characteristic variety) $CV_e \subset T_e(T^*\mathbb{R}^3)$ is the cone

$$CV_e := \{\xi \in T_e(T^*\mathbb{R}^3) : g_\omega(\xi, \xi) = 0\}. \tag{3.29}$$

Let $L \subset T^*\mathbb{R}^3$ be some generalized solution to the (2.11) and let $e \in L$. We denote the pull-back of the characteristic variety to L as follows:

$$CV_e^L := \{\xi \in T_eL : h_\omega(\xi, \xi) = 0\}, \tag{3.30}$$

where $h_\omega = g_\omega|_L$ is the pull-back metric on L.

It easily follows from definition 3.8 that the characteristic variety CV_e^L is a full-fledged cone if e is an hyperbolic point, a degenerate cone if e is a parabolic point and the zero vector, $\{0\} \subset T_eL$, if e is an elliptic point. The characteristic variety is the basic ingredient to build the characteristic surfaces within a generalized solution L. We understand a characteristic surface $\mathcal{C} \subset L$ as the enveloping surface of characteristic varieties CV_e^L as e varies across $\mathcal{C}$, as the following definition clarifies.

Definition 3.13. A surface $\mathcal{C} \subset L$ is called *characteristic* if at any point $e \in \mathcal{C}$, the tangent space $T_e\mathcal{C}$ comprises one (and only one) light-like direction.

We remark that definition 3.13 closely parallels the notion of characteristics in general relativity, where they are identified with light-like surfaces [24]. Definition 3.13 may be considered as a straightforward generalization to nonlinear PDEs of the classical notion of characteristics for linear PDEs (see appendix A). To motivate this statement, fix coordinates $\{q^1, q^2, q^3\}$ on L and consider the surface

$$\mathcal{C} = \{(q^1, q^2, q^3) \in L : F(q^1, q^2, q^3) = 0\}. \tag{3.31}$$

Further, consider the following vector based at points on $\mathcal{C}$,

$$dF^\sharp = h^{ij} \frac{\partial F}{\partial q^i} \frac{\partial}{\partial q^j}, \tag{3.32}$$

where summation on repeated indices is implied and h^{ij} denotes the components of the inverse metric h_ω^{-1}. It is straightforward to check the equivalence of the following statements: (i) $dF^\sharp$ is a

tangent vector to $\mathcal{C}$, (ii) $dF^\sharp$ is a light-like vector and (iii) F satisfies

$$h_\omega^{-1}(dF, dF) = 0. \tag{3.33}$$

Equation (3.33) is the analogue of the eikonal equation (A 3) in the linear setting, where the matrix coefficient is replaced by the inverse of h_ω. We shall further elaborate on this analogy next. Note that equation (3.33) is only well defined away from parabolic points, where h_ω is not invertible and has only non-trivial solutions on hyperbolic branches of L. Any hyperbolic branch is regular by virtue of proposition 3.9, so it may be described as the graph of a function, e.g. $P^*(x, y, z)$, to the extent a classical solution is (we may take $\{x, y, z\}$ as local coordinates on hyperbolic branches). Therefore, the linearization of (2.11) about a hyperbolic branch of L is well defined, and, building on equation (3.9), we may write

$$h_\omega^{-1} = \frac{1}{2\det(A)} A, \tag{3.34}$$

where $A = \mathrm{adj}(\mathrm{Hess}(\mathrm{P}^*))$ is the coefficient matrix of the linearized equation (2.11) about P^*. Since $\det(A) \neq 0$ on hyperbolic points, we may get rid of this term in equation (3.33), and write

$$\nabla F \cdot A \nabla F = \sum_{i,j=1}^{3} a_{ij}(x, y, z) \frac{\partial F}{\partial x^i} \frac{\partial F}{\partial x^j} = 0, \tag{3.35}$$

where $(x^1, x^2, x^3) = (x, y, z)$. The analogy with (A 3) should now be apparent.

Equation (3.33) is a nonlinear PDE of the first order and, as (A 3), is solved by the classical methods of wave optics (e.g. Appendix 4 of [25]). Namely, the solution surface is understood as foliated by light-like curves (i.e. the light rays of wave optics) which satisfy a set of Hamilton's canonical equations of motion. We briefly recall the main steps of the solution procedure for completeness of exposition [21,25]. Consider the cotangent bundle T^*L with coordinates $\{q^1, q^2, q^3, p_1, p_2, p_3\}$ and symplectic structure

$$\Omega = dp_i \wedge \mathrm{d}q^i. \tag{3.36}$$

Equation (3.33) is thus interpreted as the zero level set of the Hamiltonian function $\mathcal{H} : T^*L \to \mathbb{R}$,

$$\mathcal{H}(p, q) = h(q)^{ij} p_i p_j, \tag{3.37}$$

under the identification $p_i = \partial F / \partial q^i$. Characteristic curves of (3.33) are defined by [25] as the integral curves of the Hamiltonian vector field $\xi_\mathcal{H}$,

$$-d\mathcal{H} = \iota_{\xi_\mathcal{H}} \Omega, \tag{3.38}$$

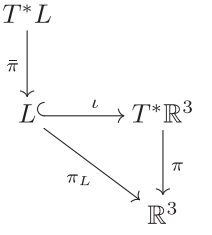
and satisfy the Hamilton's canonical equations:

$$\dot{q}^i = \frac{\partial \mathcal{H}}{\partial p_i} = 2(h_\omega)^{ij} p_j \quad \text{and} \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q^i} = -\frac{\partial (h_\omega)^{jk}}{\partial q^i} p_j p_k. \tag{3.39}$$

Initial conditions are not free, but are subject to the condition

$$\mathcal{H}(p(0), q(0)) = 0. \tag{3.40}$$

Integral curves of (3.39) are called *bicharacteristics* [16]. Once projected to L along the cotangent bundle, $\bar{\pi} : T^*L \to L$, bicharacteristics foliate the characteristic surfaces $\mathcal{C} \subset L$. The following commutative diagram summarizes the relations introduced so far.



royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 479: 20220682

Remark 3.14. Light rays are equivalently described by the Lagrangian

$$\mathcal{L} = h_{ij}(q)\dot{q}^i\dot{q}^j, \tag{3.41}$$

related to the Hamiltonian (3.37) by the classical Legendre transform. The equivalent of (3.40) reads

$$\mathcal{L}(q(0), \dot{q}(0)) = 0. \tag{3.42}$$

The Euler–Lagrange equations associated with (3.41) plus condition (3.42) precisely yield the light-like geodesics of h_ω. Therefore, characteristic surfaces $\mathcal{C}$ are foliated by light-like geodesics of h_ω, and this offers an alternative approach to computing them.

In the next section, we provide an example of a generalized solution to the SG equations and show the interaction of characteristics, elliptic–hyperbolic transition and singularities.

4. Exact solutions

Not many exact solutions to the full SG system (2.9) are known [26,27], and even fewer are the generalized ones. A common assumption often encountered in the literature is that of uniform potential vorticity, which helps finding particular solutions and possesses physical relevance. Under this assumption, equation (2.11) decouples from system (2.9) and is interpreted as a MAE for the unknown geopotential. The choice $\epsilon q_g = 1$ brings (2.11) to the form

$$\det \mathrm{Hess}(P) = 1, \tag{4.1}$$

widely studied in the literature. A two-parameter family of classical solutions to (4.1) is introduced in [18] and generalized in [28]. Once a particular solution to (4.1) is selected, one is able to build a full solution to (2.9), as we show in §4.

(a) Construction of exact solutions

We build generalized solutions to (4.1) by solving (2.32), which, under the assumption $\epsilon q_g = 1$, takes the form

$$\frac{\partial^2 T}{\partial x^2}\frac{\partial^2 T}{\partial y^2} - \left(\frac{\partial^2 T}{\partial x \partial y}\right)^2 + \frac{\partial^2 T}{\partial Z^2} = 0. \tag{4.2}$$

Although time dependence is still possible for constant vorticity flows, we shall restrict to stationary solutions, which nevertheless show some interesting features. We look for analytical solutions to (4.2) in the form

$$T = T^{(0)}(Z) + T^{(1)}_\alpha(Z)x^\alpha + \frac{1}{2}T^{(2)}_{\alpha\beta}(Z)x^\alpha x^\beta + \frac{1}{3!}T^{(3)}_{\alpha\beta\gamma}(Z)x^\alpha x^\beta x^\gamma + \dots, \tag{4.3}$$

where $(x^1, x^2) = (x, y)$ and summation on repeated indices is implied. Several classes of finite dimensional reductions of (4.3) are possible. Third-order truncation of the aforementioned series provides a wide class of fully polynomial solutions whose coefficients satisfy

$$\frac{d^2 T^{(3)}_{\alpha\beta\gamma}}{dZ^2} = 0, \tag{4.4}$$

$$\frac{d^2 T^{(2)}_{11}}{dZ^2} + 2\det(T^{(3)}_{1\alpha\beta}) = 0, \quad \frac{d^2 T^{(2)}_{12}}{dZ^2} + T^{(3)}_{111}T^{(3)}_{222} - T^{(3)}_{112}T^{(3)}_{122} = 0, \quad \frac{d^2 T^{(2)}_{22}}{dZ^2} + 2\det(T^{(3)}_{2\alpha\beta}) = 0, \tag{4.5}$$

$$\frac{d^2 T^{(1)}_1}{dZ^2} + T^{(2)}_{22}T^{(3)}_{111} - 2T^{(2)}_{12}T^{(3)}_{112} + T^{(2)}_{11}T^{(3)}_{222} = 0, \quad \frac{d^2 T^{(1)}_2}{dZ^2} + T^{(2)}_{22}T^{(3)}_{112} - 2T^{(2)}_{12}T^{(3)}_{122} + T^{(2)}_{11}T^{(3)}_{222} = 0 \tag{4.6}$$

and

$$\frac{d^2 T^{(0)}}{dZ^2} + \det(T^{(2)}_{\alpha\beta}) = 0. \tag{4.7}$$

A straightforward inspection of equations (4.4)–(4.7) shows that the coefficient functions are polynomials of a definite degree in Z. Specifically, $T^{(3)}_{\alpha\beta\gamma}$ has degree 1, $T^{(2)}_{\alpha\beta}$ has degree 4, $T^{(1)}_{\alpha}$ has degree 7 and $T^{(0)}$ has degree 10. Retaining terms of the fourth order and higher in the expansion (4.3) leads to a wider class of exact solutions, though they are generally not polynomial in the vertical variable.

Polynomial solutions are particularly valuable for their ability to encode the local behaviour of more complex solutions and, in particular, the singular structure. The simplest non-trivial Lagrangian singularity is the fold (A_2), and the germ of a Lagrangian submanifold with this feature is canonically described by a generating function:

$$T^*(x,y,Z)=\frac{Z^3}{6} \tag{4.8}$$

(see, e.g. [29] for a list of low-dimensional canonical forms of elementary catastrophes). We build an example solution to (4.2) by deformation of (4.8) through the addition of a polynomial term. One of the simplest choices is

$$T=\frac{y^2}{2}-\frac{x^2Z}{2}+\frac{Z^3}{6}. \tag{4.9}$$

(b) Lagrangian submanifold and projection

The generating function (4.9) determines a Lagrangian submanifold according to

$$L=\left\{(x,y,z,X,Y,Z)\in T^*\mathbb{R}^3 : X=\frac{\partial T}{\partial x}=xZ, Y=\frac{\partial T}{\partial y}=y, z=-\frac{\partial T}{\partial Z}=\frac{x^2}{2}-\frac{Z^2}{2}\right\}. \tag{4.10}$$

The restriction $\pi_L:=\pi|_L$ of the bundle projection (2.20) to L in local coordinates is

$$\pi_L(x,y,Z)=\left(x,y,-\frac{\partial T}{\partial Z}\right)=\left(x,y,\frac{x^2}{2}-\frac{Z^2}{2}\right) \tag{4.11}$$

and shows that the fold singularity occurs across the plane

$$\Sigma L=\left\{(x,y,Z)\in L : 0=\det(\mathrm{d}\pi_L)=-\frac{\partial^2 T}{\partial Z^2}=-Z\right\}. \tag{4.12}$$

The projection of the singular locus to the physical space identifies the caustics,

$$\pi_L(\Sigma L)=\left\{(x,y,z)\in\mathbb{R}^3 : z=\frac{x^2}{2}\right\}, \tag{4.13}$$

and figure 1 provides a view of them. Due to the nature itself of the fold singularity, the caustics bound the solution domain in the physical space. In other words, the solution is only defined in the domain $z\le x^2/2$ of $\mathbb{R}^3(x,y,z)$.

Remark 4.1. A different approach is used in [4], where the authors put the focus on the singularities of the geopotential graph. The Chynoweth–Sewell relations (3.24)–(3.26) yield the graph of the multi-valued geopotential $P(x,y,z)$ as parametrized by Z,

$$\left.\begin{aligned}P&=\frac{y^2}{2}-\frac{Z^3}{3}\\ \text{and}\qquad z&=\frac{x^2}{2}-\frac{Z^2}{2},\end{aligned}\right\} \tag{4.14}$$

and figure 2 shows a section of it for $y=$ constant. The cusped edge is a distinctive feature of the A_2 singularity in the *Legendrian* context.

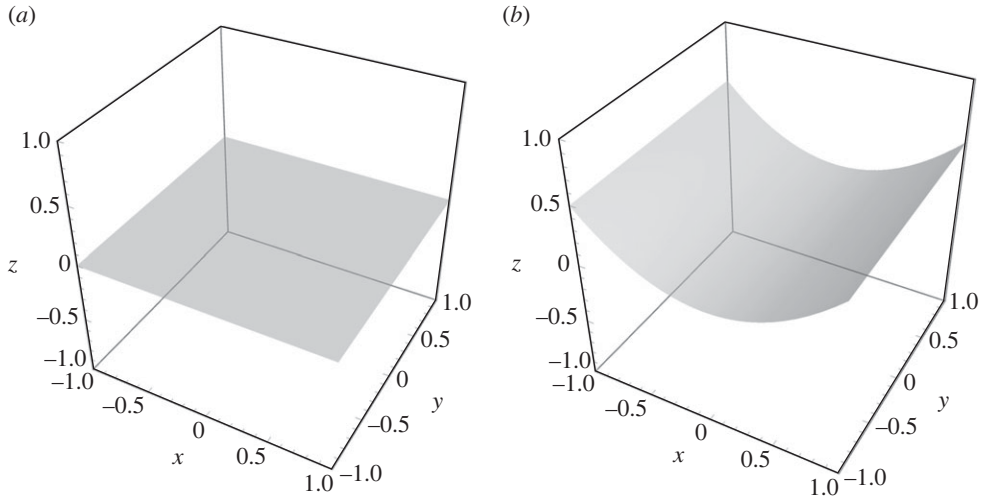


Figure 1. (*a*) The singular locus ΣL defined in (4.12) within the generalized solution L given by (4.10). (*b*) The caustics $\pi_L(\Sigma L)$ defined in (4.13) showing the typical appearance of the A_2 singularity.

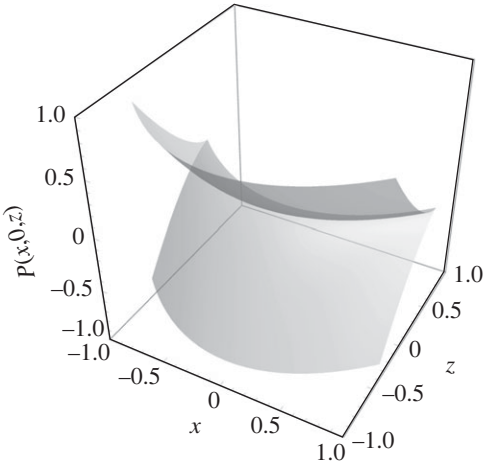


Figure 2. A slice through the graph of the multi-valued geopotential (4.14) for $y = 0$.

(c) Lychagin–Rubtsov metric

Next we examine the pseudo-Riemannian geometry of the solution. The Lychagin–Rubtsov metric on L in local coordinates is

$$\begin{aligned} h_\omega &= 2(T_{xx}\,\mathrm{d}x^2 + 2T_{xy}\,\mathrm{d}x\,\mathrm{d}y + T_{yy}\,\mathrm{d}y^2 - T_{ZZ}\,\mathrm{d}Z^2) \\ &= 2(-Z\,\mathrm{d}x^2 + \mathrm{d}y^2 - Z\,\mathrm{d}Z^2). \end{aligned} \tag{4.15}$$

This immediately implies, according to definition 3.8, that the problem is elliptic for $Z < 0$ and hyperbolic for $Z > 0$. Characteristic surfaces in the hyperbolic region are determined by the light-like geodesics of (4.15) (see remark 3.14). General geodesics satisfy

$$\ddot{x} = -\frac{\dot{x}\dot{Z}}{Z}, \quad \ddot{y} = 0, \quad \ddot{Z} = \frac{\dot{x}^2 - \dot{Z}^2}{2Z}. \tag{4.16}$$

The first two equations can be immediately integrated once and yield two of constants of the motion,

$$\dot{x}Z = C_1 \quad \text{and} \quad \dot{y} = C_2. \tag{4.17}$$

By using (4.17) in the third of (4.16), the problem of finding general geodesics is reduced to the single equation

$$2Z^3\ddot{Z} = C_1^2 - Z^2\dot{Z}^2. \tag{4.18}$$

A further simplification is available for light-like geodesics, which are subject to the additional constraint

$$-Z\dot{x}^2 + \dot{y}^2 - Z\dot{Z}^2 = 0. \tag{4.19}$$

Using (4.17) in (4.19) gives the separable equation:

$$C_1^2 - C_2^2 Z + Z^2\dot{Z}^2 = 0, \tag{4.20}$$

with implicit solution,

$$\pm s = \frac{2\sqrt{C_2^2 Z - C_1^2}(2C_1^2 + C_2^2 Z)}{3C_2^4} - \frac{2\sqrt{C_2^2 Z_0 - C_1^2}(2C_1^2 + C_2^2 Z_0)}{3C_2^4}. \tag{4.21}$$

We shall remark at this point that (4.19) is compatible with the geodesics equations (4.16), as can be verified by taking a derivative of (4.19) and using (4.16) to eliminate the second derivative terms. Indeed, any solution of (4.20) is a geodesic curve. This can be checked by taking the derivative of (4.20) with respect to the curve parameter to get

$$-C_2^2 + 2Z\dot{Z}^2 + 2Z^2\ddot{Z} = 0. \tag{4.22}$$

Then, multiplication by Z and the use of (4.19) to eliminate C_2 leads back to equation (4.18). Next, we use the first constant of the motion in (4.17) to get $x(s)$,

$$\frac{\mathrm{d}x}{\mathrm{d}Z} = \frac{C_1}{Z}\frac{\mathrm{d}s}{\mathrm{d}Z} = \pm\frac{C_1}{\sqrt{C_2^2 Z - C_1^2}}. \tag{4.23}$$

Once integrated, equations (4.23) and (4.21) and the second of equation (4.17) allow us to write the light-like geodesics as follows:

$$\left.\begin{aligned} &x \pm \frac{2\dot{x}_0 Z_0}{\dot{y}_0^2}\sqrt{\dot{y}_0^2 Z - \dot{x}_0^2 Z_0^2} = \text{constant} \\ \text{and}\quad &y \pm \frac{2}{3\dot{y}_0^3}(2\dot{x}_0^2 Z_0^2 + \dot{y}_0^2 Z)\sqrt{\dot{y}_0^2 Z - \dot{x}_0^2 Z_0^2} = \text{constant}, \end{aligned}\right\} \tag{4.24}$$

where we have used $C_1 = \dot{x}_0 Z_0$, $C_2 = \dot{y}_0$. To see how characteristics interact with the parabolic boundary ΣL, we assume next that the geodesic starting point (x_0, y_0, Z_0) belongs to ΣL, i.e. $Z_0 = 0$. This results in

$$\left.\begin{aligned} &x = x_0 \\ \text{and}\quad &(y - y_0)^2 = \frac{4}{9}Z^3. \end{aligned}\right\} \tag{4.25}$$

Thus, characteristics intersecting the parabolic boundary form semicubical cusps at the intersection point, as expected from the literature on the subject [5]. Figure 3 provides a view of the characteristic surfaces near the singular locus.

(d) Full semigeostrophic solution

In the remainder of this section, we explicitly reconstruct the full SG solution from the knowledge of the generating function (4.9). Thanks to the simple structure of the solution (4.9), we are able to

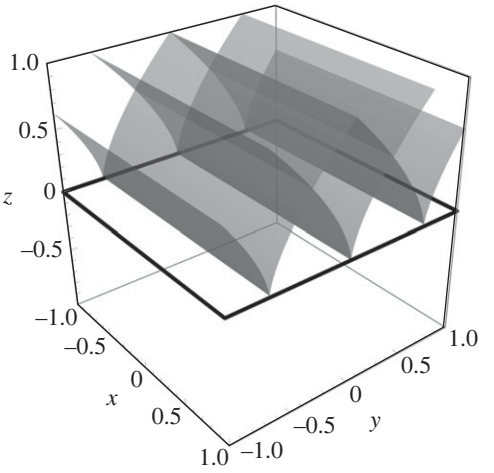


Figure 3. Characteristic surfaces (grey) intersecting the boundary of the hyperbolic region, $Z = 0$. Intersection points are found along parallel lines along the parabolic plane $Z = 0$ where the characteristics form semicubical cusps.

explicitly (piecewise) invert the relation (4.14) and write

$$\epsilon\theta = Z = \pm\sqrt{x^2 - 2z} \quad \text{and} \quad P = \frac{y^2}{2} \mp \frac{1}{3}(x^2 - 2z)^{3/2}. \tag{4.26}$$

Once the geopotential P is known, absolute momentum and potential temperature are obtained by derivation as follows:

$$M = \frac{\partial P}{\partial x} = \mp x\sqrt{x^2 - 2z}, \qquad N = \frac{\partial P}{\partial y} = y. \tag{4.27}$$

Next, the geostrophic wind is found as follows:

$$u_g = q_g(y - N) = 0 \quad \text{and} \quad v_g = q_g(M - x) = q_g(\mp x\sqrt{x^2 - 2z} - x), \tag{4.28}$$

where we have used $\epsilon q_g = 1$. The momentum balance equations plus the transport of potential temperature yield a system of algebraic equations for the unknown components of the velocity field,

$$\left.\begin{aligned} &M_x u + M_y v + M_z w = u_g, \\ &N_x u + N_y v + N_z w = v_g \\ \text{and} \quad &\theta_x u + \theta_y v + \theta_z w = 0. \end{aligned}\right\} \tag{4.29}$$

Since $N = y$, it easily follows that $v = v_g$. Moreover, $M_y = \theta_y = 0$ and $u_g = 0$, which imply that $u = w = 0$. Indeed, the first and the last equations in (4.29) form a linear homogeneous system with non-degenerate coefficient matrix as follows:

$$\frac{\partial(M, \theta)}{\partial(x, y)} = \det \mathrm{Hess}(P) = 1. \tag{4.30}$$

In conclusion, the flow field corresponding to (4.9) is a purely geostrophic meridional wind,

$$u = 0, \quad v = v_g = q_g(\mp x\sqrt{x^2 - 2z} - x), \quad w = 0. \tag{4.31}$$

To restore single-valuedness of the solution, Chynoweth and Sewell appealed to the convexity principle of [6]. Namely, only convex branches of the multivalued graph of P are retained while concave ones are discarded. Although the application of this principle requires some attention in

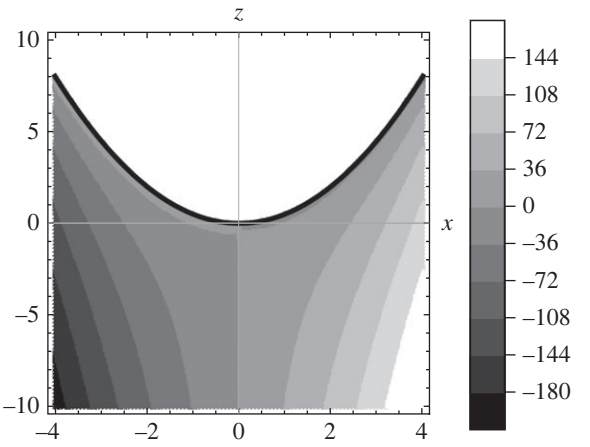


Figure 4. Magnitude of the geostrophic wind (4.31) on a section normal to the flow. The fluid domain is bounded from above by the caustics (4.13).

the general case, it is straightforward in this example. The admissible branch of P is found to be,

$$P = \frac{y^2}{2} + \frac{1}{3}(x^2 - 2z)^{3/2}, \tag{4.32}$$

which corresponds to the elliptic branch of the multi-valued solution, $Z < 0$. This corresponds to the velocity field

$$u = 0, \quad v = v_g = q_g(x\sqrt{x^2 - 2z} - x), \quad w = 0, \tag{4.33}$$

and represents a geostrophic wind in the northern hemisphere directed poleward. Figure 4 shows the wind magnitude on a section normal to the flow.

This example shares qualitative features with a larger class of exact solutions, i.e. two-dimensional stationary flows. These flows are characterized by the independence of the geopotential ϕ of one of the horizontal coordinates (in this case, y), which results in a vanishing zonal component of the geostrophic wind. Under stationary conditions, flows in this class are purely geostrophic (either zonal or meridional).

5. Conclusions and future directions

The Lychagin–Rubtsov metric has been much studied in the context of Monge–Ampère geometry, but its pull-back to generalized solutions, realized as Lagrangian submanifolds, has hitherto been unexplored. We have explored this feature from the point of view of PDE theory in the physically and mathematically important example of the SG equations.

In particular, we have shown connections between the signature of the pull-back metric on solutions, the symbol type of the MAE, and its role in describing the characteristic surfaces of hyperbolic equations. We recognize the pull-back metric as a tool for studying singularities, which complements and extends the earlier work of Kossowski [13], where a version of the Lychagin–Rubtsov metric on $T^*\mathbb{R}^2$ was the primary object of interest.

Several questions are still open. We illuminated the meaning of the light-like geodesics in hyperbolic regime, but the potential role of space-like and time-like geodesics, and the elliptic regime in this context, remain to be explored.

Another intriguing question is the geometrical and physical meaning of the curvature of the Lagrangian submanifolds, and its relationship with singularities. This aspect has been explored in the work of Napper *et al.* on Navier–Stokes equations [30], and its implications for SG theory is a matter for future research. We focused on the kinematic aspects of the SG equations, considering

time as a fixed parameter, but the system dynamics is important. Considering time-dependent solutions leads to a one-parameter family of metrics, i.e. a notional geometric flow, whose properties are unknown. We might speculate a relation between such a geometric flow and the onset of dynamic singularities.

Data accessibility. No additional research data beyond the data presented and cited in this work are needed to validate the research findings in this work. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

Authors' contributions. R.D.: conceptualization, formal analysis, investigation, methodology and writing—original draft; G.O.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, supervision and writing—review and editing; I.R.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision and writing—review and editing; V.R.: conceptualization, formal analysis, investigation, methodology, resources, supervision and writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. R.D. and G.O. were supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie (grant no. 778010 IPaDEGAN). R.D. and G.O. thank the financial support of the project MMNLP (Mathematical Methods in Non Linear Physics) of the INFN.

Acknowledgements. We would like to thank T. Bridges, L. Napper and M. Wolf for many useful discussions. R.D. and G.O. also gratefully acknowledge the auspices of the GNFM Section of INdAM under which part of this work was carried out. This work is part of R.D.'s dual PhD program Bicocca-Surrey.

Appendix A. Characteristics of linear PDEs

We recall here some classical terminology from the theory of linear PDEs (e.g. [21]). Let a second-order linear PDE in n independent variables $x = (x_1, \dots, x_n)$ has a principal part

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}. \tag{A 1}$$

The x-depending quadratic form

$$\sigma(x, \boldsymbol{\xi}) = \sum_{i,j=1}^{n} a_{ij}(x) \xi_i \xi_j, \quad \boldsymbol{\xi} := (\xi_1, \dots, \xi_n) \tag{A 2}$$

is called the *principal symbol* of the equation. A vector $\boldsymbol{\xi}$ based at x is called *characteristic* if $\sigma(x, \boldsymbol{\xi}) = 0$ and the set of characteristic vectors at x is called the *Fresnel cone*. An implicitly defined hypersurface $F(x_1, \dots, x_n) = 0$ is called *characteristic surface* or simply *characteristic* if its normal vector is characteristic. In other words, F satisfies the *eikonal* equation:

$$\sigma(x, \nabla F) = \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial F}{\partial x_i} \frac{\partial F}{\partial x_j} = 0. \tag{A 3}$$

References

1. Cullen MJP, Purser RJ. 1984 An extended Lagrangian theory of semigeostrophic frontogenesis. *J. Atmos. Sci.* **41**, 1477–1497. (doi:10.1175/1520-0469(1984)041<1477:AELTOS>2.0.CO;2)
2. Hoskins BJ, Bretherton FP. 1972 Atmospheric frontogenesis models: mathematical formulation and solution. *J. Atmos. Sci.* **29**, 11–37. (doi:10.1175/1520-0469(1972)029<0011:AFMMFA>2.0.CO;2)
3. Hoskins BJ. 1975 The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.* **32**, 233–242. (doi:10.1175/1520-0469(1975)032<0233:TGMAAT>2.0.CO;2)

4. Chynoweth S, Sewell MJ. 1989 Dual variables in semigeostrophic theory. *Proc. R. Soc. Lond. A* **424**, 155–186. (doi:10.1098/rspa.1989.0074)
5. Landau LD, Lifshitz EM. 1987 *Fluid mechanics—course of theoretical physics*, vol. 6, 2nd edn. Oxford, UK: Pergamon Press.
6. Shutts GJ, Cullen MJP. 1987 Parcel stability and its relation to semigeostrophic theory. *J. Atmos. Sci.* **44**, 1318–1330. (doi:10.1175/1520-0469(1987)044<1318:PSAIRT>2.0.CO;2)
7. Holt MW, Shutts GJ. 1990 An analytical model of the growth of a frontal discontinuity. *Q. J. R. Meteorol. Soc.* **116**, 269–286. (doi:10.1002/qj.49711649203)
8. Kushner A, Lychagin V, Rubtsov V. 2006 *Contact geometry and nonlinear differential equations*. Cambridge, UK: Cambridge University Press.
9. Delahaies S, Roulstone I. 2010 Hyper-Kähler geometry and semi-geostrophic theory. *Proc. R. Soc. A.* **466**, 195–211. (doi:10.1098/rspa.2009.0139)
10. Roulstone I, Banos B, Gibbon JD, Roubtsov VN. 2009 A geometric interpretation of coherent structures in Navier-Stokes flows. *Proc. R. Soc. Lond. A* **465**, 2015–2021. (doi:10.1098/rspa.2008.0483)
11. Banos B, Roubtsov VN, Roulstone I. 2016 Monge–Ampère structures and the geometry of incompressible flows. *J. Phys. A: Math. Theor.* **49**, 244003 (17p). (doi:10.1088/1751-8113/49/24/244003)
12. Vinogradov AM, Kupershmidt BA. 1977 The structures of hamiltonian mechanics. *Russ. Math. Surv.* **32**, 177. (doi:10.1070/RM1977v032n04ABEH001642)
13. Kossowski M. 1991 Local existence of multivalued solutions to analytic symplectic Monge-Ampère equations (the nondegenerate and type changing cases). *Ind. Univ. Math. J.* **40**, 123–148. (doi:10.1512/iumj.1991.40.40006)
14. Ishikawa G, Machida Y. 2006 Singularities of improper affine spheres and surfaces of constant Gaussian curvature. *Int. J. Math.* **17**, 269–293. (doi:10.1142/S0129167X06003485)
15. Ishikawa G, Machida Y. 2006 Extra singularities of geometric solutions to Monge-Ampère equation of three variables. *Kyoto Univ. Res. Inf. Repos.* **1502**, 41–53.
16. Lychagin V. 1985 Singularities of multivalued solutions of nonlinear differential equations, and nonlinear phenomena. *Acta Appl. Math.* **3**, 135–173. (doi:10.1007/BF00580702)
17. Lychagin VV, Rubtsov VN. 1983 Local classification of Monge–Ampère differential equations. *Dokl. Akad. Nauk SSSR* **272**, 34–38.
18. Banos B. 2002 Nondegenerate Monge–Ampère Structures in dimension 6. *Lett. Math. Phys.* **62**, 1–15. (doi:10.1023/A:1021655609692)
19. Oliver M. 2006 Variational asymptotics for rotating shallow water near geostrophy: a transformational approach. *J. Fluid Mech.* **551**, 197–234. (doi:10.1017/S0022112005008256)
20. Roulstone I, Norbury J. 1994 A Hamiltonian structure with contact geometry for the semi-geostrophic equations. *J. Fluid Mech.* **272**, 211–234. (doi:10.1017/S0022112094004441)
21. Courant R, Hilbert D. 1962 *Methods of mathematical physics*, vol. **II**. Partial differential equations. New York: Wiley Interscience.
22. Harvey R, Lawson H. 1982 Calibrated geometry. *Acta Math.* **148**, 47–157. (doi:10.1007/BF02392726)
23. Duzhin SV. 2004 Infinitesimal classification of systems of two first order partial differential equations in two variables. *J. Math. Sci.* **119**, 30–34. (doi:10.1023/B:JOTH.0000008738.73271.8e)
24. Ehlers J, Newmann ET. 2000 The theory of caustics and wave front singularities with physical applications. *J. Math. Phys.* **41**, 3344. (doi:10.1063/1.533316)
25. Arnold VI. 1978 *Mathematical methods of classical mechanics*. New York: Springer.
26. Birkett HR, Thorpe AJ. 1997 Superposing semi-geostrophic potential-vorticity anomalies. *Q. J. R. Meteorol. Soc.* **123**, 2157–2163.
27. Shutts G. 1991 Some exact solutions to the semi-geostrophic equations for uniform potential vorticity flows. *Geophys. Astrophys. Fluid Dyn.* **57**, 99–114. (doi:10.1080/03091929108225230)
28. Loftin J, Yau S-T, Zaslow E. 2005 Affine manifolds, SYZ geometry and the 'Y' vertex. *J. Differential Geom.* **71**, 129–158. (doi:10.4310/jdg/1143644314)
29. Arnold VI, Gusein-Zade SM, Varchenko AN. 2012 *Singularities of differentiable maps*, vol. **1**. Classification of critical points, caustics and wave fronts. Basel: Birkhäuser.
30. Napper L, Roulstone I, Rubtsov V, Wolf M. 2023 Monge–Ampère geometry and vortices. (https://arxiv.org/abs/2302.11604)