Gradients and canonical transformations

by Gaetano Zampieri (Torino)

Abstract. The main aim of this paper is to give some counterexamples to global invertibility of local diffeomorphisms which are interesting in mechanics. The first is a locally strictly convex function whose gradient is non-injective. The interest in this function is related to the Legendre transform. Then I show two non-injective canonical local diffeomorphisms which are rational: the first is very simple and related to the complex cube, the second is defined on the whole $\mathbb{R}^4$ and is obtained from a recent important example by Pinchuk. Finally, a canonical transformation which is also a gradient (of a convex function) is provided.

1. Introduction. To motivate our results, let us start from the autonomous Lagrange equations of classical mechanics

$$(1.1) \qquad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(q,\dot{q}) - \frac{\partial L}{\partial q}(q,\dot{q}) = 0, \qquad L \in C^2(\Omega; \mathbb{R}), \ \Omega = \Omega^\circ \subseteq \mathbb{R}^{2n}.$$

We assume the open set Ω to be *connected*, and the *Lagrangian function* L to be *locally strictly convex* with respect to $\dot{q}$; more precisely, the following $n \times n$ Hessian matrix is assumed to be positive definite:

$$(1.2) \qquad \frac{\partial^2 L}{\partial \dot{q}^2}(q,\dot{q}) > 0 \quad \text{ for all } (q,\dot{q}) \in \Omega.$$

So, near any point $(q,\dot{q})$, we can invert $p = \partial L/\partial \dot{q}$ with respect to $\dot{q}$, and plugging the local inverse function for $\dot{q}$ into the formula $p \cdot \dot{q} - L(q,\dot{q})$ (where "$\cdot$" is the scalar product), we have the local *Hamiltonian* function $H(q,p)$. The function H is called the *Legendre transform* of L with respect to $\dot{q}$. A natural question is now whether the gradient of a function with a positive definite Hessian matrix, at each point of a connected open set, can fail to be injective. The answer is *yes*. In Section 2 we show an analytic example.

[153]

Notice that the gradient is injective if the domain is convex; Section 5 has a proof of this elementary known fact.

In a recent paper, Pucci and Serrin [PS] proved an elegant sufficient condition for the existence of global Hamiltonian functions on non-convex sets (with a degree of smoothness weaker than above). I refer the reader to [PS] also for a discussion of the previous literature on this topic where, as far as I know, an example like the one in Section 2 is not shown. Hopefully, further sufficient conditions for global Legendre transforms will appear in the future; their importance is also related to other theories both in ODEs and PDEs.

By means of the function $H(q, p)$ we can pass from the second order system (1.1) to the equivalent *Hamiltonian* system

$$(1.3) \qquad \dot{q} = \frac{\partial H}{\partial p}(q, p), \quad \dot{p} = -\frac{\partial H}{\partial q}(q, p).$$

Now, let us forget the local origin of (1.3), and assume for the sequel to have a Hamiltonian function $(q, p) \mapsto H(q, p)$ of class C^2 defined on the whole open connected set $D \subseteq \mathbb{R}^{2n}$. A C^2 diffeomorphism $f : D \to f(D)$ is called a *canonical transformation* if, for any H as above, the system (1.3) is transformed by f into the Hamiltonian system defined by the transformed Hamiltonian $K = H \circ f^{-1}$. We can check at once that this last property is equivalent to

$$(1.4) \quad f'(q, p) \begin{pmatrix} O & I_n \\ -I_n & O \end{pmatrix} f'(q, p)^T = \begin{pmatrix} O & I_n \\ -I_n & O \end{pmatrix} \quad \text{for all } (q, p) \in D$$

where $f'(q, p)$ is the Jacobian matrix of f at (q, p), O is the $n \times n$ zero matrix, I_n is the $n \times n$ identity matrix, and the exponent ^{T} means transposition. From (1.4) we have for the Jacobian determinant

$$(1.5) \qquad \det f'(q, p) = 1 \quad \text{for all } (q, p) \in D$$

and we see at once that (1.5) is equivalent to (1.4) in the particular case of dimension 2.

More generally, we consider *canonical local diffeomorphisms*, namely f as above but possibly failing to be injective. The main point is not to be more general, but to consider the very practical situation where condition (1.4) is easily checked. We ask: is injectivity a corollary? The answer is *no*. In Section 3 we see a very simple rational homogeneous example in dimension 2. Can we give examples on the whole $\mathbb{R}^{2n}$? In Section 3 we see a *rational* example in dimension 4. To build it we use a famous recent polynomial constructed by Pinchuk [P]. However, our example is not polynomial and the question whether a *polynomial* example exists or not is equivalent to the *Jacobian conjecture*, open since 1939 (see the references in [GTZ]).

A final question: can a canonical transformation be the gradient of a strictly convex function? We may consider affine canonical transformations. Even non-linear functions can do the job. Section 4 shows an example, defined on a bounded rectangle of $\mathbb{R}^2$. To get it we used the technique in Nitsche [N] starting from Scherk's minimal surface $z = \ln \cos x - \ln \cos y$ (see [G], p. 190). On the whole $\mathbb{R}^{2n}$ these non-linear examples are impossible by a classical theorem of Jörgens, generalized to any dimension by Calabi and Pogorelov [Po].

2. Non-injective locally convex gradient. Our first and main example is the following homogeneous function of degree 2:

$$u : \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\} \to \mathbb{R},$$

(2.1)
$$(x, y) \mapsto (x^2 + y^2)\left(1 + \frac{y^2}{\left(x + \sqrt{x^2 + y^2}\right)^2}\right)$$
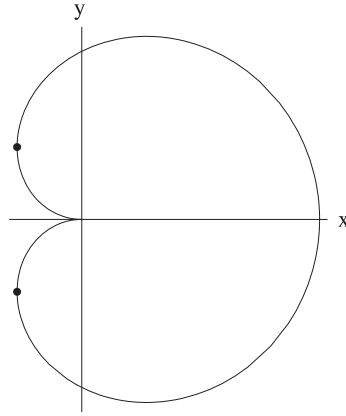
The following picture shows the surface $z = u(x, y)$:
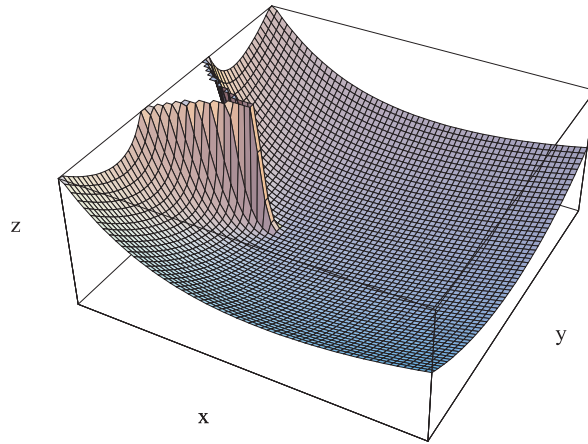
The second picture displays a level curve. We check at once that $u(r \cos \theta, r \sin \theta) = 2r^2/(1 + \cos \theta)$. So the level curves have the simple polar equation $r = a\sqrt{1 + \cos \theta}$ and resemble cardioids.

The gradient as a column vector is

(2.2)
$$\nabla u(x, y) = \begin{pmatrix} \partial_x u(x, y) \\ \partial_y u(x, y) \end{pmatrix}$$

$$= \frac{2\sqrt{x^2 + y^2}}{\left(x + \sqrt{x^2 + y^2}\right)^2} \begin{pmatrix} 2x^2 - y^2 + 2x\sqrt{x^2 + y^2} \\ 3xy + 2y\sqrt{x^2 + y^2} \end{pmatrix}.$$

The function $\nabla u : \mathbb{R}^2 \setminus \{(x,0) : x \leq 0\} \to \mathbb{R}^2$ is non-injective. Indeed, the second picture suggests the existence of two points, on each level curve, with the same $x < 0$ and opposite y, where the gradients are parallel to the x-axis and coincide. The equation $\partial_y u(x,y) = 0$, and $y \neq 0$, give $y = \pm\sqrt{5}\,x/2$. Now, for any $x < 0$, we have

$$\nabla u(x, \sqrt{5}\,x/2) = \frac{9}{5\sqrt{5}} \begin{pmatrix} \sqrt{5}\,(8x - 7\sqrt{x^2}) \\ 2\,(x + \sqrt{x^2}) \end{pmatrix} = \begin{pmatrix} 27x \\ 0 \end{pmatrix},$$

(2.3)

$$\nabla u(x, -\sqrt{5}\,x/2) = \frac{9}{5\sqrt{5}} \begin{pmatrix} \sqrt{5}\,(8x - 7\sqrt{x^2}) \\ -2(x + \sqrt{x^2}) \end{pmatrix} = \begin{pmatrix} 27x \\ 0 \end{pmatrix},$$

so the non-injectivity of the gradient is proved.

However, $\nabla u : \mathbb{R}^2 \setminus \{(x,0) : x \leq 0\} \to \mathbb{R}^2$ is locally strictly convex since it is defined in a connected open set, its derivative at $(1,0)$ (as well as at any other point $(x,0)$ with $x > 0$) is

(2.4) $\qquad u''(1,0) = \begin{pmatrix} \partial_{xx}u(1,0) & \partial_{xy}u(1,0) \\ \partial_{yx}u(1,0) & \partial_{yy}u(1,0) \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 5/2 \end{pmatrix}$

and the eigenvalues of $u''(x,y)$ must always be strictly positive since their product, namely $\det u''(x,y)$, is equal to

(2.5) $\qquad \dfrac{4(x^2 + y^2)}{(x + \sqrt{x^2 + y^2})^6}(40x^4 + 44x^2y^2 + 7y^4 + 8x(5x^2 + 3y^2)\sqrt{x^2 + y^2}),$

which has no zeros in $\mathbb{R}^2 \setminus \{(x,0) : x \leq 0\}$. Indeed, such zeros would be also zeros of

(2.6) $\qquad (40x^4 + 44x^2y^2 + 7y^4)^2 - (8x(5x^2 + 3y^2)\sqrt{x^2 + y^2})^2$

which is equal to $y^6(40x^2 + 49y^2)$. So y would vanish, and for $y = 0$ the expression in (2.5) becomes $160x^5/(x + \sqrt{x^2})^5 \neq 0$ for $(x,0) \in \mathbb{R}^2 \setminus \{(x,0) : x \leq 0\}$.

3. Non-injective and canonical. Next, let us consider the homogeneous rational function of degree 1,

$$(3.1) \quad f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2, \quad (x,y) \mapsto \left(\frac{x^3 - 3xy^2}{\sqrt{3}\,(x^2 + y^2)}, \frac{3x^2 y - y^3}{\sqrt{3}\,(x^2 + y^2)} \right).$$

It is a canonical local diffeomorphism, namely its Jacobian determinant is constantly equal to 1:

$$(3.2) \qquad \det f'(x,y) = 1 \quad \text{for all } (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}.$$

We can check this fact directly or remark that the numerators in (3.1) are the components of the complex cube, so in polar coordinates we just have $r \exp(i3\theta)/\sqrt{3}$. This last form should also clarify how f was built. Of course f is not injective, for instance $f(-\sqrt{3}/2, -1/2) = f(\sqrt{3}/2, -1/2)$.

Our next example is constructed starting from the Pinchuk polynomial [P]. Here we do not need the specific form of that function, we just recall that it is a function $g : \mathbb{R}^2 \to \mathbb{R}^2$ with polynomial components such that $\det g'(x) \neq 0$ for all $x \in \mathbb{R}^2$ but non-injective. There is a classic way to build our map on it: we just consider

$$(3.3) \qquad F : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R}^2, \quad (x,y) \mapsto (g(x), g'(x)^{-T} y)$$

($^{-T}$ denotes the inverse of the transpose matrix). We check at once that condition (1.4) is satisfied. So we have a canonical local diffeomorphism which fails to be injective. By the polynomiality of g, the function F is rational.

Canonical transformations of the form $(h(x), h'(x)^{-T} y)$ are called *Mathieu* in the literature.

4. Gradient and canonical. Now, consider the function

$$(4.1) \qquad \begin{aligned} &]{-\pi/2}, \pi/2[^2 \to \mathbb{R}^2, \\ &(x,y) \mapsto (\sinh^{-1}(\cos y \tan x), \sinh^{-1}(\cos x \tan y)). \end{aligned}$$

Its Jacobian matrix can be written as

$$(4.2) \qquad \frac{1}{\sqrt{1 + \tan^2 x + \tan^2 y}} \begin{pmatrix} \sec^2 x & -\tan x \tan y \\ -\tan x \tan y & \sec^2 y \end{pmatrix}.$$

By the symmetry of this matrix, the function in (4.1) is a gradient. Let v be its primitive which vanishes at the origin. We have

$$(4.3) \qquad \det v''(x,y) = 1 \quad \text{for all } (x,y) \in \,]{-\pi/2}, \pi/2[^2.$$

This was precisely our goal.

5. Strict convexity. To be self-contained we end with the following well known proposition which, for instance, shows the injectivity of the function (4.1) at once.

PROPOSITION. *Let $f : \Omega \to \mathbb{R}^n$ be a $C^1(\Omega)$ map with $\Omega \subseteq \mathbb{R}^n$ open and convex. If the quadratic form $\xi \mapsto \xi \cdot f'(x)\,\xi$ is positive definite at any $x \in \Omega$, then f is injective. In particular, this holds for $f = \nabla u$ with $u : \Omega \to \mathbb{R}$ such that $u''(x) > 0$ for all $x \in \Omega$.*

P r o o f. Let $x_0, x_1 \in \Omega$ be two distinct points, and s the segment $s : [0,1] \to \Omega$, $\theta \mapsto \theta x_1 + (1 - \theta)x_0$. Moreover, let

$$g : [0,1] \to \Omega, \quad \theta \mapsto (x_1 - x_0) \cdot (f(s(\theta)) - f(x_0)).$$

We have $g'(\theta) = (x_1 - x_0) \cdot f'(s(\theta))(x_1 - x_0) > 0$ by hypothesis. So the map g is strictly increasing and, in particular, $g(1) \neq g(0) = 0$. Finally $g(1) = (x_1 - x_0) \cdot (f(x_1) - f(x_0)) \neq 0$ implies $f(x_1) \neq f(x_0)$. ∎

Acknowledgments. The author thanks Franco Cardin who asked him whether one can invert the gradient of a locally strictly convex function on a non-convex set. Mathematical chats with Franco are always as interesting as pleasant.

The pictures were produced using the application *Mathematica* by Wolfram Research Inc.

References

[A]   V. I. A r n o l' d, *Mathematical Methods of Classical Mechanics*, Springer, 1978.
[G]   E. G i u s t i, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, 1984.
[GTZ]   G. G o r n i, H. T u t a j - G a s i ń s k a and G. Z a m p i e r i, *Drużkowski matrix search and D-nilpotent automorphisms*, Indag. Math. 10 (1999), 235–245.
[N]   J. C. C. N i t s c h e, *Elementary proof of Bernstein's theorem on minimal surfaces*, Ann. of Math. 66 (1957), 543–544.
[P]   S. P i n c h u k, *A counterexample to the real Jacobian conjecture*, Math. Z. 217 (1994), 1–4.
[Po]   A. V. P o g o r e l o v, *The Minkowski Multidimensional Problem*, Wiley, 1978.
[PS]   P. P u c c i and J. S e r r i n, *On the derivation of Hamilton's equations*, Arch. Rational Mech. Anal. 125 (1994), 297–310.

Dipartimento di Matematica
Università di Torino
via Carlo Alberto 10
I-10123 Torino, Italy
E-mail: zampieri@dm.unito.it