Solving the Social Choice problem under equality constraints

Crespo, Juan A. and Sanchez-Gabites, J.J

Universidad Autónoma de Madrid

27 July 2016

SOLVING THE SOCIAL CHOICE PROBLEM UNDER EQUALITY CONSTRAINTS

JUAN A. CRESPO† AND J. J. SÁNCHEZ-GABITES‡

Abstract. Suppose that a number of equally qualified agents want to choose collectively an element from a set of alternatives defined by equality constraints. Each agent may well prefer a different element, and the social choice problem consists in deciding whether it is possible to design a rule to aggregate all the agents' preferences into a social choice in an egalitarian way. In this paper we obtain criteria that solve this problem in terms of conditions that are explicitly computable from the constraints. As a theoretical consequence, we show that the only way to avoid running into a social choice paradox consists in designing (if possible) the set of alternatives satisfying certain optimality condition on the constraints, that is, in the natural way from the point of view of economics.

Keywords. Social choice, optimization, rational design.

JEL classification codes. D71, C60, D63.

Introduction

Suppose that an element needs to be selected out of a *set of alternatives* X by a number of agents, each of which may well want to choose a different one. In order to make a collective decision it is necessary to provide some rule to aggregate the individual preferences of the agents into a social one. Broadly speaking, the *social choice problem* consists in deciding whether this aggregation process can be performed in a socially acceptable manner or, more formally, whether there exist *aggregation*

The authors warmly thank Carlos Hervés, Francisco Marhuenda, Juan Pablo Rincón, and María del Carmen Roldán for their critical reading of the paper.

† Juan A. Crespo is supported by the Ministerio de Economía y Competitividad (Spain) grant ECO2013-42710-P and the Comunidad de Madrid grant MadEco-CM (S2015/HUM-3444).

‡ J. J. Sánchez-Gabites is supported by the Ministerio de Economía y Competitividad (Spain) grant MTM2015-63612-P and the Comunidad de Madrid grant MadEco-CM (S2015/HUM-3444).

rules that satisfy certain axioms which are believed to capture whatever is meant by "socially acceptable".

There are several approaches to the social choice problem in the literature. In the Arrowian model [1] each agent orders *all* the elements of X and the aggregation rule yields an ordering of the alternatives that, in order to be socially acceptable, is essentially required to be compatible with the orderings established by the agents. By contrast, in the model of Chichilnisky [5, 6] each agent just reveals her *preferred* alternative from X and the aggregation rule is simply a function that takes as an input the bundle of individual selections and produces as an output a single element of X, the social or collective choice. Again, for this aggregation rule to be socially acceptable it is required to satisfy certain axioms: it should be *anonymous* in the sense that all the agents are equally considered; *unanimous*, which means that if all agents happen to select the same alternative from X then this has to be also the social choice; and *continuous*, a condition about which we shall say a few words shortly. These axioms are described formally in Section 1.

In this paper we want to analyze the social choice problem over sets of alternatives that are defined, as is often the case in quantitative economics, by means of a collection of equality constraints $g_i(\vec{x}) = c_i$; that is,

$$X = \{\vec{x} \in \mathbb{R}^n : g_i(\vec{x}) = c_i \text{ for } i = 1, \ldots, m\}. \tag{1}$$

Such sets usually consist of a continuum of alternatives, so it does not seem reasonable to require that the agents order all the alternatives as in the Arrowian model, but just state their preferred one. Moreover, the continuity axiom becomes almost unavoidable since it is natural to assume that an agent cannot distinguish between two sufficiently close alternatives and, as a consequence, switching from one to the other should cause only a small change in the aggregation function (see a detailed discussion following [16, Remark 2.3.1, p. 6]). These observations strongly suggest that we work within the model of Chichilnisky, and then our social choice problem boils down to the following question: *given a set of alternatives X as above, is it possible to find an* *anonymous, unanimous and continuous aggregation rule over X?* Notice that the aggregation rule should be defined whatever the alternative chosen by each of the agents, since there are no assumptions on how they perform their individual choices. In particular, the agents are in principle independent from each other and may even be completely irrational. It is important to keep this in mind to fully appreciate the results obtained here, at least in their theoretical consequences.

The social choice problem was analyzed and solved, for a wide class of sets of alternatives, by Eckmann [9, 10] in a mathematical guise[1] and later on, in a social choice context, by Chichilnisky and Heal [7] and Weinberger [22]. They showed that, under certain hypotheses, the social choice problem has a solution if and only if the set of alternatives X satisfies a mathematical condition called *contractibility*. Although this solution is completely satisfactory from a mathematical point of view, deciding whether a given set is contractible or not is still a very hard mathematical problem. Also, the very notion of contractibility itself may probably be alien to most readers. It seemed to us that these facts render the result of Chichilnisky and Heal and Weinberger difficult to apply both in specific examples and in theoretical investigations in economics, and our purpose in this paper is to obtain elementary and explicitly computable criteria to analyze the social choice problem over sets X defined by equality constraints. For the present expository purposes we shall just state our main theorem (Theorem 3 in Section 1):

Main Theorem. *Let X be defined as in* (1). *A necessary condition for the social choice problem over X to have a solution is as follows: for any constraint $g_i(\vec{x}) = c_i$ such that the set Y_i determined by the remaining constraints is bounded, c_i must be either the global maximum or the global minimum value that g_i attains over Y_i.*

[1]Social choice functions may be regarded as some sort of generalized means. This point of view is quite independent of any particular interpretation of the variables, and in this guise the social choice problem already attracted mathematicians –for purely theoretical reasons– in the first part of the past century. After the seminal paper of G. Aumann [2], several solutions for spaces satisfying certain regularity conditions have been published [9, 11, 14]. The interested reader may enjoy Eckmann's account of the history of the problem [10].

The condition that Y_i is bounded guarantees that g_i indeed attains a global maximum and a global minimum value on Y_i. What the theorem says is that unless c_i happens to coincide with either of these; that is, unless the constraint $g_i(\vec{x}) = c_i$ is *optimal with respect to the others*, the social choice problem over X has no solution. Notice that in general this optimality condition may have to be satisfied by several constraints simultaneously, since several of the Y_i may be bounded (however, each of the corresponding constraints $g_i(\vec{x}) = c_i$ would have to be optimal over a *different* set Y_i), and notice also that the theorem tells us nothing about those constraints for which Y_i is not bounded. Finally, observe how the language used in the statement of the theorem is completely elementary, making no reference to contractibility whatsoever.

In addition to its applicability to specific examples, this result has a theoretical interpretation that we feel valuable because it shows that there exists a surprising relationship between the social choice problem and economics. We will now discuss these two aspects in turn.

A) *Application to specific examples*. The necessary optimality condition afforded by the main theorem can be used as a criterion to show that a given social choice problem has no solution. When this is the case it is customary to speak of a *social choice paradox*, since any procedure to perform collective choices will violate at least one of the three axioms (unanimity, anonimity and continuity) laid out earlier and, in that sense, will fail to be socially acceptable.

Example 1 below is included to illustrate how computations may proceed in a typical case. We have deliberately chosen an example without any particular interpretation because at this point we want to emphasize that the main theorem is purely mathematical in nature and therefore completely independent of the interpretation (if any) of the social choice problem under consideration.

Example 1. Suppose that we want to analyze the social choice problem over the set $X \subseteq \mathbb{R}^3$ defined as

$$X = \{(x, y, z) \in \mathbb{R}^3 : x^4 + 2(y^2 + 1)z^2 + y^2 = 1 \ , \ 2x^2 - 2y^2 = 1\}.$$

The set defined by the first constraint alone (that is, removing the second constraint) is easily seen to be bounded. A computation using the method of Lagrange multipliers shows that the second constraint is not optimal with respect to the first and therefore, according to our main theorem, the social choice problem over X has no solution. ■

Example 1 is therefore an instance of a social choice paradox, and some playing around with the mathematics would lead to many more examples. None of these would be very interesting, however, since they would not *mean* anything. By contrast, as soon as one considers examples whose variables and constraints have some meaningful interpretation, the optimality criterion provided by the main theorem acquires in turn an interesting interpretation which, in particular, brings economics into the picture in a somewhat unexpected way. This is what we discuss now.

B) *Theoretical consequences of the main theorem*. Suppose, for the purpose of illustration, that in Example 1 the first constraint is exogenous in the sense that it cannot be operated upon, but some agent (either the agents performing the choice themselves or another, completely unrelated agent) can fix the value c_2 of the second constraint; that is, the social choice problem can be *designed* to some extent. Then we know that only setting that second constraint to be optimal with respect to the first will allow for the possibility of social choice on X. Thus *optimization is a necessity in designing a social choice problem if we want it to have a solution.* We emphasize that this is a purely mathematical conclusion, insofar as Example 1 was also purely mathematical.

Now let us set ourselves in a context where the variables and the constraints have an economical interpretation. For definiteness, assume that $\vec{x}$ represents the bundle of production factors used by some firm. In general, of course, there will exist constraints on the bundles that can be used, maybe owing to technological limitations on the production process, the necessity to satisfy a given demand, etc. These constraints determine a set X from which the bundle of production factors has to be chosen. If the firm is run by several agents, each one is entitled to her own preference and so they are faced with a social choice problem over X. Will this problem have a solution, as seems desirable? Let us consider, in turn, two particular cases:

(i) Imagine first that the constraints represent technological limitations on the productive process that are enforced by several laws of Nature. These constraints are exogenous, in the sense that they cannot be operated upon. There is no reason to expect that any of these constraints will be optimal with respect to the others[2], and so the main theorem implies that the social choice problem over X has no solution.

(ii) Suppose instead that there are only two constraints $P(\vec{x}) = d$ and $U(\vec{x}) = u$ which represent the requirements that the firm satisfies a given demand d and obtains certain level of utility u. We take the external demand d to be fixed, so the first constraint is exogenous, but assume that the agents can choose the level of utility u. Then our main theorem implies that if the agents want to allow themselves to perform a collective choice of the bundle of production factors to be used then they must fix their utility level u to be either the best (or, quite paradoxically, the worst) attainable while satisfying the given demand d. If there were additional technological constraints as in (i), the conclusion would still be the same: the agents should fix an optimal level of utility u that satisfies the demand d and accords to the technological constraints.

The conclusion of the previous paragraph is, by itself, hardly surprising for anyone and almost axiomatic for an economist. The puzzling point is *how the conclusion is reached*: we have just applied a mathematical theorem, as we did in Example 1, and nowhere in our argumentation have we made any assumptions with an economical content. As mentioned at the beginning of the paper when describing the social

[2]More quantitatively, since $g_i|_{Y_i}$ has exactly two global optimal values (or just one, if the maximum and the minimum happen to coincide), the probability that c_i actually agrees with any one of these is zero.

choice problem in the abstract, there are no conditions on how the agents choose their preferred alternatives or how their individual utilities relate to each other[3] or to the utility $U(\vec{x})$ of the firm. In sum, there is *no* economics beyond the fact that the variables and constraints have an economical interpretation. Our starting point was just a bare problem in social choice but, nevertheless, in attempting to solve it we have been led, by the mathematics alone, to the familiar condition that the level of utility u should be optimal. This conclusion *is* economics, emerging spontaneously. Mathematically it is nothing but the optimality condition from our main theorem again, but now interpreted in an economical context, and to emphasize this we shall refer to it as the condition of *rationality in the design* of the set of alternatives X. This choice of terminology seems appropriate, given that optimization is one of the distinguishing features of rationality. Summarizing, we may state the following "principle of rational design":

A social choice problem that is not designed rationally has no solution.

Conversely, another result in this paper (Theorem 5) establishes the following:

A social choice problem that is designed rationally has, generically[4]*, a solution.*

We call these "principles" because they provide a general idea about how a social choice problem should be designed (when possible) but, as always, their application to each particular case has to be exercised with due care using the precise statements of the theorems and observing that their hypotheses are satisfied.

Let us emphasize once again that we are considering rationality in the *design* of the social choice problem, not rationality of the agents that make the collective decision. In (i) above the latter may well be rational but, still, the social choice problem they face has no solution because there was no rationality involved in its design (one could even object to the use of the term "design" in that example, given that the constraints came from laws of Nature). In (ii) the same agents making the collective decision

[3]In fact, their individual utilities appear nowhere in the social choice problem.

[4]The strict converse to the main theorem is not true in general, as shown in Example 12 in Section 4, thus the word "generically". Its precise meaning will be explained in Section 4; for the moment it may be understood as "for almost every social choice problem that is rationally designed".

could participate, to some extent, in the design of the social choice problem and therefore avoid running into a social choice paradox. This is probably the situation where the principle of rational design more clearly reproduces results that are familiar in economics, as just seen. Still, yet another situation can arise: that in which the problem can be designed but not by the agents making the decision themselves. In that case we may be led to conclusions that cannot be interpreted in such a straightforward way. As an illustration, consider the following:

(iii) Revisit example (i) above, but now suppose that the productive process is such that the agents can, by combining the production factors in different proportions, vary the level of pollution generated by the firm. Assume that the latter is fixed by the government, thereby adding a further constraint to those already in (i). By the principle of rational design, unless the allowed level of pollution is the minimum or the maximum allowed by the technological constraints, the agents running the firm will again face a social choice paradox.

Now there is no economic reason[5] to explain why the pollution level should be minimized or maximized (subject to the technological constraints) but, still, this is a necessary condition to allow for a socially acceptable aggregation rule. Thus we see that there is a variety of situations where the principle of rational design can be applied, sometimes leading to familiar conclusions and sometimes not.

Finally, it might be interesting to remark that the fact that no economics was involved in our reasoning actually leaves a subtle trace: in (ii) we concluded that the utility u had to be the global maximum or the global minimum allowed by the constraint on the demand, but we could not discriminate between the two. Mathematics took us that far, but no more. This owes to the fact that, mathematically, both types of extrema are on the same footing and there is no reason to discard one in favour of the other. It is only economics that tells us that (in this case) the maximum is desirable while the minimum is not. Something similar happens in (iii), where again

[5]At least within the variables and constraints contemplated in the social choice problem.

the choice between the maximum or the minimum should be guided by considerations which lie beyond mathematics.

Having seen how the results in this paper can be used both as practical criteria to analyze specific examples and also as tools for more theoretical investigations, we finish this Introduction with a brief account of how the paper is organized. In Section 1 we formally describe the model of social choice and the main theorem introduced above. We also state an auxiliary result (Proposition 2) which provides a useful criterion by itself. Since the proof of the main theorem is somewhat involved, we have thought it convenient to introduce first an outline of the main ideas that come into play (we do this in Section 2) and then start with the proof proper (Section 3). Some technical lemmas are postponed to appendices A and B. Up to Section 2 we will be able to introduce the necessary mathematical background along the way, but later on some notions from algebraic topology (homology theory and some duality results) and differential geometry will be needed. Suitable references will be included where appropriate.

1. Statement of results

In this section we review very briefly the basic elements of the social choice model of Chichilnisky and give the formal statements of our results.

1.1. **Basic definitions of the social choice model.** Let $k \geq 2$ be the number of agents performing the collective choice over the set of alternatives X. As argued in the Introduction, for us the preference of the ith agent will simply consist of a single element $p_i \in X$ (the favourite alternative of the agent) and an aggregation rule will therefore be a mapping $F(p_1, p_2, \ldots, p_k) = p$, where $p \in X$ is the collective choice. Notice that the agents have absolute freedom in choosing their preferred alternative, so F should be defined for any tuple $(p_1, p_2, \ldots, p_k) \in X \times \overset{(k)}{\cdots} \times X$; that is, it should be a mapping $F : X \times \overset{(k)}{\cdots} \times X \longrightarrow X$. Also, in order for the aggregation rule F to be socially acceptable it is required to satisfy the axioms of anonimity, unanimity and continuity which, formally, read as follows:

- *Anonimity*: $F(p_1, p_2, \ldots, p_k)$ should be independent of the ordering of the p_i.
- *Unanimity*: $F(p, p, \ldots, p) = p$ for every $p \in X$.
- *Continuity*: F is a continuous mapping (notice that it makes sense to speak of continuity since X is a subset of some $\mathbb{R}^n$).

An aggregation rule F that satisfies the above three axioms is called a *social choice function* over the given set of alternatives X. With this terminology, the social choice problem is stated as follows: *given a set of alternatives X, is it possible to find a social choice function over X?*

When X is simple enough the existence of social choice functions can sometimes be easily established directly. For example, when X is an interval of the real line then the mean, the maximum and the minimum

$$F(p_1, \ldots, p_k) := \frac{1}{k} \sum_{i=1}^{k} p_i \quad , \quad F(p_1, \ldots, p_k) = \max_{1 \leq i \leq k} p_i \quad , \quad F(p_1, \ldots, p_k) = \min_{1 \leq i \leq k} p_i$$

are all well defined social choice functions. The mean is, more generally, a social choice function over any convex set $X \subseteq \mathbb{R}^n$, but it can no longer be assured to be a suitable social choice function over sets X defined by equality constraints, which are the ones of interest to us. The reason is that, as soon as some of the constraints defining X are nonlinear[6], the mean of two elements from X does not need to belong to X. As another example, most auction methods (first price, second price, etc.) also satisfy the above axioms.

Recall that we speak of a social choice paradox whenever X does not admit a social choice function. The classical example of such a set of alternatives X is the circumference, a result obtained by Chichilnisky [5] when considering linear preferences on the commodity space of bundles of two collective goods (the paper by Baigent [3] contains a clear exposition of the main ideas and techniques involved in the proof). Together with higher dimensional spheres and the Möbius band [4], these seem to be

[6]In passing, let us observe that if all the constraints defining X are linear then X is a convex subset of $\mathbb{R}^n$ and therefore the social choice problem over X has a solution; namely, the mean. This does not contradict our results because neither X nor any of the sets Y obtained by removing any one of the constraints are bounded, being linear subspaces of $\mathbb{R}^n$.

the only examples ever considered in the literature. The methods presented in this paper, because of their simplicity, greatly enlarge the variety of examples of social choice paradoxes.

1.2. **Our results.** As mentioned earlier, we are interested in the social choice problem over sets of alternatives $X \subseteq \mathbb{R}^n$ that are defined by a collection of equality constraints. Thus, let there be a collection of smooth maps $g_1, \ldots, g_m : \mathbb{R}^n \longrightarrow \mathbb{R}$ and values $c_1, \ldots, c_m \in \mathbb{R}$ which determine the set of alternatives X as

$$X = \{\vec{x} \in \mathbb{R}^n : g_i(\vec{x}) = c_i \text{ for every } 1 \leq i \leq m\}.$$

The constraints can be completely arbitrary (in particular, they do not need to be linear) but we shall always assume that the set of alternatives X they determine is bounded, which is a reasonable requisite in most problems that try to capture some aspect of reality. Of course, for the social choice problem to make sense X should be nonempty. Also, if X is finite then the problem has a somewhat trivial answer in the affirmative (see Section 4), so the case of interest is when X is actually infinite. A convenient way of encapsulating these considerations consists in requiring that the number of constraints m is strictly smaller than the dimension of the ambient space n; that is, $m < n$.

We will first consider an auxiliary base case (which, however, has some interest in itself) assuming that the g_i satisfy the standard *constraint qualifications*; that is, the gradients of the g_i are linearly independent at each $\vec{x} \in X$. In this case we will definitely run into a social choice paradox:

Proposition 2. *Let the set of alternatives X be bounded, $m < n$, and assume that the g_i satisfy the constraint qualifications. Then the social choice problem over X has no solution.*

The classical case when X is a sphere can be analyzed very easily using Proposition 2. An $(n-1)$–dimensional sphere is described by a single $(m = 1)$ implicit equation

$x_1^2 + \ldots + x_n^2 = 1$ in $\mathbb{R}^n$ which evidently satisfies the constraint qualifications. Hence the social choice problem over spheres has no solution for $n \geq 2$.

Now we move on to the precise statement of our main theorem. Observe that Proposition 2 entails that, in order to have any hope of solving the social choice problem over X, it is necessary that the g_i do not satisfy the constraint qualifications. The simplest case is when $m-1$ of the constraints (say, for definiteness, the first $m-1$) do satisfy them and it is only the addition of the remaining constraint what spoils this condition. Thus, we shall assume that the gradients of $g_1, \ldots, g_{m-1}$ are linearly independent at each point of the feasible set they determine

$$Y = Y_m = \{\vec{x} \in \mathbb{R}^n : g_1(\vec{x}) = c_1, \ldots, g_{m-1}(\vec{x}) = c_{m-1}\}.$$

Having chosen the last constraint as the one on which we are going to focus, in the sequel we shall safely omit the subindex from Y_m and simply write Y without risk of confusion. Then our main theorem reads as follows:

Theorem 3. *Let the set Y be bounded and connected, $m < n$, and assume that the first $m-1$ constraints satisfy the constraint qualifications.*[7] *If the social choice problem over X has a solution then the last constraint must be optimal with respect to the remaining ones.*

The connectedness assumption on Y means that it consists only of a single "piece" and is included just for convenience: if Y is not connected, that is, if it consists of several disjoint pieces, then the conclusion of the theorem is that c_m must be the global optimum value of g_m restricted to one of those pieces.

The optimality condition provided by Theorem 3 is necessary, but in general not sufficient, to avoid a social choice paradox (see Example 12 in Section 4). However, there is a second order condition which once again only involves notions familiar from optimization theory and turns out to be enough to guarantee that the optimality

[7]Also, as a technical hypothesis, it is necessary to assume that $g_m|_Y$ has only finitely many critical values. We shall explain this in detail later on.

condition is indeed sufficient. Construct the Lagrangian function

$$L(\lambda_1, \ldots, \lambda_{m-1}, \vec{x}) = g_m(\vec{x}) - \sum_{i=1}^{m-1} \lambda_i \left(g_i(\vec{x}) - c_i\right),$$

which corresponds to the optimization problem of finding the critical points of g_m subject to the first $m-1$ constraints. The Hessian matrix of L with respect to all its variables (the λ_i and the x_i) is called the *bordered Hessian* of the optimization problem. Then:

Theorem 4. *Let the set Y be bounded and connected, $m < n$, and assume that the first $m-1$ constraints satisfy the constraint qualifications. Suppose that the last constraint is indeed optimal with respect to the others. In addition, assume that the bordered Hessian has a nonzero determinant at each point in X. Then the social choice problem over X has a solution.*

Applying this theorem may be difficult in practice, but in fact its interest is mainly theoretical. This stems from the fact that the condition concerning the nonzero determinant of the bordered Hessian is satisfied *generically*, in a sense to be explained later on in Section 4. Thus we can state the following result:

Theorem 5. *Let the set Y be bounded and connected, $m < n$, and assume that the first $m-1$ constraints satisfy the constraint qualifications. Then a necessary and, generically, sufficient condition for the social choice problem over X to have a solution is that the last constraint be optimal with respect to the remaining ones.*

2. A rough outline of the proof of Theorem 3

The proof of Theorem 3 is somewhat complicated and we have thought it convenient to include an outline of the basic ideas and difficulties that it involves.

Since the spaces of alternatives X of interest to us go beyond the simple cases of convex sets or spheres, there is no straightforward way of deciding whether there exists a social choice function over and we need to resort to the beautiful result of Chichilnisky and Heal already mentioned in the Introduction, and whose precise statement is the following [7, Theorem 1, p. 82]:

Theorem. Let the space of alternatives X be a parafinite CW complex. A necessary and sufficient condition for the existence of social choice functions on X, for every number of agents, is that each component of X is contractible.

In addition to the difficulty of deciding whether a given space is contractible or not, another issue that arises in applying the theorem of Chichilnisky and Heal concerns the condition that X should be a parafinite CW complex. Without entering into the details, this technical condition means that X is sufficiently well behaved from a topological point of view. For instance, any manifold M (with or without boundary) is indeed a parafinite CW complex because it can be triangulated, as shown by Whitehead [23] or Whitney [24, Theorem 12A, p. 124]. Unfortunately, a set X defined by a collection of constraints that do not satisfy the constraint qualifications (as in Theorem 3) may have a very complicated structure and in particular the characterization of Chichilnisky and Heal may not be applicable. To circumvent this difficulty we will introduce the concept of *homotopy social choice functions* in Section 3.

We shall need the following two lemmas concerning the contractibility of manifolds:

Lemma 6. *Let M be a compact manifold of dimension $d \geq 1$ and without boundary. Then none of the components of M is contractible.*

Lemma 7. *Let M be a compact contractible manifold of dimension $d \geq 2$. Then its boundary ∂M is nonempty (by the previous lemma) and connected.*

In proving these results it seems unavoidable to make use of homology with real coefficients, which is a powerful tool from algebraic topology. Since this machinery might not be familiar to the reader, we have postponed the proofs to Appendix A.

Let us continue by explaining briefly the general scheme that will be followed to prove Theorem 3. We will consider the whole family of social choice problems that arise as c_m runs in the real numbers, thus changing the set of alternatives X. To emphasize that c_m now plays the role of a parameter we shall replace it by u and reflect this explicitly in the notation for X, letting

$$X_u = \{\vec{x} \in \mathbb{R}^n : g_i(\vec{x}) = c_i \text{ for } 1 \leq i \leq m-1 \text{ and } g_m(\vec{x}) = u\}.$$

This can be equivalently described as

$$X_u = \{\vec{x} \in Y : g_m(\vec{x}) = u\},$$

where Y is the set defined by the first $(m-1)$ constraints as introduced earlier.

Recall that a point $\vec{x} \in Y$ is called a *critical point* of the restricted map $g_m|_Y$ if the gradient $\nabla g_m(\vec{x})$ is a linear combination of $\{\nabla g_1(\vec{x}), \ldots, \nabla g_{m-1}(\vec{x})\}$; that is, the gradient of g_m is a linear combination of the gradients of the constraints that define Y (notice that this is the classical necessary condition from the theory of Lagrange multipliers for $g_m|_Y$ to reach a local extremum value at $\vec{x}$). In that case $u = g_m(\vec{x})$ is said to be a *critical value* of $g_m|_Y$. Equivalently, u is a critical value of $g_m|_Y$ if X_u contains a critical point of $g_m|_Y$. Otherwise u is said to be a *regular* value of $g_m|_Y$.

(1) Consider first the case when u is a regular value of $g_m|_Y$. This amounts to saying that the gradients of $g_1, \ldots, g_m$ are all linearly independent at each $\vec{x} \in X_u$; that is, they satisfy the constraint qualifications. This is precisely the situation considered in Proposition 2. Geometrically, this condition guarantees that (if nonempty) X_u is a differentiable manifold of dimension $d = n - m$ without boundary (see for instance [8, Theorem 2.3, p. 213]) and in particular it is indeed a parafinite CW complex as mentioned earlier. Thus we can directly apply the theorem of Chichilnisky and Heal and reduce the problem to deciding whether the components of X_u are contractible. In Proposition 2 we assumed X_u to be bounded and infinite. The first condition, together with the fact that X_u is closed in $\mathbb{R}^n$ (because it is the preimage of $(c_1, c_2, \ldots, u)$ via the continuous map $g = (g_1, \ldots, g_m) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$), entails that X_u is compact. The second condition implies that the dimension d of X_u is at least 1, since otherwise X_u would consist only of finitely many points. Therefore we can apply Lemma 6 to $M = X_u$ to learn that none of the components of X_u is contractible, and we conclude

that the social choice problem over X_u has no solution. This proves Proposition 2, settling the case when u is a regular value of $g_m|_Y$.

It follows from the previous paragraph that only when u is a critical value of $g_m|_Y$ there is some hope for the social choice problem over X_u to have a solution. Saying that u is a critical value of $g_m|_Y$ means that there exists a critical point $\vec{x} \in X_u$, which by definition satisfies the classical necessary condition from the theory of Lagrange multipliers for $g_m|_Y$ to reach a local extremum value at $\vec{x}$. However, it is well known that this is not a sufficient condition, in that $\vec{x}$ could very well be a saddle point (so that u would be neither a local maximum value nor a local minimum value) or, even if it is indeed a local extremum, it does not need to be a global one. This is why Theorem 3 is not a straightforward consequence of Proposition 2: we have to rule out all the possibilities just described and conclude that u must in fact be a global optimum of $g_m|_Y$.

To illustrate this let us refer back to Example 1. The method of Lagrange multipliers shows that there are actually six critical points and three critical values of $g_2|_Y$, which are $u = \pm 2$ (the global optima of $g_2|_Y$) and $u = 0$. For these three values of u Proposition 2 tells us nothing about the social choice problem on X_u, since the constraints do not satisfy the constraint qualifications. A finer analysis is needed to show that also for $u = 0$ the social choice problem over X_u has no solution. This finer analysis is precisely the content of Theorem 3.

(2) Let us sketch now how the proof of Theorem 3 goes. We need to prove that, if there exists a social choice function over X_u, then u is either the global maximum $u_{\max}$ or the global minimum $u_{\min}$ of $g_m|_Y$ (these global optima exist as consequence of the assumption that Y is bounded, and hence compact). We already know by (1) that u must be a critical value of $g_m|_Y$ and, in particular, this implies that now we cannot guarantee that X_u is a manifold (nor a CW complex, in fact) and therefore we cannot apply the theorem of Chichilnisky and Heal directly. To overcome this difficulty we need to make a rather lengthy detour. The argument will be by contradiction, so suppose that there exists a social choice function over X_u but $u_{\min} < u < u_{\max}$. Pick

two numbers u_1 and u_2 such that $u_{\min} < u_1 < u < u_2 < u_{\max}$ and consider the auxiliary set

$$X_{[u_1,u_2]} = \{\vec{x} \in Y : u_1 \leq g_m(\vec{x}) \leq u_2\}.$$

With a suitable choice of u_1 and u_2 the set $X_{[u_1,u_2]}$ can be shown to be a compact manifold with boundary. In fact, its boundary $\partial X_{[u_1,u_2]}$ is the union of the two disjoint sets

$$X_{u_1} = \{\vec{x} \in Y : g_m(\vec{x}) = u_1\} \quad \text{and} \quad X_{u_2} = \{\vec{x} \in Y : g_m(\vec{x}) = u_2\}.$$

We will prove that:

(i) The social choice function that exists over X_u by assumption can "almost" be extended to another one defined on all $X_{[u_1,u_2]}$ (Proposition 9).

(ii) As a consequence of (i) and the theorem of Chichilnisky and Heal applied to the manifold $X_{[u_1,u_2]}$, it follows that the latter must be contractible (Lemma 8).

(iii) But, since $u_{\min} < u_1 < u < u_2 < u_{\max}$, both sets X_{u_1} and X_{u_2} are nonempty and therefore the boundary of $X_{[u_1,u_2]}$ is not connected, having at least two pieces. This entails that $X_{[u_1,u_2]}$ cannot be contractible (Lemma 7).

A contradiction arises between (ii) and (iii), proving that the assumption that $u_{\min} < u < u_{\max}$ is untenable and so u must be a global optimum value of $g_m|_Y$. The word "almost" in (i) owes to the following: while it may not always be possible to extend a social choice function originally defined only on X_u to a social choice function defined on the larger set $X_{[u_1,u_2]}$, it is always possible to extend it at the homotopy level. This is what prompts the definition of homotopy social choice function mentioned earlier, a notion which is slightly weaker than that of a true social choice function, but still good enough for our purposes.

3. Proof of Theorem 3

As a technical assumption we require that $g_m|_Y$ has at most finitely many critical values[8], a condition that will be fulfilled in any problem with a reasonable economical interpretation. For instance, whenever g_m is an analytic function (polynomials being the simplest case) this condition is automatically satisfied.

Note. For the sake of brevity, from now on we shall sometimes abbreviate "social choice function" as SCF.

3.1. **Homotopy social choice functions.** Let us first recast the conditions of unanimity and anonimity in a slightly different –but well known– equivalent way. Suppose F is an SCF for k agents on a space of alternatives X. Denote by the letter Δ the diagonal map

$$\Delta : X \longrightarrow X^k \quad ; \quad \Delta(p) = (p, p, \ldots, p)$$

and by the letter P any permutation map $P : X^k \longrightarrow X^k$. The conditions of unanimity (U) and anonimity (A) on F can then be equivalently stated as

(U) $F \circ \Delta = \mathrm{id}$,
(A) $F \circ P = F$.

As it turns out, the argument given by Chichilnisky and Heal to prove that the existence of social choice functions for any number of agents implies that the preference space is contractible works equally well if (U) and (A) only hold at the homotopy level; that is, if they are replaced by

(HU) $F \circ \Delta \simeq \mathrm{id}$,
(HA) $F \circ P \simeq F$ for any permutation $P : X^k \longrightarrow X^k$.

The reason is that, when one considers the maps F^*, Δ^* and P^* induced by F, Δ and P between homotopy groups, unanimity (U) and its homotopical counterpart (HU) yield the same relation $F^* \circ \Delta^* = \mathrm{id}$, and the same goes for (A) and (HA)

[8]It might be convenient to recall the distinction between critical points and critical values: since many critical points may correspond to the same critical value, $g_m|_Y$ may well have infinitely many critical points in spite of having only finitely many critical values as required.

(namely, $F^* \circ P^* = F^*$). Since it is only these relations that are needed to conclude that X is contractible, our assertion follows.

For the sake of brevity let us call a continuous map $F : X^k \longrightarrow X$ a *homotopy SCF* (for k agents) if it satisfies conditions (HU) and (HA) above. These functions are not to be interpreted in any economical sense, but just as mathematical objects that will be useful to prove Theorem 3. Our discussion may be summed up in the following

Lemma 8. *Let M be a compact manifold. Assume there exist homotopy social choice functions over M for any number of agents. Then M is contractible.*

3.2. **An extension result.** Now we are going to establish the extension result that is key in proving Theorem 3; namely, that an SCF on X_u can be extended to a *homotopical* SCF on $X_{[u_1,u_2]}$.

Let u be a critical value of $g_m|_Y$ and suppose that it is not a global optimum, so that $u_{\min} < u < u_{\max}$. As a consequence of the technical assumption that $g_m|_Y$ only has finitely many critical values we may choose $u_1, u_2 \in [u_{\min}, u_{\max}]$ such that $u_1 < u < u_2$ and u is the only critical value of $g_m|_Y$ on the interval $[u_1, u_2]$.

Proposition 9. *Let u_1 and u_2 be chosen as above. If there exists a social choice function F for k agents on X_u, then there exists a homotopy social choice function F' for k agents on $X_{[u_1,u_2]}$.*

The proof of the proposition needs Lemmas 10 and 11, which we state now. Their proofs are postponed to Appendix B, since they are slightly technical.

Lemma 10. *There exist a neighbourhood U of X_u in $X_{[u_1,u_2]}$ and a continuous function $F_U : U^k \longrightarrow X_{[u_1,u_2]}$ with the properties*

(1) $F_U(p, \ldots, p) = p$,

(2) $F_U(p_1, \ldots, p_k)$ *is independent of the ordering of the p_i.*

Notice that F_U is close to being an SCF on U (it is certainly unanimous and anonymous), but it does not qualify as such because its target space is $X_{[u_1,u_2]}$ rather than U.

Lemma 11. *Given any neighbourhood U of X_u in $X_{[u_1,u_2]}$ there exists a continuous mapping $r : X_{[u_1,u_2]} \longrightarrow X_{[u_1,u_2]}$ such that:*

(1) *$r(p) \in U$ for every $p \in X_{[u_1,u_2]}$,*
(2) *r is homotopic to the identity in $X_{[u_1,u_2]}$.*

We are going to put these two results together to prove Proposition 9.

Proof of Proposition 9. Apply Lemma 10 to find U and F_U; then apply Lemma 11 to the U just obtained to get r. Define $F' : X^k_{[u_1,u_2]} \longrightarrow X_{[u_1,u_2]}$ by

$$F'(p_1, \ldots, p_k) := F_U(r(p_1), \ldots, r(p_k)).$$

Notice that the definition is correct: all the $r(p_i)$ belong to U and therefore it makes sense to evaluate F_U on the k–tuple $(r(p_1), \ldots, r(p_k))$.

We claim that F' is a homotopy SCF on $X_{[u_1,u_2]}$. Clearly F' is insensitive to the ordering of its arguments because the same is true of F_U, so (HA) holds. Also, composing F' with the diagonal map $\Delta(p) = (p, \ldots, p)$ yields

$$F' \circ \Delta(p) = F_U(r(p), \ldots, r(p)) = r(p),$$

and since $r \simeq \mathrm{id}$ in $X_{[u_1,u_2]}$, we see that

$$F' \circ \Delta \simeq \mathrm{id}.$$

This establishes property (HU) and shows that F' is indeed a homotopy SCF. □

3.3. **The proof of Theorem 3.** We are finally ready to prove Theorem 3.

Proof of Theorem 3. The hypothesis that Y is connected implies that $g_m(Y)$ is a connected subset of $\mathbb{R}$, and so it must be an interval. Thus $g_m(Y)$ is the interval $[u_{\min}, u_{\max}]$. As already mentioned we reason by contradiction, so assume that $u_{\min} < u < u_{\max}$ and there exists an SCF over X_u. By Proposition 2 u must be a critical value of $g_m|_Y$. As we did before, we may choose u_1 and u_2 in the interval $[u_{\min}, u_{\max}]$ such that $u_1 < u < u_2$ and u is the only critical value of $g_m|_Y$ in the interval $[u_1, u_2]$. Notice that $X_{[u_1,u_2]}$ is a manifold, because both u_1 and u_2 are regular values of $g_m|_Y$.

Its dimension is the same as that of Y, which is $d = n - (m - 1) \geq 2$ because of the condition $m < n$. As in the proof of Proposition 2, $X_{[u_1,u_2]}$ is compact.

By Proposition 9 the SCF that exists over X_u can be extended to a homotopical SCF on $X_{[u_1,u_2]}$ and, as a consequence of Lemma 8, it follows that $X_{[u_1,u_2]}$ is contractible. Then by Lemma 7 its boundary $\partial X_{[u_1,u_2]}$ has to be connected. However, this boundary is the disjoint union of X_{u_1} and X_{u_2}, both of which are nonempty because both u_1 and u_2 belong to $g_m(Y)$. This contradiction finishes the proof. □

4. On the sufficiency of the optimality condition

Let us begin by presenting an example that illustrates how the optimality condition of Theorem 3 is generally not sufficient to avoid a social choice paradox:

Example 12. Consider the set of alternatives

$$X = \{(x, y, z) \in \mathbb{R}^3 : (x^2 + y^2 + z^2 + 3)^2 - 16(x^2 + y^2) = 0\ ,\ z = 1\}.$$

It turns out that the second constraint is optimal with respect to the first one, so the necessary condition provided by the main theorem is satisfied.[9] Substituting $z = 1$ in $(x^2 + y^2 + z^2 + 3)^2 - 16(x^2 + y^2) = 0$ and rearranging terms yields the implicit equation $x^2 + y^2 = 4$, so X is actually a circumference of radius 2 centered at $(0, 0, 1)$ and contained in the plane $z = 1$. But circumferences are the prototype of a set of alternatives that does not admit a social choice function, so we conclude that the social choice problem over X has no solution. ■

In spite of the previous example, it is often the case that when a constraint is optimal with respect to the others the set of alternatives X actually reduces to a *finite* number of points, and in this case there *do exist* social choice functions over X. To see why, begin by labelling the alternatives (that is, the elements of X) in any order. Then, given the bundle of individual preferences $(p_1, \ldots, p_k)$, simply let $F(p_1, \ldots, p_k)$ be that alternative, among those that appear in $(p_1, \ldots, p_k)$, having

[9] The set Y_2 defined by the second constraint alone is not bounded, so the main theorem does not require the first constraint to be optimal with respect to the second one.

the highest label. It is easy to check that F satisfies the three axioms of anonimity, unanimity and continuity[10] and is therefore a social choice function over X. (It is, however, questionable to what extent such a function is actually of interest in the realm of social choice.) This very simple observation lies at the heart of both Theorem 4 and Theorem 5.

4.1. **The proof of Theorem 4.** The strategy consists in showing that, under the hypotheses of the theorem, X is indeed finite and therefore admits an SCF as just shown.

Proof of Theorem 4. Consider the constrained optimization problem

$$(P): \begin{cases} \text{optimize} & g_m(\vec{x}) \\ \text{subject to} & g_1(\vec{x}) = c_1 \\ & \quad\vdots \\ & g_{m-1}(\vec{x}) = c_{m-1} \end{cases}$$

Let $\vec{x}_0$ be a point in X. Since by assumption the constraint $g_m(\vec{x}) = c_m$ is optimal with respect to the remaining ones, $\vec{x}_0$ is a solution to (P) and, in particular, it must satisfy the standard first order conditions of constrained optimization. Although we know that $\vec{x}_0$ is a global, and hence local, optimum of $g_m|_Y$, let us pretend for a second that we ignore this and classify $\vec{x}_0$ using the second order criterion appropriate to optimization under constraints.

Denote $D^2_{\vec{x}}L$ the matrix of second partial derivatives of the Lagrangian L with respect to $\vec{x}$ and consider the quadratic form q that results from restricting $D^2_{\vec{x}}L$ to the nullspace of the Jacobian matrix of the constraints $(g_1, \ldots, g_{m-1})$. Of course, all the derivatives should be evaluated at the point of interest, $\vec{x}_0$. Then (see [8, Theorem 8.9, p. 154]), depending on whether q is indefinite, negative definite, or positive definite, we conclude that $g_m|_Y$ has either a saddle point, a strict local maximum, or a strict

[10]Notice the vacuous role of the continuity axiom, owing to the fact that any function is continuous on any finite subset of Euclidean space.

local minimum at $\vec{x}_0$. If q is only semidefinite, we cannot conclude anything about the nature of $\vec{x}_0$.

According to the theory of restricted quadratic forms, q can be classified in terms of the matrix $D^2_{\vec{x}}L$ bordered with the Jacobian of the g_i, and this matrix is nothing but the bordered Hessian. The assumption that the latter has a nonzero determinant implies that q is either indefinite or definite, but not semidefinite [20, Theorem 16.4, p. 389]. Thus the second order criterion allows us to classify $\vec{x}_0$ as either a saddle point or a strict (local) optimum. However, as mentioned in the first paragraph, $\vec{x}_0$ is certainly not a saddle point because it is a global optimum, so it must be a *strict* optimum. The word "strict" is crucial here: it implies that $\vec{x}_0$ has a neighbourhood U in Y such that $g_m|_U$ attains the value c_m at $\vec{x}_0$ and only at $\vec{x}_0$. Since this is true for every $\vec{x}_0 \in X$ and X is compact, it follows that $g_m(\vec{x}) = c_m$ only has finitely many solutions on Y or, otherwise stated, that X is actually a finite set. □

It is worthwhile to interpret the proof of Theorem 4 in the context of Morse functions, since we will then be able to generalize it to a proof of Theorem 5. This is the goal of the following two subsections.

4.2. **Morse functions.** Hidden in the proof of Theorem 4 is the concept of nondegeneracy of a critical point, whose abstract definition is as follows. Let M be a compact, boundariless manifold, and let $h : M \longrightarrow \mathbb{R}$ be a smooth function. A critical point $p \in M$ for h is said to be *nondegenerate* if the matrix of second partial derivatives of h at p has a nonzero determinant[11]. The map h itself is called a *Morse function* if all its critical points are nondegenerate.

In the case of Theorem 4 the manifold M is Y, which is a smooth compact submanifold of $\mathbb{R}^n$, and h is the restriction of the globally defined map g_m to Y, that is, $h = g_m|_Y$. The matrix of second partial derivatives of $h = g_m|_Y$ is the matrix of the restricted quadratic form q (this is not entirely obvious, since q was constructed

[11]One would express h in local coordinates around p and construct the matrix of second partial derivatives of this local expression. Whether or not this matrix has a nonzero determinant turns out to be independent of the coordinates chosen, and this makes the above definition valid. The interested reader can find more information about this in [17, Chapter 2, pp. 33ff.]

from the Lagrangian rather than directly from g_m, but see equation (3) in [8, p. 290]). With this language, the condition that the bordered Hessian has a nonzero determinant at each $\vec{x} \in X$ amounts to requiring that every critical point of $g_m|_Y$ is nondegenerate, that is, that $g_m|_Y$ is a Morse function.

By their very definition, nondegenerate critical points of a map h have the property that the quadratic term in the Taylor expansion of h around them is either definite or indefinite, but not semidefinite. In particular, regardless of the nature of the critical point (whether a saddle or a local optimum), it is isolated in the sense that it has neighbourhood U that contains no other critical point. As a consequence, a Morse function on a compact manifold M can only have finitely many critical points altogether (see for instance [17, Corollary 2.19, p. 47]). This is reminiscent of the proof of Theorem 4 and, in fact, we may now rephrase the latter as follows:

Theorem 4′. *Let the set Y be bounded and connected, $m < n$, and assume that the first $m-1$ constraints satisfy the constraint qualifications. Suppose that the last constraint is indeed optimal with respect to the others. In addition, assume that $g_m|_Y$ is a Morse function. Then the social choice problem over X has a solution.*

Proof. Every point in X is a critical point of $g_m|_Y$ because of the assumption about the optimality of the last constraint with respect to the others. Since Morse functions on compact manifolds have only finitely many critical points, it follows that X is finite and so the social choice problem over X has a solution. □

Yet another property of Morse functions of interest to us is that they are “generic” in the set $\mathcal{C}^\infty(M, \mathbb{R})$ of all smooth functions $h : M \longrightarrow \mathbb{R}$ (together with the above reformulation of Theorem 4, these will be the main ingredients in the proof of Theorem 5). More precisely, combining [17, Lemma 2.26, p. 52] and [17, Theorem 2.20, p. 47] one has:

Theorem. For a compact, boundariless, smooth manifold M, the set of Morse functions is open and dense in $\mathcal{C}^\infty(M, \mathbb{R})$ in the strong $\mathcal{C}^2$–topology.

The strong $\mathcal{C}^2$–topology on $\mathcal{C}^\infty(M,\mathbb{R})$ can be most easily described by saying that two functions $h_1, h_2 \in \mathcal{C}^\infty(M,\mathbb{R})$ are ϵ–close when the functions themselves, together with their derivatives up to second order, differ by no more than ϵ at each point of M. We will write $\|h_1 - h_2\| < \epsilon$ to denote this.[12] Equivalently, a sequence h_n converges to h if and only if the maps h_n, together with their partial derivatives up to second order, converge uniformly to h and its corresponding partial derivatives.

4.3. **The proof of Theorem 5.** Let us begin by explaining more carefully the statement of the theorem. Consider once more the bounded set Y defined by the first $m-1$ constraints alone. These should be thought of as being fixed once and for all, and we imagine that the last constraint $g_m(\vec{x}) = c_m$ is a parameter so that the map g_m and the number c_m can vary, yielding a whole family of sets of alternatives $X_{g_m}^{c_m}$. However, since we are only interested in the case when the last constraint is optimal with respect to the others, for each map g_m there are only two possible choices of c_m: either the global maximum $c_m^{\max}$ or the global minimum $c_m^{\min}$ of $g_m|_Y$. The content of Theorem 5 is that for most choices of g_m both possibilities lead to a set of alternatives where the social choice problem has a solution.

Let us formalize this idea. For any smooth map $g_m : \mathbb{R}^n \longrightarrow \mathbb{R}$ denote $c_m^{\max} = \max\ g_m|_Y$ and $c_m^{\min} = \min\ g_m|_Y$ (since Y is bounded by assumption, these two numbers are well defined) and consider the two sets of alternatives

$$X_{g_m}^{\max} = \{\vec{x} \in \mathbb{R}^n : g_1(\vec{x}) = c_1, \ldots, g_{m-1}(\vec{x}) = c_{m-1}, g_m(\vec{x}) = c_m^{\max}\}$$

and

$$X_{g_m}^{\min} = \{\vec{x} \in \mathbb{R}^n : g_1(\vec{x}) = c_1, \ldots, g_{m-1}(\vec{x}) = c_{m-1}, g_m(\vec{x}) = c_m^{\min}\}.$$

With this notation, the precise statement of Theorem 5 is the following:

Theorem 5′. *Let the set Y be bounded and connected, $m < n$, and assume that the first $m-1$ constraints satisfy the constraint qualifications. There is a set $\mathcal{M} \subseteq$*

[12]Since the derivatives depend on the coordinates chosen to compute them, this definition has to be set up with some care, but we have no need to go any further into these details. The interested reader is referred to [17, p. 51].

$\mathcal{C}^\infty(\mathbb{R}^n,\mathbb{R})$ which is open and dense (with the strong $\mathcal{C}^2$–topology) and such that, when g_m belongs to this set, the social choice problems over both $X_{g_m}^{\max}$ and $X_{g_m}^{\min}$ admit a solution.

Proof. Consider the set

$$\mathcal{M} = \{g \in \mathcal{C}^\infty(\mathbb{R}^n,\mathbb{R}) : g|_Y \text{ is a Morse function}\}.$$

According to Theorem 4′, whenever g_m belongs to $\mathcal{M}$ the social choice problem over both $X_{g_m}^{\max}$ and $X_{g_m}^{\min}$ has a solution. Therefore, we only need to show that $\mathcal{M}$ is open and dense in $\mathcal{C}^\infty(\mathbb{R}^n,\mathbb{R})$. Since we will work in $\mathcal{C}^\infty(Y,\mathbb{R})$ and in $\mathcal{C}^\infty(\mathbb{R}^n,\mathbb{R})$ simultaneously, we denote their respective distances by $\|\cdot\|_Y$ and $\|\cdot\|$.

(1) Openness. Pick a map $g \in \mathcal{M}$; that is, a smooth map $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $g|_Y$ is a Morse function. Since Morse functions form an open subset of $\mathcal{C}^\infty(Y,\mathbb{R})$ as discussed at the end of the previous section, there exists $\epsilon > 0$ such that any other smooth map $h_0 : Y \longrightarrow \mathbb{R}$ satisfying $\|g|_Y - h_0\|_Y < \epsilon$ is also a Morse function. Now, given any smooth $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\|g - h\| < \epsilon$, setting $h_0 = h|_Y$ one evidently has $\|g|_Y - h_0\|_Y \le \|g - h\| < \epsilon$ and so $h|_Y$ is also a Morse function, that is, $h \in \mathcal{M}$. Hence $\mathcal{M}$ is open in $\mathcal{C}^\infty(\mathbb{R}^n,\mathbb{R})$.

(2) Density. We have to show that given any $g \in \mathcal{C}^\infty(\mathbb{R}^n,\mathbb{R})$ and any $\epsilon > 0$ there is $h \in \mathcal{C}^\infty(\mathbb{R}^n,\mathbb{R})$ such that $\|g - h\| < \epsilon$ and $h|_Y$ is a Morse function. Consider the restriction $g|_Y$. By the density of Morse functions in $\mathcal{C}^\infty(Y,\mathbb{R})$ mentioned earlier there is a Morse function $h_0 : Y \longrightarrow \mathbb{R}$ such that $\|g - h_0\|_Y < \epsilon$. Extend h_0 to a smooth $\tilde{h}_0 : \mathbb{R}^n \longrightarrow \mathbb{R}$ and find an open neighbourhood U of Y in $\mathbb{R}^n$ such that $\|g - \tilde{h}_0\| < \epsilon$ on U. Choose an even smaller neighbourhood V of Y such that $\bar{V} \subseteq U$ and a smooth bump function $\Theta : \mathbb{R}^n \longrightarrow [0,1]$ such that $\Theta|_V \equiv 1$ and $\Theta \equiv 0$ outside U. Finally, let $h := \Theta \cdot \tilde{h}_0 + (1-\Theta)\cdot g$. This is a smooth function defined on $\mathbb{R}^n$ that coincides with h_0 on Y; therefore, $h|_Y$ is a Morse function. As for the distance between g and h, we have:

(i) on V the equality $h = \tilde{h}_0$ holds, so

$$\|g - h\|_V = \|g - \tilde{h}_0\|_V < \epsilon,$$

(ii) on $U - V$

$$\|g - h\|_{U-V} = \|g - \Theta \cdot \tilde{h}_0 - (1 - \Theta) \cdot g\|_{U-V} = (\sup_{U-V} |\Theta|) \cdot \|g - \tilde{h}_0\|_{U-V} < \epsilon,$$

(iii) outside U the equality $\Theta \equiv 0$ holds, so

$$\|g - h\|_{U^c} = \|g - g\|_{U^c} = 0.$$

Therefore $\|g - h\| < \epsilon$ on all of $\mathbb{R}^n$, proving that $\mathcal{M}$ is dense in $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$. □

5. Appendix A: proofs of Lemmas 6 and 7

In this appendix we establish Lemmas 6 and 7. As mentioned earlier, we need to use some homology theory. Since this topic is rather elaborate we cannot even recall here the basic definitions, so we refer the interested reader to the book by Hatcher [13] and limit ourselves to state the results that we need.

To any space U we may assign a sequence of (real) vector spaces $H_j(U; \mathbb{R})$ for $j = 0, 1, 2, \ldots$ which capture some geometric information about U. These are called the j–dimensional *homology groups* of U (even though they are actually vector spaces) with coefficients in the real numbers. For simplicity we shall just speak of the homology groups of U and denote them $H_j(U)$, supressing $\mathbb{R}$ from the notation.

The following properties hold:

(a) The dimension of $H_0(U)$ is the number of path connected components of U.

(b) If U is contractible, then $H_0(U) = \mathbb{R}$ and $H_j(U) = \{0\}$ for every $j \geq 1$.

(c) Poincaré duality: if U is a compact, boundariless manifold of dimension d, then $H_j(U) = H_{d-j}(U)$ for every j.

With these properties, the proof of Lemma 6 reduces to a simple computation:

Proof of Lemma 6. Let U be a connected component of M. Then U is itself a compact manifold of dimension $d \geq 1$ and without boundary. We have $H_0(U; \mathbb{R}) = \mathbb{R}$ by (a)

above, because U is connected. By Poincaré duality $H_d(U;\mathbb{R}) = H_0(U;\mathbb{R}) = \mathbb{R}$, and it follows from (b) that U is not contractible, because it has a homology group of dimension ≥ 1 (namely, its d–dimensional homology group) which is nonzero. $\square$

The proof of the second lemma requires slightly more sophisticated tools. We need to use *relative homology groups*, which are defined not just for a space U but for a pair (U, U_0) formed by a space U and a subset U_0 of U. That is, for each $j = 0, 1, 2, \ldots$ there is a real vector space $H_j(U, U_0)$ called the j–dimensional relative homology group of the pair (U, U_0) with coefficients in the real numbers. There is a relation between the relative homology of a pair (U, U_0) and the homology groups of both U and U_0, which is expressed by a so-called long exact sequence as follows. For each dimension j there are linear maps $H_j(U_0) \longrightarrow H_j(U)$, $H_j(U) \longrightarrow H_j(U, U_0)$ and $H_j(U, U_0) \longrightarrow H_{j-1}(U_0)$ that fit into a sequence

$$\ldots \longrightarrow H_2(U) \longrightarrow H_2(U, U_0) \longrightarrow H_1(U_0) \longrightarrow H_1(U) \longrightarrow H_1(U, U_0) \longrightarrow 0$$

(which continues to the left in the same fashion) having the property of being *exact*: the image of the map entering any one of the terms of the sequence coincides with the kernel of the map connecting that term to the one to its right.

In addition to this, we shall also make use of

(d) Lefschetz duality: if M is a compact manifold (with boundary) of dimension d, then $H_j(M, \partial M) = H_{d-j}(M)$ for every j.

Proof of Lemma 7. By Lefschetz duality $H_1(M, \partial M) = H_{d-1}(M)$. The latter homology group is zero by (b), because M is contractible and $d - 1 \geq 1$, so $H_1(M, \partial M)$ is zero too. Also, again by (b) we have that $H_0(M) = \mathbb{R}$.

Consider the following portion of the long exact sequence for the pair $(M, \partial M)$:

$$H_1(M, \partial M) \xrightarrow{\Delta} H_0(\partial M) \longrightarrow H_0(M) \longrightarrow 0.$$

Since $H_1(M, \partial M) = \{0\}$, the image of Δ is zero and, by the exactness of the sequence, the map connecting $H_0(\partial M)$ to $H_0(M)$ has zero kernel. It is therefore injective. In a similar fashion one proves that it is surjective, now analyzing what happens with the arrow $H_0(M)$ to 0. Hence it is an isomorphism, which shows that $H_0(\partial M) = H_0(M) = \mathbb{R}$. Using (a) we conclude that ∂M is connected. □

6. Appendix B: proofs of Lemmas 10 and 11

In this section, and to unclutter the notation, we shall denote the elements of X_u and Y with letters $p, q, \ldots$ instead of vectors $\vec{x}, \vec{y}, \ldots$ as before.

6.1. **Proof of Lemma 10.** It is both notationally and conceptually simpler to prove a slightly more general result, from which Lemma 10 follows letting $M = X_{[u_1,u_2]}$ and $Z = X_u$:

Lemma 13. *Let M be a compact manifold and $Z \subseteq M$ a closed subset of M. Suppose $F : Z^k \longrightarrow Z$ is an SCF over Z. Then there exist a neighbourhood U of Z in M and a continuous map $F_U : U^k \longrightarrow M$ such that F_U is unanimous and anonymous.*

Proof. Think of F as a mapping $F : Z^k \longrightarrow M$ and extend it setting $F(p, \ldots, p) = p$ for every $p \in M$. Now its domain is

$$D := Z^k \cup \{(p, \ldots, p) : p \in M\},$$

which is a compact subset of M^k. Clearly F is still continuous on this new larger domain D.

Consider the quotient space obtained from M^k by identifying, via an equivalence relation $\sim$, each k–tuple $(p_1, \ldots, p_k)$ with all of its permutations. We shall denote $\pi : M^k \longrightarrow M^k/\sim$ the canonical projection. The set D projects onto a compact subset $\pi(D)$ of $M^k/\sim$. In turn the map F, due to its invariance under permutation of its arguments, descends to a continuous map

$$\bar{F} : \pi(D) \longrightarrow M.$$

Now we make use of the following extension result: every continuous map from a closed subset of a metric space into a manifold M can be extended to a continuous map defined on a neighbourhood W of the subset (see for instance [15, Proposition 8.3, p. 47]). Applying this result to the closed subset $\pi(D)$ of the metric space $M^k/\sim$ and the map $\bar{F}$ we see that the latter can be extended continuously to a neighbourhood W of $\pi(D)$ in $M^k/\sim$. For notational ease the extension will still be denoted $\bar{F}$.

The set $\pi^{-1}(W)$ is a neighbourhood of D in M^k, so in particular it is a neighbourhood of Z^k. It is easy to see that there exists a neighbourhood U of Z in M such that $U^k \subseteq \pi^{-1}(W)$. Then the map

$$F_U : U^k \longrightarrow M \quad ; \quad F_U(p_1, \ldots, p_k) = (\bar{F} \circ \pi)(p_1, \ldots, p_k)$$

provides the desired extension: it is clearly continuous and unanimous, and it is also anonymous because any two permutations of a k–tuple $(p_1, \ldots, p_k)$ are projected by π onto the same element of $M^k/\sim$. □

6.2. **Proof of Lemma 11.** The construction of the map r is rather indirect: we shall define a tangent vectorfield on Y, consider the flow φ that it generates and then use φ to define r. This approach is closely related to Morse theory, and a quick glance at the book by Milnor [18, pp. 12 and 13] may be useful. Some acquaintance with differential geometry is required to follow the argument.

Recall that $Y \subseteq \mathbb{R}^n$ is a differentiable manifold defined by the constraints $g_i(p) = c_i$ for $i = 1, 2, \ldots, m-1$. At any point $p \in Y$ their gradients are all orthogonal to Y or, otherwise stated, the tangent space to Y at p is the subspace of $\mathbb{R}^n$ orthogonal to all the $\{\nabla g_i(p) : 1 \leq i \leq m-1\}$. Denote $V(p)$ the projection of $\nabla g_m(p)$ onto that tangent space, thus obtaining a tangent vectorfield $p \mapsto V(p)$ on Y. This vectorfield $V(p)$ can be given a very rough but rather helpful intuitive interpretation: inasmuch as $\nabla g_m(p)$ tells us the direction along which f increases most quickly, its projection $V(p)$ tells us in what direction we should advance to obtain the quickest increase of g_m *while remaining in* Y.

Assertion 1. $V(p)$ is zero precisely when p is a critical point of $g_m|_Y$.

Proof. Notice that $V(p)$ is zero precisely when $\nabla g_m(p)$ is orthogonal to Y at p; that is to say, precisely when $\nabla g_m(p)$ is a linear combination of the gradients $\{\nabla g_i(p) : 1 \leq i \leq m-1\}$ or, equivalently, when p is a critical point of $g_m|_X$. □

Assertion 2. The scalar product $\nabla g_m(p) \cdot V(p)$ is always nonnegative and it is actually positive when p is not a critical point of $g_m|_Y$.

Proof. Observe that by construction the angle between $\nabla g_m(p)$ and $V(p)$ is at most ninety degrees, so the scalar product $\nabla g_m(p) \cdot V(p)$ is always nonnegative. Together with Assertion 1, this proves the result. □

Using $V(p)$ we define a new tangent vectorfield $W : p \mapsto (u - g_m(p))V(p)$. Let $\varphi : Y \times \mathbb{R} \longrightarrow Y$ be the flow generated on Y by the vectorfield W. Notice that, since Y is compact, φ is globally defined.

At this point it may be helpful to have a look at Figure 1. Panel (a) shows a very simple set of alternatives Y in $\mathbb{R}^3$ defined by a single restriction, so Y is a surface. The remaining constraint g_m in this case is taken to be $g_m(x,y,z) = z$. The Z axis is represented vertically, so the level set X_u is simply the intersection of Y with the horizontal plane at height u. The gradient of $g_m|_X$ is the vertical vector $(0,0,1)$, which is perpendicular to Y precisely at the two points p and q (and possibly others not shown in the picture); these are, then, critical points[13]. Since p belongs to X_u, we see that u is not a regular a value of $g_m|_Y$ and we cannot expect X_u to be manifold. And indeed, as shown in the drawing, X_u is an "eight-figure" (two circumferences having a single point in common) so it is not a manifold. By contrast, u_1 and u_2 are regular values of $g_m|_X$ since X_{u_1} and X_{u_2} do not contain critical points, and they are both manifolds.

Let us focus our attention on the set $X_{[u_1,u_2]}$, which is the whole region of Y comprised between heights u_1 and u_2. It is also a manifold, this time with boundary. Panel (b) shows $X_{[u_1,u_2]}$ together with an sketch of $W(p)$ and φ. We mentioned earlier

[13]Notice that q is a local minimum but p is neither a local maximum nor a local minimum (it is a saddle point). This is related to the fact that being a critical point is a necessary but not sufficient condition for being a local optimum, as already highlighted in Section 2.

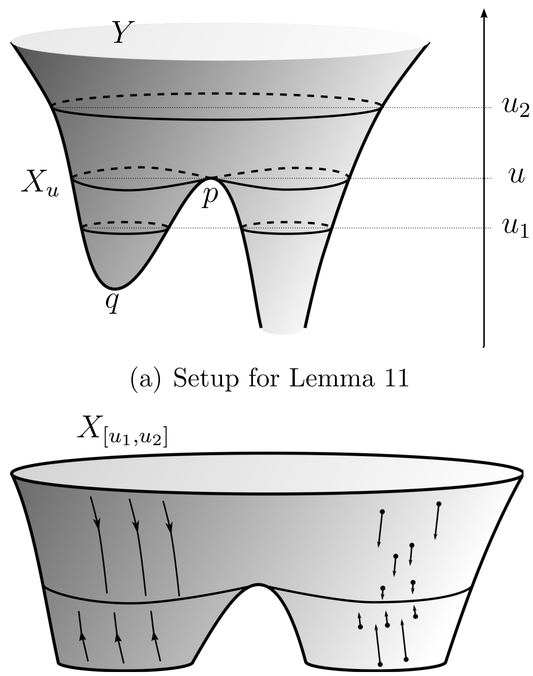
that $V(p)$ points, at each $p \in X$, in the direction that we should follow, moving within Y, to obtain the quickest increase of g_m. Given that $g_m(x, y, z) = z$, in our case $V(p)$ points in the direction of quickest ascent along Y from p. The vector field $W(p)$ is obtained multiplying $V(p)$ by the modulating factor $u - g_m(p)$, which is zero precisely on X_u, negative above X_u and positive below X_u. Taking into account these signs, $W(p)$ is zero on X_u, points in the direction of quickest *descent* if p is above X_u, and points in the direction of quickest ascent if p is below X_u. The small arrows to the right of panel (b) in Figure 1 are intended to convey this idea. If we follow the directions of these arrows, starting at any point p, it seems clear that we will move towards X_u advancing ever more slowly, since $W(p)$ (which is our speed) becomes smaller the closer we get to X_u. Unless $p \in X_u$, in which case we would actually stay still since $W(p) = 0$, we would approach X_u asymptotically but never get there. In any case, there will be a finite time t_p at which we will enter any prescribed neighbourhood U of X_u and never leave it again. The map r that we are looking for will essentially be defined as $r(p) =$ the point we reach at time t_p. In our trip from p to $r(p)$ we might follow a simple path like the ones shown to the left side of the drawing or, possibly, a much more complicated one which approaches X_u spiralling around it or in some other strange fashion.

Let us go back to mathematics again. The following proposition collects some properties of φ that are the formal counterparts of the ideas just described:

Proposition 14. *The flow φ has the following properties:*

(1) *For every $q \in X_{[u_1,u_2]}$ and every $t \geq 0$, the point $\varphi(q,t)$ belongs to $X_{[u_1,u_2]}$ too.*
(2) *For every neighbourhood U of X_u in $X_{[u_1,u_2]}$ there exists $T > 0$ such that $\varphi(q,t) \in U$ for every $q \in X_{[u_1,u_2]}$ and every $t \geq T$.*

In the parlance of dynamical systems, (1) means that $X_{[u_1,u_2]}$ is positively invariant under φ and (2) states that X_u is a stable attractor in $X_{[u_1,u_2]}$. As a preparation to prove the proposition we are going to investigate some qualitative properties of the trajectories of φ.



(a) Setup for Lemma 11

(b) The vector field W and the flow φ

FIGURE 1.

Fix a point $q \in X_{[u_1,u_2]}$ and let γ be the trajectory of φ with initial condition $\gamma(0) = q$ (in terms of the flow, $\gamma(t) = \varphi(q,t)$). More explicitly, $\gamma : \mathbb{R} \longrightarrow Y$ is a smooth curve in Y which satisfies $\gamma(0) = q$ and

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t}(t) = W(\gamma(t))$$

for every $t \in \mathbb{R}$ (that is, γ is an integral curve of the vectorfield W). We are interested in the behaviour of $\gamma(t)$ for $t \geq 0$, and for definiteness we consider the case $u < f(q) \leq u_2$.

Assertion 3. The inequality $u < g_m(\gamma(t))$ holds for every $t \in \mathbb{R}$.

Proof. Each point of X_u is a zero of W and therefore a fixed point of the flow φ. Since the trajectory γ goes through the point q, which does not belong to X_u, it follows that $\gamma(t) \notin X_u$ for every $t \in \mathbb{R}$. Consider the map $t \mapsto (g_m \circ \gamma)(t)$. By what we have just seen, it never attains the value u, so (as it continuous) it must be the case that

$g_m(\gamma(t))$ is either always $> u$ or $< u$. Since we have taken $g_m(q) > u$, it follows that $g_m(\gamma(t)) \neq u$ for every $t \in \mathbb{R}$. □

Assertion 4. The inequality $g_m(\gamma(t)) \leq u_2$ holds for every $t \geq 0$.

Proof. Using the chain rule we compute the time derivative of the map $t \mapsto (g_m \circ \gamma)(t)$ as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}(g_m \circ \gamma)(t) = \nabla g_m(\gamma(t)) \cdot \frac{\mathrm{d}\gamma}{\mathrm{d}t}(t) = \nabla g_m(\gamma(t)) \cdot W(\gamma(t)),$$

and since by definition $W(p) = (u - g_m(p))V(p)$, we have $\nabla g_m(p) \cdot W(p) = (u - g_m(p))\nabla g_m(p) \cdot V(p)$, so

$$\frac{\mathrm{d}}{\mathrm{d}t}(g_m \circ \gamma)(t) = (u - g_m(\gamma(t)))\nabla g_m(\gamma(t)) \cdot V(\gamma(t)). \tag{2}$$

The right hand side is the product of two factors. The first is $u - f_m(\gamma(t))$, which is strictly negative by the previous assertion, and the second is $\nabla g_m(\gamma(t)) \cdot V(\gamma(t)) \geq 0$ which is nonnegative by Assertion 2. Thus the derivative of $t \mapsto (g_m \circ \gamma)(t)$ is nonpositive, and so the map is nonincreasing. In particular, since at $t = 0$ we have $(g_m \circ \gamma)(t) = g_m(q) \leq u_2$, this same inequality holds for all $t \geq 0$. □

Assertion 5. For any $p \in X_{[u_1,u_2]}$ such that $u < g_m(p) \leq u_2$ the inequality

$$(u - g_m(p))\nabla f(p) \cdot V(p) < 0$$

holds true.

Proof. Since $u - g_m(p) < 0$, we only need to prove that $\nabla g_m(p) \cdot V(p) > 0$. By Assertion 2, this scalar product is always nonnegative and it is zero precisely when p is a critical point of $g_m|_X$. Now, the choice of u_1 and u_2 guarantees that the critical points that $g_m|_Y$ may have in the set $X_{[u_1,u_2]}$ are all contained in X_u. Since $p \notin X_u$, the assertion follows. □

Assertion 6. $g_m(\gamma(t)) \to u$ as $t \to +\infty$.

Proof. Notice that $t \mapsto (g_m \circ \gamma)(t)$ must indeed converge to some u_* as $t \to +\infty$ because according to the computation in Assertion 4 it is a monotonous nonincreasing function bounded below by u. Let us assume that u_* is strictly larger than u and arrive at a contradiction.

Set $D(p) = (u - g_m(p))\nabla g_m(p) \cdot V(p)$ for brevity. Let $X_{u_*} = \{p \in Y : g_m(p) = u_*\}$. This set is closed in Y, so it is compact. Also, the previous assertion says that $D(p) < 0$ for every $q \in X_{u_*}$. Since X_{u_*} is compact and D is continuous, there is an $\epsilon < 0$ such that $D(p) < \epsilon$ for every $p \in X_{u_*}$. In fact, more is true: there is a neighbourhood U of X_{u_*} in $X_{[u_1,u_2]}$ where the same inequality holds; that is, $D(p) < \epsilon$ for every $p \in U$.

We are almost finished. Since $g_m(\gamma(t)) \to u_*$ as $t \to +\infty$, there exists $T > 0$ such that $\gamma(t) \in U$ for every $t > T$. By the mean value theorem, for any t there exists ξ_t between t and $t + 1$ such that

$$g_m(\gamma(t+1)) - g_m(\gamma(t)) = \frac{\mathrm{d}}{\mathrm{d}t}(g_m \circ \gamma)(\xi_t) = D(\gamma(\xi_t)), \tag{3}$$

where in the last equality we have used equation (2). Let us consider what happens in the above expression when $t \to +\infty$. Since ξ_t lies between t and $t + 1$, as soon as $t > T$ we also have $\xi_t > T$ and therefore $\gamma(\xi_t) \in U$, which entails $D(\gamma(\xi_t)) < \epsilon$. Thus the right hand side of (3) is bounded away from 0 (recall that $\epsilon < 0$). However, its left hand side converges to 0 as $t \to +\infty$ because both summands converge to u_*. This contradiction finishes the proof. □

Proof of Proposition 14. (1) We have seen that for an initial condition $q = \gamma(0)$ satisfying $u < g_m(q) \leq u_2$, the trajectory $\gamma(t)$ remains in the set $\{p \in Y : u < g_m(p) \leq u_2\}$ for every $t \geq 0$. Evidently, if the initial condition q satisfies $u_1 \leq g_m(q) < u$, similar arguments show that $\gamma(t)$ remains in the set $\{p \in Y : u_1 \leq g_m(p) < u\}$ for all $t \geq 0$. The remaining case, $g_m(q) = u$, is very simple: q is then a zero of the vectorfield W and so $\gamma(t) = q$ for every $t \in \mathbb{R}$. Summing up, for an initial condition $q \in X_{[u_1,u_2]}$

the trajectory γ remains in the set $X_{[u_1,u_2]}$. Part (1) of the proposition is just a restatement of this, since in terms of the flow the trajectory γ with initial condition $q = \gamma(0)$ is simply $\gamma(t) = \varphi(q,t)$.

(2) Find u_1' and u_2' such that $u_1 < u_1' < u < u_2' < u_2$ and $X_{[u_1',u_2']} \subseteq U$. In accordance with the notation we have been using so far, denote

$$X_{(u_1',u_2')} = \{p \in Y : u_1' < g_m(p) < u_2'\},$$

which is an open subset of $X_{[u_1,u_2]}$ by continuity of g_m. In fact it is a neighbourhood of X_u in $X_{[u_1,u_2]}$, so by Assertion 6 for each $q \in X_{[u_1,u_2]}$ there exists $t_q \geq 0$ such that $\varphi(q,t_q) \in X_{(u_1',u_2')}$. Now the continuity of φ guarantees that q has an open neighbourhood U_q in $X_{[u_1,u_2]}$ such that $\varphi(U_q \times \{t_q\}) \subseteq X_{(u_1',u_2')}$. In particular $\varphi(U_q \times \{t_q\}) \subseteq X_{[u_1',u_2']}$, and by part (1) of this proposition (applied to $X_{[u_1',u_2']}$ rather than $X_{[u_1,u_2]}$) we see that $\varphi(U_q \times \{t\}) \subseteq X_{[u_1',u_2']}$ for every $t \geq t_q$. The U_q cover the compact set $X_{[u_1,u_2]}$, so a finite family of them cover it too, say $U_{q_1}, U_{q_2}, \ldots, U_{q_r}$. Let T be the maximum of $t_{q_1}, t_{q_2}, \ldots, t_{q_r}$. Then whenever $t \geq T$ we have that $\varphi(q,t) \in X_{[u_1',u_2']}$ for every $t \in X_{[u_1,u_2]}$, proving the proposition. ☐

We are finally ready to prove Lemma 11. For the convenience of the reader, we restate it here:

Lemma. *Given any neighbourhood U of X_u in $X_{[u_1,u_2]}$ there exists a continuous mapping $r : X_{[u_1,u_2]} \longrightarrow X_{[u_1,u_2]}$ such that:*

(1) *$r(p) \in U$ for every $p \in X_{[u_1,u_2]}$,*
(2) *r is homotopic to the identity in $X_{[u_1,u_2]}$.*

Proof. According to Proposition 14 there exists $T \geq 0$ such that $\varphi(p,t) \in U$ for every $p \in X_{[u_1,u_2]}$ and every $t \geq T$. Let r be defined by $r(p) := \varphi(p,T)$. By construction $r(p) \in U$, so indeed satisfies condition (1). Also, r is homotopic to the identity: the flow $\varphi(p,t)$, for $0 \leq t \leq T$, provides a suitable homotopy. Thus the lemma is proved. ☐

7. Concluding remarks

Many of the sets of interest in economics are naturally described as subsets of Euclidean space defined by a number of constraints. Solving the social choice problem over such a set X can be very hard even with the aid of the classical characterization of Chichilnisky and Heal (i.e., that X should be contractible) because: (i) the contractibility condition is probably unfamiliar to someone without a specific mathematical background in topology, (ii) deciding whether a given set X is contractible is in general very difficult, and even more so because describing X in terms of constraints makes it difficult to gain any geometric intuition about it.

Motivated by this, we have provided several criteria that are easy to check and solve the social choice problem over sets X defined in terms of constraints. Besides their practical use, these criteria also have two interesting theoretical consequences. The first one is that, generically, the social choice problem over a set X defined by equality constraints $g_i(\vec{x}) = c_i$ has no solution. Thus, if the constraints g_i come from some natural or random process (in a nontechnical sense of the word), with probability one the social choice problem over X will have no solution. In fact (and this is the second consequence), the c_i have to be very finely tuned indeed if we want a social choice function over X to exist: namely, one of the c_i has to be either the global maximum or the global minimum of its constraint g_i over the set of alternatives defined by the remaining constraints $g_j = c_j$, $j \neq i$. Thus, if one of the constraint values is not fixed but can be operated upon by some agent, it must be carefully chosen to be a global optimum. We call this the principle of rational design. In this sense, the only way to avoid a social choice paradox consists in designing the set of alternatives (if possible) in a way that is natural in economics. An interesting point to observe here is that the need for optimization emerges unexpectedly and not as a consequence of any assumption concerning rationality, utility functions, or any other element related to economics, of which there are none in the social choice problem under consideration.

References

[1] K. J. Arrow. *Social choice and individual values*. Yale University Press, 1951.

[2] G. Aumann. Beiträge zur Theorie der Zerlegungsräume. *Math. Ann.*, 1:249–294, 1936.

[3] N. Baigent. *Topological Theories of social choice*, In Handbook of social choice and Welfare: volume 2, (K. J. Arrow, A. K. Sen and K. Suzumura, editors), Chapter 18. Elsevier, 2011.

[4] J. C. Candeal and E. Induráin. The Moebius strip and a social choice paradox. *Economics Letters*, 45: 407–412, 1994.

[5] G. Chichilnisky. On fixed point theorems and social choice paradoxes. *Economics Letters*, 3: 347–351, 1979.

[6] G. Chichilnisky. Social choice and the topology of spaces of preferences. *Advances in Mathematics*, 37: 165–176, 1980.

[7] G. Chichilnisky and G. Heal. Necessary and sufficient conditions for a resolution of the social choice paradox. *Journal of Economic Theory*, 31:68–87, 1983.

[8] A. de la Fuente. *Mathematical Methods and Models for Economists*. Cambridge University Press, 2000.

[9] B. Eckmann. Raüme mit Mittelbildungen. *Comment. Math. Helv.*, 28: 329–340, 1954.

[10] B. Eckmann. Social choice and topology. A case of pure and applied mathematics. *Expo. Math.*, 22: 385–393, 2004.

[11] B. Eckmann, T. Ganea and P. Hilton. *Generalized means*. Studies in mathematical analysis and related topics, 82–92. Stanford University Press, 1962.

[12] C. H. Edwards. *Advanced calculus of several variables*. Academic Press, 2015.

[13] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[14] P. Hilton. A new look at means on topological spaces. *Internat. J. Math. Math. Sci.* 20(4): 617–620, 1997.

[15] S. Hu. *Theory of retracts*. Wayne State University Press, 1965.

[16] L. Lauwers. Topological social choice. *Mathematical Social Sciences*, 40:1–39, 2000.

[17] Y. Matsumoto. *An introduction to Morse theory*. Translations of Mathematical Monographs, 1997.

[18] J. Milnor. *Morse theory*. Annals of Mathematical Studies. Princeton University Press, 1963.

[19] J. Milnor. *Topology from the differentiable viewpoint*. Princeton University Press, 1997.

[20] C. P. Simon and L. Blume. *Mathematics for economists*. W. W. Norton & Company, 1994.

[21] E. H. Spanier. *Algebraic topology*. McGraw–Hill Book Co., 1966.

[22] S. Weinberger. On the topological social choice model. *Journal of Economic Theory*, 115(2): 377-384, 2007.

[23] J. H. C. Whitehead. On C^1-Complexes. *Ann. of Math.*, 41(2): 809–824, 1940.

[24] H. Whitney. *Geometric integration theory*. Princeton University Press, 1957.

(Juan A. Crespo) Facultad de Ciencias Económicas, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain

E-mail address: juan.crespo@uam.es

(J. J. Sánchez-Gabites) Facultad de Ciencias Económicas, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain

E-mail address: JaimeJ.Sanchez@uam.es