UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Evolution in games with a continuous action space

van Veelen, M.; Spreij, P.

DOI
10.1007/s00199-008-0338-8
Publication date
2009
Document Version
Final published version
Published in
Economic Theory

Link to publication

Citation for published version (APA):
van Veelen, M., & Spreij, P. (2009). Evolution in games with a continuous action space. *Economic Theory*, *39*(3), 355-376. https://doi.org/10.1007/s00199-008-0338-8

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Sep 2026

Econ Theory (2009) 39:355–376
DOI 10.1007/s00199-008-0338-8

OpenAccess

RESEARCH ARTICLE

Evolution in games with a continuous action space

Matthijs van Veelen · Peter Spreij

Received: 3 September 2007 / Accepted: 1 February 2008 / Published online: 28 February 2008
© The Author(s) 2008

Abstract Allowing for games with a continuous action space, we investigate how evolutionary stability, the existence of a uniform invasion barrier, local superiority and asymptotic stability relate to each other. This is done without restricting the populations of which we want to investigate the stability to monomorphic population states or to strategies with finite support.

Keywords Evolutionary stability · Uniform invasion barrier · Local superiority · Asymptotic stability · Continuous action space · Topology

JEL Classification C70

1 Introduction

The recent literature on evolutionary dynamics for populations playing games with a continuous action space shows that generalizing results from games with a finite action space is not always straightforward and sometimes not even possible. This can

The authors have benefitted from careful and precise comments by Gerard van der Laan, Jan van Mill, the editor and an anonymous referee. Matthijs van Veelen gratefully acknowledges financial support by the Netherlands' Organization for Scientific Research (NWO).

M. van Veelen (✉)
Department of Economics and Econometrics, Universiteit van Amsterdam,
Roetersstraat 11, 1018 WB Amsterdam, The Netherlands
e-mail: C.M.vanVeelen@uva.nl

P. Spreij
Korteweg-de Vries Institute for Mathematics, Universiteit van Amsterdam,
Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands

be learned from Oechssler and Riedel (2001, 2002) and from Cressman (2005), Cressman and Hofbauer (2005) and Cressman et al. (2006). These authors have therefore chosen to focus for an important part on pure strategies—or monomorphic population states—as candidates for stability in population dynamics. This is done for very good reasons, because such restrictions on the candidates for stability allow useful results to be established and still encompass a wide range of interesting equilibria of games. In this paper, however, we choose to avoid making such restrictions on the candidates for asymptotic stability and the aim is to push the limit of how far we can get without them a little further.

An important part of the theory for games with a finite action space consists of connecting static properties that a strategy can have—such as evolutionary stability, the existence of a uniform invasion barrier and local superiority—to notions that apply to population dynamics, the most important of which is asymptotic stability in the replicator dynamics. The main results for finite action spaces can be summarized very shortly; the three static properties mentioned above are equivalent and they imply asymptotic stability in the replicator dynamics (see Weibull 1996). In Sect. 3 we look at the equivalents of those concepts for games with a continuous action space and investigate how they are related. This is done by bringing together results from the literature with a number of missing links.

Finally there is an appendix that concerns some topological details. This is needed, for in Bomze (1991), which is an important and often quoted paper in this field, it is tacitly assumed that auxiliary results on Kullback–Leibler I-divergences borrowed from Csiszár (1962, 1964, 1967) use modern, standard topology definitions. The appendix first indicates why this cannot be the case, and then it shows how the ingredients of Csiszár's proofs can be used to prove the auxiliary results when using modern standard topology textbook definitions too.

2 Definitions

The action space S is a complete separable metric space with metric $\rho(x, y)$. A strategy is a probability measure on $(S, \mathcal{B})$ with Borel σ-field $\mathcal{B}$ and $\mathcal{P}[S, \mathcal{B}]$ denotes the set of all such probability measures. We will consider pairwise contests and random matching unless explicitly mentioned otherwise; players are drawn at random from a population to play a symmetric 2-player game, which is characterized by a function $A : S \times S \to \mathbb{R}$. Here $A(x, y)$ is the payoff to player 1 if she plays x and the opponent plays y. (Please note that symmetry only implies that the payoff to player 2 is $A(y, x)$ and not that $A(y, x) = A(x, y)$, which would make it a doubly symmetric game). The expected payoff to a player playing strategy P against a player playing Q is given by the function $u[P, Q] = \int \int A(x, y)\, \mathrm{d}P(x)\, \mathrm{d}Q(y)$.

Symmetric games, pairwise contests and random matching is exactly the setting chosen by Weibull (1996) for finite action spaces, but one could think of more general forms of this function $u[P, Q]$ that are not necessarily bilinear. This would allow for games such as the sex ratio game as described by Sigmund (1987), but it should be noted that any expected payoff function that is not bilinear implies a departure from the setting of pairwise contests, as Bomze and Pötscher (1989) show. Some of the

results in this paper however do not use the bilinearity of the function $u[P, Q]$ and therefore remain valid in a more general setting where $u[P, Q] = \int F_Q(x)\, dP(x)$, $F_Q : S \to \mathbb{R}$, $Q \in \mathcal{P}[S, \mathcal{B}]$ (see for instance Sect. 3 arrow 2 and Proposition 16). The counterexamples on the other hand are all bilinear, which makes them useful for disproving implications even in the setting of pairwise contests.

The following definitions are close copies of their discrete counterparts.

Definition 1 $\epsilon_P(Q)$ is an invasion barrier for P against Q, $Q \neq P$, if $u\big[P, (1-\epsilon)P + \epsilon Q\big] > u[Q, (1-\epsilon)P + \epsilon Q]$ for all $\epsilon \in (0, \epsilon_P(Q))$.

It will be useful to have a notion for the largest invasion barrier, so we will let $\overline{\epsilon}_P(Q)$ be the supremum of $\epsilon_P(Q) \in (0, 1)$ if there is an invasion barrier in $(0, 1)$ and 0 if there is not.

Definition 2 P is an evolutionary stable strategy if there exists an invasion barrier $\epsilon_P(Q) \in (0, 1)$ for every strategy $Q \neq P$.

An equivalent definition of evolutionary stability is that both $u[P, P] \geq u[Q, P]$ for all Q and $u[P, P] = u[Q, P] \Rightarrow u[P, Q] > u[Q, Q]$ for all $Q \neq P$ hold.

Definition 3 P has a uniform invasion barrier if there exists an $\overline{\epsilon} \in (0, 1)$ such that $u[P, (1-\epsilon)P + \epsilon Q] > u[Q, (1-\epsilon)P + \epsilon Q]$ for all $Q \neq P$ and all $\epsilon \in (0, \overline{\epsilon})$.

A strategy that has a uniform invasion barrier is also referred to as *uninvadable*.

For the third stability concept we will need a definition of a *vicinity*. A vicinity is defined with the help of a function $d : \mathcal{P}[S, \mathcal{B}] \times \mathcal{P}[S, \mathcal{B}] \to [0, \infty]$, of which we only assume that $d(P, P) = 0$ and $d(P, Q) > 0$ for $P \neq Q$, which implies that d not necessarily is a distance. V is a vicinity of P with respect to d if and only if there is a $\delta > 0$ such that $V = \{Q \in \mathcal{P}[S, \mathcal{B}] \mid d(Q, P) < \delta\}$. If this function d is in fact a distance, then the set of all vicinities of all population states P makes a base for a topology. In that case we can read *neighbourhood* whenever the word *vicinity* is used. However, one of the examples that follow is not a distance, and therefore it is better to avoid using neighbourhoods, that are only defined once there is a topology. The details of the distinction between vicinities and neighbourhoods can be found in the appendix.

Definition 4 P is locally superior with respect to d if it has a vicinity V with respect to d such that $u[P, Q] > u[Q, Q]$ for all $Q \neq P$ in V.

Please note that there is a minor difference with the discrete counterpart (Hofbauer, Schuster & Sigmund, 1979, see also Weibull 1996) where the choice of a function d is not an issue, and *neighbourhoods* are used.

Whether or not a strategy is locally superior obviously depends on what we consider to be a vicinity. In this section, we will consider a few possible functions d that define different types of vicinities, but the question what vicinities we allow for is not just of technical interest. After all, it would be nice if we could see a vicinity of a population state P as a set of states that can be reached from P by mutations or by a shock with relatively large probability.

The first function d we consider is the Prohorov metric. This metric is defined as follows (see for instance Bickel et al. 1993).

$$d_{pr}(Q,P)=\inf\left\{\varepsilon>0\;\middle|\;\begin{array}{l}P(B)\leq Q(B^{\varepsilon})+\varepsilon\\ Q(B)\leq P(B^{\varepsilon})+\varepsilon\end{array}\quad\text{for all } B\in\mathcal{B}\right\},$$

where $B^{\varepsilon}=\{x\in S\mid\exists\, y\in B \text{ such that } \rho(x,y)<\varepsilon\}$. It is worth knowing that the Prohorov metric metrizes weak convergence, that is, $d_{pr}(P_n,P)\to 0$ if and only if $\int g(x)\,\mathrm{d}P_n(x)\to\int g(x)\,\mathrm{d}P(x)$ for all bounded and continuous g on S. The topology it induces is called the weak topology.

The second option is the variational distance,

$$\begin{aligned}d_{vd}(Q,P)&=2\sup\{|P(B)-Q(B)|\;B\in\mathcal{B}\}\\&=\int\left|\frac{\mathrm{d}P}{\mathrm{d}\mu}-\frac{\mathrm{d}Q}{\mathrm{d}\mu}\right|\mathrm{d}\mu,\\&\text{for any } \mu \text{ dominating both } P \text{ and } Q,\end{aligned}$$

which induces the strong topology. To illustrate the difference between the vicinities we make with those two functions, we can take $S=[0,1]$ and $P=\delta_0$ the point measure at 0. Now compare the collection $V_{pr}(P)$ of all vicinities of P in the Prohorov metric to $V_{vd}(P)$ that contains all vicinities of P in the variational distance. It is not too hard to see that these two collections of vicinities share no elements apart from the whole set $\mathcal{P}[S,\mathcal{B}]$; just take δ_y, the point measures at y. For $y\in(0,1]$ we find that $d_{vd}(\delta_y,P)=2$ whereas $d_{pr}(\delta_y,P)=y$. Therefore any vicinity of P in the Prohorov metric contains probability measures δ_y (choose $y<\delta$ for a vicinity $V_{pr}(P,\delta)=\{Q\in\mathcal{P}[S,\mathcal{B}]\mid d_{pr}(Q,P)<\delta\}$) while none of these δ_y is contained in any of the vicinities of P in the variational distance, apart from, of course, the whole set $\mathcal{P}[S,\mathcal{B}]$ which is a vicinity in the variational distance for $\delta>2$ and in the Prohorov metric for $\delta>1$.

Going in the opposite direction is more successful. Because $d_{pr}(Q,P)\leq\frac{1}{2}d_{vd}(Q,P)$, which follows directly from the definitions, we find that $V_{vd}(P,2\delta)\subset V_{pr}(P,\delta)$.

A third possibility is to use the Kullback–Leibler I-divergence (see also Bomze 1991).[1]

$$d_{kl}(Q,P)=H_P(Q)=\begin{cases}\displaystyle\int\limits_{\mathrm{supp}(P)}\log\left(\frac{\mathrm{d}P(x)}{\mathrm{d}Q(x)}\right)\mathrm{d}P(x) & \text{if } P\ll Q\\ \infty & \text{if not}\end{cases}$$

This function does not satisfy the triangle inequality and is not symmetric. In the appendix we show that the set of all vicinities of all distributions $P\in\mathcal{P}[S,\mathcal{B}]$ cannot be a base for a topology. With the Kullback–Leibler I-divergence we therefore have to resort to using vicinities rather than neighbourhoods.

To see where the vicinities we make with this Kullback–Leibler I-divergence differ from the other two, take $S=[0,1]$ and P uniform on S. Again we make a

[1] Names that are also frequently used are Kullback–Leibler relative-entropy measure and Kulback–Leibler distance. Since it is neither a distance in the topological sense nor a measure in the measure theoretic sense, we chose to use the word I-divergence, as Csiszár (1962, 1964, 1967) does.

collection $V_{kl}(P)$ of all vicinities of P in the Kullback–Leibler I-divergence and compare this to $V_{vd}(P)$. If we then consider distributions P_y uniform on $[y, 1]$, we find that for all $y \in (0, 1)$ the Kullback–Leibler I-divergence $d_{kl}\left(P_y, P\right) = \infty$ while $d_{vd}\left(P_y, P\right) = 2y$. This implies that any vicinity of P in the variational distance contains probability measures P_y (just choose $y < \delta/2$ for a vicinity $V_{vd}(P, \delta) = \{Q \in \mathcal{P}[S, \mathcal{B}] \mid d_{vd}(Q, P) < \delta\}$) while none of these P_y is contained in any of the vicinities of P in the Kullback–Leibler I-divergence since $d_{kl}\left(P_y, P\right) = \infty$.

Bomze (1991) uses Reiss (1998) to show that $[d_{vd}(Q, P)]^2 \leq d_{kl}(Q, P)$ and therefore we know that $V_{kl}\left(P, \delta^2\right) \subset V_{vd}(P, \delta)$.

Summing up the relation between these three types of vicinities in an intuitive way, one can say that Q being close to P in the variational distance implies that Q is close to P in the Prohorov metric, but not the other way round, and that if Q is close to P in the Kullback-Leibler sense, Q will also be close to P in the variational distance, but not vice versa.

In this paper we follow Cressman and Hofbauer (2005) in sticking to the term local superiority and mention what types of vicinities are used. In the literature, two of those kinds of local superiorities have names of their own; Bomze (1990, 1991) calls a strategy that is locally superior in the variational distance *strongly uninvadable* and Oechssler and Riedel (2002) term a strategy that is locally superior in the Prohorov metric *evolutionary robust*. Local superiority in the Kullback-Leibler I-divergence is a new definition, but Bomze (1991) introduced vicinities of this type.

For the derivation of the replicator dynamics we can simply follow the standard case. A population at time t is characterized by a measure $R(t)$, where $R(t)(B)$ is the magnitude of the part of the population playing a strategy in $B \in \mathcal{B}$ at time t. By definition the expected payoff of a strategy in the game determines the expected growth of the part of the population that plays it, so we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(R(t)(B)\right) = \int_B \int_S A(x, y)\, P(t)(\mathrm{d}y)\, R(t)(\mathrm{d}x) \quad \forall B \in \mathcal{B},$$

where $P(t)(C) = \frac{R(t)(C)}{R(t)(S)} \quad \forall C \in \mathcal{B}$, which can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(P(t)(B)\right) = \int_B \int_S A(x, y)\, P(t)(\mathrm{d}y)\, P(t)(\mathrm{d}x) - P(t)(B) \int_S \int_S A(x, y)\, P(t)(\mathrm{d}y)\, P(t)(\mathrm{d}x) \quad \forall B \in \mathcal{B}.$$

Oechssler and Riedel (2001, Theorem 2) show that when the strong topology is used, boundedness of A guarantees existence and uniqueness of trajectories and positive invariance of $\mathcal{P}[S, \mathcal{B}]$ under the replicator dynamics. They also indicate (Remark 2, p. 145) that for the weak topology, A needs to be continuous to achieve this.

In the more general, non-bilinear case the replicator dynamics are defined as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(P\left(t\right)\left(B\right)\right) = \int_B F_{P(t)}\left(x\right) P\left(t\right)\left(\mathrm{d}x\right) - P\left(t\right)\left(B\right) \int_S F_{P(t)}\left(x\right) P\left(t\right)\left(\mathrm{d}x\right) \quad \forall B \in \mathcal{B}.$$

Conditions under which this is well defined can be found in Bomze (1991, Lemma 1). Please note that the space on which the last two differential equations are defined is the set $\mathcal{P}\left[S, \mathcal{B}\right]$, since $P(t)\left(S\right) = 1$. All points along a trajectory through a point $P\left(0\right)$ have the same atoms, but the limiting distribution may have a different set of atoms. More formally, if we define $P\left(\infty\right)$ as the weak limit point of $P\left(t\right)$, and $A_{P(t)} = \{x \in S \mid P\left(t\right)\left(\{x\}\right) > 0\}$ and $A_{P(\infty)} = \{x \in S \mid P\left(\infty\right)\left(\{x\}\right) > 0\}$ then neither $A_{P(0)} \subset A_{P(\infty)}$ nor $A_{P(\infty)} \subset A_{P(0)}$ necessarily holds.

The states that interest us most are the asymptotically stable states; states that guarantee a pull back to status quo after any small perturbation of the population state. Like local superiority, this concept is dependent on what we consider small perturbations, or in other words: which type of vicinities we use. Assuming that the whole set $\mathcal{P}\left[S, \mathcal{B}\right]$ is positively invariant under the dynamics, the following two types of stability are defined (see also Definition 6.5 of Weibull 1996; Definition 6 of Oechssler and Riedel 2001).

Definition 5 A state $P \in \mathcal{P}\left[S, \mathcal{B}\right]$ is Lyapounov stable with respect to d if every vicinity V with respect to d of P contains a vicinity V^0 of P such that $Q\left(t\right) \in V$ for all $Q\left(0\right) \in V^0$ and $t \geq 0$.

Definition 6 A state $P \in \mathcal{P}\left[S, \mathcal{B}\right]$ is asymptotically stable with respect to d if it is Lyapounov stable with respect to d and $\exists \epsilon > 0$ s.t. $d\left(Q\left(0\right), P\right) < \epsilon \Rightarrow d\left(Q\left(t\right), P\right) \to 0$ as $t \to \infty$.

3 From evolutionary to asymptotic stability and back

An interesting question is how the different concepts defined in Sect. 2 relate to each other. We know that for the discrete, finite case evolutionary stability (*ESS*), the existence of a uniform invasion barrier (*UIB*) and local superiority (*LS*) are equivalent and imply asymptotic stability (*AS*). For finite action spaces (Fig. 1), there is no need to specify a function d for local superiority and asymptotic stability. With a finite action space, any vicinity of a probability measure P with respect to any of the functions d mentioned above also contains a vicinity of P with respect to any other of the three. Hence, as long as the action space is finite, local superiority is equivalent across the three functions mentioned above, and the same naturally also holds for asymptotic stability.

Fig. 1 Finite action space

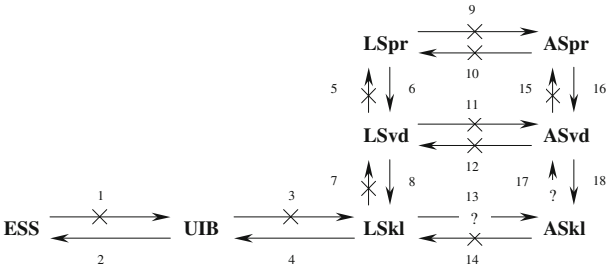

Fig. 2 Continuous action space

A continuous action space breaks up a few of those links as we will see just now. The examples that show which implications do not hold are of course negative results that suggest that local superiority might be the most useful of the three static concepts, but all is not lost for evolutionary stability; there are assumptions that pave the way from evolutionary stability to local superiority.

The implications under examination are drawn as arrows in Fig. 2. In this section we will check them one at a time, trying to find out which of them hold and which do not. For some arrows we will only mention known results or counterexamples, but since local superiority in the Kullback–Leibler I-divergence is a new definition, there are a few new arrows to consider. On top of that there are propositions that are concerned with restoring implications under extra assumptions.

We start with the relation between evolutionary stability and the existence of a uniform invasion barrier.

1 ESS $\nRightarrow$ UIB

Oechssler and Riedel (2001, p. 148) give an example of a payoff function for which the equilibrium strategy is *ESS* but not *UIB*, which in their terminology means that it is not uninvadable. Their example is a game $A(x, y) = -x^4 + 4xy$ and one can easily check that if δ_x is the Dirac measure in x, then δ_0 is an *ESS*, but that for $x \neq 0$ the maximum invasion barrier of δ_x is $\frac{x^2}{4}$, and this goes to 0 if x goes to 0.

One remarkable feature of their counterexample is that if we take the limiting distribution of the sequence of strategies for which the maximum invasion barrier goes to 0, it coincides with the *ESS* itself. One may wonder whether examples can be constructed where this is not the case, and it turns out that under very mild assumptions this is not possible, as the following proposition shows.

Proposition 7 *Assume that A is continuous and that S is compact. If P is an ESS and $\{Q_n\}_{n\in\mathbb{N}}$ is a sequence of strategies for which the maximum invasion barrier tends to 0, then Q_n converges weakly to P.*

Proof For showing that Q_n converges weakly to P, it will suffice to show that every subsequence has a further subsequence that weakly converges to P (see Billingsley 1968, Theorem 2.3).

Since S is compact, the sequence $\{Q_{n_k}\}_{k\in\mathbb{N}}$ is naturally tight. Let $\{Q_{n_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{Q_n\}_{n\in\mathbb{N}}$. By Prohorov's theorem (see Billingsley 1968, Theorem 6.1)

there exists a subsequence $\left\{Q_{n_{k_i}}\right\}_{i\in\mathbb{N}}$ and a probability measure Q such that $Q_{n_{k_i}}$ weakly converges to Q. We now show that $Q = P$.

Since P is an *ESS*, we know by definition that for every $Q_{n_{k_i}} \neq P$ either $u[P, P] > u\left[Q_{n_{k_i}}, P\right]$ or $u[P, P] = u\left[Q_{n_{k_i}}, P\right]$ and $u\left[P, Q_{n_{k_i}}\right] > u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right]$. In the latter case the invasion barrier is 1. This also holds if $u[P, P] > u\left[Q_{n_{k_i}}, P\right]$ and $u\left[P, Q_{n_{k_i}}\right] \geq u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right]$. The fact that $\overline{\epsilon}_P(Q_n)$ decreases to 0 therefore implies that there is an N such that $\overline{\epsilon}_P\left(Q_{n_{k_i}}\right) < 1$ for $i > N$, and therefore that from this N onwards $u[P, P] > u\left[Q_{n_{k_i}}, P\right]$ and $u\left[P, Q_{n_{k_i}}\right] < u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right]$. For any such $Q_{n_{k_i}}$, the maximum invasion barrier $\overline{\epsilon}_P\left(Q_{n_{k_i}}\right)$ follows from the equation

$$\frac{u[P, P] - u\left[Q_{n_{k_i}}, P\right]}{u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right] - u\left[P, Q_{n_{k_i}}\right]} = \frac{\overline{\epsilon}_P\left(Q_{n_{k_i}}\right)}{1 - \overline{\epsilon}_P\left(Q_{n_{k_i}}\right)}$$

Since $\overline{\epsilon}_P(Q_n) \to 0$, it also holds that

$$\lim_{i\to\infty} \frac{u[P, P] - u\left[Q_{n_{k_i}}, P\right]}{u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right] - u\left[P, Q_{n_{k_i}}\right]} = \lim_{i\to\infty} \frac{\overline{\epsilon}_P\left(Q_{n_{k_i}}\right)}{1 - \overline{\epsilon}_P\left(Q_{n_{k_i}}\right)} = 0.$$

Weak convergence of $Q_{n_{k_i}}$ implies that $u\left[Q_{n_{k_i}}, P\right] \to u[Q, P]$, $u\left[P, Q_{n_{k_i}}\right] \to u[P, Q]$ and $u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right] \to u[Q, Q]$ since A is continuous and S compact. Convergence of $u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right]$ follows from Theorem 3.2 in Billingsley (1968) and separability of S, where separability is implied by compactness, see Pervin (1964), page 105.

But then we have found that P is not an *ESS* if $Q \neq P$, since $u[P, Q] \leq u[Q, Q]$ and $u[P, P] = u[Q, P]$, which follows directly from the fact that $u\left[Q_{n_{k_i}}, Q_{n_{k_i}}\right] - u\left[P, Q_{n_{k_i}}\right]$ is bounded. From this contradiction, it follows that $Q = P$. □

2 UIB ⇒ ESS

The other direction is obvious; a strategy with a uniform invasion barrier is an *ESS*.

3 UIB ⇏ $\mathbf{LS}_{kl}$

The following game is not very elegant, but it does serve as an example of a game with a strategy that has a uniform invasion barrier without being locally superior in the

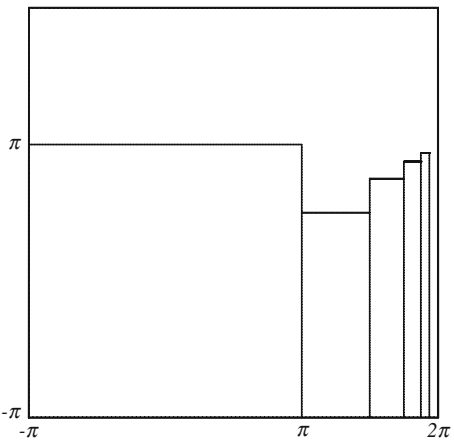

Fig. 3 Partition of the domain of $A(x, y)$

Kullback-Leibler I-divergence. Let S be the interval $[-\pi, 2\pi]$ and take the sequence $\{a_n\}_{n\in\mathbb{N}}$ with $a_n = \left(1 - \frac{1}{2^n}\right)\pi$. Define the payoff function of the game as

$$A = \begin{cases} \left|\sin\frac{1}{2}(x-y)\right| & \text{for } x \in [-\pi, \pi] \quad \text{and } y \in [-\pi, \pi] \\ \frac{3(\pi - a_n) + 2\cos\frac{1}{2}a_n}{(\pi + a_n)(a_n - a_{n-1})} & \text{for } x \in \left[\pi + a_{n-1}, \pi + a_n\right] \text{ and } y \in [-\pi, a_n] \\ 0 & \text{elsewhere} \end{cases}$$

where the different parts of the domain are pictured in Fig. 3.
Take as an equilibrium strategy P the uniform distribution on $[-\pi, \pi]$, in other words,

$$\frac{dP(x)}{d\lambda(x)} = \begin{cases} \frac{1}{2\pi} & \text{for } x \in [-\pi, \pi] \\ 0 & \text{elsewhere} \end{cases}$$

where λ is the Lebesque measure. Then we construct a sequence of signed measures $\{H_n\}_{n\in\mathbb{N}}$ with $H_n = H_n^+ - H_n^-$ which is defined by

$$\frac{dH_n^+(x)}{d\lambda(x)} = \begin{cases} \frac{1}{2\pi} & \text{for } x \in \left[\pi + a_{n-1}, \pi + a_n\right] \\ 0 & \text{elsewhere} \end{cases},$$
$$\frac{dH_n^-(x)}{d\lambda(x)} = \begin{cases} \frac{1}{2\pi} & \text{for } x \in [a_n, \pi] \\ 0 & \text{elsewhere} \end{cases}$$

This is a sequence of differences that we use to make a sequence $\{Q_n\}_{n\in\mathbb{N}}$ of strategies, where $Q_n = P + \frac{1}{2}H_n$. Note that

$$\begin{aligned} \lim_{n\to\infty} d_{kl}(Q_n, P) &= \lim_{n\to\infty} \int \log\left(\frac{\mathrm{d}P(x)}{\mathrm{d}Q_n(x)}\right) \mathrm{d}P(x) \\ &= \lim_{n\to\infty} \left(\int_{-\pi}^{a_n} \log 1 \mathrm{d}x + \int_{a_n}^{\pi} \log 2 \mathrm{d}x \right) = 0 \end{aligned}$$

If we make an obvious generalization of the function u to allow for the arguments to be measures that do not necessarily integrate to 1, we can take a shortcut in the algebra by computing the following integrals:

$$u\left[H_n^-, P - H_n^-\right] = \int_{-\pi}^{a_n}\int_{a_n}^{\pi} \left|\sin\frac{1}{2}(x-y)\right| \frac{1}{2\pi}\mathrm{d}x\frac{1}{2\pi}\mathrm{d}y = \frac{2}{\pi^2}\cos\frac{1}{2}a_n$$

$$u\left[H_n^-, H_n^-\right] = \int_{a_n}^{\pi}\int_{a_n}^{\pi} \left|\sin\frac{1}{2}(x-y)\right| \frac{1}{2\pi}\mathrm{d}x\frac{1}{2\pi}\mathrm{d}y = \frac{1}{\pi^2}\left((\pi - a_n) - 2\cos\frac{1}{2}a_n\right)$$

$$u\left[H_n^+, P\right] = u\left[H_n^+, P - H_n^-\right] = \int_{-\pi}^{a_n}\int_{\pi+a_{n-1}}^{\pi+a_n} \frac{3(\pi - a_n) + 2\cos\frac{1}{2}a_n}{(\pi + a_n)(a_n - a_{n-1})}\frac{1}{2\pi}\mathrm{d}x\frac{1}{2\pi}\mathrm{d}y$$

$$= \frac{1}{4\pi^2}\left(3(\pi - a_n) + 2\cos\frac{1}{2}a_n\right).$$

Note that furthermore, due to the parts of the payoff function that are zero, $u\left[H_n^+, H_n^-\right] = u\left[H_n^-, H_n^+\right] = u\left[H_n^+, H_n^+\right] = 0$. Now we can write

$$\begin{aligned} u[Q_n, Q_n] - u[P, Q_n] &= u\left[P + \frac{1}{2}H_n, P + \frac{1}{2}H_n\right] - u\left[P, P + \frac{1}{2}H_n\right] \\ &= u[P, P] + \frac{1}{2}\{u[H_n, P] + u[P, H_n]\} \\ &\quad + \frac{1}{4}u[H_n, H_n] - u[P, P] - \frac{1}{2}u[P, H_n] \\ &= \frac{1}{2}u[H_n, P] + \frac{1}{4}u[H_n, H_n] \\ &= \frac{1}{2}\left\{u\left[H_n^+, P\right] - u\left[H_n^-, P\right]\right\} + \frac{1}{4}\left\{u\left[H_n^+, H_n^+\right]\right. \\ &\quad \left. - u\left[H_n^+, H_n^-\right] - u\left[H_n^-, H_n^+\right] + u\left[H_n^-, H_n^-\right]\right\} \end{aligned}$$

Filling in the integrals found above, and using $u[H_n^-, P] = u[H_n^-, P - H_n^-] + u[H_n^-, H_n^-]$, we get

$$\begin{aligned} &\frac{1}{2}\left\{\frac{1}{4\pi^2}\left(3(\pi - a_n) + 2\cos\frac{1}{2}a_n\right) - \frac{1}{4\pi^2}(4(\pi - a))\right\} \\ &\quad + \frac{1}{4}\left\{\frac{1}{\pi^2}\left((\pi - a) - 2\cos\frac{1}{2}a_n\right)\right\} \\ &\quad = \frac{1}{8\pi^2}\left((\pi - a_n) - 2\cos\frac{1}{2}a_n\right) > 0 \end{aligned}$$

which implies that the strategy P is not locally superior in the Kullback–Leibler I-divergence.

To show that the strategy P has a uniform invasion barrier, it will do to look at the sequence $\left\{\widetilde{Q}_n\right\}_{n\in\mathbb{N}}$ with $\widetilde{Q}_n = P + H_n$. For this sequence we can write

$$\begin{aligned}\frac{\epsilon_{\widetilde{Q}_n}}{1-\epsilon_{\widetilde{Q}_n}} &= \frac{u\left[P,\,P\right]-u[\widetilde{Q}_n,\,P]}{u\left[\widetilde{Q}_n,\,\widetilde{Q}_n\right]-u\left[P,\,\widetilde{Q}_n\right]} = \frac{-u\left[H_n,\,P\right]}{u\left[H_n,\,P+H_n\right]} \\ &= \frac{-\left(\left(3\left(\pi-a_n\right)+2\cos\frac{1}{2}a_n\right)-4\left(\pi-a_n\right)\right)}{\left(3\left(\pi-a_n\right)+2\cos\frac{1}{2}a_n\right)-8\cos\frac{1}{2}a_n} = \frac{\left(\pi-a_n\right)-2\cos\frac{1}{2}a_n}{3\left(\pi-a_n\right)-6\cos\frac{1}{2}a_n} = \frac{1}{3}\end{aligned}$$

From this equation, it follows that $\epsilon_P\left(\widetilde{Q}_n\right) = \frac{1}{4}$, which is an invasion barrier for all strategies $\widetilde{Q}_n$. Since the payoff function is zero on $[\pi, 2\pi] \times [-\pi, 2\pi]$, as well as on $\left[a_{n-1}, a_n\right] \times [a_n, \pi]$ for all $n > 0$, no mutant can perform better against P than these $\widetilde{Q}_n$'s, which implies that $\frac{1}{4}$ is a uniform invasion barrier.[2]

4 $\mathbf{LS}_{kl} \Rightarrow$ UIB

This implication is relatively easy and basically not very different from the finite case:

Proposition 8 *A strategy P which is locally superior in the Kullback-Leibler I -divergence has a uniform invasion barrier.*

Proof First observe that

$$\begin{aligned}d_{kl}\left((1-\epsilon)\,P+\epsilon Q,\,P\right) &= \int\limits_{\mathrm{supp}(P)} \log\left(\frac{\mathrm{d}P\left(x\right)}{\mathrm{d}\left[(1-\epsilon)\,P+\epsilon Q\right](x)}\right)\mathrm{d}P\left(x\right) \\ &< \int\limits_{\mathrm{supp}(P)} \log\left(\frac{\mathrm{d}P\left(x\right)}{\mathrm{d}\left[(1-\epsilon)\,P\right](x)}\right)\mathrm{d}P\left(x\right) \\ &= \int\limits_{\mathrm{supp}(P)} \log\frac{1}{(1-\epsilon)}\mathrm{d}P\left(x\right) = \log\frac{1}{(1-\epsilon)}\end{aligned}$$

P being locally superior in the Kullback-Leibler I-divergence means that there is a vicinity $V_{kl}\left(P,\delta\right)$ of P such that for all $R \in V_{kl}\left(P,\delta\right)$

$$u\left[P,\,R\right] > u\left[R,\,R\right]$$

Now take $\overline{\epsilon} = 1 - e^{-\delta}$. Using our first observation, we find that for all $\epsilon < \overline{\epsilon}$ and all $Q \in \mathcal{P}\left[S,\mathcal{B}\right]$ the following holds: $d_{kl}\left((1-\epsilon)\,P+\epsilon Q,\,P\right) < \delta$. The local superiority

[2] In game A we could replace $3\left(\pi-a_n\right)+2\cos\frac{1}{2}a_n$ by anything between $4\left(\pi-a_n\right)$ and $8\cos\frac{1}{2}a_n$. We would only have to adapt the Q_n 's accordingly; if we take $(1-\beta)\cdot 4\left(\pi-a_n\right)+\beta\cdot 8\cos\frac{1}{2}a_n$, $\beta \in (0,1)$, then we have to choose an α that exceeds β in $Q_n = P + \alpha H_n$.

then tells us that

$$\begin{aligned} & u\left[P,(1-\epsilon)\,P+\epsilon Q\right]-u\left[(1-\epsilon)\,P+\epsilon Q,(1-\epsilon)\,P+\epsilon Q\right]>0 \\ \Leftrightarrow\ & u\left[\epsilon P,(1-\epsilon)\,P+\epsilon Q\right]-u\left[\epsilon Q,(1-\epsilon)\,P+\epsilon Q\right]>0 \\ \Leftrightarrow\ & u\left[P,(1-\epsilon)\,P+\epsilon Q\right]-u\left[Q,(1-\epsilon)\,P+\epsilon Q\right]>0 \end{aligned}$$

But then we have shown that $\overline{\epsilon}$ is a uniform invasion barrier. □

There are different kinds of assumptions we can make under which equivalence of the existence of a uniform invasion barrier and local superiority can be restored. One is to restrict the support of the candidate for local superiority to a finite set. This is done in Oechssler and Riedel (2001, Lemma 1, part 2), where they show the following. (Note that they refer to a strategy as uninvadable when we say it has a uniform invasion barrier and that their strong uninvadability equals what we call local superiority in the variational distance).

Proposition 9 *Assume that the support of P is finite. Then P is locally superior in the variational distance if P has a uniform invasion barrier.*

This proposition actually does a bit more than repair only the third arrow from Fig. 2; it gives a condition under which both non-implications **3** and **7** are turned into implications in one go.

There is however also a quite different restriction on the support of P under which equivalence can be restored too. Under the assumption that the support equals the whole action space S, we can in fact go from evolutionary stability directly to local superiority in the Prohorov metric (**1**, **3**, **5** and **7** in one step).

Proposition 10 *Assume that A is continuous and that the support of P equals the action space S. Then P is locally superior in the Prohorov metric if P is an ESS.*

Proof This follows directly from the second definition of evolutionary stability, which states that P is evolutionary stable if both $u\left[P,P\right]\geq u\left[Q,P\right]\ \forall\ Q$ and $u\left[P,P\right]=u\left[Q,P\right]\Rightarrow u\left[P,Q\right]>u\left[Q,Q\right]\ \forall Q\neq P$ hold. If the support of P equals S then, by Lemma 11, $u\left[P,P\right]$ must equal $u\left[Q,P\right]$ and then the second part of the definition tells us that P must even be globally superior, and therefore that P will also be locally superior in the Prohorov metric. □

Here the following lemma, however intuitively clear, is used:

Lemma 11 *If (P,P) is a Nash equilibrium of A, $supp\,(Q)\subset supp\,(P)$ and A is continuous, then $u\left[Q,P\right]=u\left[P,P\right]$.*

Proof (P,P) being a Nash equilibrium of A—that we have assumed to be a symmetric game throughout the paper—means that $u\left[P,P\right]\geq u\left[Q,P\right]$ for all Q. As a consequence, $u\left[P,P\right]\geq u\left[\delta_x,P\right]=\int A\left(x,y\right)P\left(\mathrm{d}y\right)$, where δ_x is the Dirac measure in x. If we define $f\left(x\right)$ as the difference, then $f\left(x\right)=u\left[P,P\right]-u\left[\delta_x,P\right]\geq 0$. But also $\int_S f\left(x\right)P\left(\mathrm{d}x\right)=\int_S\left(u\left[P,P\right]-u\left[\delta_x,P\right]\right)P\left(\mathrm{d}x\right)=0$, so that we can conclude that $\left[P\left\{x:f\left(x\right)>0\right\}=0\right.$ and $\left.\right]P\left\{x:f\left(x\right)=0\right\}=1$. The set $\left\{x:f\left(x\right)=0\right\}$ is closed

because $f(x)$ is continuous. Together with the assumed inclusion, this implies that $\text{supp}(Q) \subset \text{supp}(P) \subset \{x : f(x) = 0\}$.

We now have that $0 \leq \int_S f(x)\,\mathrm{d}Q = \int_{\text{supp}(Q)} f(x)\,\mathrm{d}Q \leq \int_{\{x:f(x)=0\}} f(x)\,\mathrm{d}Q = 0$, which reads $u[P, P] = u[Q, P]$. □

The assertion of the lemma remains valid if the two assumptions that $\text{supp}(Q) \subset \text{supp}(P)$ and that A is continuous are replaced with the single assumption that Q is absolutely continuous w.r.t. P. Inspection of the proof of the lemma shows that the only place where the original two assumptions are used, is where we obtain the conclusion that $Q\{x : f(x) = 0\} = 1$ from $P\{x : f(x) = 0\} = 1$. This is obviously also guaranteed under the alternative assumption.

We conjecture that the following theorem also holds, but we were not able to prove or disprove it:

Conjecture 12 *If $P(\text{int}(\text{supp}(P))) = 1$, A is continuous and S is compact, then P has a UIB if P is an ESS.*

Because Proposition 7 states that $\overline{\epsilon}_P(Q_n) \to 0$ implies that Q_n converges weakly to P, it would be enough to show that, coversely, if $P(\text{int}(\text{supp}(P))) = 1$, and Q_n converges weakly to P, then $\overline{\epsilon}_P(Q_n) = -\frac{u[Q_n - P, P]}{u[Q_n - P, Q_n - P]}$ cannot go to 0. This contradiction would prove the conjecture.

Now we will turn to how the three types of local superiority relate to each other.

5 $\mathbf{LS}_{vd} \nRightarrow \mathbf{LS}_{pr}$

Oechssler and Riedel (2002, Example 4) show that a strategy with a *UIB* need not be locally superior in the Prohorov metric. Here we focus only on the step from local superiority in the variational distance and local superiority in the Prohorov metric, but luckily δ_0 (the Dirac measure in 0) in their counterexample—$A(x, y) = -x^2 + xy$—is also locally superior in the variational distance.

6 $\mathbf{LS}_{pr} \Rightarrow \mathbf{LS}_{vd}$

Since $V_{vd}(P, 2\delta) \subset V_{pr}(P, \delta)$, a vicinity in the Prohorov metric contains one in the variational distance, so it is obvious that local superiority in the Prohorov metric implies local superiority in the variational distance.

7 $\mathbf{LS}_{kl} \nRightarrow \mathbf{LS}_{vd}$

For a counterexample we go back to the one we used at arrow **3** and modify it slightly. We change the game into

$$A = \begin{cases} \left|\sin\frac{1}{2}(x-y)\right| & x \in [-\pi, \pi], \quad y \in [-\pi, \pi] \\ \frac{8\cos\frac{1}{2}a_n}{(\pi+a_n)(a_n-a_{n-1})} & x \in \left[\pi + a_{n-1}, \pi + a_n\right], \; y \in [-\pi, \; a_n] \\ 0 & \text{elsewhere} \end{cases}$$

If we look at the same equilibrium strategy P, that is P uniform on $[-\pi, \pi]$, and at mutant strategies $Q_n(\alpha) = P + \alpha H_n$, with the H_n's as before, we see that

$$\begin{aligned} & u[Q_n(\alpha), Q_n(\alpha)] - u[P, Q_n(\alpha)] \\ & = \alpha u[H, P] + \alpha^2 u[H, H] \\ & = \alpha \left(u\left[H^+, P - H^-\right] + u\left[H^-, P\right]\right) + \alpha^2 \left(u\left[H^-, H^-\right]\right) \\ & = \alpha \left(\frac{2}{\pi^2} \cos \frac{1}{2} a_n - \frac{1}{\pi^2} (\pi - a_n) \right) + \alpha^2 \left(\frac{1}{\pi^2} (\pi - a_n) - \frac{2}{\pi^2} \cos \frac{1}{2} a_n \right) \end{aligned}$$

and this is smaller than 0 if α is smaller than 1. But for $\alpha = 1$, which makes the only sequence $\{Q_n(\alpha)\}_{\pi \in \mathbb{N}}$ that goes to 0 in the variational distance but not in the Kullback-Leibler I-divergence, this expression equals 0. Therefore P is not locally superior in the variational distance, although it is locally superior in the Kullback-Leibler I-divergence.

8 $\mathbf{LS}_{vd} \Rightarrow \mathbf{LS}_{kl}$

Similar to **6**, $V_{kl}\left(P, \delta^2\right) \subset V_{vd}(P, \delta)$ implies that a strategy that is local superior in the variational distance is also local superior in the Kullback-Leibler I-divergence.

This completes the implications within the set of static definitions. When we go on to characteristics of strategies as population states in the dynamics, it is useful to note a few things. The first is that there are again different types of asymptotic stability, depending on which type of vicinity we choose. This time, however, not all of them are equally useful without further restrictions or assumptions. The reason for that is that population states that are close by in the variational distance or in the Prohorov metric may put probability zero on a set that has positive probability in the equilibrium. If a shock takes the population to such a state, it is obvious that selective dynamics cannot bring the population back. This observation is made in Oechssler and Riedel (2002) and in Cressman (2005) and Propositions 13 and 14 make that precise. Although they are relatively obvious observations, they will turn out to be useful for proving some implications later on.

Proposition 13 *If a strategy P is asymptotically stable in the variational distance, then the support is finite.*

Proof Suppose the support of P is not finite. Then, by Lemma 15 below, for any $\epsilon > 0$ one can find a set $C \in \mathcal{B}$ such that $0 < P(C) \leq \frac{\epsilon}{2}$. Define Q as $Q(B) = \frac{1}{1-P(C)} P(B \backslash C) \ \forall \ B \in \mathcal{B}$. For this Q we have that $d_{vd}(Q, P) = 2P(C) \leq \epsilon$ and $d_{vd}(Q(t), P) \geq d_{vd}(Q, P) > 0$ for all $t \geq 0$ if we take $Q(0) = Q$. The latter follows from $Q(0)(C) = 0$, which implies that $Q(t)(C) = 0$ for all $t \geq 0$. Strategy P therefore is not asymptotically stable in the variational distance. □

Proposition 14 *If a strategy P is asymptotically stable in the Prohorov distance, then the support of P is* (1) *finite and* (2) *consists of isolated points only.*

Proof Suppose the support of P is not finite. Then with Lemma 15 we know that for any $\epsilon > 0$ one can find an open set C such that $0 < P(C) < \epsilon$. Define Q as $Q(B) =$

$\frac{1}{1-P(C)} P(B \backslash C) \ \forall B \in \mathcal{B}$. For this Q we have that $d_{pr}(Q, P) \leq \frac{1}{2} d_{vd}(Q, P) = P(C) < \epsilon$. For the open set C, we now have that $Q(t)(C) = 0$ for all $t \geq 0$ if we take $Q(0) = Q$. Hence $\liminf Q(t)(C) = 0 < P(C)$, which by the Portmanteau theorem contradicts that $Q(t)$ converges weakly to P. Strategy P therefore is not asymptotically stable in the Prohorov metric.

If the support of P is finite, but does not consist of isolated points only, then there is a point $x \in S$ for which $P(x) > 0$ and $\rho(x, S \backslash \{x\}) = 0$. Then for any $0 < \epsilon < \min\{P(x), \rho(x, \text{supp}(P) \backslash \{x\})\}$, we can define a strategy Q_ϵ such that the only difference between P and Q_ϵ is that Q_ϵ has probability mass 0 in x and probability mass $P(x)$ in a point y with $0 < \rho(x, y) < \epsilon$. For this Q_ϵ we have that $0 < d_{pr}(Q_\epsilon, P) < \epsilon$ and $d_{pr}(Q(t), P) \geq d_{pr}(Q_\epsilon, P)$ for all $t \geq 0$ if we take $Q(0) = Q_\epsilon$. Strategy P therefore is not asymptotically stable in the Prohorov metric. □

Lemma 15 *The support of P is a finite set if and only if $\exists\ \varepsilon > 0$ such that for all $F \in \mathcal{B}$ either $P(F) = 0$ or $P(F) \geq \varepsilon$.*

Proof ($\Rightarrow$) Let $F_0 = \{x_1, \ldots, x_n\}$ with $P(F_0) = 1$ and $P(\{x_i\}) > 0$ for all i. Take $\varepsilon = \min\{P(\{x_1\}), \ldots, P(\{x_n\})\} > 0$. Note that we do not need the separability of S here.

($\Leftarrow$) Let A_P be the set of atoms of P. In general A_P is at most countable, but it follows from the assumption that A_P is even a finite set. If $P(A_P) = 1$, then that proves the assertion. If $0 \leq P(A_P) < 1$ we define Q as $Q(B) = \frac{1}{1-P(A_P)} P(B \backslash A_P) \ \forall B \in \mathcal{B}$. This way P is renormalized on the reduced space $S \backslash A_P$ by taking the conditional probability measure on this set, to again obtain a probability measure, but now without possible atoms. By separability of S we can find an $S_0 = \{x_1, x_2, \ldots\}$ that is a countable dense subset of S. Take such an x_n and consider a sequence of open balls $B_n^k, k = 1, 2, \ldots$ with centers x_n that decreases to $\{x_n\}$. Then, by continuity of the probability measure, we have that $P\left(B_n^k\right) \downarrow 0$ for $k \to \infty$. By the assumption there then exists an index k_n such that $P\left(B_n^{k_n}\right) = 0$. Since S_0 is dense, it holds that $S = \bigcup_{k=1}^{\infty} B_n^{k_n}$ and hence it would follow that $Q(S) = 0$. But then it cannot be that $P(A_P) < 1$ and it must be that $P(A_P) = 1$ after all. □

In general, two ways to get more positive results have been tried successfully. The first one is to focus on strategies with finite support in combination with the variational distance. This is done in for instance Oechssler and Riedel (2001) and Cressman (2005). The second way to get positive results is to make additional assumptions concerning starting points of trajectories, as is done in for instance Bomze (1990, 1991), Oechssler and Riedel (2002) and Cressman (2005). In Bomze (1990, Theorem 2, 1991, Theorem 4) it is assumed that starting points are in a vicinity that is defined with the Kullback-Leibler I-divergence. In Oechssler and Riedel (2002, Conjecture 1) and Cressman (2005, Definition 1), it is assumed that starting points have a support that includes the support of the candidate for stability. In this paper, we would like to let these two different assumptions enter in a slightly different way.

The additional assumption in Bomze (1990, 1991) implies the introduction of the Kullback-Leibler I-divergence as a function d that defines vicinities of probability distributions. When this is used to (further) restrict the possible starting points on top

of—or instead of—demanding that starting points be close in for instance the variational distance, then we think it is natural to treat this restriction no different from the restrictions we impose by using the Prohorov metric or the variational distance. Because we think of them as functions that reflect how the likelihood of different mutations or shocks compare to each other, we would prefer to give the Kullback-Leibler I-divergence a similar interpretation. This restriction is therefore naturally incorporated in the definition of asymptotic stability in the Kullback-Leibler I-divergence.[3]

Restricting the starting points of trajectories to those that have a support that includes the support of the candidate for stability we would like to treat slightly differently. This restriction is made with the correct argument that if we allow for starting points to violate this assumption, then there are interesting games in which no equilibrium strategy can ever be asymptotically stable in the variational distance nor in the Prohorov metric. Another way of formulating this is that for those games no positive results can be proven. Propositions 13 and 14 can be seen as formalizing that argument. The restriction itself, however, we think is not naturally thought of as excluding relatively unlikely mutations and including the more likely ones. We therefore would like to formally separate the reasons why starting points of trajectories should not have a support that is a strict subset of the support of the candidate for stability, from the reasons why other starting points are excluded, which is that it takes a relatively unlikely shock to shift the population to them.

Summarizing, we have chosen to allow for the possibility to let the restriction to vicinities in the Kullback-Leibler I-divergence enter the definition of asymptotic stability through the function d, while we have left the second restriction out of the definition of asymptotic stability itself. Leaving it out does not at all imply that this restriction can or should not be used in formulating useful positive stability results. It only reflects that we do not see it as a representation of relative likelihoods of shocks or mutations. A final thing worth mentioning here is that the use of asymptotic stability in the Kullback-Leibler I-divergence makes excluding starting points of trajectories on grounds of their support superfluous, because $d_{\mathrm{KL}}(Q, P) = \infty$ for any Q for which $\mathrm{supp}(P)$ is not included in $\mathrm{supp}(Q)$. Given that we emphatically want to allow for equilibrium distributions that have densities (see for instance the examples at arrows **3** and **7**), for which Propositions 13 and 14 exclude asymptotic stability in the Prohorov metric and the variational distance, we chose to place a bit more emphasis on asymptotic stability in the Kullback-Leibler I-divergence than on restrictions concerning the support, which in that case becomes redundant.

9 $\mathbf{LS}_{pr} \not\Rightarrow \mathbf{AS}_{pr}$

Take $S = [0, 1]$ and

$$A(x, y) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x > 0 \end{cases}$$

[3] Van Veelen (2001) contains a simple finger exercise with mutations that explores the possibilities of such an interpretation, which gives vicinities that are somewhere in between the ones that come with Kullback-Leibler I -divergence and the variational distance.

If we take probability measure P with all mass at $x = 0$, then $u[P, Q] = P(0) = 1 > Q(0) = u[Q, Q]$ for all $Q \neq P$. This P is therefore locally superior in the Prohorov metric, but it is obviously not asymptotically stable in the Prohorov metric (see Proposition 14 and realise that $x = 0$ is not an isolated point in S).

10 $\mathbf{AS}_{pr} \nRightarrow \mathbf{LS}_{pr}$

12 $\mathbf{AS}_{vd} \nRightarrow \mathbf{LS}_{vd}$

14 $\mathbf{AS}_{kl} \nRightarrow \mathbf{LS}_{kl}$

The counterexample for the finite case still serves as an example that these implications do not hold. For this distorted Rock Scissors Paper game, I refer to Weibull (1996, p. 102). The only important thing is that we have to choose the action space S properly. If we locate the three actions at 1, 2 and 3, respectively, one has to make sure that these are isolated points in the action space if we use the Prohorov metric—take for instance $S = \{1, 2, 3\}$—because otherwise the equilibrium would not be asymptotically stable. For the other two cases, S can be $\mathbb{R}$ if we define $A(x, y) = 0$ if $x \notin \{1, 2, 3\}$ or $y \notin \{1, 2, 3\}$.

11 $\mathbf{LS}_{vd} \nRightarrow \mathbf{AS}_{vd}$

Take as an example $S = [-\pi, \pi]$ and $A(x, y) = \left|\sin\left(\frac{1}{2}(x - y)\right)\right|$. The strategy P with uniform density $f(x) = \frac{1}{2\pi}$ is globally superior, but any strategy Q_δ with density

$$g_\delta(x) = \begin{cases} 0 & x \in [-\pi, -\pi + \delta) \\ \frac{1}{2\pi - \delta} & x \in [-\pi + \delta, \pi] \end{cases}$$

is within variational distance ϵ if $\delta < \epsilon/2$, and yet $d_{vd}(Q(t), P) \geq 2\delta$ for all $t \geq 0$ if we take $Q(0) = Q_\delta$.

13 $\mathbf{LS}_{kl} \Rightarrow ?\mathbf{AS}_{kl}$

From Theorem 4 in Bomze (1991) and Theorem 2 in Bomze (1990) we know that if P is local superior in the variational distance and we start in an appropriately small Kullback-Leibler vicinity of P, then (1) the Kullback-Leibler I-divergence decreases monotonically and (2) the variational distance goes to zero. Because the proofs only use the local superiority in the Kullback-Leibler vicinity, this could be summarized as follows:

Proposition 16 *Assume that replicator dynamics are well-defined. If a strategy P is locally superior in the Kullback-Leibler I-divergence, then for $d_{kl}(Q(0), P)$ small enough, $d_{kl}(Q(t), P)$ decreases monotonically and $d_{vd}(Q(t), P) \to 0$ as t increases.*

In order to arrive at asymptotic stability at arrow **13**, it remains to be proven that also $d_{kl}\left(Q\left(t\right), P\right)$ not only decreases monotonically, but actually goes to 0.

15 $\mathbf{AS}_{vd} \nRightarrow \mathbf{AS}_{pr}$

The counterexample at arrow **9** also serves as a counterexample here; if we take $S = [0, 1]$ and

$$A\left(x, y\right) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x > 0 \end{cases}$$

then probability measure P with all mass at $x = 0$ is asymptotically stable in the variational distance, but not in the Prohorov metric.

16 $\mathbf{AS}_{pr} \Rightarrow \mathbf{AS}_{vd}$

From Proposition 14 we know that if P is asymptotically stable in the Prohorov metric, then its support must consist of isolated points only. In that case the two asymptotic stabilities coincide.

17 $\mathbf{AS}_{kl} \nRightarrow ?\mathbf{AS}_{vd}$

Any strategy that is asymptotically stable in the Kullback-Leibler I-divergence and that has a density would do as a counterexample (see Proposition 13). A proof of the implication at arrow **13** would certainly help establishing this, for then we could take the example from arrow **11**.

18 $\mathbf{AS}_{vd} \Rightarrow \mathbf{AS}_{kl}$

From Propositions 13 we know that if P is asymptotically stable in the variational distance, then its support must consist of a finite number of points. In that case the two asymptotic stabilities coincide.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

4 Appendix

The appendix is meant to serve the reader that wants to check up on the reason why the Kullback-Leibler type vicinities do not come with a straightforward topology, as the Prohorov metric or the variational distance do.

The definition of a Kullback-Leibler vicinity is due to Bomze (1991), and on page 81 of the article, he refers to Csiszár (1967) for the argument why they, as it is stated there, do not define a topology. The relevant theorem from Csiszár (1967) is Theorem 3, which proves this point for a general set of divergences, of which the Kullback-Leibler

I-divergence is a special case. In order to indicate that there is a problem concerning the interpretation of this theorem, we quote the remark that directly follows the theorem in Csiszár (1967). The general set of f-divergences we encounter there is defined with functions $f(u)$, and for $f(u) = u \log u$ we have our case, which is, in terms of the theorem, the I-divergence of order 1.

Remark The corresponding theorem for I-divergences of order $\alpha > 0$ has been proved in Csiszár (1962); the idea of the proof is similar also in the general case, but the lack of knowledge of the concrete form of $f(u)$ causes some difficulties.

In order that the Fréchet (V)-space E be a topological space (in the sense that the given neighbourhood systems of the points of E are basis for the neighbourhood systems of the points of E in some topology on E) the following property is clearly necessary:

(B) For all points $e \in E$, each neighbourhood U of e contains another neighbourhood U' of e such that each point $e' \in U'$ possesses a neighbourhood which is a subset of U.

Note to Remark: This property is well known to be sufficient, too, provided that each point of E possesses at least one neighbourhood and is contained in all of its neighbourhoods.

The idea this remark conveys is to some extent clear, but a few things are strange if we just use standard topology textbook definitions. One is that for sufficiency, it is demanded that each point is contained in all of its neighbourhoods. This is a rather odd requirement, for points by definition are contained in their neighbourhoods (see Definition 18). It is also remarkable that in (B) there apparently are neighbourhoods, while the purpose of this property is to prove that something is not a topological space. This is strange, because in standard topology terminology, neighbourhoods can only be defined once there is a topology (again, see Definition 18).

The solution to this problem lies in the topology textbook that Csiszár (1967) refers to, which is General Topology by Sierpinski (1952). There we find that the order of defining neighbourhoods and topologies is the reverse of the order we are used to now. The starting point there is a system of neighbourhoods called a Fréchet (V)-space. These neighbourhoods may be different from our notion of a neighbourhood; a neighbourhood of a point e for instance need not contain e itself. A topological space is then defined as a Fréchet (V)-space that satisfies four conditions. Modern textbook definitions, as they are given in Definitions 17 and 18, go in the opposite direction; first a topological space is defined, and once we have it, we can check whether or not a set U is a neighbourhood of a point e. It is also worth noting that Sierpinski's set of four conditions includes what we now call the first separation axiom, which is not included in the standard definition of a topological space.

All of this implies that we can run into trouble if we read the proof of Theorem 3 in Csiszár (1967), or the proof of the special case in Csiszár (1962), and assume that the terminology there refers to modern textbook definitions. If we do, then there are some parts of the proofs that would need no proving and others that would just not be true. The ingredients however can without much complications be rearranged so that they make a correct proof when using modern textbook definitions. Therefore we will start

with some important observations and then indicate how the sequences constructed in Csiszár (1962) can be used to prove the point.

4.1 Definitions

We begin with modern standard definitions of a topological space, a neighbourhood, a base and a subbase, as well as a well-known theorem.

Definition 17 A topological space is a pair $(X, \mathcal{T})$, with X a set and $\mathcal{T}$ a topology, that is, a set of subsets of X that satisfies the following axioms:

1. $\emptyset, X \in \mathcal{T}$;
2. If $O_i \in \mathcal{T}\ \forall\, i \in I$, then $\bigcup_{i\in I} O_i \in \mathcal{T}$;
3. If $O_1, ..., O_n \in \mathcal{T}$, then $\bigcap_{i=1}^{n} O_i \in \mathcal{T}$.

Definition 18 U is a neighbourhood of $P \in X$ if there is a $O \in \mathcal{T}$ such that $P \in O \subset U$.

Definition 19 Let $(X, \mathcal{T})$ be a topological space. A collection $\mathcal{B}$ is a base for $\mathcal{T}$ if

1. $\mathcal{B} \subset \mathcal{T}$,
2. for every $O \in \mathcal{T}$ there is a $\mathcal{B}' \subset \mathcal{B}$ such that $O = \bigcup \{B : B \in \mathcal{B}'\}$.

Definition 20 Let $(X, \mathcal{T})$ be a topological space. A collection $\mathcal{B}$ is a subbase for $\mathcal{T}$ if the set of finite intersections from $\mathcal{B}$ is a base for $\mathcal{T}$.

Theorem 21 *Let X be a set and $\mathcal{B}$ a collection of subsets of X. Then there is a unique topology $\mathcal{T}$ on X such that $\mathcal{B}$ is a base for $\mathcal{T}$ if and only if*

1. $\bigcup \{B : B \in \mathcal{B}\} = X$
2. *for all $B_1, B_2 \in \mathcal{B}$ and all $x \in B_1 \cap B_2$ there exists a $B \in \mathcal{B}$ such that $x \in B \subset B_1 \cap B_2$*

4.2 Vicinities

For a probability measure P and a function d, one can define the set of vicinities of P as follows:

$$V_d(P) = \{V_d(P, \delta)\}_{\delta>0}$$

where

$$V_d(P, \delta) = \{Q \in \mathcal{P}[S, \mathcal{B}] \mid d(Q, P) < \delta\}$$

We will look at the union of such sets over all probability measures in $\mathcal{P}[S, \mathcal{B}]$

$$F^d = \bigcup_{P\in\mathcal{P}[S,\mathcal{B}]} V_d(P)$$

A collection of this type is sometimes called a Fréchet (V) -space.

With the three functions d from Sect. 2, we have now defined F^{pr}, F^{vd} and F^{kl}. Each of these collections is a subbase for a topology on $\mathcal{P}[S, \mathcal{B}]$, and the topologies generated by the subbases F^{pr}, F^{vd} and F^{kl} are denoted by $\mathcal{T}(F^{pr})$, $\mathcal{T}\left(F^{vd}\right)$ and $\mathcal{T}\left(F^{kl}\right)$ respectively. Because the Prohorov metric and the variational distance are actual distances, we know—using Theorem 21 and the triangle inequality—that F^{pr} and F^{vd} are also bases for $\mathcal{T}(F^{pr})$ and $\mathcal{T}\left(F^{vd}\right)$, respectively. The Kullback-Leibler I-divergence however is not a distance, and F^{kl} is not a base for $\mathcal{T}\left(F^{kl}\right)$, as we will see below.

We will do two things now. First we will show that $\mathcal{T}(F^{pr}) \subset \mathcal{T}\left(F^{vd}\right) \subset \mathcal{T}\left(F^{kl}\right)$ and that these inclusions are strict. Then we will finally show how the sequences of distributions that are constructed by Csiszár (1962) imply that condition 2 in Theorem 21 is not met, and hence F^{kl} is not a base for a topology.

4.3 Inclusions

Because F^{pr} is a base for $\mathcal{T}(F^{pr})$, we know that if $O \in \mathcal{T}(F^{pr})$ and $P \in O$, this implies that there is a $B \in F^{pr}$ such that $P \in B \subset O$. In other words, there is a Q and a $\delta > 0$ such that $P \in V_{pr}(Q, \delta) \subset O$. Because the Prohorov metric is a distance, we can use the triangle inequality to show that there also is an $\epsilon > 0$ such that $P \in V_{pr}(P, \epsilon) \subset V_{pr}(Q, \delta) \subset O$. But then, since we know that $V_{vd}(P, 2\epsilon) \subset V_{pr}(P, \epsilon)$, we also know that $P \in V_{vd}(P, 2\epsilon) \subset V_{pr}(P, \epsilon) \subset O$. This can be done for any $P \in O$, and therefore also $O \in \mathcal{T}\left(F^{vd}\right)$. Summarizing, we have found that $\mathcal{T}(F^{pr}) \subset \mathcal{T}\left(F^{vd}\right)$.

A similar argument can be employed to show that $\mathcal{T}\left(F^{vd}\right) \subset \mathcal{T}\left(F^{kl}\right)$; we then use that the variational distance is a distance too and that $V_{kl}\left(P, \delta^2\right) \subset V_{vd}(P, \delta)$.

The converse inclusions are typically not true with continuous action spaces S. To illustrate that $\mathcal{T}(F^{pr})$ does not contain $\mathcal{T}\left(F^{vd}\right)$, we take $S = [0, 1]$ and $P = \delta_0$ the point measure at 0. For all $y \in (0, 1]$ we find that $d_{vd}\left(\delta_y, P\right) = 2$, while $d_{pr}\left(\delta_y, P\right) = y$. Therefore none of these probability measures δ_y are elements of $V_{vd}(P, 1) \in F^{vd}$. For any $B \in F^{pr}$ such that $P \in B$, however, there is a $y \in (0, 1]$ close enough to 0 for which $\delta_y \in B$. Hence for $P \in V_{vd}(P, 1)$ there is no $B \in F^{pr}$ such that $P \in B \subset V_{vd}(P, 1)$. This therefore constitutes a counterexample; $V_{vd}(P, 1) \in \mathcal{T}\left(F^{vd}\right)$, but $V_{vd}(P, 1) \notin \mathcal{T}(F^{pr})$.

To illustrate that $\mathcal{T}\left(F^{vd}\right)$ does not contain $\mathcal{T}\left(F^{kl}\right)$, we take $S = [0, 1]$ and P uniform on S. If we take distributions P_y uniform on $[y, 1]$, we find that for all $y \in (0, 1)$ the Kullback–Leibler I-divergence $d_{kl}\left(P_y, P\right) = \infty$, while $d_{vd}\left(P_y, P\right) = 2y$. Therefore none of these probability measures P_y are elements of $V_{kl}(P, 1) \in F^{kl}$. For any $B \in F^{vd}$ such that $P \in B$, however, there is a $y \in (0, 1]$ close enough to 0 for which $P_y \in B$. Hence for $P \in V_{kl}(P, 1)$ there is no $B \in F^{vd}$ such that $P \in B \subset V_{kl}(P, 1)$. This therefore constitutes a counterexample; $V_{kl}(P, 1) \in \mathcal{T}\left(F^{kl}\right)$, but $V_{kl}(P, 1) \notin \mathcal{T}\left(F^{vd}\right)$.

4.4 Not a base

The central ingredient of the proof of Satz 5 (Theorem 5) of Csiszár (1962) is the construction of a probability distribution P and sequences of distributions $\{Q_m\}_{m \in \mathbb{N}}$ and

$\left\{R_{m,n}\right\}_{n>m,m\in\mathbb{N}}$ for which $\lim_{m\to\infty} d_{kl}\left(P,\, Q_m\right) = 0$ and $\lim_{n\to\infty} d_{kl}\left(Q_m,\, R_{m,n}\right) = 0\,\forall\, m$ but $d_{kl}\left(P,\, R_{m,n}\right) = \infty\,\forall\, m, n$. We can then read Theorem 3 from Csiszár (1967) applied to the Kullback-Leibler I-divergence, together with the remark quoted above, as follows: F^{kl} is not a base for a topology. This is definitely true, also when using modern textbook Definition 19, since from the counterexample in Csiszár (1962), it follows that $V_{kl}\left(P, \delta\right), \delta > 0$ and $V_{kl}\left(Q_m, \epsilon\right), \epsilon > 0$, for m large enough, can feature as B_1 and B_2, and, for n large enough, $R_{m,n}$ would be an x for which condition 2 in Theorem 21 is not met.

References

Bickel, P.J., Klaassen, C.A., Ritov, Y., Wellner, J.A.: Efficient and Adaptive Estimation for Semiparametric Models. Baltimore: Johns Hopkins University Press (1993)

Billingsley, P.: Convergence of Probability Measures. New York: Wiley (1968)

Bomze, I.M.: Dynamical aspects of evolutionary stability. Monatshefte Math **110**, 189–206 (1990)

Bomze, I.M.: Cross entropy minimization in uninvadable states of complex populations. J Math Biol **30**, 73–87 (1991)

Bomze, I.M., Pötscher, B.M.: Game Theoretical Foundations of Evolutionary Stability. Berlin: Springer (1989)

Cressman, R.: Stability of the replicator equation with continuous strategy space. Math Soc Sci **50**, 127–147 (2005)

Cressman, R., Hofbauer, J.: Measure dynamics on a one-dimensional continuous trait space: theoretical foundations for adaptive dynamics. Theor Pop Biol **67**, 47–59 (2005)

Cressman, R., Hofbauer, J., Riedel, F.: Stability of the replicator equation for a single-species with a multi-dimensional trait space. J Theor Biol **239**, 273–288 (2006)

Csiszár, I.: Informationstheoretische Konvergenzbegriffe im Raum der Wahrscheinlichkeitsverteilungen. Magyar Tudomanyos Akademia Matematikai Kutato Intezetenek Közlemenyei **7**, 137–158 (1962)

Csiszár, I.: Über topologische und metrische Eigenschaften der relativen Information der ordnung α. In: Transactions of the third Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, Prague, pp. 63–73 (1964)

Csiszár, I.: On topological properties of f-divergences. Stud Sci Math Hungarica **2**, 329–339 (1967)

Csiszár, I.: I-divergence geometry of probability distributions and minimization problems. Ann Probab **3**, 46–158 (1975)

Hofbauer, J., Schuster, P., Sigmund, K.: A note on evolutionary stable strategies and game dynamics. J Theor Biol **81**, 609–612 (1979)

Oechssler, J., Riedel, F.: On the dynamic foundation of evolutionary stability in continuous models. J Econ Theory **107**, 223–252 (2002)

Oechssler, J., Riedel, F.: Evolutionary dynamics on infinite strategy spaces. Econ Theory **17**, 141–162 (2001)

Pervin, W.J.: Foundations of General Topology. Textbooks in Mathematics. New York: Academic Press (1964)

Reiss, R-D.: Approximate Distributions of Order Statistics. Berlin: Springer (1998)

Seymour, R.: Dynamics for infinite dimensional games. Working Paper, University College London (2000)

Sigmund, K.L.: Game dynamics, mixed strategies and gradient systems. Theor Pop Biol **32**, 114–126 (1987)

Sierpinski, W.: General Topology. Toronto: University of Toronto Press (1952, 1966)

Van Veelen, M.: Evolution in games with a continuous action space. Tinbergen discussion paper 2001-068/1 (2001)

Weibull, J.W.: Evolutionary Game Theory. Cambridge MA.: MIT Press (1996)

Williams, D.: Probability with Martingales. Cambridge: Cambridge University Press (1991)