

# Unifying evolutionary dynamics: a set theory exploration of symmetry and interaction

Francisco Berkemeier <sup>1,2,✉</sup> and Karen M. Page <sup>3,4</sup>

<sup>1</sup>Department of Genetics, University of Cambridge, CB2 3EH, Cambridge, United Kingdom

<sup>2</sup>Department of Pathology, University of Cambridge, CB2 1QP, Cambridge, United Kingdom

<sup>3</sup>Department of Mathematics, University College London, WC1E 6BT, London, United Kingdom

<sup>4</sup>Institute for the Physics of Living Systems, University College London, WC1E 6BT, London, United Kingdom

✉Corresponding author: [fp409@cam.ac.uk](mailto:fp409@cam.ac.uk)

**Within the expansive landscape of evolutionary dynamics, symmetry features embedded in well-established models significantly influence the interpretation of individual interaction patterns. Such symmetries are determined through interaction kernel functions, which serve as mathematical models for characterizing the complexity of interactions between individuals, each with distinct phenotypes. By incorporating analytical tools from logic and set theory, we aim to provide a deeper understanding of these functions, relevant to mechanisms of evolution. We prove that the kernels introduced in Champagnat et al.'s unifying framework exist provided birth and death rates are symmetric with respect to non-focal traits. The kernels may nevertheless be highly challenging to construct, thereby indicating a complex underlying mathematical infrastructure within unified evolutionary dynamics. We show how interaction kernels for asymmetric frameworks arising in evolutionary graph theory can be derived by incorporating individuals' graph labels into their phenotypes. These insights invite new avenues for research, providing a fresh understanding of the interactions between individuals in broader biological contexts.**

Evolutionary Dynamics | Interaction Kernels | Symmetry | Axiom of Choice

## Introduction

In the mid-19th century, Charles Darwin put forth the transformative theory of natural selection, suggesting that species undergo changes over time (Darwin et al., 1958; Darwin, 1987). This theory posits that each species possesses unique, inheritable genetic variations that have slowly evolved from a shared ancestor over a vast period of time, ultimately constructing an intricate tree of life. This tree manifests through a continual and varied sequence of branching events, including reproduction, death, and mutation, giving birth to new species in a phenomenon referred to as speciation. These fundamental processes permeate all of biology. Evolution shaped by natural selection promotes advancement and enhancement. More specifically, the traits that contribute to increased fitness within a species' phenotype are likely to become more preva-

lent. Darwinian fitness can be loosely understood as the capacity of an individual to live long enough to procreate successfully numerous times in order to sustain the population or species.

In early debates, biometricians W. F. R. Weldon and Karl Pearson presented arguments in favour of Darwin's theory, asserting that even minor variations could wield considerable influence on evolutionary progress (Darwin, 1859; Martins, 2007). In contrast, Francis Galton and William Bateson proposed that evolution fundamentally occurs in leaps and bounds, facilitated by the sudden appearance of markedly different individuals capable of passing on their phenotypes to future generations (Norton, 1973; Olby, 1989). As the scientific understanding evolved, the concepts of Mendelian genetics and Darwinian evolution were harmoniously integrated by Haldane, Fisher, and Wright (Haldane, 1956; Fisher et al., 1958; Haldane and Jayakar, 1963; Wright, 1969). This synthesis expanded the mathematical framework for modelling and interpreting these systems. Such mathematical articulation of evolutionary dynamics laid the foundation for groundbreaking theories, including Hamilton's kin selection (Hamilton, 1970), Kimura's theory of neutral evolution (Kimura and Harper, 1970), and Maynard Smith's evolutionary game theory (Smith, 1982). Overall, the cornerstone of mathematical models for adaptation and co-evolution of biological populations within Darwinian dynamics, focusing on selection and mutation, lies in a comprehensive understanding of population genetics combined with evolutionary game theory.

Processes that enhance fitness are typically instrumental in shaping evolutionary dynamics by posing an optimisation problem amenable to modelling. Over time, these dynamics often reach a predictable equilibrium after resolving the problem at hand. However, stochastic effects arise in these systems due to genetic mutations and environmental changes. To analyse frequency-dependent selection across the phenotypic space, the integration of evolutionary game theory becomes an essential part of the mathematical and computational approach to biology.

The triad of evolution – reproduction, mutation, and selection – has been the subject of extensive research within the domain of evolutionary dynamics. Notably,

a number of studies have attempted to integrate various evolutionary dynamics modelling methodologies in multiple ways. A unified framework for modelling evolutionary processes was created in the deterministic scenario, taking into account a continuous relative fitness function of strategically interacting individuals. Here, the authors examined broad categories of evolutionary dynamics and compared them to Nash equilibria (Kreps, 1989; Joosten, 1996). Leveraging the equivalence between the replicator-mutator equation, which delineates the dynamics of population distribution (Hadeler, 1981), and the Price equation, describing its moments (Price et al., 1970; Price, 1972), numerous mathematical models of evolutionary dynamics were demonstrated to fall under a cohesive, unified framework (Page and Nowak, 2002). These equations have been revealed to spawn a multitude of modelling contexts, such as adaptive dynamics (Nowak and Sigmund, 1990; Metz et al., 1995; Dieckmann and Law, 1996), evolutionary game dynamics (Hofbauer et al., 1998), the Lotka-Volterra ecological equation (Volterra, 1926), and the quasi-species equation pertinent to molecular evolution (Eigen et al., 1989). Evolutionary dynamics were also studied under different timescales and rescaling limits by N. Champagnat and colleagues (Champagnat et al., 2006). It was discovered that the architectural framework of macroscopic models within evolutionary dynamics is considerably shaped by the temporal aspects of individual events. In this context, the authors presented a detailed model that encapsulates the stochastic dynamics of birth, mutation, and death in continuous time, influenced by each individual's trait values and their interactions with others. This culminated in a general algorithm for an efficient numerical simulation of the individual-level model. Using the stochastic point process, an array of macroscopic models of adaptive evolution was generated, each differing based on the assumed renormalisation (such as population size, mutation rate, and mutation step), thereby influencing their deterministic or stochastic nature under precise time rescaling.

Early in the discussion of this unifying description of evolutionary dynamics, microscopic population point processes were introduced as dependent on the interaction between different individuals, characterised by their phenotypes. Such a dependence was written in terms of interaction kernels, which provide the basis of the process construction. Our study aims to identify the constraints of such a modelling framework. Specifically, we strive to comprehend how symmetry with respect to (w.r.t.) non-focal phenotypes, reliant on these interaction kernels, is upheld within this framework. Our investigation leads us to a deep theoretical discourse, resting on the principles of the Axiom of Choice and Zorn's Lemma. Although the biological interpretation of the Axiom of Choice offers a fascinating philosophical topic on its own, it falls outside the purview of this work.

We revisit the point process modelling notions in Champagnat et al. (2006), along with the primary instruments

from logic and set theory employed in validating the central theorems in this work. Ultimately, we highlight some implications of these theorems with examples concerning evolutionary games defined on graphs.

## Theoretical framework

Here we present the main theoretical background, including the relevant concepts presented in Champagnat et al. (2006), as well as the set theory elements essential to the main discussion.

### Evolutionary dynamics

The explanation of the population point process in Champagnat et al. (2006) makes use of several mathematical techniques and ideas. The comprehensive exposition below closely follows the discussion presented in that paper.

Each individual's phenotype is represented as a real-valued vector of trait values. The trait space  $X$  satisfies  $X \subseteq \mathbb{R}^\ell$ , where  $\ell$  is the number of traits. A counting metric that counts the number of individuals displaying various phenotypes is used to characterise the population as a whole, evolving according to a Markov process (Gagniuc, 2017). The Markov property presupposes that the population's dynamics at time  $t$  depend exclusively on the population's current state.

Here, we take into account a population in which individuals can reproduce and die at rates that are determined by their unique traits as well as by their interactions with others who have similar or different traits. These events occur continuously and at random. Reproduction is almost faithful, but there is a chance that a mutation will result in an offspring with a phenotype that is different from that of its ancestor. For the purpose of our discussion, however, we do not account for mutations.

Birth and death rates of focal individuals with phenotype  $x \in \mathbb{R}^\ell$  are defined as functions of  $x$  and their interactions with other individuals. These interactions are measured by an *interaction kernel*, which is a function of how phenotypically different they are. For a population of size  $n$ , this dependence is explicitly given by

$$b(x, \sum_{i=1}^n V(x - x_i)) \quad (1)$$

$$d(x, \sum_{i=1}^n U(x - x_i)), \quad (2)$$

where  $b$  and  $d$  are the birth and death rates, respectively, and  $V$  and  $U$  are the interaction kernels affecting reproduction and mortality. We may now question the generality of this approach and which ways of defining such rates are not being considered by writing the dependence as above.

Symmetry w.r.t. non-focal phenotypes is a first property that is worth discussing. Indeed, both birth and death rates with trait dependence given by Eq. (1)-Eq. (2) are

symmetric w.r.t. non-focal traits, which has intricate implications derived from set theory that we explore in the following sections.

### Elements of set theory

We are interested in determining what are the properties of a real function  $h$  such that there exist functions  $f$  and  $g$  satisfying

$$f\left(x, \sum_{i=1}^n g(x - y_i)\right) = h(x, y_1, \dots, y_n) \quad (3)$$

for all  $x, y_1, \dots, y_n \in \mathbb{R}^\ell$  and  $n \geq 2$ . We claim that Eq. (3) holds, for some functions  $f$  and  $g$ , if and only if  $h$  is symmetric w.r.t.  $y_1, \dots, y_n$ , that is, its value is the same independently of the order of its arguments. For example, a function  $f(x_1, x_2)$  taking two arguments is a symmetric function if and only if  $f(x_1, x_2) = f(x_2, x_1)$  for all  $x_1$  and  $x_2$  such that  $(x_1, x_2)$  and  $(x_2, x_1)$  are in the domain of  $f$ . More relevant to our case, we can see that a function of the form

$$\sum_{i=1}^n g(x - x_i) \quad (4)$$

is also symmetric w.r.t.  $x_1, \dots, x_n$ . To begin the thought process behind the main theorems in this work, we present some examples with  $n = 2$ .

**Example 1** (A functional equation). Suppose

$$f(x + y) = ax + by \quad (5)$$

holds for some function  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $a, b \in \mathbb{R}$ . Given the function  $\bar{f} : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined as  $\bar{f}(x, y) \equiv f(x + y)$  we see that it is symmetric w.r.t.  $x$  and  $y$ , since  $x + y = y + x$ . Then, it follows that

$$ax + by = \bar{f}(x, y) = \bar{f}(y, x) = ay + bx \quad (6)$$

and thus  $a = b$ . Hence, if  $a \neq b$ , Eq. (5) does not hold for any function  $f$ . In some sense, symmetry imposed by  $\bar{f}(x, y)$  is extended to  $ax + by$ . A more straightforward way to see this is by noting that  $f(0 + 0) = f(1 - 1)$  implies  $a = b$ .

**Example 2** (Symmetry necessity). Let  $h(x, y, z) = ax + by + cz$ , where  $a, b, c \in \mathbb{R}$ . Then Eq. (3) becomes

$$f(x, g(x - y) + g(x - z)) = ax + by + cz. \quad (7)$$

However, from Example 1,  $\bar{f}(x, y, z) \equiv f(x, g(x - y) + g(x - z))$  is symmetric w.r.t.  $y$  and  $z$  and so

$$\begin{aligned} ax + by + cz &= f(x, g(x - y) + g(x - z)) \\ &= f(x, g(x - z) + g(x - y)) = ax + bz + cy \end{aligned} \quad (8)$$

thus  $f$  and  $g$  exist if and only if  $b = c$ . In particular, take  $g(w) = w$  and  $f(u, v) = au + b(2u - v)$ . For  $n > 2$ , the reasoning is similar. This example aims to suggest that

some kind of symmetry in  $h$  in Eq. (3) is necessary. In fact, since  $f$  is symmetric w.r.t.  $y_1, \dots, y_n$ ,  $h$  is necessarily symmetric w.r.t. the same variables, thus we now focus on the reverse implication.

**Example 3** (Symmetry sufficiency). We now illustrate that symmetry of  $h$  might not only be necessary, but also sufficient to guarantee the existence of  $f$  and  $g$  such that Eq. (3) holds. Since isolated dependence on  $x$  is granted from  $f$ , we drop it from  $h$  and assume the non-linear case  $h(y, z) = y^2 + z^2$ , which is symmetric w.r.t.  $y$  and  $z$ . Thus

$$f(x, g(x - y) + g(x - z)) = y^2 + z^2. \quad (9)$$

Here, the answer is not trivial and further machinery is needed. With a suitable change of variables, we can rewrite Eq. (9) as

$$f(x, g(y) + g(z)) = (y + x)^2 + (z + x)^2. \quad (10)$$

Since we cannot explicitly impose dependence of  $g$  on variable  $x$ , the only way to satisfy Eq. (10) is to find a function  $g$  such that the sum  $g(y) + g(z)$  uniquely determines the unordered pair  $(y, z)$ . This would then guarantee explicit dependence of  $f$  on  $(y, z)$ , simplifying the problem. Note that the unordered property is a consequence of symmetry. Ideally, if  $y, z \in \mathbb{R}$ , we would like  $g$  to be a function between  $\mathbb{R}$  and any set with the cardinality of the continuum (cardinality of  $\mathbb{R}$ , usually denoted by  $2^{\aleph_0}$ ), such that any two sums in that set give distinct values. Such a set and bijection are not trivial to construct and their existence relies on Zorn's Lemma, discussed next.

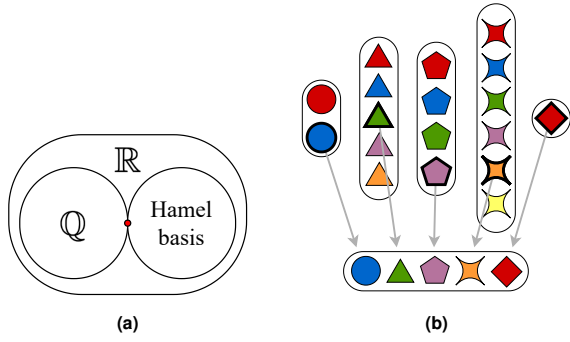
#### On the existence of a basis and the Axiom of Choice.

By interpreting the real numbers as a vector space over some set, constructing a basis for  $\mathbb{R}$  and defining  $g$  in Eq. (10) as the bijection between  $\mathbb{R}$  and its basis leads to the unique description of the unordered pair  $(y, z)$ , since  $g(y)$  and  $g(z)$  are linearly independent over the set we build our basis on. In general, this would then allow us to take the right-hand side of Eq. (3) as any symmetric function w.r.t. non-focal traits.

One way to do this is to consider  $\mathbb{R}$  as a vector space over the rationals, where a basis is a collection  $\mathcal{B}$  of real numbers such that every real number is a rational linear combination of numbers  $\beta_i \in \mathcal{B}$  in precisely one way, that is, every real number  $\alpha$  has a unique representation of the form

$$\alpha = \sum_{i=1}^m q_i \beta_i, \quad (11)$$

where  $q_i$  is rational and  $m$  depends on  $\alpha$ . This set is also known as the *Hamel basis* for  $\mathbb{R}$  (Hamel, 1905). Note that any Hamel basis contains, at most, one rational number (Figure 1a). Considered as a vector space over  $\mathbb{Q}$ ,  $\mathbb{R}$  has uncountable dimension, so it cannot be spanned by countably many vectors. Hamel bases have been studied in the context of additive functions (Jones,



**Figure 1. Hamel bases and Axiom of Choice.**

(a) Hamel bases contain, at most, one rational number (in red). (b) The choice function takes exactly one element from each set  $\alpha$  (Yorgey, 2014).

1942; Boros and Daróczy, 2006), including the discussion of the general solution of Cauchy’s functional equation (Aczél and Erdős, 1965). The existence of such a basis depends on the *Axiom of Choice* or, more directly, *Zorn’s Lemma* (Zermelo, 1904; Kuratowski, 1922; Zorn, 1935). In the following, we are then considering an axiomatic system under Zermelo-Fraenkel set theory with the Axiom of Choice (ZFC) (Zermelo, 1908; Fraenkel et al., 1973).

The Axiom of Choice states that for any family  $\mathcal{A}$  of nonempty sets, there exists a set that contains exactly one element in common with each of the nonempty sets. In other words, for any family  $\mathcal{A}$  of nonempty sets, there exists a *choice function*

$$F : \mathcal{A} \rightarrow \bigcup_{\alpha \in \mathcal{A}} \alpha \quad (12)$$

such that for every  $\alpha \in \mathcal{A}$ ,  $F(\alpha) \in \alpha$  (Figure 1b). These statements are equivalent to a third “version” of the Axiom of Choice:

**Zorn’s Lemma** (Axiom of Choice). *If  $X$  is a partially ordered set such that every chain in  $X$  has an upper bound, then  $X$  has a maximal element.*

It is a known result that Zorn’s Lemma, which relies on the assumption of the Axiom of Choice, yields that every vector space has a basis (Halpern, 1966; Bleicher, 1967). In the Supplementary Information, we sketch the proof of that result for  $\mathbb{R}$ .

Going back to Example 3, with  $g$  as the bijection between  $\mathbb{R}$  and its Hamel basis  $\mathcal{B}$ , we have that  $(g(y), g(z)) \equiv (y_{\mathcal{B}}, z_{\mathcal{B}}) \in \mathcal{B} \times \mathcal{B}$ . Hence, since  $\mathcal{B}$  is a basis of  $\mathbb{R}$ ,  $y_{\mathcal{B}}$  and  $z_{\mathcal{B}}$  are linearly independent over the rationals and the sum  $g(y) + g(z) = y_{\mathcal{B}} + z_{\mathcal{B}}$  uniquely defines a real number. Thus it is possible to rewrite the dependence  $f$  in Eq. (10) as

$$f(x, g(y) + g(z)) \equiv h(x, y + x, z + x), \quad (13)$$

which is symmetric w.r.t.  $y$  and  $z$ . A numerical formulation of  $f$  and  $g$  is discussed later in Example 4, where we explicitly solve Example 3. Next, we present the main implications of this argument.

## Results

Recall the general phenotype-dependent expressions of birth and death rates in Champagnat et al. (2006),

$$b(x, \sum_{i=1}^n V(x - x_i)) \quad (14)$$

$$d(x, \sum_{i=1}^n U(x - x_i)). \quad (15)$$

Motivated by the discussion in the previous section, we present the main theorems on symmetry derived from a phenotype-dependence like Eq. (14)-Eq. (15).

### Main theorems

Under the Axiom of Choice and via a set theory argument, one can show that the birth and death rates can be taken to be any symmetric functions.

**Theorem 1** (General symmetry).

*Assume the Axiom of Choice. For all  $x, x_1, \dots, x_n \in \mathbb{R}^\ell$  and  $n \geq 2$ , there exist functions  $f$  and  $g$  such that*

$$f(x, \sum_{i=1}^n g(x - x_i)) = h(x, x_1, \dots, x_n) \quad (16)$$

*holds if and only if  $h$  is symmetric w.r.t.  $x_1, \dots, x_n$ .*

*Proof.* ( $\Rightarrow$ ) First, assume that, for some functions  $f, g$  and  $h$ , Eq. (16) holds for all  $x, x_1, \dots, x_n \in \mathbb{R}^\ell$ . Since  $f$  is symmetric w.r.t.  $x_1, \dots, x_n$ ,  $h$  is symmetric w.r.t. the same variables. ( $\Leftarrow$ ) Now, take  $\ell = 1$  for simplicity and let  $h$  be any symmetric function. Considering a suitable change of variables, rewrite Eq. (16) as

$$f(x, \sum_{i=1}^n g(x_i)) = h(x, x_1 + x, \dots, x_n + x). \quad (17)$$

Let  $g$  be the bijection between  $\mathbb{R}$  and its Hamel basis,  $\mathcal{B}$ , which exists according to Zorn’s Lemma and has the cardinality of the continuum. Then,  $\sum_{i=1}^n g(x_i)$  uniquely determines the unordered  $n$ -tuple  $(x_1, \dots, x_n)$ , since  $\{g(x)\}_x$  are linearly independent over  $\mathbb{Q}$ . Hence, with  $w = \sum_{i=1}^n g(x_i)$ , define

$$f(x, w) = h(x, x_1 + x, \dots, x_n + x). \quad (18)$$

Reverting the change of variables, we get that Eq. (16) holds for all  $x, x_1, \dots, x_n \in \mathbb{R}$ . A similar argument yields the same result for any dimension  $\ell$ . ■

According to Theorem 1 and by interpreting  $f$  and  $g$  as the event rate and respective interaction kernel, Champagnat’s approach is only considering birth and death rates which are symmetric w.r.t. the trait values of the non-focal individuals, leaving a plethora of cases to be explored. We now ask which of them could be of relevant interest to evolutionary dynamics. Furthermore, the highly non-constructive nature of the proof of symmetry sufficiency ( $\Leftarrow$ ) may be rather uninteresting to the applied mathematician. Therefore, in addition to the previous limitation, Champagnat’s framework is limited to a very particular type of symmetry.

## Constructible symmetry

A way to avoid the non-constructive scenario and the Axiom of Choice is to relax slightly our phenotype requirements. If we consider the traits to be vectors of rational numbers, a more direct approach may be taken. Indeed, from a computational perspective, all we really manage in simulations are approximations to real numbers, which can be seen as rational numbers (to machine error, in non-symbolic mathematics (Zazkis, 2005; Obersteiner and Hofreiter, 2017; Kristiansen, 2017; Georgiev et al., 2021)). Therefore, on rational domains, we have a weaker but constructible version of Theorem 1.

### Theorem 2 (Constructible symmetry).

For all  $x, x_1, \dots, x_n \in \mathbb{Q}^\ell$  and  $n \geq 2$ , there exist constructible functions  $f$  and  $g$  such that

$$f\left(x, \sum_{i=1}^n g(x - x_i)\right) = h(x, x_1, \dots, x_n) \quad (19)$$

holds if and only if  $h$  is symmetric w.r.t.  $x_1, \dots, x_n$ .

*Proof.* ( $\Rightarrow$ ) Similar to the proof of Theorem 1. ( $\Leftarrow$ ) Take  $\ell = 1$  for simplicity. Considering the change of variables Eq. (17), we are now interested in finding a constructible function  $g$  such that  $\sum_{i=1}^n g(x_i)$  is uniquely defined for each  $n$ -tuple  $(x_1, \dots, x_n)$ , i.e. such that the sum is injective w.r.t. the same tuple. Take  $\gamma : \mathbb{Q} \rightarrow \mathbb{N}$  to be the bijection between  $\mathbb{Q}$  and  $\mathbb{N}$ . We avoid determining an explicit analytical formula for  $\gamma$ , arguing instead that this is directly programmable with a Cantor pairing function (Szudzik, 2006; Lisi, 2007). If  $g(x_i) = n^{\gamma(x_i)}$ , the sum

$$\sum_{i=1}^n g(x_i) = \sum_{i=1}^n n^{\gamma(x_i)} \quad (20)$$

is unique for each unordered tuple of integers  $(\gamma(x_1), \dots, \gamma(x_n))$  and thus for  $(x_1, \dots, x_n)$ . To see this, let  $k_i \equiv \gamma(x_i) \in \mathbb{N}$ . We are then looking for an injective function  $\sigma : \mathbb{N} \rightarrow \mathbb{N}^n$  such that

$$\sigma\left(\sum_{i=1}^n n^{k_i}\right) = p(k_1, \dots, k_n), \quad (21)$$

where  $p : \mathbb{N}^n \rightarrow \mathbb{N}^n$  is any permutation of  $(k_1, \dots, k_n)$ . Then, due to the uniqueness of base- $n$  representations, we have that such  $\sigma$  exists. In particular, consider the number  $S_n \equiv \sum_{i=1}^n n^{k_i}$  written in base- $n$ , i.e.,

$$S_n = \sum_{i=1}^n n^{k_i} = (s_1 s_2 \dots s_m)_n. \quad (22)$$

Assume first that  $\{k_i\}$  are not all equal to a specific value. Then, we have that  $\sum_{j=1}^m s_j = n$  and the number  $s_j$  gives precisely the number of terms with exponent  $j = k_i$  that appear in  $\sum_{i=1}^n n^{k_i}$ . If, however, the sum is itself a power of  $n$  ( $k_i = k$ ,  $\forall i$ , for some  $k$ ), then we take  $k_i = \log_n(S_n/n)$ ,  $\forall i$ . Hence, we have a unique description of the  $n$ -tuple via the sum. Writing  $f$  in terms

of  $\sigma$  and applying  $\gamma^{-1}$  to  $p(k_1, \dots, k_n)$  implies an explicit dependence on the tuple  $(x_1, \dots, x_n) \in \mathbb{Q}^n$ . A similar argument yields the same result for any dimension  $\ell$ . ■

Following Theorem 2, a natural extension would be to consider any countable subset of  $\mathbb{R}$ , beyond  $\mathbb{Q}$ . For example, the subset of real algebraic numbers, which is countable and contains  $\mathbb{Q}$ , could be of particular interest, since these represent roots of non-zero polynomials with rational coefficients, which may appear when modelling certain traits. The previous argument would then rely entirely on the constructible nature of the bijection between any countable set and  $\mathbb{N}$ . For computational purposes, however, rational traits seem to be enough.

### Further examples

**Example 4** (Solving Example 3). Consider the one-trait scenario ( $\ell = 1$ ) in a population with three individuals. Let  $h : \mathbb{Q} \rightarrow \mathbb{R}$  be defined as  $h(x, y, z) = y^2 + z^2$ . We now provide explicit functions  $f$  and  $g$  such that

$$f(x, g(x - y) + g(x - z)) = y^2 + z^2 \quad (23)$$

for all  $x, y, z \in \mathbb{Q}$ . Again, a suitable change of variables allows us to rewrite the previous identity as

$$f(x, g(y) + g(z)) = (y + x)^2 + (z + x)^2. \quad (24)$$

Let  $\gamma : \mathbb{Q} \rightarrow \mathbb{N}$  be the bijection between  $\mathbb{Q}$  and  $\mathbb{N}$ ,  $g(x) = 2^{\gamma(x)}$  and fix  $(k_1, k_2) \equiv (\gamma(y), \gamma(z))$ . We aim to find an injective function  $\sigma : \mathbb{N} \rightarrow \mathbb{N}^2$  such that

$$\sigma(2^{k_1} + 2^{k_2}) = (\max\{k_1, k_2\}, \min\{k_1, k_2\}). \quad (25)$$

Notice first that

$$2^{\max\{k_1, k_2\}} < 2^{k_1} + 2^{k_2} \leq 2^{\max\{k_1, k_2\} + 1} \quad (26)$$

implies

$$\max\{k_1, k_2\} < \log_2(2^{k_1} + 2^{k_2}) \leq \max\{k_1, k_2\} + 1, \quad (27)$$

and so  $\log_2(2^{k_1} + 2^{k_2})$  falls between two successive integers. Hence, with  $k \equiv 2^{k_1} + 2^{k_2}$  and  $k_1 \neq k_2$ , we have

$$\begin{cases} \max\{k_1, k_2\} = \lfloor \log_2(k) \rfloor, \\ \min\{k_1, k_2\} = \lfloor \log_2(k - 2^{\lfloor \log_2(k) \rfloor}) \rfloor, \end{cases} \quad (28)$$

where  $\lfloor \cdot \rfloor$  is the floor function, that is,  $\lfloor x \rfloor$  gives the greatest integer less than or equal to  $x$ . If  $k_1 = k_2$ , we have that  $k = 2^{k_1+1}$  and thus  $\log_2(k/2) = k_1 \in \mathbb{N}$ . Combining these two cases, we may define  $\sigma$  as

$$\sigma(k) = \left( \begin{array}{c} \lfloor \log_2(k) \rfloor - \chi_{\mathbb{N}}(\log_2(k/2)) \\ \lfloor \log_2(k - 2^{\lfloor \log_2(k) \rfloor}) - \chi_{\mathbb{N}}(\log_2(k/2)) \end{array} \right) \quad (29)$$

$$\equiv (\sigma_1(k), \sigma_2(k))^T, \quad (30)$$

where  $\chi_{\mathbb{N}}$  is the indicator function of set  $\mathbb{N}$ . Hence, Eq. (25) is satisfied. Finally, with  $g(w) = 2^{\gamma(w)}$ , we define  $f$  as

$$f(u, v) = (\gamma^{-1}[\sigma_1(v)] + u)^2 + (\gamma^{-1}[\sigma_2(v)] + u)^2. \quad (31)$$

Hence,

$$f(x, g(y) + g(z)) = f(x, 2^{\gamma(y)} + 2^{\gamma(z)}) \quad (32)$$

$$= \sum_{i=1,2} \left( \gamma^{-1}[\sigma_i(2^{\gamma(y)} + 2^{\gamma(z)})] + x \right)^2 \quad (33)$$

$$= (y + x)^2 + (z + x)^2 \quad (34)$$

$$= h(x, y + x, z + x). \quad (35)$$

Reverting the change of variables, we have found  $f$  and  $g$  such that Eq. (23) holds for all rationals.

We now aim to find “reasonable” examples, in the sense of having some level of literature background and application, of birth and death rates that not only contrast with Champagnat’s framework, but also suggest a potentially more general approach to unifying evolutionary dynamics. Based on these examples, or class of examples, we propose an expansion of the modelling in that paper.

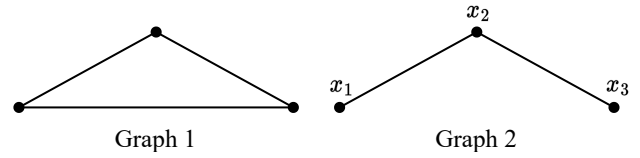
### Evolutionary games on graphs

In evolutionary dynamics, games dependent on cooperativity in a graph structure are often accepted as realistic approximations to describing the dynamics of populations. Thus we now aim to frame evolutionary graph theory in Champagnat’s framework and understand its limitations through examples motivated by other works (Ohtsuki et al., 2006; Allen et al., 2013, 2017).

**Asymmetric interactions.** As discussed below, different degrees and weights of vertices representing individuals might suggest some overall asymmetric growth rates, regarding traits and interactions between specific individuals. For example, the graphs in Figure 2 are a simple first case where Champagnat’s approach could fail, due to the asymmetric nature of the incomplete Graph 2. Considering the focal individual with trait  $x_1$ , its dynamics would only be influenced by the individual with trait  $x_2$ , as suggested in Graph 2, leading to an interaction kernel equation of the form

$$f(x_1, g(x_1 - x_2) + g(x_1 - x_3)) = h(x_1, x_2), \quad (36)$$

which is not possible due to the symmetry w.r.t.  $x_2$  and  $x_3$  imposed by the left-hand side. Considering the degree of each vertex as an extra trait, however, could solve the mathematical conundrum in Eq. (36), guaranteeing non-focal trait symmetry in some sense, which we discuss next for some examples. Notice also that, even if a specific permutation of the unordered tuple  $(x_2, x_3)$  was picked, information on individuals’ labels is required to handle asymmetry, which we discuss in the following.



**Figure 2. Two evolutionary graphs.**

In a population in which interaction is defined by incomplete graphs, asymmetry can be overturned by a redefinition of the phenotypic landscape.

**Incomplete graphs.** Consider any graph  $G$  of order  $n$  with vertices and edges representing different individuals and their potential interactions, respectively. We assume that the graph is fixed for the duration of the evolutionary dynamics and that the edges are equally weighted. In the one-shot continuous Prisoner’s Dilemma (Poundstone, 1993), a player with cooperativity  $x$  pays a cost  $C(x)$  to produce a benefit  $B(x)$  for the other player. Here, we consider the game with more players. Therefore, the payoff of a player  $i$  with cooperativity  $x_i$  is

$$P_i = \sum_{j:ij \in E} [B(x_j) - C(x_i)], \quad (37)$$

where  $E$  is the edge set of the graph  $G$  and  $ij$  is the edge connecting players  $i$  and  $j$ . If  $G$  is complete, the sum in Eq. (37) is over all  $j$ . The fitness of player  $i$ ,  $f_i$ , is given by a constant term, denoting the baseline fitness, plus the payoff that arises from the game, that is,

$$f_i \equiv 1 - \phi + \phi P_i, \quad (38)$$

where  $\phi$  measures the intensity of selection. For simplicity, and without loss of generality, we consider strong selection,  $\phi = 1$ , meaning that the payoff is large compared to the baseline fitness. Hence,  $f_i = P_i$ . Without delving into much detail on the biological meaning of fitness and its connection to event rates (Wassersug and Wassersug, 1986; Smith et al., 1989; Bertram and Masel, 2019), we assume that, independently of the chosen update rule (see note below), both birth and death rates of a focal individual  $i$  depend, in some way, on its fitness. Therefore, taking the cooperativity as a trait, we expect the following dependence

$$\begin{aligned} b(x_i, \sum_{j=1}^n V(x_i - x_j)) &= \tilde{b}(x_i, f_i) \\ &= \tilde{b}(x_i, \sum_{j:ij \in E} [B(x_j) - C(x_i)]), \end{aligned} \quad (39)$$

where  $\tilde{b}$  is related to  $b$  and  $V$  in a specific way. Similarly for the death rate  $d$ . The question is now whether  $f_i$  is symmetric w.r.t. the other cooperativities  $x_j$ ,  $j \neq i$ . In general, this is not necessarily true, as seen in the example of Graph 2 (Figure 2), by noting that, for example,  $f_1 = B(x_2) - C(x_1) \neq B(x_3) - C(x_1)$ ,  $\forall x_2, x_3$ . However, a phenotype redefinition based on the graph’s connectivity could hint at a possible mathematical solution. Notice that, since we have access to the edge set

$E$ , rewriting Eq. (37) as

$$\begin{aligned} f_i = P_i &= \sum_{j:ij \in E} [B(x_j) - C(x_i)] \\ &= \sum_{j=1}^n [B(x_j) - C(x_i)] \chi_E(ij) \end{aligned} \quad (40)$$

suggests that some extra trait regarding the edge set is needed to guarantee symmetry. Due to Theorem 1, Eq. (39) is always possible as long as Eq. (40) is symmetric w.r.t. non-focal traits. However, the edge set  $E$  must be somehow obtained from the phenotypes, in order to be consistent with the dependence in Eq. (1)-Eq. (2). In other words, we want to define the set  $\{j : ij \in E\}$  in terms of a certain trait of the phenotype. We now question whether there is a way of relabelling the graph such that, for any vertex  $i$ , with corresponding new label  $p_i \in \mathbb{N}$ , there exists a set  $L$  for which

$$\chi_L(p_i - p_j) = \begin{cases} 1 & \text{if } ij \in E, \\ 0 & \text{if } ij \notin E. \end{cases} \quad (41)$$

The most straightforward thing to do is take the labels  $p_1, \dots, p_n$  to be so far apart that no difference  $p_i - p_j$  is repeated, which condition is reminiscent of a *Sidon sequence* (Erdos and Turán, 1941; O'Bryant, 2004). Indeed, we can take  $p_i = 2^i$ . If  $2^a - 2^b = 2^c - 2^d$ , then we also have  $2^a + 2^b = 2^c + 2^d$ , and by the uniqueness of binary representations, we must have  $a = c$  and  $b = d$ . Now, let

$$L = \{2^i - 2^j : ij \in E\}, \quad (42)$$

where  $E$  is the edge set of the graph. Then Eq. (41) holds. Hence, considering  $p_i = 2^i$  as a second trait in the phenotype, we can rewrite Eq. (40) as

$$\begin{aligned} f_i &= \sum_{j:ij \in E} [B(x_j) - C(x_i)] \\ &= \sum_{j=1}^n [B(x_j) - C(x_i)] \chi_L(p_i - p_j). \end{aligned} \quad (43)$$

Therefore, it becomes clear that fitness  $f_i$ , written as a function of two traits (cooperativity and vertex degree), is symmetric w.r.t. similar non-focal traits and thus Champagnat's framework can be extended to incomplete graphs. Going back to the example Eq. (36),  $h$  can now be redefined as

$$f \left[ \begin{pmatrix} x_1 \\ p_1 \end{pmatrix}, g \left[ \begin{pmatrix} x_1 - x_2 \\ p_1 - p_2 \end{pmatrix} \right] + g \left[ \begin{pmatrix} x_1 - x_3 \\ p_1 - p_3 \end{pmatrix} \right] \right] \quad (44)$$

$$\equiv h \left[ \begin{pmatrix} x_1 \\ p_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ p_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ p_3 \end{pmatrix} \right] \quad (45)$$

$$\equiv h \left[ \begin{pmatrix} x_1 \\ p_1 \end{pmatrix}, x_2 \chi_L(p_1 - p_2), x_3 \chi_L(p_1 - p_3) \right], \quad (46)$$

which works for any other focal individual. More generally, we may consider weighted graphs.

**Weighted graphs.** In Allen et al. (2017), the population structure is represented by a weighted, undirected graph  $G$  with edge weights  $w_{ij}$ . The *weighted degree* of a vertex  $i$  is defined as  $w_i = \sum_{j \in G} w_{ij}$ . These weights determine the frequency of game interaction and the probability of replacement between vertices  $i$  and  $j$ , considering a death-birth update rule. Here, the notion of cooperation in a game is simplified to two types of individuals, characterised by

$$s_i = \begin{cases} 1 & \text{if } i \text{ is a cooperator,} \\ 0 & \text{if } i \text{ is a defector.} \end{cases} \quad (47)$$

Again, we can think of  $s_i$  as a trait. Considering the donation game, the payoff matrix is defined as

$$\begin{array}{cc} & C & D \\ C & (b-c & -c) \\ D & (b & 0) \end{array}, \quad (48)$$

where  $C$  and  $D$  represent cooperators and defectors, respectively, and  $b, c > 0$  are the benefit and cost values. Then, the edge-weighted average payoff of vertex  $i$  is given by

$$P_i = -cs_i + b \sum_{j \in G} \frac{w_{ij} s_j}{w_i}. \quad (49)$$

The fitness of individual  $i$ , which is interpreted as birth rate, is then given by  $f_i = 1 + \delta P_i$ , where  $\delta > 0$  quantifies the strength of selection. Since the weights  $w_{ij}$  are fixed (or interpreted as extra traits), we have that, as long as at least two weights are different, Eq. (49) is not symmetric w.r.t non-focal types  $s_j$ . Hence, evolutionary dynamics on weighted graphs are also not covered by Champagnat's framework, unless we consider, again, the connectivity between individuals as traits, which relaxes the definition of asymmetry.

We have then concluded that, from an evolutionary graph theory perspective, the framework in Champagnat et al. (2006) is restricted to symmetric games on well-mixed populations, unless some assumption is made on the phenotype definition to include the edge set asymmetries of the interaction graph.

**A note on update rules.** Evolutionary dynamics depend on update rules upon which these games are settled. In Ohtsuki et al. (2006), for example, a few update rules are considered: death-birth, birth-death and "imitation" updating. However, it may be quite a challenge to adapt them in Champagnat's framework. The key difference is that Champagnat uses the values of a sequence of random variables uniformly distributed in  $[0, 1]$  to select the type of birth or death events, based on the event rates, a methodology not that dissimilar to Gillespie's algorithm (Gillespie, 1976, 1977). Therefore, they consider a birth-mutation-death process that follows no particular order and where there is no explicit connection between two consecutive events, which may, in some sense, be more realistic. Setting extremely high or low event rates

in an artificial manner and introducing specific traits for positioning and number of neighbours might approximate common update rules in evolutionary graph dynamics. Nonetheless, our symmetry argument remains independent of this adaptation.

### Solving asymmetry

In the previous examples, we have shown that the inclusion of a structure-dependent trait in the phenotypic landscape of the population allows for an adjustment of asymmetric frameworks to be accounted for by Champagnat's symmetric framework. While seemingly a very artificial argument, incomplete graphs revealed the necessity for the existence of a well-defined trait based on the graph portrait of interaction, in order to guarantee symmetry. The *a priori* knowledge of the interaction structure is key in order to apply the main framework. We call this trait addition the *Sidon-extension* of non-focal phenotypes to  $\mathbb{R}^{\ell+1}$ .

The previous examples Eq. (37) and Eq. (49) are but a small class of very specific *linear* asymmetry. We will see, however, that any asymmetric functions w.r.t non-focal traits can still fit into Champagnat's framework under a suitable transformation.

**Example 5** (Symmetrizations). Consider the following asymmetric and non-linear function  $h(y, z)$ . In general, such a function could not be represented by an interaction kernel as in Champagnat et al. (2006). However, a symmetrization of  $h$  based on the relative positioning of individuals (or phenotypes) would extend the previous arguments to any asymmetric function. Pairwise adding extra positional traits  $p_z > p_y$ , we may write a symmetric version of  $h$  as follows

$$h_S(\mathbf{y}, \mathbf{z}) = h(y, z) + H(p_y - p_z)(h(z, y) - h(y, z)) \quad (50)$$

where  $\mathbf{y} = (y, p_y)^T$ ,  $\mathbf{z} = (z, p_z)^T$ , and  $H$  is the Heaviside step function. Indeed  $h_S$  is now symmetric w.r.t.  $\mathbf{y}$  and  $\mathbf{z}$ , and Theorem 1 applies. A relatively straightforward generalisation to any  $n$  and  $\ell$  is then possible. Alternative and general symmetrization techniques (e.g., Steiner symmetrization) could further refine the definition of  $h_S$  and consequent applicability of Champagnat's framework to general functions (Brock, 1995; Brock and Solynin, 2000). For the purpose of this work, however, we confine our approach to a simple application of boolean functions to construct our generalisation.

An important note is that, while the primary phenotypic space in  $\mathbb{R}^\ell$  fails in the case of incomplete or weighted graphs, its Sidon-extension guarantees the symmetry imposed by the interaction kernels. This is only possible because we are considering vector-wise (or, individual-based) symmetry w.r.t. non-focal traits, and the inclusion of a multiplicative boolean factor in the definition of a general  $h$ . Without this assumption, Champagnat's framework generally fails in asymmetric scenarios. We may then write a generalised version of Theorem 1.

### Theorem 3 (Symmetric asymmetry).

Assume the Axiom of Choice, and let

$$h_S(\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_n) = h(x, x_1, \dots, x_n) + S(\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_n) \quad (51)$$

for some symmetrizing function  $S$ , where  $\mathbf{x}_i \in \mathbb{R}^{\ell+1}$  is a Sidon-extension of vector  $x_i$  with unique parameter  $p_i$ ,  $\forall i$ . Then, for all  $x, x_1, \dots, x_n \in \mathbb{R}^\ell$  and  $n \geq 2$ , there exist functions  $f, g$  and  $S$  such that

$$f(x, \sum_{i=1}^n g(x - x_i)) = h_S(\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_n) \quad (52)$$

holds for any function  $h$ .

*Proof.* Let  $h : \mathbb{R}^{n+1} \times \mathbb{R}^\ell \rightarrow \mathbb{R}$  be any function. Take  $\{p_i\}$  to be a Sidon sequence (labelling traits) and consider the list of all distinct pairwise combinations of  $p_i$ -values. There will be  $\binom{n}{2}$  such pairs. For each pair  $(p_i, p_j)$ , we define a switch term that accounts for the difference between the original function and the function with  $x_i$  and  $x_j$  switched. If the original function is  $h(x, x_1, \dots, x_n)$ , the switch term for  $(x_i, x_j)$  maps  $x_i \rightarrow x_j$  and  $x_j \rightarrow x_i$ . By using Heaviside step functions  $H(p_i - p_j)$  for each pair, we determine whether or not to apply the switch term. Finally, summing up all these terms weighted by their respective Heaviside functions we guarantee the existence of a function  $S$  that symmetrizes  $h_S$ , for any  $h$ . The existence of  $f$  and  $g$  is then a direct consequence of Theorem 1. ■

## Discussion

In this work, we have explored the realms of logic and set theory to extrapolate an analysis of the symmetry of the interaction kernel definition within a well-established model of evolutionary dynamics (Champagnat et al., 2006).

By invoking the Axiom of Choice, and utilising a mathematical discourse grounded in Zorn's Lemma, we secured the existence of a Hamel basis. The bijection of this basis to the real numbers enabled us to uniquely determine the sum terms in the dependence of the interaction kernel. The inherently non-constructive nature of this approach further propelled our efforts to examine the symmetry argument across countable sets. This analysis ultimately culminated in two versions of our central theorems pertaining to the overall symmetry of interaction kernels in relation to non-focal traits within an individual's phenotype (Theorems 1 and 2).

This revelation led us to infer that the framework outlined in Champagnat et al. (2006), although limited to birth and death rate functions that maintain symmetry w.r.t. non-focal traits, can be adapted to accommodate any form of symmetry. Essentially, we illustrated that the event rates could indeed adopt any symmetric functions. Nevertheless, the feasibility of constructing such functions is heavily contingent on the trait values – failing in the real case but possible, in general, within any



countable subset of  $\mathbb{R}^\ell$ . In the latter scenario, we presented a practical exemplification in the case of nonlinear symmetry.

Ultimately, we ventured into an examination of the symmetry of evolutionary games on graphs. We analysed incomplete and weighted population dynamics graphs in cooperative games, as well as the case of general individual-based asymmetric interaction. We concluded that symmetry can be ensured, provided that individual labelling is incorporated as a component of the phenotype of different individuals (Theorem 3).

Our exploration into the subtleties of symmetry within interaction kernels and the application of set theory notably enhances the modelling and understanding of evolutionary dynamics. By uncovering the roles and limitations of symmetry in shaping interactions between individuals and phenotypic variations, we open a new pathway for a deeper and more nuanced understanding of biological systems. The set theory framework further contributes to the mathematical foundations of these models, allowing for more sophisticated and detailed interpretations. This research serves as a crucial stepping stone in bringing about more refined and predictive models of evolutionary processes, which will be instrumental in furthering our understanding of adaptive evolution and cooperative behaviours within complex biological networks.

#### ACKNOWLEDGEMENTS

We thank Prof. Nadia Sidorova and Dr. Rubén Pérez-Carrasco for their helpful discussions and comments that greatly improved the manuscript. This research was made possible by a UCL Studentship Award from the Department of Mathematics and the Leverhulme Trust Research Project Grant RPG-2022-028.

#### COMPETING INTERESTS

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

## References

- Charles Darwin, Alfred Russel Wallace, et al. Evolution by natural selection. *Evolution by Natural Selection*, 1958.
- Charles Darwin. *Charles Darwin's natural selection: being the second part of his big species book written from 1856 to 1858*. Cambridge University Press, 1987.
- Charles Darwin. *On the Origin of Species by Means of Natural Selection*. Murray, London, 1859. or the Preservation of Favored Races in the Struggle for Life.
- Lilian Al-Chueyr Pereira Martins. Weldon, pearson, bateson and the biometric-mendelian controversy: a dispute between evolutionists. *Filosofia Unisinos/Unisinos Journal of Philosophy*, 8(2):170–190, 2007.
- Bernard J Norton. The biometric defense of darwinism. *Journal of the History of Biology*, 6(2):283–316, 1973.
- Robert Olby. The dimensions of scientific controversy: the biometric—mendelian debate. *The British Journal for the History of Science*, 22(3):299–320, 1989.
- John Burdon Sanderson Haldane. The relation between density regulation and natural selection. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, 145(920):306–308, 1956.
- RA Fisher et al. The genetical theory of natural selection. *The Genetical Theory of Natural Selection*, (2nd Ed), 1958.
- JBS Haldane and SD Jayakar. Polymorphism due to selection depending on the composition of a population. *Journal of Genetics*, 58(3):318–323, 1963.
- Sewall Wright. The theory of gene frequencies. *Evolution and the Genetics of Populations*, 2:290–344, 1969.
- William D Hamilton. Selfish and spiteful behaviour in an evolutionary model. *Nature*, 228(5277):1218–1220, 1970.
- Motoo Kimura and Charles Athiel Harper. *An introduction to population genetics theory*. 1970.
- John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.
- David M Kreps. Nash equilibrium. In *Game Theory*, pages 167–177. Springer, 1989.
- Reinoud Joosten. Deterministic evolutionary dynamics: a unifying approach. *Journal of Evolutionary Economics*, 6(3):313–324, 1996.
- Karl P Hader. Stable polymorphisms in a selection model with mutation. *SIAM Journal on Applied Mathematics*, 41(1):1–7, 1981.
- George R Price et al. Selection and covariance. *Nature*, 227:520–521, 1970.
- George R Price. Extension of covariance selection mathematics. *Annals of Human Genetics*, 35(4):485–490, 1972.
- Karen M Page and Martin A Nowak. Unifying evolutionary dynamics. *Journal of Theoretical Biology*, 219(1):93–98, 2002.
- MA Nowak and K Sigmund. The evolution of reactive strategies in iterated games. *Acta Applicandae Math*, 20:247–265, 1990.
- Johan AJ Metz, Stefan AH Geritz, Géza Meszéna, Frans JA Jacobs, and Joost S Van Heerwaarden. *Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction*. WP-95-099, 1995.
- Ulf Dieckmann and Richard Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34(5):579–612, 1996.
- Josef Hofbauer, Karl Sigmund, et al. *Evolutionary games and population dynamics*. Cambridge University Press, 1998.
- Vito Volterra. *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. Società anonima tipografica "Leonardo da Vinci", 1926.
- Manfred Eigen, John McCaskill, and Peter Schuster. The molecular quasi-species. *Advances in Chemical Physics*, 75:149–263, 1989.
- Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical Population Biology*, 69(3):297–321, 2006.
- Paul A Gagniu. *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons, 2017.
- Georg Hamel. Eine basis aller zahlen und die un stetigen lösungen der funktionalgleichung:  $f(x+y) = f(x) + f(y)$ . *Mathematische Annalen*, 60(3):459–462, 1905.
- FB Jones. Measure and other properties of a hamel basis. *Bulletin of the American Mathematical Society*, 48(6):472–481, 1942.
- Zoltán Boros and Zoltán Daróczy. A composite functional equation with additive solutions. *Publicationes Mathematicae Debrecen*, 69(1-2):245–253, 2006.
- J Aczél and P Erdős. The non-existence of a hamel-basis and the general solution of cauchy's functional equation for nonnegative numbers. *Publicationes Mathematicae Debrecen*, 12:259–263, 1965.
- Ernst Zermelo. Beweis, daß jede menge wohlgeordnet werden kann. *Mathematische Annalen*, 59(4):514–516, 1904.
- Casimir Kuratowski. Une méthode d'élimination des nombres transfinis des raisonnements mathématiques. *Fundamenta Mathematicae*, 3(1):76–108, 1922.
- Max Zorn. A remark on method in transfinite algebra. *Bulletin of the American Mathematical Society*, 41(10):667–670, 1935.
- Ernst Zermelo. Untersuchungen über die grundlagen der mengenlehre. i. *Mathematische Annalen*, 65(2):261–281, 1908.
- Abraham Adolf Fraenkel, Yehoshua Bar-Hillel, and Azriel Levy. *Foundations of Set Theory*. Elsevier, 1973.
- Brent Yorgey. *Combinatorial species and labelled structures*. PhD Thesis, 2014.
- James D Halpern. Bases in vector spaces and the axiom of choice. *Proceedings of the American Mathematical Society*, 17(3):670–673, 1966.
- MN Bleicher. Some theorems on vector spaces and the axiom of choice. *Journal of Symbolic Logic*, 32(2), 1967.
- Rina Zazkis. Representing numbers: Prime and irrational. *International Journal of Mathematical Education in Science and Technology*, 36(2-3):207–217, 2005.
- Andreas Obersteiner and Veronika Hofreiter. Do we have a sense for irrational numbers? *Journal of Numerical Cognition*, 2(3):170–189, 2017.
- Lars Kristiansen. On subrecursive representability of irrational numbers. *Computability*, 6(3):249–276, 2017.
- Ivan Georgiev, Lars Kristiansen, and Frank Stephan. Computable irrational numbers with representations of surprising complexity. *Annals of Pure and Applied Logic*, 172(2):102893, 2021.
- Matthew Szudzik. An elegant pairing function. In *Wolfram Research (ed.) Special NKS 2006 Wolfram Science Conference*, pages 1–12, 2006.
- Meri Lisi. Some remarks on the cantor pairing function. *Le Matematiche*, 62(1):55–65, 2007.
- Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502, 2006.
- Benjamin Allen, Martin A Nowak, and Ulf Dieckmann. Adaptive dynamics with interaction structure. *The American Naturalist*, 181(6):E139–E163, 2013.
- Benjamin Allen, Gabor Lippner, Yu-Ting Chen, Babak Fotouhi, Naghmeh Momeni, Shing-Tung Yau, and Martin A Nowak. Evolutionary dynamics on any population structure. *Nature*, 544(7649):227, 2017.
- William Poundstone. *Prisoner's Dilemma/John Von Neumann, game theory and the puzzle of the bomb*. Anchor, 1993.
- JD Wassersug and RJ Wassersug. Fitness fallacies. *Natural History*, 95(3):34–37, 1986.
- John Maynard Smith et al. *Evolutionary Genetics*. Oxford University Press, 1989.
- Jason Bertram and Joanna Masel. Density-dependent selection and the limits of relative fitness. *Theoretical Population Biology*, 129:81–92, 2019.
- Paul Erdos and Pál Turán. On a problem of sidon in additive number theory, and on some related problems. *Journal of the London Mathematical Society*, 16(4):212–215, 1941.
- Kevin O'Bryant. A complete annotated bibliography of work related to sidon sequences. *arXiv preprint math/0407117*, 2004.
- Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.
- Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- Friedemann Brock. Continuous steiner-symmetrization. *Mathematische Nachrichten*, 172(1):25–48, 1995.
- Friedemann Brock and Alexander Solynin. An approach to symmetrization via polarization. *Transactions of the American Mathematical Society*, 352(4):1759–1796, 2000.

# Supplementary Information

## A note on Zorn's Lemma

In this section, we sketch a proof that  $\mathbb{R}$  has a basis, under the Axiom of Choice and following Zorn's Lemma, and discuss the existence of the critical bijection  $g$  between  $\mathbb{R}$  and its Hamel basis. See [Halpern \(1966\)](#); [Bleicher \(1967\)](#) for further reading.

First, we recall some definitions from set theory ([Harzheim, 2005](#))

- A *partially ordered* set is a set  $X$  together with an ordering  $\leq$  of the elements of  $X$  that is transitive (if  $x \leq y$  and  $y \leq z$  then  $x \leq z$ ) and antisymmetric (if  $x \leq y$  and  $y \leq x$  then  $x = y$ );
- A *totally ordered* subset of  $X$  is a set  $Y \subseteq X$  such that,  $\forall y, z \in Y$ , either  $y \leq z$  or  $z \leq y$ ;
- A *chain*  $Y$  in a partially ordered set  $X$  is a totally ordered subset of  $X$ ;
- An *upper bound* for a subset  $Y$  of a partially ordered set  $X$  is an element  $u$  such that  $y \leq u$  for every  $y \in Y$ ;
- Finally, a *maximal element* in a partially ordered set  $X$  is an element  $x_0$  such that the only element  $x \in X$  satisfying  $x_0 \leq x$  is  $x_0$  itself.

The main result follows.

**Theorem 4** ( $\mathbb{R}$  has a basis). *Under the Axiom of Choice, there exists a basis of  $\mathbb{R}$ .*

*Proof.* (Sketch) Recall that  $\mathcal{B}$  is defined as the collection of real numbers such that every real number is a rational linear combination of numbers  $\beta_i \in \mathcal{B}$  in precisely one way, that is, every real number  $\alpha$  has a unique representation of the form

$$\alpha = \sum_{i=1}^m q_i \beta_i, \quad (S1)$$

where  $q_i$  is rational and  $m$  depends on  $\alpha$ . It suffices to show that  $\mathcal{B}$  exists.

With  $\mathcal{B}$  as a candidate  $\mathbb{Q}$ -basis of  $\mathbb{R}$ , the objects we are interested in are subsets of  $\mathbb{R}$  that are linearly independent over  $\mathbb{Q}$ . First, we note that a maximal linearly independent subset of  $\mathbb{R}$ , where by "maximal" we mean that the set is not contained in any larger linearly independent set, spans the whole of  $\mathbb{R}$ . If it did not, we could just pick an element that did not belong to its linear span and add it to the linearly independent set, redefining maximality. Although seemingly an interminable process, Zorn's lemma tells us that it ends. Thus, we are looking at the set of all linearly independent subsets of  $\mathbb{R}$ , with the partial order  $\subset$ .

In order to apply Zorn's Lemma, we then need to check that every chain has an upper bound. Imagine that we have a collection  $Y$  of linearly independent subsets of  $\mathbb{R}$  and that, for any two of those subsets, one is contained in the other. By definition, the upper bound of such a chain is a set that contains all the sets in  $Y$ , so it has to contain their union. If we take the upper bound as the union, it should be linearly independent. Indeed, if  $\alpha_1, \dots, \alpha_n$  belong to the union, then each  $\alpha_i$  belongs to some linearly independent set  $\beta_i \in Y$ . Since  $Y$  is a chain, one of these sets  $\beta_i$  contains all the others. If that is  $\beta_j$ , then the linear independence of  $\beta_j$  implies that no non-trivial linear combination of  $\alpha_1, \dots, \alpha_n$  can be zero, which proves that the union of the sets in  $Y$  is linearly independent. Therefore, by Zorn's Lemma, there is a maximal linearly independent set. Such a set forms a basis for  $\mathbb{R}$ . ■

Now, in order to define  $g$  in Eq. (3) as the bijection between  $\mathbb{R}$  and its Hamel basis, it is important to show that such a basis has the cardinality of the continuum. Let the cardinality of some Hamel basis be  $\kappa$ . Since the cardinality of the set of coefficients is  $\aleph_0$  (cardinality of  $\mathbb{Q}$ ), we have  $\aleph_0 \kappa$  elements using each basis element of the form  $q_1 \beta_1$ ,  $\aleph_0^2 \kappa^2$  elements using two basis elements  $q_1 \beta_1 + q_2 \beta_2$ , and so on. Hence, the span of the basis has cardinality given by

$$\aleph_0 \kappa + \aleph_0^2 \kappa^2 + \aleph_0^3 \kappa^3 + \dots = \kappa, \quad (S2)$$

where we used the fact that  $\aleph_0$  is the smallest infinite cardinal (the product of two infinite cardinals is equal to whichever is bigger, thus  $\aleph_0 \kappa = \kappa$ ). Since the span is  $\mathbb{R}$ , we have  $\kappa = 2^{\aleph_0}$  and the bijection exists.

## References

- James D Halpern. Bases in vector spaces and the axiom of choice. *Proceedings of the American Mathematical Society*, 17(3):670–673, 1966.  
MN Bleicher. Some theorems on vector spaces and the axiom of choice. *Journal of Symbolic Logic*, 32(2), 1967.  
Egbert Harzheim. *Ordered sets*, volume 7. Springer Science & Business Media, 2005.