UNIVERSITY OF Southampton

University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: Christoph Thies (2022) "Causal models of multilevel selection", University of Southampton, Faculty of Engineering and Physical Sciences, School of Electronics and Computer Science, PhD Thesis.

**University of Southampton**

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

# Causal models of multilevel selection

*by*
Christoph Thies

February 2022

*A thesis for the degree of*
*Doctor of Philosophy*

# University of Southampton

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science



## Abstract

Thesis for the degree of Doctor of Philosophy
**Causal models of multilevel selection**
**Christoph Thies**

In social evolution, fitness of an individual depends not only on the phenotype of the individual itself but also on the phenotype of its social environment. When measuring the strength of selection in empirical data, this leads to the question of how selection is assigned to the individual and the population level. Two methods for carrying out this assignment have been primarily discussed in the literature. The multilevel Price approach uses a multilevel expansion of the Price equation to express the components of selection on the two levels in terms of covariances. Contextual analysis, on the other hand, uses linear regression to assign fitness components to the individual and the population level. However, the two methods generally do not agree in their results, and discussions which one is preferable have been inconclusive. In this thesis, I argue that the root of the problem lies in viewing the two approaches as correlational. While both are equally valid as correlational models, underlying an empirical scenario is a causal process that may or may not match the given approach. To find the correct approach I therefore suggest to regard contextual analysis and the Price approach as causal models, i.e., as processes that generate the given data. In order to implement the approaches as process models I view the process of selection as a transformation from metapopulations to populations. I show that transformations of this kind as well as other aspects of biological systems can be expressed in terms of monoidal categories. More precisely, probability monads can be used to capture essential features of metatpopulations and allow a convenient graphical representation. Using this formalism, I construct process models of multilevel selection that correspond to the multilevel Price

equation and contextual analysis, thus endowing the correlational models with a causal structure. The parts that make up the models can be rearranged, composed, and refined in a consistent manner, and the diagrammatic formalism of string diagrams allows intuitive manipulation of the models. Finally, in the last chapter I discuss the two approaches for the case where fitness on the two levels combines additively. Using an empirical example, I show how the correct approach can be determined by intervening on the empirical system.

# Contents

# Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: Thies and Watson (2021)

Signed:..... Date:.....

# Acknowledgements

**Thank you:**

**Richard Watson** For sharing insights and believing in me.

**Miguel Brun-Usan, Jamie Caldwell, Laura Mears, Freddie Nash, David Prosser, and Alfredo Rago** For learning together and invaluable support.

**Markus Brede, Patrick Doncaster, and other staff and students at University of Southampton** For help, interesting discussions, and an open and supportive atmosphere.

**Samir Okasha, Will Ratcliff, and Michael Wade** For invaluable feedback on earlier drafts of parts of the thesis.

**Tobias Fritz, Nathaniel Virgo, and others on the category theory Zulip** For comments and advice regarding category theory.

**Gillian Brown, Kevin Laland, and Amanda Seed** For giving me confidence in myself.

**Natalie Feiner, Heikki Helanterä, Tobias Uller, and others in the EES research community** For support and kindness.

**Hannah Brückmann, Riccarda Cappeller, and other deskmates at 'Der Salon' in Neukölln** For helpful conversations and making me feel welcome.

# Preface

The central claim of my thesis is that statistical equations of population genetics that describe multilevel selection (Heisler and Damuth, 1987; Okasha, 2006; Gardner, 2015) can be represented in terms of mathematical structures that allow causal modelling on a high level of abstraction (Fong, 2013; Pearl, Glymour, and Jewell, 2016). The correspondence yields a representation of the associated biological observables and processes as the spaces and functions of *symmetric monoidal categories* (Jacobs and Zanasi, 2018; Schmid, Selby, and Spekkens, 2020). The models constructed in this way are intuitive, yet mathematically rigorous and can be subjected to structured logical or statistical inference (Jacobs, 2019). They are useful for the design as well as for the computer-aided analysis of experiments (Patterson, 2020).

From a mathematical point of view, the thesis is delivered when its arguments trace a valid correspondence. However, the use of categories in multilevel selection theory is unusual[1] and therefore requires a perhaps unusual amount of patience on the side of the reader. I see the value of this thesis not in the reported mathematical facts themselves[2] but in the view on evolution they reveal. Therefore, it is my goal to make this view accessible to the reader rather than to provide a complete and rigorous mathematical account of the ideas expressed in this thesis.

In order to keep the picture simple, also on the biological side all but the most fundamental features have been left out. Biological examples and descriptions are therefore gross simplifications presented in order to make the abstract concept realised in the mathematics tangible. This is not to ignore biological complexity but to abstract from it so that the structure can be described more easily. These simplifications may make it seem like the thesis is not doing justice to population genetics – the field of mathe-

[1] But not rare in theoretical biology, see for instance Ehresmann and Vanbremeersch (2007), Tuyéras (2017), and Leinster (2021) for applications of category theory to biological and evolutionary phenomena.

[2] The mathematical structures employed as well as the operations they are subjected to in this thesis have a long history in category theory and other fields, see, for instance, Giry (1982), Riehl (2017), and Jacobs and Zanasi (2018). This manuscript contains no 'new' mathematics.

matical biology that formalises genetical evolution and underlies conceptualisations such as the Price equation and contextual analysis. Population genetics and related research in mathematical biology have advanced far from the traditional and basic parts that feature in this manuscript. However, I do not attempt to add any knowledge to the field of population genetics. Instead, I intend to add a point of view on the fundamental evolutionary processes that are described by population genetics. Some perspectives may inevitably depart from stances that are taken for granted by experts in the field of population genetics. Fortunately, the unambiguous language of category theory allows scrutinising the premises and consequences of arguments to any level of mathematical detail, and hence the reflections of population genetic equations in monoidal categories described in this thesis can be challenged and adapted in a rigorous manner.

More generally, it is my goal to be as explicit and precise as possible when using concepts such as 'selection' or 'replication'. Since this requires a clearly communicated picture of these concepts, many statements take the form of definitions of concepts that have been discussed and defined multiple times across the literature. Distilling a complicated and possibly contentious concept into a brief and formal definition cannot do justice to all its aspects. It is my goal to make *some* aspects accessible to formal assessment through the mathematical tools described in this thesis.

Finally, multilevel selection itself is within a circle of ideas that provoked often hot and extended debates involving the role of statistics and causality in explanations of evolutionary processes (for instance Nowak, Tarnita, and Wilson (2010), Abbot et al. (2011), Gardner, West, and Wild (2011), and Allen, Nowak, and Wilson (2013)). In light of this tradition, having to defend the assumptions, simplifications, and omissions in this thesis is daunting. However, I regard the reconciliation the results of this thesis offer with respect to the measurement of multilevel selection as a point in favour of a causal stance (Okasha, 2015; Pearl and Mackenzie, 2018; Otsuka, 2019). Making this stance explicit by putting it centre stage – as in the functional representation of multilevel selection developed in this thesis – may facilitate a mathematically meaningful debate on the role of causality in multilevel selection theory.

**Note:** Some of the material in this thesis has been posted on a public forum for discussing category theory. An archive of this forum is available under the following link: `https://mattecapu.github.io/ct-zulip-archive/`. I discussed this work in topic What is normalisation? and in topic Categorical Radon-Nikodym.

# 1 Selection and the ladder of causation

## 1.1 Introduction

Multilevel selection theory (MLS) is a conceptual approach for understanding the evolution of biological systems that span two or more levels of organisation. It is based on the idea that properties of populations of individuals may affect natural selection so that multiple populations that vary phenotypically form a metapopulation on which population selection acts (Damuth and Heisler, 1988).[1] Instantiating Darwin's principle of natural selection two or more times in this manner has been shown to integrate with traditional Darwinian selection on the level of the individual (Frank, 1998; Rice, 2004; Okasha, 2006; Gardner, 2015). In social evolution (Débarre, Hauert, and Doebeli, 2014), evolutionary transitions in individuality (Clarke, 2016), the evolution of holobionts and mutualisms (Lloyd and Wade, 2019), and in coevolution generally (Bapteste and Papale, 2020), multilevel selection frequently plays a role in explaining the evolution of higher level function.

However, there is some debate about the interpretation of MLS, more precisely about the relation between causal explanations of multilevel selection processes and the statistical analysis of observations of such processes (Okasha, 2006; Gardner, 2015; Goodnight, 2015; Otsuka, 2016a, 2019; Huneman, 2021). In particular, there is no agreement on which measurements are considered evidence for an instance of multilevel selection in an evolutionary scenario. For this, a statistical structure is required that, given observations, can be used to judge the claim that MLS is effective (i.e., that natural selection occurs on more than one level) in a given instance. In the theoretical literature on MLS, two such statistical structures have been primarily discussed: the multilevel Price equation and contextual analysis. While both approaches are formally valid, they may disagree in their verdict on whether MLS is in effect, in which case one of the two must clash with intuition about the system (Okasha, 2006).

In the case of one level – that of the individual – the general *Price equation* defines

[1] Damuth and Heisler (1988) refer to individuals in groups where we refer to individuals in populations.

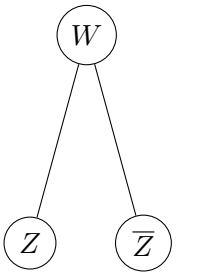
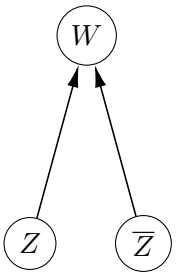
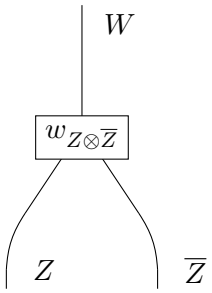
the distinction between selection with respect to a phenotype $Z$ and transmission, and thereby selection itself (Gardner, 2020). In order to instantiate the Price equation for a collection of observations of a system so that the strength of selection can be inferred, the character under selection along with unit and 'arena' of selection are specified (Gardner, 2015). In single-level systems, the latter two are realised by the genetic individual and the population it is part of, that is a collection of individuals within the same selective environment. A system with two levels, however, by definition comprises two units of selection, each with their arena of selection and phenotype under selection. As we will see in the course of this thesis, the essential difference between the Price approach to multilevel selection (PA) and contextual analysis (CA) is their respective choices of arena of selection for the individual level. The Price approach is based on the assumption that individual selection acts within populations, such that its total effect on the metapopulation is analogous to that of transmission bias in the single-level Price equation. In this view, the units under population selection themselves – characterised by the population phenotype $\overline{Z}$ – change during reproduction (more appropriately imagined as *persistence* for populations (Bapteste and Papale, 2020; Papale, 2020)). This change of composition within each population occurs by differential reproduction of the individuals in an environment created by the population phenotype $\overline{Z}$ that entails selection with respect to the individual phenotype $Z$. The arena of selection for the individual level in contextual analysis as discussed in Okasha (2006, 2015), on the other hand, is given by the population that comprises all individuals in the metapopulation. A complete episode of multilevel selection in this view is population selection within a metapopulation together with individual selection that acts across the metapopulation and with no regard for the population structure.

When the Price approach and contextual analysis are considered as statistical tools for estimating parameters based on observations of multilevel selection, it is not clear which of the two is preferable (Okasha, 2006). On the first rung of Pearl's ladder of causation (Table 1.1) (Pearl and Mackenzie, 2018; Bareinboim et al., 2020) this question cannot be settled, as both are valid operations that may be applied to associational data recorded during an experiment. However, the description in the previous paragraph holds that properties of the population an individual is part of have causal effects on within-population selection in the Price approach but not in contextual analysis.[2]

This commitment to a causal account of the process of multilevel selection has the

[2] We refer here to the 'basic' form of contextual analysis discussed in Okasha (2006), see Section "The multilevel Price equation and contextual analysis".

**Table 1.1:** Pearl's ladder of causation (Bareinboim et al., 2020).

| | | |
|---|---|---|
| Observing | Correlation: Given correlation between observables (these appear as linkages between variables measured on the system), the first level allows introducing corresponding associations between the formal representatives of these variables as in the diagram on the right. This diagram represents the assumption that $Z$ and $\overline{Z}$ (individual and population phenotype) are both associated with $W$ (total fitness of an individual) but not directly associated with each other (over and above the fact that the individual phenotypes $Z$ in a population determine $\overline{Z}$). Probabilistically, this means that $Z$ and $\overline{Z}$ are conditionally independent given $W$ (see Section 1.3.4 in Pearl, Glymour, and Jewell (2016)). Observations allow the construction of contingency tables. The Price approach to MLS and contextual analysis may be regarded as operating on this level to yield estimates of covariances and linear regression parameters, resp., as discussed in Section 1.4. The parameters may be used, e.g., to predict allele frequency dynamics. |  |
| Intervening | Causality: On the second level, causality is assumed in the sense that the value of one variable has an effect on the observed value of another variable (but not neccessarily the other way around). This directionality lets us decorate the edges in diagrams on the first level with tips so that arrows point in the direction of causation (the diagram on the first rung is the *skeleton* of the decorated diagram to the right (Pearl, 2009, Section 1.2; Fong, 2013, Section 3.2)). The resulting diagrams on the second level encode assumptions about the effects of interventions. |  |
| Reasoning | Representability: Finally on the third level, representable processes can be expressed in terms of functions between the variables involved. Varying the input to these functions as well as the functions themselves allows counterfactual reasoning. The functional representation of processes is the focus of Section 1.6 in this chapter and of Chapter 2. |  |

benefit that directed acyclic graphs (DAG) and the formalism associated with the second rung of the ladder of causation may be used to represent models corresponding to the Price approach and contextual analysis. Differences between the two approaches are then reflected in the associated directed acyclic graphs. These graphs are more expressive than the associational models on the first rung because causality adds directedness to associations. In particular, they admit the do operator for reasoning about interventions on the system that fix variables to a certain value without affecting other (parent) variables (Pearl, Glymour, and Jewell, 2016). Considered as models on the second rung of the ladder, the Price approach and contextual analysis rest on different assumptions that are encoded in the associated causal graphs. In a given biological scenario, none, one, or both of these graphs may be causally adequate (cf. Okasha (2015)). Where both are correct, no inconsistencies occur as in this case no individual selection but only population selection is acting on the metapopulation. In Section "A functional representation of multilevel selection" we will see that interventions formalised by the do operator allow a distinction between the two directed acyclic graphs (also referred to as *causal graphs* or *causal structures*) that represent the Price approach and contextual analysis, and if the assumptions underlying the causal graph hold, these interventions must be reflected in empirical measurements. Thus, commitment to a causal interpretation of the two conflicting views allows designing experiments that specifically test which of the models is appropriate.

While causality in the sense of directed association is the central premise for the second rung of the ladder of causation, it is not sufficient for the third. *Counterfactual reasoning* requires a functional representation of the processes that create the causal relationships between the variables. A functional representation of a causal graph is a set of functions that determine the values of the dependent variables in terms of the independent variables, like $w_{Z\otimes\overline{Z}}$ determines the value of individual fitness $W$ in terms of individual phenotype and population phenotype on the top level in Table 1.1. Contextual analysis and, more generally, structural causal models provide such a representation (Heisler and Damuth, 1987; Pearl, Glymour, and Jewell, 2016). This representation gives meaning to varying the variables $Z, \overline{Z}$, and $W$ in the right column of Table 1.1 and allows answering questions like "What would have been the outcome of selection (i.e., the number of offspring) for a given individual with phenotypic value $z\colon Z$ within a population with phenotypic value $\overline{z}\colon \overline{Z}$ if that individual had expressed the phenotypic value $z'\colon Z$ instead?". Given the functional representation $w_{Z\otimes\overline{Z}}$, an answer to this question is simply

the fitness $w_{Z\otimes\overline{Z}}(z', \overline{z})\colon W$.[3]

The functional representation that characterises the third rung of the ladder of causation allows counterfactual reasoning in yet another way, different from varying the variables to assess the outcome under counterfactual conditions. When a process is assumed to be faithfully represented by a function that computes the outcome of the process for any input of conditions, then not only unobserved combinations of variables can be counterfactually imagined. Like their input, the functions themselves become objects within a space of multiple objects over which the assumptions may vary. This switch of perspective from functions as fixed objects to functions as varying objects corresponds to the distinction between *regression coefficients* that approximate a fixed statistical association and *structural coefficients* that parameterise a meaningfully varying family of functions (Pearl, Glymour, and Jewell, 2016; Section 3.8.1).

In multilevel selection specifically, varying functions that determine the outcome of selection corresponds to varying the selection pressures the metapopulation is subject to. For the study of evolutionary transitions in individuality, which are characterised by changes in the reproductive organisation of organisms, a functional viewpoint seems indispensable (Watson and Thies, 2019). Along with the arguments that aim to clarify the relationship between the two approaches to multilevel selection we describe mathematical structures (in particular monoidal categories and probability monads) that facilitate the construction of statistical models of selection. An essential feature of the models constructed in this way is their focus on composition that allows – and demands – a complete specification of all processes and interactions involved.

## 1.2 Individual selection and the Price equation

The evolutionary models discussed in this thesis concern the action of natural selection on a population of individuals – possibly itself situated within a metapopulation of populations. The mathematical structures we use to model processes associated with this action are those that constitute the formal frame of the *Price equation* and the theory of natural selection this equation describes (Price, 1970, 1995; Gardner, 2020). The Price equation captures the action of selection in a *covariance form*

$$\Delta_S \mathrm{E}_{i\colon I}(z_i) = \mathrm{Cov}_{i\colon I}(\nu_i, z_i) \tag{1.1}$$

[3] We can give this answer because we assume the fitness function $w_{Z\otimes\overline{Z}}$ to be defined on the product $Z \otimes \overline{Z}$ (see Section "Monoidal categories" for the definition of the product $Z \otimes \overline{Z}$) so that any pair $(z, \overline{z})\colon Z \times \overline{Z}$ may be assigned a fitness value $w_{Z\otimes\overline{Z}}(z, \overline{z})$.

(equation (1) in Price (1970), see §2 in Grafen (2000) for a brief overview of the full Price equation and the concept of relative fitness involved). In this equation, the finite set $I$ represents a population of individuals $i : I$ with phenotypic values $z_i : Z$. $\mathrm{E}_{i:\,I}(z_i)$ denotes the mean value of the phenotype $Z$ among the individuals indexed by $i : I$. Moreover, each individual has assigned a relative fitness $\nu_i : \mathbb{R}_{\geq 0}$ that represents the rate of change in representation ('growth rate', if this representation increases) from the parent individual in the parent population to its offspring in the offspring population[4]. The numerical covariance[5] of relative fitness with the associated phenotypic values $z_i : Z$, taken over the population $I$, is written as $\mathrm{Cov}_{i:\,I}(\nu_i, z_i)$ in equation (1.1). Finally, $\Delta_S \mathrm{E}_{i:\,I}(z_i)$ denotes the change in mean phenotypic value from that of the population $I$ to that of a population $J$ comprising the offspring individuals so that $J$ is the outcome of the action of selection on $I$.

Equation (1.1) describes natural selection in the absence of mutation, recombination, phenotypic plasticity, and other mechanisms that may entail phenotypic changes from parent to offspring (*property changes* in Price (1995)). Effects of this kind are captured by the second term on the right hand side of the full Price equation

$$\Delta \mathrm{E}_{i:\,I}(z_i) = \mathrm{Cov}(\nu_i, z_i) + \mathrm{E}_{i:\,I}(\nu_i \Delta z_i) \tag{1.2}$$

(equation (4) in Price (1970) and equation (2.1) in Gardner (2020)), where $\Delta z_i$ denotes the difference between the character value of parent $i$ and the average character value of its offspring (Okasha, 2006; Gardner, 2020). The second summand in equation (1.2) is often referred to as *transmission* term of the Price equation because it refers to the effect of changes during reproduction of an individual in its offspring, that is during transmission from parent to offspring generation. Following Gardner (2020), we consider the Price equation without transmission term (1.1) as population genetic *definition of selection*.

The Price equation expresses a relation between variables that refer to biological sys-

[4] To give an example of relative fitness, if an individual in a population of 3 individuals has assigned 2 offspring in an offspring population of 3, then its relative fitness is 2 because the representation, or frequency, in the population doubled from parent to offspring. If in the previous example the magnitude of the offspring population is 6 instead of 3, then the relative fitness of the individual is 1 because the representation remains unchanged from $\frac{1}{3}$ to $\frac{2}{6} = \frac{1}{3}$.

[5] The covariance of a set of value pairs $(x_i, y_i) : (\mathbb{R} \times \mathbb{R})^{|I|}$ is given by

$$\mathrm{Cov_{i:I}}(x_\mathrm{i}, y_\mathrm{i}) = \mathrm{E_{i:I}}(x_\mathrm{i} y_\mathrm{i}) - \mathrm{E_{i:I}}(x_\mathrm{i})\mathrm{E_{i:I}}(y_\mathrm{i}).$$

.

tems[6]. The variables $z_i$ represent outcomes of phenotype measurements that may be taken in population $I$ in advance of the process referred to by the equation. The variables $\Delta z_i$ implicitly refer to additional phenotype measurements on the individuals' offspring in population $J$. Finally, the variables $\nu_i$ encode the change from the probability of randomly drawing an individual in $I$ to drawing one of that individual's offspring in $J$. The Price equation describes a linkage between these variables that follows from their definition and may therefore be regarded a tautology (Frank, 2012). Conceptually, however, the Price equation formally separates evolutionary progression under natural selection from one generation to the next into the two components selection and transmission.

If, in an asexually reproducing species, we are concerned with the action of selection alone, we may assume that offspring is an exact copy of its single parent in a selection process represented by equation (1.1). The transformation of population $I$ to population $I'$ is generally assumed to follow a *law*[7]

$$\mathsf{interact}_Z : Z \to \mathbb{N} \tag{1.3}$$

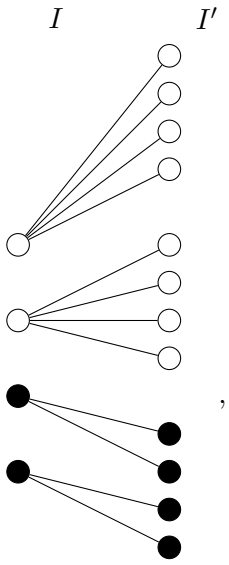
that determines the (expected) number of offspring of an individual according to its phenotypic value (Wade and Kalisz, 1990), more precisely the "average number of successful gametes" (Grafen (2000), page 1223). The definition of the function (1.3) is motivated by the idea that "selection acts through offspring number" (Levin and Grafen, 2019, page 1069). We assume this number to be a whole number so that each circle represents one individual in the diagrams below.[8] This means that each individual $i\colon I$ in the parent

[6] Price (1995) suggested wider significance of 'selection mathematics' but we consider only natural and experimental selection of biological systems.

[7] We attach no meaning to the term *law* other than that a law determines the outcome of a class of processes, here episodes of selection of any population of phenotypes in $Z$. The name $\mathsf{interact}_Z$ is explained in Section "Interactors and replicators".

[8] This assumption will be relaxed later on and $\mathsf{interact}$ is considered a function $\mathsf{interact} : Z \to \mathbb{R}$.

generation is replaced by $\mathsf{interact}_Z z_i : \mathbb{N}$ copies of itself in the offspring population $I'$,




, (1.4)

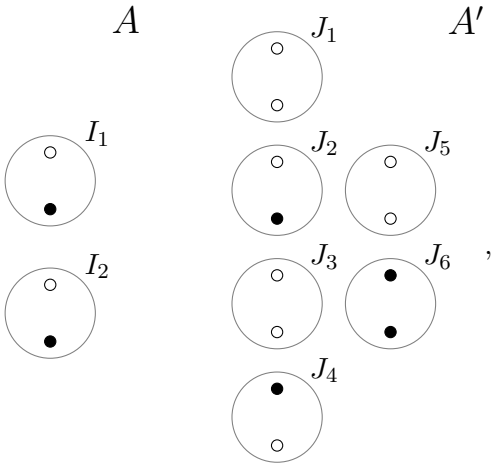
where $\mathsf{interact}_Z\bigcirc = 4 : \mathbb{N}$ and $\mathsf{interact}_Z\bullet = 2 : \mathbb{N}$ (see below for the notation). Since selection is realised in interaction of the phenotype with the environment, this law depends on the environment. We do not indicate this dependence because the selective environment is assumed to remain unchanged. Absolute fitness, that is the number of offspring of an individual ($\mathsf{interact}_Z z_i : \mathbb{N}$ from definition (1.3) above), and relative fitness ($\nu_i : [0,1] \subset \mathbb{R}$ in equation (1.1) above), are assumed deterministic (cf. Pearl (2018), Section 2)[9].

The Price equation describes the transformation of a parent population to its offspring population in terms of the change in mean phenotypic value the transformation entails. However, the mean phenotypic value may be one among other properties of the population $I$ and its offspring population $I'$. To give an example, let the empty circles in diagram (1.4) be encoded with the phenotypic value 0 and the filled circles with 1. Then the Price equation traces the change of mean phenotypic value from $\frac{1}{2}$ to $\frac{1}{3}$. If populations are represented as frequencies, or distributions, of values $\bigcirc$ and $\bullet$, so that population $I$ in diagram (1.4) is assigned $\frac{2}{4}|\bigcirc\rangle + \frac{2}{4}|\bullet\rangle = \frac{1}{2}|\bigcirc\rangle + \frac{1}{2}|\bullet\rangle$ and $I'$ is assigned $\frac{8}{12}|\bigcirc\rangle + \frac{4}{12}|\bullet\rangle = \frac{2}{3}|\bigcirc\rangle + \frac{1}{3}|\bullet\rangle$[10], then the process is fully described by the change in mean phenotypic value, that is the Price equation with respect to the phenotype $Z$.

[9] In the end of the thesis, we will see how stochasticity with respect to phenotype and fitness may be implemented in the models discussed.

[10] The intuitively meaningful representation of populations, or distributions, of a discrete phenotype such as $Z = \{\bullet, \bigcirc\}$ in terms of symbolic sums $p_\bullet|\bullet\rangle + p_\bigcirc|\bigcirc\rangle$ with $p_\bullet, p_\bigcirc : [0,1], p_\bullet + p_\bigcirc = 1$ is detailed in Jacobs (2019), Section 1.5, and referred to as *ket notation*.

If, however, for example, the populations $I$ and $I'$ in (1.4) are characterised by organisation into subpopulations (such that they are better denoted as metapopulations $A$ and $A'$)




, (1.5)

where $A$ and $A'$ denote metapopulations with subpopulations $I_1, I_2 : A$ and $J_1, \ldots, J_6 : A'$, then equation (1.1) does not describe the process completely because the transformation $\frac{1}{2}|\circ\rangle + \frac{1}{2}|\bullet\rangle \mapsto \frac{2}{3}|\circ\rangle + \frac{1}{3}|\bullet\rangle$ may be realised by multiple offspring metapopulations on the right hand side of (1.5). This issue is sometimes referred to as *dynamic insufficiency* of the Price equation (Frank, 2012; van Veelen et al., 2012). The term may suggest that the Price equation is 'insufficient' to describe the evolutionary dynamics when such organisation or other additional properties are involved, but that is not the case. The Price equation may be applied to any character, including contextual characters[11], that can be observed on individuals (Damuth and Heisler, 1988), and changes in means of higher moments of the phenotypic distribution yield the complete dynamics (Gardner, West, and Barton, 2007). Nevertheless, the Price equation describes the transformation in terms of changes in univariate means of phenotypic values in the population. As we will see below, our approach, in contrast, is to give a functional description of the transformation, or process, itself. The components of the transformation, such as the change in mean phenotypic value as in equation (1.1), can then be derived from the functional description.

To avoid confusion, it should be noted that the Price equation features in two roles in this thesis. First, we take the Price equation as *definition of selection* in the sense of Gardner (2020). The concept of selection that is formalised in this thesis aims to be the

[11] Contextual characters of an individual are properties that depend on the context the individual is situated in. For example, properties of a population are contextual properties of the individuals it comprises, see Table 1 in Damuth and Heisler (1988).

concept of selection formalised in the Price equation. Second, we use the (multilevel) Price equation as *model of multilevel selection*. To do so (below in Section 2), we apply the structure of the Price equation without transmission bias (1.1) recursively to the higher as well as to the lower level of the system subject to multilevel selection. How the effects of the two levels on individual fitness interact is the essential difference between the Price approach to multilevel selection (**PA**) and contextextual analysis (**CA**). The distinction between CA and PA is the problem that motivated this thesis and is discussed in the next section (see also Section "The multilevel Price equation and contextual analysis").

## 1.3 Contextual analysis and the Price approach

When individual traits have effects on other individuals, individual fitness depends not only on self but also on the social environment, i.e., interaction partners. Social evolution theory deals with this problem by regarding the social environment as an external factor that, together with direct fitness effects of a trait, determines evolutionary dynamics with respect to selection (Frank, 1998). By assuming a certain correlation between trait value of an individual and average trait value of its social environment, e.g., through relatedness, Hamilton's rule can be formulated and answers the question of whether a trait with direct and indirect effects increases or decreases in frequency given the organisation of the population, that is the corresponding parameter of relatedness $r$ (Frank, 1997).

Multilevel selection theory (MLS) posits the social environment as a unit, that is the group or population an individual is situated in, that can be subject to selection acting at a level above that of individuals (Wilson, 1975; Wade, 1976, 1978; Uyenoyama and Feldman, 1980; Wilson and Sober, 1989). The theory thus promotes the concept of a population from a mere collection of individuals targeted by similar selection pressures to a unit that has a causal role in the selection process. More precisely, MLS theory understands a population as a unit whose interaction with the selective environment – through properties of the population as a whole – causally affects the fitness of its individual subunits (Wade and Kalisz, 1990). This means that individual fitness is a composite quantity determined by two factors: the individual effect of the trait and an effect on the population that an individual is a part of, and via this population effect, on the individual itself. The MLS view is not in opposition with *kin selection theory* (Hamilton, 1964) but merely highlights that selection at the population level may be part of a causal mechanism resulting in individual fitness differences and should be part

of a causal account of the system.

The distinction between individual effects and population effects of individual traits presents MLS with a problem: how can the presence of a population effect be detected empirically/statistically and how can the strength of the population effect be quantified in comparison to the individual effect of the trait. After all, the claim that population effects determine individual fitness can only be of use if such effects can be detected empirically. To give an example, Eldakar, Wilson, et al. (2010) claim that the fitness of male water striders, *Aquarius remigis*, organised into patches (that is populations of water striders) depends on two components that are both affected by an aggressiveness trait individually expressed by the males. The individual component is given by the positive effect of aggressiveness on fitness mediated by mating success which is higher for more aggressive males that secure more mating opportunities than less aggressive males (Watters and Sih, 2005). The population component of individual fitness, on the other hand, arises from a different causal pathway and represents a negative effect of aggressiveness on fitness. Since the harassment experienced by females on a patch reflects the cumulative male aggression level on that patch and females tend to avoid harassment by escaping their current patch, the trait has a negative effect on patch productivity by decreasing the number of females on the patch and therefore the reproductive resources of all males on that patch. If such a decomposition into causes of individual fitness is to be useful, this decomposition must be empirically accessible in the sense that fitness is quantitatively given as a function of an individual component and a population component. This is possible only with a valid method of measuring the decomposition in empirical data.

Two methods for carrying out a quantitative decomposition of individual fitness into an individual component and a population component have received particular attention in the literature (Heisler and Damuth, 1987; Goodnight, Schwartz, and Stevens, 1992; Frank, 1998; Okasha, 2006; Sober, 2011; Gardner, 2015; McLoone, 2015): contextual analysis and the multilevel Price equation. However, the partitions of individual fitness given by the two methods are different in general. In particular, there are cases in which the Price approach claims the absence of population effects while contextual analysis claims their presence and vice versa.

The inconsistency between the two approaches is problematic because proponents of MLS argue that the distinction between individual effects and population effects is not just a statistical exercise but reflects a separation of causal pathways in the biological system under study as described above. While one causal pathway emanating from the individual trait is proposed to affect only individual aspects of fitness (the fitness

of the bearer), a different pathway is claimed to relate the trait with properties of the population as a whole and hence with a population component of individual fitness. Since the desired decomposition must reflect the underlying biological reality, two methods of decomposition that yield different answers cannot both be correct (Sober, 2011).

## 1.4 Correlation

Motivated by the inconsistencies of CA and PA, we look more closely at the concepts of correlation and causation using Pearl, Glymour, and Jewell (2016) and the ladder of causation (Table 1.1).

### 1.4.1 The multilevel Price equation and contextual analysis

The Price equation (1.2) (see Price equation examples on Wikipedia for instructive examples of its use) states an identity between properties of two distributions over a phenotype, for instance between metapopulation $A$ and metapopulation $A'$ over the phenotype $Z$ in diagram (1.5). More specifically, (1.2) expresses the change in mean phenotypic value from the parent to the offspring population in terms of covariances and means involving the two successive populations (see Section "Individual selection and the Price equation"). The purpose of this expression is a formal definition of selection: "Price's equation is *definining* evolutionary change – or, more properly, its component parts. Importantly, Price's equation provides a completely general, formal definition of selection.... Price's equation highlights that there are two conceptual components to evolution, namely selection and transmission." (Gardner, 2020, page 2). In summary,

> Price's equation describes evolutionary change as that which owes to changes in the frequencies of things (i.e. selection) and that which does not (i.e. everything else, collected under the umbrella-term 'transmission').
>
> Gardner (2020): "Price's equation made clear"

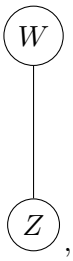
Hence the correlation, or covariance, in the first component of the Price equation (1.2)

$$\mathrm{Cov}(\nu_i, z_i) \tag{1.6}$$

is regarded as formal counterpart of selection. In instantiations of the Price equation with recorded data, the associated numeric term is taken to be indicative of selection.

This interpretation of the observed data, that is the measurement outcomes giving rise to the observables $Z, \overline{Z}$, and $W$ above, amounts to accepting the corresponding

correlations, or *associations*, between the variables that represent that data on the formal side. Specifically, the component (1.6) of equation (1.2) refers to an association between $Z$ and $W$ (see the first rung of the ladder of causation in Section 1.4),




, (1.7)

that is between individual phenotype and overall individual fitness, as required for Darwinian evolution according to Lewontin's principle of differential fitness (Lewontin, 1970; see Section "Units of selection").

Given a metapopulation, such as $A$ in diagram (1.5), the *multilevel Price equation* is a decomposition of individual selection into a component of population selection and a component of individual selection relative to the individual's ambient population (Gardner, 2015; equation (5))

$$\Delta \mathrm{E}_I(Z) = \mathrm{Cov}_I(\overline{W}, \overline{Z}) + \mathrm{E}_L \mathrm{Cov}_{K(L)}(W, Z), \tag{1.8}$$

where the index $I$ covers all $n : \mathbb{N}$ individuals in the metapopulation, the index $L$ covers all $p : \mathbb{N}$ populations, and the index K(L) covers for each population the $m : \mathbb{N}$ individuals within that population (assuming populations of homogeneous magnitude $m$ for simplicity; the total number of individuals in the metapopulation is $n = mp$). A derivation of equation (1.8) is given in the Appendix.

In the multilevel Price equation (1.8), the *selection* component

$$\mathrm{Cov}_I(\overline{W}, \overline{Z}) \tag{1.9}$$

refers to selection of *populations* and belongs to the association between overall individual fitness $W$ and phenotype of the ambient population $\overline{Z}$ that formalises the analog of individual selection (1.6) for populations and is mediated by $\overline{W}$ as population selection (cf. Sections "Mediators and conditional independence" and "Population selection"). The *transmission* component refers to transmission bias on the population level, formalised as (averaged) individual selection *within* subpopulations

$$\mathrm{E}_L \mathrm{Cov}_{K(L)}(W, Z). \tag{1.10}$$

The associations (1.9) and (1.10) furnish the basis of multilevel selection according to the *Price approach* (Okasha, 2006).

Contextual analysis, on the other hand, represents associations between variables in terms of regression coefficients of structural equations that involve these variables (Heisler and Damuth, 1987; Damuth and Heisler, 1988). A standard approach to formalising the association that features in Table 1.1 is the linear regression

$$W = c_1 Z + c_2 \overline{Z}, \tag{1.11}$$

where $c_1, c_2 : \mathbb{R}$ are regression coefficients. In a regression analysis of empirical data, that is in an instantiation of the structural equation (1.11) with data, the associated numerical term $c_1$, for instance, is taken to be indicative of individual selection (Okasha, 2006).

The Price equation expresses the associations between variables in terms of covariances and contextual analysis expresses these associations in terms of regression coefficients. However, generally for linear models, covariances and regression coefficients are linked by equations such as $b = \frac{\mathrm{Cov}(x,y)}{Var(x)}$ for the structural equation $y = a + bx$. We will see in Section "Measuring multilevel Selection" that for additive fitness functions the relationship between the coefficients of the Price equation and contextual analysis with respect to multilevel selection is given by a transformation of variables so that the multilevel Price equation (1.8) may be seen as equivalent to the linear regression

$$W = c_1'(Z - \overline{Z}) + c_2' \overline{Z} \tag{1.12}$$

with regression coefficients (cf. equation (1.11))

$$c_1' = c_1 = \frac{\mathrm{Cov}_I(W, Z - \overline{Z})}{\mathrm{Var}_I(Z - \overline{Z})} \tag{1.13}$$

$$c_2' = c_1 + c_2 = \frac{\mathrm{Cov}_I(W, \overline{Z})}{\mathrm{Var}_I \overline{Z}}. \tag{1.14}$$

For a detailed discussion of contextual analysis with transformed variables see Section "Discussion", particularly Table 1, in Heisler and Damuth (1987).

### 1.4.2 Genotypes and phenotypes

We maintain a conceptual distinction between *phenotypes* and *phenotypic values*. By *phenotype* we refer to an observable characteristic of the type of organism, i.e., species, that makes up the studied population such as the phenotype 'neck length' in a population

of giraffes. 'Phenotype' is synonymous with 'character' and 'trait'. An individual can be assigned a *phenotypic value* that represents the value of the phenotype as it is realised in this individual. For a given species, we define the phenotype $Z$ as the space of possible phenotypic values $z : Z$ which an individual of that species may possess. For a given individual, we define the phenotypic value $z$ as an instance of the phenotype of the species $Z$. For the phenotype giraffe neck length, $Z$ may be a subset of the real numbers that contains the values that might be measured on a sampled giraffe. An *instance* of this type, or a phenotypic value, is a real number that corresponds to the neck length (e.g., in metres) of a sampled giraffe.

Analogously, we make a distinction between *genotypes* and *genotypic values*. While genotype refers to a class of genomes such as a genetic species (Baker and Bradley (2006); page 648), genotypic value refers to an instance of the genotype, that is the genome of an individual of a genetic species.

The *types*[12] – such as genotype and phenotype in the above sense – we consider in this thesis may all be imagined as discrete sets. This may seem like a limitation of the models constructed here, as phenotypes are often conceptualised as real numbers $\mathbb{R}$ (or vectors of real numbers) that admit a metric, differential operators, and other operations that rely on topological properties of the real numbers. In this thesis, we make no use of topological or other properties of phenotype spaces other than that they are not empty. Moreover, since our models ultimately serve to describe experiments with actual objects, measurements, and values – all, for the time being, in finite multiplicity – the phenotype and genotype spaces in this thesis may be imagined as finite. For the same reasons, all distributions – and distributions of distributions – are finitely supported.

We assume that the individual phenotype, more precisely the phenotypic value displayed by an individual, can be measured deterministically (without disturbing the system in any way) and is given as a point in a space $Z$, that is an element of the set that makes up the space $Z$. For example, the space could be a finite set that describes a phenotypic property

$$Z = \{\text{wrinkly}, \text{smooth}\} \cong \{\bullet, \circ\},$$

or a set of real numbers

$$Z \subset \mathbb{R},$$

that stand for midpoints of bins designed to partition the space of possible outcomes of measurements of adult neck lengths of giraffes.

[12] See Bartosz Milewski's notes on functional programming for a useful way of thinking about types.

In order to reflect the distinction between individual phenotype and population phenotype, we choose a different name for the population phenotype $\overline{Z}$. $Z$ and $\overline{Z}$ are not assumed to be related in any specific way. However, in view of the discussion about emergent traits (see Okasha (2006) and Lenton et al. (2021)) we do assume that a population phenotype is *determined* by the collection of individual phenotypes the population comprises, more precisely, by the *distribution* of individual phenotypes that make up the population. With the space $PZ$ that contains as points the distributions on $Z$, that is formal sums such as $\frac{1}{2}|\bigcirc\rangle + \frac{1}{2}|\bullet\rangle : PZ$ and $\frac{2}{3}|\bigcirc\rangle + \frac{1}{3}|\bullet\rangle : PZ$ that represent populations $I$ and $I'$, resp., in diagram (1.4) (see also Section "Populations and probability monads" below), a population phenotype is a phenotype $\overline{Z}$ together with a map $PZ \to \overline{Z}$ that yields a population phenotype for a distribution of individual phenotypes, namely the distribution that describes the composition of the population in terms of its constituent individuals.

In applications of the Price equation, the population phenotype of interest is often given by the average individual phenotype within the population, for instance when *genetic value* is considered as character under selection (Okasha, 2006; Gardner, 2015). Also for the expression of the Price equation in terms of covariances (1.2), the population phenotype is assumed to be given as average individual phenotype. For the general models constructed in this thesis, the functional form of the population phenotype $PZ \to \overline{Z}$ does not need to be specified and hence this assumption is not required.

## 1.5 Causality

While the processes referred to in our models are causal, the formalisation in terms of covariances and regression coefficients in Section "Correlation" is correlational. This formal representation does without assumptions beyond conditional independence, and associated models can be represented as graphs (Pearl, 2009, Chapter 1). For example, the fact that individual fitness is correlated with the phenotypic values of both the individual and the population it is part of may be represented as the graph on the first rung of the ladder of causation described in Pearl and Mackenzie (2018)(see Table 1.1).

The second rung of the ladder of causation is characterised by the admission of causality. Accepting causality means allowing hypotheses about *causal relations* between variables, that is cause-effect relations. Cause-effect relations allow formalisation as *channels*, that is functions between spaces of probability distributions on sets of possible process outcomes (Jacobs and Zanasi, 2018, Section 2). For instance, in the ket notation of Sec-

tion "Individual selection and the Price equation", the channel that computes the causal effect of the observable $z_I : Z$ (phenotypic composition of population $I$, see function (2.60) in Section "Units of selection" below) on the observable $z_{I'} : Z$ (phenotypic composition of population $I'$) in diagram (1.4) is the channel that realises, or implements, the transformation $\frac{1}{2}|\circ\rangle + \frac{1}{2}|\bullet\rangle \mapsto \frac{2}{3}|\circ\rangle + \frac{1}{3}|\bullet\rangle$.

The directedness of causal relations is reflected in a refined representation of associations as *directed* graphs, so that the skeleton on the lowest rung in Table 1.1 is refined to the directed acyclic graph (DAG) (1.17) with channels, or cause-effect relations, as edges.[13] The causal assumptions underlying the DAG (1.17) yield expectations with regard to the system that are not implied by the correlations considered on the first rung of the ladder of causation. In particular, accepting the graphs constructed in this way as models of reality allows predicting the effect of *interventions*. Interventions and their consequences for the causal structure on which they operate are formalised by the do operator that is put in perspective in Section "Inference, intervention, and the do operator" below (Pearl, Glymour, and Jewell, 2016, Chapter 2; Pearl, 2019).

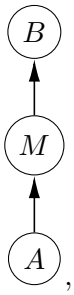
### 1.5.1 Mediators and conditional independence

Given a segment between two nodes $A$ and $B$


, (1.15)

[13] For instance, the wetness of the ground and rainfall may be correlated. However, this correlation alone does not give licence to causal statements. The latter require causal assumptions: rainfall causes the ground to be wet, but a wet ground doesn't cause rainfall.

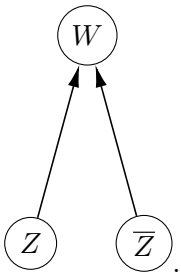
a mediator of this association, that is a node $M$ along the segment between the two nodes




, (1.16)

specifies the dependence of $B$ on $A$ in the sense that the two variables become *conditionally independent* given $M$ (see Pearl (1988), Section 3.1.3, and Fong (2013), Section 3.1). In other words, $A$ has no effect on $B$ other than the effect mediated by $M$, or $A$ has no direct effect on $B$, or, given $M$, knowledge of $A$ does not add to the knowledge of $B$. Diagram (1.16) demonstrates the conditional independence of $A$ and $B$ given $M$: if $M$ is removed, the effect of $A$ on $B$ is void, and the two observables are on disjoint and causally independent connected components of the graph.

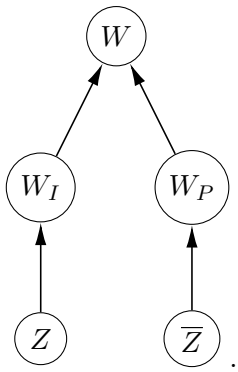
In this thesis, the central example for a causal structure is the dependence of individual fitness on individual and population phenotype $Z$ and $\overline{Z}$




. (1.17)

In order to reflect the two modes of reproduction and, more appropriately for populations, persistence on the individual and population level in the formal model (Doolittle and Booth, 2017; Doolittle and Inkpen, 2018; Papale, 2020), we introduce mediators $W_I$ and

$W_P$ that represent reproduction on the individual and population level, resp.




. (1.18)

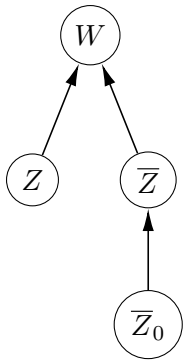
In light of the discussion on mediators above, this means that if, for instance, differences with respect to population reproduction $W_P$ between individuals are somehow nullified, differences in $\overline{Z}$ between individuals do not entail differences in overall fitness $W$ between individuals.

### 1.5.2 Inference, intervention, and the do operator

The directions of the edges in the diagrams above require commitment to a causal stance because they are entailed by causal assumptions regarding the effects of intervening on variables in the system. Consider again the empirical example of Section "Contextual analysis and the Price approach", where water striders sojourn in patches on the water surface and thus form populations within a metapopulation that spans across the habitat (Eldakar, Wilson, et al., 2010). With the sex ratio on a patch as population phenotype $\overline{Z}$, the causal structure in diagram (1.17) encodes the hypothesis that the sex ratio of a patch has an effect on overall individual fitness of its inhabitants but that no other property of the population as a whole has an effect on individual fitness (individual properties proper are subsumed in $Z$ and their effect travels along the channel represented by the left upward pointing arrow in diagram (1.17)). This assumption allows to infer that when sex ratio is fixed across a collection of populations then other properties of a specific individual's ambient population have no effect on its fitness.

The direction of the right arrow in diagram (1.17) corresponds to the assumption that controlling the value of $\overline{Z}$ implies controlling the value of $W$ – to the extent of the influence of $\overline{Z}$ on determining $W$ in concert with $Z$ in the diagram. This is in contrast with a hypothetical fourth variable $\overline{Z}_0$ that holds the population phenotype at some

earlier point in time. Since $\overline{Z}_0$ determines $\overline{Z}$ and not the other way around




(1.19)

intervening on $\overline{Z}$ has no effect on $\overline{Z}_0$.

The separation of the two legs in DAG (1.17) (or (1.19)) corresponds to the assumption of separate causal pathways, so that intervening on one of them leaves the other one unaffected (see Pearl (2009), Section 1.3.1, and the *principle of autonomy* described therein). An *intervention* in this context means that a variable in the DAG is set to a specific value, regardless of its parents' values. Practically, this may correspond to preparing an experiment with certain parameters according to the intended intervention. The sex ratios of populations of water striders, for example, may be adjusted by controlling the flow of individuals between patches (Eldakar, Dlugos, et al., 2009). Formally, interventions are represented by the *do operator* and may rely on mathematical functions associated to the edges of the DAG (Pearl, 2019). When representability may be assumed so that the DAG can be refined to a string diagram, intuitive operations on string diagrams correspond to inferential operations such as interventions and counterfactural reasoning (Jacobs, Kissinger, and Zanasi, 2019). The representation of causal structures in terms of string diagrams is the focus of Chapter "A functional representation of multilevel selection" below.

## 1.6 Representability

### 1.6.1 Structural Causal Models (SCM)

Above in Section "Correlation", the parameters of the regression were seen to encode the association between variables in a causal structure. In this thesis, we interpret the linear regression associated with contextual analysis (1.11) and that associated with the Price approach (1.12) as structural causal models (SCM) of the process select associated with an episode of multilevel selection in a metapopulation.

In addition to the variables that describe the system, an SCM includes a set of functions "that determine or simulate how values are assigned to each variable" (Pearl, 2018, page 4), for CA and PA these are linear regression equations. The encoding of the proposed causal relation in terms of numeric functions, such as linear equations with coefficients in the real numbers $\mathbb{R}$, has the benefit that the relations can be varied as the functions are varied along their parametrisation, that is the regression coefficients. Conversely, parameters of the SCM are estimated from data.

The focus of this thesis lies on abstract properties of the descriptions of biological systems discussed in this thesis. Of special interest in this respect is the dependency structure implied by the functional representation. In the following sections, this dependency structure is discussed in terms of monoidal categories and string diagrams. The abstract representation of CA and PA in Section "Models of multilevel selection" rests on this representation.

### 1.6.2 Category theory

In the previous section we discussed the representation of functional relations in terms of SCM such as linear regressions. Category theory gives an abstract view on the spaces and functions that implement a given process. Since this view does not rely on numerical specification, *string diagrams* can be used to represent the corresponding functional dependencies abstractly. String diagrams allow intuitive, yet rigorous manipulation in a similar spirit to the do operator and related graphical operations on Bayesian networks, or directed acyclic graphs (DAG) (see Section "Inference, intervention, and the do operator"), that correspond to valid mathematical operations on the underlying spaces and functions.

A fundamental idea realised in the mathematical structure 'category' is that of a collection of spaces with functions, or *morphisms*, between them. The following definition is copied from Riehl (2017) word-for-word for convenience. It not only defines the structure 'category' but also introduces some fundamental concepts.

> A **category** consists of
>
> - a collection of **objects** $X, Y, Z, \ldots$
> - a collection of **morphisms** $f, g, h, \ldots$
>
> so that
>
> - Each morphism has specified **domain** and **codomain** objects; the notation

> $f : X \to Y$ signifies that $f$ is a morphism with domain $X$ and codomain $Y$.
>
> - Each object has a designated **identity morphism** $1_X : X \to X$.
> - For any pair of morphisms $f, g$ with the codomain of $f$ equal to the domain of $g$, there exists a specified **composite morphism** $gf$ whose domain is equal to the domain of $f$ and whose codomain is equal to the codomain of $g$, i.e.,:
>
> $$f : X \to Y, g : Y \to Z \rightsquigarrow gf : X \to Z.$$
>
> This data is subject to the following two axioms:
>
> - For any $f : X \to Y$, the composites $1_Y f$ and $f 1_X$ are both equal to $f$.
> - For any composable triple of morphisms $f, g, h$, the composites $h(gf)$ and $(hg)f$ are equal and henceforth denoted by $hgf$ .
>
> $$f : X \to Y, g : Y \to Z, h : Z \to W \rightsquigarrow hgf : X \to W.$$
>
> That is, the composition law is associative and unital with the identity morphisms serving as two-sided identities.
>
> Riehl (2017): *Category Theory in Context*

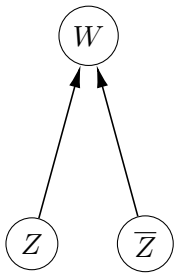
While this definition is very abstract and fundamental, it captures intuitions about functions and their domains (and codomains) that seem to be too simple to require formal definition when thinking of a typical example such as the category of finite sets FinSet.[14] Nevertheless, category theory has proven to be a useful tool for reasoning about mathematical structures and models (Riehl, 2017). Applications of category theory to areas outside of mathematics have received increased interest in recent years (Fong and Spivak, 2019). In particular, monoidal categories (see next section) have been successfully applied to problems in computer science, engineering, machine learning, and the wider real world (Baez and Stay, 2009).

[14] The category of finite sets FinSet has finite sets as *objects* and usual functions between finite sets as *morphisms*. The axioms to be satisfied by FinSet to qualify as category in the sense of the definition above are quickly verified: each finite set has an identity function that may be safely ignored, and three functions with appropriate domains and codomains can be composed associatively.

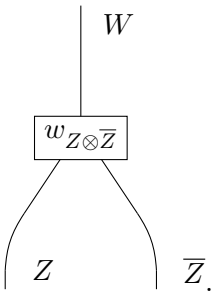
### 1.6.3 Monoidal categories

Bayesian networks (Fong, 2013, Chapter 3) encode cause-and-effect relations between variables as edges in directed acyclic graphs (DAG) so that parent nodes have causal effects on their offspring nodes. While this representation highlights dependencies (more precisely, *conditional independencies*) between variables, it ignores subtler aspects of functional dependencies that are consequences of *how* the independent variables are composed to yield the dependent variables (cf. Section "Mediators and conditional independence"). Symmetric monoidal categories with copy and discard functions (see below in this section), called *Markov categories* due to the similarity of their morphisms with Markov kernels on measurable spaces (Fritz, 2019) or *CD categories* (Jacobs, Kissinger, and Zanasi, 2019), allow drawing a more refined picture of the processes to be modelled. In particular, processes with several input variables may be composed of several processes that occur concurrently. String diagrams, a graphical formalism for functional relations in monoidal categories, extend Bayesian networks by representing this refined view completely and rigorously (see Jacobs, Kissinger, and Zanasi (2019), Section 3 "Bayesian Networks as String Diagrams").

In string diagrams, causal relations, or channels, are represented explicitly as functions (Jacobs and Zanasi, 2018, Chapter 3: "Directed graphical models") such as $w_{Z\otimes\overline{Z}} : Z \otimes \overline{Z} \to W$ in Table 1.1. The function $w_{Z\otimes\overline{Z}}$ refines the Bayesian network




to the string diagram




In this thesis we use basic properties of *monoidal categories*, i.e., categories in which

spaces can be paired to give product spaces (see Baez and Stay (2009) for an introduction to monoidal categories that is both rigorous and intuitive). More precisely, the spaces, or objects, of a monoidal category $\mathsf{C}$ with monoidal product $\otimes : \mathsf{C} \times \mathsf{C} \to \mathsf{C}$ and unit object $I : \mathsf{C}$ behave like a monoid. This means that spaces can be multiplied ("composed") associatively

$$(A \otimes B) \otimes C \cong A \otimes (B \otimes C) \text{ for all } A, B, C : \mathsf{C}$$

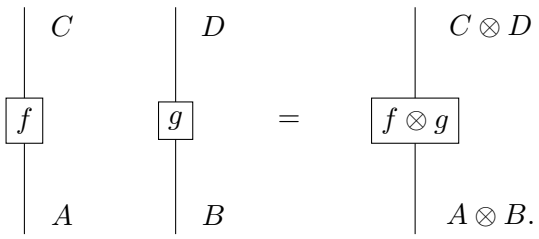
and that multiplication with the unit (a space that is, in a sense, trivial; see equation (1.25) below) leaves objects unchanged

$$I \otimes A \cong A \otimes I \cong A \text{ for all } A : \mathsf{C}.$$

Here, the symbol $\cong$ means that the spaces on its left and right are (naturally) isomorphic (see Fong (2013) and Fong and Spivak (2019) for rigorous definitions). Moreover, a monoidal category is *symmetric* if the product of spaces is commutative

$$A \otimes B \cong B \otimes A \text{ for all } A, B : \mathsf{C}.$$

Monoidal categories moreover satisfy certain axioms with respect to the interplay of morphisms and products. These follow from the rigorous definition in terms of categorical notions (see Baez and Stay (2009)) and are "automatic" in the graphical representation in terms of string diagrams. In a simplifying picture, categories may be thought of as vertices (objects) with arrows between them (morphisms) so that arrows can be composed in sequence when the codomain of one coincides with the domain of the other. In monoidal categories, arrows can moreover be composed parallely: $f : A \to C$ and $g : B \to D$ may be composed to $f \otimes g : A \otimes B \to C \otimes D$ so that



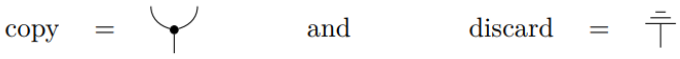

(1.20)

The symmetric monoidal categories of interest for this thesis are moreover equipped with copy and discard morphisms (Jacobs and Zanasi, 2018)

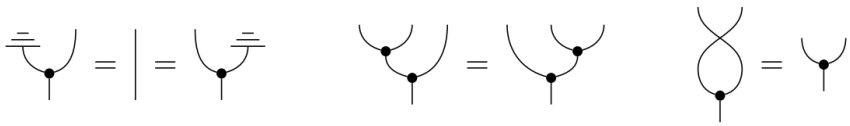
$$\mathsf{copy}_A : A \to A \otimes A \tag{1.21}$$

$$\mathsf{discard}_A : A \to I \tag{1.22}$$

for $A : \mathsf{C}$ and may be represented as (diagrams below copied from Jacobs, Kissinger, and Zanasi (2019)[15])




In this representation, algebraic properties of the functions find intuitive expression




Moreover, the functions satisfy various properties with respect to the symmetric monoidal structure of $\mathsf{C}$ described above.

The categories of interest for this thesis are based on the category of measurable spaces and stochastic maps $\mathsf{Stoch}$ (see Fong (2013), Chapter 2). The objects, or spaces, we consider contain the *states* of the system as points (Jacobs and Zanasi, 2018; Section 2). Since our systems are (meta)populations of individuals, states are (distributions of) distributions of individuals. For instance, given a population of full and empty circles such as $I$ in diagram (1.4) we can define a random variable with coordinate space $Z = \{\bullet, \circ\}$ that consists of the outcomes of sampling one individual from population $I$. The population $I$ itself may be seen as a distribution on $Z$, namely the distribution according to which the random variable given by sampling an individual from $I$ is distributed. We may represent this distribution as formal sum $\frac{1}{2}|\bullet\rangle + \frac{1}{2}|\circ\rangle : PZ$ where $PZ$ denotes the space of distributions, or states, on the set $Z$ (see Jacobs and Zanasi (2018) where $PZ$ is denoted by $\mathcal{D}(Z)$ and Section "Populations and probability monads" where the notation is explained).

The monoidal product of these spaces of distributions on sets are spaces of distributions on the corresponding cartesian product, so that, for example, for $A = \{a_1, a_2\}$ and $B = \{b_1, b_2\}$, the space $P(A \otimes B)$ contains distributions on the set

$$A \times B = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_2, b_2)\}. \tag{1.23}$$

In other words, when the outcomes of two random variables are considered *together*, that is as a product $A \otimes B$, this composition of observations yields outcomes that are

[15] © Bart Jacobs and Fabio Zanasi; licensed under Creative Commons License CC-BY.

combinations of outcomes in $A$ and $B$, that is pairs of outcomes as in (1.23). In ket notation, corresponding product states may be written as

$$\sum_{i,j} c_{ij}|a_i, b_j\rangle : P(A \otimes B). \tag{1.24}$$

The unit object of these categories of spaces of outcomes of probabilistic processes is the set with one element $I = \{*\}$, the outcome that is always entailed and that does not alter the outcome when observed in combination with other observables[16]

$$A \times I = \{(a_1, *), (a_2, *)\} \cong A \cong I \times A = \{(*, a_1), (*, a_2)\} : \mathsf{Set}. \tag{1.25}$$

Symmetric monoidal categories allow representing not only processes, such as experiments where input is transformed to output, but also the inferential formalism to reason about such processes in terms of Bayesian or frequentist inference, see Jacobs (2019) and Schmid, Selby, and Spekkens (2020).

[16] $\mathsf{Set}$ denotes the category of sets with sets as objects and maps between sets as morphisms.

# 2 A functional representation of multilevel selection

## 2.1 Multilevel selection

The concept of selection in Price (1995) delineates a process in which properties (character under selection, or phenotype) of an entity (unit under selection) in a population of similar entities (arena of selection) determine the multiplicity of reproduction of that entity in an offspring population (target of selection, or the entity's fitness). For individual selection, this entity is an individual in a population such that the individual's trait value determines its number of offspring. This dependence is formalised in the selection component (1.1) of the Price equation (1.2). The transmission component of (1.2) is assumed to be void on the individual level in this thesis.

### 2.1.1 Population selection

As argued by Hull (1980), if we accept both the viewpoint that genes are units of selection (Dawkins) and the viewpoint that organisms (populations of genes) are units of selection, then the question whether populations of individuals can be selected is settled: they can because organisms are selected and organisms are populations of individual genes. A question that remains is whether a *specific* population of individuals of some type may be regarded as individual unit of selection itself (Clarke, 2016). This question is discussed in more detail in Section "Interactors and replicators" below.

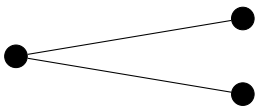
Our mathematical implementation of population selection as analogue of individual selection finds its biological justification in the viewpoint that "the theory of natural selection embodies a set of principles that apply more generally than to selection solely among organisms" (Damuth and Heisler, 1988; page 408). Accordingly, for population selection we define the unit of selection as the population (of individuals) within an arena of selection given by a *metapopulation* of populations (Gardner, 2020). A metapopulation of populations is a collection of multiple populations that – like individuals within a pop-

ulation in non-social evolution – do not interact causally (see Uyenoyama and Feldman (1980), page 395; Wade (2016)). The character under selection is given by a population phenotype $\overline{Z}$ that is a function of the individuals within the population, usually a function of the phenotypic composition of the population (this and related functions are discussed in Section "A functional representation of multilevel selection" below). As for individual selection with law (1.3), we assume that population selection is stable in its functional form (Wade and Kalisz, 1990) so that we may assume a law

$$\mathsf{interact}_{\overline{Z}} : \overline{Z} \to \mathbb{N} \tag{2.1}$$

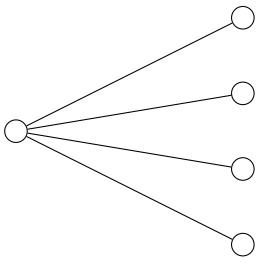
that determines the number of offspring to be expected for a population according to its phenotypic character in $\overline{Z}$. Representing reproduction of populations in terms of whole numbers of offspring populations may seem strange – unless reproduction is realised, for instance, by division or budding of populations as in planaria (Fields and Levin, 2018). In particular, adopting a view of fitness as determining persistence rather than reproduction (Borrelli et al., 2015; Bapteste and Papale, 2020; Lenton et al., 2021) requires the ability to represent non-integral fitness. Nevertheless, at present we regard population fitness as a map from the population phenotype to the whole numbers because it allows us to draw diagrams such as diagram (1.5). We will later consider fitness with values in nonnegative real numbers so that $\mathsf{interact}_{\overline{Z}} : \overline{Z} \to \mathbb{R}_{\geq 0}$.

When we discussed individual selection in Section "Individual selection and the Price equation", the assumption of perfect heredity on the individual level allowed us to realise offspring as perfect copy of its parent so that



(2.2)

and



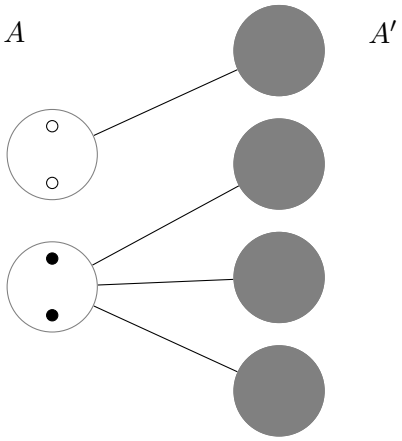
(2.3)

represent the selection process from the individuals' perspective as the population transforms from the parent population to the offspring population in diagram (1.4).

However, the offspring of individual $i \colon I$ in the Price equation (1.2) includes all individuals in the offspring population $J$ that have $i$ as (single) parent. Selection as in equation

(1.1) determines the number of offspring of $i$ in $J$ but not the phenotypic (or genotypic) values of the offspring individuals. The change in phenotypic value that occurs between a parent and each of its offspring is conceptually different from selection and is captured (as average) by the transmission term of equation (1.2).[1]

Unlike for individuals, we do not assume perfect heredity for populations. Selection diagrams similar to diagrams (2.2) and (2.3) but in which the circles stand for populations rather than for individuals should represent offspring without specifying their phenotypic value as this value is not determined by selection. Accordingly, an example of a diagram of population selection analogous to diagram (1.4) for individual selection is




(2.4)

where absolute population fitness of a population composed of two empty circles is 1 and that of a population composed of two full circles is 3.

### 2.1.2 String diagrams

In Section "Monoidal categories" we introduced string diagrams as graphical representation of morphisms of monoidal categories. For the CD categories[2] featuring in this thesis, string diagrams describe transformations of variables. These transformations can

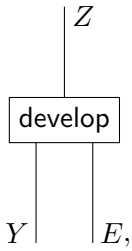
[1] The Price equation decomposes total change in phenotypic composition from a parent population to its offspring population into a component that represents a change in quantity of the entities involved (selection) and a component that represents a change in quality of the entities involved (transmission). The equation thereby defines selection (see Section "Individual selection and the Price equation"). The transmission component may represent various mechanisms such as mutation, recombination, and phenotypic plasticity, all of which we assume to be absent on the individual level. In population genetic models with more sophisticated modes of replication, this second component of the Price equation can often be ignored when transmission bias is not absent but undirected.

[2] CD categories are symmetric monoidal categories with copy and discard Jacobs, Kissinger, and Zanasi (2019).

be thought of as processes that connect their input with their output by transforming the former into the latter. The representation in terms of string diagrams is useful because *dependencies* can be sorted out graphically in these diagrams. In causal models, a dependency represents causal influence of a variable on another variable.

In this section, we explain a graphical formalism that expresses the categorical structures introduced above in terms of string diagrams. String diagrams allow rigorous diagrammatic reasoning about models that are formulated in terms of symmetric monoidal categories as described in Section "Monoidal categories". The rules and assumptions associated with the structure of the category ensure that the behaviour of mathematical models formulated in terms of objects and morphisms of that category corresponds to the behaviour one would intuitively associate with strings on a two-dimensional plane. The description of string diagrams in this section does not cover mathematical details. A rigorous introduction to almost all of the mathematical concepts used in this thesis and to many more is Jacobs (2019).

The formalism described here is associated with computational structures that allow providing process descriptions such as




(2.5)

where genes $Y$ and environment $E$ of a developing organism jointly determine the phenotypic outcome, with a rigorous mathematical context that allows constructing concrete statistical models (Patterson, 2020; Schmid, Selby, and Spekkens, 2020). The string diagrams used in this thesis represent experimental procedures or other well-defined processes. Strings, more precisely, string segments, represent observables that may be measured deterministically and without changing the outcome of other measurements.[3]

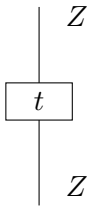
[3] We will see below in Section "Populations and probability monads" how non-deterministic observables, that is *random variables*, may be represented in string diagrams as *tubes* that comprise the samples.

The diagram


(2.6)

shows an observable with state space $Z$. As an example, in the following we may imagine $Z$ to be space of measurements of a phenotype such as giraffe neck length in metres, or a phenotype with two values such as $Z = \{\bullet, \circ\} : \mathsf{Set}$.

During the experiment, an observable may undergo transformations, for example when the system it refers to is subject to a causal process that changes its configuration. We represent a transformation as function on the state space $Z$ of the variable that takes an instance, i.e., a state value, and returns a transformed instance, i.e., the new state of the observable. Diagrammatically,




(2.7)

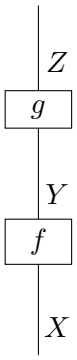
where $t\colon Z \to Z$ is a function associated with the transformation or process (that we picture as deterministic in this section). In diagram (2.7), $t$ could represent development from infant to adult giraffe such that the observable represented by the string below $t$ is infant neck length while the observable represented by the string above $t$ is adult neck length of the same individual. Both observables have outcomes in the same space $Z$ so that processes such as development may be conceptualised as transformations $Z \to Z$ in this formalism.

Functions in string diagrams may also refer to transformations *between* types (or observables) that rather than describing how a variable changes during a process ('over time' as in diagram (2.7)) describe how variables connected through a process are related. A function $\mathsf{develop}\colon Y \to Z$ gives a description of instances of type $Y$ in terms of type $Z$. It does so by providing the means, or algorithm[4], to convert any instance of $Y$ into the corresponding instance in $Z$. In a model of genetic evolution with development in a constant environment, develop may deterministically assign a phenotype in $Z$ to any

[4]This algorithm is the formal mirror image of the actual biological process the function refers to.

genotype in $Y$ (a *genotype-phenotype map* between a genotype space and a phenotype space as in diagram (2.69), see Stadler et al. (2001)).

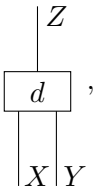
The string diagrams in this thesis do not capture time but only the sequence of processes along strands of the diagram. This sequence proceeds along each strand of the diagram from bottom to top. The process represented in




(2.8)

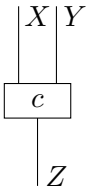
has an instance of $X$ as input, an instance of $Y$ as intermediate, and an instance of $Z$ as output, each depending on its predecessor with $X$ assumed as given. We consider only *connected* string diagrams in which all strands are connected to at least one other strand through a process.

Two strands of a diagram may be connected through functions that depend on both. In




, (2.9)

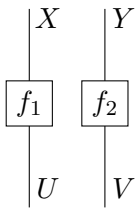
the output in $Z$ depends on both strands $X$ and $Y$. Similarly, strands may be joined because they originate from the same process




(2.10)

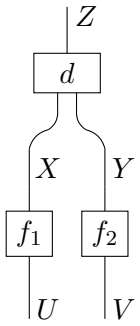
so that both depend on the same input.

A process as in diagram (2.9) synchronises strands in the sense that while



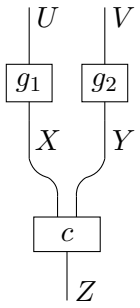

(2.11)

does not specify the order in which $f_1 : U \to X$ and $f_2 : V \to Y$ occur, it is implied in




(2.12)

that both conclude before the process $d$ begins. Likewise, a process as in diagram (2.10) anchors strands in the sense that in




(2.13)

both processes $g_1 : X \to U$ and $g_2 : Y \to V$ begin after $c$ has concluded.

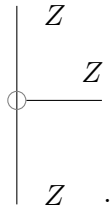
A string segment between boxes represents an observable that is available for measurement after the process that produced it has concluded and before the process that transforms it has started. Output states of processes are available along string segments to serve as input states for other processes. Using variables in multiple processes requires copying, implemented for all $X : \mathsf{C}$ by a copy function $\mathsf{copy}_X : X \to X \otimes X$ (see discussion following its definition in (1.21)). Variables must be available in multiple instances in a string diagram when the corresponding physical quantities are involved in multiple causal processes, and copying keeps track of dependencies in the diagram: copies depend on all variables and functions their 'parents' depend on, and 'siblings', that is instances originating from the same copy process, become independent after 'birth' in the sense

that transformations of one sibling leave others unchanged. This property reflects the *principle of autonomy* in Pearl (2009) that rests on causal assumptions about the system (cf. Section "Inference, intervention, and the do operator"):

> The ability of causal networks to predict the effects of actions of course requires a stronger set of assumptions in the construction of those networks, assumptions that rest on causal (not merely associational) knowledge and that ensure the system would respond to interventions in accordance with the principle of autonomy. These assumptions are encapsulated in the [...] definition of causal Bayesian networks.
>
> Pearl (2009): *Causality: Models, Reasoning, and Inference*

Diagram (2.8) specifies the dependencies of three observables over $X, Y,$ and $Z$, each represented as string segment annotated with the name of the space in which the observable takes values.[5] While the value of $Z$ is determined by $Y$, the value of $Y$ is determined by $X$. A diagram like (2.8) describes the general process of $gf$ ($gf : X \to Z$ denotes the composite function as in Section "Category theory") transforming any input from $X$ to output in $Z$. An instance of $X$ to which $gf$ is applied is not specified. In a concrete experiment this process is instantiated with a value, or instance, of $X$ (cf. diagram (2.16) below).

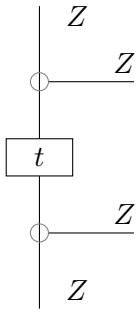
Since the observables considered in this thesis are classical, we can measure any string segment in the diagram to determine the value of the corresponding observable. In order to specify such measurements of an observable obtained during an experiment, we make use of an operation described in Schmid, Selby, and Spekkens (2020). This operation "allows us to directly gain knowledge from a classical causal system" (Schmid, Selby, and Spekkens (2020), page 15) and corresponds to deterministic measurement of a classical observable. In a diagram, this operation on a string segment is depicted like this:




. (2.14)

[5] We may refer to these simply as 'observables $X, Y$ and $Z$' when the diagram allows this. Diagram (2.7), for example, does not, as it shows two observables over $Z$, one that may be observed before $t$ and one that may be observed after $t$.

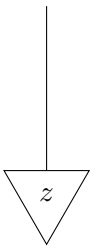
The outcome of the measurement operation depends on the segment in the diagram at which it is performed. In the diagram




(2.15)

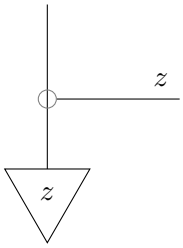
the instance obtained with a measurement before the transformation $t$ occurs is generally different from the value measured after the transformation. However, if $t$ is deterministic, the input to $t$ *determines* the upper measurement in diagram (2.15).

'Open' strings like those entering the diagrams above from the bottom can be 'grounded' by attaching an instance like this



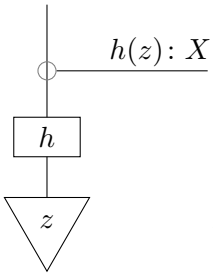

(2.16)

such that measurement of a variable that has been instantiated in this manner as in



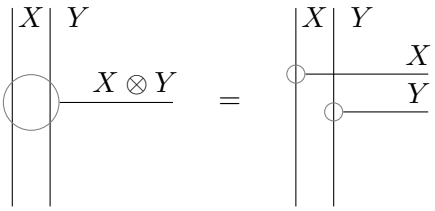

(2.17)

yields the instance itself. For a deterministic map $h\colon Z \to X$ we then also have




. (2.18)

Measuring an observable over a product $X \otimes Y$ models an experiment where two variables/factors are observed in combination ('simultaneously', cf. Section "Monoidal categories"). This is defined by




(2.19)

(equation (91) in Schmid, Selby, and Spekkens (2020)) and thus amounts to picking string segments (that is observables) that are measured together.

### 2.1.3 Populations and probability monads

The Price equation (1.2) decomposes the change in mean phenotype from a parent generation to an offspring generation. In this thesis, we understand this succession of generations as transformations of populations that follows laws such as (1.3) and (2.1) that transform an input population – the parent generation – to an output population – the offspring population. In order to refer to this transformation as function, we need to define spaces of which populations are points, so that the transformation that corresponds to the selective regime encoded in the fitness laws may be computed for any population. The transformation traced by the Price equation (1.2) then becomes an instance of this function with the given parent population as input.

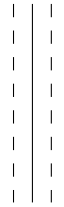
In Section "Monoidal categories" we introduced the state space $PZ$ of distributions on the set $Z : \mathsf{Set}$ (see Jacobs and Zanasi (2018) where $PZ$ is denoted by $\mathcal{D}(Z)$). The spaces $Z$ and $PZ$ are related in a canonical way that holds for any set of outcomes $Z : \mathsf{Set}$. The relation constitutes a *functor* of categories $P : \mathsf{C} \to \mathsf{C}$ called *probability monad* that associates to any space $Z$ the space of distributions on that space $PZ$ and to any mapping of spaces $g : Y \to Z$ in the domain of the monad $P$ a mapping $Pg : PY \to PZ$ in its codomain (see Definition 1.3.1. in Riehl (2017) for a rigorous definition of functors). The notion summarises the properties required for a mapping between categories to be consistent with the morphisms in the domain and codomain of the functor. Given a function of sets $g : Y \to Z$, for example, the probability monad induces a map of distributions on $Y$ to distributions on $Z$ that substitutes elements of $Y$ by their image elements in $Z$ to yield a formal expression in $PZ$. With $Y = \{\bullet, \circ\}$ and $Z = \{\text{wrinkly}, \text{smooth}\}$, the map

$$g : Y \to Z \text{ with } g\bullet = \text{smooth and } g\circ = \text{wrinkly}, \tag{2.20}$$

for instance, gives the transformation

$$\frac{2}{3}|\bullet\rangle + \frac{1}{3}|\circ\rangle \mapsto \frac{2}{3}|\text{smooth}\rangle + \frac{1}{3}|\text{wrinkly}\rangle. \tag{2.21}$$

Following an idea of Virgo (2020), we represent the monad explicitly in string diagrams. To this end, we draw monad variables in $PZ$, that is variables on the set of distributions over $Z$, as tubes that encompass a string

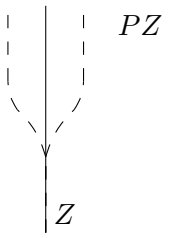



(2.22)

Associated with the functor $P$ are maps $\eta_Z : Z \to PZ$ and $\mu_Z : PPZ \to PZ$ that make $P$ a *monad* in the sense of Fritz and Perrone (2020), Section 1[6] (see Fritz and Perrone (2020), Section 6.1 "The idea of probability monads", for more details on probability monads). The functions $\eta_Z$ and $\mu_Z$ formalise general operations on formal expressions. $\eta$ is sometimes called the *unit* of the monad and includes $Z : \mathsf{Set}$ itself in the space of formal expressions on $Z$ by sending elements of $Z$ to formal sums with one term, so that in ket notation

$$\eta_Z : Z \to PZ, z \mapsto |z\rangle. \tag{2.23}$$

In string diagrams, the monad unit may be drawn as




(2.24)

The second monad function,

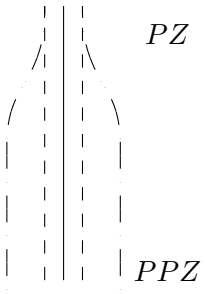
$$\mu_Z : PPZ \to PZ, \tag{2.25}$$

[6] The monad $T : \mathsf{C} \to \mathsf{C}$ considered in Fritz and Perrone (2020) is more general than the probability monad considered here. Using the ket notation for formal sums, $PZ$ is the space of formal expressions of elements of $Z$ with coefficients in $[0, 1] \subset \mathbb{R}$ that sum to 1. All formal expressions considered in this thesis are finite.

realises reducing sums-of-sums to sums, for instance[7],

$$\mu_Z\left(\frac{1}{2}\left(\frac{2}{3}|\bullet\rangle+\frac{1}{3}|\circ\rangle\right)+\frac{1}{2}\left(\frac{1}{3}|\bullet\rangle+\frac{2}{3}|\circ\rangle\right)\right)=\frac{1}{2}|\bullet\rangle+\frac{1}{2}|\circ\rangle : PZ. \tag{2.26}$$

This operation may also be drawn as string diagram:



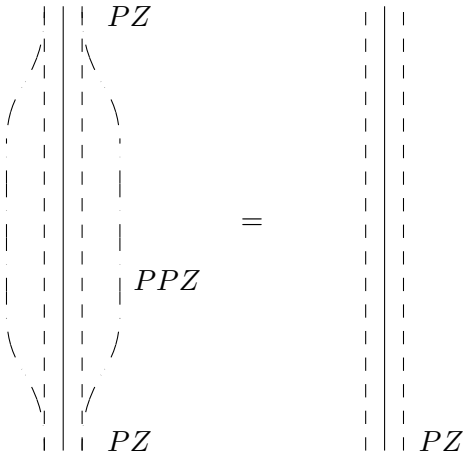

(2.27)

The operation $\mu_Z : PPZ \to PZ$ adds and multiplies coefficients with values in $[0,1] \subset \mathbb{R}$ and does not involve algebraic operations on elements of $Z$ itself. For metapopulations, $\mu$ amounts to joining the populations in the metapopulation in $PPZ$ to *one* population in $PZ$.

Taken together, the two monad functions are consistent in the sense that

$$PZ \xrightarrow{\eta_{PZ}} PPZ \xrightarrow{\mu_Z} PZ = PZ \xrightarrow{1_{PZ}} PZ, \tag{2.28}$$

where $1_{PZ} : PZ \to PZ$ denotes the identity on $PZ$, or, as string diagram,



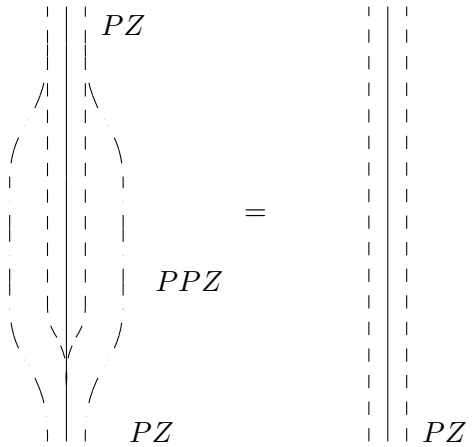

(2.29)

and

$$PZ \xrightarrow{P\eta_Z} PPZ \xrightarrow{\mu_Z} PZ = PZ \xrightarrow{1_{PZ}} PZ \tag{2.30}$$

[7] Unlike the example equation (1) in Fritz and Perrone (2020), our example refers to the probability monad that contains distributions, that is *normalised* formal sums (the coefficients sum to 1). The reduction $\mu : PPZ \to PZ$ preserves normalisation.

with associated string diagram




(2.31)

(see equations (2) in Fritz and Perrone (2020)). For the first equality (2.28) note that $PZ$ itself is a space in $\mathsf{C}$ and the functor $P$ can be applied to it. The associated map $\eta_{PZ} : PZ \to PPZ$ sends a distribution $\omega = p_\bullet|\bullet\rangle + p_\circ|\circ\rangle : PZ$ to the distribution of distributions $|\omega\rangle = |p_\bullet|\bullet\rangle + p_\circ|\circ\rangle\rangle : PPZ$.

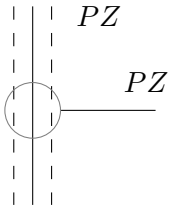
Elements of $PZ$ are distributions and as such express uncertainty via a sampling map

$$\mathsf{sample} : PZ \to Z \tag{2.32}$$

that formalises a sampling operation, such as drawing a random sample from a population of individuals, or algorithmically sampling from a distribution that is given in terms of mathematical formulae. Nevertheless, variables with values in $PZ$ – like those with values in $Z$ obtained by measuring the phenotype of an individual – are considered deterministic. In particular, given a population of individual phenotypes $I$ as in diagram (1.4), we may consider the variable with values in $PZ$ that represents the population $I$. It is given by sorting the finitely many individuals according to their phenotypic value and constructing a corresponding normalised formal sum on those phenotypic values in $Z$ that are represented in $I$. The population $I$ in diagram (1.4), for instance, is assigned the distribution $\frac{2}{4}|\circ\rangle + \frac{2}{4}|\bullet\rangle = \frac{1}{2}|\circ\rangle + \frac{1}{2}|\bullet\rangle : PZ$. Moreover, given a population $I$, there is the *random* variable that represents sampling one individual from $I$. If we denote by $\omega_I : PZ$ the distribution that corresponds to the population of phenotypes $I$, this random variable is distributed according to $\omega_I$.

The aspect of determinism is manifest also in the graphical formalism. We can represent the variable given by determining the phenotypic distribution of a population in

terms of diagrams and the measurement operation introduced in Section "String diagrams" as follows



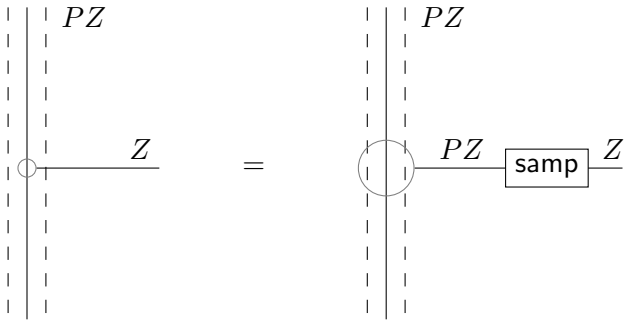

(2.33)

However, the diagram also comprises the random *inner* variable given by sampling one individual from the population and determining its phenotype. It is given by the string within the tube in (2.33) and can be obtained in two equivalent ways:

1. by sampling one individual from the population, or

2. by sampling from the distribution given by the variable that holds the phenotypic distribution of the population.

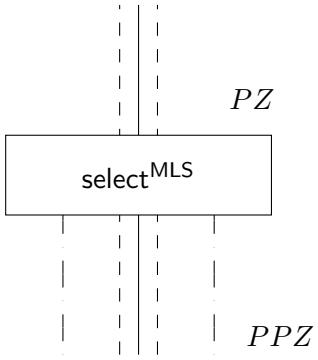
In the diagrammatic representation, the tube that holds the deterministic variable indicates that the string within represents a random variable. The two operationalisations of $\mathsf{sample} : PZ \to P$ take the equivalent forms




(2.34)

When discussing diagram (1.5) we noted that the Price equation is sometimes referred to as 'dynamically insufficient' even though the Price equation can in principle describe the dynamics with means of higher moments of the distribution (Gardner, West, and Barton, 2007). This is because the instantiation (1.2) for the phenotypic distribution does indeed not map the metapopulation structure. In particular, while the organisation of the metapopulation $A : PPZ$ in (1.5) is required for computing (1.2), the populations in the offspring metapopulation ($J_1, \ldots, J_6 : PZ$ in (1.5)) are not explicit in (1.2). This mirrors the fact that (1.2) refers to a multilevel selection [1] model in which populations replicate only due to the replication of the individuals within (or genes within

organisms). The populations themselves do not replicate, and there is no notion of direct populational heredity. Replication on the population level is included in models of multilevel selection [2]. For this reason, we understand the multilevel selection process $\mathsf{select}^{MLS}$ described by the multilevel Price equation and contextual analysis as transformation from metapopulations to populations




(2.35)

### 2.1.4 Selection and absolute fitness

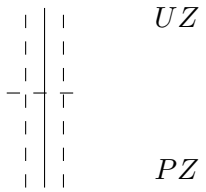
In order to implement the selection process using the absolute fitness function $\mathsf{interact} : Z \to \mathbb{R}_{\geq 0}$, we express the result of individual interactions with the environment giving rise to absolute fitness in terms of the multiset monad (Jacobs, 2019, Section 1.4). The multiset monad over the set of phenotypic values $Z$ is the unnormalised version of the probability monad considered in the previous section and denoted by $UZ$. Elements of $UZ$ are finite sums of elements of $Z$ with nonnegative real coefficients but without the requirement that the coefficients sum to 1. The elements of $UZ$ represent populations of $Z$ values, given in absolute terms. The monad functions

$$\eta_Z : Z \to UZ \text{ and } \mu_Z : UUZ \to UZ \tag{2.36}$$

are defined analogous to those of the probability monad ((2.23) and (2.25)) with $\eta_Z$ sending an element of $Z$ to a sum with one term and $\mu_Z$ sending a sum of sums to a simple sum.

The multiset monad is linked to the probability monad by inclusion and normalisation. A distribution in $PZ$ is automatically a sum of elements of $Z$ with nonnegative real coefficients, that is an element of $UZ$. In string diagrams, we denote inclusion $UZ \to PZ$

by a dashed horizontal line as follows




(2.37)

Conversely, given a (nonempty) element of $UZ$, normalisation yields an element of $PZ$:

$$\mathsf{norm} : UZ \to PZ, \sum_j d_j|Z_j\rangle \mapsto \frac{1}{\sum_k d_k} \sum_j d_j|Z_j\rangle. \tag{2.38}$$

With these definitions, the selection process $\mathsf{select} : PZ \to PZ$ is given by unnormalised selection

$$\mathsf{select}_{UZ} : PZ \to UZ, \sum_j d_j|Z_j\rangle \mapsto \sum_j d_j \mathsf{interact}(Z_j)|Z_j\rangle$$

followed by normalisation $\mathsf{norm} : UZ \to PZ$ so that

$$\mathsf{select}_{PZ} : PZ \xrightarrow{\mathsf{select}_{UZ}} UZ \xrightarrow{\mathsf{norm}} PZ. \tag{2.39}$$

### 2.1.5 Soft selection and hard selection

The crucial difference between the Price approach to multilevel selection and contextual analysis (see Section "The multilevel Price equation and contextual analysis") lies in the distinction between hard and soft selection (Wade, 1985; Goodnight, Schwartz, and Stevens, 1992). In Figure 2.1, individual fitness within a metapopulation of three populations of individuals with given phenotype is considered in three different selection regimes "Group selection"[8], "Soft selection", and "Hard selection". The dotted lines in the graphs show population fitness (that is average individual fitness within the population (multilevel selection [1], see Section "Interactors and replicators" below)) as it varies with average population phenotype. The three short lines in each graph show relative fitness within the respective population dependent on individual phenotype. While "Group Selection" refers to population selection, "Soft Selection" and "Hard Selection" are models of individual selection that differ in the choice of arena of selection.

**Population selection** When only population selection acts but no individual selection (panel "Group selection"), fitness within the populations is homogeneous, in other

[8] 'Group selection' is a different name for 'population selection' (see Section "Population selection").

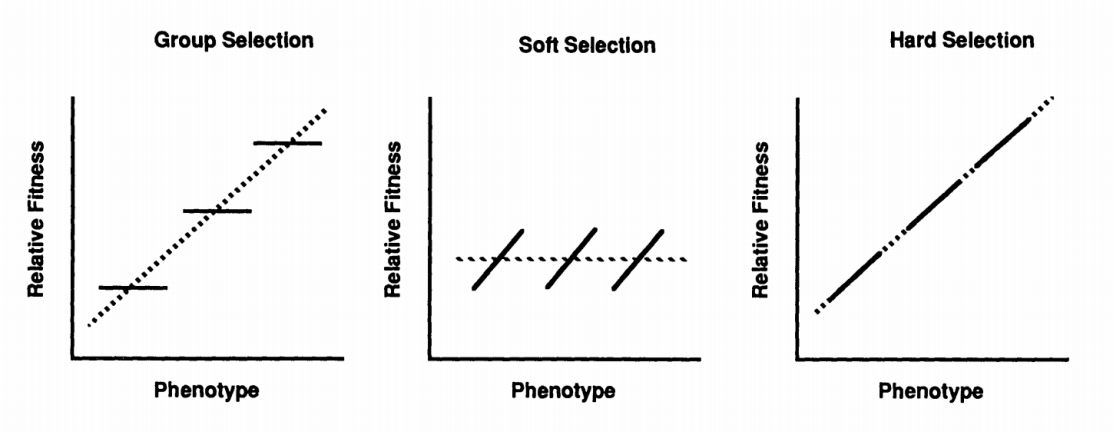



**Figure 2.1:** Fig. 2 from Goodnight, Schwartz, and Stevens (1992).

words fitness does not depend on individual phenotype directly (but only indirectly through its contribution to the population phenotype).

**Soft selection** When only soft selection acts on the metapopulation, individual fitness is independent of population phenotype. Within the populations, however, individual selection acts with respect to the individual phenotype, that is the arena of selection of each individual is its population. In particular, individual fitness of an individual depends on the *distribution* of phenotypes that makes up the population it is part of because population regulation occurs locally within the population so that populations sizes remain constant under soft selection (Wade, 1985).

**Hard selection** Hard selection refers to individual selection across the metapopulation that acts without regard to populations. The arena of selection for hard selection is the metapopulation. Despite the absence of population selection the dotted line for population fitness goes up due to cross-level by-products, see Section "Cross-level by-products".

We can formalise these selective scenarios using the ket notation and probability monads as described in Section "Populations and probability monads". To this end, we represent populations such as $I$ in diagram (1.4) as elements of $PZ$ by assigning to a population the normalised distribution that characterises its phenotypic composition. Individual selection then becomes a function

$$\mathsf{select}_{PZ} : PZ \to PZ \tag{2.40}$$

that sends a parent population to its offspring population. In the generational progression in diagram (1.4), for example,

$$\mathsf{select}_{PZ}\omega_I = \omega_{I'} : PZ, \tag{2.41}$$

where $\omega_I, \omega_{I'} : PZ$ are the distributions that represent the populations $I$ and $I'$, resp. Using the function $\mathsf{interact} : Z \to \mathbb{R}_{\geq 0}$ that assigns to a phenotypic value the expected absolute fitness of its carrier, that is the number of offspring, $\mathsf{select}$ decomposes into a function

$$\mathsf{select}_{PZ} = PZ \xrightarrow{\cdot\mathsf{interact}} UZ \xrightarrow{\mathsf{norm}} PZ, \tag{2.42}$$

$$\sum_j d_j|Z_j\rangle \mapsto \sum_j d_j\mathsf{interact}(Z_j)|Z_j\rangle \mapsto \frac{1}{\sum_k d_k\mathsf{interact}(Z_k)} \sum_j d_j\mathsf{interact}(Z_j)|Z_j\rangle. \tag{2.43}$$

Analogously, we may represent metapopulations as elements of $PPZ$, that is distributions of distributions on $Z$, or populations of populations of individual phenotypes. The function

$$\mathsf{select}_{PPZ} : PPZ \to PPZ \tag{2.44}$$

then formalises a complete episode of multilevel selection. In the remainder of this section, we describe how the assumptions of the three models of selection above are reflected in the structure of these selection maps.

1. Population selection Population selection is characterised by the property that it leaves the within-population distributions unchanged. In terms of the ket notation, if a metapopulation $\alpha : PPZ$ is composed of several populations $\omega_1, \ldots, \omega_n : PZ$ in varying proportions so that

$$\alpha = \sum_{i=1}^{n} c_i|\omega_i\rangle : PPZ, \text{ with } c_i : [0,1] \subset \mathbb{R}, \sum_{i=1}^{n} c_i = 1, \tag{2.45}$$

then, since population selection has no individual fitness effects within populations,

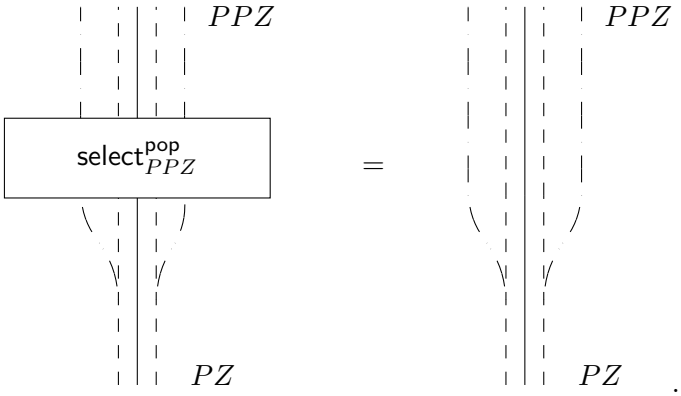
$$\mathsf{select}^{\mathsf{pop}}_{PPZ}\alpha = \sum_{i=1}^{n} c'_i|\omega_i\rangle, \text{ with } c'_i : [0,1] \subset \mathbb{R}, \sum_{i=1}^{n} c'_i = 1 : PPZ. \tag{2.46}$$

In other words, the metapopulation changes under the action of population selection $\mathsf{select}^{\mathsf{pop}}_{PPZ} : PPZ \to PPZ$ because the proportions in which populations are represented in the metapopulation change. The phenotypic values of the populations,

that is the composition of the populations $\omega_i$ that make up the metapopulation $\alpha$, however, do not change. In particular, population selection in a metapopulation with one population phenotype – like individual selection in a population comprising only one phenotype (possibly in several instances, that is individuals) – has no effect:

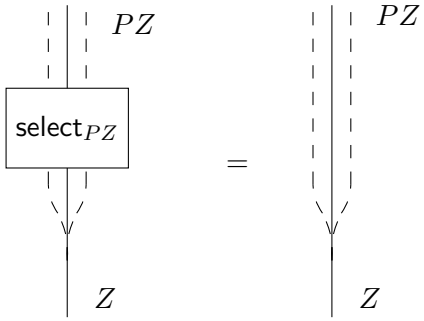
$$\mathsf{select}^{\mathsf{pop}}_{PPZ}|\omega\rangle = |\omega\rangle : PPZ, \text{ for all } \omega : PZ. \tag{2.47}$$

In terms of string diagrams, this property may be expressed as




. (2.48)

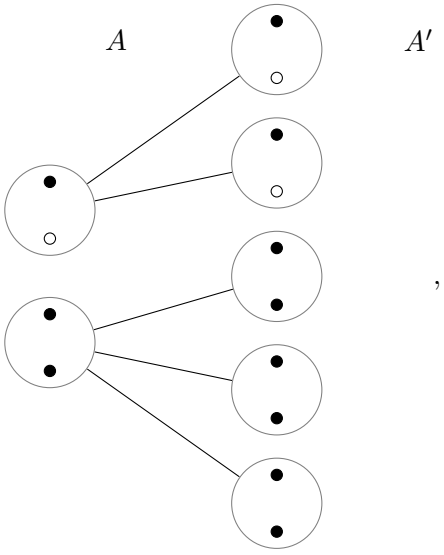
It is analogous to




(2.49)

for individual selection in populations and reflects the assumption that there is no transmission bias on the population level (2.48) (or on the individual level (2.49); see Section "Individual selection and the Price equation", equation (1.1)). Finally, in terms of metapopulations drawn as in diagram (1.5) population selection is

instantiated, for example, as follows:




, (2.50)

where $\mathsf{interact}_{\overline{Z}\overline{z}\bullet\circ} = 2 : \mathbb{N}$ and $\mathsf{interact}_{\overline{Z}\overline{z}\bullet\bullet} = 3 : \mathbb{N}$.

2. Soft selection The arena of soft selection is the population within a metapopulation. The effects of soft selection on an individual therefore depend only on the phenotypic composition of the population it is situated in but not on that of other populations in the metapopulation. Changes to the composition of a metapopulation under soft selection are due to within-population changes so that in ket notation with $\alpha = \sum_{i=1}^{n} c_i|\omega_i\rangle : PPZ$ as above

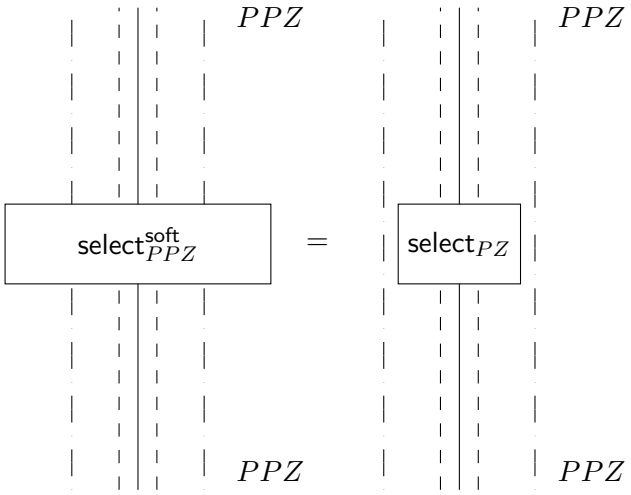
$$\mathsf{select}^{\mathsf{soft}}_{PPZ}\alpha = \sum_{i=1}^{n} c_i|\omega_i'\rangle, \text{ with } \omega_i' : PZ, i = 1, \ldots, n, \tag{2.51}$$

where for each $i = 1, \ldots, n$ the new population distribution $\omega_i'$ arises as result of a transformation of individual selection internal to the populations $\mathsf{select}_{PZ} : PZ \to PZ$ with $\mathsf{select}_{PZ}\omega_i = \omega_i'$. Since soft selection preserves the size of the population absolute individual fitness within a population under soft selection sums to 1 so that the interaction function for soft selection $\mathsf{interact}_Z : Z \otimes PZ \to \mathbb{R}_{\geq 0}$ depends on the composition of the population and has the property

$$\sum_j d_j \mathsf{interact}_Z(Z_j, \omega) = 1 \tag{2.52}$$

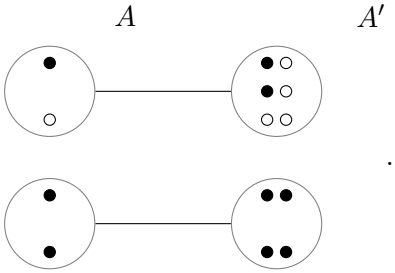
for a population $\omega = \sum_j d_j|Z_j\rangle$. Condition (2.51) may be drawn in terms of string

diagrams as follows




(2.53)

In terms of populations, the action of soft selection may be represented as




(2.54)

3. Hard selection Hard selection differs from the other two modes of selection $PPZ \to PPZ$ in that it pays no regard to the metapopulation structure and therefore realises non-social individual selection in multilevel selection (that is selection that does not depend on the social environment; see 'non-social trait' (Okasha, 2006)). Hard selection cannot be implemented as function $\mathsf{select}^{\mathsf{hard}}_{PPZ} : PPZ \to PPZ$ off-hand because it is not clear how individuals should be assigned back to populations after they replicated in the arena given by the metapopulation.[9]

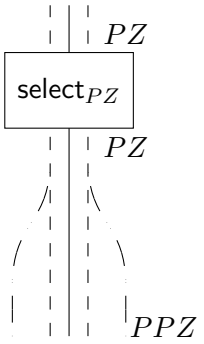
   The function $\mu_Z : PPZ \to PZ$ (2.25) from Section "Populations and probability monads" maps distributions of distributions to distributions – or metapopulations to populations – by considering the normal-weighted sum of the distributions as

[9] It is for this reason that we (and the Price equation more generally) consider an episode of multilevel selection as a map $PPZ \to PZ$ that maps metapopulations to single populations without further structure rather than as a map $PPZ \to PPZ$ that maps metapopulations to metapopulations.

distribution – or by regarding all individuals in the metapopulation as members of *one* population (see example (2.26)). Using $\mu_Z$ we can represent hard selection as

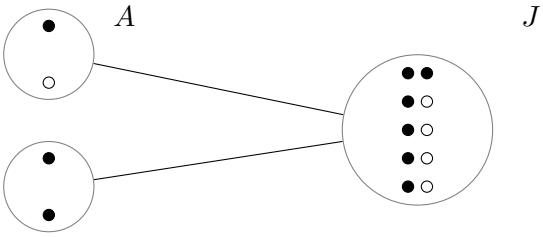
$$\mathsf{select}^{\mathsf{hard}}_{PPZ} = \mathsf{select}_{PZ}\mu_Z : PPZ \to PZ \tag{2.55}$$

with a selection function $\mathsf{select}_{PZ} : PZ \to PZ$ that realises individual selection in a population. This composition of functions is given as string diagram




(2.56)

that, in terms of populations, takes the form




(2.57)

### 2.1.6 Interactors and replicators

Population selection is involved in a range of natural selective scenarios that differ by significance and role of both population reproduction and individual reproduction. In biofilms, loosely interacting bacteria form populations with well-defined properties that affect reproduction of the individual bacteria within (Ereshefsky and Pedroso, 2013). Individual cells may act as social niche constructors and grow colonies that not only create a favourable environment for the initial cells and their offspring but that also reproduce by fission or budding (Ryan, Powers, and Watson, 2016). In planaria, populations of niche constructing cells are moreover able to reproduce sexually (Fields and Levin, 2018) (see Section "Epistatic effects" below). Finally, in animals with strict germline/soma separation – such as giraffes – reproduction on the level of individual cells is tightly coordinated across the population (that is the giraffe's organism) and ensures near-identical reproduction on the population level. Disruption of this population-wide coordination of individual level reproduction may manifest, for example, as cancer (Aktipis et al., 2015).

In addition to capturing these and similar examples, the concept of multilevel selection (of which individual selection and population selection are components) should allow a diachronic view on *transitions in individuality* (Okasha, 2006). During evolutionary transitions in individuality, reproductive properties of populations change in parallel with reproductive conditions, i.e., selection pressures, experienced by the individuals within (Watson and Thies, 2019).

Hull (1980) makes a useful distinction between the aspects of *interaction* and *replication* of a biological system that together warrant reproduction, or persistence, of the system. The distinction allows a stringent description of the selection process that motivates our approach to population selection and multilevel selection more generally:

> **replicator** an entity that passes on its structure directly in replication
>
> **interactor** an entity that directly interacts as a cohesive whole with its environment in such a way that replication is differential
>
> With the aid of these two technical terms, the selection process itself can be defined:
>
> **selection** a process in which the differential extinction and proliferation of interactors cause the differential perpetuation of the replicators that produced them
>
> Hull (1980): "Individuality and Selection"

With this definition, *reproduction* – that is the individual process that entails selection on the population level and, more generally, the succession of generations referred to by the full Price equation (1.2) – may be understood as composed of the two corresponding aspects replication and interaction. We will see later in this Chapter 2 that this additional structure of the process is mirrored in its formal represention: while the interaction of the individual with its environment determines the output of its replication (*interactor*) the replication of the individual takes place, thus determining the outcome of future interaction (*replicator*). Replicator and interactor are interleaved across parent and offspring as indicated in Figure 2.2 so that each of the two aspects is dependent on the premises provided by the other.

If a system comprises replicators and interactors, the difference between the two is functional and not necessarily physical, that is replicator and interactor do not have to be realised in disjoint subsystems or interact through distinct causal/functional pathways. The replicator assumes functions that serve to replicate an operational genotypic

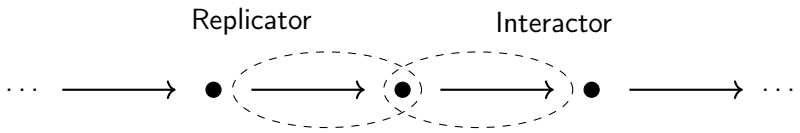



**Figure 2.2:** In a sequence of parent and offspring individuals, the interactor ensures that the replicator it developed from is re-instantiated as viable replicator. The replicator entails a copy of the interactor whence it originated.

instance, that is to produce a copy of the parent. The interactor, on the other hand, assumes functions that serve to maintain the entity in a way that ensures successful replication. An interactor interacts with the environment to ensure survival and successful operation of the replicator(s) it forms a unit with.

**Replicator** A genome[10], that is a carrier of a genotype, is a paradigmatic biological replicator because it passes on its structure directly and largely intact during meiosis. An organism, that is a carrier of a phenotype, is a paradigmatic interactor because – from the viewpoint of genetic selection – it serves no other purpose but to facilitate the replication of the genome from which it developed (Dawkins, 2006).

While genes are clearly replicators, to regard populations that are recurring temporal assemblies of replicators as replicators in their own right may stretch the concept too far (Lloyd and Wade, 2019). Even if lineages of populations can be identified, along which structure is transmitted from parent to offspring, this structure is often passed on neither directly nor intact. In the space between these examples, things are less clear. While genetical organisms such as asexually reproducing bacteria form lineages of near identical replicators from one perspective, bacteria also are "extremely temporary manifestations" of the genes that are passed on during bacterial reproduction and that form lineages of replicators in a much stricter sense (Hull, 1980, Section "Levels of Replication"). For sexually reproducing organisms, the difference between these two viewpoints becomes greater still.

**Interactor** An interactor entails its own reproduction, or persistence (Papale, 2020), – more precisely, the replication of the parts that entail the interactor – through interaction with the environment. A paradigmatic example of an interactor is the organism of a genetic individual. An organism may be viewed as temporary and

[10] All organisms in this thesis replicate asexually by cloning, so that offspring is genetically identical to its parent. In particular, mutation and recombination are not considered.

non-reproducing vehicle for the reproduction of those genes that developed into the organism and that the organism functions to pass on to its offspring.

While organisms are clearly interactors, the question in which sense genes are interactors is difficult. Entities ancestral to genes in early evolution that were selected for, e.g., stability rather than reproduction (Borrelli et al., 2015), may deserve the term 'interactor', but genes within the organism of an animal hardly interact with an external environment directly. Genetic replication in the context of an organism is regulated on the cell or organism level. The (conceptual) change of the environment of a gene from the selective environment experienced by an interactor to the developmental environment experienced by a replicator is characteristic of evolutionary transitions in individuality (Watson and Thies, 2019).

The difference between interactors and replicators parallels the distinction between multilevel selection [1] and multilevel selection [2] within multilevel selection (Table 2.1) (see Damuth and Heisler (1988), Section "Two meanings of multilevel selection", and Okasha (2006)). In multilevel selection [1], the population an individual is part of together with individual phenotypic properties determine the individual's fitness among the individuals in a metapopulation of populations. Population fitness in multilevel selection [1] refers to a functional component of individual fitness, or causal factor in the process that determines individual fitness. This factor is associated with a mechanism that justifies viewing populations as interactors and is therefore shared by all individuals within a population as contextual character – the population phenotype $\overline{Z}$. Like the individual phenotype through the function (1.3), $\overline{Z}$ affects individual fitness through the function (2.1). However, reproduction of a population as in Section "Population selection" should be viewed abstractly as increasing the representation of all individuals in a population homogeneously within the metapopulation and not as physical replication of populations.

In correspondence with the distinction between multilevel selection [1] and multilevel selection [2] there are two ways in which reproduction occurs on the population level.

**Multilevel selection [1]** Individual reproduction within a population in a metapopulation entails reproduction of the population in the sense that the *metapopulation* is transformed in a manner compatible with the Price equation (1.2) with populations as individuals. The latter means that selection and transmission can be distinguished for *populations* given replication of the (lower level) individuals that constitute them. This form of population reproduction may be usefully referred to

**Table 2.1:** Forms of multilevel selection and role of populations in the selection process. Individuals are always assumed to be interactors *and* replicators in the models considered in this thesis.

| | Populations are | |
|---|---|---|
| | **interactors** | **replicators** |
| Individual selection | No | No |
| Multilevel selection [1] | Yes | No |
| Multilevel selection [2] | Yes | Yes |

as *persistence* (rather than reproduction (Papale, 2020)) and allows the definition of models of multilevel selection [1].

**Multilevel selection [2]** Populations themselves replicate as units, that is populations are to be considered individuals in their own right. Moreover, the individuals that make up the populations may undergo parallel Darwinian dynamics on the individual level (Godfrey-Smith, 2009).

This thesis is focused on multilevel selection [1] as this is the context in which the inconsistency of the Price approach to multilevel selection and contextual analysis is manifest (Okasha, 2006).

### 2.1.7 Population phenotype

The population phenotype $\overline{Z}$ reflects properties of populations like the individual phenotype $Z$ reflects properties of individuals. While $Z$ is a function of the genotype an individual develops from (and the environment, see Section "Genotypes and phenotypes"), $\overline{Z}$ is assumed to be a function of the individual phenotype, manyfold expressed by the individuals the population comprises. We will discuss the structure of this function from a mathematical point of view in the course of this Chapter 2. Note that we do not assume the population phenotype to be the average of the individual phenotypes expressed in the population but allow the former to be an arbitrary function of the latter (see Section "Population phenotype and individual phenotypes").

In the following two sections, we discuss two different ways in which population phenotype may be a function of individual phenotype. These two ways correspond broadly to the two viewpoints in the 'Wright-Fisher controversy' described in Wade and Goodnight (1998). The distinctive issues are summarised in the table reproduced in Figure 2.3. We

1538 M. J. WADE AND C. J. GOODNIGHT

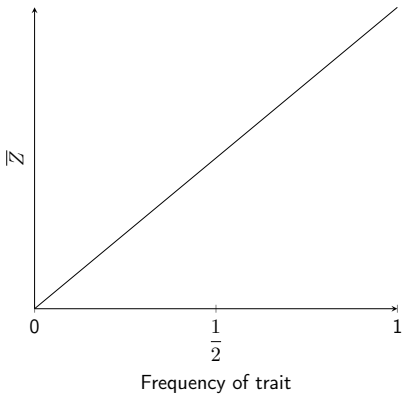
TABLE 1. The fundamental differences in emphasis that underlie the Wright-Fisher controversy.

| | Wright | Fisher |
|---|---|---|
| Central problem of evolutionary theory | Origin of adaptive novelty in a constantly changing environment | Refinement of existing adaptation in a stable or slowly changing environment |
| Major processes of evolutionary change | Combination of local natural selection, random genetic drift, migration, and interdemic selection | Mutation and natural selection |
| Ecological context of evolution | Small, subdivided populations | Large, panmictic populations |
| Genetic basis of evolutionary change | Epistasis and pleiotropy; context-dependence of allelic effects | Additive genetic effects; context-independence of allelic effects |
| Process of speciation | Inevitable by-product of local adaptation in epistatic systems | Disruptive or locally divergent selection |

**Figure 2.3:** Table 1 from Wade and Goodnight (1998) is reproduced here to provide context for the notions of additivity and epistasis but is not discussed in detail.

focus on the penultimate row that refers to the question how phenotypic properties of parts may give rise to phenotypic properties of a corresponding whole.

1. Additive effects The evolution of cooperation, or altruism, is the focus of much research in social evolution (Hamilton, 1964; Wilson, 1975, 1980; Axelrod and Hamilton, 1981). Associated models refer to the evolution of an altruistic trait, that is an individual trait that decreases individual fitness of cooperators (that is bearers of the altruistic trait) while increasing fitness of a cooperator's ambient population (see Section "Population selection") (for a definition, see Uyenoyama and Feldman (1980), page 381; West, Griffin, and Gardner (2007)). This fitness effect is assumed to be proportional to an additive phenotypic effect of the altruistic trait on the population phenotype such that both change in the same way as the frequency of cooperators varies in the population. The graph




(2.58)

is a simple example of such an additive effect. In this example, population fitness is proportional to $\overline{Z}$ such that (1) a cooperator has a fitness-increasing effect on population fitness compared to a defector, and (2) individual fitness of a cooperator is lower than that of a defector in the same population.

Many examples of such traits can be found among bacteria, such as *Pseudomonas aeruginosa* where cooperators secrete and share compounds within spatially separated populations of individuals (Weigert and Kümmerli, 2017). These compounds change the population phenotype by improving the nutritional environment of each individual within. However, the costly production of these compounds puts cooperators at a disadvantage relative to non-cooperators in their population.

From the viewpoint of traditional individualistic evolution, given the conflict between the levels of selection inherent in altruistic traits, natural systems whose function relies on cooperation of their parts require explanation. Often short term fitness benefits of not cooperating are in conflict with long term fitness benefits of cooperating.

To give a biological example reported in Eldakar, Wilson, et al. (2010), male water striders *Aquarius remigis* organised into patches – that is populations such that the complete experimental setup comprises a metapopulation – are subject to individual selection and population selection that both are affected by an aggressiveness trait individually expressed in males. Aggressiveness has a positive effect on individual fitness that is higher for more aggressive males that secure more mating opportunities than less aggressive males. However, aggressiveness has the following negative effect on *male* population fitness: since the aggression experienced by females reflects the cumulative aggressiveness of males on their patch and females tend to avoid aggression by leaving their current patch, the trait has a negative effect on male fitness by decreasing the number of females, thus diminishing the reproductive resources of all males on a patch.

The assumptions associated with the two columns in Figure 2.3, each taken to the extreme, lead to opposite predictions for the outcome of this experiment. Individual aggressiveness in the metapopulation *increases* with selection if individual selection is stronger than population selection. For instance, females may respond to the population phenotype 'patch aggressiveness' only weakly, such that the slope of the population selection differential is gentle, while aggressive males are much more successful at reproduction than their non-aggressive patch mates, such that

the slope of the individual selection differential is steep. However, individual aggressiveness in the metapopulation *decreases* with selection if population selection dominates individual selection. For instance, if patches are inhabited by genetically closely related individuals such that individual selection within patches is ineffective and if sufficient variation of the population phenotype in the metapopulation allows population selection then, in total, selection may favour cooperators over defectors. Cooperative behaviour within family groups of many animal species has been explained with this effect (Abbot et al., 2011).

The configuration of selection described in the end of the preceding paragraph is a basic model of *kin selection* (Gardner, West, and Wild, 2011). The assumption of additivity of phenotypic effects (Levin and Grafen, 2019) defines a class of systems for which the causal interpretation/functional representation admits an expression of the total effect on the population phenotype in terms of the separate effects of the phenotypic properties of each of the individuals in the population. This viewpoint has been very successful in explaining many issues in social evolution (Gardner, West, and Wild, 2011). Moreover, additivity allows "to adopt inclusive fitness as an organizing framework for understanding social behavior" (Levin and Grafen, 2019, page 1066). This organising framework, described by Grafen (2006), admits viewing inclusive fitness as quantity that is maximised by natural selection, a principle that is an "indispensable" approximation for understanding social evolution.[11] Nonadditive phenotypic effects, referred to as epistatic effects, lie outside the scope of inclusive fitness theory and are discussed in the next section.

2. Epistatic effects

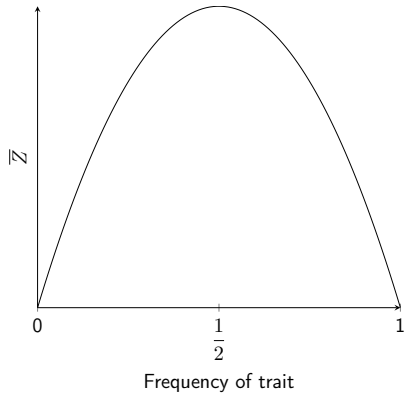
When the effect of the phenotype of an individual on the phenotype of its population depends on the phenotypic composition of the population, the assumption of additivity described in the previous section is not justified. Epistatic, or nonadditive, effects arise from functional complementarity between instances of the phenotype (i.e., individuals) (Wade and Goodnight, 1998; Manrubia, 2021). Similar to the altruistic trait that defines cooperators and defectors in Section "Additive effects" we may consider a synergistic trait whose effect on the population pheno-

[11] Inclusive fitness is indispensable for understanding social evolution in a similar way that general relativity is indispensable for understanding physics. Both theories have been applied successfully over many years and are crucial to our understanding. More widely within social evolution and physics, however, the scope of each theory is limited, and other theories are better suited to describe phenomena that lie outside this scope.

type arises in interaction with its counterpart expressed by bearers of variants of the trait. By the assumption of nonadditivity, the effect of the trait cannot be determined independent of the population it is expressed in. The phenotypic properties of bearer and non-bearer of the trait correspond to complementary parts of a shared function – expression of the population phenotype.

In the following paragraphs, two population phenotypes – population fitness and population size – are discussed briefly as examples. The upshot of the examples is that the functional description of epistatic effects requires individual phenotypic effects to be dependent on the composition of the population as reflected in function graph (2.59). Additive models of epistatic effects, such as those suitable for additive effects of cooperation discussed in Section "Additive effects", miss this dependency (Smith and Inglis, 2021).

a) Population fitness To give an abstract example, Michod (2006) describes a model where the fitness of an evolutionary unit is a function of the two components fecundity (reproduction) and viability (survival). Simplifying the original model, we assume that the individual phenotype can take the values 'survival' or 'reproduction' depending on whether the individual contributes to survival or reproduction of the population it is part of[12]. Then a population phenotype $\overline{Z}$ that is the product of the total individual contributions to fecundity and viability of the population may depend on the composition of the population as follows




(2.59)

If the phenotype of the population depends on the degree to which the individuals within the population collectively execute the complementary functions

[12] In Michod (2006) the two functions are considered in several trade-off relations, as in Fig. 3 in Michod (2006).

that give rise to this phenotype, the population must contain individual contributions to both survival and reproduction, like somatic cells and germ cells in a multicellular organism (Grochau-Wright et al., 2017). In graph (2.59), $\overline{Z}$ is maximised if survival and reproduction are represented in equal parts in the population (the value $\frac{1}{2}$ was chosen for illustration but is not assumed generally). Often germ cells are rare among all cells of a multicellular organism since somatic cells constitute most of the physical structure of the organism while germ cells only constitute a part of the reproductive system. It is crucial, however, that both cell types are present and that a graph corresponding to graph (2.59) reaches non-expression of the trait at the ends of the interval (Watson and Thies, 2019). For, without germ cells there is no faithful reproduction of the population (that is the multicellular organism) but only individual reproduction that manifests, for instance, as growth of the organism. Conversely, if all cells of a developing organism enter an embryonic state and forego reproduction until they get the opportunity to germinate, there is no organismal structure to facilitate reproduction through these germs. In the terms discussed in Section "Interactors and replicators", in the former case the multicellular organism may be an interactor (if the fact that the cells are organised in a multicellular organism affects their fitness) but it isn't a replicator (because there are no germ cells to effect reproduction of the organism as a whole). In the latter case the population may be a replicator (if the germ cells germinate without soma) but it isn't an interactor (as there is no physical structure to interact with the selective environment).

b) Population size Wade (1976) experimentally selected populations of flour beetles, *Tribolium castaneum*, for the population phenotype given by population size, or *magnitude*, at a certain 'age' of the population after germinating from a propagule of 16 individuals.[13] Because of the strong response observed, these experiments furnish a paradigmatic empirical example of a response to population selection.

   In a metapopulation of 48 populations of initially 16 individuals each, populations were assigned to one of four experimental treatment groups, each of which was subjected to a different regime of artificial selection on the individual and the population level. Figure 2.4 gives an overview of the four

[13] Wade (1976) uses the terms 'population' and 'individual' in our sense. However, our 'population selection' is referred to as 'group selection'.

conditions of the experiment. In all conditions, 48 propagules were assem-

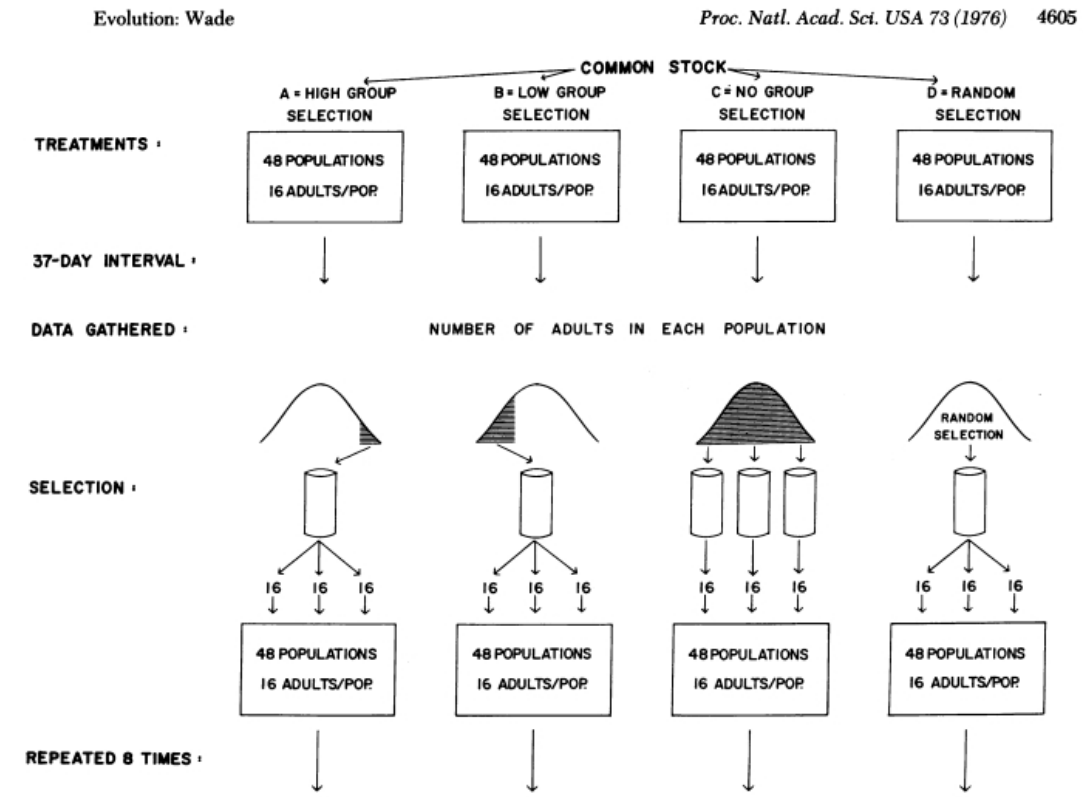



**Figure 2.4:** The selection pattern in condition A realises population selection for large population size and (individual) soft selection for high reproductive rate within the population. Condition B selects for small size on the population level but like condition A for high reproductive rate on the individual level. Conditions C and D realise *soft selection* for high reproductive rate without and with random population selection. resp. (see Section "Soft selection and hard selection"). Figure reproduced from Wade (1976).

bled from a parent metapopulation of 48 populations to found an offspring metapopulation of 48 populations. Each propagule consisted of 16 individuals randomly chosen from *one* of the populations in the parent generation, hence individuals in all conditions underwent soft selection (for individual reproductive rate) (see Section "Soft selection and hard selection"). The conditions differ in the choice of parent populations that contribute propagules to the offspring generation. The pattern of this choice determines strength and di-

rection of population selection as described in Section "Population selection".

In conditions A and B, each propagule was assembled from individuals randomly chosen from the population with the largest magnitude (condition A) or the smallest magnitude (condition B) – and the population with the next largest (smallest) magnitude after the previous has been portioned into propagules. Conditions A and B represent population selection for large and small population magnitude at day 37 after germination, resp.

In a control treatment (condition C), population selection is avoided. To this end, *one* propagule is assembled from each of the parent populations so that population fitness is independent of population phenotype. Soft selection continues to act in condition C, because the chance for an individual to be selected into the propagule of its population depends on its (relative) fitness within that population. Finally in condition D, population selection is randomised by randomly picking populations for selection. Since unlike in condition C selected populations contribute more propagules when they reach a larger population size, population selection (for large magnitude) is in effect, as well as soft selection as before. The response to experimental selection over several generations of metapopulations is shown in Figure 2.5.

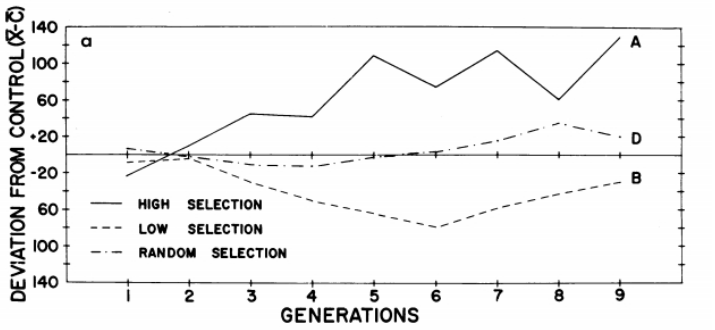



**Figure 2.5:** "For generations three through six, $A > C = D > B$ with $P << 0.005$; for generations seven through nine, $A > D > C > B$ with $P << 0.005$." Figure reproduced from Wade (1976).

The strong response to population selection (compare $A$ to the horizontal axis that represents the control treatment) is explained with the selection of several traits – such as fecundity, developmental time, and cannibalism rate – *in combination*. Under condition C (individual selection, no population se-

lection), mean adult population size declined due to increased cannibalism of adult *Tribolium* on eggs and pupae, a trait selected for by individual selection within the populations. Population selection for large and small population magnitude (conditions A and B) lead to strong responses in both directions that are possible due to the selection of combinations of individual traits in propagules (Wade, 1976; Wade and Griesemer, 1998). These trait combinations with epistatic effects give rise to favourable population phenotypes that are inherited on the population level because the propagules of 16 individuals transmit a sufficient part of the within-population variation responsible for the epistatic effect to the offspring populations.

3. Population phenotype and individual phenotypes As mentioned in the beginning of this Section, we assume the population phenotype $\overline{Z}$ to be a function of the distribution of individual phenotypes in the population. In additive scenarios it is often appropriate to use the average of the individual phenotypes in the population as population phenotype, for instance if the level of cooperation is a good indicator of the effect of cooperation within a population on its fitness. In the models described in this thesis, in order to cater for epistatic effects and other effects that arise from the interaction between individual phenotypes in a population, we do not assume the population phenotype to be the average of the individual phenotypes occuring in the population. When population fitness depends on combinations of individual phenotypes the population phenotype may be represented more adequately by functions that reflect the mechanism by which it arises from individual phenotypes. In the experiments with flour beetles described above, for instance, a suite of individual traits is thought to contribute to population fitness by creating a social environment, i.e., a population phenotype, that allows population size to increase. The quality of this environment depends on the combination of several individual phenotypes such as developmental time and cannibalism rate. Farine, Montiglio, and Spiegel (2015) discuss role and relevance of *group phenotypic composition* (GPC) as factor that is equivalent to our notion of population composition and that, through expression of a population phenotype, affects population-level outcomes. For instance, niche partitioning within populations allows larger population sizes due to reduced competition for resources. A GPC yielding an optimal population phenotype in this respect would consist of different individuals specialising in specific resources so that the overlap between the niches is minimal.

Therefore the variance of the distribution of individual phenotypes rather than its average best reflects the dependence of population phenotype on individual phenotypes. In fire ants, the acceptance of multiple queens into the colony, a potentially import factor for the reproductive output of the colony, depends on a "threshold frequency" of tolerant workers (Farine, Montiglio, and Spiegel, 2015). The presence of "keystone individuals" in a population as factor for population outcomes, for instance through collective decision making, is another example of a population phenotype that is not given as average over individual phenotypes.

## 2.2 Units of selection

So far we have treated the entities that undergo the process described by the Price equation (1.2) in an abstract manner as circles in diagrams or as fractions of populations (that is point values of distributions over discrete types). In this section we describe in more detail how the mathematical structures ('actors' and their interactions) relate to the biological entities and processes they describe. This section sets the biological stage for the following formalisation. The crucial maps are introduced out of the definition of evolution given by Lewontin (1970).

### 2.2.1 Variation, selection, and heredity

The *unit of selection*, "the entity upon which selection acts" (Gardner, 2015; page 306), is each of the individuals that make up a population, each holder of the index $i$ in equation (1.2), and each circle in diagrams (1.4) and (1.5). Following the account of Godfrey-Smith (2009) (and its extension in Papale (2020)), we consider a unit of selection an object of a type, such as an individual of a biological species, that is involved in a natural process realising the three Darwinian principles of evolution:

1. Different individuals in a population have different morphologies, physiologies, and behaviors (**phenotypic variation**).
2. Different phenotypes have different rates of survival and reproduction in different environments (**differential fitness**).
3. There is a correlation between parents and offspring in the contribution of each to future generations (**fitness is heritable**).[14]

[14] It seems not settled whether this definition of Darwin's principle of heredity or the definition in Uyenoyama and Feldman (1981) on page 393 is preferable. The latter definition holds that the

Lewontin (1970): "The Units of Selection"

In the following sections, we discuss how these three principles correspond to elements of our description.

1. Phenotypic variation Selection occurs with respect to a "character under selection" (Gardner, 2015). For any individual, this character, through expression in the individual, determines the outcome of that individual's interaction with the selective environment, that is its absolute fitness, or number of offspring. We refer to the character under selection as phenotype $Z$. An instance of the phenotype is a phenotypic value expressed in an individual such as the values $z_i : Z$ in equation (1.2).

   The expression of character values in the individuals within any population $K$ corresponds to a map

   $$z_K : K \to Z \tag{2.60}$$

   that assigns to each individual in population $K$ its phenotypic value in $Z$. We may think of this map as determining the trait value of an individual such as measuring a giraffe's neck length in metres. For an individual $i\colon I$ in a given population $I$, we may write $i$'s phenotypic value as $z_i = z_I(i) : Z$.

   The expression of a phenotype in a biological individual considered in this thesis is a function of genotype and environment. This distinction is important because while selection *depends* on the genotype only indirectly through the phenotype[15], it *determines* the genotypic composition of the offspring population directly because inheritance is genetic. This issue is discussed in more detail in Section "Genotypes and phenotypes" below.

   For a population to "undergo evolutionary change" (Lewontin, 1970) in the sense of changing allele frequencies, phenotypic variation within the population is required, for otherwise no difference between individuals in the population can be made by selection and no directed change in genotypic composition can ensue.

*phenotypes* of parent and offspring are correlated rather than their fitnesses. Under constant selective conditions, heredity of phenotype implies heredity of fitness but not the other way around. In this thesis, we use the stronger notion and assume that the phenotype/genotype itself is heritable.

[15] Selection 'sees', or processes, the phenotypic properties of an individual through the individual's interaction with the selective environment. Genetic differences between individuals without phenotypic differences are without consequence for selection (see Section "Interactors and replicators").

For any given population $K$ and using function 2.60, the first of Lewontin's principles may be put simply as

$$z_K \neq \mathsf{const.} \tag{2.61}$$

2. Differential fitness Absolute fitness of a phenotype is an assignment of offspring number to phenotypic values, that is a function $w\colon Z \to \mathbb{N}$ as in equation (1.3). This function is an abstraction of the process of reproduction occurring over an individual's lifetime that expresses the total number of offspring an individual entails in terms of the individual's phenotypic properties. In this thesis, we assume the selective environment[16] that facilitates this entailment to be constant across populations, including populations in parent-offspring relation. Consequently, we do not indicate the environment in the notation.

   A population of individuals under selection is transformed to its offspring population according to a fitness law (1.3). Directed evolutionary change, or an evolutionary response to selection, is a change in allele frequencies so that those alleles increase in representation that are associated with successful phenotypic properties of the individuals they constitute. A population of individuals undergoes directed evolutionary change over the succession of its offspring populations if fitness differs between individual instances of the phenotype in a manner that brings forth appropriate transformations.

   We assume so far that absolute individual fitness (1.3), that is the number of offspring of an individual, is independent of the population the individual is situated in. Relaxing this assumption marks the transition from 'standard' Darwinian evolution of individuals to *social evolution* (Frank, 1998; Okasha, 2006; West, Griffin, and Gardner, 2007; Wade, 2016). In the former, individuals, or rather clones, compete independently for representation in the population by individually interacting with the selective environment. This interaction, mediated by the phenotype, entails the reproduction of the individual, including the multiplicity of this reproduction (that is absolute fitness of the individual). The outcome of this interaction does not depend on properties of the population by assumption. In social evolution, on the other hand, absolute fitness of individuals is at least to some degree determined by phenotypic interactions with *other individuals* so that absolute individual fitness

[16] The term *selective environment* refers to aspects of the individual's environment that determine its fitness. This is in contrast to the developmental environment that interacts with the genotype to yield the phenotype.

becomes a function of contextual characters such as the phenotypic composition of the population as a whole. These interactions may be competitive in nature, for example when individuals of the same species share a limited resource or in predator-prey interactions. However, as we saw in Section "Epistatic effects", interactions in social evolution may also be cooperative or synergistic (Watson and Thies, 2019). It is these latter cases in which the outcomes of social evolution often differ from what might be expected from the Darwinian catchphrase of the 'struggle for existence' among competing variants of individuals.

It should be noted that also in the case that absolute individual fitness is independent of the population the outcome of evolution depends on the composition of the population since relative fitness determines adaptation of the population. It is often taken for granted that a beneficial mutation can fixate in a population of asexually reproducing individuals due to the fitness benefit it confers on the lineage of the individual it arises in. This assumes that the mutation maintains its fitness advantage over other individuals in the population until it fixates. Subsequently, other beneficial mutations which again confer a fitness advantage over the new background can spread. In this way, beneficial mutations can accumulate, thus adapting the population over time. However, if after one beneficial mutation in one individual another beneficial mutation occurs in another individual, *clonal interference* may slow down or even hinder adaptation. This is due to the fact that the second mutation decreases the fitness advantage of the first relative to the population. The models constructed in this thesis can describe clonal interference since selection is implemented based on relative fitness in the population.

3. Fitness is heritable The third principle of Lewontin's description of the Darwinian process of evolution refers to the continuity with respect to fitness (and phenotype) between parent and offspring. This continuity, that for asexual genetic inheritance in a constant environment is realised as near identity[17], allows a response to selection in the sense that fitter alleles – with respect to the conditions given by the constant selective environment and the constant developmental environment – increase in representation over several generations of individuals.

[17] How near depends on the organism since in addition to mutation, other mechanisms may lead to discrete changes in genotype from parent to offspring. With lateral gene transfer as discussed in Doolittle and Bapteste (2007), for example, the outcome of asexual individual reproduction becomes a function of the population.

The notion of heritability of fitness (or phenotype) relies on two structures. First, through the notion of heritability itself, it postulates a parent-offspring relation. When reproduction is asexual, as assumed in this thesis, this relation may be represented succinctly as a map

$$\mathsf{lin}\colon I' \to I \tag{2.62}$$

that assigns each individual of a population its parent in the parent population (Hull, 1980). We refer to this map as *lineage map*. When it is applied repeatedly it traces out the lineage of its initial argument – any individual – to the first ancestor.

Since we consider individual selection without transmission bias (that is no mutation, no phenotypic plasticity/constant developmental environment), phenotypic heredity is perfect and we have

$$\begin{array}{ccc} I & \xrightarrow{z_I} & Z \\ {\scriptstyle \mathsf{lin}}\uparrow & \nearrow_{z_{I'}} & \\ I', & & \end{array} \tag{2.63}$$

that is $z_I \mathsf{lin} = z_{I'}\colon I' \to Z$, or offspring and parent are phenotypically identical (the maps $z_I$ and $z_{I'}$ are defined as in function (2.60); see Section "Genotypes and phenotypes" below).

The second part of Lewontin's third principle, that is the postulated correlation between parent and offspring fitness, may be represented as an association

$$\begin{array}{ccccc} I & \xrightarrow{z_I} & Z & \xrightarrow{w} & \mathbb{N} \\ {\scriptstyle \mathsf{lin}}\uparrow & & & & \\ I' & \xrightarrow{z_{I'}} & Z & \xrightarrow{w} & \mathbb{N} \end{array} \tag{2.64}$$

that shows two ways of assigning each individual in the offspring population a fitness value: given an individual in $I'$, we may consider its own fitness and that of its parent.

If these two quantities are uncorrelated, that is statistically independent, high and low fitness of a parent may equally likely be followed by high or low fitness of its offspring. Any corresponding increase or decrease in representation of a phenotypic value is hence nullified over time, and directed evolutionary change is not to be expected.

Our assumption on ideal phenotypic inheritance from parent to offspring (2.63) lets us change diagram (2.64) to

$$\begin{array}{ccccc} I & \xrightarrow{z_I} & Z & \xrightarrow{w} & \mathbb{N} \\ {\scriptstyle \mathsf{lin}}\uparrow & & {\scriptstyle 1_Z}\uparrow & & \\ I' & \xrightarrow{z_{I'}} & Z & \xrightarrow{w} & \mathbb{N} \end{array} \tag{2.65}$$

because $1_Z z_{I'} = z_{I'} = z_I \mathsf{lin}$. In this thesis, we assume the selective environment to be constant across populations and over 'time', that is over the succession of populations in parent-offspring relation. This means that for any population $K$, absolute fitness of its constituent individuals is determined through $K \xrightarrow{z_K} Z \xrightarrow{w} \mathbb{N}$ by the same law $w\colon Z \to \mathbb{N}$. This means in particular that the two instances in diagrams (2.64) and (2.65) are the same, so that (2.65) may be collapsed to

$$\begin{array}{ccccc} I & \xrightarrow{z_I} & Z & \xrightarrow{w} & \mathbb{N} \\ {\scriptstyle \mathsf{lin}}\uparrow & \nearrow_{z_{I'}} & & & \\ I'. & & & & \end{array} \tag{2.66}$$

In this case of ideal inheritance (2.63) and constant selective environment, both of which we assume throughout the thesis on the individual level, Lewontin's third principle of heritable fitness realises the postulated correlation as identity, that is fitness is inherited unchanged.

### 2.2.2 Individual heredity

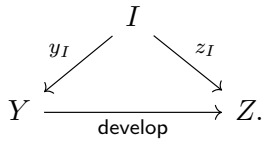
In equation (2.60) we defined the map $z_K\colon K \to Z$ that for a population $K$ determines the phenotypic values of the individuals within. The phenotype $Z$ describes phenotypic properties (such as giraffe neck length) that determine the outcome of an individual's interaction with the selective environment, that is the individual's absolute fitness in that environment. The phenotype itself, however, is not inherited. Instead, individuals comprise a genome, and heredity is the result of genetic reproduction (that is assumed asexual in this thesis). Phenotypic properties of an individual are the outcome of a genotype-environment-phenotype map that determines the phenotypic value in $Z$ given the genotypic value in $Y$, that is the genome, and properties of the developmental environment. Since we assume the developmental environment constant, we refer to this map more simply as the genotype-phenotype map

$$\mathsf{develop}\colon Y \to Z. \tag{2.67}$$

With the map

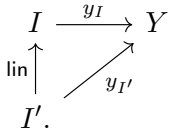
$$y_K \colon K \to Y \tag{2.68}$$

that assigns to an individual in any population $K$ its genotype (as in function (2.60) for the phenotype), the identity of the bearer of genotype and phenotype implies the diagram



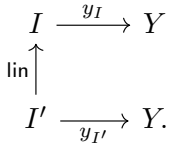

(2.69)

Similar to diagram (2.63), we may represent perfect genetic inheritance diagrammatically as



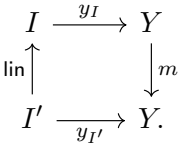

(2.70)

If we consider genotypic inheritance more generally, the parent offspring relation implies no more than an association




(2.71)

From this viewpoint, mutation then appears as genotypic transformation $m \colon Y \to Y$ in the diagram




(2.72)

Other changes in genotypic value from parent to offspring, such as recombination or lateral gene transfer (Doolittle and Bapteste, 2007; Bapteste and Papale, 2020), may be represented similarly as functions that depend on the genetic composition of the population in addition to the genetic composition of the parent individual.

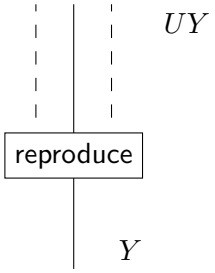
The Price equation (1.2) describes the transformation of a parent population to an offspring population in terms of change in the phenotypic composition of the population[18]. It partitions this change into reproduction and phenotypic change: the former represents the change in representation of parent compared to (multiple) offspring in

[18] If the population has no further structure, the phenotypic composition of a population may be represented by mean and higher moments of the distribution given by the phenotypic values of the individuals that make up the population.

their respective populations due to fitness differences between individuals and the latter the change in phenotype from parent to offspring due to recombination, mutation and other mechanisms that entail phenotypic differences between parent and offspring.

In the case of genetic reproduction of individuals, underlying this transformation of phenotypic composition is a transformation of *genotypic composition* of the population, and it is the latter that is required for a response to selection as described in the preceding sections of this chapter. This is because the genotype, not the phenotype, is passed from parent to offspring[19]. For this reason, in applications of the Price equation to genetic evolution the phenotype is often represented as *genetical character*, or breeding value, that is as function $\mathsf{develop}\colon Y \to Z$ that yields the phenotype in question in terms of the genotype (Gardner, West, and Barton, 2007; Gardner, 2015). Since this function is generally not invertible the relation between the genotypic and phenotypic transformations between parent and offspring population may be complicated. In this thesis, however, we assume that develop is invertible, such that phenotypic composition and genotypic composition may be regarded as identical. This assumption is justified, for instance, in models of monoploid organisms with one locus and two or more alleles at this locus that each correspond to a different phenotypic value. If develop is invertible we may identify $Y$ and $Z$ in diagram (2.69) and regard genotypic and phenotypic composition of a population as the same.

In non-social evolution, reproduction of populations is a process that derives an offspring population from a parent population. It is effected by the reproduction of the individuals that make up the population and occurs according to a law (1.3) that determines the offspring population of each individual, given as element of the unnormalised monad over the genotype $UY$ (see Section "Selection and absolute fitness"), according to its genotype

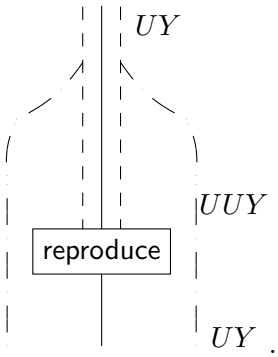



, (2.73)

Reproduction of the population is a superposition of individual reproduction: the off-

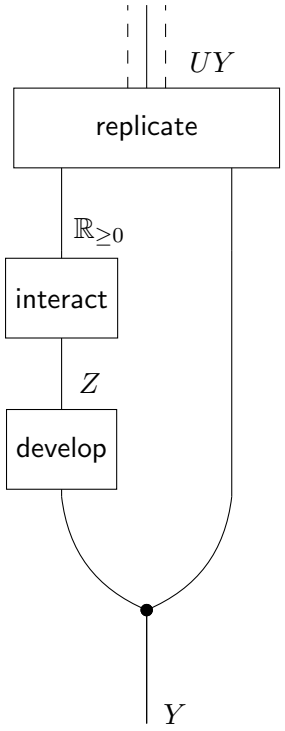
[19]Parental effects and other mechanisms of phenotypic inheritance do not feature in this thesis.

spring populations of the individuals are combined to yield the offspring generation




. (2.74)

In this thesis, reproduction is by cloning with multiplicity determined by offspring number via a law $\mathsf{interact} : Z \to \mathbb{R}_{\geq 0}$ that depends on the individual phenotype. With this law we can define a complete episode of reproduction as described in Section "Interactors and replicators" as follows:




. (2.75)

# 3 Models of multilevel selection

In this final chapter we put together the elements described above to construct models of multilevel selection. Our approach rests on the premise that selection in biological populations may be understood as differential replication of the units under selection. For the basic units termed *individuals* in this thesis, the generational progression from parent to offspring described by the Price equation (1.2) is given by selection (1.6) because replication on the individual level is identical, that is transmission bias is absent. For *populations*, we understand selection as differential replication, or persistence, as well. However, unlike for individuals, transmission effects are not void for populations. Instead, transmission effects in transformations from parent to offspring metapopulation are assumed to arise from individual selection across the metapopulation (contextual analysis) or within the populations (Price approach). Using the functions discussed in Section "Soft selection and hard selection" we describe the processes $\mathsf{select}^{\mathsf{CA}} : \mathsf{PPZ} \to \mathsf{PZ}$ and $\mathsf{select}^{\mathsf{PA}} : \mathsf{PPZ} \to \mathsf{PZ}$ in terms of probability monads and associated string diagrams.

## 3.1 Contextual analysis

Episodes of multilevel selection described by contextual analysis correspond to a function $\mathsf{select}^{\mathsf{CA}} : PPZ \to PZ$ that is composed of hard selection on the individual level and population selection on the population level that represent separate causal pathways. Since the metapopulation that forms the input of the process is involved in both, the variables are first copied and then concurrently subjected to hard selection and population selection that take place across the metapopulation as a whole so that the outcome of each is considered in absolute terms. Finally, the two metapopulations are combined, or superimposed, with the function

$$\mathsf{mix} : UZ \otimes UZ \to UZ, \alpha_1 \otimes \alpha_2 \mapsto \alpha_1 + \alpha_2 \tag{3.1}$$

that mixes two distributions linearly by combining them into one.

The representation of the functional structure underlying the model of contextual analysis is given in Figure 3.1.

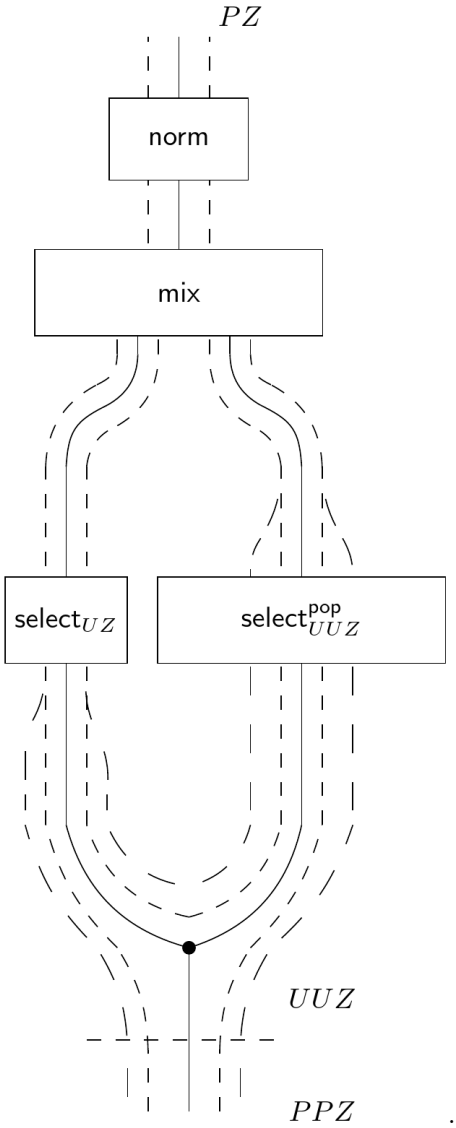



**Figure 3.1:** The process corresponding to contextual analysis. The string diagram is read from bottom to top. The normalised metapopulation is considered unnormalised before being copied at the dot. The two copies each undergo separate causal processes: selection on the individual level on the left and selection at the population level on the right. On the left, since individual selection is realised as hard selection in contextual analysis, the metapopulation is first collapsed to one population that is then subjected to individual selection. On the right, population selection may change the relative frequencies of the populations in the metapopulation while leaving the composition of the populations unchanged. The metapopulation is then collapsed before it is combined with the result of the left side in the function mix. Finally in norm, the population is normalised to a distribution in $PZ$.

Algebraically, for a metapopulation $\omega = \sum_i c_i|\omega_i\rangle : PPZ$ with $\omega_i = \sum_j d_{ij}|Z_j\rangle : PZ$ such that the population phenotype of population $\omega_i$ is given by $\overline{Z}_i : \overline{Z}$ the process is given by

$$\begin{aligned}\omega &= \sum_i c_i|\omega_i\rangle \mapsto \sum_i c_i|\omega_i\rangle \otimes \sum_i c_i|\omega_i\rangle \mapsto \sum_{ij} c_i d_{ij} \mathsf{interact}_Z(Z_j)|Z_j\rangle \otimes \sum_i c_i \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)|\omega_i\rangle \\ &= \sum_{ij} c_i d_{ij} \mathsf{interact}_Z(Z_j)|Z_j\rangle \otimes \sum_{ij} c_i d_{ij} \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)|Z_j\rangle \\ &\mapsto \mathsf{norm} \sum_{ij} c_i d_{ij} (\mathsf{interact}_Z(Z_j) + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i))|Z_j\rangle,\end{aligned}$$

absolute fitness of an individual with phenotype $Z_j$ within a population with phenotype $\overline{Z}_i$ is therefore

$$W_{ij} = \mathsf{interact}_Z(Z_j) + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i).$$

For phenotypic values in the real numbers and a linear choice of the $\mathsf{interact}$ functions such that $\mathsf{interact}_Z(Z_j) = \alpha + \beta_1 Z_j$ and $\mathsf{interact}_{\overline{Z}}(\overline{Z}_i) = \beta_2 \overline{Z}_i$ with parameters $\alpha, \beta_1, \beta_2 : \mathbb{R}$ this gives the multiple regression of contextual analysis

$$W_{ij} = \alpha + \beta_1 Z_j + \beta_2 \overline{Z}_i. \tag{3.2}$$

## 3.2 Price approach

The multilevel Price equation describes episodes of multilevel selection that are composed of soft selection $\mathsf{select}^{\mathsf{soft}}_{PPZ} : PPZ \to PPZ$ on the individual level and population selection $\mathsf{select}^{\mathsf{pop}}_{PPZ} : PPZ \to PPZ$ on the population level. Again the causal pathways of the individual and population level are assumed separate. We describe a multiplicative and an additive model that both respect the separation into individual effects (action of $\mathsf{interact}_Z$) and population effects (action of $\mathsf{interact}_{\overline{Z}}$) in the sense that the between-population covariance and the average within-population covariance of the multilevel Price equation are zero if $\mathsf{interact}_{\overline{Z}}$ or $\mathsf{interact}_Z$, resp., are constant, that is if population selection or individual selection is absent.

### 3.2.1 Multiplicative model

Since $\mathsf{select}^{\mathsf{soft}}_{PPZ}$ acts on the inner distributions only (see (2.53)) and $\mathsf{select}^{\mathsf{pop}}_{PPZ}$ leaves the inner distributions unchanged (see (2.48)), the functional structure underlying the Price approach may be represented as shown in Figure 3.2.

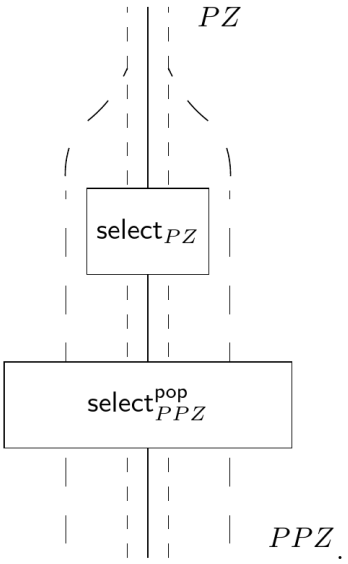



**Figure 3.2:** The multiplicative process corresponding to the Price approach. Population selection acts in the metapopulation, leaving the composition of the populations unchanged. Then soft individual selection acts within each population leaving the relative size of the populations unchanged.

To see that the process above separates individual selection and population selection according to the multilevel Price equation, we write down the process explicitly. Let the finitely many individual phenotypic values be denoted by $Z_1, \ldots, Z_n : Z$ and a metapopulation be given by $\sum_i c_i|\omega_i\rangle : PPZ$ with $c_i : \mathbb{R}, \sum_i c_i = 1$ and populations $\omega_i = \sum_j d_{ij}|Z_j\rangle : PZ$ with population phenotypes $\overline{Z}_i : \overline{Z}$, where the index $i$ covers the populations in the metapopulation and the index $j$ the individual phenotypic values. With $\mathsf{interact}_{\overline{Z}} : \overline{Z} \to \mathbb{R}_{\geq 0}$, the population selection process transforms the metapopulation as follows

$$\sum_i c_i|\omega_i\rangle \mapsto \frac{1}{\overline{W}} \sum_i c_i \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)|\omega_i\rangle : PPZ$$

with average population fitness $\overline{W} = \sum_i c_i \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)$.

The inner process of individual selection within populations (soft selection) is implemented by $\mathsf{interact}_Z : Z \otimes PZ \to \mathbb{R}_{\geq 0}$ (see Section "Soft selection and hard selection") and transforms the populations such that

$$\omega_i \mapsto \mathsf{select}\omega_i = \mathsf{select}(\sum_j d_{ij}|Z_j\rangle) = \sum_j d_{ij} \mathsf{interact}_Z(Z_j, \omega_i)|Z_j\rangle : PZ$$

with $\sum_j d_{ij} \mathsf{interact}_Z(Z_j, \omega_i) = 1$ for all $i$ (Equation (2.52)). The complete process in Figure 3.2 where individual and population selection occur in a nested manner is then

given by

$$\omega = \sum_i c_i|\omega_i\rangle \mapsto \frac{1}{\overline{W}} \sum_i c_i \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)| \sum_j d_{ij} \mathsf{interact}_Z(Z_j, \omega_i)|Z_j\rangle\rangle$$
$$\mapsto \frac{1}{\overline{W}} \sum_{i,j} c_i d_{ij} \mathsf{interact}_{\overline{Z}}(\overline{Z}_i) \mathsf{interact}_Z(Z_j, \omega_i)|Z_j\rangle : PZ$$

so that in a population with phenotype $\overline{Z}_i$ absolute fitness of an individual with phenotype $Z_j$ is

$$W_{ij} = \mathsf{interact}_{\overline{Z}}(\overline{Z}_i) \mathsf{interact}_Z(Z_j, \omega_i) \tag{3.3}$$

and average individual fitness in population $\omega_i$ is

$$W_i = \sum_j d_{ij} \mathsf{interact}_{\overline{Z}}(\overline{Z}_i) \mathsf{interact}_Z(Z_j, \omega_i) = \mathsf{interact}_{\overline{Z}}(\overline{Z}_i).$$

Denoting the average individual phenotype in population $i$ by $\hat{Z}_i$ the population covariance term of the multilevel Price equation is given by

$$\mathrm{Cov}(W_i, \hat{Z}_i) = \sum_i c_i (W_i - \overline{W}) \hat{Z}_i$$

and is independent of individual selection. The average individual covariance term amounts to

$$\mathrm{E}_I(\mathrm{Cov}(W_{ij}, Z_j)) = \sum_i c_i \left( \sum_j d_{ij} \mathsf{interact}_{\overline{Z}}(\overline{Z}_i) \mathsf{interact}_Z(Z_j, \omega_i) Z_j - W_i \hat{Z}_i \right)$$
$$= \sum_i c_i \mathsf{interact}_{\overline{Z}}(\overline{Z}_i) \left( \sum_j d_{ij} (\mathsf{interact}_Z(Z_j, \omega_i) - 1) Z_j \right)$$

and therefore measures the strength of individual selection in the populations, weighted by population size after population selection.

### 3.2.2 Additive model

The additive model differs from contextual analysis only in the arena of individual selection. For a metapopulation $\omega = \sum_i c_i|\omega_i\rangle : PPZ$ with populations $\omega_i = \sum_j d_{ij}|Z_j\rangle : PZ$ it is given by

$$\omega \mapsto \sum_{ij} c_i d_{ij} \mathsf{interact}_Z(Z_j, \omega_i)|Z_j\rangle \otimes \sum_i c_i \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)|\omega_i\rangle$$
$$\mapsto \sum_{ij} c_i d_{ij} (\mathsf{interact}_Z(Z_j, \omega_i) + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i))|Z_j\rangle$$

hence total individual fitness for an individual with phenotypic value $Z_j$ in population $\omega_i$ is

$$W_{ij} = \mathsf{interact}_Z(Z_j, \omega_i) + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)$$

The model is shown diagrammatically in Figure 3.3. We calculate the components of the multilevel Price equation. With the average individual fitness in population $\omega_i$

$$W_i = \sum_j d_{ij}(\mathsf{interact}_Z(Z_j, \omega_i) + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)) = 1 + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)$$

and average population fitness $\overline{W} = \sum_i c_i \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)$ the population covariance is given by

$$\begin{aligned}\mathrm{Cov}(W_i, \hat{Z}_i) &= \sum_i c_i W_i \hat{Z}_i - (1 + \overline{W}) \sum_i c_i \hat{Z}_i = \sum_i c_i (1 + \mathsf{interact}_{\overline{Z}} - 1 - \overline{W}) \hat{Z}_i \\ &= \sum_i c_i (\mathsf{interact}_{\overline{Z}}(\overline{Z}_i) - \overline{W}) \hat{Z}_i\end{aligned}$$

and is independent of individual selection effected by $\mathsf{interact}_Z : Z \otimes PZ \to \mathbb{R}_{\geq 0}$. The average within-population covariance for the additive model, on the other hand, is independent of population selection since

$$\begin{aligned}\mathrm{E}_I \mathrm{Cov}(W_{ij}, Z_j) &= \sum_i c_i \left( \sum_j d_{ij} W_{ij} Z_j - W_i \hat{Z}_i \right) \\ &= \sum_i c_i \left( \sum_j d_{ij} \mathsf{interact}_Z(Z_j, \omega_i) Z_j + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i) \hat{Z}_i - (1 + \mathsf{interact}_{\overline{Z}}(\overline{Z}_i)) \hat{Z}_i \right) \\ &= \sum_i c_i \left( \sum_j d_{ij} \mathsf{interact}_Z(Z_j, \omega_i) Z_j - \hat{Z}_i \right) = \sum_{ij} c_i d_{ij} (\mathsf{interact}_Z(Z_j, \omega_i) - 1) Z_j.\end{aligned}$$

## 3.3 Parameter estimation

The generative models of multilevel selection described in the previous section specify the suggested processes completely given the functions $\mathsf{interact}_Z : Z \to \mathbb{R}_{\geq 0}$ (or $\mathsf{interact}_Z : Z \otimes PZ \to \mathbb{R}_{\geq 0}$) and $\mathsf{interact}_{\overline{Z}} : \overline{Z} \to \mathbb{R}_{\geq 0}$. This means that for any metapopulation of phenotypes in $Z$ a population of phenotypes can be computed that is the outcome of an episode of selection on the metapopulation using the model. If, conversely, observed data is given the models may be fitted to the data by estimating the $\mathsf{interact}$ functions. The data required for estimation are individual and population phenotype as well as absolute

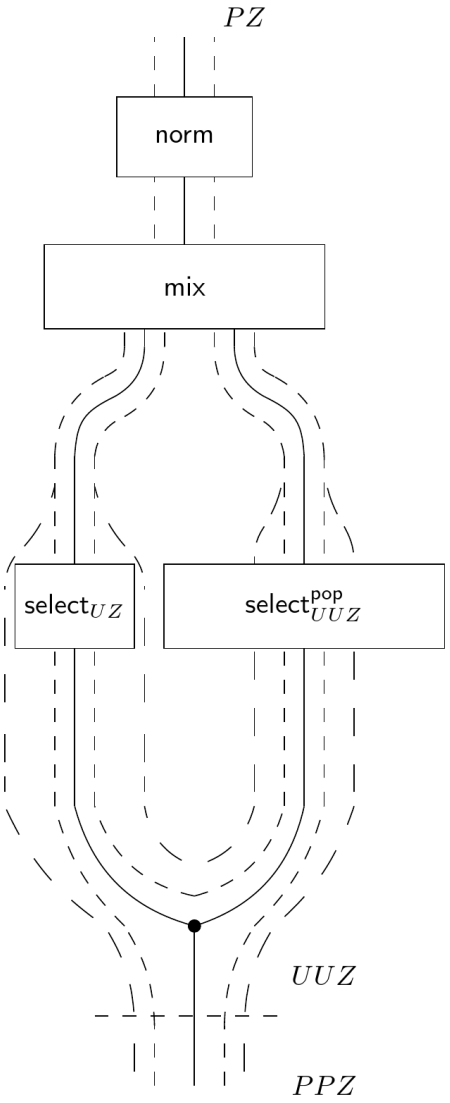



**Figure 3.3:** The additive process corresponding to the Price approach. The metapopulation is copied to undergo to separate causal processes. On the left, the relative sizes of the population remain unchanged while soft individual selection acts in each of the populations. On the right, population composition remains unchanged while population selection acts on the populations. Both metapopulations are then collapsed to population and combined additively with the function mix. Finally, the population is normalised.

fitness of the individuals organised in populations. The models of the multilevel Price equation above moreover require the distribution of phenotypes of each population.

In order to estimate the functions $\mathsf{interact}_Z$ and $\mathsf{interact}_{\overline{Z}}$ the functions may be represented as parameterised families of functions that reflect the proposed dependency of offspring number on individual and population phenotype. For example, if in the model of contextual analysis above fitness is proposed to depend linearly on the numerical phenotypes $Z$ and $\overline{Z}$ corresponding parameterised families are given by $\mathsf{interact}_Z(Z;\alpha,\beta) = \alpha + \beta Z, \mathsf{interact}_{\overline{Z}}(\overline{Z};\gamma) = \gamma\overline{Z}$ with parameters $\alpha, \beta, \gamma : \mathbb{R}$.

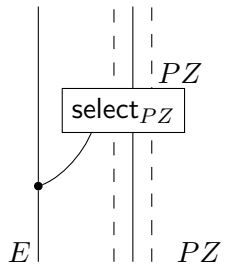
These specifications yield a model of the observed data phenotypes and fitness, and Bayesian inference can be used to estimate the parameters. Bayesian inference requires the definition of prior distributions for the parameters. These encode knowledge about the parameters as well as assumptions that are made prior to analysing the recorded data. For instance, regression parameters as in the example above are often assumed to be normally distributed around zero. After a model for the observed data such as (3.2) or (3.3) is specified and prior distributions for the parameters are defined, Bayesian updating is used to obtain posterior estimates for the parameters based on the observed data. Krapu and Borsuk (2019) discuss software frameworks that implement several algorithms for computing and analysing posterior distributions obtained from models such as (3.2) or (1.12). Point estimates of the parameters can be computed as *maximum a posteriori* (MAP) points for instance in Python using PyMC3 (Salvatier, Wiecki, and Fonnesbeck, 2016).

## 3.4 Outlook

### 3.4.1 Environmental effects and niche construction

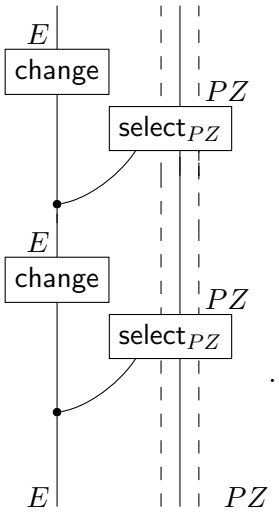
Throughout the thesis we assumed the absence of environmental effects on selection (or a constant selective environment). The environment can be included in the models if we assume that the environmental properties relevant for selection can be encoded in an environment type $E$ (similar to the phenotype $Z$) so that it can be treated like other variables in the models. An episode of selection that depends on the environment then

may be represented as




(3.4)

where the selection function is extended to a function $\mathsf{select}_{PZ} : PZ \otimes E \to PZ$ that depends on the environment. For an environment that undergoes changes according to a function change, a series of selection episodes is then given by




(3.5)

Niche construction is a process during which a population of organisms modifies the environment so that the environmental modifications influence selection pressures on the population or other recipient organisms (Laland, Matthews, and Feldman, 2016). Since properties of the organisms determine how the environment is modified the environmental change is then dependent on the population. Figure 3.4 shows how a population of organisms that modify their selective environment may be represented in terms of string diagrams. The diagram shows the feedback that is characteristic for niche construction: the population modifies the environment that in turn determines the selection pressures on the population. The population after selection again modifies the environment thus creating a feedback loop that allows the population to adapt both to the niche construction activity and the constructed niche.

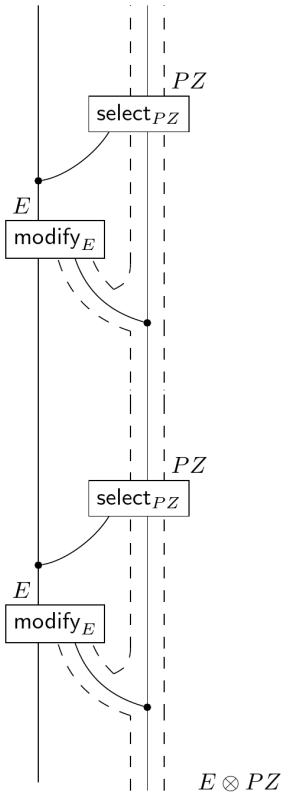
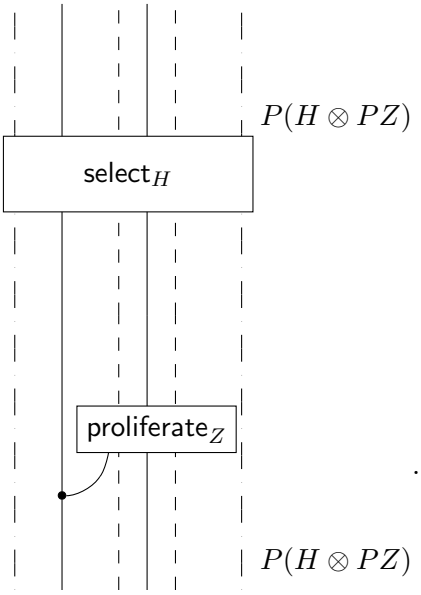



**Figure 3.4:** The diagram shows a population of organisms $PZ$ and their environment $E$. The environment undergoes changes that are caused by the organisms. The function $\text{modify}_E : E \otimes PZ \to E$ takes the environment and the population as input and outputs the modified environment. The selection function that implements selection in the population depends on the modified environment.

### 3.4.2 Holobionts

The term holobiont describes "an individual host and its microbial community, including viruses and cellular microorganisms" (Theis et al. (2016), page 1). Often the microbiome contributes to the host's fitness by providing metabolic processes that cannot be carried out by the host (Roughgarden et al., 2018). Conversely, phenotypic properties of the host determine the composition of its microbial community. Central to the question of coadaptation of host and microbiome is the mode with which microbes are transmitted. When transmission is vertical the microbiome is passed on directly from parent to offspring host. A simple model of a population of holobionts with vertical transmission is given by




. (3.6)

The diagram shows a population of hosts $H$ and their microbial communities $PZ$. The microbial communities proliferate dependent on the host phenotype in $\mathsf{proliferate} : H \otimes PZ \to PZ$. Fitness of the holobiont in $\mathsf{select}_H$ is then determined by host and microbiome phenotype and the offspring generation is produced, keeping the host-microbiome associations.

Transmission is horizontal when the offspring host acquires its microbiome not from its parent but from a pool of microbes that is shared between hosts. A model of holobiont selection with horizontal transmission similar to the model in Roughgarden et al. (2018) is given in Figure 3.5.

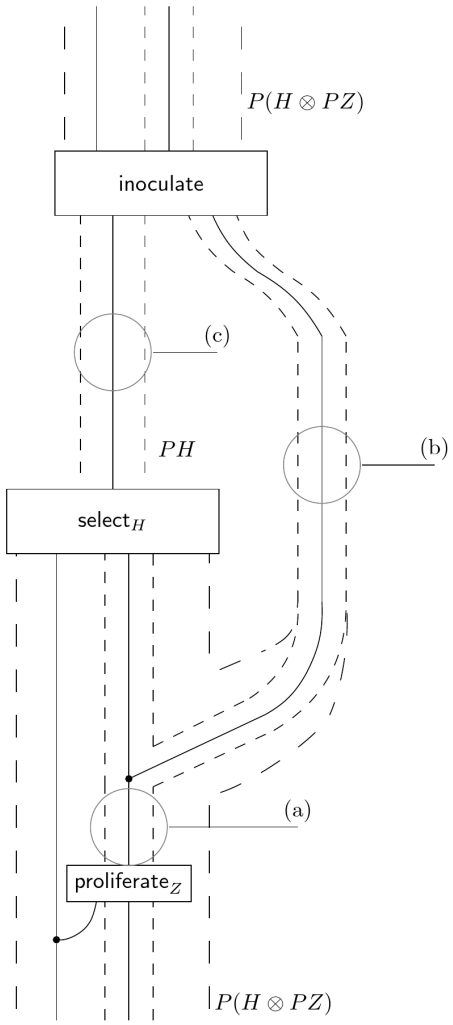



**Figure 3.5:** Selection cycle with horizontal transmission of a population of holobionts (hosts with phenotype $H$ with their microbial communities $PZ$). During the process proliferate : $H \otimes PZ \to PZ$ the microbes proliferate and undergo selection, and this process may depend on properties of the host phenotype. (a) points to the microbial community after proliferation. After this step, samples of the microbe populations of all hosts in the host population are released into a common pool at (b). Selection in the host population depends on the microbe populations as their composition affects their host's fitness. The hosts in the host population at (c) do not have associated microbe populations because transmission is strictly horizontal in this model. The hosts are inoculated with microbes from the common pool in the process inoculate.

# 4 Measuring multilevel selection

In the previous chapters we constructed generative models of multilevel selection that implement the causal structures associated with the multilevel Price equation and contextual analysis. In this chapter we discuss the use of these two approaches to multilevel selection for measuring the strength of selection at the two levels given empirical data collected in a metapopulation that is subject to selection. More precisely, the aim of this chapter is to show that the essential difference between the Price approach and contextual analysis lies in the causal structure each method posits as underlying the observed measurements of individual fitness. Briefly, while contextual analysis assumes that the individual component is determined by direct fitness effects of the trait only, the Price approach sees the individual component as a result of within-population competition and duly assumes it to be affected by the trait values of other members of the population. Put differently, contextual analysis assumes that the individual effect of the trait is absolute in the sense that it is independent of the social environment. The Price approach, on the other hand, assumes that the trait affects the competitiveness of its bearer so that its fitness effect is relative in the sense that it depends on the social environment. This difference leads to different remainders to be explained by population effects and thereby to different measurements of the strength of population selection. Recognising that the difference between the two approaches arises from a difference in the underlying model of reality enables us to see how to determine which of the two approaches is correct in a given case, i.e., the one whose underlying model reflects the causal structure of the system that is being studied. In particular, the applicability of the two approaches depends on the biological scenario at hand and cannot be made on theoretical grounds.

## 4.1 Model

### 4.1.1 Fitness and selection

The evolutionary model in which we frame our arguments is as simple as possible whilst being able to support the features we set out to discuss. Individuals are defined by their

allele at a biallelic locus, with the two alleles representing the presence and absence of a trait, which also defines their phenotype (denoted by $Z$) and replicate asexually without mutation. A population of individuals is partitioned into non-overlapping populations of equal size such that an individual interacts equally with all members of its population (the assumptions on population size and disjointness are made for convenience only). The absolute fitness of an individual determines per capita growth rate and is a function of its own trait as well as the population trait, but not a function of other properties of the metapopulation (absence of, e.g., global frequency dependence; moreover we do not assume a mechanism of global population regulation). The population phenotype is defined as the average phenotype of the individuals within the population. The following considerations concern on bout of selection so that individuals reproduce within their given social environment. We therefore take the populations for granted and do not consider the mechanism of their formation. Taking a causalist stance, we assume that the fitness function is deterministic rather than a statistical abstraction from data (Otsuka, 2016a) and stable in its functional form (i.e., the selective environment that determines fitness in interaction with the phenotype is not changing (Wade and Kalisz, 1990)). Moreover, we assume that the fitness function is additive such that

$$w = c_1 Z + c_2 \overline{Z}, \tag{4.1}$$

where $w$ denotes the fitness of an individual with phenotype $Z$ and population phenotype $\overline{Z}$, and $c_1, c_2 \in \mathbb{R}$ denote the coefficients of the functional representation of fitness in this simple hypothetical example (fitness and individual phenotype are centered at the population mean). This notation corresponds to the method of direct fitness or neighbour-modulated fitness in KS (Taylor, Wild, and Gardner, 2007). $c_1 Z$ represents the direct fitness effect of the trait on its bearer, $c_2 \overline{Z}$ the indirect fitness effect on trait bearers' interaction partners. However, in contrast to approaches using inclusive fitness, we emphasise a causal viewpoint in this thesis by regarding equations such as Equation (4.1) as structural equations that mirror causal assumptions about the system rather than as regression equations. The assumption of additivity in Equation (4.1) is a gross simplification that has been criticised because it ignores synergistic effects and therefore only applies to rare cases (Allen, Nowak, and Wilson, 2013; Van Cleve and Akçay, 2014). The aim of this chapter, however, is not biological generality but to demonstrate causal distinctions made by contextual analysis and the Price approach. The additive fitness function (4.1) suffices to show that the two approaches to multilevel selection discussed in this chapter yield causally nonequivalent structural equations in this case.

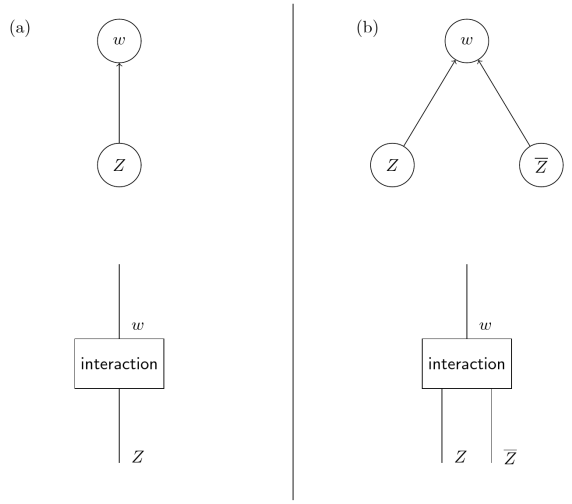



**Figure 4.1:** The crucial difference between nonsocial and social evolution lies in the factors that determine fitness. The diagrams in the upper row are causal graphs that show the dependence of individual fitness on phenotypic factors. The lower row are string diagrams (Jacobs, Kissinger, and Zanasi, 2019) that explicitly represent the process that yields a variable given another variable. In nonsocial evolution (a) individual fitness depends on the individual phenotype only while in social evolution (b) individual fitness depends on the social partners' phenotype in addition to the focal individual's phenotype (Taylor, Wild, and Gardner, 2007). In both cases individual fitness arises from the interaction of the phenotype(s) with the selective environment.

The upper row of diagrams in Figure 4.1 shows the crucial difference between nonsocial and social evolution in terms of causal graphs. While individual fitness $w$ depends only on the individual phenotype $Z$ in nonsocial evolution, social evolution introduces a dependency on social partners' phenotype $\overline{Z}$ (Wolf, Brodie, and Moore, 1999). In Pearl's causal modelling framework, causal graphs as in the upper row of Figure 4.1 represent the graphical counterpart of structural equations (Pearl, 2012). These structural equations model the causal process that yields the output variables, here $w$, given the input variables, here $Z$ and $\overline{Z}$. In contrast to multiple regression these structural equation reflect assumptions about the causal structure of the system. To make the causal process explicit we make use of string diagrams (Jacobs, Kissinger, and Zanasi, 2019) as described in Section String diagrams to represent the causal graphs (i.e., Bayesian networks) in the lower row of diagrams in Figure 4.1. These diagrams are read from bottom to top and depict the variables as strings and the structural equations that transform the variables (i.e., the modelled processes) as boxes. The process of interaction of the phenotype with the selective environment is represented as box interaction that determines individual fitness based on individual and population phenotype. The diagrams in Figure 4.1 are graphical and nonparametric versions of the structural equation 4.1. Both the diagrams and the structural equation describe general features of the causal process proposed to determine the fitness of individuals with phenotype $Z$ within a population with phenotype $\overline{Z}$.

The process of selection in a population is given by the change in trait frequency according to the Price equation without mutation

$$\hat{w}\Delta\hat{Z} = \mathrm{Cov}(w, Z), \tag{4.2}$$

where $\hat{w}$ and $\hat{Z}$ denote average fitness and average individual phenotype across the population. Note that we do not assume that populations themselves replicate or can be assigned population fitness over and above the fitness of the individuals that constitute a population. Our model is therefore of MLS1 type in the sense of Heisler and Damuth (1987), i.e., the focus of the analysis is on individuals, population trait and population fitness are averages of the corresponding quantities of the individuals within the population.

The starting point for the analysis of selection in a population in terms of MLS is the observation that an aspect of selection acts on populations as a whole. This means that individual selection is in part determined by the population trait $\overline{Z}$ because selection favours populations with high (or low) population trait. In particular, this aspect

of selection is the same for all members of a population and is captured by the process by which some populations contribute more offspring to the next generation than others due to differential proliferation and extinction (Uyenoyama and Feldman, 1980; Wade and Griesemer, 1998). Note that it makes no difference to the change in trait frequency whether (an aspect of) selection acts on the population as a whole or on all population members individually but in the same way. However, the aim of MLS is not only the prediction of outcomes but also the attainment of a causal understanding of the selection process (Sober and Wilson, 1994). This understanding comprises selective processes at the individual and the population level: individual fitness not only *depends* on a population trait in addition to the individual trait but fitness also *arises* as consequence of a process that affects the population as a whole in addition to a process that affects each individual specifically. Acknowledging the latter of these dual viewpoints is characteristic of multilevel selection theory as opposed to kin selection theory. The neighbour-modulated approach to kin selection, for instance, formalises the fitness effects of the social environment as factor that alters individual fitness but doesn't view the population as interacting with its own selective environment (Taylor, Wild, and Gardner, 2007).

Explanations for the evolution of cooperative traits, i.e., individual traits that are costly for their bearers in comparison with nonbearers, often rely on the interplay between two processes of this kind. In microcolonies of the bacterium *Pseudomonas aeruginosa*, for instance, the production of siderophores puts individuals at a fitness disadvantage because the process of producing the metabolite binds resources that could otherwise be used for reproduction (Weigert and Kümmerli, 2017). However, the secreted siderophores are shared within the microcolony and thus increase colony fitness due to their iron-scavenging function. Total fitness of individuals results as the combination of the two processes acting on the individual directly and on its population.

In Figure 4.2, the causal graph for social evolution in Figure 4.1 (b) is refined to a causal graph and a corresponding string diagram that explicitly represent the distinct processes of interaction with the selective environments on the individual level and the population level. The box "combine" in the lower diagram (b) corresponds to a function that combines the outcomes of the two processes into total individual fitness. For a complete specification of the model in terms of structural equations this function must be specified in addition to structural equations for the processes on the individual and the population levels. As we will argue in the following section, the causal models corresponding to contextual analysis and the Price approach combine their respective structural equations for

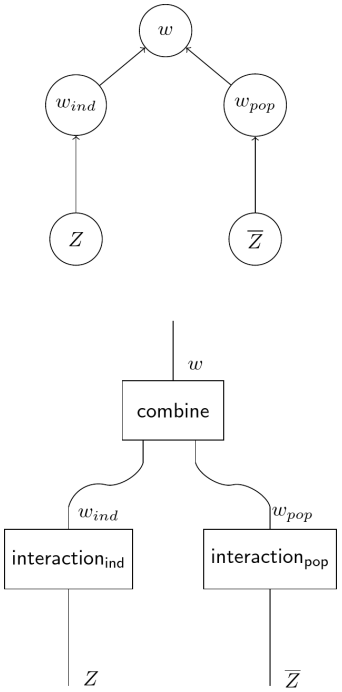



**Figure 4.2:** When multilevel selection (MLS1) operates, individual fitness may not only depend on the individual trait and the population trait as in Figure 4.1 (b). A model of MLS1 usually involves two distinct processes whose combined outcomes yield individual fitness. The individual component of fitness $w_{ind}$ is the outcome of the individual's interaction with its selective environment as in nonsocial evolution. But also the population as a whole and through its population phenotype $\overline{Z}$ interacts with a selective environment in a causally distinct process that yields the population component of fitness $w_{pop}$. Total fitness arises as combination of the two processes.

the individual and population level interaction with the selective environment additively, so that

$$w = w_{\text{ind}} + w_{\text{pop}}. \tag{4.3}$$

The purpose of the decomposition in Equation (4.3) is to explicitly and formally acknowledge the basic tenet of MLS that fitness (here at the individual level) is determined not only by how the individual interacts with its selective environment but also by how the individual's population interacts with the selective environment on a level above that of the individual. We introduce $w_{pop}$ to formally capture fitness effects that result from the interaction of the population as a whole with the selective environment. The quantities $w_{ind}$ and $w_{pop}$ are proxies for the effects of causal processes, the former for processes that affect individual fitness specifically for each individual, the latter for processes that affect the population. The decomposition in Equation (4.3) is additive because of the simple fitness function chosen in Equation (4.1). While an MLS analysis always rests on a decomposition of fitness into contributions from various levels, this decomposition is, generally, not additive. The difference between the two approaches with respect to the arena of individual selection, which is the global population for contextual analysis and the local population for the Price approach, holds more generally, regardless of whether individual and population selection are combined additively.

Since the Price equation is linear in the fitness argument, the decomposition expressed in Equation (4.3) corresponds to a decomposition of the strength of selection itself

$$\hat{w}\Delta\hat{Z} = \hat{w}(\Delta\hat{Z})_{ind} + \hat{w}(\Delta\hat{Z})_{pop} = \text{Cov}(w_{\text{ind}}, Z) + \text{Cov}(w_{\text{pop}}, Z)$$

In order to make quantitative statements about the strengths of population selection vs. individual selection, an MLS analysis must determine the components in this decomposition. However, while individual trait and fitness as well as aggregates thereof can be measured directly, individual effect and population effect, or their covariance with the individual trait, are generally not amenable to direct measurement. The multilevel Price equation and contextual analysis are two methods of obtaining $w_{\text{ind}}$ and $w_{\text{pop}}$ by statistical means given individual traits and fitnesses (Okasha, 2006).

### 4.1.2 Contextual analysis and the Price approach

Equation (4.1) partitions individual fitness into the effects of the individual trait and the population trait. It describes how two phenotypic traits combine to yield another trait of the individual, namely absolute fitness. Contextual analysis (Heisler and Damuth, 1987;

Okasha, 2006; note that here and the following we refer by 'contextual analysis' to the standard structural equation with the untransformed variables $Z$ and $\overline{Z}$, as is customary in discussions on the issues reported here) takes effects of the population trait in Equation (4.1) as indicating population selection. Strictly speaking, $c_2 \neq 0$ in Equation (4.1) implies the potential of the trait to undergo population selection conditional on the existence of population-trait variation between populations (Wolf, Brodie, and Moore (1999); see McLoone (2015) for a discussion of this difference). We regard population effects on fitness as more fundamental than a concept of population selection itself as the former do not depend on properties of a metapopulation but reflect causal processes that increase or decrease reproductive success of an individual situated in a population context vis-\'{a}-vis a specific selection regime that in turn determines individual fitness. Population effects can lead to population selection if, in a specific metapopulation, they generate fitness differences between individuals. This requires $\mathrm{Var}(\overline{Z}) \neq 0$, for if $\mathrm{Var}(\overline{Z}) = 0$ all individuals have the same population trait and are therefore subject to the same population effects. Given a population of individuals and a fitness function that yields individual fitness as superposition of fitness effects of the variables that causally determine fitness, the Price equation yields the effect of selection on this population, that is the change in mean phenotype over one generation (Figure 4.3).

The Price approach to multilevel selection (Price, 1972, Okasha, 2006, Gardner, 2015) rests on the partition of selection itself given by the multilevel expansion of the Price equation (4.2)

$$\hat{w}\Delta\hat{Z} = \mathrm{Cov}(W, \overline{Z}) + \mathrm{E}[\mathrm{Cov}_{\mathrm{wg}}(w, Z)] \tag{4.4}$$

and posits that a population is undergoing group selection if the first term in Equation (4.4) is non-zero. In light of our remarks regarding group effects and group selection above, the Price approach and contextual analysis therefore decompose different quantities and are not directly comparable. However, this difference is superficial as the partition of fitness effects given by contextual analysis corresponds to a partition of selection and the partition of selection given by the multilevel Price equation corresponds to a partition of fitness effects. Contextual analysis, i.e., the functional representation of fitness in Equation (4.1), determines selection according to Equation (4.2) for a population that is partitioned into groups: given a population of individuals $i \in 1, \ldots, n$ with fitnesses

$$w_i = c_1 Z_i + c_2 \overline{Z}_i,$$

where $\overline{Z}_i$ is the trait of the group the $i^{\mathrm{th}}$ individual is part of, the change in mean trait

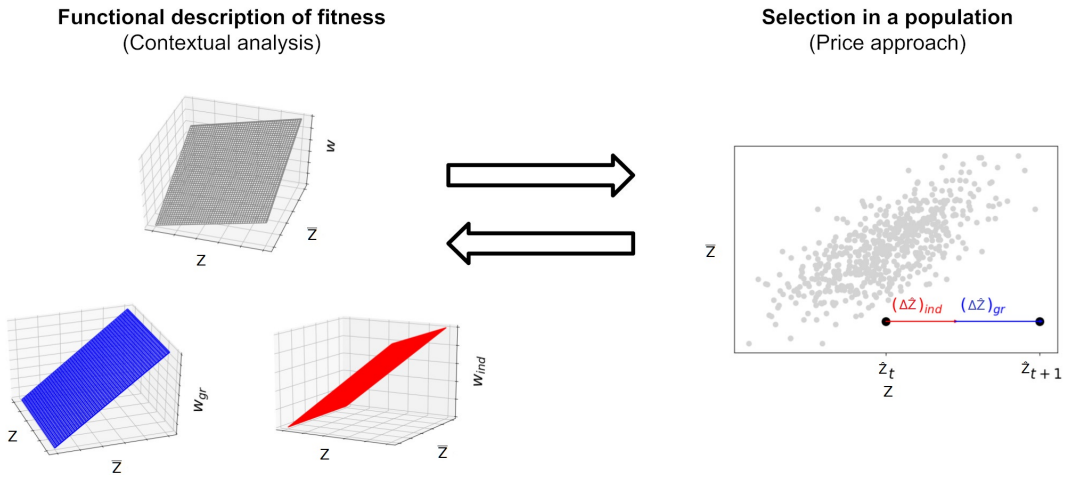



**Figure 4.3:** In the left panel, contextual analysis decomposes the fitness function into an individual and a population component that depend on the individual trait and the population trait, resp. In the right panel, the multilevel Price equation decomposes selection, that is a change in mean phenotype in the population, into a change due to individual selection and a change due to population selection. The standard Price equation without transmission bias maps a functional description of fitness to a process of selection in a population of individuals whose fitness is defined by this functional description. Since the Price equation is given by a covariance and therefore linear in the fitness function, a decomposition of fitness into the summands individual and population effects yields a corresponding and unique decomposition of selection in a population into the summands individual and population selection, each contributing a change in mean phenotype. In Section 4.1.2, we show that, conversely, the coefficients of a decomposition of selection into individual and population selection using the Price approach correspond to coefficients of a structural equation for fitness, given the assumption that fitness is linear as in Equation (4.1).

value in the population follows from Equation (4.2) as (Okasha, 2004)

$$\hat{w}\Delta\hat{Z} = c_1\mathrm{Var}(Z) + c_2\mathrm{Var}(\overline{Z}). \tag{4.5}$$

Thus the decomposition of fitness into individual and group effects given by contextual analysis corresponds to a decomposition of selection whose components, according to contextual analysis, represent the component of individual selection $c_1\mathrm{Var}(Z)$ and the component of group selection $c_2\mathrm{Var}(\overline{Z})$. Conversely, the components of individual selection and group selection according to the Price approach for a population with non-vanishing variance within and between groups correspond to a decomposition of individual fitness into a component of individual effects and group effects. To see how, note that with $w = c_1 Z + c_2\overline{Z}$ (Equation (4.1)),

$$\mathrm{Cov}(W, \overline{Z}) = (c_1 + c_2)\mathrm{Var}(\overline{Z}) \tag{4.6}$$

(Okasha (2006); page 89). Using Equation (4.5) and Equation (4.6), the decomposition according to Equation (4.4) is

$$\begin{aligned}\hat{w}\Delta\hat{Z} &= \mathrm{Cov}(W, \overline{Z}) + \mathrm{E}[\mathrm{Cov}_{\mathrm{wg}}(w, Z)] = (c_1 + c_2)\mathrm{Var}(\overline{Z}) + \mathrm{E}[\mathrm{Cov}_{\mathrm{wg}}(w, Z)] \\ &= c_1\mathrm{Var}(Z) + c_2\mathrm{Var}(\overline{Z})\end{aligned}$$

and therefore

$$\mathrm{E}[\mathrm{Cov}_{\mathrm{wg}}(w, Z)] = c_1(\mathrm{Var}(Z) - \mathrm{Var}(\overline{Z})).$$

Hence the decomposition of fitness

$$w = c_1'(Z - \overline{Z}) + c_2'\overline{Z} \tag{4.7}$$

corresponds to the decomposition of selection

$$\mathrm{Cov}(w, Z) = \mathrm{Cov}(c_1'(Z - \overline{Z}) + c_2'\overline{Z}, Z) = c_1'(\mathrm{Var}(Z) - \mathrm{Var}(\overline{Z})) + c_2'\mathrm{Var}(\overline{Z}),$$

where $c_1' = c_1$ and $c_2' = c_1 + c_2$.

Through this correspondence both contextual analysis and the Price approach yield decompositions of fitness effects as well as of selection (see Table 4.1). Note that the possibility of conducting contextual analysis with respect to the variables $Z - \overline{Z}$ and $\overline{Z}$ rather than $Z$ and $\overline{Z}$ – the former choice of variables being equivalent to the Price approach, the latter to contextual analysis – is discussed in Heisler and Damuth (1987) along with examples of circumstances under which this might be causally adequate.

**Table 4.1:** The decompositions of contextual analysis and the Price approach as individual and group fitness effects $w = w_{\text{ind}} + w_{\text{gr}}$ and as components of selection. The parameters of contextual analysis and the Price approach are linked by the equations $c_1' = c_1$ and $c_2' = c_1 + c_2$.

| | **Fitness** | **Selection** |
|---|---|---|
| | Individual component + group component | Individual selection + group selection |
| | $w = w_{ind} + w_{gr}$ | $\Delta\hat{Z} = (\Delta\hat{Z})_{ind} + (\Delta\hat{Z})_{gr}$ |
| **Contextual analysis** | $c_1 Z + c_2 \overline{Z}$ | $c_1 \text{Var}(Z) + c_2 \text{Var}(\overline{Z})$ |
| **Price approach** | $c_1'(Z - \overline{Z}) + c_2' \overline{Z}$ | $c_1'(\text{Var}(Z) - \text{Var}(\overline{Z})) + c_2' \text{Var}(\overline{Z})$ |

It should be noted that while the interpretation of contextual analysis as structural equation based on causal assumptions is natural, the Price equation and its multilevel expansion are usually viewed as mere statistical identities. However, a causal interpretation of the coefficients of the multilevel Price equation as indicators of selection on the individual and group level as in Okasha (2006) requires the assumption of a causal process that gives rise to the measured coefficients. As demonstrated in this section above for the additive fitness function given by Equation (4.1) the structural equation for a process corresponding to the multilevel Price equation is given by Equation (4.7).

### 4.1.3 Causal intuitions underlying an MLS analysis

A core idea of social evolution is that an individual trait of social organisms has fitness effects not only on its bearer but also on the social environment of the bearer. Common to paradigmatic examples of group selection is an individual trait with effects that change individual fitness homogeneously across the group such that these effects are best viewed as group effects (Sober, 1980). For the water striders described in Eldakar, Wilson, et al. (2010) the exodus of females from patches with high levels of aggressiveness is a group effect of the trait 'aggressiveness' in males. This group effect is negative because group productivity is assumed to decrease with the number of females on a patch as females provide reproductive resources. On the other hand, aggressiveness has a positive individual effect because aggressive males secure more mating opportunities. Whether contextual analysis or the Price approach is appropriate depends on details of this latter

mechanism. If, for instance, less aggressive males generally have lower reproduction rates due to female behaviour and independent of other males on the local patch the contextual analysis model may be more appropriate. If, on the other hand, male reproduction is subject to competition within patches where, for example, successful reproduction depends on the ability of males to guard their mates the model suggested by the Price approach may correspond more closely to the actual process. In either case, the causal interpretation of the trait refers to proximate fitness effects of the trait and involves the individual as well as the group it is in but not other groups or the population as a whole. Therefore the causal interpretation takes place on the fitness side rather than on the selection side in Figure 4.3.

Since we assume that fitness is an effect of the individual/population trait an individual exhibits, we can read the equations in the left column of Table 4.1 as structural equations that determine fitness given the traits. By the assumption on the additivity of interactions these equations are linear. The interpretation of structural equations is aided by the use of causal graphs, more precisely, directed acyclic graphs with causal rather than correlational interpretation (Pearl, 2009). Figure 4.4 shows the causal graphs corresponding to the structural equations in Table 4.1. Since the components $w_{\text{ind}}$ and $w_{\text{gr}}$ reflect quantities that refer to processes occurring in the biological system studied, the causal graphs constitute models of the underlying reality. For example, the population effect of the aggressiveness trait in water striders is given by the propensity of females to remain on the focal patch and this propensity is a function of mean male aggressiveness in the patch (this function is linear by assumption), i.e., the population trait $\overline{Z}$. The non-equivalence of the causal graphs (a) and (b) in Figure 4.4 reflects a difference in how the individual/population components of individual fitness depend on individual/population trait. It should be noted that the factors $Z$ and $\overline{Z}$ are not strictly independent as suggested by omitted arrows between $Z$ and $\overline{Z}$ in Figure 4.4. Since the population phenotype is generated collectively by all individuals within a group, $Z$ does affect $\overline{Z}$. The arrows are omitted in Figure 4.4 because our arguments focus on that part of the causal structure that determines fitness. Details of how the interaction of individual phenotypes gives rise to the population phenotype are not relevant for the present discussion.

The model of fitness underlying contextual analysis (panel (a) in Figure 4.4) is based on the assumption that the individual component and the population component of fitness are determined only by the individual trait and the population trait, resp. This means that fitness differences within populations, i.e., differences in the individual com-

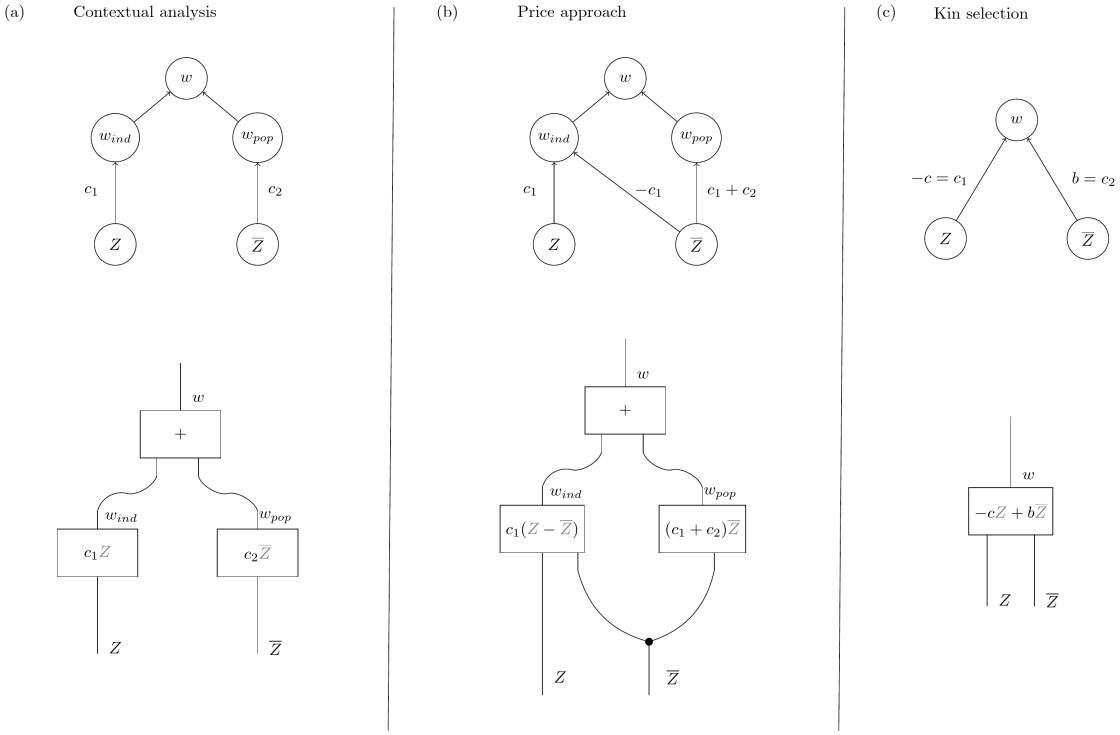



**Figure 4.4: Upper row:** Causal graphs showing the interdependence of the variables $Z$ (individual trait), $\overline{Z}$ (population trait, i.e., mean of individual trait in subpopulation), $w$ (individual fitness), $w_{\mathrm{ind}}$ (individual component of individual fitness), and $w_{\mathrm{pop}}$ (population component of individual fitness). Population fitness $W$ is given as average over the individual fitnesses in a population. (a) Contextual analysis assumes an absolute individual effect of the trait. (b) In the Price approach, the trait is assumed to have a relative effect in the sense that the trait affects fitness depending on the trait expression of other members of the population. (c) In contrast, kin selection theory acknowledges the possibility of indirect effects in addition to direct effects but makes no further assumptions on the causal structure. In KS, it is customary to denote the direct effect of the trait on its bearer by $-c$ and the indirect effect by $b$. The parameter of relatedness $r$ represents the correlation between $Z$ and $\overline{Z}$ and is not pictured in the graph because we focus on selection rather than on properties of population composition. Also the effect of individual phenotype on population phenotype has been omitted, see text. **Lower row:** String diagrams making the processes that yield the output variables given the input variables explicit. Inside the boxes are the structural equations that mirror the respective processes numerically. In diagrams (a) and (b), the fitness effect of the population process interaction$_{\mathsf{pop}}$ is proportional to the population phentype. In diagram (b) for the Price approach, the population phenotype $\overline{Z}$ is "copied" at the black dot because it is involved in both processes. In interaction$_{\mathsf{ind}}$ instantiated with the structural equation $c_1(Z - \overline{Z})$, the population phenotype renders the effect of the individual phenotype relative to the population so that individuals, via their individual phenotype, compete within populations.

ponent, are due to the individual trait and independent of the population trait. In that sense contextual analysis assumes the individual effects of the trait to be absolute, i.e., independent of population context. In contrast, the Price approach assumes that the population trait also affects the individual component of fitness in a specific way (see the path coefficients in Figure 4.4). This effect of the population trait on the individual component is equivalent to the assumption that fitness differences within populations are due to competition between population members in which the individual trait determines competitiveness of an individual. Indeed, the functional representation of fitness according to the Price approach from Table 4.1

$$w = c_1'(Z - \overline{Z}) + c_2'\overline{Z}$$

shows that the individual component sums to zero over each population and that individuals with higher-than-average trait have a positive individual component (negative if $c_1' < 0$). In other words, the trait affects individual fitness not by generally increasing or decreasing its bearer's fitness but by increasing or decreasing its bearer's competitive ability within the population. We discuss examples of these differences in the next section.

## 4.2 Results

### 4.2.1 Cases of disagreement

When comparing the Price approach and contextual analysis it should be kept in mind that both aim to quantify group selection and therefore start with the intuitive identification of an effect of the trait on a population-level property that affects fitness of all individuals within a population homogeneously. In the water strider example, an MLS analysis is based on the assertion, intuitively acquired by inspection of the empirical system, that the population mean of the trait 'aggressiveness' affects the number of females in a population and therefore the productivity of the population as a whole. This assertion is independent of the subsequent choice of statistical approach to quantifying the strength of population selection. Contextual analysis and the Price approach therefore agree on the nature of the population effect on fitness $w_{\mathrm{gr}}$ and on the mechanism bringing forth this effect, though not on its magnitude. The difference between the two approaches lies in the question of which factors affect the individual component of fitness, i.e., which factors are responsible for within-population differences in fitness.

The problem cases for contextual analysis and the Price approach discussed by Okasha (2006) and others (Heisler and Damuth, 1987; Sober, 2011; Goodnight, 2015) reveal issues with the two approaches because the intuition about the level on which fitness effects occur is inconsistent with the verdict of one of the approaches with respect to the strengths of individual and population selection. This intuition is best understood in terms of fitness effects and not in terms of selection because it is based on a mechanism that mediates an effect of the trait on the population component of absolute fitness and is therefore independent of composition and organisation of the population as a whole. Changing a patch of water striders to exhibit a lower level of the population trait 'aggressiveness' increases population fitness because less females will flee the patch. This causal explanation of the biological scenario is the core of an MLS analysis and it is independent of other patches and selection dynamics in the overall population. We conclude that the intuition with regard to the levels on which selection acts is about the mechanisms and not about frequency changes in the population. Accordingly, the following discussion is couched in terms of the left-hand side of Figure 4.3.

In the following examples, we determine the coefficients $c_1, c_2$ of the kin selection model (Figure 4.4 (c)) and discuss their interpretation in terms of the refined models provided by contextual analysis and the Price approach (models (a) and (b) in Figure 4.4).

1. Non-social trait A trait is non-social if the fitness of an individual does not depend on the trait values of its interaction partners (population mates) (Okasha, 2006) so that $c_2 = 0$ and $w = c_1 Z$ ($c_1 \neq 0$ unless the trait is altogether neutral) in Figure 4.4. Intuitively, a trait of this type cannot be subject to population selection, because it has no fitness effects on its bearer's interaction partners and therefore cannot affect the population component of fitness. However, the causal graph that represents the assumptions of the Price approach (Figure 4.4 (b)) shows an effect of population trait on population component of fitness with weight $c_1 + c_2 = c_1$ and therefore detects population selection where intuitively there is none. Population effects of this type have been called cross-level by-products (Okasha, 2006) and will be discussed in a later section. Note that the causal graph underlying contextual analysis correctly shows the absence of population effects.

2. Soft selection The tension between the Price approach and contextual analysis is reversed in the case of soft selection (Wade, 1985; Goodnight, Schwartz, and Stevens, 1992; Débarre and Gandon, 2011). Briefly, soft selection occurs in a metapopulation if mean individual fitness is homogeneous across populations, i.e.,

if all populations have the same reproductive output. Soft selection models situations in which individuals of each population share a fixed resource and the trait under soft selection determines how an individual fares in the within-population competition for this resource. The population trait determines competitiveness of the population, i.e., mean competitiveness of its members, in the sense that an individual has lower fitness in a competitive population than in a population with low population trait. Soft selection is intuitively considered to be free of population selection (Wade, 1985; Okasha, 2006; Sober, 2011). The trait has no effect on the population level because changing the trait value of an individual in a population has no homogeneous fitness effect within the population as the change has no consequences for mean population fitness but merely changes the outcome of the within-population competition. It is easy to see that a kin selection model of soft selection takes the form $w = c_1 Z - c_1 \overline{Z}$, i.e., $c_2 = -c_1$, with $c_1 > 0$ (resp. $c_1 < 0$) if a higher trait value implies higher (resp. lower) competitiveness. The interpretation of these parameters according to the Price approach yields that the edge from $\overline{Z}$ to $w_{\text{pop}}$ has weight $c_1 + c_2 = 0$ in the causal graph (b) in Figure 4.4. The Price approach correctly detects the absence of population selection in this example. However, contextual analysis mistakes the effect of the population trait on fitness as an effect on the population component of fitness according to the causal graph (a) in Figure 4.4. Though most researchers that engaged with the problem of inconsistency between contextual analysis and the Price approach seem to agree that no population selection occurs in soft selection, some have argued to the contrary. Goodnight, Schwartz, and Stevens (1992) regard soft selection as an example of population selection since an individual's fitness depends on the trait of the population of which it is a member. We agree that individual fitness depends on the population trait but this effect of the population trait on fitness is an individual effect (the diagonal arrow in Figure 4.4 (b) targets $w_{\text{ind}}$) that represents within-population competition. In soft selection, there is no population effect since the trait does not influence population fitness.

3. Genotypic selection with meiotic drive Okasha (2004) introduces 'frameshifting' as a desirable property of a general theory of multilevel selection. The theory is capable of frameshifting if it formalises features of population selection in such a way that they hold by analogy whenever the hierarchy given by the population/individual relation is instantiated at other levels of organisation. The treatment of genotypic

selection with meiotic drive in MLS terms is relevant in that context because it tests the ability of MLS to frameshift to levels below the level of organisms. Following Wilson (1990), Okasha (2004) discusses diploid population genetics as an example of multilevel selection where alleles correspond to individuals and diploid genotypes to populations. In this analogy, population effects on allelic fitness are due to genotypic fitness, i.e., organismic fitness of the organism with a specific genotype, and individual effects are due to meiotic drive that creates within-population fitness differences between alleles.

Given the intuition that individual selection as well as population selection is at work in genotypic selection with meiotic drive, the expectation with respect to a decomposition of fitness into individual and population effects is clearly that population selection is present when genotypic fitnesses differ whereas individual selection is brought about by unfair meiosis. However, it is easy to see using specific fitness functions that contextual analysis doesn't agree with intuition in this case. When only meiotic drive is acting while there is no difference in fitness between genotypes, for example, the situation is analogous to soft selection that was shown above to be captured by the Price approach rather than contextual analysis. Furthermore, Okasha (2006) gives the example of two alleles $A$ and $B$ such that genotypic fitnesses are given by $w_{AA} = 4, w_{AB} = 3,$ and $w_{BB} = 2$ while meiotic drive causes 2 of the 3 gametes produced by an $AB$ organism to be $A$. Then fitness of an $A$ allele is 2 and that of a $B$ allele is 1, independent of the genetic background. Thus, despite unfair meiosis and dependence of fitness on the population trait, contextual analysis concludes the absence of population selection, $c_1 \neq 0, c_2 = 0$. The Price approach, in contrast, reaches the correct conclusion that individual fitness is given by $w = \underbrace{c_1(Z - \overline{Z})}_{\text{ind. component}} + \underbrace{c_1\overline{Z}}_{\text{population component}}$ and therefore that both components of selection are non-zero.

Okasha's conclusion that the "covariance approach [i.e., the Price approach] appears to frameshift down quite well, the contextual approach very badly" (Okasha (2004), page 498) is thus readily explained by the viewpoint developed so far: unfair meiosis corresponds to the zero-sum game of within-population competition. This is precisely the causal structure assumed by the Price approach.

### 4.2.2 Cross-level by-products

A core assumption of MLS theory is that a trait an individual expresses may affect properties of its population as a whole and therefore population fitness (i.e., mean individual fitness in a population). This effect is captured by the population component $w_{\text{pop}}$ of individual fitness. However, population fitness, in MLS1, is the average individual fitness in a population and therefore comprises not only the population component but also the average individual component $w_{\text{ind}}$. This is problematic because the part of population fitness that entails selection on the population property caused by the trait is $w_{\text{pop}}$ only. The contribution of $w_{\text{ind}}$ to population fitness is called a cross-level by-product (Okasha, 2006) because it represents fitness of the individuals that constitute the population, i.e., the lower level, rather than fitness that is a property of the population as a whole, i.e., the higher level. Intuitively, a population with many individually fit members seems more fit than a population with few individually fit members even when the population component $w_{\text{pop}}$ and therefore the fitness with respect to population selection that is to be quantified is the same for both populations. The non-social trait case discussed above is a good example of this effect. Since there is no population property for population selection to act on in this case, population fitness comprises solely of individual fitness from the level below and therefore consists entirely of cross-level by-products.

To see how contextual analysis and the Price approach handle cross-level by-products assume that individual fitness is given by the expression $w = c_1 Z + c_2 \overline{Z}$. The decomposition of population fitness $W = (c_1 + c_2)\overline{Z}$ into a component due to population effects and a component due to individual effects depends on the causal structure and therefore differs between the two approaches. While contextual analysis partitions population fitness into individual and population component as $W = \underbrace{c_1 \overline{Z}}_{\text{ind. component}} + \underbrace{c_2 \overline{Z}}_{\text{population component}}$ the decomposition according to the Price approach yields only a population component, $W = (c_1 + c_2)\overline{Z}$. For a non-social trait ($c_1 \neq 0, c_2 = 0$) the Price approach mistakenly traces population fitness entirely back to a non-existing population effect, whereas contextual analysis correctly assigns population fitness to the individual effect. The fact that contextual analysis handles cross-level by-products correctly in the non-social trait case has led Okasha to conclude that contextual analysis is "on balance preferable" (Okasha (2006), page 99) to the Price approach. However, it should be noted that in the soft selection case ($c_1 = -c_2$) contextual analysis decomposes population fitness as

$$W = \underbrace{c_1 \overline{Z}}_{W_{\text{ind}}} - \underbrace{c_1 \overline{Z}}_{W_{\text{pop}}}$$

and hence detects cross-level by-products of magnitude $c_1\overline{Z}$ even though cross-level by-products are absent since the individual components of fitness sum to zero in each population.

In their study on multilevel selection in water striders, Eldakar, Wilson, et al. (2010) choose contextual analysis for quantifying population selection because contextual analysis controls for "potential cross-level byproducts" (Eldakar, Wilson, et al. (2010), page 3186). However, as we have seen, contextual analysis does not correctly account for cross-level by-products automatically. Which of the two approaches is correct depends on the kind of individual selection that acts on the system, i.e., the causal structure underlying fitness. In this case, the causal graphs (a) and (b) in Figure 4.4 both seem possible. Recall that aggressiveness in male water striders is hypothesised to have an effect on the individual component of fitness (aggressive males secure more mating opportunities than non-aggressive males) and on the population component of fitness (patches with higher aggression levels have fewer females). Contextual analysis assumes that the individual component is independent of the population trait: in addition to the population component shared by all males in a population each male has an individual component that is determined by its trait and independent of the population trait. Another, perhaps more plausible, assumption underlies the Price approach: the population trait determines the number of females on a patch and this reproductive resource is distributed to the males according to their competitiveness. We will discuss an experimental intervention that would reveal the correct underlying causal structure in the next section.

### 4.2.3 Determining the preferable approach

Several authors have discussed the question which of the two approaches is preferable in general (Okasha, 2006; Sober, 2011; McLoone, 2015). However, even the most extensive discussion of this question (Okasha, 2006) has been inconclusive in the sense that in light of the problematic cases discussed above neither can be endorsed unreservedly. We argued that a general preference cannot be justified as the essential difference between the two approaches lies in non-equivalent assumptions about the causal structure of the biological system which, as the problematic cases demonstrate, may be either of the two. However, our reduction of the difference between the Price approach and contextual analysis to a difference between their respective causal graphs has the benefit that experimental interventions that reveal the correct causal structure and with it the correct approach can easily be derived from the causal graphs (Pearl, 2009). Note that while we argue that the suggested interventions in principle reveal the correct structure we do not claim

that such interventions are feasible for a given biological system. Moreover, while the two approaches discussed here are the main approaches to measuring the strength of population selection, it may well be possible that neither is suitable in a given scenario. We will discuss this and other limitations of this work below.

Imagine that we have a biological system such as a population of water striders in Eldakar et al. in which intuitive inspection suggests that individual fitness depends on an individual component and a population component as in Figure 4.4 (a) and (b). Analysis reveals proposed causal pathways for individual trait and population trait to affect individual fitness via the two components. In particular, such an analysis comprises a hypothesis on the mechanism that mediates the effect of the population trait on the population component of individual fitness. For the water strider example the population trait is mean aggressiveness on a patch, the population component is proportional to the number of females on a patch, and the mechanism that mediates the effect of the former on the latter is female exodus determined by the females' preference for low aggressiveness patches. Choosing contextual analysis or the Price approach for quantification goes hand in hand with the commitment to regard Figure 4.4 (a) or (b), resp., as the causal structure underlying the observed phenomena. The causal structures posited by the two approaches differ in that the Price approach assumes an effect of the population trait on the individual component of fitness. This assumption is reflected in the diagonal arrow in Figure 4.4 (b) that is missing in panel (a). The two arrows emanating from $\overline{Z}$ in (b) represent two distinct cause-effect relations between the population trait and individual fitness. But given the hypothesis on the mechanism that mediates the effect of the population trait on the population component of fitness these two distinct cause-effect relations correspond to two distinct mechanisms through which the population trait affects fitness. Consequently, it is in principle possible to separate the effects by intervening on one of the mechanisms but not the other. This intervention translates to removing the vertical arrows from $\overline{Z}$ to $w_{\text{pop}}$ in Figure 4.4 (a) and (b) so that the system is described by the modified graphs in Figure 4.5. But in the system with suppressed population effects the two causal structures in Figure 4.5 (a) and (b) can be distinguished on the basis of the observable quantities $Z$, $\overline{Z}$, and $w$. In particular, contextual analysis predicts individual fitness to be independent of population membership when the system is being intervened on in this way. The Price approach, however, predicts continued dependence of fitness on population trait due to within-population competition. As these predictions cannot both be true, the intervention allows the identification of one of the two approaches as being in accord with experimental observations.

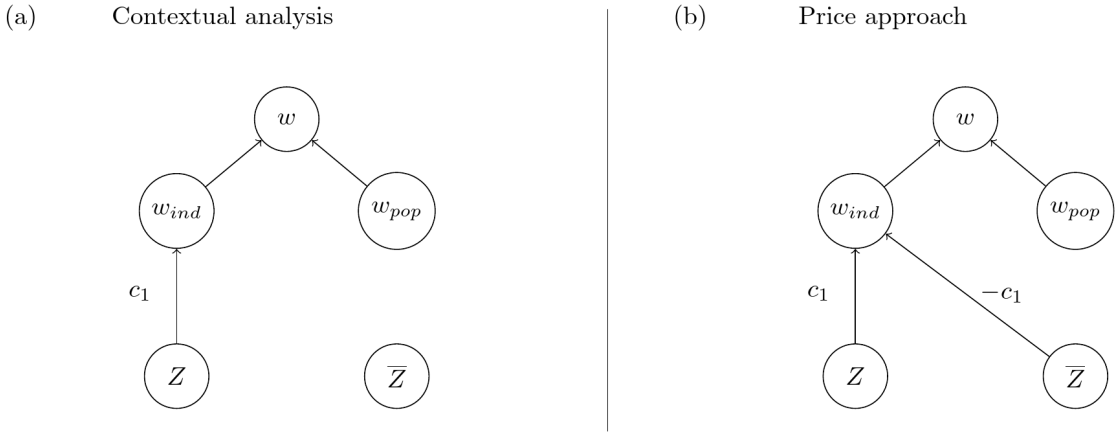



**Figure 4.5:** Modified causal graphs when suppressing the effect of the population trait on the population component. Contextual analysis predicts fitness to be independent of the population trait when the population effect is suppressed while the Price approach does not.

Corresponding to the causal models expressed in the graphs are mechanisms corresponding to each of the arrows in the model. In the water strider example, given the mechanisms corresponding to the arrows in Figure 4.4, it is now easy to see how a decision for one of the two approaches may be reached. Since the effect of population trait on fitness is mediated by female exodus, the effect can be suppressed by preventing females from leaving patches, i.e., by removing female dispersal between patches (Eldakar, Dlugos, et al., 2009). It is crucial that this intervention leaves the diagonal arrow in Figure 4.5 (b) intact. This is because the diagonal arrow represents a different causal pathway, namely the within-patch competition for females which is not affected by preventing females from leaving the patch. An informed decision for contextual analysis can then be reached if fitness is independent of mean aggressiveness on a patch when female dispersal is removed, i.e., if the diagonal arrow in Figure 4.5 (b) was not part of the underlying causal structure in the unperturbed system. The Price approach is more appropriate if fitness still depends on patch composition under this experimental condition.

Both the Price approach and contextual analysis serve the purpose to determine the quantities $w_{ind}$ and $w_{gr}$ in Equation (4.3), or equivalent quantities (see Table 4.1), from the more easily measurable variables individual trait and individual fitness. In order to achieve this, both approaches require assumptions that can be conveniently represented in terms of causal graphs as in Figure 4.4. We have shown above how, in principle, it is possible to determine which of the two approaches is more appropriate. However, we have seen that the causal structures posited are highly contrived. It seems therefore very

well possible that neither of the two approaches is suitable for determining the level-specific strength of selection. This is the case when neither of the causal graphs (a) and (b) in Figure 4.4 represents the causal structure underlying the biological phenomenon in question.

## 4.3 Conclusion

Population selection refines kin selection by splitting individual fitness into two components, i.e., by assuming that fitness is determined by two additional factors that are themselves determined by the variables individual trait and population trait. The causal graphs in Figure 4.4 show that this means that population selection adds a layer to the causal structure of selection assumed by kin selection. This addition constitutes a proper refinement of kin selection and corresponds to avoiding averaging over the causes of individual fitness (the 'averaging fallacy' described by Sober and Wilson (1999)). From this viewpoint, the tension between contextual analysis and the Price approach can be seen as an instance of the purely formal problem of connecting an additional layer of nodes to an existing graph. The connection schemes proposed by contextual analysis and the Price approach, i.e., the coefficients of the paths targeting $w_{\text{ind}}$ and $w_{\text{gr}}$ in Figure 4.4, are two solutions to this problem. Since omitted paths in a causal graph represent hypotheses about the absence of effects the correct approach is the approach whose hypotheses are satisfied in the biological system at hand. Phrasing the problem in terms of causal graphs demonstrates that, even in the additive case, other refinements are in principle possible and could apply to scenarios in which the individual component is given neither by soft selection (Price approach) nor by hard selection (contextual analysis) but by intermediate selection regimes (Débarre and Gandon, 2011). Casting an MLS analysis in terms of refinements of causal graphs gives a formal argument for the non-equivalence of MLS and kin selection. We have argued that the refinement introduced by MLS is non-trivial (see difficulties with Price approach and contextual analysis) and provides a view on the system that is tailored to the levels of organisation in the system. This view is crucial when cause-effect relations that pertain to a specific level are manipulated or undergo change, such as during an evolutionary transition in individuality, and the system-level consequences of such alterations are to be predicted. Strengthening the formal core of MLS not only facilitates the application of MLS in evolutionary science but also aids in assessing benefits, limitations, and formal requirements of this approach to empirical and theoretical biological scenarios.

# 5 Summary and Conclusion

In this thesis I have discussed multilevel selection, its representation in terms of mathematical models, and how these models correspond to instances of multilevel selection in biological systems. Motivated by the inconsistency between contextual analysis and the Price approach to multilevel selection, I argued that causal models are required to avoid inconsistency and constructed causal models of multilevel selection that correspond to contextual analysis and the multilevel Price equation.

Using the ladder of causation, I illustrate in the first chapter how causality and representability extend the notion of correlation and argued that a complete understanding of a system that allows reasoning about counterfactual configurations requires a functional model of the system. Equations of evolutionary change, such as the multilevel Price equation and contextual analysis, are often correlational but their application to empirical systems is based on a causal interpretation.

In the second chapter, I show how models of selection and other biological processes may be implemented functionally thus yielding causal models. Probability and multiset monads are used to represent populations and metapopulations so that multilevel selection can be implemented as function that takes a metapopulation as input and outputs a population after selection. This realises a functional representation of the evolutionary process of multilevel selection and allows the construction of more involved models by composing basic functions sequentially, parallely, and recursively. Moreover, I show how string diagrams can be used to conveniently represent the functions involved. This diagrammatic formalism allows an intuitive representation of the constructed models.

I introduce complete models of multilevel selection that correspond to contextual analysis and the multilevel Price equation in the third chapter. These models are formulated in terms of probability monads and represented graphically as well as algebraically. The models highlight the conceptual differences between contextual analysis and the multilevel Price equation.

Finally in the fourth chapter, I use a functional viewpoint to clarify the relation of contextual analysis and the Price approach when measuring multilevel selection in em-

pirical data. The two approaches yield conflicting results for the strength of selection on the two levels, and I argue that this is because they correspond to non-equivalent causal models. In order to get results that agree with the understanding of the system, the causal structure of the actual biological system must be matched with the causal structure of the model. I outline how the correct causal structure may be determined through intervention on the system.

There are several limitations of the thesis that concern the models considered and the factors included in them. The models in this thesis assume perfect heredity so that offspring are clones of their single parent, ignoring mutation. Also, the models do not include notions of phenotypic plasticity or development and, consequently, do not allow a distinction between genotype and phenotype. Moreover, reproduction is assumed asexual throughout. Another limitation of the thesis concerns the structure of the metapopulations considered. Applications of multilevel selection, for instance in theories on the evolution of cooperation, often require specific properties of the metapopulation such as assortment of genotypic values according to Hamilton's parameter of relatedness $r$. While any structuring of the metapopulation into populations is possible with the models described above, I have not included a mechanism by which populations are formed. Such a mechanism would allow to iterate selection so as to assess the response to selection.

An obvious avenue for further research is extending the models constructed in this thesis to include other factors and mechanisms that are relevant to evolution such as mutation and sexual reproduction. Since selection acts on the phenotype it is natural to include the environment so as to model development or phenotypic plasticity. In applications of multilevel selection populations often feature additional structure such as class structure in insect societies (Gardner, 2015). The formalism of monads used in this thesis lends itself to represent further structure of the metapopulation. In this way, class structure or other population structure could be included in the models. Otsuka (2016b, 2019) advocates a causal formulation of evolutionary equations in order to make assumptions regarding the causal structure of the system explicit. The functional approach used in this thesis could be used for modelling evolutionary phenomena other than selection. The graphical representation in terms of string diagrams allows the sequential, parallel, and recursive composition of simple components to obtain complex models.

# Appendix

## 1 The multilevel Price equation

The multilevel Price equation is a recursive application of the Price equation (1.2) with populations as units of selection (i.e., as evolutionary individuals). We assume ideal transmission on the individual level – that is no mutation, recombination, or phenotypic plasticity – and the organisation of individuals into disjoint populations. The populations constitute a metapopulation and are assumed to be of homogeneous magnitude for simplicity. With the selection component (1.6) of the Price equation for populations, we have ($i : I$ with $|I| = n$ indexes the individuals, $l : L$ with $|L| = p$ indexes the populations, and $k : K$ with $|K| = m$ indexes the individuals within the populations of magnitude $m : \mathbb{N}$; there are $n$ individuals in $p$ populations of $m$ individuals each)

$$\Delta \mathrm{E}_I(Z) = \mathrm{Cov}_I(W, Z) = \mathrm{Cov}_I(\overline{W}, \overline{Z}) + \mathrm{E}_L \mathrm{Cov}_{K(L)}(W, Z). \tag{1}$$

The first equality follows because we assume transmission effects on the individual level to be void. For the second equality,

1. assume without loss of generality that $\mathrm{E}_I Z = 0$ (which implies $\mathrm{E}_L \overline{Z} = 0$),

2. remember that $\mathrm{E}_I W = 1$ because $W$ here denotes *relative* fitness (which implies $\mathrm{E}_L \overline{W} = 1$), and

3. note that

$$\mathrm{Cov}_L(\overline{W}, \overline{Z}) = \mathrm{Cov}_I(W, \overline{Z}).$$

To see 3., denote the $n$ data points recorded for each of the variables $W, \overline{W}, Z,$ and $\overline{Z}$

by $w_i, z_i,$ and $\overline{z}_i$ for $i : I$. With $\mathrm{Cov}_I(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mathrm{E}_I X)(y_i - \mathrm{E}_I Y)$, we have

$$\begin{aligned}\mathrm{Cov}_L(\overline{W},\overline{Z}) &= \frac{1}{p}\sum_{l=1}^{p}(\overline{w}_l - \underbrace{\mathrm{E}_L\overline{W}}_{=1})(\overline{z}_l - \underbrace{\mathrm{E}_L\overline{Z}}_{=0}) = \frac{1}{p}\sum_{l=1}^{p}(\overline{w}_l - 1)\overline{z}_l \overset{(1.)}{=} \frac{1}{p}\sum_{l=1}^{p}\overline{w}_l\overline{z}_l = \\ &= \frac{1}{p}\sum_{l=1}^{p}\overline{z}_l\frac{1}{m}\sum_{k=1}^{m}\overline{w}_{l_k} = \frac{1}{n}\sum_{l=1}^{p}\sum_{k=1}^{m}\overline{w}_{l_k}\overline{z}_l \\ &= \mathrm{Cov}_I(W,\overline{Z}).\end{aligned}$$

This gives

$$\begin{aligned}\mathrm{E}_L\mathrm{Cov}_{K(L)}(W,Z) &= \frac{1}{p}\sum_{l=1}^{p}\mathrm{Cov}_l(W,Z) = \frac{1}{p}\sum_{l=1}^{p}\frac{1}{m}\sum_{k=1}^{m}(w_{l_k} - \overline{w}_l)(z_{l_k} - \overline{z}_l) \\ &= \frac{1}{n}\sum_{l=1}^{p}\sum_{k=1}^{m}[w_{l_k}z_{l_k} - \overline{w}_l z_{l_k} - w_{l_k}\overline{z}_l + \overline{w}_l\overline{z}_l] \\ &= \mathrm{Cov}_I(W,Z) + \mathrm{Cov}_L(\overline{W},\overline{Z}) - \underbrace{\frac{2}{p}\sum_{l=1}^{p}\overline{w}_l\overline{z}_l}_{=2\mathrm{Cov}_L(\overline{W},\overline{Z})} \\ &= \mathrm{Cov}_I(W,Z) - \mathrm{Cov}_L(\overline{W},\overline{Z})\end{aligned}$$

to prove the second equality and thus motivate the *multilevel Price equation* (1) as expansion of the Price equation for individuals in metapopulations.

# Bibliography


Abbot, P. et al. (2011). "Inclusive fitness theory and eusociality". In: *Nature* 471.7339, E1–E4. DOI: 10.1038/nature09831 (↑ xii, 55).

Aktipis, A. et al. (2015). "Cancer across the tree of life: cooperation and cheating in multicellularity". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 370.1673. DOI: 10.1098/rstb.2014.0219 (↑ 48).

Allen, B., M. A. Nowak, and E. O. Wilson (2013). "Limitations of inclusive fitness". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.50, pp. 20135–20139. DOI: 10.1073/pnas.1317588110 (↑ xii, 84).

Axelrod, R. and W. D. Hamilton (1981). "The evolution of cooperation". In: *Science* 211.4489, pp. 1390–1396. DOI: 10.1126/science.7466396 (↑ 53).

Baez, J. C. and M. Stay (2009). "Physics, Topology, Logic and Computation: A Rosetta Stone". In: *Physics, topology, logic and compIn New Structures for Physics, ed. Bob Coecke, Lecture Notes in Physics vol. 813, Springer, Berlin, 2011, pp. 95-174*. DOI: 10.1007/978-3-642-12821-9_2 (↑ 22, 24).

Baker, R. J. and R. D. Bradley (2006). "Speciation in Mammals and the Genetic Species Concept". In: *Journal of Mammalogy* 87.4, pp. 643–662. DOI: 10.1644/06-mamm-f-038r2.1 (↑ 15).

Bapteste, E. and F. Papale (2020). "Modeling the evolution of interconnected processes: It is the song and the singers". In: *BioEssays*, p. 2000077. DOI: 10.1002/bies.202000077 (↑ 1, 2, 28, 67).

Bareinboim, E. et al. (2020). *On Pearl's Hierarchy and the Foundations of Causal Inference*. Tech. rep. ACM Special Volume in Honor of Judea Pearl (provisional title). URL: https://causalai.net/r60.pdf (↑ 2, 3).

Borrelli, J. J. et al. (2015). "Selection on stability across ecological scales". In: *Trends in Ecology & Evolution* 30.7, pp. 417–425. DOI: 10.1016/j.tree.2015.05.001 (↑ 28, 51).

Clarke, E. (2016). "A levels-of-selection approach to evolutionary individuality". In: *Biology & Philosophy* 31.6, pp. 893–911. DOI: 10.1007/s10539-016-9540-4 (↑ 1, 27).

Damuth, J. and I. L. Heisler (1988). "Alternative formulations of multilevel selection". In: *Biology & Philosophy* 3.4, pp. 407–430. DOI: 10.1007/BF00647962 (↑ 1, 9, 14, 27, 51).

Dawkins, R. (2006). *The Selfish Gene*. Oxford University Press. ISBN: 978-0-19-929114-4 (↑ 50).

Débarre, F. and S. Gandon (2011). "Evolution in Heterogeneous Environments: Between Soft and Hard Selection". In: *The American Naturalist* 177.3, E84–E97. DOI: 10.1086/658178 (↑ 97, 104).

Débarre, F., C. Hauert, and M. Doebeli (2014). "Social evolution in structured populations". In: *Nature Communications* 5, p. 3409. DOI: 10.1038/ncomms4409 (↑ 1).

Doolittle, W. F. and E. Bapteste (2007). "Pattern pluralism and the Tree of Life hypothesis". In: *Proceedings of the National Academy of Sciences* 104.7, pp. 2043–2049. DOI: 10.1073/pnas.0610699104 (↑ 64, 67).

Doolittle, W. F. and A. Booth (2017). "It's the song, not the singer: an exploration of holobiosis and evolutionary theory". In: *Biology & Philosophy* 32.1, pp. 5–24. DOI: 10.1007/s10539-016-9542-2 (↑ 18).

Doolittle, W. F. and S. A. Inkpen (2018). "Processes and patterns of interaction as units of selection: An introduction to ITSNTS thinking". In: *Proceedings of the National Academy of Sciences*, p. 201722232. DOI: 10.1073/pnas.1722232115 (↑ 18).

Ehresmann, A. C. and J.-P. Vanbremeersch (2007). *Memory Evolutive Systems; Hierarchy, Emergence, Cognition*. Studies in Multidisciplinarity. Elsevier Science Inc. ISBN: 0444522441 (↑ xi).

Eldakar, O. T., M. J. Dlugos, et al. (2009). "Population Structure Mediates Sexual Conflict in Water Striders". In: *Science* 326.5954, p. 816. DOI: 10.1126/science.1180183 (↑ 20, 103).

Eldakar, O. T., D. S. Wilson, et al. (2010). "THE ROLE OF MULTILEVEL SELECTION IN THE EVOLUTION OF SEXUAL CONFLICT IN THE WATER STRIDER *AQUARIUS REMIGIS*". In: *Evolution* 64.11, pp. 3183–3189. DOI: 10.1111/j.1558-5646.2010.01087.x (↑ 11, 19, 54, 93, 101).

Ereshefsky, M. and M. Pedroso (2013). "Biological individuality: the case of biofilms". In: *Biology & Philosophy* 28.2, pp. 331–349. DOI: 10.1007/s10539-012-9340-4 (↑ 48).

Farine, D. R., P.-O. Montiglio, and O. Spiegel (2015). "From Individuals to Groups and Back: The Evolutionary Implications of Group Phenotypic Composition". In: *Trends in Ecology & Evolution* 30.10, pp. 609–621. DOI: 10.1016/j.tree.2015.07.005 (↑ 60, 61).

Fields, C. and M. Levin (2018). "Are Planaria Individuals? What Regenerative Biology is Telling Us About the Nature of Multicellularity". In: *Evolutionary Biology* 45.3, pp. 237–247. DOI: 10.1007/s11692-018-9448-9 (↑ 28, 48).

Fong, B. (2013). *Causal theories: A categorical perspective on Bayesian networks*. URL: https://arxiv.org/abs/1301.6201 (↑ xi, 3, 18, 23–25).

Fong, B. and D. I. Spivak (2019). *An Invitation to Applied Category Theory*. Cambridge University Press. DOI: 10.1017/9781108668804 (↑ 22, 24).

Frank, S. A. (1997). "THE PRICE EQUATION, FISHER'S FUNDAMENTAL THEOREM, KIN SELECTION, AND CAUSAL ANALYSIS". In: *Evolution* 51.6, pp. 1712–1729. DOI: 10.1111/j.1558-5646.1997.tb05096.x (↑ 10).

– (1998). *Foundations of Social Evolution*. Princeton University Press. ISBN: 9780691059341 (↑ 1, 10, 11, 63).

– (2012). "Natural selection. IV. The Price equation". In: *Journal of Evolutionary Biology*. DOI: 10.1111/j.1420-9101.2012.02498.x (↑ 7, 9).

Fritz, T. (2019). "A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics". In: *Adv. Math. 370, 107239 (2020)*. DOI: 10.1016/j.aim.2020.107239 (↑ 23).

Fritz, T. and P. Perrone (2020). "Monads, Partial Evaluations, and Rewriting". In: *Electronic Notes in Theoretical Computer Science* 352, pp. 129–148. DOI: 10.1016/j.entcs.2020.09.007 (↑ 37–39).

Gardner, A. (2015). "The genetical theory of multilevel selection". In: *Journal of Evolutionary Biology* 28.2, pp. 305–319. DOI: 10.1111/jeb.12566 (↑ xi, 1, 2, 11, 13, 16, 61, 62, 68, 90, 106).

– (2020). "Price's equation made clear". In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 375.1797, p. 20190361. DOI: 10.1098/rstb.2019.0361 (↑ 2, 5, 6, 9, 12, 27).

Gardner, A., S. A. West, and N. H. Barton (2007). "The Relation between Multilocus Population Genetics and Social Evolution Theory". In: *The American Naturalist* 169.2, pp. 207–226. DOI: 10.1086/510602 (↑ 9, 40, 68).

Gardner, A., S. A. West, and G. Wild (2011). "The genetical theory of kin selection". In: *Journal of Evolutionary Biology* 24.5, pp. 1020–1043. DOI: 10.1111/j.1420-9101.2011.02236.x (↑ xii, 55).

Giry, M. (1982). "A categorical approach to probability theory". In: *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, pp. 68–85. DOI: 10.1007/bfb0092872 (↑ xi).

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. Oxford University Press (↑ 52, 61).

Goodnight, C. J. (2015). "Multilevel selection theory and evidence: a critique of Gardner, 2015". In: *Journal of Evolutionary Biology* 28.9, pp. 1734–1746. DOI: 10.1111/jeb.12685 (↑ 1, 97).

Goodnight, C. J., J. M. Schwartz, and L. Stevens (1992). "Contextual Analysis of Models of Group Selection, Soft Selection, Hard Selection, and the Evolution of Altruism". In: *The American Naturalist* 140.5, pp. 743–761. DOI: 10.1086/285438 (↑ 11, 42, 43, 97, 98).

Grafen, A. (2000). "Developments of the Price equation and natural selection under uncertainty". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 267.1449, pp. 1223–1227. DOI: 10.1098/rspb.2000.1131 (↑ 6, 7).

– (2006). "Optimization of inclusive fitness". In: *Journal of Theoretical Biology* 238.3, pp. 541–563. DOI: 10.1016/j.jtbi.2005.06.009 (↑ 55).

Grochau-Wright, Z. I. et al. (2017). "Genetic basis for soma is present in undifferentiated volvocine green algae". In: *Journal of Evolutionary Biology* 30.6, pp. 1205–1218. DOI: 10.1111/jeb.13100 (↑ 57).

Hamilton, W. D. (1964). "The genetical evolution of social behaviour. I". In: *Journal of Theoretical Biology* 7.1, pp. 1–16. DOI: 10.1016/0022-5193(64)90038-4 (↑ 10, 53).

Heisler, I. L. and J. Damuth (1987). "A Method for Analyzing Selection in Hierarchically Structured Populations". In: *The American Naturalist* 130.4, pp. 582–602. DOI: 10.1086/284732 (↑ xi, 4, 11, 14, 86, 89, 92, 97).

Hull, D. L. (1980). "Individuality and Selection". In: *Annual Review of Ecology and Systematics* 11.1, pp. 311–332. DOI: 10.1146/annurev.es.11.110180.001523 (↑ 27, 49, 50, 65).

Huneman, P. (2021). "Inclusive Fitness teleology and Darwinian explanatory pluralism a theoretical sketch and an application to current controversies". In: *Biosemiotics: The Natural Foundations of Symbolism*. Ed. by E. Pagni and R. S. Theisen. Springer (↑ 1).

Jacobs, B. (2019). "Structured Probabilistic Reasoning". Forthcoming book. URL: cs.ru.nl/B.Jacobs/PAPERS/ProbabilisticReasoning.pdf (↑ xi, 8, 26, 30, 41).

Jacobs, B., A. Kissinger, and F. Zanasi (2019). "Causal Inference by String Diagram Surgery". In: *Foundations of Software Science and Computation Structures*. Springer International Publishing, pp. 313–329. DOI: 10.1007/978-3-030-17127-8_18 (↑ 20, 23, 25, 29, 85, 86).

Jacobs, B. and F. Zanasi (2018). "The Logical Essentials of Bayesian Reasoning". In: *CoRR*. URL: http://arxiv.org/abs/1804.01193v2 (↑ xi, 16, 23–25, 36).

Krapu, C. and M. Borsuk (2019). "Probabilistic programming: A review for environmental modellers". In: *Environmental Modelling & Software* 114, pp. 40–48. DOI: 10.1016/j.envsoft.2019.01.014 (↑ 78).

Laland, K., B. Matthews, and M. W. Feldman (2016). "An introduction to niche construction theory". In: *Evolutionary Ecology* 30.2, pp. 191–202. DOI: 10.1007/s10682-016-9821-z (↑ 79).

Leinster, T. (2021). *Entropy and Diversity: The Axiomatic Approach*. Cambridge University Press. ISBN: 9781108965576 (↑ xi).

Lenton, T. M. et al. (2021). "Survival of the Systems". In: *Trends in Ecology & Evolution* 36.4, pp. 333–344. DOI: 10.1016/j.tree.2020.12.003 (↑ 16, 28).

Levin, S. R. and A. Grafen (2019). "Inclusive fitness is an indispensable approximation for understanding organismal design". In: *Evolution* 73.6, pp. 1066–1076. DOI: 10.1111/evo.13739 (↑ 7, 55).

Lewontin, R. C. (1970). "The Units of Selection". In: *Annual Review of Ecology and Systematics* 1.1, pp. 1–18. DOI: 10.1146/annurev.es.01.110170.000245 (↑ 13, 61, 62).

Lloyd, E. A. and M. J. Wade (2019). "Criteria for Holobionts from Community Genetics". In: *Biological Theory*. DOI: 10.1007/s13752-019-00322-w (↑ 1, 50).

Manrubia, S. (2021). *The Simple Emergence of Complex Molecular Function*. URL: http://arxiv.org/abs/2105.11784v1 (↑ 55).

McLoone, B. (2015). "Some Criticism of the Contextual Approach, and a Few Proposals". In: *Biological theory* 10.2, pp. 116–124. DOI: 10.1007/s13752-015-0209-z (↑ 11, 90, 101).

Michod, R. E. (2006). "The group covariance effect and fitness trade-offs during evolutionary transitions in individuality". In: *Proceedings of the National Academy of Sciences* 103.24, pp. 9113–9117. DOI: 10.1073/pnas.0601080103 (↑ 56).

Nowak, M. A., C. E. Tarnita, and E. O. Wilson (2010). "The evolution of eusociality". In: *Nature* 466.7310, pp. 1057–1062. DOI: 10.1038/nature09205 (↑ xii).

Okasha, S. (2004). "Multi-Level Selection, Covariance and Contextual Analysis". In: *The British journal for the philosophy of science* 55.3, pp. 481–504. DOI: 10.1093/bjps/55.3.481 (↑ 92, 98, 99).

Okasha, S. (2006). *Evolution and the Levels of Selection*. Clarendon Press. ISBN: 9780199267972. DOI: 10.1093/acprof:oso/9780199267972.001.0001 (↑ xi, 1, 2, 6, 11, 14, 16, 47, 49, 51, 52, 63, 89, 90, 92, 93, 97–101).

– (2015). "The Relation between Kin and Multilevel Selection: An Approach Using Causal Graphs". In: *The British Journal for the Philosophy of Science* 67.2, pp. 435–470. DOI: 10.1093/bjps/axu047 (↑ xii, 2, 4).

Otsuka, J. (2016a). "A critical review of the statisticalist debate". In: *Biology & Philosophy*. DOI: 10.1007/s10539-016-9528-0 (↑ 1, 84).

– (2016b). "Causal Foundations of Evolutionary Genetics". In: *The British journal for the philosophy of science* 67.1, pp. 247–269. DOI: 10.1093/bjps/axu039 (↑ 106).

– (2019). *The Role of Mathematics in Evolutionary Theory*. Elements in the Philosophy of Biology. Cambridge University Press. ISBN: 9781108672115. DOI: 10.1017/9781108672115 (↑ xii, 1, 106).

Papale, F. (2020). "Evolution by means of natural selection without reproduction: revamping Lewontin's account". In: *Synthese*. DOI: 10.1007/s11229-020-02729-6 (↑ 2, 18, 50, 52, 61).

Patterson, E. (2020). *The algebra and machine representation of statistical models*. URL: https://arxiv.org/abs/2006.08945 (↑ xi, 30).

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Elsevier Science and Technology. ISBN: 9780934613736 (↑ 18).

– (2009). *Causality: Models, Reasoning, and Inference*. 2nd ed. New York: Cambridge University Press. ISBN: 9780521895606 (↑ 3, 16, 20, 34, 94, 101).

– (2012). "The Causal Foundations of Structural Equation Modeling". In: *Handbook of Structural Equation Modeling*. Ed. by R. H. Hoyle. New York: Guilford Press, pp. 68–91. DOI: 10.21236/ada557445. URL: http://dx.doi.org/10.21236/ada557445 (↑ 86).

– (2018). *Causal and Counterfactual Inference*. Tech. rep. R-485. Forthcoming section in *The Handbook of Rationality*, MIT Press, available at http://ftp.cs.ucla.edu/pub/stat_ser/r485.pdf (↑ 8, 21).

– (2019). "On the Interpretation of do(x)". In: *Journal of Causal Inference* 7.1. DOI: 10.1515/jci-2019-2002 (↑ 17, 20).

Pearl, J., M. Glymour, and N. P. Jewell (2016). *Causal Inference in Statistics: A Primer*. Wiley. ISBN: 978-1-119-18684-7 (↑ xi, 3–5, 12, 17).

Pearl, J. and D. Mackenzie (2018). *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books. ISBN: 978-0-46-509760-9 (↑ xii, 2, 16).

Price, G. R. (1970). "Selection and Covariance". In: *Nature* 227.5257, pp. 520–521. DOI: 10.1038/227520a0 (↑ 5, 6).

– (1972). "Extension of covariance selection mathematics". In: *Annals of human genetics* 35.4, pp. 485–490. DOI: 10.1111/j.1469-1809.1957.tb01874.x (↑ 90).

– (1995). "The nature of selection". In: *Journal of Theoretical Biology* 175.3. (Written circa 1971, published posthumously), pp. 389–396. DOI: 10.1006/jtbi.1995.0149 (↑ 5–7, 27).

Rice, S. H. (2004). *Evolutionary theory: mathematical and conceptual foundations*. Sinauer Associates. ISBN: 978-0878937028 (↑ 1).

Riehl, E. (2017). *Category Theory in Context*. Dover Publications. ISBN: 9780486809038 (↑ xi, 21, 22, 36).

Roughgarden, J. et al. (2018). "Holobionts as Units of Selection and a Model of Their Population Dynamics and Evolution". In: *Biological theory* 13.1, pp. 44–65. DOI: 10.1007/s13752-017-0287-1 (↑ 81).

Ryan, P. A., S. T. Powers, and R. A. Watson (2016). "Social niche construction and evolutionary transitions in individuality". In: *Biology & Philosophy* 31.1, pp. 59–79. DOI: 10.1007/s10539-015-9505-z (↑ 48).

Salvatier, J., T. V. Wiecki, and C. Fonnesbeck (2016). "Probabilistic programming in Python using PyMC3". In: *PeerJ Computer Science* 2, e55. DOI: 10.7717/peerj-cs.55 (↑ 78).

Schmid, D., J. H. Selby, and R. W. Spekkens (2020). *Unscrambling the omelette of causation and inference: The framework of causal-inferential theories* (↑ xi, 26, 30, 34, 36).

Smith, J. and R. F. Inglis (2021). "Evaluating Kin and Group Selection As Tools for Quantitative Analysis of Microbial Data". In: *Proceedings of the Royal Society B: Biological Sciences* 288.1951, p. 20201657. DOI: 10.1098/rspb.2020.1657 (↑ 56).

Sober, E. (1980). "Holism, Individualism, and the Units of Selection". In: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1980.2, pp. 93–121. DOI: 10.1086/psaprocbienmeetp.1980.2.192588 (↑ 93).

– (2011). "Realism, Conventionalism, and Causal Decomposition in Units of Selection: Reflections on Samir Okasha's Evolution and the Levels of Selection". In: *Philosophy and phenomenological research* 82.1, pp. 221–231. DOI: 10.1111/j.1933-1592.2010.00471.x (↑ 11, 12, 97, 98, 101).

Sober, E. and D. S. Wilson (1994). "A Critical Review of Philosophical Work on the Units of Selection Problem". In: *Philosophy of science* 61.4, pp. 534–555. DOI: 10.1086/289821 (↑ 87).

– (1999). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press. ISBN: 9780674930476 (↑ 104).

Stadler, B. M. et al. (2001). "The Topology of the Possible: Formal Spaces Underlying Patterns of Evolutionary Change". In: *Journal of Theoretical Biology* 213.2, pp. 241–274. DOI: 10.1006/jtbi.2001.2423 (↑ 32).

Taylor, P. D., G. Wild, and A. Gardner (2007). "Direct fitness or inclusive fitness: how shall we model kin selection?" In: *Journal of evolutionary biology* 20.1, pp. 301–309. DOI: 10.1111/j.1420-9101.2006.01196.x (↑ 84, 85, 87).

Theis, K. R. et al. (2016). "Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes". In: *mSystems* 1.2. DOI: 10.1128/mSystems.00028-16 (↑ 81).

Thies, C. and R. A. Watson (2021). "Identifying Causes of Social Evolution: Contextual Analysis, the Price Approach, and Multilevel Selection". In: *Frontiers in Ecology and Evolution* 9. DOI: 10.3389/fevo.2021.780508. URL: http://dx.doi.org/10.3389/fevo.2021.780508 (↑ vii).

Tuyéras, R. (2017). *Category Theory for Genetics*. URL: https://arxiv.org/abs/1708.05255 (↑ xi).

Uyenoyama, M. K. and M. W. Feldman (1980). "Theories of kin and group selection: A population genetics perspective". In: *Theoretical Population Biology* 17.3, pp. 380–414. DOI: 10.1016/0040-5809(80)90033-7 (↑ 10, 28, 53, 87).

– (1981). "On relatedness and adaptive topography in kin selection". In: *Theoretical Population Biology* 19.1, pp. 87–123. DOI: 10.1016/0040-5809(81)90036-8 (↑ 61).

Van Cleve, J. and E. Akçay (2014). "PATHWAYS TO SOCIAL EVOLUTION: RECIPROCITY, RELATEDNESS, AND SYNERGY". In: *Evolution*, pp. 2245–2258. DOI: 10.1111/evo.12438. URL: http://dx.doi.org/10.1111/evo.12438 (↑ 84).

Van Veelen, M. et al. (2012). "Group selection and inclusive fitness are *not* equivalent; the Price equation vs. models and statistics". In: *Journal of Theoretical Biology* 299, pp. 64–80. DOI: 10.1016/j.jtbi.2011.07.025 (↑ 9).

Virgo, N. (2020). Post in Zulip forum. URL: https://mattecapu.github.io/ct-zulip-archive/stream/238032-Categorical-Probability-and-Statistics-2020-workshop/topic/Categorical.20Radon-Nikodym.html#205870566 (↑ 37).

Wade, M. J. and J. R. Griesemer (1998). "Populational heritability: empirical studies of evolution in metapopulations". In: *The American Naturalist* 151.2, pp. 135–147. DOI: 10.1086/286107 (↑ 60, 87).

Wade, M. J. (1976). "Group selection among laboratory populations of *Tribolium*". In: *Proceedings of the National Academy of Sciences of the United States of America* 73.12, pp. 4604–4607. DOI: 10.1073/pnas.73.12.4604 (↑ 10, 57–60).

– (1978). "A Critical Review of the Models of Group Selection". In: *The Quarterly review of biology* 53.2, pp. 101–114. DOI: 10.1086/410450 (↑ 10).

– (1985). "Soft Selection, Hard Selection, Kin Selection, and Group Selection". In: *The American Naturalist* 125.1, pp. 61–73. DOI: 10.1086/284328 (↑ 42, 43, 97, 98).

– (2016). *Adaptation in Metapopulations*. University of Chicago Press. DOI: 10.7208/chicago/9780226129877.001.0001 (↑ 28, 63).

Wade, M. J. and C. J. Goodnight (1998). "PERSPECTIVE: THE THEORIES OF FISHER AND WRIGHT IN THE CONTEXT OF METAPOPULATIONS: WHEN NATURE DOES MANY SMALL EXPERIMENTS". In: *Evolution* 52.6, pp. 1537–1553. DOI: 10.1111/j.1558-5646.1998.tb02235.x (↑ 52, 53, 55).

Wade, M. J. and S. Kalisz (1990). "THE CAUSES OF NATURAL SELECTION". In: *Evolution* 44.8, p. 1947. DOI: 10.2307/2409605 (↑ 7, 10, 28, 84).

Watson, R. A. and C. Thies (2019). "Are Developmental Plasticity, Niche Construction, and Extended Inheritance Necessary for Evolution by Natural Selection? The Role of Active Phenotypes in the Minimal Criteria for Darwinian Individuality". In: *Evolutionary Causation*. The MIT Press. DOI: 10.7551/mitpress/11693.003.0011 (↑ 5, 49, 51, 57, 64).

Watters, J. V. and A. Sih (2005). "The mix matters: behavioural types and group dynamics in water striders". In: *Behaviour* 142.9-10, pp. 1417–1431. DOI: 10.1163/156853905774539454 (↑ 11).

Weigert, M. and R. Kümmerli (2017). "The physical boundaries of public goods cooperation between surface-attached bacterial cells". In: *Proceedings of the Royal Society B: Biological Sciences* 284.1858, p. 20170631. DOI: 10.1098/rspb.2017.0631 (↑ 54, 87).

West, S. A., A. S. Griffin, and A. Gardner (2007). "Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection". In: *Journal of Evolutionary Biology* 20.2, pp. 415–432. DOI: 10.1111/j.1420-9101.2006.01258.x (↑ 53, 63).

Wilson, D. S. (1975). "A theory of group selection". In: *Proceedings of the National Academy of Sciences of the United States of America* 72.1, pp. 143–146. DOI: 10.1073/pnas.72.1.143 (↑ 10, 53).

Wilson, D. S. (1980). *The natural selection of populations and communities*. Menlo Park, California: Benjamin/Cummings Pub. Co. ISBN: 0805395601 (↑ 53).

Wilson, D. S. (1990). "Weak Altruism, Strong Group Selection". In: *Oikos* 59.1, p. 135. DOI: 10.2307/3545133 (↑ 99).

Wilson, D. S. and E. Sober (1989). "Reviving the Superorganism". In: *Journal of Theoretical Biology* 136.3, pp. 337–356. DOI: 10.1016/S0022-5193(89)80169-9 (↑ 10).

Wolf, J. B., E. D. Brodie III, and A. J. Moore (1999). "Interacting Phenotypes and the Evolutionary Process. II. Selection Resulting from Social Interactions". In: *The American naturalist* 153.3, pp. 254–266. DOI: 10.1086/303168 (↑ 86, 90).

# Index