ON THE DIFFUSION APPROXIMATION OF WRIGHT–FISHER MODELS WITH SEVERAL ALLELES AND LOCI AND ITS GEOMETRY

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
angenommene

DISSERTATION

zur Erlangung des akademischen Grades
DOCTOR RERUM NATURALIUM
(Dr. rer. nat.)
im Fachgebiet
MATHEMATIK

vorgelegt von
Dipl.-Math. Julian Hofrichter

Die Annahme der Dissertation wurde empfohlen von:
1. Prof. Dr. Jürgen Jost, MPI für Mathematik in den Naturwissenschaften, Leipzig
2. Prof. Dr. Angela Stevens, Westfälische Wilhelms-Universität, Münster/W.

Die Verleihung des akademischen Grades erfolgt mit Bestehen der Verteidigung am 22. Juli 2014 mit dem Gesamtprädikat
magna cum laude.

This work was typeset using LaTeX and KOMA-Script.

Acknowledgements

I would like to thank all people who have contributed to the success of this thesis. First and foremost, my cordial thanks go to my supervisor Prof. Dr. Jürgen Jost for giving me the opportunity to work on this interesting and versatile research project as well as for his approachability and his constant support.

My gratitude also goes to the Max Planck Institute for Mathematics in the Sciences and the International Max Planck Research School for financial support and for providing an excellent work environment.

Furthermore, I wish to thank my fellow students at the Max Planck Institute for many constructive and exhilarating discussions: In particular, this includes Dat Tran Dat, Benjamin Eltzner, Jan Schlemmer, David Dubray and André Schlichting. I would like to especially thank Benjamin Eltzner for carefully reading the entire manuscript and Mark Sampson for first hand expertise in English grammar.

Lastly, I would like to acknowledge the moral support I received from many people, notably my friends and family.

The author's work was funded from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement № 267087.

Contents

1 Introduction

Since the pioneering work of R. A. Fisher and S. G. Wright, a mathematical conception of population genetics has proven very useful. Population genetics describe biological processes at the level of *genes*, which appear as the biological unit of coding, function and inheritance (cf. [12]), and may therefore yield an account of the genetic configuration of a given population and its evolution over time. Also, the same type of processes may be observed in other settings where evolution is examined, for example in evolutionary models of linguistics (cf. [6]).

The key point in these models consists of the fact that evolution is assumed to be (partially) random and hence relates to the mathematical concept of stochastic processes. First applications of this entered the debate in the 1920s and 1930s in the context of a study of genetic drift, i. e. the change in the genetic configuration which is only due to random effects. Such models featuring a fixed population size and non-overlapping generations are usually referred to as Wright–Fisher models, introduced by Fisher (implicitly) and Wright in 1922 resp. 1931 (cf. [11], [33]). The simplest application is a model with one locus at which two alleles may occur. One may then ask what the probabilistic configuration of the model is at a given time.

Technically, such processes are Markov chains as the probabilities for genetic changes only depend on the current genetic configuration. In order to match the observations of evolutionary biology, in addition to genetic drift, Wright–Fisher models also need to incorporate other evolutionary effects such as natural selection and mutation as well as recombination in multi-loci models. In this generalised setting, a diffusion limit of the considered processes proves to be handier, i. e. instead of a discrete state space and a discrete time parameter, the configuration is then given by a real-valued random variable, while time is continuous. This so-called diffusion approximation is a standard tool and may be found in many textbooks (e. g. in [10]); its application to population genetics is notably due to M. Kimura in the

1950s (cf. [17]).

Continuous models may then in turn be described by means of analytical concepts, i. e. differential equations for the corresponding probability density function referred to as the Kolmogorov equations (cf. [20]). In the absence of natural selection and other effects, these equations describe a heat-equation type evolution, thus the evolution of the genetic configuration is similar to the diffusion of heat in a metal. The inclusion of other evolutionary mechanisms such as natural selection, mutation or recombination then corresponds to a deterministic component which biases the evolution into a certain direction (this, in terms of physics, is referred to as 'drift') in contrast to the undirected random component ('diffusion'). Due to the analogy to physics, this type of differential equations also appears in the physics literature by the notion of Fokker-Planck equations. For the purposes of this thesis, we will stick with the genetics nomenclature and use the notion of the Kolmogorov equations with the Kolmogorov forward equation describing the future evolution of a given configuration, whereas the Kolmogorov backward equation gives probabilities conditional upon ending up in a specified state at some future time.

The present thesis is basically located within the context of the diffusion approximation of Wright–Fisher models and the Kolmogorov equations describing their evolution. The main subject of this thesis is twofold: The first half is concerned with an analysis of recombinational Wright–Fisher models as well as with stating the corresponding Kolmogorov equations for their diffusion limit. Triggered by a result of M. Kimura together with T. Ohta in 1969 (cf. [25]), which appears in the literature without much further background, a full account of the diffusion approximation of recombinational Wright–Fisher models is developed, also with a view towards information geometry. The other half of the thesis addresses analytical solution schemes for such Kolmogorov equations arising from Wright–Fisher models. However, for simplification, the treatment in that part is limited to models with only one locus and consequently without recombination as well as other evolutionary effects, thus may be subsumed under 'neutral evolution'.

For this setting, analytic solutions already exist in the literature, for example for the two-allelic case as well as for the tri-allelic case by M. Kimura (cf. [16], [18]) and for the general case with n alleles by R. A. Littler and E. D. Fackerell (cf. [22]) as well as recently in [30]. All these, however, appear to be incomplete to a certain extent in the sense that a systematic account of the boundary, which – other than in standard heat equation theory – requires a specific treatment, and hence a full description of the dynamics is missing. In the analysis presented here, we are able to develop such an extended solution scheme including the boundary, predicated on M. Kimura's findings in [16]. This may be applied both to the forward and the backward setting, leading to interesting results.

For the Kolmogorov backward equation, our analysis also reconfirms some results by R. A. Littler for the stationary case (corresponding to $t = -\infty$; cf. [21], [23]); the presented approach provides a systematic derivation for these results, which is missing in the mentioned sources. Additionally, another important aspect, that is the uniqueness of these stationary solutions, obviously has not been covered by Littler nor elsewhere – yet, this gap may be filled via a specifically designed blow-up scheme for the domain as will be presented in this thesis.

The outline of the present thesis is thus as follows: In chapter 2, a systematic approach to the diffusion approximation of recombinational two- or more loci Wright–Fisher models is presented. As a point of departure we choose a specific Kolmogorov backward equation for the diffusion approximation of a recombinational two-loci Wright–Fisher model, to which – with the help of some information geometrical methods, i. e. by calculating the sectional curvatures of the corresponding statistical manifold (which is the domain equipped with the corresponding Fisher metric) – we are able to identify the underlying Wright–Fisher model. Accompanying this, for all methods and tools involved a suitable introduction is presented. Furthermore, the considerations span a separate analysis for the two most common underlying models (RUZ and RUG) as well as a comparison of the two models. Finally, transferring corresponding results for a simpler model described by Antonelli and Strobeck in [5], solutions of the Kolmogorov equations are contrasted with Brownian motion in the same domain, leading to interesting insights.

In chapter 3, the perspective of the diffusion approximation of recombinational Wright–Fisher models is further developed as the model underlying the Ohta–Kimura formula is subsequently extended by an integration of the concepts of natural fitness and mutation. Simultaneously, the corresponding extensions of the Ohta–Kimura formula are stated. Crucial for this is the development of a suitable fitness scheme, which is accomplished by a multiplicative aggregation of fitness values for pairs of gametes/zygotes. Furthermore, the model is generalised to have an arbitrary number of alleles and – in the following step – an arbitrary number of loci respectively. The latter involves an increased number of recombination modes, for which the concept of recombination masks as introduced in [27] is also implemented into the model. Another generalisation in terms of coarse-graining is performed via an application of schemata; this also affects the previously introduced concepts, specifically mask recombination, which are adapted accordingly. Eventually, a geometric analysis of linkage equilibrium states of the multi-loci Wright–Fisher models is carried out, relating to the concept of hierarchical probability distributions in information geometry (cf. [3]), which concludes our considerations of recombinational Wright–Fisher models and their extensions.

In chapter 4, we usher in our discussion of analytical solution schemes for the Kolmogorov equations corresponding to the diffusion approximation of Wright–Fisher models, which represents the second part of the thesis. To this end, we start with the simplest setting of a 1-dimensional Wright–Fisher model, for which we recall the solution strategy for the corresponding Kolmogorov forward equation given by M. Kimura. From this, we are able to construct a unique extended solution which also accounts for the dynamics of the model on lower-dimensional entities of the state space, i. e. configurations of the model where one of the alleles no longer exists in the population, utilising the concept of the (boundary) flux of a solution; a discussion of the moments of the distribution confirms our findings. A similar treatment is then carried out for the corresponding Kolmogorov backward equation, yielding analogous results of existence and uniqueness for an extended solution. For the latter in particular, a corresponding account of the configuration on the boundary turns out to be crucial, which is also reflected in the probabilistic interpretation of the backward solution. Additionally, the long-term behaviour of solutions is analysed, and a comparison between such solutions of the forward and the backward equation is made.

In chapter 5, we basically aim to transfer the results obtained in the previous chapter to the subsequent increasingly complicated setting of a Wright–Fisher model with 1 locus and an arbitrary number of alleles: With solution schemes for the interior of the state space (i. e. not encompassing the boundary) already existing in the literature, we develop an extension scheme for a successive determination of the solution on lower-dimensional entities of the domain. This scheme, again, makes use of the concept of (boundary) flux of solutions, and we may therefore show that this extended solution fulfils additional properties regarding the completeness of the diffusion approximation with respect to the boundary. These properties may be formulated in terms of the moments of the distribution, and we illustrate their connection to the underlying Wright–Fisher model. Altogether, stipulating such a moments condition, we show existence and uniqueness of an extended solution on the entire domain. Furthermore, the corresponding Kolmogorov backward equation is examined, for which we similarly present a (backward) extension scheme which allows extending a solution in a domain (perceived as a boundary instance of a larger domain) to all adjacent higher-dimensional entities of the larger domain along a certain path. This generalises the integration of boundary data observed in the previous chapter; in total, we may show the existence of a solution of the Kolmogorov backward equation in the entire domain for arbitrary boundary data.

Of particular interest to our discussion are stationary solutions of the Kolmogorov backward equation as they describe eventual hit probabilities for a certain target set of the model (in accordance with the probabilistic interpretation of solutions of the backward equation). The presented backward extension scheme allows the construction of solutions for all relevant cases. Eventually, in chapter 6, the hitherto missing uniqueness assertion for this type of solutions is established by means of a specific iterated transformation which resolves the critical incompatibilities of solutions by a successive blow-up while the domain is converted from a simplex into a cube. Then – under certain additional assumptions on the regularity of the transformed solution – the uniqueness directly follows from general principles. Lastly, several other aspects of the blow-up scheme are discussed; in particular, it is illustrated in what way the required extra regularity relates to reasonable additional properties of the underlying Wright–Fisher model.

2 A recombinational two-loci Wright–Fisher model and its diffusion approximation

In this chapter we are concerned with setting up a two-loci Wright–Fisher model with recombination and giving its diffusion approximation. However, we proceed indirectly and choose the differential equation for the transition probability density for such a model as our point of departure and subsequently identify and analyse the underlying Wright–Fisher models. This in particular involves an application of information geometrical tools, which will likewise be introduced.

2.1 The Ohta–Kimura formula

The differential equation for the transition probability density $f\colon \Omega_{(p,q,D)} \times (0,\infty) \longrightarrow \mathbb{R}$ of the diffusion approximation of a recombinational two-loci Wright–Fisher model given by T. Ohta and M. Kimura as presented in [10], p. 228 reads

$$\begin{aligned}\frac{\partial f}{\partial t} = &\frac{1}{4}p(1-p)\frac{\partial^2 f}{(\partial p)^2} + \frac{1}{4}q(1-q)\frac{\partial^2 f}{(\partial q)^2} + \frac{1}{2}D\frac{\partial^2 f}{\partial p \partial q} + \frac{1}{2}D(1-2p)\frac{\partial^2 f}{\partial p \partial D} \\ &+ \frac{1}{2}D(1-2q)\frac{\partial^2 f}{\partial q \partial D} + \frac{1}{4}\{pq(1-p)(1-q) + D(1-2p)(1-2q) - D^2\}\frac{\partial^2 f}{(\partial D)^2} \\ &- \frac{1}{2}D(1+2NR)\frac{\partial f}{\partial D}\end{aligned} \tag{2.1}$$

for $f(\,\cdot\,,t) \in C^2(\Omega_{(p,q,D)})$ for every $t>0$ and $f(p,q,D,\,\cdot\,) \in C^1((0,\infty))$ for $(p,q,D) \in \Omega_{(p,q,D)}$ and with

$$\begin{aligned}&\Omega_{(p,q,D)} := \\ &\left\{(p,q,D) \in \mathbb{R}^3 \middle| 0 < p,q < 1, \max(p+q-1,0) - pq < D < \min(p,q) - pq\right\}.\end{aligned} \tag{2.2}$$

Thus, the coefficient matrix of the 2nd order derivatives equals

$$\begin{aligned}&\big(a^{ij}(p,q,D)\big)\\ &:= \frac{1}{4}\begin{pmatrix} p(1-p) & D & D(1-2p) \\ D & q(1-q) & D(1-2q) \\ D(1-2p) & D(1-2q) & pq(1-p)(1-q)+D(1-2p)(1-2q)-D^2 \end{pmatrix}.\end{aligned} \tag{2.3}$$

Since a derivation of this formula is not provided in the literature, we will aim to make up for this in the following: Calculating the sectional curvatures of the statistical manifold corresponding to the given formula will help us identify the underlying probability distribution, which will then lead us to the corresponding Wright–Fisher model. In the following, we start with the introduction of some basic concepts.

2.2 Aspects of information geometry

2.2.1 Riemannian metrics

We briefly recapitulate the definitions (cf. [13], pp. 13 f.):

2.1 Definition. A *Riemannian manifold* (M, g) is a differentiable manifold M equipped with a *Riemannian metric* g, i. e. a scalar product on each tangent space T_pM which depends smoothly on the base point $p \in M$.

In local coordinates $x = (x^1, \dots, x^d)$, a Riemannian metric is given by a matrix $(g_{ij}(x))$. When changing coordinates to $\tilde{x}$, then this matrix transforms as

$$\tilde{g}_{kl}(\tilde{x}) = \sum_{i,j} g_{ij}(x) \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial x^j}{\partial \tilde{x}^l} \tag{2.4}$$

with $(\tilde{g}_{kl}(\tilde{x}))$ denoting the representation of the metric in the $\tilde{x}$-coordinates. This is due to the requirement that the scalar product of two tangent vectors be independent of the coordinate representation. Correspondingly, we have for $(g^{kl}(x))$, which denotes the inverse of $(g_{ij}(x))$,

$$\tilde{g}^{kl}(\tilde{x}) = \sum_{i,j} g^{ij}(x) \frac{\partial \tilde{x}^k}{\partial x^i} \frac{\partial \tilde{x}^l}{\partial x^j}. \tag{2.5}$$

Now, when stating a parabolic differential equation in some domain $\Omega_T := \Omega \times (0,T) \subset \mathbb{R}^n \times (\mathbb{R}_+ \cup \{\infty\})$ in the general form

$$\frac{\partial}{\partial t} u(x,t) = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} u(x,t) + \sum_{i=1}^{n} b^i(x) \frac{\partial}{\partial x^i} u(x,t) + c(x) u(x,t) \qquad \text{with } (x,t) \in \Omega_T, \tag{2.6}$$

we have for its transformation behaviour under coordinates changes:

2.2 Lemma. *The parabolic partial differential equation* (2.6) *transforms under a change of the spatial coordinates* $\Omega \longrightarrow \tilde{\Omega}, x \longmapsto \tilde{x}$ *into*

$$\frac{\partial}{\partial t} \tilde{u}(\tilde{x},t) = \sum_{k,l=1}^{n} \tilde{a}^{kl}(\tilde{x}) \frac{\partial^2}{\partial \tilde{x}^k \partial \tilde{x}^l} \tilde{u}(\tilde{x},t) + \sum_{k=1}^{n} \tilde{b}^k(\tilde{x}) \frac{\partial}{\partial \tilde{x}^k} \tilde{u}(\tilde{x},t) + \tilde{c}(\tilde{x}) \tilde{u}(\tilde{x},t) \tag{2.7}$$

with $\tilde{u}\big(\tilde{x}(x),t\big) = u(x,t)$ *and*

$$\tilde{a}^{kl}(\tilde{x}) = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial \tilde{x}^k}{\partial x^i} \frac{\partial \tilde{x}^l}{\partial x^j} \qquad \textit{for } k,l = 1,\dots,n, \tag{2.8}$$

$$\tilde{b}^k(\tilde{x}) = \sum_{i=1}^{n} b^i(x) \frac{\partial \tilde{x}^k}{\partial x^i} + \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 \tilde{x}^k}{\partial x^i \partial x^j} \qquad \textit{for } k = 1,\dots,n, \tag{2.9}$$

$$\tilde{c}(\tilde{x}) = c(x). \tag{2.10}$$

Proof. Let $\tilde{x}$ be a change of coordinates and $\tilde{u}$ such that $u(x,t) = \tilde{u}(\tilde{x}(x),t)$. Then we have by the chain rule

$$\sum_{i,j=1}^{n} a^{ij} \frac{\partial^2}{\partial x^i \partial x^j} u = \sum_{i,j=1}^{n} \sum_{k=1}^{n} a^{ij} \frac{\partial}{\partial x^i} \Big(\frac{\partial \tilde{x}^k}{\partial x^j} \frac{\partial}{\partial \tilde{x}^k} \tilde{u} \Big)$$

$$
\begin{aligned}
&= \sum_{i,j=1}^{n} \Big(\sum_{l,k=1}^{n} a^{ij} \frac{\partial \tilde{x}^l}{\partial x^i} \frac{\partial \tilde{x}^k}{\partial x^j} \frac{\partial^2}{\partial \tilde{x}^l \partial \tilde{x}^k} \tilde{u} + \sum_{k=1}^{n} a^{ij} \frac{\partial^2 \tilde{x}^k}{\partial x^i \partial x^j} \frac{\partial}{\partial \tilde{x}^k} \tilde{u} \Big) \\
&= \sum_{l,k=1}^{n} \sum_{i,j=1}^{n} a^{ij} \frac{\partial \tilde{x}^l}{\partial x^i} \frac{\partial \tilde{x}^k}{\partial x^j} \frac{\partial^2}{\partial \tilde{x}^l \partial \tilde{x}^k} \tilde{u} + \sum_{k=1}^{n} \sum_{i,j=1}^{n} a^{ij} \frac{\partial^2 \tilde{x}^k}{\partial x^i \partial x^j} \frac{\partial}{\partial \tilde{x}^k} \tilde{u} \qquad (2.11)
\end{aligned}
$$

and

$$
\sum_{i=1}^{n} b^i \frac{\partial}{\partial x^i} u = \sum_{i=1}^{n} \sum_{k=1}^{n} b^i \frac{\partial \tilde{x}^k}{\partial x^i} \frac{\partial}{\partial \tilde{x}^k} \tilde{u} = \sum_{k=1}^{n} \sum_{i=1}^{n} b^i \frac{\partial \tilde{x}^k}{\partial x^i} \frac{\partial}{\partial \tilde{x}^k} \tilde{u}. \qquad (2.12)
$$

Now putting $\tilde{a}^{lk}$, $\tilde{b}^k$ and $\tilde{c}$ as in equation (2.8), we have

$$
\sum_{i,j=1}^{n} a^{ij} \frac{\partial^2}{\partial x^i \partial x^j} u + \sum_{i=1}^{n} b^i \frac{\partial}{\partial x^i} u + cu = \sum_{l,k=1}^{n} \tilde{a}^{lk} \frac{\partial^2}{\partial \tilde{x}^l \partial \tilde{x}^k} \tilde{u} + \sum_{k=1}^{n} \tilde{b}^k \frac{\partial}{\partial \tilde{x}^k} \tilde{u} + \tilde{c}\tilde{u}. \qquad \square
$$

As the coefficient matrix of the second order derivatives (a^{ij}) – additionally to being positive definite – thus transforms like the coefficients (g^{ij}) of the inverse of the Riemannian metric $g = (g_{ij})$ under changes of coordinates (i. e. twice contravariantly), this allows interpreting a parabolic PDE on a domain Ω in terms of a Riemannian manifold (Ω, g) with the (inverse of the) corresponding metric g given by the coefficients $(g^{ij}) = (a^{ij})$. In fact, in section 2.2.4 we will show that for equation (2.1) (a^{ij}) coincides with (the inverse of) a Fisher metric on $\Omega_{(p,q,D)}$.

Here, $\Omega_{(p,q,D)} \subset \mathbb{R}^3$ clearly is a differentiable manifold, and it may be checked that $g = (g_{ij})$ given by $(g^{ij}) = (a^{ij})$ defines a scalar product (i. e. (g_{ij}) is symmetric and positive definite in $\Omega_{(p,q,D)}$) with its coefficients depending continuously on the base point. However, inverting (g^{ij}) yields a quite lengthy expression, which we state here only for completeness:

$$
(g_{i1})_{(p,q,D)} = 4 \det(g^{ij})^{-1} \begin{pmatrix} -(p-1)p(q-1)^2q^2-(2p-1)(q-1)q(2q-1)D+(-1-3(q-1)q)D^2 \\ -(p-1)p(q-1)qD+D^3 \\ D(q(-1-2p(q-1)+q-2D)+D) \end{pmatrix}, \qquad (2.13)
$$

$$
(g_{i2})_{(p,q,D)} = 4 \det(g^{ij})^{-1} \begin{pmatrix} -(p-1)p(q-1)qD+D^3 \\ -(p-1)^2p^2(q-1)q-(p-1)p(2p-1)(2q-1)D+(-1-3(p-1)p)D^2 \\ D(p(-1+p+2q-2pq-2D)+D) \end{pmatrix}, \qquad (2.14)
$$

$$(g_{i3})_{(p,q,D)} = 4\det(g^{ij})^{-1}\begin{pmatrix} D(q(-1-2p(q-1)+q-2D)+D) \\ D(p(-1+p+2q-2pq-2D)+D) \\ (p-1)p(q-1)q-D^2 \end{pmatrix}. \tag{2.15}$$

2.2.2 Curvature

To understand the geometrical properties of the problem given, we will calculate the sectional curvatures of the manifold $(\Omega_{(p,q,D)}, g)$ with (g^{ij}) being given by the coefficients of the 2nd order derivatives.

In accordance with [13], p. 164, we define:

2.3 Definition. The *sectional curvature* of the plane spanned by the (linearly independent) tangent vectors $X = \xi^i \frac{\partial}{\partial x^i}$, $Y = \eta^j \frac{\partial}{\partial x^j} \in T_x M$ of the Riemannian manifold (M, g) is

$$K(X \wedge Y) := \frac{R_{ijkl}\xi^i\eta^j\xi^k\eta^l}{(g_{ik}g_{jl} - g_{ij}g_{kl})\xi^i\eta^j\xi^k\eta^l}. \tag{2.16}$$

In the preceding formulae as well as throughout this section, the *summation convention* is employed, meaning that it is summed over all indices which appear both as upper and as lower index. Furthermore, R denotes the *Riemann curvature tensor*, whose components are given by

$$R^k_{lij} = \partial_i\Gamma^k_{jl} - \partial_j\Gamma^k_{il} + \Gamma^k_{ir}\Gamma^r_{jl} - \Gamma^k_{js}\Gamma^s_{il}. \tag{2.17}$$

The appearing Γ are the *Christoffel symbols* of the (unique) Levi–Civita connection for (M, g), which may be expressed in terms of the metric. Correspondingly, the Christoffel symbols of 1st kind are given by

$$\Gamma_{ijl} := \frac{1}{2}\big(\frac{\partial}{\partial x^j}g_{il} + \frac{\partial}{\partial x^i}g_{jl} - \frac{\partial}{\partial x^l}g_{ij}\big), \tag{2.18}$$

whereas those of 2nd kind are given by

$$\Gamma^k_{ij} := g^{kl}\Gamma_{ijl} = \frac{1}{2}g^{kl}\big(\frac{\partial}{\partial x^j}g_{il} + \frac{\partial}{\partial x^i}g_{jl} - \frac{\partial}{\partial x^l}g_{ij}\big) \tag{2.19}$$

respectively (cf. [13], p. 18).

However, equation (2.16) requires the purely covariant formulation of the curvature tensor, which is obtained by lowering the upper index, hence

$$R_{klij} = g_{km} R^m_{lij}. \tag{2.20}$$

The resulting formula may be simplified to some extent, which also reduces the calculative effort. Doing so, we first evaluate the derivatives occurring in R^m_{lij} (cf. equation (2.17)):

$$\begin{aligned}\partial_i \Gamma^m_{jl} - \partial_j \Gamma^m_{il} &= \partial_i (g^{ms} \Gamma_{jls}) - \partial_j (g^{mr} \Gamma_{ilr}) \\ &= (\partial_i g^{ms}) \Gamma_{jls} + g^{ms} (\partial_i \Gamma_{jls}) - (\partial_j g^{mr}) \Gamma_{ilr} - g^{mr} (\partial_j \Gamma_{ilr}).\end{aligned}$$

Now exploiting that, for the Levi-Civita connection ∇ of (M, g), we have $\nabla_i (g^{ms}) = 0$, i. e. $\partial_i g^{ms} = -\Gamma^m_{ai} g^{as} - \Gamma^s_{bi} g^{mb}$, thus

$$\begin{aligned}\partial_i \Gamma^m_{jl} - \partial_j \Gamma^m_{il} = &-\Gamma^m_{ai} g^{as} \Gamma_{jls} + \Gamma^m_{aj} g^{ar} \Gamma_{ilr} - \Gamma^s_{bi} g^{mb} \Gamma_{jls} + \Gamma^r_{bj} g^{mb} \Gamma_{ilr} \\ &+ g^{mr} (\partial_i \Gamma_{jlr} - \partial_j \Gamma_{ilr}).\end{aligned}$$

As $g^{as} \Gamma_{jls} = \Gamma^a_{jl}$ and $\Gamma^m_{ai} = \Gamma^m_{ia}$, the first two terms on the right-hand side cancel with the last two terms of equation (2.17), and now lowering the upper index m yields

$$R_{klij} = g_{km} R^m_{lij} = -\Gamma^s_{ki} \Gamma_{jls} + \Gamma^r_{kj} \Gamma_{ilr} + \partial_i \Gamma_{jlk} - \partial_j \Gamma_{ilk}$$

as $g_{km} g^{mb} = \delta^b_k$. Expressing the appearing Christoffel symbols of 2nd kind through those of 1st kind resp. employing their definition for the last two terms, we eventually arrive at

$$R_{klij} = \frac{1}{2} \left(\partial_i \partial_l g_{jk} + \partial_j \partial_k g_{il} - \partial_i \partial_k g_{jl} - \partial_j \partial_l g_{ik} \right) + g^{rs} \left(\Gamma_{kjs} \Gamma_{ilr} - \Gamma_{kir} \Gamma_{jls} \right), \tag{2.21}$$

which we will use for calculating the curvatures in the following (note $R_{klij} = R_{ijkl}$). We additionally note:

2.4 Remark. If the metric g is scaled with a factor λ, then the corresponding sectional curvatures are scaled by $\frac{1}{\lambda}$.

Proof. In equation (2.16), g appears twice in the denominator, while R_{klij} in the enumerator only contains either $\partial\partial g$ or terms of the form $g^{-1}\partial g \partial g$, which both deliver only one factor λ. □

2.2.3 Geometrical properties corresponding to the Ohta–Kimura formula

Analysing the geometry of the given problems results in:

2.5 Proposition. *The sectional curvatures of* $(\Omega_{(p,q,D)}, g)$ *with the inverse metric* (g^{ij}) *given by the coefficient matrix* (a^{ij}) *of the 2nd order derivatives of the Ohta–Kimura formula* (2.1) *are all (constantly) equal to* $\frac{1}{16}$.

Proof. We calculate the sectional curvature for a plane spanned by two coordinate axes $K_{ij} := K(\frac{\partial}{\partial x^i} \wedge \frac{\partial}{\partial x^j})$, which is according to equations (2.16) and (2.21) given by

$$\begin{aligned} K_{ij} &= \frac{R_{ijji}}{g_{ii}g_{jj} - (g_{ij})^2} \\ &= \frac{\frac{1}{2}\left(\partial_i\partial_i g_{jj} + \partial_j\partial_j g_{ii} - \partial_i\partial_j g_{ji} - \partial_j\partial_i g_{ij}\right) + g^{rs}\left(\Gamma_{jjs}\Gamma_{iir} - \Gamma_{jir}\Gamma_{jis}\right)}{g_{ii}g_{jj} - (g_{ij})^2} \\ &= \dots \\ &= \frac{1}{16}. \end{aligned}$$

□

The preceding result is essentially independent of any prefactors (i. e. $\frac{1}{2}$) added to the metric as in accordance with remark 2.4, this does not affect its property of describing a manifold of constant curvature.

Hence, as dealing with a manifold with constant curvature, it is appropriate to shift to more natural coordinates, i. e. coordinates which are adapted to the geometrical situation. Since for $n \in \mathbb{N}$ the *standard n-sphere*

$$S^n = \left\{ (x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} \,\middle|\, \sum_{i=1}^{n+1} \left(x^i\right)^2 = 1 \right\} \subset \mathbb{R}^{n+1} \tag{2.22}$$

is a manifold of constant (positive) curvature, we may adopt such sphere coordinates, which are given (for the positive sector S^n_+) by

$$\Big\{(x^1,\dots,x^n)\in\mathbb{R}^n \Big| x^i>0, \sum_{i=1}^{n}\left(x^i\right)^2<1\Big\} \tag{2.23}$$

as only n coordinates are needed as x^{n+1} is suppressible and may be calculated from the other coordinates. In these coordinates, the standard metric g^n of S^n (which is the metric induced on S^n from Euclidean metric on $\mathbb{R}^{n+1}$; cf. [12], pp. 137 f.) is given by

$$(g^n_{ij}(x)) = \begin{pmatrix} 1+\frac{(x^1)^2}{(x^{n+1})^2} & \frac{x^1x^2}{(x^{n+1})^2} & \cdots & \frac{x^1x^n}{(x^{n+1})^2} \\ \frac{x^1x^2}{(x^{n+1})^2} & 1+\frac{(x^2)^2}{(x^{n+1})^2} & \cdots & \frac{x^2x^n}{(x^{n+1})^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{x^1x^n}{(x^{n+1})^2} & \frac{x^2x^n}{(x^{n+1})^2} & \cdots & 1+\frac{(x^n)^2}{(x^{n+1})^2} \end{pmatrix} \tag{2.24}$$

with $x^{n+1} := \sqrt{1-\sum_{i=1}^{n+1}(x^i)^2}$; calculating the sectional curvatures of (S^n, g^n) directly verifies the property of being a manifold of positive curvature.

Furthermore, we may exploit the natural bijection between the positive sector of the n-sphere and the (open) n-dimensional *standard orthogonal simplex*

$$\Delta_n := \Big\{(x^1,\dots,x^n)\in\mathbb{R}^n \Big| x^i>0, \sum_{i=1}^{n} x^i<1\Big\}, \tag{2.25}$$

which is given by

$$\begin{aligned} S^n_+ &\longrightarrow \Delta_n, \\ (y^i) &\longmapsto (x^i) \quad \text{with } x^i := \left(y^i\right)^2 \text{ for } i=1,\dots,n. \end{aligned} \tag{2.26}$$

2.6 Lemma. *Under a change of coordinates* $(p,q,D)\longmapsto x=(x^1,x^2,x^3)$ *with*

$$x^1 := pq + D, \quad x^2 := p(1-q) - D, \quad x^3 := q(1-p) - D, \tag{2.27}$$

the domain $\Omega_{(p,q,D)}$ *is mapped onto* Δ_3*, while the coefficient matrix* $(a^{ij}(p,q,D))$ *of* *the 2nd order derivatives in equation* (2.1) *transforms into*

$$(\overline{a}^{ij}(x)) = \frac{1}{4}\begin{pmatrix} x^1(1-x^1) & -x^1x^2 & -x^1x^3 \\ -x^1x^2 & x^2(1-x^2) & -x^2x^3 \\ -x^1x^3 & -x^2x^3 & x^3(1-x^3) \end{pmatrix}, \quad x \in \Delta_3 \tag{2.28}$$

and

$$(\overline{a}_{ij}(x)) = 4\begin{pmatrix} (x^1)^{-1}+(x^4)^{-1} & (x^4)^{-1} & (x^4)^{-1} \\ (x^4)^{-1} & (x^2)^{-1}+(x^4)^{-1} & (x^4)^{-1} \\ (x^4)^{-1} & (x^4)^{-1} & (x^3)^{-1}+(x^4)^{-1} \end{pmatrix}, \quad x \in \Delta_3 \tag{2.29}$$

with $x^4 := 1 - \sum_{i=1}^3 x^i$ *respectively.*

Proof. The assertion on $\Omega_{(p,q,D)}$ may be checked straightforwardly. Furthermore, we have

$$\left(\frac{\partial x^i}{\partial p}, \frac{\partial x^i}{\partial q}, \frac{\partial x^i}{\partial D}\right)_{i=1,2,3} = \begin{pmatrix} q & p & 1 \\ 1-q & -p & -1 \\ -q & 1-p & -1 \end{pmatrix}. \tag{2.30}$$

Applying the formula for $\overline{a}^{ij}$ given in lemma 2.2 yields the desired result. □

Transforming further into sphere coordinates by

$$(x^i) \longmapsto (y^i) \quad \text{with } y^i := \sqrt{x^i} \text{ for } i = 1, 2, 3, \tag{2.31}$$

thus applying the natural bijection $\Delta_3 \longrightarrow S^3_+$ of equation (2.26), we have $\frac{\partial y^i}{\partial x^j} = \frac{1}{2\sqrt{x^j}}\delta^i_j$, from which we obtain analogously

$$(\hat{a}^{ij}(y)) = \frac{1}{16}\begin{pmatrix} 1-(y^1)^2 & -y^1y^2 & -y^1y^3 \\ -y^1y^2 & 1-(y^2)^2 & -y^2y^3 \\ -y^1y^3 & -y^2y^3 & 1-(y^3)^3 \end{pmatrix}, \quad y \in S^3_+ \tag{2.32}$$

and

$$(\hat{a}_{ij}(y)) = 16 \begin{pmatrix} 1 + \frac{(y^1)^2}{(y^4)^2} & \frac{y^1 y^2}{(y^4)^2} & \frac{y^1 y^3}{(y^4)^2} \\ \frac{y^1 y^2}{(y^4)^2} & 1 + \frac{(y^2)^2}{(y^4)^2} & \frac{y^2 y^3}{(y^4)^2} \\ \frac{y^1 y^3}{(y^4)^2} & \frac{y^2 y^3}{(y^4)^2} & 1 + \frac{(y^3)^2}{(y^4)^2} \end{pmatrix}, \quad y \in S^3_+ \tag{2.33}$$

with $y^4 := \sqrt{1 - \sum_{i=1}^3 (y^i)^2}$ respectively.

Thus, $(\hat{a}_{ij})$ resp. $(\bar{a}_{ij})$ coincide (up to the prefactor 16) with the standard metric g^3 of the 3-sphere $S^3_+ \subset \mathbb{R}^4$ (cf. equation (2.24)), which again confirms the statement of proposition 2.5 (cf. also remark 2.4). Moreover, $\bar{a}^{ij}(x)$ is (up to scaling and the missing prefactor N) the covariance matrix of the multinomial distribution $\mathcal{M}(N; p^1, \ldots, p^4)$ with parameters $p^i = x^i$, $i = 1, 2, 3$; $p^4 = 1 - \sum_{i=1}^3 x^i$: This already hints at its coincidence with the Fisher metric of the multinomial distribution on Δ_3 as will be illustrated in the following section.

2.2.4 The Fisher metric

Generally, the *Fisher information metric* of a suitable family of probability distributions on some domain Ω parametrised by $\theta = (\theta_1, \ldots, \theta_n) \in \Theta \subset \mathbb{R}^n$ with probability density functions $p(\theta) \colon \Omega \longrightarrow \mathbb{R}$ is given by (cf. [2], p. 27)

$$g_{ij}(\theta) := \mathrm{E}_{p(\theta)}\left(\frac{\partial}{\partial \theta_i} \log p(\theta) \frac{\partial}{\partial \theta_j} \log p(\theta)\right) \quad \text{for } i, j \in \{1, \ldots, n\},\ \theta \in \Theta, \tag{2.34}$$

wherein the expectation is to be taken with respect to $p(\theta)$ as indicated. It may be checked that this defines a Riemannian metric (cf. definition 2.1) on the parameter space Θ. Then, (Θ, g) is called a *statistical manifold*.

In the very common case of an *exponential family* with probability density function

$$p(\omega, \theta) = g(\omega) \exp\left(\sum_{i=1}^n \theta_i X_i(\omega) - \psi(\theta)\right) \tag{2.35}$$

with $X_1, \ldots, X_n$ being the corresponding *observables* (cf. [24]), we have for the Fisher metric

$$g_{ij}(\theta) = \mathrm{E}\left(\left(X_i - \frac{\partial}{\partial\theta_i}\psi\right)\left(X_j - \frac{\partial}{\partial\theta_j}\psi\right)\right) = \mathrm{E}(X_iX_j) - \mathrm{E}(X_i)\,\mathrm{E}(X_j),$$

$$i, j = 1, \dots, n, \quad (2.36)$$

as

$$\frac{\partial}{\partial\theta_i}\psi(\theta) = \mathrm{E}(X_i), \quad i = 1, \dots, n,$$

thus already the property of being the covariance matrix of the corresponding distribution.

For the multinomial distribution $\mathcal{M}(N; p^1, \dots, p^{n+1})$, $p \in \Delta_n$, which also forms an exponential family in their *natural parameters*

$$\theta_i := \log\left(\frac{p^i}{1 - \sum_{j=1}^n p^j}\right), \quad i = 1, \dots, n \tag{2.37}$$

and with $g(\omega) = \binom{N}{X_1(\omega), \dots, X_n(\omega), N - \sum_j X_j(\omega)}$ and $\psi(\theta) = N \log\left(1 + \sum_{j=1}^n \exp(\theta_j)\right)$, the Fisher metric g thus defines a Riemannian metric on the space of natural parameters Θ of the multinomial distribution with $(g_{ij}(\theta))$ coinciding with the corresponding covariance matrix (cf. equation (2.36)). However, this is only true for (g_{ij}) given in the θ-coordinates, whereas the Ohta–Kimura formula and its derivations are rather formulated in terms of the p-coordinates.

To formulate the Fisher metric of the multinomial distribution in the coordinates $p^1, \dots, p^n$ (the parameter p^{n+1} does not appear as coordinate due to $\sum_{i=1}^{n+1} p^i = 1$), we may change the coordinates adequately with the metric transforming in accordance with equation (2.4). Hence, we obtain

$$\bar{g}_{kl}(p) = \sum_{i,j=1}^n g_{ij}(\theta)\frac{\partial\theta_k}{\partial p^i}\frac{\partial\theta_l}{\partial p^j} = N^2 \sum_{i,j=1}^n g_{ij}(g_{ki})^{-1}(g_{lj})^{-1} = N^2 g^{kl}(\theta(p))$$

$$\text{for } k, l = 1, \dots, n,\ p \in \Delta_n \quad (2.38)$$

as we have

$$N\frac{\partial p^i}{\partial \theta_j} = \frac{\partial}{\partial \theta_j}\operatorname{E}(X_i) = \frac{\partial^2}{\partial \theta_i \partial \theta_j}\psi(\theta) = \operatorname{E}(X_i X_j) - \operatorname{E}(X_i)\operatorname{E}(X_j) \equiv g_{ij}(\theta), \qquad (2.39)$$

with the first equality holding particularly for the multinomial distribution. Thus, the Fisher metric on the parameter space in p-coordinates Δ_n is the inverse of the metric on the parameter space in θ-coordinates Θ (up to the factor N^2; cf. [12]). For the inverse metric, this further implies

$$\bar{g}^{kl}(p) = \frac{1}{N^2} g_{kl}(\theta(p)) = \frac{1}{N} p^k(\delta^k_l - p^l) \quad \text{for } k,l = 1,\ldots,n,\, p \in \Delta_n, \qquad (2.40)$$

signifying that the observed matrix $(\bar{a}^{ij}(x))$ in equation (2.28) is (up to a factor $\frac{1}{N}$) the inverse of the Fisher metric of the multinomial distribution on its parameter space in coordinates $p \equiv x$, i. e. Δ_3. Reversing the transformations of the preceding section and ignoring the factor N, we finally have:

2.7 Lemma. *The coefficients of the 2nd order derivatives of the Ohta–Kimura formula* (2.1) *equal (up to a constant factor) the components of the inverse of the Fisher metric of the multinomial distribution on* $\Omega_{(p,q,D)}$.

Likewise, we may now reformulate proposition 2.5:

2.8 Proposition. $\Omega_{(p,q,D)}$ *equipped with the Fisher metric of the multinomial distribution carries the geometrical structure of a manifold of constant positive curvature* $\equiv \frac{1}{16}$.

This probabilistic link via the Fisher metric hints on how the Ohta–Kimura formula comes about: When also transforming the 1st order derivatives into the x-coordinates in accordance with lemma 2.6, then – because of equation (2.30) and $\frac{\partial^2 x^1}{\partial p \partial q} = 1$, $\frac{\partial^2 x^2}{\partial p \partial q} = \frac{\partial^2 x^3}{\partial p \partial q} = -1$ as the only non-vanishing 2nd order coordinate derivatives – the formula reads

$$\frac{\partial}{\partial t}u(x,t) = \frac{1}{4}\sum_{i,j=1}^{3} x^i(\delta^i_j - x^j)\frac{\partial^2}{\partial x^i \partial x^j}u(x,t) + \sum_{i=1}^{3}(-\delta^i_1 + \delta^i_2 + \delta^i_3)NRD(x)\frac{\partial}{\partial x^i}u(x,t)$$
$$\text{for } (x,t) \in (\Delta_3)_\infty. \qquad (2.41)$$

In the following section, we will present a model from which this formula may be derived.

2.3 A recombinational two-loci Wright–Fisher model

2.3.1 General setup

To reflect gene dynamics, we employ a standard recombinational Wright–Fisher model as described for example in [10], pp. 67 f. A population of a fixed number $N \in \mathbb{N}_+$ of individuals *(zygotes)* is considered, which are intended to be the carriers of the genetic information that we are looking at. For simplicity, it is assumed that these individuals are *monoecious*, i. e. all belong to the same sex. The genetic configuration of the individuals is as follows:

- Each individual carries *two loci* of genetic information.
- Each locus contains two alleles *(diploidy)*.

Here, these alleles are chosen from a set of two alleles for each locus, e. g. at the A-locus the alleles A_1 and A_2 may appear, while for the B-locus we have alleles B_1 and B_2, and hence we write $\substack{A_1B_1\\A_1B_2}$ for a given zygote (with the two-lined notation indicating its diploidy). Altogether this makes up 4×4 possible genetic configurations for the individuals, but we will stipulate a symmetric identification, i. e. not to distinguish between individuals with the configuration $\substack{A_{i_1}B_{j_1}\\A_{i_2}B_{j_2}}$ and $\substack{A_{i_2}B_{j_2}\\A_{i_1}B_{j_1}}$. Hence, 10 different genetic configurations remain.

For means of reproduction, though, the zygotes produce *gametes*, which consist of again two loci, but with only one allele in each locus *(haploidy)*. Correspondingly, every zygote may be viewed as being formed by two gametes. Here, we have four different types of gametes, which are of course A_1B_1, A_1B_2, A_2B_1 and A_2B_2. Which type of gamete is produced by some zygote is dependent on whether *recombination* takes place or not. Without recombination, a zygote $\substack{A_{i_1}B_{j_1}\\A_{i_2}B_{j_2}}$ may either produce a gamete $A_{i_1}B_{j_1}$ or a gamete $A_{i_2}B_{j_2}$ (assumingly with equal probability of $\frac{1}{2}$), whereas a recombinational event may produce also the gametes $A_{i_1}B_{j_2}$ and $A_{i_2}B_{j_1}$ (again with equal probability).

Without recombination, the concurrence of the alleles at the two loci in the gamete produced is already determined: This is usually referred to as *linkage* between the two loci, since the alleles appearing there are 'linked' by their incidence at the zygote considered. Recombination opposes this behaviour, as it allows for combinations of alleles which may not be present at the input zygote. The recombinational events themselves are assumed to occur probabilistically: The *recombination rate* R may depend on the biological configuration and expresses to which extent the linkage between the two loci considered is lifted: $R = \frac{1}{2}$ signifies that no effect of linkage remains at all, whereas on the other hand, $R = 0$ implies that we effectively have a model with only one locus but with four different alleles at this locus instead. To avoid the latter case, $R > 0$ is assumed here. For later uses we also need to require that R depends on the population size N and that $R(N) \in \mathcal{O}(N^{-1})$.

2.3.2 Reproduction

Reproduction takes place in the model in two steps. In the first step, the current population (i. e. a collection of zygotes) produces a collection of gametes – possibly including recombination. In the next step, two gametes each are combined to build a new zygote, i. e. belonging to the next generation. There are different possibilities to handle the first step; the two most common models will be introduced in the following. For the second step, in both models multinomial sampling of the gametes that were produced in the first step is applied.

Random union of gametes (RUG)

In the *RUG model* (as proposed by Karlin and McGregor in [15]), a given population is viewed as a collection of gametes rather than as a collection of zygotes, which actually contain the gametes. Random union now signifies that two of these gametes are sampled randomly and combined. At this stage, recombination is integrated into the model: Every such double gamete now produces a new gamete, possibly under recombination (cf. fig. 2.1). These new gametes are then in the next step employed to form the next generation of zygotes.

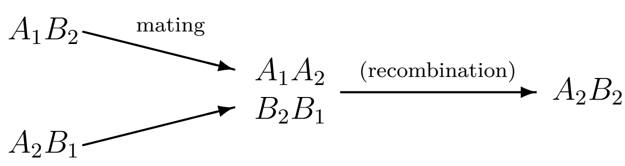

Figure 2.1: Random union of gametes

Random union of zygotes (RUZ)

In the *RUZ model* (as proposed by Watterson in [31]), mating really takes place between the zygotes of the current generation, in the sense that first two zygotes are sampled randomly, which then produces a gamete each, possibly under recombination. These two gametes are then conflated into a new double gamete. From these double gametes, again the zygotes of the next generation are formed.

Randomness of the transition

In order to clarify whether different parts of the transition are really random, a closer look as executed in [4] is required: The multinomial sampling in the second step – in a more precise sense – has to be seen as sampling from an infinite pool of gametes/pool of gamete pairs whose frequencies are determined by (infinitely many) RUG/RUZ-steps from the given population of zygotes. Correspondingly, also the mating interstages actually form a pool of gamete pairs (RUG)/gametes (RUZ) (as illustrated in the figures 2.2 and 2.3).

Hence, in the first step, we actually have deterministic behaviour, and recombinational effects only emerge in terms of altering the gamete frequencies in the pool, but no (random) fluctuations occurs. Randomness only applies to the second step, and consequently the properties of multinomial sampling determine the characteristics of the model to a large extent as will be seen in section 2.4.

2.3.3 Mathematical formulation of the reproduction mechanism

For both models, the quantitative analysis is best done in terms of gamete numbers as there are only four of them. Moreover, as the population size is fixed to N, it is appropriate to rather shift to frequencies, and hence we obtain four gamete frequencies, which will be denoted by $c = (c_1, \dots, c_4)$ for the gametes A_1B_1, A_1B_2, A_2B_1, A_2B_2. Due to their frequency property, i. e. $\sum_{i=1}^{4} c_i = 1$, actually already three frequencies yield an exhaustive description; the fourth one is often carried along for notational reasons, however. For the RUG model, the formulation in terms of gametes arises naturally as the population is viewed anyway by the gametes contained in a given zygote population, but it may similarly be applied also to the RUZ model (at a different stage, however).

Finally, we wish to construct a stochastic process[1] $(C(t))_{t\in\mathbb{N}} = (C_1(t), \dots, C_4(t))_{t\in\mathbb{N}}$ both for the RUG and the RUZ model depicting the evolution of gamete frequencies under successive transitions as described. Again, actually only three coordinates are needed as all four coordinates always sum up to 1. It is evident that the transitions only depend on the current state of the process (i. e. not on the previous history nor explicitly on the time parameter), thus such a process is *Markovian* and invariant under time shifts *(homogeneous in time)*. In the following, we will determine the distribution of $C(t+1)$ with the current situation $C(t) = c$ given in terms of its moments up to order 3.

The RUG model

The frequencies in the gamete pool (cf. figure 2.2), denoted here by $c' = (c'_1, \dots, c'_4)$, – thus after mating and recombination – may be calculated as follows (cf. [15]): A gamete A_1B_1 is produced

- with relative frequency 1 when sampling A_1B_1 twice,
- with relative frequency $\frac{1}{2}$ when sampling either A_1B_1 and A_1B_2 or A_1B_1 and A_2B_1,
- with relative frequency $\frac{1}{2}$ when sampling A_1B_1 and A_2B_2 in case of no recombination,
- with relative frequency $\frac{1}{2}$ when sampling A_1B_2 and A_2B_1 in case of recombination,

[1] The capitalisation of frequencies ($C(t)$) indicates that these are now random variables; their values, however, are still denoted by small letters. This notation scheme will essentially be kept for all other appearing random variables.

thus we have (with additional factors 2 by symmetry for sampling different alleles and with $c = (c_1, \dots, c_4)$ given):

$$\begin{aligned} c_1' =& c_1^2 + 2c_1c_2 \cdot \frac{1}{2} + 2c_1c_3 \cdot \frac{1}{2} \\ &+ 2c_1c_4 \cdot \frac{1}{2}(1-R) + 2c_2c_3 \cdot \frac{1}{2}R \\ =& c_1 - R(c_1c_4 - c_2c_3). \end{aligned} \tag{2.42}$$

For the remaining frequencies, we similarly obtain:

$$c_2' = c_2 + R(c_1c_4 - c_2c_3), \tag{2.43}$$

$$c_3' = c_3 + R(c_1c_4 - c_2c_3), \tag{2.44}$$

$$c_4' = c_4 - R(c_1c_4 - c_2c_3). \tag{2.45}$$

The appearing expression

$$c_1c_4 - c_2c_3 =: D(c) \tag{2.46}$$

is called the *coefficient of linkage disequilibrium*: If $D(c) = c_1 - (c_1 + c_2)(c_1 + c_3) \equiv 0$, the population is said to be in *linkage equilibrium*, signifying that the product of the gross frequencies of the alleles A_1 and B_1 equals the frequency of the gamete A_1B_1, i. e. no special concurrence of the alleles at the different loci exists (cf. 'linkage' on p. 20; simultaneously, an analogous statement holds true for all other combinations of alleles at different loci, e. g. A_2 and B_2).

Clearly, this coefficient of linkage disequilibrium is also time dependent through the frequencies $c(t)$, but if no confusion is to be feared, we may drop the arguments of D. For $D' := D(c') = c_1'c_4' - c_2'c_3'$ we have

$$D' = (c_1 - RD)(c_4 - RD) - (c_2 + RD)(c_3 + RD) \tag{2.47}$$

$$= (c_1c_4 - c_2c_3) - RD(c_1 + c_2 + c_3 + c_4) \tag{2.48}$$

$$= (1 - R)D. \tag{2.49}$$

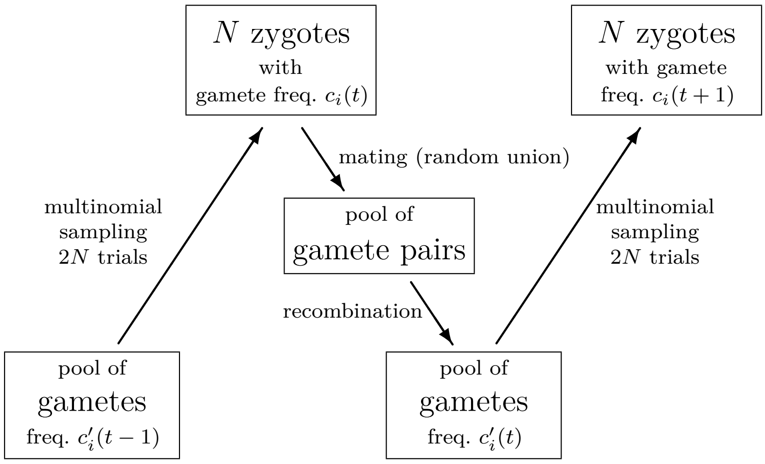

Figure 2.2: The RUG model

Hence, as $R > 0$, we have $D' < D$ due to the recombinational events, signifying that the effect of linkage is receding in the population. However, this perspective will be further developed in section 2.6.

The analysis of the multinomial sampling step is also straightforward. The gamete frequencies of the next generation $(C_1(t+1), \dots, C_4(t+1))$ given $C(t) = c$ are $\mathcal{M}(2N, (c_1', \dots, c_4'))$ distributed, and for the expectation values and second resp. third moments, we thus have:

$$\mathrm{E}(C_i(t+1)|C(t) = c) = c_i' = c_i \pm RD, \tag{2.50}$$

$$\begin{aligned} \mathrm{E}((C_iC_j)(t+1)|C(t) = c) &= \left(1 - \frac{1}{2N}\right) c_i'c_j' + \frac{1}{2N} c_i' \delta_j^i \\ &= \left(1 - \frac{1}{2N}\right)(c_i \pm RD)(c_j \pm RD) + \frac{1}{2N}(c_i \pm RD)\delta_j^i, \end{aligned} \tag{2.51}$$

$$\begin{aligned} \mathrm{E}((C_iC_jC_k)(t+1)|C(t) = c) &= \left(1 - \frac{3}{2N} + \frac{2}{(2N)^2}\right) c_i'c_j'c_k' + \frac{1}{(2N)^2} c_i' \delta_j^i \delta_k^j \\ &\quad + \left(\frac{1}{2N} - \frac{1}{(2N)^2}\right)\left(c_i'c_k'\delta_j^i + c_i'c_j'\delta_k^j + c_j'c_k'\delta_i^k\right). \end{aligned} \tag{2.52}$$

For the increments $\delta C_i := C_i(t+1) - C_i(t)$, this yields (the notation $\mathrm{E}_{\delta t}$ indicates that this is the expectation of the increment of the process within time δt):

$$\mathrm{E}_1(\delta C_i | C(t) = c) = \pm RD \tag{2.53}$$

and

$$\begin{aligned}
&\mathrm{E}_1(\delta C_i \delta C_j | C(t) = c)\\
&\qquad = \mathrm{E}((C_i C_j)(t+1) | C(t) = c) + \mathrm{E}_1(\delta C_i | C(t) = c)\, \mathrm{E}_1(\delta C_j | C(t) = c)\\
&\qquad\quad - \mathrm{E}(C_i(t+1) | C(t) = c)\, \mathrm{E}(C_j(t+1) | C(t) = c)\\
&\qquad = \frac{1}{2N}\big(c_i' \delta_j^i - c_i' c_j'\big) \pm (RD)^2\\
&\qquad = \frac{1}{2N}\big((c_i \pm RD)\delta_j^i - (c_i \pm RD)(c_j \pm RD)\big) \pm (RD)^2\\
&\qquad = \frac{1}{2N}\big(c_i \delta_j^i - c_i c_j \pm (\delta_j^i \mp c_i \mp c_j) RD \mp (RD)^2\big) \pm (RD)^2
\end{aligned} \tag{2.54}$$

as well as

$$\begin{aligned}
&\mathrm{E}_1(\delta C_i \delta C_j \delta C_k | C(t) = c)\\
&\quad = \mathrm{E}((C_i C_j C_k)(t+1) | C(t) = c) - c_i c_j c_k\\
&\qquad - \mathrm{E}((C_i C_j)(t+1) | C(t) = c) c_k + \mathrm{E}(C_k(t+1) | C(t) = c) c_i c_j\\
&\qquad - \mathrm{E}((C_i C_k)(t+1) | C(t) = c) c_j + \mathrm{E}(C_j(t+1) | C(t) = c) c_i c_k\\
&\qquad - \mathrm{E}((C_j C_k)(t+1) | C(t) = c) c_i + \mathrm{E}(C_i(t+1) | C(t) = c) c_j c_k\\
&\quad = \left(1 - \frac{3}{2N} + \frac{2}{(2N)^2}\right) c_i' c_j' c_k' + \left(\frac{1}{2N} - \frac{1}{(2N)^2}\right)\big(c_i' c_k' \delta_j^i + c_i' c_j' \delta_k^j + c_j' c_k' \delta_i^k\big)\\
&\qquad + \frac{1}{(2N)^2} c_i' \delta_j^i \delta_k^j - \left(1 - \frac{1}{2N}\right)(c_i' c_j' c_k + c_i' c_k' c_j + c_j' c_k' c_i)\\
&\qquad - \frac{1}{2N}\big(c_i' \delta_j^i c_k + c_i' \delta_i^k c_j + c_j' \delta_k^j c_i\big) + c_i c_j c_k' + c_i c_j' c_k + c_i' c_j c_k - c_i c_j c_k\\
&\quad = (c_i' - c_i)(c_j' - c_j)(c_k' - c_k) + \frac{1}{4N^2}\big(2 c_i' c_j' c_k' - c_i' c_k' \delta_j^i - c_i' c_j' \delta_k^j - c_j' c_k' \delta_i^k + c_i' \delta_j^i \delta_k^j\big)\\
&\qquad + \frac{1}{2N}\big(c_i' c_j' (1 - \delta_j^i)(c_k - c_k') + c_i' c_k' (1 - \delta_k^i)(c_j - c_j') + c_j' c_k' (1 - \delta_k^j)(c_i - c_i')\big)\\
&\quad = \pm (RD)^3 + \frac{RD}{2N}\big(\pm c_i' c_j' (1 - \delta_j^i) \pm c_i' c_k' (1 - \delta_k^i) \pm c_j' c_k' (1 - \delta_k^j)\big)
\end{aligned}$$

$$+\frac{1}{4N^2}\big(2c'_i c'_j c'_k - c'_i c'_k \delta^i_j - c'_i c'_j \delta^j_k - c'_j c'_k \delta^k_i + c'_i \delta^i_j \delta^j_k\big). \tag{2.55}$$

The RUZ model

The RUZ model as described by Watterson in [31] comes along somewhat more complicated, but is effectively not very different from the RUG model. However, the assessment of the model in terms of gamete frequencies now refers rather to the gamete pool (cf. figure 2.3) than to the population itself, which is mainly due to technical reasons.

For a population of zygotes ${}^{i}_{j}$ containing gametes $i, j = 1, \ldots, 4$, the relative frequency with which a zygote[2] containing gametes i and j produces a gamete k is given by a table $(\rho^{\substack{i\\j}}_k)$, which directly reflects the considerations above for the RUG model, i. e.

$$\begin{array}{r|cccc}
\rho^{\substack{i\\j}}_k & k=1 & k=2 & k=3 & k=4 \\
\hline
(i,j)=(1,1) & 1 & 0 & 0 & 0 \\
(1,2) & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
(1,3) & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
(1,4) & \frac{1}{2}(1-R) & \frac{1}{2}R & \frac{1}{2}R & \frac{1}{2}(1-R) \\
(2,2) & 0 & 1 & 0 & 0 \\
(2,3) & \frac{1}{2}R & \frac{1}{2}(1-R) & \frac{1}{2}(1-R) & \frac{1}{2}R \\
(2,4) & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
(3,3) & 0 & 0 & 1 & 0 \\
(3,4) & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
(4,4) & 0 & 0 & 0 & 1
\end{array} \tag{2.56}$$

for $i \le j$; for $i > j$, the values for $\rho^{\substack{i\\j}}_k$ are given by $\rho^{\substack{i\\j}}_k = \rho^{\substack{j\\i}}_k$. With given zygote frequencies $x_{\substack{i\\j}}$, the relative frequency with which a gamete k is obtained after uniformly sampling a zygote hence equals

$$c'_k = \sum_{i,j=1}^{4} x_{\substack{i\\j}} \rho^{\substack{i\\j}}_k \quad \text{for } k = 1, \ldots, 4, \tag{2.57}$$

[2]For simplicity, here we do not use the symmetric identification of zygotic types. This is without any persisting effect as the results will be formulated in terms of gamete frequencies.

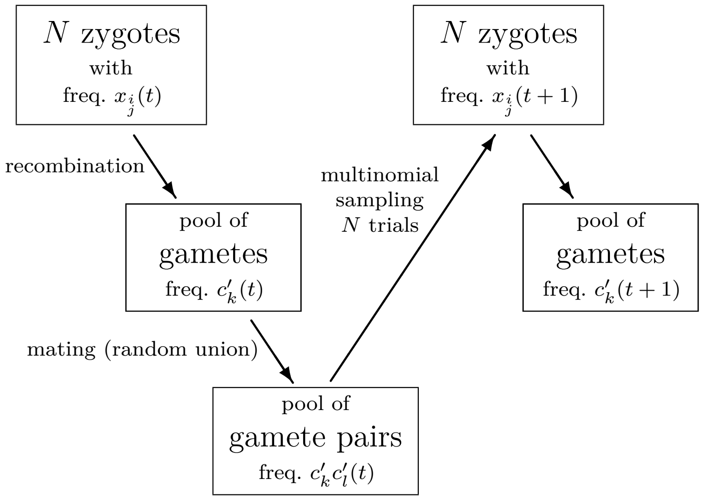

Figure 2.3: The RUZ model

and after random mating, we obtain a pool of gamete pairs (k, l) with frequencies $(c'_k c'_l)$ (for systematic reasons, variables referring to the pool of gametes/gamete pairs are primed throughout).

In the next step, multinomial sampling with these pair frequencies taken as parameters then results in a probability distribution for $(X_{\substack{i\\j}}(t+1))$, from which again the distribution of the frequencies $(C'_k(t+1))$ in the gamete pool may be calculated in accordance with equation (2.57); the corresponding moment generating function for $(C'_k(t+1))$ is thus given by (cf. [31])

$$M_{C'}(\theta) = \mathrm{E}\left(e^{\sum_{k=1}^4 C'_k(t+1)\theta_k} \middle| C'(t) = c'\right) = \left(\sum_{i,j=1}^{4} c'_i c'_j e^{\sum_{k=1}^4 \theta_k \rho_k^{\substack{i\\j}}}\right)^N \quad \text{with } \theta = (\theta_1, \dots, \theta_4) \in \mathbb{R}^4, \quad (2.58)$$

which may be used to calculate the moments of the process.

However, this formula actually corresponds to swapping the first and the second step of the transition, i. e. first sampling multinomially from a pool of gamete pairs with frequencies $(c'_i c'_j)$ to obtain a new population of zygotes $(X_{\substack{i\\j}}(t+1))$ and thereafter letting it produce (infinitely many) new gametes with frequencies $(C'_k(t+1))$ in accordance with equation (2.57) (from which again a pool of gamete pairs with frequencies $(C'_k C'_l(t+1))$ may be obtained). Both $(X_{\substack{i\\j}}(t+1))$ and $(C'_k(t+1))$ are random variables then, with the latter depending deterministically on $(X_{\substack{i\\j}})$.

Hence, we are able to calculate the moments of $(C'_k(t+1))$ directly from properties of the multinomial distribution (yielding the same results as when using the moment generating function in equation (2.58)):

$$\begin{aligned}
\mathrm{E}(C'_k(t+1)|C'(t)=c') &= \mathrm{E}\left(\sum_{i,j} X_{\substack{i\\j}}(t+1)\rho^{\substack{i\\j}}_k \middle| C'(t)=c'\right)\\
&= \sum_{i,j} \mathrm{E}(X_{\substack{i\\j}}(t+1)|C'(t)=c')\rho^{\substack{i\\j}}_k\\
&= \sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k
\end{aligned} \tag{2.59}$$

and

$$\begin{aligned}
\mathrm{E}((C'_k C'_l)(t+1)|C'(t)=c') &= \sum_{i,j}\sum_{m,n} \mathrm{E}((X_{\substack{i\\j}} X_{\substack{m\\n}})(t+1)|C'(t)=c')\rho^{\substack{i\\j}}_k \rho^{\substack{m\\n}}_k\\
&= \sum_{i,j}\sum_{m,n}\left(\left(1-\frac{1}{N}\right)c'_i c'_j c'_m c'_n + \frac{1}{N} c'_i c'_j \delta^{mn}_{ij}\right)\rho^{\substack{i\\j}}_k \rho^{\substack{m\\n}}_k\\
&= \left(1-\frac{1}{N}\right)\sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \sum_{m,n} c'_m c'_n \rho^{\substack{m\\n}}_k\\
&\quad + \frac{1}{N}\sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k
\end{aligned} \tag{2.60}$$

as well as

$$\begin{aligned}
&\mathrm{E}((C'_k C'_l C'_m)(t+1)|C'(t)=c')\\
&\quad= \sum_{i,j}\sum_{p,q}\sum_{r,s} \mathrm{E}((X_{\substack{i\\j}} X_{\substack{p\\q}} X_{\substack{r\\s}})(t+1)|C'(t)=c')\rho^{\substack{i\\j}}_k \rho^{\substack{p\\q}}_k \rho^{\substack{r\\s}}_k\\
&\quad= \sum_{i,j}\sum_{p,q}\sum_{r,s}\left(\left(1-\frac{3}{N}+\frac{2}{N^2}\right)c'_i c'_j c'_p c'_q c'_r c'_s + \left(\frac{1}{N}-\frac{1}{N^2}\right)\times\right.\\
&\qquad \left.\left(c'_i c'_j c'_r c'_s \delta^{ij}_{pq} + c'_i c'_j c'_p c'_q \delta^{pq}_{rs} + c'_p c'_q c'_r c'_s \delta^{rs}_{ij}\right) + \frac{1}{N^2} c'_i c'_j \delta^{ij}_{pq}\delta^{pq}_{rs}\right)\rho^{\substack{i\\j}}_k \rho^{\substack{p\\q}}_k \rho^{\substack{r\\s}}_k
\end{aligned}$$

$$
\begin{aligned}
&= \Big(1 - \frac{3}{N} + \frac{2}{N^2}\Big) \sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \sum_{p,q} c'_p c'_q \rho^{\substack{p\\q}}_k \sum_{r,s} c'_r c'_s \rho^{\substack{r\\s}}_k + \Big(\frac{1}{N} - \frac{1}{N^2}\Big) \times \\
&\Big(\sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k \sum_{r,s} c'_r c'_s \rho^{\substack{r\\s}}_k + \sum_{p,q} c'_p c'_q \rho^{\substack{p\\q}}_k \rho^{\substack{p\\q}}_k \sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \\
&+ \sum_{r,s} c'_r c'_s \rho^{\substack{r\\s}}_k \rho^{\substack{r\\s}}_k \sum_{p,q} c'_p c'_q \rho^{\substack{p\\q}}_k \Big) + \frac{1}{N^2} \sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k .
\end{aligned}
\tag{2.61}
$$

The sum appearing in the first equation evaluates analogously to the RUG model to

$$
\sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k = c'_k \pm RD' \tag{2.62}
$$

(with $D' = D(c')$, cf. p. 23), while for the sum appearing in the second equation, we get the more complicated result

$$
\sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k = \frac{1}{2} \left(c'_k \delta^k_l + c'_k c'_l \pm R \left(\begin{Bmatrix} c'_1 c'_4 \\ c'_2 c'_3 \end{Bmatrix} + \begin{Bmatrix} c'_2 c'_3 \\ c'_1 c'_4 \end{Bmatrix} \right) \mp R^2 S' \right) \tag{2.63}
$$

with $S' := c'_1 c'_4 + c'_2 c'_3 =: f_1(c')$ and the braces indicating that in each place either the upper or the lower term is chosen dependent on the indices k and l: If the gametes k and l share exactly one allele, at both places the upper entry is selected as well as the upper sign in $\pm$ and $\mp$ is selected; in all other cases, at one of the braces, the upper entry is selected, whereas at the other place, the lower entry is selected (or conversely) as well as the lower sign both in $\pm$ and $\mp$ is selected. The last sum $\sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k$ may also be computed analogously, however, this is not necessary for our purposes as will turn out later and hence not carried out here.

For the increment expectations, we obtain consequently:

$$
\mathrm{E}_1(\delta C'_k | C'(t) = c') = \pm RD', \tag{2.64}
$$

$$
\begin{aligned}
\mathrm{E}_1(\delta C'_k \delta C'_l | C'(t) = c') &= -\frac{1}{N} \sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \sum_{m,n} c'_m c'_n \rho^{\substack{m\\n}}_k + \frac{1}{N} \sum_{i,j} c'_i c'_j \rho^{\substack{i\\j}}_k \rho^{\substack{i\\j}}_k \pm (RD')^2 \\
&= -\frac{1}{N} (c'_k \pm RD')(c'_l \pm RD') \\
&\quad + \frac{1}{2N} \left(c'_k \delta^k_l + c'_k c'_l \pm R \left(\begin{Bmatrix} c'_1 c'_4 \\ c'_2 c'_3 \end{Bmatrix} + \begin{Bmatrix} c'_2 c'_3 \\ c'_1 c'_4 \end{Bmatrix} \right) \mp R^2 S' \right) \pm (RD')^2
\end{aligned}
$$

$$= \frac{1}{2N}\Big(c'_k\delta^k_l - c'_kc'_l \pm R\Big(2(\mp c'_k \mp c'_l)D' + \begin{Bmatrix} c'_1c'_4 \\ c'_2c'_3 \end{Bmatrix} + \begin{Bmatrix} c'_2c'_3 \\ c'_1c'_4 \end{Bmatrix}\Big)$$
$$\mp R^2(S' \pm D'^2)\Big) \pm (RD')^2 \tag{2.65}$$

and

$$\begin{aligned}
&\mathrm{E}_1(\delta C'_k\delta C'_l\delta C'_m|C'(t)=c')\\
&\quad= \Big(1-\frac{3}{N}+\frac{2}{N^2}\Big)\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k \sum_{p,q} c'_pc'_q\rho^{\substack{p\\q}}_k \sum_{r,s} c'_rc'_s\rho^{\substack{r\\s}}_k + \Big(\frac{1}{N}-\frac{1}{N^2}\Big)\times\\
&\quad\Big(\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k\rho^{\substack{i\\j}}_k \sum_{r,s} c'_rc'_s\rho^{\substack{r\\s}}_k + \sum_{p,q} c'_pc'_q\rho^{\substack{p\\q}}_k\rho^{\substack{p\\q}}_k \sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k\\
&\quad+\sum_{r,s} c'_rc'_s\rho^{\substack{r\\s}}_k\rho^{\substack{r\\s}}_k \sum_{p,q} c'_pc'_q\rho^{\substack{p\\q}}_k\Big) + \frac{1}{N^2}\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k\rho^{\substack{i\\j}}_k\rho^{\substack{i\\j}}_k - c'_kc'_lc'_m\\
&\quad-\Big(\Big(1-\frac{1}{N}\Big)\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k \sum_{r,s} c'_rc'_s\rho^{\substack{r\\s}}_k + \frac{1}{N}\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k\rho^{\substack{i\\j}}_k\Big)c'_m + c'_kc'_l(c'_m \pm RD')\\
&\quad-\Big(\Big(1-\frac{1}{N}\Big)\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k \sum_{r,s} c'_rc'_s\rho^{\substack{r\\s}}_k + \frac{1}{N}\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k\rho^{\substack{i\\j}}_k\Big)c'_l + c'_kc'_m(c'_l \pm RD')\\
&\quad-\Big(\Big(1-\frac{1}{N}\Big)\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k \sum_{r,s} c'_rc'_s\rho^{\substack{r\\s}}_k + \frac{1}{N}\sum_{i,j} c'_ic'_j\rho^{\substack{i\\j}}_k\rho^{\substack{i\\j}}_k\Big)c'_k + c'_lc'_m(c'_k \pm RD')\\
&\quad= \pm(RD')^3 - \frac{R}{N}f_1(c') + \frac{1}{N^2}f_2(c')
\end{aligned} \tag{2.66}$$

with $f_1(c'), f_2(c')$ denoting some functions purely depending on (c'_1, c'_2, c'_3).

Regarding the notation used here, in the long run, i. e. if multiple transitions are performed, it is not crucial whether the model is viewed in terms of the evolution of gamete frequencies either in the population or in the gamete pool. For this reason, when generalising the models in section 2.4 ff., we may formulate the RUZ model analogously to the RUG model by using unprimed variables.

RUG and RUZ in comparison

Except for interchanging the mating/recombination step and the sampling step and consequently being actually formulated in terms of (c_i) resp. (c'_i), the differences between the RUG model and RUZ model are only subtle as the procedures applied for the two steps are basically analogous. The only systematic discrepancy is that, in the RUG model, a given population is merely viewed as a set of gametes, whereas in the RUZ model, this set of gametes is supposed to carry an additional structure, i. e. that each two gametes are paired into a zygote.

To assess the effect of this difference, we may formulate the RUG model equivalently to the RUZ model by also swapping the sampling step and the mating/recombination step and check whether the transition probabilities for one full transition of the pool of gametes (c'_i) are identical. By doing so, we may reuse $(\rho^{\substack{i\\j}}_k)$ (cf. table (2.56)), but this time with (C_iC_j) with $(C_i) \sim \mathcal{M}(2N, (c'_i))$ as input instead of $(X_{\substack{i\\j}})$ in the RUZ model, and for the frequencies (C'_k) in the gamete pool, we thus have

$$C'_k = \sum_{i,j=1}^{4} C_iC_j\rho^{\substack{i\\j}}_k \quad \text{for } k = 1, \ldots, 4, \tag{2.67}$$

yielding

$$M_{C'}(\theta) = \mathrm{E}\left(e^{\sum_{k=1}^{4} C'_k\theta_k}\right) = \mathrm{E}\left(e^{\sum_{i,j=1}^{4} C_iC_j \sum_{k=1}^{4} \rho^{\substack{i\\j}}_k\theta_k}\right) \quad \text{with } \theta = (\theta_1, \ldots, \theta_4) \in \mathbb{R}^4 \tag{2.68}$$

when calculating the moment generating function.

This, however, is different from the corresponding result for the RUZ model in equation (2.58) as the distribution of (C_iC_j) in the RUG model is generally different from that of $(X_{\substack{i\\j}}) \sim \mathcal{M}(N, (c'_ic'_j))$ in the RUZ model: For example for $N = 1$ and (c'_i) given, we have $P_{C^2}(C_1C_1 > 0) = {c'_1}^2 + 2c'_1(c'_2 + c'_3 + c'_4) = c'_1(2 - c'_1)$ in contrast to $P_X(X_{\substack{1\\1}} > 0) = {c'_1}^2$. Translated back to the model, this example reveals the effect of the additional zygotic structure of the RUZ population: From a population comprising the gamete 1 and another gamete different from 1 in the RUG model, still the gamete 1 may be sampled twice for mating, whereas from a population of just one zygote other than $\genfrac{}{}{0pt}{}{1}{1}$ in the RUZ model, no zygote $\genfrac{}{}{0pt}{}{1}{1}$ may be generated. However, in general this effect is only slight as it diminishes with increasing population size.

Moreover, assuming that no recombination is allowed, the RUZ model effectively becomes forgetful about the zygotic structure, and then both models agree: Without recombination, $(\rho^{\substack{i\\j}}_k)$ for both models turns into $\rho^{\substack{i\\j}}_k = \frac{1}{2}(\delta^i_k + \delta^j_k)$, and equation (2.57) then reads

$$C'_k = \sum_{i,j=1}^{4} X_{\substack{i\\j}} \rho^{\substack{i\\j}}_k = \frac{1}{2}\Big(\sum_{j=1}^{4} X_{\substack{k\\j}} + \sum_{i=1}^{4} X_{\substack{i\\k}}\Big) \equiv C_k \quad \text{for } k = 1, \ldots, 4, \tag{2.69}$$

which is just the frequency of gamete k in the (multinomially sampled) population $(X_{\substack{i\\j}})$. Hence, the RUZ model then only depends on the gamete frequencies in the population irrespective of their belonging with a certain zygote (the same then trivially also holds for the RUG model). We thus obtain

$$\begin{aligned} M_{C'(X)}(\theta) = \mathrm{E}\left(e^{\sum_i C'_i(X)\theta_i}\right) &= \mathrm{E}\left(e^{\frac{1}{2}\sum_i(\sum_j X_{\substack{i\\j}} + \sum_j X_{\substack{j\\i}})\theta_i}\right) \\ &= \mathrm{E}\left(e^{\frac{1}{2}\sum_{i,j} X_{\substack{i\\j}}(\theta_i+\theta_j)}\right) \\ &= \Big(\sum_{i,j} c'_i c'_j e^{\frac{\theta_i+\theta_j}{2N}}\Big)^N \end{aligned} \tag{2.70}$$

when calculating the moment generating function for the transition probability of the RUZ model.

Analogously, for the RUG model, equation (2.68) without recombination turns into

$$\begin{aligned} M_{C'}(\theta) &= \mathrm{E}\left(e^{\frac{1}{2}\sum_k(\sum_j C_k C_j + \sum_i C_i C_k)\theta_k}\right) \\ &= \mathrm{E}\left(e^{\sum_k C_k\theta_k}\right) \\ &= \Big(\sum_i c'_i e^{\frac{\theta_i}{2N}}\Big)^{2N} = \Big(\sum_{i,j} c'_i c'_j e^{\frac{\theta_i+\theta_j}{2N}}\Big)^N, \end{aligned} \tag{2.71}$$

implying that the transition probabilities for both models agree, as multinomially sampling N zygotes from a pool of gamete pairs with frequencies $(c'_i c'_j)$ yields the same distribution for the gametes as directly sampling $2N$ gametes from a pool of gametes with frequencies (c'_i).

Again with recombination, the zygotic structure affecting the recombination still does not alter the increment expectations of the RUZ model in comparison with those of the RUG model (cf. equations (2.53) and (2.64)), but it becomes visible in the increment product expectation (cf. equations (2.54) and (2.65)), where the

R-terms differ. However, as stated, this effect diminishes with increasing population size, and correspondingly, the non-agreeing terms are scaled by a prefactor $\frac{1}{N}$.

In accordance with these observations, it is without much systematic difference which of the two models is used as long as N is sufficiently large. This is in particular the case for the diffusion approximation presented in the next section, where we have $N \to \infty$ and consequently may show that both models yield the same results (cf. proposition 2.9). For this reason, in further investigations – in particular for generalisation as in chapter 3 – we may only consider the RUG model as it is the simpler[3] and more transparent one.

2.4 The diffusion approximation

In order to allow for a better analytic treatment, often a continuous limit process of discrete processes (as arising from the Wright–Fisher model) is considered. This limit process is usually referred to as *diffusion approximation* (cf. [29], pp. 129 f.) and accomplished by scaling both the underlying space and the time parameter of the discrete process such that a process with continuous state space (after suitable normalisation if necessary) and continuous time parameter is obtained as scaling limit.

In the setting of the Wright–Fisher model, the scaling parameter is given by the size N of the population considered; thus, the process $(C(t))_{t\in\mathbb{N}}$ may be marked with an index N to indicate the size of underlying space resp. that it takes its values in the state space $\Delta_{\frac{3}{N}}$ given by

$$\Delta_{\frac{3}{N}} := \left\{(c_1, c_2, c_3) \in \mathbb{N}^3_{N^{-1}} \middle| \sum_{i=1}^{3} c_i < 1\right\} \tag{2.72}$$

with $\mathbb{N}_{N^{-1}} = \left\{\frac{k}{N} \middle| k \in \mathbb{N}\right\}$[4]. As this state space already is given by frequencies, it does not need to be rescaled itself. However, we may still additionally rescale time

[3] Cp. Ockham's razor.

[4] This does not really apply to the RUZ model as there the state space is given by the frequencies in the (infinite) gamete pool, which may be set to any arbitrary value. Nevertheless, as a result of the transition with a finite population as interstage, analogously only discrete values may occur.

by N, thus having N transitions within one time step. We denote such a doubly rescaled process by $(\hat{C}_N(t))_{t\in\mathbb{N}_{N^{-1}}}$ with accordingly $\hat{C}_N(t) := C_N(Nt)$, $t \in \mathbb{N}_{N^{-1}}$, while the continuous limit process is denoted by $(X(t))$ with corresponding state space $\Delta_3 = \{(x^1, x^2, x^3) \in \mathbb{R}^3 \big| x^i > 0, i = 1, 2, 3; \sum_{i=1}^{3} x^i < 1\}$. Throughout our further analysis, we always use variables with *upper* indices when referring to a continuous limit processes, whereas variables with lower index refer to a discrete process.

2.4.1 The Kolmogorov equations

Eventually, the dynamics of such a continuous process may be described by certain differential equations for its probability density, the Kolmogorov equations (cf. [20], pp. 445 ff.).

Generally, for a continuous process $X(t) = (X^i(t))_{i=1,\dots,n}$ with values in $\Omega \subset \mathbb{R}^n$ and $t \in (t_0, t_1) \subset \mathbb{R}$ satisfying certain conditions, in particular

$$\lim_{\delta t\to 0} \frac{1}{\delta t} \mathrm{E}_{\delta t}\big(\delta X^i \big| X(t) = x\big) =: \mu^i(x,t), \quad i = 1,\dots,n \tag{2.73}$$

and

$$\lim_{\delta t\to 0} \frac{1}{\delta t} \mathrm{E}_{\delta t}\big(\delta X^i \delta X^j \big| X(t) = x\big) =: \sigma^{ij}(x,t), \quad i,j = 1,\dots,n \tag{2.74}$$

existing for all $(x,t) \in \Omega \times (t_0, t_1)$ and further

$$\lim_{\delta t\to 0} \frac{1}{\delta t} \mathrm{E}_{\delta t}\big((\delta X)^\alpha \big| X(t) = x\big) = 0 \tag{2.75}$$

for all $(x,t) \in \Omega \times (t_0, t_1)$ and all multi-indices $\alpha := (\alpha_1, \dots, \alpha_n)$ with $|\alpha| \geq 3$, and with probability density $f(p,s,x,t) := \frac{\partial^n}{\partial x^1 \cdots \partial x^n} P(X(t) \leq x | X(s) = p)$ with $s < t$, we have:

- the *Kolmogorov forward equation* (also known as the *Fokker–Planck equation*)

$$\frac{\partial}{\partial t} f(p,s,x,t) = -\sum_{i=1}^{n} \frac{\partial}{\partial x^i} \big(\mu^i(x,t) f(p,s,x,t)\big) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x^i \partial x^j} \big(\sigma^{ij}(x,t) f(p,s,x,t)\big), \tag{2.76}$$

- the *Kolmogorov backward equation*

$$-\frac{\partial}{\partial s}f(p,s,x,t)=\sum_{i=1}^{n}\mu^i(p,s)\frac{\partial}{\partial p^i}f(p,s,x,t)+\frac{1}{2}\sum_{i,j=1}^{n}\sigma^{ij}(p,s)\frac{\partial^2}{\partial p^i\partial p^j}f(p,s,x,t) \tag{2.77}$$

with $(p,s),(x,t)\in\Omega\times(t_0,t_1)$ in each case.

The probability density function f as given here depends on two points in the state space (p,s) and (x,t) although either Kolmogorov equation only involves derivatives with respect to one of them (correspondingly, f needs to be of class C^2 with respect to the relevant spatial variables in Ω and of class C^1 with respect to the relevant time variable in (t_0,t_1)). In accordance with the probabilistic setting, the first order derivatives are commonly called *drift terms* (with μ^i being the *drift coefficients*), while the second order derivatives are usually named *diffusion terms* (with *diffusion coefficients* σ^{ij}).

Note that, for the forward equation, the (generic) condition $\{X(s)=p\}$ is often replaced with $X(s)\sim\rho$ for some fixed time s (usually $s=0$) and some probability distribution $\rho\in\mathfrak{P}(\mathbb{R}^n,\mathfrak{B}_n(\mathbb{R}^n))$, while simultaneously the equation is only considered in the domain $\Omega\times(s,t_1)$ with $s<t_1$. If so, we may simply write $f(x,t)$ and indicate the presence of such an initial distribution by separately stating an *initial condition*, which needs to be attained continuously by f.

Similarly, for the backward equation, we may integrate the density function f over some suitable *target set* $A\in\mathfrak{B}_n(\mathbb{R}^n)$ for some fixed time t (usually $t=0$) and only consider the differential equation in the domain $\Omega\times(t_0,t)$ with $t_0<t$. Then, we obtain a backward equation for $u(p,s):=P(X(t)\in A|X(s)=p)$, $s\in(t_0,t)$. In that case, the presence of such a target set is separately stated by a *final condition*, which again needs to be attained continuously by f; more generally, the final condition may also express $X(t)\sim\tau$ for some probability distribution $\tau\in\mathfrak{P}(\mathbb{R}^n,\mathfrak{B}_n(\mathbb{R}^n))$. Furthermore, for notational consistency, we may then also use t resp. $-t$ instead of s, whereas the difference in the spatial parameter (x vs. p) is maintained throughout the remainder to provide a formal distinction between the two different types of solution.

Technically, assessing the model in terms of such a differential equation for its probability density function rather than as a stochastic process signifies that we shift from considerations of a random evolution of the process itself to such of the *deterministic evolution* of a function which then encodes the randomness of the process. In general, this leads to important ramifications, in particular regarding the boundary behaviour and the hierarchicality of the process; this issue specifically arises within more general considerations and will be addressed in chapters 4 and 5. For the moment, we may ignore boundary concerns as we only consider the interior of the state space.

2.4.2 The Kolmogorov equations in the context of diffusion approximation

In order to formulate the Kolmogorov equations for a stochastic process obtained through diffusion approximation, under the given assumptions, one only needs to determine the drift/diffusion coefficients, i. e. the *infinitesimal increment expectations/variances* $\mu^i(x,t)$ resp. $\sigma^{ij}(x,t)$ of the process, which may be calculated – in accordance with the type of convergence – as a scaling limit of the underlying discrete process. Hence, when scaling δt with $\frac{1}{N}$ as $N \to \infty$, we get

$$\begin{aligned}\lim_{N\to\infty} N\,\mathrm{E}_{\frac{1}{N}}\big((\delta X)^\alpha \big| X(t)=x\big) &= \lim_{N\to\infty} N\,\mathrm{E}_{\frac{1}{N}}\big((\delta \hat{C}_N)^\alpha \big| \hat{C}_N(t_N)=c_N\big) \\ &= \lim_{N\to\infty} N\,\mathrm{E}_1\big((\delta C_N)^\alpha \big| C_N(Nt_N)=c_N\big)\end{aligned} \tag{2.78}$$

for all multi-indices $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ and $c_N \in \Delta_{\frac{3}{N}}$, $t_N \in \mathbb{N}_{N-1}$ for $N \in \mathbb{N}_+$ such that $c_N \to x \in \Delta_3$ and $t_N \to t \in \mathbb{R}_+$ as $N \to \infty$. However, in the further course, we will suppress the index of the conditional value c_N for notational simplicity.

Here, we may now use the results of the equations (2.53), (2.54) for the RUG model and (2.64), (2.65) for the RUZ model respectively. As the processes are time homogeneous, i. e. the transitions do not depend on the time, we omit the parameter t and also just write $\{C = c\}$ as the condition for the increment expectations.

First, we have for the (simple) increment expectations identically

$$\mu^i(x) = \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni} \big| C_N = c\big) = \lim_{N\to\infty} (-\delta^i_1 + \delta^i_2 + \delta^i_3) N R(N) D(x)$$

$$\text{for } x \in \Delta_3 \text{ and } i = 1,2,3 \quad (2.79)$$

for both models, from which it becomes clear that we have to require $R = R(N)$ with $R(N) \in \mathcal{O}(N^{-1})$ in order to assure the finiteness of the given limit (cf. section 2.3.1).

Presuming this, we obtain for the expectation of increment products

$$\begin{aligned}
\sigma^{ij}(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\delta C_{Nj}\big|C_N = c\big) \\
&= \lim_{N\to\infty} \frac{1}{2}\big(c_i\delta^i_j - c_ic_j \pm R(N)D\big(\delta^i_j \pm c_i \pm c_j \pm R(N)D\big)\big) \\
&\qquad \pm \lim_{N\to\infty} N(R(N)D)^2 \\
&= \frac{1}{2}x^i\big(\delta^i_j - x^j\big) \quad \text{for } x \in \Delta_3 \text{ and } i,j = 1,2,3
\end{aligned} \tag{2.80}$$

for the RUG model and similarly

$$\begin{aligned}
\sigma^{kl}(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Nk}\delta C_{Nl}\big|C_N = c\big) \\
&= \lim_{N\to\infty} \frac{1}{2}\Big(c_k\delta^l_k - c_kc_l + R(N)\Big(2(\mp c_k \mp c_l)D + \begin{Bmatrix} c_1c_4 \\ c_2c_3 \end{Bmatrix} + \begin{Bmatrix} c_2c_3 \\ c_1c_4 \end{Bmatrix}\Big)\Big) \\
&\qquad \mp \lim_{N\to\infty} \frac{1}{2}\big(R^2(N)(S \pm D^2) \pm 2N(R(N)D)^2\big) \\
&= \frac{1}{2}x^k\big(\delta^k_l - x^l\big) \quad \text{for } x \in \Delta_3 \text{ and } k,l = 1,2,3
\end{aligned} \tag{2.81}$$

for the RUZ model, which is the same result. Hence, in the diffusion limit, recombinational effects only survive in the (simple) increment expectations, which means in terms of the Kolmogorov equations that they only influence the drift terms, while the diffusion terms are unaffected. Note that all coefficients do not depend on the time parameter explicitly as already mentioned.

Furthermore, one may check that the expectation for higher increment products (i. e. $|\alpha| \geq 3$) all vanish with both models as required by equation (2.75): For the threefold increment expectations, we have

$$\lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\delta C_{Nj}\delta C_{Nk}\big|C_N = c\big)$$

$$= \lim_{N\to\infty} \Big(\pm N(R(N)D)^3 + \frac{1}{2}R(N)D\big(\pm c'_i c'_j(1-\delta^i_j) \pm c'_i c'_k(1-\delta^i_k) \pm c'_j c'_k(1-\delta^j_k)\big) + \frac{1}{4N}\big(2c'_i c'_j c'_k - c'_i c'_k \delta^i_j - c'_i c'_j \delta^j_k - c'_j c'_k \delta^k_i + c'_i \delta^i_j \delta^j_k\big)\Big) = 0$$

for the RUG model (cf. equation (2.55)) and similarly for the RUZ model (cf. equation (2.66))

$$\lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C'_{Nk}\delta C'_{Nl}\delta C'_{Nm}\big|C'_N = c'\big) = \lim_{N\to\infty}\Big(\pm N(R(N)D')^3 + \mathcal{O}(N^{-1})\Big) = 0.$$

The according result for the m-fold increment expectations with $m \geq 4$ may be shown analogously.

Now we may formulate the Kolmogorov equations for both the RUZ and the RUG model (which are identical – cf. also section 2.3.3) – altogether, we have:

2.9 Proposition. *The diffusion approximation of a two-loci 2-allelic recombinational Wright–Fisher model with N individuals and recombination rate $R(N) \in \mathcal{O}(N^{-1})$ may be described by the Kolmogorov forward equation for its transition probability density*[5] $f\colon (\Delta_3)_\infty \longrightarrow [0,1]$ *of its gametic configuration* $x = (x^1, \ldots, x^3) \in \Delta_3$

$$\begin{aligned}\frac{\partial}{\partial t} f(p,s,x,t) = \frac{1}{4}\sum_{i,j=1}^{3} \frac{\partial^2}{\partial x^i \partial x^j}\Big(x^i\big(\delta^i_j - x^j\big)f(p,s,x,t)\Big) \\ + \sum_{i=1}^{3}(\delta^i_1 - \delta^i_2 - \delta^i_3)\frac{\partial}{\partial x^i}\Big(\big(\lim_{N\to\infty} NR(N)D(x)\big)f(p,s,x,t)\Big) \quad (2.82)\end{aligned}$$

as well as by the Kolmogorov backward equation

$$\begin{aligned}-\frac{\partial}{\partial s} f(p,s,x,t) = \frac{1}{4}\sum_{i,j=1}^{3} p^i\big(\delta^i_j - p^j\big)\frac{\partial^2}{\partial p^i \partial p^j} f(p,s,x,t) \\ + \sum_{i=1}^{3}(-\delta^i_1 + \delta^i_2 + \delta^i_3)\Big(\lim_{N\to\infty} NR(N)D(p)\Big)\frac{\partial}{\partial p^i} f(p,s,x,t) \quad (2.83)\end{aligned}$$

with $(p,s),(x,t) \in \Omega \times \mathbb{R}$ in each case. Thus, recombination only affects the drift

[5] For the required regularity of f, cf. section 2.4.1; this also applies to further statements.

terms, but does not alter the diffusion terms. Furthermore, the given formulae do not depend on whether the RUG or the RUZ model is chosen as underlying model.

The above Kolmogorov backward equation (2.83) agrees with equation (2.41), obtained from transforming the Ohta–Kimura formula (2.1) into the 'simplex' coordinates (cf. equation (2.25)), as may be seen when suppressing the coordinates (x, t) (by stating some final condition for $t = 0$) and in turn subsequently replacing (p, s) by $(x, -t)$, yielding

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{4} \sum_{i,j=1}^{3} x^i \big(\delta^i_j - x^j\big) \frac{\partial^2}{\partial x^i \partial x^j} f(x,t) + \sum_{i=1}^{3} (-\delta^i_1 + \delta^i_2 + \delta^i_3) \Big(\lim_{N\to\infty} NR(N)D(x) \Big) \frac{\partial}{\partial x^i} f(x,t) \quad \text{with } (x,t) \in (\Delta_3)_\infty . \tag{2.84}$$

However, in equation (2.41) as well as in equation (2.1), the limit sign with $NR(N)D$ is missing.

In the following, as the RUG model and the RUZ model lead to identical Kolmogorov equations in the diffusion limit, further results will be formulated without necessarily stating which model has been used – which will usually be the RUG model as already mentioned.

2.5 Alternative coordinates

Triggered by the choice of coordinates in the Ohta–Kimura formula, we will have a closer look at the inverse of the corresponding transformation (cf. lemma 2.6), which turns out to be

$$p = x^1 + x^2, \quad q = x^1 + x^3, \quad D = x^1(1 - x^1 - x^2 - x^3) - x^2 x^3 \tag{2.85}$$

with D coinciding with the definition for $D(x)$ in equation (2.46) taking into account $x^4 \equiv 1 - x^1 - x^2 - x^3$.

In terms of the model presented and by the definition of the c_i (cf. p. 22), it is obvious that $p = x^1 + x^2$ resp. $c_1 + c_2$ prior to the diffusion limit correspond to the (total) frequency of allele A_1 appearing at the first locus, whereas $q = x^1 + x^3$ resp. $c_1 + c_3$ corresponds to the (total) frequency of the allele B_1, which appears at the second locus.

The *allele frequencies p and q*, however, do not yet yield a full set of coordinates; for instance, this could be achieved by adding one of the frequencies $x^1, \dots, x^4$ as third coordinate. Alternatively, the coefficient of linkage disequilibrium D, measuring the concurrence of the depicted alleles in the actual gamete population (cf. p. 23), is commonly employed as third coordinate. This allele-focused view yields an efficient description as the allele frequencies are not altered by recombinational effects both in the RUG and in the RUZ model (cf. for example equations (2.53) and (2.64)). In the diffusion limit, this leads to non-existing drift terms for p and q as seen in equation (2.1); only for the D-coordinate, by definition measuring the linkage disequilibrium between the two loci, a drift term is observed, which also reflects recombinational effects leading to a drift component in direction of decrementing D.

However, in the next section, we will present a refined approach towards the interpretation of the drift terms, whereas the allelic perspective will be further developed for generalised models in the next chapter.

2.6 A comparison with Brownian motion

In order to gain a deeper insight into the behaviour of the process, we wish to compare it with Brownian motion on the same domain as done in [5] for an analogous model without recombination. The fundamental idea is that, for a Riemannian manifold (M, g), the *Laplace–Beltrami operator*

$$\Delta_g := \sum_{i,j} g^{ij}(x)\Big(\frac{\partial^2}{\partial x^i \partial x^j} - \sum_k \Gamma^k_{ij}(x)\frac{\partial}{\partial x^k}\Big) \tag{2.86}$$

is invariant under all isometries of (M, g). Consequently, the probability density f of a stochastic process fulfilling the differential equation (i. e. Kolmogorov backward equation)

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{2}\Delta_g f(x,t) \tag{2.87}$$

is likewise invariant under all isometries, and the corresponding process is thus said to be *spatially homogeneous* or *optimally random* as no direction is favoured. The fundamental solution of equation (2.87) is defined to be the probability density of *(standard) Brownian motion* on (M, g), and hence we wish to compare equation (2.87) with the differential equation describing the dynamics of the considered model. This may easily be done by reformulating a generic parabolic PDE

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{2}\sum_{i,j} g^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j} f(x,t) + \sum_i \mu^i(x)\frac{\partial}{\partial x^i} f(x,t) \tag{2.88}$$

with $g^{ij}(x) \equiv \sigma^{ij}(x)$ as

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{2}\Delta_g f(x,t) + \sum_k \Big(\frac{1}{2}\sum_{i,j} g^{ij}(x)\Gamma^k_{ij}(x) + \mu^k(x)\Big)\frac{\partial}{\partial x^k} f(x,t). \tag{2.89}$$

The new additional drift terms

$$\sum_{i,j} \frac{1}{2} g^{ij}(x)\Gamma^k_{ij}(x) =: ch^k(x) \tag{2.90}$$

may be dubbed *Christoffel forces* (cf. [5]) as they originate from the geometrical properties of (M, g). For our model given in the x-coordinates (cf. equation (2.41)), we obtain

$$ch^k(x) = \frac{1}{2}\Big(x^k - \frac{1}{4}\Big) \quad \text{for } k = 1, 2, 3, \tag{2.91}$$

which means that we may interpret the process as a modification of Brownian motion on (the positive sector of) the sphere S^3 with an additional drift

$$(\mu^k + ch^k)(x) = \pm \lim_{N\to\infty} NR(N)D + \frac{1}{2}\Big(x^k - \frac{1}{4}\Big). \tag{2.92}$$

The first terms represents recombinational effects, whereas the second terms describes the influence of the geometry of the model, i. e. a drift pointing from the centroid $(\frac{1}{4}, \dots, \frac{1}{4})$ to the boundary. This may be understood biologically as follows: If a gamete type is (already) sparsely/densely represented in the population, then it is even more likely that this gamete type dies out/wins over other gamete types.

When wishing to assess the interplay between geometrical and recombinational effects, one may switch back to the (p, q, D) coordinates. Computing the Christoffel forces then yields

$$ch^p(p, q, D) = \frac{1}{2}\Big(p - \frac{1}{2}\Big) \qquad \text{and} \qquad ch^q(p, q, D) = \frac{1}{2}\Big(q - \frac{1}{2}\Big), \tag{2.93}$$

whereas for the coefficient of linkage disequilibrium, we obtain

$$ch^D(p, q, D) = -\frac{1}{2}\Big(p - \frac{1}{2}\Big)\Big(q - \frac{1}{2}\Big) + D. \tag{2.94}$$

As we have $\mu^p = \mu^q = 0$ (cf. equation (2.1)), in these coordinates there is no effect of recombination on the frequencies p and q, while we still have geometrical influences. The situation is similar to the one described before: If for example p is less/greater than $\frac{1}{2}$, meaning that the allele A_1 is sparsely/densely represented in the population, then its frequency is even more likely to decrease/increase compared with Brownian motion on $(\Omega_{(p,q,D)}, g)$ with $\frac{1}{2}g^{ij} = a^{ij}$ as in equation (2.3).

For the drift into D-direction, we have:

2.10 Lemma. *In comparison with Brownian motion on $(\Omega_{(p,q,D)}, g)$, irrespective of recombinational effects, the model features additional positive D-drift for $(p, q, D) \in \Omega_{(p,q,D)}$ close to either $(0, 1, 0)$ or $(1, 0, 0) \in \partial\Omega_{(p,q,D)}$, whereas for (p, q, D) close to $(0, 0, 0)$ or $(1, 1, 0)$, we have additional negative D-drift.*

However, in a neighbourhood of $(\frac{1}{2}, \frac{1}{2}, D)$, the D-drift is predominantly governed by the value of D, leading to additional incremental D-drift plus a decremental component caused by recombinational effects, which may flip the direction of the net drift if the recombination rate $R(N)$ is sufficiently high.

Proof. For the drift coefficient of D, we have with the Christoffel forces

$$(\mu^D + ch^D)(p, q, D) = \Big(-\lim_{N\to\infty} NR(N) + \frac{1}{2}\Big)D - \frac{1}{2}\Big(p - \frac{1}{2}\Big)\Big(q - \frac{1}{2}\Big)$$

$$= -\lim_{N\to\infty} NR(N)D + \frac{1}{2}\Big(D - \Big(p - \frac{1}{2}\Big)\Big(q - \frac{1}{2}\Big)\Big) \qquad (2.95)$$

as $\mu^D(p,q,D) = \big(-\lim_{N\to\infty} NR(N) - \frac{1}{2}\big)D$ as may be seen by transforming equation (2.83) appropriately. Ignoring recombinational effects, we now have both incremental or decremental additional D-drift – depending on $D \gtrless \big(p - \frac{1}{2}\big)\big(q - \frac{1}{2}\big)$. The term $\big(p - \frac{1}{2}\big)\big(q - \frac{1}{2}\big)$ becomes extremal at the corners of $\Omega_{(p,q,D)}$ (equalling $\pm\frac{1}{4}$ there), while D, being restricted by $\max(p+q-1,0) - pq < D < \min(p,q) - pq$ (cf. equation (2.2)), continuously vanishes there. At the centre of the p-q-plane, however, D ranges from $-\frac{1}{4}$ to $\frac{1}{4}$, while the other term is zero, thus yielding additional incremental D-drift.

The latter effect is overruled by recombinational effects as soon as we have $\lim_{N\to\infty} NR(N) > \frac{1}{2}$, then leading to additional decremental D-drift. Analogously, the first effect may also be exceeded by recombinational effects if $\lim_{N\to\infty} NR(N)D$ is sufficiently big. □

This is quite a surprising result as without recombination, there is no stable point in $\Omega_{(p,q,D)}$ with respect to the additional drift into D-direction in comparison with Brownian motion. In particular, the naive interpretation of equation (2.1) without taking into account Christoffel forces suggests an ubiquitous additional drift towards the (then stable) linkage equilibrium $\{D = 0\}$ with the recombinational effects only amplifying an inherent property of the model. This now turns out to be inappropriate as only the recombination – and this only dependent on D – provides a drift stabilising the linkage equilibrium, which is opposing the native behaviour of the model.

3 Generalisations of the recombinational Wright–Fisher model

3.1 Integration of other evolutionary mechanisms

As is well known, there are several other effects different from recombination which may influence a given population. In the following we will introduce the two most important concepts, which are natural selection and mutation: For natural selection, a comprehensive discussion is presented, whereas mutation is only outlined briefly as it leads to structurally analogous effects on the model.

3.1.1 Extension by natural selection

Being a fundamental principle in population genetics, we now wish to integrate *natural selection* into the considered model, which means that the reproductive success of a given individuals (or pairs of individuals) is determined by its *fitness value* in comparison with the average fitness value of the population.

In the simplest case, in a population of $N \in \mathbb{N}_+$ individuals with the individuals $i = 1, \dots, N$ present at a frequencies $c = (c_i)$, every individual i is assigned a fitness value $w_i \in \mathbb{R}$. Then the average fitness value of the population is given by $\bar{w} = \sum_{i=1}^{N} w_i c_i$, and the individual frequencies after selection evaluate to

$$c_i' = \frac{w_i}{\bar{w}} c_i. \tag{3.1}$$

However, it is more plausible to assign fitness values rather to pairs of individuals than to single ones as the reproductive success is likely to depend on the concurrence of the parents. Hence, one may assign fitness values $w_{\substack{i\\j}}$ to a pair of individuals $\substack{i\\j}$.

For our model, a simple adaption of this idea would be to assign fitness values to pairs of gametes (in the sense of the RUG model), i. e.

$$
\begin{array}{c|cccc}
w_{\substack{i\\j}} & A_1B_1 & A_1B_2 & A_2B_1 & A_2B_2 \\ \hline
A_1B_1 & w_{\substack{1\\1}} & w_{\substack{1\\2}} & w_{\substack{1\\3}} & w_{\substack{1\\4}} \\
A_1B_2 & w_{\substack{2\\1}} & w_{\substack{2\\2}} & w_{\substack{2\\3}} & w_{\substack{2\\4}} \\
A_2B_1 & w_{\substack{3\\1}} & w_{\substack{3\\2}} & w_{\substack{3\\3}} & w_{\substack{3\\4}} \\
A_2B_2 & w_{\substack{4\\1}} & w_{\substack{4\\2}} & w_{\substack{4\\3}} & w_{\substack{4\\4}}.
\end{array}
\tag{3.2}
$$

Naturally, we assume $w_{\substack{i\\j}} = w_{\substack{j\\i}}$, and consequently we may also view (3.2) as fitness table for single zygotes (with symmetric identification), which would be applied for the RUZ model correspondingly. If it is additionally stipulated that the fitness of $\substack{A_1B_2\\A_2B_1}$ agrees with the one of $\substack{A_1B_1\\A_2B_2}$ (hence $w_{\substack{2\\3}} = w_{\substack{1\\4}}$), then (3.2) can be shortened into a table which gives the fitness in terms of the genetic configuration at each locus of a zygote, i. e.

$$
\begin{array}{c|ccc}
w_{\substack{i\\j}} & B_1B_1 & B_1B_2 & B_2B_2 \\ \hline
A_1A_1 & w_{\substack{1\\1}} & w_{\substack{1\\2}} & w_{\substack{2\\2}} \\
A_1A_2 & w_{\substack{1\\3}} & w_{\substack{1\\4}} & w_{\substack{2\\4}} \\
A_2A_2 & w_{\substack{3\\3}} & w_{\substack{3\\4}} & w_{\substack{4\\4}}.
\end{array}
\tag{3.3}
$$

3.1.2 An additive fitness scheme and its application

A simple scheme to fill table (3.3) is given by

$$
\begin{array}{c|ccc}
w & B_1B_1 & B_1B_2 & B_2B_2 \\ \hline
A_1A_1 & 1 & 1+\frac{s_2}{2} & 1+s_2 \\
A_1A_2 & 1+\frac{s_1}{2} & 1+\frac{s_1}{2}+\frac{s_2}{2} & 1+\frac{s_1}{2}+s_2 \\
A_2A_2 & 1+s_1 & 1+s_1+\frac{s_2}{2} & 1+s_1+s_2,
\end{array}
\tag{3.4}
$$

signifying that

- for the first locus A_2 is fitter than A_1 by $s_1 \in \mathbb{R}$,
- for the second locus B_2 is fitter than B_1 by $s_2 \in \mathbb{R}$,
- the effects of both loci add.

Due to the last point, such a scheme is called *additive*. The fitness values for heterozygotes is calculated by *averaging* the values of the corresponding homozygotes.

One may show that the diffusion approximation of the recombinational Wright–Fisher RUG/RUZ model presented in section 2.3 now extended by selection in accordance with the fitness scheme (3.4) – when requiring that s_1, s_2 depend on N with $s_i \in \mathcal{O}(N^{-1})$, $i = 1, 2$, thus $\lim_{N\to\infty} Ns_i(N) =: \sigma_i \in \mathbb{R}$ – yields drift coefficients (cf. section 2.4.2)

$$\mu^1(x) = -\frac{\sigma_1}{2}x^1(x^3 + x^4) - \frac{\sigma_2}{2}x^1(x^3 + x^4) - \lim_{N\to\infty} NR(N)D(x), \tag{3.5}$$

$$\mu^2(x) = -\frac{\sigma_1}{2}x^2(x^3 + x^4) + \frac{\sigma_2}{2}x^2(1 - x^2 - x^4) + \lim_{N\to\infty} NR(N)D(x), \tag{3.6}$$

$$\mu^3(x) = +\frac{\sigma_1}{2}x^3(1 - x^3 - x^4) - \frac{\sigma_2}{2}x^3(x^2 + x^4) + \lim_{N\to\infty} NR(N)D(x) \tag{3.7}$$

with $x^4 = 1 - x^1 - x^2 - x^3$ in the x-coordinates for either model, while all diffusion coefficients remain unchanged. Thus analogous to recombination, selection in the diffusion limit also only leads to (deterministic) drift: Carriers of the alleles which are assumed to be fitter get a positive drift component, whereas carriers of the opposing, less fit alleles receive a negative drift component.

In the (p, q, D)-coordinates, one obtains

$$\mu^p(p, q, D) = -\frac{\sigma_1}{2}p(1 - p) - \frac{\sigma_2}{2}D, \tag{3.8}$$

$$\mu^q(p, q, D) = -\frac{\sigma_1}{2}D - \frac{\sigma_2}{2}q(1 - q), \tag{3.9}$$

$$\mu^D(p, q, D) = -\frac{\sigma_1}{2}D(1 - 2p) - \frac{\sigma_2}{2}D(1 - 2q) - \frac{1}{2}D(1 + 2\lim_{N\to\infty} NR(N)), \tag{3.10}$$

for either model with the diffusion coefficients remaining unaffected again. This agrees with the result stated in [25].

3.1.3 Multiplicative fitness schemes

An additive fitness scheme as given in equation (3.4) is somewhat unsatisfactory as in particular the combination of alleles may be relevant for the fitness value and not the individual contribution of each allele. A more general fitness scheme is thus given by filling table (3.3) via

$$\begin{array}{c|ccc} w & B_1B_1 & B_1B_2 & B_2B_2 \\ \hline A_1A_1 & 1+2a & 1+a+b & 1+2b \\ A_1A_2 & 1+a+c & * & 1+b+d \\ A_2A_2 & 1+2c & 1+c+d & 1+2d \end{array} \tag{3.11}$$

with $a, b, c, d \in \mathbb{R}$, which corresponds to assigning every combination of homozygotes at the two loci an independent value. The fitness of heterozygotes is still calculated by averaging; at the position in the table marked with an asterisk, we do not have a unique value, which is due to the fact that here we do not necessarily have $w_{\frac{2}{3}} = w_{\frac{1}{4}}$. This becomes evident when translating (3.11) into the fitness table for gamete pairs

$$\begin{array}{c|cccc} w_{\frac{i}{j}} & A_1B_1 & A_1B_2 & A_2B_1 & A_2B_2 \\ \hline A_1B_1 & 1+2a & 1+a+b & 1+a+c & 1+a+d \\ A_1B_2 & 1+b+a & 1+2b & 1+b+c & 1+b+d \\ A_2B_1 & 1+c+a & 1+c+b & 1+2c & 1+c+d \\ A_2B_2 & 1+d+a & 1+d+b & 1+d+c & 1+2d; \end{array} \tag{3.12}$$

thus, for $*$ we have either $1 + a + d$ or $1 + b + c$.

This fitness scheme is also obtained when assigning fitness values w_i to the gametes by

$$\begin{array}{c|c} \text{gam.} & w_i \\ \hline A_1B_1 & 1+a \\ A_1B_2 & 1+b \\ A_2B_1 & 1+c \\ A_2B_2 & 1+d \end{array} \tag{3.13}$$

and subsequently *multiplying* the fitness values of two gametes to get the fitness of them as a pair, e. g.

$$w_{\frac{1}{2}} := w_1 w_2 = 1 + a + b + ab. \tag{3.14}$$

The herein occurring products of the form ab are assumed to be negligible (and hence do not appear in table (3.12)), which is the case anyway in the diffusion limit, where we have $a, \ldots, d \in \mathcal{O}(N^{-1})$ (cf. below).

Note that the presented scheme is neither limited to a certain number of alleles nor to a certain number of loci as it is essentially independent of the coaction of loci and alleles and only takes gamete types resp. pairs of gamete types into account. Having said this, any number of gametes may be assigned fitness values as in table (3.13), yielding fitness values for pairs of gametes resp. zygotes analogous to table (3.11). This is exploited for generalisations of the current situation in sections 3.2 and 3.3.

3.1.4 A recombinational Wright–Fisher model with a multiplicative fitness scheme

We consider again the RUG and the RUZ model as presented in section 2.3 and enhance them by selection in accordance with the fitness scheme (3.12). For notational simplicity, we change the designation of the fitness values from $1+a, \ldots, 1+d$ to $1+a_1, \ldots, 1+a_4$ and assume that $a_1, \ldots, a_4$ depend on N with $a_i \in \mathcal{O}(N^{-1})$, $i = 1, \ldots, 4$, thus with $\lim_{N\to\infty} N a_i(N) =: \alpha_i \in \mathbb{R}$, $i = 1, \ldots, 4$.

In the RUG model, the selection step is employed between mating and recombination (cf. fig. 3.1). Consequently, we obtain for the gamete frequencies in the pool (cf. equation (2.42)):

$$
\begin{aligned}
c_i' =& \sum_{j=1}^{4} \frac{w_{\frac{i}{j}}}{\bar{w}} c_i c_j \pm R\Big(\frac{w_{\frac{1}{4}}}{\bar{w}} c_1 c_4 - \frac{w_{\frac{2}{3}}}{\bar{w}} c_2 c_3\Big) \\
=& \frac{1}{\bar{w}} \Big(\sum_{j=1}^{4} (1 + a_i + a_j) c_i c_j \pm R\big((1 + a_1 + a_4) c_1 c_4 - (1 + a_2 + a_3) c_2 c_3\big)\Big) \\
=& \frac{1}{\bar{w}} \big(c_i + a_i c_i + \bar{a} c_i \pm R(D + (a_1 + a_4) c_1 c_4 - (a_2 + a_3) c_2 c_3)\big)
\end{aligned}
\tag{3.15}
$$

with $\bar{a} := \sum_{j=1}^{4} a_j c_j$ and

$$
\bar{w} = \sum_{i,j} w_{\frac{i}{j}} c_i c_j = \sum_{i,j=1}^{4} (1 + a_i + a_j) c_i c_j = 1 + 2\bar{a}, \tag{3.16}
$$

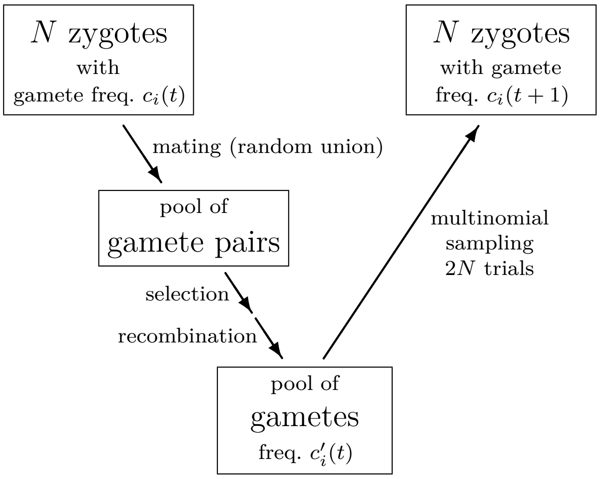

Figure 3.1: The RUG model with selection

being the average fitness of the population.

In the diffusion limit $N \to \infty$ (cf. section 2.4), we have $\bar{a} \to 0$ and thus $\bar{w} \to 1$, while we have $Ra_i \in \mathcal{O}(N^{-2})$ for all i. Consequently, we obtain for the drift coefficients (cf. equations (2.79))

$$\begin{aligned}
\mu^i(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\big|C_N = c\big) = \lim_{N\to\infty} N(c'_i - c_i) \\
&= \lim_{N\to\infty} \frac{N}{\bar{w}}\big(c_i + a_i(N)c_i + \bar{a}(N)c_i - (1+2\bar{a}(N))c_i \\
&\qquad\qquad \pm R(N)(D(c) + (a_1(N)+a_4(N))c_1c_4 - (a_2(N)+a_3(N))c_2c_3)\big) \\
&= (\alpha_i - \bar{\alpha})x^i \pm \lim_{N\to\infty} NR(N)D(x) \qquad (3.17)
\end{aligned}$$

with $\bar{\alpha} := \lim_{N\to\infty} \sum_{j=1}^4 Na_j(N)c_j = \sum_{j=1}^4 \alpha_j c_j$, whereas the diffusion coefficients are (cf. equation (2.80))

$$\begin{aligned}
\sigma^{ij}(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\delta C_{Nj}\big|C_N = c\big) \\
&= \lim_{N\to\infty} \frac{1}{2}(c'_i\delta^i_j - c'_ic'_j) + N(c'_i - c_i)(c'_j - c_j) \\
&= \frac{1}{2}(x^i\delta^i_j - x^ix^j). \qquad (3.18)
\end{aligned}$$

Again, one may check that the expectations for higher increment products all vanish as described in section 2.4.2, which is crucial for the validity of the diffusion approximation.

We thus have in generalisation of proposition 2.9 (cf. also equation (2.84)):

3.1 Lemma. *The diffusion approximation of a two-loci two-allelic recombinational Wright–Fisher model with recombination rate R encompassing a multiplicative fitness scheme $\alpha \in \mathbb{R}^4$ may be described by the Kolmogorov backward equation for its transition probability density $f\colon (\Delta_3)_\infty \longrightarrow [0,1]$ of its gametic configuration $x = (x^1, x^2, x^3) \in \Delta_3$*

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{4} \sum_{i,j=1}^{3} x^i \big(\delta^i_j - x^j\big) \frac{\partial^2}{\partial x^i \partial x^j} f(x,t) + \sum_{i=1}^{3} \big((\alpha_i - \bar{\alpha}) x^i \pm \lim_{N\to\infty} N R(N) D(x)\big) \frac{\partial}{\partial x^i} f(x,t) \quad (3.19)$$

with $(x,t) \in (\Delta_3)_\infty$.

Analogous to recombinational effects, selection effects only affect the drift terms.

In comparison with an additive fitness scheme, the effect this time is even more transparent: The additional limit fitness α_i of a gamete i is compared with the average additional limit fitness of the population $\bar{\alpha}$ and leads to positive or negative drift if it is higher resp. lower.

The same statement as in lemma 3.1 may also be deduced by using the RUZ model extended by selection. However, in the RUZ model, the selection step is employed at a different stage than in the RUG model, i. e. applies to the zygotes before they produce a new gamete under possible recombination (cf. 3.2). Calculating the expectation of the gamete frequencies in the next generation $C_k(t)$ leads – when including selection – to the same result as in the RUG model (cf. equations (2.59) and (2.62)):

$$\begin{aligned} \mathrm{E}_1(\delta C_k | C = c) &= \frac{1}{\bar{w}} \sum_{i,j} (1 + a_i + a_j) c_i c_j \rho^{\substack{i\\j}}_k - c_k \\ &= \frac{1}{\bar{w}} \big((a_k - \bar{a}) c_k \pm R(D + (a_1 + a_4) c_1 c_4 - (a_2 + a_3) c_2 c_3)\big) \end{aligned} \quad (3.20)$$

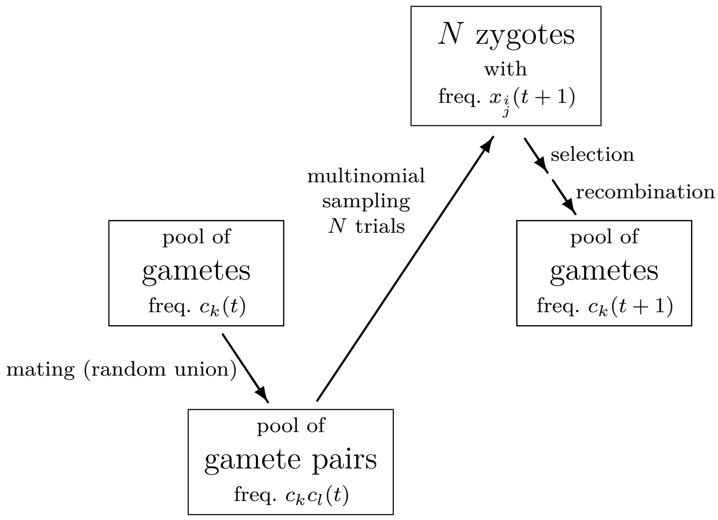

Figure 3.2: The RUZ model with selection

For the expectations of the increment products, we get again a different result, but in the diffusion limit, the value coincides with the one obtained for the RUG model as (cf. equation (2.65))

$$\begin{aligned}\mathrm{E}_1(\delta C_k \delta C_l | C = c) = &-\frac{1}{N\bar{w}^2}\sum_{i,j}(1+a_i+a_j)c_ic_j\rho_k^{\substack{i\\j}}\sum_{m,n}(1+a_m+a_n)c_mc_n\rho_k^{\substack{m\\n}} \\ &+\frac{1}{N\bar{w}}\sum_{i,j}(1+a_i+a_j)c_ic_j\rho_k^{\substack{i\\j}}\rho_k^{\substack{i\\j}} + \mathrm{E}_1(\delta C_k)\,\mathrm{E}_1(\delta C_l) \qquad (3.21)\end{aligned}$$

and hence (cf. equation (2.81))

$$\begin{aligned}\sigma^{kl}(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Nk}\delta C_{Nl}\big|C_N = c\big) \\ &= \lim_{N\to\infty}\Big(-\sum_{i,j}c_ic_j\rho_k^{\substack{i\\j}}\sum_{m,n}c_mc_n\rho_k^{\substack{m\\n}} + \sum_{i,j}c_ic_j\rho_k^{\substack{i\\j}}\rho_k^{\substack{i\\j}}\Big) \\ &= \frac{1}{2}x_k\big(\delta_l^k - x_l\big). \qquad (3.22)\end{aligned}$$

Furthermore, all expectations of the higher increment products also vanish for the RUZ model.

Lastly, it is noted that the assertion of lemma 3.1 is consistent with the results for the additive fitness scheme from the preceding section: By putting $a := 0$, $b := \frac{s_2}{2}$, $c := \frac{s_1}{2}$, $d := \frac{s_1+s_2}{2}$, implying $\alpha = 0$, $\beta = \frac{\sigma_2}{2}$, $\gamma = \frac{\sigma_1}{2}$, $\delta = \frac{\sigma_1+\sigma_2}{2}$, we again obtain drift coefficients as stated in equation (3.5).

3.1.5 Extension by mutation

Lastly, the integration of the evolutionary mechanism of mutation is sketched: In contrast to the generally directional effect of natural selection, *mutation* represents a variational force to the genetic configuration (cf. [10], pp. 11 f.). This is usually modelled via spontaneous shifts in the population which randomly occur at a certain stage within the reproduction process; here we assume this to take place right at the beginning of the reproduction cycle (cf. also below).

In order to obtain a quantitative description, analogous to natural selection, an appropriate *mutation scheme* needs to be selected: For example, it may be stipulated that only certain shifts and only into certain directions may occur. However, in the most general setting, any gamete i may mutate into any other gamete j and vice versa. Each event is then assigned a corresponding mutation rate $b_{i,j} \in [0,1]$, $i \neq j$, e. g. in the current setting, gamete A_1B_1 may mutate into A_2B_2 at rate $b_{1,4}$. One may also consider a mutation scheme at the level of alleles wherein then every shift of alleles is assigned a certain rate. Altogether, by multiplication this again amounts to a corresponding mutation rate for each gamete mutation. For this reason, in the following the gamete mutation scheme is employed as this imposes less restrictions and in conjunction with the RUG model is the most handy. However, one might also use the RUZ model and a corresponding mutation scheme at the level of zygotes and show similarly as above that they lead to equivalent results.

To begin with, we will equip the 2-loci 2-allelic recombinational model as presented in chapter 2 with the mutation scheme $b_{i,j}$, $i, j = 1, \ldots, 4$, $i \neq j$. To assess the diffusion limit of the model, we again need to assume that the mutation scheme depends on the population size N such that $b_{i,j} \in \mathcal{O}(N^{-1})$ and $\lim_{N\to\infty} N b_{i,j}(N) =: \beta_{i,j} \in \mathbb{R}$ for $i, j = 1, \ldots, 4$, $i \neq j$. If mutation – as already hinted – is assumed to occur on the gametes of a given population before the mating step (in the sense of the RUG model), i. e. directly after sampling gametes from the given zygote population as depicted in figure 3.3, we obtain for the gamete frequencies c_i^m after the mutation

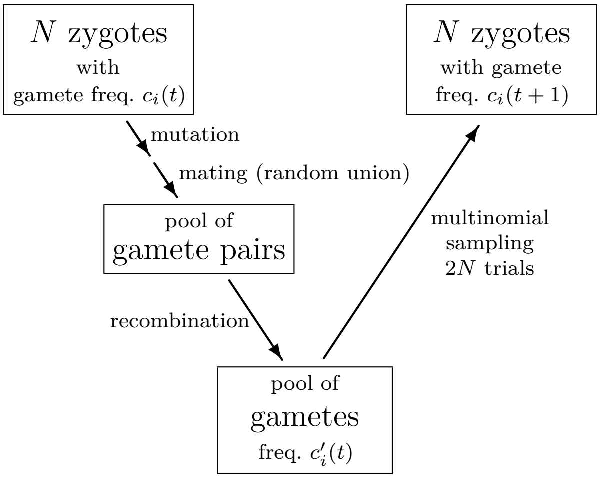

Figure 3.3: The RUG model with mutation

step

$$c_i^m = \Big(1 - \sum_{\substack{j=1 \\ j\neq i}}^{4} b_{i,j}\Big) c_i + \sum_{\substack{j=1 \\ j\neq i}}^{4} b_{j,i} c_j \tag{3.23}$$

and for the gamete frequencies in the pool (cf. equation (2.42))

$$c'_i = c_i^m \pm R\big(c_1^m c_4^m - c_2^m c_3^m\big). \tag{3.24}$$

In the diffusion limit $N \to \infty$ (cf. section 2.4), we consequently obtain for the drift coefficients (cf. equations (2.79))

$$\begin{aligned}
\mu^i(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni} \big| C_N = c\big) = \lim_{N\to\infty} N(c'_i - c_i) \\
&= \lim_{N\to\infty} N\Big(- \sum_{j\neq i} b_{i,j} c_i + \sum_{j\neq i} b_{j,i} c_j\Big) \pm \lim_{N\to\infty} NR\Bigg(c_1 c_4 - c_2 c_3 \\
&\quad + c_1\Big(- \sum_{j=1}^{3} b_{4,j} c_4 + \sum_{j=1}^{3} b_{j,4} c_j\Big) + c_4\Big(- \sum_{j=2}^{4} b_{1,j} c_1 + \sum_{j=2}^{4} b_{j,1} c_j\Big)
\end{aligned}$$

$$
\begin{aligned}
&+\Big(-\sum_{j=1}^{3} b_{4,j}c_4+\sum_{j=1}^{3} b_{j,4}c_j\Big)\Big(-\sum_{j=2}^{4} b_{1,j}c_1+\sum_{j=2}^{4} b_{j,1}c_j\Big)\\
&-c_2\Big(-\sum_{j\neq 3} b_{3,j}c_3+\sum_{j\neq 3} b_{j,3}c_j\Big)-c_3\Big(-\sum_{j\neq 2} b_{2,j}c_2+\sum_{j\neq 2} b_{j,2}c_j\Big)\\
&-\Big(-\sum_{j\neq 3} b_{3,j}c_3+\sum_{j\neq 3} b_{j,3}c_j\Big)\Big(-\sum_{j\neq 2} b_{2,j}c_2+\sum_{j\neq 2} b_{j,2}c_j\Big)\Big)\\
=&-\sum_{j\neq i}\beta_{i,j}x^i+\sum_{j\neq i}\beta_{j,i}x^j\pm\lim_{N\to\infty}NR(N)D(x) \qquad (3.25)
\end{aligned}
$$

as we particularly have $Rb_{i,j}\in\mathcal{O}(N^{-2})$. Hence, for the diffusion coefficients, we obtain as previously (cf. equation (2.80))

$$
\begin{aligned}
\sigma^{ij}(x)&=\lim_{N\to\infty}N\,\mathrm{E}_1\big(\delta C_{Ni}\delta C_{Nj}\big|C_N=c\big)\\
&=\lim_{N\to\infty}\frac{1}{2}(c_i'\delta_j^i-c_i'c_j')+N(c_i'-c_i)(c_j'-c_j)\\
&=\frac{1}{2}(x^i\delta_j^i-x^ix^j). \qquad (3.26)
\end{aligned}
$$

Again, one may check that the expectation for the higher increment products all vanish as described in section 2.4.2, which is crucial for the validity of the diffusion approximation.

We thus have in generalisation of proposition 2.9 (cf. also equation (2.84)):

3.2 Lemma. *The diffusion approximation of a two-loci two-allelic recombinational Wright–Fisher model with recombination rate $R(N)\in\mathcal{O}(N^{-1})$ encompassing a mutation scheme $\beta\in\mathbb{R}^{12}$ may be described by the Kolmogorov backward equation for its transition probability density $f\colon(\Delta_3)_\infty\longrightarrow[0,1]$ of its gametic configuration $x=(x^1,x^2,x^3)\in\Delta_3$ being*

$$
\begin{aligned}
\frac{\partial}{\partial t}f(x,t)=&\frac{1}{4}\sum_{i,j=1}^{3}x^i\big(\delta_j^i-x^j\big)\frac{\partial^2}{\partial x^i\partial x^j}f(x,t)\\
&+\sum_{i=1}^{3}\Big(-\sum_{\substack{j=1\\ j\neq i}}^{4}\beta_{i,j}x^i+\sum_{\substack{j=1\\ j\neq i}}^{4}\beta_{j,i}x^j\pm\lim_{N\to\infty}NR(N)D(x)\Big)\frac{\partial}{\partial x^i}f(x,t) \qquad (3.27)
\end{aligned}
$$

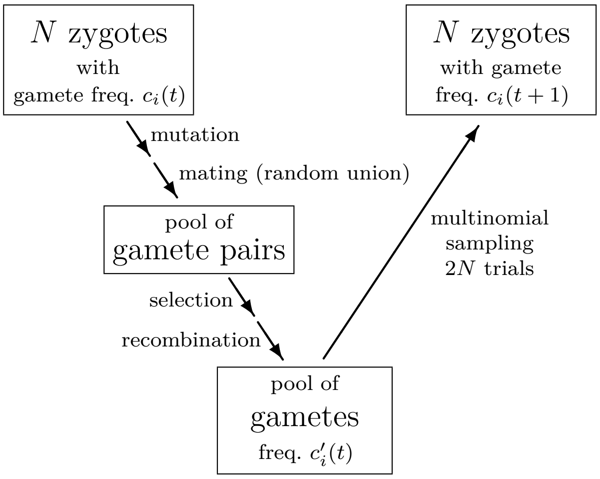

Figure 3.4: The RUG model with mutation and selection

with $(x,t) \in (\Delta_3)_\infty$ *and* $x^4 = 1 - \sum_1^3 x^i$.

Analogous to recombination and selection, mutation effects only affect the drift terms.

3.1.6 A recombinational Wright–Fisher model with mutation and selection

Furthermore, the model presented may be equipped with both mutation and natural selection (cf. section 3.1.1), in which the multiplicative fitness scheme α_i as introduced in section 3.1.4 is applied for selection. Again, we need to specify at which stage of the model both mechanism are incorporated: At first, mutation is assumed to appear after the mating step (as previously), whereas natural selection is performed on the newly formed gamete pairs (cf. figure 3.4). Consequently, we obtain

$$
\begin{aligned}
c_i' =& \sum_{j=1}^{4} \frac{w_{ij}}{\bar{w}} c_i^m c_j^m \pm R\Big(\frac{w_{14}}{\bar{w}} c_1^m c_4^m - \frac{w_{23}}{\bar{w}} c_2^m c_3^m\Big) \\
=& \frac{1}{\bar{w}}\Big(\sum_{j=1}^{4}(1 + a_i + a_j)\Big(- \sum_{\substack{l=1\\ l\neq i}}^{4} b_{i,l} c_i + \sum_{\substack{l=1\\ l\neq i}}^{4} b_{l,i} c_l\Big)\Big(- \sum_{\substack{l=1\\ l\neq j}}^{4} b_{j,l} c_j + \sum_{\substack{l=1\\ l\neq j}}^{4} b_{l,j} c_l\Big)
\end{aligned}
$$

$$\pm R\big(c_1c_4 - c_2c_3 + \dots\big)\Big) \tag{3.28}$$

for the gamete frequencies in the pool (cf. equation (2.42)) with c_i^m being the gamete frequencies after the mutation step (cf. equation (3.23)).

Alternatively, when interchanging the mutation and selection step by moving selection ahead of mutation, we obtain for the gamete frequencies in the pool the diverging expression

$$c_i' = \Big(1 - \sum_{\substack{j=1\\ j\neq i}}^{4} b_{i,j}\Big)c_i^s + \sum_{\substack{j=1\\ j\neq i}}^{4} b_{j,i}c_j^s \pm R\bigg(\Big(1 - \sum_{j=2}^{4} b_{1,j}\Big)c_1^s + \sum_{j=2}^{4} b_{j,1}c_j^s \dots\bigg) \tag{3.29}$$

with $c_i^s = \frac{1+a_i}{\bar{w}}c_i$ being the gamete frequencies after the selection step with here $\bar{w} := 1 + \sum_{j=1}^{4} a_j c_j$. In this setting, natural selection needs to be applied at the level of single gametes (instead of pairs of gametes), which still provides an equivalent result as described in section 3.1.3 since the applied selection scheme effectively is a haploid one (cf. also table (3.13)).

However, the chosen order of the two evolutionary mechanisms does not influence the final result as in the diffusion limit $N \to \infty$ (cf. section 2.4), we accordingly obtain for both variants as the drift coefficients (cf. equations (2.79))

$$\begin{aligned}\mu^i(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\big|C_N = c\big) = \lim_{N\to\infty} N(c_i' - c_i)\\ &= \Big(\alpha_i - \bar{\alpha} - \sum_{\substack{j=1\\ j\neq i}}^{4} \beta_{i,j}\Big)x^i + \sum_{\substack{j=1\\ j\neq i}}^{4} \beta_{j,i}x^j \pm \lim_{N\to\infty} NR(N)D(x)\end{aligned} \tag{3.30}$$

with $\bar{\alpha} := \lim_{N\to\infty} \sum_{j=1}^{4} Na_j(N)c_j = \sum_{j=1}^{4} \alpha_j c_j$, whereas the diffusion coefficients are (cf. equation (2.80))

$$\begin{aligned}\sigma^{ij}(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\delta C_{Nj}\big|C_N = c\big)\\ &= \lim_{N\to\infty} \frac{1}{2}(c_i'\delta_j^i - c_i'c_j') + N(c_i' - c_i)(c_j' - c_j)\\ &= \frac{1}{2}(x^i\delta_j^i - x^ix^j).\end{aligned} \tag{3.31}$$

We thus have in further generalisation of proposition 2.9 (cf. also equation (2.84)):

3.3 Lemma. *The diffusion approximation of a two-loci two-allelic recombinational Wright–Fisher model with recombination rate R encompassing a multiplicative fitness scheme $\alpha \in \mathbb{R}^4$ and a mutation scheme $\beta \in \mathbb{R}^{12}$ may be described by the Kolmogorov backward equation for its transition probability density $f\colon (\Delta_3)_\infty \longrightarrow [0,1]$ of its gametic configuration $x = (x^1, x^2, x^3) \in \Delta_3$ being*

$$\begin{aligned}\frac{\partial}{\partial t} f(x,t) &= \frac{1}{4}\sum_{i,j=1}^{3} x^i\big(\delta^i_j - x^j\big)\frac{\partial^2}{\partial x^i \partial x^j} f(x,t) \\ &+ \sum_{i=1}^{3}\Bigg(\Big(\alpha_i - \bar{\alpha} - \sum_{\substack{j=1\\ j\neq i}}^{4}\beta_{i,j}\Big)x^i + \sum_{\substack{j=1\\ j\neq i}}^{4}\beta_{j,i}x^j \pm \lim_{N\to\infty} NR(N)D(x)\Bigg)\frac{\partial}{\partial x^i} f(x,t) \qquad (3.32)\end{aligned}$$

with $(x,t) \in (\Delta_3)_\infty$ and $x^4 = 1 - \sum_1^3 x^i$.

Having stated the above result, we may conclude our discussion of mutation at this point and only carry natural selection onwards to our further considerations. However, mutational effects generalise as straightforwardly as selectional effects, and furthermore, the following results may analogously be completed by mutation if necessary.

3.2 Two-loci multi-allelic models

The generalisation of the models considered in section 2.3 to arbitrarily many alleles and giving its diffusion approximation is quite straightforward. However, in the remainder we will only consider the RUG model as it is the more transparent one (cf. also p. 33).

3.2.1 Quantitative analysis

To adapt the notation to the modified situation, we write ij for a gamete containing alleles i and j (and hence c_{ij} for the corresponding frequency) with i and j now running from 1 to some arbitrary $n \geq 2$ which denotes the number of alleles admitted.

Correspondingly, we now have an (n^2-1)-dimensional state space (by dropping $c_{nn} = 1 - \sum_{(i,j)\neq(n,n)} c_{ij}$) which is given by (cf. equation (2.72))

$$\Delta_{\frac{n^2-1}{N}} := \left\{ (c_{11}, \ldots, c_{1n}, c_{21}, \ldots, c_{n1}, \ldots, c_{nn-1}) \in \mathbb{N}^{n^2-1}_{N^{-1}} \Big| \sum_{(i,j)\neq(n,n)} c_{ij} < 1 \right\} \quad (3.33)$$

for the discrete approach resp. its limit space for means of diffusion approximation being

$$\Delta_{n^2-1} = \left\{ (x^{11}, \ldots, x^{1n}, x^{21}, \ldots, x^{n1}, \ldots, x^{nn-1}) \in \mathbb{R}^{n^2-1} \big| x^{ij} > 0; \sum_{(i,j)\neq(n,n)} x^{ij} < 1 \right\}. \quad (3.34)$$

For a quantitative treatment, we calculate again the frequencies c'_{ij} in the gamete pool after mating and recombination with the gamete frequencies of the current population c_{ij} given; selection is excluded at the moment. A gamete ij is produced if

- the gamete ij is sampled twice with relative frequency 1,
- gametes ij and il or kj are sampled with relative frequency $\frac{1}{2}$ (plus a factor 2 by symmetry),
- gametes ij and kl are sampled with relative frequency $\frac{1}{2}(1-R)$ (plus a factor 2 by symmetry),
- gametes il and kj are sampled with relative frequency $\frac{1}{2}R$ (plus a factor 2 by symmetry)

with $i, j, k, l \in \{1, \ldots, n\}$ and $k \neq i$, $l \neq j$. Consequently, we obtain

$$c'_{ij} = \sum_{k,l=1}^{n} c_{ij}c_{kl} + R\Big(\sum_{\substack{k\neq i \\ l\neq j}} (c_{il}c_{kj} - c_{ij}c_{kl}) \Big). \quad (3.35)$$

Putting

$$D_{ij}(c) := \sum_{k\neq i, l\neq j} (c_{il}c_{kj} - c_{ij}c_{kl}) \quad (3.36)$$

and using the frequency property, we get

$$c'_{ij} = c_{ij} + RD_{ij}(c), \tag{3.37}$$

which is analogous to the formula obtained in the 2-allelic case (cf. equation (2.42)). However, now we have n^2 *coefficients of linkage disequilibrium* D_{ij} dependent on the chosen gamete type, whereas in the 2-allelic case only one coefficient D suffices for gamete frequencies (with different sign, cf. also section 3.5). This proves to be a special case of the general situation as for $n = 2$ we have

$$D_{11} = D_{22} = -D_{12} = -D_{21} = c_{12}c_{21} - c_{11}c_{22} \equiv c_2c_3 - c_1c_4 = -D. \tag{3.38}$$

Correspondingly, while in the 2-allelic case the linkage equilibrium is equivalent to $D \equiv 0$, here the *(full) linkage equilibrium* between the two loci considered is assumed to be only achieved if $D_{ij} \equiv 0$ for all $i, j = 1, \dots, n$.

In the diffusion limit, where we may also write $D^{ij}(x)$ with upper indices for notational consistency, we obtain consequently as drift coefficients

$$\mu^{ij}(x) = \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Nij}\big|C_N = c\big) = \lim_{N\to\infty} NR(N)D^{ij}(x), \tag{3.39}$$

whereas the diffusion coefficients again show no effect of recombination, thus

$$\begin{aligned}
\sigma^{ij,kl}(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Nij}\delta C_{Nkl}\big|C_N = c\big) \\
&= \lim_{N\to\infty} \frac{1}{2}\big((c_{ij} + R(N)D_{ij})\delta^{ij}_{kl} - (c_{ij} + R(N)D_{ij})(c_{kl} + R(N)D_{kl})\big) \\
&\quad + \lim_{N\to\infty} NR^2(N)D_{ij}D_{kl} \\
&= \frac{1}{2}x^{ij}(\delta^{ij}_{kl} - x^{kl}).
\end{aligned} \tag{3.40}$$

With the results of the preceding section for fitness schemes generalising straightforwardly, thus $\alpha_{ij} \in \mathbb{R}$ being the additional limit fitness of allele ij and $\bar{\alpha} = \sum_{i,j} \alpha_{ij}c_{ij}$ being the average additional limit fitness of the population (cf. pp. 49 f.), we altogether have:

3.4 Lemma. *For $n \geq 2$, the diffusion approximation of a two-loci n-allelic recombina-*

tional Wright–Fisher model with recombination rate R encompassing a multiplicative fitness scheme $\alpha = (\alpha_{ij}) \in \mathbb{R}^{n^2}$ may be described by the Kolmogorov backward equation for its transition probability density $f\colon (\Delta_{n^2-1})_\infty \longrightarrow [0,1]$ of its gametic configuration $x = (x^{11}, \dots, x^{1n}, x^{21}, \dots, x^{n1}, \dots, x^{nn-1}) \in \Delta_{n^2-1}$ being

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{4} \sum_{\substack{(i,j)\neq(n,n)\\(k,l)\neq(n,n)}} x^{ij}\big(\delta^{ij}_{kl} - x^{kl}\big) \frac{\partial^2}{\partial x^{ij} \partial x^{kl}} f(x,t) + \sum_{(i,j)\neq(n,n)} \big((\alpha_{ij} - \bar{\alpha})x^{ij} \pm \lim_{N\to\infty} NR(N) D^{ij}(x)\big) \frac{\partial}{\partial x^{ij}} f(x,t) \quad (3.41)$$

with $(x,t) \in (\Delta_{n^2-1})_\infty$.

3.2.2 Alternative coordinates

The formula for the coefficients of linkage disequilibrium D_{ij} may still be simplified, i. e.

$$\begin{aligned} D_{ij}(c) &= \sum_{l\neq j} c_{il} \sum_{k\neq i} c_{kj} - c_{ij}\Big(1 - \sum_{l\neq j} c_{il} - \sum_{k\neq i} c_{kj} - c_{ij}\Big) \\ &= \Big(c_{ij} + \sum_{k\neq i} c_{kj}\Big)\Big(c_{ij} + \sum_{l\neq j} c_{il}\Big) - c_{ij} \\ &= \Big(\sum_{i=1}^{n} c_{ij}\Big)\Big(\sum_{j=1}^{n} c_{ij}\Big) - c_{ij}. \end{aligned} \quad (3.42)$$

This may be rendered even more handy by introducing a bullet $\bullet$ notation, which indicates that it is summed over all alleles at the corresponding locus, e. g. $c_{i\bullet} := \sum_{j=1}^{n} c_{ij}$ for $i = 1, \dots, n$ with $c_{nn} := 1 - \sum_{(i,j)\neq(n,n)} c_{ij}$. The above formula then reads

$$D_{ij}(c) = c_{\bullet j} c_{i\bullet} - c_{ij}. \quad (3.43)$$

As $c_{i\bullet}$ corresponds to the gross frequency of (all carriers of) the allele i at the first locus etc., generalising the coordinates p, q from section 2.5, it becomes evident that now the linkage equilibrium between the two loci considered is comprised by

n^2 *linkage equilibria with respect to the alleles*[1] *i and j* (correspondingly $D_{ij} \equiv 0$), which are achieved if the product of the total frequencies of the alleles considered equals the frequency of the gamete carrying both these alleles (cf. also the analogue situation in the 2-allelic case on p. 23). However, only $(n-1)^2$ coefficients of linkage disequilibrium are independent as we have

$$\sum_{i=1}^{n} D_{ij}(c) = \sum_{i=1}^{n} c_{\bullet j} c_{i \bullet} - \sum_{i=1}^{n} c_{ij} = c_{\bullet j} - c_{\bullet j} = 0 \tag{3.44}$$

and analogously $\sum_{j=1}^{n} D_{ij}(c) \equiv 0$. Moreover, equation (3.44) implies that – analogously to the 2-allelic model – the allele frequencies $c_{i\bullet}$ and $c_{\bullet j}$, $i, j \in \{1, \ldots, n\}$ are not affected by recombinational effects (cf. equation (3.37)).

Following these observations, it appears beneficial to reformulate equation (3.41) in terms of coordinates reflecting the allele frequencies. Primarily, we may take the $2(n-1)$ coordinates $c_{i\bullet}, c_{\bullet j}$ resp. $x^{i\bullet}, x^{\bullet j}$ with $i, j = 1, \ldots, n-1$ (note $\sum_{i=1}^{n} c^{i\bullet} = 1$) and, analogous to the approach presented in section 2.5, add the $(n-1)^2$ coefficients of linkage disequilibrium D^{ij} with $i, j = 1, \ldots, n-1$ to constitute a full set of $n^2 - 1$ coordinates.

Transforming equation (3.41) into such $(x^\bullet, D)$-coordinates then yields drift coefficients (not taking into account selective effects)

$$\mu^{x^{i\bullet}}(x^\bullet, D) = \sum_{(k,l) \neq (n,n)} \lim_{N\to\infty} N R(N) D^{kl} \frac{\partial x^{i\bullet}}{\partial x^{kl}} = 0 \quad \text{for } i = 1, \ldots, n-1 \tag{3.45}$$

and likewise $\mu^{x^{\bullet j}}(x^\bullet, D) = 0$ for $j = 1, \ldots, n-1$ as expected by equation (3.44). Furthermore, we have

$$\begin{aligned}
\mu^{D^{ij}}(x^\bullet, D) &= \sum_{\substack{(k,l) \\ \neq (n,n)}} \lim_{N\to\infty} N R(N) D^{kl} \frac{\partial D^{ij}}{\partial x^{kl}} + \frac{1}{4} \sum_{\substack{(k,l)\neq(n,n) \\ (r,s)\neq(n,n)}} x^{kl} \big(\delta^{kl}_{rs} - x^{rs}\big) \frac{\partial^2 D^{ij}}{\partial x^{kl} \partial x^{rs}} \\
&= \lim_{N\to\infty} N R(N) \Big(\sum_{l=1}^{n} D^{il} x^{\bullet j} + \sum_{k=1}^{n} D^{kj} x^{i\bullet} - D^{ij} \Big)
\end{aligned}$$

[1]Correspondingly, the coefficients of linkage disequilibrium may rather be perceived as being indexed by combinations of alleles than by gamete types; this observation will become even clearer in section 3.5.

$$
\begin{aligned}
&+ \frac{1}{4}\sum_{l,r=1}^{n} x^{il}\big(\delta^{il}_{rj} - x^{rj}\big) + \frac{1}{4}\sum_{k,s=1}^{n} x^{kj}\big(\delta^{kj}_{is} - x^{is}\big) \\
&= -\lim_{N\to\infty} NR(N)D^{ij} - \frac{1}{2}D^{ij} \quad \text{for } i,j = 1,\ldots,n-1 \qquad (3.46)
\end{aligned}
$$

via

$$
\begin{aligned}
\frac{\partial D^{ij}}{\partial x^{kl}} &= x^{\bullet j}\delta^i_k + x^{i\bullet}\delta^j_l - \delta^i_k\delta^j_l \quad \text{for } i,j \neq n \text{ and } (k,l) \neq (n,n), \\
\frac{\partial^2 D^{ij}}{\partial x^{kl}\partial x^{rs}} &= \delta^i_k\delta^j_s + \delta^j_l\delta^i_r \quad \text{for } i,j \neq n \text{ and } (k,l),(r,s) \neq (n,n).
\end{aligned}
$$

Thus, recombination leads to (additional) decremental D^{ij}-drift for $i,j = 1,\ldots,n-1$ via the term $-\lim_{N\to\infty} NR(N)D^{ij}$, whereas the allele frequencies do not receive an additional drift component, generalising the result for the previous model. Furthermore, there is an additional drift component resulting from the choice of coordinates (currently $-\frac{1}{2}D^{ij}$); for a complete assessment of the drift situation, however, a comparison with Brownian motion as in section 2.6 would be necessary. To this end, also the diffusion terms need to be stated, yielding a full generalisation of equation (2.1); this, however, is not pursued right here due to the bulkiness of the corresponding expressions – but will be done for means of determining geometric properties of the linkage equilibrium states in section 3.5.1.

3.3 Multi-loci models

When considering recombinational models with $k \geq 3$ loci, there exist far more relations between different (subsets of) loci. In particular, different modes of recombination are possible, and for a certain recombinational event, it is hence necessary to define which loci actually recombine. This may be done by introducing the concept of *recombination masks* as in [27] with each mask corresponding to one particular mode of recombination; the full recombinational action is then assumed to be the *additive* effective of all active mask as will be described in the following.

3.3.1 Recombination masks

As stated in [27], a recombination mask specifies – if recombination occurs – which allele at a given locus is sampled. Assuming pairs of gametes from which recombination is performed, the recombination mask may be given as a $2 \times k$ matrix with entries 1 and $*$, e. g.

$$m = \begin{pmatrix} 1 & * & 1 & \dots & 1 & 1 & * \\ * & 1 & * & \dots & * & * & 1 \end{pmatrix}, \tag{3.47}$$

signifying that, at the first locus, the allele is sampled from the one gamete, while at the second locus, the allele is sampled from the other gamete and so on. Actually, giving one row of the matrix is sufficient, however, for notational reasons we will often need both rows: Then, the upper row is denoted by $\overline{m}$, whereas the lower row is denoted by $\underline{m}$. Furthermore, all matrices which can be transformed into one another by interchanging the upper and the lower row are identified (as they correspond to the same mode of recombination) and the trivial ones are excluded, which means that altogether $2^{k-1} - 1$ masks exist.

This *space of all recombination masks with k slots* is denoted by M_k, out of which of course only certain masks may be active in a given model. Correspondingly, it is assumed that every active mask $m \in M_k$ is assigned a *recombination rate* $R_m > 0$, denoting the rate at which recombination events governed by this mask occur, whereas non-active masks receive the recombination rate 0. Necessarily, we have $\sum_{m \in M_k} R_m \leq 1$. Again, these individual recombination rates may also depend on the population size N.

Under the effect of each mask m, recombination basically is a *2-locus event*, independently of the total number of loci k: All those loci at which the allele is sampled from the same gamete are linked by that mask and hence may be interpreted as a single locus, at which an accordingly increased number of alleles may occur; the corresponding holds true for the remaining loci, at which the alleles are sampled from the other gamete. More complex recombinational actions (i. e. recombination between three or more different subsets of loci) are included in the model via the additivity of single mask recombination events as already stated. For each individual mask, though, the considerations of the preceding section 3.2 are likewise applicable as is described in the following sections.

3.3.2 Further notation

With regard to a systematic treatment, we will need to introduce some additional notation: For a model with k loci, at which n alleles each may occur with $k, n \geq 2$, we number the loci with $1, \ldots, k$ and correspondingly write $i = i_1 \ldots i_k$ for a gamete containing alleles i_1 at locus 1, i_2 at locus 2 and so forth until k (and hence, $c_{i_1 \ldots i_k}$ for the corresponding frequency), where $i_1, \ldots, i_k$ now run from 1 to n. This yields an $(n^k - 1)$-dimensional state space (by dropping $c_{n \ldots n} = 1 - \sum_{(i_1,\ldots,i_k) \neq (n,\ldots,n)} c_{i_1 \ldots i_k}$) which is given by (cf. equation (3.33))

$$\begin{aligned}\Delta_{\frac{n^k-1}{N}} := \Big\{ \big(& c_{1\ldots 11}, \ldots, c_{1\ldots 1n}, c_{1\ldots 21}, \ldots, c_{1\ldots n1}, \ldots, c_{1\ldots nn}, \ldots, \\ & c_{2\ldots 11}, \ldots, c_{n\ldots nn-1} \big) \in \mathbb{N}^{n^k-1}_{N^{-1}} \Big| \sum_{(i_1,\ldots,i_k) \neq (n,\ldots,n)} c_{i_1 \ldots i_k} < 1 \Big\} \end{aligned} \tag{3.48}$$

for the discrete approach resp. the corresponding limit space for means of diffusion approximation being

$$\begin{aligned}\Delta_{n^k-1} = \Big\{ \big(& x^{1\ldots 11}, \ldots, x^{1\ldots 1n}, x^{1\ldots 21}, \ldots, x^{1\ldots n1}, \ldots, x^{1\ldots nn}, \ldots, \\ & x^{2\ldots 11}, \ldots, x^{n\ldots nn-1} \big) \in \mathbb{R}^{n^k-1} \Big| x^{i_1 \ldots i_k} > 0; \sum_{(i_1,\ldots,i_k) \neq (n,\ldots,n)} x^{i_1 \ldots i_k} < 1 \Big\}. \end{aligned} \tag{3.49}$$

Now, for some mask $m \in M_k$ (cf. equation (3.47)), in slight alteration of the notation, we denote the set of loci marked by an entry 1 in $\overline{m}$ also by $\overline{m} = \{\overline{m}_1, \ldots, \overline{m}_l\} \subset \{1, \ldots, k\}$ with $1 \leq l \leq k$ (corresponding to the index i of section 3.2.1), whereas the set of the remaining loci marked by an entry 1 in $\underline{m}$ is also denoted by $\underline{m} = \{\underline{m}_1, \ldots, \underline{m}_{k-l}\} = \{1, \ldots, k\} \setminus \{\overline{m}_1, \ldots, \overline{m}_l\}$ (corresponding to the index j previously), i. e.

$$\begin{array}{c} m = \begin{pmatrix} 1 & * & 1 & \ldots & 1 & 1 & * \\ * & 1 & * & \ldots & * & * & 1 \end{pmatrix}. \\ \begin{array}{ccccccc} \uparrow & \uparrow & \uparrow & & \uparrow & \uparrow & \uparrow \\ \overline{m}_1 & \underline{m}_1 & \overline{m}_2 & \ldots & \overline{m}_{l-1} & \overline{m}_l & \underline{m}_{k-l} \end{array} \end{array} \tag{3.50}$$

Correspondingly, for the alleles which m samples from the one (upper) gamete, we may thus write $i_{\overline{m}} = i_{\overline{m}_1} \dots i_{\overline{m}_l}$ as well as we may write $i_{\underline{m}} = i_{\underline{m}_1} \dots i_{\underline{m}_{k-l}}$ for the alleles sampled from the other (lower) gamete. Eventually, for the allelic configuration of a gamete produced by recombination governed by m, we may write $\langle i_{\overline{m}}, i_{\underline{m}} \rangle = \langle i_{\overline{m}_1} \dots i_{\overline{m}_l}, i_{\underline{m}_1} \dots i_{\underline{m}_{k-l}} \rangle$ with the angles

$$\langle \cdot , \cdot \rangle \tag{3.51}$$

indicating that the alleles are sorted into the correct ordering regarding the locus they are associated with. In the given example (cf. equation (3.50)), this would yield $\langle i_{\overline{m}}, i_{\underline{m}} \rangle = i_{\overline{m}_1} i_{\underline{m}_1} i_{\overline{m}_2} \dots i_{\overline{m}_{l-1}} i_{\overline{m}_l} i_{\underline{m}_{k-l}}$.

3.3.3 Quantitative analysis of multi-loci mask recombination

For some gamete $i = i_1 \dots i_k$ with corresponding frequency $c_i = c_{i_1 \dots i_k}$ and considering only one single mask $m \in M_k$, we consequently obtain in the RUG model after mating and recombination (cf. equation (3.37))

$$c_i' = c_i + R_m D_i^m, \tag{3.52}$$

whereas taking into account all masks $m \in M_k$, we obtain in consequence of the assumed additivity

$$c_i' = c_i + \sum_{m \in M_k} R_m D_i^m. \tag{3.53}$$

The appearing *coefficients of m-linkage disequilibrium* D_i^m are given by (cf. equation (3.42))

$$D_i^m := \sum_{\substack{(\tilde{i}_{\overline{m}_1},\dots,\tilde{i}_{\overline{m}_l}) \\ \neq (i_{\overline{m}_1},\dots,i_{\overline{m}_l})}} \sum_{\substack{(\tilde{i}_{\underline{m}_1},\dots,\tilde{i}_{\underline{m}_{k-l}}) \\ \neq (i_{\underline{m}_1},\dots,i_{\underline{m}_{k-l}})}} \left(c_{\langle \tilde{i}_{\overline{m}}, i_{\underline{m}} \rangle} c_{\langle i_{\overline{m}}, \tilde{i}_{\underline{m}} \rangle} - c_{\langle i_{\overline{m}}, i_{\underline{m}} \rangle} c_{\langle \tilde{i}_{\overline{m}}, \tilde{i}_{\underline{m}} \rangle} \right). \tag{3.54}$$

As linkage generally includes *all* loci, these coefficients effectively assess the disequilibrium of linkage between two subset of loci generated by the mask m, i. e. $\overline{m}$ and $\underline{m}$, referred to as the *m-linkage* here (for a more detailed discussion cf. also section 3.5.2).

Correspondingly, since the construction is symmetric in $\overline{m}$ and $\underline{m}$, the coefficients of m-linkage disequilibrium are well-defined with respect to the identification of upper and lower row of the mask m that we have stipulated (cf. p. 64). Instead of D_i^m, we may thus also write $D_i^{\overline{m}}$ or $D_i^{\underline{m}}$ – the required counterpart $\underline{m}$ resp. $\overline{m}$ is always given as the row of same dimension with complementary entries.

In the diffusion limit – with $D^{m,i}(x)$ formulated with upper indices for notational consistency – we obtain consequently as drift coefficients

$$\mu^i(x) = \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\big|C_N = c\big) = \lim_{N\to\infty} N \sum_{m\in M_k} R_m(N) D^{m,i}(x), \tag{3.55}$$

whereas the diffusion coefficients again show no effect of recombination as we have

$$\begin{aligned}\sigma^{i,j}(x) &= \lim_{N\to\infty} N\,\mathrm{E}_1\big(\delta C_{Ni}\delta C_{Nj}\big|C_N = c\big)\\ &= \lim_{N\to\infty} \frac{1}{2}\Big(c_i + \sum_{m\in M_k} R_m(N) D_i^m\Big)\delta_j^i\\ &\quad - \lim_{N\to\infty}\Big(c_i + \sum_{m\in M_k} R_m(N) D_i^m\Big)\Big(c_j + \sum_{m\in M_k} R_m(N) D_j^m\Big)\\ &\quad + \lim_{N\to\infty} N \sum_{m,\widetilde{m}\in M_k} R_m(N) R_{\widetilde{m}}(N) D_i^m D_j^{\widetilde{m}}\\ &= \frac{1}{2}x^i(\delta_j^i - x^j).\end{aligned} \tag{3.56}$$

With the results of section 3.1.1 for fitness schemes generalising also to multi loci, thus $\alpha_i = \alpha_{i_1\ldots i_k} \in \mathbb{R}$ being the additional limit fitness of allele $i = i_1 \ldots i_k$ and $\bar{\alpha} = \sum_i \alpha_i c_i$ being the average additional limit fitness of the population (cf. pp. 49 f.), we altogether have:

3.5 Theorem. *For $k, n \geq 2$, the diffusion approximation of a k-loci n-allelic Wright–Fisher model encompassing a multiplicative fitness scheme $\alpha = (\alpha_i) \in \mathbb{R}^{n^k}$ and recombination governed by recombination masks $m \in M_k$ with corresponding recombination rate R_m may be described by the Kolmogorov backward equation for its transition probability density $f\colon (\Delta_{n^k-1})_\infty \longrightarrow [0,1]$ of its gametic configuration $x = \big(x^{1\ldots 11}, \ldots, x^{1\ldots 1n}, x^{1\ldots 21}, \ldots, x^{1\ldots n1}, \ldots, x^{1\ldots nn}, \ldots, x^{2\ldots 11}, \ldots, x^{n\ldots nn-1}\big) \in \Delta_{n^2-1}$*

being

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{4} \sum_{i,j \neq (n,\dots,n)} x^i \big(\delta^i_j - x^j\big) \frac{\partial^2}{\partial x^i \partial x^j} f(x,t) + \sum_{i \neq (n,\dots,n)} \Big((\alpha_i - \bar{\alpha}) x^i + \lim_{N\to\infty} N \sum_{m \in M_k} R_m(N) D^{m,i}(x) \Big) \frac{\partial}{\partial x^i} f(x,t) \quad (3.57)$$

with $i = i_1 \dots i_k$, $j = j_1 \dots j_k$ *and* $(x,t) \in (\Delta_{n^k-1})_\infty$.

3.3.4 Alternative coordinates

Finding alternative coordinates adapted to the recombination interaction like $(x^\bullet, D)$ in the two-loci multi-allelic model (cf. section 3.2.2) is somewhat more intricate with the increased number of loci. However, at first the formula (3.54) for D^m_i may analogously be shortened into (cf. equation (3.43))

$$\begin{aligned} D^m_i(c) &= \Big(\sum_{i_{\overline{m}}=(1,\dots,1)}^{(n,\dots,n)} c_{\langle i_{\overline{m}}, i_{\underline{m}} \rangle} \Big) \Big(\sum_{i_{\underline{m}}=(1,\dots,1)}^{(n,\dots,n)} c_{\langle i_{\overline{m}}, i_{\underline{m}} \rangle} \Big) - c_i \\ &=: c_{\langle i_{\underline{m}}, \bullet \rangle} c_{\langle i_{\overline{m}}, \bullet \rangle} - c_i, \end{aligned} \quad (3.58)$$

wherein $\bullet$ as in $c_{\langle i_{\underline{m}}, \bullet \rangle}$ now signifies similarly to section 3.2.2 that it is summed over all corresponding alleles at the loci not belonging to $\underline{m}$, whereas the alleles $i_{\underline{m}}$ remain fixed, i. e.

$$c_{\langle i_{\underline{m}}, \bullet \rangle} := \sum_{i_{\overline{m}}=(1,\dots,1)}^{(n,\dots,n)} c_{\langle i_{\underline{m}}, i_{\overline{m}} \rangle} \quad (3.59)$$

with $c_{n\dots n} = 1 - \sum_{(i_1,\dots,i_k) \neq (n,\dots,n)} c_{i_1 \dots i_k}$; the significance of $c_{\langle i_{\overline{m}}, \bullet \rangle}$ is analogous.

The form of $D^{m,i}(x)$ shown is again symmetric in $\overline{m}$ and $\underline{m}$. Moreover, it is directly analogous to the one of the previous coefficients of linkage disequilibrium $D^{ij}(x)$ as may be seen when comparing equations (3.43) and (3.58), completing the structural analogy between two-loci recombination and multi-loci recombination via a certain mask alluded to in the beginning of the section.

Consequently, also the corresponding *allele tuple frequencies* $c_{\langle \bullet, i_{\underline{m}} \rangle}$ and $c_{\langle i_{\overline{m}}, \bullet \rangle}$

remain unaffected by the recombination governed by the mask m as we have

$$\sum_{i_{\overline{m}}=(1,\dots,1)}^{(n,\dots,n)} D_i^m(c) = \sum_{i_{\overline{m}}=(1,\dots,1)}^{(n,\dots,n)} c_{\langle i_{\underline{m}},\bullet\rangle} c_{\langle i_{\overline{m}},\bullet\rangle} - \sum_{i_{\overline{m}}=(1,\dots,1)}^{(n,\dots,n)} c_i = c_{\langle i_{\underline{m}},\bullet\rangle} - c_{\langle i_{\underline{m}},\bullet\rangle} = 0 \quad (3.60)$$

and analogously $\sum_{i_{\underline{m}}} D_i^m(c) = 0$. This, however, only holds for the contemplated mask m, whereas the action of any other mask may still alter $c_{\langle \bullet, i_{\underline{m}}\rangle}$ and $c_{\langle i_{\overline{m}},\bullet\rangle}$.

When searching for more general invariants of mask recombination, again the allele frequencies $(c_{\langle i_j,\bullet\rangle})$ with $i_j = 1,\dots,n-1$ for $j = 1,\dots,k$ come into focus as we have

$$\mu^{x^{\langle i_j,\bullet\rangle}}(x^\bullet, D) = \sum_{m\in M_k} \sum_{l\neq(n,\dots,n)} \lim_{N\to\infty} N R_m(N) D^{m,l} \frac{\partial x^{\langle i_j,\bullet\rangle}}{\partial x^l} = 0 \quad (3.61)$$

via $\frac{\partial x^{\langle i_j,\bullet\rangle}}{\partial x^l} = \delta_{l_j}^{i_j}$ and either $\{j\}^c \supset \overline{m}$ or $\{j\}^c \supset \underline{m}$ for the corresponding drift coefficients of the diffusion approximation (not taking into account selective effects). Thus, all allele frequencies remain unaffected by recombinational drift effects.

Similarly to the two-loci model, we may wish to additionally evaluate the influence of recombination on the drift behaviour of the coefficients of linkage disequilibrium. In doing so, we need to anticipate some concepts which are actually only introduced in sections 3.4 and 3.5: As we do not get a very enlightening result (cf. equation (3.46)) for the coefficients of m-linkage disequilibrium nor do they render a full alternative coordinate system (there are only $2^{k-1} - 1$ masks, whereas we have $n^k - 1$ gamete frequencies), we rather calculate drift coefficients for the $\binom{k}{2}(n-1)^2$ *coefficients of generalised 2-linkage disequilibrium*

$$D_2^{\langle i_{j_1}, i_{j_2},\bullet\rangle} \quad \text{with } j_1, j_2 \in \{1,\dots,k\}, j_1 \neq j_2;\ i_{j_1}, i_{j_2} = 1,\dots,n-1, \quad (3.62)$$

adapting the concept of twofold linkage interactions as with D^{ij} in equation (3.43) to the multi-loci setting (cf. section 3.5.2). Moreover, these coefficients reflect the action of all (relevant) masks as will be shown below.

Thus, transforming equation (3.57) into the yet incomplete $(x^\bullet, D_2)$-coordinates yields (again not taking into account selective effects)

$$
\begin{aligned}
\mu^{D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle}}(x^\bullet, D) &= \sum_{\widetilde{m}\in M_k} \sum_{l\neq(n,\dots,n)} \lim_{N\to\infty} N R_{\widetilde{m}}(N) D^{\widetilde{m},l} \frac{\partial D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle}}{\partial x^l} \\
&\quad + \frac{1}{4} \sum_{\substack{l\neq(n,\dots,n)\\ r\neq(n,\dots,n)}} x^l \big(\delta_r^l - x^r\big) \frac{\partial^2 D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle}}{\partial x^l \partial x^r} \\
&= \sum_{\widetilde{m}\in M_k} \lim_{N\to\infty} N R_{\widetilde{m}}(N) \sum_l \Big(D^{\widetilde{m},l} \delta_{l_{j_1}}^{i_{j_1}} x^{\langle i_{j_2}, \bullet\rangle} + D^{\widetilde{m},l} \delta_{l_{j_2}}^{i_{j_2}} x^{\langle i_{j_1}, \bullet\rangle} \\
&\quad - D^{\widetilde{m},l} \delta_{l_{j_1}}^{i_{j_1}} \delta_{l_{j_2}}^{i_{j_2}} \Big) + \frac{1}{2} \big(x^{\langle i_{j_1}, i_{j_2}, \bullet\rangle} - x^{\langle i_{j_1}, \bullet\rangle} x^{\langle i_{j_2}, \bullet\rangle} \big) \\
&= - \sum_{m\in M_2} \lim_{N\to\infty} N R_m(N) D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle} - \frac{1}{2} D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle}
\end{aligned}
\tag{3.63}
$$

with $R_m := \sum_{\{\widetilde{m}\in M_k | \widetilde{m}^{*\langle i_{j_1}, i_{j_2}, *\rangle} = m\}} R_{\widetilde{m}}$ for $m \in M_2$ via

$$
\begin{aligned}
\frac{\partial D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle}}{\partial x^l} &= x^{\langle i_{j_2}, \bullet\rangle} \delta_{l_{j_1}}^{i_{j_1}} + x^{\langle i_{j_1}, \bullet\rangle} \delta_{l_{j_2}}^{i_{j_2}} - \delta_{l_{j_1}}^{i_{j_1}} \delta_{l_{j_2}}^{i_{j_2}}, \\
\frac{\partial^2 D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle}}{\partial x^l \partial x^r} &= \delta_{l_{j_1}}^{i_{j_1}} \delta_{r_{j_2}}^{i_{j_2}} + \delta_{r_{j_1}}^{i_{j_1}} \delta_{l_{j_2}}^{i_{j_2}}
\end{aligned}
$$

and by equation (3.60). Thus, the coefficients of generalised 2-linkage disequilibrium $D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle}$ receive an additional decremental drift component caused by recombinational events with masks $\widetilde{m}$ bisecting the loci j_1 and j_2, i. e. $j_1 \in \overline{m}$ and $j_2 \in \underline{m}$ or conversely, plus an extra component resulting from the choice of coordinates $(-\frac{1}{2} D_2^{\langle i_{j_1}, i_{j_2}, \bullet\rangle})$. The property of bisecting the defined loci may also be formulated by anticipating the schema terminology and notation presented in section 3.4: Thus, drift is only caused by masks which are not converted into the trivial mask by reduction with the schema class $\langle i_{j_1}, i_{j_2}, *\rangle$ (cf. also pp. 74 f. for a more detailed discussion of the interplay of schema classes and masks).

However, as $(x^{\bullet\bullet}, D_2)$ does not yet form a full set of coordinates, we may add the coefficients of generalised l-linkage disequilibrium, $3 \leq l \leq k$ anticipated from section 3.5.3 (for a definition cf. equation (3.120)). For these coefficients, we similarly obtain as drift coefficients

$$\begin{aligned}\mu^{D_l^{\langle i_{j_1},\dots,i_{j_l},\bullet\rangle}}(x^\bullet, D_2,\dots,D_k) &= -\sum_{\widetilde{m}\in M_k}\lim_{N\to\infty} NR_{\widetilde{m}}(N)\sum_r D^{\widetilde{m},r}\delta^{i_{j_1}\dots i_{j_l}}_{r_{j_1}\dots i_{r_l}} + \dots \\ &= -\sum_{m\in M_l}\lim_{N\to\infty} NR_m(N) D^{m,\langle i_{j_1},\dots,i_{j_l},\bullet\rangle} \qquad (3.64)\\ &\quad + \textit{terms originating from the choice of coordinates}\end{aligned}$$

with $R_m := \sum_{\{\widetilde{m}\in M_k|\widetilde{m}^{*\langle i_{j_1},\dots,i_{j_l},*\rangle}=m\}} R_{\widetilde{m}}$ for $m \in M_l$, similarly signifying that also the coefficients of l-linkage disequilibrium receive a recombinational drift from all mask recombination events whose mask bisects the defined loci $j_1,\dots,j_l$, i. e. is not converted into the trivial mask when reduced by the schema class $\langle i_{j_1},\dots,i_{j_l},*\rangle$ (plus additional terms coming from the choice of coordinates). However, for a full assessment of the drift situation, again a comparison with Brownian motion as in section 2.6 would be necessary – this would imply even more calculative effort than with the two-loci model and is hence not pursued here.

3.4 Schemata

In many situation of mathematical modelling, it has proven useful to shift from a detailed, low-level assessment of a given system to a more abstract, higher-level view – a technique which is commonly known as *coarse-graining*. In the context of genetic models, this may be implemented by reducing the information of the genetic configuration via bundling certain gamete frequencies. In doing so, it stands to reason to particularly coarse-grain the information of certain loci, i. e. to bundle gametes which have a certain number of alleles in common. These clusters of alleles are then dubbed *schemata* as presented in [27], a concept which is already foreshadowed in the context of recombination, i. e. with the introduction of recombination masks m, whose halves $\overline{m}$ and $\underline{m}$ may be directly interpreted as a *schema class* as will be illustrated.

3.4.1 Notation

Returning to the setting of a k-loci n-allelic Wright–Fisher model with the loci denoted by $1, \dots, k$ and the corresponding alleles by $i_1, \dots, i_k \in \{1, \dots, n\}$, thus with state space $\Delta_{\frac{n^k-1}{N}}$ (resp. Δ_{n^k-1} in case of diffusion approximation; cf. equations (3.48) and (3.49)), a *schema* of order $l < k$ may be given by specifying a certain allele for l loci, whereas for the remaining $k - l$ loci, no allele is specified (cf. [27], pp. 9 f.). Denoting the non-specified alleles by $\bullet$, we may for example consider a schema ($k = 5$, $n = 5$)

$$i_1 \bullet i_3 i_4 i_5 \tag{3.65}$$

with $i_1, i_3, i_4, i_5 \in \{1, \dots, 5\}$, comprising the gametes $i_1 1 i_3 i_4 i_5$, $i_1 2 i_3 i_4 i_5$, $i_1 3 i_3 i_4 i_5$, $i_1 4 i_3 i_4 i_5$, $i_1 5 i_3 i_4 i_5$.

Furthermore, for a systematic treatment, we introduce the notion of a *schema class* s (of length k and order $l < k$) being given by a $1 \times k$-matrix with entries 1 and $*$ with the entry 1 appearing exactly l times, thus e. g. ($k = 5$, $l = 4$)

$$s = \big(1, *, 1, 1, 1\big). \tag{3.66}$$

Similar to the recombination masks of section 3.3.1, these entries are sequentially referring to the loci of a gamete, signifying that, at a locus marked with 1, the allele is taken into account, whereas it is not taken into account at loci marked with $*$. Correspondingly, a schema defining an allele exactly at every locus marked with 1 is said to correspond to the given schema class s. In a slight alteration of the notation, the set of loci marked by entry 1 is also denoted by $s = \{s_1, \dots, s_l\} \subset \{1, \dots, k\}$ with the corresponding set of alleles denoted by $i_s = i_{s_1} \dots i_{s_l}$. The remaining loci are denoted by $s^c = \{s^c_1, \dots, s^c_{k-l}\}$, and the corresponding set of alleles subsequently by $i_{s^c} = i_{s^c_1} \dots i_{s^c_{k-l}}$.

Now, a schema corresponding to schema class s may be denoted by $\langle i_s, \bullet \rangle$ with $i_s \in \{1, \dots, n\}^l$ giving the alleles at the specified loci and the bullet $\bullet$ indicating that the allele at the remaining loci (i. e. s^c) are undefined. The angles (cf. equation (3.51)) signify that the alleles are fitted into the correct sequential ordering in accordance with the loci they are associated with. Consequently, the frequency of a schema

$\langle i_s, \bullet \rangle$ corresponding to some schema class s is then determined by summing over all gamete frequencies $c_i \in \Delta_{\frac{n^k-1}{N}}$ comprised by that schema, thus

$$c_{\langle i_s, \bullet \rangle} := \sum_{i_{s^c}=(1,\dots,1)}^{(n,\dots,n)} c_{\langle i_s, i_{s^c} \rangle} \tag{3.67}$$

with again $c_{n\dots n} = 1 - \sum_{(i_1,\dots,i_k)\neq(n,\dots,n)} c_{i_1\dots i_k}$.

3.4.2 Quantitative analysis

We may now transfer the results of the preceding section to a k-loci n-allelic recombinational Wright–Fisher model in a coarse-grained view by schemata. This may for example be achieved by selecting a certain schema class s (of order $l < k$) and assessing the model by giving the frequencies of all schemata relating to s. The corresponding state space is then given by $\Delta_{\frac{n^l-1}{N}}$ (by dropping $c_{\langle n\dots n_s, \bullet \rangle} = 1 - \sum_{i_s \neq (n,\dots,n)} c_{\langle i_s, \bullet \rangle}$) resp. Δ_{n^l-1} in case of diffusion approximation; this also reflects the loss of information caused by coarse-graining as the ungrained model comes with the wider state space $\Delta_{\frac{n^k-1}{N}}$ resp. Δ_{n^k-1}.

Directly resorting to the diffusion approximation of the model, theorem 3.5 may then be adapted to the schema view by transforming coordinates $(x_i) \mapsto (\tilde{x}_{\langle i_s, \bullet \rangle})$ correspondingly. In doing so, we have

$$\frac{\partial \tilde{x}_{\langle i_s, \bullet \rangle}}{\partial x_j} = \delta^{i_s}_{j_s}, \tag{3.68}$$

and consequently the coefficients of the first order derivatives in equation (3.57) transform as (cf. lemma 2.2)

$$\begin{aligned} \tilde{b}^{\langle i_s, \bullet \rangle}(\tilde{x}) &= \sum_j b^j(x) \delta^{i_s}_{j_s} \\ &= \sum_j \Big((\alpha_j - \bar{\alpha}) x^j + \lim_{N\to\infty} N \sum_{m \in M_k} R_m(N) D^{m,j}(x) \Big) \delta^{i_s}_{j_s} \\ &\equiv (\alpha_{\langle i_s, \bullet \rangle}(x) - \bar{\alpha}(x)) \tilde{x}^{\langle i_s, \bullet \rangle} + \lim_{N\to\infty} N \sum_{m \in M_k} R_m(N) \sum_{i_{s^c}} D^{m,i}(x) \end{aligned} \tag{3.69}$$

for $i_s \in \{1, \dots, n\}^l \setminus \{(n, \dots, n)\}$ when putting

$$\alpha_{\langle i_s, \bullet \rangle}(x) := \frac{\sum_{i_{s^c}} \alpha_i x^i}{\tilde{x}^{\langle i_s, \bullet \rangle}} \tag{3.70}$$

for the fitness term. This expression may be interpreted as the additional limit fitness of the schema $\langle i_s, \bullet \rangle$. However, as indicated, it is still dependent on the individual gamete fitness as the schema fitness of course evaluates the weighted average fitness of all gametes present in that schema. This may be overcome by also lifting the assignment of fitness values to the level of schemata with the downside that the individual fitness of alleles can no longer be taken into account, thus defining $\alpha_{\langle i_s, \bullet \rangle}$ for all $i_s \in \{1, \dots, n\}^l$ and subsequently putting

$$\alpha_i := \alpha_{\langle i_s, \bullet \rangle} \quad \text{for all } i \in \{1, \dots, n\}^k. \tag{3.71}$$

If doing so, also the average additional limit fitness $\bar{\alpha}$ may be formulated in terms of $\tilde{x}$ as we have

$$\bar{\alpha}(x) = \sum_i \alpha_i x^i = \sum_{i_s} \alpha_{\langle i_s, \bullet \rangle} \sum_{i_{s^c}} x^i \equiv \sum_{i_s} \alpha_{\langle i_s, \bullet \rangle} \tilde{x}^{\langle i_s, \bullet \rangle}. \tag{3.72}$$

Continuing with the recombination term, we obtain (cf. equation (3.58))

$$\sum_{i_{s^c}} D^{m,i}(x) = \sum_{i_{s^c}} \Big(\sum_{i_{\overline{m}}} x_i \Big) \Big(\sum_{i_{\underline{m}}} x_i \Big) - \sum_{i_{s^c}} x_i. \tag{3.73}$$

The recombination mask m bisects the set of loci into $\overline{m}$ and $\underline{m}$. Thus, m also defines a segmentation $s^c = \overline{s}^c \,\dot{\cup}\, \underline{s}^c$ with $\overline{s}^c \subset \overline{m}$ and $\underline{s}^c \subset \underline{m}$. Regarding the mask m itself, all loci s^c are effectively removed from the mask, converting it non-injectively into a mask $m^{*s} \in M_l$, thus of dimension $2 \times l$ with then $\overline{m}^{*s} \cup \underline{m}^{*s} = s$; if e. g. $m = \left(\begin{smallmatrix} 1*1** \\ *1*11 \end{smallmatrix}\right)$ and $s = (1 * 111)$, we obtain (with the diamonds $\diamond$ indicating the entries effectively removed)

$$m^{*s} = \begin{pmatrix} 1 & \diamond & 1 & * & * \\ * & \diamond & * & 1 & 1 \end{pmatrix}. \tag{3.74}$$

Correspondingly, we also say that the mask m is *reduced* by the schema class s, yielding the reduced mask m^{*s}.

For the coefficients of m-linkage disequilibrium, we thus have

$$\begin{aligned}\sum_{i_{s^c}} D^{m,i}(x) &= \Big(\sum_{i_{\overline{s}^c}}\sum_{i_{\overline{m}}} x_i\Big)\Big(\sum_{i_{\underline{s}^c}}\sum_{i_{\underline{m}}} x_i\Big) - \tilde{x}_{\langle i_s,\bullet\rangle} \\ &= \Big(\sum_{i_{(s^c\cup i_{\overline{m}^{*s}})}} x_i\Big)\Big(\sum_{i_{(s^c\cup i_{\underline{m}^{*s}})}} x_i\Big) - \tilde{x}_{\langle i_s,\bullet\rangle} \\ &= \Big(\sum_{i_{\overline{m}^{*s}}} \tilde{x}_{\langle i_s,\bullet\rangle}\Big)\Big(\sum_{i_{\underline{m}^{*s}}} \tilde{x}_{\langle i_s,\bullet\rangle}\Big) - \tilde{x}_{\langle i_s,\bullet\rangle} =: D^{m^{*s},\langle i_s,\bullet\rangle}(\tilde{x}). \qquad (3.75)\end{aligned}$$

The form of the appearing *coefficients of m^{*s}-linkage disequilibrium* $D^{m^{*s},\langle i_s,\bullet\rangle}(\tilde{x})$ is again analogous to the previous coefficients $D^{m,i}(x)$, only with schema frequencies instead of gamete frequencies appearing as argument. This may be seen when comparing equations (3.75) and (3.58) – and is to be expected as the recombination event itself remains unaffected.

However, another effect enters into the discussion as for certain combinations of m and s, we have either $\overline{m}^{*s} = \varnothing$ or $\underline{m}^{*s} = \varnothing$ and correspondingly $\underline{m}^{*s} = s$ resp. $\overline{m}^{*s} = s$, which signifies that the mask m is converted into a trivial mask via the deletion of the unspecified loci, e. g. $\left(\begin{smallmatrix}1{*}111\\ {*}1{*}{*}{*}\end{smallmatrix}\right) \rightsquigarrow \left(\begin{smallmatrix}1{\diamond}111\\ {*}{\diamond}{*}{*}{*}\end{smallmatrix}\right)$. Correspondingly, either of the sums in equation (3.75) then is such that effectively the summation spans all loci (consequently yielding 1), whereas the other sum comprises only one summand, i. e. $\tilde{x}_{\langle i_s,\bullet\rangle}$. Consequently, $D^{m^{*s},\langle i_s,\bullet\rangle}$ then always evaluates to zero, reflecting the fact that, with the trivial mask, no linkage relation is expressed and hence the corresponding coefficient always indicates 'linkage equilibrium', i. e. $D = 0$.

With M_l being the domain of all non-trivial reduced masks of dimension l, we thus have for the full recombination term in equation (3.69)

$$\begin{aligned}&\lim_{N\to\infty} N \sum_{\widetilde{m}\in M_k} R_{\widetilde{m}}(N) \sum_{i_{s^c}} D^{\widetilde{m},i}(x) \\ &= \lim_{N\to\infty} N \sum_{\widetilde{m}\in M_k} R_{\widetilde{m}}(N) D^{\widetilde{m}^{*s},\langle i_s,\bullet\rangle}(\tilde{x}) \\ &= \lim_{N\to\infty} N \sum_{m\in M_l} R_m(N) D^{m,\langle i_s,\bullet\rangle}(\tilde{x}), \qquad (3.76)\end{aligned}$$

wherein the recombination rate R_m associated to $m \in M_l$ reflects that the reduced mask m may originate from different masks $\widetilde{m} \in M_k$, thus

$$R_m := \sum_{\{\widetilde{m} \in M_k | \widetilde{m}^{*s} = m\}} R_{\widetilde{m}} \quad \text{for } m \in M_l. \tag{3.77}$$

If a mask $\widetilde{m}$ is reduced into the trivial mask, the corresponding term does not appear in the sum as then $D^{\widetilde{m}^{*s},\langle i_s,\bullet\rangle}$ in accordance with the above considerations vanishes.

Ultimately, for the coefficients of the second order derivatives in equation (3.57), we obtain

$$\begin{aligned} \tilde{a}^{\langle i_s,\bullet\rangle,\langle j_s,\bullet\rangle}(\tilde{x}) &= \sum_{k,l} a^{k,l}(x) \delta^{i_s}_{k_s} \delta^{j_s}_{l_s} \\ &= \sum_{k,l} x^k (\delta^k_l - x^l) \delta^{i_s}_{k_s} \delta^{j_s}_{l_s} \\ &= \tilde{x}^{\langle i_s,\bullet\rangle} \big(\delta^{i_s}_{j_s} - \tilde{x}^{\langle j_s,\bullet\rangle} \big). \end{aligned} \tag{3.78}$$

Altogether, the adapted version of theorem 3.5 then reads:

3.6 Corollary. *For $k, n \geq 2$, $l \in \{1, \dots, k-1\}$ and a schema class s of length k and order l, the diffusion approximation of a k-loci n-allelic Wright–Fisher model encompassing a fitness scheme $\alpha = (\alpha_{\langle i_s,\bullet\rangle}) \in \mathbb{R}^{n^l}$ and recombination governed by recombination masks $\widetilde{m} \in M_k$ with corresponding recombination rate $R_{\widetilde{m}}$ may be described by the Kolmogorov backward equation for its transition probability density $f \colon (\Delta_{n^l-1})_\infty \longrightarrow [0,1]$ of its schemata configuration $x = (x^{\langle 1\dots 11_s,\bullet\rangle}, \dots, x^{\langle 1\dots nn-1_s,\bullet\rangle}) \in \Delta_{n^l-1}$ being*

$$\begin{aligned} \frac{\partial}{\partial t} f(x,t) = \frac{1}{4} &\sum_{i_s,j_s \neq (n,\dots,n)} x^{\langle i_s,\bullet\rangle} \big(\delta^{i_s}_{j_s} - x^{\langle j_s,\bullet\rangle} \big) \frac{\partial^2}{\partial x^{\langle i_s,\bullet\rangle} \partial x^{\langle j_s,\bullet\rangle}} f(x,t) \\ &+ \sum_{i_s \neq (n,\dots,n)} (\alpha_{\langle i_s,\bullet\rangle}(x) - \bar{\alpha}(x)) x^{\langle i_s,\bullet\rangle} \frac{\partial}{\partial x^{\langle i_s,\bullet\rangle}} f(x,t) \\ &+ \lim_{N\to\infty} N \sum_{m \in M_l} R_m(N) \sum_{i_s \neq (n,\dots,n)} D^{m,\langle i_s,\bullet\rangle}(\tilde{x}) \frac{\partial}{\partial x^{\langle i_s,\bullet\rangle}} f(x,t) \end{aligned} \tag{3.79}$$

with $i_s = i_{s_1} \dots i_{s_l}$, $j_s = j_{s_1} \dots j_{s_l}$, $R_m = \sum_{\{\widetilde{m} \in M_k | \widetilde{m}^{*s} = m\}} R_{\widetilde{m}}$ and $(x,t) \in (\Delta_{n^l-1})_\infty$.

3.5 The geometry of linkage equilibrium states

Returning to the Ohta–Kimura formula (2.1), we may put $D = 0$, yielding

$$(a^{ij}(p,q,0)) = \frac{1}{4}\begin{pmatrix} p(1-p) & 0 & 0 \\ 0 & q(1-q) & 0 \\ 0 & 0 & p(p-1)q(q-1) \end{pmatrix} \tag{3.80}$$

with $(p,q,0) \in \Omega_{(p,q,0)} = \{(p,q,0) \in \mathbb{R}^2 \times \{0\} \big| 0 < p,q < 1\}$ as coefficient matrix of the second order derivatives and

$$(a_{ij}(p,q,0)) = 4\begin{pmatrix} \frac{1}{p(1-p)} & 0 & 0 \\ 0 & \frac{1}{q(1-q)} & 0 \\ 0 & 0 & \frac{1}{p(p-1)q(q-1)} \end{pmatrix} \tag{3.81}$$

for its inverse (cf. equations (2.13) f.). As described in section 2.2, $(a_{ij}(p,q,D))$ may be interpreted as a metric (i. e. the Fisher metric of the multinomial distribution) on $\Omega_{(p,q,D)}$: Thus, when dropping the third coordinate $D = 0$, $(a_{ij}(p,q,0))$ turns into a *product metric* (cf. definition 3.7 below) on $\Delta_1 \times \Delta_1$, in which each factor Δ_1 is equipped with the (inverse) metric $g(x) = \frac{1}{4}x(1-x)$, $x \in \Delta_1$ – corresponding (up to the prefactor) to the standard metric of the 1-dimensional sphere $S^1_+ \subset \mathbb{R}^2_+$ (cf. section 2.2.3). Hence, the state space $\Omega_{(p,q,0)}$ resp. a corresponding restriction of Δ_3 of the diffusion approximation of the two-loci two-allelic recombinational Wright–Fisher model *in linkage equilibrium* (cf. p. 23) equipped with the Fisher metric of the multinomial distribution (cf. lemma 2.7) carries (independently of the chosen coordinate representation) the geometrical structure of

$$S^1_+ \times S^1_+ \subset S^3_+, \tag{3.82}$$

which is known as the *Clifford Torus* (after William K. Clifford, who in [8] first described $S^1 \times S^1$ as a closed, (locally) Euclidean surface embedded in an elliptic 3-space (cf. [28], p. 373)).

We wish to extend this observation to more general Wright–Fisher models as presented previously in this chapter. First, we give a definition of the product metric (cf. [9], p. 42):

3.7 Definition. Let (M_1, g_1) and (M_2, g_2) be Riemannian manifolds and consider the Cartesian product $M_1 \times M_2$; let $\pi_1 \colon M_1 \times M_2 \longrightarrow M_1$ and $\pi_2 \colon M_1 \times M_2 \longrightarrow M_2$ be the natural projections. Then the *product metric* $g_1 \times g_2$ on $M_1 \times M_2$ is defined as

$$(g_1 \times g_2)(u,v)_{(p,q)} = g_1\big(d\pi_1(p,q)(u), d\pi_1(p,q)(v)\big)_p + g_2\big(d\pi_2(p,q)(u), d\pi_2(p,q)(v)\big)_q \quad \text{for all } (p,q) \in M_1 \times M_2, u, v \in T_{(p,q)}(M_1 \times M_2) \tag{3.83}$$

with $d\pi_1$, $d\pi_2$ being the *derivatives* of the natural projections (cf. [13], p. 6).

The following lemma may be checked directly:

3.8 Lemma. *Let (M_1, g_1) and (M_2, g_2) be Riemannian manifolds and let $g_1 \times g_2$ as in definition 3.7. Then $(M_1 \times M_2, g_1 \times g_2)$ is a Riemannian manifold.*

Taking the representation in local coordinates of the metrics, i. e. $(g_{1ij}(x_1))_{i,j=1,\ldots,n}$ with $x_1 = (x_1^1, \ldots, x_1^n)$ and $(g_{2kl}(x_2))_{k,l=n+1,\ldots,n+m}$ with $x_2 = (x_2^{n+1}, \ldots, x_2^{n+m})$, equation (3.83) turns into

$$\big((g_1 \times g_2)_{rs}\big)_{r,s=1,\ldots,m} = \begin{pmatrix} (g_{1ij}(x_1))_{i,j=1,\ldots,n} & 0^{n,m} \\ 0^{m,n} & (g_{2kl}(x_2))_{k,l=n+1,\ldots,n+m} \end{pmatrix} \tag{3.84}$$

with $0^{n,m}$ being the $n \times m$ null matrix. For the inverse metric, we correspondingly obtain

$$\big((g_1 \times g_2)^{rs}\big)_{r,s=1,\ldots,m} = \begin{pmatrix} ({g_1}^{ij}(x_1))_{i,j=1,\ldots,n} & 0^{n,m} \\ 0^{m,n} & ({g_2}^{kl}(x_2))_{k,l=n+1,\ldots,n+m} \end{pmatrix} \tag{3.85}$$

as for a block matrix $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ like $\big((g_1 \times g_2)_{rs}\big)_{r,s=1,\ldots,m}$, we have

$$M^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(M/A)^{-1}CA^{-1} & -A^{-1}B(M/A)^{-1} \\ -(M/A)^{-1}CA^{-1} & (M/A)^{-1} \end{pmatrix} \tag{3.86}$$

with $(M/A)^{-1} := (D - CA^{-1}B)^{-1}$ being the *Schur complement of A in M* (cf. [34], pp. 17 f.). Clearly, these considerations extend straightforwardly to product metrics with more than two factors.

3.5.1 Linkage equilibria in two-loci multi-allelic models

In order to generalise the above observations, we first need to extend the notion of linkage equilibrium to more advanced models. Generally, 'linkage' between certain loci (or sets of loci) relates to the fact that the allelic configuration at the one locus resp. loci set affects/determines the allelic configuration at the other locus resp. loci set (cf. also p. 20), i. e. in the two-loci 2-allelic model, each allele at the one locus is 'linked' with both other alleles at the other locus. However, as the allele frequencies at each locus actually form a 1-dimensional space, for a given allele only one parameter suffices to describe the degree of linkage disequilibrium with respect to both other alleles at the other locus as well as both relations for the other allele at the same locus. Hence, only one parameter is sufficient to describe $2 \cdot 2$ linkage relations; this fact is algebraicly expressed in equation (3.38).

With $n \geq 3$ alleles and two loci (cf. also section 3.2), we have $n \cdot n$ linkage relations, which – as the allele frequencies at each locus form an $(n-1)$-dimensional space – effectively reduces to $(n-1)^2$ linkage relations, which are correspondingly expressed by $(n-1)^2$ coefficients of linkage disequilibrium D^{ij}, $i, j = 1, \ldots, n-1$ defined in equation (3.36). If all these coefficients vanish, the population is defined to be in *linkage equilibrium*, signifying that each allele at the one locus is in linkage equilibrium with all other alleles at the other locus and conversely.

To determine the geometrical structure of the two-loci model in linkage equilibrium, the corresponding state space Δ_{n^2-1} (cf. equation (3.34)) needs to be transformed appropriately, i. e. with all D^{ij} appearing as coordinate. This may be achieved by transforming into the alternative coordinates $(x^\bullet, D) = (x^\bullet, D)_{i,j=1,\ldots,n-1}$ comprising $2n-2$ allele frequencies $x^{i\bullet}$ and $x^{\bullet j}$ and all coefficients of linkage disequilibrium D^{ij} (cf. section 3.2.2). This implies

$$\frac{\partial x^{i\bullet}}{\partial x^{kl}} = \delta^i_k, \quad \frac{\partial x^{\bullet j}}{\partial x^{kl}} = \delta^j_l, \quad \frac{\partial D^{ij}}{\partial x^{kl}} = -\delta^i_k \delta^i_l + \delta^i_k x^{\bullet j} + x^{i\bullet} \delta^j_l \tag{3.87}$$

for $i, j \neq n$ and $(k, l) \neq (n, n)$, and transforming the (inverse) metric given by the coefficients of the second order derivatives of the corresponding Kolmogorov equation (3.41), i. e. $(a^{ij,kl}(x)) = \big(x^{ij}(\delta^{ij}_{kl} - x^{kl})\big)$, accordingly yields

$$\begin{aligned} a^{x^{i\bullet}, x^{\bullet l}}(x^\bullet, D) &= \sum_{m,n} \sum_{r,s} x^{mn}(\delta^{mn}_{rs} - x^{rs}) \frac{\partial x^{i\bullet}}{\partial x^{mn}} \frac{\partial x^{\bullet l}}{\partial x^{rs}} \\ &= \sum_{n} \sum_{r} x^{in}(\delta^{in}_{rl} - x^{rl}) \\ &= x^{il} - x^{i\bullet} x^{\bullet l} = D^{il} \\ &\equiv a^{x^{\bullet l}, x^{i\bullet}}(x^\bullet, D) \quad \text{for } i, l = 1, \dots, n-1 \end{aligned} \tag{3.88}$$

(with the last equality being due to the undirectedness of linkage) and similarly

$$a^{x^{i\bullet}, x^{k\bullet}}(x^\bullet, D) = x^{i\bullet}(\delta^i_k - x^{k\bullet}) \quad \text{for } i, k = 1, \dots, n-1, \tag{3.89}$$

$$a^{x^{\bullet j}, x^{\bullet l}}(x^\bullet, D) = x^{\bullet j}(\delta^j_l - x^{\bullet l}) \quad \text{for } j, l = 1, \dots, n-1. \tag{3.90}$$

For the further components of the metric, we have

$$\begin{aligned} a^{x^{i\bullet}, D^{kl}}(x^\bullet, D) &= \sum_{m,n} \sum_{r,s} x^{mn}(\delta^{mn}_{rs} - x^{rs}) \frac{\partial x^{i\bullet}}{\partial x^{mn}} \frac{\partial D^{kl}}{\partial x^{rs}} \\ &= \sum_{n} x^{in} \Big(-(\delta^{in}_{kl} - x^{kl}) + \sum_{r} (\delta^{in}_{rl} - x^{rl}) x^{k\bullet} + \sum_{s} (\delta^{in}_{ks} - x^{ks}) x^{\bullet l} \Big) \\ &= -\delta^i_k D^{kl} + x^{i\bullet} D^{kl} + x^{k\bullet} D^{il} \\ &\equiv a^{D^{kl}, x^{i\bullet}}(x^\bullet, D) \quad \text{for } i, k, l = 1, \dots, n-1 \end{aligned} \tag{3.91}$$

and analogously

$$a^{x^{\bullet j}, D^{kl}}(x^\bullet, D) \equiv a^{D^{kl}, x^{\bullet j}}(x^\bullet, D) = -\delta^j_l D^{kl} + x^{\bullet j} D^{kl} + x^{\bullet l} D^{kj} \quad \text{for } j, k, l = 1, \dots, n-1, \tag{3.92}$$

thus these entries also vanish in linkage equilibrium, signifying that the corresponding coordinate representation of the inverse metric in linkage equilibrium entirely turns into a block matrix, i. e.

$$(a^{(x^\bullet,D)}) = \begin{pmatrix} (a^{x^{i\bullet},x^{k\bullet}}) & 0^{n-1,n-1} & 0^{n-1,(n-1)^2} \\ 0^{n-1,n-1} & (a^{x^{\bullet j},x^{\bullet l}}) & 0^{n-1,(n-1)^2} \\ 0^{(n-1)^2,n-1} & 0^{(n-1)^2,n-1} & (a^{D^{ij},D^{kl}}) \end{pmatrix}, \tag{3.93}$$

with the remaining entries being

$$a^{D^{ij},D^{kl}}(x^\bullet, D) = (x^{i\bullet} - \delta^i_k)x^{k\bullet}(x^{\bullet j} - \delta^j_l)x^{\bullet l} \quad \text{for } i,j,k,l = 1,\dots,n-1, \tag{3.94}$$

hence giving the generalisation of the metric representation corresponding to the Ohta–Kimura formula in linkage equilibrium (cf. equation (3.80)) to an arbitrary number of alleles.

Moreover, $(a^{(x^\bullet,D)}(x^\bullet, D))$ may be inverted in accordance with equation (3.86), exhibiting the product structure of $(a_{(x^\bullet,D)}(x^\bullet, D))$. We thus have:

3.9 Lemma. *In linkage equilibrium, for all $n \geq 2$ the corresponding restriction of the state space Δ_{n^2-1} of the diffusion approximation of a two-loci n-allelic Wright–Fisher model equipped with the Fisher metric of the multinomial distribution is a $(2n-2)$-dimensional manifold and carries the geometric structure of*

$$S^{n-1}_+ \times S^{n-1}_+ \subset S^{2n-1}_+. \tag{3.95}$$

3.5.2 Linkage equilibria in three-loci multi-allelic models

Having more than two loci, the situation gets significantly more complicated as now linkage does not only need to be considered between pairs of loci, but also in relations of higher order. In particular it needs to be clarified what exactly is to be understood by the term ‘linkage equilibrium’ in this extended setting. To keep the calculative effort manageable, we first analyse a three-loci model; in doing so, we will give a definition of linkage equilibrium in the current setting plus an adaption of lemma 3.9.

When wishing to analyse the geometry of the corresponding state space Δ_{n^3-1} (cf. equation (3.48) with $k = 3$) restricted to the – yet to be determined – linkage equilibrium states, again we need to transform the state space appropriately, for which in turn suitable coordinates are required: This may be done somewhat tentatively by first adapting the coordinate scheme of the two-loci model as far as applicable and subsequently extending it to also fit the three-loci model.

Hence, analogous to the two-loci model, the configuration at each locus will be assessed by the corresponding allele frequencies $x^{i_1\bullet\bullet}, x^{\bullet i_2\bullet}, x^{\bullet\bullet i_3}$ with $i_1, i_2, i_3 = 1, \ldots, n-1$, yielding $(n-1)^3$ coordinates; the coefficients of linkage disequilibrium D^{ij} are transferred into $3(n-1)^2$ *coefficients of generalised 2-linkage disequilibrium* $D_2^{i_1 i_2\bullet}, D_2^{i_1\bullet i_3}, D_2^{\bullet i_2 i_3}$, $i_1, i_2, i_3 = 1, \ldots, n-1$ with

$$\begin{aligned} D_2^{i_1 i_2\bullet}(x) &:= x^{i_1\bullet\bullet}x^{\bullet i_2\bullet} - x^{i_1 i_2\bullet}, \\ D_2^{i_1\bullet i_3}(x) &:= x^{i_1\bullet\bullet}x^{\bullet\bullet i_3} - x^{i_1\bullet i_3}, \\ D_2^{\bullet i_2 i_3}(x) &:= x^{\bullet i_2\bullet}x^{\bullet\bullet i_3} - x^{\bullet i_2 i_3}, \end{aligned} \tag{3.96}$$

measuring the linkage disequilibrium with respect to any pair of loci (the corresponding twofold interactions are structurally analogous to those of the two-loci model, giving rise to the notion of 'generalised 2-linkage'). The structure observed in the equations (3.96) coincides with those of the coefficients of linkage disequilibrium for masks $D^{m,i}$ defined in equation (3.54) resp. (3.58) and (3.75) with the argument being all schemas of order 2 (cf. section 3.4.1) and the corresponding mask $m \in M_2$ bisecting the two defined loci of the given schema.

However, $(x^{\bullet\bullet}, D_2)$ does not yet form a full set of coordinates nor is linkage between more than two loci (threefold interactions) taken into account. For this reason, we introduce the *coefficients of generalised 3-linkage disequilibrium* by extending the structure (3.96) into

$$D_3^{i_1 i_2 i_3}(x) := x^{i_1\bullet\bullet}x^{\bullet i_2\bullet}x^{\bullet\bullet i_3} - x^{i_1 i_2 i_3} \quad \text{for } i_1, i_2, i_3 = 1, \ldots, n-1, \tag{3.97}$$

now taking into account all three loci and hence employing the product of all corresponding allele frequencies, which is no longer related to any mask recombination event. By adding these $(n-1)^3$ coefficients as coordinates, $(x^{\bullet\bullet}, D_2, D_3)$ now forms a complete set of $n^3 - 1$ coordinates, which – as will turn out – renders a suitable description of the linkage equilibrium.

Thus, on changing coordinates from (x) to $(x^{\bullet\bullet}, D_2, D_3)$, which implies

$$\frac{\partial x^{i_1\bullet\bullet}}{\partial x^{j_1j_2j_3}} = \delta^{i_1}_{j_1}, \ \frac{\partial x^{\bullet i_2\bullet}}{\partial x^{j_1j_2j_3}} = \delta^{i_2}_{j_2}, \ \frac{\partial x^{\bullet\bullet i_3}}{\partial x^{j_1j_2j_3}} = \delta^{i_3}_{j_3}, \ \frac{\partial D_2^{i_1i_2\bullet}}{\partial x^{j_1j_2j_3}} = -\delta^{i_1}_{j_1}\delta^{i_2}_{j_2} + \delta^{i_1}_{j_1}x^{\bullet i_2\bullet} + x^{i_1\bullet\bullet}\delta^{i_2}_{j_2},$$
$$\frac{\partial D_2^{i_1\bullet i_3}}{\partial x^{j_1j_2j_3}} = -\delta^{i_1}_{j_1}\delta^{i_3}_{j_3} + \delta^{i_1}_{j_1}x^{\bullet\bullet i_3} + x^{i_1\bullet\bullet}\delta^{i_3}_{j_3}, \quad \frac{\partial D_2^{\bullet i_2i_3}}{\partial x^{j_1j_2j_3}} = -\delta^{i_2}_{j_2}\delta^{i_3}_{j_3} + \delta^{i_2}_{j_2}x^{\bullet\bullet i_3} + x^{\bullet i_2\bullet}\delta^{i_3}_{j_3},$$
$$\frac{\partial D_3^{i_1i_2i_3}}{\partial x^{j_1j_2j_3}} = -\delta^{i_1}_{j_1}\delta^{i_2}_{j_2}\delta^{i_3}_{j_3} + \delta^{i_1}_{j_1}x^{\bullet i_2\bullet}x^{\bullet\bullet i_3} + x^{i_1\bullet\bullet}\delta^{i_2}_{j_2}x^{\bullet\bullet i_3} + x^{i_1\bullet\bullet}x^{\bullet i_2\bullet}\delta^{i_3}_{j_3},$$

we may also transform the (inverse) metric given by the coefficient matrix of the 2nd order derivatives in the corresponding Kolmogorov equation (3.57), i. e. $(a^{ij}(x)) = \big(x^i(\delta^i_j - x^j)\big)$ with $i = i_1i_2i_3$ and $j = j_1j_2j_3$, by which we obtain at first

$$\begin{aligned} a^{x^{i_1\bullet\bullet}, x^{\bullet j_2\bullet}}(x^{\bullet\bullet}, D_2, D_3) &= D_2^{i_1j_2\bullet}, \\ a^{x^{i_1\bullet\bullet}, x^{\bullet\bullet j_3}}(x^{\bullet\bullet}, D_2, D_3) &= D_2^{i_1\bullet j_3}, \\ a^{x^{\bullet i_2\bullet}, x^{\bullet\bullet j_3}}(x^{\bullet\bullet}, D_2, D_3) &= D_2^{\bullet i_2j_3} \end{aligned} \tag{3.98}$$

and

$$\begin{aligned} a^{x^{i_1\bullet\bullet}, x^{j_1\bullet\bullet}}(x^{\bullet\bullet}, D_2, D_3) &= x^{i_1\bullet\bullet}(\delta^{i_1}_{j_1} - x^{j_1\bullet\bullet}), \\ a^{x^{\bullet i_2\bullet}, x^{\bullet j_2\bullet}}(x^{\bullet\bullet}, D_2, D_3) &= x^{\bullet i_2\bullet}(\delta^{i_2}_{j_2} - x^{\bullet j_2\bullet}), \\ a^{x^{\bullet\bullet i_3}, x^{\bullet\bullet j_3}}(x^{\bullet\bullet}, D_2, D_3) &= x^{\bullet\bullet i_3}(\delta^{i_3}_{j_3} - x^{\bullet\bullet j_3}) \end{aligned} \tag{3.99}$$

as well as

$$\begin{aligned} a^{x^{i_1\bullet\bullet}, D_2^{j_1j_2\bullet}}(x^{\bullet\bullet}, D_2, D_3) &= \sum_{k_1,k_2,k_3}\sum_{l_1,l_2,l_3} x^{k_1k_2k_3}(\delta^{k_1k_2k_3}_{l_1l_2l_3} - x^{l_1l_2l_3})\frac{\partial x^{i_1\bullet\bullet}}{\partial x^{k_1k_2k_3}}\frac{\partial D_2^{j_1j_2\bullet}}{\partial x^{l_1l_2l_3}} \\ &= \sum_{k_2,k_3,l_3} x^{i_1k_2k_3}\Big(-\delta^{i_1k_2k_3}_{j_1j_2l_3} + x^{j_1j_2l_3} + \sum_{l_2}(\delta^{i_1k_2k_3}_{j_1l_2l_3} - x^{j_1l_2l_3})x^{\bullet j_2\bullet} \\ &\quad + \sum_{l_1}(\delta^{i_1k_2k_3}_{l_1j_2l_3} - x^{l_1j_2l_3})x^{j_1\bullet\bullet}\Big) \\ &= \delta^{i_1}_{j_1}D_2^{j_1j_2\bullet} - x^{i_1\bullet\bullet}D_2^{j_1j_2\bullet} - x^{j_1\bullet\bullet}D_2^{i_1j_2\bullet} \\ &\equiv a^{D_2^{j_1j_2\bullet}, x^{i_1\bullet\bullet}}(x^{\bullet\bullet}, D_2, D_3) \end{aligned} \tag{3.100}$$

and analogously

$$
\begin{aligned}
a^{x^{i_1\bullet\bullet},D_2^{j_1\bullet j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{j_1\bullet j_3},x^{i_1\bullet\bullet}}(x^{\bullet\bullet},D_2,D_3) = \delta^{i_1}_{j_1}D_2^{j_1\bullet j_3} - x^{i_1\bullet\bullet}D_2^{j_1\bullet j_3} - x^{j_1\bullet\bullet}D_2^{i_1\bullet j_3},\\
a^{x^{\bullet i_2\bullet},D_2^{j_1j_2\bullet}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{j_1j_2\bullet},x^{\bullet i_2\bullet}}(x^{\bullet\bullet},D_2,D_3) = \delta^{i_2}_{j_2}D_2^{j_1j_2\bullet} - x^{\bullet i_2\bullet}D_2^{j_1j_2\bullet} - x^{\bullet j_2\bullet}D_2^{j_1i_2\bullet},\\
a^{x^{\bullet i_2\bullet},D_2^{\bullet j_2j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{\bullet j_2j_3},x^{\bullet i_2\bullet}}(x^{\bullet\bullet},D_2,D_3) = \delta^{i_2}_{j_2}D_2^{\bullet j_2j_3} - x^{\bullet i_2\bullet}D_2^{\bullet j_2j_3} - x^{\bullet j_2\bullet}D_2^{\bullet i_2j_3},\\
a^{x^{\bullet\bullet i_3},D_2^{j_1\bullet j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{j_1\bullet j_3},x^{\bullet\bullet i_3}}(x^{\bullet\bullet},D_2,D_3) = \delta^{i_3}_{j_3}D_2^{j_1\bullet j_3} - x^{\bullet\bullet i_3}D_2^{j_1\bullet j_3} - x^{j_1\bullet\bullet}D_2^{j_1\bullet i_3},\\
a^{x^{\bullet\bullet i_3},D_2^{\bullet j_2j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{\bullet j_2j_3},x^{\bullet\bullet i_3}}(x^{\bullet\bullet},D_2,D_3) = \delta^{i_3}_{j_3}D_2^{\bullet j_2j_3} - x^{\bullet\bullet i_3}D_2^{\bullet j_2j_3} - x^{\bullet\bullet j_3}D_2^{\bullet j_2i_3}
\end{aligned}
\tag{3.101}
$$

with all indices $i_1, i_2, i_3, j_1, j_2, j_3$ running from 1 to $n-1$. This all corresponds to the results for the two-loci model analysed in the preceding section. Thus again, when all coefficients of generalised 2-linkage disequilibrium vanish, the coordinate representation of the inverse metric has (at least partially) the structure of a block matrix as all components $(a^{x^{\bullet\bullet},D_2})$ vanish and $(a^{x^{\bullet\bullet},x^{\bullet\bullet}})$ is a block matrix itself.

However, this condition (perceivable as linkage equilibrium with respect to all pairs of loci) does not yet yield a product metric on the corresponding restriction of the state space as the coordinate representation of the inverse metric is not yet entirely a block matrix: When calculating the metric components for allele frequencies and coefficients of generalised 2-linkage disequilibrium – which do not share any defined locus and hence refer to all three loci e. g. $x^{i_1\bullet\bullet}$ and $D_2^{\bullet j_2j_3}$ – these entities do not yet vanish, rather the coefficients of generalised 3-linkage disequilibrium (in the given example $D_3^{i_1j_2j_3}$) come in. We have

$$
\begin{aligned}
a^{x^{i_1\bullet\bullet},D_2^{\bullet j_2j_3}}(x^{\bullet\bullet},D_2,D_3) &= \sum_{k_1,k_2,k_3}\sum_{l_1,l_2,l_3} x^{k_1k_2k_3}(\delta^{k_1k_2k_3}_{l_1l_2l_3} - x^{l_1l_2l_3})\frac{\partial x^{i_1\bullet\bullet}}{\partial x^{k_1k_2k_3}}\frac{\partial D_2^{\bullet j_2j_3}}{\partial x^{l_1l_2l_3}}\\
&= \sum_{k_2,k_3,l_1} x^{i_1k_2k_3}\Big(-\delta^{i_1k_2k_3}_{l_1j_2j_3} + x^{l_1j_2j_3} + \sum_{l_2}(\delta^{i_1k_2k_3}_{l_1l_2j_3} - x^{l_1l_2j_3})x^{\bullet j_2\bullet}\\
&\quad + \sum_{l_3}(\delta^{i_1k_2k_3}_{l_1j_2l_3} - x^{l_1j_2l_3})x^{\bullet\bullet j_3}\Big)\\
&= D_3^{i_1j_2j_3} - x^{i_1\bullet\bullet}D_2^{\bullet j_2j_3} - x^{\bullet j_2\bullet}D_2^{i_1\bullet j_3} - x^{\bullet\bullet j_3}D_2^{i_1j_2\bullet} \qquad (3.102)\\
&\equiv a^{D_2^{\bullet j_2j_3},x^{i_1\bullet\bullet}}(x^{\bullet\bullet},D_2,D_3)
\end{aligned}
$$

and likewise

$$\begin{aligned}
a^{x^{\bullet i_2 \bullet}, D_2^{j_1 \bullet j_3}}(x^{\bullet\bullet}, D_2, D_3) &\equiv a^{D_2^{j_1 \bullet j_3}, x^{\bullet i_2 \bullet}}(x^{\bullet\bullet}, D_2, D_3) \\
&= D_3^{j_1 i_2 j_3} - x^{j_1 \bullet\bullet} D_2^{\bullet i_2 j_3} - x^{\bullet i_2 \bullet} D_2^{j_1 \bullet j_3} - x^{\bullet\bullet j_3} D_2^{j_1 i_2 \bullet}, \quad (3.103) \\
a^{x^{\bullet\bullet i_3}, D_2^{j_1 j_2 \bullet}}(x^{\bullet\bullet}, D_2, D_3) &\equiv a^{D_2^{j_1 j_2 \bullet}, x^{\bullet\bullet i_3}}(x^{\bullet\bullet}, D_2, D_3) \\
&= D_3^{j_1 j_2 i_3} - x^{j_1 \bullet\bullet} D_2^{\bullet j_2 i_3} - x^{\bullet j_2 \bullet} D_2^{j_1 \bullet i_3} - x^{\bullet\bullet i_3} D_2^{j_1 j_2 \bullet}. \quad (3.104)
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
&a^{x^{i_1 \bullet\bullet}, D_3^{j_1 j_2 j_3}}(x^{\bullet\bullet}, D_2, D_3) \\
&\quad = \sum_{k_1,k_2,k_3} \sum_{l_1,l_2,l_3} x^{k_1 k_2 k_3} (\delta^{k_1 k_2 k_3}_{l_1 l_2 l_3} - x^{l_1 l_2 l_3}) \frac{\partial x^{i_1 \bullet\bullet}}{\partial x^{k_1 k_2 k_3}} \frac{\partial D_3^{j_1 j_2 j_3}}{\partial x^{l_1 l_2 l_3}} \\
&\quad = \sum_{k_2,k_3} x^{i_1 k_2 k_3} \Big(- \delta^{i_1 k_2 k_3}_{j_1 j_2 j_3} + x^{j_1 j_2 j_3} + \sum_{l_2,l_3} (\delta^{i_1 k_2 k_3}_{j_1 l_2 l_3} - x^{j_1 l_2 l_3}) x^{\bullet j_2 \bullet} x^{\bullet\bullet j_3} \\
&\qquad + \sum_{l_1,l_3} (\delta^{i_1 k_2 k_3}_{l_1 j_2 l_3} - x^{l_1 j_2 l_3}) x^{j_1 \bullet\bullet} x^{\bullet\bullet j_3} + \sum_{l_1,l_2} (\delta^{i_1 k_2 k_3}_{l_1 l_2 j_3} - x^{l_1 l_2 j_3}) x^{j_1 \bullet\bullet} x^{\bullet j_2 \bullet} \Big) \qquad (3.105) \\
&\quad = (\delta^{i_1}_{j_1} - x^{i_1 \bullet\bullet}) D_3^{j_1 j_2 j_3} - x^{j_1 \bullet\bullet} (x^{\bullet j_2 \bullet} D_2^{i_1 \bullet j_2} - x^{\bullet\bullet j_3} D_2^{i_1 j_2 \bullet}) \\
&\quad \equiv a^{D_3^{j_1 j_2 j_3}, x^{i_1 \bullet\bullet}}(x^{\bullet\bullet}, D_2, D_3)
\end{aligned}$$

and analogously

$$\begin{aligned}
a^{x^{\bullet i_2 \bullet}, D_3^{j_1 j_2 j_3}}(x^{\bullet\bullet}, D_2, D_3) &\equiv a^{D_3^{j_1 j_2 j_3}, x^{\bullet i_2 \bullet}}(x^{\bullet\bullet}, D_2, D_3) \\
&= (\delta^{i_2}_{j_2} - x^{\bullet i_2 \bullet}) D_3^{j_1 j_2 j_3} - x^{\bullet j_2 \bullet} (x^{j_1 \bullet\bullet} D_2^{\bullet i_2 j_2} - x^{\bullet\bullet j_3} D_2^{j_1 i_2 \bullet}), \\
&\qquad (3.106) \\
a^{x^{\bullet\bullet i_3}, D_3^{j_1 j_2 j_3}}(x^{\bullet\bullet}, D_2, D_3) &\equiv a^{D_3^{j_1 j_2 j_3}, x^{\bullet\bullet i_3}}(x^{\bullet\bullet}, D_2, D_3) \\
&= (\delta^{i_3}_{j_3} - x^{\bullet\bullet i_3}) D_3^{j_1 j_2 j_3} - x^{\bullet\bullet j_3} (x^{j_1 \bullet\bullet} D_2^{\bullet j_2 i_3} - x^{\bullet j_2 \bullet} D_2^{j_1 \bullet i_3}). \\
&\qquad (3.107)
\end{aligned}$$

Thus, the *linkage equilibrium with respect to all three loci* is defined to be a state where all coefficients of generalised 2-linkage and generalised 3-linkage disequilibrium vanish, i. e. both all twofold and all threefold linkage interactions between the loci are in equilibrium. Consequently, we then have for the coordinate representation of the inverse metric

$$(a^{(x^{\bullet\bullet},D_2,D_3)}) = \begin{pmatrix} (a^{x^{\bullet\bullet},x^{\bullet\bullet}}) & 0^{n-1,(n-1)^2} & 0^{n-1,(n-1)^3} \\ 0^{(n-1)^2,n-1} & (a^{D_2,D_2}) & (a^{D_2,D_3}) \\ 0^{(n-1)^3,n-1} & (a^{D_3,D_2}) & (a^{D_3,D_3}) \end{pmatrix} \tag{3.108}$$

with $(a^{x^{\bullet\bullet},x^{\bullet\bullet}})$ being a block matrix itself, i. e.

$$(a^{x^{\bullet\bullet},x^{\bullet\bullet}}) = \begin{pmatrix} (a^{x^{i_1\bullet\bullet},x^{j_1\bullet\bullet}}) & 0^{n-1,n-1} & 0^{n-1,n-1} \\ 0^{n-1,n-1} & (a^{x^{\bullet i_2\bullet},x^{\bullet j_2\bullet}}) & 0^{n-1,n-1} \\ 0^{n-1,n-1} & 0^{n-1,n-1} & (a^{x^{\bullet\bullet i_3},x^{\bullet\bullet j_3}}) \end{pmatrix}, \tag{3.109}$$

and the remaining entries (a^{D_2,D_2}) equalling

$$\begin{aligned} a^{D_2^{i_1i_2\bullet},D_2^{j_1j_2\bullet}}(x^{\bullet\bullet},D_2,D_3) &= (x^{i_1\bullet\bullet}-\delta^{i_1}_{j_1})x^{j_1\bullet\bullet}(x^{\bullet i_2\bullet}-\delta^{i_2}_{j_2})x^{\bullet j_2\bullet}, \\ a^{D_2^{i_1\bullet i_3},D_2^{j_1\bullet j_3}}(x^{\bullet\bullet},D_2,D_3) &= (x^{i_1\bullet\bullet}-\delta^{i_1}_{j_1})x^{j_1\bullet\bullet}(x^{\bullet\bullet i_3}-\delta^{i_3}_{j_3})x^{\bullet\bullet j_3}, \\ a^{D_2^{\bullet i_2i_3},D_2^{\bullet j_2j_3}}(x^{\bullet\bullet},D_2,D_3) &= (x^{\bullet i_2\bullet}-\delta^{i_2}_{j_2})x^{\bullet j_2\bullet}(x^{\bullet\bullet i_3}-\delta^{i_3}_{j_3})x^{\bullet\bullet j_3} \end{aligned} \tag{3.110}$$

as well as

$$\begin{aligned} a^{D_2^{i_1i_2\bullet},D_2^{j_1\bullet j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{j_1\bullet j_3},D_2^{i_1i_2\bullet}}(x^{\bullet\bullet},D_2,D_3) \\ &= -\delta^{i_1}_{j_1}D_3^{i_1i_2i_3}+x^{i_1\bullet\bullet}D_3^{j_1i_2j_3}+x^{j_1\bullet\bullet}D_3^{i_1i_2j_3}, \end{aligned} \tag{3.111}$$

$$\begin{aligned} a^{D_2^{i_1i_2\bullet},D_2^{\bullet j_2j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{\bullet j_2j_3},D_2^{i_1i_2\bullet}}(x^{\bullet\bullet},D_2,D_3) \\ &= -\delta^{i_2}_{j_2}D_3^{i_1i_2i_3}+x^{\bullet i_2\bullet}D_3^{i_1j_2j_3}+x^{\bullet j_2\bullet}D_3^{i_1i_2i_3}, \end{aligned} \tag{3.112}$$

$$\begin{aligned} a^{D_2^{i_1\bullet i_3},D_2^{\bullet j_2j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{\bullet j_2j_3},D_2^{i_1\bullet i_3}}(x^{\bullet\bullet},D_2,D_3) \\ &= -\delta^{i_3}_{j_3}D_3^{i_1i_2i_3}+x^{\bullet\bullet i_3}D_3^{i_1i_2j_3}+x^{\bullet\bullet j_3}D_3^{i_1i_2i_3} \end{aligned} \tag{3.113}$$

and (a^{D_3,D_2}) resp. (a^{D_2,D_3}) being

$$\begin{aligned} a^{D_3^{i_1i_2i_3},D_2^{\bullet j_2j_3}}(x^{\bullet\bullet},D_2,D_3) &\equiv a^{D_2^{\bullet j_2j_3},D_3^{i_1i_2i_3}}(x^{\bullet\bullet},D_2,D_3) \\ &= x^{i_1\bullet\bullet}(x^{\bullet i_2\bullet}-\delta^{i_2}_{j_2})x^{\bullet j_2\bullet}(x^{\bullet\bullet i_3}-\delta^{i_3}_{j_3})x^{\bullet\bullet j_3}, \end{aligned} \tag{3.114}$$

$$\begin{aligned}
a^{D_3^{i_1 i_2 i_3}, D_2^{j_1 \bullet j_3}}(x^{\bullet\bullet}, D_2, D_3) &\equiv a^{D_2^{j_1 \bullet j_3}, D_3^{i_1 i_2 i_3}}(x^{\bullet\bullet}, D_2, D_3) \\
&= (x^{i_1\bullet\bullet} - \delta^{i_1}_{j_1}) x^{j_1\bullet\bullet} x^{\bullet i_2\bullet} (x^{\bullet\bullet i_3} - \delta^{i_3}_{j_3}) x^{\bullet\bullet j_3}, \qquad (3.115) \\
a^{D_3^{i_1 i_2 i_3}, D_2^{j_1 j_2 \bullet}}(x^{\bullet\bullet}, D_2, D_3) &\equiv a^{D_2^{j_1 j_2 \bullet}, D_3^{i_1 i_2 i_3}}(x^{\bullet\bullet}, D_2, D_3) \\
&= (x^{i_1\bullet\bullet} - \delta^{i_1}_{j_1}) x^{j_1\bullet\bullet} (x^{\bullet i_2\bullet} - \delta^{i_2}_{j_2}) x^{\bullet j_2\bullet} x^{\bullet\bullet i_3} \qquad (3.116)
\end{aligned}$$

and eventually (a^{D_3,D_3}) equalling

$$\begin{aligned}
&a^{D_3^{i_1 i_2 i_3}, D_3^{j_1 j_2 j_3}}(x^{\bullet\bullet}, D_2, D_3) \\
&= \begin{cases}
x^{i_1\bullet\bullet} x^{j_1\bullet\bullet} x^{\bullet i_2\bullet} x^{\bullet j_2\bullet} x^{\bullet\bullet i_3} x^{\bullet\bullet j_3} \\
x^{i_1\bullet\bullet} x^{j_1\bullet\bullet} x^{\bullet i_2\bullet} x^{\bullet j_2\bullet} x^{\bullet\bullet i_3} x^{\bullet\bullet j_3} \big(2 - \frac{1}{\delta^{i_1}_{j_1} x^{j_1\bullet\bullet} + \delta^{i_2}_{j_2} x^{\bullet j_2\bullet} + \delta^{i_3}_{j_3} x^{\bullet\bullet j_3}}\big) \\
x^{i_1\bullet\bullet} x^{j_1\bullet\bullet} x^{\bullet i_2\bullet} x^{\bullet j_2\bullet} x^{\bullet\bullet i_3} x^{\bullet\bullet j_3} \big(2 - \frac{x^{i_1\bullet\bullet} + x^{\bullet i_2\bullet}}{x^{i_1\bullet\bullet} x^{\bullet i_2\bullet}} \delta^{i_1 i_2}_{j_1 j_2} \\
\qquad - \frac{x^{i_1\bullet\bullet} + x^{\bullet\bullet i_3}}{x^{i_1\bullet\bullet} x^{\bullet\bullet i_3}} \delta^{i_1 i_3}_{j_1 j_3} - \frac{x^{\bullet i_2\bullet} + x^{\bullet\bullet i_3}}{x^{\bullet i_2\bullet} x^{\bullet\bullet i_3}} \delta^{i_2 i_3}_{j_2 j_3}\big)
\end{cases} \qquad (3.117)
\end{aligned}$$

if either none, one or two of the indices i_1, i_2, i_3 and j_1, j_2, j_3 coincide resp.

$$\begin{aligned}
&a^{D_3^{i_1 i_2 i_3}, D_3^{i_1 i_2 i_3}}(x^{\bullet\bullet}, D_2, D_3) \\
&= x^{i_1\bullet\bullet} x^{\bullet i_2\bullet} x^{\bullet\bullet i_3} (1 - x^{i_1\bullet\bullet} x^{\bullet i_2\bullet} - x^{\bullet i_2\bullet} x^{\bullet\bullet i_3} - x^{i_1\bullet\bullet} x^{\bullet\bullet i_3} + 2 x^{i_1\bullet\bullet} x^{\bullet i_2\bullet} x^{\bullet\bullet i_3}) \qquad (3.118)
\end{aligned}$$

if all three indices coincide.

Thus, the block matrix $(a^{(x^{\bullet\bullet}, D_2, D_3)})$ may be inverted in accordance with equation (3.86), demonstrating the product structure of $(a_{(x^{\bullet\bullet}, D_2, D_3)})$, and we consequently obtain for the three-loci model transferring the assertion of lemma 3.9:

3.10 Lemma. *In linkage equilibrium, for all $n \geq 2$ the corresponding restriction of the state space Δ_{n^3-1} of the diffusion approximation of a 3-loci n-allelic Wright–Fisher model equipped with the Fisher metric of the multinomial distribution is a $(3n-3)$-dimensional manifold and carries the geometric structure of*

$$S^{n-1}_+ \times S^{n-1}_+ \times S^{n-1}_+ \subset S^{3n-1}_+. \qquad (3.119)$$

3.5.3 Outlook and further considerations

It appears plausible that also general models with an arbitrary number of loci and alleles feature a product structure in linkage equilibrium when equipped with their Fisher metric. Clearly, the presented definition of linkage equilibrium in the context of three loci (cf. p. 85) may be extended to any number of loci via defining suitable coefficients of linkage disequilibrium with respect to any tuple of loci. This may be achieved by pursuing the forms (3.96) and (3.97): For a model with $k \geq 2$ loci and $n \geq 2$ alleles at each locus, we introduce the *coefficients of generalised l-linkage disequilibrium* for $2 \leq l \leq k$ by putting

$$D_l^{\langle i_{j_1},\dots,i_{j_l},\bullet\rangle}(x) := \prod_{m=1}^{l} x^{\langle i_{j_m},\bullet\rangle} - x^{\langle i_{j_1},\dots,i_{j_l},\bullet\rangle} \tag{3.120}$$

for $i_{j_1},\dots,i_{j_l} = 1,\dots,n-1$ and every subset $\{j_1,\dots,j_l\} \subset \{1\dots,k\}$ with $j_r \neq j_s$ for $r \neq s$, measuring the l-fold linkage interactions for every subset of l loci.

Taking the $\binom{n}{l}(n-1)^l$ coefficients of generalised l-linkage disequilibrium, $l = 2,\dots,k$ plus the $n^k - 1$ gamete frequencies as coordinates yields a full alternative description of the $k(n-1)$-dimensional model as we have

$$\sum_{l=1}^{k} \binom{k}{l} (n-1)^l = (n-1+1)^k - 1 = (n-1)^k - 1. \tag{3.121}$$

It may be added that – in accordance with the given scheme – we may also formulate coefficients of generalised 1-linkage disequilibrium

$$D_l^{\langle i_{j_1},\bullet\rangle}(x) := \prod_{m=1}^{1} x^{\langle i_{j_m},\bullet\rangle} - x^{\langle i_{j_1},\bullet\rangle} \equiv 0 \quad \text{for all } i_{j_1} = 1,\dots,n \text{ with } j_1 \in \{1,\dots,k\}, \tag{3.122}$$

which though provide no information about the state of the model as each locus is trivially in (full) linkage equilibrium with itself. Instead, the allele frequencies are used as non-interaction coordinates.

Correspondingly to the three-loci case, we may then define the model to be in linkage equilibrium if for all l the coefficients of generalised l-linkage disequilibrium vanish and conjecture that the corresponding restriction of the state space Δ_{n^k-1} when equipped with the Fisher metric of the multinomial distribution carries the geometrical structure of

$$\underbrace{S_+^{n-1} \times \ldots \times S_+^{n-1}}_{k} \subset S_+^{kn-1}. \tag{3.123}$$

To prove this, however, in the light of the bulkiness of the calculations involved, one may strive for more advanced tools; potentially a perception of the coefficients of generalised l-linkage disequilibrium as coefficients of l-tuple interactions and the corresponding stages of linkage equilibria as a hierarchical foliation of the state space as presented in [3] might be helpful.

Other interesting observations in this context include the following: All considerations with respect to linkage equilibria actually do not take into account the recombinational structure of the model (if present at all) as the calculations only relate to the diffusion coefficients of the corresponding Kolmogorov equations of the diffusion approximation of the model – which are, as has been shown previously, independent of recombination. Without recombination, however, the assignment of alleles to loci becomes effectless as a k-loci n-allelic model and a (1-locus) nk-allelic model may be identified.

Conversely, this signifies that – as soon as any loci structure comes in – the concept of linkage between different loci or subsets of loci generates the question of linkage equilibria, which themselves directly relate to the concept of recombination via the common structure of coefficients of linkage disequilibrium. Moreover, the formulation of these coefficients also directly implies the concept of coarse-graining and schemata as introduced in section 3.4.

4 Analytic aspects of the diffusion approximation of the 1-dimensional Wright–Fisher model

Having dealt with various generalisations of recombinational Wright–Fisher models in chapter 3, we now wish to solve the corresponding Kolmogorov equations, i. e. determine the associated dynamics and check whether this provides a complete and valid description of the behaviour of the corresponding model. Finding such a solution analytically, however, can prove to be quite challenging. For this reason, additional features of the models like recombination, mutation or natural selection will not be considered in the remainder, nor any coarse-graining. Without recombination, the assignment of alleles to a certain locus is no longer significant, and hence any Wright–Fisher model may be interpreted as a 1-locus model with a corresponding number of alleles. As a result, we will start our analysis with the simplest case of a 1-locus 2-allelic haploid model.

4.1 The Kolmogorov forward equation

To state the problem, we return to the setting of proposition 2.9. Setting the number of loci at 1 and changing the notation into the form $u\colon (\Delta_1)_\infty \longrightarrow \mathbb{R}$ for the solution and $f\colon \Delta_1 \longrightarrow \mathbb{R}$ for the separately stated initial condition for $t = 0$ thus yields as the Kolmogorov forward equation (cf. section 2.4.1) for the diffusion approximation of the 1-locus 2-allelic Wright–Fisher model[1]

$$\begin{cases} \frac{\partial}{\partial t}u(x,t) = \frac{1}{2}\frac{\partial^2}{(\partial x)^2}\big(x(1-x)u(x,t)\big) & \text{in } (\Delta_1)_\infty = (0,1)\times(0,\infty) \\ u(x,0) = f(x) & \text{in } \Delta_1,\ f\in\mathcal{L}^2(\Delta_1) \end{cases} \tag{4.1}$$

with the regularity requirement $u(\,\cdot\,,t)\in C^2(\Delta_1)$ for each fixed $t>0$ and $u(x,\,\cdot\,)\in C^1((0,\infty))$ for each fixed $x\in\Delta_1$.

Subsequently, we will aim at finding a solution of the given Kolmogorov forward equation and develop a presumably complete solution scheme – from which we will advance to an analogous treatment of the corresponding Kolmogorov backward equation in section 4.2.

4.1.1 The solution scheme by M. Kimura

An explicit solution of the above Kolmogorov forward equation (4.1) was already provided by M. Kimura in [16]. In the following, we will put Kimura's solution strategy into perspective and analyse the obtained solution with special regard to its interaction with the boundary of the given domain as this will turn out to be a key point for our subsequent analysis.

In solving the given problem, Kimura uses a separation ansatz $u(x,t)=v(x)w(t)$, converting the first line of equation (4.1) into

$$\begin{aligned} v(x)\frac{\partial}{\partial t}w(t) &= \frac{1}{2}w(t)\frac{\partial^2}{(\partial x)^2}\big(x(1-x)v(x)\big) \\ \Rightarrow \qquad \frac{1}{w}\frac{\partial}{\partial t}w(t) &= \frac{1}{2v}\frac{\partial^2}{(\partial x)^2}\big(x(1-x)v(x)\big) =: -\lambda \end{aligned} \tag{4.2}$$

with $\lambda\geq 0$ as we have

$$0 \leq \int_0^1 \left(\frac{\partial}{\partial x}\big(x(1-x)v(x)\big)\right)^2 dx = -\int_0^1 x(1-x)v(x)\frac{\partial^2}{(\partial x)^2}\big(x(1-x)v(x)\big)\,dx$$

[1]Since a haploid model is sufficient for the intended setting without recombination, the corresponding equation lacks a factor $\frac{1}{2}$ when compared to those from the diploid model in proposition 2.9. This applies to all further considerations and will not be stated separately.

$$= \frac{\lambda}{2} \int_0^1 \underbrace{x(1-x)v^2(x)}_{\geq 0}\, dx \tag{4.3}$$

for $v \in C^2(\overline{\Delta}_1)$.

The problem then reduces to finding solutions of $\frac{\partial}{\partial t} w = -\lambda w$, which is easily done by putting $w(t) := w(0)e^{-\lambda t}$, and of

$$\frac{1}{2}\frac{\partial^2}{(\partial x)^2}\big(x(1-x)v\big) = -\lambda v(x), \quad x \in \Delta_1. \tag{4.4}$$

We already note that no boundary values on $\partial\Delta_1 = \{0,1\}$ need to be stated for the differential equation (4.4) as the solution finds those itself (this observation will become even more relevant in our analysis of the corresponding higher-dimensional problem in chapter 5).

Expanding equation (4.4) further into

$$x(1-x)\frac{\partial^2}{(\partial x)^2}v(x) + 2(1-2x)\frac{\partial}{\partial x}v(x) + (2\lambda - 2)v(x) = 0, \quad x \in \Delta_1 \tag{4.5}$$

and substituting $z = 1 - 2x$ yields

$$(1-z^2)\frac{\partial^2}{(\partial z)^2}\tilde{v}(z) - 4z\frac{\partial}{\partial z}\tilde{v}(z) + (2\lambda - 2)\tilde{v}(z) = 0, \quad z \in (-1,1) \tag{4.6}$$

with $\tilde{v}(z(x)) = v(x)$ (cf. [16]). A generalised form of this equation is known as the *Gegenbauer differential equation*

$$(1-z^2)\frac{\partial^2}{(\partial z)^2}X(z) - (2\alpha+1)z\frac{\partial}{\partial z}X(z) + n(n+2\alpha)X(z) = 0 \quad \text{for } z \in \mathbb{R} \tag{4.7}$$

with $n \in \mathbb{N}$[2] and $\alpha \in \mathbb{R}$. Its solutions are the *Gegenbauer polynomials* $C_n^\alpha(z)$, which are for example given by the recurrence relation

$$nC_n^\alpha(z) = 2(n+\alpha-1)zC_{n-1}^\alpha(z) - (n+2\alpha-2)C_{n-2}^\alpha(z), \quad n \geq 2 \tag{4.8}$$

[2] Throughout this work, we usually assume $\mathbb{N} \equiv \mathbb{N}_0 = \{0,1,2,\dots\}$ and $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$.

and $C_0^\alpha(z) = 1$ and $C_1^\alpha(z) = 2\alpha z$.

The Gegenbauer polynomials obey an orthogonality relation w. r. t. the product $\langle \cdot, \cdot \rangle_\alpha = \int\limits_{-1}^{+1} (1-z^2)^{\alpha - \frac{1}{2}}\, dz$

$$\left\langle C_n^\alpha, C_m^\alpha \right\rangle_\alpha = \delta_n^m c_n^\alpha \tag{4.9}$$

with $c_n^\alpha = 2^{1-2\alpha}\pi \frac{\Gamma(n+2\alpha)}{(n+\alpha)\Gamma^2(\alpha)\Gamma(n+1)}$.

In order to fit them to the given problem, we have $\alpha = \frac{3}{2}$ and $2\lambda - 2 = n(n+3)$, therefore one puts

$$\lambda_n := \frac{(n+1)(n+2)}{2}, \quad n \in \mathbb{N}. \tag{4.10}$$

The product then takes the form $\langle \cdot, \cdot \rangle_{\frac{3}{2}} = \int\limits_{-1}^{+1} (1-z^2)\, dz$ with the orthogonality relation being

$$\left\langle C_n^{\frac{3}{2}}, C_m^{\frac{3}{2}} \right\rangle_{\frac{3}{2}} = \delta_n^m \frac{4\lambda_n}{2n+3} \tag{4.11}$$

as $\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$ and consequently $c_n^{\frac{3}{2}} = \frac{(n+2)!}{(n+\frac{3}{2})n!} = \frac{2(n+1)(n+2)}{2n+3} = \frac{4\lambda_n}{2n+3}$.

Transforming back to the x coordinate, $C_n^{\frac{3}{2}}(1-2x)$, $n \in \mathbb{N}$ are obtained as solutions of the eigenvalue equation (4.4) and hence

$$C_n^{\frac{3}{2}}(1-2x)e^{-\lambda_n t}, \quad n \in \mathbb{N} \tag{4.12}$$

as solutions of the differential equation in (4.1).

In order to also fulfil the initial condition, one may use that the linear combinations of the Gegenbauer polynomials $C_n^{\frac{3}{2}}(z)$, $n \in \mathbb{N}$ are dense in $C^\infty([-1,1])$ (cf. the Stone–Weierstrass theorem in [1], p. 88) and consequently also in $\mathcal{L}^2(-1,1)$, for which reason the initial condition $f(x)$ – after transforming it into $\tilde{f}(z)$ with $f(x) = \tilde{f}(z(x))$ – may be decomposed in terms of these polynomials, i. e.

$$\tilde{f}(z) = \sum_{n \in \mathbb{N}} \alpha_n C_n^{\frac{3}{2}}(z) \quad \text{with } \alpha_n = \frac{1}{c_n} \langle \tilde{f}, C_n^{\frac{3}{2}} \rangle_{\frac{3}{2}} = \frac{2n+3}{4\lambda_n} \langle \tilde{f}, C_n^{\frac{3}{2}} \rangle_{\frac{3}{2}} \tag{4.13}$$

and consequently

$$u(x,t) := \sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n} \langle \tilde{f}, C_n^{\frac{3}{2}} \rangle_{\frac{3}{2}} C_n^{\frac{3}{2}}(1-2x) e^{-\lambda_n t}, \quad (x,t) \in (\Delta_1)_\infty \tag{4.14}$$

is obtained. Thus, for any given data (i. e. any initial condition), there exists an (explicit) solution of equation (4.1). Moreover, this solution also is the unique solution as will be shown in lemma 4.2 below.

In the special case $f(x) = \delta_p(x)$ for some $p \in \Delta_1$, thus $\tilde{f}(z) = \delta_r(z)$ with $r := 1-2p \in (-1,1)$, one gets

$$\langle \tilde{f}, C_n^{\frac{3}{2}} \rangle_{\frac{3}{2}} = \int_{-1}^{+1} (1-z^2) C_n^{\frac{3}{2}}(z) \delta_r(z)\, dz = (1-r^2) C_n^{\frac{3}{2}}(r) \tag{4.15}$$

and continuing

$$\tilde{f}(z) = \sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n} (1-r^2) C_n^{\frac{3}{2}}(r) C_n^{\frac{3}{2}}(z), \tag{4.16}$$

thus

$$u(x,t) := \sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n} (1-r^2) C_n^{\frac{3}{2}}(r) C_n^{\frac{3}{2}}(1-2x) e^{-\lambda_n t}, \quad (x,t) \in (\Delta_1)_\infty \tag{4.17}$$

as a solution of the original problem (4.1) with initial condition $f = \delta_p$ (cf. [16]).

4.1.2 Properties of Kimura's solution

Concerning the regularity of the obtained solution, we have:

4.1 Lemma. *The solution given in equation* (4.14) *continuously extends to the boundary and is of class C^∞ with respect to $x \in \overline{\Delta}_1$ for $t > 0$.*

Proof. Clearly, all Gegenbauer polynomials (and all their derivatives) continuously extend to the boundary, and furthermore, we have

$$\left\| \frac{\partial^k}{\partial x^k} C_n^{\frac{3}{2}}(1-2x) \right\|_{C^0(\overline{\Delta}_1)} \le \lambda_n^{k+1} \quad \text{for } k, n \ge 0 \tag{4.18}$$

and also

$$\left\|(1-z^2)C_n^{\frac{3}{2}}(z)\right\|_{\mathcal{L}^2(\Delta_1)} \leq \sqrt{\lambda_n} \quad \text{for } n \geq 0. \tag{4.19}$$

Thus consequently for arbitrary $k \geq 0$, we obtain

$$\begin{aligned} &\left\| \sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n} \langle \tilde{f}, C_n^{\frac{3}{2}} \rangle_{\frac{3}{2}} \frac{\partial^k}{\partial x^k} C_n^{\frac{3}{2}}(1-2x) e^{-\lambda_n t} \right\|_{C^0(\overline{\Delta}_1)} \\ &\leq \|f\|_{\mathcal{L}^2(\Delta_1)} \sum_{n\in\mathbb{N}} \lambda_n^{k+\frac{3}{2}} e^{-\lambda_n t} \leq \|f\|_{\mathcal{L}^2(\Delta_1)} \int_0^\infty x^{k+\frac{3}{2}} e^{-xt} dx < \infty \quad \text{for } t > 0, \end{aligned} \tag{4.20}$$

meaning that the series and all its derivatives converge uniformly on $\overline{\Delta}_1$. Analogously, one may prove that the solution also is of class C^∞ with respect to $t \in (0,1)$. □

Moreover, it may be shown that the solution found in equation (4.14) actually (uniquely) solves the given problem:

4.2 Lemma. *For a given initial condition $f \in \mathcal{L}^2(\Delta_1)$, Kimura's solution (cf. equation (4.14)) is the unique solution of the Kolmogorov forward equation (4.1) in Δ_1 within the class of solutions which are square-integrable with respect to x for every $t \geq 0$.*

Proof. As we have uniform convergence in the spatial variable, we may differentiate the series term by term with respect to x; for t, uniform convergence of the termwise differentiated series may be shown similarly as for x. Hence, the series fulfils equation (4.1) as all summands do so by construction.

Furthermore, the property of matching the initial condition f follows directly from the decomposition of f in terms of the Gegenbauer polynomials (which coincides with the corresponding decomposition of $u(\,\cdot\,,0)$ by definition). Since the linear span of the Gegenbauer polynomials is dense in $\mathcal{L}^2([-1,1])$, the equality $\left\langle \tilde{u}(\,\cdot\,,0), C_n^{\frac{3}{2}} \right\rangle_{\frac{3}{2}} = \left\langle \tilde{f}, C_n^{\frac{3}{2}} \right\rangle_{\frac{3}{2}}$ for all $n \in \mathbb{N}$ already implies $\tilde{u}(\,\cdot\,,0) = \tilde{f}$ and hence $u(\,\cdot\,,0) = f$.

To show the uniqueness, assume that $u' \colon (\Delta_1)_\infty \longrightarrow \mathbb{R}$ with $u'(\,\cdot\,,t) \in \mathcal{L}^2(\Delta_1)$ for all $t \in [0,\infty)$ is another solution of equation (4.1). As the linear span of the Gegenbauer polynomials is dense in $\mathcal{L}^2$, we may thus for all $t \geq 0$ decompose $u'(\,\cdot\,,t)$

as

$$u'(x,t) := \sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n} \langle \tilde{u}'(z,t), C_n^{\frac{3}{2}}(z)\rangle_{\frac{3}{2}} C_n^{\frac{3}{2}}(1-2x), \quad x \in \Delta_1. \tag{4.21}$$

The property of u' of solving equation (4.1) then implies

$$\frac{\partial}{\partial t}\langle \tilde{u}'(z,t), C_n^{\frac{3}{2}}(z)\rangle_{\frac{3}{2}} = -\lambda_n \langle \tilde{u}'(z,t), C_n^{\frac{3}{2}}(z)\rangle_{\frac{3}{2}}, \quad t > 0 \tag{4.22}$$

for all $n \in \mathbb{N}$, yielding

$$\langle \tilde{u}'(z,t), C_n^{\frac{3}{2}}(z)\rangle_{\frac{3}{2}} = e^{-\lambda_n t}\langle \tilde{u}'(z,0), C_n^{\frac{3}{2}}(z)\rangle_{\frac{3}{2}} = e^{-\lambda_n t}\langle \tilde{f}, C_n^{\frac{3}{2}}\rangle_{\frac{3}{2}} \tag{4.23}$$

as u' also fulfils the initial condition. Consequently, u' already agrees with u. $\square$

4.3 Remark. In accordance with standard PDE theory (cf. e. g. [14]), any solution $u' \colon (\Delta_1)_\infty \longrightarrow \mathbb{R}$ of the Kolmogorov forward equation (4.1) is sufficiently regular, in particular $u'(\,\cdot\,,t) \in \mathcal{L}^2(\Delta_1)$ for every $t \geq 0$. Hence, we may just say that Kimura's solution (4.14) is the unique solution of equation (4.1).

4.1.3 Moments and an extension scheme to the boundary

Due to the degeneracy of the differential operator at the boundary, i. e. its lack of uniform ellipticity when approaching the boundary, we cannot prescribe any boundary values (in the classical sense), as already stated. Instead, a solution finds boundary values itself as it extends continuously to the boundary (lemma 4.1).

Hence, – starting at some p in the interior Δ_1 – we first have $u(0) = u(1) = 0$ for u as in equation (4.14), while at the course of time both boundary values become positive. For t approaching infinity, the solution vanishes everywhere, including at the boundary. This is of course not the kind of behaviour that one would expect – calculations with the underlying discrete model suggest that the total number of individuals (and the expectation value of the gametic configuration) should be preserved. In particular, we see some accumulation at the boundary (i. e. in the stages of homozygosity) for $t \to \infty$ in the discrete model, which should also be reflected in the continuous model.

More generally, this leads us to a discussion of the moments of the distribution, both in the discrete and in the continuous case (with the numbers of individuals in the discrete case corresponding to the total probability mass of the distribution, i. e. the zeroth moment) and their mutual relation. In [16], Kimura also tackled the issue of moments of the distribution; the following considerations are based on his work.

Thus, if we denote the discrete process with N individuals by $(C(t))_{t\in\mathbb{N}}$ (cf. also section 2.3.3), for its moments

$$\overline{m}_k(t) := \mathrm{E}\left(C(t)^k\right), \quad k, t \in \mathbb{N}, \tag{4.24}$$

we then have

$$\begin{aligned}\mathrm{E}\left(C(t+1)^k \middle| C(t)\right) &= \mathrm{E}\left((C(t)+\delta C)^k \middle| C(t)\right) \\ &= C(t)^k + \sum_{l=1}^{k} \binom{k}{l} C(t)^{k-l}\, \mathrm{E}_1\left((\delta C)^l | C(t)\right).\end{aligned} \tag{4.25}$$

Using $\mathrm{E}_1(\delta C|C(t)) = 0$, $\mathrm{E}_1((\delta C)^2|C(t)) = \frac{1}{N}C(t)(1-C(t))$, $\mathrm{E}_1((\delta C)^l|C(t)) \in \mathcal{O}(N^{-2})$ for $l \geq 3$ (cf. also equations (2.53)–(2.55) accordingly for a haploid model and with $R = 0$), this yields

$$\begin{aligned}\overline{m}_k(t+1) - \overline{m}_k(t) &= \mathrm{E}\left(\mathrm{E}\left(C(t+1)^k \middle| C(t)\right) - C(t)^k\right) \\ &= -\frac{k(k-1)}{2N}\left(\mathrm{E}\left(C(t)^k\right) - \mathrm{E}\left(C(t)^{k-1}\right)\right) + \mathcal{O}(N^{-2})\chi_{\{k\geq 3\}},\end{aligned} \tag{4.26}$$

describing the (expected) incremental change of the moments within 1 time step for $k \geq 1$, whereas $\overline{m}_0(t) \equiv 1$. In particular, the expectation value is thus preserved as is the total number of individuals.

Turning our attention to the diffusion limit process $(X(t))_{t\in\mathbb{R}_+}$, we of course wish to obtain a corresponding description. In order to do so, the natural assumption on the diffusion approximation would be that the moments of the limit process coincide with the limits of the moments of the (suitably time-rescaled) discrete processes

$(\hat{C}_N(t_N))_{t_N \in \mathbb{N}_{N^{-1}}}$ (cf. section 2.4), i. e.

$$\lim_{N\to\infty} \mathrm{E}(\hat{C}_N(t_N)^k) = \mathrm{E}(X(t)^k) \tag{4.27}$$

for $k \in \mathbb{N}$, $t \in \mathbb{R}_+$ and $(t_N)_{N\in\mathbb{N}}$ such that $t_N \to t$ as $N \to \infty$.[3]

Stipulating equation (4.27) and formally defining

$$\bar{\mu}_k(t) := \mathrm{E}(X(t)^k), \quad k \in \mathbb{N},\ t \geq 0, \tag{4.28}$$

for the moments of the limit process, then equation (4.26) – after replacing t by Nt_N with $t_N \in \mathbb{N}_{N^{-1}}$ for $N \in \mathbb{N}$ such that $t_N \to t$ for some $t \in \mathbb{R}_+$ as $N \to \infty$ and dividing by $\delta t = \frac{1}{N}$ – for $N \to \infty$ turns into a differential equation for the moments $\bar{\mu}_k(t)$

$$\frac{\partial}{\partial t}\bar{\mu}_k(t) = -\frac{k(k-1)}{2}\Big(\bar{\mu}_k(t) - \bar{\mu}_{k-1}(t)\Big) \quad \text{for } k \geq 1,\ t > 0, \tag{4.29}$$

the *moments evolution equation* (cf. [16], p. 145). Here, again the expectation value is preserved as is the total mass, which may additionally be formulated as $\frac{\partial}{\partial t}\bar{\mu}_0(t) \equiv 0$.

To proceed, we may now choose a different approach than Kimura in [16] and rather pose the question of how the obtained solution (cf. u in equation (4.14)) can be related to the moments evolution equation, in particular how the moments $\bar{\mu}_k(t)$ may be formulated in terms of the solution.

Preliminarily, one would consider

$$\int_0^1 u(x,t)x^k\,dx, \quad k \in \mathbb{N},\ t \geq 0 \tag{4.30}$$

for $\bar{\mu}_k(t)$, which – however – does not fulfil the conditions stated and hence may not be interpreted as limit of the moments of the discrete processes $\overline{m}_k(t)$. This is rooted in their definition in equation (4.24), which also involves configurations of $C(t)$ where an allele frequency may be zero; in the diffusion limit, this corresponds

[3] This is effectively a stronger condition than the one in equation (2.78) since the expression does not involve any contribution of the boundary $\partial\Delta_{\frac{1}{N}}$, i. e. $\lim_{N\to\infty} N\,\mathrm{E}_{\frac{1}{N}}\big(\delta\hat{C}_N|_{\partial\Delta_{\frac{1}{N}}}\big|\hat{C}_N(t_N) = c_N\big) = 0$ for $c_N \in \Delta_{\frac{1}{N}}$.

to the boundary $\partial\Delta_1$ of the state space of the continuous process Δ_1. For this reason, the definition of $\bar{\mu}_k(t)$ should likewise account for the boundary by extending expression (4.30) in an appropriate way.

Kimura instead solved the moments evolution equation (4.29) explicitly in [16] and was able to construct values u_0 on $\partial\Delta_1$ from this solution, which were interpreted as probabilities for the process to have reached $\{0\}$ resp. $\{1\}$. Thus, in terms of the probability density function of the process, these values may be viewed as point masses on the boundary. Explicitly, Kimura obtained (cf. [16], p. 146)

$$u_0(0,t) = p - \frac{1}{2}\sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)e^{-\lambda_n t} \tag{4.31}$$

and

$$u_0(1,t) = 1 - p - \frac{1}{2}\sum_{n\in\mathbb{N}}(-1)^n \frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)e^{-\lambda_n t} \tag{4.32}$$

as boundary probabilities.

Our further analysis reveals that these values directly show some strong connection to the solution given; indeed, we may formulate them in terms of the solution u, i. e. more precisely in terms of the (outward normal) flux $G_u^\perp$ of the continuous extension of the solution to the boundary. Generally, the *flux* $G_u\colon (\Delta_1)_\infty \longrightarrow \mathbb{R}$ of a solution $u\colon (\Delta_1)_\infty \longrightarrow \mathbb{R}$ of equation (4.1) is defined such that the continuity equation

$$\frac{\partial}{\partial t}u(x,t) = -\frac{\partial}{\partial x}G_u(x,t), \quad (x,t) \in (\Delta_1)_\infty. \tag{4.33}$$

is satisfied. Since u satisfies $\frac{\partial}{\partial t}u(x,t) = \frac{1}{2}\frac{\partial^2}{(\partial x)^2}\big(x(1-x)u(x,t)\big)$, we thus have

$$G_u(x,t) = -\frac{1}{2}\frac{\partial}{\partial x}\big(x(1-x)u(x,t)\big), \quad (x,t) \in \big(\Delta_1\big)_\infty. \tag{4.34}$$

This concept naturally extends also to the boundary $\partial\Delta_1 = \{0,1\}$ if u is extendable there with the extension being of class C^2 with respect to $x \in \overline{\Delta}_1$, which is the case for Kimura's solution by lemma 4.1. We thus have for the outward normal component of the flux $G_u^\perp := G_u \cdot \nu$ on the boundary (with the extension of u also denoted by u)

$$G_u^{\perp}(x,t) = \mp\frac{1}{2}(1-2x)u(x,t) + x(1-x)\frac{\partial}{\partial x}u(x,t) = \frac{1}{2}u(x,t), \quad (x,t) \in \big(\partial\Delta_1\big)_\infty. \tag{4.35}$$

Kimura's boundary values may then be formulated as follows:

4.4 Lemma. *For $u_0\colon (\partial\Delta_1)_\infty \longrightarrow \mathbb{R}$ as in equation (4.31) and the continuous extension of Kimura's solution $u(x,t)$ to $(\partial\Delta_1)_\infty$, we have*

$$u_0(x,t) = \int_0^t G_u^{\perp}(x,\tau)\,d\tau = \frac{1}{2}\int_0^t u(x,\tau)\,d\tau, \quad (x,t) \in (\partial\Delta_1)_\infty. \tag{4.36}$$

Proof. We have

$$\begin{aligned}\frac{1}{2}\int_0^t u(x,\tau)\,d\tau &= \frac{1}{2}\int_0^t \sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)C_n^{\frac{3}{2}}(1-2x)e^{-\lambda_n\tau}\,d\tau \\ &= \frac{1}{2}\sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)C_n^{\frac{3}{2}}(1-2x)\big(\lambda_n^{-1} - \lambda_n^{-1}e^{\lambda_n t}\big),\end{aligned} \tag{4.37}$$

wherein we may first put $\varepsilon > 0$ as lower integration limit to assure uniform convergence and then let $\varepsilon \searrow 0$. Now taking into account $C_n^{\frac{3}{2}}(1-2x)\big|_{\{x=1\}} = (-1)^n\lambda_n$ for $n \in \mathbb{N}$, we obtain for $x = 1$

$$\frac{1}{2}\int_0^t u(1,\tau)\,d\tau = p - \frac{1}{2}\sum_{n\in\mathbb{N}}(-1)^n\frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)e^{-\lambda_n t} = u_0(1,t) \tag{4.38}$$

as we have with $\frac{(-1)^n}{2} = \int_0^1 xC_n^{\frac{3}{2}}(1-2x)\,dx$ for $n \in \mathbb{N}$

$$\begin{aligned}&\frac{1}{2}\sum_{n\in\mathbb{N}}\frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)(-1)^n \\ &= \int_0^1 x\sum_{n\in\mathbb{N}}\frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)C_n(1-2x)\,dx = \int_0^1 x\delta_p(x)\,dx = p.\end{aligned} \tag{4.39}$$

Similarly, by using $C_n^{\frac{3}{2}}(1-2x)\big|_{\{x=0\}} = \lambda_n$ for all n and exploiting the symmetry of the Gegenbauer polynomials, i. e. $C_n^{\frac{3}{2}}(r) = (-1)^n C_n^{\frac{3}{2}}(-r)$, we also get

$$\frac{1}{2}\int_0^t u(0,\tau)\,d\tau = 1 - p - \frac{1}{2}\sum_{n\in\mathbb{N}} \frac{2n+3}{4\lambda_n}(1-r^2)C_n^{\frac{3}{2}}(r)e^{-\lambda_n t} = u_0(0,t). \tag{4.40}$$

□

As may be seen from the proof, $1-p$ and p are the probabilities to eventually reach (and end up in) the boundary points 0 resp. 1, while the series express the probabilities to do so not earlier than at time t.

An important observation now is that Kimura's boundary values exactly fit the solution in the interior in the sense that the total mass of the probability density function of the process is preserved, thus

$$u_0(0,t) + \int_0^1 u(x,t)\,dx + u_0(1,t) = 1 \quad \text{for all } t \geq 0 \tag{4.41}$$

as Kimura showed in [16]. It is straightforward to check that for this construction an analogous assertion about the expectation value also holds, i. e.

$$\int_0^1 xu(x,t)\,dx + u_0(1,t) = p \quad \text{for all } t \geq 0. \tag{4.42}$$

Thus, already the 0th and the 1st moment evolve as described by equation (4.29).

Hence, this so far successful integration of the boundary leads us to the introduction of an accordingly refined concept of solution: If we explicitly include the boundary into the description of the process with specifically prescribed values there, we may define an *extended solution*[4]

$$\hat{u}\colon (\overline{\Delta}_1)_\infty \longrightarrow \mathbb{R}, \quad \hat{u}(x,t) := u(x,t)\chi_{\Delta_1}(x) + u_0(x,t)\chi_{\partial\Delta_1}(x) \tag{4.43}$$

[4]Likewise, also an extended initial condition $f\colon \overline{\Delta}_1 \longrightarrow \mathbb{R}$ might be prescribed, signifying that the process carries additional boundary weight right from the beginning, which, however, evolves exactly as a 0-dimensional process (i. e. stays constant) and is without further effect on the model. For this reason, we will usually assume $f|_{\partial\Delta_1} \equiv 0$ or that f is extended that way if it is only given on Δ_1.

of equation (4.1) in the sense that $\hat{u}|_{\Delta_1}$ solves equation (4.1), while $\hat{u}$ describes the behaviour of the process both in the interior and on the boundary; its moments may correspondingly be defined as

$$\bar{\mu}_k(t) = \mathrm{E}\left(X(t)^k\right) := \int_{\Delta_1} u(x,t)x^k\,dx + u_0(x,t)x^k\chi_{\partial\Delta_1}(x), \quad k \in \mathbb{N},\ t \geq 0, \tag{4.44}$$

extending the preliminary conception of expression (4.30). We note that this new formulation involves an integration adapted to the varying dimensionality of the domain of the process and its boundary, i. e. the additional boundary values are interpreted as 0-dimensional masses (this concept will be further developed and generalised in chapter 5).

Now, for $\hat{u}$ and its moments as defined in equation (4.44), the aforementioned conservation laws read:

$$\bar{\mu}_0(t) = \int\limits_{\Delta_1} u(x,t)\,dx + u_0(x,t)\chi_{\partial\Delta_1}(x) = 1 \;\text{ and }\; \bar{\mu}_1(t) = \int\limits_{\Delta_1} u(x,t)x\,dx + u_0(1,t) = p \tag{4.45}$$

for all $t \geq 0$. Furthermore, this formulation of moments in equation (4.44) likewise yields *all* 'right' moments of the process in the sense that their evolution is as to be expected by the underlying discrete model, confirming the choice for the moments in equation (4.44); more generally, based upon Kimura's work [16] and the assumption on the diffusion approximation that the moments of the limit process match the limit of those of the underlying discrete model (cf. equation (4.27)), we may then formulate the following result:

4.5 Proposition. *For a given initial condition $f \in \mathcal{L}^2(\Delta_1)$, the Kolmogorov forward equation corresponding to the diffusion approximation of the 1-dimensional Wright–Fisher model* (4.1) *always allows a unique extended solution $\hat{u}\colon \left(\overline{\Delta}_1\right)_\infty \longrightarrow \mathbb{R}$ in the sense that $\hat{u}|_{\Delta_1}$ solves equation* (4.1) *and that its moments $\bar{\mu}_k(t) = \int_{\Delta_1} \hat{u}(x,t)x^k\,dx + \hat{u}(x,t)x^k\chi_{\partial\Delta_1}(x)$, $k \in \mathbb{N}, t \geq 0$ (cf. equation* (4.44)*) fulfil the moments evolution equation* (4.29)*.*

Proof. What remains to be shown is that for the extended solution $\hat{u}$ as defined in equation (4.43) and u as in equation (4.14) and u_0 as in lemma 4.4 *all* moments $\bar{\mu}_k(t)$

evolve in accordance with the moments equation (4.29). For arbitrary $\varphi \in C^\infty(\overline{\Delta}_1)$ we get by integrating by parts twice (with $u|_{\partial\Delta_1}$ denoting the continuous extension of u in opposition to the 'proper' boundary value u_0 and G_u being the flux of u (cf. equation (4.34))):

$$\begin{aligned}\int_{\Delta_1} \frac{\partial^2}{(\partial x)^2}\big(x(1-x)u\big)\varphi\,dx &= \frac{\partial}{\partial x}\big(x(1-x)u\big)\varphi - x(1-x)u\frac{\partial}{\partial x}\varphi\Big|_0^1 \\ &\qquad + \int_{\Delta_1} ux(1-x)\frac{\partial^2}{(\partial x)^2}\varphi\,dx \\ &= -2G_u(x,t)\varphi\big|_0^1 + \int_{\Delta_1} ux(1-x)\frac{\partial^2}{(\partial x)^2}\varphi\,dx. \end{aligned} \tag{4.46}$$

Consequently, for $\hat{u}$ we obtain (cf. lemma 4.4)

$$\begin{aligned}\int_{\Delta_1} u_t\varphi\,dx + (u_0)_t\varphi\chi_{\partial\Delta_1} &= \int_{\Delta_1} \frac{1}{2}\frac{\partial^2}{(\partial x)^2}\big(x(1-x)u\big)\varphi\,dx + G_u^\perp\varphi\chi_{\partial\Delta_1} \\ &= \int_{\Delta_1} \frac{1}{2}ux(1-x)\frac{\partial^2}{(\partial x)^2}\varphi\,dx. \end{aligned} \tag{4.47}$$

In particular for $\varphi = x^k$, $k \geq 2$, we have

$$\frac{1}{2}x(1-x)\frac{\partial^2}{(\partial x)^2}x^k = -\frac{k(k-1)}{2}\big(x^k - x^{k-1}\big), \tag{4.48}$$

while for $\varphi = 1$ and $\varphi = x$ the right-hand side vanishes altogether. Hence, the moment formula (4.29) for $\bar{\mu}_k(t)$, $k \in \mathbb{N}$ follows.

Conversely, if the moments evolution equation is assumed to hold for the moments corresponding to some $\hat{u}'\colon (\overline{\Delta}_1)_\infty \longrightarrow \mathbb{R}$ with $\hat{u}'|_{\Delta_1}$ solving equation (4.1), then we immediately have $\hat{u}' \equiv \hat{u}$ on Δ_1 by lemma 4.2 and consequently also on $\overline{\Delta}_1$ by the moments equation. □

Thus, the extended solution $\hat{u}$, which is piecewise defined both in the interior and on the boundary, yields a complete account of the dynamics of the process. For $t \to \infty$, the interior solution uniformly vanishes in Δ_1 (for a proof cf. the analogous lemma 4.7), and we obtain as limit for $t = \infty$

$$\hat{u}(x,\infty) = (1-p)\chi_{\{0\}}(x) + p\chi_{\{1\}}(x), \quad x \in \overline{\Delta}_1, \tag{4.49}$$

which may either be derived from the explicit formulae for u_0 (cf. equation (4.31) f.) or as a consequence of the moments evolution equation (4.29). Consequently, all mass is eventually accumulated at the boundary points as desired.

4.2 The Kolmogorov backward equation

After having obtained a solution of the Kolmogorov forward equation for the 1-dimensional Wright–Fisher model including a corresponding extension scheme for the boundary of the domain, we now wish to treat the corresponding Kolmogorov backward equation (cf. section 2.4.1) by similar strategies. The equation again may be stated by making use of proposition 2.9 applied to the current setting of 1 locus and 2 alleles and separately stating the final condition f for $t = 0$, which corresponds to a certain (generalised) target set. At first again only considering the interior Δ_1, this yields

$$\begin{cases} -\frac{\partial}{\partial t}u(p,t) = \frac{1}{2}p(1-p)\frac{\partial^2}{(\partial p)^2}u(p,t) & \text{in } (\Delta_1)_{-\infty} = \Delta_1 \times (-\infty, 0) \\ u(p,0) = f(p) & \text{in } \Delta_1,\ f \in \mathcal{L}^2(\Delta_1) \end{cases} \tag{4.50}$$

for $u(\,\cdot\,,t) \in C^2(\Delta_1)$ for each fixed $t \in (-\infty, 0)$ and $u(x,\,\cdot\,) \in C^1((-\infty,0))$ for each fixed $x \in \Delta_1$. Note that this equation is given in terms of the initial value p and that the time parameter t now is negative.

4.2.1 A solution scheme by Gegenbauer polynomials

The procedure described in paragraph 4.1.1 may be analogously applied to the Kolmogorov backward equation (4.50) in Δ_1, for the notation used see also there. The separation ansatz yields

$$p(1-p)\frac{\partial^2}{(\partial p)^2}v(p) + 2\kappa v(p) = 0, \quad p \in \Delta_1 \tag{4.51}$$

as a differential equation in p, which needs to be solved for the eigenvalues κ. Again changing coordinates by dint of $r = 1 - 2p$, this equation reads

$$(1 - r^2)\frac{\partial^2}{(\partial r)^2}\tilde{v}(r) + 2\kappa\tilde{v}(r) = 0. \tag{4.52}$$

Again, the Gegenbauer polynomials $C_n^{\alpha}(r)$, $n \geq 0$ (cf. p. 93) emerge as solutions, requiring $\alpha = -\frac{1}{2}$ and $2\kappa = n(n + 2\alpha)$. Hence, we put

$$\kappa_n := \frac{n(n-1)}{2}, \quad n \in \mathbb{N}, \tag{4.53}$$

whereas the product takes the form $\langle \cdot, \cdot \rangle_{-\frac{1}{2}} = \int\limits_{-1}^{+1} \frac{1}{1-r^2}\, dr$, and the orthogonality relation is

$$\left\langle C_n^{-\frac{1}{2}}, C_m^{-\frac{1}{2}} \right\rangle_{-\frac{1}{2}} = \delta_n^m \frac{1}{2n-1}\frac{1}{\kappa_n} \quad \text{for } n, m \geq 2 \tag{4.54}$$

as $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$, and thus $c_n^{-\frac{1}{2}} = \frac{(n-2)!}{(n-\frac{1}{2})n!} = \frac{1}{2n-1}\frac{2}{n(n-1)} = \frac{1}{2n-1}\frac{1}{\kappa_n}$.

Consequently, we obtain $C_n^{-\frac{1}{2}}(1 - 2p)$, $n \in \mathbb{N}$ as solutions of the eigenvalue equation (4.51) and hence

$$C_n^{-\frac{1}{2}}(1 - 2p)e^{\kappa_n t}, \quad n \in \mathbb{N} \tag{4.55}$$

as solutions of the differential equation in (4.50).

Note that the Gegenbauer polynomials appearing in the forward equation and those for the backward equation are linked through

$$p(1-p)C_n^{\frac{3}{2}}(1 - 2p) = \frac{1}{4}(n + 1)(n + 2)C_{n+2}^{-\frac{1}{2}}(1 - 2p) \quad \text{for } n \geq 0, \tag{4.56}$$

which is a consequence of the relation between the forward operator and the backward operator (cf. also the more general lemma 5.1); accordingly, we have $\lambda_n \equiv \kappa_{n+2}$. By equation (4.56), all $C_n^{-\frac{1}{2}}(1 - 2p)$ vanish on $\partial\Delta_1$ for $n \geq 2$.

Next, we wish to decompose the final condition $f(p)$ – after transforming it into $\tilde{f}(r)$ with $f(p) = \tilde{f}(r(p))$ – in terms of the Gegenbauer polynomials, thus obtaining a series

$$\tilde{f}(r) = \sum_{n\in\mathbb{N}} \beta_n C_n^{-\frac{1}{2}}(r) \tag{4.57}$$

with

$$\beta_n = \frac{1}{c_n}\langle \tilde{f}, C_n^{-\frac{1}{2}}\rangle_{-\frac{1}{2}} = (2n-1)\kappa_n \langle \tilde{f}, C_n^{-\frac{1}{2}}\rangle_{-\frac{1}{2}}. \tag{4.58}$$

However, some care is required here as for $n = 0$ and $n = 1$ the coefficients $\langle \tilde{f}, C_n^{-\frac{1}{2}}\rangle_{-\frac{1}{2}}$ do not need to exist for all f; on top of that, the eigenvalue κ_n is zero. Nevertheless, this may be ignored for the moment: Since the linear combinations of $(C_n^{\frac{3}{2}})_{n\geq 0}$ are dense in $C^\infty([-1,1])$, the linear span of $(C_n^{-\frac{1}{2}})_{n\geq 2}$ is already dense in $C_c^\infty([-1,1])$ by equation (4.56) (for a definition of C_c^∞ cf. equation (5.18) on p. 120) and consequently also in $\mathcal{L}^2((-1,1))$. Hence, it suffices to decompose $\tilde{f}$ in terms of the $C_n^{-\frac{1}{2}}$ with $n \geq 2$, yielding

$$f(p) := \sum_{n\geq 2}(2n-1)\kappa_n \langle \tilde{f}, C_n^{-\frac{1}{2}}\rangle_{-\frac{1}{2}} \tag{4.59}$$

and consequently

$$u(p,t) := \sum_{n\geq 2}(2n-1)\kappa_n \langle \tilde{f}, C_n^{-\frac{1}{2}}\rangle_{-\frac{1}{2}} C_n^{-\frac{1}{2}}(1-2p)e^{\kappa_n t}, \quad (p,t)\in(\Delta_1)_{-\infty} \tag{4.60}$$

as solution of equation (4.50). Its convergence and regularity (i. e. being of class C^∞ with respect to $p \in \Delta_1$) may be proven similarly to the forward case in section 4.1.2 using

$$\left\| \frac{\partial^k}{\partial p^k} C_n^{-\frac{1}{2}}(1-2p) \right\|_{C^0(\overline{\Delta}_1)} \leq \kappa_n^{k+1} \quad \text{for } k\geq 0,\, n\geq 2 \tag{4.61}$$

and

$$\left\| \frac{1}{1-r^2} C_n^{-\frac{1}{2}}(r) \right\|_{\mathcal{L}^2(\Delta_1)} \leq 1 \quad \text{for } n \geq 2. \tag{4.62}$$

However, for a given final condition f, in general this is not the unique solution as is already observable at its decomposition in terms of eigenfunctions in equation (4.59): For example, for $f = 1 \in \mathcal{L}^2(\Delta_1)$ as final condition, we obtain both u as in equation (4.60) and $u' \equiv 1$ as a solution of equation (4.50): Although we have $u = u' \equiv f$ on $\Delta_1 \times \{0\}$, their further evolution disagrees as u vanishes everywhere, whereas u' remains constant.

Analogous to the forward case, we expect that a proper integration of the boundary into the model may resolve this defect.

4.2.2 The uniqueness of solutions and an extension scheme from the boundary

Striving for an inclusion of the boundary into our model, we may at first observe that – as in the forward case – a solution u of the backward equation as given in equation (4.60) is smoothly extendable to the boundary $\partial\Delta_1 = \{0, 1\}$. However, this is quite effectless as all summands in equation (4.60) vanish on $\partial\Delta_1$, and consequently the extension of the solution there also vanishes for all $t > 0$, nor does it solve the uniqueness issue. Yet, this is consistent with the probabilistic interpretation of the solution, meaning that there is no positive probability to reach the interior when starting at the boundary (cf. below).

Nevertheless, the concept of solution still may be extended to the boundary if we reverse our angle of view and rather look for an extension of the yet to be determined specific boundary values into the interior than for a continuation the other way round. Also, this will rather match the backward setting with its reversed sense of time.

Such an extension may be carried out by (at first formally) augmenting the domain of equation (4.50) so that it comprises all of $\overline{\Delta}_1$. Thus, we may state an *extended Kolmogorov backward equation* by

$$\begin{cases} -\frac{\partial}{\partial t}u(p,t) = \frac{1}{2}p(1-p)\frac{\partial^2}{(\partial p)^2}u(p,t) & \text{in } \left(\overline{\Delta}_1\right)_{-\infty} = [0,1] \times (-\infty, 0) \\ u(p,0) = f(p) & \text{in } \overline{\Delta}_1,\ f|_{\Delta_1} \in \mathcal{L}^2(\Delta_1) \end{cases} \tag{4.63}$$

for $u(\,\cdot\,,t) \in C^2(\overline{\Delta}_1)$ for each fixed $t \in (-\infty, 0)$ and $u(x,\,\cdot\,) \in C^1((-\infty,0))$ for each fixed $x \in \overline{\Delta}_1$. This also allows us to prescribe an *extended final condition*

$f = f_0\chi_{\partial_0\Delta_1} + f_1\chi_{\Delta_1}$, which is defined on all of $\overline{\Delta}_1$, i.e. the boundary points may also belong to the target set considered (in contrast to the forward case, this extended condition is crucial and will also affect the solution in the interior as will be described below). The dynamics of the boundary values are those of a 0-dimensional Wright–Fisher process, hence they stay constant; this may formally be expressed by $\frac{\partial}{\partial t}u|_{\partial\Delta_1} = 0$, which coincides with the restriction of the first line of equation (4.63) to $\partial\Delta_1$. For this reason, it is also justified by the considered model to formulate equation (4.63) to hold on the boundary as well since this reflects the evolution of the process both in the interior and on the boundary of the domain.

A solution scheme for the extended Kolmogorov backward equation with a given extended final condition f is now as follows: At first, we solve equation (4.63) only in $\partial\Delta_1$ for the final condition $f\chi_{\partial_0\Delta_1} = f_0$ (separately for each $\Delta_0 \subset \partial_0\Delta_1$) and consecutively extend these solutions to Δ_1 by the requirement that each extension yields a solution of equation (4.50) in $\left(\overline{\Delta}_1\right)_{-\infty}$, thus in particular

(i) extends to the boundary $\partial_0\Delta_1$ such that it coincides with the previous solution in the respective $\Delta_0 \subset \partial_0\Delta_1$ resp. vanishes on the other $\Delta_0 \subset \partial_0\Delta_1$ and is of class C^2 with respect to the spatial variables in $\overline{\Delta}_1$,

(ii) fulfils $-\frac{\partial}{\partial t}u(p,t) = \frac{1}{2}p(1-p)\frac{\partial^2}{(\partial p)^2}u(p,t)$ in $(\Delta_1)_{-\infty}$.

A solution u_0 of equation (4.63) respectively restricted to the relevant $\Delta_0 \subset \partial_0\Delta_1$ is – of course – trivial, i.e. $u_0 \equiv f_0$, but for the continuation to Δ_1, we note that the eigenvalues need to be matched in terms of the chosen separation ansatz for being a solution: As on $\partial_0\Delta_1$ the spectrum of the differential operator $\frac{1}{2}p(1-p)\frac{\partial^2}{(\partial p)^2}$ is only $\{0\}$, the continuation to Δ_1 needs to be established by eigenfunctions corresponding to the eigenvalue 0. We may also use the Gegenbauer polynomials to this end, this time the ones corresponding to the eigenvalue 0, i.e. $C_0^{-\frac{1}{2}}(1-2p) = 1$ and $C_1^{-\frac{1}{2}}(1-2p) = 2p-1$. Using condition (i) above, their associated coefficients β_0, β_1 may be determined by

$$f_0(0) = \beta_0 C_0^{-\frac{1}{2}}(1) + \beta_1 C_1^{-\frac{1}{2}}(1) = \beta_0 - \beta_1 \tag{4.64}$$

$$\text{and} \quad f_0(1) = \beta_0 C_0^{-\frac{1}{2}}(-1) + \beta_1 C_1^{-\frac{1}{2}}(-1) = \beta_0 + \beta_1 \tag{4.65}$$

such that f_0 is matched by their continuous extension to the boundary. Thus,

$$\bar{u}_0\colon (\Delta_1)_{-\infty} \longrightarrow \mathbb{R} \quad \text{with} \quad \bar{u}_0(p,t) := \beta_0 + \beta_1(2p_1 - 1) \tag{4.66}$$

defines an extension of u_0 in $\partial_0\Delta_1$ to a solution of equation (4.63) in $\overline{\Delta}_1$. Rewriting this as

$$\bar{u}_0(p,t) \equiv (\beta_0 + \beta_1)p_1 + (\beta_0 - \beta_1)(1 - p_1), \tag{4.67}$$

it becomes apparent that the right term is the contribution of the solution in $\Delta_0^{(\{0\})} = \{0\}$ and the left term is the one of $\Delta_0^{(\{1\})} = \{1\}$ with both extensions satisfying the conditions stated above. We note that this continuation to Δ_1 by eigenfunctions for the eigenvalue 0 is unique for given boundary data by the maximum principle and that it is of course also definable on $\Delta_1 \times (-\infty, 0]$; then $\bar{f}_0 := \bar{u}_0(\,\cdot\,, 0)$ may be interpreted as an extension of the final condition f_0 in $\partial_0\Delta_1$ to Δ_1.

The observed extension $\bar{u}_0$ at this stage, however, actually is the solution to a different problem, i. e. to the extended Kolmogorov backward equation (4.63) with final condition $f_0\chi_{\partial_0\Delta_1} + \bar{f}_0\chi_{\Delta_1}$ (instead of $f_0\chi_{\partial_0\Delta_1} + f_1\chi_{\Delta_1}$). Then, $\bar{f}_0$ is perceivable as an induced additional target set on the entire Δ_1, which is generated by the eventual target set resp. f_0 in $\partial_0\Delta_1$ via the extension conditions (i) and (ii) above. This induced target set in turn modifies the already existing final condition f_1 in Δ_1, and consequently a (proper) solution u_1 in Δ_1 needs to cater to this modified final condition, signifying that the attraction of the eventual target set in $\partial_0\Delta_1$ via the induced target set in Δ_1 is handed over to the 1-dimensional process governing the correspondent dynamics.

Thus, in the next step, a (proper) solution u_1 of equation (4.63) restricted to Δ_1 is determined as in the previous section 4.2.1, but this time for the correspondingly adapted final condition

$$f_1'(p) := f_1(p) - \bar{f}_0(p), \quad p \in \Delta_1. \tag{4.68}$$

Restricting the choice of eigenfunctions to those corresponding to a strictly positive eigenvalue, this construction is unique for arbitrary $(f_1 - \bar{f}_0) \in \mathcal{L}^2(\Delta_1)$, which may be shown as in the forward case in lemma 4.2. Also, with such a choice of eigenfunctions, there is no interference of the solution u_1 with the boundary data since the resulting solution smoothly vanishes on the boundary (as already stated). Thus, the disjointness of the effective spectra of the operator in the interior and on the boundary prevents any ambiguity at this point.

Altogether, this amounts to an *extended solution* $u\colon \big(\overline{\Delta}_1\big)_{-\infty} \longrightarrow \mathbb{R}$ of equation (4.50) defined by

$$\begin{aligned} u(p,t) :=& u_0(p,t)\chi_{\partial_0\Delta_1}(p) + \big(\bar{u}_0(p,t) + u_1(p,t)\big)\chi_{\Delta_1}(p) \\ =& \beta_0 + \beta_1(2p-1) + \sum_{n\geq 2}(2n-1)\kappa_n\langle \tilde{f}', C_n^{-\frac{1}{2}}\rangle_{-\frac{1}{2}} C_n^{-\frac{1}{2}}(1-2p)e^{\kappa_n t} \end{aligned} \tag{4.69}$$

with $u(\,\cdot\,,t) \in C^\infty(\overline{\Delta}_1)$ for $t<0$ and

$$u(p,0) = \begin{cases} u_0 & = f_0 \quad \text{in } \partial_0\Delta_1 \\ \bar{u}_0 + u_1 & = f_1 \quad \text{in } \Delta_1, \end{cases} \tag{4.70}$$

thus fulfilling the extended final condition $f = f_0\chi_{\partial_0\Delta_1} + f_1\chi_{\Delta_1}$ in $\overline{\Delta}_1$. If we have $f_0 = 0$, then this solution agrees in Δ_1 with the one obtained before in equation (4.60) for $f := f_1$. Otherwise, we obtain another solution which evolves differently.

Thus, as already hinted, in contrast to the forward case, the specification of boundary data is required to obtain a unique solution. We have:

4.6 Proposition. *For a given final condition $f\colon \overline{\Delta}_1 \longrightarrow \mathbb{R}$, $f|_{\Delta_1} \in \mathcal{L}^2(\Delta_1)$, the extended Kolmogorov backward equation* (4.63) *corresponding to the diffusion approximation of the 1-dimensional Wright–Fisher model always allows a unique solution $u\colon \big(\overline{\Delta}_1\big)_{-\infty} \longrightarrow \mathbb{R}$ with $u(\,\cdot\,,t) \in C^\infty(\overline{\Delta}_1)$ for each fixed $t \in (-\infty,0)$ and $u(x,\,\cdot\,) \in C^\infty((-\infty,0))$ for each fixed $x \in \overline{\Delta}_1$.*

In the light of these considerations, we may now resolve the previous uniqueness counterexample for the Kolmogorov backward equation (4.50) in Δ_1 presented on p. 108. There, the initial condition $f=1$ did not seem to lead to a unique solution: This may now be overcome by shifting to the extended Kolmogorov backward equation and specifying whether the extended initial condition on $\overline{\Delta}_1$ is either $f=1$ or $f=\chi_{\Delta_1}$ corresponding to $f_0=1$ resp. $f_0=0$ (or something else). In all cases, a corresponding (extended) solution u is obtained as the unique solution (only in the first case, we have $u \equiv 1$).

4.2.3 Long-term behaviour and a probabilistic interpretation

Analogous to the forward case, we are also interested in the behaviour of the backward solution in the long run and will check whether it matches our previous findings in the forward case (cf. equation (4.49)).

Hence, for the extended solution $u = u_0\chi_{\partial_0\Delta_1} + (\bar{u}_0 + u_1)\chi_{\Delta_1}$ (cf. equation (4.69)), we may directly observe that for $t \to -\infty$ all eigenmodes corresponding to a positive eigenvalue (i. e. u_1) vanish, whereas those corresponding to the eigenvalue zero (i. e. $\bar{u}_0$) survive with the presence of the latter merely determined by the values of the extended final condition on the boundary. This construction of $\bar{u}_0$ as presented in the preceding section is equivalent to finding a solution of the *homogeneous* or *stationary* formulation of the extended Kolmogorov backward equation (4.63) being[5]

$$\begin{cases} \frac{1}{2}p(1-p)\frac{\partial^2}{(\partial p)^2}u(p) = 0, & p \in \Delta_1 \\ u(p) = f(p), & p \in \partial\Delta_1 \end{cases} \tag{4.71}$$

for $u \in C^2(\overline{\Delta_1})$ with u_0 now appearing as an (arbitrary) boundary condition f; this may be perceived as the Kolmogorov backward equation for $t = -\infty$. As already pointed out, a solution of such an equation is always unique by the maximum principle.

These considerations may be reflected by the following lemma:

4.7 Lemma. *For $t \to -\infty$, a solution of the extended Kolmogorov backward equation* (4.63) *for a given final condition $f = f_0\chi_{\partial_0\Delta_1} + f_1\chi_{\Delta_1}$, $f_1 \in \mathcal{L}^2(\Delta_1)$ converges uniformly on $\overline{\Delta_1}$ to the unique solution of the corresponding stationary equation* (4.71) *with boundary condition f_0. The space of such solutions is spanned by 1 and p.*

Proof. It remains to show that $u = \bar{u}_0 + u_1$ as given in equation (4.69) (and all its derivatives) converge uniformly on $\overline{\Delta_1}$ towards $\bar{u}_0$. This may be done by observing

[5] As in equation (4.63), one might equivalently formulate the equation to also hold on the boundary. However, we rather stick with the usual notation for boundary value problems.

that for $t > t_1 > 0$ we have

$$\begin{aligned} \left\| \frac{\partial^k}{\partial p^k} u_1(p,t) \right\|_{C^0(\overline{\Delta}_1)} &= \left\| \sum_{n \geq 2} (2n-1)\kappa_n \langle \tilde{f}', C_n^{-\frac{1}{2}} \rangle_{-\frac{1}{2}} \frac{\partial^k}{\partial p^k} C_n^{-\frac{1}{2}} (1-2p) e^{\kappa_n t} \right\|_{C^0(\overline{\Delta}_1)} \\ &\leq 3\|f\|_2 \sum_{n \geq 2} \kappa_n^{k+2} e^{-\kappa_n t} \leq 3\|f\|_2 \underbrace{\int_2^\infty x^{k+2} e^{-x t_1} dx}_{<\infty} \cdot e^{-2(t-t_1)} \end{aligned} \tag{4.72}$$

for all $k \geq 0$, which thus vanishes for $t \to \infty$. $\square$

In consideration of the probabilistic interpretation – which is that when the final condition f is chosen to be the characteristic function of some appropriate target set A, then the solution $u(p,t)$ gives the probability to end up in A at time 0 when starting in p at time $t < 0$ – this signifies that, if the boundary is not included into A, the eventual hit probability (i. e. when starting at $t = -\infty$) vanishes everywhere, whereas all interior points are without effect. This matches the expected behaviour of the process from the analysis of the Kolmogorov forward equation in section 4.1.3, where in the long run the probability density function of the process vanishes uniformly in the interior (cf. equation (4.49)).

However, adding some boundary points to the target set A corresponding to $f_0 \neq 0$ in $\partial_0 \Delta_1$ should change the observed behaviour of the backward solution decisively. Indeed, this provokes a positive eventual hit probability in the entire interior, which is again consistent with the model: Even if starting with a high frequency of a certain allele, eventually this allele may become extinct. Analogously, the observed behaviour of the (extended) forward solution in equation (4.49), where e. g. a process starting in $p \in \Delta_1$ eventually ends up in $\{1\}$ with probability p, is matched; the corresponding stationary backward solution $u(p, -\infty) = p$ yields the same result. Furthermore, in case both boundary points belong to the target set, the hit probabilities generated by each one overlap; in a more refined scenario, one could also assign arbitrary values to the boundary points, and the solution would then give a weighted mixed eventual hit probability.

Still, we wish to point out a subtle difference in the possible interpretations of an extended solution originating from $\partial_0 \Delta_1$ (as in section 4.2.2), which may simultaneously appear as a stationary solution in Δ_1. While the stationary solution is time-independent in the strict sense, the obtained solution in proposition 4.6 is time-dependent, and this in principle also holds if it only consists of the (time-independent) component $\bar{u}_0$. As stated, for $t = 0$ this may be interpreted as the probability density of an additional (generalised) target set induced by $\partial_0 \Delta_1$, spanning the entire Δ_1 as the boundary at this time is not reachable from anywhere in the interior. For $t < 0$, it would then indicate the joint hit probability for both the induced and the eventual target set since the boundary has become reachable meanwhile. As t progresses further, the attraction of the component of the target set in Δ_1 diminishes, whereas that of $\partial_0 \Delta_1$ increases as the boundary is more and more likely to be reached over time. For $t = -\infty$, then a solution as in proposition 4.6 merely expresses the attraction of the boundary target set component, which coincides with the interpretation of u as a stationary solution.

We also note that accounting for the boundary in the backward scenario again reveals the hierarchical structure of the process already observed in the analysis of the Kolmogorov forward equation. Here, this again expresses itself in the different dimensionality of the boundary and the interior as well as in a correspondingly different reception of the final condition on them: The value of f at boundary points substantially influences the whole process, whereas the value of f at single points in Δ_1 is without effect as this is a $\lambda\!\!\lambda_1$-null set. This hierarchicality as well as the fact that the boundary values determine the solution in the interior (in a manner of speaking "extend to the interior") will also be observed in higher dimensions (cf. chapter 5.4).

5 Analytic aspects of the diffusion approximation of the multidimensional Wright–Fisher model

5.1 Preliminaries

Having analysed the diffusion approximation of the 1-dimensional Wright–Fisher model in the preceding chapter 4, we now wish to generalise the solution techniques found, in particular for the extension scheme, by applying them to the diffusion approximation of the multidimensional Wright–Fisher model. First, we again aim to solve the corresponding Kolmogorov equations. Again, recombination, mutation or natural selection will be left out, and we will consider a 'pure' 1-locus n-allelic haploid model; as already stated, without recombination this also covers multi-loci models.

To begin with, we will explore the structure of the now bigger domain and introduce some new notation with respect to it.

5.1.1 The simplex

As we are considering frequencies, this directly leads to simplices as corresponding state space. Recapitulating the definition from section 2.2.3, the (open) *n-dimensional standard orthogonal simplex* is given by

$$\Delta_n := \Big\{(x^1,\dots,x^n) \in \mathbb{R}^n \Big| x^i > 0 \text{ for } i = 1,\dots,n \text{ and } \sum_{i=1}^{n} x^i < 1\Big\}, \tag{5.1}$$

which we usually refer to when talking of the 'simplex'. This is a projection of the fully regular (open) *n-dimensional standard simplex*

$$\Sigma^n := \Big\{(x^1,\dots,x^{n+1})\in\mathbb{R}^{n+1}\Big|x^i>0 \text{ for } i=1,\dots,n+1 \text{ and } \sum_{i=1}^{n+1}x^i=1\Big\} \tag{5.2}$$

onto $\mathbb{R}^n$ by suppressing the $(n+1)$-st coordinate. In terms of the frequencies of a Wright–Fisher model, there is no difference between the two simplices; therefore, we will primarily use the first formulation, which is the handier one.

5.1.2 The boundary structure of the simplex

In the following, we will introduce a suitable notation for different instances of the boundary of the simplex. Beforehand, we explicitly denote the coordinate indices for Δ_n by providing the corresponding coordinate index set $I_n := \{i_0,\dots,i_n\}\subset\{0,\dots,n\}$, $i_j\neq i_l$ for $j\neq l$ with usually $i_0\equiv 0$ corresponding to $x^0 := 1-\sum_{i=1}^n x^i$ as upper index of Δ_n, thus

$$\Delta_n^{(I_n)} = \big\{(x^1,\dots,x^n)\in\mathbb{R}^n\big|x^i>0 \text{ for all } i\in I_n\big\}. \tag{5.3}$$

This is particularly useful for boundary instances of the simplex (cf. below) or if for other purposes a certain ordering $(i_j)_{j=0,\dots,n}$ of the coordinate indices is needed. If no index set is stated for Δ_n, we usually have $I_n\equiv\{0,\dots,n\}$ as in equation (5.1).

Now assessing the boundary structure of the simplex, we first recall that the simplex Δ_n is open in the standard topology on $\mathbb{R}^n$ (which we always assume when writing Δ_n); its closure $\overline{\Delta_n}$ is given by (again using the index set notation)

$$\overline{\Delta_n^{(I_n)}} = \big\{(x^1,\dots,x^n)\in\mathbb{R}^n\big|x^i\geq 0 \text{ for all } i\in I_n\big\} \tag{5.4}$$

including $x^0 = 1-\sum_{i=1}^n x^i$. The boundary $\partial\Delta_n = \overline{\Delta_n}\setminus\Delta_n$ consists of various subsimplices of descending dimensions called *faces*, starting from the $(n-1)$-dimensional *facets* down to the *vertices* (which represent 0-dimensional faces). All appearing subsimplices of dimension $k\leq n-1$ are isomorphic to the k-dimensional standard orthogonal simplex Δ_k and hence (in slight abuse of notation) will be denoted by

$\Delta_k^{(I_k)}$ with usually $I_k := \{i_0, \ldots, i_k\} \subset I_n$, $i_j \neq i_l$ for $j \neq l$ and $i_0 \equiv 0$ again listing the $k+1$ coordinate indices with strictly positive value:

$$\Delta_k^{(I_k)} := \Big\{(x^1, \ldots, x^n) \in \overline{\Delta_n^{(I_n)}} \Big| x^i > 0 \text{ for } i \in I_k;\, x^i = 0 \text{ for } i \in I_n \setminus I_k\Big\}. \tag{5.5}$$

Generally, for a given $k \leq n-1$, there are of course $\binom{n+1}{k+1}$ different (unordered) subsets I_k of I_n, each of which corresponds to a certain boundary face $\Delta_k^{(I_k)}$. We therefore introduce the *k-dimensional boundary $\partial_k\Delta_n$ of Δ_n* by putting

$$\partial_k\Delta_n^{(I_n)} := \bigcup_{I_k \subset I_n} \Delta_k^{(I_k)} \subset \partial\Delta_n^{(I_n)} \quad \text{for } 0 \leq k \leq n-1. \tag{5.6}$$

For reasons of notational systematics, we sometimes also write $\partial_n\Delta_n$ for Δ_n, though this does not comprise a boundary in the current setting (however, it might be imagined as a boundary face of some Δ_{n+1}). Of course, the concept of the k-dimensional boundary also applies to simplices which are themselves boundary instances of some $\Delta_l^{(I_l)}$, $I_l \subset I_n$ for $0 \leq k < l \leq n$, thus

$$\partial_k\Delta_l^{(I_l)} = \bigcup_{I_k \subset I_l} \Delta_k^{(I_k)} \subset \partial\Delta_l^{(I_l)}. \tag{5.7}$$

Translating this notation to the setting of the Wright–Fisher model, we have: Δ_n corresponds to the state of all $n+1$ alleles being present, whereas $\partial_k\Delta_n$ represents the state of exactly (any) $k+1$ alleles being present in the population. The individual $\Delta_k^{(\{i_0,\ldots,i_k\})}$ comprising $\partial_k\Delta_n$ correspond to the state of exactly the alleles $i_0, \ldots, i_k$ being present in the population. Likewise, $\partial_{k-1}\Delta_k^{(\{i_0,\ldots,i_k\})}$ corresponds to the state of exactly one further allele out of $i_0, \ldots, i_k$ being eliminated from the population.

5.1.3 Geometrical properties of the simplex

In contrast to the fully regular standard simplex Σ^n (e.g. for $n = 3$, the *regular tetrahedron*), the standard orthogonal simplex Δ_n as defined in equation (5.1) has an orthogonal corner at the coordinate origin and volume

$$\mathrm{vol}_n(\Delta_n) = \int\limits_{\Delta_n} d\lambda\!\!\lambda_n = \frac{1}{n!} \quad \text{for all } n \in \mathbb{N} \tag{5.8}$$

with $\lambda\!\!\lambda_n$ being the n-dimensional Lebesgue measure on $\mathbb{R}^n$. The subsimplices $\Delta_{n-1}^{(I_n \setminus \{l\})}$, $l = 1, \ldots, n$ are in fact $(n-1)$-dimensional standard orthogonal simplices spanned by $n-1$ coordinate axes respectively, whereas the $(n+1)$-st subsimplex

$$\Delta_{n-1}^{(I_n \setminus \{0\})} = \Big\{ (x^1, \ldots, x^n) \in \mathbb{R}^n \Big| x^i > 0 \text{ for } i = 1, \ldots, n \text{ and } \sum_{i=1}^{n} x^i = 1 \Big\} \tag{5.9}$$

lies opposite the origin coinciding with Σ^{n-1} as defined in equation (5.2).

When determining the $(n-1)$-dimensional volume of $\Delta_{n-1}^{(I_n \setminus \{0\})}$, the measure with respect to which we integrate in equation (5.8) needs to be replaced by the one induced on $\Delta_{n-1}^{(I_n \setminus \{0\})}$ by the Lebesgue measure of $\mathbb{R}^n$; however, this will still be denoted by $\lambda\!\!\lambda_{n-1}$ as it is apparent from the domain of integration which version is used. We then obtain that $\mathrm{vol}_{n-1}\big(\Delta_{n-1}^{(I_n \setminus \{0\})}\big)$ is $\sqrt{n}$ times the volume of the other subsimplices $\Delta_{n-1}^{(I_n \setminus \{l\})} \cong \Delta_{n-1}$ as follows from the *generalised Pythagorean theorem*

$$\sum_{l=1}^{n} \mathrm{vol}_{n-1}^2 \big(\Delta_{n-1}^{(I_n \setminus \{l\})}\big) = \mathrm{vol}_{n-1}^2 \big(\Delta_{n-1}^{(I_n \setminus \{0\})}\big), \tag{5.10}$$

thus $\mathrm{vol}_{n-1}\big(\Delta_{n-1}^{(I_n \setminus \{0\})}\big) = \sqrt{n}\, \mathrm{vol}_{n-1}\big(\Delta_{n-1}^{(I_n \setminus \{1\})}\big)$.

This difference in volume of the boundary faces, however, is only due to the choice of coordinates of the standard orthogonal simplex, in which the 0th coordinate is implicitly formulated in terms of the remaining coordinates. With the regular simplex Σ^n, none of the faces is distinct, and the same also applies to the Wright–Fisher models that we are looking at: For the corresponding process, none of the alleles is special, and hence all boundary faces are equivalent. Correspondingly, for example in equation (5.52), a correction factor $\frac{1}{\sqrt{n}}$ to $\mathrm{vol}_{n-1}\big(\Delta_{n-1}^{(I_n \setminus \{0\})}\big)$ appears.

5.1.4 Products and further notation

Analogous to $\mathrm{vol}_n(\Delta_n)$, we also introduce a *product* of functions $u, v \in \mathcal{L}^2(\Delta_n)$ by putting

$$(u,v)_n := \int\limits_{\Delta_n} u(x)v(x)\,\lambda\!\!\lambda_n(dx) \tag{5.11}$$

with $\lambda\!\!\lambda_n$ again denoting the Lebesgue measure. It is emphasised that only the interior Δ_n is assumed as the domain of integration; the index – if no confusion is to be expected – may be omitted. Of course, as n was arbitrary, the product may also be used on some $\Delta_k^{(I_k)} \subset \partial\Delta_n$; yet, this requires a slight modification: Analogous to $\mathrm{vol}_{n-1}\big(\Delta_{n-1}^{(I_n\setminus\{l\})}\big)$, if integrating over some $\Delta_k^{(I_k)}$ with $0 \notin I_k$, the measure needs to be replaced with the one induced on $\Delta_k^{(I_k)}$ by the Lebesgue measure of the containing $\mathbb{R}^{k+1}$ – this measure, however, will still be denoted by $\lambda\!\!\lambda_k$ as it is clear from the domain of integration $\Delta_k^{(I_k)}$ with either $0 \in I_k$ or $0 \notin I_k$ which version is actually used.

The products $(\,\cdot\,,\cdot\,)_k$ on $\partial_k\Delta_n^{(I_n)}$ for $k = 0,\dots,n$ may then be utilised to construct a *hierarchical product* on the closure of the simplex $\overline{\Delta}_n$ which is adapted to the hierarchical structure of the boundary instances of the simplex. Hence, for functions $u, v\colon \overline{\Delta}_n \longrightarrow \mathbb{R}$ with $u, v|_{\partial_k\Delta_n^{(I_n)}} \in \mathcal{L}^2\big(\partial_k\Delta_n^{(I_n)}\big)$ for all $k = 0,\dots,n$, we may put

$$[u,v]_n := \sum_{k=0}^{n} (u,v)_k \tag{5.12}$$

with $(u,v)_k$ – deviantly to its proper definition in equation (5.11) – in this context denoting the integral over the full k-dimensional boundary $\partial_k\Delta_n$ of Δ_n (cf. equation (5.6)), thus

$$[u,v]_n = \sum_{k=0}^{n} (u,v)_k = \sum_{k=0}^{n} \int\limits_{\partial_k\Delta_n} u(x)v(x)\,\lambda\!\!\lambda_k(dx) = \sum_{k=0}^{n} \sum_{I_k\subset I_n} \int\limits_{\Delta_k^{(I_k)}} u(x)v(x)\,\lambda\!\!\lambda_k(dx) \tag{5.13}$$

with $\lambda\!\!\lambda_k$ again denoting either the Lebesgue measure of $\mathbb{R}^k$ or – if the domain of integration is some $\Delta_k^{(I_k)}$ with $0 \notin I_k$ – the measure induced on $\Delta_k^{(I_k)}$ by Lebesgue measure of the containing $\mathbb{R}^{k+1}$.

Corresponding to the hierarchical product $[\,\cdot\,,\cdot\,]_n$, we may restate the integrability criterion for functions on $\overline{\Delta}_n$ by defining a corresponding space

$$\mathcal{L}^2\Big(\bigcup_{k=0}^{n} \partial_k\Delta_n\Big) := \Big\{ f\colon \overline{\Delta}_n \longrightarrow \mathbb{R} \,\Big|\, f|_{\partial_k\Delta_n} \text{ is } \lambda\!\!\lambda_k\text{-measurable and} \\ \int_{\partial_k\Delta_n} |f(x)|^2 \,\lambda\!\!\lambda_k(dx) < \infty \text{ for all } k = 0, \dots, n \Big\} \tag{5.14}$$

(with the modified measure $\lambda\!\!\lambda_k$ if applicable), whereas for the top-dimensional simplex, we simply have

$$\mathcal{L}^2(\Delta_n) := \Big\{ f\colon \Delta_n \longrightarrow \mathbb{R} \,\Big|\, f \text{ is } \lambda\!\!\lambda_n\text{-measurable and } \int_{\Delta_n} |f(x)|^2 \,\lambda\!\!\lambda_n(dx) < \infty \Big\}, \tag{5.15}$$

coinciding with the usual definition. Furthermore, we also define for $k \in \mathbb{N} \cup \{\infty\}$

$$C_0^k(\overline{\Delta}_n) := \big\{ f \in C^k(\overline{\Delta}_n) \big| f|_{\partial\Delta_n} = 0 \big\}, \tag{5.16}$$

$$C_0^k(\Delta_n) := \big\{ f \in C^k(\Delta_n) \big| \exists\, \bar{f} \in C_0^k(\overline{\Delta}_n) \text{ with } \bar{f}|_{\Delta_n} = f \big\} \tag{5.17}$$

as well as

$$C_c^k(\overline{\Delta}_n) := \big\{ f \in C^k(\overline{\Delta}_n) \big| \operatorname{supp}(f) \subsetneq \Delta_n \big\}. \tag{5.18}$$

5.2 The Kolmogorov operators

With the simplex notation and the products in place, we may now start our analysis: Again picking up the setting of proposition 2.9, we may formulate the Kolmogorov forward equation for the diffusion approximation of a now n-allelic 1-locus Wright–Fisher model

$$\begin{cases} \frac{\partial}{\partial t} u(x,t) = L_n u(x,t) & \text{in } (\Delta_n)_\infty = \Delta_n \times (0,\infty) \\ u(x,0) = f(x) & \text{in } \Delta_n,\ f \in \mathcal{L}^2(\Delta_n) \end{cases} \tag{5.19}$$

for $u(\,\cdot\,,t) \in C^2(\Delta_n)$ for each fixed $t \in (0,\infty)$ and $u(x,\,\cdot\,) \in C^1((0,\infty))$ for each fixed $x \in \Delta_n$ and with

$$L_n u(x,t) := \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x^i \partial x^j}\big(x^i(\delta^i_j - x^j)u(x,t)\big) \tag{5.20}$$

being the *forward operator*. Analogously, we introduce the *backward operator*

$$L_n^* u(x,t) := \frac{1}{2}\sum_{i,j=1}^{n} \big(x^i(\delta^i_j - x^j)\big)\frac{\partial^2}{\partial x^i \partial x^j}u(x,t), \tag{5.21}$$

appearing in the corresponding Kolmogorov backward equation (a detailed discussion of this equation will be carried out in chapter 5.4). The definitions of the operators given in equations (5.20) and (5.21) also apply to the closure $\overline{\Delta}_n$; this is just noted here as we will also consider extensions of the solution and the differential equation to the boundary.

For relations between the two operators, we immediately have the following two lemmas:

5.1 Lemma. *L_n and L_n^* are (formal) adjoints with respect to the product $(\,\cdot\,,\cdot\,)_n$ in the sense that*

$$(L_n u, \varphi)_n = (u, L_n^*\varphi)_n \quad \text{for } u \in C^2(\overline{\Delta}_n),\ \varphi \in C^2_0(\overline{\Delta}_n). \tag{5.22}$$

Proof. The assertion directly follows from proposition 5.7 below. □

5.2 Lemma. *For an eigenfunction $\varphi \in C^2(\overline{\Delta}_n)$ of L_n and $\omega_n := \prod_{k=1}^n x^k\big(1 - \sum_{l=1}^n x^l\big)$, we have: $\omega_n\varphi \in C^2_0(\overline{\Delta}_n)$ is an eigenfunction of L_n^* corresponding to the same eigenvalue and conversely.*

Proof. Looking for a function ω_n with $L_n^*(\omega_n u) = \omega_n L_n(u)$ (and hence for L_n-eigenfunctions φ consequently $L_n^*(\omega_n\varphi) = \omega_n L_n(\varphi) = -\lambda\omega_n\varphi$), we have on the one hand

$$L_n u = -\frac{n(n+1)}{2}u + \sum_i (1-(n+1)x^i)\frac{\partial}{\partial x^i}u + \frac{1}{2}\sum_{i,j} x^i(\delta^i_j - x^j)\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}u \tag{5.23}$$

and

$$\begin{aligned} L_n^*(\omega_n u) &= \frac{1}{2}\sum_{i,j} x^i(\delta^i_j - x^j)\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}u \\ &= \frac{1}{2}\sum_{i,j} x^i(\delta^i_j - x^j)\Big(\Big(\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}\omega_n\Big)u + 2\frac{\partial}{\partial x^j}\omega_n\frac{\partial}{\partial x^i}u + \omega_n\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}u\Big) \end{aligned} \tag{5.24}$$

on the other hand. Thus, it is sufficient if such a function ω_n satisfies

$$\begin{cases} \sum_{i,j} x^i(\delta^i_j - x^j)\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}\omega_n = -n(n+1)\omega_n \\ \sum_j x^i(\delta^i_j - x^j)\frac{\partial}{\partial x^j}\omega_n = (1-(n+1)x^i)\omega_n \quad \text{for all } i, \end{cases} \tag{5.25}$$

which is the case for $\omega_n = \prod_{k=1}^n x^k\big(1-\sum_{l=1}^n x^l\big)$ as may easily be verified by direct computation. □

We continue with some further observations on the operators, in particular with regard to the boundary of Δ_n: The operator L_n^*, if restricted to subsimplices $\Delta_k^{(I_k)} \cong \Delta_k$ in $\partial\Delta_n^{(I_n)}$ of any dimension k, then again is the adjoint of the differential operator L_k corresponding to the evolution of a $(k+1)$-allelic process in Δ_k:

5.3 Lemma. *For $0 \le k < n$ and $I_k \subset \{0,\dots,n\}$, $|I_k| = k$, we have*

$$L_n^*\big|_{\Delta_k^{(I_k)}} = L_k^*. \tag{5.26}$$

Proof. For $I_k \subset \{1,\dots,n\}$, $|I_k| = k$ and $\Delta_k^{(I_k)} := \big\{(x^1,\dots,x^n)\big|x^i > 0 \text{ for } i \in I_k,\ x^i = 0 \text{ for } i \in \{0,\dots,n\}\setminus I_k\big\}$ with $x^0 = 1-\sum_i^n x^i$, we directly have:

$$\begin{aligned} L_n^*\big|_{\Delta_k^{(I_k)}} &= \frac{1}{2}\sum_{i,j=1}^n \big(x^i(\delta^i_j - x^j)\big)\frac{\partial^2}{\partial x^i\partial x^j}\Big|_{\Delta_k^{(I_k)}} \\ &= \frac{1}{2}\sum_{i,j\in I_k} \big(x^i(\delta^i_j - x^j)\big)\frac{\partial^2}{\partial x^i\partial x^j} \equiv L_k^*. \end{aligned} \tag{5.27}$$

By symmetry, this also holds for I_k with $0 \in I_k$, hence for arbitrary I_k. □

We may therefore omit the index k in L_k^* whenever convenient, in particular when considering domains where (parts of) the boundary are included.

For the operator L_n, we do not have such a restriction property: If restricted to some $\Delta_k^{(I_k)}$, it does not correspond to L_k describing a $(k+1)$-allelic process in Δ_k. This becomes clear when carrying out the differentiations, i. e.

$$L_n u(x,t) = -\frac{n(n+1)}{2}u(x,t) + \sum_{i=1}^{n}(1-(n+1)x^i)\frac{\partial}{\partial x^i}u(x,t) + L_n^* u(x,t), \qquad (x,t) \in (\Delta_n)_\infty. \tag{5.28}$$

This expanded equation may be interpreted as follows: The 2nd order derivatives $(= L_n^*)$ represent the (undirected, but generally with different absolute value for different directions) diffusion term, while the 1st order derivatives may be interpreted as (directed) drift from the centroid $\left(\frac{1}{n+1}, \dots, \frac{1}{n+1}\right)$ of Δ_n to the boundary. The reaction term $-\frac{n(n+1)}{2}u$ expresses the total loss of mass, which is due to the absorption at the boundary being (heuristically seen) proportional to $\frac{\mathrm{vol}_{n-1}(\partial_{n-1}\Delta_n)}{\mathrm{vol}_n(\Delta_n)} = \frac{(n+1)n!}{(n-1)!}$ (cf. also proposition 5.7 below).

Now restricting equation (5.28) to e. g. $\Delta_{n-1}^{(I_n\setminus\{m\})}$ for some $m \neq 0$ (assuming the appropriate extendibility of the solution u) yields

$$\begin{aligned} u_t(x,t) = L_n u(x,t) = -\frac{n(n+1)}{2}u(x,t) + \sum_{\substack{i=1\\ i\neq m}}^{n}(1-(n+1)x^i)\frac{\partial}{\partial x^i}u(x,t) \\ + \frac{\partial}{\partial x^m}u(x,t) + L_{n-1}^* u(x,t), \quad (x,t) \in \left(\Delta_{n-1}^{(I_n\setminus\{m\})}\right)_\infty, \end{aligned} \tag{5.29}$$

which may be contrasted with the proper equation for $\Delta_{n-1}^{(I_n\setminus\{m\})}$ with L_{n-1} (i. e. equation (5.28) with n replaced by $n-1$), thus

$$u_t(x,t) = L_{n-1}u(x,t) = -\frac{n(n-1)}{2}u(x,t) + \sum_{\substack{i=1\\ i\neq m}}^{n}(1-nx^i)\frac{\partial}{\partial x^i}u(x,t) + L_{n-1}^* u(x,t), \qquad (x,t) \in \left(\Delta_{n-1}^{(I_n\setminus\{m\})}\right)_\infty \tag{5.30}$$

Thus, while the diffusion terms agree in both cases $(= L_{n-1}^*)$, the drift components are notably different: Apart from that the extension of the n-dimensional process features a drift component normal to $\Delta_{n-1}^{(I_n\setminus\{m\})}$ (i. e. $\frac{\partial}{\partial x^m}u(x,t)$), also the projected drift onto

$\Delta_{n-1}^{(I_n\setminus\{m\})}$ does not agree with the one generated by L_{n-1}: The n-dimensional process has a diverging drift centred about $\left(\frac{1}{n+1},\ldots,\frac{1}{n+1}\right)\in\Delta_{n-1}^{(I_n\setminus\{m\})}$, whereas the drift of the $(n-1)$-dimensional process centres about $\left(\frac{1}{n},\ldots,\frac{1}{n}\right)$, the centroid of $\Delta_{n-1}^{(I_n\setminus\{m\})}$. Also, the reaction terms expressing the total loss over time do not agree, which is due to the higher number of options to lose an allele that the higher-dimensional solution has (cf. also the heuristics in the preceding paragraph).

This lacking restriction property of the forward operator (and the corresponding process) seemingly is a notable distinction from the behaviour of the underlying discrete process. However, this is rather a technical artefact than a gap in the diffusion approximation: If considering paths, a process that has already lost one or more alleles is equivalent to a process starting with the same reduced number of alleles. Of course, this applies both to a discrete process as well as to the corresponding diffusion limit process. However, in the context of the Kolmogorov equations, we are rather observing probability densities than paths: Then, the evolution of the process on a boundary instance (i. e. at a stage where the process would already have lost some alleles) is generally not independent from higher dimensional entities of the domain (cf. also section 5.3.2). In this conception, this holds for both the discrete and the continuous limit processes; merely, probability densities are rarely used when setting up a discrete Wright–Fisher model and only arise as a technical necessity for formulating the Kolmogorov equations (cf. also section 2.4.1).

5.3 Solution schemes for the Kolmogorov forward equation

As in the 1-dimensional case, we are eventually striving for a solution of the Kolmogorov forward equation (5.19) which is preferably valid and complete in terms of the underlying model. Again, solution schemes already exist: As early as 1956, Kimura presented a solution scheme for the 3-allelic case ($n=2$) in [18]. Another approach by separation of variables was presented by Baxter, Blythe and McKane in [7], this time for an arbitrary number of alleles. A recent work by T. D. Tran in [30] provides a solution scheme in terms of a generalisation of the Gegenbauer polynomials used for solving the 1-dimensional Kolmogorov forward equation:

5.4 Proposition (cf. [30], p. 55)**.** *For $n \in \mathbb{N}_+$ and each multi-index $\alpha = (\alpha^1, \dots, \alpha^n)$ with $|\alpha| = \alpha^1 + \cdots + \alpha^n = m \geq 0$,*

$$C_{m,\alpha}(x) := x^\alpha + \sum_{|\beta|=0}^{m-1} a_{m,\beta} x^\beta, \quad x \in \Delta_n \tag{5.31}$$

with $a_{m,\beta}$ inductively defined by $a_{m,\beta} := \delta^\alpha_\beta$ for $|\beta| = m$ and

$$a_{m,\beta} := -\frac{\sum_{i=1}^n (\beta^i + 2)(\beta^i + 1) a_{m,(\beta^1,\dots,\beta^i+1,\dots,\beta^n)}}{(m - |\beta|)(m + |\beta| + 2n + 1)} \quad \text{for all } 0 \leq |\beta| \leq m-1, \tag{5.32}$$

is an eigenfunction of L_n in $\overline{\Delta}_n$ corresponding to the eigenvalue $\lambda_m^{(n)} = \frac{(m+n)(m+n+1)}{2}$.

Knowing the eigenfunctions, it is straightforward to construct a solution of the corresponding equation (for details cf. [30]), leading to the following result, which effectively coincides with those of the other sources mentioned (i. e. [18], [7]):

5.5 Proposition. *For $n \in \mathbb{N}$ and an initial condition $f \in \mathcal{L}^2(\Delta_n)$, the Kolmogorov forward equation corresponding to the diffusion approximation of the n-dimensional Wright–Fisher model* (5.19) *always allows a unique solution $u\colon \big(\Delta_n\big)_\infty \longrightarrow \mathbb{R}$ with $u(\,\cdot\,, t) \in C^\infty(\Delta_n)$ for each fixed $t \in (0, \infty)$ and $u(x, \,\cdot\,) \in C^\infty((0,\infty))$ for each fixed $x \in \Delta_n$. Furthermore, this solution (and all its spatial derivatives) may be extended continuously to the boundary $\partial\Delta_n$.*

However, this result primarily only covers the existence (and uniqueness) of a solution in the interior Δ_n and does not provide any proper assertions about the behaviour of the process on the boundary. Yet, as seen from the discussion in chapter 4, the appropriate inclusion of the boundary in terms of a probability density describing its entire evolution is crucial for a complete account of the model. Thus, our main goal will be to establish a solution scheme including the boundary as already done in the 1-dimensional setting. This will again depend on the extendibility of the solution to the boundary of corresponding regularity, i. e. an extension of class C^2 with respect to the spatial variables at least. For this, the corresponding statement of proposition 5.5 is needed; also, this regularity, which is based on the regularity of the generalised Gegenbauer polynomials, is in concordance with standard PDE theory (cf. e. g. [14]).

5.3.1 Moments and the weak formulation of the Kolmogorov forward equation

As in the 1-dimensional model, we may observe that a solution of equation (5.19) in Δ_n lacks conservation properties: As the smallest eigenvalue of L_n is $\lambda_0^{(n)} = \frac{n(n+1)}{2} > 0$, a solution vanishes everywhere in Δ_n for $t \to \infty$, which in particular implies that the total mass and the expectation values are not preserved. However, again these properties are an important aspect of the validity of the model, and we will hence strive to fulfil them. Eventually, this will lead us to introducing a suitable extended solution on the entire $\overline{\Delta}_n$ as in the 1-dimensional setting (cf. section 4.1.3); a justification of this concept will again be obtained through considerations of moments of the process, for which we analogously require that they coincide with the limit of the corresponding moments of the underlying (suitably rescaled) discrete processes (cf. equation (4.27)).

This assumption again allows us to also establish an *n-dimensional generalisation of the moments evolution equation* (cf. the 1-dimensional case in equation (4.29)) being

$$\frac{\partial}{\partial t}\bar{\mu}_\alpha(t) = -\frac{|\alpha|(|\alpha|-1)}{2}\bar{\mu}_\alpha(t) + \sum_{i=1}^{n} \frac{\alpha_i(\alpha_i-1)}{2}\bar{\mu}_{\alpha-e_i}(t) \tag{5.33}$$

for $\alpha = (\alpha_1, \dots, \alpha_n)$, $|\alpha| \geq 1$, whereas $\frac{\partial}{\partial t}\bar{\mu}_0(t) = 0$ (with e_i denoting the multi-index $(0, \dots, 0, 1, 0, \dots, 0)$ with 1 appearing at the i-th position) for the moments of the n-dimensional process, which we may formally write as

$$\bar{\mu}_\alpha(t) = \mathrm{E}\left(X(t)^\alpha\right). \tag{5.34}$$

This may be derived analogously to its 1-dimensional correspondent, equation (4.29) in section 4.1.3, primarily for the discrete model $(C(t))_{t\in\mathbb{N}}$ using $\mathrm{E}_1(\delta C|C(t)) = 0$, $\mathrm{E}_1(\delta C_i \delta C_j|C(t)) = \frac{1}{N}C(t)(\delta_j^i - C(t))$, $\mathrm{E}_1((\delta C)^\alpha|C(t)) \in \mathcal{O}(N^{-2})$ for $|\alpha| \geq 3$ (cf. also equations (2.53)–(2.55) accordingly for a haploid model and with $R = 0$) and by shifting to the diffusion limit subsequently.

While we obviously have

$$\overline{m}_\alpha(t) = \mathrm{E}\left(C(t)^\alpha\right) \tag{5.35}$$

for the moments of the underlying discrete process, it is again not quite apparent what exactly corresponds to the moments $\bar{\mu}_\alpha(t)$ of the diffusion limit process $(X(t))_{t\in\mathbb{R}_+}$. However, analogous to the 1-dimensional setting, the discrete moments as defined in equation (5.35) also span configurations where certain allele frequencies may be zero; in the diffusion limit, this corresponds to the boundary $\partial\Delta_n$ of the state space of the continuous process Δ_n. For this reason, the expectation with respect to $(X(t))_{t\in\mathbb{R}_+}$ should explicitly take into account the boundary $\partial\Delta_n$, and hence, when rewriting equation (4.29) by application of a (generic) product $[\,\cdot\,,\cdot\,]$

$$\bar{\mu}_\alpha(t) = \mathrm{E}\left(X(t)^\alpha\right) = \left[U(t), x^\alpha\right], \tag{5.36}$$

this should involve an integration over $\overline{\Delta}_n$ in extension of $(\,\cdot\,,\cdot\,)_n$ (cf. equation (5.11)); in this context the capitalised $U\colon (\overline{\Delta}_n)_\infty \longrightarrow \mathbb{R}$ shall denote an extended solution which is assumed to be the probability density function of $(X(t))_{t\in\mathbb{R}_+}$ on the entire $\overline{\Delta}_n$ (thus in particular $U|_{\Delta_n}$ is a solution of the Kolmogorov forward equation (5.19) in Δ_n).

Yet, even with the exact form of $[\,\cdot\,,\cdot\,]$ remaining unclear, we may already illustrate the coherence between the moments evolution equation (5.33) and the Kolmogorov backward operator L^* in $\overline{\Delta}_n$ as defined in equation (5.21). Since L^* has polynomial coefficients, it maps polynomials to polynomials, and we have

$$\begin{aligned}
L^* x^\alpha &= \frac{1}{2}\sum_{i,j=1}^{n}\left(x^i(\delta^i_j - x^j)\right)\frac{\partial^2}{\partial x^i \partial x^j} x^\alpha \\
&= \frac{1}{2}\sum_{i=1}^{n}\alpha_i(\alpha_i - 1)(x^{\alpha-e_i} - x^\alpha) - \frac{1}{2}\sum_{i\neq j}\alpha_i\alpha_j x^\alpha \\
&= \frac{1}{2}\sum_{i=1}^{n}\alpha_i(\alpha_i - 1)x^{\alpha-e_i} - \frac{1}{2}|\alpha|(|\alpha|-1)x^\alpha \qquad \text{for } x \in \overline{\Delta}_n,
\end{aligned} \tag{5.37}$$

which yields, using the notation of equation (5.34),

$$\big[U(t), L_n^* x^\alpha\big] = \frac{1}{2}\sum_{i=1}^{n} \alpha_i(\alpha_i - 1)\bar{\mu}_{\alpha-e_i}(t) - \frac{1}{2}|\alpha|(|\alpha| - 1)\bar{\mu}^\alpha(t) \tag{5.38}$$

with the right-hand side being equal to that of equation (5.33). Thus, if the moments equation is fulfilled for some probability density function U, we may equivalently write

$$\frac{\partial}{\partial t}\bar{\mu}_\alpha(t) = \left[\frac{\partial}{\partial t}U(t), x^\alpha\right] = \big[U(t), L_n^* x^\alpha\big] \quad \text{for } t \in (0, \infty). \tag{5.39}$$

A (formally) more general version of this equation with x^α replaced by a generic test function φ, thus

$$\left[\frac{\partial}{\partial t}U(t), \varphi\right] = \big[U(t), L_n^* \varphi\big] \quad \text{for } \varphi \in C^\infty(\overline{\Delta}_n) \text{ and all } t \in (0, \infty), \tag{5.40}$$

may be perceived as a *weak formulation* of the Kolmogorov forward equation (5.19). Simultaneously, we may rewrite the initial condition[1] weakly as

$$\big[U(\,\cdot\,, 0), \varphi\big] = \big[f, \varphi\big] \quad \text{for all } \varphi \in C^\infty(\overline{\Delta}_n). \tag{5.41}$$

The advantage of such a weak formulation is that – in addition to accounting for all $\overline{\Delta}_n$ – there is no explicit regularity requirement towards the boundary (yet, we will need that its restriction to interior instances is continuously extendable to the corresponding boundary). Only, an integrability condition applies, which is at least $U, \frac{\partial}{\partial t}U, f \in \mathcal{L}^2(\Delta_n)$ since a corresponding requirement on $\partial\Delta_n$ still needs to be formulated.

Continuing our analysis of relations to the moments evolution equation, the weak equation (5.40) obviously implies the moments evolution equation (5.33), but the

[1] As the integration is over $\overline{\Delta}_n$, f may now also be formulated as an extended initial condition on the entire $\overline{\Delta}_n$. Then, $f|_{\partial\Delta_n} \neq 0$ would correspond to the process (partially) already starting on certain boundary instances. However, these parts of the process exactly evolve as a proper process of corresponding dimension, and hence do not yield any further insight into the nature of the process. For this reason, we will usually assume $f|_{\partial\Delta_n} \equiv 0$ or that f is extended that way if it is only given on Δ_n.

converse is also true: If the moments equation is given, the weak equation already holds for all $\varphi = x^\alpha$ and consequently also for all polynomials. For an arbitrary $\varphi \in C^\infty\big(\overline{\Delta}_n\big)$ we may find polynomials $P_m(x)$, $m = (m_1, \ldots, m_n)$ with

$$\big\|P_m - \varphi\big\|_{C^0(\overline{\Delta}_n)} \to 0 \quad \text{for } \min_{1\le i\le n} m_i \to \infty \tag{5.42}$$

as well as

$$\left\|\frac{\partial}{\partial x^i} P_m - \frac{\partial}{\partial x^i}\varphi\right\|_{C^0(\overline{\Delta}_n)} \to 0 \quad \text{and} \quad \left\|\frac{\partial^2}{\partial x^i \partial x^j} P_m - \frac{\partial^2}{\partial x^i \partial x^j}\varphi\right\|_{C^0(\overline{\Delta}_n)} \to 0 \tag{5.43}$$

for all $i, j = 1, \ldots, n$ and $\min_{1\le i\le n} m_i \to \infty$ (this may be shown as a generalisation of the result for 1 dimension in [32]; the 2-dimensional case is illustrated in [19]). As $\lambda\!\!\lambda_n(\Delta_n)$ (and precautionarily also $\lambda\!\!\lambda_k(\partial_k\Delta_n)$ for $0 \le k \le n-1$) are all finite, we obtain

$$\Big|\big[U, L^* P_m\big] - \big[U, L^*\varphi\big]\Big| \to 0 \quad \text{for } \min_{1\le i\le n} m_i \to \infty \tag{5.44}$$

and also

$$\left|\left[\tfrac{\partial}{\partial t}U, P_m\right] - \left[\tfrac{\partial}{\partial t}U, \varphi\right]\right| \to 0 \quad \text{for } \min_{1\le i\le n} m_i \to \infty \tag{5.45}$$

and consequently that the weak equation holds for all $\varphi \in C^\infty\big(\overline{\Delta}_n\big)$.

Summarising our findings, we have (preliminarily, as the exact product and hence the specific requirements on U are not identified yet):

5.6 Lemma. *A probability density function $U\colon \big(\overline{\Delta}_n\big)_\infty \longrightarrow \mathbb{R}$ with the corresponding moments fulfilling the moments evolution equation* (5.33) *also solves the weak formulation of the Kolmogorov forward equation* (5.70) *and conversely.*

As by lemma 5.1 a solution U of the weak equation restricted to the interior Δ_n also solves the Kolmogorov forward equation (5.19) in Δ_n and vice versa (each time assuming sufficiently regular extendibility of the solution in Δ_n to $\partial\Delta_n$), we rather focus on solutions of the weak equation as then no regularity of the solution itself up to the boundary is required, i. e. in particular functions which are defined piecewise on boundary instances are also in focus.

5.3.2 The boundary flux and a hierarchical extension of solutions

In order to construct suitable boundary values as required for an extended solution $U\colon \left(\overline{\Delta}_n\right)_\infty \longrightarrow \mathbb{R}$, the very useful concept of boundary flux already used for the 1-dimensional model in section 4.1.3 is also applied here. Restating the definition – this time for Δ_n –, the flux $G_u\colon (\Delta_n)_\infty \longrightarrow \mathbb{R}^n$ of a solution $u\colon (\Delta_n)_\infty \longrightarrow \mathbb{R}^n$ of equation (5.19) is given in terms of its components

$$G_u^i(x,t) := -\frac{1}{2}\sum_{j=1}^n \frac{\partial}{\partial x^j}\left(x^i(\delta_j^i - x^j)u(x,t)\right) = -\frac{1}{2}\sum_{j=1}^n \frac{\partial}{\partial x^j}\left(a^{ij}u(x,t)\right), \quad i=1,\dots,n, \tag{5.46}$$

thus implying $\operatorname{div} G_u = \sum_{i=1}^n \frac{\partial}{\partial x^i} G_u^i = -L_n u = -u_t$. Again, this concept directly extends to boundary instances of $\overline{\Delta}_n$ if u is extendable to the boundary such that the extension is of class C^2 with respect to the spatial variables (which is the case for a solution as in proposition 5.5).

With the flux at hand, we may now state the generalised form of lemma 5.1, which gives the adjointness relation for the Kolmogorov operators L_n and L_n^* with non-vanishing boundary terms:

5.7 Proposition. *For $n \in \mathbb{N}_+$ and $u, \varphi \in C^2\left(\overline{\Delta}_n\right)$, we have*

$$(L_n u, \varphi)_n = -\int\limits_{\partial_{n-1}\Delta_n} \varphi\, G_u \cdot \nu \, d\lambda_{n-1} + (u, L_n^*\varphi)_n \tag{5.47}$$

with G_u being the flux of u and ν the outward unit surface normal to $\partial\Delta_n$.

Proof. We use the integration by parts formula

$$\int\limits_\Omega \frac{\partial u}{\partial x^i}\varphi \, d\Omega = \int\limits_{\partial\Omega} \varphi u \, \nu^i \, d\partial\Omega - \int\limits_\Omega u \frac{\partial \varphi}{\partial x^i}\, d\Omega, \tag{5.48}$$

holding for a domain Ω with piecewise continuous boundary $\partial\Omega$, $u, \varphi \in C^1(\overline{\Omega})$ and

ν^i being the i-th component of the outward unit surface normal to $\partial\Omega$. This yields

$$
\begin{aligned}
(L_n u, \varphi)_n &= -\int\limits_{\Delta_n} \sum_i \frac{\partial}{\partial x^i} G_u^i \varphi \, d\lambda_n \\
&= -\int\limits_{\partial\Delta_n} \sum_i G_u^i \nu^i \varphi \, d\lambda_{n-1} + \int\limits_{\Delta_n} \sum_i G_u^i \frac{\partial}{\partial x^i} \varphi \, d\lambda_n.
\end{aligned}
\tag{5.49}
$$

$\partial\Delta_n \setminus \bigcup_{k=0}^{n-2} \partial_k \Delta_n$ clearly is a null set with respect to λ_{n-1}, and we may hence replace the domain of integration at the first summand by $\partial_{n-1}\Delta_n$. For the second term, we apply the integration by parts formula again (also using the modified domain of integration):

$$
\begin{aligned}
\int\limits_{\Delta_n} \sum_i G_u^i \frac{\partial}{\partial x^i} \varphi \, d\lambda_n = &-\int\limits_{\partial_{n-1}\Delta_n} \frac{1}{2} \sum_{i,j} x^i (\delta_j^i - x^j) u \nu^j \frac{\partial}{\partial x^i} \varphi \, d\lambda_{n-1} \\
&+ \int\limits_{\Delta_n} \frac{1}{2} \sum_{i,j} a^{ij} u \frac{\partial^2}{\partial x^i \partial x^j} \varphi \, d\lambda_n.
\end{aligned}
\tag{5.50}
$$

For the appearing boundary integral over $\partial_{n-1}\Delta_n = \bigcup_{l=0}^{n} \Delta_{n-1}^{(I_n \setminus \{l\})}$, we have $\nu^j = -\delta_l^j$ on $\Delta_{n-1}^{(I_n \setminus \{l\})}, l = 1, \ldots, n$ and $\nu^j = \frac{1}{\sqrt{n}}$ on $\Delta_{n-1}^{(I_n \setminus \{0\})}$, which yields

$$
\sum_j x^i (\delta_j^i - x^j) u \nu^j = -x^i (\delta_l^i - x^l) u = 0 \qquad \text{on } \Delta_{n-1}^{(I_n \setminus \{l\})} = \big\{ x^l = 0 \big\} \tag{5.51}
$$

and

$$
\begin{aligned}
\sum_j x^i (\delta_j^i - x^j) u \nu^j &= \frac{1}{\sqrt{n}} \sum_j x^i (\delta_j^i - x^j) u \\
&= \frac{1}{\sqrt{n}} x^i \Big(1 - \sum_j x^j \Big) u = 0 \quad \text{on } \Delta_{n-1}^{(I_n \setminus \{0\})} = \Big\{ 1 - \sum_j x^j = 0 \Big\},
\end{aligned}
\tag{5.52}
$$

thus the second integral over $\partial_{n-1}\Delta_n$ vanishes. Altogether, we have

$$
\begin{aligned}
(L_n u, \varphi)_n &= -\int\limits_{\partial_{n-1}\Delta_n} \sum_i G_u^i \nu^i \varphi \, d\lambda\!\!\lambda_{n-1} + \int\limits_{\Delta_n} u \frac{1}{2} \sum_{i,j} a^{ij} \frac{\partial^2}{\partial x^i \partial x^j} \varphi \, d\lambda\!\!\lambda_n && (5.53)\\
&= -\int\limits_{\partial_{n-1}\Delta_n} G_u \cdot \nu \varphi \, d\lambda\!\!\lambda_{n-1} + (u, L_n^* \varphi)_n. && \square
\end{aligned}
$$

Thus, if only employing $(\,\cdot\,,\cdot\,)_n$ for the generic product $[\,\cdot\,,\cdot\,]$ in equation (5.36), we do not yet get that the weak formulation of the Kolmogorov forward equation and hence the moments equation (5.33) is satisfied by a solution u of the Kolmogorov forward equation (5.19) – which is of no surprise as this does not account for the boundary at all. However, integrating the flux G_u on $\partial_{n-1}\Delta_n$ over time as boundary values (as in the 1-dimensional setting), already yields a limited version of the desired moments condition: For $\varphi\colon \overline{\Delta_n} \longrightarrow \mathbb{R}$ being a polynomial of degree less than 2, thus with

$$
L^* \varphi = 0, \tag{5.54}
$$

proposition 5.7 yields for a solution u of equation (5.19) resp. for the flux G_u of its continuous extension to $\partial\Delta_n$

$$
\frac{\partial}{\partial t}(u, \varphi)_n = (L_n u, \varphi)_n = -\int\limits_{\partial_{n-1}\Delta_n} G_u \cdot \nu \varphi \, d\lambda\!\!\lambda_{n-1}, \tag{5.55}
$$

which may be integrated over t, yielding

$$
(u(t), \varphi)_n = -\int\limits_{\partial_{n-1}\Delta_n} \int\limits_0^t G_u(\tau) \cdot \nu \varphi \, d\tau d\lambda\!\!\lambda_{n-1} + (u(0), \varphi)_n. \tag{5.56}
$$

For $\varphi = 1$ and $\varphi = x^i$, $i = 1, \ldots, n$, this results in:

5.8 Lemma. *For $n \in \mathbb{N}_+$ and a solution $u\colon (\Delta_n)_\infty \longrightarrow \mathbb{R}$ of the Kolmogorov forward* *equation* (5.19) *(cf. proposition 5.5), we have*

$$\int\limits_{\Delta_n} u(t)\, d\lambda\!\!\lambda_n + \int\limits_{\partial_{n-1}\Delta_n} \int\limits_0^t G_u(\tau)\cdot\nu \, d\tau d\lambda\!\!\lambda_{n-1} = \int\limits_{\Delta_n} u(0)\, d\lambda\!\!\lambda_n \qquad \equiv 1, \tag{5.57}$$

$$\int\limits_{\Delta_n} x^i u(t)\, d\lambda\!\!\lambda_n + \int\limits_{\partial_{n-1}\Delta_n} x^i \int\limits_0^t G_u(\tau)\cdot\nu \, d\tau d\lambda\!\!\lambda_{n-1} = \int\limits_{\Delta_n} x^i u(0)\, d\lambda\!\!\lambda_n \qquad \equiv p^i, \quad i = 1,\dots,n \tag{5.58}$$

for all $t \geq 0$*; the right equalities hold for* $u(0,x) = \delta_p(x)$*,* $p \in \Delta_n$*.*

Thus, we already obtain conservation laws for the total mass and the first moments if the accumulated flux on the boundary is added to the solution as an extra $(n-1)$-dimensional weight on the boundary. This accumulated flux may be calculated explicitly as the normal component of the flux $G^\perp = G \cdot \nu$ becomes on $\Delta_{n-1}^{(I_n \setminus \{l\})} \subset \partial_{n-1}\Delta_n$, $l = 1,\dots,n$

$$\begin{aligned} G\cdot\nu = -G^l &= \frac{1}{2}\sum_j \frac{\partial}{\partial x^j}(x^l(\delta_j^l - x^j)u) \\ &= \frac{1}{2}\sum_{j\neq l}(-x^l u - x^l x^j \frac{\partial}{\partial x^j}u) - \frac{1}{2}(1-2x^l)u - \frac{1}{2}x^l(1-x^l)\frac{\partial}{\partial x^l}u \\ &= \frac{1}{2}u \end{aligned} \tag{5.59}$$

and on $\Delta_{n-1}^{(I_n \setminus \{0\})} \subset \partial_{n-1}\Delta_n$

$$\begin{aligned} G\cdot\nu &= \frac{1}{\sqrt{n}}\sum_i G^i = -\frac{1}{2\sqrt{n}}\sum_{i,j}\frac{\partial}{\partial x^j}(x^i(\delta_j^i - x^j)u) \\ &= -\frac{1}{2\sqrt{n}}\Big(\sum_{\substack{i,j\\ j\neq i}}(-x^i u - x^i x^j \frac{\partial}{\partial x^i}u) + \sum_i \big((1-2x^i)u + x^i(1-x^i)\frac{\partial}{\partial x^i}u\big)\Big), \end{aligned}$$

which equals, accounting for $\sum_i x^i = 1$ and $\sum_{j,j\neq i} x^j = 1 - x^i$ on $\Delta_{n-1}^{(I_n\setminus\{0\})}$,

$$= -\frac{1}{2\sqrt{n}}\big((-n+1)u + (n-2)u\big) = \frac{1}{2\sqrt{n}}u. \tag{5.60}$$

Thus, if we introduce a rescaled version $\lambda\!\!\lambda_k^\star$ of the measure $\lambda\!\!\lambda_k$ (cf. p. 119) for $k = 1, \dots, n$ such that

$$\int\limits_{\Delta_k^{(I_k)}} d\lambda\!\!\lambda_k^\star = \frac{1}{k!} \quad \text{for all index sets } I_k \subset I_n, \tag{5.61}$$

we may rewrite equation (5.47) due to the generalised Pythagorean theorem (cf. equation (5.10)) as

$$(L_n u, \varphi) = -\frac{1}{2} \int\limits_{\partial_{n-1}\Delta_n} \varphi\, u\, d\lambda\!\!\lambda_{n-1}^\star + (u, L_n^* \varphi) \tag{5.62}$$

and reformulate the preceding lemma correspondingly:

5.9 Lemma. *For $n \in \mathbb{N}_+$ and a solution $u\colon (\Delta_n)_\infty \longrightarrow \mathbb{R}$ of the Kolmogorov forward equation* (5.19) *(cf. proposition 5.5), we have*

$$\int\limits_{\Delta_n} u(t)\, d\lambda\!\!\lambda_n + \frac{1}{2} \int\limits_{\partial_{n-1}\Delta_n} \int\limits_0^t u\, d\tau d\lambda\!\!\lambda_{n-1}^\star = \int\limits_{\Delta_n} u(0)\, d\lambda\!\!\lambda_n \qquad \equiv 1, \tag{5.63}$$

$$\int\limits_{\Delta_n} x^i u(t)\, d\lambda\!\!\lambda_n + \frac{1}{2} \int\limits_{\partial_{n-1}\Delta_n} x^i \int\limits_0^t u\, d\tau d\lambda\!\!\lambda_{n-1}^\star = \int\limits_{\Delta_n} x^i u(0)\, d\lambda\!\!\lambda_n \qquad \equiv p^i, \quad i = 1, \dots, n \tag{5.64}$$

for all $t \geq 0$; the right equalities hold for $u(0, x) = \delta_p(x)$, $p \in \Delta_n$.

However, this concept of a solution in Δ_n plus accumulated flux on the boundary $\partial_{n-1}\Delta_n$ is not yet sufficient as in general it does not yield the desired evolution laws for moments of degree 2 or higher (a more detailed discussion is postponed to p. 143), nor does $\partial_{n-1}\Delta_n$ account for the full boundary $\partial\Delta_n$. Thus, we may still strive to extend the construction of boundary data to the remaining boundary instances.

To this end, we first note that the incoming flux on $\partial_{n-1}\Delta_n$ for $n \geq 2$ should not be accumulated statically (as it is only in 0 dimensions), but rather evolve as if it was an $(n-1)$-dimensional Wright–Fisher process, i. e. as a subsolution on $\partial_{n-1}\Delta_n$. We may then carry forward the construction of boundary data to the boundary instance of subsequent lower dimension by assessing the respective boundary flux of the subsolutions on each $\partial_{n-2}\Delta_{n-1}$. Continuing this scheme successively to all boundary instances of descending dimension leads us to the following definition (cf. also the corresponding definition in equation (4.43) for the 1-dimensional case):

5.10 Definition. For $\Delta_n^{(I_n)}$ with $I_n = \{0, 1, \dots, n\}$ and a solution $u\colon \big(\Delta_n^{(I_n)}\big)_\infty \longrightarrow \mathbb{R}$ of the Kolmogorov forward equation (5.19) for given $f\colon \Delta_n^{(I_n)} \longrightarrow \mathbb{R}$ as in proposition 5.5, a *hierarchical extension*

$$U\colon \big(\overline{\Delta_n^{(I_n)}}\big)_\infty \longrightarrow \mathbb{R} \quad \text{with} \quad U(x,t) := \sum_{k=0}^{n} U_k(x,t)\chi_{\partial_k \Delta_n^{(I_n)}}(x) \tag{5.65}$$

is given by

$$U_k\colon \big(\partial_k\Delta_n^{(I_n)}\big)_\infty \longrightarrow \mathbb{R} \;\; \text{with} \;\; U_k(x,t) := \begin{cases} u(x,t) & \text{for } x \in \Delta_n^{(I_n)} \equiv \partial_n\Delta_n^{(I_n)} \\ U_{k,I_k}(x,t) & \text{for } x \in \Delta_k^{(I_k)} \subset \partial_k\Delta_n^{(I_n)}, I_k \subset I_n \\ 0 & \text{else} \end{cases} \tag{5.66}$$

for all $0 \le k \le n$ and

$$U_{k,I_k}\colon \big(\Delta_k^{(I_k)}\big)_\infty \longrightarrow \mathbb{R} \quad \text{with} \quad U_{k,I_k}(x,t) := \int_0^t u_{k,I_k}^\tau(x, t-\tau)\, d\tau \tag{5.67}$$

for $0 \le k \le n-1$ and for all subsets $I_k \subset I_n$ and with $u_{k,I_k}^\tau(x,t)\colon \big(\Delta_k^{(I_k)}\big)_\infty \longrightarrow \mathbb{R}$ being a solution of

$$\begin{cases} L_k u(x,t) = \frac{\partial}{\partial t} u(x,t) & (x,t) \in \big(\Delta_k^{(I_k)}\big)_\infty \\ u(x,0) = \sum_{I_{k+1}\supset I_k} G^\perp_{U_{k+1,I_{k+1}}}(x,\tau) & x \in \Delta_k^{(I_k)} \end{cases} \tag{5.68}$$

for all $\tau > 0$ as in proposition 5.5 and with $G^\perp_{U_{k+1,I_{k+1}}}$ being the normal component of the flux of the continuous extension of $U_{k+1,I_{k+1}}$ to $\overline{\Delta_{k+1}^{(I_{k+1})}}$.

5.11 Remark. For a given solution u of equation (5.19), the induced boundary functions U_k on $\partial_k\Delta_n^{(I_n)}$ for $0 \le k \le n-1$ in general do not fulfil the equation

$\frac{\partial}{\partial t}U_k = L_k U_k$ in some $\Delta_k^{(I_k)} \subset \partial_k\Delta_n^{(I_n)}$ and consequently are not solutions of the corresponding k-dimensional problem (5.19) in $\Delta_k^{(I_k)}$, which is due to the incoming flux of probability density from the higher-dimensional entities (cf. also the considerations on pp. 123 f.). Thus in the considered forward case, the process in a domain (if interpreted as a certain boundary part of a higher-dimensional domain) is affected by entities of all higher dimensions subsequently, whereas conversely the process gives rise to boundary processes on entities of all lower dimensions. When only considering the interior of the domain of highest dimension, boundary values may be ignored (as they are a consequence of only the process itself), and the problem is fully stated by equation (5.19).

5.3.3 An application of the hierarchical conception

As demonstrated in the preceding section, the hierarchical extension scheme of a solution in Δ_n to $\overline{\Delta}_n$ proceeds successively from the interior to boundary instances of subsequent lower dimension, while on every such a k-dimensional boundary instance a corresponding subsolution exists. The idea now is to also utilise this concept for the formulation of the moments and the related weak equation: We may thus define an integration over $\overline{\Delta}_n$ resp. redefine the (generic) product $[\,\cdot\,,\cdot\,]$ on $\overline{\Delta}_n$ from equation (5.34) such that it takes into account the descending dimensionality of boundary instances appropriately. This is essentially rendered by an integration with respect to a measure of corresponding dimension on lower-dimensional boundary instances, i. e. as in the product $[\,\cdot\,,\cdot\,]_n$ on $\overline{\Delta}_n$ (cf. equation (5.12)).

When applying this concept, we obtain for the formulation of the moments of the process in equation (5.36)

$$\bar{\mu}_\alpha(t) := \sum_{k=0}^{n} \int\limits_{\partial_k\Delta_n} U(x,t)x^\alpha \, \lambda\!\!\!\lambda_k(dx) \equiv \big[U, x^\alpha\big]_n, \quad t \geq 0,\ \alpha = (\alpha_1, \dots, \alpha_n), \qquad (5.69)$$

whereas the weak formulation of the Kolmogorov forward equation (5.40) turns into

$$\left.\begin{cases} \left[\dfrac{\partial}{\partial t}U(t), \varphi\right]_n = \big[U(t), L^*\varphi\big]_n \quad \text{for } t \in (0,\infty) \\ \big[U(\,\cdot\,,0), \varphi\big]_n = \big[f, \varphi\big]_n \end{cases}\right\} \text{ for all } \varphi \in C^\infty(\overline{\Delta}_n) \qquad (5.70)$$

with the integrability requirement then being $U(\,\cdot\,,t), \frac{\partial}{\partial t}U(\,\cdot\,,t), f \in \mathcal{L}^2\big(\bigcup_{k=0}^n \partial_k \Delta_n\big)$ for $t \geq 0$.

Even with the modified formulation in equation (5.70), we note that lemma 5.6 still holds as in the corresponding proof, no specific form of the product on $\overline{\Delta}_n$ was required. Now, we may actually give its accurate formulation:

5.12 Lemma. *A function $U\colon \big(\overline{\Delta}_n\big)_\infty \longrightarrow \mathbb{R}$, $U(\,\cdot\,,t), \frac{\partial}{\partial t}U(\,\cdot\,,t) \in \mathcal{L}^2\big(\bigcup_{k=0}^n \partial_k \Delta_n\big)$ for $t \geq 0$ with corresponding moments $[U(t), x^\alpha]_n$, $\alpha = (\alpha_1, \ldots, \alpha_n)$, $t \geq 0$ fulfilling the moments evolution equation* (5.33) *also solves the weak formulation of the Kolmogorov forward equation* (5.70) *and conversely.*

For the hierarchically extended solution and the product $[\,\cdot\,,\cdot\,]_n$, we may now prove that in continuation of lemma 5.1 and proposition 5.7 the following assertion holds:

5.13 Proposition. *A hierarchical extension $U\colon \big(\overline{\Delta_n^{(I_n)}}\big)_\infty \longrightarrow \mathbb{R}$ (cf. definition 5.10) of a solution u of the Kolmogorov forward equation* (5.19) *in Δ_n fulfils*

$$\left[\frac{\partial}{\partial t}U(t), \varphi\right]_n = \big[U(t), L^*\varphi\big]_n \tag{5.71}$$

for $\varphi \in C^\infty\big(\overline{\Delta_n^{(I_n)}}\big)$ and for all $t \in (0, \infty)$.

Proof. By proposition 5.7 we have for $U_n \equiv u$ and arbitrary $\varphi \in C^\infty\big(\overline{\Delta_n^{(I_n)}}\big)$

$$\left(\frac{\partial}{\partial t}U_n, \varphi\right)_n = \big(L_n U_n, \varphi\big)_n = -\int\limits_{\partial_{n-1}\Delta_n^{(I_n)}} \varphi\, G_{U_n}^\perp \, d\lambda\!\!\lambda_{n-1} + \big(U_n, L_n^*\varphi\big)_n \tag{5.72}$$

with $G_{U_n}^\perp = G_{U_n} \cdot \nu$ denoting the (normal) flux of the continuous extension of U_n to $\partial_{n-1}\Delta_n^{(I_n)}$. The appearing boundary integral may be expressed in terms of the evolution of the boundary function U_{n-1} that lives on $\partial_{n-1}\Delta_n^{(I_n)}$. As this implies a hierarchical dependence on the particular subprocesses, we directly start our consideration for arbitrary $k \in \{1, \ldots, n\}$. Then we have by proposition 5.7 and for all $I_k \subset I_n$

$$\int\limits_{\Delta_k^{(I_k)}} (L_k U_{k,I_k})\varphi \, d\lambda\!\!\lambda_k = -\int\limits_{\partial_{k-1}\Delta_k^{(I_k)}} \varphi\, G_{U_{k,I_k}}^\perp \, d\lambda\!\!\lambda_{k-1} + \int\limits_{\Delta_k^{(I_k)}} U_{k,I_k} L_k^*\varphi \, d\lambda\!\!\lambda_k \tag{5.73}$$

with $G_{U_{k,I_k}}$ again denoting the flux of the continuous extension of U_{k,I_k} to $\partial_{k-1}\Delta_k^{(I_k)}$ (not to be confused with the proper boundary function U_{k-1} on $\partial_{k-1}\Delta_n^{(I_n)}$). This may be summarised for the whole k-dimensional boundary $\partial_k\Delta_n^{(I_n)}$ of $\Delta_n^{(I_n)}$ by summing over all $\Delta_k^{(I_k)} \subset \partial_k\Delta_n^{(I_n)}$ resp. all subsets $I_k \subset I_n$. This yields (because of $\bigcup_{I_k\subset I_n} \Delta_k^{(I_k)} = \partial_k\Delta_n^{(I_n)}$ and the definition of U_k)

$$\int\limits_{\partial_k\Delta_n^{(I_n)}} (L_kU_k)\varphi \, d\lambda_k = \sum_{I_k\subset I_n} \int\limits_{\partial_{k-1}\Delta_k^{(I_k)}} \varphi \, G^{\perp}_{U_{k,I_k}} \, d\lambda_{k-1} + \int\limits_{\partial_k\Delta_n^{(I_n)}} U_k L_k^*\varphi \, d\lambda_k, \tag{5.74}$$

which may be rewritten by transforming the boundary term using $\bigcup_{I_k\subset I_n} \partial_{k-1}\Delta_k^{(I_k)} = \bigcup_{I_{k-1}\subset I_n} \Delta_{k-1}^{(I_{k-1})}$ and employing the product notation:

$$\big(L_kU_k, \varphi\big)_k = \sum_{I_{k-1}\subset I_n} \int\limits_{\Delta_{k-1}^{(I_{k-1})}} \varphi \sum_{I_k\supset I_{k-1}} G^{\perp}_{U_{k,I_k}} \, d\lambda_{k-1} + \big(U_k, L_k^*\varphi\big)_k. \tag{5.75}$$

Now, the sum of fluxes appearing herein may be expressed in terms of the evolution of the associated boundary function $U_{k-1,I_{k-1}}$ on $\Delta_{k-1}^{(I_{k-1})}$ for every $I_{k-1} \subset I_n$. By the chain rule, we have on $\Delta_{k-1}^{(I_{k-1})}$

$$\begin{aligned}
\frac{\partial}{\partial t} U_{k-1,I_{k-1}}(x,t) &= \frac{\partial}{\partial t} \int\limits_0^t u^{\tau}_{k-1,I_{k-1}}(x,t-\tau)\, d\tau \\
&= u^{\tau}_{k-1,I_{k-1}}(x,t-\tau)\big|_{\tau=t} + \int\limits_0^t \frac{\partial}{\partial t} u^{\tau}_{k-1,I_{k-1}}(x,t-\tau)\, d\tau \\
&= u^{t}_{k-1,I_{k-1}}(x,0) + \int\limits_0^t L_{k-1} u^{\tau}_{k-1,I_{k-1}}(x,t-\tau)
\end{aligned} \tag{5.76}$$

by the solution property of $u^{\tau}_{k-1,I_{k-1}}$. Interchanging L_{k-1} with the τ-integration and substituting $u^{t}_{k-1,I_{k-1}}(x,0)$ by the initial values as prescribed altogether yields:

$$-\sum_{I_k\supset I_{k-1}} G^{\perp}_{U_{k,I_k}}(x,t) = -\frac{\partial}{\partial t} U_{k-1,I_{k-1}}(x,t) + L_{k-1}U_{k-1,I_{k-1}}(x,t). \tag{5.77}$$

Multiplying this with φ, integrating over $\Delta_{k-1}^{(I_{k-1})}$ and summing over all $I_{k-1} \subset I_n$ results in

$$
\begin{aligned}
&- \sum_{I_{k-1} \subset I_n} \int_{\Delta_{k-1}^{(I_{k-1})}} \varphi \sum_{I_k \supset I_{k-1}} G^{\perp}_{U_{k,I_k}} \, d\lambda\!\!\lambda_{k-1} \\
&= - \sum_{I_{k-1} \subset I_n} \int_{\Delta_{k-1}^{(I_{k-1})}} \varphi \, \frac{\partial}{\partial t} U_{k-1,I_{k-1}} \, d\lambda\!\!\lambda_{k-1} + \sum_{I_{k-1} \subset I_n} \int_{\Delta_{k-1}^{(I_{k-1})}} \varphi \, L_{k-1} U_{k-1,I_{k-1}} \, d\lambda\!\!\lambda_{k-1} \\
&= - \left(\frac{\partial}{\partial t} U_{k-1}, \varphi \right)_{k-1} + \big(L_{k-1} U_{k-1}, \varphi \big)_{k-1}
\end{aligned} \tag{5.78}
$$

because of $\bigcup_{I_{k-1} \subset I_n} \Delta_{k-1}^{(I_{k-1})} = \partial_{k-1} \Delta_n^{(I_n)}$. Combining this with equation (5.75), we get

$$
\big(L_k U_k, \varphi \big)_k = - \left(\frac{\partial}{\partial t} U_{k-1}, \varphi \right)_{k-1} + \big(L_{k-1} U_{k-1}, \varphi \big)_{k-1} + \big(U_k, L_k^* \varphi \big)_k, \tag{5.79}
$$

which – by assumption – holds for all $k \in \{1, \dots, n\}$. Hence, this formula may be iterated over k, yielding

$$
\begin{aligned}
\left(\frac{\partial}{\partial t} U_n, \varphi \right)_n &= \big(L_n U_n, \varphi \big)_n \\
\Leftrightarrow \quad \left(\frac{\partial}{\partial t} U_n, \varphi \right)_n + \left(\frac{\partial}{\partial t} U_{n-1}, \varphi \right)_{n-1} &= \big(U_n, L_n^* \varphi \big)_n + \big(L_{n-1} U_{n-1}, \varphi \big)_{n-1} \\
&\;\;\vdots \\
\Leftrightarrow \quad \sum_{k=0}^{n} \left(\frac{\partial}{\partial t} U_k, \varphi \right)_k &= \sum_{k=1}^{n} \big(U_k, L_k^* \varphi \big)_k + \big(L_0 U_0, \varphi \big)_0,
\end{aligned} \tag{5.80}
$$

wherein the last summand on the right-hand side may (formally) be replaced by $\big(U_0, L_0^* \varphi \big)_0$ as they both vanish due to $L_0 = L_0^* = 0$, thus proving the assertion. □

By lemma 5.12 we immediately obtain (going beyond the results of lemma 5.8 and lemma 5.9):

5.14 Corollary. *All moments* $\bar{\mu}_\alpha(t)$, $t \geq 0$ *as defined in equation* (5.69) *of a hierarchical extension* $U \colon \big(\overline{\Delta_n^{(I_n)}}\big)_\infty \longrightarrow \mathbb{R}$ *(cf. definition 5.10) of a solution* u *of* *the Kolmogorov forward equation* (5.19) *in* Δ_n *satisfy the moments evolution equation* (5.33).

Proof. For $\varphi = 1$ and $\varphi = x^i$, we have $L^*(\varphi) = 0$, thus by equation (5.71)

$$\sum_{k=0}^{n} \left(\frac{\partial}{\partial t} U_k, \varphi \right)_k = 0. \qquad \square$$

Thus, the hierarchical extension of a solution of the Kolmogorov forward equation (5.19) via the flux of the solution yields the 'right' boundary values on the entire $\partial\Delta_n$ in the sense that all moments of the process defined via the hierarchical product $[\,\cdot\,,\cdot\,]_n$ in equation (5.69) do behave like the limit of the moments underlying discrete processes (cf. equation (4.27)), which as well confirms the specific choice of $[\,\cdot\,,\cdot\,]_n$. In particular for the total mass and the expectation value, we again have the desired conservation laws (by putting $\varphi = 1$ resp. $\varphi = x^i$ and hence $L^*(\varphi) = 0$):

$$\sum_{k=0}^{n} \int_{\partial_k \Delta_n} U_k(x,t)\, d\lambda_k = \sum_{k=0}^{n} \int_{\partial_k \Delta_n} U_k(x,0)\, d\lambda_k \equiv \int_{\Delta_n} u(x,0)\, d\lambda_n \tag{5.81}$$

and

$$\sum_{k=0}^{n} \int_{\partial_k \Delta_n} x^i U_k(x,t)\, d\lambda_k = \sum_{k=0}^{n} \int_{\partial_k \Delta_n} x^i U_k(x,0)\, d\lambda_k \equiv \int_{\Delta_n} u(x,0)\, d\lambda_n, \quad i = 1, \dots, n \tag{5.82}$$

for $t \geq 0$.

Moreover, based on the choice of the product $[\,\cdot\,,\cdot\,]_n$, we may show that any extension of a solution of the Kolmogorov forward equation (5.19) to $\overline{\Delta}_n$ yielding the correct moments already coincides with the hierarchical extension as in definition 5.10, which is due to lemma 5.12 and the following proposition:

5.15 Proposition. *For an initial condition* $f \in \mathcal{L}^2(\Delta_n)$*, a solution* $U\colon \left(\overline{\Delta}_n\right)_\infty \longrightarrow \mathbb{R}$ *of the weak Kolmogorov forward equation* (5.70) *is uniquely defined on* $\overline{\Delta}_n$.

5.16 Corollary. *For an initial condition* $f \in \mathcal{L}^2(\Delta_n)$*, a solution* $U\colon \left(\overline{\Delta}_n\right)_\infty \longrightarrow \mathbb{R}$ *of the weak Kolmogorov forward equation* (5.70) *coincides with the hierarchical extension*

$U\colon \left(\overline{\Delta}_n\right)_\infty \longrightarrow \mathbb{R}$ (cf. definition 5.10) of a solution u of the (strong) Kolmogorov forward equation (5.19) *in Δ_n.*

For the proof of proposition 5.15, we need the following lemma:

5.17 Lemma. *The linear span of $\left\{\omega_n\varphi \in C_0^\infty\left(\overline{\Delta}_n\right)\middle|\varphi \text{ eigenfunction of } L_n\right\}$ is dense in $C_c^\infty(\overline{\Delta}_n)$.*

Proof. From proposition 5.4 we already see that the linear combinations of the eigenfunctions of L_n are dense in $C^\infty(\overline{\Delta}_n)$. Dividing a function $f \in C_c^\infty(\overline{\Delta}_n)$ by ω_n (cf. lemma 5.2) again yields a function in $C_c^\infty(\overline{\Delta}_n) \subset C_0^\infty(\overline{\Delta}_n)$ as ω_n is in $C_0^\infty(\overline{\Delta}_n)$ itself and positive in the interior Δ_n. □

Proof of proposition 5.15. Assume that $U'\colon \left(\overline{\Delta}_n\right)_\infty \longrightarrow \mathbb{R}$ is another solution of equation (5.70) for a given initial condition f. We will subsequently show the accordance of U and U' on all $\partial_k\Delta_n \subset \overline{\Delta}_n$ for $k = n, \ldots, 0$. Starting with $\partial_n\Delta_n \equiv \Delta_n$, we have: For an eigenfunction $\varphi \in C^\infty(\overline{\Delta}_n)$ of L_n (corresponding to the eigenvalue λ), we obtain by lemma 5.2 that $\psi := \omega_n\varphi$ is an eigenfunction of L_n^* corresponding to the eigenvalue λ and – by nature of ω_n – $\psi \in C_0^\infty(\overline{\Delta}_n)$. For such a ψ, the weak Kolmogorov forward equation (5.70) then reduces to

$$\left(\frac{\partial}{\partial t}U, \psi\right)_n = (U, L_n^*\psi)_n \ \equiv -\lambda(U, \psi)_n \tag{5.83}$$

and

$$\left(\frac{\partial}{\partial t}U', \psi\right)_n = (U', L_n^*\psi)_n \equiv -\lambda(U', \psi)_n \tag{5.84}$$

respectively. Consequently, by t-integration we have

$$(U(t), \psi)_n = e^{-\lambda t}(U(0), \psi)_n, \tag{5.85}$$

$$(U'(t), \psi)_n = e^{-\lambda t}(U'(0), \psi)_n, \tag{5.86}$$

from which we obtain via $U(0) = U'(0) = f$

$$(U(t), \psi)_n = (U'(t), \psi)_n \quad \text{for } t \geq 0 \tag{5.87}$$

and for all eigenfunctions ψ. Since the linear span of these functions is dense in $C_c^\infty(\overline{\Delta}_n)$ (cf. lemma 5.17), U and U' agree in Δ_n.

Now, we proceed inductively. Assuming that the accordance of U and U' is already shown for all $\partial_k\Delta_n \subset \overline{\Delta}_n$ with $k > m$, then for an eigenfunction $\varphi\colon \overline{\Delta}_m \longrightarrow \mathbb{R}$ of L_m (corresponding to the eigenvalue λ), $\psi := \omega_m\varphi$ again is an eigenfunction of L_m^* corresponding to eigenvalue λ and – by nature of ω_m – $\psi \in C_0^\infty(\overline{\Delta}_m)$. From any such $\psi\colon \overline{\Delta}_m \longrightarrow \mathbb{R}$, a function $\overline{\psi}\colon \overline{\Delta_n^{(I_n)}} \longrightarrow \mathbb{R}$ may be composed, e. g. by copying ψ to $\Delta_m^{(I_m)} \subset \partial_m\Delta_n$ for all $I_m \subset I_n$ and employing convex combination of the boundary values to spread to all higher dimensional (boundary) instances subsequently while putting $\overline{\psi} := 0$ on all lower dimensional boundary instances. Of course, $\overline{\psi}$ is generally not an eigenfunction of L^* in $\overline{\Delta}_n$, but we still have $(L^*\overline{\psi})\big|_{\Delta_m^{(I_m)}} = L_m^*\psi = -\lambda\psi$ for all $\Delta_m^{(I_m)} \subset \partial\Delta_n$.

For such a $\overline{\psi}$, the weak Kolmogorov forward equation (5.70) is rendered into

$$\left(\frac{\partial}{\partial t}U, \overline{\psi}\right)_m = -\big(U, L_n^*\overline{\psi}\big)_m + \sum_{k=m+1}^{n}\left(\left(\frac{\partial}{\partial t}U, \overline{\psi}\right)_k - \big(U, L_k^*\overline{\psi}\big)_k\right) \tag{5.88}$$

and

$$\left(\frac{\partial}{\partial t}U', \overline{\psi}\right)_m = -\big(U', L_n^*\overline{\psi}\big)_m + \sum_{k=m+1}^{n}\left(\left(\frac{\partial}{\partial t}U', \overline{\psi}\right)_k - \big(U', L_k^*\overline{\psi}\big)_k\right) \tag{5.89}$$

with the sums on the right agreeing as $U = U'$ on all $\partial_k\Delta_n$ with $k > m$, hence

$$\left(\frac{\partial}{\partial t}(U - U'), \overline{\psi}\right)_m = \big(U' - U, L_n^*\overline{\psi}\big)_m \equiv \lambda\big(U' - U, \overline{\psi}\big)_m, \tag{5.90}$$

which yields – analogously to our considerations above – $U = U'$ in $\partial_m\Delta_n$ on account of the completeness of the $\overline{\psi}$'s and the initial condition. □

Thus, with the additional assumptions that the moments of the process coincide with the limits of the moments of the underlying discrete processes, we altogether have as a generalisation of the corresponding result in proposition 4.5 for the 1-dimensional model:

5.18 Theorem. *For $n \in \mathbb{N}$ and a given initial condition $f \in \mathcal{L}^2(\Delta_n)$, the Kolmogorov forward equation corresponding to the diffusion approximation of the n-dimensional Wright–Fisher model* (5.19) *always allows a unique extended solution $U\colon \left(\overline{\Delta}_n\right)_\infty \longrightarrow \mathbb{R}$ in the sense that $U|_{\Delta_n}$ is a solution of equation* (5.19) *and that its moments $\bar{\mu}_\alpha(t) := \left[U(t), x^\alpha\right]_n$, $t \geq 0$ (cf. equation* (5.69)*) fulfil the n-dimensional moments evolution equation* (5.33)*.*

Finally, it may be illustrated that the presented hierarchical extension scheme for solutions is qualitatively different from the extension scheme in 1 dimension as the obtained boundary values are no longer static as already observed on pp. 134 f. However, for an attempt of also equipping the n-dimensional setting with static boundary values, one might consider a (simple) extension $\hat{u}\colon (\overline{\Delta}_n)_\infty \longrightarrow \mathbb{R}$ of a solution u of the Kolmogorov forward equation (5.19) in Δ_n that has the t-integral of the outward normal component of the flux $\int_0^t G_u(\tau) \cdot \nu \, d\tau$ added as a supplementary $(n-1)$-dimensional mass on the boundary $\partial_{n-1}\Delta_n$ (and vanishes elsewhere on $\partial\Delta_n$). For such a solution, we have already shown in lemma 5.8 that at least total mass and expectation values are preserved.

However, we may now demonstrate that in general all higher moments do not evolve as required by the moments evolution equation (5.33): For the t-derivative of $\hat{u}$, we then have correspondingly

$$\hat{u}_t = \begin{cases} u_t = L_n u & \text{in } \Delta_n \\ G_u \cdot \nu & \text{in } \partial_{n-1}\Delta_n \\ 0 & \text{else} \end{cases} \tag{5.91}$$

and thus by virtue of proposition 5.7

$$\begin{aligned} [\hat{u}_t, \varphi]_n = (L_n u, \varphi)_n + \int\limits_{\partial_{n-1}\Delta_n} \varphi \, G_u \cdot \nu \, d\lambda\!\!\lambda_{n-1} = (u, L_n^* \varphi)_n \quad \text{for all } \varphi \in C^\infty(\overline{\Delta}_n) \\ \equiv [\hat{u}, L^* \varphi]_n - (\hat{u}, L_{n-1}^* \varphi)_{n-1}, \end{aligned} \tag{5.92}$$

signifying that $\hat{u}$ only fulfils the weak formulation of the Kolmogorov forward equation (5.70) for such a φ with $(\hat{u}, L^* \varphi)_{n-1}$ vanishing. Specifying the assertion of lemma 5.8, a sufficient condition for this would be $L_n^*\varphi|_{\partial_{n-1}\Delta_n} = 0$, which is the case for $\varphi \in \{1\} \cup \{x^i | i = 1, \dots, n\} \cup \mathcal{B}_n$ with $\mathcal{B}_n := \left\{\omega_n\psi \in C_0^\infty(\overline{\Delta}_n) \middle| \psi \text{ eigenfunction of } L_n\right\}$ (for ω_n, cf. lemma 5.2), but not generally for all $\varphi \in C^\infty(\overline{\Delta}_n)$. Hence, $\hat{u}$ is not a solution of the weak equation – which is no surprise as it is not a hierarchically extended solution on $\overline{\Delta}_n$ as in definition 5.10, which we have shown to be the unique solution. In the 1-dimensional case, though, the linear span of $\{1\} \cup \{x\} \cup \mathcal{B}_1$ is dense in $C^\infty(\overline{\Delta}_1)$ (cf. lemma 5.17), signifying that equation (5.70) is fulfilled for all $\varphi \in C^\infty(\overline{\Delta}_1)$ and that hence $\hat{u}$ is a solution of the corresponding weak equation (cf. also the proof of proposition 4.5). This again is in accordance with the model as in 1 dimension the full dynamics of the model (thus accumulation at the boundary) are captured by $\hat{u}$ (cf. also pp. 102 et seq.).

5.3.4 Conclusion and outlook

The presented methods allow a complete and valid solution of the diffusion approximation of an n-dimensional Wright–Fisher model in the sense that this model behaves analogously to what one would expect from the underlying discrete model, in particular including all results found for the 1-dimensional case. By the stated extension scheme to the boundary resp. its equivalent, the weak formulation of the Kolmogorov forward equation, it is also possible to construct a solution on the closure of the simplex explicitly (this is carried out in [30]), from which it is e. g. possible to calculate fixation probabilities for certain alleles or exit times from a given domain. For a most universal application, however, it would be expedient to generalise the model by incorporating further evolutionary mechanism as selection, mutation or recombination (and eventually coarse-graining). The corresponding Kolmogorov equations (for a diploid model) have been presented in chapter 3, and one would expect that a full solution of these generalised equations again involves a hierarchical account of the boundary structure of the simplex.

5.4 The Kolmogorov backward equation

Following our discussion of the Kolmogorov forward equation for the diffusion approximation of an n-dimensional Wright–Fisher model in the preceding section, we will now analyse its counterpart, the corresponding Kolmogorov backward equation (for the definitions cf. section 2.4.1). The equation again may be stated by making use of proposition 2.9 applied to the current setting of 1 locus and n alleles and separately stating the final condition f for $t = 0$, which corresponds to a certain (generalised) target set. At first again only considering the (open) simplex Δ_n, this yields

$$\begin{cases} -\frac{\partial}{\partial t} u(p,t) = L_n^* u(p,t) & \text{in } (\Delta_n)_{-\infty} = \Delta_n \times (-\infty, 0) \\ u(p,0) = f(p) & \text{in } \Delta_n,\ f \in \mathcal{L}^2(\Delta_n) \end{cases} \tag{5.93}$$

for $u(\,\cdot\,,t) \in C^2(\Delta_n)$ for each fixed $t \in (-\infty, 0)$ and $u(x,\,\cdot\,) \in C^1((-\infty,0))$ for each fixed $x \in \Delta_n$ and with the n-dimensional backward operator as in equation (5.21), i. e.

$$L_n^* u(p) := \frac{1}{2} \sum_{i,j=1}^{n} \big(p^i(\delta^i_j - p^j)\big) \frac{\partial^2}{\partial p^i \partial p^j} u(p). \tag{5.94}$$

In accordance with our considerations in section 4.2 for the 1-dimensional model, the backward solution $u(p,t)$ now expresses the probability of having started in some $p \in \Delta_n$ at the negative time t conditional upon being in a certain state $u(p,0) = f(p)$ at time $t = 0$, i. e. having reached the corresponding target set.

5.4.1 Solution schemes for the Kolmogorov backward equation

Since the Kolmogorov operator L_n and L_n^* are linked through the adjointness relation given in lemma 5.1, the generalised Gegenbauer polynomials in Δ_n introduced in proposition 5.4 also occur as eigenfunctions of the Kolmogorov backward operator as is illustrated in lemma 5.2. All these eigenfunctions acquired by the adjointness relation are in $C_0^\infty(\Delta_n)$, but L_n^* in Δ_n possesses even more eigenfunctions (in particular for smaller eigenvalues) as all eigenfunctions of L_k^* in Δ_k for some $0 \le k < n$ also occur as eigenfunctions of L_n^* by e. g. constant extension.

With the eigenfunctions given, the construction of a solution of equation (5.93) in Δ_n is rather straightforward. However, the – in comparison with the forward case – larger set of eigenfunctions causes ambiguities when decomposing a final condition in terms of eigenfunctions with differently evolving solutions (analogous to the 1-dimensional case), for which reason the choice of eigenfunctions needs to be restricted to the 'proper' eigenfunctions in the domain, i. e. those in $C_0^\infty(\Delta_n)$, which are derived from eigenfunctions of L_n; this is also sufficient as their linear span is already dense in $\mathcal{L}^2(\Delta_n)$ (cf. proposition 5.4 and lemma 5.17). Correspondingly, a solution by proper eigenfunctions will be called a *proper solution of the Kolmogorov backward equation in* Δ_n, and hence, the existence and uniqueness of a solution observed in the forward case directly translates into a corresponding backward counterpart of proposition 5.5; equivalent results may also be found in the literature (e. g. [22]):

5.19 Proposition. *For $n \in \mathbb{N}$ and a given final condition $f \in \mathcal{L}^2(\Delta_n)$, the Kolmogorov backward equation corresponding to the diffusion approximation of the n-dimensional Wright–Fisher model* (5.93) *always allows a unique proper solution $u\colon \big(\Delta_n\big)_{-\infty} \longrightarrow \mathbb{R}$ with $u(\,\cdot\,,t) \in C_0^\infty(\Delta_n)$ for each fixed $t \in (-\infty,0)$ and $u(x,\,\cdot\,) \in C^\infty((-\infty,0))$ for each fixed $x \in \Delta_n$.*

The notion of a proper solution in Δ_n in terms of the $C_0^\infty(\Delta_n)$-eigenfunctions also corresponds with the probabilistic interpretation of the solution as the hit probability for (subsets of) the interior Δ_n is obviously higher when starting rather in the middle of the domain than close to the boundary. This already hints at the problem that we are going to address next: the inclusion of the boundary into the model, i. e. setting up a solution scheme for the entire $\overline{\Delta}_n$. This will also clarify the role of the non-proper eigenfunctions of L_n^*, which will be interpreted in terms of an extension of proper solutions in sub-dimensional boundary instances of the domain.

5.4.2 Inclusion of the boundary and the extended Kolmogorov backward equation

Again, an inclusion of the boundary cannot be done by an extension of the proper solutions obtained in proposition 5.19 (as has been carried out in the forward case) since the continuous extension of such a solution always vanishes on the boundary by definition. However, the concept of solution still may be extended to the boundary if we reverse our angle of view and rather look for an extension of the yet to be determined, specific boundary values into the interior than for a continuation the other way round as has been implemented in the 1-dimensional case (cf. section 4.2.2). Likewise, this approach will rather match the backward setting with its reversed sense of time.

Such an extension may be accomplished by (at first formally) augmenting the domain of equation (5.93) such that it comprises the entire $\overline{\Delta}_n$. Thus, we may state an *extended Kolmogorov backward equation* by

$$\begin{cases} -\frac{\partial}{\partial t}u(p,t) = L^* u(p,t) & \text{in } \left(\overline{\Delta}_n\right)_{-\infty} = \overline{\Delta}_n \times (-\infty, 0) \\ u(p,0) = f(p) & \text{in } \overline{\Delta}_n,\ f \in \mathcal{L}^2\left(\bigcup_{k=0}^n \partial_k \Delta_n\right) \end{cases} \tag{5.95}$$

for (preliminarily at least) $u(\,\cdot\,,t)|_{\Delta_n} \in C^2(\Delta_n)$ for each fixed $t \in (-\infty,0)$ and $u(x,\,\cdot\,) \in C^1((-\infty,0))$ for each fixed $x \in \overline{\Delta}_n$.

Likewise, we may also prescribe an *extended final condition* f which is defined on $\overline{\Delta}_n$, i. e. any boundary instance may also belong to the target set considered. With n dimensions, the boundary now has a hierarchical structure, i. e. it comprises collections of simplices of different dimension. On each of them we may observe an (extended) solution as all these entities represent a certain state of the model. Correspondingly, the associated integrability criterion is such that the extended final condition needs to be of class $\mathcal{L}^2$ on every (boundary) instance of the domain.

Regarding the dynamics of such boundary functions, we note that now the configuration on the boundary is no longer static in general (which distinguishes this problem also from usual final-boundary value problems; this is only true for 0-dimensional entities), but is again subject to the same type of evolution, merely in different dimension. Hence, the dynamics may be formulated by using L_k^* with k being the corresponding dimension resp. by L_n^* restricted to the corresponding domain – as this just matches the degeneracy behaviour of L_n^* (cf. lemma 5.3). Hence, the index may be omitted, and we may just write L^* (for dimension 0, i. e. the vertices, no evolution is present, and we formally put $L^* = L_0^* := 0$ there as already described in section 4.2.2). For this reason, it is also justified by the considered model to formulate equation (5.93) to hold on the boundary as well since this exactly captures the dynamics on all lower dimensional entities of the domain.

The corresponding regularity requirement with respect to the spatial variables for a solution of equation (5.95) is that it needs to be at least of class C^2 in every boundary instance (actually, a solution typically always is of class C^∞, which likewise applies to each boundary instance). In principle, this conception would still allow for piece-wise defined solutions on $\overline{\Delta}_n$ as observed in the forward case. However, in terms of the probabilistic interpretation, a solution should be continuous when approaching a boundary since small portions of an allele – which would else be non-existent – should not affect hit probabilities substantially; this, however, only holds true for a single allele in question and hence for boundary relations with exactly one dimension in difference, i. e. for $\Delta_k^{(I_k)}$ and a boundary face $\Delta_{k-1} \subset \partial_{k-1}\Delta_k^{(I_k)}$. Correspondingly, with the restriction property of L^* given, a natural assumption would be that L^*u is continuous up to the boundary, for which $u \in C^2(\Delta_{k-1} \cup \Delta_k^{(I_k)})$ with respect to the spatial variables – respectively $u \in C^2(\partial_{k-1}\Delta_k^{(I_k)} \cup \Delta_k^{(I_k)})$ if all relevant boundary faces shall be accounted for – is sufficient.

Globally, we may thus require that such a property applies to all possible boundary transitions within $\overline{\Delta}_n$ and define correspondingly for $l \in \mathbb{N} \cup \{\infty\}$

$$u \in C_p^l\big(\overline{\Delta}_n\big) \;:\Leftrightarrow\; u|_{\Delta_d^{(I_d)} \cup \partial_{d-1}\Delta_d^{(I_d)}} \in C^l(\Delta_d^{(I_d)} \cup \partial_{d-1}\Delta_d^{(I_d)}) \quad \text{for all } I_d \subset I_n,\ 1 \leq d \leq n \tag{5.96}$$

with respect to the spatial variables, implying that L^* is continuous at all boundary transitions within $\overline{\Delta}_n$. For later purposes, we also introduce a special version for ascending chains of (sub-)simplices with a more specific boundary condition and hence put for index sets $I_k \subset \ldots \subset I_n$ and again for $l \in \mathbb{N} \cup \{\infty\}$

$$u \in C_{p_0}^l\Big(\bigcup_{d=k}^{n} \Delta_d^{(I_d)}\Big) \;:\Leftrightarrow\; \begin{cases} u|_{\Delta_d^{(I_d)}} \text{ is extendable to } \bar{u} \in C^l(\Delta_d^{(I_d)} \cup \partial_{d-1}\Delta_d^{(I_d)}) \text{ with} \\ \bar{u}|_{\partial_{d-1}\Delta_d^{(I_d)}} = u\chi_{\Delta_{d-1}^{(I_{d-1})}}\chi_{\{d>k\}} \text{ for all } \max(1,k) \leq d \leq n \end{cases} \tag{5.97}$$

with respect to the spatial variables. We note that such a function may straightforwardly be completed into a function defined on the entire $\overline{\Delta}_n$ by putting $u := 0$ on

$\overline{\Delta}_n \setminus \big(\bigcup_{d=k}^{n} \Delta_d^{(I_d)}\big)$ for corresponding t; however, such an extension is generally not of class $C_p^l\big(\overline{\Delta}_n\big)$ w. r. t. the spatial variables.

These specific regularity properties are better adapted to the model than a global regularity property for $\overline{\Delta}_n$ (as this would limit the choice of solutions excessively otherwise), and we may hence use a regularity as in equality (5.97) as the standard requirement w. r. t. the spatial variables for solutions of the extended Kolmogorov backward equation (5.95), with $l = 2$ as needed for L^* (this will also apply when talking of a 'smooth' function throughout the remainder of this chapter; likewise, 'smoothly extendable' is meant to signify that a continuous extension exists and is at least of class C^2). For the sake of completeness, we also list a corresponding property for the final condition being

$$f \in \mathcal{L}^2\Big(\bigcup_{d=k}^{n} \Delta_d^{(I_d)}\Big) :\Leftrightarrow \begin{cases} f|_{\Delta_d^{(I_d)}} \text{ is } \lambda\!\!\lambda_d\text{-measurable and} \\ \int_{\Delta_d^{(I_d)}} |f(x)|^2 \, \lambda\!\!\lambda_d(dx) < \infty \text{ for all } d = k, \dots, n, \end{cases} \tag{5.98}$$

which is a refined version of the previous definition in equation (5.14).

Within the extended setting, we again expect the solution to depend crucially on the situation on the boundary: Whenever some boundary instance is included into the target set, this alters the probability of having started there for all accessible boundary instances of higher dimension as the process inevitably tends towards corresponding (lower-dimensional) boundary entities.

5.4.3 An extension scheme for solutions of the Kolmogorov backward equation

Having stated the extended Kolmogorov backward equation, we will now present an extension scheme for solutions of the Kolmogorov backward equation (5.93) in boundary instances of the considered domain (this may of course be a boundary instance itself and consequently may allow successive extensions), which will eventually lead us to a construction of solutions of the extended Kolmogorov backward equation (5.95). As in the 1-dimensional case (cf. p. 109), we need to determine which requirements are to be met by such an extension, and hence we define for an extension from some simplex of arbitrary dimension $d-1$ to the simplex of subsequent higher dimension (however, $d = 1$ requires a somewhat specific treatment; a probabilistic analysis of these conditions will be provided in the next section):

5.20 Definition (extension constraints)**.** Let I_d be an index set with $|I_d| = d+1 \geq 2$, $0, s \in I_d$ and $\Delta_d^{(I_d)} = \{(p^i)_{i\in I_d\setminus\{0\}} | p^i > 0 \text{ for } i \in I_d\}$ with $p^0 := 1 - \sum_{i\in I_d\setminus\{0\}} p^i$. For $d \geq 2$ and a solution $u\colon \big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)_{-\infty} \longrightarrow \mathbb{R}$ of the Kolmogorov backward equation (5.93), i. e. $u(\,\cdot\,, t) \in C^\infty\big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)$ for $t < 0$, $u(x, \,\cdot\,) \in C^\infty((-\infty, 0))$ for $x \in \Delta_{d-1}^{(I_d\setminus\{s\})}$ and

$$-\frac{\partial}{\partial t} u = L^* u \quad \text{in } \big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)_{-\infty}, \tag{5.99}$$

a function $\bar{u}\colon \big(\Delta_d^{(I_d)}\big)_{-\infty} \longrightarrow \mathbb{R}$ with $\bar{u}(\,\cdot\,, t) \in C^\infty\big(\Delta_d^{(I_d)}\big)$ for $t < 0$ and $\bar{u}(x, \,\cdot\,) \in C^\infty((-\infty, 0))$ for $x \in \Delta_d^{(I_d)}$ is said to be an extension of u in accordance with the *extension constraints* if

(i) for $t < 0$ $\bar{u}(\,\cdot\,, t)$ is continuously extendable to the boundary $\partial_{d-1}\Delta_d^{(I_d)}$ such that it coincides with $u(\,\cdot\,, t)$ in $\Delta_{d-1}^{(I_d\setminus\{s\})}$ resp. vanishes on the remainder of $\partial_{d-1}\Delta_d^{(I_d)}$ and is of class C^∞ with respect to the spatial variables in $\Delta_d^{(I_d)} \cup \partial_{d-1}\Delta_d^{(I_d)}$,

(ii) it is a solution of the corresponding Kolmogorov backward equation in $\big(\Delta_d^{(I_d)}\big)_{-\infty}$, i. e. $-\frac{\partial}{\partial t}\bar{u} = L^*\bar{u}$ in $\big(\Delta_d^{(I_d)}\big)_{-\infty}$.

For $d = 1$, this analogously applies to functions u with $-\frac{\partial}{\partial t}u = 0$ (in accordance with $L_0^* \equiv 0$), and consequently the equation in condition (ii) is replaced with $L^*\bar{u} = 0$. Furthermore, an extension which encompasses multiple extension steps is said to be in accordance with the extension constraints, if this holds for every extension step.

5.21 Remark. In case of $d \geq 2$, if u for $t < 0$ extends smoothly to the boundary $\partial_{d-2}\Delta_{d-1}^{(I_d\setminus\{s\})}$ such that this extension vanishes everywhere on $\partial_{d-2}\Delta_{d-1}^{(I_d\setminus\{s\})}$, the above definition corresponds to $(u\chi_{\Delta_{d-1}^{(I_d\setminus\{s\})}} + \bar{u}\chi_{\Delta_d^{(I_d)}}) \in C_{p_0}^\infty(\Delta_{d-1}^{(I_d\setminus\{s\})} \cup \Delta_d^{(I_d)})$ with respect to the spatial variables for $t < 0$ (cf. equality (5.97)) except for the Kolmogorov backward equation solution property.

In the following, we will mainly be concerned with the existence of such extensions which comply with definition 5.20; the issue of their uniqueness will be dealt with in the next chapter. Corresponding to the chosen separation ansatz (on which the result 5.19 is based), the foundation for the construction of an extension as desired is such of the eigenmodes; the following lemma generalises the existence procedure applied for extensions from $\partial_0\Delta_1$ in the 1-dimensional case to arbitrary dimension:

5.22 Lemma (extension of eigenfunctions). *Let I_d be an index set with $|I_d| = d+1 \geq 2$, $0, s \in I_d$ and $\Delta_d^{(I_d)} = \{(p^i)_{i\in I_d\setminus\{0\}} | p^i > 0$ for $i \in I_d\}$ with $p^0 := 1 - \sum_{i\in I_d\setminus\{0\}} p^i$. For $d \geq 2$ and an eigenfunction $\psi \in C^\infty\big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)$ of L_{d-1}^* for the eigenvalue $\kappa \geq 0$, i. e.*

$$L_{d-1}^*\psi = -\kappa\psi \quad \text{ in } \Delta_{d-1}^{(I_d\setminus\{s\})} \subset \partial\Delta_d^{(I_d)}, \tag{5.100}$$

a linear interpolation $\bar{\psi} = \bar{\psi}^{r,s} \colon \Delta_d^{(I_d)} \longrightarrow \mathbb{R}$ of ψ from $\Delta_{d-1}^{(I_d\setminus\{s\})}$ (source face) towards $\Delta_{d-1}^{(I_d\setminus\{r\})} \subset \partial_{d-1}\Delta_d^{(I_d)}$ for some $r \in I_d \setminus \{s\}$ (target face) is given by

$$\bar{\psi}^{r,s}(p) := \psi(\pi^{r,s}(p)) \cdot \frac{p^r}{p^s + p^r} \quad \text{for } p \in \Delta_d^{(I_d)} \tag{5.101}$$

with $\pi^{r,s}(p^1, \dots, p^d) = (\tilde{p}^1, \dots, \tilde{p}^d)$ such that $\tilde{p}^s = 0$, $\tilde{p}^r = p^s + p^r$ and $\tilde{p}^i = p^i$ for $i \in I_d \setminus \{s, r\}$.

Then $\bar{\psi}$ features regularity corresponding to that of as ψ in $\Delta_d^{(I_d)}$ (i. e. is of class C^∞) and fulfils

$$L_d^*\bar{\psi} = -\kappa\bar{\psi} \quad \text{in } \Delta_d^{(I_d)}. \tag{5.102}$$

Moreover, $\bar{\psi}$ extends smoothly to $\Delta_{d-1}^{(I_d\setminus\{s\})}$ and $\Delta_{d-1}^{(I_d\setminus\{r\})}$, and there we have

$$\bar{\psi}|_{\Delta_{d-1}^{(I_d\setminus\{s\})}} = \psi, \qquad \bar{\psi}|_{\Delta_{d-1}^{(I_d\setminus\{r\})}} = 0. \tag{5.103}$$

If furthermore ψ extends smoothly to $\Delta_{d-2}^{(I_d\setminus\{s,q\})} \subset \partial_{d-2}\Delta_d^{(I_d\setminus\{s\})}$ for some $q \in I_d\setminus\{r, s\}$, then $\bar{\psi}$ likewise extends smoothly to $\Delta_{d-1}^{(I_d\setminus\{q\})}$. In particular, $\bar{\psi}$ fulfils the extension constraint 5.20 (i) if ψ extends smoothly to $\partial_{d-2}\Delta_{d-1}^{(I_d\setminus\{s\})} \setminus \Delta_{d-2}^{(I_d\setminus\{r,s\})}$ and vanishes there.

For $d = 1$, the preceding statements analogously hold for arbitrary $\psi \colon \Delta_0^{(I_1\setminus\{s\})} \longrightarrow \mathbb{R}$ as eigenfunction of $L_0^* \equiv 0$ for the eigenvalue 0; then, $\bar{\psi}$ is of class C^∞ in $\Delta_1^{(I_1)}$, and such an extension is always in accordance with the extension constraint 5.20 (i).

Since the eigenfunctions are the building blocks for a solution scheme, the preceding lemma directly extends to solutions of the Kolmogorov backward equation:

5.23 Proposition (extension of solutions)**.** *Let I_d be an index set with $|I_d| = d+1 \geq 2$, $0, s \in I_d$ and $\Delta_d^{(I_d)} = \{(p^i)_{i\in I_d\setminus\{0\}} | p^i > 0$ for $i \in I_d\}$ with $p^0 := 1 - \sum_{i\in I_d\setminus\{0\}} p^i$. For $d \geq 2$, a given final condition $f \in \mathcal{L}^2\big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)$ and a given extension target face index $r \in I_d \setminus \{s\}$, a solution $u\colon \big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)_{-\infty} \longrightarrow \mathbb{R}$ of the Kolmogorov backward equation* (5.93)*, $u(\,\cdot\,,t) \in C^\infty\big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)$ for $t < 0$ and $u(x,\,\cdot\,) \in C^\infty((-\infty,0))$ for $x \in \Delta_{d-1}^{(I_d\setminus\{s\})}$, may be extended to a function*

$$\bar{u} = \bar{u}^{r,s}\colon \big(\Delta_d^{(I_d)}\big)_{-\infty} \longrightarrow \mathbb{R} \tag{5.104}$$

with $\bar{u}(\,\cdot\,,t) \in C^\infty\big(\Delta_d^{(I_d)}\big)$ for $t < 0$ and $\bar{u}(x,\,\cdot\,) \in C^\infty((-\infty,0))$ for $x \in \Delta_d^{(I_d)}$ as well as fulfilling

$$-\frac{\partial}{\partial t}\bar{u} = L^*\bar{u} \quad \textit{in } \big(\Delta_d^{(I_d)}\big)_{-\infty}. \tag{5.105}$$

Furthermore, for $t < 0$ $\bar{u}(\,\cdot\,,t)$ smoothly extends to the boundary in $\Delta_{d-1}^{(I_d\setminus\{s\})}$ with

$$\bar{u}(\,\cdot\,,t)\big|_{\Delta_{d-1}^{(I_d\setminus\{s\})}} = u, \quad \textit{in particular} \quad \bar{u}(\,\cdot\,,0)\big|_{\Delta_{d-1}^{(I_d\setminus\{s\})}} = f\big|_{\Delta_{d-1}^{(I_d\setminus\{s\})}} \tag{5.106}$$

and in $\Delta_{d-1}^{(I_d\setminus\{r\})}$ with $\bar{u}(\,\cdot\,,t)\big|_{\Delta_{d-1}^{(I_d\setminus\{r\})}} = 0$. If furthermore $u(\,\cdot\,,t)$ for $q \in I_d \setminus \{r,s\}$ extends smoothly to $\Delta_{d-2}^{(I_d\setminus\{q,s\})} \subset \partial_{d-2}\Delta_d^{(I_d\setminus\{s\})}$ for some t, then $\bar{u}(\,\cdot\,,t)$ likewise extends smoothly to $\Delta_{d-1}^{(I_d\setminus\{q\})}$. In particular, $\bar{u}$ fulfils the extension constraints 5.20 if $u(\,\cdot\,,t)$ extends smoothly to $\partial_{d-2}\Delta_{d-1}^{(I_d\setminus\{s\})} \setminus \Delta_{d-2}^{(I_d\setminus\{r,s\})}$ and vanishes there for $t < 0$.

For $d = 1$, the antecedent analogously holds for functions $u\colon \big(\Delta_0^{(I_1\setminus\{s\})}\big)_{-\infty} \longrightarrow \mathbb{R}$ with $u(x,\cdot\,) \in C^\infty((-\infty,0))$ and $\frac{\partial}{\partial t}u = 0$; then, $\bar{u}(\,\cdot\,,t)$ is of class C^∞ in $\Delta_1^{(I_1)}$ for every t as well as $\bar{u}(x,\,\cdot\,) \in C^\infty((-\infty,0))$ for $x \in \Delta_d^{(I_d)}$ with $\frac{\partial}{\partial t}\bar{u} = 0$, and equation (5.105) *holds correspondingly. Furthermore, this extension always is in accordance with the extension constraints 5.20.*

5.24 Remark. The extension of a solution of the Kolmogorov backward equation for a final condition $f \in \mathcal{L}^2\big(\Delta_{d-1}^{(I_d\setminus\{s\})}\big)$ as in proposition 5.23 is also applicable for $t = 0$, yielding an analogously extended final condition $\bar{f} = \bar{f}^{r,s} \in \mathcal{L}^2\big(\Delta_d^{(I_d)}\big)$. We then have $\bar{u}(\,\cdot\,,0) \equiv \bar{f}$ in $\Delta_d^{(I_d)}$ by continuous extension as we have $u(\,\cdot\,,0) = f$ in $\Delta_{d-1}^{(I_d \setminus \{s\})}$; however, for $d \geq 2$ this extension of f in general does not have the boundary regularity described due to the missing regularity of f (and hence in general does not satisfy the extension boundary constraint 5.20 (i)).

In addition to the preceding proposition, it should be noted that $\bar{u}$ does not necessarily extend continuously to the entire $\overline{\Delta}_d$, in particular not to the remaining boundary parts of dimension $d-2$ and less. This is due to the fact that on instances of $\partial_{d-2}\Delta_d^{(I_d)}$, which are shared boundaries of higher-dimensional faces of the simplex, continuous extensions from each of those faces regularly may exists, but do not necessarily correspond.

Proof of lemma 5.22. The regularity assertion for $\bar{\psi}$ in $\Delta_d^{(I_d)}$ follows from the regularity of π and of the projection and from $\frac{p^r}{p^s+p^r}$ being of class C^∞ on $\Delta_d^{(I_d)}$. The boundary behaviour is similarly straightforward as $\pi^{r,s} = \mathrm{id}$ and $\frac{p^r}{p^s+p^r} = 1$ on $\Delta_{d-1}^{(I_d \setminus \{s\})}$, whereas $\frac{p^r}{p^s+p^r} = 0$ on $\Delta_{d-1}^{(I_d \setminus \{r\})}$. Both boundary extensions are smooth in the sense described, which is again due to the regularity of the projection and of $\frac{p^r}{p^s+p^r}$ when approaching $\Delta_{d-1}^{(I_d \setminus \{s\})}$ resp. $\Delta_{d-1}^{(I_d \setminus \{r\})}$. Analogous considerations yield the assertion for other boundary faces of $\partial_{d-1}\Delta_d^{(I_d)}$: The projection $\pi^{r,s}$ maps $\partial_{d-1}\Delta_d^{(I_d)} \setminus \big(\Delta_{d-1}^{(I_d \setminus \{r\})} \cup \Delta_{d-1}^{(I_d \setminus \{s\})}\big)$ smoothly onto $\partial_{d-2}\Delta_{d-1}^{(I_d \setminus \{s\})}$, which together with $\frac{p^r}{p^s+p^r}$ being of class C^∞ on $\partial_{d-1}\Delta_d^{(I_d)}$ (via $p^s + p^r > 0$) yields the stated regularity; the value of this boundary extension of $\bar{\psi}$ of course coincides with the one of the corresponding extension of ψ.

To prove equation (5.102), w. l. o. g. let $I_d = \{0, 1, \dots, d\}$; summation indices, however, run from 1 to d if nothing differing is stated. To begin with, we have

$$\begin{aligned} L_d^*\left(\psi(\pi^{r,s}(p)) \cdot \frac{p^r}{p^s+p^r}\right) = {} & \left(L_d^*\psi(\pi^{r,s}(p))\right)\frac{p^r}{p^s+p^r} \\ & + \sum_{i,j} p^i(\delta^i_j - p^j)\left(\frac{\partial}{\partial p^i}\psi(\pi^{r,s}(p))\right)\left(\frac{\partial}{\partial p^j}\frac{p^r}{p^s+p^r}\right) \\ & + \frac{1}{2}\psi(\pi^{r,s}(p))\sum_{i,j} p^i(\delta^i_j - p^j)\left(\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\frac{p^r}{p^s+p^r}\right). \end{aligned} \tag{5.107}$$

Next, we will show that the first summand equals $-\kappa\bar{\psi}$, whereas the two other summands vanish on $\Delta_d^{(I_d)}$.

For the first summand, we use $L^*_{d-1}\psi = -\kappa\psi$ in $\Delta_{d-1}^{(I_d\setminus\{s\})}$, which holds by assumption. To extend this statement to $\Delta_d^{(I_d)}$, the interplay of the projection needs to be analysed, for which several cases are distinguished. That is, for $s \neq 0$, $r = 0$, the projection $\pi^{0,s}$ yields $\tilde{p}^s = 0$ and $\tilde{p}^i = p^i$ for $i \in \{1, \dots, d\} \setminus \{s\}$, hence $\frac{\partial \tilde{p}^m}{\partial p^i} = \delta_i^m(1-\delta_s^m)$, and we have

$$\begin{aligned} L_d^*\psi(\pi^{0,s}(p)) &= \frac{1}{2}\sum_{i,j} p^i(\delta^i_j - p^j)\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\psi(\pi^{0,s}(p)) \\ &= \frac{1}{2}\sum_{m,n}\sum_{i,j} p^i(\delta^i_j - p^j)\delta^m_i(1-\delta^m_s)\delta^n_j(1-\delta^n_s)\frac{\partial}{\partial \tilde{p}^m}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) \\ &= \frac{1}{2}\sum_{m,n\neq s} \tilde{p}^m(\delta^m_n - \tilde{p}^n)\frac{\partial}{\partial \tilde{p}^m}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) = L^*_{d-1}\psi(\tilde{p}) \equiv -\kappa\psi(\tilde{p}). \end{aligned} \tag{5.108}$$

If $s = 0$, $r \neq 0$ and hence $\Delta_{d-1}^{(I_d\setminus\{0\})} = \big\{(\tilde{p}^1, \dots, \tilde{p}^d)\big|\tilde{p}^i > 0 \text{ for } i = 1, \dots, d, \sum_{i=1}^d \tilde{p}^i = 1\big\}$, we have $\tilde{p}^i = p^i$ for $i \in \{1, \dots, d\} \setminus \{r\}$ and $\tilde{p}^r = p^r + p^0$, thus $\frac{\partial \tilde{p}^m}{\partial p^i} = \delta^m_i - \delta^m_r$. We get:

$$\begin{aligned} L_d^*\psi(\pi^{r,0}(p)) &= \frac{1}{2}\sum_{m,n}\sum_{i,j} p^i(\delta^i_j - p^j)(\delta^m_i - \delta^m_r)(\delta^n_j - \delta^n_r)\frac{\partial}{\partial \tilde{p}^m}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) \\ &= \frac{1}{2}\sum_{m,n} p^m(\delta^m_n - p^n)\frac{\partial}{\partial \tilde{p}^m}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) - \frac{1}{2}\sum_n\sum_i p^i(\delta^i_n - p^n)\frac{\partial}{\partial \tilde{p}^r}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) \\ &\quad - \frac{1}{2}\sum_m\sum_j p^m(\delta^m_j - p^m)\frac{\partial}{\partial \tilde{p}^r}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) + \frac{1}{2}\sum_{i,j} p^i(\delta^i_j - p^j)\frac{\partial^2}{(\partial \tilde{p}^r)^2}\psi(\tilde{p}) \\ &= \frac{1}{2}\sum_{m,n} p^m(\delta^m_n - p^n)\frac{\partial}{\partial \tilde{p}^m}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) - \frac{1}{2}\sum_n p^0p^n\frac{\partial}{\partial \tilde{p}^r}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) \\ &\quad - \frac{1}{2}\sum_m p^mp^0\frac{\partial}{\partial \tilde{p}^m}\frac{\partial}{\partial \tilde{p}^r}\psi(\tilde{p}) + \frac{1}{2}p^0(1-p^0)\frac{\partial^2}{(\partial \tilde{p}^r)^2}\psi(\tilde{p}). \end{aligned} \tag{5.109}$$

When replacing the remaining p-coordinates by $\tilde{p}$ (except for p^0, which is missing in $\Delta_{d-1}^{(I_d\setminus\{0\})}$) via $p^i = \tilde{p}^i - p^0\delta^i_r$ for $i = \{1, \dots, d\}$, the expression transforms into:

$$L_d^*\psi(\pi^{r,0}(p)) = \frac{1}{2}\sum_{m,n\neq r} \tilde{p}^m(\delta^m_n - \tilde{p}^n)\frac{\partial}{\partial \tilde{p}^m}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p}) + \frac{1}{2}\sum_{n\neq r}(-\tilde{p}^r + p^0)\tilde{p}^n\frac{\partial}{\partial \tilde{p}^r}\frac{\partial}{\partial \tilde{p}^n}\psi(\tilde{p})$$

$$
\begin{aligned}
&+\frac{1}{2}\sum_{m\neq r}\tilde{p}^m(-\tilde{p}^r+p^0)\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^r}\psi(\tilde{p})+\frac{1}{2}(\tilde{p}^r-p^0)(1-\tilde{p}^r+p^0)\times\\
&\frac{\partial^2}{(\partial\tilde{p}^r)^2}\psi(\tilde{p})-\frac{1}{2}\sum_{n\neq r}p^0\tilde{p}^n\frac{\partial}{\partial\tilde{p}^r}\frac{\partial}{\partial\tilde{p}^n}\psi(\tilde{p})-\frac{1}{2}\sum_{m\neq r}\tilde{p}^m p^0\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^r}\psi(\tilde{p})\\
&-p^0(\tilde{p}^r-p^0)\frac{\partial^2}{(\partial\tilde{p}^r)^2}\psi(\tilde{p})+\frac{1}{2}p^0(1-p^0)\frac{\partial^2}{(\partial\tilde{p}^r)^2}\psi(\tilde{p})\\
=&\frac{1}{2}\sum_{m,n}\tilde{p}^m(\delta^m_n-\tilde{p}^n)\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^n}\psi(\tilde{p})=L^*_{d-1}\psi(\tilde{p})\equiv-\kappa\psi(\tilde{p}).
\end{aligned}
\tag{5.110}
$$

The next-to-last equality is due to the fact that in $\Delta_{d-1}^{(I_d\setminus\{0\})}$ one coordinate is obsolete and consequently ψ is formulated in $d-1$ coordinates (which may be chosen freely). It is straightforward to show that, independently of the choice of the omitted coordinate r, we have $L^*_{d-1}=\frac{1}{2}\sum_{m,n\neq r}\tilde{p}^m(\delta^m_n-\tilde{p}^n)\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^n}$ on $\Delta_{d-1}^{(I_d\setminus\{0\})}$.

Lastly, if $s\neq0$, $r\neq0$, the projection $\pi^{r,s}$ yields $\tilde{p}^s=0$, $\tilde{p}^r=p^s+p^r$ and $\tilde{p}^i=p^i$ for the remaining indices, hence $\frac{\partial\tilde{p}^m}{\partial p^i}=\delta^m_i(1-\delta^m_s)+\delta^m_r\delta^i_s$. Then we have:

$$
\begin{aligned}
L^*_d\psi(\pi^{r,s}(p))&=\frac{1}{2}\sum_{i,j}p^i(\delta^i_j-p^j)\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\psi(\pi^{r,s}(p))\\
&=\frac{1}{2}\sum_{\substack{m,n,\\i,j}}p^i(\delta^i_j-p^j)(\delta^m_i(1-\delta^m_s)+\delta^m_r\delta^i_s)(\delta^n_j(1-\delta^n_s)+\delta^n_r\delta^j_s)\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^n}\psi(\tilde{p})\\
&=\frac{1}{2}\sum_{m,n\neq s}p^m(\delta^m_n-p^n)\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^n}\psi(\tilde{p})-\frac{1}{2}\sum_{n\neq s}p^s p^n\frac{\partial}{\partial\tilde{p}^r}\frac{\partial}{\partial\tilde{p}^n}\psi(\tilde{p})\\
&\quad-\frac{1}{2}\sum_{m\neq s}p^m p^s\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^r}\psi(\tilde{p})+\frac{1}{2}p^s(1-p^s)\frac{\partial^2}{(\partial\tilde{p}^r)^2}\psi(\tilde{p}).
\end{aligned}
\tag{5.111}
$$

Replacing the p-coordinates works as shown in the preceding case, and thereupon we obtain

$$
L^*_d\psi(\pi^{r,s}(p))=\frac{1}{2}\sum_{m,n\neq s}\tilde{p}^m(\delta^m_n-\tilde{p}^n)\frac{\partial}{\partial\tilde{p}^m}\frac{\partial}{\partial\tilde{p}^n}\psi(\tilde{p})=L^*_{d-1}\psi(\tilde{p})\equiv-\kappa\psi(\tilde{p}),\tag{5.112}
$$

thus in total

$$
L^*_d\psi(\pi^{r,s}(p))=L^*_{d-1}\psi(\tilde{p})\equiv-\kappa\psi(\tilde{p})=-\kappa\psi(\pi^{r,s}(p))\tag{5.113}
$$

for arbitrary r, s, which is the desired equality result for the first summand.

To show that the two remaining summands vanish, an analogous case-by-case analysis is necessary. If $s = 0$, $r \neq 0$, we have $\frac{p^r}{p^0+p^r} = \frac{p^r}{1-\sum_{l\neq r} p^l}$. Due to (remember $\frac{\partial \tilde{p}^m}{\partial p^i} = \delta_i^m - \delta_r^m$)

$$\frac{\partial}{\partial p^r}\psi(\pi^{r,0}(p)) = \sum_m \frac{\partial \tilde{p}^m}{\partial p^r}\frac{\partial}{\partial \tilde{p}^m}\psi(\tilde{p}) = 0, \tag{5.114}$$

the second summand equalling

$$\begin{aligned}\sum_{i\neq r} p^i \left(\frac{\partial}{\partial p^i}\psi(\pi^{r,0}(p))\right) &\underbrace{\sum_j (\delta_j^i - p^j)\left(\frac{\partial}{\partial p^j}\frac{p^r}{1-\sum_{l\neq r}p^l}\right)} \\ &= \big(1-\sum_{j\neq r}p^j\big)\frac{p^r}{\big(1-\sum_{l\neq r}p^l\big)^2} - p^r\frac{1}{1-\sum_{l\neq r}p^l} = 0\end{aligned} \tag{5.115}$$

along with the third summand equalling

$$\begin{aligned}&\frac{1}{2}\psi(\pi^{r,0}(p))\sum_{i\neq r}\left(\sum_{j\neq r}p^i(\delta_j^i - p^j)\left(\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\frac{p^r}{1-\sum_{l\neq r}p^l}\right)\right. \\ &\qquad\qquad\left. - 2p^ip^r\left(\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^r}\frac{p^r}{1-\sum_{l\neq r}p^l}\right)\right) \\ &= \frac{1}{2}\psi(\pi^{r,0}(p))\sum_{i\neq r}\left(p^i(1-\sum_{j\neq r}p^j)\frac{2p^r}{\big(1-\sum_{l\neq r}p^l\big)^3} - 2p^ip^r\frac{1}{\big(1-\sum_{l\neq r}p^l\big)^2}\right) = 0\end{aligned} \tag{5.116}$$

vanish.

Similarly, if $s \neq 0$, $r = 0$, thus $\frac{p^0}{p^s+p^0} = \frac{1-\sum_l p^l}{1-\sum_{l\neq s}p^l}$ and again (with $\frac{\partial \tilde{p}^m}{\partial p^i} = \delta_i^m(1-\delta_s^m)$)

$$\frac{\partial}{\partial p^s}\psi(\pi^{0,s}(p)) = \sum_m \frac{\partial \tilde{p}^m}{\partial p^s}\frac{\partial}{\partial \tilde{p}^m}\psi(\tilde{p}) = 0, \tag{5.117}$$

the second summand equalling

$$\sum_{i\neq s} p^i \left(\frac{\partial}{\partial p^i}\psi(\pi^{0,s}(p))\right) \underbrace{\sum_j (\delta^i_j - p^j)\left(\frac{\partial}{\partial p^j}\frac{1-\sum_l p^l}{1-\sum_{l\neq s} p^l}\right)}$$
$$= \big(1-\sum_j p^j\big)\frac{-1}{1-\sum_{l\neq s} p^l} + \big(1-\sum_{j\neq s} p^j\big)\frac{1-\sum_l p^l}{\big(1-\sum_{l\neq s} p^l\big)^2} = 0 \tag{5.118}$$

vanishes, and the third summand via

$$\begin{aligned}
&\sum_{i,j} p^i(\delta^i_j - p^j)\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\frac{1-\sum_l p^l}{1-\sum_{l\neq s} p^l}\\
&= \sum_{i,j} p^i(\delta^i_j - p^j)\left(\frac{(\delta^i_s - 1)+(\delta^j_s - 1)}{\big(1-\sum_{l\neq s} p^l\big)^2} + 2(1-\delta^i_s)(1-\delta^j_s)\frac{1-\sum_l p^l}{\big(1-\sum_{l\neq s} p^l\big)^3}\right)\\
&= -2\frac{\big(\sum_{i\neq s} p^i\big)\big(1-\sum_{j\neq s} p^j\big)}{\big(1-\sum_{l\neq s} p^l\big)^2} + 2\Big(\sum_{i\neq s} p^i\Big)\Big(1-\sum_{j\neq s} p^j\Big)\frac{1-\sum_l p^l}{\big(1-\sum_{l\neq s} p^l\big)^3} = 0
\end{aligned} \tag{5.119}$$

also does.

Ultimately, if $s \neq 0$, $r \neq 0$, we have

$$p^j \frac{\partial}{\partial p^j}\frac{p^r}{p^s + p^r} = \frac{p^s p^r}{(p^s+p^r)^2}(\delta^j_r - \delta^j_s). \tag{5.120}$$

Using this property for the second summand, we obtain

$$\begin{aligned}
&\sum_{i,j} p^i(\delta^i_j - p^j)\left(\frac{\partial}{\partial p^i}\psi(\pi^{r,s}(p))\right)\left(\frac{\partial}{\partial p^j}\frac{p^r}{p^s+p^r}\right)\\
&= \sum_i \left(\frac{\partial}{\partial p^i}\psi(\pi^{r,s}(p))\right) p^i \left(\sum_j \delta^i_j\left(\frac{\partial}{\partial p^j}\frac{p^r}{p^s+p^r}\right) - \sum_j p^j\left(\frac{\partial}{\partial p^j}\frac{p^r}{p^s+p^r}\right)\right)\\
&= \sum_i \frac{\partial}{\partial p^i}\psi(\pi^{r,s}(p))\frac{p^s p^r}{(p^s+p^r)^2}(\delta^i_r - \delta^i_s) = 0.
\end{aligned} \tag{5.121}$$

The last equality is due to the fact that the sum over i in the last line vanishes in conjunction with the symmetry of π in the coordinates p^s and p^r, i. e. we have $\frac{\partial \tilde{p}^m}{\partial p^i} = \delta_i^m(1-\delta_s^m) + \delta_r^m \delta_s^i$ and consequently

$$\frac{\partial}{\partial p^s}\psi(\pi^{r,s}(p)) = \frac{\partial}{\partial \tilde{p}^r}\psi(\tilde{p}) = \frac{\partial}{\partial p^r}\psi(\pi^{r,s}(p)). \tag{5.122}$$

For the third summand, we use

$$\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\frac{p^r}{p^s+p^r} = 2\frac{\delta_j^i(\delta_s^i p^r - \delta_r^i p^s)}{(p^s+p^r)^3} + \frac{\delta_s^i \delta_r^j (1-\delta_j^i)(p^r-p^s)}{(p^s+p^r)^3} \tag{5.123}$$

and thereon get

$$\sum_{i,j} p^i(\delta_j^i - p^j)\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\frac{p^r}{p^s+p^r} = \frac{2p^s(1-p^s)p^r - 2p^r(1-p^r)p^s - 2p^s p^r(p^r-p^s)}{(p^s+p^r)^3} = 0. \tag{5.124}$$

Altogether, we have

$$L_d^*\bar{\psi} = L_d^*\left(\psi(\pi^{r,s}(p))\cdot\frac{p^r}{p^s+p^r}\right) = -\,\kappa\psi(\pi^{r,s}(p))\frac{p^r}{p^s+p^r} = -\kappa\bar{\psi} \tag{5.125}$$

for arbitrary $r, s \in I_d$, thus proving equation (5.102). □

5.4.4 A probabilistic interpretation of the extension scheme

In order to understand the nature of the presented extensions as well as their probabilistic implications, we will at first illustrate the significance of the extension constraints 5.20: If in the setting of a Wright–Fisher model an (evolving) probability density for ending up in a target set is given on the space of $d-1$ alleles, we wish to determine how the attraction of the mentioned target set also extends to the space of d alleles as of course it is also reachable by a d-allelic process by correspondingly losing one allele. As already stated, a natural assumption for such an extension is that the probability density at the transition from the d-allelic domain to the $(d-1)$-allelic domain stays regular, i. e. small alterations of the allelic configuration should only affect the probability in a controlled way, which is formulated in condition (i) in the extension constraints. Moreover, for the transition to domains of a different set of $d-1$ alleles, the corresponding probability should also stay regular with the additional requirement that in the limit it vanishes on those other $(d-1)$-allelic domains; this is also part of condition (i). As a possible extension is so far only confined towards the boundary of the domain, we also wish to link the evolution of the original probability density and its extension by requiring that both are subject to the same type of evolution in the corresponding domain, i. e. are governed by the corresponding Kolmogorov backward equation in the relevant formulation, which is condition (ii).

The extension proposition 5.23 then states that any (proper) solution of the Kolmogorov backward equation, which describes the evolving attraction of some target set given via the final condition f, may be extended to a corresponding solution of the Kolmogorov backward equation in the domain of subsequent higher dimension with both conditions above applying. However, as may be observed by remark 5.24, this actually yields the solution to a somewhat altered problem, namely the attraction generated by the target set itself plus an induced (generalised) target set in the bigger domain which are given by f and its corresponding extension $\bar{f}$. If one wishes to return to the original problem, thus the attraction of the original target set only located in the $(d-1)$-allelic domain, the induced target set needs to be compensated for by a proper solution in (the interior of) the d-allelic domain for a corresponding final condition. This just signifies that the influence of the target set on the $(d-1)$-allelic domain is handed over to the d-allelic process and evolves accordingly (which will be fully applied in section 5.4.6). In particular, the evolution of the attraction of each component may be adverse as is e. g. observed for $d=1$ (there, the attraction of $\partial_0\Delta_1$ increases over time, whereas that of Δ_1 diminishes, cf. section 4.2.2).

However, as may also be seen in proposition 5.23, the given extension scheme involves a potential ambiguity regarding the choice of the extension target face index r. Aiming for a clarification of its role, we will look at a suitable example, namely the simplest case, where we start the extension from some vertex. As a solution in a 1-allelic domain is always constant (due to $L_0=0$), this likewise applies to its extension, and hence we obtain a stationary solution (this type of solutions will be discussed in more detail in section 5.4.7). Yet, this still reveals the relevant characteristics of the extension scheme.

A simple example

We will demonstrate an application of the extension proposition 5.23 for Δ_2 and a vertex in $\partial_0\Delta_2$ as an (eventual) target set: In a first step, a solution in e. g. the vertex $\{1\}$ can be shown to extend to the 1-dimensional simplex $\Delta_1^{(\{0,1\})} = (0,1)$. We have $u \equiv 1$ on $\{1\} = \Delta_0^{(\{1\})}$ and consequently $\bar{u}^{1,0}(p^1) = 1 \cdot \frac{p^1}{p^0+p^1} \equiv p^1$ in $\Delta_1^{(\{0,1\})}$ by linear extension towards $\{0\} = \Delta_0^{(\{0\})}$ (note $p^0 = 1 - p^1$), thus reproducing the result from section 4.2.3. Clearly, at this level ($d = 1$) there are no ambiguities, and the considered extension is in accordance with the extension constraints 5.20. As a next step, however, the function $\bar{u}$ itself may be extended to $\Delta_2^{(\{0,1,2\})}$ (from boundary part $\Delta_1^{(\{0,1\})}$ and now $p^0 := 1 - p^1 - p^2$), in principle offering two different choices for the extension direction: An extension towards $\Delta_1^{(\{1,2\})}$ leads to

$$\left(\bar{u}^{1,0}\right)^{0,2}(p^1,p^2) = \bar{u}(\pi^{0,2}(p^1,p^2))\frac{p^0}{p^0+p^2} = p^1\frac{1-p^1-p^2}{1-p^1} \quad \text{in } \Delta_2^{(\{0,1,2\})}, \tag{5.126}$$

whereas an extension towards $\Delta_1^{(\{0,2\})}$ would yield

$$\left(\bar{u}^{1,0}\right)^{1,2}(p^1,p^2) = \bar{u}(\pi^{1,2}(p^1,p^2))\frac{p^1}{p^2+p^1} = (p^1+p^2)\frac{p^1}{p^1+p^2} \equiv p^1 \quad \text{in } \Delta_2^{(\{0,1,2\})}. \tag{5.127}$$

Here, this example already illustrates the mentioned ambiguity (which in this case is ruled out by the extension constraints, however), namely that the extensions generally are not unique as there are several possible extension target faces for $d \geq 2$. Their distinction is primarily seen by their behaviour towards the boundary: In the given example, the latter function $\left(\bar{u}^{1,0}\right)^{1,2}$ extends continuously to the full closure $\overline{\Delta}_2$, while the other solution $\left(\bar{u}^{1,0}\right)^{0,2}$ – although being of class C^∞ in the interior – does not extend continuously to the boundary in $\{1\}$ (cf. proposition 5.23). On the other hand, $\left(\bar{u}^{1,0}\right)^{0,2}$ smoothly vanishes on all other faces in $\partial_1\Delta_2$ except for the (last) extension source face, thus particularly complying with the extension constraints 5.20, which does not hold true for $\left(\bar{u}^{1,0}\right)^{1,2}$.

This difference is particularly relevant when interpreting the solutions in terms of probability. As stated, we want to determine the eventual 'attraction' by $u \equiv 1$ on $\{1\}$, i. e. $\{1\}$ as eventual target set, which in the 1-dimensional setting is done by the function $\bar{u}^{1,0}$, hence describing how probably the allele 0 is lost over 1 eventually. Adding one further dimension, the two extensions correspond to either asking for the probability to lose allele 2 over 0 $\left[(\bar{u}^{1,0})^{0,2}\right]$ resp. to lose allele 2 over 1 $\left[(\bar{u}^{1,0})^{1,2}\right]$: The first extension $(\bar{u}^{1,0})^{0,2}$, in compliance with the extension constraints, vanishes on $\Delta_1^{(\{1,2\})}$, whereas positive boundary values only occur on $\Delta_1^{(\{0,1\})}$. This corresponds to the extension of $\{1\}$ as ultimate target set in competition to allele 0, and consequently, $(\bar{u}^{1,0})^{0,2}$ gives the probability to end up in vertex $\{1\}$ ultimately under the additional assumption to pass through $\Delta_1^{(\{0,1\})}$ (and not through $\Delta_1^{(\{1,2\})}$) before as (transitional) first level target set generated by $\{1\}$, thus how probable a loss of alleles by

$$\{0,1,2\} \longrightarrow \{0,1\} \longrightarrow \{1\} \tag{5.128}$$

is. (This also gives an explanation why $(\bar{u}^{1,0})^{0,2}$ is not smoothly extendable into $\{1\}$, i. e. the boundary values on $\Delta_1^{(\{0,1\})}$ and $\Delta_1^{(\{1,2\})}$ are not compatible as on the former they correspond to $\{1\}$ as target set, whereas on the latter they correspond to the empty target set.) Correspondingly, we may also contract the notation of $(\bar{u}^{1,0})^{0,2}$ into $\bar{u}^{1,0,2}$.

On the other hand, $(\bar{u}^{1,0})^{1,2}(p) = p_1$ only vanishes on $\Delta_1^{(\{0,2\})}$, thus has positive values on both other faces, which corresponds to the influence of $\{1\}$ as target set in competition to *both* the alleles 0 and 2. Consequently, it gives the probability to end up in $\{1\}$ ultimately with no certain first level target specified.

In the remainder, we will only consider those pathwise extensions (like $\bar{u}^{1,0,2}$) corresponding to the predication of the extension constraints 5.20, i. e. extensions which correspond to a successive loss of alleles. For these, the *probabilistic interpretation of proposition 5.23*, in addition to what has already been described, is that, in the context of a Wright–Fisher model, an extension of a probability distribution from a $(d-1)$-allelic domain to a d-allelic domain is always such that the potential loss of the extra allele is modelled as if it was in competition with just 1 other allele r dependent on the index chosen *(fibration property)*. Thus, we say that allele s is lost over allele r.

However, as already stated, in general this target face index r is not uniquely determined, and – depending on the situation at hand – its choice may decide whether an extension is in accordance with the extension boundary constraint 5.20 (i). For a simple extension from a 0-dimensional domain (as observed in the preceding example) or if the starting distribution smoothly vanishes towards all boundaries of subsequent lower dimension, an extension is always in accordance with the extension constraints, but particularly in the case of an iteratively continued extension, the choice of the target face index $r(d)$ is crucial as will be shown in the following section. Conversely, demanding the extension constraints also uniquely determines the index $r(d)$ from the second step for non-empty target sets.

5.4.5 Iterated extensions

A repeated application of proposition 5.23 yields the existence of iterated extensions (generalising the corresponding result for $n = 2$ in [21] and the (less explicit) result in [23] without derivation):

5.25 Proposition (pathwise extension of solutions)**.** *Let $k, n \in \mathbb{N}$ with $0 \leq k < n$, $\{i_k, i_{k+1}, \dots, i_n\} \subset I_n := \{0, 1, \dots, n\}$ with $i_i \neq i_j$ for $i \neq j$ and $I_k := I_n \setminus \{i_{k+1}, \dots, i_n\}$, and let u_{I_k} be a proper solution of the Kolmogorov backward equation* (5.95) *in $\Delta_k^{(I_k)}$ for some final condition $f \in \mathcal{L}^2\big(\Delta_k^{(I_k)}\big)$ as in proposition 5.19. For $d = k+1, \dots, n$ and $I_d := I_k \cup \{i_{k+1}, \dots i_d\}$, an extension of $\bar{u}_{I_k}^{i_k,\dots,i_{d-1}}$ in $\big(\Delta_{d-1}^{(I_{d-1})}\big)_{-\infty}$ to $\bar{u}_{I_k}^{i_k,\dots,i_d}$ in $\big(\Delta_d^{(I_d)}\big)_{-\infty}$ as by proposition 5.23 is in accordance with the extension constraints 5.20 if (and for $d \geq k+2$ and $[f] \neq 0$ in $L^2\big(\Delta_k^{(I_k)}\big)$ also only if) putting $r(d) = i_{d-1}$ for the extension target face index, and we respectively have*

$$\bar{u}_{I_k}^{i_k,\dots,i_d}(p,t) = u_{I_k}(\pi^{i_k,\dots,i_d}(p), t) \prod_{j=k}^{d-1} \frac{p^{i_j}}{\sum_{l=j}^{d} p^{i_l}}, \quad (p,t) \in \big(\Delta_d^{(I_d)}\big)_{-\infty} \tag{5.129}$$

with $p^0 = 1 - \sum_{i \in I_d \setminus \{0\}} p^i$ and $\pi^{i_k,\dots,i_d}(p) = (\tilde{p}^1, \dots, \tilde{p}^n)$ such that $\tilde{p}^{i_k} = p^{i_k} + \ldots + p^{i_d}$, $\tilde{p}^{i_{k+1}} = \ldots = \tilde{p}^{i_d} = 0$ and $\tilde{p}^j = p^j$ for $j \in I_d \setminus \{i_k, \dots, i_d\}$.

Correspondingly, the resulting assembling of all extensions to a function $\bar{U}_{I_k}^{i_k,\dots,i_n}$ in $\Big(\bigcup_{k \leq d \leq n} \Delta_d^{(I_d)}\Big)_{-\infty}$ by putting

$$\bar{U}_{I_k}^{i_k,\dots,i_n}(p,t) := u_{I_k}(p,t)\chi_{\Delta_k^{(I_k)}}(p) + \sum_{k+1 \leq d \leq n} \bar{u}_{I_k}^{i_k,\dots,i_d}(p,t)\chi_{\Delta_d^{(I_d)}}(p)$$

$$= u_{I_k}(p,t)\chi_{\Delta_k^{(I_k)}}(p) + \sum_{k+1\le d\le n} u_{I_k}(\pi^{i_k,\dots,i_d}(p),t)\prod_{j=k}^{d-1}\frac{p^{i_j}}{\sum_{l=j}^{d} p^{i_l}}\chi_{\Delta_d^{(I_d)}}(p) \quad (5.130)$$

with $p^0 = 1-\sum_{i\in I_n\setminus\{0\}} p^i$ *is in* $C^\infty_{p_0}\Big(\bigcup_{k\le d\le n}\Delta_d^{(I_d)}\Big)$ *with respect to the spatial variables for* $t<0$ *as well as in* $C^\infty((-\infty,0))$ *with respect to* t*, and we have*

$$\begin{cases} L^*\bar U_{I_k}^{i_k,\dots,i_n} = -\frac{\partial}{\partial t}\bar U_{I_k}^{i_k,\dots,i_n} & \text{in } \Big(\bigcup_{k\le d\le n}\Delta_d^{(I_d)}\Big)_{-\infty} \\ \bar U_{I_k}^{i_k,\dots,i_n}(\,\cdot\,,0) = \bar F_{I_k}^{i_k,\dots,i_n} & \text{in } \bigcup_{k\le d\le n}\Delta_d^{(I_d)} \end{cases} \quad (5.131)$$

with $\bar F_{I_k}^{i_k,\dots,i_n} \in \mathcal{L}^2\Big(\bigcup_{k\le d\le n}\Delta_d^{(I_d)}\Big)$ *being an analogous extension of the final condition* $f=f_{I_k}$ *in* $\Delta_k^{(I_k)}$ *as by remark 5.24; in particular, we have* $\bar U_{I_k}^{i_k,\dots,i_n}\big|_{\Delta_k^{(I_k)}}(\,\cdot\,,0) = f$ *in* $\Delta_k^{(I_k)}$*.*

5.26 Corollary. *For* $n\in\mathbb{N}_+$*,* $k=0$ *and* $u_{\{i_0\}}\equiv 1$ *in* $\Delta_0^{(\{i_0\})}\subset\partial_0\Delta_n$*, equation* (5.129) *resp. equation* (5.130) *restricted to* Δ_n *and with the* t*-coordinate suppressed coincides with Littler's formula in* Δ_n *(cf. [23]):*

$$\bar U_{\{i_0\}}^{i_0,i_1\dots,i_n}\big|_{\Delta_n}(p) \equiv \bar u_{\{i_0\}}^{i_0,i_1\dots,i_n}(p) = p^{i_0}\cdot\frac{p^{i_1}}{1-p^{i_0}}\cdot\ldots\cdot\frac{p^{i_{n-1}}}{1-\sum_{l=0}^{n-2}p^{i_l}}. \quad (5.132)$$

Proof of proposition 5.25. The result is basically an application of proposition 5.23, which yields the regularity and the solution property (cf. equation (5.131)) in every $\Delta_d^{(I_d)}$. It only remains to be shown that the boundary behaviour in each extension step is in accordance with the extension constraints 5.20 as well as the formula (5.129), which is both done inductively.

Clearly, a proper solution u_{I_k} of the Kolmogorov backward equation in $\big(\Delta_k^{(I_k)}\big)_{-\infty}$ as in proposition 5.19 satisfies equation (5.129) and is of class $C_0^\infty\big(\Delta_k^{(I_k)}\big)$ w. r. t. the spatial variables for $t<0$ (which in particular signifies that it is smoothly extendable to $\partial_{k-1}\Delta_k^{(I_k)}$). Extending u_{I_k} to $\big(\Delta_{k+1}^{(I_{k+1})}\big)_{-\infty}$ via proposition 5.23 with $s(k+1)=i_{k+1}$ and $r(k+1)=i_k$ yields a function $\bar u_{I_k}^{i_k,i_{k+1}}$ of type (5.129), which for $t<0$ smoothly extends to all boundary faces $\partial_k\Delta_{k+1}^{(I_{k+1})}$ and vanishes there except for $\Delta_k^{(I_k)}$ (where it coincides with u_{I_k}) by the assumed boundary behaviour of u_{I_k}. We may thus assume that for $k<d-1<n$ an assembled extension $\bar U_{I_k}^{i_k,\dots,i_{d-1}}$ (corresponding to

equation (5.130)) in $C^\infty_{p_0}\big(\bigcup_{k\le m\le d-1}\Delta_m^{(I_m)}\big)$ with respect to the spatial coordinates exists whose top-dimensional component $\bar{U}_{I_k}^{i_k,\dots,i_{d-1}}\big|_{\left(\Delta_{d-1}^{(I_{d-1})}\right)_{-\infty}} =: \bar{u}_{I_k}^{i_k,\dots,i_{d-1}}$ satisfies equation (5.129).

We may then perform an extension of $\bar{u}_{I_k}^{i_k,\dots,i_{d-1}}$ in $\big(\Delta_{d-1}^{(I_{d-1})}\big)_{-\infty}$ to $\bar{u}_{I_k}^{i_k,\dots,i_d}$ in $\big(\Delta_d^{(I_d)}\big)_{-\infty}$ via proposition 5.23 with $s(d)=i_d$ and $r(d)=i_{d-1}$. By the assumed boundary behaviour of $\bar{u}_{I_k}^{i_k,\dots,i_{d-1}}$ (i. e. $\bar{U}_{I_k}^{i_k,\dots,i_{d-1}}$ being of class $C^\infty_{p_0}$), $\bar{u}_{I_k}^{i_k,\dots,i_d}$ smoothly extends to all boundary faces $\partial_{d-1}\Delta_d^{(I_d)}\setminus\Delta_{d-1}^{(I_d\setminus\{i_{d-1}\})}$ and vanishes there except for $\Delta_{d-1}^{(I_{d-1})}$ (where it coincides with $\bar{u}_{I_k}^{i_k,\dots,i_{d-1}}$) for $t<0$. By putting $r(d)=i_{d-1}$, this particularly also holds for $\Delta_{d-1}^{(I_d\setminus\{i_{d-1}\})}$, which in turn would otherwise be violated if $f\neq 0$ almost everywhere as may be seen from the proof of proposition 5.23. Then, the boundary behaviour is in accordance with the extensions constraints 5.20, and we correspondingly have $\bar{U}_{I_k}^{i_k,\dots,i_d} := \bar{U}_{I_k}^{i_k,\dots,i_{d-1}} + \bar{u}_{I_k}^{i_k,\dots,i_d}\chi_{\Delta_d^{(I_d)}} \in C^\infty_{p_0}\big(\bigcup_{k\le m\le d}\Delta_m^{(I_m)}\big)$ w. r. t. the spatial variables for $t<0$.

To show equation (5.129), we obtain for $\bar{u}_{I_k}^{i_k,\dots,i_d}$ by equation (5.101) when plugging in the formula (5.129) for $\bar{u}_{I_k}^{i_k,\dots,i_{d-1}}$

$$\begin{aligned}
\bar{u}_{I_k}^{i_k,\dots,i_d}(p,t) &= \bar{u}_{I_k}^{i_k,\dots,i_{d-1}}(\pi^{i_{d-1},i_d}(p),t)\frac{p^{i_{d-1}}}{p^{i_{d-1}}+p^{i_d}}\\
&= u_{I_k}(\pi^{i_k,\dots,i_{d-1}}(\pi^{i_{d-1},i_d}(p)),t)\prod_{j=k}^{d-2}\frac{(\pi^{i_{d-1},i_d}(p))^{i_j}}{\sum_{l=j}^{d-1}(\pi^{i_{d-1},i_d}(p))^{i_l}}\frac{p^{i_{d-1}}}{p^{i_{d-1}}+p^{i_d}}\\
&= u_{I_k}(\pi^{i_k,\dots,i_d}(p),t)\prod_{j=k}^{d-1}\frac{p^{i_j}}{\sum_{l=j}^{d}p^{i_l}} \qquad \text{in } \big(\Delta_d^{(I_d)}\big)_{-\infty}
\end{aligned} \tag{5.133}$$

as $(\pi^{i_{d-1},i_d}(p))^{i_j}=p^{i_j}$ for $i_j=i_k,\dots,i_{d-2}$ and $(\pi^{i_{d-1},i_d}(p))^{i_{d-1}}=p^{i_{d-1}}+p^{i_d}$. If some index i_j equals zero (w. l. o. g. $i_0=0$) corresponding to $(\pi^{i_{d-1},i_d}(p))^0$, this expression gets replaced by $p^0\in\Delta_d^{(I_d)}$ as we have $(\pi^{i_{d-1},i_d}(p))^0 = 1-\sum_{j=1}^{d-1}(\pi^{i_{d-1},i_d}(p))^{i_j} = 1-\sum_{j=1}^{d}p^{i_j}\equiv p^0$. Furthermore, $\pi^{i_k,\dots,i_{d-1}}(\pi^{i_{d-1},i_d}(p)) = \pi^{i_k,\dots,i_d}(p)$ directly follows from the definitions, thus proving equation (5.129) for $\bar{u}_{I_k}^{i_k,\dots,i_d}$ in $\Big(\bigcup_{k\le m\le d}\Delta_m^{(I_m)}\Big)_{-\infty}$. □

5.27 Remark. Geometrically, the choice of the extension target face indices $s(d)=i_d$ and $r(d)=i_{d-1}$ signifies that the extension source face $\Delta_{d-1}^{(\{i_0,\dots,i_{d-2},i_{d-1}\})}$ and the target face $\Delta_{d-1}^{(\{i_0,\dots,i_{d-2},i_d\})}$ are adjacent faces to the highest degree, as they share $d-1$ vertices (for $d\ge 2$). Furthermore, their intersection $\Delta_{d-2}^{(\{i_0,\dots,i_{d-2}\})}$ is the extension source face of the previous step.

Sticking to the preceding probabilistic interpretation, $\bar{u}_{I_k}^{i_k,i_{k+1},\dots,i_n}$ depicts the iterated 'attraction' of an (analogously extended) target set in $\Delta_k^{(I_k)}$ along a corresponding *extension path* specified by $i_k,\dots,i_n$ resp. the corresponding index sets $I_k \subset \ldots \subset I_n$. Thus, $\bar{U}_{I_k}^{i_k,i_{k+1},\dots,i_n}$ gives the total probability for all paths in $\overline{\Delta}_n$ starting in $\Delta_n^{(I_n)}$, passing through the (sub)simplices

$$\Delta_{n-1}^{(I_{n-1})} \longrightarrow \Delta_{n-2}^{(I_{n-1})} \longrightarrow \ldots \longrightarrow \Delta_{k+1}^{(I_{k+1})} \longrightarrow \Delta_k^{(I_k)} \tag{5.134}$$

and reaching the eventual target set, which, in the setting of the Wright–Fisher model, corresponds to eventually losing $n-k$ of originally n alleles in such a manner that from dimension $n-1$ down to 0 exactly the allele sets

$$I_n \longrightarrow I_{n-1} \longrightarrow \ldots \longrightarrow I_{k+1} \longrightarrow I_k \tag{5.135}$$

are present until reaching the eventual target set. The loss of the corresponding allele at each stage is over the allele which is lost next, thus allele i_d is lost over i_{d-1}. Merely in the last step, i. e. the loss of allele i_{k+1}, the index i_k determines which of the alleles in I_k is the one i_{k+1} is lost over.

However, the corresponding extensions in proposition 5.25 are not satisfactory to the extent that they lack a global (pathwise) regularity property on the entire $\overline{\Delta}_n$, i. e. are not in C_p^∞ w. r. t. the spatial variables, as this applies only along the corresponding extension path. Outside this path, generally no continuous (or even smooth) extensions exist. This is caused by the incompatibilities involved by this construction (cf. also section 5.4.4): For example on $\Delta_{k+1}^{(\tilde{I}_{k+1})}$ with $\tilde{I}_{k+1} := I_k \cup \{\tilde{\imath}_k\}$ and $\tilde{\imath}_k \in I_n \setminus I_{k+1}$, a positive hit probability for the target set in $\Delta_k^{(I_k)}$ by a direct loss of allele $\tilde{\imath}_k$ would exist, yet the considered solution necessarily vanishes on $\Delta_{k+1}^{(\tilde{I}_{k+1})}$ as this is a boundary face of $\Delta_{k+2}^{(I_{k+2})}$ outside the specified path.

This defect may be overcome by mounting these extensions into a global solution covering all possible extensions paths, each one of them corresponding to a certain ordering of the indices in $I_n \setminus I_k$. While this mounting construction is rather straightforward, there is an ambiguity at another stage: For a given extension path and a non-empty target set, in the first extension step, the extension target face is not defined by the extension boundary condition (i) in definition 5.20 (except for $k = 0$; cf. proposition 5.25), consequently all indices in I_k may serve as target face index, and hence all choices of this also need to be taken into account. Thus, additionally summing over all possible first stage extensions and normalising, we obtain the following result:

5.28 Proposition (global extension of solutions)**.** *Let $k, n \in \mathbb{N}$ with $0 \leq k < n$, $I_k \subset I_n := \{0, 1, \ldots, n\}$ with $|I_k| = k + 1$, and let u_{I_k} be a proper solution of the Kolmogorov backward equation (5.95) in $\Delta_k^{(I_k)}$ for some final condition $f \in \mathcal{L}^2\big(\Delta_k^{(I_k)}\big)$ as in proposition 5.19. Then an assembling of all pathwise extensions of u_{I_k} as by proposition 5.25 into a function $\bar{U}_{I_k} \in \big(\overline{\Delta}_n\big)_{-\infty}$ by putting*[2]

$$\begin{aligned}\bar{U}_{I_k}(p,t) &:= u_{I_k}(p,t)\chi_{\Delta_k^{(I_k)}}(p) \\ &+ \frac{1}{|I_k|} \sum_{i_k \in I_k} \sum_{k+1 \leq d \leq n} \sum_{i_{k+1} \in I_n \setminus I_k} \cdots \sum_{\substack{i_d \in I_n \setminus (I_k \cup \\ \{i_{k+1}, \ldots, i_{d-1}\})}} \bar{u}_{I_k}^{i_k, \ldots, i_d}(p,t)\chi_{\Delta_d^{(I_k \cup \{i_{k+1}, \ldots, i_d\})}}(p) \end{aligned} \tag{5.136}$$

for $(p,t) \in \big(\bigcup_{I_k \subset I_d \subset I_n} \Delta_d^{(I_d)}\big)_{-\infty}$ and $\bar{U}_{I_k}(p,t) := 0$ in the remainder of $\big(\overline{\Delta}_n\big)_{-\infty}$ is in $C_p^\infty\big(\overline{\Delta}_n\big)$ with respect to the spatial variables for $t < 0$ as well as in $C^\infty((-\infty, 0))$ with respect to t. Furthermore, $\bar{U}_{I_k}$ is a solution of the corresponding Kolmogorov backward equation in $\big(\overline{\Delta}_n\big)_{-\infty}$ and for $t = 0$ matches an analogously assembled extension $\bar{F}_{I_k}$ of $f = f_{I_k}$ in $\Delta_k^{(I_k)}$ as final condition in $\overline{\Delta}_n$ (cf. remark 5.24).

Proof. The asserted global regularity directly follows from properties of the applied extension scheme as stated in lemma 5.22 and proposition 5.25 and the construction of $\bar{U}_{I_k}$, which is such that potential discontinuities are ruled out by assembling all extensions along arbitrary paths. The solution property and the compliance with the analogously constructed final condition likewise straightforwardly extend from proposition 5.25. □

Shifting again to the probabilistic interpretation, $\bar{U}_{I_k}$ now depicts the full iterated 'attraction' of some eventual target set in $\Delta_k^{(I_k)}$ and its (successively) induced target sets in $\Delta_d^{(I_d)} \subset \overline{\Delta}_n$ with $I_d \supset I_k$, which may now be reached along arbitrary paths.

[2]The last sum actually only comprises a single summand; this notation is used to illustrate the choice of the index i_d, however.

Thus, $\bar{U}_{I_k}$ gives the total probability for all paths from $\Delta_n^{(I_n)}$ to eventually $\Delta_k^{(I_k)}$ – with no assumptions on possible interstages made. In the setting of the Wright–Fisher model, this corresponds to eventually losing $n-k$ of previously n alleles irrespective of any order of loss.

Since $\bar{U}_{I_k}$ represents the most general extension of a given solution u_{I_k} in $\Delta_k^{(I_k)}$ to $\overline{\Delta}_n$, we are now in a position to develop a general solution scheme for solutions of the extended Kolmogorov backward equation (5.95) in the next section.

5.4.6 Construction of general solutions via the extension scheme

For a given final condition $f = \sum_{d=0}^{n} f_d \chi_{\partial_d \Delta_n} \in \mathcal{L}^2\big(\bigcup_{d=0}^{n} \partial_d \Delta_n\big)$, the following extension scheme allows us to construct a solution of the extended Kolmogorov backward equation (5.95) which captures the full dynamics of the process on the entire $\big(\overline{\Delta}_n\big)_{-\infty}$ analogously to such in the 1-dimensional case. The main ingredient for this are the global extensions of a (proper) solution of the Kolmogorov backward equation in every instance of the domain as in proposition 5.28; these globally extended solutions just need to be superposed in a way that eventually the given final condition is met in the entire $\overline{\Delta}_n$ (cf. also section 5.4.4 for a probabilistic interpretation).

Thus, first equation (5.95) is solved in each $\big(\Delta_0^{(\{i_0\})}\big)_{-\infty} \subset (\partial_0 \Delta_n)_{-\infty}$ for the final condition f_0, and afterwards, these solutions are successively extended to $\big(\overline{\Delta}_n\big)_{-\infty}$ by means of proposition 5.28, which analogously generates a successively extended final condition in $\overline{\Delta}_n$ for $t=0$. Subsequently, a (proper) solution in each $\big(\Delta_1^{(I_1)}\big)_{-\infty} \subset (\partial_1 \Delta_n)_{-\infty}$ for the final condition f_1 minus the extension of f_0 is determined, which is then successively extended to $\big(\overline{\Delta}_n\big)_{-\infty}$ (again likewise generating an analogously extended final condition). This procedure is repeated until after finding a (proper) solution in $(\Delta_n)_{-\infty}$ an extended solution in the entire $\big(\overline{\Delta}_n\big)_{-\infty}$ is determined.

A solution of the extended Kolmogorov backward equation (4.63) restricted to some $\big(\Delta_0^{(\{i_0\})}\big)_{-\infty} \subset (\partial_0 \Delta_n)_{-\infty}$ is – of course – trivial, i. e. $u_{\{i_0\}}(p,t) = f_0(p)$ for $(p,t) \in \big(\Delta_0^{(\{i_0\})}\big)_{-\infty}$, and by proposition 5.28 we obtain $\bar{U}_{\{i_0\}}$ as an extension to

$\left(\overline{\Delta}_n\right)_{-\infty}$. Summing over all $\Delta_0^{(\{i_0\})}$ yields

$$\bar{U}_0 := \sum_{\{i_0\}\subset I_n} \bar{U}_{\{i_0\}} \qquad \text{in } \left(\overline{\Delta}_n\right)_{-\infty} \tag{5.137}$$

with $\bar{U}_0$ in $C_p^\infty\left(\overline{\Delta}_n\right)$ with respect to the spatial variables as well as in $C^\infty((-\infty,0))$ with respect to t and

$$\begin{cases} L^*\bar{U}_0 = -\frac{\partial}{\partial t}\bar{U}_0 & \text{in } \left(\overline{\Delta}_n\right)_{-\infty} \\ \bar{U}_0(\,\cdot\,,0) = \bar{F}_0' & \text{in } \overline{\Delta}_n \end{cases} \tag{5.138}$$

with $\bar{F}_0'$ being a corresponding superposed global extension of all $f_0' \equiv f_0$ in $\partial_0\Delta_n$ as described above for the $u_{\{i_0\}}$ (cf. also remark 5.24), in particular we have $\bar{U}_0|_{\partial_0\Delta_n}(\,\cdot\,,0) = f_0$.

For the next step, proper solutions in $(\partial_1\Delta_n)_{-\infty}$ are determined and likewise extended to $\left(\overline{\Delta}_n\right)_{-\infty}$. However, as this extension procedure will be repeated for all d-dimensional instances of $(\Delta_n)_{-\infty}$ for $d = 1,\dots,n$, we directly assume that suitable solutions in $\left(\bigcup_{m=0}^{d-1}\partial_m\Delta_n\right)_{-\infty}$ already have successively been determined and extended to $\left(\overline{\Delta}_n\right)_{-\infty}$, thus $\sum_{m=0}^{d-1}\bar{U}_m$ solves the extended Kolmogorov backward equation (5.95) in $\left(\overline{\Delta}_n\right)_{-\infty}$ and matches the final condition f for $t=0$ in $\bigcup_{m=0}^{d-1}\partial_m\Delta_n$ (still, with $\bar{U}_0(\,\cdot\,,0),\dots,\bar{U}_{d-1}(\,\cdot\,,0)$ in $\overline{\Delta}_n$ respectively matching a corresponding superposed global extension $\bar{F}_m'$ of the final condition f_m' in $\partial_m\Delta_n$ modified as below). Then, a proper solution u_{I_d} by proposition 5.19 in each $\left(\Delta_d^{(I_d)}\right)_{-\infty} \subset (\partial_d\Delta_n)_{-\infty}$, $I_d \subset I_n$ is determined which matches the modified final condition

$$f_d' := f_d - \sum_{m=0}^{d-1} \bar{F}_m'|_{\partial_d\Delta_n} \quad \text{in } \partial_d\Delta_n, \tag{5.139}$$

correspondingly restricted to the relevant $\Delta_d^{(I_d)}$. For each I_d, the solution u_{I_d} is then extended to $\left(\overline{\Delta}_n\right)_{-\infty}$ via proposition 5.28 each leading to a function $\bar{U}_{I_d}$. Clearly, these extensions do not interfere with the solutions on lower dimensional entities by definition.

Summing over the extensions of all u_{I_d}, $I_d \subset I_n$, we obtain

$$\bar{U}_d := \sum_{I_d \subset I_n} \bar{U}_{I_d} \quad \text{in } \big(\overline{\Delta}_n\big)_{-\infty} \tag{5.140}$$

as the global extension of all (proper) solutions in $(\partial_d \Delta_n)_{-\infty}$. By proposition 5.28 and the linearity of the differential equation, $\bar{U}_d$ is in $C_p^\infty\big(\overline{\Delta}_n\big)$ w. r. t. the spatial variables as well as in $C^\infty((-\infty, 0)$ with respect to t and solves the extended Kolmogorov backward equation and for $t = 0$ matches a corresponding superposed global extension $\bar{F}'_d$ of the final condition f'_d in $\partial_d \Delta_n$, thus in particular $\bar{U}_d(\,\cdot\,, 0)|_{\partial_d \Delta_n} = f'_d$. Consequently, the sum of all up to now extended solutions also is in $C_p^\infty\big(\overline{\Delta}_n\big)$ w. r. t. the spatial variables as well as in $C^\infty((-\infty, 0)$ with respect to t and fulfils

$$\begin{cases} L^*\Big(\sum_{m=0}^d \bar{U}_m\Big) = -\frac{\partial}{\partial t}\Big(\sum_{m=0}^d \bar{U}_m\Big) & \text{in } \big(\overline{\Delta}_n\big)_{-\infty} \\ \Big(\sum_{m=0}^d \bar{U}_m\Big)\big|_{\bigcup_{m=0}^d \partial_m \Delta_n}(\,\cdot\,, 0) = f|_{\bigcup_{m=0}^d \partial_m \Delta_n} & \text{in } \bigcup_{m=0}^d \partial_m \Delta_n. \end{cases} \tag{5.141}$$

Repeating the preceding step successively one eventually arrives at $\sum_{m=0}^{n-1} \bar{U}_m$. For the remaining $(\Delta_n)_{-\infty}$, at last a (proper) solution $u_{I_n} =: \bar{U}_n$ by proposition 5.19 is determined matching the modified final condition

$$f'_n := f_n - \sum_{m=0}^{n-1} \bar{F}'_m|_{\Delta_n} \quad \text{in } \Delta_n. \tag{5.142}$$

Then the sum of all globally extended (proper) solutions in all instances of the domain

$$\bar{U} := \sum_{j=0}^{n} \bar{U}_j \tag{5.143}$$

is in $C_p^\infty\big(\overline{\Delta}_n\big)$ w. r. t. the spatial variables as well as in $C^\infty((-\infty, 0))$ with respect to t and fulfils

$$\begin{cases} L^*\bar{U} = -\frac{\partial}{\partial t}\bar{U} & \text{in } \big(\overline{\Delta}_n\big)_{-\infty} \\ \bar{U}(\,\cdot\,, 0) = f & \text{in } \overline{\Delta}_n, \end{cases} \tag{5.144}$$

thus is a solution of the extended Kolmogorov backward equation (5.95).

Altogether, we have the following existence result:

5.29 Theorem. *For a given final condition* $f \in \mathcal{L}^2\big(\bigcup_{d=0}^{n} \partial_d \Delta_n\big)$*, the extended Kolmogorov backward equation* (5.93) *corresponding to the n-dimensional Wright–Fisher model in diffusion approximation always allows a solution* $\bar{U}\colon \big(\overline{\Delta}_n\big)_{-\infty} \longrightarrow \mathbb{R}$ *with* $\bar{U}(\,\cdot\,,t) \in C_p^{\infty}\big(\overline{\Delta}_n\big)$ *for each fixed* $t \in (-\infty, 0)$ *and* $\bar{U}(x,\,\cdot\,) \in C^{\infty}((-\infty,0))$ *for each fixed* $x \in \overline{\Delta}_n$*.*

In the following chapter, we will be able to show that for $f \in \mathcal{L}^2\big(\partial_0 \Delta_n\big)$ – and under some additional regularity assumptions – the solution obtained, i. e. $\bar{U}_0$, also is the unique solution.

5.4.7 The stationary Kolmogorov backward equation

As in the 1-dimensional case, we are also interested in the long-term behaviour of the process, i. e. which alleles are eventually lost and in which order (cf. section 4.2.3). This will lead us to a stationary version of the Kolmogorov backward equation; solutions thereof have already appeared implicitly in the preceding section as extensions of solutions in $\partial_0 \Delta_n$ since the corresponding operator L_0^* only possesses the eigenvalue 0. Even with the extended setting presented in section 5.4.2 available, we will at first tentatively limit our view to some interior simplex Δ_n, thus restrict a given extended solution to Δ_n if necessary.

Then, for a solution in Δ_n, we may argue again that all eigenmodes of the solution corresponding to a positive eigenvalue vanish for $t \to -\infty$, while those corresponding to the eigenvalue zero are preserved (cf. the analogous lemma 4.7 in the 1-dimensional case). Thus, it may be shown that a solution of the Kolmogorov backward equation (5.93) in Δ_n converges uniformly to a solution of the corresponding *homogeneous* or *stationary Kolmogorov backward equation*

$$\begin{cases} L^* u(p) = 0 & \text{in } \Delta_n \\ u(p) = f(p) & \text{in } \partial \Delta_n \end{cases} \tag{5.145}$$

for $u \in C^2(\Delta_n)$ and with boundary condition f, which needs to be attained smoothly in a certain sense.

In this formulation, the equation appears as a boundary value problem (for some suitably chosen boundary function f, assuring the uniqueness of a solution). However, as may be expected from previous considerations, the role of the boundary here is different from usual boundary value problems and again requires some extra care: On the one hand, a proper solution in Δ_n always converges to the trivial stationary solution (i. e. constantly equalling 0), which is linked to the fact that their (continuous) extension to the boundary also vanishes at all negative times. On the other hand, any solution which extends to $\partial\Delta_n$ is already strongly constrained by the degeneracy behaviour of the differential operator if suitable regularity assumptions on the solution in $\overline{\Delta}_n$ (cf. also equality (5.96)) apply:

5.30 Lemma (stem lemma)**.** *For a solution $u \in C^\infty(\Delta_n)$ of equation* (5.145) *with extension $\bar{u} \in C_p^\infty\big(\overline{\Delta}_n\big)$, we have*

$$L^*\bar{u} = 0 \qquad \text{in } \overline{\Delta}_n. \tag{5.146}$$

Proof. The statement is proven iteratively: Assuming that we have $L_k^*\bar{u} = 0$ for all $\Delta_k^{(I_k)} \subset \partial_k\Delta_n$, we show that this property extends to each $\Delta_{k-1}^{(I_{k-1})} \subset \partial_{k-1}\Delta_k^{(I_k)}$ for every $\Delta_k^{(I_k)}$, hence we obtain $L_{k-1}^*\bar{u} = 0$ on $\partial_{k-1}\Delta_n$. A repeated application then yields equation (5.146).

W. l. o. g. let $\Delta_k^{(I_k)}$ and $\Delta_{k-1}^{(I_{k-1})} \subset \partial_{k-1}\Delta_k^{(I_k)}$ with $I_k \setminus I_{k-1} = \{i_k\}$. Then for the operator L_k^* in $\Delta_k^{(I_k)}$, we have

$$L_k^* = L_{k-1}^* + p^{i_k}\Bigg(\sum_{i_j \in I_k\setminus\{0\}} (\delta_{i_k}^{i_j} - p^{i_j})\frac{\partial}{\partial p^{i_j}}\frac{\partial}{\partial p^{i_k}}\Bigg) \tag{5.147}$$

with L_{k-1}^* being the restriction of L_k^* to $\Delta_{k-1}^{(I_{k-1})}$.

Now, choosing some $p \in \Delta_{k-1}^{(I_{k-1})}$ and a sequence $(p_l)_{l\in\mathbb{N}}$ in $\Delta_k^{(I_k)}$ with $p_l \to p$ and applying the above formula to $\bar{u}$ at $p_l \in \Delta_k^{(I_k)}$, the big bracket is controlled by $p_l^{i_k} \to 0$ while approaching p and – with the derivatives of $\bar{u}$ inside being bounded on a closed neighbourhood of p by reason of the regularity of $\bar{u}$ – is continuous up to p. Likewise, all derivatives of $\bar{u}$ within $\Delta_{k-1}^{(I_{k-1})}$ are continuously matched by the corresponding ones in $\Delta_k^{(I_k)}$, thus $L_{k-1}^*(\bar{u}(p_l))$ is also continuous up to the boundary in p (as the corresponding coefficients are, too). Hence, the whole expression is continuous up to the boundary in p with $L^*_{k-1}\bar{u}(p) \equiv L^*_k\bar{u}(p) = 0$, and since p was arbitrary, this applies to all of $\Delta_{k-1}^{(I_{k-1})}$. □

This signifies that – assuming the stated pathwise regularity – the values of $\bar{u}$ resp. f on $\partial\Delta_n = \bigcup_{k=0}^{n-1} \partial_k\Delta_n$ for equation (5.145) may not be chosen freely but have to be solutions of the corresponding version of the stationary Kolmogorov backward equation (5.145) in each $\Delta_k^{(I_k)} \subset \partial_k\Delta_n$ for all $k = 0, \ldots, n-1$. Hence, the boundary value problem in equation (5.93) is rather restated as an *extended homogeneous* or *extended stationary Kolmogorov backward equation*[3]

$$\begin{cases} L^*\bar{u}(p) = 0 & \text{in } \overline{\Delta}_n \setminus \partial_0\Delta_n \\ \bar{u}(p) = f(p) & \text{in } \partial_0\Delta_n \end{cases} \tag{5.148}$$

for $\bar{u} \in C^2_p\big(\overline{\Delta}_n\big)$ with the only 'free' boundary values remaining the ones on the vertices $\partial_0\Delta_n$. The values on $\partial_0\Delta_n$, however, suffice as boundary information determining a solution uniquely if we also assume global continuity of the solution. Consequently, a stationary solution and the stationary component of a global extension as in the preceding section are also identical:

5.31 Proposition. *A solution $\bar{u} \in C^\infty_p\big(\overline{\Delta}_n\big) \cap C^0\big(\overline{\Delta}_n\big)$ of the extended stationary Kolmogorov backward equation* (5.148) *for some boundary condition $f_0\colon \partial_0\Delta_n \longrightarrow \mathbb{R}$ is uniquely defined and coincides with (the projection of) a solution of the extended Kolmogorov backward equation* (5.95) *in $\big(\overline{\Delta}_n\big)_{-\infty}$ to $\overline{\Delta}_n$ for a final condition $f \in \mathcal{L}^2\big(\bigcup_{d=0}^n \partial_d\Delta_n\big)$ with $f \equiv f_0\chi_{\partial_0\Delta_n}$ as by theorem 5.29. Furthermore, the space of solutions is spanned by $p^1, \ldots, p^n$ and 1.*

Proof. The first assertion may be shown by a successive application of the maximum principle: In every instance of the domain $\Delta_k^{(I_k)} \subset \partial_k\Delta_n$ for all $1 \le k \le n$, the operator L^* is locally uniformly elliptic, and hence, $\bar{u}|_{\Delta_k^{(I_k)}}$ is uniquely defined by its values on $\partial\Delta_k^{(I_k)}$ by virtue of the maximum principle. Applying this consideration successively for $\partial_0\Delta_n, \ldots, \partial_n\Delta_n = \Delta_n$ yields the desired global uniqueness.

[3] As already stated, it is without effect whether $\partial_0\Delta_n$ is added to the domain of definition of the differential equation or not. Although $\partial_0\Delta_n$ has been included in equation (5.146), this is not done here for formal reasons.

Next, we will show that a final condition $f = \chi_{\Delta_0^{(\{i_0\})}}$ for some $i_0 \in I_n$ gives rise to an extended solution $\bar{U}(p,t) = \bar{U}(p) = p^{i_0}$ in $\left(\overline{\Delta}_n\right)_{-\infty}$ resp. $\overline{\Delta}_n$ proving the second assertion. With f as described, the extended solution (cf. theorem 5.29) is solely given by $\bar{U} \equiv \bar{U}_{i_0}$, i. e.

$$\begin{aligned}\bar{U}_{\{i_0\}}(p,t) &= u_{\{i_0\}}(p,t)\chi_{\Delta_0^{(\{i_0\})}}(p) \\ &\quad + \sum_{1\le d\le n}\sum_{i_1\in I_n\backslash\{i_0\}}\cdots\sum_{i_d\in I_n\backslash\{i_0,\ldots,i_{d-1}\}}\bar{u}_{\{i_0\}}^{i_0,\ldots,i_d}(p,t)\chi_{\Delta_d^{(\{i_0,\ldots,i_d\})}}(p) \qquad (5.149)\end{aligned}$$

(cf. equation (5.136)). Considering an arbitrary $\Delta_d^{(I_d)} \subset \overline{\Delta}_n$, $I_d \subset I_n$, we obtain for the restriction of $\bar{U}_{i_0}$ to $\Delta_d^{(I_d)}$ using equation (5.130)

$$\begin{aligned}\bar{U}_{\{i_0\}}(p,t)\big|_{\Delta_d^{(I_d)}} &= \sum_{i_1\in I_d\backslash\{i_0\}}\cdots\sum_{\substack{i_d\in\\ I_d\backslash\{i_0,\ldots,i_{d-1}\}}}\bar{u}_{\{i_0\}}^{i_0,\ldots,i_d}(p,t) \\ &= \sum_{i_1\in I_d\backslash\{i_0\}}\cdots\sum_{\substack{i_d\in\\ I_d\backslash\{i_0,\ldots,i_{d-1}\}}}u_{\{i_0\}}(\pi^{i_0,\ldots,i_d}(p),t)\prod_{j=0}^{d-1}\frac{p^{i_j}}{\sum_{l=j}^{d}p^{i_l}} \qquad (5.150)\end{aligned}$$

with $u_{\{i_0\}}(\pi^{i_0,\ldots,i_d}(p),t) \equiv 1$ as $\pi^{i_0,\ldots,i_d}(p) \in \Delta_0^{(\{i_0\})}$ for all $p \in \Delta_d^{(I_d)}$ and $u_{\{i_0\}} = f = 1$ in $\left(\Delta_0^{(\{i_0\})}\right)_{-\infty}$ by assumption. Since we have $\sum_{l=0}^{d} p^{i_l} = 1$ in $\Delta_d^{(I_d)}$, we may replace the expression $\sum_{l=j}^{d} p^{i_l}$ by $1 - \sum_{l=0}^{j-1} p^{i_l}$ and rearrange the sum (by also suppressing the last sum as the index i_d does no longer occur), which yields altogether

$$\begin{aligned}&\bar{U}_{\{i_0\}}(p,t)\big|_{\Delta_d^{(I_d)}} = \\ &p^{i_0}\left(\sum_{\substack{i_1\in\\ I_d\backslash\{i_0\}}}\frac{p^{i_1}}{1-p^{i_0}}\cdots\left(\sum_{\substack{i_j\in\\ I_d\backslash\{i_0,\ldots,i_{j-1}\}}}\frac{p^{i_j}}{1-\sum_{l=0}^{j-1}p^{i_l}}\cdots\left(\sum_{\substack{i_{d-1}\in\\ I_d\backslash\{i_0,\ldots,i_{d-2}\}}}\frac{p^{i_{d-1}}}{1-\sum_{l=0}^{d-2}p^{i_l}}\right)\right)\right). \qquad (5.151)\end{aligned}$$

As we have $\frac{p^{i_j}+\ldots+p^{i_d}}{1-\sum_{l=0}^{j-1}p^{i_l}} = 1$ for $j = d-1,\ldots,1$, the whole expression reduces to $\bar{U}_{\{i_0\}}(p,t)\big|_{\Delta_d^{(I_d)}} = p^{i_0}$. Since $\Delta_d^{(I_d)}$ was arbitrary, we obtain $\bar{U}_{\{i_0\}}(p,t) \equiv \bar{U}_{\{i_0\}}(p) = p^{i_0}$ in the entire $\overline{\Delta}_n$. □

Regarding the probabilistic interpretation, the extended setting (5.148) also matches the considerations of section 5.4.2 as equation (5.148) may be viewed as the limit equation for $t \to -\infty$ of the extended Kolmogorov backward equation (5.95) (which may be shown as previously). This is also reflected in proposition 5.31: For $t \to -\infty$ and any solution, the only target set with persisting attraction are of course the vertices (respectively corresponding to configurations of the model where all but one allele are extinct), and hence the stationary solutions match the stationary components of the global extensions as in theorem 5.29, which in turn result from a non-vanishing final condition in $\partial_0\Delta_n$. Then, every $\Delta_0^{(\{i\})} \subset \partial_0\Delta_n$ may give rise to a solution (component) p^i – in particular yielding a positive target hit probability on the entire Δ_n for all times. However, it is still noted that even the stationary component of solutions as in theorem 5.29 may in principle be perceived as time-dependent and also describing the transitional attraction of target sets in the entire $\overline{\Delta}_n$ induced by a given ultimate target set in $\partial_0\Delta_n$ (cf. also section 4.2.3).

In total, proposition 5.31 under the given restrictions thus already yields a full description of the stationary model in the entire $\overline{\Delta}_n$. However, dropping the global continuity assumption, a much wider class of (stationary) solutions may be observed as described in the preceding section, and the goal of the following chapter will be to maintain the uniqueness of solutions even for this bigger class.

6 A regularising blow-up scheme for solutions of the extended Kolmogorov backward equation

6.1 Motivation and preliminary considerations

While the preceding chapter was dealing with the construction of extensions of solutions of the Kolmogorov backward equation to higher-dimensional entities of the domain in accordance with the constraints 5.20, which resulted in the depicted extension scheme, we now wish to address the issue of the uniqueness of such extensions.

Standard theory which caters to questions of this kind, however, is inapplicable as the operator does not meet usual regularity requirements with its degeneracy behaviour. For example, the (universal) uniqueness result in [26], pp. 177 f. for combined Dirichlet–Poisson problems is not applicable as the square root of the operator L^* does not fulfil a Lipschitz continuity condition at the boundary, which is mandatory. Alternatively, in order to resolve the uniqueness issue, we will pursue a strategy here which is specifically adapted to the situation at hand. This eventually proves successful if we stipulate some additional regularity properties, and correspondingly we may show that for the stationary components of the extension of solutions as in proposition 5.25 for a given extension path as well as the global extensions as they appear in proposition 5.28 and in theorem 5.29 are unique.

This strategy utilised here is aimed at gaining global regularity in the closure of the domain by resolving any incompatibilities between different boundary faces, which are typical for the considered pathwise iteration of the extension of solutions in accordance with the extension constraints 5.20 (the first extension step, however, does not yet cause incompatibilities). That will be achieved by an appropriate transformation of the relevant part of the domain (i. e. the simplex Δ_n) which transports the whole problem to the corresponding image domain of a product of a simplex and a cube. Simultaneously, the iteratively extended solutions of proposition 5.25 are turned into corresponding solutions of the transformed equation, which are then of sufficient global regularity, in particular are globally continuous. For generic iteratively extended solutions in accordance with the extension constraints 5.20, this does not yet yield a corresponding regularity, however, their transformation image may assumingly be extended that way. Alternatively, such an assumption may also be based on certain reasonable properties of the underlying model as will be discussed in section 6.5.

With such regularised solutions at hand, the further advancement for proving the uniqueness in the stationary case (corresponding to $k = 0$ in proposition 5.25) is straightforward: Analogous to a globally continuous solution of the original problem in Δ_n (cf. section 5.4.7), such a solution is uniquely defined by its values on the vertices of the domain by virtue of the maximum principle, which will be detailed in section 6.4. It then only remains to be shown that the extension constraints 5.20 already uniquely confine sufficient boundary data.

In order to illustrate the motivation for the regularisation scheme using the example of $\bar{U}_{I_k}^{i_k,\dots,i_n}$ in $\overline{\Delta_n^{(I_n)}}$ (cf. equation (5.130)), an assessment of the geometrical situation of the respective incompatibilities reveals that for every $t < 0$ the critical area for the top-dimensional component $\bar{u}_{I_k}^{i_k,\dots,i_n}$ resp. its continuous extension actually only consists of the domain where we have $p^{i_n} + p^{i_{n-1}} = 0$, hence $\overline{\Delta_{n-2}^{(I_{n-2})}}$. On all other boundary instances of arbitrary dimension, $\bar{u}_{I_k}^{i_k,\dots,i_n}$ as in equation (5.129) is continuously extendable and of class C^∞ with respect to the spatial variables there. Thus, at first there is only one connected component of the boundary gap which needs to be addressed.

However, as will turn out, the full hierarchical solution $\bar{U}_{I_k}^{i_k,\dots,i_n}$ actually comprises a nested incompatibility in $\overline{\Delta_{n-2}^{(I_{n-2})}}$ in the sense that also $\bar{u}_{I_k}^{i_k,\dots,i_{n-1}}$ does not extend continuously to $\overline{\Delta_{n-3}^{(I_{n-3})}}$ and so forth until $\bar{u}_{I_k}^{i_k,i_{k+1},i_{k+2}}$ not extending continuously to $\overline{\Delta_k^{(I_k)}}$. This implies that the desired transformation needs to affect all relevant dimensions, which will be accomplished by an iterative advancement: In each step, we will remove one dimension from the simplex and convert it into a dimension of the corresponding cube component, i. e. the corresponding coordinate is released from the simplex property $\sum_i p^i \leq 1$. In doing so, the solution gains the required regularity at the corresponding level with each iteration, i. e. eventually each of its components is transformed such that it extends smoothly to the boundary. Thus, after $n-k-1$ of these steps, the relevant component of $\overline{\Delta_n^{(I_n)}}$ is converted into a cube of dimension $n-k-1$, and we may show that the correspondingly transformed solution is sufficiently regularised, in particular meaning that it now smoothly extends to the full boundary.

6.2 The cube and further notation

Prior to describing the blow-up transformation in full detail in the following section, we wish to introduce some additional notation for the appearing cubes and their boundary instances: In conjunction to the definitions for Δ_n in section 5.1.1, we define for $n \in \mathbb{N}$ an *n-dimensional cube* $\square_n$ as

$$\square_n := \big\{(p^1,\ldots,p^n)\big|p^i \in (0,1) \text{ for } i=1,\ldots,n\big\}. \tag{6.1}$$

Analogous to Δ_n, if we wish to denote the corresponding coordinate indices explicitly, this may be done by providing the coordinate index set $I_n' := \{i_1,\ldots,i_n\} \subset \{1,\ldots,n\}$, $i_j \neq i_l$ for $j \neq l$ as upper index of $\square_n$, thus

$$\square_n^{(I_n')} = \big\{(p^1,\ldots,p^n)\big|p^i \in (0,1) \text{ for } i \in I_n'\big\}. \tag{6.2}$$

This is particularly useful for boundary instances of the cube (cf. below) or if for other purposes a certain ordering $(i_j)_{j=0,\ldots,n}$ of the coordinate indices is needed. For $\square_n$ itself and if no ordering is needed, the index set may be omitted (in such a case it may be assumed $I_n' \equiv \{1,\ldots,n\}$ as in equation (6.1)). Please note that a primed index set is always assumed to not contain index 0 (resp. $i_0 = 0$, which we usually stipulate in case of orderings) as the cube does not encompass a 0th coordinate.

In the standard topology on $\mathbb{R}^n$, $\square_n$ is open (which we always assume when writing $\square_n$), and its closure $\overline{\square}_n$ is given by (again using the index set notation)

$$\overline{\square_n^{(I'_n)}} = \big\{(p^1,\dots,p^n)\big|p^i \in [0,1] \text{ for } i \in I'_n\big\}. \tag{6.3}$$

Similarly to the simplex, the boundary $\partial\square_n$ of $\square_n$ consists of various subcubes (faces) of descending dimensions, starting from the $(n-1)$-dimensional facets down to the vertices (which represent 0-dimensional cubes). All appearing subcubes of dimension $0 \le k \le n-1$ are isomorphic to the k-dimensional standard cube $\square_k$ and hence will be denoted by $\square_k$ if it is irrelevant or given by the context which subcube exactly shall be addressed. However, we may state $\square_k^{(I'_k)}$ with the index set $I'_k := \{i_1,\dots,i_k\} \subset I'_n$, $i_j \neq i_l$ for $j \neq l$ stipulating that I'_k lists all k 'free' coordinate indices, whereas the remaining coordinates are fixed at zero, i. e.

$$\square_k^{(I'_k)} := \big\{(p^1,\dots,p^n)\big|p^i \in (0,1) \text{ for } i \in I'_k;\ p^i = 0 \text{ for } i \in I'_n \setminus I'_k\big\} \tag{6.4}$$

down until $\square_0^{(\varnothing)} := (0,\dots,0)$ for $k=0$.

For a given k, there are of course $\binom{n}{k}$ different (unordered) subsets I'_k of I'_n, each of which corresponds to a certain boundary face $\square_k^{(I'_k)}$. Moreover, for each subset I'_k with k elements, altogether $2^{(n-k)}$ subcubes of dimension k exist in $\partial\square_n$, which are isomorphic to $\square_k^{(I'_k)}$ (including $\square_k^{(I'_k)}$) depending on the (respectively fixed) values of the coordinates with indices not in I'_k. Thus, if necessary, we may rather state a certain boundary face $\square_k$ of $\partial\square_n$ for $0 \le k \le n-1$ by only giving the values of the $n-k$ fixed coordinates, i. e. with indices in $I'_n \setminus I'_k$, which may be either 0 or 1, hence

$$\square_k = \big\{p^{j_1} = b_1,\dots,p^{j_{n-k}} = b_{n-k}\big\} \tag{6.5}$$

with $j_1,\dots,j_{n-k} \in I'_n$, $i_r \neq i_s$ for $r \neq s$ and $b_1,\dots,b_{n-k} \in \{0,1\}$ chosen accordingly. In particular for dimension $n-1$, it is noted that we have $n-1$ faces, which each appear twice; in zero dimension, there are 2^n vertices. If we wish to indicate the total *k-dimensional boundary of* $\square_n$, i. e. the union of all k-dimensional faces belonging to $\overline{\square}_n$, we may write $\partial_k\square_n$ for $k = 0,\dots,n$ with analogously $\partial_n\square_n := \square_n$.

Lastly, when writing products of simplex and cube which do not span all considered dimensions, we indicate the value of the missing coordinates by curly brackets marked with the corresponding coordinate index, i. e. for $I_n = \{i_0, i_1,\dots,i_n\}$ and $I_k \subset I_n$

with $i_{k+1} \notin I_k$ we have e. g.

$$\Delta_k^{(I_k)} \times \{1\}^{(\{i_{k+1}\})} \times \square_{n-k-1}^{(I'_n \setminus (I'_k \cup \{i_{k+1}\}))}$$
$$:= \left\{(p^{i_1}, \dots, p^{i_n}) \middle| p^i > 0 \text{ for } i \in I_k, p^{i_{k+1}} = 1, p^j \in (0,1) \text{ for } j \in I'_n \setminus (I'_k \cup \{i_{k+1}\})\right\} \tag{6.6}$$

with $p^{i_0} = p^0 = 1 - \sum_{j=1}^k p^{i_j}$. If coordinates are fixed at 0, the corresponding entry may be omitted, e. g. we may just write $\Delta_k^{(I_k)}$ for $\underline{\Delta_k^{(I_k)} \times \{0\}^{(I_n \setminus I_k)}}$.

Furthermore, we also introduce a (closed) cube $\overline{\square_k^{(I'_k)}}$ with a removed base vertex $\square_0^{(\varnothing)}$ somewhat inexactly denoted by $\overline{\boxtimes_k^{(I'_k)}}$, i. e.

$$\overline{\boxtimes_k^{(I'_k)}} := \overline{\square_k^{(I'_k)}} \setminus \square_0^{(\varnothing)} = \left\{p^{i_1}, \dots, p^{i_k} \in [0,1] \middle| \sum_{j=1}^k p^{i_j} > 0\right\}. \tag{6.7}$$

For functions defined on the cube, the pathwise smoothness required for an application of the yet to be introduced corresponding Kolmogorov backward operator (cf. p. 182) may be defined as with the simplex in equality (5.96); hence, we put

$$\tilde{u} \in C_p^l(\overline{\square}_n) \;:\Leftrightarrow\; \tilde{u}|_{\square_d \cup \partial_{d-1}\square_d} \in C^l(\square_d \cup \partial_{d-1}\square_d) \text{ for every } \square_d \subset \overline{\square}_n \tag{6.8}$$

with respect to the spatial variables, implying that the operator is continuous at all boundary transitions within $\overline{\square}_n$. This concept may likewise apply to subsets of $\overline{\square}_n$ where needed.

6.3 The blow-up transformation and its iteration

In this section, we will present the blow-up transformation in full detail and state all necessary results. We start with the findings for the basic transformation (which is already formulated to hold in arbitrary dimension as required later) and will later advance to the results for a suitably iterated application of the blow-up transformation:

6.1 Lemma (blow-up transformation)**.** *Let $I_d = \{0, 1, \dots, d\}$. A blow-up transfor-*

mation Φ^r_s *with* $r, s \in I_d \setminus \{0\}$ *mapping*

$$\overline{\Delta^{(I_d)}_d} \setminus \overline{\Delta^{(I_d\setminus\{r,s\})}_{d-2}} = \left\{(p^1,\ldots,p^d)\middle|p^i \geq 0 \text{ for } i \in I_d, p^r + p^s > 0\right\} \tag{6.9}$$

with $p^0 := 1 - \sum_{i\in I_d\setminus\{0\}} p^i$ C^∞*-diffeomorphically onto*

$$\begin{aligned}&\left(\overline{\Delta^{(I_d\setminus\{s\})}_{d-1}} \setminus \overline{\Delta^{(I_d\setminus\{r,s\})}_{d-2}}\right) \times \overline{\Box^{(\{s\})}_1}\\ &\qquad = \left\{(\tilde{p}^1,\ldots,\tilde{p}^d)\middle|\tilde{p}^i \geq 0 \text{ for } i \in I_d \setminus \{s\}, \tilde{p}^r > 0; \tilde{p}^s \in [0,1]\right\}\end{aligned} \tag{6.10}$$

with $\tilde{p}^0 := 1 - \sum_{i\in I_d\setminus\{0,s\}} \tilde{p}^i$ *and altogether*

$$\overline{\Delta^{(I_d)}_d} \longmapsto \left(\overline{\Delta^{(I_d\setminus\{s\})}_{d-1}} \times \overline{\Box^{(\{s\})}_1}\right) \setminus N_r \tag{6.11}$$

with

$$N_r := \overline{\Delta^{(I_d\setminus\{r,s\})}_{d-2}} \times \{0\}^{(\{r\})} \times \overline{\boxtimes^{(\{s\})}_1}, \tag{6.12}$$

appearing as an additional $(d-1)$*-dimensional face of* $\overline{\Delta^{(I_d\setminus\{s\})}_{d-1}} \times \overline{\Box^{(\{s\})}_1}$, *is given by*

$$\tilde{p}^i := \quad p^i \qquad \text{for } i \neq r, s, \tag{6.13}$$

$$\tilde{p}^r := \quad p^r + p^s, \tag{6.14}$$

$$\tilde{p}^s := \begin{cases} \frac{p^s}{p^r+p^s} & \text{for } p^r + p^s > 0 \\ 0 & \text{for } p^r + p^s = 0. \end{cases} \tag{6.15}$$

6.2 Corollary. *While we obtain* $N_r = \overline{\Delta^{(I_d\setminus\{r,s\})}_{d-2}} \times \overline{\boxtimes^{(\{s\})}_1}$ *as an additional* $(d-1)$*-dimensional face with* Φ^r_s*, the existing* $(d-1)$*-dimensional faces of* $\overline{\Delta^{(I_d)}_d}$ *including their boundaries are mapped as follows:*

$$\overline{\Delta^{(I_d\setminus\{s\})}_{d-1}} \longmapsto \overline{\Delta^{(I_d\setminus\{s\})}_{d-1}} \times \{0\}^{(\{s\})}, \tag{6.16}$$

$$\overline{\Delta^{(I_d\setminus\{r\})}_{d-1}} \setminus \overline{\Delta^{(I_d\setminus\{r,s\})}_{d-2}} \longmapsto \left(\overline{\Delta^{(I_d\setminus\{s\})}_{d-1}} \setminus \overline{\Delta^{(I_d\setminus\{r,s\})}_{d-2}}\right) \times \{1\}^{(\{s\})} \tag{6.17}$$

and

$$\overline{\Delta_{d-1}^{(I_d\setminus\{i\})}} \setminus \overline{\Delta_{d-3}^{(I_d\setminus\{i,r,s\})}} \longmapsto \left(\overline{\Delta_{d-2}^{(I_d\setminus\{i,s\})}} \setminus \overline{\Delta_{d-3}^{(I_d\setminus\{i,r,s\})}}\right) \times \overline{\square_1^{(\{s\})}} \quad \textit{for } i \in I_d \setminus \{r,s\}. \tag{6.18}$$

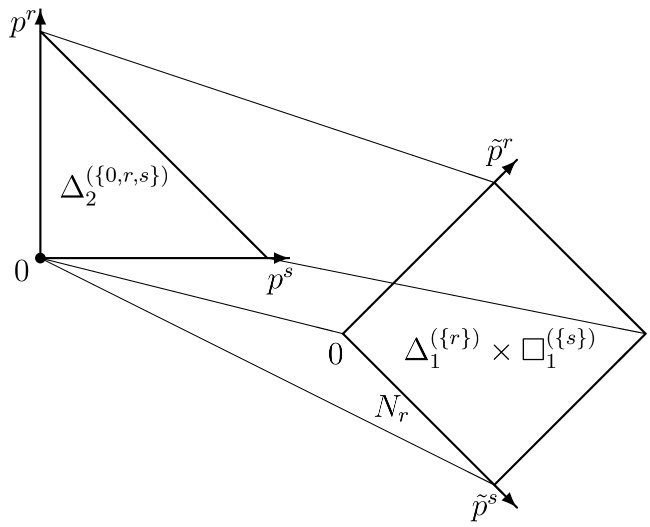

Figure 6.1: An illustration of the blow-up transformation for $d = 2$

6.3 Remark. If the $\tilde{p}^s$ in lemma 6.1 is chosen differently with

$$\tilde{p}^s := \frac{p^r}{p^r + p^s}, \tag{6.19}$$

this flips the orientation of the $\tilde{p}^s$-coordinate in $\square_1^{(\{s\})}$ as all appearances of $\tilde{p}^s$ now need to be replaced by $1 - \tilde{p}^s$. This, however, does not affect the statements of lemma 6.1, whereas in corollary 6.2 the images of $\overline{\Delta_{d-1}^{(I_d\setminus\{r\})}} \setminus \overline{\Delta_{d-2}^{(I_d\setminus\{r,s\})}}$ and $\overline{\Delta_{d-1}^{(I_d\setminus\{s\})}} \setminus \overline{\Delta_{d-2}^{(I_d\setminus\{r,s\})}}$ are interchanged. Thus, unless stated differently, in the following we will always assume that the $\tilde{p}^s$-coordinate is chosen with an orientation as given in lemma 6.1.

Proof of lemma 6.1. The transformation corresponds geometrically to a scaling of the domain into $\tilde{p}^s$-direction with the scaling factor given by $\frac{1}{\tilde{p}^r}$. The assertion about the transformation domains is straightforward since we have $0 \leq \frac{p^s}{p^r+p^s} \leq 1$ on $\overline{\Delta_d^{(I_d)}} \setminus \overline{\Delta_{d-2}^{(I_d\setminus\{r,s\})}}$. Likewise, the C^∞-diffeomorphism property follows from Φ_s^r being smoothly differentiable as long as $\tilde{p}^r = p^r + p^s > 0$ and the smoothness of the inverse transformation $(\Phi_s^r)^{-1}$, given by

$$p^r = \tilde{p}^r(1 - \tilde{p}^s), \tag{6.20}$$

$$p^s = \tilde{p}^r \tilde{p}^s, \tag{6.21}$$

$$p^i = \tilde{p}^i \quad \text{for } i \neq r, s. \tag{6.22}$$

By this, it also becomes obvious that $(\Phi_s^r)^{-1}$ maps $\left(\overline{\Delta_{d-1}^{(I_d \setminus \{s\})}} \setminus \overline{\Delta_{d-2}^{(I_d \setminus \{r,s\})}}\right) \times \overline{\square_1^{(\{s\})}}$ onto $\overline{\Delta_d^{(I_d)}} \setminus \overline{\Delta_{d-2}^{(I_d \setminus \{r,s\})}}$. □

The next lemma is concerned with the transformation behaviour of the operator L_n^*; all considerations apply to L_n^* in its domain Δ_n as well as – considering the restrictability of L_n^* (cf. lemma 5.3) – in the closure $\overline{\Delta}_n$ resp. to the transformed operator $\tilde{L}_n^*$ in the subsequent transformation images of the domain (the domain in question may not be stated explicitly – this will be done in proposition 6.5):

6.4 Lemma. *Let $I'_n := \{1, \dots, n\}$ be an index set with $r, s \in I'_n$ and let $\{i_1, \dots, i_n\}$ be an ordering of I'_n such that $r, s \in \{i_1, \dots, i_m\}$ for some $m \leq n$. When changing coordinates $(p^i)_{i \in I'_n} \mapsto (\tilde{p}^i)_{i \in I'_n}$ by Φ_s^r, the operator*

$$L_n^* = \frac{1}{2} \sum_{i,j=1}^{n} a^{ij}(p) \frac{\partial}{\partial p^i} \frac{\partial}{\partial p^j} \tag{6.23}$$

with $a^{ij}(p) = p^i(\delta_j^i - p^j)$ for $i, j \in \{i_1, \dots, i_m\}$, $a^{ij} = 0$ else for $i \neq j$ is transformed into

$$\tilde{L}_n^* = \frac{1}{2} \sum_{k,l=1}^{k} \tilde{a}^{kl}(\tilde{p}) \frac{\partial}{\partial \tilde{p}^k} \frac{\partial}{\partial \tilde{p}^l} \tag{6.24}$$

with $\tilde{a}^{kl}(\tilde{p}) = \tilde{p}^k(\delta_l^k - \tilde{p}^l)$ for $k, l \in \{i_1, \dots, i_m\} \setminus \{s\}$, $\tilde{a}^{ss}(\tilde{p}) = \frac{\tilde{p}^s(1-\tilde{p}^s)}{\tilde{p}^r}$, $\tilde{a}^{sl} = \tilde{a}^{ls} = 0$ for $l \neq s$ and $\tilde{a}^{kl}(\tilde{p}) = a^{kl}(p)$ (with the coordinates yet to be replaced) for all remaining indices. This also holds if the $\tilde{p}^s$-coordinate is chosen with opposite orientation (cf. remark 6.3).

Proof. Under a change of coordinates $(p^i) \mapsto (\tilde{p}^i)$, the coefficients of the 2nd order derivatives a^{ij} transform as (cf. lemma 2.2)

$$\tilde{a}^{kl} = \sum_{i,j} a^{ij} \frac{\partial \tilde{p}^k}{\partial p^i} \frac{\partial \tilde{p}^l}{\partial p^j}, \tag{6.25}$$

while we may get additional first order derivatives with coefficients $\sum_{i,j} a^{ij} \frac{\partial^2 \tilde{p}^k}{\partial p^i \partial p^j}$.

For the transformation at hand, we have (cf. equations (6.14) and (6.13))

$$\frac{\partial \tilde{p}^k}{\partial p^i} = \delta_i^k + \delta_r^k \delta_i^s \quad \text{for } k \neq s \tag{6.26}$$

and (cf. equation (6.15))

$$\frac{\partial \tilde{p}^s}{\partial p^i} = \frac{p^r}{(p^r + p^s)^2} \delta_i^s - \frac{p^s}{(p^r + p^s)^2} \delta_i^r = \frac{1 - \tilde{p}^s}{\tilde{p}^r} \delta_i^s - \frac{\tilde{p}^s}{\tilde{p}^r} \delta_i^r. \tag{6.27}$$

Utilising this, we obtain

$$\tilde{a}^{kl}(\tilde{p}) = \sum_{i,j} a^{ij}(p)(\delta_i^k + \delta_r^k \delta_i^s)(\delta_j^l + \delta_r^l \delta_j^s) \tag{6.28}$$

for $k, l \neq s$, yielding

$$\begin{aligned} \tilde{a}^{kl}(\tilde{p}) &= a^{kl}(p) + a^{kt}(p)\delta_r^l + a^{sl}(p)\delta_r^k + a^{ss}(p)\delta_r^k \delta_r^l \\ &= p^k(\delta_l^k - p^l) - p^k p^s \delta_r^l - p^s p^l \delta_r^k + p^s(1 - p^s)\delta_r^k \delta_r^l \\ &= \tilde{p}^k(\delta_l^k - \tilde{p}^l) \end{aligned} \tag{6.29}$$

for $k, l \in \{i_1, \dots, i_m\} \setminus \{s\}$ using the given form of the a^{ij}, whereas for all other index pairs not containing the index t, we always have

$$a^{kt}(p)\delta_r^l = a^{sl}(p)\delta_r^k = a^{ss}(p)\delta_r^k \delta_r^l = 0 \tag{6.30}$$

and hence

$$\tilde{a}^{kl}(\tilde{p}) = \sum_{i,j} a^{ij}(p)\delta_i^k \delta_j^l = a^{kl}(p), \tag{6.31}$$

thus proving the last statement. Furthermore, we have for arbitrary $l \neq s$

$$\begin{aligned}
\tilde{a}^{sl}(\tilde{p}) &= \sum_{i,j} a^{ij}(p)\bigg(\frac{1-\tilde{p}^s}{\tilde{p}^r}\delta_i^s - \frac{\tilde{p}^s}{\tilde{p}^r}\delta_i^r\bigg)(\delta_j^l + \delta_r^l\delta_j^s) \\
&= \frac{1-\tilde{p}^s}{\tilde{p}^r}(a^{sl}(p) + a^{ss}(p)\delta_r^l) - \frac{\tilde{p}^s}{\tilde{p}^r}(a^{rl}(p) + a^{rt}(p)\delta_r^l) \\
&= \Big(- \frac{1-\tilde{p}^s}{\tilde{p}^r}\tilde{p}^r\tilde{p}^s\tilde{p}^l + \frac{\tilde{p}^s}{\tilde{p}^r}(1-\tilde{p}^s)\tilde{p}^r\tilde{p}^l\Big)\chi_{\{i_1,\dots,i_m\}}(l) \\
&\quad - \frac{\tilde{p}^s}{\tilde{p}^r}\tilde{p}^r(1-\tilde{p}^s)\delta_r^l + \Big(\frac{1-\tilde{p}^s}{\tilde{p}^r}\tilde{p}^r\tilde{p}^s(1-\tilde{p}^r\tilde{p}^s) + \frac{\tilde{p}^s}{\tilde{p}^r}\tilde{p}^r(1-\tilde{p}^s)\tilde{p}^r\tilde{p}^s\Big)\delta_r^l = 0
\end{aligned} \tag{6.32}$$

as well as $\tilde{a}^{ls} = 0$ $(l \neq s)$ by symmetry and eventually

$$\begin{aligned}
\tilde{a}^{ss}(\tilde{p}) &= \sum_{i,j} a^{ij}(p)\bigg(\frac{1-\tilde{p}^s}{\tilde{p}^r}\delta_i^s - \frac{\tilde{p}^s}{\tilde{p}^r}\delta_i^r\bigg)\bigg(\frac{1-\tilde{p}^s}{\tilde{p}^r}\delta_j^s - \frac{\tilde{p}^s}{\tilde{p}^r}\delta_j^r\bigg) \\
&= a^{ss}(p)\bigg(\frac{1-\tilde{p}^s}{\tilde{p}^r}\bigg)^2 + a^{rr}(p)\bigg(\frac{\tilde{p}^s}{\tilde{p}^r}\bigg)^2 - 2a^{sr}(p)\frac{\tilde{p}^s(1-\tilde{p}^s)}{(\tilde{p}^r)^2} \\
&= \tilde{p}^s(1-\tilde{p}^r\tilde{p}^s)\frac{(1-\tilde{p}^s)^2}{\tilde{p}^r} + (1-\tilde{p}^s)(1-\tilde{p}^r+\tilde{p}^r\tilde{p}^s)\frac{(\tilde{p}^s)^2}{\tilde{p}^r} \\
&\quad - 2\tilde{p}^r\tilde{p}^s(1-\tilde{p}^s)\frac{\tilde{p}^s(1-\tilde{p}^s)}{\tilde{p}^r} = \frac{\tilde{p}^s(1-\tilde{p}^s)}{\tilde{p}^r},
\end{aligned} \tag{6.33}$$

by which the form of all $\tilde{a}^{kl}$ is shown.

When checking for possible additional first order derivatives, it is obvious that the second order coordinate derivatives do not vanish at first glance only for $\tilde{p}^s$. But we have (cf. equation (6.27))

$$\frac{\partial}{\partial p^j}\frac{\partial}{\partial p^i}\tilde{p}^s = \frac{2}{(p^r+p^s)^3}(p^s\delta_i^r - p^r\delta_i^s)(\delta_j^r + \delta_j^s) + \frac{1}{(p^r+p^s)^2}(\delta_i^s\delta_j^r - \delta_i^r\delta_j^s) \tag{6.34}$$

and subsequently

$$\begin{aligned}
\sum_{i,j} a^{ij}\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}\tilde{p}^s &= \frac{2}{(p^r+p^s)^3}\big(p^s(a^{rr}+a^{rs}) - p^r(a^{sr}+a^{ss})\big) + \frac{1}{(p^r+p^s)^2}(a^{sr}-a^{rs}) \\
&= \frac{2}{(p^r+p^s)^3}\big(p^sp^r(1-p^r-p^s) + p^rp^s(p^r-1+p^s)\big) = 0,
\end{aligned} \tag{6.35}$$

for which again the specified form of the appearing a^{ij} is needed.

If $\tilde{p}^s$ is chosen with different orientation as in remark 6.3, instead of equation (6.27) we then have

$$\frac{\partial \tilde{p}^s}{\partial p^i} = \frac{\tilde{p}^s}{\tilde{p}^r}\delta^s_i - \frac{1-\tilde{p}^s}{\tilde{p}^r}\delta^r_i, \tag{6.36}$$

signifying that in the respective formulae the indices r and s are swapped, which in turn is matched by the corresponding inverse transformation now yielding $p^r = \tilde{p}^r\tilde{p}^s$ and $p^s = \tilde{p}^r(1-\tilde{p}^s)$. □

Combining the preceding results, we obtain for an iterated application of the blow-up transformation:

6.5 Proposition. *Let $k, n \in \mathbb{N}$ with $0 \le k \le n-2$, $\{i_k, i_{k+1}, \ldots, i_n\} \subset I_n := \{0, 1, \ldots, n\}$ with $i_i \neq i_j$ for $i \neq j$ and $I_d := I_n \setminus \{i_{d+1}, \ldots, i_n\}$ for $d = k, \ldots, n-1$. A repeated blow-up transformation $\Phi^{r_{n-k-1}}_{s_{n-k-1}} \circ \ldots \circ \Phi^{r_1}_{s_1}$ with $\Phi^{r_m}_{s_m}$ as in lemma 6.1 with $r_m = i_{n-m}$ and $s_m = i_{n-m+1}$ for $m = 1, \ldots, n-k-1$ maps $\overline{\Delta^{(I_{k+1})}_{k+1}}$ onto itself and*

$$\Delta^{(I_d)}_d \longmapsto \Delta^{(I_{k+1})}_{k+1} \times \square^{(I_d \setminus I_{k+1})}_{d-k-1} \quad \text{for } d = k+2, \ldots, n \tag{6.37}$$

and altogether

$$\overline{\Delta^{(I_n)}_n} \longmapsto \left(\overline{\Delta^{(I_{k+1})}_{k+1}} \times \overline{\square^{(I_n \setminus I_{k+1})}_{n-k-1}}\right) \setminus \bigcup_{j=k+1}^{n-1} N_j. \tag{6.38}$$

The $n-k-1$ additional $(n-1)$-dimensional faces $N_{k+1}, \ldots, N_{n-1}$ of $\overline{\Delta^{(I_{k+1})}_{k+1}} \times \overline{\square^{(I_n \setminus I_{k+1})}_{n-k-1}}$ are given by

$$N_{k+1} = \overline{\Delta^{(I_k)}_k} \times \{0\}^{(\{i_{k+1}\})} \times \overline{\boxtimes^{(I_n \setminus I_{k+1})}_{n-k-1}} \tag{6.39}$$

and

$$N_j = \overline{\Delta^{(I_{k+1})}_{k+1}} \times \overline{\square^{(I_{j-1} \setminus I_{k+1})}_{j-k-2}} \times \{0\}^{(\{i_j\})} \times \overline{\boxtimes^{(I_n \setminus I_j)}_{n-j}} \tag{6.40}$$

for $j = k+2, \ldots, n-1$. Simultaneously, the operator $L^* = \sum p^i(\delta^i_j - p^j)\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}$ in

$\overline{\Delta_n^{(I_n)}}$ *is transformed into*[1]

$$\tilde{L}^* = \frac{1}{2}\sum_{j,l=1}^{k+1} \tilde{p}^{i_j}(\delta_l^j - \tilde{p}^{i_l})\frac{\partial}{\partial \tilde{p}^{i_j}}\frac{\partial}{\partial \tilde{p}^{i_l}} + \frac{1}{2}\sum_{j=k+2}^{n} \frac{\tilde{p}^{i_j}(1-\tilde{p}^{i_j})}{\prod_{l=k+1}^{j-1} \tilde{p}^{i_l}} \frac{\partial^2}{(\partial \tilde{p}^{i_j})^2} \tag{6.41}$$

in $\left(\overline{\Delta_{k+1}^{(I_{k+1})}} \times \overline{\square_{n-k-1}^{(I_n \setminus I_{k+1})}}\right) \setminus \bigcup_{j=k+1}^{n-1} N_j$.

If in any step the coordinate $\tilde{p}^{s_j}$ *is chosen with alternative orientation (cf. remark 6.3), all appearances of* $\tilde{p}^{s_j}$ *in the above formulae are replaced by* $(1-\tilde{p}^{s_j})$.

Thus, the iterated blow-up translates the (extended) Kolmogorov backward equation in $\overline{\Delta}_n$ into a corresponding differential equation in $\left(\overline{\Delta_{k+1}^{(I_{k+1})}} \times \overline{\square_{n-k-1}^{(I_n \setminus I_{k+1})}}\right) \setminus \bigcup_{j=k+1}^{n-1} N_j$. For the successively extended solutions of the Kolmogorov backward equation introduced in the preceding chapter, the transformation behaviour is as follows:

6.6 Proposition. *Let* $k, n \in \mathbb{N}$ *with* $0 \le k \le n-2$, $\{i_k, i_{k+1}, \ldots, i_n\} \subset I_n := \{0, 1, \ldots, n\}$ *with* $i_i \neq i_j$ *for* $i \neq j$ *and* $I_d := I_n \setminus \{i_{d+1}, \ldots, i_n\}$ *for* $d = k, \ldots, n-1$, *and let* u_{I_k} *in* $\left(\Delta_k^{(I_k)}\right)_{-\infty}$ *and* $\bar{U}_{I_k}^{i_k,\ldots,i_n}$ *in* $\left(\bigcup_{k\le d\le n} \Delta_d^{(I_d)}\right)_{-\infty}$ *as in proposition 5.25. Then a repeated blow-up transformation* $\Phi_{s_{n-k-1}}^{r_{n-k-1}} \circ \ldots \circ \Phi_{s_1}^{r_1}$ *with* $\Phi_{s_m}^{r_m}$ *as in lemma 6.1 with* $r_m = i_{n-m}$ *and* $s_m = i_{n-m+1}$ *for* $m = 1, \ldots, n-k-1$ *converts*

$$\begin{aligned} \bar{U}_{I_k}^{i_k,\ldots,i_n}(p,t) &:= u_{I_k}(p,t)\chi_{\Delta_k^{(I_k)}}(p) + \sum_{k+1\le d\le n} \bar{u}_{I_k}^{i_k,\ldots,i_d}(p,t)\chi_{\Delta_d^{(I_d)}}(p) \\ &= u_{I_k}(p,t)\chi_{\Delta_k^{(I_k)}}(p) + \sum_{k+1\le d\le n} u_{I_k}(\pi^{i_k,\ldots,i_d}(p),t)\prod_{j=k}^{d-1}\frac{p^{i_j}}{\sum_{l=j}^{d} p^{i_l}}\chi_{\Delta_d^{(I_d)}}(p) \end{aligned} \tag{6.42}$$

[1]Please note that on boundary instances of $\square_{n-k-1}^{(I_n \setminus I_{k+1})}$, i.e. $\tilde{p}^{i_l} = 0$ for some $l \in I_n \setminus I_{k+1}$, the corresponding summands are assumed not to appear in the right sum in equation (6.41), which may be interpreted as a result of a successive restriction. The given domain is the maximal domain for the operator as it is not defined on the exception set $\bigcup_{j=k+1}^{n-1} N_j$ (however, cf. also lemma 6.9 for the stationary case).

on $\left(\bigcup_{k\leq d\leq n}\Delta_d^{(I_d)}\right)_{-\infty}$ *into*

$$\tilde{U}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_n}(\tilde{p},t) := u_{I_k}(\tilde{p},t)\chi_{\Delta_k^{(I_k)}}(\tilde{p}) + \sum_{k+1\leq d\leq n}\tilde{u}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_d}(\tilde{p},t)\chi_{\Delta_{k+1}^{(I_{k+1})}\times\square_{d-k-1}^{(I_d\setminus I_{k+1})}}(\tilde{p}) \quad (6.43)$$

on $\left(\bigcup_{k\leq d\leq n}\Delta_{k+1}^{(I_{k+1})}\times\square_{n-k-1}^{(I_n\setminus I_{k+1})}\right)_{-\infty}$ *with*

$$\tilde{u}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_d}(\tilde{p},t) := \bar{u}_{I_k}^{i_k,i_{k+1}}(\tilde{\pi}^{i_{k+1}}(\tilde{p}),t)\prod_{j=k+2}^{d}(1-\tilde{p}^{i_j}) \quad \textit{for } d=k+2,\dots,n \quad (6.44)$$

with $\tilde{\pi}^{i_{k+1}}(\tilde{p}^{i_j}) := \tilde{p}^{i_j}$ *for* $i_j\in I_{k+1}$, $\tilde{\pi}^{i_{k+1}}(\tilde{p}^{i_j}) := 0$ *else. The transformed functions* $\tilde{u}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_d}$ *smoothly extend to* $\left(\overline{\Delta_{k+1}^{(I_{k+1})}}\times\overline{\square_{d-k-1}^{(I_d\setminus I_{k+1})}}\right)_{-\infty}$ *respectively; consequently also* $\tilde{U}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_n}$ *smoothly extends to* $\left(\overline{\Delta_{k+1}^{(I_{k+1})}}\times\overline{\square_{n-k-1}^{(I_n\setminus I_{k+1})}}\right)_{-\infty}$. *Furthermore, it may be simplified to*

$$\tilde{U}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_n}(\tilde{p},t) \equiv \tilde{u}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_n}(\tilde{p},t) \quad \textit{in } \left(\overline{\Delta_{k+1}^{(I_{k+1})}}\times\overline{\square_{n-k-1}^{(I_n\setminus I_{k+1})}}\right)_{-\infty}. \quad (6.45)$$

If in any step the coordinate $\tilde{p}^{s_j}$ *is chosen with alternative orientation (cf. remark 6.3), all appearances of* $\tilde{p}^{s_j}$ *in the above formulae need to be replaced with* $(1-\tilde{p}^{s_j})$.

For the stationary components, we have in particular:

6.7 Corollary. *For* $k=0$ *and w. l. o. g.* $i_0=0$, *the transformed function of proposition 6.6 in equation* (6.45) *simplifies to*

$$\tilde{U}_{\{i_0\}}^{i_0,i_1;i_2,\dots,i_n}(\tilde{p}) = u_{\{i_0\}}(1)\cdot\prod_{j=1}^{n}(1-\tilde{p}^{i_j}) \quad \textit{in } \overline{\square_n^{(I'_n)}}, \quad (6.46)$$

while in accordance with proposition 6.5 the domain is mapped

$$\Delta_d^{(I_d)} \longmapsto \square_d^{(I'_d)} \quad \textit{for } d=0,\dots,n \quad (6.47)$$

and altogether

$$\overline{\Delta_n^{(I_n)}} \longmapsto \overline{\square_n^{(I'_n)}} \setminus \bigcup_{j=1}^{n-1} N_j. \tag{6.48}$$

The $n-1$ *additional* $(n-1)$*-dimensional faces* $N_1, \dots, N_{n-1}$ *of* $\partial \square_n^{(I'_n)}$ *are given by*

$$N_1 = \{0\}^{(\{i_1\})} \times \overline{\boxtimes_{n-1}^{(I'_n \setminus I'_1)}} \tag{6.49}$$

and

$$N_j = \overline{\square_{j-1}^{(I'_{j-1})}} \times \{0\}^{(\{i_j\})} \times \overline{\boxtimes_{n-j}^{(I'_n \setminus I'_j)}} \tag{6.50}$$

for $j = 2, \dots, n-1$*, whereas the operator* $L^* = \sum p^i (\delta^i_j - p^j) \frac{\partial}{\partial p^i} \frac{\partial}{\partial p^j}$ *in* $\overline{\Delta_n^{(I_n)}}$ *is transformed into*

$$\tilde{L}^* = \frac{1}{2} \sum_{j=1}^{n} \frac{\tilde{p}^{i_j}(1-\tilde{p}^{i_j})}{\prod_{l=1}^{j-1} \tilde{p}^{i_l}} \frac{\partial^2}{(\partial \tilde{p}^{i_j})^2} \quad \text{in } \overline{\square_n^{(I'_n)}} \setminus \bigcup_{j=1}^{n-1} N_j. \tag{6.51}$$

Proof of propositions 6.5 and 6.6. We prove the assertions of both propositions in parallel: Aiming to transform $\bar{U}_{I_k}^{i_k,\dots,i_n}$ into a function that does not feature any incompatibilities and hence is of sufficient regularity with respect to the entire closure of the (transformed) domain, we show that the full blow-up via a repeated application of the coordinate transformation Φ^r_s of lemma 6.1 with the indices r and s to be picked as shown in each step yields the desired result for $\tilde{U}_{I_k}^{i_k,i_{k+1};i_{k+2},\dots,i_n}$, whereas the transformation behaviour of the domain and the operator is as stated in proposition 6.5. Note that in the designation of any domains, we will usually suppress the t-component throughout this proof, e. g. write $\Delta_n^{(I_n)}$ instead of $\left(\Delta_n^{(I_n)}\right)_{-\infty}$, for notational simplicity.

Starting with the top-dimensional component of $\bar{U}_{I_k}^{i_k,\dots,i_n}$, which is

$$\bar{u}_{I_k}^{i_k,\dots,i_n}(p,t) = \bar{u}_{I_k}^{i_k,\dots,i_{n-1}}(\pi^{i_{n-1},i_n}(p),t) \cdot \frac{p^{i_{n-1}}}{p^{i_{n-1}} + p^{i_n}}$$

$$= u_{I_k}(\pi^{i_k,\dots,i_{n-1}}(\pi^{i_{n-1},i_n}(p)),t)\prod_{j=k}^{n-2}\frac{p^{i_j}}{\sum_{l=j}^{n}p^{i_l}}\cdot\frac{p^{i_{n-1}}}{p^{i_{n-1}}+p^{i_n}} \quad \text{in } \Delta_n^{(I_n)} \tag{6.52}$$

with $p^{i_0} \equiv p^0 = 1 - \sum_{j=1}^{n} p^{i_j}$ (if $i_0 \neq 0$, one may change the coordinates, i. e. permute the vertices correspondingly), we initially put[2] $r_1 := i_{n-1}$ and $s_1 := i_n$. Changing coordinates $(p^i) \mapsto (\tilde{p}^i)$ by $\Phi_{s_1}^{r_1}$ maps $\Delta_n^{(I_n)}$ onto $\Delta_{n-1}^{(I_{n-1})} \times \Box_1^{(\{i_n\})}$ and $\overline{\Delta_{n-1}^{(I_{n-1})}}$ onto $\overline{\Delta_{n-1}^{(I_{n-1})}} \times \{0\}^{(\{i_n\})}$, whereas the entire domain $\overline{\Delta_n^{(I_n)}}$ is transformed into $\left(\overline{\Delta_{n-1}^{(I_{n-1})}} \times \overline{\Box_1^{(\{i_n\})}}\right) \setminus N_{n-1}$ with

$$N_{n-1} := \overline{\Delta_{n-2}^{(I_{n-2})}} \times \{0\}^{(\{i_{n-1}\})} \times \overline{\boxtimes_1^{(\{i_n\})}} \tag{6.53}$$

being an additional $(n-1)$-dimensional face of $\overline{\Delta_{n-1}^{(I_{n-1})}} \times \overline{\Box_1^{(\{i_n\})}}$ (cf. lemma 6.1). Simultaneously, the $(n-2)$-dimensional incompatibility at $\overline{\Delta_{n-2}^{(I_{n-2})}}$ of the continuous extension of $\bar{u}_{I_k}^{i_k,\dots,i_n}$ to $\partial_{n-1}\Delta_n^{(I_n)}$ is removed as the transformation yields

$$\begin{aligned}\tilde{u}_{I_k}^{i_k,\dots,i_{n-1};i_n}(\tilde{p},t) &:= \bar{u}_{I_k}^{i_k,\dots,i_{n-1}}(\tilde{\pi}^{i_{n-1}}(\tilde{p}),t)\cdot(1-\tilde{p}^{i_n}) \\ &= u_{I_k}(\pi^{i_k,\dots,i_{n-1}}(\tilde{\pi}^{i_{n-1}}(\tilde{p})),t)\prod_{j=k}^{n-2}\frac{\tilde{p}^{i_j}}{\sum_{l=j}^{n}\tilde{p}^{i_l}}\cdot(1-\tilde{p}^{i_n}) \\ &\qquad\qquad \text{in } \Delta_{n-1}^{(I_{n-1})} \times \Box_1^{(\{i_n\})}\end{aligned} \tag{6.54}$$

by equation (6.20) et seq. (note $\tilde{\pi}^{i_{n-1}}(\tilde{p}) = \pi^{i_{n-1},i_n}(p)$). Hence, the complete function $\bar{U}_{I_k}^{i_k,\dots,i_n}$ is transformed into

$$\begin{aligned}\tilde{U}_{I_k}^{i_k,\dots,i_{n-1};i_n}(p,t) := &\sum_{k\le d\le n-1}\bar{u}_{I_k}^{i_k,\dots,i_d}(p,t)\chi_{\Delta_d^{(I_d)}}(p) \\ &+ \tilde{u}_{I_k}^{i_k,\dots,i_{n-1};i_n}(p,t)\chi_{\Delta_{n-1}^{(I_{n-1})}\times\Box_1^{(I_n\setminus I_{n-1})}}(p)\end{aligned} \tag{6.55}$$

with the transformed top-dimensional component $\tilde{u}_{I_k}^{i_k,\dots,i_{n-1};i_n}(\tilde{p},t)$ smoothly extending

[2] Alternatively, one could also put $r_1 := i_n$ and $s_1 := i_{n-1}$, which would correspond to inverting the orientation of the $\tilde{p}^{s_1}$-coordinate in accordance with remark 6.3 (cf. also below) plus subsequently swapping the coordinate indices i_n and i_{n-1}, thus $\tilde{p}^{i_n}$ would get replaced with $1-\tilde{p}^{i_{n-1}}$ and $\tilde{p}^{i_{n-1}}$ with $\tilde{p}^{i_n}$.

to $\Delta_{n-1}^{(I_{n-1})} \times \overline{\square_1^{(\{i_n\})}}$ with

$$\tilde{u}_{I_k}^{i_k,\ldots,i_{n-1};i_n}(\tilde{p},t)\big|_{\Delta_{n-1}^{(I_{n-1})}\times\{0\}^{(\{i_n\})}} \equiv \bar{u}_{I_k}^{i_k,\ldots,i_{n-1}}(\tilde{p},t) \quad \text{in } \Delta_{n-1}^{(I_{n-1})} \times \{0\}^{(\{i_n\})}. \tag{6.56}$$

As $\bar{u}_{I_k}^{i_k,\ldots,i_{n-1}}$ itself smoothly extends to $\partial_{n-2}\Delta_{n-1}^{(I_{n-1})}$, thus $\tilde{u}_{I_k}^{i_k,\ldots,i_{n-1};i_n}$ now smoothly extends to the entire $(\partial_{n-2}\Delta_{n-1}^{(I_{n-1})}) \times \overline{\square_1^{(\{i_n\})}}$, in particular to $\Delta_{n-2}^{(I_{n-2})} \times \overline{\square_1^{(\{i_n\})}} \subset \overline{N}_{n-1}$ (however, $\bar{u}_{I_k}^{i_k,\ldots,i_{n-1}}$ resp. its continuous extension to $\partial_{n-2}\Delta_{n-1}^{(I_{n-1})}$ still has an incompatibility at $\overline{\Delta_{n-3}^{(I_{n-3})}}$).

The operator $L^* = \frac{1}{2}\sum_{i,j=1}^n p^i(\delta_j^i - p^j)\frac{\partial}{\partial p^i}\frac{\partial}{\partial p^j}$ in $\overline{\Delta_n^{(I_n)}}$ transforms into (cf. lemma 6.4)

$$\tilde{L}^* = \frac{1}{2}\sum_{j,l\neq n} \tilde{p}^{i_j}(\delta_l^j - \tilde{p}^{i_l})\frac{\partial}{\partial \tilde{p}^{i_j}}\frac{\partial}{\partial \tilde{p}^{i_l}} + \frac{1}{2}\frac{\tilde{p}^{i_n}(1-\tilde{p}^{i_n})}{\tilde{p}^{i_{n-1}}}\frac{\partial}{\partial \tilde{p}^{i_n}}\frac{\partial}{\partial \tilde{p}^{i_n}} \tag{6.57}$$

on $\left(\overline{\Delta_{n-1}^{(I_{n-1})}} \times \overline{\square_1^{(\{i_n\})}}\right) \setminus N_{n-1}$ since we have $\tilde{a}^{kl}(\tilde{p}) = p^k(\delta_l^k - p^l) = \tilde{p}^k(\delta_l^k - \tilde{p}^l)$ for $k,l \neq i_{n-1}, i_n$. If $\tilde{p}^{i_n}$ is chosen with alternative orientation (cf. remark 6.3), then $\tilde{p}^{i_n}$ needs to be replaced by $(1-\tilde{p}^{i_n})$ everywhere.

As already indicated, the transformed solution is still not smoothly extendable to the full boundary of the transformed domain: Its $(n-2)$-dimensional incompatibility is resolved, but its lower-dimensional incompatibilities persist. Thus, the highest-dimensional incompatibility now is of dimension $n-3$, and hence the situation is ready for another application of the blow-up transformation, yielding a corresponding situation afterwards.

Thus, an iterative advancement is necessary to resolve all incompatibilities. For this purpose, we assume that after the m-th step ($m = 1, \ldots, n-k-2$) an already transformed function $\tilde{U}_{I_k}^{i_k,\ldots,i_{n-m};i_{n-m+1},\ldots,i_n}$ with (note that we again associate coordinates p resp. $\tilde{p}$ etc. to the domain before/after the $(m+1)$-th transition; furthermore, we will use the convention $\bar{u}_{I_k}^{i_k} \equiv u_{I_k}$ to simplify the notation)

$$\begin{aligned}\tilde{U}_{I_k}^{i_k,\ldots,i_{n-m};i_{n-m+1},\ldots,i_n}(p,t) &= \sum_{k\leq d\leq n-m} \bar{u}_{I_k}^{i_k,\ldots,i_d}(p,t)\chi_{\Delta_d^{(I_d)}}(p) \\ &+ \sum_{n-m+1\leq d\leq n} \tilde{u}_{I_k}^{i_k,\ldots,i_{n-m};i_{n-m+1},\ldots,i_d}(p,t)\chi_{\Delta_{n-m}^{(I_{n-m})}\times\square_{d-n+m}^{(I_d\setminus I_{n-m})}}(p)\end{aligned} \tag{6.58}$$

with

$$\tilde{u}_{I_k}^{i_k,\dots,i_{n-m};i_{n-m+1},\dots,i_d}(p,t) = \bar{u}_{I_k}^{i_k,\dots,i_{n-m}}(\tilde{\pi}^{i_{n-m}}(p),t) \prod_{j=n-m+1}^{d} (1-p^{i_j}) \tag{6.59}$$

for $d = n-m+1,\dots,n$ and

$$\begin{aligned} \bar{u}_{I_k}^{i_k,\dots,i_{n-m}}(p,t) &= \bar{u}_{I_k}^{i_k,\dots,i_{n-m-1}}(\pi^{i_{n-m-1},i_{n-m}}(p),t) \cdot \frac{p^{i_{n-m-1}}}{p^{i_{n-m-1}}+p^{i_{n-m}}} \\ &= u_{I_k}(\pi^{i_k,\dots,i_{n-m-1}}(\pi^{i_{n-m-1},i_{n-m}}(p)),t) \prod_{j=k}^{n-m-2} \frac{p^{i_j}}{\sum_{l=j}^{n} p^{i_l}} \cdot \frac{p^{i_{n-m-1}}}{p^{i_{n-m-1}}+p^{i_{n-m}}} \end{aligned} \tag{6.60}$$

in $\Delta_{n-m}^{(I_{n-m})}$. The corresponding total domain as an image of $\overline{\Delta_n^{(I_n)}}$ is given by

$$\left(\overline{\Delta_{n-m}^{(I_{n-m})}} \times \overline{\square_m^{(I_n \setminus I_{n-m})}}\right) \setminus \bigcup_{j=n-m}^{n-1} N_j \tag{6.61}$$

with previously additional $(n-1)$-dimensional faces

$$N_{n-m} = \overline{\Delta_{n-m-1}^{(I_{n-m-1})}} \times \{0\}^{(\{i_{n-m}\})} \times \overline{\boxtimes_m^{(I_n \setminus I_{n-m})}} \tag{6.62}$$

and

$$N_j = \overline{\Delta_{n-m}^{(I_{n-m})}} \times \overline{\square_{j-n+m-1}^{(I_{j-1} \setminus I_{n-m})}} \times \{0\}^{(\{i_j\})} \times \overline{\boxtimes_{n-j}^{(I_n \setminus I_j)}} \tag{6.63}$$

for $j = n-m+1,\dots,n-1$.

The functions $\tilde{u}_{I_k}^{i_k,\dots,i_{n-m};i_{n-m+1},\dots,i_d}$ smoothly extend each to $\Delta_{n-m}^{(I_{n-m})} \times \overline{\square_{d-n+m}^{(I_d \setminus I_{n-m})}}$, and we have

$$\tilde{u}_{I_k}^{i_k,\dots,i_{n-m};i_{n-m+1},\dots,i_d}\big|_{\Delta_{n-m}^{(I_{n-m})} \times \square_m^{(I_{d-1} \setminus I_{n-m})}} = \tilde{u}_{I_k}^{i_k,\dots,i_{n-m};i_{n-m+1},\dots,i_{d-1}} \tag{6.64}$$

for $d = n-m+2,\dots,n$ and

$$\tilde{u}_{I_k}^{i_k,\dots,i_{n-m};i_{n-m+1}}\big|_{\Delta_{n-m}^{(I_{n-m})}} = \bar{u}_{I_k}^{i_k,\dots,i_{n-m}}. \tag{6.65}$$

With $\bar{u}_{I_k}^{i_k,\dots,i_{n-m}}$ being smoothly extendable to $\partial_{n-m-1}\Delta_{n-m}^{(I_{n-m})}$, also the functions $\tilde{u}_{I_k}^{i_k,\dots,i_{n-m};i_{n-m+1},\dots,i_d}$ smoothly extend to $\left(\partial_{n-m-1}\Delta_{n-m}^{(I_{n-m})}\right) \times \overline{\square_{d-n+m}^{(I_d \setminus I_{n-m})}}$, in particular all additional faces are covered.

Furthermore, we assume the operator L^* to be of the corresponding form

$$L^* = \frac{1}{2}\sum_{j,l=1}^{n-m} p^{i_j}(\delta_l^j - p^{i_l})\frac{\partial}{\partial p^{i_j}}\frac{\partial}{\partial p^{i_l}} + \frac{1}{2}\sum_{j=n-m+1}^{n} \frac{p^{i_j}(1-p^{i_j})}{\prod_{l=n-m}^{j-1} p^{i_l}} \frac{\partial^2}{(\partial p^{i_j})^2} \tag{6.66}$$

on $\left(\overline{\Delta_{n-m}^{(I_{n-m})}} \times \overline{\square_m^{(I_n \setminus I_{n-m})}}\right) \setminus \bigcup_{j=n-m}^{n-1} N_j$.

For the $(m+1)$-th blow-up step going to be applied now, we first notice that $\bar{u}_{I_k}^{i_k,\dots,i_{n-m}}$ resp. its continuous extension to $\partial_{n-m-1}\Delta_{n-m}^{(I_{n-m})}$ still has an incompatibility at $\overline{\Delta_{n-m-2}^{(I_{n-m-2})}} \subset \Delta_{n-m}^{(I_{n-m})}$, corresponding to $p^{i_{n-m}} + p^{i_{n-m-1}} = 0$. Consequently, this may be resolved by a blow-up transformation $\Phi_{s_{m+1}}^{r_{m+1}}$ with $r_{m+1} = i_{n-m-1}$ and $s_{m+1} = i_{n-m}$ (note that, due to the stipulation $i_0 = 0$, we always have $r_{m+1}, s_{m+1} \neq 0$), mapping the simplex part of the domain (cf. lemma 6.1)

$$\Delta_{n-m}^{(I_{n-m})} \longmapsto \Delta_{n-m-1}^{(I_{n-m-1})} \times \square_1^{(\{i_{n-m}\})} \tag{6.67}$$

resp.

$$\overline{\Delta_{n-m-1}^{(I_{n-m-1})}} \longmapsto \overline{\Delta_{n-m-1}^{(I_{n-m-1})}} \times \{0\}^{(\{i_{n-m}\})} \tag{6.68}$$

and altogether

$$\overline{\Delta_{n-m}^{(I_{n-m})}} \longmapsto \overline{\Delta_{n-m-1}^{(I_{n-m-1})}} \times \overline{\square_1^{(\{i_{n-m}\})}} \setminus N_{n-m-1} \tag{6.69}$$

with

$$N_{n-m-1} := \overline{\Delta_{n-m-2}^{(I_{n-m-2})}} \times \{0\}^{(\{i_{n-m-1}\})} \times \overline{\boxtimes_1^{(\{i_{n-m}\})}} \tag{6.70}$$

being an additional $(n-m-1)$-dimensional face of $\overline{\Delta_{n-m-1}^{(I_{n-m-1})}} \times \overline{\square_1^{(\{i_{n-m}\})}}$.

From this – when gradually adding the cubic part $\square_m^{(I_n \setminus I_{n-m})}$ with coordinates

$p^{i_{n-m+1}}, \ldots, p^{i_n}$ – equation (6.67) turns into

$$\Delta_{n-m}^{(I_{n-m})} \times \square_{d-n+m}^{(I_d \setminus I_{n-m})} \longmapsto \Delta_{n-m-1}^{(I_{n-m-1})} \times \square_{d-n+m+1}^{(I_d \setminus I_{n-m-1})} \quad \text{for } d \geq n-m, \tag{6.71}$$

and by applying equation (6.69) to the previous image of the initial domain $\overline{\Delta_n^{(I_n)}}$ in equation (6.61), we obtain for the transformed total domain

$$\big(\overline{\Delta_{n-m-1}^{(I_{n-m-1})}} \times \overline{\square_{m+1}^{(I_n \setminus I_{n-m-1})}}\big) \setminus \bigcup_{j=n-m-1}^{n-1} \tilde{N}_j \tag{6.72}$$

with $\tilde{N}_{n-m}, \ldots, \tilde{N}_{n-1}$ being the images of the previous additional faces: The faces $N_{n-m+1}, \ldots, N_{n-1}$ are only affected indirectly as they contain the full $\overline{\Delta_{n-m}^{(I_{n-m})}}$ as a factor, and hence only the i_{n-m}-th coordinate is moved from the simplex to the cubic fraction, thus

$$\tilde{N}_j = \overline{\Delta_{n-m-1}^{(I_{n-m-1})}} \times \overline{\square_{j-n+m}^{(I_{j-1} \setminus I_{n-m-1})}} \times \{0\}^{(\{i_j\})} \times \overline{\boxtimes_{n-j}^{(I_n \setminus I_j)}} \tag{6.73}$$

for $j = n-m+1, \ldots, n-1$, whereas $N_{n-m} \equiv \tilde{N}_{n-m}$ is virtually not affected as only $p^{i_{n-m}} = 0$ is transformed into $\tilde{p}^{i_{n-m}} = 0$. For the 'new' additional $(n-1)$-dimensional face $\tilde{N}_{n-m-1}$ (resulting from N_{n-m-1}), we may – having added the remaining dimensions – relax the condition $\tilde{p}^{i_{n-m}} > 0$ in equation (6.70), which ensures $N_{n-m-1} \neq \overline{\Delta_{n-m-2}^{(I_{n-m-2})}}$, into $\sum_{j=n-m}^{n} \tilde{p}^{i_j} > 0$ and hence obtain

$$\tilde{N}_{n-m-1} := \overline{\Delta_{n-m-2}^{(I_{n-m-2})}} \times \{0\}^{(\{i_{n-m-1}\})} \times \overline{\boxtimes_{m+1}^{(I_n \setminus I_{n-m-1})}}. \tag{6.74}$$

Simultaneously, $\bar{u}_{I_k}^{i_k,\ldots,i_{n-m}}$ and $\tilde{u}_{I_k}^{i_k,\ldots,i_{n-m};i_{n-m+1},\ldots,i_d}$, $d = n-m+1, \ldots, n$ get transformed into

$$\tilde{u}_{I_k}^{i_k,\ldots,i_{n-m-1};i_{n-m},\ldots,i_d}(\tilde{p}, t) = \bar{u}_{I_k}^{i_k,\ldots,i_{n-m-1}}(\tilde{\pi}^{i_{n-m-1}}(\tilde{p}), t) \prod_{j=n-m}^{d} (1 - \tilde{p}^{i_j}) \tag{6.75}$$

in $\Delta_{n-m-1}^{(I_{n-m-1})} \times \square_{d-n+m+1}^{(I_d \setminus I_{n-m-1})}$ for $d \geq n-m$, and hence

$$\tilde{U}^{i_k,\dots,i_{n-m-1};i_{n-m},\dots,i_n}_{I_k}(p,t) := \sum_{k\leq d\leq n-m-1} \bar{u}^{i_k,\dots,i_d}_{I_k}(p,t)\chi_{\Delta^{(I_d)}_d}(p)$$
$$+ \sum_{n-m\leq d\leq n} \tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m},\dots,i_d}_{I_k}(p,t)\chi_{\Delta^{(I_{n-m-1})}_{n-m-1}\times\square^{(I_d\setminus I_{n-m-1})}_{d-n+m+1}}(p). \quad (6.76)$$

The transformed functions $\tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m},\dots,i_d}_{I_k}$ then each smoothly extend to $\Delta^{(I_{n-m-1})}_{n-m-1}\times\overline{\square^{(I_d\setminus I_{n-m-1})}_{d-n+m+1}}$, and we have

$$\tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m},\dots,i_d}_{I_k}\big|_{\Delta^{(I_{n-m-1})}_{n-m-1}\times\square^{(I_{d-1}\setminus I_{n-m-1})}_{m+1}} = \tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m},\dots,i_{d-1}}_{I_k} \quad (6.77)$$

for $d = n-m+1,\dots,n$ and

$$\tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m}}_{I_k}\big|_{\Delta^{(I_{n-m-1})}_{n-m-1}} = \bar{u}^{i_k,\dots,i_{n-m-1}}_{I_k}. \quad (6.78)$$

With $\bar{u}^{i_k,\dots,i_{n-m-1}}_{I_k}$ being smoothly extendable to $\overline{\partial_{n-m-2}\Delta^{(I_{n-m-1})}_{n-m-1}}$, the functions $\tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m},\dots,i_d}_{I_k}$ also smoothly extend to $\left(\partial_{n-m-2}\Delta^{(I_{n-m-1})}_{n-m-1}\right)\times\overline{\square^{(I_d\setminus I_{n-m-1})}_{d-n+m+1}}$, by which all additional faces are covered; in particular, $\tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m}}_{I_k}$ smoothly extends to N_{n-m-1} resp. eventually $\tilde{u}^{i_k,\dots,i_{n-m-1};i_{n-m},\dots,i_n}_{I_k}$ extends to $\tilde{N}_{n-m-1}$ (however, $\bar{u}^{i_k,\dots,i_{n-m-1}}_{I_k}$ resp. its continuous extension to $\partial_{n-m-2}\Delta^{(I_{n-m-1})}_{n-m-1}$ still has an incompatibility at $\overline{\Delta^{(I_{n-m-3})}_{n-m-3}}$).

To analyse the transformation behaviour of the operator, we first note that the requirements of lemma 6.4 on a^{ij} are met as for $i,j\in\{i_1,\dots,i_{n-m}\}$ we have $a^{ij}(p) = p^i(\delta^i_j - p^j)$ by equation (6.66), while all other non-diagonal coefficients vanish. Hence, by the lemma, we have for $i,j\in\{i_1,\dots,i_{n-m}\}$

$$\tilde{a}^{ij}(\tilde{p}) = \tilde{p}^i(\delta^i_j - \tilde{p}^j), \quad (6.79)$$

while for $\tilde{a}^{i_ji_j}$ with $j = n-m+1,\dots,n$ we obtain

$$\tilde{a}^{i_ji_j}(\tilde{p}) = a^{i_ji_j}(p) = \frac{p^{i_j}(1-p^{i_j})}{\prod_{l=n-m}^{j-1} p^{i_l}} = \frac{\tilde{p}^{i_j}(1-\tilde{p}^{i_j})}{\prod_{l=n-m-1}^{j-1} \tilde{p}^{i_l}}. \quad (6.80)$$

Likewise, $\tilde{a}^{i_{n-m}i_{n-m}}$ takes the form

$$\tilde{a}^{i_{n-m}i_{n-m}}(\tilde{p}) = \frac{\tilde{p}^{i_{n-m}}(1-\tilde{p}^{i_{n-m}})}{\tilde{p}^{i_{n-m-1}}}, \tag{6.81}$$

whereas all other coefficients vanish. Altogether, this yields

$$\tilde{L}^* = \frac{1}{2}\sum_{j,l=1}^{n-m-1} \tilde{p}^{i_j}(\delta^j_l - \tilde{p}^{i_l})\frac{\partial}{\partial\tilde{p}^{i_j}}\frac{\partial}{\partial\tilde{p}^{i_l}} + \frac{1}{2}\sum_{j=n-m}^{n} \frac{\tilde{p}^{i_j}(1-\tilde{p}^{i_j})}{\prod_{l=n-m-1}^{j-1}\tilde{p}^{i_l}}\frac{\partial^2}{(\partial\tilde{p}^{i_j})^2} \tag{6.82}$$

on $\left(\overline{\Delta^{(I_{n-m-1})}_{n-m-1}} \times \overline{\square^{(I_n\setminus I_{n-m-1})}_{m+1}}\right) \setminus \bigcup_{j=n-m-1}^{n-1} N_j$. If $\tilde{p}^{i_{n-m}}$ is chosen with alternative orientation (cf. remark 6.3), then $\tilde{p}^{i_{n-m}}$ needs to be replaced by $(1-\tilde{p}^{i_{n-m}})$ everywhere.

Thus, after the $(m+1)$-th blow-up step, domain, solution and operator are of analogous form as before, just with the index m replaced by $m+1$. Eventually, after $n-k-1$ blow-up steps domain, solution and operator have attained the asserted form of the corresponding statements. In particular, the remaining u_{I_k} as a proper solution smoothly extends to the entire boundary of $\Delta^{(I_k)}_k$, and hence so does $\overline{\bar{u}^{i_k,i_{k+1}}_{I_k}}$ in $\Delta^{(I_k)}_{k+1}$, implying that each $\tilde{u}^{i_k,i_{k+1};i_{k+2},\dots,i_d}_{I_k}$ smoothly extends to $\overline{\Delta^{(I_{k+1})}_{k+1}} \times \overline{\square^{(I_d\setminus I_{k+1})}_{d-k-1}}$, and eventually $\tilde{U}^{i_k,i_{k+1};i_{k+2},\dots,i_n}_{I_k}$ smoothly extends to $\overline{\Delta^{(I_{k+1})}_{k+1}} \times \overline{\square^{(I_n\setminus I_{k+1})}_{n-k-1}}$. Moreover, the restriction property in equations (6.77) and (6.78) yields equation (6.45). $\square$

Proof of corollary 6.7. In the given setting, we have $\bar{u}^{i_0,i_1}_{\{i_0\}}(\tilde{p}) = u_{\{i_0\}}(\tilde{p}^{i_0}+\tilde{p}^{i_1})\frac{\tilde{p}^{i_0}}{\tilde{p}^{i_0}+\tilde{p}^{i_1}} = u_{\{i_0\}}(1)(1-\tilde{p}^{i_1})$ in $\overline{\Delta^{(\{i_0,i_1\})}_1} = \overline{\square^{(\{i_1\})}_1}$ (and $\overline{\Delta^{(\{i_0\})}_0} = \{0\}^{(\{i_0\})}$), which proves the asserted form of the (simplified) solution, the domain and the additional faces. $\square$

However, the global smoothness of the transformed solution of proposition 5.25 observed in the preceding corollary does not necessarily hold for other functions in question, i. e. arbitrary iteratively extended solutions U in accordance with the extension constraints 5.20 (this corresponds to U particularly being of class $C^\infty_{p_0}$, cf. also remark 5.21). However, we still have a weaker global regularity assertion for the transformed function $\tilde{U}$ on the entire image of the simplex (only formulated for the stationary component corresponding to the setting of corollary 6.7):

6.8 Lemma. *Let $n \geq 2$, $I_d := \{i_0, i_1, \dots, i_d\} \subset \{0, 1, \dots, n\}$ for $d = 0, \dots, n$ with $i_i \neq i_j$ for $i \neq j$ and $u_{\{i_0\}}\colon \Delta^{(\{i_0\})}_0 \longrightarrow \mathbb{R}$. Then an iterated extension*

$U = \sum_{d=0}^{n} u_d \in C^\infty_{p_0}\big(\bigcup_{d=0}^{n} \Delta_d^{(I_d)}\big)$ of $u_{\{i_0\}}$ in accordance with the extension constraints 5.20 is transformed by a successive blow-up transformation $\Phi^{r_{n-1}}_{s_{n-1}} \circ \ldots \circ \Phi^{r_1}_{s_1}$ as in proposition 6.5 into a function $\tilde{U} = \sum_{d=0}^{n} \tilde{u}_d \colon \bigcup_{d=0}^{n} \Box_d^{(I'_d)} \longrightarrow \mathbb{R}$ with extension to all faces $\big\{\tilde{p}^{i_1} = 1\big\}, \ldots, \big\{\tilde{p}^{i_n} = 1\big\}$ (perceivable as boundary instance of any $\Box_d^{(I'_d)} \subset \overline{\Box_n^{(I'_n)}}$) which is of class C^∞_p and vanishes on the mentioned faces.

Proof. By lemma 6.1 and proposition 6.5 resp. corollary 6.7, the full blow-up transformation respectively maps

$$\bigcup_{d=0}^{n} \Delta_d^{(I_d)} \longmapsto \bigcup_{d=0}^{n} \Box_d^{(I'_d)} \tag{6.83}$$

C^∞-diffeomorphically (cf. equation (6.47)). By the $C^\infty_{p_0}$-regularity of U, u_n in $\Delta_n^{(I_n)}$ smoothly connects with u_{n-1} in $\Delta_{n-1}^{(I_{n-1})}$, and consequently so does $\tilde{u}_n$ in $\Box_n^{(I'_n)}$ with $\tilde{u}_{n-1}$ in $\Box_{n-1}^{(I'_{n-1})}$; an analogous statement holds for all lower dimensions. Thus it remains to be shown that $\tilde{U}$ extends those faces of $\overline{\Box_n^{(I'_n)}}$ given by $\{\tilde{p}^{i_j} = 1\}$ for $j = 1, \ldots, n$ such that the extension is of class C^∞_p.

In anticipation of lemma 6.18 on p. 210, the interior of $\{\tilde{p}^{i_j} = 1\} \subset \overline{\Box_n^{(I'_n)}}$ corresponds to $p^{i_{j-1}} = 0$ and $p^{i_l} > 0$ for $l \neq j-1$ in $\overline{\Delta_n^{(I_n)}}$, thus to $\Delta_{n-1}^{(I_n \setminus \{i_{j-1}\})}$, which is a boundary face of $\Delta_n^{(I_n)}$ outside the assumed extension path defined by the (ordered) I_n. Hence by the $C^\infty_{p_0}$-regularity, the relevant continuous extension of U needs to be zero there, and this is attained smoothly when coming from the interior $\Delta_n^{(I_n)}$. Considering the diffeomorphism properties of the transformation, this also applies to the cube.

An analogous observation holds for subcubes $\Box_{d-1}^{(I'_{d-1})} \subset \overline{\Box}_n$, $d = 1, \ldots, n$: The interior of its face $\{\tilde{p}^{i_j} = 1\}$ corresponds to $\Delta_{d'-1}^{(I_{d-1} \setminus \{i_{j-1}\})} \subset \overline{\Delta_{d-1}^{(I_{d-1})}}$ when transformed back to the simplex (cf. equation (6.83) and lemma 6.18). This is again outside the assumed extension path, in particular if starting in $\Delta_{d-1}^{(I_{d-1})}$, and hence the corresponding boundary extension of u_{d-1} needs to smoothly attain zero there by the $C^\infty_{p_0}$-regularity, which likewise applies analogously to the cube. □

6.4 The uniqueness of solutions of the stationary Kolmogorov backward equation

We will now start the discussion of the main application of the blow-up scheme, which is the uniqueness of iteratively extended solutions of the Kolmogorov backward equation in accordance with the extension constraints 5.20. However, as already mentioned, in the presented work, this is limited to the stationary components. First, we will discuss the uniqueness of solutions of the (correspondingly transformed) stationary Kolmogorov backward equation on the cube, which is basically analogous to our considerations for the simplex in section 5.4.7. After that, we will state the main result of this chapter by applying the uniqueness result for the cube to the transformed iteratively extended solutions (assuming sufficient regularity if necessary), thus substantially broadening the previous uniqueness result of proposition 5.31.

Regarding the uniqueness of stationary solutions on the cube with the transformed Kolmogorov backward operator given by equation (6.51), we have in conjunction to lemma 5.30 for the simplex:

6.9 Lemma (stem lemma, cube version)**.** *For a solution $u \in C^\infty(\square_n)$ of the stationary Kolmogorov backward equation $\tilde{L}_n^* u = 0$ in $\square_n$ with*

$$\tilde{L}_n^* := \frac{1}{2} \sum_{i=1}^{n} \frac{\tilde{p}^i(1-\tilde{p}^i)}{\prod_{j=1}^{i-1} \tilde{p}^j} \frac{\partial^2}{(\partial \tilde{p}^i)^2} \tag{6.84}$$

and with extension $\bar{u} \in C_p^\infty(\overline{\square}_n)$, we have

$$\tilde{L}^* \bar{u} = 0 \quad \text{in } \overline{\square}_n, \tag{6.85}$$

i. e.

$$\tilde{L}_d^* \bar{u} = 0 \quad \text{with} \quad \tilde{L}_d^* := \frac{1}{2} \sum_{\substack{i=\hat{\imath}(d)+1 \\ i \neq i_m}}^{n} \frac{\tilde{p}^i(1-\tilde{p}^i)}{\prod\limits_{\substack{j=\hat{\imath}(d)+1 \\ j \neq i_m}}^{i-1} \tilde{p}^j} \frac{\partial^2}{(\partial \tilde{p}^i)^2} \tag{6.86}$$

in $\square_d = \left\{\tilde{p}^{i_1} = b_{i_1}, \ldots, \tilde{p}^{i_{n-d}} = b_{i_{n-d}}\right\} \subset \partial_d \square_n$ for all $1 \leq d \leq n-1$ and all

$i_1, \dots, i_{n-d} \in \{1, \dots, n\}$, $i_k \neq i_l$ for $k \neq l$ with $\hat{\imath} = \hat{\imath}(d) := \underset{i_1,\dots,i_{n-d}}{\arg\max}\{b_{i_m} = 0\}$ resp. $\hat{\imath}(d) := 0$ if $b_{i_m} = 1$ for all i_m.

Proof. The statement is proven iteratively: Assuming that equation (6.86) holds in some (arbitrary) domain $\Box_{d+1} \subset \partial_{d+1}\Box_n$, we show that a corresponding formula also holds for any $\Box_d \subset \partial_d\Box_{d+1} \subset \partial_d\Box_n$. A repeated application of the argument then yields the assertion.

Let $\Box_{d+1} = \left\{\tilde{p}^{i_1} = b_1, \dots, \tilde{p}^{i_{n-d-1}} = b_{n-d-1}\right\}$ and $\Box_d = \left\{\tilde{p}^{i_1} = b_1, \dots, \tilde{p}^{i_{n-d}} = b_{n-d}\right\}$ with $i_{n-d} \neq i_1, \dots, i_{n-d-1}$ and $b_{n-d} \in \{0, 1\}$. If we have $i_{n-d} < \hat{\imath}(d+1)$, then as $\tilde{p}^{i_{n-d}} \to 0$ resp. $\tilde{p}^{i_{n-d}} \to 1$, the value of the operator in equation (6.86) applied to $\bar{u}$ – with the occurring derivatives and the coefficients being continuous – depends continuously on $\tilde{p}$ up to the boundary, thus equation (6.86), which already has the corresponding form for $\Box_d$ (note $\hat{\imath}(d) \equiv \hat{\imath}(d+1)$), also holds on $\Box_d$.

If we rather have $i_{n-d} > \hat{\imath}(d+1)$ and $b_{n-d} = 1$, then, when choosing some $\tilde{p} \in \Box_d$ and a sequence $(\tilde{p}_l)_{l\in\mathbb{N}}$ in $\Box_{d+1}$ with $\tilde{p}_l \to \tilde{p}$, the expression

$$\frac{1}{2}\frac{\tilde{p}_l^{i_{n-d}}(1 - \tilde{p}_l^{i_{n-d}})}{\prod_{\substack{j=\hat{\imath}(d)+1 \\ j\neq i_m}}^{i_{n-d}-1} \tilde{p}_l^j} \frac{\partial^2}{(\partial \tilde{p}_l^{i_{n-d}})^2}\bar{u}(\tilde{p}_l) \tag{6.87}$$

is controlled by $(1 - \tilde{p}_l^{i_{n-d}})$ while approaching $\tilde{p}$ and – with the derivatives of $\bar{u}$ being bounded on a closed neighbourhood of $\tilde{p}$ by reason of the regularity of $\bar{u}$ – is continuous up to $\tilde{p}$. Analogous to the previous case, all other summands of the operator in equation (6.86) are also continuous on the boundary, thus proving that the corresponding form of equation (6.86) (with the i_{n-d}-th summand deleted) holds in $\Box_d$ (again $\hat{\imath}(d) \equiv \hat{\imath}(d+1)$).

If instead $i_{n-d} > \hat{\imath}(d+1)$ and $b_{n-d} = 0$, then we may multiply the whole equation (6.86) by $\tilde{p}^{i_{n-d}}$. If now $\tilde{p}^{i_{n-d}} \to 0$, then by a similar argument as above all derivatives of the operator that do not contain $\tilde{p}^{i_{n-d}}$ in the denominator of their coefficient continuously vanish, whereas the values of all other summands are also continuous up to the boundary. Thus, equation (6.86) holds on $\Box_d$ with the index $\hat{\imath}(d+1)$ replaced by $\hat{\imath}(d) = i_{n-d}$. □

The obtained equation (6.85) may again be perceived as an extended version of the stationary Kolmogorov backward equation on the cube (cf. also equation (5.148), although the domains do not fully correspond), and we have (cf. proposition 5.31):

6.10 Proposition. *A solution $\bar{u} \in C_p^\infty(\overline{\square}_n) \cap C^0\big(\overline{\square}_n\big)$ of the extended stationary Kolmogorov backward equation*

$$\tilde{L}^* \bar{u} = 0 \quad \text{in } \overline{\square}_n \tag{6.88}$$

with $\tilde{L}^*$ as in equation (6.86) *is uniquely determined by its values on $\partial_0 \square_n$.*

Proof. The uniqueness may be shown by a successive application of the maximum principle: In every instance of the domain $\square_d \subset \partial_d \square_n$ for all $1 \leq d \leq n$, the solution $\bar{u}|_{\square_d}$ is uniquely defined by its values on $\partial \square_d$: If equation (6.86) comprises d derivative terms, this follows directly from Hopf's maximum principle as the operator is locally uniformly elliptic on $\square_d$; if it only comprises $d' < d$ derivative terms, analogous considerations apply for each d'-dimensional fibre of $\square_d$ (with corresponding boundary part), thus giving the uniqueness of a solution on every fibre first and after assembling also on all $\square_d$. Applying this consideration successively for $\partial_0 \square_n, \ldots, \partial_n \square_n = \square_n$ yields the desired global uniqueness. $\square$

With the blow-up scheme and its regularising effect on the solution at hand, we may now show that the preceding uniqueness result may also be conveyed to the simplex $\overline{\Delta}_n$, assuming some additional regularity. As already indicated, it then only remains to be shown that there is sufficient and unique boundary data for any extension, and we eventually have:

6.11 Theorem. *Let $n \in \mathbb{N}_+$, $I_d := \{i_0, i_1, \ldots, i_d\} \subset \{0, 1, \ldots, n\}$ for $d = 0, \ldots, n$ with $i_i \neq i_j$ for $i \neq j$ and $u_{\{i_0\}} \colon \Delta_0^{(\{i_0\})} \longrightarrow \mathbb{R}$ be given. Then an extension $\bar{U}_{\{i_0\}}^{i_0,\ldots,i_n} \colon \bigcup_{0 \leq d \leq n} \Delta_d^{(I_d)} \longrightarrow \mathbb{R}$ as in proposition 5.25 is unique within the class of extensions U which satisfy the extension constraints 5.20, i. e.*

(i) are of class $C_{p_0}^\infty\big(\bigcup_{0 \leq d \leq n} \Delta_d^{(I_d)}\big)$ with $U|_{\Delta_0^{(\{i_0\})}} = u_{\{i_0\}}$ and

(ii) solve the stationary Kolmogorov backward equation (5.148) *in $\bigcup_{0 \leq d \leq n} \Delta_d^{(I_d)}$,*

as well as, in case $n \geq 2$, whose

(iii) transformation image $\tilde{U}\colon \bigcup_{d=0}^{n} \square_d^{(I'_d)} \longrightarrow \mathbb{R}$ *by a successive blow-up transformation* $\Phi_{s_{n-1}}^{r_{n-1}} \circ \ldots \circ \Phi_{s_1}^{r_1}$ *as in proposition 6.5 has an extension to the entire boundary* $\partial \square_n^{(I'_n)}$ *which is of class* $C_p^\infty\big(\overline{\square_n^{(I'_n)}}\big) \cap C^0\big(\overline{\square_n^{(I'_n)}}\big)$.

Consequently, also the global extension $\bar{U}_{\{i_0\}}$ *as in proposition* 5.28 *resp. also in theorem 5.29 is unique.*

Proof. The assertion for the trivial case $n = 1$ directly follows, as $\bar{U}_{\{i_0\}}^{i_0,i_1}$ is as already sufficiently regular in $\overline{\Delta_1^{(I_1)}} \equiv \overline{\square_1^{(I'_1)}}$ for an application of the maximum principle, in particular globally continuous. For $n \geq 2$, any function U which is a solution of the stationary Kolmogorov backward equation (5.148) in $\overline{\Delta_n^{(I_n)}}$ by a full blow-up transformation of the domain transforms into a function $\tilde{U}$, which solves the stationary Kolmogorov backward equation (6.41) in $\bigcup_{d=0}^{n} \square_d^{(I'_d)}$ (cf. proposition 6.5 resp. corollary 6.7 and lemma 6.8). Furthermore, with the assumed regularity after a full blow-up, it has an extension to $\overline{\square_n^{(I'_n)}}$ which is pathwise smooth as well as globally continuous and by lemma 6.9 solves the stationary Kolmogorov backward equation $\tilde{L}^* \bar{\tilde{U}} = 0$ in $\overline{\square_n^{(I'_n)}}$ with $\tilde{L}^*$ as in equation (6.86). Hence, the uniqueness result of proposition 6.10 applies and proves the uniqueness of the transformed function (and, regarding the injectivity of the blow-up, also the uniqueness of U) – for specified boundary data on the entire $\partial_0 \square_n^{(I'_n)}$. Thus, we only need to show that this boundary data is uniquely determined by the assumptions made.

This is straightforward: In accordance with lemma 6.8, $\tilde{U}$ resp. its corresponding continuous extension vanishes on $\{\tilde{p}^{i_j} = 1\} \subset \partial \square_n^{(I'_n)}$, $j = 1, \ldots, n$. As by assumption (iii) the continuous extendability applies to the entire $\overline{\square_n^{(I'_n)}}$, $\tilde{U}$ resp. its extension even vanishes on

$$\overline{\{\tilde{p}^{i_1} = 1\}}, \ldots, \overline{\{\tilde{p}^{i_n} = 1\}}. \tag{6.89}$$

In particular, this signifies that $\tilde{U}$ resp. its extension vanishes on any vertex $\square_0 \subset \partial_0 \square_n^{(I'_n)}$ – which may always be written as

$$\square_0 = \big\{\tilde{p}^{i_j} = b_j \quad \text{for } j = 1, \ldots, n\big\} \quad \text{with correspondingly } b_j \in \{0,1\} \text{ –} \tag{6.90}$$

except for the vertex $\square_0^{(\varnothing)} = \{(0, \ldots, 0)\}$, where it attains the value $u_{\{i_0\}}$ as stated previously. Thus, the (transformed) boundary data given on all vertices is identical for any extension in question, and since $\bar{U}^{i_0,\ldots,i_n}_{\{i_0\}} \colon \bigcup_{0\le d\le n} \Delta_d^{(I_d)} \longrightarrow \mathbb{R}$ as in proposition 5.25 satisfies the extension constraints and has an extension to the entire boundary $\partial\square_n^{(I'_n)}$ which is in $C_p^\infty(\overline{\square_n^{(I'_n)}}) \cap C^0\big(\overline{\square_n^{(I'_n)}}\big)$ (this may be seen directly from equation (6.46)), it also is the unique extension. $\square$

6.5 Some combinatorial and model related aspects of the blow-up scheme

Having stated the main result, we now wish to illuminate some combinatorial aspects of the presented blow-up scheme as well as give a justification of the regularity assumption in theorem 6.11 (iii) in terms of the model.

As, with both the simplex and the cube, a stationary solution on the entire domain which is pathwise smooth and globally continuous is uniquely determined by its values on the vertices, the increased number of dimensions of the linear space of solutions from $n+1$ in proposition 5.31 to $(n+1)!$ in proposition 5.25 resp. corollary 5.26 should be linked to the increase in the number of vertices that $\square_n$ encompasses (i. e. 2^n) in comparison with that of Δ_n (i. e. $n+1$). In accordance with proposition 6.10, one would at first suspect that the 2^n vertices of the domain $\square_n$ imply a 2^n-dimensional solution space (which is different from $(n+1)!$). However, it turns out that the increase in the number of different solutions is rather due to the numerous ways to catenate cube and simplex part as we have (the result holds for arbitrary $0 \le k \le n-2$):

6.12 Lemma. *There are* $\frac{1}{2}\frac{(n+1)!}{(k+1)!}$ *possibilities to map* $\overline{\Delta_n^{(I_n)}}$ *onto* $\overline{\Delta_{k+1}^{(I_{k+1})}} \times \overline{\square_{n-k-1}^{(I_n\setminus I_{k+1})}}$ *(irrespective of any additional faces) by an iterative blow-up scheme as presented in proposition 6.5.*

Proof. The given number results from the various choices of coordinates that are made during the application of such a blow-up scheme. As is obvious from the proof of proposition 6.5, in the first blow-up step two out of $n+1$ coordinates in I_n are chosen, i. e. $p^{i_{n-1}}$ and p^{i_n}, in no specific order as they determine the current $(n-2)$-dimensional blow-up domain (swapping the coordinates only swaps the corresponding indices of the $\tilde{p}$-coordinates as well as inverts the orientation of the $\tilde{p}^{i_n}$-coordinate orientation; cf. also the footnote on p. 189). In all following blow-up steps, only one further coordinate out of the $n-d+1$ remaining coordinates is chosen since by the applied scheme the corresponding blow-up domain needs to be a subsimplex of subsequent lower dimension of the previous blow-up domain, which also defines the other blow-up index. With $n-k-1$ blow-up steps altogether, the total number of the choices made thus equals

$$\frac{1}{2}\cdot(n+1)n\cdot(n-2)\cdots(k+2)=\frac{1}{2}\frac{(n+1)!}{(k+1)!}.\ \Box$$

Furthermore, each blow-up scheme actually works for two different iteratively extended solutions, which is due to the ambiguity in the first blow-up step: Thus, for a given extension path I_n, an iterated blow-up transformation as in proposition 6.6 has the same qualitative effect on the corresponding iteratively extended function as in proposition 5.25, even if in the extension of the function the indices i_{n-1} and i_n are exchanged; that just converts the last factor of the transformed extended function from $(1-\tilde{p}^{i_n})$ into $\tilde{p}^{i_n}$. Conversely, this is also reflected when asking which iteratively extended functions may be transformed back into the desired form by a given blow-up scheme (formulated only for $k=0$):

6.13 Lemma. *In the situation of proposition 6.6 resp. corollary 6.7, out of the class of functions $\tilde{u}\in C_p^\infty(\overline{\Box_n^{(I'_n)}})\cap C^0(\overline{\Box_n^{(I'_n)}})$ which solve the extended stationary Kolmogorov backward equation* (6.85) *in $\overline{\Box_n^{(I'_n)}}$ and vanish everywhere on $\partial_0\Box_n^{(I'_n)}$ except for a single vertex $\widehat{\Box}_0\subset\partial_0\Box_n^{(I'_n)}$, only those functions with $\widehat{\Box}_0\in\overline{N}_1\cap\ldots\cap\overline{N}_{n-1}$ are transformed into a function of type* (5.129) *when reversing the given blow-up transformation.*

Proof. Since being in $\partial_0\Box_n^{(I'_n)}$, we generically have $\widehat{\Box}_0\in\{0,1\}^n$; furthermore, let $\tilde{u}_0$ be the value of $\tilde{u}$ on $\widehat{\Box}_0$. At first we note that all solutions in question are of the form

$$\tilde{u}(\tilde{p})=\tilde{u}_0\prod_{j=1}^{n}\tilde{u}_{i_j}(\tilde{p}),\quad \tilde{p}\in\overline{\Box_n^{(I'_n)}}\tag{6.91}$$

with either $\tilde{u}_{i_j}(\tilde{p})=(1-\tilde{p}^{i_j})$ or $\tilde{u}_{i_j}(\tilde{p})=\tilde{p}^{i_j}$ for $j=1,\ldots,n$ as they all satisfy equation (6.85) and are of sufficient regularity – hence by proposition 6.10 are uniquely determined by their values on $\partial_0\Box_n^{(I'_n)}$. The requirement of vanishing on all vertices except for $\widehat{\Box}_0$ directly implies the depicted shape.

If $\widehat{\Box}_0 \notin \overline{N}_1 \cap \ldots \cap \overline{N}_{n-1}$, then consequently the i_l-th component of $\widehat{\Box}_0$ equals 1 for some $l \in \{1, \ldots, n-1\}$, and we correspondingly have $\tilde{u}_{i_l}(\tilde{p}) = \tilde{p}^{i_l}$. Transforming $\tilde{u}$ now reversely in particular signifies for $l \neq 1$ that, when reversing the $(n-l+1)$-th blow-up step (with $r_{n-l+1} = i_{l-1}$ and $s_{n-l+1} = i_l$), $\tilde{p}^{i_l}$ is replaced by $\frac{p^{i_l}}{p^{i_{l-1}}+p^{i_l}}$. When subsequently reversing the preceding (i. e. $(n-l)$-th) blow-up step, the numerator of $\frac{p^{i_l}}{p^{i_l}+p^{i_{l+1}}}$ is transformed into $p^{i_l} + p^{i_{l+1}}$ and hence cancels with the denominator of the transformation result of the factor $\tilde{u}_{i_{l+1}}$, being $\frac{p^{i_l}}{p^{i_l}+p^{i_{l+1}}}$ resp. $\frac{p^{i_{l+1}}}{p^{i_l}+p^{i_{l+1}}}$. If $l = 1$, reversing the last (i. e. $(n-1)$-th) blow-up step yields $\tilde{p}^{i_1} = p^{i_1} + p^{i_2}$ and $\tilde{p}^{i_2} = \frac{p^{i_2}}{p^{i_1}+p^{i_2}}$ resp. $1 - \tilde{p}^{i_2} = \frac{p^{i_1}}{p^{i_1}+p^{i_2}}$, which also cancel. Thus, at least one fraction is lost, and the fully reversely transformed solution $\bar{u}$ is not of the form as given in equation (5.129).

Solely, both choices of $\tilde{u}_{i_n}$ correspond to a function $\bar{u}$ of the form (5.129) as described in proposition 5.25 after a full reverse transformation: It is easy to see that this corresponds to endowing either of the two vertices in $\overline{N}_1 \cap \ldots \cap \overline{N}_{n-1}$ with a positive value. $\Box$

6.14 Remark. As may be seen from the yet following lemma 6.18, both possible choices of $\widehat{\Box}_0$ correspond to $\Delta_0^{(\{0\})}$ when transformed back to the simplex.

Thus, combining the results of lemma 6.12 and lemma 6.13 for $k = 0$, $\frac{(n+1)!}{2}$ different blow-up schemes and a two-dimensional space of transformable iteratively extended solutions each yield altogether an $(n+1)!$-dimensional space of functions of the desired type (5.129).

Another interesting aspect is to explain the increase in the number of vertices from $\overline{\Delta}_n$ to $\overline{\Box}_n$ when applying the blow-up scheme in terms of the incompatibilities which need to be resolved. As already stated, $\bar{U}_{\{i_0\}}^{i_0,\ldots,i_n}$ features an $(n-1)$-fold nested incompatibility, i. e. an incompatibility of the boundary extensions of the corresponding component of $\bar{U}_{\{i_0\}}^{i_0,\ldots,i_n}$ at each different level. To resolve this at the top-dimensional component, we require one extra copy of the face Δ_{n-2} as this is the intersection of the two conflictual faces, which signifies an increase in vertices by $n-1$. After $d-1$ blow-up steps, the corresponding face is of dimension $n-d-1$

and is – within the simplex part of $\Delta_{n-d+1} \times \square_{d-1}$ – again the intersection of two conflictual faces, thus also one extra copy is required. The hereby implied increase by $n-d$ vertices, however, is reflected in the $d-1$ dimensions of the cube fraction $\square_{d-1}$, yielding an additional factor 2^{d-1}. Eventually summing over all new vertices generated in the $n-1$ blow-up steps and the previously existing $n+1$ vertices, we obtain

$$\sum_{d=1}^{n-1} 2^{d-1}(n-d) + (n+1) = 2^n, \tag{6.92}$$

which is the total number of vertices in $\square_n$ (if the reflection in the cube part is ignored, then only $\sum_{d=1}^{n-1}(n-d) = \frac{n(n-1)}{2}$ new vertices would be generated).

Justification of the regularity assumption in terms of the model

Finally, we will illustrate some model related arguments in support of the *regularity assumption* as it appears in theorem 6.11 (iii) for $n \geq 2$, thus argue that the existence of a pathwise smooth extension to the entire closed cube for a transformed iteratively extended stationary solution of the Kolmogorov backward equation in accordance with the extension constraints 5.20 – which by lemma 6.8 is already pathwise smooth in the transformation image of the simplex – may also be based on some reasonable assumptions on the behaviour of the underlying Wright–Fisher model. As such an extension particularly requires that all incompatibilities are resolved (in contrast to the solution on the simplex), we may furthermore stipulate that atypical solutions whose extension still lacks global continuity in the entire closure of the cube are also excluded in terms of the underlying model, hence in total matching the condition 6.11 (iii).

Analysis of a simple example

To develop the general principle of the desired plausibilisation of the C_p^∞-regularity property, we start with some observations on the behaviour of an iteratively extended solution as in proposition 5.25 for $n = 2$ in the context of the presented blow-up. We then have e. g.

$$\bar{U}^{0,1,2}_{\{0\}} = u_{\{0\}}\chi_{\Delta_0^{(\{0\})}} + \bar{u}^{0,1}_{\{0\}}\chi_{\Delta_1^{(\{0,1\})}} + \bar{u}^{0,1,2}_{\{0\}}\chi_{\Delta_2^{(\{0,1,2\})}} \quad \text{in} \bigcup_{d=0}^{2} \Delta_d^{(\{0,\dots,d\})} \tag{6.93}$$

(for a detailed construction of such a function cf. also p. 160), and of course only the top-dimensional component $\bar{u}^{0,1,2}_{\{0\}}(p) = p^0 \cdot \frac{p^1}{1-p^0}$ in $\Delta_2^{(\{0,1,2\})}$ (with $p^0 = 1 - p_1 - p_2$) resp. its continuous extension yields incompatibilities. Hence, applying a blow-up transformation with $r = 1$, $s = 2$ transforms it via $\tilde{p}^1 := p^1 + p^2$ and $\tilde{p}^2 := \frac{p^2}{p^1+p^2}$ into $\tilde{u}^{0,1;2}_{\{0\}}(\tilde{p}) = (1 - \tilde{p}^1)(1 - \tilde{p}^2)$, which is found to be smoothly extendable to $\overline{\square_2^{(\{1,2\})}}$ (as then also $\tilde{U}^{0,1;2}_{\{0\}} = u_{\{0\}}\chi_{\square_0} + \bar{u}^{0,1}_{\{0\}}\chi_{\square_1^{(\{1\})}} + \tilde{u}^{0,1;2}_{\{0\}}\chi_{\square_2^{(\{1,2\})}}$).

The observed smooth extendability of $\tilde{u}^{0,1;2}_{\{0\}}$ in particular to the additional 1-dimensional face $N_1 := \{0\}^{(\{1\})} \times \overline{\boxtimes_1^{(\{2\})}}$ of $\overline{\square_2^{(\{1,2\})}}$ then implies that, for any sequence $(\tilde{p}_l)_{l\geq 0}$ in $\square_2^{(\{1,2\})}$ with $\lim_{l\to\infty} \tilde{p}_l := \tilde{p} \in \overline{N}_1$, i. e. $\tilde{p}^1 = 0$ and $\tilde{p}^2 = \alpha$ for some $\alpha \in [0, 1]$, the limit

$$\lim_{l\to\infty} \tilde{u}^{0,1;2}_{\{0\}}(\tilde{p}_l) \tag{6.94}$$

exists and depends smoothly on α itself, in particular up to the boundary of N_1, equalling $\{(0, 0), (0, 1)\}$. There, the limit coincides with the limit of (the continuous extension of) $\tilde{u}^{0,1;2}_{\{0\}}|_{\square_1^{(\{1\})}} \equiv \bar{u}^{0,1}_{\{0\}}$ on the face $\overline{\square_1^{(\{1\})}} \times \{0\}^{(\{2\})}$ resp. $\tilde{u}^{0,1;2}_{\{0\}} \equiv 0$ on $\overline{\square_1^{(\{1\})}} \times \{1\}^{(\{2\})}$.

Translating this back to $\Delta_2^{(\{0,1,2\})}$, we have that, for all sequences $(p_l)_{l\geq 1}$ in $\Delta_2^{(\{0,1,2\})}$ with $\lim_{l\to\infty} p^1_l = 0$, $\lim_{l\to\infty} p^2_l = 0$ and $\lim_{l\to\infty} \frac{p^2_l}{p^1_l+p^2_l} = \alpha$ for some $\alpha \in [0, 1]$, the limit

$$\lim_{l\to\infty} \bar{u}^{0,1,2}_{\{0\}}(p_l) \tag{6.95}$$

also exists and likewise depends smoothly on α. This signifies that, when approaching the incompatibility of (the continuous extension of) $\bar{u}^{0,1,2}_{\{0\}}$ at $\Delta_0^{(\{0\})} = \{0\}$ – with a certain limit ratio of p^1 and p^2 given –, we still obtain a limit for $\bar{u}(p)$, which itself depends smoothly on the chosen limit ratio and coincides with the corresponding limit of $\bar{u}^{0,1,2}_{\{0\}}|_{\Delta_1^{(\{0,1\})}} \equiv \bar{u}^{0,1}_{\{0\}}$ on the face $\Delta_1^{(\{0,1\})}$ if $\alpha \equiv 0$ resp. $\bar{u}^{0,1,2}_{\{0\}} \equiv 0$ on $\Delta_1^{(\{0,2\})}$ if

$\alpha \equiv 1$.

General considerations

While having observed the properties of a specific iteratively extended solution as by proposition 5.25 in the preceding example, one may generally argue that we may expect analogous properties for any function which in the setting of the Wright–Fisher model gives the probability for an eventual loss of alleles by

$$\{0,1,2\} \longrightarrow \{0,1\} \longrightarrow \{0\} \tag{6.96}$$

(if present) when starting in some $p_l \in \overline{\Delta_2^{(\{0,1,2\})}}$ (as does $\bar{U}_{\{0\}}^{0,1,2}$ resp. its top-dimensional component $\bar{u}_{\{0\}}^{0,1,2}$). Then the considered situation of p_l approaching the vertex 0 via $p_l^0 \to 1$ for $l \to \infty$ may be interpreted as allele 0 increasingly dominating the initial population – up to the limit case, where allele 0 is the only remaining allele. For this limit, one would expect the following behaviour for the model:

(i) As the path probability depends crucially on the ratio of the frequency of allele 2 to the joint frequency of the alleles 1 and 2 (and the ratio of the joint frequency of the alleles 1 and 2 to the joint frequency of all alleles), its value should stay regular when approaching the incompatibility, i. e. have a *finite limit*, if the limits for the mentioned ratios exist (the latter equalling zero).

(ii) This limit of the path probability should always be less or equal than the original path probability for only the allele 0 being present and should depend *smoothly* on the limit ratios. Furthermore, for boundary values of the limit ratios, it should be smoothly *compatible* with the path probability for starting with a configuration of only 2 alleles (which in the given example is $\bar{u}_{\{0\}}^{0,1}$ in $\Delta_1^{(\{0,1\})}$ resp. 0 in $\Delta_1^{(\{0,2\})}$ by definition).

These assumptions on the model directly translate into regularity properties of the corresponding solution of the Kolmogorov backward equation in $\Delta_2^{(\{0,1,2\})}$, describing the path probabilities (as is also observed with $\bar{u}_{\{0\}}^{0,1,2}$): (i) yields the existence of all limits for the vertex $\{0\}$, which are by (ii) bounded and do smoothly depend on the chosen limit approach direction towards the vertex as well as coincide with the lower dimensional components of the function if the approach limit direction agrees with the direction of the adjacent faces. For the transformation image of any such function satisfying the listed properties, this correspondingly yields that the solution by (i) continuously extends to the new face N_1 with the extension by (ii) being bounded by the value of the function in $\Box_0^{(\varnothing)}$ as well as also being smooth and smoothly compatible with the remainder of the function towards the boundary points of N_1, thus pathwise smooth.

With more than two alleles, the situation is analogous, but more intricate as one has to take into account the hierarchical structure of $\overline{\Delta}_n$: We may then expect that any function describing the path probabilities for a successive extinction of alleles fulfils adequate limit requirements on the simplex, which are formulated in the following definition in terms of the underlying model:

6.15 Definition (domination limit property)**.** Let $n \geq 2$ and $I_d := \{i_0, i_1, \dots, i_d\} \subset \{0, 1, \dots, n\}$ for $d = 0, \dots, n$ with $i_i \neq i_j$ for $i \neq j$ and w.l.o.g. $i_0 = 0$ (if $i_0 \neq 0$, the vertices may be permuted correspondingly). In the setting of the Wright–Fisher model with allele set I_n, a function $U = \sum_{d=0}^{n} u_d \in C_{p0}^{\infty}\big(\bigcup_{d=0}^{n} \Delta_d^{(I_d)}\big)$ giving the hit probabilities for the eventual target set $\{i_0\}$ along the extinction path

$$\{i_0, \dots, i_{n-1}\} \longrightarrow \{i_0, \dots, i_{n-2}\} \longrightarrow \dots \longrightarrow \{i_0, i_1\} \longrightarrow \{i_0\} \tag{6.97}$$

(cf. equation (5.135), p. 165) is said to fulfil the *domination limit property* if for all index pairs (d_1, d_2) out of $\{0, \dots, n\}$ with $d_1 + 2 \leq d_2$ and for all sequences $(p_l)_{l \geq 0}$ in $\Delta_{d_2}^{(I_{d_2})}$ with $\lim_{l\to\infty} p_l = p \in \Delta_{d_1}^{(I_{d_1})}$, i.e. $p^{i_{d_1+1}}, \dots, p^{i_{d_2}} = 0$ and particularly $p^{i_{d_1}} > 0$, the following properties hold:

(i) $u_{d_2}(p_l)$ has a *finite limit* for $l \to \infty$ if simultaneously all limits for the *pivotal allele ratios* $\frac{p_l^{i_{d_1+2}} + \dots + p_l^{i_{d_2}}}{p_l^{i_{d_1+1}} + p_l^{i_{d_1+2}} + \dots + p_l^{i_{d_2}}}, \dots, \frac{p_l^{i_{d_2}}}{p_l^{i_{d_2-1}} + p_l^{i_{d_2}}}$ exist.

(ii) $\lim_{l\to\infty} u_{d_2}(p_l)$ is less or equal in absolute value than $u_{d_1}(p)$ and depends *smoothly* on $\lim_{l\to\infty} p_l$ as well as on the given pivotal limit ratios. Furthermore, for boundary values of the pivotal limit ratios, we require *compatibility* in the sense that the limit (ii)$_\mathrm{a}$ smoothly coincides with any (other) domination limit of U for $\Delta_d^{(I_d)}$ if $\lim_{l\to\infty} \frac{p_l^{i_{d+1}}+\ldots+p_l^{i_{d_2}}}{p_l^{i_d}+p_l^{i_{d+1}}+\ldots+p_l^{i_{d_2}}}$ equals zero for some $d \geq d_1+1$,

(ii)$_\mathrm{b}$ in particular smoothly coincides with u_d in $\Delta_d^{(I_d)}$ (resp. its continuous extension) if all $\lim_{l\to\infty} \frac{p_l^{i_{d+1}}+\ldots+p_l^{i_{d_2}}}{p_l^{i_d}+p_l^{i_{d+1}}+\ldots+p_l^{i_{d_2}}}, \ldots, \lim_{l\to\infty} \frac{p_l^{i_{d_2}}}{p_l^{i_{d_2-1}}+p_l^{i_{d_2}}}$ equal zero for some $d \geq d_1$,

(ii)$_\mathrm{c}$ smoothly attains zero if at least one of the pivotal allele limit ratios in (i) equals 1.

6.16 Remark. Definition 6.15 is well-defined in the sense that the conditions (ii)$_\mathrm{a}$, (ii)$_\mathrm{b}$ and (ii)$_\mathrm{c}$ likewise apply to domination limits for higher dimensions $d_1' > d_1$; in condition (ii)$_\mathrm{b}$, the functions u_d pathwise smoothly connect by lemma 6.8, and in condition (ii)$_\mathrm{c}$, the relevant pivotal allele ratios rather appear as coordinates (cf. also the proof of proposition 6.17). Moreover, the case $d_2 = d_1 + 1$ (which would also apply for $n = 1$) is not included in the definition as this corresponds to a regular transition to a boundary instance of subsequent lower dimension.

Returning to the differential equations nomenclature, a function giving the hit probabilities of course is a solution of the corresponding Kolmogorov backward equation, whereas 'for the extinction path' corresponds to 'for the extension path' with the extension constraints 5.20 applying; the latter is somewhat redundantly also formulated by the $C_{p_0}^{\infty}$-regularity requirement. In the following, we will illustrate the significance of the further stipulations in definition 6.15.

Starting with the pivotal ratios in (i), they may be seen as determining the (previous) path probabilities until reaching the new starting set of alleles $I_{d_1} = \{i_0, \ldots, i_{d_1}\}$ resp. $\Delta_d^{(I_{d_1})}$ as the domain of u_{d_1} (plus lower-dimensional components). This path until reaching I_{d_1}, which may be perceived as a previous path, does not need to start with n alleles; instead it may start with an arbitrary number of $d_2 \leq n$ alleles, thus not accounting for earlier losses of alleles (if applicable).

Since it is always two alleles which are competing (in accordance with the fibration property (cf. section 5.4.4)), in the first step, where i_{d_2} and i_{d_2-1} are rivalling, the pivot ratio is given by

$$\frac{p_l^{i_{d_2}}}{p_l^{i_{d_2-1}} + p_l^{i_{d_2}}}, \tag{6.98}$$

whereas for the next step it is assumed that, by the loss of allele i_{d_2}, its share is transferred to allele i_{d_2-1}, thus the pivot ratio is given by

$$\frac{p_l^{i_{d_2-1}} + p_l^{i_{d_2}}}{p_l^{i_{d_2-2}} + p_l^{i_{d_2-1}} + p_l^{i_{d_2}}}. \tag{6.99}$$

This is continued until the last pivotal ratio, which is given by

$$\frac{p_l^{i_{d_1+2}} + \ldots + p_l^{i_{d_2}}}{p_l^{i_{d_1+1}} + p_l^{i_{d_1+2}} + \ldots + p_l^{i_{d_2}}}. \tag{6.100}$$

The limit for $l \to \infty$ of

$$\frac{p_l^{i_{d_1+1}} + \ldots + p_l^{i_{d_2}}}{p_l^{i_{d_1}} + p_l^{i_{d_1+1}} + \ldots + p_l^{i_{d_2}}}, \tag{6.101}$$

corresponding to the pivot ratio for allele i_{d_1+1} winning over i_{d_1}, is zero by assumption – which is a direct consequence of the chosen setting of $i_0, \ldots, i_{d_1}$ dominating the population.

Thus, the ratios contain all information about the previous path, which is passed on to the limit case – the resulting domination limit path probability may then be perceived as a corresponding modification of the 'proper' path probability u_{d_1} for starting in $\Delta_{d_1}^{(I_{d_1})}$, which is – due to the inclusion of the previous path – always less or equal than u_{d_1}. This is expressed in 6.15 (ii); equality would correspond to reaching $\Delta_{d_1}^{(I_{d_1})}$ almost surely in the previous path.

Furthermore, item (ii) in definition 6.15 lists several compatibility conditions for the obtained limit, which are needed to obtain the desired regularity result in the following; in particular, this is the required smoothness of the domination limit. Moreover, condition (ii)$_\mathrm{a}$ ensures that the limit is smoothly compatible with other domination limits at other instances of the domain, whereas condition (ii)$_\mathrm{b}$

guarantees that the 'proper' path probability u_d in $\Delta_d^{(I_d)}$ is smoothly matched if all antecedent pivotal allele ratios vanish, which corresponds to almost surely reaching $\Delta_d^{(I_d)}$ beforehand; this also holds at higher levels. If, however, a certain pivotal allele ratio becomes 1, this implies that the respective starting domain $\Delta_d^{(I_d)}$ is almost surely missed in the previous path – and condition $(\mathrm{ii})_{\mathrm{c}}$ demands that the domination limit smoothly vanishes in such a case.

For functions fulfilling the domination limit property, we now have, enhancing the statement of lemma 6.8:

6.17 Proposition. *Let $n \geq 2$ and $I_d := \{i_0, i_1, \ldots, i_d\} \subset \{0, 1, \ldots, n\}$ for $d = 0, \ldots, n$ with $i_i \neq i_j$ for $i \neq j$ and w. l. o. g. $i_0 = 0$. Then a full blow-up transformation $\Phi_{s_{n-1}}^{r_{n-1}} \circ \ldots \circ \Phi_{s_1}^{r_1}$ as in proposition 6.5 transforms a function $U = \sum_{d=0}^{n} u_d \in C_{p_0}^{\infty}\big(\bigcup_{d=0}^{n} \Delta_d^{(I_d)}\big)$ satisfying the domination limit property 6.15 into a function $\tilde{U} = \sum_{d=0}^{n} \tilde{u}_d \colon \bigcup_{d=0}^{n} \square_d^{(I'_d)} \longrightarrow \mathbb{R}$ with extension to $\overline{\square_n^{(I'_n)}}$ which is of class $C_p^{\infty}\big(\overline{\square_n^{(I'_n)}}\big)$.*

To show the assertion of the proposition, we will trace the regularity of $\tilde{U}$ towards the additional faces back to that of U in $\overline{\Delta_n^{(I_n)}}$ for approaching the incompatibilities – which is accomplished by the priorly following lemma – and thereon use the domination limit property 6.15. Note that in the following we will use a disjoint formulation of the additional faces by putting

$$N_j = \square_{j-1}^{(I'_{j-1})} \times \{0\}^{(\{i_j\})} \times \overline{\boxtimes_{n-j}^{(I'_n \setminus I'_j)}}. \tag{6.102}$$

6.18 Lemma. *In the setting of a full blow-up transformation as in proposition 6.5, for $d = 1, \ldots, n$ the additional face $N_d = \square_{d-1}^{(I'_{d-1})} \times \{0\}^{(\{i_d\})} \times \overline{\boxtimes_{n-d}^{(I'_n \setminus I'_d)}} \subset \overline{\square_n^{(I'_n)}}$ corresponds to $\Delta_{d-1}^{(I_{d-1})} \subset \overline{\Delta_n^{(I_n)}}$ with additional values existing for $\frac{p^{i_{d+1}} + \ldots + p^{i_n}}{p^{i_d} + p^{i_{d+1}} + \ldots + p^{i_n}}, \ldots, \frac{p^{i_n}}{p^{i_{n-1}} + p^{i_n}}$ (perceivable as limits of sequences as in definition 6.15). Furthermore, for $j = 1, \ldots, d-1$ the face $\{\tilde{p}^{i_j} = 1\} \subset \overline{\square_{d-1}^{(I'_{d-1})}}$ corresponds to $p^{i_{j-1}} = 0$ in $\overline{\Delta_{d-1}^{(I_{d-1})}}$, in particular its interior corresponds to $\Delta_{d-2}^{(I_{d-1} \setminus \{i_{j-1}\})}$.*

Proof. To take account of the 'additional' faces N_m of $\overline{\square_n^{(I'_n)}}$ produced during the blow-up transformations, we carry out the full blow-up transformation by proposition 6.5, yielding

$$\tilde{p}^{i_1} := p^{i_1} + \ldots + p^{i_n}, \tag{6.103}$$

$$\tilde{p}^{i_2} := \begin{cases} \frac{p^{i_2}+\ldots+p^{i_n}}{p^{i_1}+p^{i_2}+\ldots+p^{i_n}} & \text{for } p^{i_1} + \ldots + p^{i_n} > 0 \\ 0 & \text{for } p^{i_1} + \ldots + p^{i_n} = 0, \end{cases} \tag{6.104}$$

$$\vdots$$

$$\tilde{p}^{i_j} := \begin{cases} \frac{p^{i_j}+\ldots+p^{i_n}}{p^{i_{j-1}}+p^{i_j}+\ldots+p^{i_n}} & \text{for } p^{i_{j-1}} + \ldots + p^{i_n} > 0 \\ 0 & \text{for } p^{i_{j-1}} + \ldots + p^{i_n} = 0, \end{cases} \tag{6.105}$$

$$\vdots$$

$$\tilde{p}^{i_n} := \begin{cases} \frac{p^{i_n}}{p^{i_{n-1}}+p^{i_n}} & \text{for } p^{i_{n-1}} + p^{i_n} > 0 \\ 0 & \text{for } p^{i_{n-1}} + p^{i_n} = 0 \end{cases} \tag{6.106}$$

for $p \in \bigcup_{d=0}^{n} \Delta_d^{(I_d)}$ and conversely

$$p^{i_1} = \tilde{p}^{i_1}(1 - \tilde{p}^{i_2}), \tag{6.107}$$

$$\vdots$$

$$p^{i_j} = \tilde{p}^{i_1} \cdots \tilde{p}^{i_j}(1 - \tilde{p}^{i_{j+1}}), \tag{6.108}$$

$$\vdots$$

$$p^{i_{n-1}} = \tilde{p}^{i_1} \cdots \tilde{p}^{i_{n-1}}(1 - \tilde{p}^{i_n}), \tag{6.109}$$

$$p^{i_n} = \tilde{p}^{i_1} \cdots \tilde{p}^{i_n} \tag{6.110}$$

for $\tilde{p} \in \bigcup_{d=0}^{n} \square_d^{(I'_d)}$ (note that we also have $p^{i_0} = 1 - \tilde{p}^{i_1}$); however, the given equations also smoothly extend to the entire $\overline{\square_n^{(I'_n)}}$. This allows it to also transform the $N_d \subset \overline{\square}_n$ back to $\overline{\Delta}_n$, i. e. $\tilde{p}^{i_d} = 0$ implies $p^{i_d}, \ldots, p^{i_n} = 0$, whereas $0 < \tilde{p}^{i_1}, \ldots, \tilde{p}^{i_{d-1}} < 1$ leads to $p^{i_1}, \ldots, p^{i_{d-1}} > 0$. Keeping the values of $\tilde{p}^{i_{d+1}}, \ldots, \tilde{p}^{i_n}$ yields the pivotal allele (limit) ratios $\frac{p^{i_{d+1}}+\ldots+p^{i_n}}{p^{i_d}+p^{i_{d+1}}+\ldots+p^{i_n}}, \ldots, \frac{p^{i_n}}{p^{i_{n-1}}+p^{i_n}}$. If however $\tilde{p}^{i_j} = 1$, this corresponds to $p^{i_{j-1}} = 0$ (and $p^{i_1}, \ldots, p^{i_{j-1}}, p^{i_{j+1}} \ldots, p^{i_d} > 0$ if $0 < \tilde{p}^{i_1}, \ldots, \tilde{p}^{i_{j-1}}, \tilde{p}^{i_{j+1}}, \ldots, \tilde{p}^{i_d} < 1$ and $\tilde{p}^{i_{d+1}} = 0$). $\square$

Proof of proposition 6.17. By lemma 6.8, we already have that $\tilde{U}$ has a (vanishing) extension to all faces $\{\tilde{p}^{i_j} = 1\}$ (and naturally their intersections as well) which is of class C_p^∞. Striving to demonstrate the corresponding extendability to the additional faces, we first note that $\square_{n-1}^{(I'_{n-1})} \equiv \square_{n-1}^{(I'_{n-1})} \times \{0\}^{(\{i_n\})}$, for which the extendability has already been shown, may be perceived as N_n – we will use this notation as a reference in the following.

For the first 'true' additional face

$$N_{n-1} = \square_{n-2}^{(I'_{n-2})} \times \{0\}^{(\{i_{n-1}\})} \times \overline{\boxtimes_1^{(\{i_n\})}}, \tag{6.111}$$

we have: In accordance with lemma 6.18, a sequence $(\tilde{p}_l)_{l\geq 0}$ in $\square_n^{(I'_n)}$ with $\lim_{l\to\infty} \tilde{p}_l = \tilde{p} \in N_{n-1}$ corresponds to a sequence $(p_l)_{l\geq 0}$ in $\Delta_n^{(I_n)}$ with limit $p \in \Delta_{n-2}^{(I_{n-2})}$ and pivotal limit ratio $\lim_{l\to\infty} \frac{p_l^{i_n}}{p_l^{i_{n-1}}+p_l^{i_n}} \equiv \tilde{p}^{i_n}$. Thus, this corresponds to the situation of a domination limit as presented in definition 6.15. Since U resp. its components are assumed to fulfil the domination limit property on $\overline{\Delta_n^{(I_n)}}$, $\lim_{l\to\infty} u_n(p_l)$ exists and depends smoothly on $\lim_{l\to\infty} p_l$ and $\lim_{l\to\infty} \frac{p_l^{i_n}}{p_l^{i_{n-1}}+p_l^{i_n}} \equiv \tilde{p}^{i_n}$. Hence, $\lim_{l\to\infty} \tilde{u}_n(\tilde{p}_l)$ likewise exists and depends smoothly on $\lim_{l\to\infty} \tilde{p}_l$, yielding the smooth extendability of $\tilde{u}_n$ up to N_{n-1}. By definition 6.15 (ii)$_\mathrm{b}$, for $\tilde{p}^{i_n} \to 0$ this domination limit smoothly connects with $\tilde{u}_{n-2}$ in $\square_{n-2}^{(I'_{n-2})} \subset \overline{N}_{n-1}$ (which itself smoothly connects with $\tilde{u}_{n-1}$ in $\square_{n-1}^{(I'_{n-1})} \equiv N_n$ as indicated before) as well as smoothly attains zero for $\tilde{p}^{i_j} \to 1$ for $j = 1, \ldots, n-2, n$ resp. corresponding combinations of boundary values of coordinates (note that $\tilde{u}_{n-2}$ also smoothly vanishes for $p^{i_j} \to 1$ for $j = 1, \ldots, n-2$ also by lemma 6.8). Thus $\tilde{U}$ is pathwise smoothly extendable also to N_{n-1} (and its intersections with N_n and the faces $\{\tilde{p}^{i_j} = 1\}$).

For the remaining faces, the situation is still somewhat more complicated: If the extendability to

$$N_d = \square_{d-1}^{(I'_{d-1})} \times \{0\}^{(\{i_d\})} \times \overline{\boxtimes_{n-d}^{(I'_n \setminus I'_d)}} \quad \text{for } d = n-2, \ldots, 1 \tag{6.112}$$

is to be proven, we may consider sequences[3] $(\tilde{p}_l)_{l\geq 0}$ in $\square_n^{(I'_n)}$ with $\lim_{l\to\infty} \tilde{p}_l = \tilde{p} \in N_d$,

[3]Alternatively, one might as well consider sequences $(\tilde{p}_l)_{l\geq 0}$ in $\square_{d'}^{(I'_{d'})}$ for some $d' > d$, which, however, by definition 6.15 (ii)$_\mathrm{a}$ yields equivalent results; thus the index n may be replaced by d' in the following. This independence of the starting domain also applies to other references to

which this time correspond to sequences $(p_l)_{l\geq 0}$ in $\Delta_n^{(I_n)}$ with $\lim_{l\to\infty} p_l = p \in \Delta_{d-1}^{(I_{d-1})}$ plus pivotal limit ratios

$$\lim_{l\to\infty} \frac{p_l^{i_{d+1}} + \ldots + p_l^{i_n}}{p_l^{i_d} + p_l^{i_{d+1}} + \ldots + p_l^{i_n}} \equiv \tilde{p}_n^{i_{d+1}}, \ldots, \lim_{l\to\infty} \frac{p_l^{i_n}}{p_l^{i_{n-1}} + p_l^{i_n}} \equiv \tilde{p}_n^{i_n} \tag{6.113}$$

(cf. again lemma 6.18). Thus, this likewise corresponds to the situation of a domination limit as presented in definition 6.15. Since U resp. its components are assumed to fulfil the domination limit property on $\overline{\Delta_n^{(I_n)}}$, $\lim_{l\to\infty} u_n(p_l)$ exists and depends smoothly on $\lim_{l\to\infty} p_l$ and the pivotal limit ratios. Hence, $\lim_{l\to\infty} \tilde{u}_n(\tilde{p}_l)$ likewise exists and depends smoothly on $\lim_{l\to\infty} \tilde{p}_l$, which yields the smooth extendability of $\tilde{u}_n$ up to N_d.

In accordance with definition 6.15 (ii)$_\mathrm{a}$, on those boundary parts of N_d with $\tilde{p}^{i_j} = 0$ for $j \geq d+1$, this domination limit smoothly connects with the one for N_j obtained previously; if multiple coordinates equal zero, then this holds for the smallest coordinate index j with $\tilde{p}^{i_j} = 0$ with the compatibility between N_j and other $N_{\bar{\jmath}}$ having been shown analogously in a previous step. If in particular there is an index $\hat{\jmath} \geq d$ such that $\tilde{p}^{i_{\hat{\jmath}}}, \ldots, \tilde{p}^{i_n} = 0$, then by condition 6.15 (ii)$_\mathrm{b}$ the domination limit smoothly connects with $\tilde{u}_{\hat{\jmath}-1}$ in $\square_{\hat{\jmath}-1}^{(I'_{\hat{\jmath}-1})} \subset \overline{N}_{\hat{\jmath}}$ resp. a corresponding domination limit of $\tilde{u}_{\hat{\jmath}-1}$; since the $\tilde{u}_d$ themselves connect smoothly (cf. lemma 6.8), this holds for all possible choices of $\hat{\jmath}$.

Similarly, if (additionally) $\tilde{p}^{i_k} \to 1$ for suitable $k = d+1, \ldots, n$, by (ii)$_\mathrm{c}$ the domination limit smoothly attains the value 0, however, this holds as well for $k = 1, \ldots, d-1$: As $\tilde{u}_{d-1}$ in $\square_{d-1}^{(I'_{d-1})}$ smoothly vanishes for $\tilde{p}^{i_k} \to 1$ for $k = 1, \ldots, d-1$ (cf. again lemma 6.8 and lemma 6.18) and any domination limit for $\Delta_{d-1}^{(I_{d-1})}$ is always less or equal in absolute value than u_{d-1} in $\Delta_{d-1}^{(I_{d-1})}$ (corresponding to $\tilde{u}_{d-1}$ in $\square_{d-1}^{(I'_{d-1})}$) by definition 6.15 (ii), it follows that $\tilde{U}(\tilde{p})$ for $\tilde{p} \in N_d$ is bounded by $\tilde{U}(\tilde{p}') \equiv \tilde{u}_{d-1}(\tilde{p}')$ with $\tilde{p}' \in \square_{d-1}^{(I'_{d-1})}$ being the projection image of $\tilde{p}$ by removing the coordinates with index strictly greater than $d-1$, thus also $\tilde{U}$ needs to vanish smoothly for $\tilde{p}^{i_k} \to 1$.

Hence, $\tilde{U}$ is extendable also to N_d and its intersections with $N_{d+1}, \ldots, N_n$ and all faces $\{\tilde{p}^{i_j} = 1\}$ such that this extension is pathwise smooth. Eventually, this yields the C_p^∞-extendability of $\tilde{U}$ to the entire $\overline{\square_n^{(I'_n)}}$. $\square$

an existing domination limit.

Bibliography

[1] C. D. Aliprantis and O. Burkinshaw. *Principles of Real Analysis*. Academic Press, London, 3rd edition, 1998.

[2] S.-I. Amari. *Differential Geometry in Statistical Inference*, chapter Differential Geometrical Theory of Statistics, pages 19–72. Number 10 in Lecture Notes – Monographs Series. Institute of Mathematical Statistics, Hayward, CA, 1987.

[3] S.-I. Amari. Information Geometry on Hierarchy of Probability Distributions. *Information Theory, IEEE Transactions on*, 47(5):1701–1711, 2001.

[4] P. L. Antonelli, K. Morgan, and G. M. Lathrop. The Geometry of Random Drift III. Recombination and Diffusion. *Advances in Applied Probability*, 9(2):260–267, 1977.

[5] P. L. Antonelli and C. Strobeck. The Geometry of Random Drift I. Stochastic Distance and Diffusion. *Advances in Applied Probability*, 9(2):238–249, 1977.

[6] G. J. Baxter, R. A. Blythe, W. Croft, and A. J. McKane. Utterance Selection Model of Language Change. *Physical Review E*, 73(4):046118, 2006.

[7] G. J. Baxter, R. A. Blythe, and A. J. McKane. Exact Solution of the Multi-Allelic Diffusion Model. *Mathematical Biosciences*, 209(1):124–170, 2007.

[8] W. K. Clifford. A Preliminary Sketch of Biquaternions. *Proceedings of the London Mathematical Society*, 4(64, 65):318–395, 1871.

[9] M. P. Do Carmo. *Riemannian Geometry*. Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, translation of the 2nd edition, 1992.

[10] W. J. Ewens. *Mathematical Population Genetics*. Springer, New York, 2nd edition, 2004.

[11] R. A. Fisher. On the Dominance Ratio. *Proceedings of the Royal Society of Edinburgh*, 42:321–341, 1922.

[12] J. Jost. Mathematical Methods in Biology and Neurobiology, monograph. *to appear*, 2008.

[13] J. Jost. *Riemannian Geometry and Geometric Analysis*. Universitext. Springer, Berlin, Heidelberg, 6th edition, 2011.

[14] J. Jost. *Partial Differential Equations*, volume 214 of *Graduate Texts in Mathematics*. Springer, Berlin, Heidelberg, 3rd edition, 2013.

[15] S. Karlin and J. McGregor. Rates and Probabilities of Fixation for Two Locus Random Mating Finite Populations without Selection. *Genetics*, 58(1):141–159, 1968.

[16] M. Kimura. Solution of a Process of Random Genetic Drift with a Continuous Model. *Proceedings of the National Academy of Sciences of the United States of America*, 41(3):144–150, 1955.

[17] M. Kimura. Stochastic Processes and Distribution of Gene Frequencies under Natural Selection. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 20, pages 33–53. Cold Spring Harbor Laboratory Press, 1955.

[18] M. Kimura. Random Genetic Drift in a Tri-Allelic Locus; Exact Solution with a Continuous Model. *Biometrics*, 12(1):57–66, 1956.

[19] E. H. Kingsley. Bernstein Polynomials for Functions of Two Variables of Class C(k). *Proceedings of the American Mathematical Society*, 2(1):64–71, 1951.

[20] A. N. Kolmogorov. Über die Analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104:415–458, 1931.

[21] R. A. Littler. Loss of Variability at One Locus in a Finite Population. *Mathematical Biosciences*, 25(1-2):151–163, 1975.

[22] R. A. Littler and E. D. Fackerell. Transition Densities for Neutral Multi-Allele Diffusion Models. *Biometrics*, 31(1):117–123, 1975.

[23] R. A. Littler and A. J. Good. Ages, Extinction Times, and First Passage Probabilities for a Multiallele Diffusion Model with Irreversible Mutation. *Theoretical Population Biology*, 13(2):214–225, 1978.

[24] M. K. Murray and J. W. Rice. *Differential Geometry and Statistics*, volume 48 of *Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC, 1993.

[25] T. Ohta and M. Kimura. Linkage Disequilibrium due to Random Genetic Drift. *Genetics Research*, 13(01):47–55, 1969.

[26] B. Øksendal. *Stochastic Differential Equations*. Universitext. Springer, Berlin, Heidelberg, 6th edition, 2003.

[27] P. F. Stadler and C. R. Stephens. Landscapes and Effective Fitness. *Comments on Theoretical Biology*, 8(4):389–431, 2003.

[28] J. Ratcliffe. *Foundations of Hyperbolic Manifolds*, volume 149 of *Graduate Texts in Mathematics*. Springer, New York, 2nd edition, 2006.

[29] S. H. Rice. *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates, Sunderland, Massachusetts, 2004.

[30] T. D. Tran. *Information Geometry and the Wright-Fisher Model of Mathematical Population Genetics*. PhD thesis, University of Leipzig, 2012.

[31] G. A. Watterson. The Effect of Linkage in a Finite Random-Mating Population. *Theoretical Population Biology*, 1(1):72–87, 1970.

[32] S. Wigert. Sur les polynomes orthogonaux et l'approximation des fonctions continues. *Arkiv för matematik, astronomi och fysik*, 22B:1–4, 1923.

[33] S. G. Wright. Evolution in Mendelian Populations. *Genetics*, 16(2):97, 1931.

[34] F. Zhang. *The Schur Complement and its Applications*, volume 4 of *Numerical Methods and Algorithms*. Springer, Boston, MA, 2005.

List of Figures

Index of notation

Index

Bibliographische Daten

On the Diffusion Approximation of Wright–Fisher Models with Several Alleles and Loci and its Geometry
(Die Diffusionsnäherung von Wright-Fisher-Modellen mit mehreren Allelen und Loci und ihre Geometrie)
Hofrichter, Julian
Universität Leipzig, Dissertation, 2014

234 Seiten, 8 Abbildungen, 34 Referenzen

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialen oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, den 21. Februar 2014

gez. Julian Hofrichter