Abstract

The quasispecies model introduced by Eigen in 1971 has close connections with the isometry group of the space of binary sequences relative to the Hamming distance metric. Generalizing this observation we introduce an abstract quasispecies model on a finite metric space $X$ together with a group of isometries $\Gamma$ acting transitively on $X$. We show that if the domain of the fitness function has a natural decomposition into the union of $t$ $G$-orbits, $G$ being a subgroup of $\Gamma$, then the dominant eigenvalue of the evolutionary matrix satisfies an algebraic equation of degree at most $t \cdot \text{rk}_\mathbb{Z} R$, where $R$ is the orbital ring that is defined in the text. The general theory is illustrated by three detailed examples. In the first two of them the space $X$ is taken to be the metric space of vertices of a regular polytope with the natural “edge” metric, these are the cases of a regular $m$-gon and of a hyperoctahedron; the final example takes as $X$ the quotient rings $\mathbb{Z}/p^n\mathbb{Z}$ with $p$-adic metric.

Keywords Quasispecies model · Finite metric space · Dominant eigenvalue · Mean population fitness · Isometry group · Regular polytope

Mathematics Subject Classification 15A18 · 92D15 · 92D25

1 Introduction

The quasispecies model, initially put forward by Eigen (1971) to comprehensively study the problem of the origin of life, is now a classical object of modern evo-
volutionary theory. More pertinent for the present paper, this model possesses a rich internal mathematical structure, as first was noted in Dress and Rumschitzki (1988) and Rumschitzki (1987), where intriguing connections between evolutionary dynamics on sequence space and tensor products of representation spaces were pointed out. This mathematical framework, interesting on its own, facilitates understanding why some versions of Eigen’s model can be solved exactly and why for some other innocently looking versions numerical computations and subtle approximations are required. In Semenov and Novozhilov (2016) we noticed and used similar connections to introduce and analyze a special case of Eigen’s model, in which two different types of sequences are present; we also formulated, using geometric language, an abstract mathematical model, which we called the generalized quasispecies or Eigen model. The goal of this paper is to present in detail, expand, and elaborate on this generalized model with the ultimate objective to outline a proper mathematical framework in which many peculiarities of the Eigen model, including the notorious error threshold, can be understood from an algebraic point of view.

Eigen’s model is quite special in bringing together abstract mathematics and biology. Even more uniquely, it also has very tight connections with statistical mechanics. The complexity and richness of the original Eigen’s model can be emphasized by the fact that it is equivalent to the famous Ising model in statistical mechanics (Leuthäusser 1986, 1987). The Ising model can be solved exactly only in some special cases, and hence any progress in understanding the conditions to solve Eigen’s model may yield insights in the analysis of the Ising model.

In what follows we neither aim for the most general formulation of the quasispecies model keeping the mutations symmetric and independent, nor do we present the most abstract version of our model, using for the most part as the specific examples of the underlying metric spaces regular polytopes with natural “edge” metrics. A more abstract example of the quotient rings \( \mathbb{Z}/p^n\mathbb{Z} \) with \( p \)-adic metric is included to illustrate the generality of the proposed algebraic framework. In this way the presentation, in our opinion, can be accessible to theoretical biologists, physicists, and mathematicians alike.

The rest of the text is organized as follows. In Sect. 2 we recall the classical Eigen’s model, provide a concise description of the main mathematical advances of its analysis and show in which way the hyperoctahedral group of isometries of the space of two-letter sequences with the Hamming distance naturally appears in the analysis of this model. This sets the stage for an abstract formulation of the generalized Eigen’s model on an arbitrary finite metric space in Sect. 3. In the same section we also review the necessary algebraic background and introduce what we call an orbital ring that allows identifying those spectral problems for which progress can be achieved. While the contribution of this paper is mainly mathematical, at the end of Sect. 3 we discuss potential biological applications of the proposed framework. Section 4 contains the main theoretical result, which is an explicit equation for the dominant eigenvalue. In Sect. 5 we apply the abstract theory developed so far to three specific cases, namely, to the regular \( m \)-gon, to the hyperoctahedral mutational landscapes, and finally to the quotient rings \( \mathbb{Z}/p^n\mathbb{Z} \) with \( p \)-adic metric. Short section (Sect. 6) is devoted to the discussion of open problems and future directions. Finally, “Appendix” contains

 Springer
technical proofs, calculations, and further algebraic properties of the studied structures necessary to make our text self-contained.

2 The quasispecies model

The quasispecies model (Eigen 1971; Eigen et al. 1988) is a system of ordinary differential equations that describes the changes with time of the vector of frequencies of different types of individuals in a population. To be specific, the individuals are defined to be sequences of a fixed length, say \( N \), composed of a two-letter alphabet \( \{0, 1\} \), hence we have \( 2^N =: l \) different types of sequences. Sequences can reproduce and mutate; the former process is incorporated into the diagonal matrix \( W = \text{diag}(w_0, \ldots, w_{l-1}) \), which is called the fitness landscape, and the latter one is described by the stochastic matrix \( Q \), which is called the mutation landscape. The entry \( w_i \geq 0 \) of \( W \) is the fitness of the sequence of type \( i \), the entry \( q_{ij} \in [0, 1] \) of \( Q \) is interpreted as the probability that, upon reproduction, the sequence of type \( j \) begets the sequence of type \( i \). It is readily shown that the asymptotic state of the vector \( \hat{p} = (\hat{p}_0, \ldots, \hat{p}_{l-1})^\top \in \mathbb{R}^l \) of frequencies of different types of sequences is the positive normalized eigenvector corresponding to the dominant (or leading) eigenvalue \( \overline{w} \) of the eigenvalue problem

\[
QW\hat{p} = \overline{w}\hat{p}.
\]

The dominant eigenvalue and the corresponding eigenvector exist under some very mild technical conditions on \( W \) and \( Q \) due to the Perron–Frobenius theorem. The leading eigenvalue \( \overline{w} \) is called the mean population fitness and is given by \( \overline{w} = \sum_{i=0}^{l-1} w_i \hat{p}_i \).

(We note that there exists an equally popular evolutionary model, which is usually called the Crow–Kimura model, whose properties are close to the problem (2.1), see, e.g., Baake and Gabriel (1999); Bratus et al. (2014); Semenov and Novozhilov (2015). Much more on the history and analysis of the various quasispecies models can be found in Baake and Gabriel (1999); Bratus et al. (2017); Jain and Krug (2007); Schuster (2015).)

To make further progress one needs to specify matrices \( W \) and \( Q \). In the simplest symmetric case we can assume that mutation at a given site of a sequence is independent from other mutations, and the mutation probability, which we denote \( 1 - q \), such that \( q \) is the fidelity, i.e., the probability of the error free reproduction, is the same for any site. Then

\[
q_{ij} = q^{N-H_{ij}}(1 - q)^{H_{ij}}, \quad i, j = 0, \ldots, l - 1,
\]

where \( H_{ij} \) is the Hamming distance between sequences of types \( i \) and \( j \) (we use the lexicographical order to index the sequences, such that sequence \( i \) is given by the binary representation of length \( N \) of the integer \( i \)). Thus the model has the natural geometry of the binary hypercube \( X = \{0, 1\}^N \), see Fig. 1.
For matrix $W$ it is possible to have different choices. One of the most frequently used is the so-called single peaked landscape (SPL), which is defined as

$$W_{SPL} := \text{diag}(w + s, w, \ldots, w), \quad w \geq 0, \ s > 0.$$  

It turns out that it is impossible, however, to calculate $\bar{w}$ and $p$ exactly in this case for finite values of $N$, and the first analysis of the quasispecies model with SPL relied heavily on numerical calculations (see Swetina and Schuster 1982 and Fig. 2). Note that numerically it is not straightforward to solve the eigenvalue problem (2.1), even for moderate values of $N$, because the dimension of the matrices is $2^N \times 2^N$. To overcome this difficulty, Swetina and Schuster (1982) considered only the so-called permutation invariant fitness landscapes for which the fitness of a given sequence is
determined by the Hamming distance from the master (zero) sequence. In this way the dimensionality of the problem can be reduced to \((N + 1) \times (N + 1)\). Indeed, instead of following the frequencies of all possible types of sequences, one only needs to take stock of the frequencies of different classes. By definition, two sequences belong to the same class if they have the same number of zeros and ones (e.g., sequences [0001] and [1000] are of different types but of the same class, obviously class with \(k\) zeros contains exactly \(\binom{N}{k}\) different types). SPL is an example of a permutation invariant fitness landscape.

Figure 2 shows the phenomenon of the notorious error threshold. On the left panel the dependence of the leading eigenvalue on the reproduction fidelity is shown, on the right panel for each fixed \(q\) the leading eigenvector \(p(q)\) of the eigenvalue problem (2.1) is shown, i.e., the frequencies of different classes of sequences at the equilibrium. If the fidelity of reproduction \(q\) is close to 1, then the whole population is dominated by one master sequence with the highest fitness (note that the graph on the right panel of Fig. 2 has the range \((0, 0.4)\)). If \(q\) decreases, the frequencies of classes with one letter 1, two letters 1, etc grow. Finally, after some critical mutation probability, that in general depends on the fitness landscape, the distribution of different classes shown in the right panel of Fig. 2 becomes abruptly binomial (which implies that the distribution of different types of sequences becomes uniform). We note that this phenomenon depends on the fitness landscape \(W\); for some \(W\) it does not manifest itself (Jain and Krug 2007; Wiehe 1997; Wilke 2005).

It turns out that it is possible to exactly calculate \(\overline{w}\) and \(p\) for Eigen’s model in the case when the contributions to the overall fitness of different sites are independent (Dress and Rumschitzki 1988; Rumschitzki 1987), and the mathematical reason for this is the decomposition of the Eigen evolutionary matrix \(QW\) as

\[
QW = Q_0W_1 \otimes Q_0W_3 \otimes \cdots \otimes Q_0W_N,
\]

where

\[
Q_0 = \begin{bmatrix}
q & 1 - q \\
1 - q & q
\end{bmatrix}, \quad W_k = \begin{bmatrix}
1 & 0 \\
0 & s_k
\end{bmatrix}, \quad k = 1, \ldots, N,
\]
is the contribution of the \(k\)th site to the fitness, and \(\otimes\) is the Kronecker product. Biologically, this case describes the absence of epistasis. More generally, as it was first noted in Dress and Rumschitzki (1988), the exact solution is in principle can be given if the structure of matrix \(W\) is related to the group of isometries of the binary hypercube \(X = \{0, 1\}^N\) (see also below).

Around the same time (at the end of 1980s) another major breakthrough about Eigen’s model was achieved: It was shown that the quasispecies model (2.1) is equivalent to the Ising model of statistical physics (Leuthäusser 1986, 1987), which caused a stream of papers that used methods of statistical physics to analyze (2.1) for various choices of \(W\) (see Baake and Wagner 2001 and references therein). Without going into the details (see, e.g., Thompson 1972 for an introduction to the Ising model), we mention that the Ising model is formulated for a given undirected graph, where the vertices can be in one of two states, and the edges represent the interactions between the vertices. In the classical two-dimensional Ising model that was solved by Onsager (1944) the graph is the lattice \(\mathbb{Z}^2\). The solution is given in the limit when the number of vertices approach infinity, and originally was obtained by analyzing the so-called transfer matrix, which, as was shown in Leuthäusser (1986, 1987), is exactly equivalent to the evolutionary Eigen matrix \(QW\). Moreover, the error threshold in Eigen’s model is the phase transition in the Ising model.

Eventually the methods of statistical physics led to the maximum principle for the quasispecies model (Baake and Georgii 2007; Hermisson et al. 2002) (see also Saakian and Hu 2006) that provides an efficient way of calculating the dominant eigenvalue \(\bar{w}\) in the case of permutation invariant fitness landscapes and under some “continuity” condition on the limit of entries \(W\) when \(N \to \infty\). Recently, the explicit expressions for the quasispecies distribution \(\hat{p}\) for the permutation invariant fitness landscapes were obtained (Cerf and Dalmau 2016a, b).

Summarizing, we remark that, notwithstanding all the progress in the analysis of Eigen’s model (2.1) outlined above, there are a great deal of open questions. In particular, we still lack analytical tools to tackle “non-continuous” fitness landscapes (but see Semenov and Novozhilov 2016), most of the existing approaches work only with permutation invariant landscapes, and there exist no necessary and sufficient conditions for the existence of the error threshold, to mention just a few. Most importantly, from our point of view, the existing analysis of Eigen’s model is almost exclusively concentrated on the case of the binary cube geometry (Fig. 1), which is supported by the biological motivation for the model (because the RNA and DNA molecules are literally polynucleotide sequences). Mathematically, however, nothing precludes us from considering an abstract model on a finite metric space \(X\) with some natural metric, thus changing the mutational landscape of Eigen’s model. We introduced such an abstract model originally in Semenov and Novozhilov (2016) and the rest of the present paper is devoted to a detailed analysis of this model. To conclude this section we remark that although the main contribution of the present text is to provide an abstract mathematical generalization of the classical Eigen’s model, potential biological applications, which are the subject of work in progress, are briefly discussed in Sect. 3.4.
3 Generalized Eigen’s model and an orbital ring

Here we first define what we call the generalized algebraic Eigen’s problem, illustrate it with several examples, and also define an orbital ring that facilitates spectral analysis of the generalized counterpart of the introduced above mutation matrix $Q$. This section is concluded with potential biological applications of the present framework.

3.1 Groups of isometries and a generalized algebraic Eigen’s problem

Let $(X, d)$ be a finite metric space. We will assume that the metric $d : X \times X \rightarrow \mathbb{N}_0$ is an integer-valued function. Consider a group $\Gamma \leq \text{Iso}(X)$ of isometries of $X$ and suppose that $\Gamma$ acts transitively on $X$, that is, $X$ is a single $\Gamma$-orbit (we consider the left action).

Since $\Gamma$ acts transitively on $X$ we may fix an arbitrary point $x_0 \in X$ and consider the function $d_{x_0} : X \rightarrow \mathbb{N}_0$ such that $d_{x_0}(x) = d(x, x_0)$. By definition,

\[
\text{diam}(X) := \max\{d_{x_0}(x) \mid x \in X\}
\]

is called the diameter of $X$. The number $N = \text{diam}(X)$ does not depend on the choice of $x_0$.

Below we give two general examples of such metric spaces, for additional examples see Semenov and Novozhilov (2016). Two special cases of the first example and the second example are treated in more details in Sect. 5.

Example 3.1 (Distance-regular, distance-transitive graphs, and regular polytopes) Let $X = L^{(0)}$ be the set of vertices of a simple, undirected, connected graph $L$. The distance $d(x, y)$ between two vertices is defined as the length of the shortest path from $x$ to $y$. In particular, each edge is supposed to be of length $e = 1$. In what follows we will call this metric $d$ the edge metric. We remark that the probabilistic properties of some random fitness landscapes on fixed graphs $L$ were considered previously in, e.g., Stadler and Happel (1999) and Stadler and Tinhofer (1999). Since we are particularly interested in the corresponding isometry group of a finite metric space $X$, we need to add some additional structure to the considered graphs. In particular, the natural candidates are the so-called distance-regular graphs. Graph $L$ of diameter $N$ is called distance-regular if there are constants $c_i, a_i, b_i$ (the intersection numbers) such that for all $i = 0, \ldots, N$ and vertices $x, y$ at distance $i = d(x, y)$, among the neighbors of $y$, there are $c_i$ at distance $i - 1$ from $x$, $a_i$ at distance $i$, $b_i$ at distance $i + 1$ (for more details see van Dam et al. 2016; Brouwer et al. 1989).

In the context of the studied quasispecies model it is even more natural to consider distance-transitive graphs (see, e.g., Biggs 1993), for which an additional property holds: if $x_1, y_1$ and $x_2, y_2$ are two pairs of vertices at distance $d(x_1, y_1) = d(x_2, y_2) = i$ then for all $i = 0, \ldots, N$ there is a graph automorphism $\gamma$ such that $\gamma x_1 = x_2$, $\gamma y_1 = y_2$. It follows that for distance-transitive graphs $L$ the group $\Gamma = \text{Aut}(L)$ acts isometrically and transitively on the associated space $X = L^{(0)}$ with the edge metric $d$. Moreover, the following symmetry of action takes place: for each $x, y \in X$ there is $\gamma \in \Gamma$ such that $\gamma x = y$, $\gamma y = x$. Finally, for a fixed vertex $x_0 \in X$ the stabilizer
(or the isotropy group) \( \Gamma_0 = \text{St}_\Gamma(x_0) = \{ \gamma \in \Gamma \mid \gamma x_0 = x_0 \} \) acts transitively on each sphere \( S_I(x_0) = \{ x \in X \mid d(x, x_0) = i \} \) with radius \( i \) and center at \( x_0 \).

Examples of distance-regular graphs are manifold and include, for instance, Cayley graphs \( L = \text{Cay}(G, H, S) \) associated with a finite group \( G \) generated by a set \( S = S^{-1} \) and its subgroup \( H \) (see Semenov and Novozhilov 2016, Section 6), and 1-skeletons of \( n \)-dimensional regular polytopes \( P \). The last example is arguably most geometrically appealing and we describe it in some more detail.

Let \( X \) be the set of vertices of an \( n \)-dimensional regular polytope \( P \) (see, e.g., Coxeter 1973), all edges of which have an integer length \( e \). For example we can consider a regular \( m \)-gon \( (m \geq 3) \) on the plane, a tetrahedron, cube, octahedron, dodecahedron, or icosahedron in the 3-dimensional space and so on, equipped with the edge metric: the distance between \( x \) and \( y \) is the minimal number of edges of \( P \) connecting \( x \) and \( y \) multiplied by \( e \). For the \( n \)-dimensional unit cube the edge metric is the same as the Hamming metric. If \( L = P^{(1)} \), 1-skeleton of \( P \), and \( e = 1 \), then \( L \) is a distance-transitive graph. The full group of isometries \( \Gamma = \text{Iso}(P) \) acts on \( P \) and, consequently, on \( X \) and \( L \). For instance, let \( P \) be an icosahedron or dodecahedron. Then \( \Gamma \cong A_5 \times S_2 \) is the full symmetry group of order 120, where \( A_5 \triangleleft S_5 \) is the alternating group of order 60, \( S_2 \) is the symmetric group of order 2.

While the previous example is visually appealing, we remark that the abstract framework we introduce below includes a lot of potentially interesting finite metric spaces that are \textit{not} associated with any distance-regular graphs. One such series of examples follows (it was announced originally in Semenov and Novozhilov 2016, Section 6, further details can be found in Sect. 5.3 below).

**Example 3.2** (Quotient rings \( \mathbb{Z}/p^n\mathbb{Z} \) with \( p \)-adic metric) Let \( p \) be a fixed prime and \( n \) be a fixed natural number. Consider the quotient ring \( X_{p^n} = \mathbb{Z}/p^n\mathbb{Z} = \{0, 1, \ldots, p^n-1\} \) (all residues viewed as numbers modulo \( p^n \)). The set \( X_{p^n} \) is equipped with the (scaled) \( p \)-adic metric \( d \): for all \( x, y \in X_{p^n} \) we have

\[
d(x, y) = \begin{cases} 
0, & \text{if } x = y, \\
\left| x - y \right|_p, & \text{if } x \neq y \text{ and } p^k \mid (x - y), \ p^{k+1} \nmid (x - y).
\end{cases}
\]

In the trivial case \( n = 1 \) the space \( (X_p, d) \) is just the metric space associated to a simplex with \( p \) vertices (with the edge metric) or to a complete (distance-transitive) graph with \( p \) vertices. If, however, \( n \geq 2 \) then \( (X_{p^n}, d) \) is not generated by any distance-regular graph in the sense of van Dam et al. (2016) because for \( (X_{p^n}, d) \) we have the values of \( d(x, y) \) equal only to either 0 or the power of \( p \), meanwhile for the distance-regular graphs the values of \( d(x, y) \) are all integers from 0 to the diameter of the graph (the exceptional case \( p = n = 2 \) is not hard to treat and left to the reader).

Now consider a quadruple \( (X, d, \Gamma, w) \), where \( (X, d) \) is a finite metric space of diameter \( N \) with integer distances and cardinality \( l = |X| \), a group \( \Gamma \trianglelefteq \text{Iso}(X) \) is a fixed group and a \textit{fitness function} \( w : X \rightarrow \mathbb{R}_{\geq 0} \). The fitness function is often represented by the vector-column \( w = (w_x) \) with non-negative real entries called \textit{fitnesses} which are indexed by \( x \in X \) (for an appropriate ordering of \( X \)).
Definition 3.3 The quadruple \((X, d, \Gamma, w)\) is called a \textit{homogeneous} \(\Gamma\)-landscape. It is called \textit{symmetric} if for any points \(y, z \in X\) there is an isometry \(\gamma = \gamma(y, z) \in \Gamma\) such that \(\gamma y = z\) and \(\gamma z = y\).

Consider also the diagonal matrix \(W = \text{diag}(w_x)\) of order \(l\) called the \textit{fitness matrix}, the symmetric distance matrix \(D = (d(x, y))\) of the same order with integer entries, and the symmetric matrix \(Q = ((1 - q)^{d(x,y)}q^{N-d(x,y)})\) for \(q \in [0, 1]\). Finally, we introduce the \textit{distance polynomial}

\[
P_X(q) = \sum_{x \in X} (1 - q)^{d(x,x_0)}q^{N-d(x,x_0)}, \quad x_0 \in X. \tag{3.1}
\]

Since \(\Gamma\) acts transitively on \(X\) this polynomial is independent of the choice of \(x_0 \in X\) and is the sum of entries in each row (column) of \(Q\), therefore making the matrix \((1/P_X(q))Q\) to be doubly stochastic.

The following key definition generalizes the classical quasispecies Eigen’s problem.

Definition 3.4 The problem to find the leading (dominant) eigenvalue \(\bar{w} = \bar{w}(q)\) of the matrix \(1/P_X(q) \cdot Q \cdot W\) and the corresponding positive eigenvector \(\hat{p} = \hat{p}(q)\) satisfying

\[
Q \cdot \hat{p} = P_X(q) \bar{w} \cdot \hat{p}, \quad \hat{p}_x = \hat{p}_x(q) > 0, \quad \sum_{x \in X} \hat{p}_x(q) = 1, \tag{3.2}
\]

will be called the \textit{generalized algebraic quasispecies} or Eigen’s problem.

Note that in (3.2)

\[
\bar{w} = \sum_{x \in X} w_x \hat{p}_x. \tag{3.3}
\]

Due to the Perron–Frobenius theorem a solution of this problem always exists. Also note that the uniform distribution vector

\[
\hat{p} = \frac{1}{|X|}(1, \ldots, 1) \top = \frac{1}{l}(1, \ldots, 1) \top \tag{3.4}
\]

provides a solution of (3.2) in the case of constant fitnesses \(w_x \equiv w > 0\). By construction matrix \(1/P_X(q) \cdot Q\) is symmetric and doubly stochastic. It will be called \textit{normalized mutation matrix} to distinguish it from the mutation matrix \(Q\) whose entries, however, do not represent probabilities. We emphasize that here and below the term “mutation” should be understood in the general sense, as “change of state.”

Problem (3.2) turns into the classical Eigen’s quasispecies problem if \(X = \{0, 1\}^n\) is the \(n\)-dimensional binary cube with the Hamming metric \(H\), and the group \(\Gamma = \text{Iso}(X)\) named in 1930 by A. Young the \textit{hyperoctahedral} group. \(\Gamma\) is isomorphic as an abstract group to the Weyl group of the root system of type \(B_n\) or \(C_n\) and is acting on the binary cube. In this case \(P_X(q) \equiv 1\). The case when \(X\) is the set of vertices of an \(n\)-dimensional simplex with the isometry group \(\text{Iso}(X) \cong S_{n+1}\) is treated in detail in Section 6 of Semenov and Novozhilov (2016).

\(\text{Springer}\)
Here we continue with a general analysis of the generalized quasispecies problem. The first step is to list some properties of the distance polynomial, which by construction is strictly positive on \([0, 1]\) (provided the parameter \(N\) is strictly equal to \(\text{diam} \ X\)) and possesses the following properties, which are checked by direct calculations:

1. 
\[
P_X(1) = 1, \quad P_X\left(\frac{1}{2}\right) = \frac{|X|}{2^N} = \frac{l}{2^N}. \tag{3.5}
\]

2. 
\[
P_X(q) = \sum_{k=0}^{N} f_k (1-q)^k q^{N-k} \in \mathbb{Z}[q], \tag{3.6}
\]

where the non-negative integers \(f_k = f_k(X) = \#\{x \in X \mid d(x, x_0) = k\}\) are the cardinalities of \(d\)-spheres in \(X\) with the center at the fixed point \(x_0\) and of radius \(k\).

**Remark 3.5** Polynomial \(S_X(t) = \sum_{k=0}^{N} f_k t^k\) is often called the *spherical growth function* of \((X, d)\). See, for instance, de la Harpe (2000, chapter IV) for details and examples.

### 3.2 Orbital ring associated with the triple \((X, d, \Gamma)\)

In this section, to study the spectral properties of the mutation matrix \(Q\), we introduce what we call an *orbital ring*. For more algebraic details and construction of similar structures we refer the reader to Brown (1982), Feit (1982), Kirillov (1976), Semenov (1994) and Serre (1996), see also “Appendix A.1”.

Let \(A\) be a \(\Gamma\)-orbit in \(X \times X\), \(\text{Orb}\) be the set of all such orbits, and let \(M_A\) be the matrix with entries \((M_A)_{x, y} = 1\) if \((x, y) \in A\) and equal to 0 otherwise. From a purely algebraic point of view it is worth mentioning that \(M_A\) can be identified with the matrix of the \(\Gamma\)-invariant \(\mathbb{Z}\)-linear endomorphism \(f_A \in \text{Hom}(\mathbb{Z}X, \mathbb{Z}X)\) such that \(f_A(y) = \sum_{(x, y) \in A} x, \mathbb{Z}X\) being a permutation \(\mathbb{Z}\Gamma\)-module (see, e.g., Brown 1982). It can be shown (see “Appendix A.1”) that 

\[
M_A M_B = \sum_{C \in \text{Orb}} \mu_{A B}^C M_C
\]

for some integer nonnegative structure constants \(\mu_{A B}^C\), and the matrices \(M_A\) compose a \(\mathbb{Z}\)-basis of a commutative ring with unity (see “Appendix A.1” for more details on this ring), thus prompting the following definition.

**Definition 3.6** The ring

\[
R = R(X, d, \Gamma) = \left\{ \sum_{A \in \text{Orb}} n_A M_A, n_A \in \mathbb{Z} \right\}
\]

is called *orbital* ring associated with the symmetric triple \((X, d, \Gamma)\).
As an abelian group $R$ is free and of rank $\text{rk}_\mathbb{Z} R = |\Gamma_0 \backslash \Gamma / \Gamma_0|$, the number of double $(\Gamma_0, \Gamma_0)$-cosets in $\Gamma$ where $\Gamma_0 = \text{St}_\Gamma(x_0)$. Since $\Gamma$ acts by isometries on $X$ then the distance function $d = d(x, y)$ is $\Gamma$-invariant with respect to the diagonal action of $\Gamma$ on the cartesian square $X \times X$, namely, $d(\gamma x, \gamma y) = d(x, y)$ for any $x, y \in X$ and $\gamma \in \Gamma$.

It is well known that the set $\text{Orb}$ of $\Gamma$-orbits $A$ in $X \times X$ is in 1-to-1 correspondence with the set $\Gamma_0 \backslash \Gamma / \Gamma_0$ of double $(\Gamma_0, \Gamma_0)$-cosets in $\Gamma$ where $\Gamma_0 = \text{St}_\Gamma(x_0)$. We will say that $\Gamma$-orbit $A$ is of degree $k$ ($\deg A = k$), $k = 0, \ldots, N$, if $d(x, y) = k$ for some (and hence for any) $(x, y) \in A$. By definition, all $\Gamma$-orbits $A$ of degree $k$ compose a subset $\text{Orb}_k \subset \text{Orb}$.

Note that the single $\Gamma$-orbit of degree 0 is the diagonal $\Delta = \{(x, x) | x \in X\} \subset X \times X$. The corresponding matrix $M_\Delta = I$, the identity matrix. Since different orbits $A$ are disjoint the matrices $M_A$ are independent over $\mathbb{Z}$. Therefore, we have the following expansion of the mutation matrix $Q$:

$$
Q = \sum_{k=0}^{N} (1 - q)^k q^{N-k} \sum_{A \in \text{Orb}_k} M_A, \quad (3.7)
$$

and the equality

$$
E = \sum_{A \in \text{Orb}} M_A = \sum_{k=0}^{N} \sum_{A \in \text{Orb}_k} M_A, \quad (3.8)
$$

where $E$ is the matrix with all the entries equal to 1.

Let us finish this section with a few words on the relation between the orbital rings introduced above and the adjacency algebras for the distance-transitive graphs (see Example 3.1). If $(X, d)$ is the metric space attached to a distance-transitive graph $L$ then the stabilizer $\Gamma_0 = \text{St}_\Gamma(x_0)$ in the group $\Gamma = \text{Aut}(L)$ acts transitively on each (non-empty) sphere $S_k(x_0)$ of radius $k$, where $k = 0, 1, \ldots, N$ and $N = \text{diam}(X)$. It follows that there are exactly $N + 1$ orbital matrices $M_0 = I, \ldots, M_N$ with entries $(M_k)_{x,y}$ equal to 1 if $d(x, y) = k$ and equal to 0 otherwise. Comparing these matrices with the distance-$k$ (adjacency) matrices for the graph $L$ (see van Dam et al. 2016, Subsection 2.5) we can conclude that they are the same. Hence, in this case, the real algebra $\mathbb{R} \otimes R(X, d, \Gamma)$ is exactly the adjacency algebra of dimension $r = N + 1$ for the graph $L$. In general, however, the orbital ring (algebra) cannot be reduced to an adjacency algebra for a distance-regular graph as it is shown in Example 3.2 and further in Sect. 5.3.

### 3.3 Spectral properties of the mutation matrix $Q$

Here we summarize the spectral properties of the mutation matrix $Q$. Our proof relies heavily on the orbital ring introduced in the previous section. Since we are especially interested in a particular choice of the transition matrix $T$ (here transition means the transition to an eigenbasis), see below Conjecture 3.9, formula 4.16, and examples in Sect. 5, we include an explicit, albeit technical, proof of this statement. We note that
in this proof, as well as in the following text, we consider matrix $Q$ as the function of parameter $q$: $Q = Q(q)$.

**Theorem 3.7** Let the triple $(X, d, \Gamma)$ be symmetric. Then there exists a non-degenerate real constant transition matrix $T = (t_{x,y})$ of order $l = |X|$ such that

1. All matrix entries of $T$ are integer algebraic (over the field of rationals $\mathbb{Q}$) numbers.
2. The column of $T$ indexed by a fixed $x_0 \in X$ is equal to $1 = (1, \ldots, 1)^\top$. If $y \neq x_0$ then $\sum_{x \in X} t_{x,y} = 0$.
3. $T^{-1}QT = \text{diag}(P_x(q))$, \hspace{1cm} (3.9)

where $P_{x_0}(q)$ is the distance polynomial $P_X(q) \in \mathbb{Z}[q]$, other eigenpolynomials $P_x(q) \in \mathbb{R}[q]$ have integer algebraic coefficients and $P_x(1) = 1$ for any $x \in X$.
4. There are at most $r = \text{rk}_\mathbb{Z} \mathbb{R} = |\text{Orb}|$ different eigenpolynomials $P_x(q)$ in (3.9).
5. For the distance matrix $D$ and $N = \text{diam}(X)$

$$T^{-1}DT = \text{diag}(\lambda_x), \quad \lambda_x = N2^{N-1}P_x(1/2) - 2^{N-2}P'_x(1/2),$$

where $P'_x(1/2)$ is the value of the derivative $dP_x/dq$ at $q = 1/2$.

**Proof** Consider the space $V = \text{Hom}_\mathbb{R}(X, \mathbb{R})$ of all linear functions $f : X \rightarrow \mathbb{R}$. Each function of $V$ can be represented as a vector-column $v = (f(x))$ (in fact, a covector). The matrix $M_A$ (see the previous section for the definition) defines a linear endomorphism $M_A : V \rightarrow V$ such that

$$M_A f(x) = \sum_{y : (x,y) \in A} f(y).$$

If $v = (f(x))$ then this endomorphism is just the multiplication $v \mapsto M_A v$.

Let us first show that each endomorphism $M_A$ commutes with $\Gamma$-action on $V = \text{Hom}_\mathbb{R}(X, \mathbb{R})$ given by the rule $\gamma f(x) = f(\gamma^{-1}x)$, $\gamma \in \Gamma$. In fact,

$$\gamma M_A f(x) = M_A f(\gamma^{-1}x) = \sum_{y : (\gamma^{-1}x,y) \in A} f(y) = \sum_{y : (x,y) \in A} f(y) = M_A \gamma f(x).$$ \hspace{1cm} (3.10)

For the following we note that it is well known that each symmetric matrix over $\mathbb{R}$ is diagonalizable, has real eigenvalues and the family of commuting symmetric matrices $M_A$, $A \in \text{Orb}$, has a common eigenbasis.

1. First, since all the entries of $M_A$ are zeroes and ones, all the real eigenvalues are integer algebraic numbers. Hence we can choose eigenvectors (vector-columns) $t_y$, $y \in X$, in a common eigenbasis with integer algebraic entries (scaling the eigenvectors by appropriate integer factors if necessary) and compose a transition matrix $T$ from these eigenvectors. Thus, the first assertion is proved.
2. Moreover, the vector $1 = (1, \ldots, 1)^\top$ (the constant function $f(x) \equiv 1$ of $V$) is an eigenvector for each $M_A$ since the sum of each row of $M_A$ is constant and hence

$$M_A 1 = s_A 1, \quad s_A = \# \{ a \in X \mid (x_0, a) \in A \}. \tag{3.11}$$

Note that the leading eigenvalue $s_A$ of $M_A$ is the cardinality of the $\Gamma_0$-orbit in $X$, $\Gamma_0 = St_{\Gamma}(x_0)$, corresponding to $\Gamma$-orbit $A$ in $X \times X$. We may index the eigenvector $1$ by $x_0$.

The subspace $V_0 = \{ f \in V = \text{Hom}_R(X, R) \mid \sum_{x \in X} f(x) = 0 \}$ is invariant for each $M_A$ since each $M_A$ is symmetric and hence has orthogonal eigenvectors. Hence, we can choose the other eigenvectors from this subspace and consequently, $\sum_{x \in X} I_{x,y} = 0$.

3. Applying the conjugation by $T$ to (3.7) we get

$$T^{-1} QT = \sum_{k=0}^{N} (1 - q)^k q^{N-k} \sum_{A \in \text{Orb}_k} T^{-1} M_A T = \text{diag}(P_x(q)) \tag{3.12}$$

for appropriate polynomials $P_x(q)$. Each $P_x(q)$ is a linear combination:

$$P_x(q) = \sum_{k=0}^{N} (1 - q)^k q^{N-k} \sum_{A \in \text{Orb}_k} \lambda_{x,A}, \quad \text{diag}(\lambda_{x,A}) = T^{-1} M_A T.$$

Since $\lambda_{x_0,A} = s_A$ [see (3.11)] then

$$P_{x_0}(q) = \sum_{k=0}^{N} (1 - q)^k q^{N-k} s_A = \sum_{k=0}^{N} (1 - q)^k q^{N-k} |S_k(x_0)|$$

is the distance polynomial, $S_k(x_0)$ is the sphere of radius $k$ centered at $x_0$. For $q = 1$ we have $Q(1) = I$ and the third assertion is proved.

4. Consider the subspace $V_{\Gamma_0} \subset V = \text{Hom}_R(X, R)$ of all functions $v = (f(x))$ which are constant on the $\Gamma_0$-orbits in $X$, that is, $\Gamma_0$-invariant functions in $V = \text{Hom}_R(X, R)$. In view of (3.10) the subspace $V_{\Gamma_0} = \text{Hom}_{\Gamma_0}(X, R)$ is $M_A$-invariant for each $A \in \text{Orb}$.

It follows that we have $r = \text{rk}_R R$-dimensional representation $\pi: R \to \text{End}_R(V_{\Gamma_0})$ such that $\pi: M_A \to M_A$, since by construction the ring $R$ is the $\mathbb{Z}$-linear span of $M_A$. It is not hard to see that the representation is faithful (consider function $f$ such that $f(x_0) = 1$ and $f(x) = 0$ if $x \neq x_0$). In what follows we identify the elements of $\pi(R)$ with the corresponding matrices and $\Gamma_0$-invariant functions $f(x)$ with the corresponding vectors $v = (f(x))$.

Since the ring $R$ is commutative we can find, similar to the proofs of 1 and 3, $\Gamma_0$-invariant eigenfunctions $v_0 = (P_0(x)), \ldots, v_{r-1} = (P_{r-1}(x))$ such that $P_0(x) \equiv 1$ and

$$Q v_j = P_j(q) v_j, \quad j = 0, \ldots, r - 1,$$
where each $P_j(q)$ is an eigenpolynomial of $Q$ coinciding with one of $P_x(q)$ in (3.9). In fact, let $t_x$ be the $x$-column of the transition matrix $T$. Then $Q t_x = P_x(q)t_x$. If $t_x$ corresponds to a $Γ_0$-invariant function in $V_{Γ_0}$ the proof is complete. Otherwise we can suppose that the $y$-component $t_{y,x}$ of $t_x$ is nontrivial and apply the operation of averaging

$$t_x \mapsto \sum_{γ ∈ Γ_y} γt_x = t, \quad Γ_y = St_Γ(y) = gΓ_0g^{-1}, \quad y = gx_0.$$ 

Note that $t$ corresponds to a $Γ_y$-invariant function and $y$-component $t$ is equal to $|Γ_y|t_{y,x}$, i.e., non-trivial. The representation $π_x : R → End_R(V_{Γ_y})$ is equivalent to $π : R → End_R(V_{Γ_0})$ since $Γ$ acts transitively on $X$. In view of (3.10) each $γt_x$ is a common eigenvector of all $M_A$ and, consequently, of $Q$, corresponding to the eigenvalue $P_x(q)$, so is $t$. Since $t \neq 0$ the proof is complete.

5. Finally, direct calculation yields the equality

$$Q'(1/2) = 2N Q(1/2) - \frac{1}{2^{N-2}} D,$$

(here $Q'$ means the differentiated matrix $dQ/dq$) whence

$$D = N2^{N-1} Q(1/2) - 2^{N-2} Q'(1/2)$$

and

$$T^{-1} DT = N2^{N-1} T^{-1} Q(1/2) T - 2^{N-2} T^{-1} Q'(1/2) T = \text{diag}(λ_x),$$

$$λ_x = N2^{N-1} P_x(1/2) - 2^{N-2} P_x'(1/2).$$

By virtue of Assertion 4 there are at most $r = rk_Z R$ different eigenvalues of $D$.

The theorem is proved. □

Example 3.8 In Semenov and Novozhilov (2016) we considered the quasispecies symmetric triple $(X, d, Γ)$ where $X = \{0, 1\}^n$ is the binary hypercube with the Hamming metric $d$ of dimension $n = N = \text{diam} X$ and $Γ$ is the hyperoctahedral group of order $2^n \cdot n!$. Let $x_0 = 0 = [0, \ldots, 0] ∈ X$ (the binary representation). Then $Γ_0 = St_Γ(x_0) ≅ S_n$ and there are exactly $r = N + 1$ $Γ_0$-orbits $A_0, \ldots, A_N$ in $X$, namely the spheres $A_k = S_k(x_0)$ of cardinalities $N \choose k$. For $Q$ there are exactly $r = N + 1$ different eigenpolynomials $P_k(q) = (2q - 1)^k$, $k = 0, \ldots, N$, of multiplicities $|A_k| = N \choose k$.

In Semenov and Novozhilov (2016) we also considered the simplicial symmetric triple $(X, d, Γ)$ where $X$ is the 0-skeleton of the regular simplex such that $|X| = n + 1$ with unit distances between different vertices. The group $Γ ≅ S_{n+1}$ and $Γ_0 ≅ S_n$. There are exactly $r = 2 Γ_0$-orbits $A_0 = \{x_0\}, A_1 = X \setminus A_0$ in $X$, namely, the spheres $A_0 = S_0(x_0), A_1 = S_1(x_0)$ of cardinalities $1, n$. For $Q$ there are $r = 2$ different eigenpolynomials, namely, $P_0(q) = q + n(1 - q)$ of multiplicity 1, $P_1(q) = 2q - 1$ of multiplicity $n$. 
Together with the calculations we present below and summarized in a table form in “Appendix C” Example 3.8 prompts us to formulate the following conjecture.

**Conjecture 3.9** For an arbitrary orbital ring $R = R(X, d, \Gamma)$ the eigenpolynomials of the corresponding mutation matrix $Q$ can be enumerated by $A \in \text{Orb}$ and there are exactly $r = \text{rk}_Z R$ different eigenpolynomials $P_A(q)$ of $Q$ of multiplicities $m_A$. It follows that

$$\sum_{A \in \text{Orb}} m_A = l = |X|.$$

*In addition, matrix $T$ in Theorem 3.7 can be chosen to be symmetric.*

### 3.4 Biological applications

Before formulating and proving our main theoretical result in the next section, we would like to pause and mention a few potentially interesting biological applications of our, rather abstract, mathematical framework. We also refer to Section 2 of Semenov and Novozhilov (2016) for an additional discussion. Some of these biological applications are currently work in progress.

First and foremost, introducing a next level of abstraction may lead to further understanding of the peculiarities of the classical quasispecies model itself.

General comments aside, one interesting example of permutation invariant fitness landscapes is the so-called *mesa-landscapes*, discussed in Wolff and Krug (2009). For these landscapes several (two or more) classes of binary sequences (see the discussion in Sect. 2 on the types and classes of sequences) have fitness $w+s$ for some $w \geq 0, s > 0$, and the rest of the classes have fitness $w$. Clearly, one can generalize this scheme with assigning fitnesses as $w+s_1, w+s_2, \text{etc.}$, for $s_1 > s_2 > \cdots > 0$, and consider it as a first reasonable generalization of the SPL, which is very well understood theoretically (for the SPL we can compute basically all the quantities of interest). It turns out that already such mild generalization leads to serious analytical troubles. An attempt to use well developed theory of the maximum principle leads to incorrect predictions (see examples in Wolff and Krug 2009) because for mesa-landscapes the scaled limit, when the sequence length $N$ tends to $\infty$, becomes a discontinuous function. Therefore, the only available tool at the present point to analyze such landscapes is numerical computations. It turns out, however, that the structure of the mesa-landscapes has exactly the decomposition which we assume in the following section [namely, Eq. (4.1)], and therefore, at least in principle, the polynomial for computing the dominant eigenvalue of mesa-landscapes, is given by an explicit equation (4.20)! The analysis of this equation is by no means a simple application of our theory, but it gives a first analytical steps to provide theoretical computations of the quasispecies models with mesa-landscapes.

On a similar note, most of the quasispecies models analyzed in the literature (see Sect. 2 for references) are permutation invariant. Our general theory, based on the decomposition (4.1), works for *non-permutation invariant* fitness landscapes and therefore provide an independent theoretical tool to treat uniformly a very broad,
albeit quite special, class of fitness landscapes. Much more on this topic can be found in Section 2 of Semenov and Novozhilov (2016).

Second, it is true that our abstract constructions have direct biological counterparts. In general, we study systems that “reproduce” with differential fitnesses and “mutate” one to another with probabilities determined by a given mutational landscape. The terms “reproduce” and “mutate” must be understood here in a very general sense as “beget an identical entity” and “change a state” respectively. The biological world around us is so rich and multifaced that some of the real biological systems can be approximately described by our abstract models.

To give one specific example, consider the phenomenon of antigenic variation in pathogens. Antigenic variation is the process by which many infectious agents protect themselves from the defence response of the immune system, and therefore is of paramount importance to be understood. Specifically, pathogens periodically change molecular composition on their surfaces, thus camouflaging from the antibodies. In some of the pathogens these changes are caused by random mutations, and therefore the mutational landscape has the geometry of hypercube. In other pathogens, however, antigenic variation is mediated by some mechanisms that specifically generate diversity in surface molecules. These mechanisms do not rely on random mutations and are, in a sense, “programmed.” The “change of state” can occur basically from any specific type of a given pathogen to any other, by switching the genes responsible for the expression of the surface protein thus giving virtually endless supply of antigen variants. In our abstract terms the mutational landscape for such pathogens is not the classical hypercube, but a simplex! Therefore, our abstract construction can be used to provide theoretical predictions about such systems. We would like to mention that the computations given for the simplest possible fitness landscape with simplicial mutational landscape yield the existence of the error threshold (see Semenov and Novozhilov 2016).

Third, we remark that the classical N-dimensional cube for the Eigen’s model often leads to complicated computations (e.g., despite vigourous research even for the permutation invariant fitness landscapes there exist no necessary and sufficient conditions to guarantee that the error threshold occurs in the model). For some other mutational landscapes the computations become much less cumbersome. For instance, if we consider the so-called two-valued fitness landscape when on some specific set of indices the fitness of the individuals is given by \( w + s \) for some \( w \geq 0, s > 0 \) (this specific set must be a \( G \)-orbit in terms of our framework), and all the rest have the fitness \( s \), then we show that the polynomial to determine the mean population fitness has degree \( N \) for the classical quasispecies model, where \( N \) is the sequence length, whereas for the simplicial fitness landscape this polynomial has degree 2! In Sect. 5.2 we show that for the octahedral mutational landscape this polynomial has degree 3. This means, among other things, that all the computations of the quantities of interest (the mean population fitness, the equilibrium distribution, the critical mutation probability, etc) are much simpler in these settings and can be used, we hope, as a first step to obtain a full theoretical understanding of the classical quasispecies model and move beyond it, to incorporate recombination and take a look at the diploid world.
4 \( G \)-Invariant homogeneous symmetric \( \Gamma \)-landscapes, \( G \leq \Gamma \)

Having at our disposal the orbital ring associated with the triple \((X, d, \Gamma)\) and, correspondingly, the spectral properties of \(Q\), we are in position to consider the eigenvalue problem (3.2). To make progress we restrict ourselves to some special fitness landscapes, which are constant along \(G\)-orbits, where \(G \leq \Gamma\).

4.1 Reduced problem

Let \((X, d, \Gamma, w)\) be a homogeneous \(\Gamma\)-landscape (\(\Gamma \leq \text{Iso}(X)\)) and let \(G \leq \Gamma\) be a fixed subgroup. If \(A\) is a \(G\)-orbit then \((A, d)\) is a metric subspace of \((X, d)\) on which \(G\) acts transitively by isometries. Consider the restriction \(w|_A\). Thus, the quadruple \((A, d, G, w|_A)\) can be viewed as a homogeneous \(G\)-sublandscape of \((X, d, \Gamma, w)\).

Definition 4.1 We call a \(\Gamma\)-landscape \((X, d, \Gamma, w)\) \(G\)-invariant if the fitness function \(w\) is constant on each \(G\)-orbit \(A\) of the \(G\)-action on \(X\), that is, \(w(A) = w_A \geq 0\), where \(w(A)\) is the restriction of \(w\) on \(A\).

For instance, for the trivial subgroup \(G = \{1\}\) each homogeneous \(\Gamma\)-landscape is \(G\)-invariant. If \(G = \Gamma\) and \(\Gamma\) acts transitively on \(X\) then a \(\Gamma\)-invariant landscape is such that \(w\) is constant on \(X\).

Let a \(\Gamma\)-landscape \((X, d, \Gamma, w)\) be symmetric and \(G\)-invariant. We suppose that \(G\)-invariant fitness function \(w\) has at least two values. We will also assume that there is a decomposition

\[ X = A_0 \bigcup_{i=1}^t A_i, \tag{4.1} \]

such that \(A_0\) is a union of \(G\)-orbits on which \(w(A_0) = w \geq 0\), and each \(A_i, i = 1, \ldots, t\), is just a single \(G\)-orbit on which \(w(A_i) = w + s_i\), where \(s_i > 0\) (\(s_i\) are not necessarily different). Then fitness matrix \(W\) can be represented as

\[ W = wI + \sum_{i=1}^t s_i E_{A_i}, \tag{4.2} \]

\(I\) being the identity matrix and \(E_{A_i}\) being the projection matrix with the only nontrivial entries \(e_{aa} = 1, a \in A_i\), on the main diagonal.

We want to solve problem (3.2). In view of (4.2) Eq. (3.2) reads

\[ w Q \hat{p} + Q \sum_{i=1}^t s_i E_{A_i} \hat{p} = P_X(q) \overline{w} \hat{p}, \]

whence

\[ (\overline{w} P_X(q) - w Q) \hat{p} = Q \sum_{i=1}^t s_i E_{A_i} \hat{p}. \tag{4.3} \]
For the matrix 1-norm we have \( \|w Q\|_1 = P_X(q) w < P_X(q) w = \|P_X(q) w I\|_1 \). Consequently, the matrix \( P_X(q) w I - w Q \) is non-singular and we obtain the equality

\[
\hat{p} = (P_X(q) w I - w Q)^{-1} Q v, \quad v = \sum_{i=1}^t s_i E_{A_i} \hat{p}.
\]

Multiplying the last equality by \( \sum_{i=1}^t s_i E_{A_i} \) yields

\[
v = \sum_{i=1}^t s_i E_{A_i} \hat{p} = \sum_{i=1}^t s_i E_{A_i} (\overline{w} P_X(q) I - w Q)^{-1} Q v.
\]

Denoting

\[
M = \sum_{i=1}^t s_i E_{A_i} (\overline{w} P_X(q) I - w Q)^{-1} Q
\]

we can rewrite (4.4) as

\[
v = M v, \quad v = \sum_{i=1}^t s_i E_{A_i} \hat{p},
\]

and hence vector \( v \) is an eigenvector of \( M \) corresponding to the eigenvalue \( \lambda = 1 \). Considering \( \overline{w} \) in (4.5), (4.6) as a parameter we now concentrate on the following reduced problem: To find the eigenvector \( v \) satisfying (4.6) and corresponding to the eigenvalue \( \lambda = 1 \) of matrix \( M \) defined in (4.5).

**Remark 4.2** Expanding the right-hand side of (4.5) we get

\[
M = \frac{1}{\overline{w} P_X(q)} \sum_{i=1}^t s_i \sum_{m=0}^{\infty} \left( \frac{\overline{w}}{\overline{w} P_X(q)} \right)^m E_{A_i} Q^{m+1}.
\]

The parameter \( \overline{w} = \overline{w}(q) \) satisfies the formula

\[
\overline{w} = w + \sum_{i=1}^t s_i \sum_{a \in A_i} p_a = w + \sum_{i=1}^t s_i \|E_{A_i} P\|_1.
\]

**4.2 Equation for the leading eigenvalue \( \overline{w} \)**

Here, using the notation and results from Sect. 4.1, we show that there exists an algebraic equation of degree at most \( t \cdot \text{rk}_Z R(X, d, \Gamma) \) for \( \overline{w} \). Here \( R = R(X, d, \Gamma) \) is the orbital ring defined in Sect. 3.2.
We can rewrite (4.5), (4.6) as follows \([w, \text{ defined in (4.8), is considered to be a parameter here}]\)

\[
M = \sum_{i=1}^{t} s_i E_{A_i} L, \quad L = Q(\overline{w}P_X(q)I - w Q)^{-1}, \quad (4.9)
\]

\[
\sum_{i=1}^{t} s_i E_{A_i} \hat{p} = \sum_{i=1}^{t} s_i E_{A_i} L \sum_{k=1}^{t} s_k E_{A_k} \hat{p}. \quad (4.10)
\]

Since \(E_{A_j}\) is a projection matrix, \(E_{A_j}^2 = E_{A_j}, E_{A_j} E_{A_i} = 0\) when \(i \neq j\) and \(s_j > 0\).

If we multiply both sides of (4.10) by \(E_{A_j}\) then we obtain \(t\) equalities

\[
E_{A_j} \hat{p} = \sum_{k=1}^{t} s_k E_{A_j} L E_{A_k} \hat{p}, \quad j = 1, \ldots, t. \quad (4.11)
\]

**Lemma 4.3** Let \(\Gamma\)-landscape be symmetric and \(G\)-invariant. Then positive vector \(\hat{p}\) corresponding to the dominant eigenvalue \(\overline{w}\) is constant on the \(G\)-orbits \(A_i\):

\[
(E_{A_i} \hat{p})_x \equiv C_i > 0, \quad x \in A_i, \quad (E_{A_i} \hat{p})_x \equiv 0, \quad x \notin A_i, \quad i = 1, \ldots, t. \quad (4.12)
\]

**Proof** Take some solution \(z\) of problem (3.2). In fact, the multiplication by \(Q\) commutes [see (3.10)] with \(G\)-action \(z \mapsto g z, g z(x) = z(g^{-1}x)\). Then \(Q \overline{w} z = P_X(q) \overline{w} z\) is equivalent to \(g Q g^{-1} g \overline{w} g z = P_X(q) \overline{w} g z\), or, in view of \(G\)-invariance, \(Q \overline{w} g z = P_X(q) \overline{w} g z\). Hence, \(g z = g z(x), g \in G\), is also a solution. The averaging \(z(x) \mapsto |G|^{-1} \sum g z(x)\) provides a \(G\)-invariant solution. In view of the Perron–Frobenius theorem the averaged solution is proportional to \(z\) and moreover, is equal to \(z\) due to the last condition of (3.2).

\(\square\)

The constants \(C_i\) in (4.12) are to be normalized in such a way that

\[
\sum_{x \in A_0} \hat{p}_x + \sum_{i=1}^{t} C_i |A_i| = 1. \quad (4.13)
\]

Then in view of (4.8) we get

\[
\overline{w} = w + \sum_{i=1}^{t} s_i C_i |A_i|. \quad (4.14)
\]

In this case let \(a \in A_j\) be a given point. The equality (4.11) implies

\[
C_j = \sum_{k=1}^{t} s_k C_k \sum_{b \in A_k} l_{ab}, \quad (l_{ab}) = L, \quad j = 1, \ldots, t. \quad (4.15)
\]
Proposition 4.4 If $A_j, A_k$ are two $G$-orbits then the inner sum $\sum_{b \in A_k} l_{ab}$ in (4.15) does not depend on the choice of $a \in A_j$.

**Proof** Note that

$$S_a, A_k = \sum_{b \in A_k} l_{ab} = 1_a^T L 1_{A_k},$$

where $1_a$ and $1_{A_k}$ are vector-columns corresponding to the characteristic functions of the sets $\{a\}$ and $A_k$ respectively. For $L$, given by (4.9) and considered as a kind of resolution [see (4.22) below], we may assert in view of (3.7) that

$$L = \sum_{A \in \text{Orb}} h_A(q) M_A,$$

$h_A(q)$ being some rational functions depending also on $w, \overline{w}$ (see below the conjugate matrix $T^{-1} L T$). From (3.10) we know that $\Gamma$- and, consequently, $G$-action commute with the multiplication by each $M_A$, hence, by $L$. Then for $g \in G$

$$S_a, A_k = 1_a^T L 1_{A_k} = (g^{-1} 1_a^T) g L g 1_{A_k} = 1_{g a}^T L 1_{A_k} = S_{g a, A_k},$$

since $1_{A_k}$ corresponds to the characteristic function of the set $A_k$, which is $G$-orbit. Since $g a, g \in G$, run over $G$-orbit $A_j$ the proposition is proved. $\square$

In what follows we use notation $F_{jk} = S_a, A_k$ for any choice $a \in A_j$. Let us conjugate $L$ by the transition matrix $T$ as introduced in Theorem 3.7. In view of Theorem 3.7 we obtain

$$F_{jk} = 1_a^T T T^{-1} L T T^{-1} 1_{A_k} = 1_a^T T T^{-1} Q T^{-1} T^{-1} T^{-1} T T^{-1} 1_{A_k}$$

$$= 1_a^T T \text{ diag} \left( \frac{P_x(q)}{w P_{X}(q) - w P_{X}(q)} \right) T^{-1} 1_{A_k}$$

$$= \sum_{c=0}^{r-1} \frac{G_{jk}^c P_c(q)}{w P_{X}(q) - w P_{X}(q)}, \quad P_0(q) = P_X(q),$$

for real algebraic numbers

$$G_{jk}^c = \sum_{x: P_x(q) = P_c(q)} \sum_{b \in A_k} t_{ax} t_x^{-1} b, \quad (t_{ax}) = T, \quad (t_x^{-1}) = T^{-1}, \quad a \in A_j, \quad (4.16)$$

since there are at most $r = \text{rk}_Z R = |\text{Orb}|$ different eigenpolynomials $P_x(q)$ in (3.9).

Thus,

$$F_{jk} = F_{jk}(q, w, \overline{w}) = \sum_{c=0}^{r-1} \frac{G_{jk}^c P_c(q)}{w P_{X}(q) - w P_{X}(q)}. \quad (4.17)$$
System (4.15) now reads

\[ C_j = \sum_{k=1}^{t} F_{jk}s_k C_k, \quad j = 1, \ldots, t. \quad (4.18) \]

For the square matrix \( F = F(w) = (F_{jk}) \) of order \( t \), the positive diagonal matrix \( S = \text{diag}(s_1, \ldots, s_t) \), and the positive vector-column \( e = (C_k) \) we consequently have

\[ e = FS\alpha. \quad (4.19) \]

Summarizing the arguments in this section we thus have proved the following

**Theorem 4.5** Let \( \Gamma \)-landscape be symmetric and \( G \)-invariant. Then the dominant eigenvalue \( w \) of problem (3.2) satisfies the equation

\[ \det(F(w) - S^{-1}) = 0. \quad (4.20) \]

In view of (4.17) this is an algebraic equation of degree at most \( t \cdot r = t \cdot \text{rk}_Z R(X, d, \Gamma) \) with coefficients depending on \( q \).

Now consider the simplest case when

\[ X = A_0 \sqcup A_1, \]

which we called *two-valued fitness landscape* in Semenov and Novozhilov (2016).

**Corollary 4.6** In conditions of Theorem 4.5 let the fitness function \( w \) have 2 values, \( w \) on \( A_0 \) and \( w + s \) on \( A_1 = X \setminus A_0 \), where \( s > 0 \), \( A_1 \) is a single \( G \)-orbit.

Then the equation for \( w \) takes the form

\[ r - 1 \sum_{c=0}^{t} \frac{G^c_{11} P_c(q)}{wP_X(q) - w P_c(q)} = \frac{1}{s}, \quad P_0(q) = P_X(q). \quad (4.21) \]

Finally, since Eq. (4.20) in principle allows one to find \( w \), we can use it to find the corresponding eigenvector \( \hat{p} \).

**Theorem 4.7** Let a homogeneous \( \Gamma \)-landscape \( (X, d, \Gamma, w) \) be symmetric and \( G \)-invariant, \( G < \Gamma \). Suppose that there is a decomposition \( X = A_0 \sqcup \bigsqcup_{i=1}^{t} A_i \) such that \( A_0 \) is a union of \( G \)-orbits on which \( w(A_0) \equiv w \geq 0 \), and each \( A_i, i = 1, \ldots, t \), is a single \( G \)-orbit on which \( w(A_i) \equiv w + s_i \), where \( s_i > 0 \). Suppose also that in (4.16) all coefficients \( G^c_{jk} \geq 0 \). Then there exists a solution \( \hat{p} = \hat{p}(q) \) of the generalized Eigen’s problem (3.2) which is constant on \( G \)-orbits.

**Proof** The (maximal) root \( w = w(q) \) of (4.20) provides a non-trivial solution of (4.19). Consider the eigenvalue \( \lambda = 1 \) of the matrix \( FS \) with non-negative entries. It follows from the Perron–Frobenius theorem that we can find a positive solution \( e = (C_k) \),
Let \( k = 1, \ldots, t \) up to the positive scalar factor. Thus, we can determine the projections \( E_{A_i} \hat{p} = C_i E_{A_i} \mathbf{1} \), where \( \mathbf{1} = (1, \ldots, 1)^\top \).

In view of (4.3) solution \( \hat{p} \) of the problem (3.2) can be reconstructed with the help of the formula

\[
\hat{p} = Q(\overline{w} P_X(q) I - w Q)^{-1} \sum_{i=1}^{t} s_i E_{A_i} \hat{p} = \sum_{i=1}^{t} s_i L E_{A_i} \hat{p}.
\] (4.22)

Lemma 4.3 implies that this solution \( \hat{p} \) is \( G \)-invariant.

The conditions (4.13) and (4.14) [which is the same as (3.3)] enable us to determine the multiplication scalar for \( C_k \) and the final expression for \( \hat{p} \). The proof is complete.

\( \square \)

**Remark 4.8** For a two-valued symmetric and \( G \)-invariant landscape \((X, d, \Gamma, w)\) satisfying the conditions of Theorem 4.7, a solution of (3.2) can be significantly simplified if \( w = 0 \) (biologically, this is the case of lethal mutations). For instance, see Semenov and Novozhilov (2016, Example 4.8).

### 5 Three examples: polygonal, hyperoctahedral, and \( p \)-adic landscapes

In this section we show how the general theory of Sects. 3 and 4 can be applied to some specific finite metric spaces \( X \). Namely, we consider first the polygonal mutational landscape and then turn to analysis of the hyperoctahedral one. Two more detailed examples of the hypercube and regular simplex can be found in Semenov and Novozhilov (2016). We would like to remark that although the classical Eigen’s model is almost exclusively based on the geometry of binary cube, other mutational landscapes can be biologically relevant. For instance, in Semenov and Novozhilov (2016) we argued that the simplicial landscape is a natural description of the switching of the antigenic variants for some bacteria. Our final example is the \( p \)-adic landscape. While the biological relevance of such an example is unclear at this point, it explicitly shows that the abstract constructions of Sect. 3 are not mere trivial generalizations of the properties of algebras related to the distance-regular graphs in the sense of van Dam et al. (2016).

#### 5.1 Polygonal landscapes

**5.1.1 Preliminaries**

Let \( X_l \) be the 0-skeleton of a regular \( l \)-gon with unit edges on a plane. We will assume that \( l \geq 3 \), the case \( l = 2 \) can be treated either directly, or as the case of 1-dimensional simplex or the case of 1-dimensional cube (i.e., a segment). Both cases were investigated in Semenov and Novozhilov (2016).

We will enumerate the points of \( X_l \) by numbers of the set \( X_l = \{0, 1, \ldots, l-1\} \) with the fixed point 0 and the counterclockwise enumeration of vertices. It is convenient
to consider these numbers as elements of the cyclic group \( \mathbb{Z}/l\mathbb{Z} \), that is, consider the integer numbers modulo \( l \). Sometimes we will refer to the classical geometric interpretation of \( X_l \) as the set of all roots of unit of degree \( l \) on the complex plane \( \mathbb{C} \), i.e., \( X_l \cong \{ 1, e^0, e, \ldots , e^{l-1} \} \), where \( e = e^{2\pi i/l} \), so that \( k \mod l \leftrightarrow e^k = e^{2\pi ki/l} \).

The metric \( d \) is the so called edge metric on \( X_l \): the distance \( d(k, j) \) between \( k \) and \( j \) is the minimal number of edges of the regular \( l \)-gon connecting successively \( k \) and \( j \). If \( X_l \cong \mathbb{Z}/l\mathbb{Z} \) then \( d(k, j) = \min \{|k - j|, l - |k - j|\} \).

For this metric if the cardinality \( l = |X_l| = 2N + 1 \) is odd then \( N = \text{diam} \ X_l \) and there are two points, namely, \( N \) and \( N + 1 \), such that \( d(0, N) = d(0, N + 1) = N \). If the cardinality \( l = |X_l| = 2N \) is even then \( N = \text{diam} \ X_l \) and there is the unique point \( N \) for which \( d(0, N) = N \). Moreover, we have the following distance polynomial:

\[
P_{X_l}(q) = \begin{cases} 
q^N + 2 \sum_{k=1}^{N} (1 - q)^k q^{N-k}, & l = 2N + 1 \text{ is odd}, \\
q^N + 2 \sum_{k=1}^{N-1} (1 - q)^k q^{N-k} + (1 - q)^N, & l = 2N \text{ is even}.
\end{cases}
\] (5.1)

It is well known that the group \( \text{Iso}(X_l) \cong D_l \) where \( D_l \) is a dihedral subgroup of order \( 2l \) which acts transitively on \( X_l \). The cyclic subgroup \( C_l < D_l \) acts also transitively on \( X_l \) but the triple \( (X_l, d, D_l) \) is symmetric in the sense of Definition 3.3 meanwhile the triple \( (X_l, d, C_l) \) is not.

In what follows we consider only the symmetric polygonal landscapes \( (X_l, d, \Gamma, w) \), \( \Gamma = D_l \). The stabilizer \( \Gamma_0 = \text{St}_\Gamma(0) \cong \mathbb{Z}/2\mathbb{Z} \). If \( X_l \cong \{ 1, e^0, e, \ldots , e^{l-1} \} \), then the unique non-trivial element of \( \Gamma_0 \) acts as the complex conjugation. For the model \( X_l \cong \mathbb{Z}/l\mathbb{Z} \) it acts by the rule \( k \mapsto l - k \mod l \).

There are exactly \( N = \text{diam}(X_l) \) \( \Gamma_0 \)-orbits in \( X \): \( A_0 = \{ 0 \} \), \( A_k = \{ k, l - k \} \), \( k = 1, \ldots , N - 1 \), and \( A_N = \{ N, N + 1 \} \) when \( l = 2N + 1 \) is odd, \( A_N = \{ N \} \) when \( l = 2N \) is even. In any case each \( A_k = S_k(0) \) is the sphere of radius \( k \) centered at \( 0 \). For the orbital ring \( R_l = R(X_l, d, \Gamma) \) (see Sect. 3.2) it means that \( r = \text{rk}_\mathbb{Z}(X_l, d, \Gamma) = N + 1 \).

Finally, explicit calculations presented in “Appendix B.1” show that the transition matrix \( T \) for this example can be found explicitly as

\[
T = T_l = (t_{ab}) := \left( \cos(2\pi ab/l) - \sin(2\pi ab/l) \right), \quad a, b \in X_l \cong \mathbb{Z}/l\mathbb{Z}.
\]

5.1.2 On single peaked and alternating \( G \)-invariant polygonal landscapes

It is known that the subgroups of the dihedral group \( D_l \) are, up to isomorphism, the following groups: dihedral groups \( D_m \) and cyclic groups \( C_m \) for \( m \) dividing \( l \). Note that \( D_l \cong \mathbb{Z}/2\mathbb{Z} \).

In this section the explicit expression of Eq. (4.21) is given for two-valued fitness landscapes \( (X_l, d, \Gamma = D_l, w) \). Specifically, we consider

1. Single peaked landscapes when \( X = A_0 \cup A_1 \) and \( A_1 \) consists of a single point, say, \( A_1 = \{ 0 \} \). Thus, \( w(0) = w + s \), \( w(x) = w \), \( x \neq 0 \). This landscape is \( G \)-invariant for \( G = \{ 1 \} \).
2. Alternating landscapes for \( l \)-gons with even \( l = 2N \) (see Fig. 3). For these landscapes \( X = A_0 \sqcup A_1 \) where \( A_1 = \{0, 2, 4, \ldots, l - 2\} \) and \( w(A_0) = w, \quad w(A_1) = w + s \). This landscape is \( G \)-invariant for the cyclic group \( G = C_N \).

1. In the case of a single peaked landscape Eq. (4.21) reads

\[
\frac{1}{w - w} + \sum_{c=1}^{N-1} \frac{2P_c(q)}{wP_X(q) - wP_c(q)} + \frac{H P_N(q)}{wP_X(q) - wP_N(q)} = \frac{l}{s}, \quad P_0(q) = P_X(q),
\]

where \( H = 2 \) if \( l = 2N + 1 \) and \( H = 1 \) if \( l = 2N \). The polynomials \( P_c(q) \) are defined in (5.1), (B.2), (B.3).

Indeed, we need to calculate the numbers \( G_{11}^c \) with the help of (4.16). But \( a = 0 \), \( A_1 = \{0\} \) and in view of Theorem B.1 \( t_{0x} = l_{x0}^{(-1)} = 1 \) for the transition matrix \( T \). Hence the result.

In Fig. 4 we present two numerical examples for the considered situation. Note the absence of non-analytical behavior of \( \overline{w} \), i.e., the absence of the error threshold (or phase transition).
Fig. 5 The leading eigenvalue of spectral problem (3.2) on regular $2N$-gon for alternating fitness landscape $w = 1, s = 1$. On the left $N = 50$, on the right $N = 200$

2. In the case of an alternating landscape with $l = 2N$ the straightforward calculation yields $G_{11}^0 = G_{11}^N = 1/2$, $G_{11}^k = 0$ for $0 < k < N$. Then Eq. (4.21) is quadratic:

$$\frac{1}{\bar{w} - w} + \frac{P_N(q)}{\bar{w}P_X(q) - wP_N(q)} = \frac{2}{s}, \quad P_0(q) = P_X(q),$$

or, for $u = w/s$, $\bar{u} = \bar{w}/s$,

$$\frac{1}{\bar{u} - u} + \frac{P_N(q)}{\bar{u}P_0(q) - uP_N(q)} = 2. \tag{5.3}$$

The polynomials $P_0(q), P_N(q) \in \mathbb{Z}[q]$:

$$P_0(q) = q^N + 2 \sum_{k=1}^{N-1} (1 - q)^k q^{N-k} + (1 - q)^N, \quad P_N(q)$$

$$= q^N + 2 \sum_{k=1}^{N-1} (-1)^k (1 - q)^k q^{N-k} + (-1)^N (1 - q)^N.$$

The solution of (5.3) is (the positive square root is taken)

$$\bar{u}(q) = \frac{(2u + 1)(P_0(q) + P_N(q)) + \sqrt{(2u + 1)^2(P_0(q) + P_N(q))^2 - 16u(u + 1)P_0(q)P_N(q)}}{4P_0(q)}, \tag{5.4}$$

numerical examples are given in Fig. 5.

5.2 Hyperoctahedral or dual Eigen’s landscapes

5.2.1 Preliminaries

Let $X_n = \{x_0, \ldots, x_k, \ldots, x_{2n-1-k}, \ldots, x_{2n-1}\}$ be the 0-skeleton of a regular $n$-dimensional hyperoctahedron which is the convex hull of the vertices (in $\mathbb{R}^n = \{(\xi_1, \ldots, \xi_n)\}$).
$x_0 = (1, 0, \ldots, 0), \ x_{2n-1} = (-1, 0, \ldots, 0), \ldots,$

$x_k = (0, \ldots, 0, 1, 0, \ldots, 0) \ (1 \text{ stands on } (k + 1)\text{th position}), \ x_{2n-1-k} = (0, \ldots, 0,\ -1, 0, \ldots, 0), \ldots$

A classical octahedron $(n = 3)$ is represented in Fig. 6.

The metric $d$ is again the edge metric on $X_n$: the distance $d(x_k, x_j)$ is defined as follows

$$d(x_k, x_j) = \begin{cases} 0, & \text{if } k = j, \\ 2, & \text{if } k + j = 2n - 1, \\ 1, & \text{otherwise}. \end{cases}$$

We have the cardinality $l = |X_n| = 2n$ and diam$(X_n) = N = 2$. The distance polynomial is $P_{X_n}(q) = q^2 + (2n - 2)(1 - q)q + (1 - q)^2$.

Since a hyperoctahedron is the dual polytope to the hypercube appearing in the classical Eigen’s model, the isometry group $\Gamma = \Gamma_n = \text{Iso}(X_n)$ is a hyperoctahedral group. $\Gamma_n$ is isomorphic as an abstract group to the Weyl group of the root system of type $B_n$ or $C_n$ and $|\Gamma_n| = 2^n \cdot n!$.

Note that the triple $(X_n, d, \Gamma_n)$ is symmetric in the sense of definition 3.3 and that the stabilizer $\Gamma_0 = \text{St}_{\Gamma_n}(x_0) \cong \text{Iso}(X_{n-1})$. Indeed, viewing $x_0$ and $x_{2n-1}$ as the “north” and “south” poles respectively, we have an isometry action of $\Gamma_0 = \text{St}_{\Gamma}(x_0)$ on the “equatorial” hyperoctahedron $X_{n-1}$ which is 1-sphere of $x_0$ with respect to the metric $d$. Then it is not hard to see that $\Gamma_0 = \text{St}_{\Gamma_n}(x_0)$ is the group of all isometries of $X_{n-1}$.
If \( n \geq 2 \) then there are exactly \( 3 = \text{diam}(X_n) + 1 \) \( \Gamma_0 \)-orbits in \( X_n \); namely, \( A_0 = \{ x_0 \} \), \( A_1 = \{ x_1, \ldots, x_{2n-2} \} \cong X_{n-1} \), and \( A_2 = \{ x_{2n-1} \} \). For the orbital ring \( R_n = R(X_n, d, \Gamma) \) (see Sect. 3.2) it means that \( r = \text{rk}_{\mathbb{Z}}(X, d, \Gamma) = 3 \).

Finally, in “Appendix B.1.1” the details of calculation of the transition matrix \( T \) and the eigenpolynomials of \( Q \) are presented.

### 5.2.2 On single peaked \( G \)-invariant hyperoctahedral landscapes

In this subsection the explicit expression of Eq. (4.21) is given for 2-fitness hyperoctahedral landscapes \((X_n, d, \Gamma_n, w)\).

Consider a single peaked landscape for which \( X_n = A_0 \sqcup A_1 \) and \( A_1 \) consists of a single point, say, \( A_1 = \{ x_0 \} \). Thus, \( w(x_0) = w + s \), \( w(x_k) = w \), \( k = 1, \ldots, 2n - 1 \). This landscape is \( G \)-invariant under the action of the trivial group \( G = \{ 1 \} \).

In the case of a single peaked landscape Eq. (4.21) reads

\[
\frac{G^0_{11}}{w - w} + \frac{G^1_{11}(2q - 1)^2}{wP_X(q) - w(2q - 1)^2} + \frac{G^2_{11}(2q - 1)}{wP_X(q) - w(2q - 1)} = \frac{l}{s},
\]

where \( P_X(q) = p_0(q) = q^2 + (2n-2)q(1-q) + (1-q)^2 \) is the distance polynomial

and, in view of the assertion 2 of Theorem B.3 and (4.16),

\[
G^0_{11} = \frac{1}{2n}, \quad G^1_{11} = \frac{n-1}{2n}, \quad G^2_{11} = \frac{n}{2n} = \frac{1}{2}.
\]

Finally, for the parameters \( u = w/s, \bar{u} = \overline{w}/s \) we obtain the (cubic) equation

\[
\frac{1}{\bar{u} - u} + \frac{(n-1)(2q - 1)^2}{\bar{u}P_X(q) - u(2q - 1)^2} + \frac{n(2q - 1)}{\bar{u}P_X(q) - u(2q - 1)} = 2n.
\]

Numerical illustrations are given in Fig. 7. Note that in this case, different from the polygonal landscapes we considered in Sect. 5.1, the numerical experiments indicate that this model possesses the error threshold.

**Fig. 7** The leading eigenvalue of spectral problem (3.2) on hyperoctahedral landscape for SPL \( w = 1, s = 1 \). On the left \( n = 100 \), on the right \( n = 1000 \).
5.3 $p$-Adic landscapes

Here we develop further Example 3.2. Recall that $p$ is a fixed prime and $n$ is a fixed natural number. Consider the quotient ring $X_{p^n} = \mathbb{Z}/p^n\mathbb{Z} = \{0, 1, \ldots, p^n - 1\}$ (all residues viewed as numbers modulo $p^n$). The set $X_{p^n}$ is equipped with the (scaled) $p$-adic metric $d$: for all $x, y \in X_{p^n}$ we have

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ p^{n-1-k} & \text{if } x \neq y \text{ and } p^k|(x - y), \ p^{k+1} \nmid (x - y). \end{cases}$$

The calculation of the diameter provides $N = \text{diam}(X_{p^n}) = p^{n-1}$. The spheres $S_{p^j}(0)$ of radius $p^j$, $j = 0, \ldots, n - 1$, centered at 0 are the following subsets:

$$S_{p^j}(0) = \{x \in X_{p^n} | p^{n-1-j}|x \text{ and } p^{n-j} \nmid x\}$$

of the cardinalities $f_1 = |S_1(0)| = p - 1$, $f_p = |S_p(0)| = p^2 - p$, $\ldots$, $f_{p^{n-1}} = |S_{p^{n-1}}(0)| = p^n - p^{n-1}$. Hence, we obtain the following formula for the distance polynomial, $X = X_{p^n}$:

$$P_X(q) = q^{p^{n-1}} + \sum_{j=0}^{n-1} (p^{j+1} - p^j)(1 - q)^{p^j}q^{p^{n-1-j}p^j}.$$  

Now let $\Gamma \leq \text{Iso}(X_{p^n})$ be the group of all affine transformations $\gamma_{a,b}: x \to ax + b$ (all numbers should be viewed modulo $p^n$) where $a \in (\mathbb{Z}/p^n\mathbb{Z})^\times$ is an invertible element (modulo $p^n$, i.e. $\gcd(a, p) = 1$) and $b \in \mathbb{Z}/p^n\mathbb{Z}$. Then $|\Gamma| = (p^n - p^{n-1})p^n$. We claim that the triple $(X, d, \Gamma)$ is homogeneous and symmetric (in the sense of Definition 3.3). In fact, if $x, y \in X_{p^n}$, $b = x + y$, and $a = -1$ then $\gamma_{a,b}x = b - x = y$, $\gamma_{a,b}y = b - y = x$. Moreover, we also claim that the stabilizer $\Gamma_0 = \text{St}_{\Gamma}(0)$ consisting of linear transformations $\gamma_{a,0}: x \to ax$ acts transitively on each sphere $S_{p^j}(0)$. Indeed, let $x = p^{n-j-1}x_1$, $y = p^{n-j-1}y_1$, where $\gcd(x_1, p) = \gcd(y_1, p) = 1$. Then $x, y \in S_{p^j}(0)$ and $x_1, y_1$ are invertible modulo $p^n$. Set $a = y_1x_1^{-1}$. Consequently,

$$\gamma_{a}x = ax = y_1x_1^{-1} \cdot p^{n-j-1}x_1 = y_1 \cdot p^{n-j-1} = y.$$  

It follows that there are exactly $n + 1$ $\Gamma_0$-orbits in $X_{p^n}$ or $n + 1$ $\Gamma$-orbits in $X_{p^n} \times X_{p^n}$. Hence, for the orbital ring $R = R(X_{p^n}, d, \Gamma)$ we obtain the rank, $r = \text{rk}_\mathbb{Z}(X_{p^n}, d, \Gamma) = n + 1$. Note that $N = \text{diam}(X_{p^n}) = p^{n-1}$ grows much faster than the rank as $n \to \infty$. 

\[ \text{Springer} \]
Let us consider a specific example of a 2-adic space $X = X_{2^3} = \{0, 1, 2, 3, 4, 5, 6, 7\}$. Then $N = \text{diam}(X) = 4$. The distance matrix is given by

$$
D = \begin{bmatrix}
0 & 4 & 2 & 4 & 1 & 4 & 2 & 4 \\
4 & 0 & 4 & 2 & 4 & 1 & 4 & 2 \\
2 & 4 & 0 & 4 & 2 & 4 & 1 & 4 \\
4 & 2 & 4 & 0 & 4 & 2 & 4 & 1 \\
1 & 4 & 2 & 4 & 0 & 4 & 2 & 4 \\
4 & 1 & 4 & 2 & 4 & 0 & 4 & 2 \\
2 & 4 & 1 & 4 & 2 & 4 & 0 & 4 \\
4 & 2 & 4 & 1 & 4 & 2 & 4 & 0 \\
\end{bmatrix}.
$$

Matrices $Q$ and $T$ are also readily calculated. It is interesting to note that $T$ is the same as in the case of the binary cube $X = \{0, 1\}^3$, i.e., for the classical Eigen’s landscape (Semenov and Novozhilov 2016). Also, we have the following eigenpolynomials of $Q$: the distance polynomial $P_X(q) = q^4 + q^3(1 - q) + 2q^2(1 - q)^2 + 4(1 - q)^4$ and the polynomials $-2q^4 + 13q^3 - 22q^2 + 16q - 4$, $-2q^4 + 5q^3 - 2q^2$, $2q^4 - q^3$ of multiplicities 1, 2 and 4 respectively.

The main goal to include this example in the current presentation is to point out that the abstract theory developed in the previous sections is more general than the case of distance-regular graphs discussed in Example 3.1. Further properties and intricacies of the $p$-adic landscapes will be discussed elsewhere.

### 6 Concluding remarks

The main contribution of the present paper is twofold. First, we introduced a generalized algebraic quasispecies model in which the standard binary hypercube of Eigen’s model is replaced with an arbitrary finite metric space $X$. Second, we showed that if the structure of the fitness landscape is related to the isometry group of $X$ then progress can be made in analytical investigation of the corresponding spectral problem. In particular, we found an explicit form of the algebraic equation for the leading eigenvalue [Eq. (4.20)].

At the same time, there are a number of open questions, which would be interesting to work on using the framework we suggest.

While the equation for $\overline{w}$ is written in the general form, in all the examples we considered here and in Semenov and Novozhilov (2016) we deal with the simplest case of two-valued fitness landscapes, when $X = A_0 \sqcup A_1$. It is important to consider examples with more complicated partition of $X$. For example, the so-called mesa landscapes (Wolff and Krug 2009) have exactly this form.

The error threshold phenomenon (see Fig. 2) was not analyzed in the present text. We remark that the error threshold was proven to exist for a simplicial mutation landscape in Semenov and Novozhilov (2016). It looks plausible to conjecture that for the considered in the present text $m$-gon landscapes the error threshold is absent whereas for the hyperoctahedral mutation landscape it does exist. Given the definition of the generalized quasispecies problem, we now have a more general question to ask:
What are the properties of a finite metric space $X$ that guarantee existence of the error threshold at least for some fitness landscapes $w$?

Finally, more detailed analysis of the connections of the considered spectral problems with the Ising model is necessary. It is generally acknowledged that 2D Ising model is a simplest physically nontrivial model that possesses a phase transition. On the other hand, for the simplicial and hyperoctahedral mutation landscapes we consider in our framework the algebraic equation for the leading eigenvalue has degree 2 and 3 respectively, and the proof of the existence of a phase transition is much simpler (see, e.g., Semenov and Novozhilov 2016). Therefore it would be very interesting to identify the Ising model counterparts of our abstract constructions.

### A Orbital rings

#### A.1 Orbital ring: some properties and examples

Here we provide more details on orbital rings defined in Sect. 3.2.

**Lemma A.1** If the triple $(X, d, \Gamma)$ is symmetric in the sense of Definition 3.3 then matrices $M_A$ are symmetric and commute pairwise. Moreover, there are integer non-negative structural constants $\mu_{AB} = \mu_{BA}$, where $A, B, C \in \text{Orb}$, such that

$$M_AM_B = \sum_{C \in \text{Orb}} \mu_{AB}^C M_C. \quad (A.1)$$

**Proof** Let $(x, z) \in A$. In view of Definition 3.3 there exists an isometry $\gamma = \gamma(x, z) \in \Gamma$ such that $\gamma x = z$ and $\gamma z = x$. Thus, $(z, x) \in A$ and $M_A$ is symmetric.

Moreover, it follows from the definition that for any $x, z \in X$ the corresponding matrix entry

$$(M_AM_B)_{x,z} = \#\{y \in X \mid (x, y) \in A, (y, z) \in B\}. \quad (A.2)$$

On the other hand, for the same transposing isometry $\gamma = \gamma(x, z) \in \Gamma$

$$\#\{y \in X \mid (x, y) \in A, (y, z) \in B\} = \#\{y \in X \mid (z, y) \in B, (y, x) \in A\} =$$
$$= \#\{y y \in X \mid (y z, y y) \in B, (y y, y x) \in A\} = \#\{y y \in X \mid (x, y y) \in B, (y y, z) \in A\}.$$

It follows that $(M_AM_B)_{x,z} = (M_BM_A)_{x,z}$ and $M_AM_B = M_BM_A$.

The pair $(x, z)$ defines a $\Gamma$-orbit $C$. For $g \in \Gamma$ we have the same as in (A.2) non-negative number

$$(M_AM_B)_{gx,gz} = \#\{y \in X \mid (gx, gy) \in A, (gy, gz) \in B\}.$$

Thus, in the product $M_AM_B$ the matrix entries with subindices $(x, z)$ and $(gx, gz)$, belonging to the same $\Gamma$-orbit $C \subset X \times X$, have the same value $\mu_{AB}^C$ for any $g \in \Gamma$. It means by the definition of the orbital matrices $M$ and the constants $\mu$ that $M_AM_B = \sum_{C \in \text{Orb}} \mu_{AB}^C M_C$.
\[ \cdots + \mu_{AB}^C M_C + \cdots. \]

Hence, (A.1) holds for some non-negative integer constants \( \mu_{AB}^C \), depending only on the orbits \( A, B, C \). The lemma is proved. \( \square \)

Moreover, we have proved

**Theorem A.2** All \( \mathbb{Z} \)-linear combinations of \( M_A, A \in \text{Orb} \), compose a commutative unital ring \( R = R(X, d, \Gamma) \) with unity \( M_\Delta = I \) called the orbital ring associated with the symmetric triple \((X, d, \Gamma)\). As \( \mathbb{Z} \)-module \( R \) is free of rank \( \text{rk}_\mathbb{Z}R = |\text{Orb}| = |\Gamma_0 \backslash \Gamma/ \Gamma_0| \).

**Example A.3** Let \( X = \{0, 1\}^2 \) be the binary square with points \( x_0 = [0, 0], x_1 = [0, 1], x_2 = [1, 0], x_3 = [1, 1] \) (binary representation of indices) with the Hamming metric \( d \). Let \( \Gamma \cong D_4 \) (the dihedral group of order 8) be the group of all isometries of \( X \).

Then the triple \((X, d, \Gamma)\) is symmetric.

The set \( \text{Orb} \) consists of three orbits (corresponding to the three orbits, namely, \( d \)-spheres \( \{x_0\} = S_0(x_0), \{x_1, x_2\} = S_1(x_0), \{x_3\} = S_2(x_0) \), of the stabilizer \( \Gamma_0 = \text{St}_\Gamma(x_0) \cong \mathbb{Z}/2\mathbb{Z} \) acting on \( X \) represented by matrices

\[
M_0 = I = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad
M_1 = \begin{bmatrix}
0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0
\end{bmatrix}, \quad
M_2 = \begin{bmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0
\end{bmatrix}.
\]

The multiplication table of these matrices in \( R = R(X, d, \Gamma) \) is as follows:

|   | \( I \)   | \( M_1 \) | \( M_2 \) |
|---|-----------|-----------|-----------|
| \( I \) | \( I \) | \( M_1 \) | \( M_2 \) |
| \( M_1 \) | \( M_1 \) | \( 2I + 2M_2 \) | \( M_1 \) |
| \( M_2 \) | \( M_2 \) | \( M_1 \) | \( I \) |

**Remark A.4** We have not yet applied the triangle inequality \( d(x, z) \leq d(x, y) + d(y, z) \). It can be used for the construction of a graded ring \( \text{gr} R = \text{gr} R(X, d, \Gamma) \).

Consider the following increasing filtration on \( R \):

\( R_{-1} = 0 < Z = R_0 < R_1 < \cdots < R_N = R \) for \( \mathbb{Z} \)-modules \( R_k = \bigoplus_{\deg A \leq k} \mathbb{Z} M_A \).

It follows from the definition and the triangle inequality that \( R_i \cdot R_j \subseteq R_{i+j} \). Hence, we can attach to the triple \((X, d, \Gamma)\) the graded ring

\[
\text{gr} R = \bigoplus_{k=0}^N R_k/R_{k-1}.
\]

For instance, in the above Example A.3 (here \( M_k \in \text{gr}_k R = R_k/R_{k-1} \) is viewed as the corresponding element of \( R_k \) modulo \( R_{k-1} \)): 

\( \text{Springer} \)
A.2 Orbital rings for polygonal landscapes

We use notation of Sect. 5.1. The orbital matrices $M_k = M_{Ak}$ have the following entries: $(M_k)_{ab} = 1$ if $d(a, b) = k$ and $(M_k)_{ab} = 0$ otherwise (for a square $X_4$, see Example A.3).

The multiplication in the commutative ring $R_l = R(X_l, d, D_l)$ is slightly different for the cases of odd and even $l$. In both cases we have $M_0 M_k = M_k$ since $M_0 = I$ is the unity of $R_l$.

1. Case $l = 2N + 1$. It can be checked that

$$M_k^2 = 2M_0 + M_{2k} \text{ if } 0 < 2k \leq N, \quad M_k^2 = 2M_0 + M_{l-2k} \text{ if } N < 2k \leq 2N = l - 1.$$ 

Also we have for $0 < k < j \leq N$:

$$M_k M_j = M_{j-k} + M_{j+k} \text{ if } 0 < j + k \leq N, \quad M_k M_j = M_{j-k} + M_{l-j-k} \text{ if } N < j + k < 2N = l - 1.$$ 

2. Case $l = 2N$. We also check that

$$M_k^2 = 2M_0 + M_{2k} \text{ if } 0 < 2k < N, \quad M_k^2 = 2M_0 + M_{l-2k} \text{ if } N < 2k < 2N = l.$$ 

If $2k = N$ then $M_k^2 = 2M_0 + 2M_N$. If $k = N$ then $M_N^2 = M_0$.

Also we have for $0 < k < j < N$:

$$M_k M_j = M_{j-k} + M_{j+k} \text{ if } 0 < j + k < N, \quad M_k M_j = M_{j-k} + M_{l-j-k} \text{ if } N < j + k < 2N = l.$$ 

If $k + j = N$ we have $M_k M_j = M_{j-k} + 2M_N$. If $j = N$ then $M_k M_N = M_{N-k}$.

**Remark A.5** It follows that the $\mathbb{Z}$-linear mapping $\rho : R_l \rightarrow \mathbb{Z}[2 \cos(2\pi/l)]$ such that $\rho(M_0) = 1, \rho(M_k) = 2 \cos(2\pi k/l)$ ($k = 1, \ldots, N$ for the case $l = 2N + 1$ and $k = 1, \ldots, N-1$ for the case $l = 2N$) and, if $l = 2N, \rho(M_N) = -1$, is, in fact, a ring homomorphism. Note also that $\mathbb{Z}[2 \cos(2\pi/l)] = \mathbb{R} \cap \mathbb{Q}[\varepsilon]$ is the ring of integers of the real field $\mathbb{R} \cap \mathbb{Q}[\varepsilon]$, $\mathbb{Q}[\varepsilon]$ being the cyclotomic field, since $2 \cos(2\pi/l) = \varepsilon + \varepsilon^{-1}$. 

\(\varepsilon\) Springer
A.3 Orbital rings for hyperoctahedral landscapes

Here we use notation of Sect. 5.2. The orbital matrices $M_k = M_{A_k}$ have the following entries: $(M_k)_{ab} = 1$ if $d(a, b) = k$ and $(M_k)_{ab} = 0$ otherwise. For a 2-octahedron (a square) $X_2$ see Example A.3, for a 3-octahedron $X_3$ (a classical one)

$$M_0 = I = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M_1 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The multiplication table of these matrices in $R_n = R(X_n, d, \Gamma_n)$ is as follows ($n \geq 2$):

| $\times$ | $I$ | $M_1$ | $M_2$ |
| --- | --- | --- | --- |
| $I$ | $I$ | $M_1$ | $M_2$ |
| $M_1$ | $M_1$ | $(2n - 2)(I + M_2) + (2n - 4)M_1$ | $M_1$ |
| $M_2$ | $M_2$ | $M_1$ | $I$ |

B Transition matrices for the examples in Sect. 5

B.1 Transition matrix $T = T_l$ and eigenpolynomials of the matrix $Q = Q_l$ for the polygonal landscapes

Let the symmetric triple $(X_l, d, \Gamma)$, $\Gamma = D_l$, be as in the previous subsection and let the columns (rows) of all matrices under consideration be indexed by $0, \ldots, l - 1$ modulo $l$ (since $X_l \cong \mathbb{Z}/l\mathbb{Z}$). Consider the square symmetric matrix of order $l$

$$T = T_l = (t_{ab}) := \left(\cos(2\pi ab/l) - \sin(2\pi ab/l)\right), \quad a, b \in X_l \cong \mathbb{Z}/l\mathbb{Z}. \quad (B.1)$$

**Theorem B.1** The matrix $T$ satisfies the following conditions:

1. The columns $t_b$ of $T$ compose a common eigenbasis for all orbital matrices $M_k$, $k = 0, \ldots, N.$
2. If \( l = 2N + 1 \) is odd then

\[
T^{-1}M_k T = 2 \text{diag}(1, \cos(2\pi k/l), \cos(4\pi k/l), \ldots, \cos(2\pi (l - 1)k/l)), \quad k = 1, \ldots, N.
\]

If \( l = 2N \) is even then

\[
T^{-1}M_k T = 2 \text{diag}(1, \cos(2\pi k/l), \cos(4\pi k/l), \ldots, \cos(2\pi (l - 1)k/l)), \quad k = 1, \ldots, N - 1,
\]

\[
T^{-1}M_N T = \text{diag}(1, -1, 1, -1, \ldots, 1, -1).
\]

3. \( T^{-1}QT = \text{diag}(P_0(q), P_1(q), \ldots, P_{l-1}(q)) \) where \( P_0(q) = P_{X_l}(q) \) is the distance polynomial (5.1). If \( l = 2N + 1 \) is odd then

\[
p_{j}(q) = p_{l-j}(q) = q^N + \sum_{k=1}^{N} 2 \cos(2\pi kj/l) (1 - q)^k q^{N-k}, \quad j = 1, \ldots, N. \quad \text{(B.2)}
\]

If \( l = 2N \) is even then

\[
p_{j}(q) = p_{l-j}(q) = q^N + \sum_{k=1}^{N-1} 2 \cos(2\pi kj/l) (1 - q)^k q^{N-k}
\]

\[
+(-1)^j (1 - q)^N, \quad j = 1, \ldots, N. \quad \text{(B.3)}
\]

4. \( T^2 = lI \). In other words, \( T^{-1} = \frac{1}{l} T \).

Proof 1 and 2. Consider cyclic matrices \( C_k \) with only \( l \) non-trivial entries \( (C_k)_{a,a+k} = 1 \) where subindices are taken modulo \( l \). Note that \( C_0 = M_0 = I \) in any case and \( C_N = M_N \) if \( l = 2N \) is even. In the other cases \( C_k + C_{l-k} = M_k \).

Consider also the vectors \( v_j = (1, \varepsilon^j, \varepsilon^{2j}, \ldots, \varepsilon^{(l-1)j})^T, \quad j \in X_l, \varepsilon = e^{2\pi i/l} \).

Straightforward checking yields

\[
C_k v_j = \varepsilon^{kj} v_j.
\]

Note that the Vandermonde determinant \( \det(\varepsilon^{kj}) \neq 0 \). It follows that

\[
C_k v_j = \varepsilon^{kj} v_j, \quad C_{l-k} v_j = \varepsilon^{-kj} v_j, \quad C_k v_{l-j} = \varepsilon^{-kj} v_{l-j}, \quad C_{l-k} v_{l-j} = \varepsilon^{kj} v_{l-j}
\]

since \( \varepsilon^l = 1 \). If \( l = 2N + 1 \) is odd then the real vector-columns \( t_0 = v_0, t_j = \text{Re} v_j - \text{Im} v_{l-j}, t_{l-j} = \text{Re} v_{l-j} + \text{Im} v_j \) are linearly independent eigenvectors of each \( M_k \) such that
\[ M_0 t_j = t_j, \quad M_k t_j = (\epsilon^{kj} + \epsilon^{-kj})t_j = 2 \cos(2\pi kj/l) t_j, \quad k = 1, \ldots, N. \]

If \( l = 2N \) is even then the vector-columns \( t_0 = v_0, t_N = (1, -1, 1, -1, \ldots, 1, -1)^T \), \( t_j = \text{Re} \, v_j - \text{Im} \, v_{l-j}, t_{l-j} = \text{Re} \, v_{l-j} + \text{Im} \, v_j \) are linearly independent eigenvectors of each \( M_k \) such that
\[
M_0 t_j = t_j, \quad M_k t_j = \epsilon^{kj} + \epsilon^{-kj} t_j = 2 \cos(2\pi kj/l) t_j, \quad k = 1, \ldots, N - 1, \quad M_N t_j = \cos(\pi j) t_j.
\]

In any case the transition matrix \( T \) has the form (B.1). This finishes the proof of the assertions 1 and 2.

3. Recall (3.7) that \( Q = \sum_{k=0}^{N} (1 - q)k q^{N-k} M_k \). In view of the assertion 2
\[
T^{-1} QT = \sum_{k=0}^{N} (1 - q)k q^{N-k} T^{-1} M_k T = \text{diag}(p_0(q), p_1(q), \ldots, p_{l-1}(q))
\]

Comparing the diagonal entries we get the desired result.

4. Straightforward calculations with trigonometric sums. The theorem is proved. \( \square \)

**Remark B.2** Note that Conjecture 3.9 is true for the triple \((X_l, d, \Gamma)\).

**B.1.1 Transition matrix \( T = T_n \) and eigenpolynomials of the matrix \( Q = Q_n \) for the hyperoctahedral landscape**

Let the symmetric triple \((X_n, d, \Gamma_n)\) be as in the previous subsection and let the columns (rows) of all matrices under consideration be indexed by 0, \ldots, 2n - 1 corresponding to \( x_0, \ldots, x_{2n-1} \). Consider the following four matrices of order \( n \) (\( n \geq 2 \)):

\[
T_{00} = \begin{bmatrix}
1 & 1 & 1 & \ldots & 1 \\
1 & -1 & 0 & \ldots & 0 \\
1 & 0 & -1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & 0 & \ldots & -1
\end{bmatrix}, \quad T_{01} = \begin{bmatrix}
1 & \ldots & 1 & 1 & 1 \\
0 & \ldots & 0 & -1 & 1 \\
0 & \ldots & -1 & 0 & 1 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-1 & \ldots & 0 & 0 & 1
\end{bmatrix},
\]

\[
T_{10} = \begin{bmatrix}
1 & 0 & 0 & \ldots & -1 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & -1 & \ldots & 0 \\
1 & -1 & 0 & \ldots & 0 \\
1 & 1 & 1 & \ldots & 1
\end{bmatrix}, \quad T_{11} = \begin{bmatrix}
1 & \ldots & 0 & 0 & -1 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \ldots & 1 & 0 & -1 \\
0 & \ldots & 0 & 1 & -1 \\
-1 & \ldots & -1 & -1 & -1
\end{bmatrix}.
\]

Here the entries of the first row and column of the matrix \( T_{00} \) are equal to 1, the diagonal entries, except for the first one, are equal to \(-1\), and the other entries are trivial. The matrices \( T_{01} \) and \( T_{10} \) are the horizontal and vertical mirror copies of \( T_{00} \), the matrix \( T_{11} \) is the horizontal mirror copy of \( T_{10} \) multiplied by \(-1\).
Consider the square symmetric matrix of order 2\(n\)

\[
T = T_n = \begin{bmatrix} T_{00} & T_{01} \\ T_{10} & T_{11} \end{bmatrix}.
\]  

(B.4)

We also introduce the following four matrices of order \(n\):

\[
K_{00} = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & 1 - n & 1 & \ldots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 - n \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix},
\]

\[
K_{01} = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 - n \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix},
\]

\[
K_{10} = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 - n \\ 1 & 1 & 1 & \ldots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix},
\]

\[
K_{11} = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & 1 & 1 & \ldots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix}.
\]

Here the diagonal entries of the matrix \(K_{00}\), except for \(k_{00} = 1\), are equal to 1\(-n\), and the other entries are equal to 1. The matrices \(K_{01}\) and \(K_{10}\) are the horizontal and vertical mirror copies of \(K_{00}\), the matrix \(K_{11}\) is the horizontal mirror copy of \(K_{10}\) multiplied by \(-1\).

Consider the square symmetric matrix of order 2\(n\)

\[
K = K_n = \begin{bmatrix} K_{00} & K_{01} \\ K_{10} & K_{11} \end{bmatrix}.
\]  

(B.5)

**Theorem B.3** The matrix \(T = T_n\), \(n \geq 2\), satisfies the following conditions:

1. \(\det(T_n) = (−1)^n \cdot 2^n \cdot n^2\), consequently, \(T_n\) is a non-degenerate matrix.
2. \(T_n^{-1} = \frac{1}{2^n} K_n\).
3. The columns \(t_b\) (\(b = 0, \ldots, 2n - 1\)) of \(T\) compose a common eigenbasis for each orbital matrix \(M_0, M_1, M_2\).
4. More precisely,

\[
T^{-1}M_1T = \text{diag}(2n - 2, -2, \ldots, -2, 0, \ldots, 0),
\]

\(n - 1\) times \(n\) times

\[
T^{-1}M_2T = \text{diag}(1, \ldots, 1, -1, \ldots, -1),
\]

\(n\) times \(n\) times

5. Let \(Q = Q_n = \left((1 - q)^{d(x_a,x_b)} \cdot q^{2-d(x_a,x_b)}\right)\). Then

\[
T^{-1}QT = \text{diag}(P_0(q), P_1(q), \ldots, P_{2n-1}(q)) = \\
= \text{diag}(q^2 + (2n - 2)(1 - q)q \\
+ (1 - q)^2, (2q - 1)^2, \ldots, (2q - 1)^2, 2q - 1, \ldots, 2q - 1).
\]

\(n - 1\) times \(n\) times
Proof 1. Subtracting the row 0 from the row \(2n-1\), the row 1 from the row \(2n-2\), ..., the row \(n-1\) from the row \(n\) (note that the subindices of matrix entries range from 0 to \(2n-1\)) we get

\[
\det(T) = \det \begin{bmatrix} T_{00} & T_{01} \\ T_{10} & T_{11} \end{bmatrix} = \det \begin{bmatrix} T_{00} & T_{01} \\ 0 & 2T_{11} \end{bmatrix} = 2^n \det(T_{00}) \det(T_{11}).
\]

Adding the sum of columns \(2, \ldots, n\) to the first column of the matrix \(T_{00}\) we obtain the equality \(\det(T_{00}) = (-1)^{n-1}n\). In a similar way we obtain that \(\det(T_{11}) = -n\). Hence the desired result.

2. Straightforward checking shows that \(T_n K_n = 2nI\). Note that the matrix \(T_{00}\) appears as a transition matrix for a simplicial landscape, see Semenov and Novozhilov (2016, Section 6) for the inverse \(T_{00}^{-1}\) and for more details.

3 and 4. Straightforward calculations show that

\[
M_1 t_0 = (2n-2)t_0, \quad M_1 t_b = -2t_b \quad (b = 1, \ldots, n-1),
\]

\[
M_2 t_b = t_b \quad (b = 0, \ldots, n-1), \quad M_1 t_b = -t_b \quad (b = n, \ldots, 2n-1).
\]

5. Since \(Q = q^2 I + q(1-q)M_1 + (1-q)2M_2\) and in view of 3 we get

\[
T^{-1} Q T = q^2 I + q(1-q) \text{diag}(2n-2, -2, \ldots, -2, 0, \ldots, 0)_{n-1 \text{ times}} \quad n \text{ times}
\]

\[
+ (1-q)^2 \text{diag}(1, \ldots, 1, -1, \ldots, -1)_{n \text{ times}} \quad n \text{ times}
\]

\[
= \text{diag}(q^2 + (2n-2)(1-q)q_{n-1 \text{ times}} \quad n \text{ times}
\]

\[
+ (1-q)^2, (2q-1)^2, \ldots, (2q-1)^2, 2q-1, \ldots, 2q-1)_{n \text{ times}} \quad n \text{ times}
\]

The theorem is proved. \(\square\)

Remark B.4 Note that Conjecture 3.9 is true for the triple \((X_n, d, \Gamma_n)\).

C Resulting tables

In Table 1 several known homogeneous symmetric triples \((X, d, \Gamma)\) of the landscapes are presented. Here \(H_n\) is the hyperoctahedral group (the Weyl group of root system \(B_n\) or \(C_n\)) of order \(2^n n!\), \(S_n\) is the symmetric group, \(D_n\) is the dihedral group of order \(2n\), \(A_5\) is the alternating group of order 60, \(D_1 \cong S_2\), \(\Gamma_0 = \text{St}(x), x \in X\). For regular polytopes \(P\) the metric space \((X, d)\) consists of the set \(X = P^{(0)}\) of vertices, the metric \(d\) is the edge metric (see Example 3.1).

In Table 2 the eigenpolynomials and their multiplicities (in brackets) of the matrix \(Q\) are given. The first one is always the (leading) distance polynomial of multiplicity 1.
Table 1  Examples of the homogeneous symmetric triples \((X, d, \Gamma)\)

| Landscape                | \(X\)       | \(d\)                      | \(l = |X|\) | \(\text{diam } X\) | \(\Gamma\) | \(\Gamma_0\) | \(r = \text{rk}_\mathbb{Z}R(X, d, \Gamma)\) |
|-------------------------|-------------|----------------------------|------------|-------------------|------------|------------|-----------------------------------------------|
| 1. Hypercubic, or Eigen’s | \(\{0, 1\}^n\) | Hamming (edge) metric      | \(2^n\)   | \(n\)            | \(H_n\)    | \(S_n\)    | \(n + 1\)                                    |
| 2. Simplicial            | \(X_n\)     | Edge metric                | \(n + 1\) | 1                 | \(S_{n+1}\) | \(S_n\)    | 2                                            |
| 3. Polygonal            | \(X_{2n}\)  | Edge metric                | \(2n\)    | \(n\)            | \(D_{2n}\) | \(D_1\)    | \(n + 1\)                                    |
| 4. Polygonal            | \(X_{2n+1}\) | Edge metric                | \(2n + 1\) | \(n\)            | \(D_{2n+1}\)| \(D_1\)    | \(n + 1\)                                    |
| 5. Hyper-octahedral     | \(X_n\)     | Edge metric                | \(2n\)    | 2                 | \(H_n\)    | \(H_{n−1}\) | 3                                            |
| 6. Dodeca-hedral        | \(X = p^{(0)}\) | Edge metric                | 20         | 5                 | \(A_5 \times S_2\) | \(D_3\)    | 6                                            |
| 7. Icosa-hedral         | \(X = p^{(0)}\) | Edge metric                | 12         | 3                 | \(A_5 \times S_2\) | \(D_5\)    | 4                                            |
| Landscape                        | $X$               | Eigenpolynomials of $Q$ (their multiplicities)                                                                 |
|---------------------------------|-------------------|---------------------------------------------------------------------------------------------------------------|
| 1. Hypercubic, or Eigen’s       | $X_n$             | $P_X(q) \equiv 1$, (1),                                                                                     |
|                                 |                   | $P_j(q) = (2q - 1)^j \binom{n}{j}$, $j = 1, \ldots, n$                                                      |
| 2. Simplicial                   | $X_n$             | $P_X(q) = q + n(1 - q)$, (1),                                                                               |
|                                 |                   | $P_1(q) = 2q - 1$, (n)                                                                                      |
| 3. Polygonal                   | $X_n$, $n = 2N$   | $P_X(q) = q^N + 2 \sum_{k=1}^{N-1} (1 - q)^k q^{N-k} + (1 - q)^N$, (1),                                         |
|                                 |                   | $P_j(q) = q^N + \sum_{k=1}^{N-1} 2 \cos(2\pi kj/l) (1 - q)^k q^{N-k} + (-1)^j (1 - q)^N$, (2),               |
|                                 |                   | $j = 1, \ldots, N - 1,$                                                                                     |
|                                 |                   | $P_N(q) = q^N + 2 \sum_{k=1}^{N-1} (-1)^k (1 - q)^k q^{N-k} + (-1)^N (1 - q)^N$, (1),                      |
| 4. Polygonal                   | $X_n$, $n = 2N + 1$| $P_X(q) = q^N + 2 \sum_{k=1}^{N} (1 - q)^k q^{N-k}$, (1),                                                      |
|                                 |                   | $P_j(q) = q^N + \sum_{k=1}^{N} 2 \cos(2\pi kj/l) (1 - q)^k q^{N-k}$, (2),                                   |
|                                 |                   | $j = 1, \ldots, N$                                                                                          |
| Landscape       | X     | Eigenpolynomials of $Q$ (their multiplicities) |
|-----------------|-------|-----------------------------------------------|
| 5. Hyperoctahedral | $X_n$ | $P_X(q) = q^2 + (2n - 2)(1 - q)q + (1 - q)^2$, $1$, $P_1(q) = (2q - 1)^2$, $(n - 1)$, $P_2(q) = 2q - 1$, $(n)$ |
| 6. Dodecahedral | $X$   | $P_X(q) = 2q^4 - 4q^3 + 4q^2 - 2q + 1$, $1$, $P_1(q) = (-2q^4 + 4q^3 + q^2 - 3q + 1)(2q - 1)$, $4$, $P_2(q) = (3q^4 - 6q^3 + 6q^2 - 3q + 1 + \sqrt{5}(q^4 - 2q^3 + 2q^2 - q)) \times (2q - 1)$, $3$, $P_3(q) = (3q^4 - 6q^3 + 6q^2 - 3q + 1 - \sqrt{5}(q^4 - 2q^3 + 2q^2 - q)) \times (2q - 1)$, $3$, $P_4(q) = (3q^2 - 3q + 1)(2q - 1)^2$, $4$, $P_5(q) = (2q - 1)^2$, $5$ |
| 7. Icosahedral  | $X$   | $P_X(q) = -2q^2 + 2q + 1$, $1$, $P_1(q) = (q^2 - q + 1 + \sqrt{5}(q^2 - q))(2q - 1)$, $3$, $P_2(q) = (q^2 - q + 1 - \sqrt{5}(q^2 - q))(2q - 1)$, $3$, $P_3(q) = (2q - 1)^2$, $5$ |
References

Baake E, Gabriel W (1999) Biological evolution through mutation, selection, and drift: an introductory review. In: Stauffer D (ed) Annual reviews of computational physics VII. World Scientific, Singapore, pp 203–264
Baake E, Georgii H-O (2007) Mutation, selection, and ancestry in branching models: a variational approach. J Math Biol 54(2):257–303
Baake E, Wagner H (2001) Mutation-selection models solved exactly with methods of statistical mechanics. Genet Res 78(1):93–117
Biggs N (1993) Algebraic graph theory. Cambridge University Press, Cambridge (second edition, 1993)
Bratus AS, Novozhilov AS, Semenov YS (2014) Linear algebra of the permutation invariant Crow–Kimura model of prebiotic evolution. Math Biosci 256:42–57
Bratus AS, Novozhilov AS, Semenov YS (2017) Rigorous mathematical analysis of the quasispecies model: from Manfred Eigen to the recent developments. arXiv: 1712.03855
Brouwer AE, Cohen AM, Neumaier A (1989) Distance-regular graphs. Springer, Berlin
Brown KS (1982) Cohomology of groups, vol 87. Springer, Berlin
Cerf R, Dalmaj J (2016a) The quasispecies distribution. arXiv preprint arXiv:1609.05738
Cerf R, Dalmaj J (2016b) Quasispecies on class-dependent fitness landscapes. Bull Math Biol 78(6):1238–1258
Coxeter HSM (1973) Regular polytopes. Courier Corporation, Chelmsford
de la Harpe P (2000) Topics in geometric group theory. University of Chicago Press, Chicago
Dress AWM, Rumschitzki DS (1988) Evolution on sequence space and tensor products of representation spaces. Acta Appl Math 11(2):103–115
Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwis-senschaften 58(10):465–523
Eigen M, McCaskill J, Schuster P (1988) Molecular quasi-species. J Phys Chem 92(24):6881–6891
Feit W (1982) The representation theory of finite groups, vol 2. Elsevier, Amsterdam
Hermisson J, Redner O, Wagner H, Baake E (2002) Mutation-selection balance: ancestry, load, and maximum principle, Theor Popul Biol 62(1):9–46
Jain K, Krug J (2007) Adaptation in simple and complex fitness landscapes. In: Battolla U, Porto M, Eduardo Roman H, Vendruscolo M (eds) Structural approaches to sequence evolution, chap 14. Springer, Berlin, pp 299–339
Kirillov AA (1976) Elements of the theory of representations, vol 145. Springer, Berlin
Leuthäusser I (1986) An exact correspondence between Eigen’s evolution model and a two-dimensional Ising system. J Chem Phys 84(3):1884–1885
Leuthäusser I (1987) Statistical mechanics of Eigen’s evolution model. J Stat Phys 48(1):343–360
Onsager L (1944) Crystal statistics. I. A two-dimensional model with an order-disorder transition. Phys Rev 65(3–4):117
Rumschitzki DS (1987) Spectral properties of Eigen evolution matrices. J Math Biol 24(6):667–680
Saakian DB, Hu CK (2006) Exact solution of the Eigen model with general fitness functions and degradation rates. Proc Natl Acad Sci USA 103(13):4935–4939
Schuster P (2015) Quasispecies on fitness landscapes. In: Domingo E, Schuster P (eds) Quasispecies: from theory to experimental systems, Current topics in microbiology and immunology. Springer, Berlin, pp 61–120
Semenov Y (1994) Rings associated with hyperbolic groups. Commun Algebra 22(15):6323–6347
Semenov YS, Novozhilov AS (2015) Exact solutions for the selection-mutation equilibrium in the Crow–Kimura evolutionary model. Math Biosci 266:1–9
Semenov YS, Novozhilov AS (2016) On Eigen’s quasispecies model, two-valued fitness landscapes, and isometry groups acting on finite metric spaces. Bull Math Biol 78(5):991–1038
Serre J-P (1996) Linear representations of finite groups, vol 42. Springer, Berlin
Stadler PF, Happel R (1999) Random field models for fitness landscapes. J Math Biol 38(5):435–478
Stadler PF, Tinhofer G (1999) Equitable partitions, coherent algebras and random walks: applications to the correlation structure of landscapes. Match 40:215–261
Swetina J, Schuster P (1982) Self-replication with errors: a model for polynucleotide replication. Biophys Chem 16(4):329–345
Thompson CJ (1972) Mathematical statistical mechanics. Macmillan, New York
van Dam ER, Koolen JH, Tanaka H (2016) Distance-regular graphs. Electron J Comb. Dynamic Survey DS22
Wiehe T (1997) Model dependency of error thresholds: the role of fitness functions and contrasts between
the finite and infinite sites models. Genet Res 69(02):127–136
Wilke CO (2005) Quasispecies theory in the context of population genetics. BMC Evol Biol 5(1):44
Wolff A, Krug J (2009) Robustness and epistasis in mutation-selection models. Phys Biol 6(3):036007