# Applications of Diagonalization

Hsiu-Hau Lin
hsiuhau@phys.nthu.edu.tw
(Apr 12, 2010)

The notes cover applications of matrix diagonalization (Boas 3.12).

## • Quadratic curves

Consider the quadratic curve,

$$5x^2 - 4xy + 2y^2 = 30. \tag{1}$$

It can be casted into the matrix form and then brought into diagonal form,

$$(x, y) \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 30 \quad \rightarrow \quad (x', y') \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = 30. \tag{2}$$

The similarity transformation brings the quadratic curve into the canonical form $x'^2 + 6y'^2 = 30$. This is clearly an ellipse with principle axes coincide with the eigenvectors.

## • Harmonic oscillators with equal mass

Consider three identical springs attached to two equal masses in series. The elastic potential energy is

$$V(x, y) = \frac{1}{2}kx^2 + \frac{1}{2}k(x - y)^2 + \frac{1}{2}ky^2. \tag{3}$$

The corresponding equations of motions are

$$m\ddot{x} = -\partial V/\partial x = -2kx + ky, \tag{4}$$
$$m\ddot{y} = -\partial V/\partial y = kx - 2ky. \tag{5}$$

Assume the solutions are $x = x_0 e^{i\omega t}$ and $y = y_0 e^{i\omega t}$, the equations can be written in matrix form,

$$\lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{6}$$

where $\lambda = m\omega^2/k$. The potential matrix is symmetric arisen from Newton's third law. Following the standard recipe for matrix diagonalization, the eigenvalues are $\lambda = 1, 3$ corresponding to the characteristic frequencies

$$\omega_1 = \sqrt{\frac{k}{m}}, \qquad \omega_2 = \sqrt{\frac{3k}{m}}. \tag{7}$$

The eigenvectors for the coupled harmonic oscillators are

$$\boldsymbol{r}_1 = (1, 1), \qquad \boldsymbol{r}_2 = (1, -1). \tag{8}$$

The in-phase oscillation mode has a smaller frequency $\omega_1$ while the out-of-phase mode has a larger frequency $\omega_2 = \sqrt{3}\omega_1$.

## • Harmonic oscillators with different masses

Consider the same setup but the spring constants and the masses are not equal: $2k$, $2m$, $6k$, $3m$, $3k$. The potential energy is

$$
\begin{aligned}
V(x, y) &= \frac{1}{2}(2k)x^2 + \frac{1}{2}(6k)(x - y)^2 + \frac{1}{2}(3k)y^2 \\
&= \frac{1}{2}k(8x^2 - 12xy + 9y^2),
\end{aligned} \tag{9}
$$

and the kinetic energy is

$$T = \frac{1}{2}(2m)\dot{x}^2 + \frac{1}{2}(3m)\dot{y}^2 = \frac{1}{2}m(2\dot{x}^2 + 3\dot{y}^2). \tag{10}$$

There are two different ways to solve the problem. Let's try the easier one first. The equations of motion are

$$
\begin{aligned}
-2m\omega^2 x &= -k(8x - 6y), \tag{11} \\
-3m\omega^2 y &= -k(-6x + 9y). \tag{12}
\end{aligned}
$$

Divide each equation by its mass and write the equations in matrix form,

$$\lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{13}$$

with $\lambda = m\omega^2/k$. The eigenvalues and the eigenvectors are

$$
\begin{aligned}
\lambda = 1, \quad \boldsymbol{r} = (1, 1); \tag{14} \\
\lambda = 6, \quad \boldsymbol{r} = (3, -2). \tag{15}
\end{aligned}
$$

The main problem with this approach is that the eigenvectors are not orthogonal anymore. Sometimes, this is not a problem and I would recommend this approach. However, there are times that the orthogonality is important and we need another more advanced approach.

Let's write both the potential and the kinetic energy in matrix form,

$$V = \tfrac{1}{2}k\, \boldsymbol{r}^T V \boldsymbol{r} \quad \text{where} \quad V = \begin{pmatrix} 8 & -6 \\ -6 & 9 \end{pmatrix}, \tag{16}$$

$$T = \tfrac{1}{2}k\, \boldsymbol{r}^T T \boldsymbol{r} \quad \text{where} \quad T = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}. \tag{17}$$

The equation of motion can be written in matrix form,

$$\lambda T \boldsymbol{r} = V \boldsymbol{r}. \tag{18}$$

Since the matrix $T$ is positive definite, we can define its square-root matrix

$$T^{\frac{1}{2}} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{3} \end{pmatrix}. \tag{19}$$

Perform a simple scaling transformation to new coordinates,

$$\boldsymbol{R} = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{3} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = T^{\frac{1}{2}} \boldsymbol{r}. \tag{20}$$

The equation of motion can be rewritten in the standard form,

$$\lambda \boldsymbol{R} = T^{-\frac{1}{2}} V T^{-\frac{1}{2}} \boldsymbol{R}. \tag{21}$$

The matrix $T^{-\frac{1}{2}} V T^{-\frac{1}{2}}$ is now symmetric and all nice properties are recovered.

## • Quasi-species equations

In addition to the basic examples in the textbook, I would like to share with more advanced applications for evolutionary dynamics. To describe the evolution of frequencies for different species in the fitness landscape, quasi-species equations capture the essential ingredient,

$$\frac{dx_i}{dt} = \sum_j x_j f_j q_{ji} - \phi x_i, \tag{22}$$

where $x_i$ and $f_j$ are the frequency and the fitness for the $j$-th sequence and $\phi = \sum_i x_i f_i$ is the average fitness for all sequences. The mutation probability

from the $j$-th sequence to the $i$-th sequence is denoted by $q_{ji}$, satisfying the sum rule $\sum_i q_{ji} = 1$.

Assuming point mutation is the dominant process, the mutation matrix takes the simple form

$$q_{ji} = u^{d_{ji}}(1 - u)^{L - d_{ji}}, \tag{23}$$

where $u$ is the probability for point mutation and $d_{ji}$ is the distance between two sequences (minimal number of point mutations to bring one sequence into another). Since $d_{ji} = d_{ij}$, the mutation matrix is also symmetric $q_{ji} = q_{ij}$.

## ● Effective Hamiltonian

The quasi-species equations can be described by an effective Hamiltonian after a gauge transformation,

$$\Psi_i(t) = \sqrt{f_i} x_i(t) e^{W(t)}, \tag{24}$$

with $\dot{W}(t) = \phi(t)$. Making use of the normalization condition $\sum_i x_i = 1$, it is straightforward to show that

$$e^{W(t)} = \sum_i \frac{1}{\sqrt{f_i}} \Psi_i(t). \tag{25}$$

Thus, the inverse gauge transformation is

$$x_i(t) = \frac{1}{\sqrt{f_i}} \Psi_i(t) e^{-W(t)} = \left( \frac{1}{\sum_j \Psi_j / \sqrt{f_j}} \right) \frac{1}{\sqrt{f_i}} \Psi_i(t). \tag{26}$$

After some algebra, one can derive the dynamical equations for $\Psi_i(t)$,

$$\frac{d\Psi_i}{dt} = -\sum_j H_{ij} \Psi_j, \qquad H_{ij} = -\sqrt{f_i f_j} q_{ji}. \tag{27}$$

This form is the same as the usual Schrödinger equation if one replaces $t \rightarrow -it$ and sets $\hbar = 1$. The general solution for the "imaginary-time" Schrödinger equation is

$$\Psi_i(t) = \sum_n c_n \Phi_i^n e^{-E_n t}, \tag{28}$$

where $E_n$ and $\Phi_i^n$ are the eigenvalues and eigenvectors of the effective Hamiltonian $H_{ij}$. In the infinite-time limit, only the ground state will survive

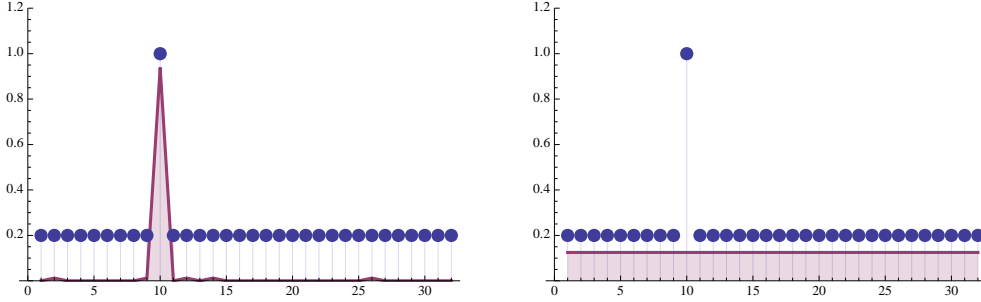$$\lim_{t \to \infty} \Psi_i(t) \longrightarrow c_0 \Phi_i^0 e^{-E_0 t}. \tag{29}$$

Figure 1: For the single-peak fitness landscape, there exists a mutation threshold $u_c$. For $u < u_c$, frequency profile is localized near the fitness peak. On the other hand, for $u > u_c$, an extended state (not necessarily uniform) emerges and the notion of quasi-species no longer exists.

The frequency in the infinite-time limit thus only depends on the ground-state wave function $\Phi_i^0$,

$$x_i^* = \lim_{t \to \infty} x_i(t) = \left( \frac{1}{\sum_j \Phi_j^0 / \sqrt{f_j}} \right) \frac{1}{\sqrt{f_i}} \Phi_i^0. \qquad (30)$$

Therefore, to compute the survival frequencies, we only need to find the ground state of the effective Hamiltonian.

Consider the simplest fitness landscape as shown in Fig. 1. There exists a sharp peak with maximum fitness $f_0 = f_M$ and the background fitness $f_i = f < f_M$ for all sequences $i \neq 0$. Note that the effective Hamiltonian is real and symmetric,

$$H_{ij} = -\sqrt{f_i f_j} u^{d_{ji}} (1 - u)^{L - d_{ji}}. \qquad (31)$$

Even though the single-peak fitness landscape is simple, the size of the sequence space is huge $N_s = 2^L$. Following Chun-Chung's suggestion, the ground state must be in the "s-wave" sector with a much small size $L$ only,

$$\Psi_i^{0s} = (1, 0, 0, ..., 0), \quad \Psi_i^{1s} = \sqrt{\frac{1}{L}} (0, 1, 1, .., 1, 0, 0, ..0),$$

$$\Psi_i^{2s} = \sqrt{\frac{2}{L(L-1)}} (0, 0, 0, .., 0, 1, 1, ...., 1, 0, 0, ..., 0), ... \qquad (32)$$

where the non-zero components in $\Psi_i^{1s}$ are with the distance $d = 1$ from the peak and similar properties hold for all other $\Psi_i^{ns}$. One can construct the $L \times L$ Hamiltonian in the s-wave sector

$$H_{nm} = \sum_{ij} \Psi_i^{ns} H_{ij} \Psi_j^{ms}, \qquad (33)$$

and numerically diagonalize the Hamiltonian to obtain the ground state. The s-wave symmetry helps tremendously and reduces the size of the relevant sequence space from $N_s = 2^L$ to just $L$!

## • Variational method

Here I am going to use a variational approach to estimate the mutation threshold. Choose two variational wave functions

$$\Psi_i^0 = (1, 0, ..., 0), \qquad \Psi_i^1 = \frac{1}{\sqrt{N_s - 1}}(0, 1, 1, ..., 1). \tag{34}$$

Construct the variational wave function for the ground state

$$\Psi_i = c_0 \Psi_i^0 + c_1 \Psi_i^1, \tag{35}$$

with the normalization constraint $c_0^2 + c_1^2 = 1$. The variational energy takes the form

$$
\begin{aligned}
E(c_1, c_2, \lambda) &= \sum_{ij} \Psi_i H_{ij} \Psi_j + \lambda(c_0^2 + c_1^2), \\
&= \sum_{n,m=0,1} c_n H_{nm}^v c_m + \lambda(c_0^2 + c_1^2),
\end{aligned}
\tag{36}
$$

where the variational Hamiltonian is

$$H_{nm}^v = \sum_{ij} \Psi_i^n H_{ij} \Psi_j^m. \tag{37}$$

The matrix elements of the variational Hamiltonian can be computed,

$$
\begin{aligned}
H_{00}^v &= -f_M (1 - u)^L, \tag{38} \\
H_{01}^v &= H_{10}^v = -\sqrt{\frac{f_M f}{N_s - 1}} \left[ 1 - (1 - u)^L \right], \tag{39} \\
H_{11}^v &= -\frac{f}{N_s - 1} \sum_{i,j \neq 0} u^{d_{ij}} (1 - u)^{L - d_{ij}} \\
&= -\frac{f}{N_s - 1} \left( \sum_{i \neq 0, j} - \sum_{i \neq 0, j = 0} \right) u^{d_{ij}} (1 - u)^{L - d_{ij}} \\
&= -f + \frac{f}{N_s - 1} \left[ 1 - (1 - u)^L \right]. \tag{40}
\end{aligned}
$$

Minimizing the variational energy $E(c_1, c_2, \lambda)$ is equivalent to diagonalizing the $2 \times 2$ variational Hamiltonian $H_{nm}^v$ which can be rewritten as

$$H_{nm}^v = C\mathbf{1} + f \begin{pmatrix} -\epsilon & -\Delta \\ -\Delta & \epsilon \end{pmatrix}, \tag{41}$$

where the constant term does not affect the frequencies $x_i(t)$ after the inverse gauge transformation,

$$C = -\frac{1}{2} \left[ f_M q_{00} + f - \frac{f}{N_s - 1}(1 - q_{00}) \right] \tag{42}$$

with $q_{00} = (1 - u)^L$ and will be dropped in the following calculations. The dimensionless parameters $\epsilon$ and $\Delta$ in the variational Hamiltonian are

$$\epsilon(u, r, L) = \frac{1}{2} \left[ r(1 - u)^L - 1 \right] + \frac{1}{2(N_s - 1)} \left[ 1 - (1 - u)^L \right], \tag{43}$$

$$\Delta(u, r, L) = \sqrt{\frac{r}{N_s - 1}} \left[ 1 - (1 - u)^L \right], \tag{44}$$

where $r = f_M/f$ is the relative fitness of the dominant sequence and $N_s = 2^L$ is the size of the sequence space. The eigenvalues of the variational Hamiltonian are

$$E_{\pm} = \pm\sqrt{\epsilon^2 + \Delta^2}. \tag{45}$$

Therefore, the ground-state energy within the variational approximation corresponds to the negative eigenvalue $E = E_-$ with the eigenvector

$$\frac{c_0}{c_1} = \frac{\Delta}{\sqrt{\epsilon^2 + \Delta^2} - \epsilon} = \frac{\sqrt{\epsilon^2 + \Delta^2} + \epsilon}{\Delta}. \tag{46}$$

## • Survival frequency and back mutations

With the coefficients $c_0$ and $c_1$ at hand, the ground state within the variational approximation is

$$\Phi_i^{v0} = \left( c_0, \frac{c_1}{\sqrt{N_s - 1}}, \frac{c_1}{\sqrt{N_s - 1}}, ...., \frac{c_1}{\sqrt{N_s - 1}} \right). \tag{47}$$
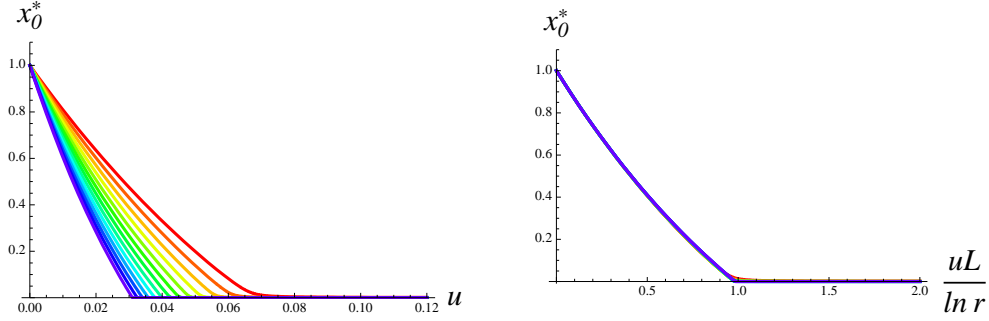
Figure 2: The frequency $x_0$ of the dominant sequence plotted versus the point mutation $u$ and the rescaled mutation $g = uL/\ln r$. The fitness landscape consists of a single sharp peak with relative fitness $r = f_M/f = 2$ and a uniform background for all other sequences. The genome length starts from $L = 10$ (red) to $L = 22$ (purple). From the left panel, it is clear that the crossover (from localized-like to extensive-like) becomes sharper as the genome length increases. For the given relative fitness $r = 2$, all curves for $x_0^*(u, r, L)$ can be collapsed onto the universal scaling function by changing the variable to $g = uL/\ln r$. It is clear that the mutation threshold is $g_c = 1$.

The survival frequency of the dominant sequence can be computed from the variational ground state,

$$
\begin{aligned}
x_0^*(u, r, L) &= \left( \frac{1}{\sum_j \Phi_j^{v0}/\sqrt{f_j}} \right) \frac{1}{\sqrt{f_0}} \Phi_0^{v0} = \frac{c_0}{c_0 + \sqrt{r(N_s - 1)c_1}} \\
&= \frac{\sqrt{\epsilon^2 + \Delta^2} + \epsilon}{\sqrt{\epsilon^2 + \Delta^2} + \epsilon + \sqrt{r(2^L - 1)}\Delta}.
\end{aligned}
\tag{48}
$$

It is convenient to introduce the rescaled variable $g$ for mutations,

$$
g = \frac{uL}{\ln r}.
\tag{49}
$$

The numerical plot for $x_0^*(u, r, L)$ is shown in Fig. 2 and it scales rather nicely when plotting versus the rescaled variable $g = uL/\ln r$ (as long as $L$ is sufficiently large) with a sharp transition at $g_c = 1$.

How good is the "no-back-mutation" assumption? A good indicator is the ratio of the back-mutating rate and the rate to stay on the dominant sequence,

$$
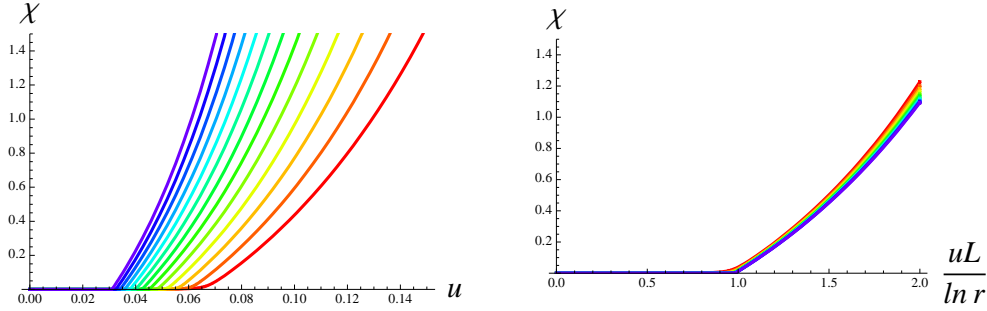\chi(u, r, L) = \frac{\sum_{j \neq 0} x_j^* f_j q_{j0}}{x_0^* f_M q_{00}}.
\tag{50}
$$

Figure 3: The back-mutation ratio $\chi(u, r, L)$ plotted versus the mutation $u$ (left panel) and the rescaled mutation $g = uL/\ln r$ (right panel). The parameters are the same as those in Fig. 2. For large enough genome length, the back-mutation rate is indeed negligibly small when compared with the rate to stay in the dominant sequence. However, beyond the mutation threshold, the ratio $\chi$ grows exponentially, hinting the enhanced back mutations destroy the localized state and prefer the extended state.

Within the variational approximation, all $x_j^* = x_1^*$ for $j \neq 0$ and the summation in the numerator can be carried out without difficulty,

$$\chi(u, r, L) = \frac{x_1^* f_1 (1 - q_{00})}{x_0^* f_0 q_{00}} = \sqrt{\frac{f}{f_M}} \frac{1 - q_{00}}{q_{00}} \frac{c_1}{\sqrt{(N_s - 1)c_0}}. \tag{51}$$

For finite genome length, the back mutations exist as shown in Fig. 3. However, the no-back-mutation assumption is reasonably good. Notice that, plotted with the rescaled mutation $g = uL/\ln r$, the data for different lengths collapse nicely into a universal curve. Above the mutation threshold $g > g_c = 1$, the back mutations grow exponentially and destroy the localized quasi-species state in the sequence space.

## • Infinite genome-length limit

Inspired by the nice collapse onto universal curve, it is interesting to take the infinite genome-length limit and derive the scaling functions for the survival frequency for the dominant species and the back-mutation ratio. Taking $L \to \infty$ limit but keeping the rescaled mutation $g$ and the relative fitness $r$ finite, the exponential factor becomes

$$\lim_{L \to \infty} (1 - u)^L = \lim_{L \to \infty} e^{L \ln(1-u)} = \lim_{L \to \infty} e^{-uL} = e^{-g \ln r} = r^{-g}.$$
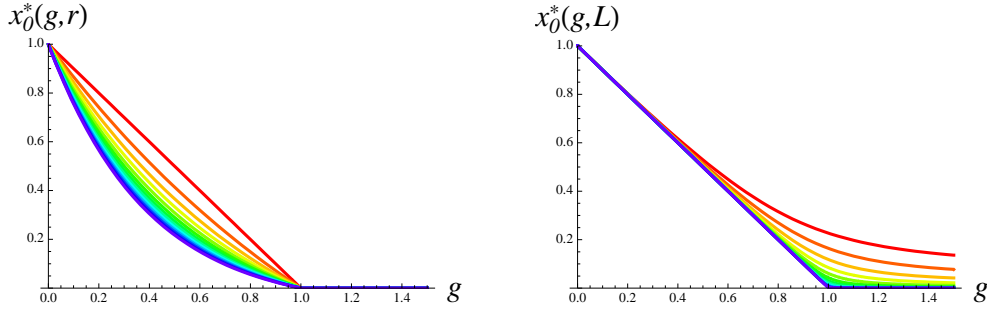
Figure 4: The universal scaling function $x_0^*(g, r)$ in the infinite-length limit (left panel) and $x_0^*(g, L)$ in the continuous limit (right panel). Taking the genome length to infinity, survival frequencies for the dominant sequence with different mutation $u$ and length $L$ collapse onto the universal scaling function $x_0^*(g, r)$. The transition is sharp and the universal scaling functions vary with different relative fitness from $r = 1^+$ (red) to $r = 13$ (purple). On the other hand, in the continuous limit ($u \to 0$ and $r \to 1^+$), the scaling function $x_0^*(g, L)$ is not necessarily sharp anymore since the genome length is finite, varying from $L = 4$ (red) to $L = 16$ (purple). Note that the scaling function in both infinite-length and continuous limit is just the linear function $x_0^*(g) = 1 - g$ below the error threshold $g_c = 1$.

In addition, the off-diagonal term $\Delta$ is negligibly small and the survival frequency for the dominant sequence is greatly simplified,

$$x_0^*(g, r) = \lim_{L \to \infty} x_0^*(u, r, L) = \frac{\frac{1}{2}\left[|r^{1-g} - 1| + (r^{1-g} - 1)\right]}{\frac{1}{2}\left[|r^{1-g} - 1| + (r^{1-g} - 1)\right] + r(1 - r^{-g})}. \quad (52)$$

Since the relative fitness $r$ is greater than unity, the numerator is zero for $g > 1$. It is straightforward to show the universal scaling function is

$$x_0^*(g, r) = \Theta(1 - g) \frac{r^{1-g} - 1}{r - 1}. \quad (53)$$

Magically, this form is exactly the same as that derived under the assumption of no back mutation.

We can also compute the back-mutation ratio in the infinite genome-length limit. Taking $L \to \infty$ but keeping $g$ and $r$ finite, the stay-on probability is $q_{00} = (1 - u)^L \to r^{-g}$. The only tricky term in $\chi(g, r)$ is

$$\lim_{L \to \infty} \frac{c_1}{\sqrt{(N_s - 1)c_0}} = \Theta(g - 1) \frac{1 - r^{1-g}}{\sqrt{r}(1 - r^{-g})}. \quad (54)$$

Therefore, the back-mutation ratio in the infinite genome-length limit is

$$\chi(g,r) = \Theta(g-1)\left(r^{g-1} - 1\right). \tag{55}$$

The universal ratio $\chi(g,r)$ is zero in the localized quasi-species state, i.e. back-mutation rate can be safely ignored.

## • Continuous limit

Following Chun-Chung's idea, the continuous limit for Eigen's model can be obtained by taking $r \to 1$ and $u \to 0$ but holding $g$ and $L$ finite. In the continuous limit,

$$\epsilon_c(g,L) \equiv \lim_{r \to 1} \frac{\epsilon}{r-1} = \frac{1}{2}\left[1 - g + \frac{g}{2^L - 1}\right], \tag{56}$$

$$\Delta_c(g,L) \equiv \lim_{r \to 1} \frac{\Delta}{r-1} = \frac{1}{\sqrt{2^L - 1}}g. \tag{57}$$

After some algebra, the survival frequency of the dominant sequence $x_0^*(g,L)$ in the continuous limit is

$$x_0^*(g,L) = \frac{\sqrt{\epsilon_c^2 + \Delta_c^2} + \epsilon_c}{\sqrt{\epsilon_c^2 + \Delta_c^2} + \epsilon_c + g}. \tag{58}$$

If we take both continuous and the infinite-length limits together, the scaling function becomes

$$\lim_{r \to 1} x_0^*(g,r) = \Theta(1-g)\lim_{r \to 1}\frac{r^{1-g} - 1}{r-1} = \Theta(1-g) \times (1-g). \tag{59}$$

It is interesting that the back mutation disappear in this limit completely,

$$\lim_{r \to 1}\chi(g,r) = \Theta(g-1)\lim_{r \to 1}(r^{g-1} - 1) = 0. \tag{60}$$