Mutative Self-Adaptation on the Sharp and Parabolic Ridge

Silja Meyer-Nieberg   Hans-Georg Beyer

11.01.2006

Contents

Part I

Introduction

Evolution Strategies

- Population-based search heuristics
- Aim: Optimization of fitness function F
- Generally: Continuous search space, $F : \mathbb{R}^N \to \mathbb{R}$
- Here: intermediate $(\mu/\mu_I, \lambda)$-ES
- μ parents create λ offspring by
  - Recombination
  - Mutation
- μ best offspring chosen as next parents

Evolution Strategies

- Population-based search heuristics
- Aim: Optimization of fitness function F
- Generally: Continuous search space, $F : \mathbb{R}^N \rightarrow \mathbb{R}$
- Here: intermediate $(\mu/\mu_I, \lambda)$-ES
- μ parents create λ offspring by
  - Recombination
  - Mutation
- μ best offspring chosen as next parents

Self-adaptation

- To travel with sufficiently high speed
  - Adaptation of mutation strength necessary
- Several methods: E.g.
  - Rechenberg: $1/5$th rule
  - Rechenberg, Schwefel: Self-adaptation
  - Hansen & Ostermeier: CSA, CMA
- Self-Adaptation
  - Adaptation left to the ES itself
  - Mutation strength subject to recombination & mutation

Self-adaptation

- To travel with sufficiently high speed
  - Adaptation of mutation strength necessary
- Several methods: E.g.
  - Rechenberg: $1/5$th rule
  - Rechenberg, Schwefel: Self-adaptation
  - Hansen & Ostermeier: CSA, CMA
- Self-Adaptation
  - Adaptation left to the ES itself
  - Mutation strength subject to recombination & mutation

Recombination

- Intermediate Recombination
- μ parents
  - Mutation strengths
    - Computation of the mean $\langle\sigma\rangle$ over all mutation strengths σ_m
  - Object vectors
    - Computation of the centroid $\langle\mathbf{y}\rangle$ over all object vectors $\mathbf{y}_m$
- Followed by mutation for each offspring

Recombination

- Intermediate Recombination
- μ parents
  - Mutation strengths
    - Computation of the mean $\langle\sigma\rangle$ over all mutation strengths σ_m
  - Object vectors
    - Computation of the centroid $\langle\mathbf{y}\rangle$ over all object vectors $\mathbf{y}_m$
- Followed by mutation for each offspring

Recombination

- Intermediate Recombination
- μ parents
  - Mutation strengths
    - Computation of the mean $\langle\sigma\rangle$ over all mutation strengths σ_m
  - Object vectors
    - Computation of the centroid $\langle\mathbf{y}\rangle$ over all object vectors $\mathbf{y}_m$
- Followed by mutation for each offspring

Mutation

For each offspring l

- Mutation of the mutation strength
  - Mutate the mean $\langle\sigma\rangle$
  - $\sigma_l = \langle\sigma\rangle\zeta$
  - ζ a random variable with $\mathrm{E}[\zeta] \approx 1$
  - Typical choice
    - Log-normal distribution $\zeta \sim \mathrm{e}^{\tau\mathcal{N}(0,1)}$
  - Learning rate τ
- Mutation of the object vector
  - Adding normally distributed random variable
  - ith coordinate $y_i^l = \langle y_i\rangle + \sigma_l\mathcal{N}_i(0,1)$

Mutation

For each offspring l

- Mutation of the mutation strength
  - Mutate the mean $\langle\sigma\rangle$
  - $\sigma_l = \langle\sigma\rangle\zeta$
  - ζ a random variable with $\mathrm{E}[\zeta] \approx 1$
  - Typical choice
    - Log-normal distribution $\zeta \sim \mathrm{e}^{\tau\mathcal{N}(0,1)}$
  - Learning rate τ
- Mutation of the object vector
  - Adding normally distributed random variable
  - ith coordinate $y_i^l = \langle y_i\rangle + \sigma_l\mathcal{N}_i(0,1)$

Ridge Functions

- Ridge functions, orthogonal representation

$$
\begin{aligned}
F_{ridge}(\mathbf{y}) &= y_1 - d\left(\sum_{i=2}^{N} y_i^2\right)^{\alpha/2} \\
&=: x - dR^{\alpha}
\end{aligned}
$$

  - Linear component + embedded sphere
- Sharp ridge: $\alpha = 1$
  - $F_{ridge}(\mathbf{y}) = x - dR$
- Parabolic ridge: $\alpha = 2$
  - $F_{ridge}(\mathbf{y}) = x - dR^2$
- Meliorating
  - Decrease R and/or increase x

Ridge Functions

- Ridge functions, orthogonal representation

$$F_{ridge}(\mathbf{y}) = y_1 - d\left(\sum_{i=2}^{N} y_i^2\right)^{\alpha/2} =: x - dR^{\alpha}$$

  - Linear component + embedded sphere
- Sharp ridge: $\alpha = 1$
  - $F_{ridge}(\mathbf{y}) = x - dR$
- Parabolic ridge: $\alpha = 2$
  - $F_{ridge}(\mathbf{y}) = x - dR^2$
- Meliorating
  - Decrease R and/or increase x

Ridge Functions

- Ridge functions, orthogonal representation

$$F_{ridge}(\mathbf{y}) = y_1 - d\Big(\sum_{i=2}^{N} y_i^2\Big)^{\alpha/2} =: x - dR^{\alpha}$$

  - Linear component + embedded sphere
- Sharp ridge: $\alpha = 1$
  - $F_{ridge}(\mathbf{y}) = x - dR$
- Parabolic ridge: $\alpha = 2$
  - $F_{ridge}(\mathbf{y}) = x - dR^2$
- Meliorating
  - Decrease R and/or increase x

Ridge Functions

- Ridge functions, orthogonal representation

$$
\begin{aligned}
F_{ridge}(\mathbf{y}) &= y_1 - d\left(\sum_{i=2}^{N} y_i^2\right)^{\alpha/2} \\
&=: x - dR^{\alpha}
\end{aligned}
$$

  - Linear component + embedded sphere
- Sharp ridge: $\alpha = 1$
  - $F_{ridge}(\mathbf{y}) = x - dR$
- Parabolic ridge: $\alpha = 2$
  - $F_{ridge}(\mathbf{y}) = x - dR^2$
- Meliorating
  - Decrease R and/or increase x

How to model the evolution dynamics

- State variables: ς, x, R
- Describing the change during one generation
- Evolution equations

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \mathrm{E}[x^{(g+1)} - x^{(g)}] + \mathcal{R}_x^{(g)} \\
R^{(g+1)} &= R^{(g)} - \mathrm{E}[R^{(g)} - R^{(g+1)}] + \mathcal{R}_R^{(g)} \\
\langle \varsigma^{(g+1)} \rangle &= \langle \varsigma^{(g)} \rangle \left(1 + \mathrm{E}[\frac{\langle \varsigma^{(g+1)} \rangle - \langle \varsigma^{(g)} \rangle}{\langle \varsigma^{(g)} \rangle}] \right) + \mathcal{R}_\sigma^{(g)}
\end{aligned}
$$

- Consist of expected change and perturbation part
- Mutation strength: relative change

How to model the evolution dynamics

- State variables: ς, x, R
- Describing the change during one generation
- Evolution equations

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \mathrm{E}[x^{(g+1)} - x^{(g)}] + \mathcal{R}_x^{(g)} \\
R^{(g+1)} &= R^{(g)} - \mathrm{E}[R^{(g)} - R^{(g+1)}] + \mathcal{R}_R^{(g)} \\
\langle \varsigma^{(g+1)} \rangle &= \langle \varsigma^{(g)} \rangle \left(1 + \mathrm{E}[\frac{\langle \varsigma^{(g+1)} \rangle - \langle \varsigma^{(g)} \rangle}{\langle \varsigma^{(g)} \rangle}] \right) + \mathcal{R}_\sigma^{(g)}
\end{aligned}
$$

- Consist of expected change and perturbation part
- Mutation strength: relative change

How to model the evolution dynamics

- State variables: ς, x, R
- Describing the change during one generation
- Evolution equations

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \mathrm{E}[x^{(g+1)} - x^{(g)}] + \mathcal{R}_x^{(g)} \\
R^{(g+1)} &= R^{(g)} - \mathrm{E}[R^{(g)} - R^{(g+1)}] + \mathcal{R}_R^{(g)} \\
\langle \varsigma^{(g+1)} \rangle &= \langle \varsigma^{(g)} \rangle \left(1 + \mathrm{E}[\frac{\langle \varsigma^{(g+1)} \rangle - \langle \varsigma^{(g)} \rangle}{\langle \varsigma^{(g)} \rangle}]\right) + \mathcal{R}_\sigma^{(g)}
\end{aligned}
$$

- Consist of expected change and perturbation part
- Mutation strength: relative change

The Evolution Equations

- To shorten the notation:
  $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \mathrm{E}[x^{(g+1)} - x^{(g)}] + \mathcal{R}_x \\
r &= R - \mathrm{E}[R - r] + \mathcal{R}_R \\
\varsigma &= \sigma\left(1 + \mathrm{E}\left[\frac{\varsigma - \sigma}{\sigma}\right]\right) + \mathcal{R}_\sigma
\end{aligned}
$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)

The Evolution Equations

- To shorten the notation: $\sigma := \langle\varsigma^{(g)}\rangle$, $\varsigma := \langle\varsigma^{(g+1)}\rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \mathrm{E}[x^{(g+1)} - x^{(g)}] + \mathcal{R}_x \\
r &= R - \mathrm{E}[R - r] + \mathcal{R}_R \\
\varsigma &= \sigma\left(1 + \mathrm{E}\left[\frac{\varsigma - \sigma}{\sigma}\right]\right) + \mathcal{R}_\sigma
\end{aligned}
$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)

The Evolution Equations

- To shorten the notation:
  $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and
  $r = R^{(g+1)}$

$$\begin{aligned} x^{(g+1)} &= x^{(g)} + \varphi_x(R, x^{(g)}, \sigma) + \mathcal{R}_x \\ r &= R - \mathrm{E}[R - r] + \mathcal{R}_R \\ \varsigma &= \sigma\left(1 + \mathrm{E}\left[\frac{\varsigma - \sigma}{\sigma}\right]\right) + \mathcal{R}_\sigma \end{aligned}$$

- x- and R-evolution: Expected changes $\Rightarrow$ progress rates
- ς-evolution: Expected change $\Rightarrow$ self-adaptation response (SAR)

The Evolution Equations

- To shorten the notation: $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$\begin{aligned} x^{(g+1)} &= x^{(g)} + \varphi_x(R, x^{(g)}, \sigma) + \mathcal{R}_x \\ r &= R - \mathrm{E}[R - r] + \mathcal{R}_R \\ \varsigma &= \sigma\left(1 + \mathrm{E}\left[\frac{\varsigma - \sigma}{\sigma}\right]\right) + \mathcal{R}_\sigma \end{aligned}$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)

The Evolution Equations

- To shorten the notation: $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \varphi_x(R, x^{(g)}, \sigma) + \mathcal{R}_x \\
r &= R - \varphi_R(R, x^{(g)}, \sigma) + \mathcal{R}_R \\
\varsigma &= \sigma\left(1 + \mathrm{E}\left[\frac{\varsigma - \sigma}{\sigma}\right]\right) + \mathcal{R}_\sigma
\end{aligned}
$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)

The Evolution Equations

- To shorten the notation: $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$\begin{aligned} x^{(g+1)} &= x^{(g)} + \varphi_x(R, x^{(g)}, \sigma) + \mathcal{R}_x \\ r &= R - \varphi_R(R, x^{(g)}, \sigma) + \mathcal{R}_R \\ \varsigma &= \sigma\left(1 + \mathrm{E}\left[\frac{\varsigma - \sigma}{\sigma}\right]\right) + \mathcal{R}_\sigma \end{aligned}$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)

The Evolution Equations

- To shorten the notation: $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \varphi_x(R, x^{(g)}, \sigma) + \mathcal{R}_x \\
r &= R - \varphi_R(R, x^{(g)}, \sigma) + \mathcal{R}_R \\
\varsigma &= \sigma\Big(1 + \psi(R, x^{(g)}, \sigma)\Big) + \mathcal{R}_\sigma
\end{aligned}
$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)

The Evolution Equations

- To shorten the notation: $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$\begin{aligned} x^{(g+1)} &= x^{(g)} + \varphi_x(R, x^{(g)}, \sigma) + \mathcal{R}_x \\ r &= R - \varphi_R(R, x^{(g)}, \sigma) + \mathcal{R}_R \\ \varsigma &= \sigma\Big(1 + \psi(R, x^{(g)}, \sigma)\Big) + \mathcal{R}_\sigma \end{aligned}$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)
- First approach: Neglecting perturbation part

The Evolution Equations

- To shorten the notation: $\sigma := \langle \varsigma^{(g)} \rangle$, $\varsigma := \langle \varsigma^{(g+1)} \rangle$, $R := R^{(g)}$, and $r = R^{(g+1)}$

$$\begin{aligned} x^{(g+1)} &= x^{(g)} + \varphi_x(R, x^{(g)}, \sigma) \\ r &= R - \varphi_R(R, x^{(g)}, \sigma) \\ \varsigma &= \sigma\Big(1 + \psi(R, x^{(g)}, \sigma)\Big) \end{aligned}$$

- x- and R-evolution: Expected changes ⇒ progress rates
- ς-evolution: Expected change ⇒ self-adaptation response (SAR)
- First approach: Neglecting perturbation part

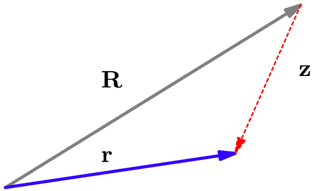
Part II

Preliminaries

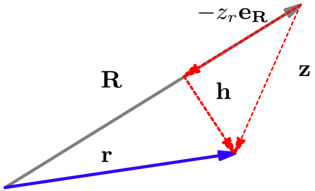
Preliminaries

- Needed: Progress rates and SAR
- Main points
  - Consider change induced by mutation
  - New vector can be decomposed into
    - x-component, i.e., first component
    - Perpendicular $(N-1)$-dimensional part
  - Second part: Similar decomposition as in the case of the sphere model
    - Part in the same manifold of previous vector $\mathbf{R}$
    - Perpendicular part

Vector Decomposition



Vector Decomposition



Preliminaries

- Needed: Progress rates and SAR
- Main points
  - Consider change induced by mutation
  - New vector can be decomposed into
    - x-component, i.e., first component
    - Perpendicular $(N-1)$-dimensional part
  - Second part: Similar decomposition as in the case of the sphere model
    - Part in the same manifold of previous vector $\mathbf{R}$
    - Perpendicular part
  - Derivation of density function (pdf) describing change by mutation
  - Main tools: Order statistics and decomposition above

Progress in Axis Direction

- Expected one-generation change in ridge direction

$$\varphi_x := \mathrm{E}[x^{(g+1)} - x^{(g)}]$$

- Progress rate for $N \to \infty$

$$\varphi_x^*(\sigma^*, R) = \frac{\sigma^*}{\sqrt{1+\alpha^2 d^2 R^{2\alpha-2}}} c_{\mu/\mu,\lambda}$$

  - with $\varphi_x^* := N\varphi_x$ and $\sigma^* := N\sigma$
  - $c_{\mu/\mu,\lambda}$ special case of the generalized progress coefficients $e_{\mu,\lambda}^{\alpha,\beta}$
  - Progress rate linear in σ^*
  - No loss part
  - Derived for $\tau = 0$

Progress towards the Axis

- Progress rate of the R-evolution

$$\varphi_R \quad := \quad \mathrm{E}[R - r]$$

- New vector $\mathbf{r} = \mathbf{R} - \langle z_R \rangle \mathbf{e}_R + \langle \mathbf{h} \rangle$
  - $\langle \mathbf{h} \rangle$ = component perpendicular to $\mathbf{R}$, $\mathbf{e_R} := \mathbf{R}/R$.
- Length $r = \|\mathbf{r}\| = \sqrt{\mathbf{r}^{\mathrm{T}}\mathbf{r}}$
- Progress rate

$$\varphi_R \quad = \quad \mathrm{E}[R - \sqrt{(R - \langle z_R \rangle)^2 + \langle h \rangle^2}]$$

Progress towards the Axis II

- Progress rate

$$\varphi_R = \mathrm{E}[R - \sqrt{(R - \langle z_R \rangle)^2 + \langle h \rangle^2}]$$

- Leads finally to

$$\varphi_R^*(\sigma^*, R) = \frac{\alpha d R^{\alpha-1} \sigma^*}{\sqrt{1 + \alpha^2 d^2 R^{2\alpha-2}}} c_{\mu/\mu,\lambda} - \frac{{\sigma^*}^2}{2R\mu}$$

  for $N \to \infty$ and $\tau = 0$

- Approximate equation for finite N and small τ
- Loss part stemming from perpendicular component

Progress towards the Axis II

- Progress rate

$$\varphi_R = \mathrm{E}[R - \sqrt{(R - \langle z_R \rangle)^2 + \langle h \rangle^2}]$$

- Leads finally to

$$\varphi_R^*(\sigma^*, R) = \frac{\alpha d R^{\alpha-1} \sigma^*}{\sqrt{1 + \alpha^2 d^2 R^{2\alpha-2}}} c_{\mu/\mu,\lambda} - \frac{{\sigma^*}^2}{2R\mu}$$

for $N \to \infty$ and $\tau = 0$

- Approximate equation for finite N and small τ
- Loss part stemming from perpendicular component

The SAR

- Self-adaptation response (SAR)
  - Expected relative change of the mutation strength

$$\psi := \mathrm{E}\left[\frac{\varsigma - \sigma}{\sigma}\right]$$

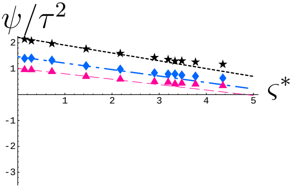
- SAR

$$\psi_\infty(\varsigma^*) = \mathcal{O}(\tau^4) + \tau^2 \left(\frac{1}{2} + e^{1,1}_{\mu,\lambda} - c_{\mu/\mu,\lambda}\varsigma^* \sqrt{\frac{\alpha^2 d^2 R^{2\alpha-2}}{R^2(1+\alpha^2 d^2 R^{2\alpha-2})}} \right)$$

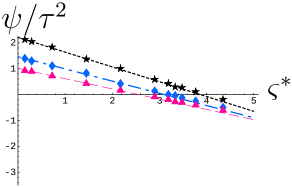
for $N \to \infty$ and $\tau \ll 1$

- Linear loss part

The SAR: Comparison with Experiments



a) $N = 30$, $d = 0.2$, sharp ridge



b) $N = 30$, $d = 0.2$, parabolic ridge

Part III

Self-Adaptation on Ridge Functions

The Evolution Equations

- Evolution Equations

$$
\begin{aligned}
x^{(g+1)} &= x^{(g)} + \frac{1}{N}\varphi_x^*(R, \sigma^*) \\
r &= R - \frac{1}{N}\varphi_R^*(R, \sigma^*) \\
\varsigma^* &= \sigma^*\Big(1 + \psi(R, \sigma^*)\Big)
\end{aligned}
$$

- Normalized with respect to N
- No influence of the x-component on ς^* and R
- $\Rightarrow$ Analyzing the system in (ς^*, R)

The Evolution Equations

- Evolution Equations

$$\begin{aligned} x^{(g+1)} &= x^{(g)} + \frac{1}{N}\varphi_x^*(R, \sigma^*) \\ r &= R - \frac{1}{N}\varphi_R^*(R, \sigma^*) \\ \varsigma^* &= \sigma^*\Big(1 + \psi(R, \sigma^*)\Big) \end{aligned}$$

- Normalized with respect to N
- No influence of the x-component on ς^* and R
- $\Rightarrow$ Analyzing the system in (ς^*, R)

The System in ς^* and R

- Analyzing the system in (ς^*, R)
- Evolution of the mutation strength
  - $\varsigma^* = \sigma^* \Big(1 + \psi(\sigma^*)\Big)$
  - $\psi(\sigma^*) =$ $\tau^2 \left(1/2 + e^{1,1}_{\mu,\lambda} - c_{\mu/\mu,\lambda}\sigma^* \sqrt{\frac{\alpha^2 d^2 R^{2\alpha-2}}{R^2(1+\alpha^2 d^2 R^{2\alpha-2})}}\right)$
- Evolution of the distance to the ridge
  - $r = R - \varphi^*_R(\sigma^*)/N$
  - $\varphi^*_R = \sqrt{\frac{\alpha^2 d^2 R^{2\alpha-2}}{1+\alpha^2 d^2 R^{2\alpha-2}}} c_{\mu/\mu,\lambda}\sigma^* - \frac{\sigma^{*2}}{2R\mu}$

The System in ς^* and R

- When does the system in ς^* and R come to a hold?
- Mutation strength
  - $\sigma^* = 0$ or $\psi(\sigma^*, R) = 0$
  - $\sigma^* = 0$ or

$$\tau^2 \left(\frac{1}{2} + e^{1,1}_{\mu,\lambda} - c_{\mu/\mu,\lambda}\varsigma^* \sqrt{\frac{\alpha^2 d^2 R^{2\alpha-2}}{R^2(1+\alpha^2 d^2 R^{2\alpha-2})}} \right) = 0$$

- Distance to the ridge
  - $\varphi^*_R(\sigma^*, R) = 0$

$$\sqrt{\frac{\alpha^2 d^2 R^{2\alpha-2}}{1+\alpha^2 d^2 R^{2\alpha-2}}} c_{\mu/\mu,\lambda}\sigma^* - \frac{{\sigma^*}^2}{2R\mu} = 0$$

- $\Rightarrow$ Closer look at the zero points

The System in ς^* and R

- When does the system in ς^* and R come to a hold?
- Mutation strength
  - $\sigma^* = 0$ or $\psi(\sigma^*, R) = 0$
  - $\sigma^* = 0$ or

$$\tau^2 \left(\frac{1}{2} + e^{1,1}_{\mu,\lambda} - c_{\mu/\mu,\lambda} \varsigma^* \sqrt{\frac{\alpha^2 d^2 R^{2\alpha-2}}{R^2(1 + \alpha^2 d^2 R^{2\alpha-2})}} \right) = 0$$

- Distance to the ridge
  - $\varphi^*_R(\sigma^*, R) = 0$

$$\sqrt{\frac{\alpha^2 d^2 R^{2\alpha-2}}{1 + \alpha^2 d^2 R^{2\alpha-2}}} c_{\mu/\mu,\lambda} \sigma^* - \frac{{\sigma^*}^2}{2R\mu} = 0$$

- $\Rightarrow$ Closer look at the zero points

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + \alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1+\alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + \alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2d c_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1}{\alpha^2 d^2 R^{2\alpha-4}} + \frac{\alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1}{\alpha^2 d^2 R^{2\alpha-4}} + R^2}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + \alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + \alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2d c_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + 4d^2 R^2}{4d^2}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + \alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + \alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} R \sqrt{\frac{1+d^2}{d^2}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

The Zero Point of the SAR

- Zero point of the SAR

$$\varsigma^*_{\psi_0} = \frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + \alpha^2 d^2 R^{2\alpha-2}}{\alpha^2 d^2 R^{2\alpha-4}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\psi_0} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\psi_0} = \begin{cases} \infty & \text{if } \alpha > 2 \\ (1/2 + e^{1,1}_{\mu,\lambda})/(2dc_{\mu/\mu,\lambda}) & \text{if } \alpha = 2 \\ 0 & \text{if } \alpha = 1 \end{cases}$$

- Differences: sharp ($\alpha = 1$) and parabolic ridge ($\alpha = 2$)

The Zero Point of the Progress Rate

- Second zero of the progress rate

$$\varsigma^*_{\varphi_{R_0}} = 2\mu c_{\mu/\mu,\lambda}\sqrt{\frac{\alpha^2 d^2 R^{2\alpha}}{1+\alpha^2 d^2 R^{2\alpha-2}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty} \varsigma^*_{\varphi_{R_0}} = 2\mu c_{\mu/\mu,\lambda} \lim_{R\to\infty} \sqrt{\frac{\alpha^2 d^2 R^{2\alpha}}{1+\alpha^2 d^2 R^{2\alpha-2}}} = \infty$$

$$\lim_{R\to 0} \varsigma^*_{\varphi_{R_0}} = 2\mu c_{\mu/\mu,\lambda} \lim_{R\to 0} \sqrt{\frac{\alpha^2 d^2 R^{2\alpha}}{1+\alpha^2 d^2 R^{2\alpha-2}}} = 0$$

The Zero Point of the Progress Rate

- Second zero of the progress rate

$$\varsigma^*_{\varphi_{R_0}} = 2\mu c_{\mu/\mu,\lambda}\sqrt{\frac{\alpha^2 d^2 R^{2\alpha}}{1+\alpha^2 d^2 R^{2\alpha-2}}}$$

- Function of R
- Limit behavior

$$\lim_{R\to\infty}\varsigma^*_{\varphi_{R_0}} = 2\mu c_{\mu/\mu,\lambda}\lim_{R\to\infty}\sqrt{\frac{\alpha^2 d^2 R^{2\alpha}}{1+\alpha^2 d^2 R^{2\alpha-2}}} = \infty$$

$$\lim_{R\to 0}\varsigma^*_{\varphi_{R_0}} = 2\mu c_{\mu/\mu,\lambda}\lim_{R\to 0}\sqrt{\frac{\alpha^2 d^2 R^{2\alpha}}{1+\alpha^2 d^2 R^{2\alpha-2}}} = 0$$

The Zero Points

- Progress rate
  - For $R \to \infty$, zero point $\to \infty$
  - For $R \to 0$, zero point $\to 0$
- SAR
  - For $R \to \infty$, zero point $\to \infty$
  - Difference between sharp and parabolic ridge
  - Sharp ridge: $R \to 0$, zero point $\to 0$
  - Parabolic ridge: $R \to 0$, zero point $\to$ positive limit
- Consequences for the sharp and parabolic ridge?

The Zero Points

- Progress rate
  - For $R \to \infty$, zero point $\to \infty$
  - For $R \to 0$, zero point $\to 0$
- SAR
  - For $R \to \infty$, zero point $\to \infty$
  - Difference between sharp and parabolic ridge
  - Sharp ridge: $R \to 0$, zero point $\to 0$
  - Parabolic ridge: $R \to 0$, zero point $\to$ positive limit
- Consequences for the sharp and parabolic ridge?

The Zero Points

- Progress rate
  - For $R \to \infty$, zero point $\to \infty$
  - For $R \to 0$, zero point $\to 0$
- SAR
  - For $R \to \infty$, zero point $\to \infty$
  - Difference between sharp and parabolic ridge
  - Sharp ridge: $R \to 0$, zero point $\to 0$
  - Parabolic ridge: $R \to 0$, zero point $\to$ positive limit
- Consequences for the sharp and parabolic ridge?

Part IV

The Parabolic Ridge

The Evolution Equations

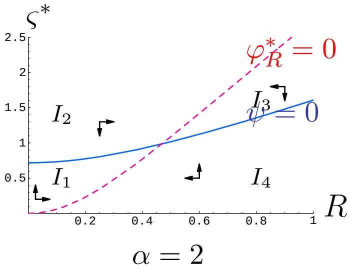
- Parabolic ridge: $\alpha = 2$, $F(\mathbf{y}) = x - dR^2$
- Evolution equations

$$
\begin{aligned}
r &= R - \frac{\varphi_R^*(R, \sigma^*)}{N} \\
&= R - \frac{1}{N}\left(\sqrt{\frac{4d^2R^2}{1+4d^2R^2}}c_{\mu/\mu,\lambda}\sigma^* - \frac{{\sigma^*}^2}{2R\mu}\right) \\
\varsigma^* &= \sigma^*\Big(1 + \psi(R, \sigma^*)\Big) \\
&= \sigma^*\left(1 + \tau^2\Big(\frac{1}{2} + e^{1,1}_{\mu,\lambda} - c_{\mu/\mu,\lambda}\sigma^*\sqrt{\frac{4d^2}{1+4d^2R^2}}\Big)\right)
\end{aligned}
$$

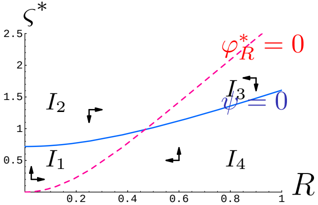
- Start: Discussion of zero points of SAR and progress rate

Zero Points as Function of R

- SAR: $\varsigma^*_{\psi_0} = \frac{1/2+e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\sqrt{\frac{1+4d^2R^2}{4d^2}}$
- Progress rate: $\varsigma^*_{\varphi_{R_0}} = 2\mu c_{\mu/\mu,\lambda}\sqrt{\frac{4d^2R^4}{1+4d^2R^2}}$
- Functions of R

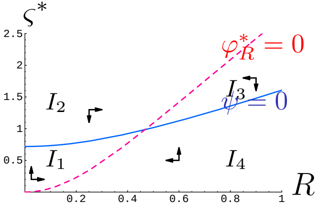


Zero Points as Function of R



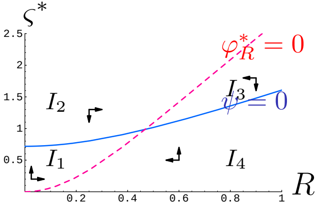
- Intersection at stationary R
- Divergence of $R \to \infty$?
  - If yes: System cannot stay in I_1 or I_2
  - In I_3 and I_4 mutation strength: smaller than zero of progress rate
  - $\Rightarrow$ Decrease of R

Zero Points as Function of R



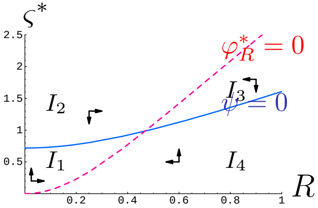
- Intersection at stationary R
- Divergence of $R \to \infty$?
    - If yes: System cannot stay in I_1 or I_2
    - In I_3 and I_4 mutation strength: smaller than zero of progress rate
    - $\Rightarrow$ Decrease of R

Zero Points as Function of R



- Intersection at stationary R
- Convergence of $R \to 0$?
  - If yes: System cannot stay in I_3 or I_4
  - In I_1 and I_2 mutation strength: higher than zero of progress rate
  - In I_1 : expected increase of mutation strength
  - $\Rightarrow$ Increase of R
- Evolution equations
  - System neither going to the origin nor diverging towards ∞

Zero Points as Function of R



- Intersection at stationary R
- Convergence of $R \to 0$?
  - If yes: System cannot stay in I_3 or I_4
  - In I_1 and I_2 mutation strength: higher than zero of progress rate
  - In I_1 : expected increase of mutation strength
  - $\Rightarrow$ Increase of R
- Evolution equations
  - System neither going to the origin nor diverging towards ∞

A Stationary State

- No divergence to ∞/No convergence to origin
- But: Evolution equations allow for stationary point
- Stationary point

$$
\begin{aligned}
r &= R \Rightarrow \varphi_R^*(R, \sigma^*) = 0 \\
\varsigma^* &= \sigma^* \Rightarrow \psi(R, \sigma^*) = 0 \bigvee \sigma^* = 0
\end{aligned}
$$

- First possibility: $\sigma^* = 0$ and arbitrary distance
- Second possibility: intersection of $\varphi_R^*(R, \sigma^*) = 0$ and $\psi(R, \sigma^*) = 0$

A Stationary State II

- Second possibility: intersection of $\varphi_R^*(R, \sigma^*) = 0$ and $\psi(R, \sigma^*) = 0$
- Solving

$$\varsigma^*_{\psi_0} = \varsigma^*_{\varphi_{R_0}}$$

- Result: Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

A Stationary State II

- Second possibility: intersection of $\varphi_R^*(R, \sigma^*) = 0$ and $\psi(R, \sigma^*) = 0$
- Solving

$$\frac{1/2 + e^{1,1}_{\mu/\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + 4d^2R^2}{4d^2}} = 2\mu c_{\mu/\mu,\lambda} \sqrt{\frac{4d^2R^4}{1 + 4d^2R^2}}$$

- Result: Stationary distance to the ridge

$$R_{st} = \frac{1}{2d} \sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

A Stationary State II

- Second possibility: intersection of $\varphi_R^*(R, \sigma^*) = 0$ and $\psi(R, \sigma^*) = 0$
- Solving

$$\frac{1/2 + e^{1,1}_{\mu/\mu,\lambda}}{c_{\mu/\mu,\lambda}} \sqrt{\frac{1 + 4d^2 R^2}{4d^2}} = 2\mu c_{\mu/\mu,\lambda} \sqrt{\frac{4d^2 R^4}{1 + 4d^2 R^2}}$$

- Result: Stationary distance to the ridge

$$R_{st} = \frac{1}{2d} \sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

A Stationary State III: Stationary distance

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Scales with $1/(2d)$
- Similar result as for CSA-ES (Beyer, 2004)
- Analogy to noisy sphere
    - Noisy sphere: $F(R) = -R^\alpha + \sigma_\epsilon \mathcal{N}(0,1)$
    - Stationary distance (e.g. Beyer,2001/Arnold, 2002) $R_{st} \propto \sigma_\epsilon$
    - Parabolic ridge: $F(R) = x - dR^2$
    - Stationary distance $R_{st} \propto 1/d$

A Stationary State III: Stationary distance

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Scales with $1/(2d)$
- Similar result as for CSA-ES (Beyer, 2004)
- Analogy to noisy sphere
  - Noisy sphere: $F(R) = -R^{\alpha} + \sigma_\epsilon \mathcal{N}(0,1)$
  - Stationary distance (e.g. Beyer,2001/Arnold, 2002) $R_{st} \propto \sigma_\epsilon$
  - Parabolic ridge: $F(R) = x - dR^2$
  - Stationary distance $R_{st} \propto 1/d$

A Stationary State III: Stationary distance

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Scales with $1/(2d)$
- Similar result as for CSA-ES (Beyer, 2004)
- Analogy to noisy sphere
  - Noisy sphere: $F(R) = -R^\alpha + \sigma_\epsilon \mathcal{N}(0,1)$
  - Stationary distance (e.g. Beyer,2001/Arnold, 2002) $R_{st} \propto \sigma_\epsilon$
  - Parabolic ridge: $F(R) = x - dR^2$
  - Stationary distance $R_{st} \propto 1/d$

A Stationary State III: Stationary distance

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Scales with $1/(2d)$
- Similar result as for CSA-ES (Beyer, 2004)
- Analogy to noisy sphere
  - Noisy sphere: $F(R) = -R^\alpha + \sigma_\epsilon \mathcal{N}(0,1)$
  - Stationary distance (e.g. Beyer,2001/Arnold, 2002) $R_{st} \propto \sigma_\epsilon$
  - Parabolic ridge: $F(R) = x - dR^2$
  - Stationary distance $R_{st} \propto 1/d$

A Stationary State IV: Stationary Mutation Strength

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

  - Can be used to obtain stationary mutation strength and stationary progress parallel to the ridge
- Stationary mutation strength
  - Consider $\varsigma^*_{\psi_0} = \frac{1/2+e^{1,1,}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\sqrt{\frac{1+4d^2R^2}{4d^2}}$
  - Leads to

$$\varsigma^*_{st} = \frac{1}{2d}\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\sqrt{\frac{2\mu c^2_{\mu/\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

A Stationary State IV: Stationary Mutation Strength

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

  - Can be used to obtain stationary mutation strength and stationary progress parallel to the ridge

- Stationary mutation strength
  - Consider $\varsigma^*_{\psi_0} = \frac{1/2+e^{1,1,}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\sqrt{\frac{1+4d^2R^2}{4d^2}}$
  - Leads to

$$\varsigma^*_{st} = \frac{1}{2d}\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\sqrt{\frac{2\mu c^2_{\mu/\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

A Stationary State V: Stationary progress rate

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Stationary mutation strength

$$\varsigma^*_{st} = \frac{1}{2d}\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\sqrt{\frac{2\mu c^2_{\mu/\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Stationary progress rate

$$\varphi^*_{x\,st} = \frac{1}{\sqrt{1 + 4d^2 R^2_{st}}} c_{\mu/\mu,\lambda} \varsigma^*_{st}$$

A Stationary State V: Stationary progress rate

- Stationary distance to the ridge

$$R_{st} = \frac{1}{2d}\sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Stationary mutation strength

$$\varsigma^*_{st} = \frac{1}{2d}\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\sqrt{\frac{2\mu c^2_{\mu/\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

- Stationary progress rate

$$\varphi^*_{x\ st} = \frac{1}{2d}\left(1/2 + e^{1,1}_{\mu,\lambda}\right)$$

Effects of Recombination

- Stationary distance $R_{st} = \frac{1}{2d}\sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Stationary mutation strength
  $\varsigma^*_{st} = \frac{1}{2d}\left(\frac{1/2+e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\sqrt{\frac{2\mu c^2_{\mu/\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Stationary progress rate $\varphi^*_{x\ st} = \frac{1}{2d}(1/2 + e^{1,1}_{\mu,\lambda})$
- Used to investigate effects of recombination
- Maximal progress occurs for non-recombinative $(1,\lambda)$-ES
- No benefit from recombination

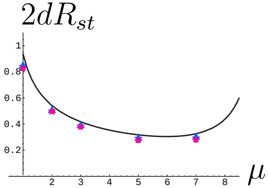
Effects of Recombination

- Stationary distance $R_{st} = \frac{1}{2d}\sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Stationary mutation strength
$$\varsigma^*_{st} = \frac{1}{2d}\left(\frac{1/2+e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\sqrt{\frac{2\mu c^2_{\mu/\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$$
- Stationary progress rate $\varphi^*_{x\ st} = \frac{1}{2d}\left(1/2 + e^{1,1}_{\mu,\lambda}\right)$
- Used to investigate effects of recombination
- Maximal progress occurs for non-recombinative $(1,\lambda)$-ES
- No benefit from recombination

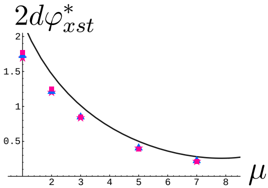
Effects of Recombination

- Stationary distance $R_{st} = \frac{1}{2d}\sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Stationary mutation strength
  $\varsigma^*_{st} = \frac{1}{2d}\left(\frac{1/2+e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\sqrt{\frac{2\mu c^2_{\mu/\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Stationary progress rate $\varphi^*_{x\ st} = \frac{1}{2d}(1/2+e^{1,1}_{\mu,\lambda})$
- Used to investigate effects of recombination
- Maximal progress occurs for non-recombinative $(1,\lambda)$-ES
- No benefit from recombination

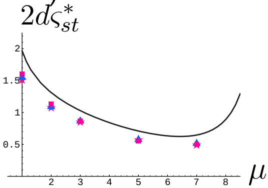
Comparison with Experiments



a) Stat. distance



b) Stat. progress



c) Mutation strength

Part V

The Sharp Ridge

Evolution Equations and Zero Points

- Sharp ridge ($\alpha = 1$), $F(\mathbf{y}) = x - dR$
- Evolution Equations

$$
\begin{aligned}
r &= R - \frac{\varphi_R^*(R, \sigma^*)}{N} \\
\varsigma^* &= \sigma^* \Big(1 + \psi(R, \sigma^*)\Big)
\end{aligned}
$$

- Zero Points

$$
\begin{aligned}
\varsigma^*_{\varphi_{R_0}} &= 2R\mu c_{\mu/\mu,\lambda} \sqrt{\frac{d^2}{1+d^2}} \\
\varsigma^*_{\psi_0} &= R\sqrt{\frac{1+d^2}{d^2}} \Big(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\Big)
\end{aligned}
$$

Evolution Equations and Zero Points

- Sharp ridge ($\alpha = 1$), $F(\mathbf{y}) = x - dR$
- Evolution Equations

$$r = R - \frac{1}{N}\left(\sqrt{\frac{d^2}{1+d^2}}c_{\mu/\mu,\lambda}\sigma^* - \frac{{\sigma^*}^2}{2R\mu}\right)$$

$$\varsigma^* = \sigma^*\left(1 + \tau^2\left(\frac{1}{2} + e^{1,1}_{\mu,\lambda} - c_{\mu/\mu,\lambda}\sigma^*\sqrt{\frac{d^2}{R^2(1+d^2)}}\right)\right)$$

- Zero Points

$$\varsigma^*_{\varphi_{R_0}} = 2R\mu c_{\mu/\mu,\lambda}\sqrt{\frac{d^2}{1+d^2}}$$

$$\varsigma^*_{\psi_0} = R\sqrt{\frac{1+d^2}{d^2}}\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)$$

Zero Points

- Zero points

$$\varsigma^*_{\varphi_R} = R\,2\mu c_{\mu/\mu,\lambda}\sqrt{\frac{d^2}{1+d^2}}$$

$$\varsigma^*_{\psi_0} = R\sqrt{\frac{1+d^2}{d^2}}\left(\frac{1/2+e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)$$

- Zero points: Linear functions in R
- Two cases
  - Intersection only in zero
  - Equal gradient, i.e., $d_{crit} = \sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$

Zero Points

- Zero points

$$\varsigma^*_{\varphi_R} = R\,2\mu c_{\mu/\mu,\lambda}\sqrt{\frac{d^2}{1+d^2}}$$

$$\varsigma^*_{\psi_0} = R\sqrt{\frac{1+d^2}{d^2}}\left(\frac{1/2+e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)$$

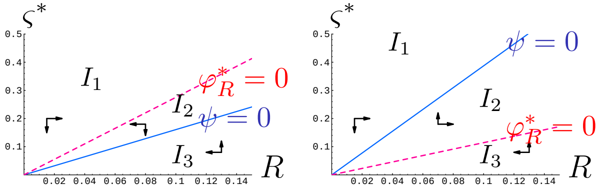
- Zero points: Linear functions in R
- Two cases
  - Intersection only in zero
  - Equal gradient, i.e., $d_{crit} = \sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$

Convergence and Divergence

- Decisive parameter:

$$d_{crit} = \sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$$

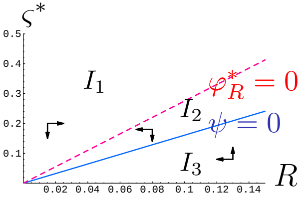
- Two cases



a) $\alpha = 1,\ d > d_{crit}$ b) $\alpha = 1,\ d < d_{crit}$

Convergence

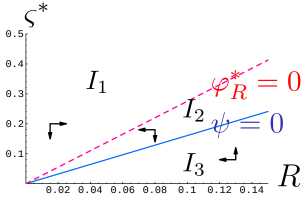
- $d_{crit} = \sqrt{\frac{1/2 + e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$
- Case: $d > d_{crit}$
- Regions I_1 and I_3 left via I_2
- System cannot leave I_2
- In I_2: Decrease of mutation strength and distance
- System approaches origin



a) $\alpha = 1$, $d > d_{crit}$

Convergence

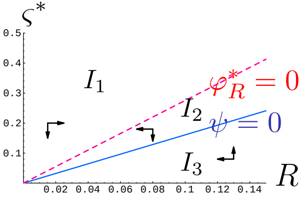
- $d_{crit} = \sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Case: $d > d_{crit}$
- Regions I_1 and I_3 left via I_2
- System cannot leave I_2
- In I_2: Decrease of mutation strength and distance
- System approaches origin



a) $\alpha = 1$, $d > d_{crit}$

Convergence

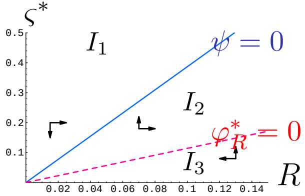
- $d_{crit} = \sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Case: $d > d_{crit}$
- Regions I_1 and I_3 left via I_2
- System cannot leave I_2
- In I_2: Decrease of mutation strength and distance
- System approaches origin



a) $\alpha = 1$, $d > d_{crit}$

Divergence

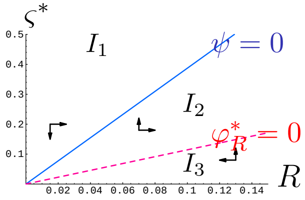
- $d_{crit} = \sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Case: $d < d_{crit}$
- Regions I_1 and I_3 left via I_2
- System cannot leave I_2
- In I_2: Increase of mutation strength and distance
- System diverges to ∞



a) $\alpha = 1$, $d < d_{crit}$

Divergence

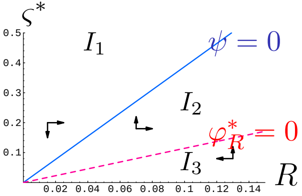
- $d_{crit} = \sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda}-1/2-e^{1,1}_{\mu,\lambda}}}$
- Case: $d < d_{crit}$
- Regions I_1 and I_3 left via I_2
- System cannot leave I_2
- In I_2: Increase of mutation strength and distance
- System diverges to ∞



a) $\alpha = 1$, $d < d_{crit}$

Divergence

- $d_{crit} = \sqrt{\frac{1/2+e^{1,1}_{\mu,\lambda}}{2\mu c^2_{\mu/\mu,\lambda} - 1/2 - e^{1,1}_{\mu,\lambda}}}$
- Case: $d < d_{crit}$
- Regions I_1 and I_3 left via I_2
- System cannot leave I_2
- In I_2: Increase of mutation strength and distance
- System diverges to ∞



a) $\alpha = 1$, $d < d_{crit}$

Divergence: Quality Change

- Sharp ridge: $F(\mathbf{y}) = x - dR$
- Divergence/Convergence depends on choice of d
- Case: No stagnation
  - Expected change of fitness?
  - Travel speed of the ES?
- Measure
  - Quality change, i.e. expected change of fitness
- Quality change
  - $\Delta Q^* = \mathrm{E}[F(\mathbf{y}^{(g+1)}) - F(\mathbf{y}^{(g)})]$

Divergence: Quality Change

- Sharp ridge: $F(\mathbf{y}) = x - dR$
- Divergence/Convergence depends on choice of d
- Case: No stagnation
  - Expected change of fitness?
  - Travel speed of the ES?
- Measure
  - Quality change, i.e. expected change of fitness
- Quality change
  - $\Delta Q^* = \mathrm{E}[x^{(g+1)} - x^{(g)}] - d\mathrm{E}[R^{(g+1)} - R^{(g)}]$

Divergence: Quality Change

- Sharp ridge: $F(\mathbf{y}) = x - dR$
- Divergence/Convergence depends on choice of d
- Case: No stagnation
  - Expected change of fitness?
  - Travel speed of the ES?
- Measure
  - Quality change, i.e. expected change of fitness
- Quality change
  - $\Delta Q^* = \varphi_x^* + d\varphi_R^*$

Divergence: Quality Change

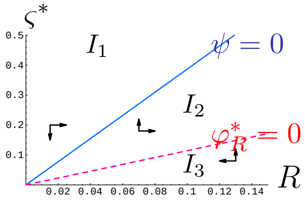
- Sharp ridge: $F(\mathbf{y}) = x - dR$
- Divergence/Convergence depends on choice of d
- Case: No stagnation
  - Expected change of fitness?
  - Travel speed of the ES?
- Measure
  - Quality change, i.e. expected change of fitness
- Quality change
  - $\Delta Q^* = \frac{1}{\sqrt{1+d^2}} c_{\mu/\mu,\lambda}\varsigma^* + \frac{d^2}{\sqrt{1+d^2}} c_{\mu/\mu,\lambda}\varsigma^* - \frac{d}{2R\mu}{\varsigma^*}^2$

Divergence: Quality Change

- Sharp ridge: $F(\mathbf{y}) = x - dR$
- Divergence/Convergence depends on choice of d
- Case: No stagnation
    - Expected change of fitness?
    - Travel speed of the ES?
- Measure
    - Quality change, i.e. expected change of fitness
- Quality change
    - $\Delta Q^* = \sqrt{1+d^2}c_{\mu/\mu,\lambda}\varsigma^* - \frac{d}{2R\mu}{\varsigma^*}^2$
- Optimizer: $\varsigma^*_{opt} = R\mu c_{\mu/\mu,\lambda}\frac{\sqrt{1+d^2}}{d}$
- Positive quality change?
- Optimizer attainable in the long run?

Quality Change

- Optimizer: $\varsigma^*_{opt} = R\mu c_{\mu/\mu,\lambda}\frac{\sqrt{1+d^2}}{d}$
- Reconsider



a) $\alpha = 1$, $d < d_{crit}$

- System in region I_2, below zero of SAR
- Zero of SAR: $\varsigma^*_{\psi_0} = R\left(\frac{1/2+e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\frac{\sqrt{1+d^2}}{d}$

Quality Change II

- Optimizer:

$$\varsigma^*_{opt} = R\,\mu c_{\mu/\mu,\lambda}\frac{\sqrt{1+d^2}}{d}$$

- Zero of SAR:

$$\varsigma^*_{\psi_0} = R\Big(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\Big)\frac{\sqrt{1+d^2}}{d}$$

- Maximal attainable mutation strength: ς^*_{opt}
- $\varsigma^*_{\psi_0} \approx \varsigma^*_{opt}$?
- Only for $\mu = 1$, i.e. no recombination
- But positive quality change

Quality Change II

- Optimizer:

$$\varsigma^*_{opt} = R\,\mu c_{\mu/\mu,\lambda}\frac{\sqrt{1+d^2}}{d}$$

- Zero of SAR:

$$\varsigma^*_{\psi_0} = R\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\frac{\sqrt{1+d^2}}{d}$$

- Maximal attainable mutation strength: ς^*_{opt}
- $\varsigma^*_{\psi_0} \approx \varsigma^*_{opt}$?
- Only for $\mu = 1$, i.e. no recombination
- But positive quality change

Quality Change II

- Optimizer:

$$\varsigma^*_{opt} = R\,\mu c_{\mu/\mu,\lambda}\frac{\sqrt{1+d^2}}{d}$$

- Zero of SAR:

$$\varsigma^*_{\psi_0} = R\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\frac{\sqrt{1+d^2}}{d}$$

- Maximal attainable mutation strength: ς^*_{opt}
- $\varsigma^*_{\psi_0} \approx \varsigma^*_{opt}$?
- Only for $\mu = 1$, i.e. no recombination
- But positive quality change

Quality Change II

- Optimizer:

$$\varsigma^*_{opt} = R\,\mu c_{\mu/\mu,\lambda}\frac{\sqrt{1+d^2}}{d}$$

- Zero of SAR:

$$\varsigma^*_{\psi_0} = R\left(\frac{1/2 + e^{1,1}_{\mu,\lambda}}{c_{\mu/\mu,\lambda}}\right)\frac{\sqrt{1+d^2}}{d}$$

- Maximal attainable mutation strength: ς^*_{opt}
- $\varsigma^*_{\psi_0} \approx \varsigma^*_{opt}$?
- Only for $\mu = 1$, i.e. no recombination
- But positive quality change

Conclusions and Outlook

- Sharp ridge
  - Using deterministic evolution equations
  - Ridge parameter d
    - Convergence to axis and stagnation
    - Divergence to axis and positive quality change
  - Recombination: optimal mutation strength not realized
- Parabolic ridge
  - Stationary distance to axis
  - Recombination not beneficial
- Future work
  - Fluctuation parts of evolution equations
  - N-dependent progress rates and SAR
  - Comparison with other adaptation schemes