Learning to cooperate via indirect reciprocity

Ulrich Berger
WU Vienna

Banff: June 17, 2010

Social dilemmas: Public Goods Game, Tragedy of the Commons, Prisoner's Dilemma

Cooperation is ubiquitous in social and economic life. But cooperation is threatened by defection of free riders. So why is there so much cooperation?

Altruistic acts (helping) decrease the payoff of the donor and increase the payoffs of the recipient. Hence altruism is not individually rational. So why do we see people helping each other?

The **helping game**

		You
Me	help	$-c, b$
	don't help	$0, 0$

$b, c > 0$

The **simultaneous helping game** . . .

		You	
		help	don't help
Me	help	$b - c, b - c$	$-c, b$
	don't help	$b, -c$	$0, 0$

The **simultaneous helping game** . . .

		You	
		help	don't help
Me	help	$b-c, b-c$	$-c, b$
	don't help	$b, -c$	$0, 0$

. . . is a **Prisoner's Dilemma**, if $b > c$.

		You	
		C	D
Me	C	$b-c$	$-c$
	D	b	0

How could altruism/helping/cooperation evolve?

Traditional answers:

- kin selection (Hamilton 1964)
- direct reciprocity / reciprocal altruism (Trivers 1971, Axelrod and Hamilton 1981)
- group selection (Wilson & Sober 1994)
- costly signaling (Gintis et al. 2001)

- **indirect reciprocity**
  (Sugden 1986, Alexander 1987, Weesie 1988)

direct reciprocity:
B helps A, A helps B, B helps A, . . .

indirect reciprocity:
B helps A, C helps B, D helps C, . . .

Indirect reciprocity is based on **reputation**.

- helping (C) increases one's reputation
- withholding help (D) decreases one's reputation

If help is preferentially directed towards those with high reputation, helping might survive evolution.

Nowak & Sigmund (1998): first formal model, *image scoring*

reputation measured by **score**

simulation results for *full score*
full score of an individual $= \#C - \#D$ since birth

Threshold strategies: k-strategist cooperates iff opponent's score is at least k.

In the long run, helping those with nonnegative full score gets established ($k = 0$ becomes fixed).

analytical results for **binary score**
- observe behavior in previous interaction as a donor

C (+1): **Good**, D (-1): **Bad**

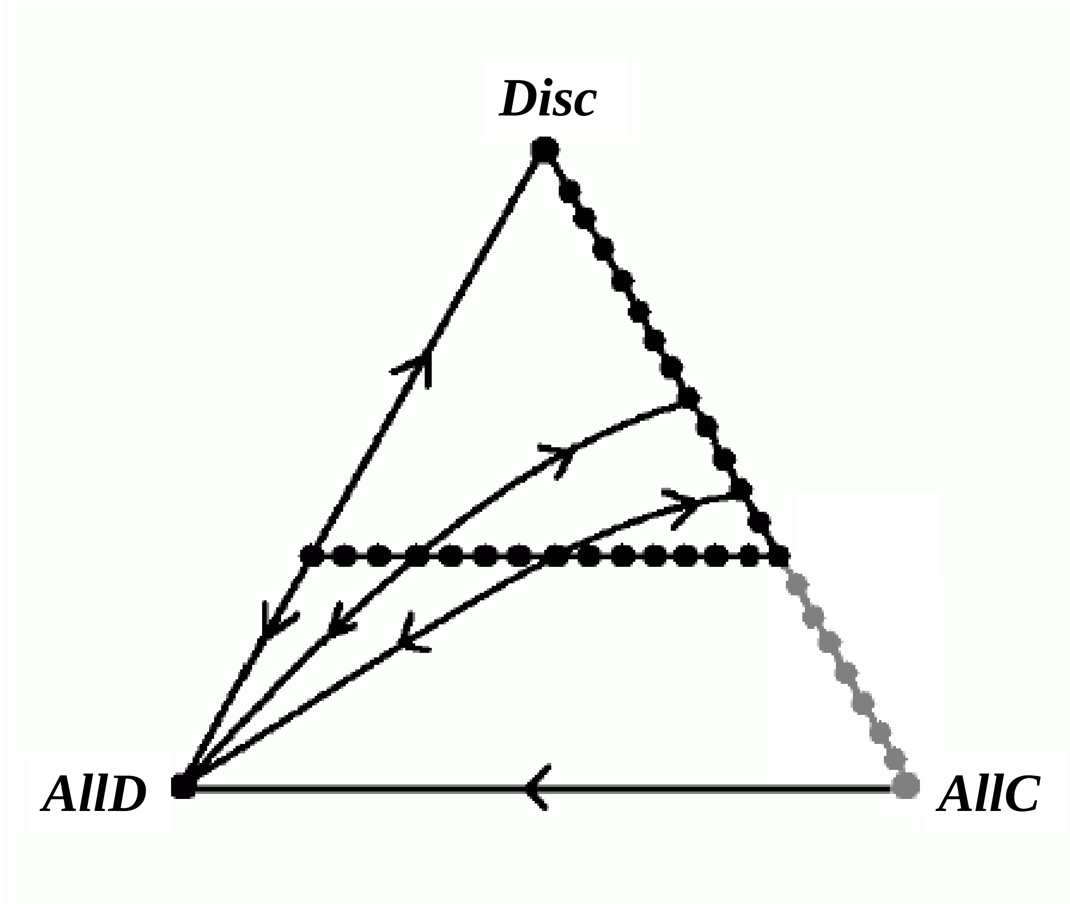
decision based on opponent's score:

AllC (cooperators, always help)

AllD (defectors, never help)

Disc (discriminators, help iff opponent is Good)

$\rightarrow$ continuum of stable equilibria where discriminators and cooperators coexist


Disc
AllD
AllC

Problem:

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Theoretical Biology 224 (2003) 115–126

Journal of Theoretical Biology

www.elsevier.com/locate/jtbi

A tale of two defectors: the importance of standing for evolution of indirect reciprocity

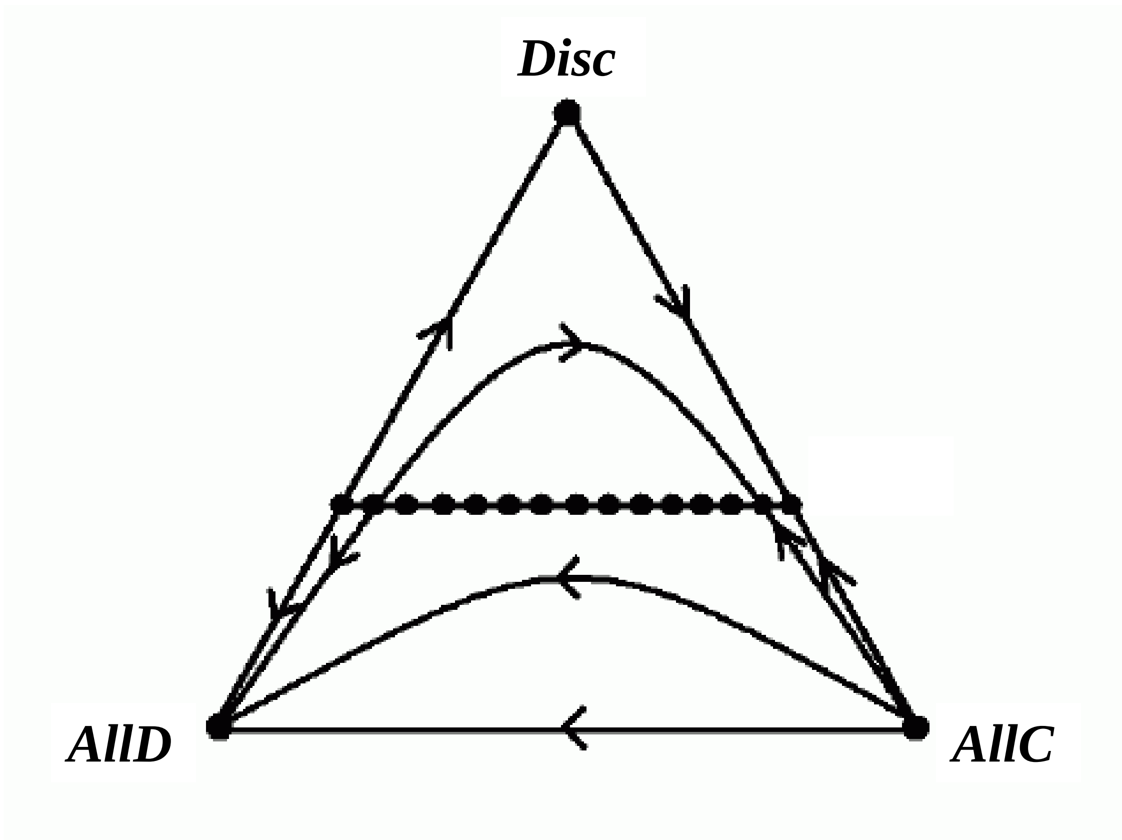
Karthik Panchanathan*, Robert Boyd

Department of Anthropology, University of California, Haines Hall 341 Box 951553, Los Angeles, CA 90095, USA

Received 14 August 2002; received in revised form 2 April 2003; accepted 2 April 2003

2. Image scoring fails when errors occur

However, Nowak and Sigmund's (1998a) claim that indirect reciprocity can be based on an image-scoring strategy hinges completely on the assumption that agents never commit errors. That is, individuals that intend to cooperate always do so and those that intend to defect always do so. In this section, by introducing errors, we show that cooperation based on image scoring cannot evolve; the defect equilibrium is reached quickly and deterministically.


Disc
AllD
AllC

Panchanathan & Boyd (2003):

“*Image scoring requires only that agents be able to acquire information as to the actions of others. [...] Our analysis shows that* **additional information is necessary to stabilize indirect reciprocity**. *Specifically,* **individuals must be able to infer motivations from observed defections**, *parsing them into those that are justified and those that are unjustified.*”

Idea: *Standing rule*
Defecting against *Bad* players is justified and does not change your reputation from *Good* to *Bad*.

Higher-order assessment rules

later literature: only higher-order assessment rules

Brandt & Sigmund (2004, 2005, 2006); Chalub et al. (2006); Ohtsuki (2004); Ohtsuki & Iwasa (2004, 2006, 2007); Ohtsuki et al. (2009); Uchida & Sigmund (2009)

Higher-order assessment rules

later literature: only higher-order assessment rules

Brandt & Sigmund (2004, 2005, 2006); Chalub et al. (2006); Ohtsuki (2004); Ohtsuki & Iwasa (2004, 2006, 2007); Ohtsuki et al. (2009); Uchida & Sigmund (2009)

Problems with higher-order assessment rules:

- high cognitive load

- high informational requirements

- weak experimental evidence

own approach:

Return to *image scoring*, but assume **multiple, private** observations of opponent's previous behavior.

- large population, asynchronous entry, infinitely many rounds, implementation errors

- two time-scales: score dynamics (fast), evolution/learning (slow)

When A encounters B and is selected as donor, he samples (recalls, is informed of) $n \geq 1$ actions chosen by B in the past.

Strategies s_k: k-$Disc$ players tolerate *at most* k defections, i.e. they intend to cooperate iff opponent cooperated at least $n-k$ out of n times.

k . . . **tolerance level**: $k=-1$ is $AllD$, $k=n$ is $AllC$

f_k . . . **cooperation function** of k-$Disc$

$f_k(p)$ = probability of playing C if opponent's past cooperation rate is p

p_i . . . equilibrium cooperation rate of incumbent i-$Disc$

p_i = maximal stable fixed point of $p' = f_i(p)$

$\pi(m|i)$ ... payoff of mutant m-$Disc$ in a population of i-$Disc$ incumbents

p_i ... incumbent's cooperation rate

$f_m(p_i)$ ... mutant's cooperation rate against incumbent

$f_m(p_i)c$ ... mutant's expected costs of helping

$f_i(f_m(p_i))$ ... incumbent's cooperation rate against mutant

$f_i(f_m(p_i))b$ ... mutant's expected revenue from being helped

$$\Rightarrow \pi(m|i) = f_i(f_m(p_i))b - f_m(p_i)c$$

What do we know about the coop. functions f_k ?

$\bar{\alpha} = 1 - \alpha$, $\alpha \ldots$ error rate $C \to D$

$f_{-1} \equiv 0$ $(AllD)$, $f_n \equiv \bar{\alpha}$ $(AllC)$

$f_k(p) = \bar{\alpha} B(k, n, 1-p), \quad B \ldots$ binomial cdf

$f_k(p) = \bar{\alpha} \sum_{h=0}^{k} \binom{n}{h} (1-p)^h p^{n-h}$

What do we know about the coop. functions f_k ?

$\bar{\alpha} = 1 - \alpha$, $\alpha \ldots$ error rate $C \to D$

$f_{-1} \equiv 0$ $(AllD)$, $f_n \equiv \bar{\alpha}$ $(AllC)$

$f_k(p) = \bar{\alpha} B(k, n, 1-p)$, $\quad B \ldots$ binomial cdf

$f_k(p) = \bar{\alpha} \sum_{h=0}^{k} \binom{n}{h} (1-p)^h p^{n-h}$

use regularized beta function

$\rightarrow \quad f_k(p) = \bar{\alpha}(n-k) \binom{n}{k} \int_0^p t^{n-k-1} (1-t)^k dt$

$$f_k(p) = \bar{\alpha}(n-k)\binom{n}{k}\int_0^p t^{n-k-1}(1-t)^k dt$$

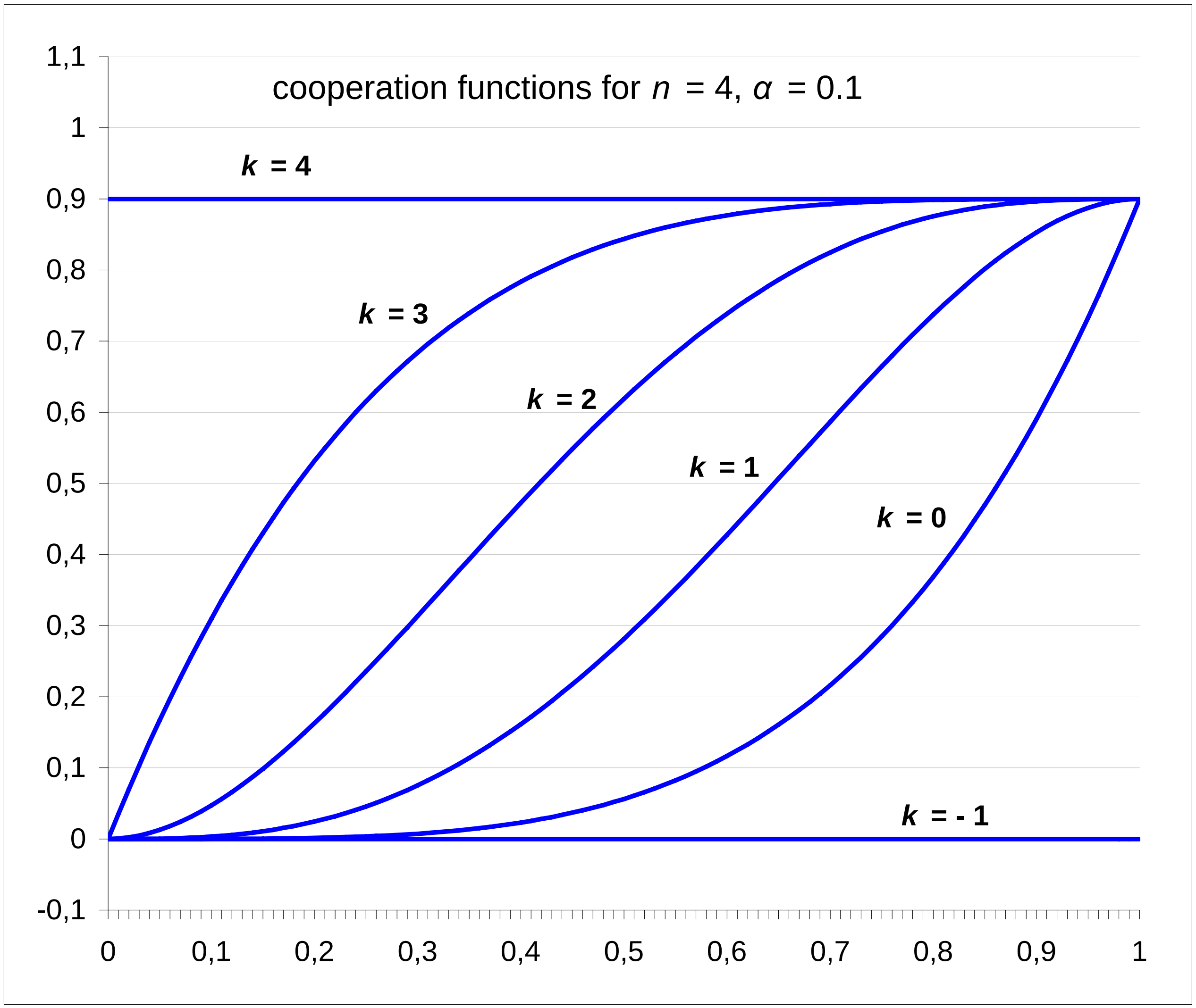
$$\Rightarrow f_0(p) = \bar{\alpha}p^n,\ f_{n-1}(p) = \bar{\alpha}(1-(1-p)^n)$$

and $f_k'(p) = \bar{\alpha}(n-k)\binom{n}{k}p^{n-k-1}(1-p)^k$

for $n \geq 3$ and $1 \leq k \leq n-2$ we get

$$f_k'(p) = \begin{cases} > 0 & \ldots \quad 0 < p < 1 \\ = 0 & \ldots \quad p = 0 \vee p = 1 \end{cases}$$

$$f_k''(p) = \begin{cases} > 0 & \ldots \quad 0 < p < \frac{n-k-1}{n-1} \\ < 0 & \ldots \quad \frac{n-k-1}{n-1} < p < 1 \\ = 0 & \ldots \quad p = 0 \vee p = 1 \vee p = \frac{n-k-1}{n-1} \end{cases}$$

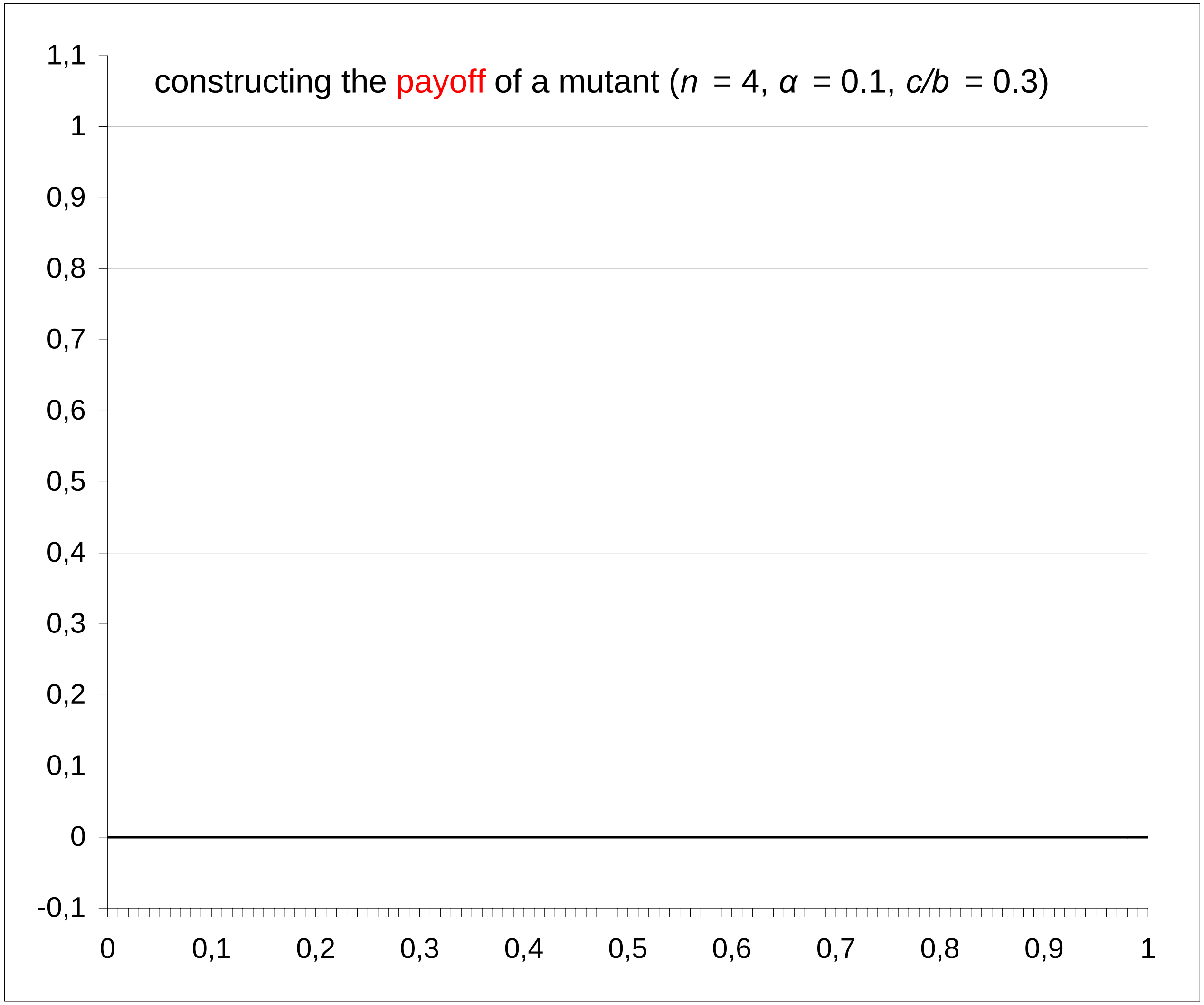


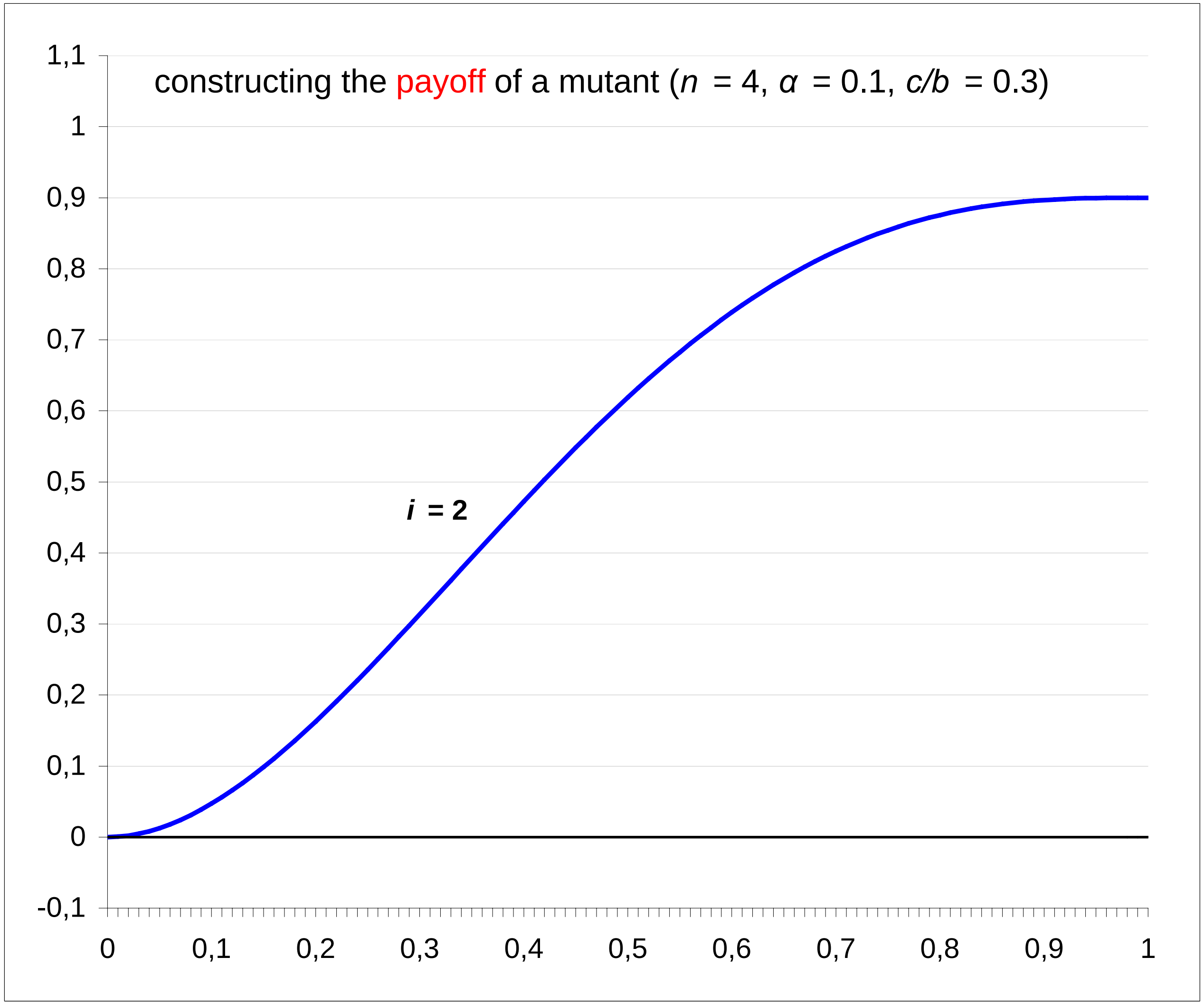
cooperation functions for n = 4, α = 0.1
k = 4
k = 3
k = 2
k = 1
k = 0
k = - 1
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1

When can a mutant invade?

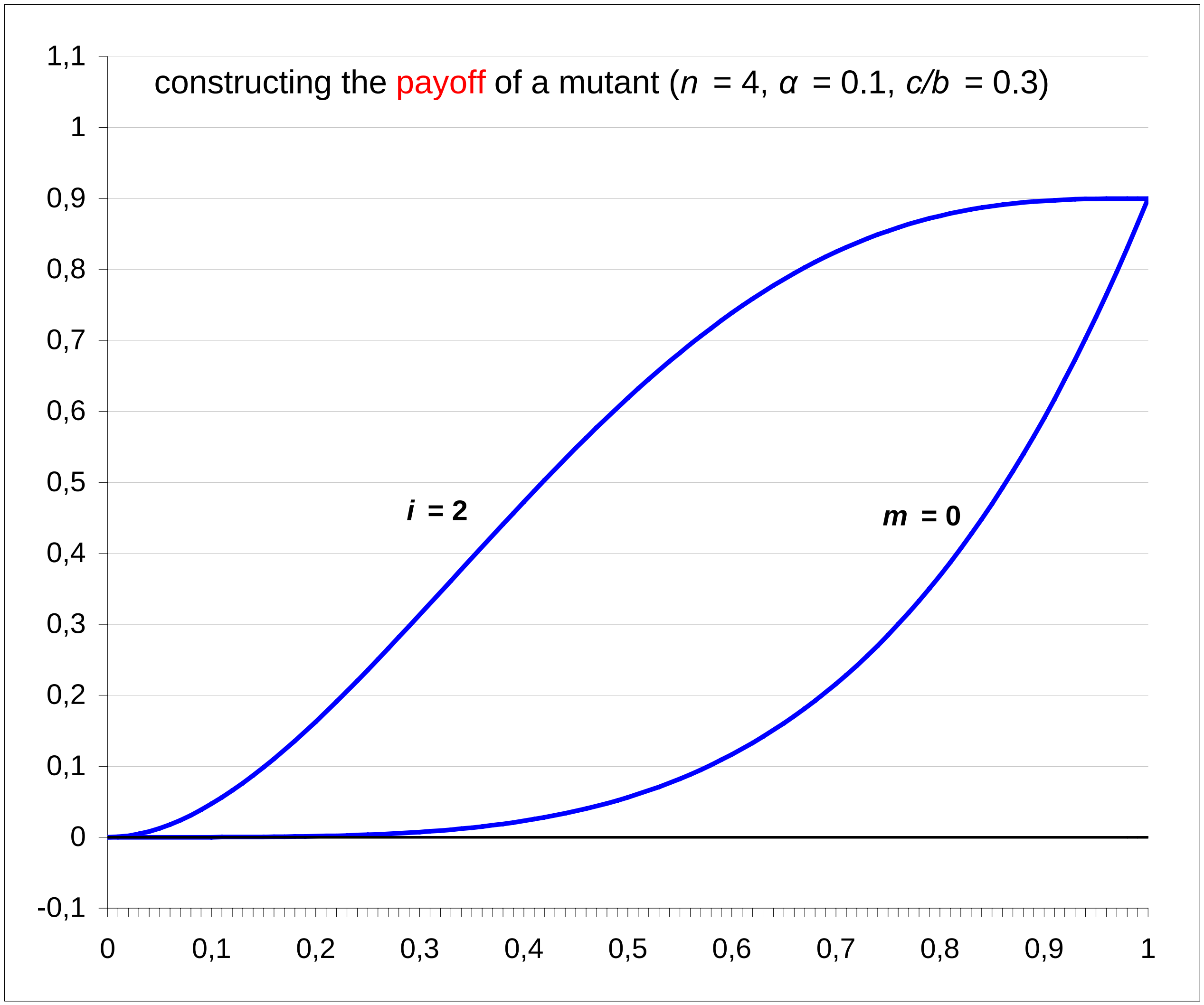
$$\pi(m|i) = f_i(f_m(p_i))b - f_m(p_i)c$$

$$\widehat{\pi}(m|i) = b^{-1}\pi(m|i) = f_i(f_m(p_i)) - f_m(p_i)c/b$$

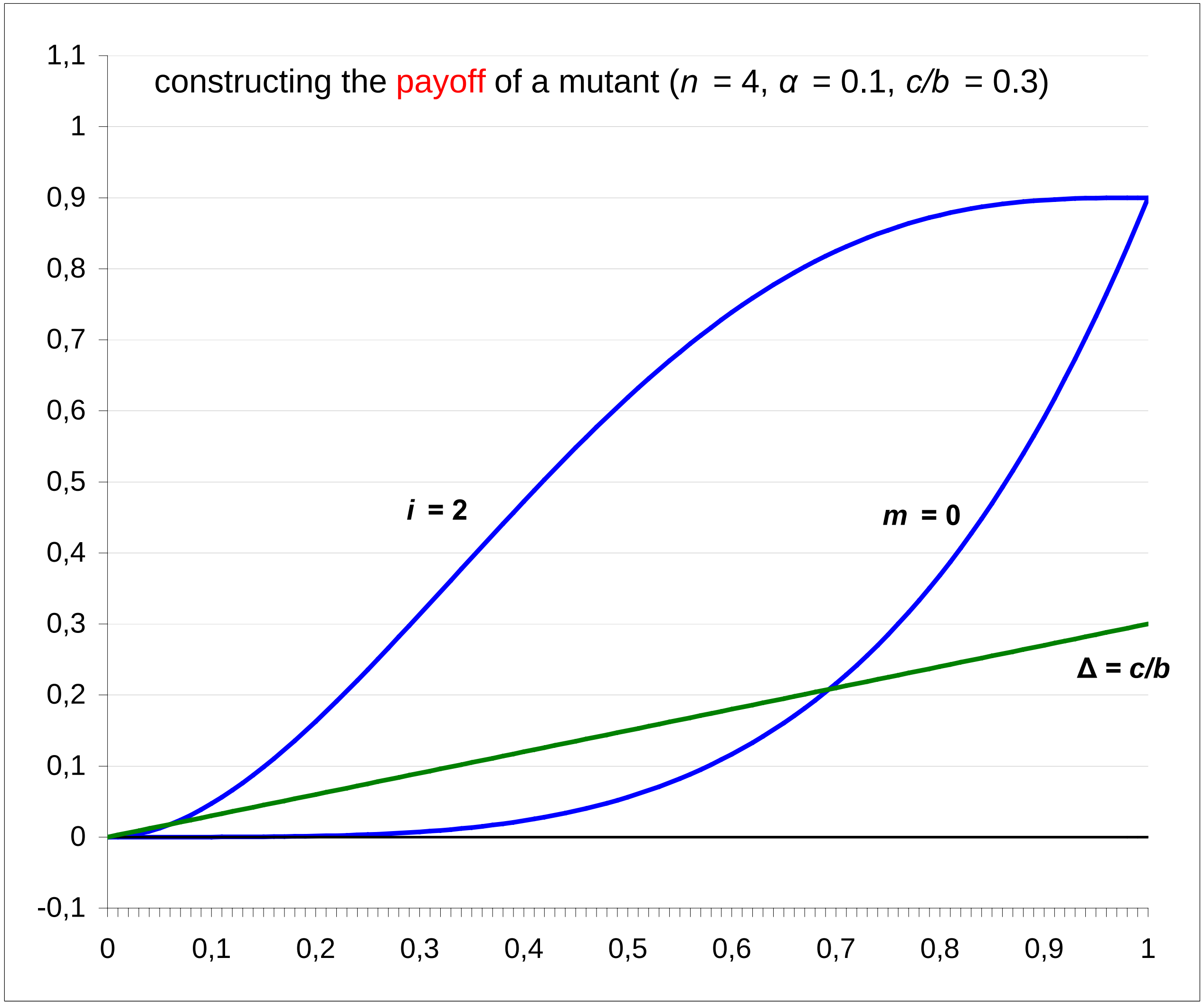

constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1

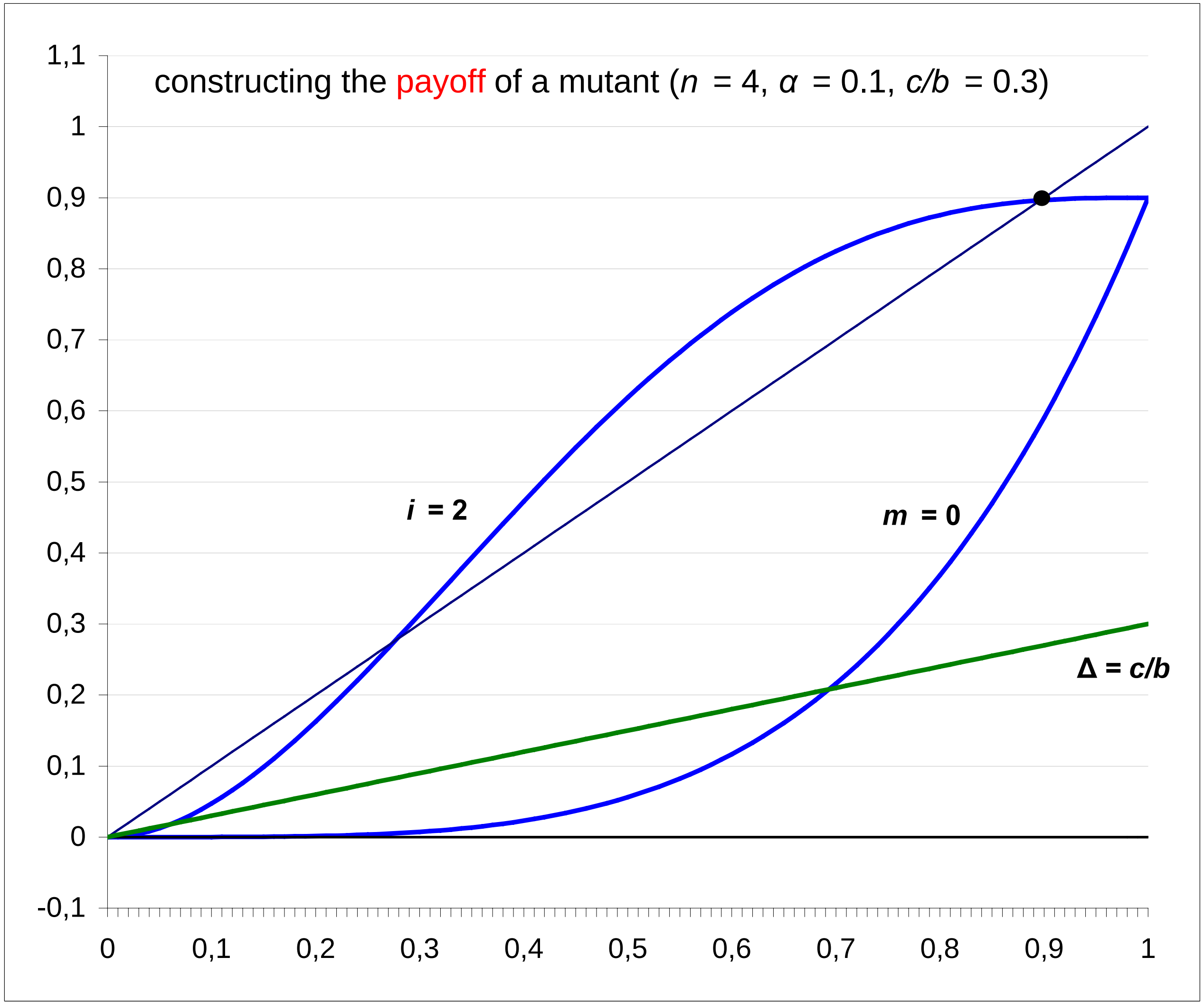


constructing the payoff of a mutant (*n* = 4, *α* = 0.1, *c/b* = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2

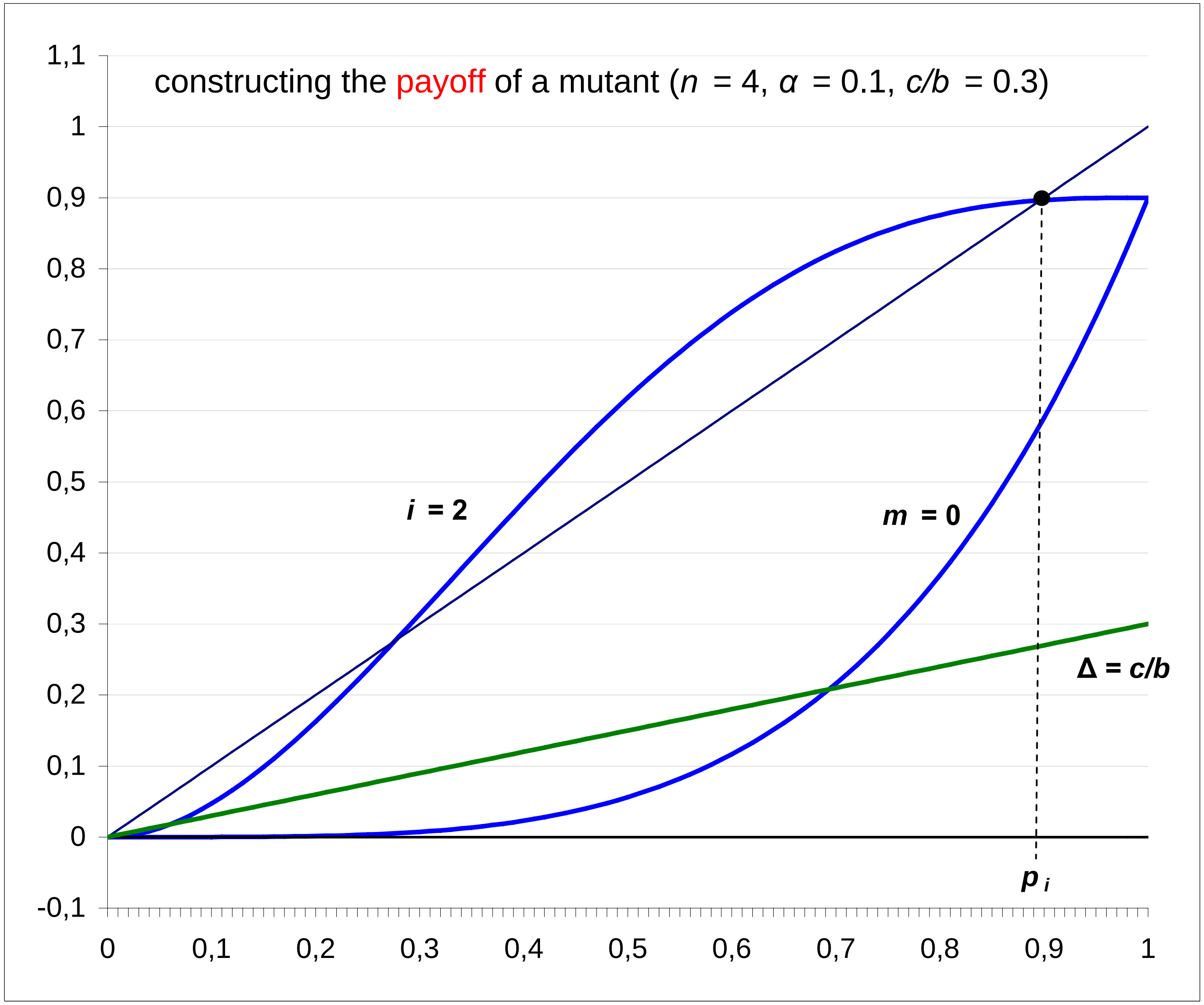


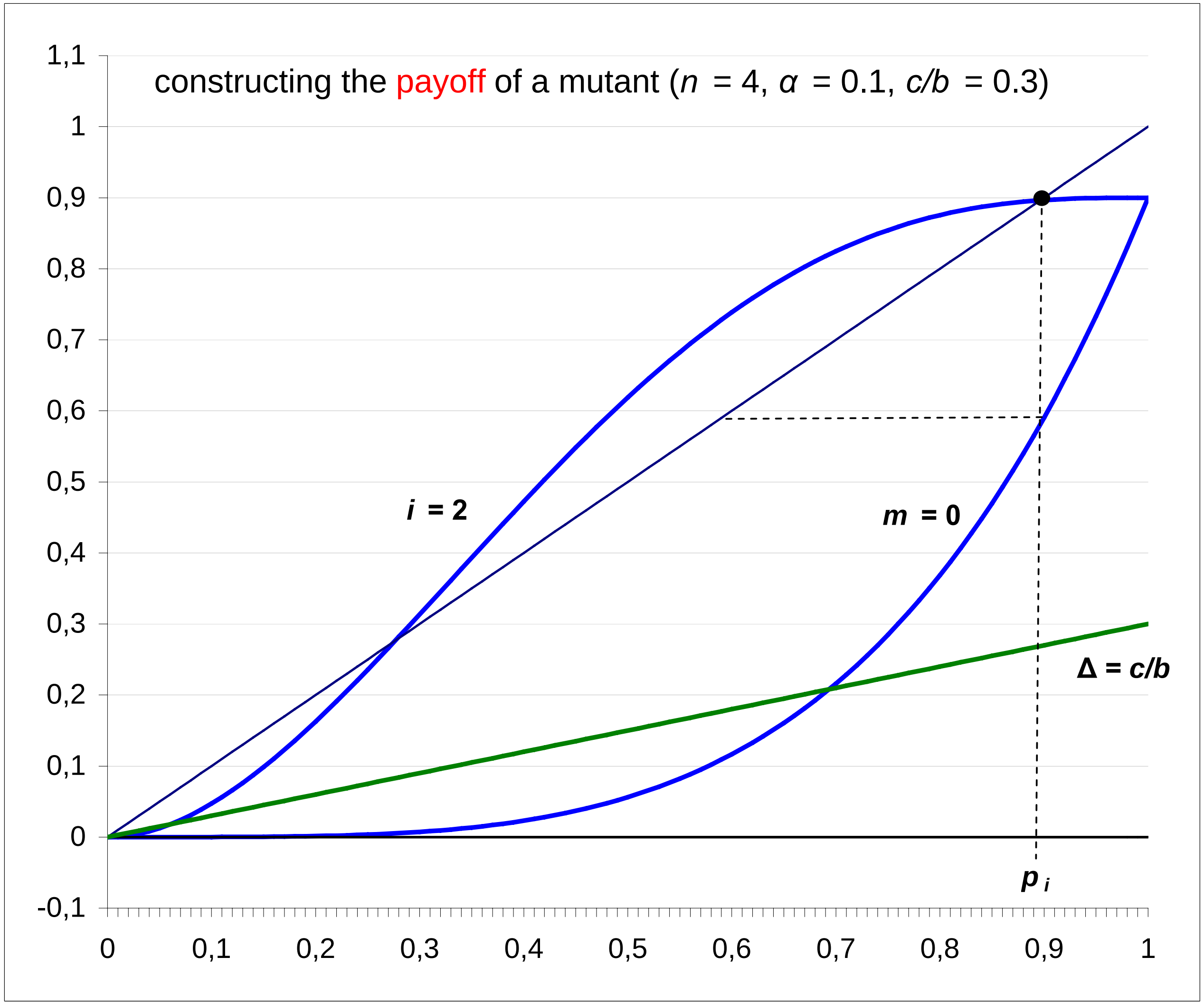
constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2
m = 0


constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2
m = 0
Δ = c/b

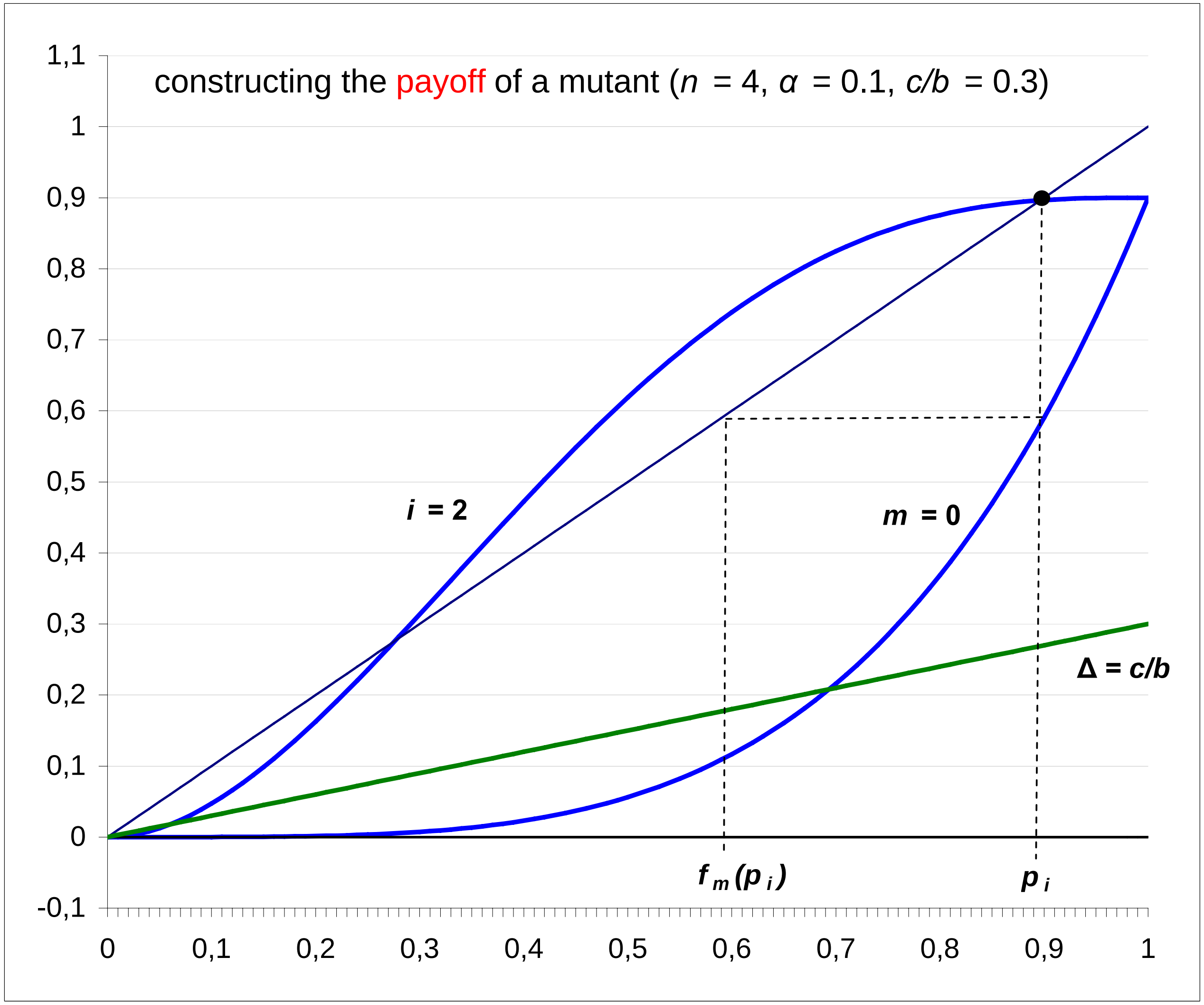


constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2
m = 0
Δ = c/b

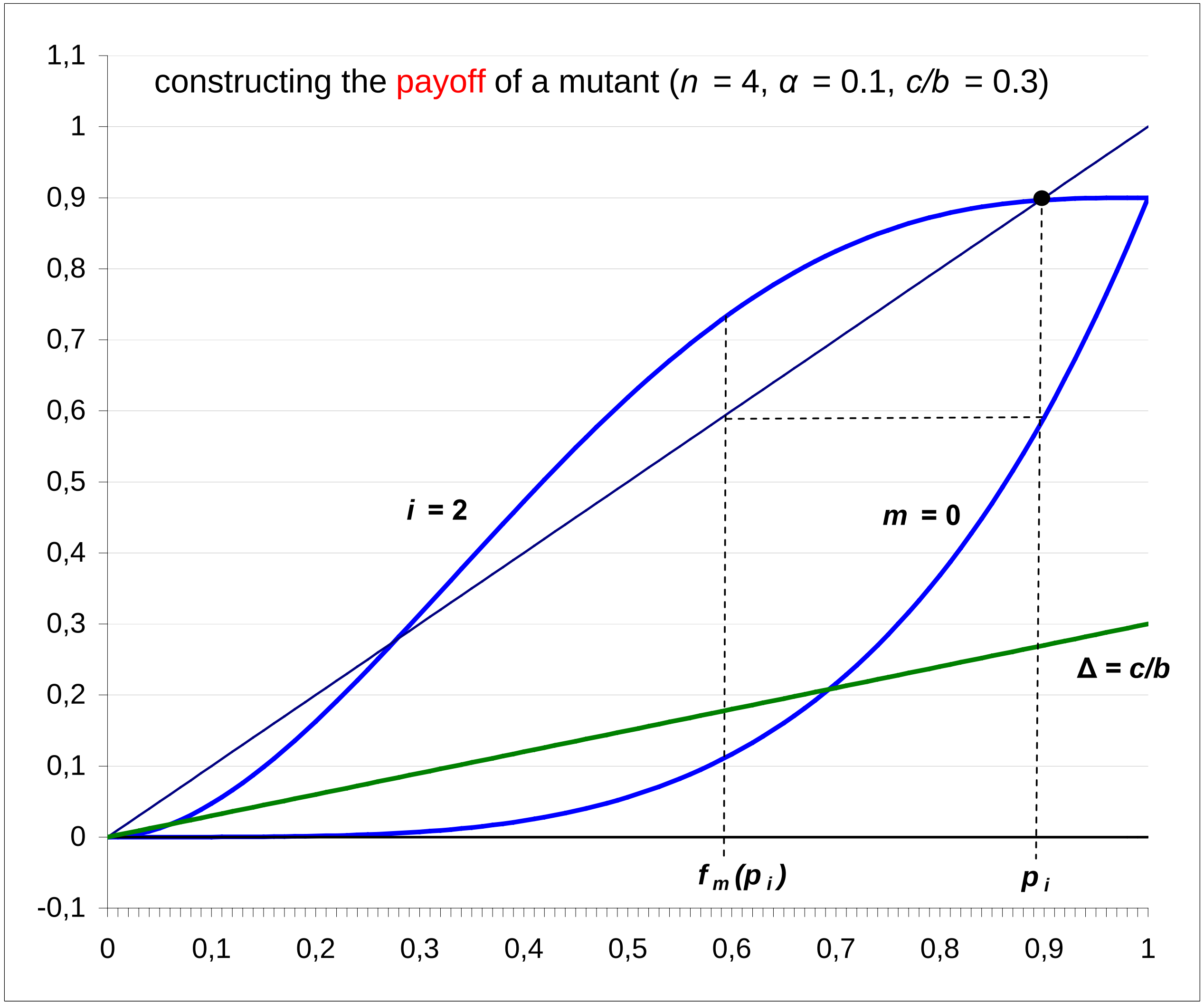

constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
i = 2
m = 0
Δ = c/b
p_i
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1



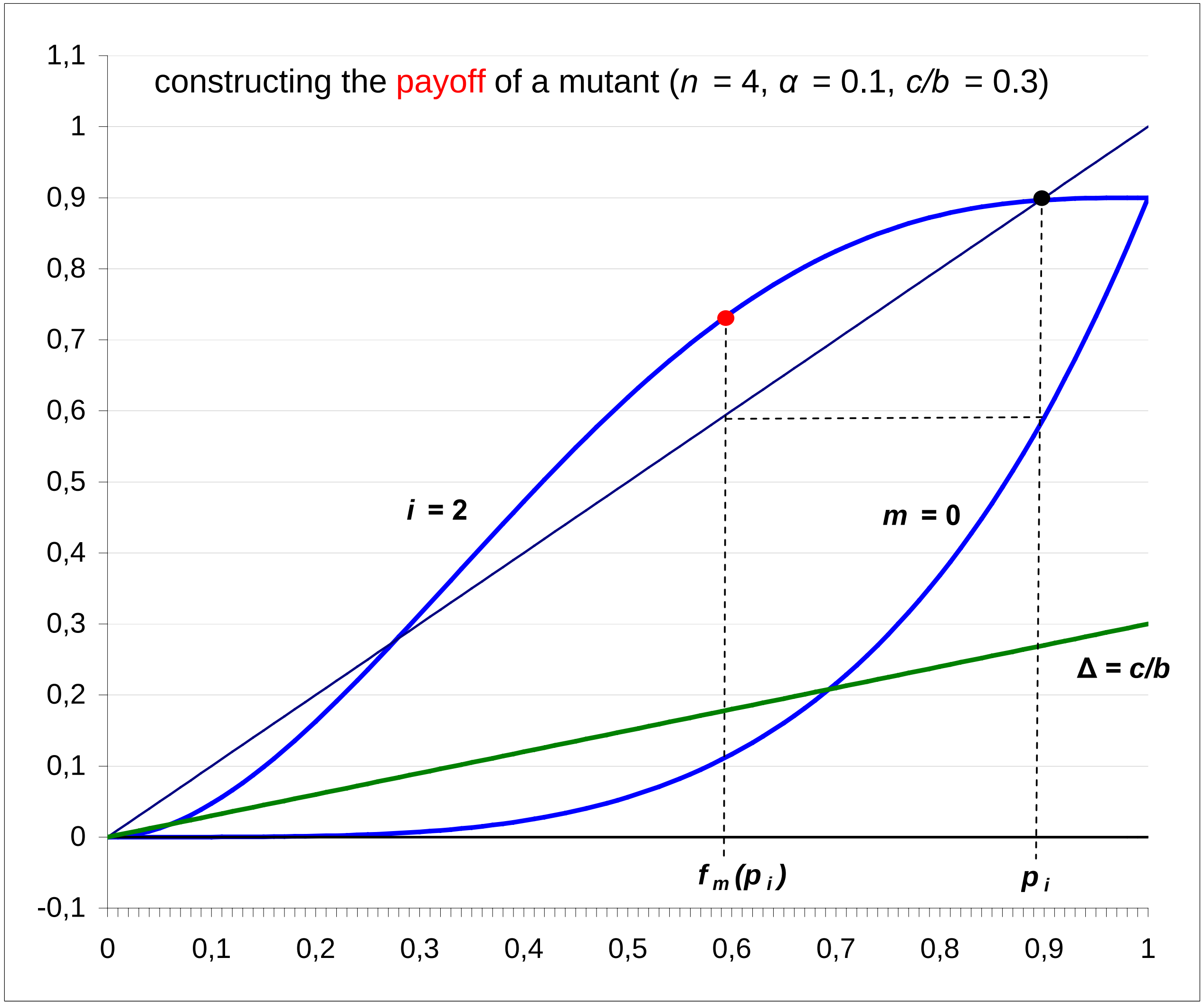
constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
i = 2
m = 0
Δ = c/b
p i
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1



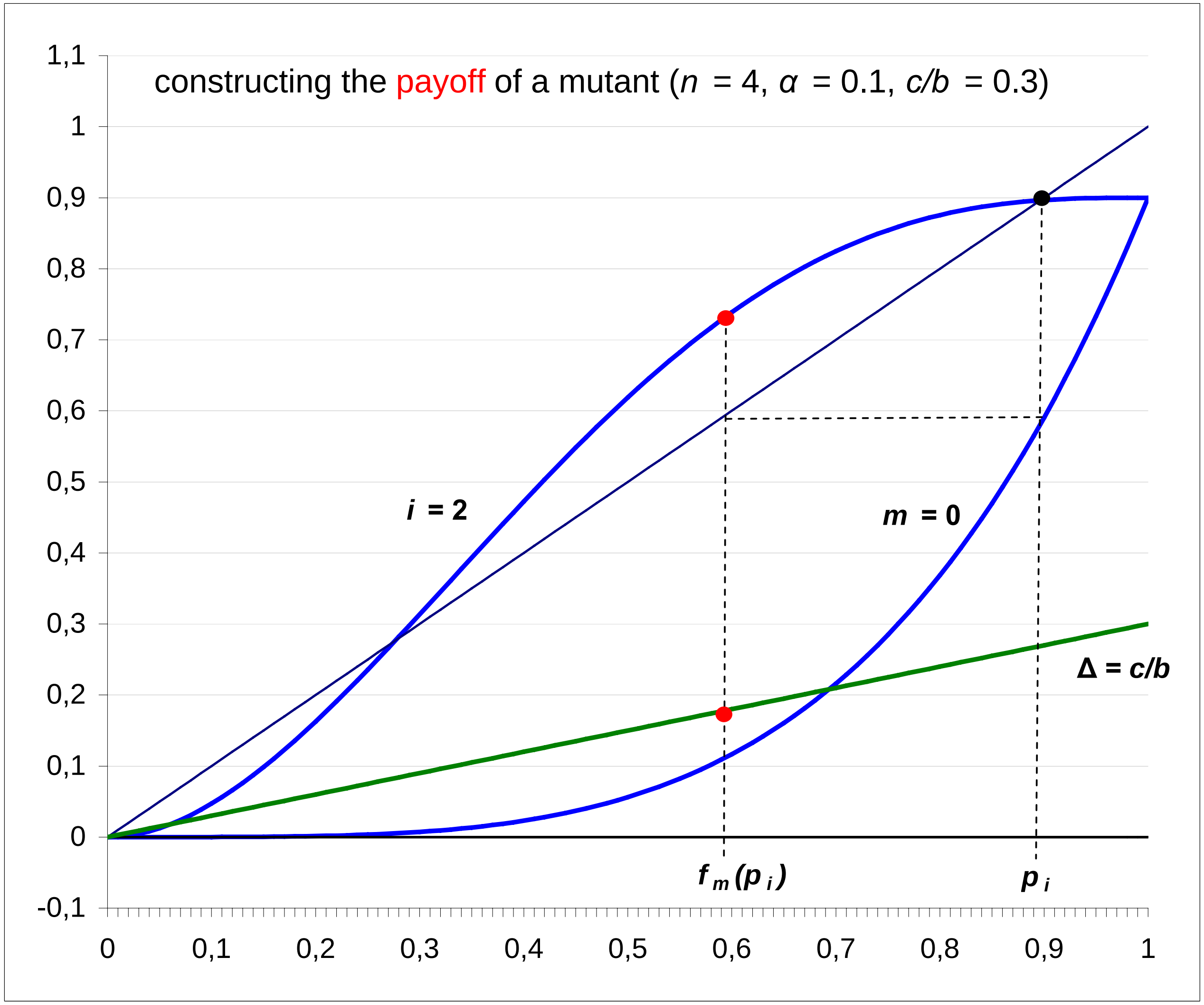
constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
i = 2
m = 0
Δ = c/b
f m (p i)
p i
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1



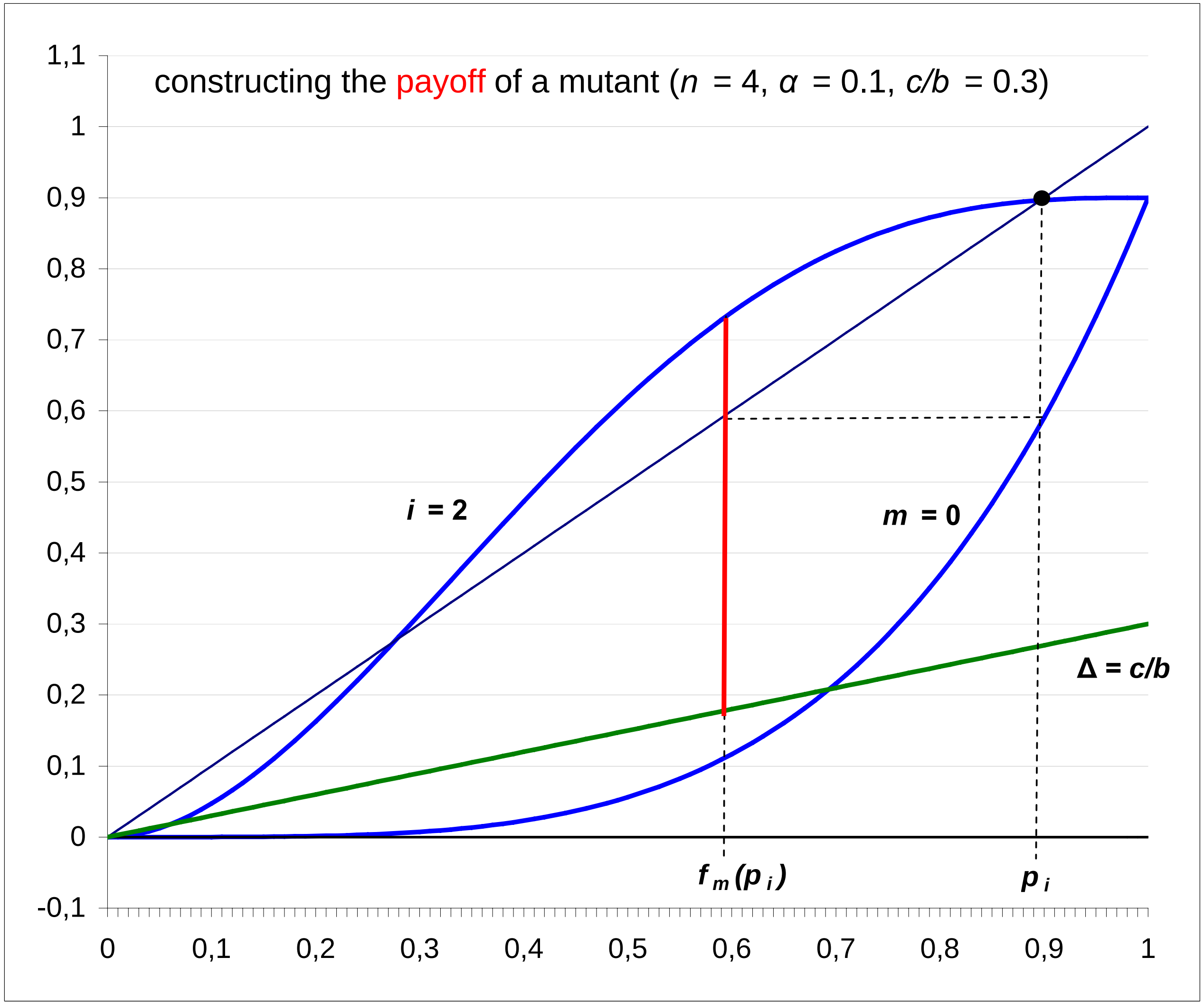
constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2
m = 0
Δ = c/b
f m (p i)
p i



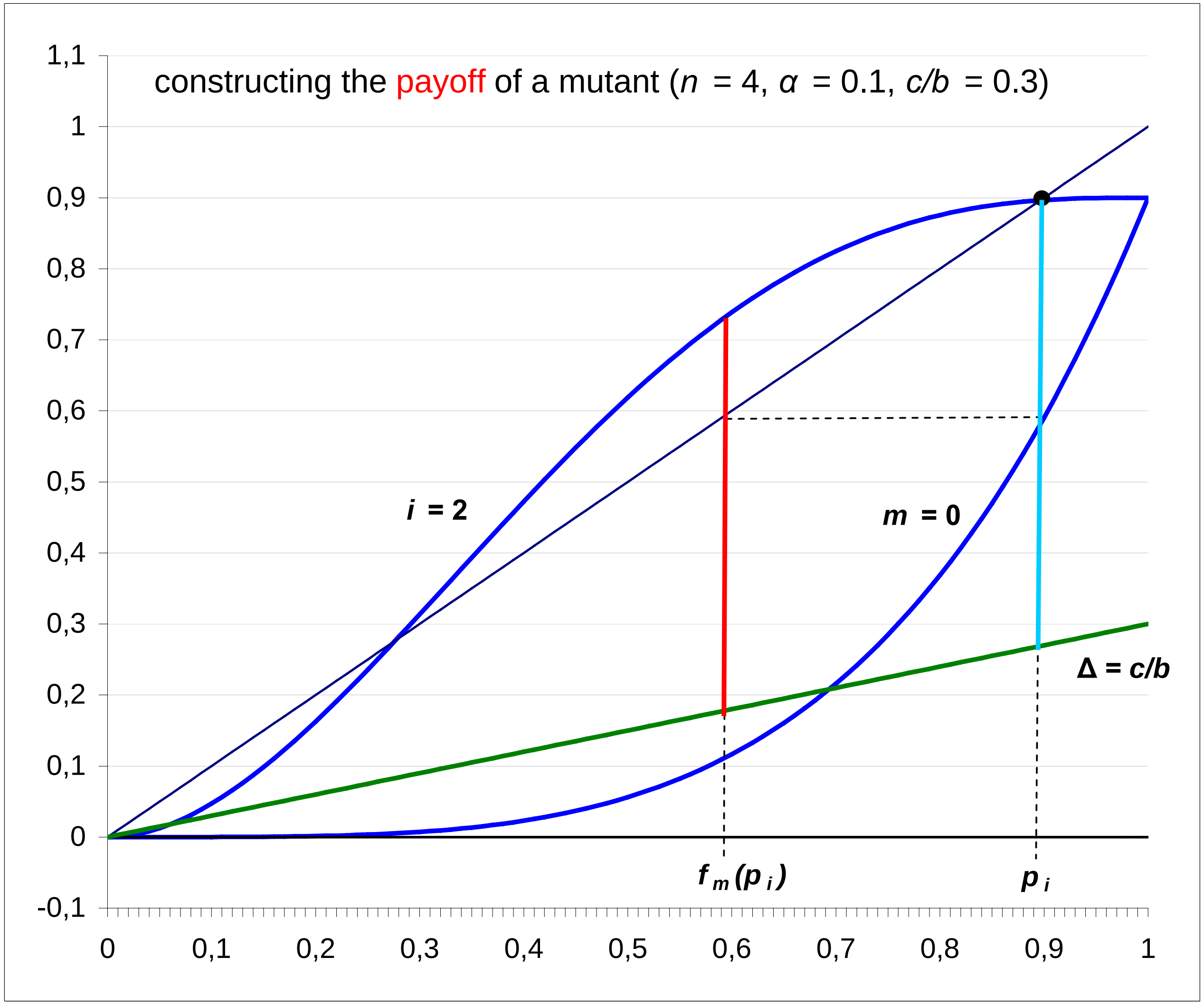
constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2
m = 0
Δ = c/b
f m (p i)
p i

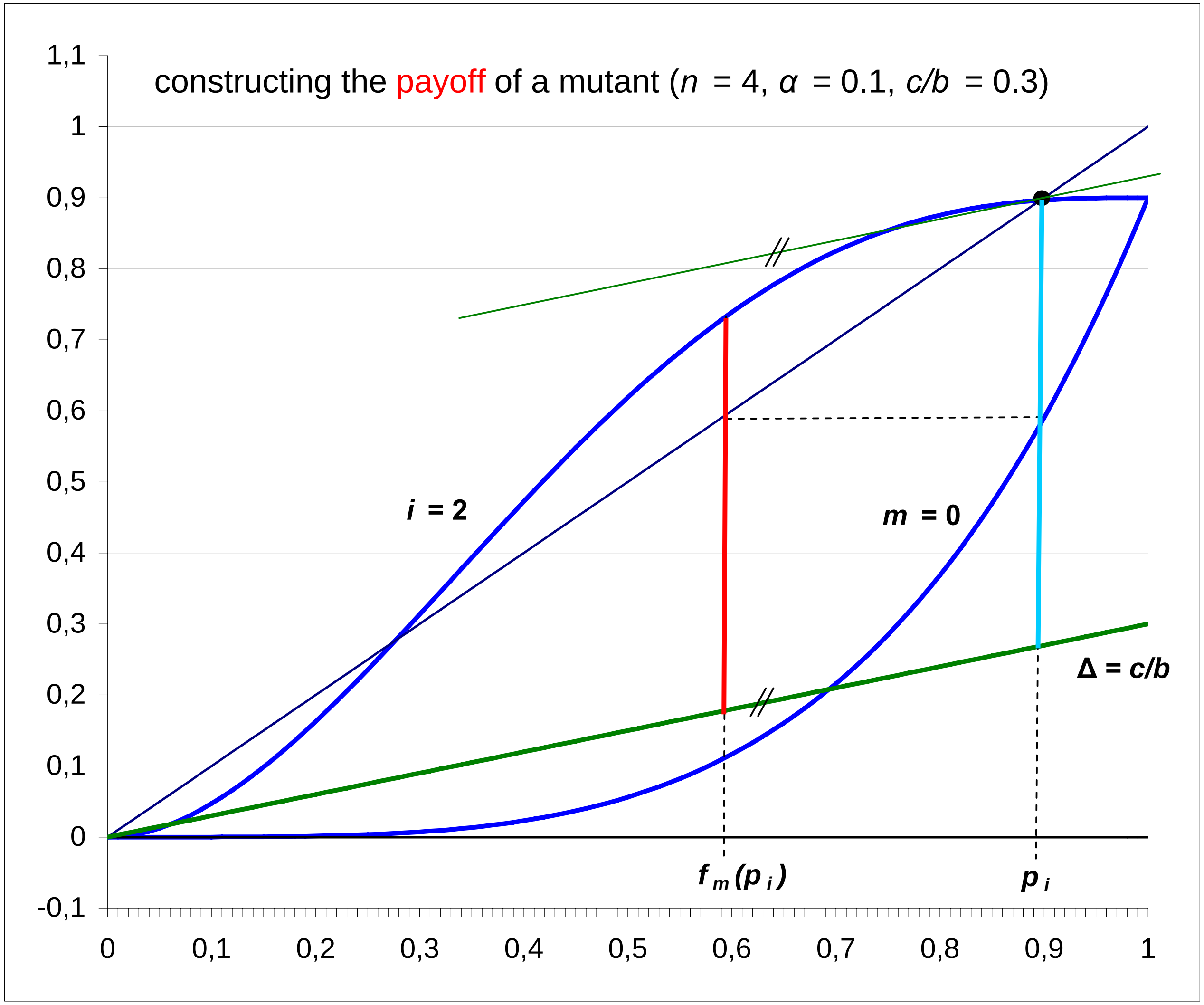


constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
i = 2
m = 0
Δ = c/b
f m (p i)
p i
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1

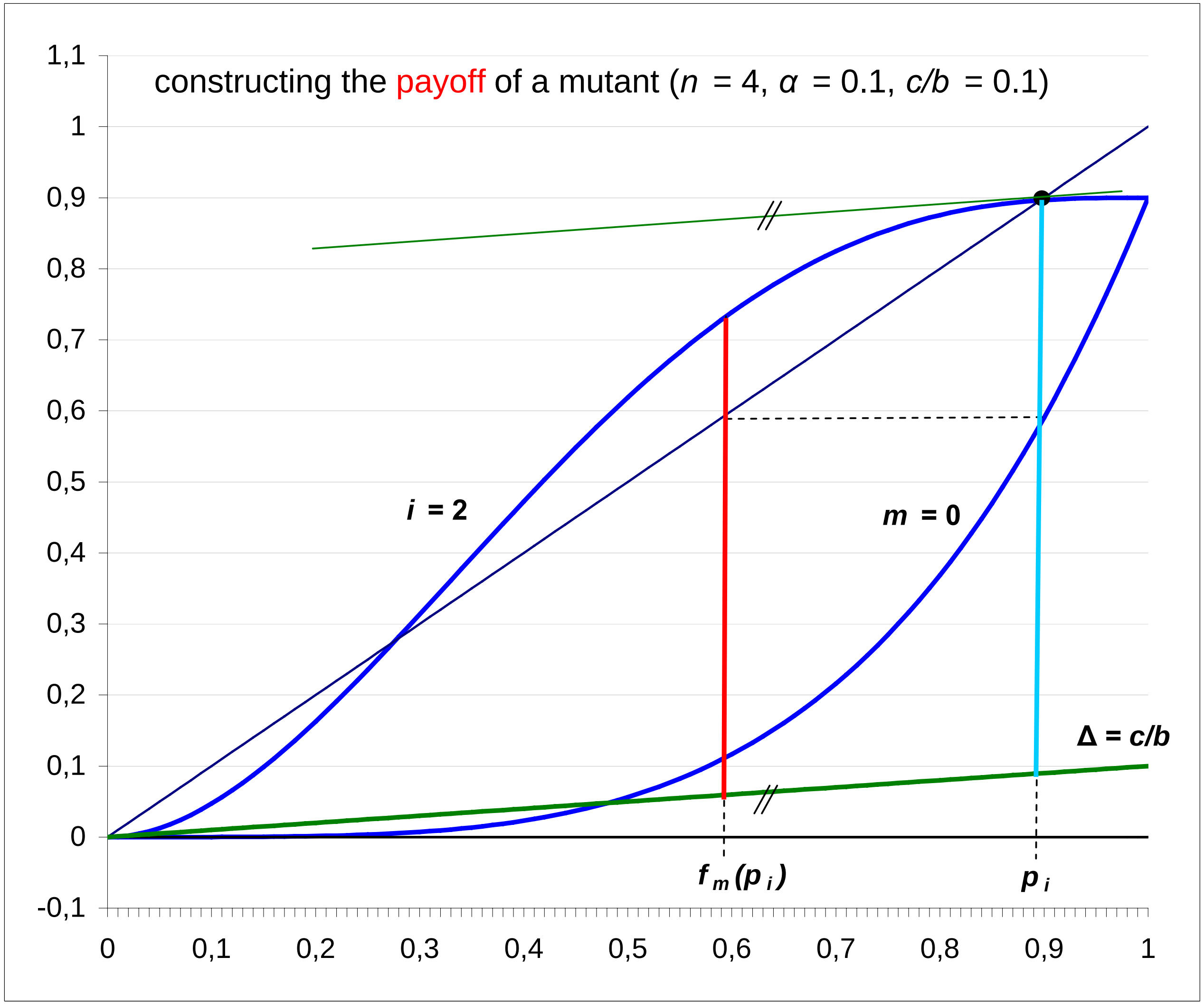


constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2
m = 0
Δ = c/b
f m (p i)
p i


constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
i = 2
m = 0
Δ = c/b
f m (p i)
p i
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1



constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.3)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
i = 2
m = 0
Δ = c/b
f m (p i)
p i
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1



constructing the payoff of a mutant (n = 4, α = 0.1, c/b = 0.1)
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1
i = 2
m = 0
$\Delta = c/b$
$f_m(p_i)$
p_i

Result:

For each k-$Disc$ with $0 \leq k \leq n-1$ which is able to sustain cooperation, there exists a cost-benefit ratio $0 < r_k < 1$ and an $\epsilon > 0$ such that k-$Disc$ is an ESS if $|c/b - r_k| < \epsilon$.

Result:

For each k-$Disc$ with $0 \leq k \leq n-1$ which is able to sustain cooperation, there exists a cost-benefit ratio $0 < r_k < 1$ and an $\epsilon > 0$ such that k-$Disc$ is an ESS if $|c/b - r_k| < \epsilon$.

Case: $n = 2$

3 strategies: $x \ldots AllC$, $y \ldots AllD$, and $z \ldots Disc$
($Disc =$ "tolerant" discriminator, $k = 1$)

assume c, b, α such that $Disc$ is ESS

$p \ldots$ **discriminator's cooperation rate** in state X,
$X = (x, y, z)$

can show: $p = 1 - \frac{1}{2\bar{\alpha}z} + \sqrt{\left(1 - \frac{1}{2\bar{\alpha}z}\right)^2 + (1-\alpha^2)\frac{x}{z}}$

$p\ldots$ **discriminator's cooperation rate** in state X, $X = (x, y, z)$

can show: $p = 1 - \frac{1}{2\bar{\alpha}z} + \sqrt{\left(1 - \frac{1}{2\bar{\alpha}z}\right)^2 + (1-\alpha^2)\frac{x}{z}}$

state-dependent payoff matrix:

	$AllC$	$Disc$	$AllD$
$AllC$	$\bar{\alpha}(b-c)$	$\bar{\alpha}[(1-\alpha^2)b-c]$	$-\bar{\alpha}c$
$Disc$	$\bar{\alpha}[b-(1-\alpha^2)c]$	$p(2-p)\bar{\alpha}(b-c)$	0
$AllD$	$\bar{\alpha}b$	0	0

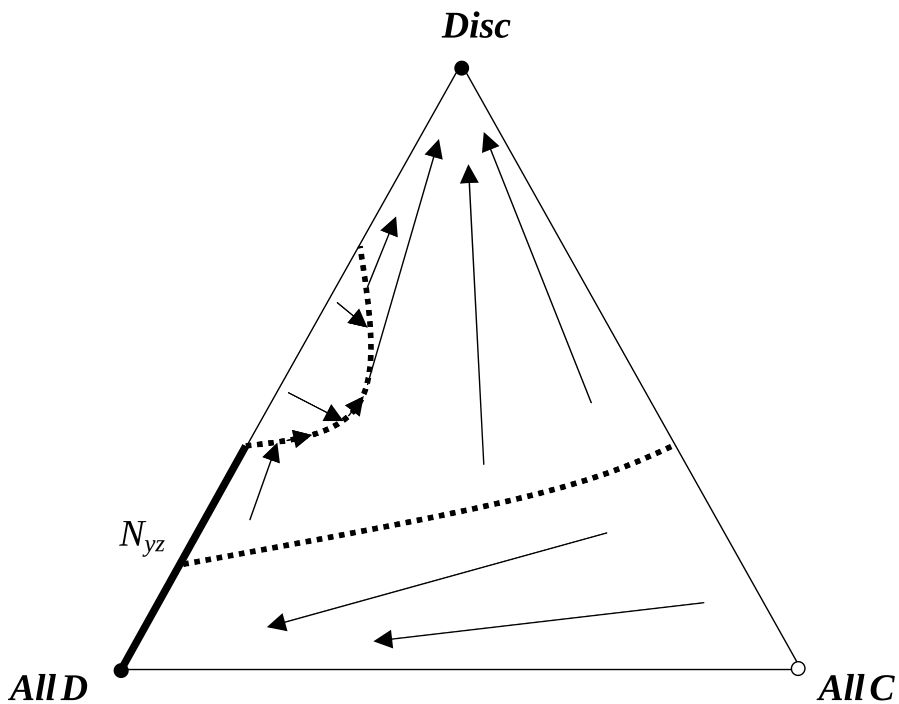
$p\ldots$ **discriminator's cooperation rate** in state X,
$X = (x, y, z)$

can show: $p = 1 - \frac{1}{2\bar{\alpha}z} + \sqrt{\left(1 - \frac{1}{2\bar{\alpha}z}\right)^2 + (1-\alpha^2)\frac{x}{z}}$

state-dependent payoff matrix:

	$AllC$	$Disc$	$AllD$
$AllC$	$\bar{\alpha}(b-c)$	$\bar{\alpha}[(1-\alpha^2)b-c]$	$-\bar{\alpha}c$
$Disc$	$\bar{\alpha}[b-(1-\alpha^2)c]$	$p(2-p)\bar{\alpha}(b-c)$	0
$AllD$	$\bar{\alpha}b$	0	0

Assume boundedly rational learning:
best response dynamics: $\dot{X} \in BR(X) - X$



Disc
N_{yz}
All D
All C

In the long run, all individuals are discriminators and (mostly) cooperate.

In the long run, all individuals are discriminators and (mostly) cooperate.

Thank you for your attention!