For a Function That Is Not Convex, Where You Start Decides the Answer

— $f(x) = x^4 - 3x^2 + 0.5x$ has a left minimum at -1.264459 ($f = -2.872457$) and a right minimum at 1.180734 ($f = -1.648425$), with a maximum between at 0.083725 — the separator is whether the second derivative is positive everywhere or negative somewhere — [Paper 771]

Yuuki YAMAGISHI (山岸 勇輝)
Yamagishi Complexity Torus Lab, Japan
September 2026

— *From thinking to syncing.* —

Abstract

Gradient descent seeks a minimum by walking downhill. If the landscape is a single valley, every start reaches the same bottom.When there are two valleys, how much the start decides the destination is counted.No new theorem or law is claimed. First, a non-convex function has two bottoms——a left minimum at -1.264459 and a right minimum at 1.180734, with a maximum at 0.083725 between (Section 2). **Second, the two bottoms differ in depth ——f is -2.872457 on the left and -1.648425 on the right; the global minimum is on the left** (Section 2). **Third, and this is the core. The start splits the destination, and nearly half reach the shallower bottom——of 1001 starts, a share 0.520480 goes left and 0.479520 right** (Section 3). **Fourth, the boundary is set by the position of the maximum——the formula gives 0.520931, differing from the grid share by 0.000452** (Section 3). **Fifth, if convex, the start does not matter——all 1001 points reach -0.226699, with a spread of 1.027×10^{-15}** (Section 3). **Sixth, the separator is whether the second derivative is positive everywhere or negative somewhere——"the minimum was found by gradient descent" does not say which minimum until one says whether the function is convex** (Section 4).
Keywords: convexity, local minimum, gradient descent, dependence on the start, basin of attraction, global minimum.

> **Scope of this paper (scope note)**: **No new theorem or law is claimed——that local minima of convex functions are global, that for non-convex functions the destination of gradient descent is set by the basin of attraction of the start, and that the basin boundary is a maximum are all known (Boyd and Vandenberghe 2004, Nocedal and Wright 2006).One dimension——only functions of one variable are considered; multidimensional landscapes with saddles are not treated.One method——only plain gradient descent with step 0.01 is used; stochastic methods, momentum and divergence at large steps are not treated.Uniform starts——1001 equally spaced points on $[-2, 2]$ are the starts.No real problem is claimed——particular learning or optimisation problems are not treated.Relation to earlier papers**: **Paper 734 showed that what caps the step size is not the size of the gradient but the curvature in the steepest direction ——its scope note stated that it was limited to quadratic functions and would not treat non-convex surfaces or saddles. There every start reached the same point. Here, on a non-convex surface, nearly half the starts reach a bottom that is not the global minimum.Paper 507 is also cited——there a convex function returned to itself after two transforms. Here a convex function sends every start to one point, and a non-convex one lets the start split the answer.**What is added is **tabulating stationary points with the signs of their second derivatives, counting the destinations of 1001 starts and checking them against the position of the maximum, showing that for a convex function the spread stays at 1.027×10^{-15}**, and **placing the separator on whether the second derivative turns negative somewhere**.

1. Notation and the Question

Setting: **take a non-convex function f and a convex function g, and descend with step $\eta = 0.01$.**

$$f(x) = x^4 - 3x^2 + 0.5x, \qquad g(x) = x^4 + x^2 + 0.5x, \qquad x_{k+1} = x_k - \eta\, f'(x_k)$$

$f''(x) = 12x^2 - 6$ turns negative around $x = 0$, while $g''(x) = 12x^2 + 2$ is positive everywhere. Starts are 1001 equally spaced points on $[-2, 2]$.

Question: **is the minimum gradient descent reaches decided by the function or by the start?**

2. Two Bottoms

Computation 2.1: **$f'(x) = 0$ was solved by bisection, giving stationary points, values and second derivatives.**

stationary point	position	f	second derivative
left minimum	-1.264459	-2.872457	13.186278
maximum	0.083725	0.020882	-5.915882
right minimum	1.180734	-1.648425	10.729604

Observation 2.2

a non-convex function has two bottoms of different depths.The left bottom is -2.872457 and the right -1.648425, so the global minimum is on the left——between the two bottoms stands a maximum at 0.083725 where the second derivative is negative.

3. The Start Splits the Destination

Computation 3.1: f **and** g **were descended from 1001 starts, and destinations counted.**

quantity	value
starts reaching the left minimum of f	521 (share 0.520480)
starts reaching the right minimum of f	480
neither	0
share from the position of the maximum, $(x_{\max}+2)/4$	0.520931
destination for g	all -0.226699 (spread 1.027×10^{-15})

Observation 3.2

this is the core: the start splits the destination, and nearly half reach the shallower bottom.The global minimum is on the left, yet a share 0.479520 of starts reached the right local minimum——gradient descent seeks whichever bottom the downhill from the start leads to, and does not know which is deeper.

Observation 3.3

the boundary is set by the position of the maximum.Starts left of the maximum at 0.083725 go left, those right of it go right; the formula share 0.520931 differs from the grid share by 0.000452, less than the grid step of 0.004——the basins of attraction are fixed in advance by the shape of the function.

Observation 3.4

if convex, the start does not matter.For g all 1001 points reached -0.226699, with a spread of 1.027×10^{-15} ——with a single bottom, every descent ends at the same place.

4. The Separator

Observation 4.1

the separator is whether the second derivative is positive everywhere or negative somewhere.Positive everywhere, the start has no effect; negative somewhere, the maximum around it divides the basins——"the minimum was found by gradient descent" does not say which minimum until one says whether the function is convex.

Observation 4.2

beside Paper 734.Paper 734 showed curvature capping the step size on quadratic functions, and stated that non-convex surfaces would not be treated——there every start reached the same point; here, on a non-convex surface, the start splits the destination.

Observation 4.3

beside Paper 507 as well.Paper 507 showed that a convex function returns to itself after two Legendre transforms——convexity guarantees return under the transform, and indifference to the start under descent.

Observation 4.4

what is not said.Multidimensional saddles, stochastic methods and real learning problems are not treated ——what is shown is that for a non-convex function the start decides the answer, with the boundary set by a maximum.

5. Roll Call and Closing

1. **left minimum -1.264459** (Computation 2.1): **the deep bottom.**
2. **right minimum 1.180734** (Computation 2.1): **the shallow bottom.**
3. **maximum 0.083725** (Computation 2.1): **the boundary.**
4. **depths -2.872457 and -1.648425** (Observation 2.2): **global on the left.**
5. **right 0.479520** (Observation 3.2): **the core.**
6. **left 0.520480** (Computation 3.1): **grid share.**
7. **formula 0.520931** (Observation 3.3): **set by the maximum.**
8. **difference 0.000452** (Observation 3.3): **below the step.**
9. **convex -0.226699** (Observation 3.4): **start does not matter.**
10. **second derivative negative somewhere or not** (Observation 4.1): **the separator; Paper 734 and Paper 507.**

Closing: **For a function that is not convex, where you start decides the answer.$f(x) = x^4 - 3x^2 + 0.5x$ has a left minimum at -1.264459 ($f = -2.872457$) and a right minimum at 1.180734 ($f = -1.648425$), with a maximum at 0.083725 between——the global minimum is on the left.Descending from 1001 points on $[-2, 2]$ splits them 0.520480 left and 0.479520 right, so nearly half reach the shallower bottom; the boundary is set by the maximum, the formula share 0.520931 differing from the grid by 0.000452——gradient descent does not know which bottom is deeper.For the convex $x^4 + x^2 + 0.5x$ all 1001 points reached -0.226699, with a spread of 1.027×10^{-15}.The separator is whether the second derivative is positive everywhere or negative somewhere——"the minimum was found by gradient descent" does not say which minimum until one says whether the function is convex.Placed among the earlier papers——Paper 734 kept to quadratics and set non-convex surfaces aside; this counts, beyond that, the start splitting the destination.To be honest——only one-dimensional functions are considered, with no saddles, stochastic methods or real problems.**

References

1. Boyd, S., & Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.
2. Nocedal, J., & Wright, S. J. (2006). *Numerical Optimization* (2nd ed.). Springer.
3. Yamagishi, Y. (2026). *Three Separate Answers Are All Decided Where Something Is Tangent* [Paper 507]. Zenodo. 10.5281/zenodo.22861868
4. Yamagishi, Y. (2026). *What Caps the Step Size Is Not the Size of the Gradient but the Curvature in the Steepest Direction* [Paper 734]. Zenodo. 10.5281/zenodo.23015057

— From thinking to syncing. —

Yuuki YAMAGISHI (山岸 勇輝)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)