Classification problem and the Kernel Trick

Sometimes a line will not be good enough to split our data. *What we need is a more complex model.*
Our points are in a line, but here we see them on a plane. To exploit that, we switch from a one-dimensional problem, the line, to a two-dimensional problem, the plane, by adding a y-axis. And then use a parabola. Let's draw the function
$y = x ^\wedge 2$. Then we lift every point to its corresponding place in the parabola. All of a sudden, our points are nicely separable because now, using the original SVM algorithm, we can find a good cut. It's the line $y=4$.

How do we bring this back to the line and find the boundary there? Our original equation is $y=x ^\wedge 2$. Our line is $y=4$.

How did these two combine? By equating them, we get a new equation:
$x ^\wedge 2=4$, which factors into two linear polynomials with solutions

- $x=2$
- $x=-2$

So, those will make our boundary. We bring this back down to the line and we have $x=2$ and $x=-2$ as the boundaries for this model. Notice that they split the data really well.
This trick is known as the Kernel Trick and it's widely used in Support Vector Machines, as well as many other algorithms in machine learning.

Quiz Question

When separating points, when does it make sense to use a polynomial solution

a. When the points can be divided vertically
b. When the points can be divided horizontally
c. When the points cannot be divided by a straight line