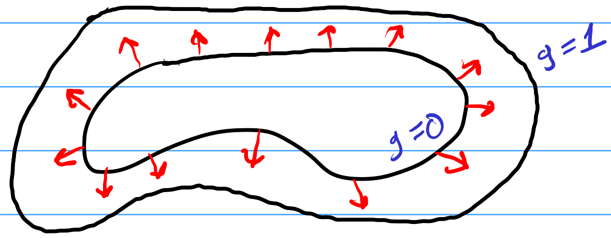
LAGRANGE MULTIPLIERS DERIVATION

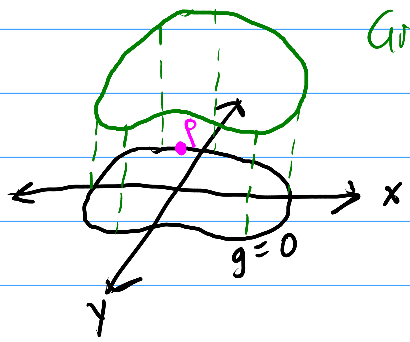
Given $f(x,y)$ and asked to minimize or maximize subject to constraint $g(x,y)=0$.

$g(x,y)=0$ is a level curve of g



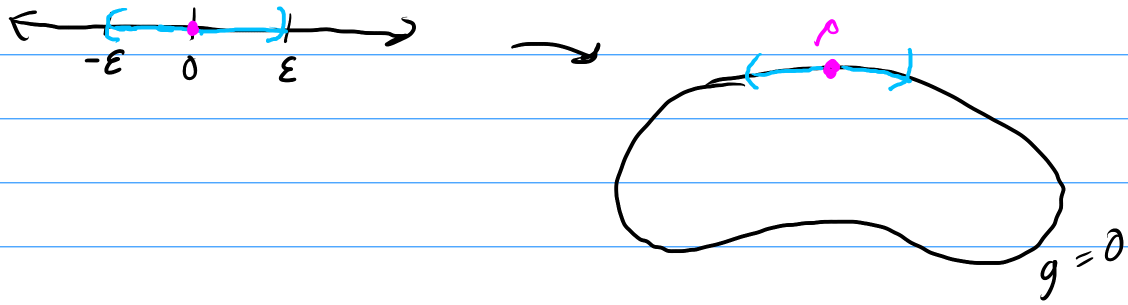
∇g is the direction of steepest ascent of g
$\hookrightarrow$ ∇g is everywhere normal to $g=0$

Now suppose p is a point on $g=0$ at which f is optimized.



Graph of f over $g=0$.

Choose a parametrization $\gamma : (-\varepsilon, \varepsilon) \longrightarrow \{g=0\}$ with $g(0)=p$. (for example, by arc length)



Since p is an extremum of f on the curve, $f(\gamma(t))$ has a critical point at $t=0$.

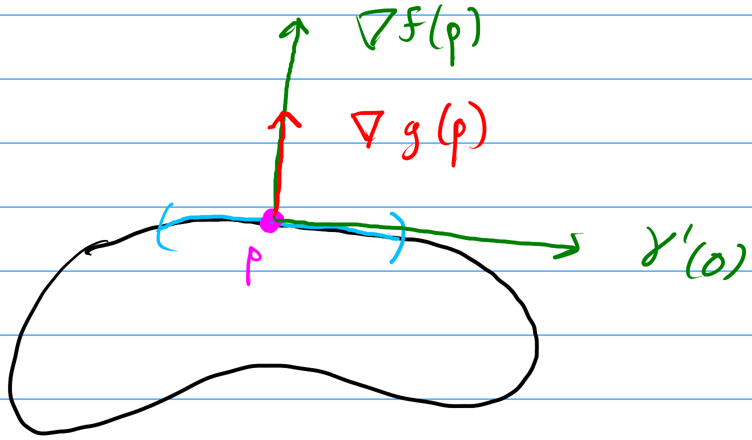
Differentiate with respect to t:

$$0 = \frac{d}{dt}\Big|_{t=0} \left[f(\gamma(t)) \right]$$
$$= \nabla f(\gamma(0)) \cdot \gamma'(0) \qquad \text{(chain rule)}$$
$$= \nabla f(p) \cdot \gamma'(0)$$

$\gamma'(0)$ is a tangent vector to the curve $\{g=0\}$ at p

$\longrightarrow$ $\nabla f(p)$ is normal to the curve

So the whole picture looks like this:



Thus we conclude $\nabla f(p) = \lambda \nabla g(p)$ for some number λ.