Recall: these give a way of maximizing or minimizing a function $f(x_1, \ldots, x_n)$ (of $n$ variables) subject to the constraint $g(x_1, \ldots, x_n) = 0$, by solving the system of $n + 1$ equations given by $g = 0$ and $\vec{\nabla} f = \lambda \vec{\nabla} g$. This does away with the need to parametrize the level set $g = 0$.

Note that this can be used to find the max/min of a function on the boundary of a region $S$ (as part of an unconstrained extremum problem), when that boundary takes the form $g = 0$.

Sometimes reality imposes more than one constraint on your variables. The simplest case is when you have a function of 3 variables $x, y, z$, and we want to maximize or minimize $f(x, y, z)$ subject to $g(x, y, z) = 0$ and $h(x, y, z) = 0$. The latter two conditions define a curve $C$ as the intersection of 2 surfaces $S_1$ and $S_2$ in $\mathbb{R}^3$ (defined by $g = 0$ resp. $h = 0$).

Lagrange multipliers can handle this situation too. The general form of a vector normal to $C$ is a linear combination of vectors normal to $S_1$ and $S_2$, which is to say $\vec{\nabla} g$ and $\vec{\nabla} h$. For the same reasons as before, if $f$ is maximized at a point $(x_0, y_0, z_0)$ on $C$, then $(\vec{\nabla} f)(x_0, y_0, z_0)$ must be normal to $C$. So you conclude that we must have

$$\vec{\nabla} f = \lambda \vec{\nabla} g + \mu \vec{\nabla} h$$

at a maximum or minimum. (Here $\lambda$ and $\mu$ are the “Lagrange multipliers”.) Together with $g = 0$ and $h = 0$, this gives a system of 5 equations in 5 variables (i.e. $x, y, z, \lambda, \mu$).

The case of $k$ constraints in $n$ variables is a straightforward generalization: there will be $k$ multipliers. But we won’t do more than 2 constraints.