Theorem 51.41. *(Theorem 28.3, Rockafellar) Let (P) be an ordinary convex program. If $x \in \mathbb{R}^n$ and $(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p$, then (λ, μ) and x have the property that*

(a) The infimum of the function $h = J + \sum_{i=1}^m \lambda_i \varphi_i + \sum_{j=1}^p \mu_j \psi_j$ is finite and equal to the optimal value of J over U, and

(b) The vector x is an optimal solution of (P) (so $x \in U$),

iff (x, λ, μ) is a saddle point of the Lagrangian $L(x, \lambda, \mu)$ of (P).

Moreover, this condition holds iff the following KKT conditions hold:

(1) $\lambda \in \mathbb{R}^m_+$, $\varphi_i(x) \leq 0$, and $\lambda_i \varphi_i(x) = 0$ for $i = 1, \ldots, m$.

(2) $\psi_j(x) = 0$ for $j = 1, \ldots, p$.

(3) $0 \in \partial J(x) + \sum_{i=1}^m \lambda_i \partial \varphi_i(x) + \sum_{j=1}^p \mu_j \partial \psi_j(x)$.

Observe that by Theorem 51.40, if the optimal value of (P) is finite and if the constraints are qualified, then Condition (a) of Theorem 51.41 holds for (λ, μ). As a consequence we obtain the following corollary of Theorem 51.41 attributed to Kuhn and Tucker, which is one of the main results of the theory. It is a generalized version of Theorem 50.18.

Theorem 51.42. *(Theorem 28.3.1, Rockafellar) Let (P) be an ordinary convex program satisfying the hypothesis of Theorem 51.40, which means that the optimal value of (P) is finite, and that the constraints are qualified. In order that a vector $x \in \mathbb{R}^n$ be an optimal solution to (P), it is necessary and sufficient that there exist Lagrange multipliers $(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ such that (x, λ, μ) is a saddle point of $L(x, \lambda, \mu)$. Equivalently, x is an optimal solution of (P) if and only if there exist Lagrange multipliers $(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p$, which, together with x, satisfy the KKT conditions from Theorem 51.41.*

Theorem 51.42 has to do with the existence of an optimal solution for (P), but it does not say anything about the optimal value of (P). To establish such a result, we need the notion of dual function.

The *dual function* G is defined by

$$G(\lambda, \mu) = \inf_{v \in \mathbb{R}^n} L(v, \lambda, \mu).$$

It is a concave function (so $-G$ is convex) which may take the values $\pm\infty$. Note that maximizing G, which is equivalent to minimizing $-G$, runs into troubles if $G(\lambda, \mu) = +\infty$ for some λ, μ, but that $G(\lambda, \mu) = -\infty$ does not cause a problem. At first glance, this seems counterintuitive, but remember that G is *concave*, not *convex*. It is $-G$ that is convex, and $-\infty$ and $+\infty$ get flipped.

Then a generalized and stronger version of Theorem 50.19(2) also holds. A proof can be obtained by putting together Corollary 28.3.1, Theorem 28.4, and Corollary 28.4.1, in Rockafellar [138]. For the sake of completeness, we state the following results from Rockafellar [138].