CHAPTER 3
Minimization of Convex Functions

Utilizing the Gâteaux variations from the preceding chapter, it is straightforward to characterize convexity for a function $J$ on a subset $\mathcal{D}$ of a linear space $\mathcal{Y}$, such that a convex function is automatically minimized by a $y \in \mathcal{D}$ at which its Gâteaux variations vanish.\(^1\) Moreover, in the presence of strict convexity, there can be at most one such $y$. A large and useful class of functions is shown to be convex. In particular, in §3.2, the role of [strongly] convex integrands $f$ in producing [strictly] convex integral functions $F$ is examined, and a supply of such $f$ is made accessible through the techniques and examples of §3.3. Moreover, the Gâteaux variations of integral functions will, in general, vanish at each solution $y$ of an associated differential equation (of Euler–Lagrange).

The resulting theory extends to problems involving convex constraining functions (§3.5), and it is used in several applications of interest including a version of the brachistochrone (§3.4), and the hanging cable (or catenary) problem of Euler. Additional applications will be found in the problem set together with extensions of this theory to other types of integral functions.

In this chapter only those conditions sufficient for a minimum are considered, and it is shown that in the presence of strict convexity they can supply a complete and satisfactory solution to the problems of interest. In particular, we may be able to ignore the difficult question of $a$ priori existence of a minimum by simply exhibiting the (unique) function which minimizes. Actually, the direct approach developed here within the framework of convexity is more basic and extends (in principle) to other problems (§3.6).

\(^1\) The definitions of functional convexity employed in this book incorporate, for convenience, some presupposed differentiability of the functions. For a convex set $\mathcal{X}$, less restrictive formulations are available, but they are more difficult to utilize. (See Problem 0.7, [F1], and [E–T].)
§3.1. Convex Functions

When \( f \in C^1(\mathbb{R}^3) \) then for \( Y = (x, y, z) \), \( V = (u, v, w) \in \mathbb{R}^3 \), we have\(^\dagger\)
\[
\delta f(Y; V) = \nabla f(Y) \cdot V; \quad \text{as in §2.4, Example 1);}
\]
moresover, \( f \) is defined to be convex (§0.8) provided that\(^\dagger\)
\[
f(Y + V) - f(Y) \geq \nabla f(Y) \cdot V = \delta f(Y; V),
\]
and strictly convex when equality holds at \( Y \) iff \( V = \emptyset \) (§0.9). We also observe that minimization of a convex function \( f \) may be particularly easy to establish, in that a point \( Y \) at which \( \nabla f(Y) = \emptyset \) clearly minimizes \( f \). (1) suggests the following:

(3.1) **Definition.** A real valued function \( J \) defined on a set \( \mathcal{D} \) in a linear space \( \mathcal{Y} \) is said to be [strictly] convex on \( \mathcal{D} \) provided that when \( y \) and \( y + v \in \mathcal{D} \) then \( \delta J(y; v) \) is defined and \( J(y + v) - J(y) \geq \delta J(y; v) \) [with equality iff \( v = \emptyset \)].

(This is to be considered as two assertions: the first is made by deleting the bracketed expressions throughout, while the second requires their presence.) \( \mathcal{D} \) itself may be nonconvex. (See 3.15)

Although "most" functions are not convex, a surprisingly large number of those of interest to us are convex—even strictly convex—as the applications will show. The following observation will prove valuable:

(3.2) **Proposition.** If \( J \) and \( \bar{J} \) are convex functions on a set \( \mathcal{D} \) then for each \( c \in \mathbb{R} \), \( c^2 J \) and \( J + \bar{J} \) are also convex. Moreover, the latter functions will be strictly convex with \( J \) (for \( c \neq 0 \)).

**Proof.**
\[
(c^2 J + \bar{J})(y + v) - (c^2 J + \bar{J})(y) = c^2(J(y + v) - J(y)) + (\bar{J}(y + v) - \bar{J}(y)) \\
\geq c^2 \delta J(y; v) + \delta \bar{J}(y; v) = \delta(c^2 J + \bar{J})(y; v), \quad \text{if } y, y + v \in \mathcal{D}, \text{ by 3.1.}
\]

This establishes the convexity of \( J + \bar{J} \) (when \( c^2 = 1 \)) and of \( c^2 J \) (when \( \bar{J} = \emptyset \)). Moreover, when \( J \) is strictly convex and \( c \neq 0 \), then there must be strict inequality except for the trivial case of \( v = \emptyset \). \( \square \)

(3.3) **Proposition.** If \( J \) is [strictly] convex on \( \mathcal{D} \) then each \( y_0 \in \mathcal{D} \) for which \( \delta J(y_0; v) = 0 \), \( \forall y_0 + v \in \mathcal{D} \) minimizes \( J \) on \( \mathcal{D} \) [uniquely].

**Proof.** If \( y \in \mathcal{D} \), then with \( v = y - y_0 \)
\[
J(y) - J(y_0) = J(y_0 + v) - J(y_0) \\
\geq \delta J(y_0; v) = 0, \quad \text{by 3.1 and hypotheses}
\]
[with equality iff \( v = \emptyset \)].

Hence \( J(y) \geq J(y_0) \) [with equality iff \( y = y_0 \)] and this is the desired conclusion. \( \square \)